Text, Script and Language in Bactria and Serindia:

Papers on cultural and linguistic interactions in pre-Islamic Central Asia

Edited by
Ching Chao-jung and Michaël Peyrot

Text, Script and Language in Bactria and Serindia

BEITRÄGE ZUR IRANISTIK

Gegründet von Georges Redard, herausgegeben von Agnes Korn

Band 55

Text, Script and Language in Bactria and Serindia: Papers on cultural and linguistic interactions in pre-Islamic Central Asia

Edited by
Ching Chao-jung and Michaël Peyrot

WIESBADEN 2025
DR. LUDWIG REICHERT VERLAG

Text, Script and Language in Bactria and Serindia:

Papers on cultural and linguistic interactions in pre-Islamic Central Asia

Edited by
Ching Chao-jung and Michaël Peyrot

WIESBADEN 2025
DR. LUDWIG REICHERT VERLAG

Published with the support of
Kyoto University Hakubi Project Funding
and the European Research Council (ERC-2022-COG 101088902)

Cover image:
Oenochoe excavated from Fayaz-tepe Buddhist monastery in Old Termez in V. A. Livshits' collection. Photo: Pavel B. Lurje (cf. Inscription Ad-1 in P. B. Lurje: The "Unknown script" of Bactria: Unpublished materials and fresh interpretations, this volume).

Bibliographic information from the German National Library
The German National Library lists this publication in the German National Bibliography; detailed bibliographic data is available online at http://dnb.dnb.de.

Printed on acid-free paper
(age-resistant – pH 7, neutral)

© Dr. Ludwig Reichert Verlag Wiesbaden 2025
Tauernstraße 11, 65199 Wiesbaden
info@reichert-verlag.de, www.reichert-verlag.de

ISBN: 978-3-7520-0924-8 (Print)
eISBN: 978-3-7520-0363-5 (Ebook)
https://doi.org/10.29091/9783752003635

The work, including all its parts, is protected by copyright.
Any use outside the narrow limits of copyright law is not allowed without the publisher's permission and is punishable by law. This applies in particular to reproductions, translations, microfilming, and storage and processing in electronic systems.
The publisher reserves the right to text and data mining in accordance with § 44b UrhG (German Copyright Act), which is hereby prohibited to third parties without the consent of the publisher.

Contents

Preface VII

General abbreviations IX

CHING Chao-jung
Bactria and Tukharistan in Chinese Buddhist scriptures:
A case study of three *Vibhāṣā* texts 1

Alessandro DEL TOMBA
A comparative study of the *Mahāvaidehaghṛta*
in Sanskrit, Khotanese, and Tocharian B 45

Federico DRAGONI
Was the Khotanese Brāhmī subscript hook borrowed
from the Kharoṣṭhī anusvāra? 105

Pavel B. LURJE
The "Unknown script" of Bactria:
Unpublished materials and fresh interpretations 131

Francesca MICHETTI
On the origin of Bactrian final -o 177

MIYAMOTO Ryoichi
Notes on Wakhsh and Rām-sēt in the Bactrian documents 235

OGIHARA Hirotoshi
A new look at ownership clauses in Tumshukese sale contracts 253

Michaël PEYROT
On the so-called "Fremdvokal" *ä* in Tocharian and Khotanese and its origins 275

Niels SCHOUBBEN
Gāndhārī light on Eastern Middle Iranian and vice versa:
Three new examples 297

Nicholas SIMS-WILLIAMS
The Bactrian inscription of Ayrtam: A minimal reading 321

List of contributors 339

Preface

The present volume collects ten papers dealing with a variety of scripts and languages in pre-Islamic Central Asia, with topics ranging from history and archaeology to philology, linguistics and palaeography. The papers are organized alphabetically by author, starting with a reconsideration of the cultural-geographical concept of "Bactria-Tukharistan" and closing with a new edition of the Ayrtam inscription.

Most of the papers not only cross several academic fields and historical regions but also widely explore documents and texts written in Bactrian, Khotanese, Tumshuqese, Sogdian, Tocharian B (i.e. Kuchean), Gāndhārī, Sanskrit, and even Chinese and Old Uyghur. Given the diverse disciplines and scholarly traditions in this collection, we have not attempted to uniformise terminology or style, and we only systematised the references and abbreviations to a limited extent.

We would like to express our most sincere thanks to Agnes Korn for initiating the idea of this volume following the panel "History and culture of pre-Islamic Afghanistan" moderated by Ching Chao-jung and Inaba Minoru of the *Tenth European Conference on Iranian Studies* (ECIS 10), which was organized by the Societas Iranologica Europaea and hosted by Leiden University (August 21–25, 2023). We took this opportunity to invite contributions of several researchers who either presented individual papers at the same conference or actively took part in the discussion. Thanks go to Pavel B. Lurje, who generously prepared the English translation of his recent Russian paper on the so-called "Unknown script" for this volume and gave a special online lecture during a monthly meeting of the collaborative research unit "Socio-economic base of ancient traffic in Inner Asia in the light of Chinese and non-Chinese sources" (Institute for Research in Humanities, Kyoto University) on May 14, 2025.

The publication of the present volume was financially made possible thanks to the support of Kyoto University Hakubi Project Funding and the European Research Council (ERC-2022-COG 101088902). We gratefully received typesetting suggestions from Alessandro Del Tomba and valuable reading notes from Agnes Korn, and we would like to extend our gratitude to Stefan Baums, Francesca Michetti and Pascale Eskes for their kind assistance in the reviewing and proofreading process. Finally, it is our great pleasure to thank Dr. Ludwig Reichert Verlag, who kindly accepted the publication of this volume in open access.

Kyoto & Leiden, November 2025
The editors

General abbreviations

AOH	*Acta Orientalia Academiae Scientiarum Hungaricae*
ARIRIAB	*Sōka daigaku kokusai bukkyō gaku kōtō kenkyūsho nenpō* 創価大学国際仏教学高等研究所年報 – *Annual Report of the International Research Institute for Advanced Buddhology at Soka University*
BAI	*Bulletin of the Asia Institute*
BEFEO	*Bulletin de l'École Française d'Extrême Orient*
BMM	*Miho Museum kenkyū kiyō* Miho Museum 研究紀要 – *Bulletin of Miho Museum*
BSOAS	*Bulletin of the School of Oriental and African Studies*
BSOS	*Bulletin of the School of Oriental Studies*
CAJ	*Central Asiatic Journal*
CRAI	*Comptes-rendus de l'Académie des inscriptions et belles-lettres*
IBK	*Indogaku bukkyōgaku kenkyū* 印度學佛教學研究 – *Journal of Indian and Buddhist Studies*
IF	*Indogermanische Forschungen*
IIJ	*Indo-Iranian Journal*
JA	*Journal Asiatique*
JAOS	*Journal of the American Oriental Society*
JIAAA	*Journal of the Inner Asian Art and Archaeology*
JIABS	*Journal of the International Association of Buddhist Studies*
JRAS	*Journal of the Royal Asiatic Society*
MSS	*Münchener Studien zur Sprachwissenschaft*
SAK	*Seinan ajia kenkyū* 西南アジア研究 – *Bulletin of the Society for Western and Southern Asiatic Studies*
SIAL	*Nairiku ajia gengo no kenkyū* 内陸アジア言語の研究 – *Studies on the inner Asian languages*
SPAW	*Sitzungsberichte der (königlich) Preußischen Akademie der Wissenschaften*
SRAA	*Silk Road Art and Archaeology*
StIr	*Studia Iranica*
TIES	*Tocharian and Indo-European Studies*
TPS	*Transactions of the Philological Society*
VDI	*Vestnik Drevnej Istorii*
WMO	*Written Monuments of the Orient*
XYWS	*Xiyu Wenshi* 西域文史 – *Literature and History of the Western Regions*
ZDMG	*Zeitschrift der Deutschen Morgenländischen Gesellschaft*

Bactria and Tukharistan in Chinese Buddhist scriptures: A case study of three *Vibhāṣā* texts*

CHING Chao-jung 慶昭蓉

This article analyzes the occurrences of the **Tukhāraka* language in the earliest Chinese translation of *Vibhāṣā* texts, T.1547 *Biposha lun* 鞞婆沙論 (tr. 383 CE), through a comparison with its counterparts in two later versions, T.1546 *Apitan piposha lun* 阿毘曇毘婆沙論 (tr. 437–439) and T.1545 *Apidamo da piposha lun* 阿毘達磨大毘婆沙論 (tr. 656–659). After an overview of the usage of "Bactria" and "Tukharistan" in Iranian studies and their known transcriptions in Chinese literature (§1), I go back to a note by Watters (1904), who connected Skt. *tukhāra-* with *Dōuqūluó* 兜呿羅 [1] in T.1509 *Da zhidu lun* 大智度論 (tr. 402–405/406, see §2). The discussion by Lévi (1897), Müller (1918), and Pelliot (1934) about this enigmatic language is then introduced, followed by a closer look at the translation process of T.1547 and the biography of Saṅghabhūti, the reciter of the Indic original (§3). In Section 4, the life of Dharmanandin from Tukharistan are reviewed, and in Section 5 I suggest, based on woodblock prints variations, which script and language he would most likely have used for noting down Saṅghabhūti's recitation, i.e. Brāhmī and Buddhist Sanskrit. The scholastic debate on the Buddha's language proficiency in T.1547 as the immediate context of the **Tukhāraka* language is then compared with the parallel exegesis in T.1546 and T.1545 (§6). Moreover, two Mahāyānic texts that mention the "language of Yuezhi" are reconsidered (§7). As a result, several hapaxes such as *Póqūlí* 婆佉梨 (EMC *ba-k^{h}ɨa-li*) in T.1547 and *Bóhēluó* 博喝羅 (EMC/LMC *pak-xat-la*) in T.1545 are proposed to be related to Balkh or Bactria, which seems to be geolinguistically distinguishable from Tukharistan in T.1547 and T.1545.

* This article is a preliminary result of my research supported by JSPS KAKENHI (Grant number 25K04495). It is an extended and revised version of Ching 2023, with elimination of the part about the *Mahāmāyūrī-vidyārājñī-sūtra*. I am grateful for the comments from Prof. Nicholas Sims-Williams, Prof. Funayama Tōru, Prof. Michaël Peyrot and all the panelists and attendants of the panel "History and culture of pre-Islamic Afghanistan" at ECIS10 (August 2023, Leiden), which was financially supported by the Murata Science and Education Foundation and Kyoto University Hakubi Project Funding. All errors and omissions remain my own.

[1] In this article, tone marks are in principle given only for transcriptions that are relevant to my main topic. Pulleyblank's EMC or LMC reconstruction is provided according to his periodisation (Pulleyblank 1991: 1–2) that the two are to be appoximately divided by 601 CE, the year of completion of the dictionary *Qieyun* 切韻. Although his theory does have minor problems (for instance, see Sims-Williams 2016: 276 on EMC *tr* ~ Sogd. *č*), these do not affect my central arguments here.

10.29091/9783752003635/001

1. Introduction

There have been differing views on the cultural and geographical distinction between Bactria and Tukharistan. The recent discovery of the ancient form τοχοαρστανο 'Tokhwarstan' in Bactrian, from the year 10 of Kanishka I (Sims-Williams 2015), has prompted scholars to explore the issue anew and in more detail.

Previously, in an historical map of Central Asia, de la Vaissière (2004: *Carte* 2) labeled the vast land between the Hissar Range and the Hindukush with the general term "*Bactriane-Tokharestan*". While archaeologists working on prehistory and the Hellenistic world usually see the drainage basin of the Amu Darya as Bactria and Northeast Afghanistan as "*Bactriane orientale*",[2] historians more often say "Tukharistan", following Chinese and Islamic Classics (Bactr. τοχοαραστανο, NP Ṭuxāristān, etc., BD II: 270; the spellings in English vary).[3] For example, the *Ḥudūd al-ʿĀlam* says that Khulm "lies between Balkh and Ṭuxāristān" and that Tukharistan was "a pleasant province consisting mostly of mountains" (Minorsky 1970: 108).[4] Therefore, Kuwayama is of the opinion that "Balkh was never included all the time in Tokharestan which only covers the modern Katagan (ancient Kadagstan) in the Surkhab valley and the Takkar area including Qalʿa-ye Zal on the Khanabad-Kunduz rivers" (p.c., email on 31 Aug. 2022).[5] Similarly, Inaba (2022: 38) states that seventh-century Tukharistan *sensu stricto* was between Balkh and Badakhshan. In one of his latest maps, de la Vaissière also separates the two regions (de la Vaissière 2024: *Carte* 2).

In Chinese chronicles, Tukharistan as a state or country (*guo* 國) appears in the *Weishu* 魏書 and the *Suishu* 隋書 as *Tǔhūluóguó* 吐呼羅國 and *Tǔhuǒluóguó* 吐火羅國, respectively (Ching & Galambos 2020: 40–42, 59–60). When Xuanzang passed there around 630, he found that Tukharistan (*Dǔhuòluó* 覩貨邏 / LMC *tuə́ʼ-xuaˋ-la*, as he transcribed Skt. *Tukhāra*-) was no longer unified but divided into 27 regions or polities (T.2087[LI]872a5–10). While Xuanzang considered Balkh (*Fúhè* 縛喝)[6] to be one of these, considerable discussion on the geographical extent of Tukharistan has arisen from

---

[2] Francfort 2022: 97. See also Bernard 2002, esp. p. 1289, fig. 1.

[3] For instance, see Shiio 1912, Kuwayama 1989, Arakawa 1997 (esp. p. 631, n. 4), Inaba 2022.

[4] See also Y. Yoshida's remark on *Fúdǐyē* 縛底耶, the capital city of Tukharistan (*Tǔhuǒluóguó* 吐火羅國) in the travelogue of Hyecho (Chin. Huichao 慧超, 704–787) from Korea as Commentary No. 145 in Kuwayama 1992: 146.

[5] Kuwayama's "Tokharestan" is narrower than Xuanzang's. In his view, the *Tǔhuǒluóguó* in the *Suishu* (SS, *j.* 83, 1853–1854) refers to the land in its strictest sense, namely the area of Qalʿa-ye Zal, known as Warwāliz or Valvālij in Islamic sources (Minorsky 1970: 109). See also Kuwayama 1992: 146, Commentary No. 144; Kuwayama 2020: 125–127; Kuwayama 2022b: 330.

[6] LMC *fĥjyak*/*fĥak-xat*, which can present Bactr. Βαχλο (Sims-Williams 2016: 279). See the attestations in T.2053[L]228a22–c1; T.2087[LI]872b29–873a24.

the fact that Termez had no king in his record (cf. de la Vaissière 2010a: 522), along with the problem that the polities he actually lists have been calculated to be 29 instead of 27 (cf. Kuwayama 2020: 124–125).

As for Balkh and Bactria, Enoki has largely complemented Bailey's list of their different transcriptions by providing the ones in Chinese.[7] In addition to *Fúhè*, he pointed out more variants in the Chinese *Tripiṭaka*, e.g. *Póqūluó* 婆佉羅,[8] *Bóqūluó* 薄佉羅,[9] *Fúkěluó* 縛渴羅,[10] and *Fúdǐyē* 縛底耶.[11] He further argued that the three foreign city names *Bótí* 薄提,[12] *Bózhī* 薄知,[13] and *Bóluó* 薄羅 [14] in the *Weishu* all denote Bactra, capital city of Bactria, during the Northern Wei 魏 period (386–535).[15]

---

[7] Bailey 1942: 22; Enoki 1958: 9–11; Enoki 1998: 63–66.

[8] EMC *ba-k^{h}ɨa-la*. The example in T.721[XVII]411a15–16 given by Enoki is wrong: It actually refers to a kind of flower and is given as *suōqūluó* 娑佉羅 in the Pilu canon preserved in the Library of Japanese Imperial Court. It either presents Skt. *bakula-/vakula-* (*Mimusops elengi*, MW: 719b; if the attested *poquluo* is correct) or a text corruption for Skt. *sakha-* (*Mimosa catechu*, MW: 1130b, in case the reading should be corrected to *suoquluo*). To my knowledge, *Póqūluó* as a state or polity is attested only in the Mahāyānic *Yuezang fen* 月藏分 (**Candragarbha-sūtra*) translated by Dharmakṣema (385–433, from India) *et al.*, see T.397[XIII]371c1, 374a25. At the time, Lévi (1905: 271, 283) restored *Póqūluó* as Skt. **vakhara-* "(?)".

[9] EMC *bak-k^{h}ɨa-la*, LMC *pɦak-k^{h}ia-la*. It is collected by Enoki from the biography of late sixth-century Dharmagupta (Damojiduo 達摩笈多) in T.2060[L]435a29. Incidentally, in the works of ZHU Fonian 竺佛念 (?– at least 413, see §3), this transcription is reserved for the Venerable Bakkula/Vak(k)ula.

[10] LMC *ɦjyak*/*ɦak-k^{h}at-la*, see T.2066[LI]2a11, 3b14. It is to be noted that *kě* 渴 (LMC *k^{h}at*) can match the [x] in Bactr. Βαχλο [vaxl] or an Indic form of it, where the plosive *t*, due to the logographic nature of Chinese writing, can theoretically present a vowel's shortness, unaccentedness or even unvoicedness. An example of *-t* rendering a relatively shorter or unaccented vowel is Dharmakṣema's rendition of Skt. *Darada-* (*Dáluótuó* 達羅陀, EMC *dat-la-da*, see Lévi 1905: 263). As for EMC/LMC *-t* used for rendering foreign consonant clusters, it is seen, for instance, in *shījī* 失雞 (LMC *ʂit-kjiaj*), the typical transcription of the Kuchean (Tocharian B) suffix *-śke* in anthroponyms (Ching 2011: 66 n. 12; Ching 2016: 42).

[11] LMC *ɦjyak*/*ɦak-tiaj´-jia*, in Hyecho's travelogue (Kuwayama 1992: 146–147).

[12] EMC *bak-dɐj*; a city in the state of Tuhuluo, cf. WS, *j.* 102, 2471; Ching & Galambos 2020: 40–42.

[13] EMC *bak-triă*/*tri*, cf. WS, *j.* 102, 2465. Concerning the city and the kingdom with the same name, Sims-Williams (2016: 276–277) supports Marquart's proposal (1901: 214) of linking it with Balkh/Bactria. As the direct source, he suggests Sogd. **vāxč* or a gentilic derivation **vāxčī*. The kingdom is termed *Baiti* 白題 by the Southern Dynasties in China, cf. de la Vaissière 2010a: 522; Yu 2012: 160–161; Ching & Galambos 2020: 42, 73, 85.

[14] EMC *bak-la*, i.e. the Kidarites' capital city after their westward migration, cf. WS, *j.* 102, 2468; Pelliot 1934: 42; Ching & Galambos 2020: 26–29.

[15] See also Tremblay 2004: 137 fn. 87, de la Vaissière 2010b: 215–216. However, Kuwayama (1985: 147–149) expresses his counterargument concerning *Bótí* and *Fúdǐyē*.

Interestingly, a Khotanese-Sanskrit text written on the verso of Pelliot Chinois 2787, a fragmentary scroll from Dunhuang containing the Chinese *Nyāyānusāriṇī* translated by Xuanzang in 653–654 (T.1582, *j.* 5), says that the family of Kanishka the Great originated "in the kingdom of Bāhlaka (Balkh), in Tokhāristān" (Khot. *bāhulaka-vaṣayä ttahvārasthaima* in Bailey's segmentation in 1942, §§158–159).[16] Since the Sanskrit counterpart *bāhūlaka-vaṣäye* (§§156–157) is in the locative singular, Bailey re-translated the Khotanese passage as "in the Bāhulaka-viṣaya Tahvārasthāna" (1965: 107), maintaining his interpretation of the Khotanese form of Tukharistan (**Tahvārasthāna*) as an explanation of the kingdom of Bactria (Skt. *Bāhulaka-*, cf. Bailey 1942: 22–23). No matter whether Khotanese Buddhists fully equated the **Tahvārasthāna* restored by Bailey with the kingdom of Balkh or not,[17] in their eyes these two terms could clarify each other (see Fig. 1). However, this could be due to Khotanese folk knowledge after the 7th century, and earlier Chinese Buddhist texts are to be re-examined.

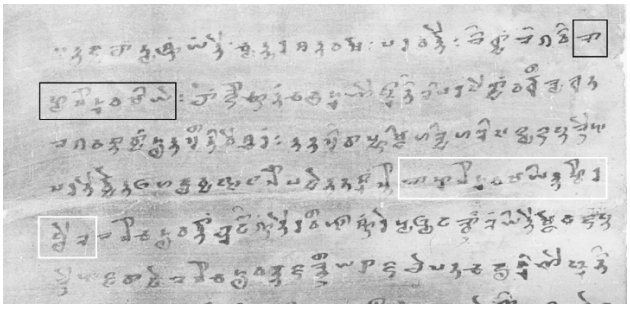


Fig. 1. Part of the Khotanese-Sanskrit text on the verso of Pelliot Chinois 2787.
Black frame: *bāhūlaka-vaṣäye* / White frame: *bāhulaka-vaṣayä ttahvārasthaima*
Microfilm image © Bibliothèque nationale de France; photoshopped by the author.

2. Tukharistan in T.1509 **Mahāprajñāpāramitā-upadeśa*

In his translation of Xuanzang's travelogue, T.2087 *Da Tang Xiyu ji* 大唐西域記, Watters (1904: 103) noted that Tukharistan is further found in T.1509 *Da zhidu lun* 大智度論 **Mahāprajñāpāramitā-upadeśa* (tr. 402–405/406), the *Treatise on the Great Virtue of*

[16] Bailey 1942: 18–19. See also Bailey 1969: 107, with slight changes of segmentation.

[17] Juxtaposition of Khotanese locative forms is uncommon. Another example is in Or.11252/2, a1: *cira kṣvā auvā* 'In Cira in the Six Villages' (Skjærvø 2003: 85–86), and Cira is known to be one of the six. The example shows that in Khotanese official documents an administratively inferior district (Cira, a village or town) appears before the prefecture or region (the "Six Villages") in which it was located. If this custom is applicable to the passage about Kanishka, then the kingdom of Balkh was a polity in Tukharistan as described by Xuanzang. Incidentally, the order is inverted in the Chinese documents issued there, e.g. *Liucheng Zhiluo* 六城質邏 '[in the] Six Villages, [at] Cira)'. On the administrative system in the kingdom of Khotan, see Rong & Wen 2009: 106–107.

Wisdom believed to be composed by the great Mahāyāna master Nāgārjuna (ca. 2nd cent. CE?), as well as in T.1547 *Biposha lun* 鞞婆沙論 **Vibhāṣā(-śāstra)*, a Sarvāstivāda compendium on doctrinal issues by an Arhat called *Sita- or *Śītapāṇi (Shituopanni 尸陀槃尼).[18] The transcription of Tukharistan in the former text is *Dōuqūluó* 兜呿羅 (EMC *təw-kʰiəʰ-la*), which is glossed by the main translator Kumārajīva (?–413) and his team as "Small/Lesser Yuezhi" (*Xiao Yuezhi* 小月氏) in small characters (in English translation placed in curly brackets):[19]

(1) *Da zhidu lun, j.* 25

弊生處者, 安陀羅、舍婆羅裸國也、**兜呿羅小月氏**、修利、安息、大秦國等。在此邊國中生, 若在大眾中, 則多怖畏。佛在迦毘羅婆中國生故無所畏。(cf. T.1509[XXV] 243a9–12)

'Inferior locations of one's birth: The states of Andhra, Śabara {the country of the naked}, **Tukhāra {Small/Lesser Yuezhi}**, Sogdiana (Xiuli 修利), Arsacid Parthia/Sassanid Iran (Anxi 安息), [and] Roman Orient (Daqin 大秦), etc. One who was born in such a country on the frontier would have a lot of terrible fears when being in the mass [of non-Buddhists]. The Buddha was born in Kapilavastu [in/of?] **Madhyadeśa*,[20] so he had no fear.' (tr. mine)

Born in the kingdom of Kucha and trained abroad – notably in Shale 沙勒 (Kashgar) and Jibin 罽賓 (Kashmir?)[21] – Kumārajīva no doubt knew the meaning of Tukhāra and its neighbouring countries. For this reason, the above passage was taken by Lévi (1933: 25–

[18] Berchert 1987: 314. On the legends about the origin and initial redaction of *Vibhāṣā* texts in Chinese and Tibetan sources, see Funayama 2021, esp. 89, 97–99, 102–104. On its school affiliation from a philosophical perspective, see the latest work by Fujimoto 2022. Sanskrit personal names reconstructed from Chinese transcriptions are asterisked when they are highly uncertain.

[19] Lévi 1933: 24–25; Pelliot 1934: 36; Lamotte 1970: 1584–1585; Ching & Galambos 2020: 19. I cite Chinese Buddhist texts with my own punctuation, segmentation, and slight revision of the Taishō edition by taking the general text structure and woodblock print variations into consideration.

[20] If *Jia-pi-luo-po-zhong-guo-sheng* 迦毘羅婆中國生 is not a mistake for *Jiapiluopoguo zhong sheng* 迦毘羅婆國中生 'born *in* the state of Kapilavastu', then the order of Kapilavastu (*Jiapiluopo* 迦毘羅婆) and *Zhongguo* 中國 'Middle land/country' reveals non-Chinese syntax (cf. fn. 17). On the range of *Zhongguo* as Central India, see Funayama 2013: fig. 2.

[21] Kumārajīva's early education is hidden behind legendary narratives. Pelliot (2002: 9) surmised that he studied abroad twice, during his childhood and during his adolescence. Jibin is Kashmir according to Zürcher (2007: 202, 226) and several other scholars, but it has been shown by Kuwayama (1990: 43–59) to be a typical name for Gandhāra even in the 4th–5th centuries on several occasions, cf. Funayama 2013: 4, 50. In the context of Kumārajīva's biography, I follow Yoshikawa and Funayama (2009: 147) in identifying Jibin more probably as Kashmir. Incidentally, the Bactrian form of Kashmir, κασπιρο, was recently discovered in a text dated to the late 4th century (Sims-Williams & Grenet 2023: 136, 141).

26) to show that Kuchean (Tocharian B) is distinct from the autochthonous language of Tukharistan. Likewise, he took it to prove that the Tukhāra people in Sanskrit texts are to be identified with the Yuezhi people who migrated westward from the Ordus Plateau or Hexi Corridor at the turn of the 2nd century BCE. As a result, von Richthofen's theory (1877: 439 fn. 5) that "Yuezhi = *Τόχαροι*" became widely accepted (see §7).

Enoki (1958: 20–22; 1998: 83–85) re-examined the passage critically in his study on the Kidarites. Pointing out the corresponding lemma in the *Fan Fanyu* 翻梵語, a Sanskrit-Chinese dictionary compiled in six-century southern China (T.2130 [LIV]1034 b16), he noticed a curious phenomenon: For Kumārajīva's translation team and their followers, Gandhāra proper and the Swat region (Uḍḍiyāna), both to the south of the Hindukush, were "Great(er) Yuezhi" (*Da Yuezhi* 大月氏 as name of country) or simply Yuezhi, whereas Tukharistan (= Bactria in Enoki's view) to the north of the Hindukush was "Small/Lesser Yuezhi".[22] This is markedly opposite to the traditional and official view since the compilation of the *Shiji* 史記 (written in the 2nd–1st cent. BCE) and the *Hanshu* 漢書 (written in the 1st–2nd cent. CE), which use "Yuezhi" not only to denote the nomads defeated by the Xiongnu but also to designate the Oxus plain and a part of the western Pamirs as the name of the people's new territory since the 130s BCE. In other words, after the decline of the Kushan Empire, Chinese official historians and some Buddhists living in China – including Kumārajīva and his team – had conflicting views on the relative mightiness, prosperity, or historicity with regard to the status of Gandhāra-Uḍḍiyāna and Bactria-Tukharistan.[23]

How to understand this discrepancy? First of all, one should note that in Kumārajīva's understanding *Dōuqūluó* is syntactically a realm or country, i.e. Tukharistan, rather than a people or tribe, although this is not so explicit due to his strategy of producing a concise translation: He would very likely refrain from adding *guo* 國 'state (in ancient Chinese feudalism); region, realm, polity, nation' to all the geographical names in the series.[24] Therefore, one may interpret "Tukhāra" in his source text – even though it no doubt comes from an earlier Indic tradition in which it usually designates a people – as a land that was formerly settled by the Tukhāras (or Tokharoi as they were called by the Greeks) and (afterwards?) ruled by the Yuezhi (as they were called by the Chinese). Before the

[22] Not to be confused with the Yuezhi's smaller branch remaining around Qinghai (China), cf. Pelliot 1934: 43. Nevertheless, Pelliot (1934: 44–45) was incorrect in identifying Kumārajīva's "Tukhāra / Small(er) Yuezhi" with Gandhāra or Northwest India.

[23] See Enoki 1958, Enoki 1998, and Ching & Galambos 2020: 26–31 on the phenomenon that Bactria-Tukharistan was termed "Great Yuezhi" and Gandhāra "Smaller Yuezhi" by Northern Wei historians. As set out by Enoki (1958: 47–48, fn. 48, 49; 1998: 82–83, fn. 99, 100), that Gandhāra was a colony of the Kidarites is not relevant for the identification of the branch of Yuezhi in Qinghai in the remote past.

[24] See Funayama 2013: 99–100 on a general evaluation of Kumārajīva's strategy.

Kidarites, it was Kujula Kadphises and his successors who assumed *Da Yuezhi* as the official Chinese name of their kingdom for conducting diplomatic and trading affairs with Eastern Han (25–220) and CAO Wei (220–265) dynasties.

Whether Kushan and Kidarite rulers were genealogically of Yuezhi stock or not is a question without final answer for the time being, but in the case of Imperial Iran, Chinese were used to call it Anxi 安息 as late as more than one century after the Sassanids overturned the Arsacids. It is according to the diplomatic protocol in the Han court that the Kushan Empire and its citizens were often called "Yuezhi 月氏" by Chinese. Accordingly, a number of immigrants from the empire's territory in the 2nd–3rd centuries CE chose ZHI 支 (for Yuezhi 月支 as a transcription variant) as their new surname.[25] This custom was observed by early Buddhists, including Dharmarakṣa, born in Dunhuang from a "Yuezhi" family (active in 266–306).[26] Nonetheless, Dharmakṣema (385–433) from India, who founded his famous school in the same city, used (or created?) a specific name for Balkh or Bactra, i.e. *Póqūluó* 婆佉羅 (see *supra*, fn. 8).

When the young Kumārajīva crossed the Pamirs and then a "mountain to the north of Yuezhi" (*Yuezhi beishan* 月氏北山, the Hindukush according to Enoki 1998: 85) before the 380s, Bactria-Tukharistan had been invaded by the Chionites and threatened by Shapur II appoximately from the 350s onwards (Grenet 2002: 220; Sims-Williams & Grenet 2023: 136; de la Vaissière 2024: 693–694). Even though it is generally assumed that Sasanian overlordship ceased there in 375–379 (Sims-Williams & Grenet 2023: 142), it is possible that the vast region known to the pious Buddhist Kumārajīva was culturally or economically inferior than Gandhāra, despite the fact that the latter was also obliged to accept the suzerainty of the Sassanids. Nevertheless, Buddhism existed in 4th-century Bactria-Tukharistan, as one can observe from the translation forum of T.1547 *Biposha lun* in China.

3. Tukharistan in T.1547 *Biposha lun* and its translation process

T.1547 *Biposha lun* was first referenced in a footnote about the ancient Sakas by Lévi (1897: 10–11, fn. 1), where he translated a passage from it:

> Le Bhagavat connaît la langue *Tchen-tan* mieux que les hommes du *Tchen-tan* ; le Bhagavat connaît la langue *Teou-k'iu-le* mieux que les hommes du *Teou-k'iu-le*.

[25] Umehara 2001: 288. See also Falk 2015: 126, Text 115; Ching & Galambos 2000: 14–15. In fact, there is no firm textual evidence to prove or reject the widespread assumption that Kujula was a Yuezhi prince by descent. Concerning the negative arguments among Sinologists, see Yu 2012: 6–9; Kuwayama 2017: 79–80; Kuwayama 2022a.

[26] Zürcher 2007: 65–70, but to say that his family is "Indo-scythian" (p. 65) is obsolete.

Confirming that this was the earliest mention of "Tukhāra" from the textual sources available then, Müller (1918: 575) retranslated it into German as follows:

> Der Buddha versteht die Sprache von Tšin-tan (= China) besser als ein geborener Chinese, und ebenso versteht er die Sprache von Tukharak besser als ein Eingeborener von Tukharak.

In the context of von Richthofen's theory, the initial motive of Lévi and his contemporaries was to identify the language of the Tokharoi, the invaders of the Graeco-Bactrian kingdom, also termed "the Indo-Scythians" in their discourse. Pelliot (1934: 34) supposed that the word in question, in pinyin *Dōuqūlè* 兜佉勒 (EMC *təw-kʰia-lək*), represents **Tukhāraka* or "*une forme iranisante* **Tukhāraga* [> *Tukhārag*]". In my opinion, although *Dōuqūlè* is in this passage again used for a realm (cf. *Zhēndàn* 真旦 for Skt. *Cīnasthāna-*; full text in §6.1), we may also consider a source form suffixed with °*ika*: Cf. Sogd. *swγδyk* 'Sogdian' transcribed as *Sùtè* 粟特, EMC *suawk-dək* (Tremblay 2004: 133); Sogd. *βγtyk/βxtyk* 'Bactrian' (de la Vaissière 2010b); Bactr. Βαμοιανιγο/ Βαμοιανιιο adj. or noun '(inhabitant) of Bamiyan' (BD II: 201a); Skt. *Bāhlika-*, *Bāhlīka-*, etc. 'Bactrian' (MW: 729c; see also Gandh. *Bahalieṇa* in CKI 60, instrumental singular of **Bahaliga* [https://gandhari.org/dictionary/bahaliga, access on 4 Oct. 2024]). In Section 5 I shall argue that the source text taught by Saṅghabhūti (Sengqiebacheng 僧伽跋澄)[27] was most probably in Sanskrit, but nevertheless the transcription of *Dōuqūlè* may reveal traces of the accent of his or Dharmanandin's native language. Until the discovery of an Indic version of T.1547, I tentatively keep to Pelliot's restoration **Tukhāraka* and flexibly use it as an adjective or noun.

The homeland of Saṅghabhūti, Jibin, again leads to ambiguity for scholars,[28] but Dharmanandin (Tanmonanti 曇摩難提)[29] was surely from Tukharistan (§4). Their collaboration in translating T.1547 is described in the biography of Saṅghabhūti in the *Gaoseng zhuan* 高僧傳 "*Biographies of eminent monks*":

(2) Biography of Saṅghabhūti (GSZ, *j.* 1)[30]

> 先是大乘之典未廣, 禪數之學甚盛, 既至長安, 咸稱法匠焉。符(for 苻)堅祕書郎趙正崇仰大法, 嘗聞外國宗習《阿毘曇毘婆沙》而跋澄諷誦, 乃四事禮供, 請譯梵文①。遂共

[27] Scholarly opinions differ on reconstructing Saṅghabhūti (adopted by Radich 2010: 122 and Fujimoto 2022: 23, etc.) or Saṅghabhadra (Zürcher 2007: 202, etc.). Here the former option is adopted. See also Yoshikawa & Funayama 2009: 110.

[28] CSZJJ, *j.* 10, T.2145[LV]73c3–8. I follow Funayama (2013: 74) in taking Jibin in his biography as "Gandhāra and/or Kashmir".

[29] It is also possible to reconstruct **Dharmanandika* (Pkt. °*diya*) or **Dharmananda* (Yoshikawa & Funayama 2009: 115).

[30] See also its earlier version in CSZJJ, *j.* 3, T.2145[LV]99a19–b9.

名德法師釋道安等集僧宣譯。**跋澄口誦經本, 外國沙門曇摩難提筆受為梵(?)文②,** 佛圖羅刹宣譯, 秦沙門敏智筆受為晉本。以偽秦建元十九年譯出, 自孟夏至仲秋方訖。(cf. T.2059[L]328b3–b10)

'Previously, the Mahāyāna scriptures were not widespread, and the studies on meditation and numerical [groups and categories of concepts][31] were popular. After [Saṅghabhūti's] arrival in Chang'an, all the people praised him [as a] professional master of the Law. ZHAO Zheng 趙正, Assistant in the Palace Library of FU Jian 苻堅 (338–385),[32] worshipped the Great Law faithfully. Having heard that [Buddhists in] foreign countries venerated and studied the **Abhidharmavibhāṣā* and that [Saṅgha]bhūti recited it, he offered him four things[33] and invited him to make a translation/transliteration/interpretation[34] [of/into/from?] the *fan* 梵 text/script/literature (*fanwen* 梵文①; Table 1). Therefore, he and the eminent master SHI Dao'an 釋道安 (312–385) convened the Buddhist community and translated [it] in front of them. **[Saṅgha]bhūti recited the original/body** (*ben* 本) **of the scripture;**[35] **a foreign *śramaṇa* named Dharmanandin wrote it down**[36] **in the script/text of *fan* 梵 (?)** (*fanwen* 梵文, var. *huwen* 胡文②; Table 2); Buddharakṣa translated it aloud [into Chinese];[37] a *śramaṇa* named Minzhi 敏智 of Qin 秦 wrote down [the result] as the version of Jin 晉 (sic!).[38] It was translated in the 19th year of the Jianyuan reign era of the illegal Qin [dynasty] (383), from the early summer to the middle of the autumn (i.e. from the 4th to the 8th Chinese month), until it was fully done. […]' (tr. mine)

In other words, Saṅghabhūti recited the source text aloud; Dharmanandin noted down his recitation into a non-Chinese (*hu* 胡 or *fan* 梵) script or text (*wen* 文); Buddharakṣa (Fotuluocha 佛圖羅刹) played the key role of orally translating Saṅghabhūti's speech into Chinese; and Minzhi produced a Chinese text as their final result. Conceivably, Dharmanandin's task was to prepare the first Indic manuscript of the treatise in China, for religious purposes and also for future examination of the translation. As for

[31] *shu* 數, lit. 'number', cf. Zürcher 2007: 203–204; Yoshikawa & Funayama 2009: 110–112.

[32] The 3rd emperor the FU Qin 苻秦 dynasty (350–394, also known as the Former Qin 前秦).

[33] These are: (1) drinks and foods, (2) clothing, (3) bedding, (4) hot water and medicines.

[34] The primary meaning of *yi* 譯 in pre-Han times is "official interpreter", who supported the communication with northern barbarians (Ching 2025). On the multivalent usage of this character by Buddhist translators as a verb 'translate; transliterate; do interpreting', see Funayama 2013, esp. pp. 75–77.

[35] *kou song jing ben* 口誦經本. The expression *kou song* 口誦 'recite orally' was confirmed by the postscript of T.194 written by a colleague of Dao'an in CSZJJ, *j.* 10, cf. T.2145[LV]71b26–27.

[36] *bi shou* 筆受, lit. 'received with the brush'.

[37] *xuan yi* 宣譯 (*yi chuan* 譯傳 in Dao'an's preface).

[38] *bi shou wei Jinben* 筆受為晉本. Dao'an's preface to T.1547 gives *bi shou wei ci Qinyan* 筆受為此秦言, lit. 'received with the brush in the language of this/our Qin'.

Buddharakṣa, this interpreter from an unknown country in the west had "widely read [Buddhist] scriptures" (*gai lan jingdian* 該覽經典) and "travelled in mainland [China] for a long time, being fluent in Han's language" (*jiu you zhongtu*, *shan xian Hanyan* 久遊中土, 善閑漢言), to the extent that "through his loud-voiced translation the *fan* 梵 text was greatly valued during the FU Qin dynasty".[39]

Usually, the character *fàn* 梵 (EMC *buam*^{h}) was reserved to render Skt. *brahman-* and *brāhmaṇa-* in Chinese Buddhist scriptures, but when denoting a script it means Brāhmī (Funayama 2013: 182; 2022: 117–124). However, due to the double meaning of *song* 誦 'to recite (from one's memory); to read (a text) aloud' and the polysemy of *wen* 文 'text, script, writing, literature, etc.' and *ben* 本 '(primary meaning) the lower trunk/stalk or root of a plant; (n.) body, base, essence, origin, version. etc.; (adj.) original, etc.; (adv.) originally', in addition to unsystematic revisions by later Buddhists based on their political and cultural ideology,[40] the above description is not crystal clear. Did Saṅghabhūti recite thoroughly from his memory or use a crib sheet with him? In the case of T.194 *Sengqieluocha suo ji jing* 僧伽羅刹所集經, it was indeed based on a manuscript brought to China himself (*ji ci jing ben* 齎此經本); the polyglot ZHU Fonian 竺佛念 was the interpreter[41] and Huisong 慧嵩 wrote the translation down in Chinese, but no one was responsible for recording his recitation in an Indic script. In the case of T.1549 *Zun Poxumi pusa suo ji lun* 尊婆須蜜菩薩所集論, a manuscript was provided by Saṅghabhūti as well, which was then held by him, Dharmanandin, and Saṅghadeva (Sengqietipo 僧伽提婆, from Jibin) together during the translation teamwork.[42] As for our T.1547 *Biposha lun*, Dao'an's preface states that *jing ben shen duo, qi ren wang shi* 經本甚多,其人忘失 (CSZJJ, *j.* 10, cf. T.2145[LV]73c11) in order to explain the situation that only forty chapters were successfully translated into Chinese. If we take *ben* 本 in this sentence as a noun rather than an adverb, then it means "the original of the scripture was of a big number (of stanzas or folios?), [but] this man (= Saṅghabhūti) forgot [and] lost [a part of it]", so the situation is obscure. In Section 5 I shall return to this problem.

---

[39] *Qi xuanyi fanwen, jian zhong Fu shi* 其宣譯梵文, 見重符(for 苻)世 in GSZ, *j.* 1, T.2059[L]326b16–18. The slight difference in CSZJJ, *j.* 13 (T.2145[LV]99b7–9) can be ignored.

[40] For example, the above paragraph demonstrates an overt inconsistency on the legitmacy of FU Jian's reign in Northern China. Likewise, the Chinese language appears variously as *Qinyan* 秦言, *Jinyan* 晉言, and *Hanyan* 漢言 in CSZJJ and GSZ.

[41] See CSZJJ, *j.* 10, T.2145[LV]71b16–23. Born in Liangzhou, ZHU Fonian was of obscure ethnicity. He was skilled in intertranslating non-Chinese and Chinese languages (*jiao yi Rong Hua* 交譯戎華), see CSZJJ, *j.* 13, T.2145[LV]99b26.

[42] See GSZ, *j.* 1 (T.2059[L]328b10–13), with text variants of *hu* and *fan* in CSZJJ, *j.* 10 and 13 (cf. T.2145[LV]72a2, 99b3). A detailed comparison is not attempted here.

4. Dharmanandin, a missionary from Tukharistan

At the translation forum sponsored by Former Qin's Imperial Secretary, Dharmanandin was more than a stenographer of the *Vibhāṣā* text. In fact, he served there as the primary teacher or so-to-speak "reciter" during the pioneering translation of the *Madhyama-āgama* and the *Ekottarika-āgama* in China. For his great achievements, Xuanzang regarded him as the leading translator in the first Chinese official institution of Buddhist translation activities (T.2053[L]266a21–22). Here is his biography in the *Biographies of eminent monks*:[43]

(3) Biography of Dharmanandin (GSZ, *j.* 1)

曇摩難提, 此云法喜, **兜佉勒人**。齠年離俗, 聰慧夙成, 研諷經典, 以專精致業。遍觀三藏, 闇誦《增一阿含經》, 博識洽聞, 靡所不綜, **是以國內遠近, 咸共推服。**少而觀方, 遍歷諸國, 常謂弘法之體, 宜宣布未聞, 故遠冒流沙, 懷寶東入, 以符氏建元中至于長安。

'Dharmanandin, in our language "law-joy"; **a man [of/born in] ****Tukhāraka***. Born smart and brilliant, he forsaked the secular world in his childhood. He studied and read scriptures, and he completed [his] studies [by gaining] special expertise. He read all the *Tripiṭaka* [and was able to] recite the *Ekottarika-āgama* in the darkness (*an song* 闇誦, i.e. without any written text in his hands); he was knowledgeable and well informed, and there was nothing that he did not master. **For this reason, [people] inside [his home] country, [living] far or nearby [his residence], all admired [his erudition].** He had travelled to other places in his youth and passed various countries; often saying that the entity [through which] the Law is spread (i.e. the sacred oral/written texts) should be openly disseminated for the ones who did not know it, he took the risk of crossing the Flowing Sand from afar and brought the treasure eastward into [China]. He arrived in Chang'an during the reign of Jianyuan (365–384) of FU [Qin].'

難提學業既優, 道聲甚盛, 符堅深見禮接。先是, 中土群經未有四《含》, 堅臣武威太守趙正欲請出經。時慕容沖已叛, 起兵擊堅, 關中擾動。正慕法情深, 忘身為道, 乃請安公等於長安城中集義學僧, 請難提譯出《中》、《增一》、二《阿含》, 并先無所出《毘曇心》、《三法度》等, 凡一百六卷。佛念傳譯, 慧嵩筆受, 自夏迄春, 綿涉兩載, 文字方具。及姚萇寇逼關內, 人情危阻, 難提乃辭還西域, 不知所終。(cf. T.2059[L]328b19–c4)

'[Dharma]nandin received such an excellent education, and his religious achievement [also] brought him great fame, that FU Jian received him very respectfully. Prior to this, there were no four *Āgama*s among the canons in the Middle Land (China). ZHAO Zheng 趙正, Governor of Wuwei, desired to invite him to issue (*chu* 出, lit. 'come out, go out, etc.') [these] scriptures. At that time, MURONG Chong 慕容沖 (359–386) had rebelled and raised an army to attack [FU] Jian (384), and [people in] the counties around the capital

[43] For an earlier, almost identical version in CSZJJ, *j.* 13, see T.2145[LV]99b11–27.

> city were jittery. [ZHAO] Zheng was deeply faithful to the Law, [so] he devoted his life (lit. 'forgot his body') to take the Path [of Buddhism]. Therefore he invited the venerable [Dao]an and others to Chang'an city, assembled monks [skilled in] exegesis, and invited [Dharma]nandin to translate [and] issue (*yi chu* 譯出) the two *Āgama*s – the *Madhyama* and the *Ekottarika* – as well as other texts which were never translated [into Chinese], such as the **Abhidharma-hṛdaya[-śāstra]* and the **Tridharmika[-śāstra]*. These amounted to 106 scrolls in total. [ZHU] Fonian transmitted them as the interpreter, and Huisong noted down [Fonian's Chinese speech]. From the summer to the spring, the [Chinese] texts were finally ready after two years. When YAO Chang 姚萇 (?–394) invaded the counties around the capital city (385), the situation became dangerous and [the forum was] impeded, so [Dharma]nandin resigned and returned to the Western Regions. No one knows about the end of his life.' (tr. mine)[44]

According to the preface (most probably by Dao'an) and the postscript of T.194 in *Chu sanzang ji ji* 出三藏記集 (CSZJJ), *j.* 10, the two *Āgama*s were translated in 384, during the war. Dharmanandin "orally recited" (*kou song* 口誦) the *Ekottarika-āgama* and perhaps also the *Huan wang jing* 幻網經 (now lost), while ZHU Fonian was his interpreter. No source tells whether Dharmanandin wrote down an Indic version of these three scriptures for his Chinese colleagues. In any case, the result was criticized to be "raw" by the school of Huiyuan 慧遠 (334–416) because he was "not fluent in Chinese".[45]

Incidentally, ZHU Fonian's preface to the *Wangzi Fayi huaimu yinyuan jing* 王子法益壞目因緣經 (T.2045) shows that he and Dharmanandin moved westward to Anding 安定 city (present-day Guyuan, Gansu) and translated that text from Dharmanandin's memory there in 391.[46] This is the latest news about him. These prefaces and hagiographies, unfortunately, do not specify Dharmanandin's scholastic affiliation nor the locales of his study. In the preface to T.1547, Dao'an called him a monk from "the country" (*qi guo* 其國) of Saṅghabhūti, so he possibly studied and stayed in Jibin for a significant period of time.[47]

---

[44] See also a partial Japanese translation in Hirakawa 1989: 29–33 and a full one in Yoshikawa & Funayama 2009: 115–118.

[45] *wei shan Jinyan* 未善晉言, cf. T.26[I]809b8. See also CSZJJ, *j.* 9 (T.2145[LV]63c26) and GSZ, *j.* 6 (T.2059[L]359b19–20). For this reason, Huiyuan and his guest, Saṅghadeva from Jibin, revised the two *Āgama*s and two treatises translated by Dharmanandin *et al*., and the results were said to be the current texts T.26, T.125, T.1550, and T.1506, respectively (see the summary in Zürcher 2007: 204).

[46] T.2045[L]172a17–b19. See also CSZJJ, *j.* 7, T.2145[LV]51b14–c16.

[47] I disagree with Enomoto's tentative suggestion (1993: 263) to include Tukharistan into the range of Jibin just because of this single mention (T.2145[LV]73c6) despite abundant sources that specify his home country. With this sort of reasoning, Tukharistan would have to be placed in India because in his preface to T.2045 ZHU Fonian claimed that he was an Indian monk or a monk from India (*Tianzhu shamen* 天竺沙門), cf. T.2045[L]172b12 and the same expression in T.2145[LV]51c9.

Several sources, including a Tibetan colophon about a *Vibhāṣā* master called *Ārya-Mūla-Sarvāstivāda-Mahā-Vinayadhara Tukhāra Vaibhāṣika Ācārya* (Fussman 2015: 192, 195), tell the long tradition of the (Mūla-)Sarvāstivādins in Bactria-Tukharistan by the 7th century, whereas a few sites such as Kara-Tepa (ca. 50–620 CE) and Fajaz-Tepa (ca. 50–400 CE) were inhabited by Mahāsāṅghika monks of different ethnic backgrounds knowing how to write various languages in their corresponding scripts: "Gāndhārī, coarse Sanskrit, Bactrian, the so-called undeciphered script and language, even Greek" (Fussman 2015: 193).[48] Dharmanandin was possibly a polyglot, but it is unknown how many languages and scripts he was able to use. Since the forum in Chang'an agreed with the term of "**Tukhāraka* language" in the translation, this might also be his mother tongue. However, it can hardly refer to the "Tocharian languages" in current Indo-European terminology (i.e. Tocharian A or Agnean, and Tocharian B or Kuchean, as suggested by Bailey 1937: 906), because the extant texts in "Tocharian" have been found only to the east of Pamirs, and there is no proof of their use on the plain of the Amu Darya.[49] Given that Kuchean (Tocharian B) writing tradition was probably still taking shape in 4th-century Kucha (Peyrot 2008: 205–206), Lévi and Bailey have good reason to terminologically distinguish the "Tocharian languages" from the languages spoken in Bactria-Tukharistan. In any case, since Buddhist remains are not abundant in Old Kunduz (Fussman 2015: 184), it does not seem necessary to limit Dharmanandin's homeland to the area of Qalʿa-ye Zal as argued by Kuwayama.

5. Text written in *hu* 胡 or *fan* 梵? The case of T.1547

The language of Saṅghabhūti's recitation is obscure. However, Dao'an and ZHAO Zheng, the founders of translation theory in China, are well known for preferring a literal translation rather than a flowery one. This is explicated in Dao'an's preface to T.1547:

(4) Dao'an's preface to the *Biposha lun* (CSZJJ, *j.* 10)
遂案本而傳, 不令有損。言游字時改倒句, 餘盡實錄也。(cf. T.2145[LV]73c21–22)
'Therefore, it is transmitted according to the original (*ben* 本) in order not to damage/lose [any text]. When the words in the [source] language are pleonastic (*you* 游, lit. 'swimming; floaty'), the syntax is modifed and rearranged (*dao* 倒, lit. 'inverted') [in our translation, but] the rest [of the content] is overall a honest record.' (tr. mine)

[48] Cf. Bonmann *et al*. 2023 and Lurje's article in this volume on the "undeciphered script and language".
[49] Sims-Williams 2002. On the problem of the terminology, see Pinault 2007: 131–136, esp. p. 134. I have no intention to challenge these conventional terms in linguistics, but a return to the initial discussion about the spoken languages in Bactria may help us to reflect the issues reactivated by the discovery of the Almosi inscriptions, since the so-called "undeciphered script" and related one(s?) possibly represent one or more different languages, including at least Bactrian (cf. Sims-Williams 2025: 209).

The same text also expresses their concern on the features of the source text and the result of translation:

(5) Dao'an's preface to the *Biposha lun* (CSZJJ, *j.* 10)

……曇無難提筆受為**梵文**①。……**胡本**②一萬一千七百五十二首盧, ……秦語為十六萬五千九百七十五字。……

'[…] Dharmanandin noted down with a brush in the ***fan* text/script** (*Fanwen* 梵文①). […] The ***hu* original/version** (*huben* 胡本②) amounts to 11,752 *śloka*s […] [it is translated into] Chinese (*Qinyu* 秦語) in 165,975 characters. […]'

趙郎謂譯人曰:「昔來出經者, 多嫌**胡言**③方質, 而改適今俗, 此政所不取也。何者? 傳**胡**④為秦, 以不閑方言, 求知辭趣耳。何嫌文質? 文質是時, 幸勿易之。」(cf. T.2145[LV]73c8–19)

'Mr. ZHAO told the translator (mainly to ZHU Fonian): "Previously, the ones who released [Buddhist] scriptures usually disliked the alienness and plainness of ***hu* language(s)** (*huyan* 胡言③) and thus adapted [their result] to our custom. This is not what Zheng, [your honest sponsor], want to have. Why? Concerning the transmission from ***hu* 胡**④ and transformation (lit. 'make', *wei* 為) [into] Qin 秦 [language/text], exactly because [we do] not have a command of the language(s) of [remote] area(s), we expect to know the [precise] wording and figurative meanings. Why do [you] dislike the plainness of the text? The text is to be plain [as it is] now; please do not change it."' (tr. mine, cf. Table 3)

The unit of *śloka* (= 32 syllables, Funayama 2021: 73, 86) evinces that the script used by Dharmanandin to record Saṅghabhūti's recitation is very possibly a kind of abugida that is relatively convenient to count the number of syllables, even when consonant clusters were heard. Theoretically it can be either Brāhmī or Kharoṣṭhī, as both were in Gandhāra used to write Sanskritised Gāndhārī in the 3rd–4th centuries, and Sanskrit was spread as a "new literary language" or even "a *lingua franca*" on Gandhāra in the beginning of the 3rd century (Strauch 2012: 133, 156–158). If we take the case of Mahāyāna Buddhism into consideration, then the language of a Mahāyānic scripture can be Prakrit (for oral transmission), broken Sanskrit mixed with Prakrit (2nd–3rd cent.), Buddhist Sanskrit (3rd–4th cent., cf. Karashima 2015: 113), or even an artificial language mixed of Sanskrit, Gāndhārī, and other Prakrits, as in the case of the Lotus Sutra analysed by Boucher (1998). Nevertheless, if Saṅghabhūti's home country was indeed Kashmir, the long-established center of the Sarvāstivādins, it must be stressed that this school was especially active with the Sanskritisation of canons (Boucher 1998: 473–474). Kharoṣṭhī was not influential in Kashmir, and it is even excluded from the Kharoṣṭhī writing area by Salomon (2024: 530). Therefore, if Saṅghabhūti was from Kashmir, the treatise he spread was no doubt in Sanskrit, or it was very Sanskritised, and the manuscript produced by Dharmanandin was conceivably prepared in Brāhmī.

In fact, T.1547 is the earliest Chinese text left to us that emphasizes the difference between the two scripts, explaining that Brāhmī is more educative:

(6) *Biposha lun*, *j.* 11

如學**梵**書已, 速學**佉樓**書; 非學佉樓, 速學梵書。(cf. T.1547[XXVIII]493b7–8)

'De même, c'est en ayant étudié l'écriture **brāhmī (*fan*)** qu'on rend plus rapide l'étude de l'écriture **kharoṣṭrī (*qulou*)** ; et ce n'est pas en étudiant l'écriture kharoṣṭrī qu'on rend plus rapide l'étude de l'écriture brāhmī.' (tr. Lévi 1904: 558, Chinese romanised in *pinyin*)

Therefore, Dao'an's translation forum was definitely aware of Brāhmī as an excellent script. This helps us to re-examine text variations in later woodblock prints concerning Saṅghabhūti's biography in GSZ (Text 2) and Dao'an's preface to T.1547 (Text 5).

Concerning Saṅghabhūti's biography, although a full investigation of extant copies is required, all the representative prints give "*fanwen* 梵文" at position ①, not only in the GSZ version but also in the earlier and more precise version of CSZJJ, *j.* 13, which is widely perceived as the main and immediate source of most of the hagiographies in GSZ (Table 1):

Text 2-①

……請譯**梵文**

'[ZHAO Zheng …] invited him (= S.) to make a translation/transliteration/interpretation [of/into/from] **the *fan* 梵 text/script/literature**.'

	Date of carving (or year of completion)	Biography of S. in GSZ, *j.* 10 = Text 2-①	Biography of S. in CSZJJ, *j.* 13
1st Korean carving	ca. 1011–1087	梵文 *fanwen*	梵文 *fanwen*
Pilu version	1148	(lost)	梵文 *fanwen*
Sixi version	1239	梵文 *fanwen*	梵文 *fanwen*
2nd Korean carving	1236–1251	梵文 *fanwen*	梵文 *fanwen*
Puning version	1277–1290	梵文 *fanwen*	梵文 *fanwen*

Table 1. Text variations of *fanwen* (①) in Saṅghabhūti's biography.

Therefore, in this sentence, the initial wording of the biography's anonymous author is in all probability *fanwen* 梵文. This clearly indicates the Indic nature of Saṅghabhūti's recitation despite the multivalence of the phrase. As to position ② in the same biography concerning the form of manuscript that Dharmanandin prepared, there are significant woodblock print variants (Table 2):

Text 2-②

……外國沙門曇摩難提筆受為**梵文** (following the carvings in Korea)

'[…] a foreign *śramaṇa* named Dharmanandin wrote it down in the script/text of *fan*'

	Biography of S. in GSZ, *j.* 10 = Text 2-②	Biography of S. in CSZJJ, *j.* 13	Dao'an's preface to T.1547 in CSZJJ, *j.* 10 = Text 5-①
1st Korean carving	梵文 *fanwen*	胡文 *huwen*	梵文 *fanwen*
Pilu vers.	(lost)	胡文 *huwen*	梵文 *fanwen*
Sixi vers.	胡文 *huwen*	胡文 *huwen*	梵文 *fanwen*
2nd Korean carving	梵文 *fanwen*	胡文 *huwen*	梵文 *fanwen*
Puning vers.	梵文 *fanwen*	梵文 *fanwen*	梵文 *fanwen*

Table 2. Text variations of fanwen (②) in Saṅghabhūti's biography.

Theoretically, the variations in Table 2 may result from several stages of text corruption (for instance, a copyist or carver was confused by other hagiographies), or from hypercorrection by Tripiṭaka editors in different periods, but it seems very likely that all the relevant occasions were uniformised to *fanwen* 梵文 in the late 13th century. The expression in Saṅghabhūti's biography is more flexible, which might reflect the attitude of the initial author and later editors, while in Dao'an's preface it is again specified with *fan* 梵. His usage in the same preface can be further illustrated as below (Table 3):

	Text 5-①	Text 5-②	Text 5-③	Text5-④
	wen 文 'text; script, etc'	*ben* 本 'version; original, etc.'	*yan* 言 'language, speech, etc.'	— (a language or text is implied)
1st Korean carving [N]	梵 *fan*	胡 *hu*	胡 *hu*	胡 *hu*
Pilu vers. [S]	梵 *fan*	胡 *hu*	胡 *hu*	胡 *hu*
Sixi vers. [S]	梵 *fan*	梵 *fan*	梵 *fan*	梵 *fan*
2nd Korean carving [N]	梵 *fan*	胡 *hu*	胡 *hu*	胡 *hu*
Puning vers. [S]	梵 *fan*	梵 *fan*	梵 *fan*	梵 *fan*

Table 3. Text variations in Dao'an's preface to T.1547.

At the first glance, this table looks rather complicated. However, it is known that the first and second Korean carvings descend from a branch of manuscript-blockprint tradition going back to the Northern Song 宋 and Liao 遼 dynasties in Northern China (960–1127 and 916–1125, respectively), while the Pilu, Sixi (usually called the "Song" print), and Puning (the "Yuan 元" print) versions were carved in Southern China during the late Southern Song (1127–1279) and then Yuan period under Mongolian rule. It is therefore clear that the northern tradition (marked with 'N' in Table 3) specifies the written form (Text 5-①) as *fan* but uses the vague term *hu* for other occasions. In contrast, the southern tradition tended to unify all as *fan*, but the earlier Pilu print still retains old features.

In other words, Table 3 shows a trend to change *hu* to *fan* around the 13th century in Southern China. This is less explicit, but nevertheless also observable in Table 2. It is

further detectable in several other essays and hagiographies in GSZ and CSZJJ, but I make no further comparison here. At least in the case of Dao'an's preface to T.1547, such a "switch from *hu* to *fan*" (Boucher 2000: 17) did not occur during the Sui (581–619) and Tang (618–907) dynasties (*contra* Yang 1998 and Boucher 2000: 18), but in the Song period when Chinese intellectuals were under great threat from the Mogolians and other northern peoples and would have avoided or disliked the usage of *hu*, originally a term denoting the Xiongnu and other steppe nomads.

If we apply Boucher's theory "*hu* = *kharoṣṭhī*, *fan* = *brāhmī*" (2000, esp. p. 18) to the three tables, then they become perplexing, because *hu* is obviously also used to denote one or more languages in this context (esp. Text 5-③). Assuming that the first Korean carving of Dao'an's preface to T.1547 is exact, then his expressions "*fanwen*" (*fan* script/text, Text 5-①) but "*huben*" (*hu* original/version, Text 5-②) are in particular need of explanation: If *ben* 本 actually refers to the "original" oral text from Saṅghabhūti's memory, then Dao'an appears to be saying that Buddhist Sanskrit (or a quite sanskritised prakrit) is *hu*; if it means the "version" freshly recorded by Dharmanandin that can serve careful length calculation, then Dao'an appears to be saying that the Brāhmī script is *hu*.[50]

However, in my opinion this is just an example of the Chinese preference for vocabulary variation in a piece of *belle-lettre*, unlike the repetiton of formulae and fixed expressions in the Indic Buddhist canons. For instance, in Dao'an's elegant foreword to the *Yinchiru jing* 陰持入經 (CSZJJ, *j.* 6, cf. T.2145[LV]44b29–45a13), Śākyamuni is mentioned four times as (1) *dasheng* 大聖 'the Great Noble one'; (2) *shixiong* 世雄 'the World's Hero'; (3) *fawang* 法王 'King of the Law'; (4) *fo* 佛, a typical transcription of "Buddha", very likely based on a Middle Indic form. Since there was no exact equivalent of "Brāhmī" and "Sanskrit" in that time, it is not surprising that a multivalent word for this rhetoric technique is used. Accordingly, I would take *hu* in the three tables to be a vague expression for "foreign" (Boucher 1998: 485b) or "Western" (Radich 2017: 7),

[50] Since *hu* 胡 became a general term for various groups in the north and west of China, including Sogdians and other Central Asians (Boucher 2000: 17 n. 26; de la Vaissière 2005), one might wonder which language and script were written down concerning the various biographies and records in CSZJJ and GSZ. In my view, however, given that Dao'an clearly gives *fan* 梵 at the position of Text 5-①, Sanskrit and/or Brāhmī must have been involved in the case of T.1547, and we may at least rule out Graeco-Bactrian and the Iranian languages written in Aramaic. Moreover, with regard to Dharmanandin's translation activities, *hu* 胡 can hardly be assumed to refer to Central Asian languages that adopted the Brāhmī script, such as Khotanese or Kuchean, since Khotanese and Kuchean Brāhmī, if in use at all, were still in an initial phase in the 4th century and in all likelihood not ready to render intricate philosophical treatises. Cf. Sander 2009; Ching *et al.* 2021: 74–79; Maggi 2022. See also Dragoni's and Peyrot's articles in this volume.

rather than a diagnostic indicator of Kharoṣṭhī or Iranian, and I adopt Funayama's proposal that "*fan* = Brāhmī and/or Sanskrit" (cf. Funayama 2013: 5, 74–75), which is easier to generalise and work with.[51]

The assumption that Dharmanandin mainly used Buddhist Sanskrit and Brāhmī throughout his career in China also helps to explain the scarcity of Bactrian Buddhist texts, because a basic knowledge of Sanskrit and Gāndhārī would be indispensable for a part of monastic populations. Concerning the *Vibhāṣā* and the *Āgama*s worked on by him, no Bactrian fragment has thus far been found, while Gāndhārī and Sanskrit fragments of the *Āgama* and *Abhidharma* genres were found in Gandhāra and to the north of it (Allon 2001; Matsuda 2002). Nevertheless, one may expect traces of his native language in some of his works, such as T.125 *Zengyi Ahan jing* 增一阿含經 (*Ekottarika-āgama*) or T.2045 *Wangzi Fayi huaimu yinyuan jing* 王子法益壞目因緣經, which seems to be the result of his last collaboration with ZHU Fonian. In fact, Radich (2017) and other researchers recently analyse T.125 stylistically and statistically, showing that it was essentially produced by Dharmanandin and ZHU Fonian, and not much affected by Saṅghadeva's revision (*contra* traditional view, cf. fn. 45; see also Radich & Anālayo 2017).

6. The **Tukhāraka* language in the *Vibhāṣā* texts

As implied by Lévi's translation (§3), the **Tukhāraka* language and its speakers in T.1547 were geographically remote, linguistically strange, but politically or culturally no less important than Chinese in the eyes of Indian or Kashmirian Buddhists. The language is mentioned in a scholastic but heated debate on the authenticity of an episode in Śākyamuni's life and the credibility of a related hymn. The background can be so summarized:[52]

> The four Heavenly Kings visited the Buddha and invited him to preach the Law. The Buddha explained the Four Noble Truths (*catvāryāryasatyāni*) to them in the **Ārya*

[51] The importance of Kharoṣṭhī cannot be denied, and Boucher (2000: esp. 17–18 n. 26) is certainly correct that the source language of early Chinese translations cannot be easily equated with any single one. But there are more exceptions to Boucher's identification "*hu* = *kharoṣṭhī*" than he tried to explain. In my view, one does not have to ascribe the *hu/fan* variations in the essays of Dao'an and his contemporaries so much to their carelessness or ignorance by calling them "confused" or "contradictory" (*contra* Boucher 2000: 16–17 fn. 25). Before a conclusion is made, late text variations, including manuscripts and woodblock prints, are to be fully analyzed.

[52] T.1545[XXVII]410a5–12 (German tr. in Chung & Wille 2002: 116–118); T.1546[XXVIII]306c16–21; T.1547[XXVIII]482c4–12. Here I omit Chinese transcriptions of the Buddha's speech in the Tanluo or Tuopoluo(?) and Mlecchа languages, which are also found in the Sanskrit *Udānavarga*, cf. Bernhard 1965: 323–326; Bernhard 1967; Bechert 1987: 388–389; Chung & Wille 2002: 116–118.

language (*shengyu* 聖語).[53] Two of them understood, but the other two did not. Then the Buddha explained the Four Truths again in the *Tuopoluo* 馱婆羅 language,[54] so one more king understood. Then the Buddha explained again in the *Miliche* 彌梨車 (Skt. *Mleccha-*) language, and then all the four kings understood his teaching.

The Sanskrit recension translated by Xuanzang into Chinese as T.1545 *Apidamo da piposha lun* 阿毘達磨大毘婆沙 **Abhidharma-mahā-vibhāṣā(-śāstra)* specifies that the episode is quoted from a Vinaya (collection of rules and precepts for monks and nuns). It has been identified with a section in the "Chapter on Medicines" (*Bhaiṣajyavastu*) of the Chinese Sarvāstivāda Vinaya.[55] As for the hymn, it is found in Mātṛceṭa's 'Praise of the Praiseworthy' (*Varṇārhavarṇa*) and several other texts.

Concerning this famous topic also known as the "one utterance" question, Hartmann (2024, esp. p. 192) points out that not all the schools of Nikāya Buddhism (or the "Small Vehicle", including the Sarvāstivādins) believed in Śākyamuni's extraodinary language proficiency, although it was defended by many Mahāyānists. Here my focus is, however, on the representative non-Aryan languages known to Chinese Buddhists from this debate, rather than the native language of Śākyamuni.

6.1. The exegesis on the Buddha's language proficiency in T.1547

As the essential source of this article, the respective text in T.1547 is translated here as full and literal as possible:[56]

(7) *Biposha lun*, *j.* 9

問曰: 世尊為四天王說四諦, 聖語, 為有力耶、無力耶? 若有力者, 何以故為二聖語說, 一曇羅國、一彌離車國語說? 若無力者, 本師偈云何通?

[53] Literally 'noble/sacred language'. According to Bernhard (1967: 149), this is Sanskrit. Hartmann (2024: 192) emphasizes that *Ārya* is to be understood here in an ethnic sense.

[54] Lévi's supposition (1905: 286) that it is Dravidian has not been substantiated, cf. Bernhard 1965: 324; Bernhard 1967: 163–164; Chung & Wille 2002: 116 n. 10.

[55] T.1435[XXIII]193a1–20. Unseen in the Chinese Mūla-sarvāstivāda Vinaya, parallel narratives are found in several Buddhist texts, including Sanskrit ones from Gilgit and Kucha, cf. Chung & Wille 2002. In the Chinese Sarvāstivāda Vinaya, the section appears as an interlude between the Buddha's initial regulation on the consumption of grapes and a later one on various fruit juices, including grape juice. These rules were perhaps made for Gandhāran and Central Asian societies being so fond of grape wine. It fits Enomoto's observation (1993: 265) that this Vinaya reflects a few philosophical and cosmological features of the Sarvāstivādins to the west of Kashmir rather than the Kashmirian orthodox.

[56] Here the present indicative and real conditional are used to translate the hymn and the exegesis to conform with its religious context, cf. the translation of the *Varṇārhavarṇa* by Hartmann (2024: 195), although historically it is unrealistic to envisage any Chinese among Śākyamuni's audiences.

‘Question: It is said that the Blessed One taught the four heavenly kings the Four Truths [in] the **Ārya* language. Was he capable [of preaching] or not? If he was capable, why did he teach two [of them] in the **Ārya* language, one in the language of the state/region of Tanluo 曇羅, [and the other] one in the language of the state/region of Mleccha? If he was not that capable, how to understand the following *gāthā* [about our] ultimate teacher?’

一音聲說法 ‘[The Buddha] teaches the Law in one utterance (lit. ‘voice [and] sound’)
悉遍成音義 [which automatically trans-]forms fully [into] all voices [and] meanings.
彼各作是念 Each of them would have this thought:
最勝為我說 “[He] preaches the most excellent [thing] for me.”’

「一音聲說法」者——是梵音也。
‘“[The Buddha] preaches the Law in one utterance”—It is the voice of Brahma(n).’[57]

「悉遍音」者——若有真旦人, 彼作是念, 謂:「佛作真旦語說法。」如是陀勒摩勒波勒佉沙婆佉梨。謂: 彼處若有**兜佉勒人**, 彼作是念, 謂:「佛作**兜佉勒語**說法。」
‘“Fully [into] all voices”—If there is a man of Cīnasthāna [in the audience], he would think: “the Buddha is preaching in the language of Cīnasthāna.” Likewise, *tuó-lè-mó-lè-bō-lè-qū-shā-pó-qū-lí*. That means: If there is **a man of *_Tukhāraka_** in that place,[58] he would think: “the Buddha is preaching in **the language of *_Tukhāraka_**.”’

「現義」者——著欲者作是念:「世尊說不淨。」恚者作是念:「世尊說慈。」……
‘“To convey (lit. ‘to present, reveal’) meanings”—One who has insatiable desires would think: “The Buddha is explaining uncleanness”; one who is of bad temper would think: “The Buddha is explaining mercy”; […]’ (corresponding to Argument I in T.1545, see infra, Text 15)

(*skipped: counterarguments from other masters*)

若通此偈者, 當何意?
‘If this *gāthā* is logical (lit. ‘if [we] pass through this *gāthā*’), what does it mean exactly?’

答曰: 世尊所說, 應機捷速。世尊語極速。為一說已, 復為一說, 如似一時。
‘Answer: What the Blessed One says is very fast depending on the situation. His speaking [speed can be] extremely quick. After making one explanation, he makes another explanation, [so] it seems that [he makes them] at the same moment.’ (corresponding to Argument III in T.1545)

---

[57] *Fanyin* 梵音, i.e. Sanskrit, cf. Hartmann 2024: 192, 201.

[58] Theoretically “that place” (*bichu* 彼處) may refer either to the assembly or to *Póqūlí*. Although the second interpretation suggests the intriguing scenario that a part of inhabitants in Bactria were **Tukhāraka* speakers, the first interpretation is more reasonable based on the content of T.1546 and 1545 (see *infra*). Moreover, *Dōuqūlè* 兜佉勒 was possibly missing at the end of the obscure *tuó-lè-mó-lè-bō-lè-qū-shā-pó-qū-lí* in this paragraph, because it appears in the next occurrence.

或曰: 世尊語音一切音, 各有境界, 應適一切音。世尊極知真旦語, 勝生真旦中者。如是陀勒摩勒波勒佉沙婆佉梨**兜佉勒**。**世尊極知兜佉勒語**, **勝生兜佉勒中者**。以是故, 說「一音聲說法　悉遍成音義」。(cf. T.1547[XXVIII]482c12–483a13)

'Another answer: The Blessed One's speech pronounces all kinds of voices; each of them has a range and limit [of pronunciation, but his speech] conforms with all the voices. The Blessed One knows the language of Cīnasthāna extremely well, better than [any]one born in Cīnasthāna. Likewise, *tuó-lè-mó-lè-bō-lè-qū-shā-pó-qū-lí-**dōu-qū-lè***. **The Blessed One knows the language of *Tukhāraka extremely well, better than [any]one born in *Tukhāraka.** For this reason, it is said that "[The Buddha] preaches in one utterance, [which automatically trans-]forms fully [into] all voices [and] meanings."'(corresponding to Argument II in T.1545) (tr. mine)

Obviously, the passage translated by Lévi (1897) is part of the last paragraph of this debate. Before it, and before the first occurrence of the enigmatic **Tukhāraka* language, the difficult passage *tuó-lè-mó-lè-bō-lè-qū-shā-pó-qū-lí* 陀勒摩勒波勒佉沙婆佉梨 appears.[59] It had become abstruse to later Buddhists as reflected in the aforementioned 6th-century dictionary:

(8) *Fan Fanyu* 翻梵語, *j.* 8

直 (var. *zhēn* 真)旦: 應云脂那, 亦云震旦, 譯曰漢地。

摩勒 (EMC *ma-lək*): 譯曰花也。

波勒 (EMC *pa-lək*): 譯曰伎也。

佉沙 (EMC *kʰɨa-ʂai*/*ʂɛː*): 譯曰不正語也。

婆佉梨 (EMC *ba-kʰɨa-li*): 應云**薄佉利** (EMC *bak-kʰɨa-liʰ*), 譯曰諂曲。(cf. T.2130 [LIV]1036b7–11)

'*Zhídàn* 直旦/*Zhēndàn* 真旦: One should say "*Zhīnà* 脂那". Also said "*Zhèndàn* 震旦". To be translated as "Land of Han 漢".

Mólè 摩勒: To be translated as "flower".

Bōlè 波勒: To be translated as "musicians/dancers".

Qūshā 佉沙: To be translated as "incorrect/improper speech".

***Póqūlí* 婆佉梨**: One should say ***Bóqūlì* 薄佉利**, to be translated as "flattery".' (tr. mine)

Probably influenced by this dictionary, Lévi (1905: 287) segmented the difficult passage as *Tuole, Mole*, *Bole, Qusha, Poquli* without fully treating these terms. However, except the one for Skt. *Cīnasthāna-*, all explanations in the above entries are not satisfactory. As far as I can see, the second and third one are based on Skt. *mālakā-* 'garland' (MW:

[59] I once wondered if it could be a transliteration of a non-Aryan (Bactrian?) phrase; Prof. Nicholas Sims-Williams kindly helped me to rule out this possibility and wisely indicated the potential link between *Po-qu-li* and Balkh (p.c. 19 February 2019, at Ryukoku University in Kyoto).

813c) or Gandh. *malaga-* ‘*idem.*’ (https://gandhari.org/dictionary/malaga, access on 4 Oct. 2024) and Skt. *bharata-* ‘actor, dancer, etc.’ (MW: 747b), respectively. *Qūshā* 佉沙 is explained by Lévi (1905: 260–261) as the typical transcription of BSkt. *Khaṣa-/Khaśa-* (cf. BHSD: 204b), which was by Medieval Chinese Buddhists usually seen as a synonym of the kingdom of Shule 疏勒 (Kashgar). Therefore, he usually rendered it simply as “*Kachgar*” for the attestations in other texts (Lévi 1905: 270; 1915: 52), but for the occurrences in this dictionary, he kept the vague meaning of the original Sanskrit word, rendering it as “*les tribus montagnardes des régions himâlayennes*” (Lévi 1915: 102, cf. Skt. *Khasa-* ‘name of a people and of its country (in the north of India)’, MW: 338c). In my opinion, the compilers possibly confused this Sanskrit demonym or toponym with *kaśmala-* ‘foul, dirty, impure’ (MW: 265b), no matter if any informant ever claimed that the language(s) spoken around Kashgar – aboriginal ones and/or local variants of Prakrit – were inappropriate or not. As to *Póqūlí*, the dictionary wrongly takes it to stand for BSkt./Pāli *sakhila-* ‘kindly in speech’ (BHSD: 544b; PTSD: 661), no doubt due to a common scribal confusion between *pó* 婆 and *suō* 娑 (EMC/LMC *sa*) in Chinese.[60]

Despite the poor quality of this dictionary, the word *Bóqūlì* 薄佉利 given as the equivalent of *Póqūlí* is intriguing. A search in the CBETA database shows that it is a hapax that occurs only once in the Chinese Sarvāstivāda Vinaya:

(9) T.1435 *Shisong lü* 十誦律, *j.* 53

優波離問佛:「若比丘作梵志形服於道行, 得何罪?」

答:「得偷蘭遮。」

「若作秦形服、大秦、安息、**薄佉利**、波羅大形服, 得何罪?」

答:「得突吉羅。如是等亦得突吉羅。」(cf. T.1435[XXIII]391a29–b3)

‘Upāli asked the Buddha: “If a monk wears (or: makes) a cloth of the style (lit. ‘form’) of Brahmins and walks on the road, which sin shall he commit?”

[The Buddha] answered: “A serious offense (*sthūlātyaya*).”

[Then Upāli asked]: “If [a monk] wears a cloth of the style of Qin 秦 (China), [of] Daqin 大秦 (Roman Orient), [of] Anxi 安息 (Parthia/Sassanid Iran), [of] ***Bóqūlì* 薄佉利**, [or even] a cloth in the form of *Bōluódà* 波羅大(?), which sin shall he commit?”

[The Buddha] answered: “A light offense (*duṣkṛta*). [To wear any clothes] like these shall all commit a light offense.”’ (tr. mine)

Assuming that *dà* 大 ‘big, large’ in *Bōluódà* 波羅大 (EMC *pa-la-da*’/*daj*^{h}) is part of the transcription rather than an indication of the size, it may be identified with Skt.

[60] Incidentally, in another chapter of the dictionary (*j.* 6), the same explanation with ‘flattery’ (*yi yue chanqu* 譯曰諂曲) is given to *shēzhìluó* 奢致羅 (EMC *ɕia-tri*^{h}*-la*; cf. T.2130[LIV]1027b10). It probably presents a Prakrit or Central Asian form related to Skt. *sakhila-*; see fn. 1 on EMC *tr-* ~ Sogd. *č.*

Pārada-/*Pārata-*,[61] a foreign people to the west of India, since Lévi (1915: 103–104) has noted an attestation of the Sanskrit word side by side with the Indo-Greeks, the Sakas, and the Bactrians (Skt. *Bāhlīka-*) in Chapter 4 of the *Rāmāyāṇa*. Accordingly, *Bóqūlì* in this passage must refer to a kind of Bactrian fashion that is visually conspicuous and different from other exotic clothing. All those distant countries in the Buddha's final answer were imaginably important destinations of Sarvāstivāda missionaries, but not all of these missionaries were successful.

Phonologically, *Póqūlí* 婆佉梨 (EMC *ba-k^{h}ia-li*) and *Bóqūlì* 薄佉利 (EMC *bak-k^{h}ia-li^h*) may transcribe a Sanskritised form of Bactr. *Βαχλιο 'Bactrian; of Bactria' (see §3 on Βαμοιανιο '(inhabitant) of Bamiyan'), or a Prakritic equivalent that no doubt contained *kh* in the word, cf. *Qūlú* 佉盧, the conventional term for Kharoṣṭhī. In the exegesis in T.1547, *Póqūlí* is used to indicate a country as well as its native language.

6.2. The exegesis on the Buddha's language proficiency in T.1546

In T.1546 *Apitan piposha lun* 阿毘曇毘婆沙論 **Abhidharma-vibhāṣā(-śāstra)*, the official translation of the Northern Liang 涼 dynasty (401–439), the hymn is so given:

(10) *Apitan piposha lun*, *j.* 41

佛以一音演說法　'The Buddha teaches the Law in one utterance,
而現種種若干義　and then numerous meanings are conveyed.
眾生皆謂獨為我　All the beings say [that: "This is] uniquely for me;
解說諸法不為他　various doctrines are expounded not for anyone else."'
(cf. T.1546[XXVIII]306c24–25; tr. mine)

Here the second *pāda* (quarter of a stanza) is translated without any mention of sound or voice (*contra* the earlier translation in T.1547, §6.1). Such an interpretation cleverly evades the question whether the Buddha was able to utter multiple sounds instantly. Affirming that the "one utterance" is in Sanskrit (*Fanyin* 梵音), the exegetical part explains the *pāda* in the following way:

(11) *Apitan piposha lun*, *j.* 41

「現種種義」者——若會中有真丹人者, 謂:「佛以真丹語為我說法。」如有釋迦人、夜摩那人、陀羅陀人、摩羅娑(for 婆)人[62]、佉沙人、**兜佉羅**人, 如是等人在會中者, 彼

[61] Cf. MW: 620a (*pārata, °taka*) and 619c (*pārada*).

[62] Here I adopt *pó* 婆 in two early woodblock prints, Pilu and Sixi, rather than the one given in the Taishō edition. Theoretically it may transcribe Mālava, an ancient country in India as so suggested by Lévi (1905: 287, cf. Sircar 1971: 205–207), but I think Skt. *mārava-* 'relating to a wilderness, being in a wilderness, etc.' (MW: 811c) could be a better candidate, see §6.4.

各各作是念:「佛以我等語, 獨為我說法。」若貪欲多者,「佛為我說不淨。」…… (cf. T.1546[XXVIII]306c26–307a2)

'"Various meanings are conveyed"—If there is a man of Cīnasthāna (*Zhēndān* 真丹) in the assembly, [he would] say: "The Buddha is teaching in the language of Cīnasthāna for me". If there is a man of *Shìjiā* 釋迦 (EMC *ɕiajk-kia*), a man of *Yèmónà* 夜摩那 (EMC *jia^{h}-ma-na^{h}*, Skt. *Yavana*-), a man of *Tuóluótuó* 陀羅陀 (EMC *da-la-da*, Skt. *Darada*-),[63] a man of *Móluópó* 摩羅婆 (EMC *ma-la-ba*), a man of *Qūshā* 佉沙 (Kashgar?), [or] **a man of *Dōuqūluó* 兜佉羅** (EMC *təw-k^{h}ia-la*, Skt. *Tukhāra*-) – if such ones are in the assembly, then each of them would think: "The Buddha is teaching the Law especially for me in the language of my [fellow country]men". If there is one with insatiable desires, [he would think:] "The Buddha is explaining the uncleanness [of one's thoughts] for me"; […]' (tr. mine) [64]

Although the exegesis on this topic is simpler than the other two versions, the occurrence of *Qūshā* as a country helps one to solve *tuó-lè-mó-lè-bō-lè-qū-shā-pó-qū-lí-**dōu-qū-lè*** in T.1547, because it implies that this commentary is also meant to explain what a Mleccha language could be. Now, *Qūshā, Póqūlí*, and *Dōuqūlè* are different countries, and the difficult passage can be further segmented. *Tuólè* 陀勒 (EMC *da-lək*) stands for Skt. *darada*-, possibly with a character missing, or with a poor transcription due to imprecise oral transmission; *mó-lè-bō-lè* 摩勒波勒 is either a copyist's mistake with a redundant *lè* 勒 (i.e. to read *Mólèbō* 摩勒波 = *Móluópó* 摩羅婆 in T. 1546), or juxtaposition of two names, *Mólè* 摩勒 (EMC *ma-lək*) and *Bōlè* 波勒 (EMC *pa-lək*). Concerning the latter solution, one may even conjecture an overcorrection of **mólèpó bōlè* 摩勒婆波勒 'the Desert people [and] the Pahlavas(?)' by later copists. Interestingly, Bactria does not appear in this version, unlike the exegeses in T.1547 and T.1545 (Table 4).

6.3. Translation forum of T.1546

The text was translated in 437–439, coincidentally during the diplomatic mission of Dong Wan 董琬 to the west of Northern Wei and before its conquest of Northern Liang.[65] Compared with the one organized by Dao'an in Chang'an, the translation forum of T.1546 was different in several ways: (1) It was based on a manuscript collected by the Chinese pilgrim Daotai 道泰 somewhere to the west of the Pamirs; (2) The translation was largely done by himself, with the help of a *Vibhāṣā* master from the western

[63] I.e. a people 'living above Peṣāwar' (MW: 470a) as well as the region inhabited by them, i.e. Dardistan (Lévi 1905: 287; 1915: 101; see also fn. 10 above).

[64] Corresponding to Argument I in T.1545, see *infra*, Text 15.

[65] On Dong Wan's travelling as far as Ferghana and neighbouring countries, see Ching & Galambos 2020: 21–24. On Sogdian inhabitants in Guzang by 439, see Ching & Galambos 2020: 31–32.

countries (*Xiyu* 西域), named Buddhavarman (Futuobamo 浮陀跋摩);[66] (3) Their translation quality was controlled by a great number of intellectuals at Guzang 姑臧, a city famous for its foreign populations on the Silk Roads (Sogd. *Kc'n*, Ancient Letter 2). The process is stated in the preface by 5th-century Daoyan 道挺 (?–?):[67]

(12) Preface to *Apitan piposha lun* (CSZJJ, *j.* 10)

有沙門道泰, 才敏自天, 冲氣疎朗, 博關奇趣, 遠參異言。往以漢土方等既備, 幽宗粗暢。其所未練, 唯三藏九部。故杖策冒嶮, 爰至葱西。**綜攬梵文**, 義承高旨, **并獲其胡本**十萬餘偈。既達涼境, 王即欲令宣譯。然懼環中之固, 將或未盡, 所以側席虛衿, 企矚明勝。

'There was a *śramaṇa* [named] Daotai, [who was] born brilliant and had an open mind; he had a wide range of interests and learnt different language(s) [of] remote [lands]. Previously, given the situation that [the collection of] *Vaitulya/Vaipulya* texts[68] had become rather complete in China [and that] the Buddha's teaching had been basically understood, what he could not fully study was the nine branches [of the scriptures of the *Śrāvaka*s in] the *Tripiṭaka*. Therefore he took a walking stick and climbed dangerous mountains until he reached the west of the Onion [Range] (Pamirs)**. He widely collected Brāhmī/Sanskrit texts** (*fanwen* 梵文) and learnt (lit. 'took') the meanings of the excellent knowledge, **and he obtained a *hu* version** (*huben* 胡本)[69] **of this [text** whose length is] more than 100,000 *śloka*s. After he arrived in the territory of Liang, the king immediately demanded of him [a Chinese] translation [to be done] openly/orally. However, being afraid that the hardness [of the truth] of nihility (lit. 'solidness of [what is] inside a ring') could not be [expounded] exhaustively, he (Daotai) humbly took a seat aside and waited for a superior master.'

時有天竺沙門浮陀跋摩, 周流敷化, 會至涼境。……遂以乙(for 丁)丑之歲四月中旬, 於涼城內苑閑豫宮寺, 請令傳譯理味; 沙門智嵩道朗等三百餘人考文詳(var. 評)義, **務存本旨**, **除煩即實**, **質而不野**。王親屢迴御駕, 陶其幽趣, **使文當理詣**, **斥**(var. 片)**言有寄。**至丁(for 己)卯歲七月上旬都訖。(cf. T.2145[LV]74a13–26)

'At that time, an Indian *śramaṇa* named Buddhavarman was travelling and teaching everywhere. Incidentally he arrived in the land of Liang. […] Therefore, in the middle of the 4th month of the *dingchou* year (*ca.* May 437), [the king] courteously ordered [them] to transmit, translate, and study [the text] at the temple/monastery [renovated from] the

[66] See the biography of Buddhavarman in GSZ, *j.* 3 (T.2059[L]339a14–28) and the bibliographical entry in CSZJJ, *j.* 2 (T.2145[LV]11c2).

[67] See also the version in GSZ, *j.* 1 (T.1546[XXVIII]1a20–b4). On the copyist's errors about the year formula, see Yoshikawa & Funayama 2009: 260–261.

[68] On *Fangdeng* 方等 (*vaitulya/vaipulya*) as a generic term for Mahāyāna canons, cf. Karashima 2015.

[69] Here *hu* 胡 'foreign, non-Chinese' of the Taishō edition is correct, since the character is verified by the first Korean carving of this chapter and it is further preserved in several woodblock prints of Daoyan's preface that reappears at the very end of T.1546 (*j.* 60, 415a1).

> Comfort Palace in the inner royal park of Liang's [capital] city. The *śramaṇa*s named Zhisong 智嵩, Daolang 道朗, etc., amounted to over 300, examined the text and elucidated (*xiang* 詳; var. *ping* 評 'criticised') its meaning **in order to keep [its] essential truth** (*benzhi* 本旨, or: the main points of the original version)**; they trimmed away the pleonasms and made it conform to reality, [so that it became] plain but not crude.** The king visited [the forum] several times in person and learnt its enjoyable purport; [he] **ordered that the wording should be appropriate, the argumentation comprehensible** (*yi* 詣, lit. 'going, visiting', that is to say 'attainable'), **and [every] argument** (?) (*chi yan* 斥言; var. *pian yan* 片言, lit. 'piece of passage') **well founded.** It was finished in the early 7th month of the *jimao* year (*ca.* August 439).' (tr. mine)

In brief, Daotai needed the explanation of a master before translating, although he had a basic understanding of Sanskrit and Brāhmī. His helper Buddhavarman is obviously good in the text as well as in Sanskrit, no matter whether it was a [new] "lingua franca" (Strauch 2012: 133) or a "church" language (Allon 2021: 112) in the Buddhist world at that time. Of course, Buddhavarman may have spoken one or more languages that facilitated his communication with Daotai, but it is hard to imagine that he was versed in Chinese.

The preface shows the team's preference for a serious and literal translation rather than a free and ornate one, even though some repetitive formulae or elements were cut out of the Indic version. Specifying the name of the king as Maoqian 茂虔, i.e. JUQU Mujian 沮渠牧犍 (r. 433–439), Buddhavarman's biography further states that it was Daotai who noted down the translation himself, and the text was scrutinised in "two rounds" (*zai zhou mang qi* 再周方訖) by "more than 300 monks skilled in exegesis" (*yixueseng sanbai yu ren* 義學僧三百餘人, cf. T.2059[L]339a14–16). In this sense, it is reasonable that Text 11 kept archaic terms such as Skt. *yavana-* 'an Ionian, Greek' (MW: 848a).

Nevertheless, the king's command prompts one to consider whether the result was further modified to conform with the actuality or folk knowledge in the 5th century. This possibility is implied by Daotai's usage of *Shìjiā* 釋迦, traditionally a transcription for the Śākyas, the clan of the Buddha. It makes us recall Buddhist legends in 7th-century Uḍḍiyāna and *Himatala (between Takhar and Badakhshan, Afghanistan) collected by Xuanzang stating that their royal houses were Śākya by descent.[70] On the other hand, since Zhisong and Daolang had previously been assistants of Dharmakṣema for his translation projects sponsored by JUQU Mujian,[71] it is especially noteworthy that Balkh

[70] T.2087[LI]838b27–884a14, 887a17–b7; 940b14–25. On his description of *Himatala, see an English translation in Ching & Galambos 2020: 100–102.

[71] GSZ, *j.* 2, T.2095[L]336a24–24, cf. Yoshikawa & Funayama 2009: 260.

or Bactria is not found in Daotai's translation, despite the fact that a new transcription *Póqūluó* 婆佉羅 was just created or promoted by Dharmakṣema's team around the 410s (see §1, fn. 8 and §2). Could it have been removed during scrutinisation of the treatise? If yes, would this have been done to reflect the unification of Bactria and Tukharistan as an immediate consequence of the Kidarites' annexation of Bactria?

According to de la Vaissière (2005), Hunnic tribes including the Hephthalites gradually started their migration from the Altai from the early 350s onwards. As for the Kidarites, this group of obscure origin was not yet very strong in the 4th century, but they seem to have conquered a vast territory as far as Balkh in a short period just before 437 (cf. Grenet 2002: 220; Bakker 2020: xv). Although it would sound appealing to link the absence of Bactria in T.1546 with the Kidarite expansion, this would be too prompt chronologically. In my opinion, one cannot eliminate another possibility, namely that Daotai faithfully translated a Brāhmī manuscript in which Balkh or Bactria was simply not found. Assuming that there is no scribal error or missing content there, it could present a local tradition that goes back to a time when the whole region was politically one realm, perhaps during the Kushan period. In fact, this version's uniquity is clear from a preceding paragraph which has no parallel in T.1545 nor in T.1547:

(13) *Apitan piposha lun*, *j.* 41

問曰: ……行者何故先見苦諦, 後乃至見道諦?

答曰: 以麁細故。……如索迦人、[72]蛇摩那人、**兜佉羅**人學射之時, 先射簸箕、草人、濕泥團等, 箭矢無不著後, 漸更學, 乃至 [73]射一毛。(cf. T.1546[XXVII]303a19–24)

'Question: […] Why shall a practitioner first observe the truth of "suffering" and afterward, and so on and so forth, until the [4th] truth of 'path'?

Answer: Because [they are from] coarse [to] fine. […] For example, when a man of *Suǒjiā* 索迦, a man of *Shémónà* 蛇摩那, [or] **a man of *Dōuqūluó* 兜佉羅** learns archery, he [shall] first shoot at dust-pans, straw-men, mud balls, etc. After he [is able to] shoot [such things] without failure, he learns gradually, and so on and so forth, until he [can] shoot at one single hair.' (tr. mine)

Here, the *Dōuqūluó* 兜佉羅 or "Tukhāra" seem to be a people as warlike as *Suǒjiā* 索迦 (EMC *sak-kia*) and *Shémónà* 蛇摩那 (EMC *zia-ma-na*^{h}). Therefore I think the latter two are uncommon transcriptions for the Sakas and another hostile group (cf. Skt. *śāmana-* 'destroying'), but it is unclear whether *Shémónà* is to be located in Chitral, a place that has been often identified – by geographical reason but without satisfactory etymology –

[72] Here *ren* 人 may also be translated as 'people' as an ethnic group.

[73] The typical expression *nai zhi* 乃至 to elide a similar passage, a repeated expression, or a series of items, is translated in this article as 'and so on and so forth, until …'. Generally it is known to be an equivalent of Pāli *pe* (Anālayo & Bucknell 2006: 216).

to *Shēmí* 賒彌 (EMC *ɕia-mjiə̆*/*mji*), a mountainous kingdom to the north of Uḍḍiyāna in the *Weishu*, to *Shāngmí* 商彌 (LMC *ʂiaŋ-mji*), a kingdom in the southwest of the Pamirs in Xuanzang's travelogue, and to *Shēmó* 奢摩 (LMC *ʂia-mua*) in Hyecho's record.[74]

6.4. The debate on the Buddha's language proficiency in T.1545

T.1545 *Apidamo da piposha lun* 阿毘達磨大毘婆沙論 is translated from a valuable copy specially prepared for Xuanzang under the command of the king of Kashmir (cf. T.2052 [L]215b28–c2). Xuanzang then translated it into 200 *juan*s as the main translator, and the result is far more copious than T.1547 (14 *juan*s) and T.1546 (100 *juan*s, only 60 are preserved). Here is the hymn in question:

(14) *Apidamo da piposha lun*, *j.* 79

佛以一音演說法　'The Buddha teaches the Law in one utterance,
眾生隨類各得解　[yet] the beings understand it according to their stocks/categories;
皆謂世尊同其語　All say that the Bhagavat's language is the same as theirs,
獨為我說種種義　"[he] explains various meanings specially for me."'
(T.1545[XXVII]410a16–17; tr. mine)

As for the exegesis, I number the respective arguments to demonstrate its structural difference from the other two versions:

(15) *Apidamo da piposha lun*, *j.* 79

(Argument I) 若至那人來在會坐, 謂:「佛為說至那音義。」如是磔迦、葉筏那、達剌陀、末鞞(var. 髀)婆、佉沙、**覩貨羅**、**博喝羅**等人來在會坐, 各各謂:「佛獨為我說自國音義。」聞已隨類, 各得領解。(*skipped: thought of the greed ones, etc.*)
'If someone of Cīna[sthāna] comes and sits down in the assembly, he would say: "The Buddha is teaching the meanings in the language of Cīna[sthāna] for [me]". Likewise, [if] someones of *Lìjiā* 磔迦, *Yèfánā* 葉筏那 (Skt. *Yavana*-), *Dálàtuó* 達剌陀 (LMC *dat-lat-tha*, Skt. *Darada*-), *Mòxiápó* 末鞞(?)婆, *Qūshā* 佉沙, ***Dǔhuòluó* 覩貨羅**, [and] ***Bóhēluó* 博喝羅**, etc., come and sit down in the assembly, each of them would say: "The Buddha is teaching the meanings only for me in the language of my home country." Having heard [his speech], each of them understands [its meaning] according to their stock/category.'

(Argument II) 復次, 如來言音, 遍諸聲境, 隨所欲語, 皆能作之。謂: 佛若作至那國語, 勝在至那中華生者。乃至若作**博喝羅語**,[75] 勝在**彼國中都**生者。以佛言音遍諸聲境故, 彼伽他作如是說。

[74] On Chitral in Chinese sources, see Kuwayama 1989: 110; Kuwayama 1992: 128; Falk 2018: 27.
[75] Corrupted as *Boluo-yu* 博羅語 in an extract of this exegesis in the *Jie shenmi jing shu* 解深密經疏, *j.* 1, written by Woncheuk (Chin. Yuance 圓測, 612/613–696), a colleague of Xuanzang from the Shilla kingdom in present-day Korea, see CBETA, X21, No. 369, 177a2.

‘Moreover, the voice of the speech of the Tathāgata covers [a great] range of various sounds. He is able to articulate (lit. ‘make’) whatever he wants to speak. That is to say: If the Buddha makes [a speech in] the language of Cīnasthāna, [then his speech will be] better than [any]one born in the central land of Cīnasthāna (*Zhina zhonghua* 至那中華); and so on and so forth, until [the simile that] if he makes [a speech in] **the language of *Bóhēluó***, [then his speech will be] better than [any]one born in **the central capital city of that country** (*bi guo zhongdu* 彼國中都). Because the voice of the Buddha’s speech covers [a great] range of various sounds, the hymn is to be explained as such.’

(Argument III) 復次, 佛語輕利, 速疾迴轉, 雖種種語, 而謂一時。謂: 佛若作至那語已, 無間復作磔迦國語, 乃至復作**博喝羅語**。以速轉故, 皆謂一時。如旋火輪, 非輪, 輪想。前頌依此, 故亦無違。(cf. T.1545[XXVII]410a18–c8)

‘Moreover, the Buddha’s speech is quick and fluent; [it is] so rapid and whirling. Although he speaks [in] various [languages, it can be] said that [it occurs] at once. That is to say: If the Buddha just made [a speech] in the language of Cīna[sthāna], then without any stop he would make it again in the language of the state of *Lìjiā*; and so on and so forth, until he would make it again in **the language of *Bóhēluó***. Because it whirls so quickly, all [his audience] will say that it occurs at once. It is like a fire-wheel (i.e. fire performance) that [actually] is not a wheel, but people would think that it is a wheel. For this [reason], the above hymn is not contradictory.’ (tr. mine)

In fact, Arguments II and III can be found in T.1546 (p. 307a26–b1), but they are short there, without providing long series of foreign languages (see Table 4).

In Xuanzang’s translation, *Lìjiā* 磔迦 (LMC *liajk-kia*) is possibly a scribal error for *Zhéjiā* 磔迦 ‘Ṭakka’ (LMC *traːjk-kia*) in Xuanzang’s travelogue, where it is reported to be a kingdom on the left bank of the Indus formerly ruled by the legendary Mihirakula. The same error is found twice in Xuanzang’s biography (*j.* 2, see T.2053[L]231c8–9, 232a6–7) in the second Korean carving and can be corrected on the basis of two earlier handwritten copies in Japan.[76] Otherwise, *Lìjiā* can be an error for *Shuòjiā* 鑠/爍迦 (EMC *ɕiak-kɨa*, LMC *ʂiak-kia*), the Saka-/Śakasthāna that is protected by a yakṣa named Śaṃkara in the *Mahāmāyūrī vidyārājñī-sūtra*.[77] Interestingly, the old capital city of *Zhéjiā*, *Shējiéluó* 奢羯羅 (EMC *ɕia-kɨat-la* for Skt. *Śākala-*, *Sāṅkala-*, etc.), so transcribed by Xuanzang in his travelogue (*j.* 4) sounds similar to the yakṣa’s name as well as to Skt. *śakara-/śakāra-* ‘a descendant of the Śakas, etc.’ (MW: 1043c). Although the Saka invasion of India is not my concern here, Xuanzang obviously chose a

[76] Utsunomiya 1979: 37–38. On the legend about Mihirakula collected by Xuanzang, see the translation by Deeg, Galambos, and Ching in Ching & Galambos 2020: 93–94. Lévi proposes to etymologise Ṭakka as “*pays des Tékins*” (1905: 300), linking it to Chin. *Chìqín* 敕勤 (EMC *tr^{h}ik-gɨn*).

[77] Lévi 1915: 55. See more text comparison in Ching 2023: 115.

transcription other than the counterpart *Shìjiā* 釋迦 in T.1546 in order to underline the foreign character of that language.

Yèfánā 葉筏那 (LMC *jiap-fĥjyat/fĥa:t-na´*) occurs twice in Xuanzang's works, in the above message and in a later chapter of the same treatise (*j.* 124, 648b18), where it is one of the sixteen leading states during Śākyamuni's lifetime and thus undoubtedly the Indo-Greek Kingdom. The country next to Dardistan is more problematic. It cannot be readily identified with Mālava in Central India, which appears in Xuanzang's travalogue as *Mólàpó* 摩臘婆 (LMC *mua-lap-pĥua*). Lévi's romanisation *Mo-ye-p'o* (pinyin *Moyepo*) for 末曄婆 seen in some woodblock prints, and his reconstruction of the consonant of 曄 as *d̩* or *l* in Medieval Chinese, are not supported by any classical dictionary.[78] Here I adopt *xiá* 鞈 in the Pilu print of 1148 CE (Library of Japanese Imperial Court), the earliest print preserved,[79] of which the LMC can be reconstructed as **xĥja:p/xja:p*. The Sanskrit word behind it, *Mòxiápó* 末鞈婆 (LMC *mua-*xĥja:p/xja:p-pĥua*), is quite certainly *mārava-*, which I have assumed to be the source of Daotai's *Móluópó* 摩羅婆 (see fn. 63), because EMC *ɣ* / LMC *xĥ* has been used to stress the guttural feature of Skt. *r* in *rājya-* 'kingdom' and other words (Funayama 2013: 188–189). The nominative plural *marucīnāḥ* in *Bṛhatsaṃhitā* 16.37, translated as 'desert [and] Chinese' or 'the Desert Cīnas' by Bakker *et al.* (2020: 277, 289), gives a hint to explore the meaning of *mārava-* (from *maru-* 'desert, etc.'). However, I would like to defer the question whether it includes the oases around the Tarim Basin or not.

In general, the language series in Xuanzang's translation corresponds with that in Daotai's except for *Bóhēluó* 博喝羅 (EMC/LMC *pak-xat-la*), another hapax in Chinese Buddhist texts, and it is this language which replaces the enigmatic **Tukhāraka* language in T.1545 for the scholastic debate. In order to better understand the text variations of the three translations (Table 4), the dating of their source texts is to be taken into consideration.

6.5. Dating the source texts of T.1545–1547

According to Xuanzang's postscript (T.1545[XXVII]410a5–8), the original *Vibhāṣā* was composed in Kashmir during the assembly of 500 sages summoned by Kanishka, 400 years after Śākyamuni's nirvāṇa, and it was carefully passed on there. The source of Daotai's copy is obscure, while nothing is known about the Arhat *Sita- or *Śītapāṇi, the author (?) of the version recited by Saṅghabhūti.

---

[78] Lévi 1905: 286–287. Further comments in Ching 2023: 114 n. 87.

[79] See the digitised image at https://db2.sido.keio.ac.jp/kanseki/bib_frame?id=007075-4004&page=2 (access on 14 Oct. 2024). The chapter is not in the surviving part of the 1st Korean carving, but such a confusion obviously resulted from the special adjectival and poetic expression *xiádié* 鞈曄 to describe a flower's petals that are growing thickly and fluttering in the air.

Fully accepting the Kashmirian Sarvāstivāda orthodoxy so much praised in Xuanzang's postscript of T.1545, Hirakawa (1974: 186) suggested that the main part of the original Sanskrit *Vibhāṣā* was primarily composed in the 2nd century and finished in the 3rd century CE, and the Sanskrit copy given to Xuanzang is the most traditional one. If so, it is expected to retain some conservative features, and Lévi's (1905: 287) suggestion that *Bóhēluó* was "*peut-être les Pahlavas*" is not unreasonable. At first glance this conforms with the indication of the "central capital city" that implies the vastness of the realm, not to mention that the Śakas, the Yavanas, and the Pahlavas are often found together in Sanskrit epics (for example, see Lévi 1904: 572) or even in the Khotanese *Book of Zambasta* (24.393, cf. Maggi 2020: 115).

Nonetheless, Xuanzang and his team are renowned for their erudition and scrupulosity. If the Brāhmī/Sanskrit copy at his disposal wrote *Pahlava-* 'the Parthians or Persians' (MW: 612c), it is strange why he did not use the common names for Parthia (i.e. Anxi 安息) or Persia (*Bōsī* 波斯 in the *Weishu* and later annals, more carefully *Bōlàsī* 波剌斯 [LMC *pua-lat-sɨ̯*] in his travelogue), nor the established terms in Buddhist texts for the Pahlavas, e.g. *Bōluópó* 鉢羅婆 (EMC *pat-la-ba*, LMC *puat-la-pɦua*; cf. Lévi 1897: 10 fn. 1). Moreover, if he intended to create a more accurate transcription of Skt. *Pahlava-*, then it is not easy to explain his neglect of the *va* over all three attestations of *Bóhēluó* in the exegesis, since he rendered Andarab in Afghanistan always as *Āndáluófú* 安呾羅縛 (LMC *ʔan-tat-la-ɦjyak*/*ɦak*) in his travelogue (*j.* 1 and 12, 5 attestations). Last but not least, one may ask why *Bóhēluó* is placed at the end of the list, rather than the typical position the Pahlavas side-by-side with the Śakas and the Yavanas.

Recently, Sasaki (2005; 2007; 2011, esp. 2007: 167–169) has argued that T.1546 and T.1547 are close to each other and reflect more archaic features, based on an in-depth analysis of their general structure and philosophical content. Radich (2010: 153) also places the recension acquired by Xuanzang later than 437, while Mitomo (2011) emphasizes the divergence of the three *Vibhāṣā* traditions that are nevertheless all from the Sarvāstivāda school. According to him, the uniquity of T.1547 is especially remarkable. Recently, Fujimoto (2022: 23–25) has argued for the significant difference between the three recensions with regard to their redactional structure. The growing consensus that Xuanzang's copy was late opens up the possibility that *Bóhēluó* is one of its relatively late features: it could reflect a sanskritised but vernacular form of Bactr. Βαχλο from Kashmir, for instance **Pahla-* or **Pāhala-*, which he could not recognize as any proper name. Conceivably, for the *Vibhāṣā* text circulating in that region, there could also have been a trend to confuse Parthia and Bactria, or to replace Parthia by Bactria. Both are terms for civilized Iranians, and such a trend is found in various

versions of the *Mahāmāyūrī-vidyārājñī-sūtra*.[80] Although unvoicing of /b/ is abnormal in Sanskrit and Gāndhārī,[81] ancient Kashmirian Buddhists may in their Sanskritisation process have been befuddled by various spellings for Balkh as well as "Kushan" (e.g. Gandh. *Kuṣana*, *Guṣaṇa*, *Khuṣaṇa*, etc.).[82] This confusion or re-interpretation might have further resulted from a shift of focus on the real language varieties *in* the Buddhist world beyond northwest India, rather than elucidating the speech of old and mystic ancient peoples. However, a satisfactory solution will come up only after a more accurate dating of the recensions.

7. The "language of Yuezhi" in T.1582 *Pusa shanjie jing*

Now we can re-examine Lévi's support for von Richthofen's equation "Yuezhi = *Τόχαροι*", a theory that is still popular today.[83] In brief, he believed that the enigmatic language in T.1547 can be equated with that of the Yuezhi 月支 in a "*passage analogue*" (Lévi 1897: 11) in T.1582 *Pusa shanjie jing* 菩薩善戒經 (translated by Guṇavarman [Qiunabamo 求那跋摩] in 431). Later, he changed his interpretation of "Yuezhi" and "Tukhāra" in Buddhist discourse by giving their primary meaning there as "Kushan" (Lévi 1905: 289, 291), and thus the **Tukhāraka* language would be that of the Kushans. While Müller was prudent about this implication,[84] Pelliot defended it as follows (my restoration in brackets):

> F. W. K. Müller (*Toχrï und Kuišan*, 576–577) estima que M[onsier] S. L[évi] supposait gratuitement l'équivalence du Tukhāra de l'une des listes au Yue-tche de l'autre, car il se serait agi, selon lui, de pays situés très loin les uns des autres et il n'y avait pas de raison de vouloir que les mêmes noms reparussent dans les deux textes. En fait, M. S. Lévi avait vu juste, car il reproduit dans le présent article ([Lévi 1933,] pp. 24–25) une liste qui figure dans une traduction de Kumārajīva (†413) […] (*skipped*: *see* §2); et une note de Kumārajīva, natif lui-même de Kučā, spécifie que les Tukhāra, ce sont les 小月氏 Siao-Yue-tche, les "Petits Yue-tche". Il n'y a donc pas de doute qu'au IV^e^ et au V^e^ siècle, les

[80] Cf. Lévi 1915: 52 (Verse No. 83), 56 (No. 96), 104. Further discussion in Ching 2023: 115–118.

[81] An exception is Niya-Prakrit *palýi* (Skt. *bali-* 'tax, etc.', Burrow 1937: 5), though the area is distant from my focus here.

[82] See '*Kuṣana*' at: https://gandhari.org/dictionary/ku%E1%B9%A3ana (access on 4 Oct. 2024).

[83] For instance, Sims-Williams 2002: 229; Benjamin 2007: ix; Grenet 2015: 205; Falk 2018: 15. However, see Yu 2015 and Kuwayama 2022a for different views.

[84] "Woraus Lévi schließt, daß Tukhāra des Vibhāṣāśāstra hier durch Yüe-tšï ersetzt wäre, ist nicht klar. Man hat hier doch bloß eine Aufzählung von allerlei weit auseinanderliegenden Ländern und Völkern vor sich, deren Sprachlaute zu den sogennanten "feinen Lauten" im Gegensatz zu den sogenannten "groben Lauten" (wie Donner u.a. Naturlaute usw.) als Beispiele citiert werden" (Müller 1918: 577).

> Chinois appellaient Yue-tche le peuple que l'Inde désignait sous le nom de Tukhāra. (Pelliot 1934: 35–36)

With regard to this argument, it is noteworthy that in 1934 Tukhāra = *Τόχαροι was* for Pelliot etymologically or phonologically equivalent with Daxia (Oxiana) and its major population conquered by the Yuezhi, rather than the invaders themselves (see esp. p. 40). This is because he had just been convinced by Haneda (1933), who supported Marquart's suggestion that Tukharistan = land of Tukhāra = Daxia 大夏, i.e. the kingless land conquered by the Yuezhi. Accordingly, Pelliot thought that the Kushans descended from the people of Daxia/ Tukhāras/Tokharoi (Pelliot 1934: 38–40). The rest of his article (esp. pp. 47–54) aims to identify the script and native language of Tukharistan based on Xuanzang's description, namely "*la langue des Kuṣanas*" (p. 65), which in turn he misidentified as Tocharian A (see especially Pelliot 1934: 53, 62).

Language diversity before the Common Era is beyond the scope of this article. However, Lévi's 1897 argument is indeed dubious. In my opinion, the Sarvāstivādin *Vibhāṣā*s should not be hastily analogised with the Mahāyānic T.1582 *Pusa shanjie jing*. Although Guṇavarman (367–431) claimed to be a prince from the royal house of Jibin, he studied Mahāyāna in Sri Lanka. It was from there that he sailed to South China and promoted his translation forum in Jiankang.[85] The text is of a specific nature: It compiles the precepts for Mahāyāna followers, including his most important sponsor, Emperor Wendi 文帝 of the LIU Song dynasty (r. 424–453), and praises the merits of the "Great Vehicle". The immediate context of the language varieties also deviates from that in the *Vibhāṣā*s: The unequal comprehension of the four heavenly kings is not stressed, but it says that a Bodhisattva (or: an advanced Mahāyānist) would acquire "divine aural comprehension" (*tian'er* 天耳) that enables him/her to understand all voices or sounds (*sheng* 聲) of deities and human beings, including sacred (or: **Ārya*) and unsacred ones, and loud and soft ones (*xisheng* 細聲). The last category is defined as follows:

(16) *Pusa shanjie jing*, *j.* 2

> 謂: 竊語聲、不了聲、陀毘羅國聲、**粟特聲、月支聲**、大秦聲、安息聲、真丹聲、佉沙聲、裸形聲、**鮮卑聲**。如是等邊地聲, 名為細聲。何以故? 嫉妒煩惱因緣得故。(cf. T.1582[XXX]972c21–24)
>
> 'That is: Whispering voices, endless subtle sounds, the language of the state of Tuopiluo 陀毘羅, **the language of Sogdia**, **the language of Yuezhi**, the language of Daqin, the language of Anxi, the language of Cīnasthāna (*Zhendansheng* 真丹聲), the language of *Qūshā* (Kashgar?), the language of naked ones, [and] **the language of Xianbei**; such languages in the frontier regions are called "lesser sounds". Why? Because they are learnt (lit. 'acquired') through the causes of jealousy and irritation.' (tr. mine)

[85] Present-day Nanjing, see Guṇavarman's biography in GSZ, *j.* 3, T.2059[L]340a15–342b10.

Guṇavarman preached to the Imperial house, noblesse, and gentry of Southern China, and the translating style of his team is rhetorically more elegant than those of Dao'an and Daotai in the north. Therefore, I would see the above passage as a quite free translation to attract more followers, and the mention of Sogdian (Sute 粟特, see §3) and Xianbei 鮮卑 possibly results from this strategy. If the "language of Yuezhi" (*Yuezhisheng* 月支聲) is also an insertion or substitution, it probably refers to Bactrian given the situation around 431, and "Yuezhi" in his usage might be Bactria-Tukharistan as well as the general population there in post-Kushan times.[86]

Very likely referring to Bactrian is the "Yuezhi 月支" in an earlier Mahāyānic list by Dharmarakṣa (active in 266–306, cf. fn. 26):

(17) *Da bao ji jing* 大寶積經, *j*. 10

一切諸人及與非人, 言語各異, 音聲不同, 辭有輕重。如來聖慧, 從其音響, 隨時而入, 皆悉化之, 立正真業。各有種號: 釋種、安息、月支、大秦、劍浮、擾動、丘慈、于闐、沙勒、禪善、焉耆[87]、前後諸國; 匈奴、鮮卑; 吳、蜀、秦地; 諸麼(for 蠻)[88]、夷、狄; 他羅多、愚民、野人…… (cf. T.310[XI]59a11–17)

'The languages of all the human and non-human beings are distinct from each other: Their voices are different, [and] there is lightness and heaviness (of tones or accents?) [in their] vocabulary. The noble wisdom of the Tathāgata enters [their minds] through these voices on timely occasions, converts all of them, [and makes them] establish the right and true vocation [in their countries, and] there are various ethnicities and appellations: The Śakas, Arsacid/Sassanid Iran (*Anxi* 安息), Yuezhi 月支 (as a country and/or a people), the Roman Orient, the Kambojas (*Jianfu* 劍浮), Raodong 擾動 (lit. 'to disturb': a text corruption?), [as well as] the various states of Qiuci 丘慈 (Kucha), Yutian 于闐 (Khotan), Shale 沙勒 (Kashgar), Shanshan 禪善, Yanqi 焉耆, the Anterior and Posterior (Jushi 車師, i.e. the semi-nomadic states around Turfan and Jimsar); Xiongnu [and] Xianbei; the regions of Wu 吳, Shu 蜀, [and] Qin 秦; various Man 蠻, Yi 夷, [and] Di 狄; the Daradas, the foolish people(s), the barbarian people(s) […]' (tr. mine)

This sinicised list invoked by Lévi (1905: 289–290) and de la Vaissière (2005: 13–15) is especially valuable. By inserting new elements into the series of ancient peoples and countries, Dharmarakṣa imbedded a Chinese perspective during the Western Jin 晉 period (266–316), and confirmed the language diversity in the five oasis kingdoms

[86] This is with an assumption that it was too early for the Buddhists in South China in 431 to be informed about the rising Kidarites, who later assumed the appellation of "Yuezhi" in their diplomatic affairs with Northern Wei after Dong Wan's mission in 437.

[87] Here *Yanqi* 焉耆 in a copy kept in Japanese imperial library is adopted.

[88] Here I take *me* 麼 that looks incomprehensible to previous researchers to be a scribal error for *man* 蠻, a typical term for noting uncivilized tribes in South China.

around the Tarim Basin that all obtained the title of *Shizhong* 侍中 (Niya-Pkt. *Jituṃg(h)a*) from the Jin court. One would ponder whether this means that one language was spoken in each of them (Early Khotanese, Early Kuchean, etc.) or that several local Prakrits (e.g. Niya-Prakrit, Kucha-Prakrit) were known to Dharmarakṣa (or to Jin Chinese) as their administrative, diplomatic, or commercial languages.

In any case, the translator no doubt wanted to extend the traditional idea of *Rong* 戎 'western barbarians' (in contrast to *Man*, *Yi*, *Di* for the southern, eastern, and northern barbarians, respectively) by providing a spectrum of foreign polities and languages known to Dunhuang citizens. It may explain the expressions of "various" or "miscellaneous" *Hu* from western countries (*Xiyu zhuhu* 西域諸胡, *Xiyu zahu* 西域雜胡) in the biography of CANG Ci 倉慈, Governor of Dunhuang in the early 3rd century (SGZ, *j.* 16, 512–513). Moreover, it is only in a post-Kushan period that Xuanying 玄應 (?–661) tentatively links Dharmarakṣa's "Yuezhi 月支" in Text 17 to *Bóqūluó* 薄佉羅, a state close to that of Badakhshan (*Boduochana* 波多叉拏), which Dharmagupta crossed in the late 6th century (see *supra* fn. 9; cf. Pelliot 1934: 36 fn. 4).

8. Conclusion

This article has discussed geolinguistic information in three Chinese *Vibhāṣā* texts T.1545–1547 in the historical context of their translation. One can envisage that various non-Aryan languages were mentioned during the controversy over the Buddha's language proficiency, but only important ones were fixed to the oral and written traditions of the treatise. For 7th-century Vaibhāṣika orthodox Kashmirians, the most representive ones are Chinese and the "*Bóhēluó* 博喝羅" language. If the latter more likely denotes Bactrian rather than a "Pahlava" language (Parthian or Middle Persian?) as I suggest based on my investigation above, then this fortuitously agrees with its current nomenclature. In this case, however, the autochthonous language of the historical Tukharistan *sensu stricto* remains unclear.

T.1547 *Biposha lun* (tr. 383), recited by Saṅghabhūti (from Jibin) and translated by Dharmanandin (from Tukharistan), ZHU Fonian and others, shows more clearly a contrast between the language of *Póqūlí* 婆佉梨 (a Sanskritised form of Bactr. *Βαχλιο?) and that of *Dōuqūlè* 兜佉勒 (**Tukhāraka* or °*ika*?). The contrast may be dialectal, since a dialect has been implied by a group of Bactrian letters apparently from south-eastern Afghanistan. Nevertheless, the existence of a distinct language cannot be ruled out.

Thus far, τοχοαρ(α)στανο as a geographical name is not yet used to define any language in the Bactrian corpus, while the endonym of this Middle Iranian language is obscure (p.c. Sims-Williams, *apud* Ching 2014: 487). In this regard, the fact that the Tukhāras are described as skilled in archery in T.1546 *Apitan piposha lun* (tr. 437–439),

and the absence of Balkh or Bactria in its exegesis are both noteworthy. The unsolved chronology of the *Vibhāṣā* recensions impedes us to make any final conclusion, but if one assumes that Daotai's translation was essentially literal without much revision, then the Indic manuscript acquired by him probably belonged to a non-Kashmirian sect that denied or neglected any political or linguistic difference between Bactria and Tukharistan. Perhaps that sect was formed when the two regions were unified in the heyday of the Kushan Empire (ca. 2nd–early 3rd cent.). On the other hand, the possibility of an immediate update about the Kidarite expansion in the early 5th century should not be entirely eliminated.

At any rate, Bactria as a marker of dress style in the Chinese Sarvāstivāda Vinaya stimulates us to reconstruct its material culture to a fuller degree. Since a region's political and cultural centers, geographical range, and common appellations can considerably change over its long history, it is wiser not to define a cultural or linguistic border by a few non-native sources. The polysemy of Chinese characters and the complications of later text modifications also deserve more caution of researchers. Despite all these restrictions, this article confirms the large language diversity in pre-Islamic Central Asia and encourages scholars to continue to investigate its earlier phase.

Abbreviations

Grammatical and language abbreviations

Bactr.	Bactrian	LMC	Late Middle Chinese, see Pulleyblank 1991
BSkt.	Buddhist Sanskrit	NP	New Persian
Chin.	Chinese	Pkt.	Prakrit
EMC	Early Middle Chinese, see Pulleyblank 1991	Skt.	Sanskrit
Gandh.	Gāndhārī	Sogd.	Sogdian

Text, woodblock-print, and shelfmark abbreviations

CKI	Catalog of Gāndhārī Texts: Inscriptions. See Baums & Glass 2002–
CSZJJ	*Chu sanzang ji ji* 出三藏記集 [*Collection of notes concerning the translation of the Tripiṭaka*], compiled by Sengyou 僧祐 (433–518) = T.2145
GSZ	*Gaoseng zhuan* 高僧傳 [*Biographies of eminent monks*], compiled by Huijiao 慧皎 (497–554) = T.2059
j.	*juan* 卷 'scroll, chapter'
Pilu	Chinese Buddhist canon of *Pilu dazang jing* 毘盧大藏經, carved in Fuzhou by 1148. A print of CSZJJ in this canon housed in the Library of Japanese Imperial Court is digitised, see: db2.sido.keio.ac.jp/kanseki/bib_frame?id=007075_1068 (access on 14 Oct. 2024).

Puning Chinese Buddhist canon *Puning zang* 普寧藏, carved in Hangzhou in 1277–1290

SGZ *Sanguo zhi* 三國志 [*Record of the Three Kingdoms*], composed by CHEN Shou 陳壽 (233–297), edition by Zhonghua shuju 1959

Sixi Chinese Buddhist canon *Sixi zang* 思溪藏, carved in Huzhou by 1239

SS *Suishu* 隋書 [*Book of Sui*], supervised by LINGHU Defen 令狐德棻 (583–666) *et al.* and completed in 636, edition by Wang *et al*. 1973

WS *Weishu* 魏書 [*Book of Wei*], compiled by WEI Shou 魏收 (507–572), edition by Zhu *et al.* 2017

1st/2nd Korean carving = The first and second carving of Chinese Buddhist Canon in Korea in ca. 1011–1087 and in 1236–1251, respectively. A partially burnt print of the 1st carving is reproduced in China as the *Gaoli dazang jing* 高麗大藏經 (2012, Beijing).

[The 2nd Korean, Puning, and Sixi prints preserved at the *Zōjōji* 増上寺 (Tokyo) have been digitised, see https://jodoshuzensho.jp/zojoji_sandaizo/search/detail_document.php (access on 16 Oct. 2024)]

References

ALLON, Mark, 2001. *Three Gāndhārī Ekottarikāgama-type sūtras: British Library Kharoṣṭhī fragments 12 and 14* (with a contribution by Andrew GLASS), Seattle–London.

ALLON, Mark, 2021. *The composition and transmission of early Buddhist texts with specific reference to sutras*, Bochum–Freiburg.

Anālayo (Bhikku) & Roderick S. BUCKNELL, 2006. 'Correspondence table for parallels to the discourses of *Majjhima Nikāya*: Toward a revision of Akanuma's *Comparative Catalogue*', *Journal of the Centre for Buddhist Studies (Sri Lanka)* 4, 215–243.

ARAKAWA Masaharu 荒川正晴, 1997. 'Tōteikoku to Sogudojin no kōeki katsudō 唐帝國とソグド人の交易活動', *Tōyōshi kenkyū* 東洋史研究 – *The Journal of Oriental Researches* 56/3, 171–204.

BAILEY, Harold W., 1937. 'Ttaugara', *BSOS* 8/4, 883–921.

BAILEY, Harold W., 1942. 'Kanaiṣka', *JRAS* 1, 14–28.

BAILEY, Harold W., 1965. 'Viśą Saṃgrāma', *Asia Major*, n.s., 11/2, 101–119.

BAILEY, Harold W., 1969. *Khotanese Texts I–III* (2nd edition), Cambridge.

BAKKER, Hans, 2020. 'Huns in Central and South Asia: A timeline', in *Huna sourcebook*, xv–xvii.

BAKKER, Hans, Csaba DEZSŐ, Gergely HIDAS & Dániel BALOGH, 2020. 'Indic sources', in *Huna sourcebook*, 275–364.

BAUMS, Stefan & Andrew GLASS, 2002–. *Catalog of Gāndhārī texts.* See https://gandhari.org/catalog.

BECHERT, Heinz, 1987. *Zur Schulzugehörigkeit von Werken der Hīnayāna-Literatur*, Göttingen.

BENJAMIN, Craig G. R., 2007. *The Yuezhi. Origin, migration and the conquest of Northern Bactria*, Turnhout.

BERNARD, Paul, 2002. 'L'oeuvre de la Délégation archéologique française en Afghanistan (1922–1982)', *CRAI* 146, 1287–1323.

BERNHARD, Franz, 1965. *Udānavarga. Band I: Einleitung, Beschreibung der Handschriften, Textausgabe, Bibliographie*, Göttingen.

BERNHARD, Franz, 1967. 'Zur Entstehung einer Dhāraṇī', *ZDMG* 117, 148–168.

BD II = SIMS-WILLIAMS, Nicholas, 2007. *Bactrian documents from Northern Afghanistan, II: Letters and Buddhist texts*, London.

BD IV = SIMS-WILLIAMS, Nicholas, 2025. *Bactrian documents IV: Documents from south of the Hindukush, I* (with a contribution by Frantz GRENET), London.

BHSD = EDGERTON, Franklin, 1953. *Buddhist Hybrid Sanskrit grammar and dictionary. Vol. 2: Dictionary*, New Haven, CT.

BONMANN, Svenja, Jakob HALFMANN, Natalie KOROBZOW & Bobomullo BOBOMULLOEV, 2023, 'A partial decipherment of the unknown Kushan script', *TPS* 121/2, 293–329.

BOUCHER, Daniel, 1998. 'Gāndhārī and the early Chinese Buddhist translations reconsidered: The case of the *Saddharmapuṇḍarīkasūtra*', *JAOS* 118/4, 471–506.

BOUCHER, Daniel, 2000. 'On *hu* and *fan* again: The transmission of "barbarian" manuscripts to China', *JIABS* 23/1, 7–28.

BURROW, Thomas, 1937. *The language of the Kharoṣṭhi documents from Chinese Turkestan*, Cambridge.

CBETA = Comprehensive Buddhist Electronic Text Archive, see online database at: https://cbetaonline.dila.edu.tw/en/.

CHING Chao-jung, 2011. 'Silk in Ancient Kucha: On the Toch. B word *kaum** found in the documents of the Tang period', *TIES* 12, 63–82.

CHING Chao-jung 慶昭蓉, 2014. 'Qiuci seng Liyan de shengping shiye (shang) — jianlun Tangren suowei 'Tuhuoluoyan' 龜茲僧利言的生平事業(上)——兼論唐人所謂"吐火羅言" – The life and career of the monk Liyan from Kucha (I): Note on Tuhuoluoyan the "Tocharian language" in Tang Chinese', *Tang yanjiu* 唐研究 – *Journal of Tang Studies* 20, 469–490.

CHING Chao-jung, 2016. 'On the names of cereals in Tocharian B', *TIES* 17, 29–64.

CHING Chao-jung 庆昭蓉, 2023. 'Hanyi fojing zhong de Bozhi guo — Poquli, Boquli and Boheluo 汉译佛经中的薄知国——婆佉梨、薄佉利与博喝罗 – On the area around Balkh in Chinese Buddhist scriptures: *Poquli, Boquli* and *Boheluo*', *XYWS* 17, 99–122.

CHING Chao-jung, 2025. 'Ancient Chinese translators in diplomacy and the historical process of acculturation', in *A cultural history of translation. Vol. 1: A cultural history of translation in classical societies* (ed. Joseph VERHEYDEN), London, 141–162.

CHING Chao-jung & Imre GALAMBOS, 2020. 'Chinese sources' (with contributions by Max DEEG & INABA Minoru), in *Huna sourcebook*, 1–112.

CHING Chao-jung, ENAMI Kazuyuki & OKADA Yoshihiro, 2021. 'Paper in eighth-century Kucha: Discovery of cotton fibres within Chinese and Kuchean documents', *CAJ* 63/1–2 (2020), 71–104.

CHUNG Jin-il & Klaus WILLE, 2002, 'Fragmente aus dem Bhaiṣajyavastu der Sarvāstivādins in der Sammlung Pelliot (Paris)', in *Sanskrit-Texte aus dem buddhistischen Kanon: Neuentdeckungen und Neueditionen IV* (ed. Jin-il CHUNG *et al.*), Göttingen, 105–125.

ENOKI Kazuo 榎一雄, 1958. 'Kidāra ōchō no nendai ni tsuite キダーラ王朝の年代について – The date of the Kidārites', *Tōyō Gakuhō* 東洋學報 – *Reports of the Investigations of the Oriental Society / Reports of the Oriental Society* 41/3, 1–52.

ENOKI Kazuo, 1998. *Studia Asiatica. The collected papers in Western languages of the late Dr. Kazuo Enoki*, Tokyo.

ENOMOTO Fumio 榎本文雄, 1993. 'Keihin — Indo bukkyō no chūshinchi no shozai 罽賓——インド仏教の一中心地の所在', in *Chi no kaikō: Bukkyō to kagaku* 知の邂逅——仏教と科学 (ed. Tsukamoto Keishō kyōju kanreki kinen ronbunshū kankōkai 塚本啓祥教授還暦記念論文集刊行会), Tokyo, 259–269.

FALK, Harry, 2018. 'The five Yabghus of the Yuezhi', *BAI* 28 (2014[2018]), 1–43.

FRANCFORT, Henri-Paul, 2022. 'Aux marches orientales de la Bactriane antique. Nouvelle perspectives', *CRAI* 164, 97–147.

FUJIMOTO Yōsuke 藤本庸裕, 2022. 'Buddha no shintai wa naze uro ka: setsuissaiubu ni yoru busshin urosetsu ronshō no ronriteki bunseki ブッダの身体はなぜ有漏か: 説一切有部による仏身有漏説論証の論理的分析 – Why is the Buddha's body impure (*sāsrava*)? A logical analysis of the Sarvāstivādin argument for the impurity of the Buddha's Body', *Journal of the Department of Oriental Philosophy School of Literature, Waseda University* 39, 22–46.

FUNAYAMA Tōru 船山徹, 2013. *Butten wa dō kanyaku sareta no ka — sūtora ga kyōten ni naru toki* 仏典はどう漢訳されたのか——スートラが経典になるとき – *Making sutras into 'Classics' (jingdian): How Buddhist scriptures were translated into Chinese*, Tokyo.

FUNAYAMA Tōru 船山徹, 2021. *Basubanzu den — Indo bukkyō shisōka Vasubandu no denki* 婆藪槃豆伝——インド仏教思想家ヴァスバンドゥの伝記 – *The biography of the Indian Master Vasubandhu*, Kyoto.

FUNAYAMA Tōru 船山徹, 2022. *Bukkyō kango gogi kaishaku — Kanji de fukameru bukkyō rikai* 仏教漢語語義解釈——漢字で深める仏教理解 – *An in-depth understanding of Buddhist terms in Chinese*, Kyoto.

FUSSMAN, Gérard, 2015. 'Kushan power and the expansion of Buddhism beyond the Soleiman mountains', in: *Kushan histories*, 153–202.

GRENET, Frantz, 2002. 'Regional interaction in Central Asia and Northwest India in the Kidarite and Hephthalite periods', in *Indo-Iranian languages and peoples* (ed. Nicholas SIMS-WILLIAMS), London, 203–224.

GRENET, FRANTZ, 2015. 'Zoroastrianism among the Kushans', in *Kushan histories*, 203–239.

HANEDA Tōru 羽田亨, 1933. ‘A propos des Ta Yue-tche et des Kouei-chouang’, *Bulletin de la Maison Franco-Japonaise* 4, 1–28.

HARTMANN, Jens-Uwe, 2024. ‘Teaching many with one utterance: The elusive concept of *ekā vāc*’, in *Siddham. Studies in Iranian philology in honour of Mauro Maggi* (ed. Gerardo BARBERA *et al*.), Wiesbaden, 191–206.

HIRAKAWA Akira 平川彰, 1974. *Indo bukkyō shi (jō kan)* インド佛教史(上巻), Tokyo.

HIRAKAWA Akira 平川彰, 1989. *Shoki daijō bukkyō no kenkyū* 初期大乘佛教の研究, vol. I, Tokyo.

Huna sourcebook = Dániel BALOGH (ed.), *Hunnic peoples in Central and South Asia: Sources for their origins and history*, Groningen, 2020.

INABA Minoru 稲葉穣, 2022. *Isurāmu no higashi, chūka no nishi — Shichi-hachi seiki no chūō ajia wo megutte* イスラームの東・中華の西——七〜八世紀の中央アジアを巡って – *The history of Central Asia between Islamdom and China from the seventh to the eighth century*, Kyoto.

KARASHIMA Seishi 辛嶋静志, 2015. ‘Who composed the Mahāyāna scriptures? The Mahāsāṃghikas and *Vaitulya* scriptures’, *ARIRIAB* 18, 113–162.

Kushan histories = Harry FALK (ed.), 2015. *Kushan histories. Literary sources and selected papers from a symposium at Berlin, December 5 to 7, 2013*, Bremen.

KUWAYAMA Shōshin 桑山正進, 1985. ‘Tohārisutān no Efutaru, Churuku to sono jōyū トハーリスターンのエフタル、テュルクとその城邑’, in *Orientogaku ronshū — Mikasano-miya Denka koki kinen* オリエント学論集——三笠宮殿下古稀記念 (ed. The Society for Near Eastern Studies in Japan), Tokyo, 140–150.

KUWAYAMA Shōshin, 1989. ‘The Hephthalites in Tokharistan and Northwest India’, *Zinbun* 24, 89–134.

KUWAYAMA Shōshin 桑山正進, 1990. *Kāpishī-Gandāra shi kenkyū* カーピシー=ガンダーラ史研究 – *Studies of history and archaeology in the Kābul Valley*, Kyoto.

KUWAYAMA Shōshin 桑山正進, 2017. ‘Kisō Kyūshūkyaku no botsunen 貴霜丘就却の歿年 – Date of the Kushan Kujula Kadphises’ demise based on Chinese textual sources’, *Tōhō gakuhō* (Kyoto) 東方學報（京都）92, 77–134.

KUWAYAMA Shōshin 桑山正進, 2020. ‘Tohāra no kyōiki to Ranshijō to Katsukoku tojō トハーラの境域と藍市城と活国都城 – Notes on Tokhāra: Towns and territory’, *SAK* 91, 119–132.

KUWAYAMA Shōshin, 2022a. ‘The Kushans, viz. the Da Yuezhi: A century-long fallacy’, in *The world of the ancient Silk Road* (ed. LIU Xinru), London–New York, 190–202.

KUWAYAMA Shōshin, 2022b. ‘The Chitral *Yaβγu* 雙靡翕侯 and a route southward in the 1st century BCE’, in *The world of the ancient Silk Road* (ed. LIU Xinru), London–New York, 322–335.

KUWAYAMA Shōshin 桑山正進 (ed.), 1992. *Echō go tenjiku koku den kenkyū* 慧超往五天竺國傳研究 – *Huichao's Wang Wu-Tianzhuguo zhuan 'Record of travels in five Indic regions'. Translation and Commentary*, Kyoto.

DE LA VAISSIÈRE, Étienne, 2004. *Histoire des marchands sogdiens* (2ème édition révisée et augmentée), Paris.

DE LA VAISSIÈRE, Étienne, 2005. 'Huns et Xiongnu', *CAJ* 49/1, 1–26.

DE LA VAISSIÈRE, Étienne, 2010a. 'De Bactres à Balkh, par le Nowbahār', *JA* 298/2, 517–533.

DE LA VAISSIÈRE, Étienne, 2010b. 'The last Bactrian kings', in *Coins, Art and Chronology, II. The first millennium C.E. in the Indo-Iranian borderlands* (ed. Michael ALRAM *et al*.), Wien, 213–218.

DE LA VAISSIÈRE, Étienne, 2024. *Asie centrale 300–850. Des routes et des royaumes*, Paris.

LAMOTTE, Étienne, 1970. *Le traité de la grande vertu de sagesse de Nāgārjuna (Mahāprajñāpāramitāśāstra)*, Louvain.

LÉVI, Sylvain, 1897. 'Notes sur les Indo-Scythes. II. Les textes historiques', *JA*, jan.-fév. 1897, 5–42.

LÉVI, Sylvain, 1904. 'Notes chinoises sur l'Inde. IV. Le pays de Kharoṣṭra et l'écriture kharoṣṭrī', *BEFEO* 4, 543–579.

LÉVI, Sylvain, 1905. 'Notes chinoises sur l'Inde. V. Quelques documents sur le bouddhisme indien dans l'Asie centrale (première partie)', *BEFEO* 5, 253–305.

LÉVI, Sylvain, 1915. 'Le catalogue géographique des *yakṣa* dans la *Mahāmāyūrī*', *JA*, 11e série, 5, 19–137.

LÉVI, Sylvain, 1933. 'Le «tokharien»', *JA* 222, 1–30.

MAGGI, Mauro, 2020. 'Khotanese sources', in *Huna sourcebook*, 113–121.

MAGGI, Mauro, 2022. 'Some remarks on the history of the Khotanese orthography and the Brāhmī script in Khotan', *ARIRIAB* 25, 149–172.

MARQUART, Josef, 1901. *Ērānšahr nach der Geographie des Ps. Moses Xorenac'i*, Berlin.

MATSUDA Kazunobu, 2002. 'Three fragments related to the *Śāriputra-Abhidharma*', in *Manuscripts in the Schøyen Collection III. Buddhist manuscripts, vol. II* (ed. Jens BRAARVIG), Oslo, 239–248.

MINORSKY, Vladimir, 1970. *Ḥudūd al-'Ālam 'The regions of the world': A Persian geography* (ed. Clifford E. BOSWORTH), Cambridge.

MITOMO Kenyō 三友健容, 2011. '*Basharon* seiritsu no sho mondai 『婆沙論』成立の諸問題', in *Indo bukkyōshi bukkyōgaku ronsō: Nakazawa Kōyū hakase koki kinen ronbunshū* インド仏教史仏教学論叢: 仲澤浩祐博士古稀記念論文集 (ed. Nakazawa Kōyū hakase koki kinen ronbunshū kankōkai), Tokyo, 21–47.

MW = MONIER-WILLIAMS, Monier, 1899. *A Sanskrit-English dictionary: Etymologically and philologically arranged with special reference to cognate Indo-European languages* (new edition, greatly enlarged and improved), Oxford.

MÜLLER, F. W. K., 1918. 'Toχrï und Kuišan (Küšän)', *SPAW* 27, 566–586.

ŌCHŌ Enichi 横超慧日, 1958. ‘Kumarajū no honyaku 鳩摩羅什の翻譯’, *Ōtani gakuhō* 大谷學報 – *The Otani gakuho. The Journal of Buddhist Studies and Humanities* 37/4, 1–25.

PELLIOT, Paul, 1934. ‘Tokharien et koutchéen’, *JA* 224, 23–106.

PELLIOT, Paul, 2002. ‘Notes sur Kumārajīva’, in: *A life journey to the East. Sinological studies in memory of Giuliano Bertuccioli (1923–2001)* (ed. Antonio FORTE & Federico MASINI), Kyoto, 1–19.

PEYROT, Michaël, 2008. *Variation and change in Tocharian B*, Amsterdam–New York.

PINAULT, Georges-Jean, 2007. ‘Sylvain Lévi dechiffreur et lecteur des textes des frontières’, in: *Sylvain Lévi (1863–1935). Études indiennes, histoire sociale* (ed. Lyne BANSAT-BOUDON & Roland LARDINOIS), Turnhout, 111–144.

PTSD = RHYS DAVIDS, T. W. & William STEDE, 1979. *The Pali Text Society’s Pali-English dictionary*, London.

PULLEYBLANK, Edwin G., 1991. *Lexicon of reconstructed pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*, Vancouver.

RADICH, Michael. 2010, ‘Embodiments of the Buddha in Sarvâstivāda doctrine: With special reference to the **Mahāvibhāṣā*’, *ARIRIAB* 13, 121–172.

RADICH, Michael, 2017. ‘On the *Ekottarikāgama* 增壹阿含經 T 125 as a work of Zhu Fonian 竺佛念’, *Zhonghua foxue xuebao* 中華佛學學報 – *Journal of Chinese Buddhist Studies* 30, 1–31.

RADICH, Michael & Bikkhu Anālayo, 2017. ‘Were the *Ekottarika-āgama* and the *Madhyama-āgama* translated by the same person? An accessment on the basis of translation style’, in *Research on the Madhyama-āgama* (ed. Dhammadinnā), Taipei, 209–237.

VON RICHTHOFEN, Ferdinand, 1877. *China. Ergebnisse eigener Reisen und darauf gegründeter Studien. Erster Band*, Berlin.

RONG Xinjiang & WEN Xin, 2009. ‘Newly discovered Chinese-Khotanese bilingual tallies’, *JIAAA* 3 (2008[2009]), 99–118.

SALOMON, Richard, 2024. ‘Kharoṣṭhī out of bounds’, *Fanxue* 梵学 – *Studia Indica* 1, 511–545.

SANDER, Lore, 2009.‘Remarks on the Formal Brāhmī Script from the Southern Silk Route’, *BAI* 19 (2005[2009]), 133–144.

SASAKI Shizuka 佐々木閑, 2005. ‘Shidahanni sen *Bibasharon* to Budabama tō yaku *Abidon-bibasha-ron* no taishō hyō 尸陀槃尼撰『鞞婆沙論』と浮陀跋摩等訳『阿毘曇毘婆沙論』の対照表’, *Bukkyō kenkyū* 仏教研究 33, 51–82.

SASAKI Shizuka 佐々木閑, 2007. ‘*Basharon* shohon no sōgo kankei『婆沙論』諸本の相互関係’, *IBK* 113, 167–173.

SASAKI Shizuka 佐々木閑, 2011. ‘*Basharon* no kōzō bunseki ni okeru ryūi ten『婆沙論』の構造分析における留意点’, *IBK* 123, 164–172.

SHIIO Benkyō 椎尾辨匡, 1912. ‘Tokara no minzoku nendai 覩貨羅の民族年代’, *Shigaku zasshi* 史学雑誌 23/6, 49–62.

SIMS-WILLIAMS, Nicholas, 2002. 'Ancient Afghanistan and its invaders: Linguistic evidence from the Bactrian documents and inscriptions', in *Indo-Iranian languages and peoples* (ed. Nicholas SIMS-WILLIAMS), London, 225–242.

SIMS-WILLIAMS, Nicholas, 2015. 'A new Bactrian inscription from the time of Kanishka', in *Kushan Histories*, 255–264.

SIMS-WILLIAMS, Nicholas, 2016. 'Bactria – Balkh: Variations on a place-name', in *Études de Linguistique iranienne. In memoriam Xavier Tremblay* (ed. Éric PIRART), Leuven–Paris–Bristol, 273–281.

SIMS-WILLIAMS, Nicholas, 2025. 'Bactrian in two scripts: Greek and Kushan', *IIJ* 68/3, 185–214.

SIMS-WILLIAMS, Nicholas & Frantz Grenet, 2023. 'A new collection of Bactrian letters on birchbark', *BAI* 31 (2022–2023 [2023]), 135–144.

SIRCAR, D. C., 1971. *Studies in the geography of ancient and medieval India* (2nd ed.), Delhi.

SKJÆRVØ, Prods Oktor, 2003. *Khotanese manuscripts from Chinese Turkestan in the British Library: A complete catalogue with texts and translations*, with contributions by Ursula SIMS-WILLIAMS, London.

STRAUCH, Ingo, 2012. 'The character of the Indian Kharoṣṭhī script and the "Sanskrit Revolution": A writing system between indentity and assimilation', in *The idea of writing: Writing across borders* (ed. Alex DE VOOGT & Joachim F. QUACK), Leiden–Boston, 131–168.

T. = TAKAKUSU Junjirō 高楠順次郎 *et al*. (ed.), 1924–1934. *Taishō shinshū daizōkyō* 大正新脩大藏經 [*Taishō Tripiṭaka*], Tokyo.

TREMBLAY, Xavier, 2004. 'La toponymie de la Sogdiane et le traitement de **xθ* et **fθ* en iranien. Essais de grammaire comparé des langues iraniennes VII', *StIr* 33, 113–149.

UMEHARA Kaoru 梅原郁, 2001. 'Zenzenkoku no kōbō: Rōran no kyojitsu 鄯善国の興亡: 楼蘭の虚實 – Prosperity and disintergration of Shan Shan 鄯善: Truth and falsehood of Lou-lan 楼蘭', in *Ryūsa shutsudo no moji shiryō* 流沙出土の文字資料 – *Written materials excavated from the sands* (ed. TOMIYA Itaru 冨谷至), Kyoto, 253–299.

UTSUNOMIYA Kiyoyoshi 宇都宮清吉 1979. *Daitō daijionji sanzōhōshi den kōi oyobi sakuin* 大唐大慈恩寺三藏法師傳考異及索引, Kyoto.

WANG Shaoying 汪紹楹, YIN Falu 陰法魯 *et al.* (ed.), 1973. *Suishu*, 6 vols., Beijing.

WATTERS, Thomas, 1904. *On Yuan Chwang's travels in India, 629–645 A.D., vol. I*, London.

YOSHIKAWA Tadao 吉川忠夫 & FUNAYAMA Tōru 船山徹, 2009. *Kōsō den* 高僧伝, Tokyo.

YU Taishan 余太山, 2012. *Yeda shi yanjiu* 嚈噠史研究, Beijing.

YU Taishan 余太山, 2015. *Guishuang shi yanjiu* 貴霜史研究, Beijing.

Zhonghua book company 中華書局 (dir.), 1959. *Sanguo zhi* 三國志, 5 vols., Beijing.

ZHU Lei 朱雷, DONG Guodong 凍國棟 *et al.* (ed.), 2017. *Weishu* 魏書, 8 vols., Beijing.

ZÜRCHER, Erik, 2007. *The Buddhist conquest of China. The spread and adaptation of Buddhism in Early Medieval China* (3rd ed. with a foreword by Stephen F. TEISER), Leiden.

T.1547 (I)	Zhendan 真旦					Tuole 陀勒	Mole 摩勒?	Bole 波勒?	Qusha 佉沙	Poquli 婆佉梨		
T.1547 (II)	Zhendan 真旦					Tuole 陀勒	Mole 摩勒?	Bole 波勒?	Qusha 佉沙	Poquli 婆佉梨	Douqule 兜佉勒	
T.1546 (I)	Zhendan 真丹			Shijia 釋迦	Yemona 夜摩那	Tuoluotuo 陀羅陀	Moluopo 摩羅婆		Qusha 佉沙		Douquluo 兜佉羅	
T.1546 (II)	Qin 秦	[Tuo]piluo [陀]毗羅	Miliche 彌梨車									
T.1546 (III)	Qin 秦		Miliche 彌梨車	Shijia 釋迦								
T.1545 (I)	Zhina 至那			Lijia 磔迦	Yefana 葉筏那	Dalatuo 達剌陀	Moxiapo 末鞞婆		Qusha 佉沙		Duhuoluo 覩貨羅	Boheluo 博喝羅
T.1545 (II)	Zhina 至那											Boheluo 博喝羅
T.1545 (III)	Zhina 至那			Lijia 磔迦								Boheluo 博喝羅
Restored in Skt.	*Cīnasthāna*	?	*Mleccha-*	*Śaka-*	*Yavana-*	*Darada-*	*Mārava-* (?)	*Pahlava-* (?)	*Khaśa-*	**Bākh(a)li-* (?)	*Tukhāra-*	**Pahla-* **Pāhala-* (?)
Approx. regions	China	(Southern India?)	People of northern frontier	Northwest India around Sagala	Gandhāra	Dardistan (around the Hindukush)	Deserted area?	Parthia	Himalaya / Pamirs (or: Kashgar)	Bactria	Tukharistan	Bactria (confused with Parthia?)

Table 4. Non-aryan countries and languages concerning the debate on the Buddha's language proficiency.

A comparative study of the *Mahāvaidehaghṛta* in Sanskrit, Khotanese, and Tocharian B*

Alessandro DEL TOMBA

The so-called *Jīvakapustaka* is a collection of 93 Āyurvedic prescriptions preserved in the incomplete tenth-century Dunhuang manuscript Ch. ii.003 of the British Library (shelf marks IOL Khot 87–110). The manuscript features sections in corrupt Sanskrit verse alongside their Late Khotanese prose renditions and is written in a distinctive form of the Khotanese documentary Brāhmī script. Its interpretation poses multiple challenges, primarily due to the peculiar script type, the complexities in deciphering the Sanskrit text, and the bewildering spelling variants used for writing Late Khotanese. This article aims to provide a new edition, translation, and commentary of the medicated ghee known as *Mahāvaidehaghṛta* (JP 11), which is unique in having a Tocharian B parallel. It also introduces a method for reconstructing the Sanskrit text by combining metrical analysis, comparison with the expanded Khotanese version, and attention to the scribe's spelling practices. The article argues that the corrupt form of the Sanskrit reflects the intervention of a Late Khotanese scribe, who adapted Sanskrit to fit the phonological and graphic norms of his own language.

1. Introduction

The medical text known by the conventional title of *Jīvakapustaka* (JP) 'The book of Jīvaka' is a collection of Āyurvedic medical prescriptions preserved in a bilingual manuscript written alternately in corrupted Sanskrit and Late Khotanese.[1] The

* For comments and suggestions, I am grateful to Ching Chao-jung 慶昭蓉, Mehrdad Derafshi, Silvia Luzzietti, Mauro Maggi, Michaël Peyrot, Ogihara Hirotoshi, Pan Tao 潘濤, and Doug Hitch. A list of the abbreviations used is provided at the end of the paper. In the translations, Sanskrit terms for ingredients, botanical elements, and specific diseases are given in non-italics to enhance readability. For the identification of plant names, I primarily used Khare 2007, Meulenbeld 1999–2002, the *International Plant Names Index* (www.ipni.org/), and the *Medicinal Plant Names Service* (mpns.science.kew.org). The research for this article was conducted as part of the project "Language contact and the transmission of medical knowledge from the Ancient Near East to Central Asia" (Grant n. RG123188B0B4B391, PI Alessandro Del Tomba), funded by Sapienza University of Rome.

[1] Since no title is extant in the manuscript, which has no colophon either, Bailey called the text *Jīvakapustaka* on account of the mention of the famous physician Jīvaka at the beginning of the text. This title has been widely adopted.

10.29091/9783752003635/002

manuscript has likely been produced in tenth-century Dunhuang in a Khotanese-speaking milieu. The Sanskrit original is in verse, while its Khotanese rendition is in prose. The beginning of the text has survived, but the manuscript ends abruptly in the middle of a Sanskrit recipe. It was discovered by Sir Aurel Stein in the Caves of the Thousand Buddhas (Qianfodong) near Dunhuang in Northwestern China. It is currently housed in the British Library (manuscript Ch. ii. 003 = IOL Khot 87–110). Presumably dating from the tenth century, the manuscript consists of 71 folios, which are numbered from 44 to 116.[2]

Sten Konow (1941) provided a first edition, translation, and glossary of the Khotanese version, dividing the surviving text into 4 chapters and 93 paragraphs. Konow's numbering of the paragraphs represents the standard way of citing. Although Konow sometimes consulted the Sanskrit version to translate the Khotanese, he considered the Sanskrit text so corrupt that it could hardly be read. In Konow's own words, "it is hardly likely that the writer of our manuscript has properly understood it, as it now stands" (1941: 6).

In the first volume of *Khotanese Texts*, Harold W. Bailey provided a diplomatic edition of the whole of the Sanskrit and Khotanese versions on facing pages, but without any attempt at interpreting the Sanskrit text.[3] Later, Chen Ming 陳明 (2005) published a tentative reconstruction of the Sanskrit based on Bailey's diplomatic edition.[4] Majid Ṭāme has completed a new edition, Persian translation, and glossary of the Khotanese in his unpublished thesis (2014). Studies on single prescriptions preserved in the *Jīvakapustaka* manuscript have been published by Ronald E. Emmerick and by Mauro Maggi.[5]

More than two decades before Konow's work, A. F. Rudolf Hoernle prepared a study of ff. 44–72r (corresponding to the first 25 paragraphs), which he intended to publish under the title *An Ancient Medical Manuscript from Turkestan*.[6] In his study, Hoernle provides an introduction to the manuscript and its script, accompanied by an edition and translation of the Khotanese version in light of his reconstruction of the corrupt Sanskrit text, which he describes as "barbarous". He marked the manuscript of his unpublished

[2] See Maggi 2009: 350–351 and 414–415. The twenty-sixth folio bears two numbers, 69 as well as 71; number 70 is omitted altogether. The twenty-seventh folio is marked with number 72.

[3] Bailey KT 1: 135–195. Facsimile in Bailey 1938: 69–141.

[4] It seems that Chen's transcription of the Sanskrit original is a mere transposition of Bailey's edition, the only differences being that Chen eliminates all unetymological anusvāras and modifies *ä* mechanically into *a*.

[5] See Emmerick 1992a, 1994, 1997 and Maggi 2022. See also Emmerick 1979a.

[6] Cf. Emmerick 1992b: 43. Between the title *An Ancient Medical Manuscript from Turkestan* and *Press-Copy*, there is the signature of the manuscript labelled as Ch. ii. 003, seemingly written with a different pen. It remains unclear whether Hoernle himself added this signature.

work as *Press-Copy*, which is currently hosted in the India Office Library. In the present study, I have extensively considered Hoernle's unpublished work (Hoernle s.d.).

In his review of Bailey's edition of the *Jīvakapustaka* manuscript, Jean Filliozat (1946–1947: 134–135) identified parallel prescriptions in the *Carakasaṃhitā* (Ca). Emmerick (1979a) further found a number of parallels in other Āyurvedic texts, including the *Suśrutasaṃhitā* (Su), the *Bhelasaṃhitā*, the *Aṣṭāṅgahṛdayasaṃhitā* (AH), and the *Siddhasāra* (Si).

Among the prescriptions that do not have any parallels in other Āyurvedic medical texts is the medicated ghee named *Mahāvaidehaghṛta*, which corresponds to paragraph 11 of the *Jīvakapustaka*. Paragraph 11 belongs to the second chapter, dedicated to medicated oils, where each section is initially written in Sanskrit and then translated into Khotanese.

In 1990, Dieter Maue identified a Tocharian B parallel of JP 11 in the combined manuscript THT 500–501 (Yarxoto), initially published by Sieg and Siegling (1953: 311). Maue's article contains an accurate reconstruction of the first three verses of the Sanskrit version of JP 11, a re-edition of the Tocharian B version with a joint of two fragments, and a table illustrating the correspondence between the medicaments cited in the Sanskrit, Khotanese, and Tocharian B versions. The article also provides notes addressing philological-linguistic problems and the identification of the ingredients. Maue's study is a significant contribution also from a broader perspective, as THT 500–501 is, as far as I know, the only Tocharian medical prescription with a precise parallel in Khotanese. Additionally, given the often-limited availability of clear Indian parallels for Tocharian medical texts, Maue's identification facilitated the identification of several previously unknown Tocharian ingredients with a high degree of certainty.[7] Later, the third fragment THT 502 was joined to the top of THT 500 by the Berlin-Brandenburgische Akademie der Wissenschaften. The surviving manuscript has been re-edited by Melanie Malzahn (*apud* CEToM).

The present article delves into philological, linguistic, and textual problems related to interpreting the Sanskrit text of JP 11, and its Khotanese and Tocharian B renditions. It begins with a new edition and translation of the three versions, with many suggestions for improvement of the reconstruction of the Sanskrit text and its Khotanese translation (§2). The commentary encompasses philological and linguistic notes on the reconstruction of the Sanskrit version (§3.1.1), a linguistic analysis of the Khotanese text, an explanation for the corrupt nature of the Sanskrit original (§3.1.2), and a linguistic analysis of the Tocharian B text, as well as a discussion of the origin of the

[7] For medical literature in Tocharian, see Filliozat 1948 and Carling 2003, 2007. For the relationships between Tocharian and Khotanese in the context of medical vocabulary, see Dragoni 2021 and 2023.

Sanskrit mantra attested at the beginning of the Tocharian B manuscript (§3.1.3). The article further addresses linguistic, textual, and semantic problems related to the medicaments in the three versions (§3.2), and comments on the diseases in the Sanskrit and Khotanese versions (§3.3).

2. Complete edition and translation of the three versions

2.1. Edition and translation of the Sanskrit version (*Jīvakapustaka* 11)

The Sanskrit section extends from the middle of the first line of the recto of folio 58 (IOL Khot 91/3) to the middle of the second line of the verso of the same leaf. Previous editions of the Sanskrit version of JP 11 are by Hoernle (s.d.: 45, 46), Bailey (first ed. KT[1] 1: 149; second ed. KT[2] 1: 149), and Chen (2005: 305–306). Chen's diphlomatic edition essentially relies on Bailey's (cf. fn. 4). Maue (1990) provides an accurate, metrically informed reconstruction of the first three verses.

The text below is divided into three sections. The first section provides a diplomatic edition of the Sanskrit text arranged metrically in alignment with the critical edition (§2.1.1). Accompanying this is an apparatus that provides notes and information on divergent readings by others. The second section offers the critical edition of the text in standard Sanskrit orthography (§2.1.2). The critical apparatus gives manuscript readings from the diplomatic edition that differ from the reconstructed text. Since the Sanskrit text follows the orthographic conventions used to write Khotanese (see §3.1.2), vowel or consonant emendations understandable in terms of the phonology and orthography of Late Khotanese are left unmarked. Explanations and comments on the reconstructed text are given in §3.1.1 of the Commentary. Preceding the apparatus is the metrical scansion of the hemistichs. Each śloka or hemistich adheres to the pathyā type, whose metrical profile is as follows: × × × × ¦ ⏑ – – × || × × × × ¦ ⏑ – ⏑ ×. The third section contains the translation of the reconstructed Sanskrit text (§2.1.3).[8]

[8] The diplomatic edition reflects the metrical arrangement of the critical edition. The beginnings of folios and lines are indicate by raised numbers, accompanied by the abbreviations r (recto) and v (verso). The apparatus gives significantly divergent readings proposed by other scholars. The critical edition is linked to the diplomatic edition through verse numbers. Its apparatus lists manuscript readings established in the diplomatic edition that differ from the reconstructed text. The following conventions are used: ‹abc› = editorial supplement; [abc] = restoration; ⁺abc = emendation; {abc} = editorial expunction; ◯ = string hole space; ××× = missing akṣaras; B = Bailey, KT[1] (1945) and KT[2] (1969); B[1] = Bailey, KT[1]; B[2] = Bailey, KT[2]; C = Chen 2005; DT = Del Tomba; H = Hoernle (s.d.); M = Maue 1990.

2.1.1. Diplomatic edition

1 [58r1]jīvantī madhūkaṃ vyāṣa śāravī : śīra caṃdanaṃ
padmaka ttraphalā dārvī sthirā sāśūmattī stha[58r2]ttā :
2 brrahattī dhāvanīś tiva śvaidaṣṭrrā sapūnarṇavā :
dva midä dva ca kākauṭī jīvakaṃ raṣabhakāṃ sthathā :
3 ātmagūpta phalā rāsnā naṃdaṃ pa[58r3]dmakesarä :
ttālīsapattraṃ maṃjaiṣṭhāṃ : ◯ drrāṃkṣā nīlopala balā .
4 ettāṃne sama bhāgauna kāraṣikāṃna samāharit、 :
kalka sucū[58r4]rṇaṃtta paṣṭvā : ghrratta prrasthaṃ vapācaye◯t、 :
5 cattūragūṇi nayisy āśanä mr̥ttranāgninā paṃcit、
nāśaye : ttamira kāca paṭalaṃ : arbudāṃnāṃ ca
6 jāttyaṃdhaunam aṇava paśr̥nte . ttena sarpaṣā
ardhabhäda adhimaṃtha karṇa-śūlaṃ . ca nāśaye :
7 ye urdhaṃjattrūṇā [58v1] raugā prrattaṃśaye grraha
mittyāstaṃbha śaṃraṃ śūlaṃ palamātträ śāṃmitte :
8 dva palaina mahābāga laghū vyādha vyahaṇittä
bālanāṃ atha vr̥aidhāṃ[58v2]nāṃm itte syād ämr̥ttaupamaṃ :
Expl. mahāvedīhaṃ ghrrattä :

1a jīvantī H DT] *jīvattī* B. **1d sāśūmattī sthattā** B DT] *sāśūmattīs thattā* H. **2a dhāvanīś tiva** B DT] *dhāvanī ściva* H. **2c midä** B DT] *midi* H. **2d jīvakaṃ** H B^2 DT] *jīvaka-* B^1; **raṣabhakāṃ sthathā** B DT] *raṣabhakāṃs thathā* H. **3b naṃdaṃ** H DT] *na‹la›daṃ* (without anusvāra) B. **3b padmakesarä** B DT] *padmakesari* H. **3c ttālīsapattraṃ** H B^2 DT] *ttālīsapattra* B^1; **maṃjaiṣṭhāṃ** H DT] *maṃjaiṣṭhā* B. **3d drrāṃkṣā** H B^2 DT] *drrākṣā* B^1. **4a ettāṃne** H B^2 DT] *ettāne* B^1. **4b kāraṣikāṃna** H B^2 DT] *kāraṣikāna* B^1; **samāharit、** DT, samāharit B] *samāharitaṃ* H. **4c sucūrṇaṃtta** H DT] *sucūrṇatta* B. **4d vapācaye t、** DT, vapācayet H] *vapācayeta* B^1, *vapācayetą* B^2; the akṣara <ye> is partially blurred by ink stains. **5ab cattūragūṇi nayisy āśanä** DT] *cattūragūṇi nayisyā śani* H, *cattūragūṇäna yisyā‹ṃ›śa nä* B; the akṣara <śa> has a peculiar shape and is made out of another sign; possibly <ca> corrected to <śa>. **5b mr̥ttranāgninā** H DT] *mr̥ttranāttinā* B; final <nā> has been added by the scribe below the following <pa>. A caret above <gni> marks the place of insertion. **5b paṃcit、** H B^2 DT] *pacit* B^1; a reading *paṃcät、* is also possible. **5d arbudāṃnāṃ** H B^2 DT] *arbūdānāṃ* B^1. **6ab paśr̥nte** DT] *paśr̥tte* H B; it seems that the scribe first wrote *paśa* and then modified <śa> in <śr̥>. **6c ardhabhäda** B DT] *ardhabhida* H; **adhimaṃtha** H B^2 DT] *adhīmaṃtha* B^1. **7a urdhaṃjattrūṇā** H B] <ṇā> appears to have been written in a more cursive way than usual. **7b prrattaṃśaye** H B^2 DT] *prrattaśaye* B^1. **7c mittyāstaṃbha** B DT] *minnyāstaṃbha* H; **śaṃraṃ** H B^2 DT] *śaraṃ* B^1. **7d palamātträ** B] *palamāttri* H; **śāṃmitte** H B^2 DT] śāmitte B^1. **8a mahābāga** H DT] *mahābhāga* B. **8b vyahaṇittä** B DT] *vyahaṇitti* H. **8c vr̥aidhāṃnāṃm** H DT] *vr̥aidhānām* B^1, *vr̥aidhāṃnām* B^2. **8d ämr̥ttaupamaṃ** H DT] *ämrrattaupamaṃ* B. **Expl**. **mahāvedīhaṃ** H B^2 DT] mahāvedīha B^1.

2.1.2. Critical edition

1 jīvantī madhukaṃ vyoṣaṃ śāriv⁺ośīra-candanam |
padmakaṃ triphalā dārvī sthirā sāṃśumatī {s}tathā ||
2 bṛhatī-dhāvanīś caiva śvadaṃṣṭrā sa-punarnavā |
dve mede dve ca kākolyau jīvakarṣabhakau {s}tathā ||
3 ātmaguptā phalā rāsnā na‹la›daṃ padmakesaram |
tālīsapatra-mañjiṣṭhā drākṣā nīlo‹t›palaṃ balā ||
4 etāni sama-bhāgāni kārṣikāṇi samāharet |
kalkaṃ su-cūrṇitaṃ piṣṭvā ghṛta-prasthaṃ vipācayet ||
5 catur-guṇe ⁺payasy aṃśaṃ ‹śa›nai‹r› ⁺mṛdva{nā}gninā pacet |
nāśayet timiraṃ kācaṃ paṭalaṃ arbudāni ca ||
6 jātyandhānām a‹kṣī›ṇ⁺īva paś⁺yante tena sarpiṣā |
ardhabhedam adhimanthaṃ karṇa-śūlaṃ ca nāśayet ||
7 ye ūrdhva-jatru⁺jāḥ rogāḥ pratiśyāya-‹gala›grahaḥ |
manyāstambhaḥ śiraḥśūlaḥ pala-mātre‹ṇa› śāmyate ||
8 dvi-palena mahābhāgaḥ laghu vyādhir vihanyate |
bālānām atha vṛddhānām iti syād amṛtopamam ||
|| mahāvaideha-ghṛtam ||

Metrics:

1ab	– – – ⏑ ¦ ⏑ – – – ‖ – ⏑ – – ¦ ⏑ – ⏑ –	1cd	– ⏑ – ⏑ ¦ ⏑ – – – ‖ ⏑ – – ⏑ ¦ ⏑ – ⏑ –
2ab	⏑ ⏑ – – ¦ ⏑ – – ⏑ ‖ ⏑ – – ⏑ ¦ ⏑ – ⏑ –	2cd	– – – – ¦ ⏑ – – – ‖ – ⏑ – ⏑ ¦ ⏑ – ⏑ –
3ab	– ⏑ – – ¦ ⏑ – – – ‖ ⏑ ⏑ – – ¦ ⏑ – ⏑ –	3cd	– – ⏑ – ¦ ⏑ – – – ‖ – – – – ¦ ⏑ – ⏑ –
4ab	– – ⏑ ⏑ ¦ ⏑ – – ⏑ ‖ – ⏑ – ⏑ ¦ ⏑ – ⏑ –	4cd	– – ⏑ – ¦ ⏑ – – – ‖ ⏑ – – – ¦ ⏑ – ⏑ –
5ab	⏑ – ⏑ – ¦ ⏑ – – – ‖ ⏑ – – – ¦ ⏑ – ⏑ –	5cd	– ⏑ – ⏑ ¦ ⏑ – – – ‖ ⏑ ⏑ – – ¦ ⏑ – ⏑ ⏑
6ab	– – – – ¦ – – – ⏑ ‖ – – – – ¦ ⏑ – ⏑ –[9]	6cd	– ⏑ – – ¦ ⏑ – – – ‖ – ⏑ – – ¦ ⏑ – ⏑ –
7ab	– – ⏑ – ¦ ⏑ – – – ‖ ⏑ – – ⏑ ¦ ⏑ – ⏑ –	7cd	– – – – ¦ ⏑ – – – ‖ ⏑ ⏑ – – ¦ ⏑ – ⏑ –
8ab	⏑ ⏑ – ⏑ ¦ ⏑ – – – ‖ ⏑ ⏑ – – ¦ ⏑ – ⏑ –	8cd	– – – ⏑ ¦ ⏑ – – – ‖ ⏑ ⏑ – ⏑ ¦ ⏑ – ⏑ –

1a madhukaṃ H DT for madhūkaṃ] *madhuko* M, *madhūkaṃ* C; **vyoṣaṃ** H M DT for vyāṣa] *vyoṣa-* C. **1b śārivośīra** H M DT for śāravī : śīra] *-śāravā uśīra* C. **1c triphalā** H C DT for ttraphalā] M *triphalaṃ*. **1c dārvī** H M DT for dārvī] *dāru-* C. **1d sthirā** H M DT] *sthirāsa* C; **tathā** H M C DT for stathā]. **2a bṛhatī** H M C DT for brrahattī]; **-dhāvanīś caiva** H C DT for dhāvanīś tiva] *dhāvanī caiva* M. **2b śvadaṃṣṭrā** H M C DT for śvaidaṣṭrrā]; **sa-punarnavā** H M C DT for sa-pūnarṇavā]. **2c dve mede** H M DT for dva midä] *dvi-medā* C; **dve ca kākolyau** DT for dva ca kākauṭī] *dve ca kākauṭe* H, *dve ca kākoḍyau* M, *dvi-kākolī* C. **2d jīvakarṣabhakau** M C DT for jīvakaṃ raṣabhakāṃ] *jīvakarṣabhakās* H; **tathā** H M C DT for stathā]. **3a ātmaguptā** H M DT for ātmagūpta] *ātmagūptā-* C; **phalā** H DT] *bālaṃ* M, *-phalā* C. **3b naladaṃ** B M C

[9] The 2nd colon of pāda a is hypermetrical.

DT] *nandanaṃ* H; **padmakesaram** C DT for padmakesarä] *padmakeśarī* H, *padmakesaraḥ* M. **3c tālīsapatra-** H C DT for ttālīspattraṃ] *tālīsapattraṃ* M; **(-)mañjiṣṭhā** H M C DT for maṃjaiṣṭhāṃ]. **3d drākṣā** H M C DT for drrāṃkṣā]; **nīlotpalaṃ** H M DT for nīlopala] *nīlotpala-* C. **4a etāni** H C DT for ettāṃne]; **sama-bhāgāni** H C DT for sama bhāgauna]. **4b kārṣikāṇi** H C DT for kāraṣikāṃna]; **samāharet** H C DT for samāharit˽]. **4c su-cūrṇitaṃ** H DT for sucūrṇaṃtta] *sa-curṇatā* C; **piṣṭvā** H C DT for paṣṭvā]. **4d ghṛta-prasthaṃ** C DT for ghrratta prrasthaṃ] *ghṛta-prasthe* H; **vipācayet** H C DT for vapācayet˽]. **5ab catur-guṇe payasy aṃśaṃ śanair** DT for cattūragūṇi nayisy āśa nä] *catur-guṇena payasā śanair* H, *catur-guṇena payasā ca* C. **5b mṛdvagninā** DT for mṛttranāgninā] *mṛdunāgninā* H, *mūtrena agninā* C; **pacet** H C DT for paṃcit˽]. **5c nāśayet** H C DT for nāśaye]; **timiraṃ** H C DT for ttamira]. **5d arbudāni** H DT for arbudāṃnāṃ] *arbudaṃ* C. **6a jātyandhānām akṣīṇīva** DT for jāttyaṃdhaunam aṇava] *jātyandho namaṇava* H, *jātyandhāni māṇavaḥ* C. **6b paśyante** DT for paśṛnte] *paśyate* H C; **ttena** H C DT for tena]; **sarpiṣā** H C DT for sarpaṣā]. **6c ardhabhedam** H DT for ardhabhäda] *ardhāvabhedaka-* C; **adhimanthaṃ** DT for adhimaṃtha] *adhīmanthaṃ* H, *adimantha* C. **6d nāśayet** H C DT for nāśaye]. **7a ūrdhvajatrujāḥ** DT for urdhaṃjattrūṇā] *ūrdhvajatruṇi* H, *ūrdhvajatrūṇāṃ* C; **rogāḥ** H DT for raugā] *rogā* C. **7b pratiśyāya-** for prrattaṃśaye] *pratiśyāyo* H, *pranāśayet* C; **‹gala›grahaḥ** B H C DT for grraha]. **7c manyāstambhaḥ** H DT for mittyāstaṃbha] *iti āstambha-* C; **śiraḥśūlaḥ** DT for śaṃraṃ śūlaṃ] *śiraḥśūlaṃ* H, *śirośūlaṃ* C. **7d pala-mātreṇa** H C DT for palamātträ]; **śāmyate** H C DT for śāṃmitte]. **8a dvipalena** H C DT for dva palaina]; **mahābhāgaḥ** DT for mahābāga] *mahābādho* H, *mahābhāgaṃ* C. **6b laghu** H C DT for laghū]; **vyādhir** H DT for vyādha] *vyādhayaḥ* C; **vyahaṇittä** DT for vihanyate] *va hanyate* H, *hanet* C. **6c vṛddhānām** H C DT for vṛaidhāṃnāṃm]. **6d iti** H C DT for itte]; **amṛtopamam** H C DT for ämṛttaupamaṃ].

2.1.3. Translation

1. (1) jīvantī, (2) liquorice, (3) the (three) hot ones [i.e. long pepper, black pepper, ginger], (4) sarsaparilla, (5) vetiver, (6) sandal, (7) bird cherry, (8) the three fruits, (9) barberry, and so also (10) tick trefoil with (11) uraria and as much **2.** (12) Indian nightshade and (13) wild eggplant, (14) caltrop with (15) hogweed, (16) the two medās and (17) the two kākolīs, as (18) jīvaka and (19) ṛṣabhaka **3.** (20) cowhage, (21) Bengal quince, (22) groundsel root, (23) Indian spikenard, (24) the filament of the lotus, (25) leaf of silver fir, (26) Indian madder, (27) grape, (28) blue water lily, (and) (29) sida root.

4. (The physician) should combine these (ingredients) into equal portions of one *kārṣa*. After having pounded a well-ground paste, one should cook one *prastha* of ghee. **5.** One should cook with gentle fire (this) portion (of ghee) gradually in milk four times as much.

It can remove (a) timira [first stage of cataract], (b) kāca [second stage of cataract], (c) paṭala [third stage of cataract], and (d) flesh nodes (in the eye). **6.** (e) The eyes of the blind from birth likewise see, thanks to this ghee. It can remove (f) hemiplegia, (g) adhimantha (eye disease), and (h) earache. **7.** With one *pala* (i) the diseases which

(are) located at the upper part of the collar bone, (j) coryza, (k) compression of the throat, (l) stiffness on the back of the neck, (m) headache are pacified.

8. With two ounces, (the patient) has great prosperity: the disease is immediately beaten. To young and old, it can be like nectar.

Expl. The Mahāvaideha ghee.

2.2. Edition and translation of the Khotanese version (*Jīvakapustaka* 11)

The Khotanese version ranges from line 2 of the verso of folio 58 (IOL Khot 91/3) to the beginning of line 5 of the recto of folio 59 (IOL Khot 92/1). Below, I give the edition of these lines followed by a literal translation. The previous editions of the Khotanese version of JP 11 are by Hoernle (s.d.: 45–46), Konow (1941: 24), Bailey (first ed. KT[1] 1: 149; second ed. KT[2] 1: 149), and Ṭāme (2014: 50–51).[10]

2.2.1. Edition

1. [58v2]jīvaką mahābūjä papala . ttūṃgarą : mīraijsya : caṃdana śārava : sumąna śārava : uśaiʾra [58v3] caṃdä pauṣṭarä halīrai vīhīle ą̄◯malai : ysālva : sālaparṇa : ‹prraśnaparṇa :› brrahattä kąṇḍārya . drāṃgūlyai : aiśta bāva : mida {mahā} [58v4] mahāmeda : kākauṭa : kṣīrakākauṭa : ◯ jīvaka raṣabhaka . ‹ā›dmagūpti : bela . lagara bāva . svarṇagūttaryāṃ būśą̄nai : padma[58v5]kyesärä : ttālīspattra : rū̧nai : gūrä : nīlautpala : sacha bāva .

2. ṣiʾ paṇa arva drrai drrai mācāṃga viśtą̄ña u samāṃ kūṭą̄ña . **3.** gvī{.}[59r1]haʾ rūṃ dva śaiga hāle . ṣvīda śau ṣaṃgä . **4.** hatsa baiśa hą̄ña tcairai : nvā daina pāchai

5. tta būri āchā jaidä pa tcaimañāṃ kāśaʾ jaidä ttamīra jaida [59r2]tcaiña paṭarä yāṣṭi . **6.** ysāna hąna hve . tcịmañī prrahīsīda vījaiṣḍi : **7.** cū hālai kamala vī vīni : pejsa tcaiña rāhaʾ : gvąña vīna jai[59r3]da : uskātta garśä bīsā āchā : haysgye ttajsīda : garśa rāhiʾ : cū paysaṇve strīsīdä kamala-rāhiʾ : **8.** śau śau sera khā[59r4]śą̄ʾña : baiśa būḍa dva sera : khvai śaira hamāve raysgä vīra āchā jaida **9.** valakāṃ u ysāḍāṃ ṣiʾ hamä ṇi mą̄ñaṃdä :

Expl. mahāve[59r5]dī rūṃ :

1 jīvaką DT, jīvakaṃ K H] *jīvaka* B, *jīvatta* Ṭ (by tacit emendation?); **mahābūjä** K B Ṭ] *mahābūji* H; **ttūṃgarą** B DT, ttūṃgaraṃ K Ṭ] *nnūṃgaraṃ* H; mīraijsya H K B Ṭ DT]; <rai> added below <mī> with a caret marking the place of insertion; **uśaiʾra** K B Ṭ DT] *uśaira* H; **caṃdä** B Ṭ DT]

[10] In the Khotanese edition, numbers identify each sentence (or part of it) within the Khotanese translation. Ingredients are numbered for easy comparison with the English translation and the Sanskrit original. The indents in both the edition and translation present the list of ingredients, their preparation, and the diseases cured by this medicated ghee. The following conventions are used: ‹abc› = editorial supplement; [abc] = restoration; {abc} = editorial deletion; ◯ = string hole space; B = Bailey, KT[1] (1945) and KT[2] (1969); B[1] = Bailey, KT[1]; B[2] = Bailey, KT[2]; DT = Del Tomba; H = Hoernle (s.d.); K = Konow 1941; M = Maue 1990, Ṭ = Ṭāme 2014.

caṃdi K H; **pauṣṭarä** K B Ṭ DT] *pauṣṭari* H; **‹prraśnaparṇa** :› DT] not supplemented by H K B Ṭ; **brrahattä** B Ṭ DT] *brrahatti* H K; **drāṃgūlyai** K B Ṭ DT] *drrāṃgūlyai* H; **aiśta** K B Ṭ DT] *aiśca* H; **mahāmeda** H K B Ṭ DT]; {mahā} has mistakenly been written twice: at the end of v3 and at the beginning of v4; **‹ā›dmagūpti** H B DT] *dmagūpti* K, *admagūpti* Ṭ (by tacit emendation?); **kyesirä** K B Ṭ DT] *kyesiri* H; **ttālīspattra** K B Ṭ DT] *ttālīsattra* H; **rūṇai** DT, rūnai H B Ṭ] *rūmnai* K (by probable typing mistake); **gūrä** K B Ṭ DT] *gūri* H; **nīlautpala** K DT, nīlôtpala H] *nīlāätpala* B, *nīlāṃtpala* Ṭ; the akṣara <lau> is peculiarly written in the JP manuscript: it is identical to <lā> but also has either two horizontal strokes or two dots at the left part of the <l>; Bailey transliterates the akṣara <lāau>, <lāai> or <lää>. **3. ṣaṃgä** K B Ṭ DT] *ṣaṃgi* H. **5. jaidä** K B Ṭ DT] *jaidi* H; **pa** K Ṭ DT] *pa‹ṃjsa›* K B, *paṃjsa* Ṭ (by tacit emendation?); **kāśa**ʾ H K B DT] *kāśa* Ṭ (by possible typing mistake); **paṭarä** K B Ṭ DT] *papari* H; the portion of the line corresponding to the sequence *paṭa* has been erased by the scribe and written over. The akṣara <ṭa> ends with a spurious stroke. **6. ysāna** K B Ṭ DT] *ysāna-* K, *ysāta* H; **hąna** DT, haṃna H Ṭ] *-haṃna* K, *hana* B; **gvąña** DT, gvaṃña H] *gvaña* K B Ṭ; **garśä** K B Ṭ DT] *garśi* H; **paysaṇve** B Ṭ DT] *paysanve* K (by possible typing mistake); **strīsīdä** K B Ṭ DT] *paysaṇvestrī sīdi* H. **8. śau śau** H K B Ṭ DT]; this portion has been erased by the scribe and written over; **rayṣgä** K B Ṭ DT] *raysgi* H. **9. hamä** K B Ṭ DT] *hami* H (possible reading); **mąñaṃdä** B^2 DT, māṃñaṃdä K Ṭ] *māṃñaṃdi* H, *māñaṃdä* B^1. **Expl. mahāvedī** DT] *mahāveṭī* H, K, B, Ṭ.

2.2.2. Translation

1. (1) Jīvaka, (2) liquorice, (3a) long pepper, (3b) ginger (and) (3c) black pepper, (4a) candana sarsaparilla, (4b) sumana sarsaparilla, (5) vetiver, (6) sandal, (7) bird cherry, (8a) chebulic myrobolan, (8b) belleric myrobalan, (8c) emblic myrobalan, (9) barberry, (10) tick trefoil, (11) (pointed leaved uraria), (12) Indian nightshade, (13) wild eggplant, (14) caltrop, (15) hogweed root, (16a) medā (and) (16b) mahāmedā, (17a) kākolī (and) (17b) kṣīra-kākolī, (18) jīvaka, (19) ṛṣabhaka, (20) cowhage, (21) Bengal quince, (22) groundsel root, (23) the fragrant (plant) of the *Suvarṇagotrikas [a nard], (24) lotus filament, (25) leaf of silver fir, (26) Indian madder, (27) grape, (28) blue water lily, (29) sida root.

2. Three mācāngas of each of these drugs must be put in and evenly pounded. **3.** Clarified butter – two and a half *śiṃga*s; milk – one *ṣaṃga*. **4.** All together should be put into a vessel (and) should be cooked with gentle fire.

5. It thus overcomes the following diseases: (a) kāca [second stage of cataract] before the eyes, it overcomes (b) timira [first stage of cataract], (c) paṭala [third stage of cataract] (and) (d) swelling (in the eyelids). **6.** (As for) (e) a man blind from birth, his eyes will be open (and) he will see. **7.** If (there are) (f) pains in half of the head, (g) severe pains in the eye, (h) pain in the ear, it will remove (them and) (i) diseases located above the throat, (when) (j) the nostrils drip, (k) pain in the throat, (l) when the jaws become stiff, (and) (m) headache. **8.** It should be drunk one ounce at a time (or) two ounces at the utmost, so

that it should be good for one: it quickly overcomes all diseases. **9.** To young and old, it is equally like nectar.

Expl. The Mahāvaideha ghee.

2.3. Edition and translation of the Tocharian B version

The three fragments THT 500, THT 501, and THT 502 are the only remnants of an individual manuscript written on the verso of a Chinese roll. A first edition of the three fragments was by Sieg and Siegling (1953: 311). Maue (1990) has shown that THT 500 and THT 501 must be joined horizontally: THT 500 and THT 501 cover, respectively, the left and right portions of lines 4 to 12 of the surviving leaf. The third fragment, THT 502, contains part of the three previous lines. The edition below is primarily based on Malzahn (*apud* CEToM), who provided a new edition with philological and linguistic commentary.[11]

2.3.1. Edition

1 (– – – – – – – a)bhijñe **:** aśoke lānt· (– – – – – – – – – – – –)
2 (– – – – – –) – na tekanmane kartse **:** || wa – (– – – – – – – – – –)
3 (– – – – – – – – – – – – – –) [p]ai [**:**] kaṣu – (– – – – – – – – – –)
4 (– –) k̲[ä]lle eśañene yamaṣle ce kr[ā]ntsa nasait yama(ṣle **:** namaś cakṣupariśodhana)
5 (rā)jāya tathāgatāyā **:** rhate samyaksaṃbuddhā(ya – –) tadyathā cakṣujñānacakṣu
6 viśodhanāya svāhā || ñake (ma)hā ⟦vairyocaṇi⟧ (ma)hāvaidyehik̲·̣ ṣalpe weñau po
7 (e)śañeṣana ⋄ tekanmane kartse ⋄ mrañco pipāl̲ tvā(ṅkaro) jivanti pisau ⋄ wasto yä-
8 (kne) – ·eñcaṣṣe· tāno ⋄ uśir̲ ⋄ candāṃ ⋄ yāstaci – – – arirāk̲ wipitāk̲ a-
9 (malā)[k̲] ⋄ śkwiśko ⋄ sālaparṇi ⋄ pṛ́śnaparṇi ⋄ (praha)ti ⋄ gandhakāri klyotañ[ä] eśpe-
10 -ṣṣe mahāmet̲·̣ ⋄ met̲ ⋄ kākoṭi ⋄ kṣirakāko(ṭi ⋄) jivak̲ rṣabhak̲ ātmagupti ⋄
11 palāṣṣai wi̯tsako enmet[r]e ⋄ rās(n)ā peparṣṣa wi(tsako) upāl̲ṣa witsako ⋄ tālis̲
12 (pattär̲ mañcaṣṭa – – – – – – – nilotp)ā(l̲ balṣa w)i(ts)ako toyna saṃtke-
13 (-nta) ///

1. lānt· SS Ma DT]; as pointed out by Malzahn, an oblique singular *lānt* should have been spelled *lānt̲* with virāma. Possible restorations include a gen.sg. *lānt(e)* or an all.sg. *lānt(aśc).* **2. -na** SS

[11] The following conventions are used: [abc] = uncertain reading; (abc) = restoration; ⟦ ⟧ = deletion by the copyist; · = indeterminable part of an akṣara; – = akṣara not readable; /// the line starts or ends with a lacuna; DT = Del Tomba; M = Maue 1990; Ma = Malzahn (*apud* CEToM); SS = Sieg & Siegling 1953.

Ma DT]; a nom.acc.pl.f. form agreeing with *tekanmane* must have been lost in the lacuna; the preceding akṣara might be read *ya* or *ṣa*; **wa-** SS Ma DT] The following akṣara might be *te.* **3. [p]ai** DT] *·ai* SS Ma; the reading is likely, but not certain. **4. kr[ā]ntsa** DT] *krāntsa* SS Ma; a long left stroke from the upper line is joined to the upper part of the akṣara <kra>; **yama(ṣle)** SS Ma DT]; **(: namaś cakṣupariśodhana)** DT]. **5. (rā)jāya** DT] *-jāya* SS Ma; **samyaksaṃbuddhā(ya – –)** DT] samyaksaṃbuddhā(– – –) Ma. **6. (ma)hā ⟦vairyocaṇi⟧ (ma)hāvaidyehiḵ·** SS Ma DT]; the manuscript reads *(ma)hāvairyocaṇi* || *(ma)hāvaidyehiḵ·*. The copyist crossed out *vairyocaṇi* with several short horizontal strokes but mistakenly left the first *mahā* undeleted. **7 tvā(ṅkaro)** SS Ma DT]. **7–8. ya(kne)** SS Ma DT]. **8. –·eñcaṣṣe** DT] – *r·eñcaṣṣe* SS, – – *r·eñcaṣṣe* M, – – *r*[*p*]*eñcaṣṣe* Ma; after restored *yä*(*kne*) at the beginning of the line, there is space for just one akṣara. A reading *rpe* is difficult because of the horizontal distance between the *r-* in ligature to the left part of the *-p*, which, in this manuscript, is written smaller than usual. This reading would therefore yield an overly broad *p*. Also, the ink traces in the upper left part of the lacuna seem to exclude reading *p*. Possible restorations are *r(m)e*, *r·(y)e*, *(nm)e,* or *(tm)e*. Compare it with <nme> in line 11. **9. (malā)[ḵ]** M Ma DT]; **(praha)ti** M Ma DT] ///*ti* SS; **10. ātmagupti** M Ma DT] *āśmagupti* SS; the horizontal stroke linking the lower part of the *t* could have been written by accident. Ching Chao-jung (p.c.) suggests that the scribe might have initially written the first stroke of *p* in anticipation of *pti*, and then immediately modified it to *tm*; **ṛṣabhaḵ** SS M Ma DT] A reading *ṛṣabhāḵ* cannot be excluded. **11. wᵢtsako** DT (Michaël Peyrot p.c.)] *wtsako* SS M Ma; *wᵢtsa* is written as a single akṣara; **wi(tsako)** SS M Ma DT]. **12. (mañcaṣṭa)** DT (tentative restoration)]; **(nilotp)ā(l balṣa)** DT (tentative restoration)]; **(w)i(ts)ako** SS M Ma DT]. **12–13. saṃtke(nta)** SS M Ma DT].

2.3.2. Translation

[1]… (higher) knowledge. (To?) king Aśoka … [2] good for diseases of … [3] costus … [4] should be ... [and] should be made into the eyes. With this application, a spell is to be cast: *namaścakṣuḥpariśodhana* [5] *rājāya tathāgatāyārhate samyaksambuddhā*(*ya* ||) *tadyathā cakṣujñānacakṣu* [6] *viśodhanāya svāhā.*

Now I am going to explain the mahāvaidehika ghee, [7] good for all diseases of the eyes: (1a) black pepper, (1b) long pepper, (and) (1c) ginger, (2) jīvantī, (3) aneth, (4) the seed of both kinds [8] of (sarsaparillas), (5) vetiver, (6) sandal, (7) (… of the) bird cherry, (8a) chebulic myrobolan, (8b) belleric myrobalan, [9] (8c) emblic myrobalan, (9) barberry, (10) tick trefoil, (11) uraria, (12) Indian nightshade, (13) wild eggplant, (14) caltrops, [10] (15) hogweed (root), (16a) mahāmedā, (16b) medā, (17a) kākolī (and) (17b) kṣīra-kākolī, (18) jīvaka, (19) ṛṣabhaka, (20) hogweed, [11] (21) root (and) *enmetre* of the *pāla*, (22) groundsel root, (23) the root of the Indian spikenard, (24) filament of the lotus, (25) (leaf of) silver fir, [12] (26) (Indian madder), (27) (grape), (28) (blue watery lily), (29) (sida) root. These medicament[13](s) …".

3. Commentary

The commentary is organised into three main sections.

The first section (§3.1) is a philological and linguistic commentary of the Sanskrit, Khotanese, and Tocharian B versions. It provides explanatory notes on the reconstruction of the Sanskrit text (§3.1.1); a linguistic analysis of the Khotanese version and a thorough examination of the errors, spelling mistakes, and orthographic peculiarities of the Sanskrit text (§3.1.2); a philological analysis of the first lines of the Tocharian manuscript and an introduction to the linguistic stage of the Tocharian B variety in which the text has been written (§3.1.3).

The second section (§3.2) deals with the correspondence between the ingredients attested in the three versions and discusses some problematic cases in detail.

The third section (§3.3) analyses the diseases cited in the Sanskrit and Khotanese versions.

3.1. Philological and linguistic commentary

3.1.1. Explanatory notes on the reconstruction of the Sanskrit text

The numerous grammatical and spelling mistakes contained in the Sanskrit manuscript indicate that the text was likely copied by someone with a limited command of the Sanskrit language, most probably the same scribe(s) who also wrote the Khotanese sections.[12]

JP 11.1

Bailey and Chen read *jīvattī* in IOL Khot 91/3 r1 (JP 11.1). This is a possible reading as it can be assumed that *jīvattī* is miswritten for *jīvaṃttī.* However, the graphic distinction between *t* and *n*, whether single or double, is a well-known difficulty of Khotanese Brāhmī, particularly in texts written in a cursive ductus (documentary script). An argument in favour of a reading *nt* is that *jīvantī* is never spelled with the anusvāra (cf. Skt. *jīvantī* in JP 4.6[48r3] [= LKhot. *jīvanta*], 33.2[78r3] [= LKhot. *jīvanta*], 37.1[81r3] [= LKhot. *jīvanti*], 43.1[86r1] [= LKhot. *jīvanta*], 47.4[88v4] [= LKhot. *jīvanta*], 54.1[96r5] [= LKhot. *jīvanta*], 77.2[107r3] [= LKhot. *jīvata*], 93.1[116v5]). The same holds true for other Late Khotanese medical texts (cf. e.g. *jīvantä* in PiŚ 106, *jīvanta* in Si 26.60).[13]

---

[12] For a discussion of the most frequent errors and spelling mistakes, see the relevant section below (§ 3.1.2).

[13] Cf. also Skt. *dantī* 'wild croton', which, in the Sanskrit version of JP, is written both with the anusvāra (cf. JP 5.1[49v5] *daṃttī* [= LKhot. *dantä*], 12.3[59v2] *daṃttī* [= LKhot. *danta*], 63.1[100r1] *daṃttī* [= LKhot. *danti*]) and without the anusvāra (cf. JP 9.6[55v2] *dantī* [= LKhot. *dantä*], JP 35.3[79v5] *dantī*

The manuscript clearly reads *madhūkaṃ* (IOL Khot 91/2 r1), which is translated *modujia guo* 末度迦果 'fruit [of] mahua tree (*Bassia Latifolia* Roxb.)' by Chen (2005: 305; see further §3.2 below). However, *madhūkaṃ* (⏑ ¦ – –) is metrically impossible and must stand for *madhukaṃ* 'liquorice' (⏑ ¦ ⏑ –). Maue (1990: 161) reconstructs a masculine nom.sg. *madhuko*, but *madhuka-* 'liquorice' is usually rather a neuter noun (cf. acc.pl.nt. *madhukāni* in e.g. AH.Utt.40.18.1).

As pointed out by Hoernle (1917: 45 fn. 1) and Maue (1990: 161 fn. 20), *śāravī* : *śīra* (IOL Khot 91/3 r1) cannot stand for *śārive* : ‹*u*›*śīra* for metrical reasons. The double dot should therefore be omitted, and the sequence should be edited as *śārivośīra-* (*śārivā-* + *uśīra-*).

Hoernle (1917: 46 fn. 6) edits *sāśūmattī sthattā* (IOL Khot 91/3 r1–2) as *sāṃśumatīs tathā* and comments *sāṃśumatīs* as a "Vernacular Sanskrit form". Chen (2005: 305) gives a similar reconstruction without commenting on final *-s*, which is indeed unclear.

JP 11.2

The manuscript reads *dhāvanīś tiva* in IOL Khot 91/2 r3 (as per Bailey, *contra* Hoernle), which can be regarded as a hypercorrect form for °*dhāvanīś caiva*. In Late Khotanese, spelling variants in clusters containing a sibilant and a stop are indeed frequent, e.g. *hīvyauṣṭā-* < OKhot. *hävyauṣcā-* 'appropriation'; *ṇaʾstyauña-* < OKhot. **nätaʾscauña-* 'limit'. Alternations between *śc* and *śt* are also attested (cf. *sūśca* ~ *sūśta*; *harśca* ~ *harśta*; *hīścä* ~ *hīśtä*, and, in the JP manuscript, JP 7.1[52v3] *attabalāś taivaṃ* for *atibalāś caiva*, JP 9.4[55r5] *mauṣparṇīś ta* for *māṣa-parṇīś ca*, etc.)

Maue (1990: 161) has *bṛhatī dhāvanī caiva*. Hoernle (s.d.: 45 fn. 2 and 46 fn. 6) has *bṛhatī dhāvanīś caiva*, where *-ś* is interpreted as a "Vernacular Sanskrit form". I take Skt. *iva* as used correlatively to *tathā* at the end of the verse, thus 'and as much (*caiva*) Indian nightshade (*bṛhatī-*) and wild eggplant (*dhāvanīs*) ... as (*tathā*) (18) jīvaka (*jīvaka-*) and (19) ṛṣabhaka (*ṛṣabhakau*)'. Thus, *bṛhatī-dhāvanīś* may be a compound inflected as a plural.

The manuscript sequence *jīvakaṃ raṣabhakāṃ sthathā* (IOL Khot 91/3 r2) has been edited by Hoernle as a plural *jīvakarṣabhakās tathā*, which is a possible reconstruction. Maue and Chen opt for a dual form *jīvakarṣabhakau* 'jīvaka and ṛṣabhaka', as word-final *-āṃ*, *-ā* frequently interchange with *-au* in Khotanese (see §3.1.2). Cf. also *bhāgauna* in IOL Khot 91/3 r3, which must stand for *bhāgāni*.

[= LKhot. *danta*]). In other Khotanese medical texts, it is consistently written without the anusvāra (cf. Si 2.27 *dantä*, Si 13.49 *dantä*, Si 26.69 *dantą*, PiŚ 117 *dāntä*).

JP 11.3

Bailey restores Skt. *naladaṃ* for ms. *naṃdaṃ* (IOL Khot 91/3 r2). Hoernle (s.d.: 46 fn. 9) conjecturally restored *nandanaṃ*. However, in Su.Ka.2.5, this unidentified plant is described as having poisonous fruits (source of so-called *sthāvara* poisons).[14] Therefore, Bailey's restoration is to be preferred (see further §3.2 below).

JP 11.4

In the manuscript, the 3sg.opt. *vipācayet* is written with final virāma (IOL Khot 91/3 r4 *vapācayet*◌̣) and not with the anusvāra (so Bailey). In the *Jīvakapustaka* manuscript, the virāma diacritic is frequently written without accuracy: the proper shape of the mark would have been a curved line above the akṣara, but in most cases it is written so slovenly as to resemble the mark of the anusvāra.

The form *sucūrṇaṃtta* (IOL Khot 91/3 r3–4) is interpreted by Chen (2005: 306) as *sa-curṇatā,* which he translates '以粉末的狀態 [in powder form]'. Following Hoernle (s.d.: 46), I rather reconstruct the form as *su-cūrṇitaṃ* 'well-ground', an attribute of *kalkaṃ* 'paste'.

JP 11.5

The restoration of the first hemistich of JP 11.5 is mine, but it has taken into account some insights by Hoernle.

In pāda a, both Hoernle and Bailey restored ms. *yisyā* as *‹pa›yasā*. Following Bailey's reading of pāda b, Chen edits the first hemistich as *catur-guṇena payasā ca mūtrena agninā pacet*, which he translates '再(將藥劑)與 4 倍的牛奶和牛尿, 用(文)火同煎 [Then (mix the potion) with four times the cow's milk and cow urine, cook it with (gentle) fire]' (2005: 307). However, the manuscript reading is actually *mṛttranāgninā* (so Hoernle), and no *mūtrena* 'with cow urine' is to be reconstructed. The parallel passage in the Khotanese version confirms that cow urine is not involved (see below). On top of that, Chen's restoration would yield a hemistich long by one syllable.

Hoernle (s.d.: 46) provides a more reasonable restoration, which takes into consideration the actual manuscript reading: *catur-guṇena payasā śanair mṛdunāgninā pacet* 'One should cook (it) gradually with milk four times as much (the ghee) with gentle fire' (my translation). As can be seen, Hoernle interpreted ms. *śanä* as the adverb *śanais* 'quietly, softly, gently, gradually', a term attested in the description of many drug preparations in medical texts (cf. a similar formulation in Su.Ka.7.30.2 *śanair mṛdvagninā pacet*). The extra syllable in Hoernle's restoration of pāda b can be easily resolved by editing *mṛttranāgninā* as *mṛdvagninā.* Although a form *mṛdunāgninā*

[14] Cf. Meulenbeld 1999–2002: 1a.206 and 1b.457 n. 608.

(attested in e.g. Ca.Cik.1.51.1 *mṛdunāgninā* and Su.Cik.3.62.1 °*mṛdunāgninā*) would align the edited text more closely with the manuscript reading, *mṛdvagninā* ‘with gentle fire, over a slow fire’ is the more prevalent variant in medical texts. Therefore, it is assumed that *mṛttranāgninā* is a copying error for *mṛdunāgninā*. The copyist error is fully justifiable from a palaeographic perspective, because the akṣara <du> looks very similar to <ttr> in the cursive ductus of Khotanese Brāhmī. This *mṛdunāgninā* should, in turn, be regarded as a normalised form of the compound *mṛdvagninā.*

The only problem with this restoration of the hemistich is that the first half-verse of 16 syllables would take the vipulā metre (type 2a), whereas all other (unproblematic) hemistichs, including the second hemistich of the same verse, adhere to the pathyā metre: *catur-guṇena ‹pa›yas{y}ā śanai‹r› ⁺mṛdvagninā pacet* ⏑ – ⏑ – ¦ ⏑ ⏑ ⏑ – || ⏑ – – – ¦ ⏑ – ⏑ – (vipulā type 2a). To reconcile this, I have interpreted ms. *cattūragūṇi* as a loc.sg. *caturguṇe*, which would metrically correspond to the first colon. The remaining manuscript sequence *nayisyāśana* is interpreted as ⁺*payasy* ⁺*aṃśaṃ ‹śa›nai‹r›*, with correction of the first akṣara and simple haplography of the two *śa*. The resulting text reads *caturguṇe payasy aṃśaṃ śanair mṛdvagninā pacet* ‘one should gradually cook the portion (of ghee) in four times as much milk with gentle fire’, which is much closer to the Khotanese rendering *gvīhaʾ rūṃ dva śaiga hāle . ṣvīda śau ṣaṃgä . hatsa baiśa hāṇña tcairai* : *nvā daina pāchai* ‘Clarified butter – two and a half *śiṃga*s, milk – one *ṣaṃga*. All together should be put into a vessel and should be cooked with gentle fire’.

The quantities of ghee and milk prescribed in the Khotanese version are thus identical to those specified in the Sanskrit version. This correspondence is based on several other instances in the *Jīvakapustaka* and the *Siddhasāra*, where 1 Khotanese *ṣaṃga* corresponds to 1 Sanskrit *āḍhaka*, while 2½ Khotanese *śaiga*s correspond to 1 Sanskrit *prastha* (see Emmerick 1979a: 239–241). Given that a single *āḍhaka* unit is equivalent to 4 *prastha*s, the Sanskrit and Khotanese versions must refer to the same quantities of milk and ghee according to different formulations: the Sanskrit version states that 1 *prastha* of ghee should be cooked in 4 portions of milk, while the Khotanese version refers more specifically to 1 *prastha* of ghee (= 2½ *śaiga*s [= 1 *prastha*]) and 1 *āḍhaka* of ghee (= 1 *ṣaṃga* [= 4 *prastha*s]).

The second hemistich of JP 11.5 is edited by Chen (2005: 307) as *nāśayet timira-kāca-paṭalam arbudaṃ ca*, which is short by one syllable. For this reason, Hoernle edits ms. *arbudāṃnāṃ* as nt.pl. *arbudāni*. While ms. *arbudāṃnāṃ* might theoretically stand for gen.pl. *arbudānāṃ* ‘of flesh nodes’ with implied ‘disease’, such a reconstruction is unwarranted on metrical grounds. As pointed out by Maggi (2018: 254 fn. 57), confusion of the diacritics *-i* and *-ā* is not rare in some Khotanese manuscripts from Dunhuang, “where spellings with *-ā* instead of *-i* are comparatively frequent”. The main formal

difference between the two diacritics is that the former is open to the left, while the latter is open to the right.

JP 11.6

I propose to restore the first hemistich of JP 11.6. Hoernle edits *jāttyaṃdhau na ma ṇa vai* as *jātyandho namaṇava* and further regards *namaṇava* as an "irremediably corrupt reading" (s.d.: 46 fn. 10). Chen has *jāty-andhāni māṇavaḥ paśyate tena sarpisā,* which he translates '(該藥驅除)生盲(眼內腫胞). (患了眼病的)年輕人用了這種酥藥, 就能看見 [(The medicaments remove) born blind (inner tumour cells in the eye). Having use this ghee, a young person (suffering from eye disease) will be able to see]'. He apparently considers *jāty-andhāni* as an attribute of implied *arbudāni* 'tumours (in the eyes)'. However, Chen's restoration faces two issues: (1) from a metrical perspective, the first hemistich would be short by one syllable; (2) the noun *māṇava-* 'a youth, lad, youngster (esp. a young Brāhman), young person' (MW: 806) does not belong to the medical vocabulary.

Other possibilities need exploration. Theoretically, the form *maṇava* could be a misspelled nom.pl. *manavāḥ* 'human beings' (with *paśr̥nte* as *paśyante* 'they see'),[15] but the problem of the missing syllable would remain. The Khotanese version has *ysāna hąna hve . tcįmañī prrahīsīda vījaiṣḍi* '(As for) a man blind from birth, his eyes will open (and) he will see'. If Sanskrit *jātyandha-* corresponds to Khot. *ysāna hąna hve*, the main difference between the Khotanese and the Sanskrit versions is that the former explicitly refers to the 'eyes', while the latter does not. As for the syllable to be restored, the nasal retroflex *ṇ* may suggest the presence of an original *r* or *ṣ* in the word. A tentative reconstruction of the sequence *jāttyaṃdhaunama ṇava* is *jātyandhānām a‹kṣī›ṇ‹ī›va* 'likewise the eyes of someone born blind', interpreting ms. *jāttyaṃdhaunama* as a substantivised gen.pl. *jāttyaṃdhānām* and restoring the missing syllable ‹kṣī› in *a‹kṣī›ṇ⁺īva* (or *a‹kṣī›ṇ⁺y eva*). A problem with this restoration is that pāda a would have an irregular scansion in the second colon, with *akṣīṇīva* corresponding to – – – ⏑ instead of the expected ⏑ – – ×.[16] I currently do not have a solution to this problem, but it would not be an isolated case in the Sanskrit *Jīvakapustaka*. For instance, pāda a of JP 6.1,

[15] The correction of *paśr̥nte* with *paśyante* is unproblematic, as the only difference between *r̥* and *y* as the second components of ligatures is that the ascending oblique stroke, characteristic of both signs, ends with a downward flourish in the case of the *r̥*.

[16] I owe thanks to Mauro Maggi, who first suggested to me the idea that a word referring to the eyes should be reconstructed in this passage. He proposes restoring *jātyandhān‹āṃ ne›⁺trāṇ‹y e›va*, with haplography of *n … n* > *n …* and emendation of ms. *m* in *tr.* The resulting text however would equally yield a hypermetrical pāda. The reconstruction of an inflected form of *akṣi-* 'eye' is made more probable by its attestation in JP 32, where the text, though very corrupt, resembles that of JP 11: 77r5–77v1 *jāttyadhā tatti akṣa‹×××›bala-varṇaṃ-kara* (Bailey KT 1: 162).

edited by Emmerick (1994), also appears to be unmetrical: "it would be no doubt that the MS. intends this reading" (Emmerick 1994: 37).

The second hemistich of JP 11.6 has been edited by Chen (2005: 307) as *ardhāvabhedaka-adimantha-karṇa-śulaṃ ca nāśayet*, which is long by two syllables. The manuscript reading can, in fact, be largely maintained. See further §3.3 below.

JP 11.7

Verse 7 has been edited and translated by Chen (2005: 307) as follows:

> *ye ūrdhvajatrūṇāṃ rogā pranāśayet gala-grahaṃ* |
> *iti āstambha-śiro-śulaṃ pala-mātreṇa śāmyate* ||
> (該藥)可以治療上述的屬於咽喉部位的那些疾病; 服用 1 婆羅劑量的(這種藥)主治咽喉緊縮、脖子僵硬、頭痛 [(This medicine) can treat the above-mentioned diseases of the throat; take 1 *pala* of (this medicine) for throat tightening, stiff neck, headache.]

Initially, I considered his interpretation of ms. *prrattaṃśaye* (IOL Khot 93/1 v1) as the verbal form *pranāśayet* 'it may destroy, it can remove' as a plausible one. Although *ye ūrdhvajatrūṇāṃ rogāḥ* would make the text much closer to the original, it would also yield an irregular scanning in the second colon of pāda a (– – ⏑ – ¦ – – – –). In addition, the nom.pl. *ye … rogā* as the subject of the 3sg.opt. *pranāśayet* is a problem. In the second hemistich, Chen emends *grraha mittyāstaṃbha* in IOL Khot 93/1 v1 as *grahaṃ iti āstambha*°, which would make the sentence much less understandable as compared to Hoernle's reconstruction (see below). Since all the nouns referring to the various diseases seem to be inflected as nominatives, a grammatically more satisfying form would have been 3pl.caus.prs.pass. *śāmyante*, but this would yield an irregular sequence in the second colon of pāda d (– – – for expected ⏑ – ×).

While *prrattaṃśaye* may stand for a verbal form, Chen's edition does not consider the Khotanese translation, which clearly indicates an additional disease. Indeed, among the counterparts of Skt. *ye ūrdhvajatrujāḥ rogāḥ* 'diseases which are located above the collar bone' (= Khot. *uskātta garśä bīsā āchā*) and *gala-grahaḥ* 'compression of the throat' (= Khot. *garśa rāhi*ˀ), the Khotanese version also includes *haysgye ttajsīda* '(when) the nostrils drip' (see §3.3). Since Khotanese amplifications in the list of diseases do not generally occur, it is more probable that ms. *prrattaṃśaye* is the missing disease of the Sanskrit version. Hoernle provides a solution to this problem. He edits the verse as follows, correcting ms. *prrattaṃśaye* (IOL Khot 91/3 v1) to nom.sg. *pratiśyāyo* 'catarrh' (Hoernle s.d.: 46):

> *ye ūrdhvajatruṇi rogāḥ pratiśyāyo galagrahaḥ* |
> *manyāstambhaḥ śiraḥśūlaṃ pala-mātreṇa śāmyate* ||

From a formal perspective, Hoernle's correction is perfectly understandable in light of the phonology and orthographic conventions of Late Khotanese: (1) the likely phonetic merger of *rr* and *r*, especially in consonant clusters (cf. §3.1.2); (2) the frequent omission of *y* in clusters with a palatal + *y*; (3) the regular representation of the dental stop [t] as *tt* in Khotanese.

In the second hemistich, Hoernle reads the manuscript as *minnyāstaṃbha*, which can be understood as a nom.sg. *manyāstambhaḥ* describing stiffness in the sternomastoid region (see §3.3). As mentioned above, the graphic distinction between <t> and <n> is a serious difficulty of Khotanese palaeography, particularly in manuscripts with a cursive ductus. It is possible that Khotanese scribes faced this difficulty, particularly when copying texts in a foreign language. Usually, however, the copyist of the Sanskrit version of the *Jīvakapustaka* applied the Khotanese orthographic convention of using double *tt* for [t] (= Skt. *t*) and *n* for [n] (= Skt. *n*). There was therefore no reason for duplicating the *n* in *minnyāstaṃbha* (for *manyāstambhaṃ*), which could consequently be read *mittyāstaṃbha* (as per Bailey).

Despite this, Sanskrit words written with double *tt* in place of Skt. *n* do occur in this manuscript. Examples include: *laipattä* for *lepana* 'pasting' (IOL Khot 87/3 r4, fol. 46, JP 2–3.13), *ttasya* for *nasya* 'errhine' (IOL Khot 88/1 r3, fol. 47, JP 2–3.18), *svastakąttāṃma* for *svastika-nāma* 'Svastika by name' (IOL Khot 88/2 v3, fol. 47, JP 2–3.21), *dvikarṣakātti* or *dvikarṣakāttä* for *dvi-karṣakāni* (IOL Khot 101/2 v5, fol. 83, JP 41.6 [= Ca.Cik.11.40]). These are probably Sanskrit words with which the Khotanese scribe was unfamiliar. Such errors are also present in the Khotanese version (cf. e.g. *uystta* [IOL Khot 94/1 v3, fol. 65, JP 19] for *uysna*, *agttamatha* [IOL Khot 94/3, 67v2, JP 22; IOL Khot 95/1, 68r3, JP 23] for *agnamatha*, etc.).

However, Hoernle's edition encounters two notable issues: (1) the 3sg.caus.prs.pass. *śāmyate* does not fit the sentence grammatically, as Hoernle interprets the various nominal forms in the verse as nominatives (nom.pl. *rogāḥ*, nom.sg. *pratiśyāyo*, nom.sg. *galagrahaḥ*, nom.sg. *manyāstambhaḥ*, nom.sg. *śiraḥśūlaṃ*); (2) the second colon of pāda a would be metrically irregular (*ye ūrdhvajatruṇi rogāḥ* – – ⏑ – ¦ ⏑ ⏑ – –).

Regarding the first issue, one can indeed consider retaining either a 3sg.caus.pass. *śāmyate* or a 3sg.prs. *śāmyati*, albeit at the expense of reconstructing an irregular agreement between the conjoined subjects of the sentence and the verbal form. Such irregularities are occasionally attested in other Āyurvedic texts *metri causa*, cf. e.g. Ca.Sū.5.59.2–60.1 *manyāstambhaḥ śiraḥśūlam arditaṃ hanusaṃgrahaḥ pīnasārdhāvabhedau ca śiraḥkampaś ca śāmyati* 'stiffness on the back of the neck (*manyāstambhaḥ*), headache (*śiraḥśūla-*), facial paralysis (*arditaṃ*), lockjaw (*hanusaṃgrahaḥ*), and cold

(*pīnasa-*), hemiplegia (*ardhāvabhedau*), and trembling of the head (*śiraḥkampaś*) are alleviated (*śāmyati* for *śāmyanti*)'.[17]

Regarding the second issue, to maintain the regular scansion – – ◡ – ¦ ◡ – – – in pāda a, I suggest correcting the text to *ye ūrdhvajatru⁺jāḥ rogāḥ* 'the diseases which (are) located at the upper part of the collar bone' (cf. Su.Cik.40.54.1 *nasyena rogāḥ śāmyanti narāṇām ūrdhvajatrujāḥ* 'with (this) errhine the diseases which are located at the upper part of the collar bone of men are alleviated').

JP 11.8

As for the final verse, Hoernle interprets the first hemistich as *dvipalena mahābādho laghu-vyādhir va hanyate* 'with two ounces, a disease causing great pain is beaten like a light disease' with the following note: "this śloka is restored conjecturally; the ms. text is excessively corrupt" (1917: 46 fn. 12). Chen has *dvi-palena mahābhāgaṃ laghu vyādhayaḥ hanet* '(服用)2 婆羅的最大份量, (患者變得)清爽, 消除(所有的)疾病 [The maximum amount of 2 *pala*s (is to be taken), (the patient becomes) refreshed, (it) eliminates (all the) diseases]', which makes pāda b short by one syllable. I propose to have two sentences in the first hemistich: (1) *dvi-palena mahābhāgaḥ* 'with two ounces (the patient) has great prosperity', corresponding to Khot. *baiśa būḍa dva sera : khvai śaira hamāve* 'two ounces at the utmost, so that it should be good for one (i.e. patient)', and (2) *laghu vyādhir vihanyate* 'the disease is quickly beaten', corresponding to Khot. *raysgä vīra āchā jaida* 'it quickly overcomes all diseases'.

The final hemistich of JP 11.8 poses no relevant problem of interpretation.

3.1.2. Linguistic analysis of the Khotanese text and the "barbarous" Sanskrit

The Khotanese section is written in regular Late Khotanese, with widespread ambiguity in nominal endings and frequent weakening of gender and number agreement.

In the nominal inflection, the nom.sg. and gen.dat.sg. of the *a*-declension are usually written *-a* < OKhot. *-ä*, *-i* (cf. nom.sg. *gvīhaʾ rūṃ* 'clarified butter (lit. cow oil)' [IOL Khot 91/3 v5–92/1 r1], *ṣvīda* 'milk' [IOL Khot 92/1 r1], *baiśa* 'all, everything' [r1], *khāśąʾña* 'to be drunk' [r3–4], gen.dat.sg. *kamala* in *hālai kamala vī* 'at half of the head' [r2], but cf. also nom.sg. *rāhiʾ* 'disease' [r3]), just as the nom.acc.pl. *-a* < OKhot. *-a* (*pejsa rāhaʾ* 'severe disease' [r2]). We also find one instance of a zero ending -Ø in

[17] As an alternative, one might consider editing ms. *śāṃmitte* as a 3sg.opt.caus. *śāmayet* and interpreting all preceding nouns as accusative forms (*ye ūrdhvajatrujāḥ rogān pratiśyāya-galagrahaṃ | manyā-stambhaṃ śiraḥ-śūlaṃ pala-mātreṇa śāmayet* 'With one ounce (the ghee) should put an end to the diseases which (are) located at the upper part of the collar bone, coryza, compression of the throat, stiffness of the neck, headache'). However, the 3sg.opt.caus. *śāmayet* does not align with the manuscript reading.

nom.sg. *rūṃ* ‘oil’ [r1] < *rūnä*.[18] The locative singular regularly shows palatalisation of the stem before the ending *-ʲa*, *-ʲä* (*hāña* [r1] to *hāna-* ‘vessel’, *garśä* [r3] to *garsa-* ‘throat’), but we also find one instance of the reanalysed ending *-ña* in *gvaña* ‘in the ear’ to *ggua-* ‘ear’. The gen.dat.pl. *-āṃ* regularly continues *-ānu* (cf. *valakāṃ u ysāḍāṃ* ‘to young and to old’ [r4]). In the feminine *ā*-declension, the nom.sg. is *-a* (*paṇa arva* ‘each drug’ [v5], *viśtāña* ‘to be put, placed’ [v5], *kūṭāña* ‘to be pounded’ [v5], *vīna* ‘pain’ [r2]), and the nom.acc.pl. is *-i* (*vīni* ‘pains’ [r2]) or *-(y)e* (*haysgye* ‘nostrils’ [r3]). The nom.acc.pl. *-e* is also attested in *paysaṇve* ‘jaw, cheek, side of the neck’ [r3] to *paysaṇuā-* (see §3.3).[19]

As for masculine *aa*-declension nominals, the nom.sg. and gen.dat.sg. are both spelled *-ai* and the nom.acc.pl.m. is regularly written *-ā* (*āchā* ‘diseases’ [r1, r3, r4], *bīsā* ‘located’ [r3]). The inflected forms of *n*-stem nouns are also expected (loc.sg. *tcaiña* ‘in the eye’ [r2], nom.acc.pl. *tcịmañī* [+ °*ī* pronoun] ‘his eyes’ [r2], gen.dat.pl. *tcaimañāṃ* ‘of the eyes’ [r1]).

Alternations between *-ä*, *-i*, *-a, -e* in word-final position are frequently attested in verbal forms, cf. 3sg.prs.act. *jaidä* [r1], *jaida* [r2] for expected *jiṃdä* < OKhot. *jändä* (SGS: 35), 3sg.prs.mid. *vījaiṣḍi* [r2] for expected *vījsäṣḍe* < OKhot. *vaj(s)iṣḍe* (SGS: 117), 3pl.prs.act. *ttajsīda* [r3] for expected *ttajsīṃdä* (SGS: 38), *prrahīsīda* [r2] for expected *prrahīsīṃdä* (SGS: 89), etc.

The noun phrase *ṣiʾ paṇa arva* displays incorrect gender agreement, since *ṣiʾ* is inflected as nom.sg.m., but *paṇa arva* ‘each medicament’ is nom.sg.f.[20] Also, in the noun phrase *nvā daina* ‘with slow fire’, the marker of the instr.abl.sg. *-na* is only on the noun, while the adjective *nvā* ‘slow, weak’ seems uninflected (< OKhot. *nuvāta-*), indicating a case of group inflection for the expected instr.abl.sg. *nvāna*.

As previously mentioned, the Sanskrit text appears to have been written down by a Late Khotanese speaker, who predominantly adhered to the orthographic system of Late Khotanese. Consequently, several spelling mistakes have been introduced into the

[18] As already shown by Dresden 1955, LKhot. *-ä* tends to be dropped word-finally after the nasals *n* and *m*.

[19] Other attestations confirm that we are dealing with a feminine noun, cf. *syaṃca paysanva* ‘left.F cheek’ (IOL Khot 37/14 r2, *Cat.* 260, KT 3: 130).

[20] Agreement with the feminine noun *arvā-* ‘medicine, medicament’ varies. Taking into account the first forty-five recipes, it unambiguously takes feminine agreement nineteen times (cf. *ṣāʾ arva* in JP 5[50v2], 5[51r1], 8[54v4–5], 9[56v1], 14[62r2], 15[52v2], 16[63r3], 19[65r3–4], 20[66r5], 21[67r1], 26[72v2], 30[75v3], 33[78v4], 34[79r5], 34[79v2], 36[80v5], 39[82v3], 41[84r5], 43[r3–4]), and masculine agreement twelve times (cf. *ṣiʾ arva* in 7[53v4], 9[56v1], 11[58v5], 16[63r4], 18[64v3], 27[73r5], 32[77v2], 32[77v3], 33[78v4], 39[82v3], 42[85v2], 42[85v3]). The spelling *ṣaʾ arva* is more ambiguous (JP 3[45v2], 8[54v5], 10[57r5], 14[62r2–3], 20[66v1], 23[68v4], 24[69v2], 25[72r2], 35[80r3], 44[v2], 44[v3]).

Sanskrit text. Examples of such mistakes, categorised by different domains, include the following:

1. spellings with *-ā-* for *-o-* (*vyāṣa* for *vyoṣa* [JP 11.1]) or with *-au-* for *-ā-* (*bhāgauna* for *bhāgāni* [JP 11.4], *jāttyaṃdhaunama* for *jātyandhānām* [JP 11.6]);
2. spellings with *-ū-* for *-u-* or with *-u-* for *-ū-* (*madhūkaṃ* for *madhukaṃ* [JP 11.1], *sāśūmattī* for *sāṃśumatī* [JP 11.1], *sapūnarṇavā* for *sapunarnavā* [JP 11.2], *ātmagūpta* for *ātmaguptā* [JP 11.3], *cattūragūṇi* for *caturguṇe* [JP 11.5], *urdhaṃjattrū°* for *ūrdhvajatru°* [JP 11.7], *laghū* for *laghu* [JP 11.8]);
3. spellings with *-āṃ* for *-au* (cf. *raṣabhakāṃ* for °*rṣabhakau* ? [JP 11.2]);
4. alternations between *au* and *o* (*kākauṭī* for *kākolyau* [JP 11.2], *raugā* for *rogān* [JP 11.7], *ämṛttaupamaṃ* for *amṛtopamam* [JP 11.8]), and between *ai*, *e* and *i* (*maṃjaiṣṭhāṃ* for *mañjiṣṭhā* [JP 11.3], *samāharit* for *samāharet* [JP 11.4], *cattūragūṇi* for *catūrgūṇe* [JP 11.5], *paṃcit* for *pacet* [JP 11.5], *palaina* for *palena* [JP 11.8], *itte* for *iti* [11.8]);
5. ubiquitous use of *tt* for *t* (cf. e.g. *ttraphalā* for *triphalā* [JP 11.1], *sāśūmattī* for *sāṃśumatī* [JP 11.1], *ettāṃne* for *etāni* [JP 11.4], *cattūragūṇi* for *caturguṇe* [JP 11.5], *ttena* for *tena* [JP 11.6], *itte* for *iti* [JP 11.8]);
6. widespread use of unetymological anusvāras (cf. e.g. *drrāṃkṣā* for *drākṣā* [JP 11.3], *ettāṃne* for *etāni* [JP 11.4], *sucūrṇaṃtta* for *sucūrṇitaṃ* [JP 11.4]);
7. use of *ä* to represent several vowels, mostly in final position (e.g. *midä* for *mede* [JP 11.2], *padmakesarä* for *padmakesaraṃ* [JP 11.3], *śanä* for *śanair* [JP 11.5], *ardhabhäda* for *ardhabheda* [JP 11.6], *palamātträ* for *palamāttre°* [JP 11.7], *ämṛttaupamaṃ* for *amṛtopamam* [JP 11.8]);
8. ubiquitous use of *-rr-* for *-r-*, particularly in clusters (cf. e.g. *brrahattī* for *bṛhatī* [JP 11.2], *śvaidaṣṭrrā* for *śvadaṃṣṭrā* [JP 11.2], *drrāṃkṣā* for *drākṣā* [JP 11.3], *prrasthaṃ* for *prasthaṃ* [JP 11.4], *grraha* for *graha* [JP 11.7]);
9. peculiar treatment of syllabic *ṛ*, which is usually changed to *r* with some supporting vowels, or which has been retained as *ṛ* alongside the supporting vowel (cf. *brrahattī* for *bṛhatī* [JP 11.2], *ghrratta* for *ghṛta* [JP 11.4], *vṛaidhāṃnāṃm* for *vṛddhānām* [JP 11.8], *ghrrattä* for *ghṛtam* [JP 11.8]);
10. deaspiration of stops (as in *mahābāga* for *mahābhāga* [JP 11.8], cf. also *gaura* for *ghora* in JP 2–3.15, *vardanaṃ* for *vardhanaṃ* in JP 4.10, and *būtvā* for *bhūtvā* in JP 2–3.7);
11. interchange of *ṇ* and *n* (cf. °*pūnarṇavā* for °*punarnavā* [JP 11.2], *kāraṣikāṃna* for *kārṣikāṇi* [JP 11.4], *haṇittä* for *hanyate* [JP 11.8]);
12. apparent metatheses (cf. *thattā* for *tathā* [JP 11.1]) and assimilations (cf. *thathā* for *tathā* [JP 11.2]);

13. omission of akṣaras, as in *śīra* for ‹*u*›*śira* [JP 11.1], *naṃdaṃ* for *na*‹*la*›*daṃ* [JP 11.3], *palamātträ* for *palamātre*‹*ṇa*› [JP 11.7]);
14. akṣaras are very frequently written without vowel diacritic, e.g. *a* for *i* (*bhāgauna* for *bhāgauni*, *sucūrṇaṃtta* for *su-cūrṇita*, *ttamira* for *timira*, *sarpaṣā* for *sarpiṣā*), *a* for *ā* (*bālanāṃ* for *bālānām*);
15. omission of final *-t* (cf. *nāśaye* for *nāśayet* [JP 11.5 and 11.6]).

Categories (1) to (11) can easily be interpreted within the framework of the phonological and orthographic systems of Late Khotanese. This demonstrates that the unconventional way of writing Sanskrit reflects an attempt by the Late Khotanese scribe to adapt Sanskrit words to his own vowel and consonant systems.

Categories (5), (6), and (7) are, to some extent, due to the spelling habits of (Late) Khotanese. Thus, (5) double *tt* stands for [t] in the bulk of Khotanese texts;[21] (6) the frequent use of so-called “unetymological anusvāras” is a well-known characteristic of many Late Khotanese manuscripts; (7) the *ä*-diacritic is a peculiarity of Turkestan Brāhmī scripts (Khotanese and Tocharian Brāhmī).[22] As far as (8) the use of *-rr-* is concerned, it falls on the border between writing conventions and phonology. While the exact phonetic nature of *rr* and *r* is still debated, they represented two different phonemes in Old Khotanese (Emmerick 2009, Hitch 2016: 37–39, Skjærvø 2022: 122–123). Conversely, they became largely interchangeable in Late Khotanese, especially word-initially and between vowels (Maggi 2024: 165 §30.3.2). In consonant clusters, the two signs alternated in Old Khotanese, but in Late Khotanese, we notice a tendency to prefer *-rr-* in clusters, except when the first consonant is *tt* [t].

As for the vowel system, the transition from Old to Late Khotanese saw the loss of vowel length as a distinct phonemic feature. This led to several mergers and variations in the accepted spellings for vowel notation.[23]

---

[21] For the history of the four orthographic systems known to have been used for Khotanese, see Maggi 2021. Among these, only the archaic orthography, used for less than a dozen of documents, does not use double *tt* for [t] in initial position and between vowels.

[22] In Late Khotanese texts written with cursive ductus, the *ä*-diacritic is not written with two dots over the akṣara, but rather with a curved line above it. It largely resembles the *i*-diacritic, but it is distinguished by the fact that the vowel stroke does not touch the consonant sign.

[23] A comprehensive linguistic account of the Late Khotanese phonological system is still a *desideratum*, from both a synchronic and a diachronic perspective. The vowel system of Late Khotanese has been variously described by Ronald E. Emmerick. In 1979b Emmerick posited a vowel system with three phonemes (/a, ɛ, ɔ/) and a number of contextual allophones. In 1989, he added /ə/ as a distinct phoneme, and, on typological grounds, described “the Late Khotanese vowel phonemes as /i a u ə/ since such a small phoneme inventory necessarily means that a wider range of phonetic realisations of the phonemes is possible” (p. 209). Also, he did not exclude the possibility of a further phonemic distinction between two front vowels /i/ and /ɛ/ in the Late Khotanese *Rāma Story*. In 1993 he established a new system

Interchangeable spellings involving *ā*, *o*, and *au* to represent /ɔ/ (1, 4) are commonly encountered in Late Khotanese texts. This /ɔ/ stems from both OKhot. /oː/ (written *o* and *au* already in Old Khotanese) and /aː/ (written OKhot. *ā*): cf. OKhot. *nāma* /naːma/ 'name' > LKhot. *nāma* /nɔma/ (Sudh, mss. A, D, P), *noma* (Sudh, ms. D), and *nauma* (Sudh, mss. C, P), or OKhot. *uysnora-*, *uysnaura-* /uznoːra/ 'living being' > LKhot. *uysnaura* (Mañj 171, KBT: 121), *uysnāra* (Mañj 234, KBT: 124).[24] Similarly, *-āṃ* and *-au* ~ *-o* often alternate in word-final position (3). This is evident in the gen.dat.pl. ending OKhot. *-ānu*, which may be written as the diachronically expected variant LKhot. *-āṃ*, as well as *-au* and *-ā*. The evolution in diachronic phonology has therefore caused a morphophonological merger with the instr.abl.pl. OKhot. *-yau* (*jsa*), *-yo* (*jsa*), which can be written *-(y)au (jsa)*, *-(y)āṃ (jsa)*, *-(y)ā (jsa)* in Late Khotanese (cf. Dresden 1955: 408; Maggi 2024: 170, §30.31.2).

The loss of vowel length has meant that *u* and *ū*, and *i* and *ī* (2) may be used interchangeably to write either the long or the short phoneme of Old Khotanese; thus, the nom.sg. of the feminine *ā*-stem noun OKhot. *ūtcā-* 'water' may be written *ūtca* or *utca* in Late Khotanese (Emmerick 1979b: 246; Maggi 2024: 166–167, §30.10 and 30.11). In the *Jīvakapustaka*, the general tendency is to use *ū* to write both Skt. *u* and *ū.*

Another set of frequently interchanging in Late Khotanese is *ai*, *e*, *i* (4), especially when these vowels continue OKhot. *ai* and long *e* /eː/ (Emmerick 1979b: 243–244, Maggi 2024: 167 §30.12). As for LKhot *ä* (7), it is employed in Late Khotanese to represent various Old Khotanese vowels, notably OKhot. *ä* /ə/ (e.g. *mästa-* 'great' > LKhot. *mästa-*, *mista-*, *mesta-*, *maista-*), word-final short *-e* /e/ (e.g. the nom.acc.pl. *-e* of *ā*-stem nouns is frequently written *-ä* or *-i*), word-final short *-u* /u/ (OKhot. *pātcu* 'then' > LKhot. *pātcä*), and word-final short *-i* /i/.[25] Also, in the latest stage of the language, *-ä* and *-a* started merging in word-final position, as observed in the *Jīvakapustaka*.

Regarding spelling mistakes that affect consonants, both (10) and (11) can be easily understood in the context of the Khotanese consonant system. Indeed, Khotanese

---

only in terms of script and phonetic value, producing an eight-vowel system (Emmerick and Pulleyblank 1993): [i, ɪ, ɛ, ə, a, ɔ, ʊ, u]. Kumamoto 1995 independently arrived at a similar reconstruction: "the vowel system of Late Khotanese does not look like [a] three vowel system [...], but one which has at least three vowels both in the front series and in the back series together with at least one central vowel". Finally, in 2009, Emmerick reconstructed six vowel phonemes, introducing new phonemic oppositions for the front and back vowels of Late Khotanese: /i, e, ə, a, o, u/. Emmerick's hesitation is symptomatic of the intrinsic difficulty linked to the great variation in the writing of Late Khotanese vowels.

[24] See Dresden 1955: 406; Emmerick 1979b: 244–246; Kumamoto 1995: 387–388; Maggi 2024: 166, §30.9. The edition of the Khotanese *Sudhanāvadāna* (Sudh) follows De Chiara 2013.

[25] Cf. Emmerick 1979b: 242–244; Maggi 2024: 167, §30.13, 30.14, and 30.16.

typically used Brāhmī signs for voiced aspirated stops only in Indo-Aryan loanwords.[26] For instance, in the Old Khotanese *Book of Zambasta,* we find alternations such as *bhadra-* ~ *badra-* 'Bhadra', *irda-* ~ *irdha-* 'ṛddhi'. Finally, in Khotanese, the retroflex nasal *ṇ* [ɳ] served as an allophone of *n* /n/ in the proximity of a retroflex segment written in imitation of a similar Sanskrit practice; therefore, interchanges between *ṇ* and *n* (11) are entirely understandable within the framework of Khotanese phonetics and orthography.

The evidence discussed above, based on a close inspection of the writing mistakes in a small portion of the manuscript and limited as they may be, indicates that the Sanskrit version of the *Jīvakapustaka* owes its "barbarous" nature – as Hoernle called it – to a Late Khotanese copyist/speaker, who adapted Sanskrit words to his own writing conventions and phonological system. Furthermore, the fact that similar mistakes do occur also in the so-called *Conversation Manual* (Pelliot Chinois 5538 v.9–87),[27] a Sanskrit-Khotanese bilingual text featuring Sanskrit words and sentences followed by a Khotanese rendering, suggests the possibility that a Khotanese scribe took dictation of the *Jīvakapustaka* manuscript (cf. Emmerick 1992a: 61).

3.1.3. Philological and linguistic analysis of the Tocharian B version

The original Tocharian B manuscript, of which only the three fragments translated above have so far been identified, must have contained a list of medical prescriptions. It has been written on the back of a Chinese roll containing the *Daoxing banre jing* 道行般若經, Lokakṣema's translation of the *Aṣṭāsāhasrikā Prajñāpāramitā* (cf. Taishō 224, 8.465c11–25; see Pan & Chen 2021: 188 fn. 3).

As definitely demonstrated by recent studies (Ching & Enami 2018; Pan & Chen 2021), several Tocharian fragments of the pustaka type appear to have been created from two physical layers of Chinese rolls, glued or sewn together, and then cut into pieces to provide new blank folios for writing Tocharian texts on both sides (Ching *et al.* 2021: 78–79). The typical procedure involved producing the pustaka type format from the Chinese roll before it was used for writing the Tocharian text. Conversely, our manuscript was directly inscribed on the verso of the Chinese roll, which shows no evidence of glue

[26] Cf. Emmerick 1989: 209, 213–215, 2009: 382–383; Emmerick and Pulleyblank 1993: 31–34; Skjærvø 2022: 122–123.

[27] Edition, translation, commentary, and glossary by Kumamoto 1988. Early edition by Bailey (KT 3: 121–124). For an analysis of the writing conventions of the Sanskrit text, see Kumamoto 1990 and 1996. The main difference between the Sanskrit of the *Jīvakapustaka* and the Sanskrit of the *Conversation Manual* is that the former mostly uses long *ā* to write Skt. *ā*, *o*, *au*, while the latter use exclusively short *a* to render both Skt. *ā* and *a*. Also, in the *Conversation Manual* the *ä*-diacritic is rare (Kumamoto 1995: 387), while in the *Jīvakapustaka* it is frequently used.

on the recto. This is of interest from a codicological standpoint, as the practice of writing non-secular Tocharian texts directly on the verso of uncut Chinese rolls is relatively uncommon[28] when compared to the prevalent practice of writing e.g. Late Khotanese literary texts (e.g. Pelliot Chinois 2801 + 2781 + P 2783, containing the *Rāma Story*) and Āyurvedic texts (e.g. Pelliot Chinois 2893 + Ch. 00265, containing the so-called *Piṇḍa-śāstra*), as well as exercises, syllabaries, and secular documents on the verso of Chinese rolls.

As for the stage of the script used to write THT 500–502, it has been regarded by Malzahn (*apud* CEToM) as classical, albeit clumsier and less elegant than in standard literary manuscripts, and with a tendency towards a cursive ductus.

Another peculiarity of this manuscript is the linguistic stage of Tocharian B it is written in. Peyrot (2008: 222) regards the language of THT 500–502 as "classical Tocharian B", while Malzahn (*apud* CEToM) analyses it as "late Tocharian B". Several spelling peculiarities, colloquial forms, and apparent scribal errors can be identified. The only form indicative of a later stage of Tocharian B is the nom.acc.pl. *toyna* of the demonstrative pronoun, as this is the late variant of classical *toy* and archaic *toṃ* (Peyrot 2008: 125; Del Tomba 2018: 346, 352). All other peculiarities cannot be regarded as proper "late" forms but rather as colloquial or irregular forms, if not actual mistakes. Examples include:

1. occasional degemination, e.g. *kaṣu* (b3) (cf. also B 499 a4, b1) for regular *kaṣṣu* (PK AS 2A b3, b5; PK AS 3B b1; W 5 b2; B 497 b7); *pipāl* (b7) (cf. also W 26 b4) for regular *pippāl*; *upāl* (b11) for regular *uppāl*;
2. (apparent) depalatalisation, e.g. *ṣalpe* (b6) (cf. also B 186 a4) for regular *ṣalype*;[29]
3. (apparent) secondary palatalisation through distant assimilations,[30] e.g. *eśañene* (b4) for regular *eśanene* 'in the eyes'; *(e)śañeṣana* (b7) for regular *eśaneṣana* 'pertaining to the eyes';
4. (apparent) assimilation of vowels, *śkwiśko* /śkʷəyśko/ (b9) for *śkwaśko* /śkʷəśko/;
5. weakening of *-i-*, e.g. *wᵢtsako* (b11) for *witsako* (cf. Peyrot 2008: 132);

[28] Examples of Tocharian non-secular documents written on a roll format are: (1) THT 993–999, containing a rare example of a bilingual *Saṃyuktāgama*-text in Sanskrit and Tocharian A ("recto" side blank); (2) SHT 768.2, a scroll in Sanskrit containing one sentence in Tocharian B ("recto" side blank); (3) THT 294, containing a *kāvya*-text in late Tocharian B (Chinese on the recto); (4) THT 496, the so-called *Love Poem* in late Tocharian B. However, Ching Chao-jung (p.c.) suggests that THT 496 and THT 993–999 are too fragmentary to determine whether they were used and kept as scrolls or whether they were simply a stack of paper sheets.

[29] As pointed out by Malzahn (*apud* CEToM), a spelling mistake cannot be excluded.

[30] Cf. Peyrot 2008: 90–91, where he describes secondary palatalisation of *n* > *ñ* before *i*.

6. irregular spellings, e.g. *klyotañ* (b9) for expected *klyotaiñ*[31] (PK AS 9B b7); *rāsnā* (b11) for expected *rasna* or ≠ *rāsna*; *upāl̲ṣa* (b11), written with the Fremdzeichen *l̲* and the virāma, instead of the expected *upāläṣṣa* (cf. a similar spelling in PK AS 3A a3);
7. incorrect agreement, e.g. masculine °*ñcaṣṣe tāno* (b8) for expected feminine °*ñcaṣṣa tāno* or the oblique feminine *palāṣṣai* (b11) in agreement with the nom.sg. *wᵢtsako* (b11);[32]
8. wrong Sanskritisations, e.g. *(ma)hāvaidyehik* (b6) for *mahāvaidehik*, where *vaidye* may be interpreted as an attempt to link the word to Skt. *vaidya-* 'versed in science, medical, expert in medical science, physician' (MW: 1022); *gandhakāri* (b9) (cf. more regular *kaṇḍakāri* in PK AS 3A a4, W 17 b4) for Sanskrit *kaṇṭakārī* shows a wrong etymological connection with Skt. *gandha-* 'scent' (Maue 1990: 163 fn. 35).

All these peculiar forms cannot be easily framed within a precise "late" or "colloquial" stage of Tocharian B. This ambiguity may explain Peyrot's classification of the manuscript's stage as "classical". If we take (1) the codicological aspect of the manuscript, (2) the various spelling mistakes, (3) the phonological and grammatical errors, and (4) the fact that the manuscript is from the site of Yarxoto in the Turfan region (where only an insignificant number of manuscripts has been found), it seems possible that the composer may have been either a non-native speaker of Tocharian B or an inadequately trained Tocharian B scribe. As Hirotoshi Ogihara (p.c.) pointed out to me, the style of Brāhmī script and format used for THT 500–502 resembles that of THT 296 and 297 (Old Uyghur book), found in Xocho. This might suggest that THT 500–502 was written by an Old Uyghur scribe.

From a thematic perspective, the surviving text can be divided into two sections. The list of medicaments needed for producing the medicated ghee named *mahāvaidyehik* begins at line b6. Unlike the Sanskrit and Khotanese versions, the name of the ghee is clarified at the beginning of the prescription. The lines preceding b6 likely dealt with a different medical recipe intended for application to the eyes (*eśañene yamaṣle*). Unfortunately, only one ingredient of this medical prescription survives in the text (i.e. *kaṣu* 'costus'), as all other ingredients are lost in the lacuna. However, the text also

[31] The fact that the nom.pl. of this noun was expected to be *klyotaiñ* is shown by the derived adjective *klyotaiṣṣe** 'pertaining to caltrop' (cf. PK AS 3A a5 *klyotaiṣṣana witsakaṃ* 'roots of caltrop'). From a late text, one would rather expect a nom.pl. *klyotaiṃ*, with development of *-ñ* > *-ṃ* (see Peyrot 2008: 78–84 and Del Tomba 2023: 166–171).

[32] Cf. Malzahn (*apud* CEToM).

mentions that a spell is to be cast for the correct application of this medicated ghee. The mantra is written in proper Sanskrit as follows:

> [b4] – – – – – – – – – – [b5] – jāya tathāgatāyārhate samyaksaṃbuddhā – (–) tadyathā cakṣu-jñānacakṣu- [b6]viśodhanāya svāhā

This spell closely corresponds to the mantra found in Vṛddhavāgbhaṭa's *Aṣṭāṅga-saṅgraha*, Sū.8.60:

> *namaś cakṣuḥ-pariśodhana-rājāya tathāgatāyārhate samyaksambuddhāya tadyathā auṃ cakṣuḥ-prajñācakṣur-jñānacakṣur-vijñānacakṣur-viśodhaya svāhā* ||
> Salutation to the lord (*rājāya*) of the purification (*pariśodhana-*) of the eye (*cakṣuḥ-*), to the Tathāgata (*tathāgatāya-*), to the Arhant (*arhate*) Samyaksambuddha (*samyaksam-buddhāya*). Hail to the purification (*viśodhaya*) of the eye of flesh (*cakṣuḥ-*), the eye of prajñā (*prajñācakṣur-*), the eye of jñāna (*jñānacakṣur-*), the eye of vijñāna (*vijñāna-cakṣur-*).

The mantra closes a section of the *Aṣṭāṅgasaṅgraha* that describes the preparation of the *sarvārthasiddha*, a powerful collyrium (*añjana*) that is said to bestow luck (*dhanyaṃ*), to provide protection from sins and poison (*pāparakṣo viṣāpaham*), to enhance vision and prolong life (*cakṣuṣyam āyuṣyaṃ*), and to vanquish enemies (*śatrughnaṃ*) (AS.Sū.8.58). In section 8.59, an extensive prose passage elaborates on the preparation of this collyrium, which, as in the Tocharian version, should be applied to washed eyes (starting with the right one) and administered while muttering the mantra addressed to the Tathāgata Arhant Samyaksambuddha (Meulenbeld 1998: 1a.488).

Based on the parallel in AS.Sū.8.60, we can now restore the Sanskrit text contained in the Tocharian fragment as follows:[33]

> (*namaś cakṣu-pariśodhana-rā*)*jāya tathāgatāyārhate samyaksaṃbuddhā*(*ya* ||) *tadyathā cakṣu-jñānacakṣu-viśodhanāya svāhā*

If compared to the parallel in the *Aṣṭāṅgasaṅgraha*, the version found in the Tocharian section appears to have slightly different forms. Interestingly, the *Aṣṭāṅgasaṅgraha* cites different superior eyes/sights in the mantra, i.e. *prajñācakṣus-* 'sight of wisdom', *jñāna-cakṣus-* 'sight of knowledge', *vijñānacakṣus-* 'sight of consciousness', while the Tocharian manuscript refers only to *jñānacakṣu-* 'eye of knowledge, inner eye, intellectual vision'. Given the general tendency of this type of mantra to favour

[33] Another interesting point is that Chapter 8 of *Aṣṭāṅgasaṅgraha*'s *Sūtrasthāna* deals with the protection of kings from poison. This might explain the presence of King Aśoka at the beginning of the fragment. As far as I know, however, Aśoka is described as suffering from a disease resulting in a fecal odour coming out of his mouth and out of the pores of his hairs. It is Aśoka's son Kunāla who is cruelly blinded by his stepmother (*Divyāvadāna* 27).

amplifications, one may speculate about the relative antiquity of the version in the Tocharian manuscript compared to the *Aṣṭāṅgasaṅgraha*. Unfortunately, no comparable instances have been found in other Āyurvedic or Buddhist texts: the mantra appears to be only attested in the *Aṣṭāṅgasaṅgraha,* and it does not find any parallel in the *Aṣṭāṅgahṛdayasaṃhitā*, where chapter 7 of *Sūtrasthāna*, corresponding to chapter 8 of AS.Sū, is notably less elaborated. The presence of this mantra in a Tocharian medical fragment provides additional evidence of the influence of the Āyurvedic tradition transmitted by the works attributed to Vāgbhaṭa in Central Asia.[34]

3.2. Some remarks on the ingredients cited in the prescription

In this section I deal with some of the medicaments cited in the three prescription versions. Below I give a table showing the correspondence between the ingredients of the three versions (adapted from Maue 1990: 161–164):[35]

SANSKRIT	KHOTANESE	TOCHARIAN B
(1) *jīvantī*	(1) *jīvaką*	(2) *jivanti*
(2) *madhuka*	(2) *mahābūjä*	(3) *pisau*
(3) *vyoṣa*	(3a) *papala*	(1b) *pipāl*
	(3b) *ttūṃgarą*	(1c) *tvā(ṅkaro)*
	(3c) *mīraijsya*	(1a) *mrañco*
(4) *śārivā*	(4a) *caṃdana śārava*	(4) *wasto yä(kne) – ·eñcaṣṣe tāno*
	(4b) *sumąna śārava*	
(5) *uśīra*	(5) *uśaiʾra*	(5) *uśir*
(6) *candana*	(6) *caṃdä*	(6) *candāṃ*
(7) *padmaka*	(7) *pauṣṭarä*	(7) *yāstaci* ///
(8) *triphalā*	(8a) *halīrai*	(8a) *arirāk*
	(8b) *vīhīle*	(8b) *wipitāk*
	(8c) *ą̄malai*	(8c) *a(malā)k*
(9) *dārvī*	(9) *ysālva*	(9) *śkwiśko*
(10) *sthirā*	(10) *sālaparṇa*	(10) *sālaparṇi*
(11) *aṃśumatī*	(11) ‹*prraśnaparṇa*›	(11) *pṛśnaparṇi*
(12) *bṛhatī*	(12) *brrahattä*	(12) *(praha)ti*
(13) *dhāvanī*	(13) *kaṇḍārya*	(13) *gandhakāri*
(14) *śvadaṃṣṭrā*	(14) *drāṃgūlyai*	(14) *klyotañä*

[34] See recently Maggi 2022. Fragments of Vāgbhaṭa's *Aṣṭāṅgahṛdayasaṃhitā* are attested in Old Uyghur (Maue 2008) and Sogdian (Reck & Wilkens 2015).

[35] In the main text, medicaments are presented according to the order in the Sanskrit version; in the table below, I give the number of the corresponding ingredients in the Khotanese and Tocharian B versions according to the relative order in the text.

(15) *punarnavā*	(15) *aiśta bāva*	(15) *eśpeṣṣe*
(16) *dve mede*	(16a) *mida*	(16b) *met*
	(16b) *mahāmeda*	(16a) *mahāmet*
(17) *dve kākolyau*	(17a) *kākauṭa*	(17a) *kākoṭi*
	(17b) *kṣīrakākauṭa*	(17b) *kṣirakāko(ṭi)*
(18) *jīvaka*	(18) *jīvaka*	(18) *jivak*
(19) *ṛṣabhaka*	(19) *raṣabhaka*	(19) *rṣabhak*
(20) *ātmaguptā*	(20) *ādmagūpti*	(20) *ātmagupti*
(21) *phalā*	(21) *bela*	(21) *palāṣṣai wi̯tsako enmet(r)e*
(22) *rāsnā*	(22) *lagara bāva*	(22) *rās(n)ā*
(23) *nalada*	(23) *svarṇagūttaryāṃ būśą̄nai*	(23) *peparṣṣa wi*(*tsako*)
(24) *padmakesara*	(24) *padmakyesärä*	(24) *upālṣa witsako*
(25) *tālīsapatra*	(25) *ttālīspattra*	(25) *tālis*(*pattär*)
(26) *mañjiṣṭhā*	(26) *rū̱nai*	(26) (*mañcaṣṭa*)
(27) *drākṣā*	(27) *gūrä*	(27) lacuna
(28) *nīlotpala*	(28) *nīlautpala*	(28) *(nilotpāl)*
(29) *balā*	(29) *sacha bāva*	(29) /// *(w)i(tsa)ko*

The ingredients attested in the three versions largely correspond (5–6, 9–10, 12–15, 18–19, 22–29); several of them, however, present specific problems to be dealt with. In the Khotanese and Tocharian versions, we sometimes find explicit enumerations of single drugs grouped in Sanskrit under hyperonyms or umbrella words (3–4, 8, 16–17). Some ingredients appear to be more problematic (1–2, 7, 11, 21), either because there are mismatches between the Sanskrit version and either the Khotanese or the Tocharian version, or because the manuscript reading is difficult.

Skt. *jīvantī* (1), Khot. *jīvaką* (1), TB *jivanti* (2)

Examples of mismatched ingredients are (1) and (2). While Skt. *jīvantī*[36] clearly corresponds to TB *jivanti*, the Khotanese section opens with *jīvaką*, a different

[36] Nadkarni 1954: 1.444 identified this plant with the *Dendrobium macraei* Lindl., an orchid, while others consider it an *Asclepiadaceae* plant, the *Leptadenia reticulata* W. & A. (so Khare 2007: 370–371, Sharma 2018: 155; see further Meulenbeld 1999–2002: 2b.101). In both the *Suśrutasaṃhitā* and the *Aṣṭāṅgasaṅgraha*, *jīvantī* is said to be beneficial for eyesight and a remover of eye-diseases: cf. Su.Sū.46.252.1 *cakṣuṣyā sarvadoṣaghnī jīvantī samudāhṛtā* '*jīvantī* is illustrated as beneficial to the eyes (and) a remover of all kinds of deranged humours for the eye', AS.Sū.7.129 *cakṣuṣyā sarvadoṣaghnī jīvantī*). Skt. *jīvantī* is included in the (ten) life-giving group (of drugs) (*jīvanīya-gaṇa*, cf. AH.Sū.15.8.2, Su.Cik.5.12, Ca.Sū.4.9, Ca.Cik.29.61–70; frequently attested in the Bower Manuscript), which also includes *medā, mahāmedā, kākolī, kṣīrakākolī, jīvaka, ṛṣabhaka*, *madhuka*, *mudga-parṇī*, *māṣa-parṇī*. The first eight drugs from this group, *jīvantī* included, are attested in our medical prescription. Hoernle refers to this group as the *daśakavarga*, a label that is not recorded in

ingredient.[37] This should be regarded as a mistake, as *jīvaka* occurs further after (18). Hoernle (s.d.: 47 fn. 2) provided a plausible explanation for this mistake. He pointed out that confusion arose between *jīvanī-* and *jīvana-*: the former is used as a synonym of Skt. *jīvantī-* (cf. e.g. Aṣṭ.Nigh.1.15.2, Dhanv.Nigh.1.136.1, Rāj.Nigh.Guḍ.37.1), while the latter is used as a synonym of Skt. *jīvaka-* (cf. e.g. Rāj.Nigh.Parp.11.1). Thus, Hoernle regarded Khot. *jīvakạ* as a mistake for Skt. *jīvanī-* (= *jīvantī-*).

Skt. *madhuka* (2), Khot. *mahābūjä* (2), TB *pisau* (3)

Khotanese *mahābūjä* clearly means 'liquorice' and thus translates Skt. *madhuka-* 'liquorice (*Glycyrrhiza glabra* Linn.)', and not Skt. *madhūka-* 'mahua tree (*Madhuca indica* J. F. Gmel.)', as suggested by Konow and Chen.[38] In the Khotanese *Siddhasāra*, *mahābhuṃja-* is consistently used to translate Skt. *madhuka-* 'liquorice'; suffice it to compare the Sanskrit original of the first two verses of Si 2.15, where both Skt. *madhuka* 'liquorice' and *madhūka* 'mahua tree' are attested, with their Khotanese rendition:[39]

Si 2.15

Skt. *śārivā padmakośīra-madhūkaṃ candana-dvayam*
kāśmaryaṃ madhukaṃ ceti śārivādir ayaṃ gaṇaḥ
(1) Indian sarsaparilla, (2) bird cherry, (3) vetiver, (4) mahua tree, (5) the two sandals, (6) white teak, and (7) liquorice plant. This is the śārivādi group.

Khot. *sumanaśāriva ṣiʾ cu sụmaṃ hīya bāva puṣṭarä . uśaiʾrä . madhūkapuṣpä . śī caṃdaṃ . u hainai caṃdaṃ kaśmīrye . u mahābāṃjä .*
(1) Indian sarsaparilla (*śārivā-*), that (is), the root of sumana, (2) bird cherry (*padmaka-*), (3) vetiver (*uśīra-*), (4) the flower of the mahua tree (*-madhūkaṃ*), (5) white sandal and red sandal (*candana-dvayam*), (6) white teak (*kāśmaryaṃ*), and (7) liquorice plant (*madhukaṃ*).

Meulenbeld's encyclopaedic work on Indian medical literature. Drugs that in older Āryurvedic texts originally belong to this group are also found in the *aṣṭavarga* drugs, with the exception of *jīvantī* and *madhuka*, and with *ṛddhi* and *vṛddhi* replacing the two *parṇī* plants. On the evolution of the *aṣṭavarga*, see Joshi 1983.

[37] Adams 2013: 280 identifies TB *jivanti* with the *Cocculus cordifolia*, possibly following Monier-Williams (423). However, in older Āyurvedic texts, the *Cocculus cordifolia* is *guḍūcī*, borrowed in Tocharian B as *gurūci* (through a Middle Indic intermediary).Ṭāme 2014: 50 emends ms. *jīvakaṃ* to *jīvatta* without providing an explanation.

[38] Konow 1941: 6, 97 *passim*, Chen 2005: 305 *passim*. Cf. Maue 1990: 161 and Ṭāme 2014: 123.

[39] In the Khotanese translation I give Sanskrit equivalents of technical terms between parentheses. In the Tibetan translation of Si 2.15, Skt. *madhūka* 'mahua' is rendered with Tib. *ma dḥu ka*, while Skt. *madhuka* 'liquorice' is translated by Tib. *shing mngar.* In the only occurrence where the Khotanese version has *madhukä* as a rendering of Skt. *madhuka-* 'liquorice' (Si 2.2), the Tibetan version also has the loanword *ma du ka* (attested only here in the entire Tibetan *Siddhasāra*).

As can be seen, Khot. *mahābāṃjä* (also spelt *mahābuṃjä*, *mahābujä* and *mahābūjä* in the *Siddhasāra*) corresponds to Skt. *madhukaṃ*, while Khot. *madhūka-puṣpä* corresponds to Skt. *madhūkaṃ*. In Si 3.22.12, Skt. *madhūka* is translated by the loanword *madhukä*. In the corrupt Sanskrit version of the *Jīvakapustaka*, *madhūka* is attested eleven times (JP 7.3 *madhūka*, 11.1 *madhūkaṃ*, 30.3 *madhūkasya*, 32.1 *madhūka*, 33.2 *madhūka*, 38.1 *madhūka*, 41.2 *madhūksyā* [!], 43.1 *madū‹ka›*, 54.1 *madhūkaṃ*, 55.1 *madhūkaṃ*, 79.2 *madhūka pūṣpayau*),[40] where it is consistently translated by Khot. *mahābāṃjä* (or its variants *mahābāṃja, mahābūjä*) except for 79.2, where we find Khot. *madūka-pūṣpa* 'flower of the mahua plant' (cf. Si 2.15). This means that the Khotanese translator could distinguish between *madhuka* 'liquorice' and *madhūka* 'mahua tree', even though the copyist of the Sanskrit text wrote both words with long *-ū-*.

As for Tocharian, TB *pisau* is usually translated 'aneth, dill', as it seems to render Skt. *puṣpāhvā-* 'aneth (*Anethum Sowa* Roxb. ex Flem.)' in the Sanskrit-Tocharian B *Yogaśataka* fragment PK AS 2B b1–6. The passage in question is as follows:

Yogaśataka 94 (PK AS 2B b2–3 and b5–6)

Skt. *puṣpāhvā-ghana-bilva-kuṣṭha-phalini-kṛṣṇā-vacā-kalkīto vasti ⋄ kāṃcika-mutra-dugdha-sahito*
'An enema by grounding anethum (*puṣpāhvā-*), nut grass (*ghana-*), Bengal quince (*bilva-*), costus (*kuṣṭha-*), perfumed cherry (**phalinī-*), long pepper (*kṛṣṇā-*), sweet flag (*vacā-*), joined with sour gruel (**kāñcika-*), cow urine (*mutra-*), milk (*dugdha-*).'

TB *pissau ⋄ pilamātti ⋄ kaṣṣu ⋄ pippāl ⋄ okaro ⋄ kāñci ⋄ kewiye miśosa malkwersa wat*
'*pissau*, Bengal quince, costus, long pepper, sweet flag, sour gruel, with cow urine or milk.'

As can be seen, the Tocharian translation omits two ingredients: there is no equivalent of Skt. *phalini-* 'perfumed cherry' and it is unclear whether *pissau* corresponds to *puṣpāhvā-* 'aneth' or to *ghana-* 'nut grass (*Cyperus rotundus* Linn.)'.[41] If *pissau* indeed translates Skt. *puṣpāhvā*, Maue (1990: 161 fn. 23) suggests that the author of the Tocharian text mixed up *madhuka* 'liquorice' with *madhurikā* 'fennel (*Foenicum vulgaris* Mill.)' (cf. Si.Nigh.138 *śatāhvā śata-puṣpā ca miśī madhurikā matā*; Nadkarni 1954: 1.557).

[40] Skt. *maṇḍūka* in JP 59.1(98r3) does not stand for *madhūka* (so Chen 2005: 401), but regularly for Skt. *maṇḍūka* 'Indian pennywort (*Hydrocotyle Asiatica* Linn.)', translated *auṣka khuysiṃ* in Khotanese.
[41] The identification of *puṣpāhvā* as 'aneth' is confirmed by the Rāj.Nigh.Śat.11.2, where *puṣpāhvā* is said to be a synonym to *śataprasūnā*, *bahalā* and *śatapatrikā*. In the Bower Manuscript, the words referring to aneth are *śatapuṣpā* 'Indian dill (Anethum graveolens Linn.)' and *śatāhvā* 'id.'.

Skt. *vyoṣa* (3), Khot. *papala*, *ttūṃgarą*, *mīraijsya* (3abc), TB *mrañco*, *pipāl*, *tvā(ṅkaro)* (1abc)

Khotanese and Tocharian enumerations of single drugs that in Sanskrit are grouped under hyperonyms are (3) and (4).

The neuter noun Skt. *vyoṣa* refers to the three hot spices, which are explicitly listed in both versions: long pepper (Khot. *papala*, TB *pipāl*), dry ginger (Khot. *ttūṃgarą*, TB *tvāṅkaro*), and black pepper (Khot. *mīraijsya*, TB *mrañco*). Khot. *papala* and TB *pipāl* are loanwords from Sanskrit *pippalī* (or a Middle Indic continuant, cf. Gāndh. *pipali*). Since Bailey (1937: 913), TB *tvāṅkaro* is considered a loanword from Khotanese. Dragoni (2021: 305–307) suggests a loanword from the Early Old Khotanese antecedent of LKhot. *ttuṃgara*(*a*)- < **tvaṃgaraa*- or **tvāṃgaraa*-. Also, it is evident that Khot. *mīriṃjsya* (variant of *mīraijsya*) and TB *mrañco* are related to Skt. *marīca*-, usually regarded as a Wanderwort (EWAia 2.321). It is further noteworthy that TB *mrañco* and Khot. *mīriṃjsya* have a nasal element in common. Dragoni suggests that Tocharian borrowed from pre-Khotanese **mriṃjsya*- (or **miriṃjsya*-).[42] However, since the Khotanese word has a Central Asian substratic origin (Dragoni 2023: 158), the possibility cannot be ruled out that both Tocharian B and Khotanese borrowed from a third unattested word.

Skt. *śārivā* (4), Khot. *caṃdana śārava*, *sumąna śārava* (4ab), TB *wasto yä(kne) – ·eñcaṣṣe tāno* (4)

While the Sanskrit version only has *śārivā*- (in the compound *śārivośīra*),[43] the Khotanese version provides clarification that both *caṃdana śārava* (= Skt. *candana-śārivā*-) and *sumąna śārava* are required. In Dhanv.Nigh.1.158–161, an informative description of the two varieties of *sārivā*s is found, where *candana-sārivā* is described as *kṛṣṇamūlī kṛṣṇā* (the black creeper) and corresponds to the black Indian sarsaparilla (*Ichnocarpus frutescens* R. Br.).[44] On the other hand, Khot. *sumana-śārava* must correspond to the other variety of the same plant, the white Indian sarsaparilla (*Hemidesmus indicus* R. Br.) (Meulenbeld 1974: 453, Khare 2007: 306). This variant is used in the Khotanese *Siddhasāra* as a translation of Skt. *śārivā* and Tib. *thal tres* (Si 2.15 and 2.30). In the former occurrence (Si 2.15), Khotanese has an explanation added: *sumana-śāriva ṣiʾ cu suṃaṃ hīya bāva* 'white Indian sarsaparilla (= *śārivā*-), that (is),

[42] Dragoni 2023: 165–167. A Khotanese origin of *mrañco* was already suggested by Blažek and Schwarz 2015 without discussion.

[43] The actual form attested in the manuscript is indeed *śāravī* (58r1), which could theoretically stand for a dual *śārive*. However, the compound *śārivośīra*- in the edited text is necessary for metrical reasons (see § 3.1.1).

[44] As pointed out by Maue 1990: 162 fn. 27, Skt. *candana-śārivā*- is also attested in the Bower Manuscript (II.241). See further Meulenbeld 1999–2002: 2b.13 n. 156.

the root of white (kind)' (Emmerick *Studies* 2: 41). At the point where one would expect a Tocharian B rendering of Skt. *śarivā-*, the manuscript contains a lacuna: b7–8 *wasto-yä*(*kne*) *– ·eñcaṣṣe tāno.* There is no doubt, however, that the text is meant to say that a seed (*tāno*) from both types (*wasto-yäkne*) of sarsaparilla (*– ·eñcaṣṣe*) is needed. Maue has observed that the vowel *-e-* before what seems to be a dual ending *-ñc* is troublesome, as one would rather expect *śāripäñc* as the dual of *śārip.* This form does not conform to the surviving akṣara(s) in the manuscript.[45]

Skt. *uśīra* (5), Khot. *uśaiʾra* (5), TB *uśir* (5) and Skt. *caṃdana* (6), Khot. *caṃdä* (6), TB *candāṃ* (6)

Sanskrit *uśīra* 'vetiver (*Vetiveria zizanioides* Linn.)' is reflected in both Tocharian B *uśir* and Khotanese *uśaiʾra* by the actual Indian word adapted to each language.

A similar case is Skt. *caṃdana* 'white sandal (*Santalum album* Linn.)', borrowed in Tocharian B as *candāṃ* /candán(ə)/. In Khotanese, the spelling *caṃdä*, for expected *caṃdaṃ*, is likely regular. In his edition of the Khotanese *Siddhasāra* (KT 1: 2–134), Bailey frequently emended ms. *caṃdä* to *caṃdaṃ.* However, it is curious how often the spelling *caṃdä* occurs in this text (8×, Si 2.13[ms. Ch], 13.32, 14.8, 15.18, 25.19, 26.29, 26.31, 26.33) compared to the more regular *caṃdaṃ* (9×, Si 2.12, 2.13[ms. P], 2.15, 2.21, 12.20, 23.15, 26.13, 26.16, 26.32[*cadaṃ*]). In other medical texts, the spelling *caṃdä* is also found (PiŚ 64 *caṃdä* vs. PiŚ 40 *caṃda*), and the same alternation is attested in the *Jīvakapustaka* (*caṃdaṃ*: 44v1, 50r5, 56v1, 63r2, 72r1, 75v3, 80r2, 81v2, 82r4, 86r3, 90v5, 97v3, 105r5, 108r4, 109r3, 110r3, 110v5; *caṃdä*: 58v3, 82r4, 97r2, 108r4, 112v1; *caṃda*: 60r5). Thus, LKhot. *caṃdä* should be regarded as a graphic variant of *caṃdaṃ.* Possibly, older *candanä* [ˈʧandanə] (stress on the first syllable) was first reduced to **candänä* through trisyllabic weakening and then to **candäṃ* [ˈʧandən] or [ˈʧandə̃].[46] Owing to the potential difficulty of writing **candäṃ*, which requires two diacritics above the second akṣara, scribes appear to have opted for alternative spellings: either *candaṃ* (with the anusvāra diacritic) or *candä* (with the diacritic of the two dots).[47]

Skt. *padmaka* (7), Khot. *pauṣṭarä* (7), TB *yāstaci* (7)

The correspondence underlying the seventh ingredient is among the most problematic cases. While Skt. *padmaka-* indeed refers to the 'bird cherry, wild Himalayan cherry (*Prunus cerasoides* D. Don.)',[48] Maue (1990: 162 fn. 30) opts for a translation 'juniper'

[45] Maue 1990: 162 fn. 28. See the relevant interlinear note in the edition (§ 2.3.1).

[46] The Sanskrit loanword Khot. *candana-* 'white sandal' was not specific of the medical jargon, as it is also attested in Buddhist literary texts.

[47] On a similar point regarding Tocharian writing conventions, see Peyrot 2021: 453–454.

[48] See Nadkarni 1954: 1.1015, Khare 2007: 520, Sharma 2018: 230.

for both the Tocharian and the Khotanese ingredients, as he follows Bailey (*Dict.* 247) in translating Khot. *puṣṭara-* as 'juniper'.

Following Maue's identification, Blažek and Schwarz (2015) considered TB *yāstaci* 'juniper' a Khotanese-mediated loanword from the Common Nuristani word for 'juniper'. They reconstruct Common Nuristani **wistar-stə* 'of juniper, belonging to juniper' → Khotanese **wästac*[*ä*] ~ **wastac*[*ä*] → Early Tocharian B **w'āstaṷë* (in their notation) > TB *yāstaci*. However, this reconstruction has some formal problems on the Tocharian as well as on the Khotanese side. Indeed, the vowel *-ā-* /á/ in the first syllable of Tocharian as representing either Khot. *-ä-* /ə/, *-i-* /i/ or Common Nuristani *-i-* would be difficult to explain. Blažek and Schwarz point out that Khotanese alternates between *-ä-*, *-i-*, and *-a-*, citing examples like Khot. *śara-* 'good', *śarā-* 'fortune', *śärā* ~ *śirā* 'good, welfare', *śśära-* 'good'.

However, it is crucial to note that these words belong to different chronological stages of Khotanese. While *-ä-* [ə] and *-i-* [i] partially interchanged in Old and Late Khotanese (particularly in the proximity of a palatal sound and in word-final position), adjectival forms like *śśära-* 'good' and *śara-* 'id.' belong to different stages: *śśära-* ~ *śśira-* is Old Khotanese, while *śara-* is Late Khotanese.[49] The same distribution of *-ä-*, *-i-*, *-a-* is shown by the Late Khotanese noun *śarā* ~ *śärā* ~ *śirā* 'goodness, welfare', which is from OKhot. *śśäratāti-* (attested as early as in the *Śūraṅgamasamādhisūtra*).[50] As Blažek and Schwarz reconstruct a very old loanword in Proto-Tocharian, the root vowel should have been a sound close to Common Nuristani *-i-*, thus either Khot. *-ä-* [ə] or *-i-* [i] and not Khot. *-a-* [a]. Furthermore, there is no pre-Khotanese reduction *-rst-* or *-st-* > Khot. *-c-*. As for Tocharian B, one should also assume an extremely old loanword, as it would have occurred before the palatalisation of (pre-)PT **w* (>**ẃ* > TB *-y*) ceased to operate. On top of these formal problems, there is no evidence that Nuristani languages have ever come into contact with either Tocharian or Khotanese.

The Sanskrit word for 'juniper' is generally *hapuṣā-*, which, as recently demonstrated by Maggi (2021–2022), must be regarded as a loanword from Iranian. In Khotanese, the word *hapūṣi* 'juniper' has been borrowed from Sanskrit, and it is attested three times, two as an ingredient (JP 26.1 [= *hapuṣā*], JP 48.4 [= *hapuṣā*]) and once in Skt. *hapuṣādya-ghṛtam* = Khot. *hapūṣāda rūṃ* (JP 26.3), the name of a medicated ghee (Maggi 2021–2022: 8 and 2022). As for Khot. *pauṣṭarä*, *puṣṭarä*, it occurs four times in the Khotanese *Siddhasāra*: Si 2.14 *puṣtarä*, 2.15 *puṣṭarä*, 23.15 *puṣṭarä*, 26.15 *puṣṭarä*,

[49] For the etymology see Emmerick *Studies* 1: 117. There are also occurrences of another noun *śirata-* 'good person' in Old Khotanese (on which see Maggi *Studies* 3: 145).

[50] For the edition and translation of the Khotanese *Śūraṅgamasamādhisūtra*, see Emmerick 1970.

where it always corresponds to Sanskrit *padmaka-* 'bird cherry'.[51] The Tibetan version consistently translates *padmaka-* as *shug pa,* which is indeed the Tibetan name for the juniper tree (hence Bailey's interpretation). Since we know from its preface that the Khotanese *Siddhasāra* has primarily been translated from Tibetan, Maue nevertheless interprets Khot. *pauṣṭarä*, *puṣṭarä* as the translation of *shug pa* 'juniper' (Boesi 2014: 82–84), which is commonly used as an apparent substitute for the bird cherry also in Buddhist texts (cf. Emmerick 1967: 155). However, in the Khotanese *Jīvakapustaka*, clearly translated from Sanskrit, the correspondence Skt. *padmaka* = Khot. *pauṣṭarä*, *puṣṭarä* strongly suggests the identification of the latter as the bird cherry.[52] Thus, Khot. *pauṣṭarä* and TB *yāstaci* can mean 'bird cherry, wild Himalayan cherry', as the Sanskrit correspondent.[53]

Further, the word division of *yāstaci* in THT 500–502 b8 is dubious as this alleged word is attested just before a lacuna of approximately three akṣaras. If *yāstaci* is indeed a single word referring to the bird cherry, one may wonder whether the word lost in the lacuna might have specified which part of the tree is to be used. According to Nadkarni (1954: 1.1015), the commonly used parts of the *padmaka* are its root bark, smaller branches, and seeds (see also Malzahn *apud* CEToM). Thus, we may tentatively restore *tāno* 'seed', *stām* 'tree', *karāk* 'branch'. As for *yāstaci*, it could theoretically be interpreted as an inflected nom.pl.m. of *yāsta-tstse* or, if a noun in its own, it could be restored as *yāstaci*(*ṣe*/*a*), i.e. a *ṣṣe*-adjective built on the oblique form of a noun *yāstaciye** (or the like; see Malzahn *apud* CEToM).

Skt. *triphalā* (8), Khot. *halīrai*, *vīhīle*, *ᾱmalai* (8abc), TB *arirāk*, *wipitāk*, *a(malā)k* (8abc)

Sanskrit *triphalā* designates the 'three fruits', which are *harītakī-* 'chebulic myrobalan (*Terminalia chebula* Retz.)', *vibhītakī-* 'belleric myrobalan (*Terminalia bellirica* Roxb.)', and *āmalakī-* 'emblic myrobalan (*Emblica officinalis* Gaertn.)'. As for the first, both Khot. *halīrai* and TB *arirāk* exhibit an unexpected adaptation if derived directly from Skt. *harītakī-*. The Tocharian B form more closely resembles Sogd. *ʾʾrʾyrʾy* /ārirē/

[51] In the Sanskrit *Siddhasāra*, the word *hapuṣā* occurs three times in three different chapters (9.14, 10.33 and 30.60), for which we unfortunately lack a Khotanese translation. The Tibetan version, on the other hand, translates *hapuṣā* with *sba ma'i 'bras bu* 'fruit of juniper'. However, a parallel of Si 9.14 is attested in JP 26, as shown by Chen 2005: 335 and Maggi 2022.

[52] In the Khotanese JP, *pauṣṭarä* occurs four times: JP 5.2(50v1) (= Skt. *padmaka*), JP 11.1(58v2) (= Skt. *padmaka*), JP 30.1(75v2) (= Skt. *padmaka*), JP 38.2(82r4) (= Skt. *padmaka*).

[53] In Tocharian B, we find a word *padmak* in W8 a5 and W13 a4. In W13, TB *padmak* is one of the ingredients used to make a powder aimed at curing eye-diseases. Following Filliozat 1948: 127, Adams 2013: 379 prefers to interpret this *padmak* as the blue or sacred lotus (Nelumbo nucifera). See also Meulenbeld 1974: 570 s.v. *padma*.

(Maue and Sims-Williams 1991: 493). Sogdian and Tocharian B may have borrowed from a common Middle Indic continuant of Skt. *harītakī-*, possibly with retroflection of *-ta-* > *-ṭa-* (cf. also Old Uyghur *arire*). On the other hand, TB *wipitāk* can directly derive from Skt. *vibhītakī-*, while Khot. *vīhīle* may be from a Middle Indic variant displaying *-bh-* > *-h-* (cf. von Hinüber 2001: 162, §189–190). The recently discovered Sogd. *βrʾyrʾy* /virīrē/ may equally derive from a comparable Middle Indic form (Reck and Benkato 2018: 69).

Skt. *dārvī* (9), Khot. *ysālva* (9), TB *śkwiśko* (9)

The correspondence between Khot. *ysālva* and Skt. *dārvī* 'barberry (*Berberis aristata* D.C.)' is confirmed by several passages from the *Siddhāsara* (Emmerick *Studies* 1: 106). In his reconstruction of the Sanskrit text, Chen (2005: 305) emends Skt. *dārvī* to *dāru*, which he translates as '天木 (雪松) [cedar, deodar]'. This is unlikely because the Khotanese translation of Skt. *dāru* is usually Khot. *divadāra, detadāra* (← Skt. *devadāru-*) in the *Siddhāsara*.

Skt. *sthirā* (10), Khot. *sālaparṇa* (10), TB *sālaparni* (10)

Khotanese *sālaparṇa* corresponds to Skt. *sthirā* also in Si 2.1. Initial *s-* in both TB *sālaparṇī* and Khot. *sālaparṇa* points to a Middle Indic continuant of Skt. *śālaparṇī-*, the *Desmodium gangeticum* DC., which is also one of the meanings of Skt. *sthirā* (cf. Dhanv.Nigh.1.87.1, Rāj.Nigh.Guḍ.37.1). A Tocharian B variant *sālavarṇi* is attested in PK AS 3A a4, W 18 b3, and W 36 b3.

Skt. *aṃśumatī* (11), Khot. *‹prraśnaparṇa›* (11), TB *pṛśnaparṇi* (11)

A notable discrepancy among the three versions is the omission of Skt. *aṃśumatī* in the Khotanese rendition. In Tocharian B, it has been translated as *pṛśnaparṇi* '*Uraria lagopoides* D.C.' (← Skt. *pṛśniparṇī*) (Meulenbeld 1974: 577, Khare 2007: 684), which is also attested as *pṛśnavarṇi* in W 36 b4 (*pṛśnavarṇi*) and W 18 b3 (*präśnavarṇi*). In other Tocharian B texts, Skt. *aṃśumatī* seems to have been borrowed as *aśamati* (cf. W 31 b4). As for Khotanese, it is plausible that the scribe inadvertently omitted this ingredient due to its formal similarity to the preceding plant name. Both *sālaparṇī* and *pṛśniparṇī* are frequently cited together in the Bower Manuscript and other recipes found in the *Jīvakapustaka*.[54] A case of comparable omission can be found in JP 33.2(78r3), where the Sanskrit text has *parṇyau dve* 'the two *parṇī* plants' (ms. *parṇau dvi*), while the Khotanese translation only has *sālaparṇa* (78v3). In the *Siddhasāra*, Khot.

[54] In several Nighaṇṭus, however, *aṃśumatī* is regarded as a synonym of *śālaparṇī* '*Desmodium gangeticum* DC' (Aṣṭ.Nigh.1.18, Dhanv.Nigh.1.87, Rāj.Nigh.Śat.17–20).

priśnaparṇä is used to translate Skt. *lāṅgulī-* 'pointed-leaved uraria'[55] in 2.1, Skt. *kalaśī-* 'id.' in 2.17, and *pṛṣṭaparṇī-* in 2.29. In this last passage, it occurs together with *sālaparṇä*, but ms. P, akin to JP 11, omitted the latter.

Skt. *bṛhatī* (12), Khot. *brrahattä* (12), TB *(praha)ti* (12)

Sanskrit *bṛhatī* 'Indian nightshade (*Solanum indicum* Linn.)' is reflected as *brrahattä* in Khotanese. In the Tocharian B manuscript, there is a lacuna of two akṣaras, but, with Maue (1990: 160 fn. 12), we may reconstruct either TB *(praha)ti* (attested in e.g. PK AS 2A a4–5, PK AS 9A b3, W 29 b2), with regular devoicing of the initial labial, or *(bräha)ti* (attested once in PK AS 3A a4).

Skt. *dhāvanī* (13), Khot. *kaṇḍārya* (13), TB *gandhakāri* (13)

Sanskrit *dhāvanī* is rendered with Khot. *kaṇḍārya* (cf. also Si 2.11) and TB *gandhakāri.* Khotanese *kaṇḍārya* is a loanword from a Middle Indic reflex of Skt. *kaṇṭakārikā*, whose literal meaning is 'producer of thorns'. Skt. *kaṇṭakārī* has been identified with the wild eggplant (*Solanum virginianum* Linn., *Solanum xanthocarpum* S. & W., or *Solanum surattense* Burm.)[56] and is said to be a synonym of *dhāvanī* in e.g. Dhanv.Nigh.1.95.1–2 and Si.Nigh.10 (Emmerick 1980–1982: 2.7). In the Khotanese *Siddhasāra*, Khot. *kaṇḍārya* corresponds to different Sanskrit and Tibetan words: Si 2.6 Khot. *kaṇḍārya* = Skt. *kaṇṭakī* = Tib. *kaṇṭa ka ri*; Si 2.11 Khot. *kaṇḍārya* = Skt. *dhāvanī* = Tib. *dha da ki* (!); Si 21.11 Khot. *kaṇḍārya* = Skt. *vyāghrā* = Tib. *bri ha tī*; Si 26.50 Khot. *kaṇḍārya* = Skt. *vyāghrī* = Tib. *bya gḥri*; Si 26.51 Khot. *kaṇḍārya* = Skt. *bṛhatī* = Tib. *bri ha ti'i 'bras bu*; Si 26.69 *kaṇḍārya* = Skt. *vyāghrī* = Tib. *bya gḥri*. As for the Tocharian version, Maue (1990: 163 fn. 35) suggests that the spelling TB *gandha*° for expected *kaṇḍa-*/*kaṇṭa-* is a wrong Sanskritisation influenced by the word *gandha-*, *gandhaka-* 'scent'. The regular noun *kaṇḍakāri* occurs in PK AS 3A a4 *brähati* ⬩ *kaṇḍakāri* and W 17 b4 *(pra)hati* ⬩ *kaṇḍakāri* (both referring to similar plants from the *Solanaceae* family). In the Sanskrit *Siddhasāra*, we find the dual form *bṛhatyau* (Si 2.1) or the compound *bṛhatī-dvaya* (Si 2.29, 5.29, 5.97), which are translated in Khotanese with Si 2.1 *vāttāka . brihatta dva : kaṇḍārya* 'the two *bṛhatī*: *vārtākī* (and) *kaṇṭakārikā*' or Si 2.29 *brihatta dva* 'the two *bṛhatī*'. Sometimes Skt. *bṛhatī-* is even rendered with '*vārtākī* (and) *kaṇṭakārikā*', as in Si 2.5 (Skt. *bṛhatī* = Khot. *kaṇḍārya u vāttāka*). Similarly, JP 5(50r1) *bṛhatī-dvaya* (ms. *brrahattī-dvaya*) is translated by Khot. *brrahatta kaṇḍārya*.

[55] In Aṣṭ.Nigh.1.17.1, Dhanv.Nigh.1.90.1, and Rāj.Nigh.Śat.37.1, *kalaśī* is said to be a synonym of *pṛśniparṇī*.

[56] See Nadkarni 1954: 1. 1150–1151, Meulenbeld 1974: 535, Khare 2007: 615, Sharma 2018: 71–73.

Skt. *śvadaṃṣṭrā* (14), Khot. *drāṃgūlyai* (14), TB *klyotañä* (14)

Sanskrit *śvadaṃṣṭrā* is indicated as a synonym of *gokṣura-* 'caltrop (*Tribulus terrestris* Linn.)'.[57] The correspondence between Khot. *drāṃgūlyai* and Skt. *śvadaṃṣṭrā-* is supported by several passages from the *Siddhasāra*. In PK AS 2A a3 (*Yogaśataka*), the TB adjective *klyotaiṣṣana* 'pertaining to caltrop' translates Skt. *ṭṛkaṭu-* (Carling 2003: 42, 47–48). Although the manuscripts of the Sanskrit *Yogaśataka* attest variant forms, there is no doubt that Skt. *ṭṛkaṭu ~ trikūṭa* must refer to the caltrop. This is demonstrated by Dhruvapala's gloss, where *trikūṭa* is mentioned as a synonym of *gokṣuraka*, and by the Tibetan translation *gze* (Filliozat 1948: 38 n.5 and 1979: 63). Typically, the fruit and root of the caltrop are employed, and it is plausible that the plural form TB *klyotañ* is used here to refer to the dried spiny fruits of the caltrop.

Skt. *punarnavā* (15), Khot. *aiśta bāva* (15), TB *eśpeṣṣe* (15)

The Tocharian B correspondent of Skt. *punarnavā* 'hogweed (*Boerhavia diffusa* Linn.)' is generally *punarnapa*/*punarṇap* but is uniquely *eśpeṣṣe* here. The Khotanese equivalent is always *aiśta* (with variants)[58], mostly followed by the Khotanese word 'root', i.e. *bāta*, *bāva* or *bā*. Following a suggestion by Adams (2013: 104), Dragoni (2021: 301) has hesitantly proposed that *eśpe* may be a loanword from a Late Khotanese compound *aiśt(a)-bā(va)* 'hogweed root'. If so, we might have here a *ṣṣe*-adjective referring to an implied part of the hogweed (possibly its roots).[59]

Skt. *dve mede* (16), Khot. *mida*, *mahāmeda* (16ab), TB *met*, *mahāmet* (16ab)

While the Sanskrit version refers to the 'two medā' (*dve mede*), the Khotanese and Tocharian B versions specify the name of both plants: *medā* (Khot. *mida*, TB *met*) and *mahāmeda* (Khot. *mahāmeda*, TB *mahāmet*).

Skt. *dve kākolyau* (17), Khot. *kākauṭa*, *kṣīrakākauṭa* (17ab), TB *kākoṭi*, *kṣirakākoṭi* (17ab)

A comparable case to that discussed in (16) is found in the rendering of Skt. *dve kākolyau* 'two kākolī' by Khot. *kākauṭa*, TB *kākoṭi*, and Khot. *kṣīrakākauṭa*, TB *kṣirakākoṭi*. The original Sanskrit manuscript of JP 11 has °*kākauṭī*, which shows a retroflex stop also attested in Khotanese and Tocharian. In Khotanese, the standard spelling of the word is with *-ṭ-*, with only a couple of occurrences of *kākaula* (with *-l-*) in JP 74(105v1) and 75(106r4). In Tocharian B, the standard variant is also °*kākoṭi*, but we find spellings with

[57] Cf. Aṣṭ.Nigh.1.20.2 *gokaṇṭako gokṣurakaḥ śvadaṃṣṭrā ca trikaṇṭakaḥ*.

[58] For an explanation of the different spellings see Dragoni 2021: 300.

[59] See Maue 1990: 163 fn. 19. For the problem of the agreement, cf. masculine °*ṣṣe tāno* for expected feminine °*ṣṣa tāno* in line b8.

-r- in medical fragments from the Weber series (°*kākori* in W6 a3, W24 a5, W28 b4, W35 a4, W39 b2). In the Sanskrit version of the *Jīvakapustaka*, the standard spelling is with *-ṭ-*, but occurrences with *-l-* can be found (*kākāl̥līm* JP 46.4[87v3], *kākaula* 74.1[105r3], *kākāl̥laṃ* 75.3[106r1], *käkaula* 78.4[108r2] ?). In the Bower Manuscript, a variant °*kākoḍī* is also attested. It seems that the variant with the retroflex stop has been borrowed in both Khotanese and Tocharian B.[60] The TB form °*kākori* attested in the Weber series conforms nicely with other Indo-Aryan loanwords in which retroflex stops have been rendered with *-r-*.[61]

Skt. *jīvaka* (18), Khot. *jīvaka* (18), TB *jīvaka* (18) and Skt. *ṛṣabhaka* (19), Khot. *raṣabhaka* (19), TB *ṛṣabhak* (19)

More recent loanwords seem to be Khot. *jīvaka*, TB *jīvaka* (both ← Skt. *jīvaka*) and Khot. *raṣabhaka*, TB *ṛṣabhak* (both ← Skt. *ṛṣabhaka*) as the Khotanese does not have reduced intervocalic *-k-* (cf. e.g. *aumalaa-*, *āmalaa-* 'emblic myrobalan' ← Skt. *āmalaka-* or *caittra-* 'leadwort' ← Skt. *citraka-*), and Tocharian B does not show devoicing and deaspiration of Skt. *bh*.

Skt. *ātmagūpta* (20), Khot. *‹ā›dmagūpti* (20), *ātmagupti* (20) and Skt. *phalā* (21), Khot. *bela* (21), TB *palāṣṣai wi̯tsako enmet(r)e* (21)

In the Sanskrit version, Maue corrects ms. *ātmagūpta phalā* to *ātmaguptā bālaṃ* (or *ātmaguptā bālā*) 'cowhage (*Mucuna prurita* Hook.) (and) fragrant sticky marrow (*Pavonia odorata* Willd.)'.[62] That the sequence must not be interpreted as a single ingredient *ātmagūpta-phalaṃ* 'cowhage fruit' is demonstrated by the Khotanese and Tocharian versions, both of which have two ingredients: Khot. *‹ā›dmagūpti* : *bela* 'cowhage (and) Bengal quince (*Aegle marmelos* Linn.)' and TB *ātmagupti* ⋄ *palāṣṣai wi̯tsako enmet(r)e* 'cowhage (and) the root and *enmetre* of *pāla*'. Nonetheless, the correspondence between the ingredients is problematic.

---

[60] This hypothesis seems more probable than inferring Sanskritisation within Khotanese and/or Tocharian B.

[61] Cf. TA *utkur* ~ *ukkur* 'uncomfortable, painful' vs. Skt. *utkuṭa-* (Carling & Pinault 2023: 63); TB *laur** 'spike, stake' vs. Skt. *lakuṭa-* ~ *laguḍa-* (Isebaert 1980: 175; MW: 893); 36 TB *maur* ~ *mahur* 'diadem, tiara' (cf. TA *prop-mahur* 'id.') vs. Skt. *makuṭa-* (BHSD 413). Other loanwords of this kind (e.g. TB *kor* ~ *koṭ*, A *kor* '10,000,000' vs. Skt. *koṭi-*; TB *kāpar*, A *kāpār* 'morsel, bite-size piece' vs. Skt. *kavaḍa-* 'mouthful'; TB *pir* 'chair, seat' vs. Skt. *pīṭha-*, see Pinault 2008: 388–389) indicate that Tocharian borrowed a number of words from a Middle Indian variety where retroflex stops were articulated as retroflex taps (cf. von Hinüber 2001: 168, § 201). The Middle Indian variety in question was probably Gāndhārī, since in this language Old Indo-Aryan [ʈ] and [ɖ] were probably pronounced as a retroflex tap [ɽ], but written as *ḍ* – or with the modified sign *ḏ* in some Niya documents (Baums 2009: 141–142; cf. already Burrow 1937: 7).

[62] Maue 1990: 161 and 163 fn. 43. He gives the stem as *bālā̆-* (cf. also *bālaka-*).

The plant name TB *pāla** /pála/ (in the derived adjective *palāṣṣe** /paláṣṣe/) is a hapax legomenon and Khot. *bela-* does not mean 'fragrant sticky marrow' but 'Bengal quince (= Skt. *bilva-*)'.[63] Maue pointed out that the first vowel in the Tocharian B word speaks against a borrowing from Khot. *bela-* or Skt. *bilva-*. Therefore, he suggests that ms. *phalā* should be corrected with either *bālaṃ* or *bālā* 'fragrant sticky marrow', borrowed as *pāla** in Tocharian B, and that Khotanese *bela* would represent a case of a mismatched ingredient. Adams is of a slightly different view. Although in the dictionary entry *palāṣṣe** he reports both 'pertaining to cowhage' and 'pertaining to Bengal quince' as possible meanings, he seems to favour the latter identification in the etymological discussion of *enmetre*, which he translates as 'bark'.[64] Nonetheless, he maintains that *pāla** 'Bengal quince' is borrowed from Skt. *bālā* 'cowhage', which is far-fetched. If we leave aside for the moment the question about the hapax legomenon *pāla** and the difficult form *phalā* in the Sanskrit manuscript, only Khotanese *bela* can be interpreted without hesitation as 'Bengal quince'. One may therefore wonder whether Skt. *phalā* is here used as a synonym of *bilva-*, rather than being a mistake for *bālā*. As a matter of fact, *pha* and *bā* are palaeographically quite different, although this type of mistake is sometimes attested in the *Jīvakapustaka* manuscript.

Among the synonyms of *bilva-*, we find *mahāphala-* (m.) in Aṣṭ.Nigh.192.2, Dhanv.Nigh.1.104.1, Nigh.Śeṣa1.63.1, and several other compounds with °*phala-*, e.g. *śrīphala-*, *lakṣmīphala-*, *gandhaphala-*, *sadāphala-*, etc. According to Hoernle, the fact that *phalā* may also be a synonym of *bilva* is supported by several parallel cases, e.g. by Skt. *dhāmārgava*, which has both synonyms *phalā* and *mahāphalā*, or by Skt. *vṛntākī*, which has both *phalā* and *raktaphalā*. Otherwise, ms. *ātmagūpta phalā* may be interpreted as a compound, inflected as nom.pl. *ātmagūptā-phalā* (with external sandhi because of following *rāsnā*). Keeping the manuscript reading has two further advantages: it not only solves the problem of the Khotanese unmatched ingredient, but it also makes pāda a of JP 11.3 perfectly metrical. Indeed, emending *phalā* to *bālā* (or *bālaṃ*) would yield an unmetrical sequence of four long syllables in the second colon of pāda a, corresponding to *bālā* (or *bālaṃ*) *rāsnā* (for expected ⏑ – – ×). Accordingly, the use of (°)*phalā-*, (°)*phala-* as a synonym of *bilva* could have served the needs of the composer of the Sanskrit text to respect the metre.

[63] For the correspondence between Khot. *bela-* and Skt. *bilva-*, see e.g. Si 2.8, 2.28, etc.

[64] Adams 2013: 384 s.v. *palāṣṣe** and 92 s.v. *enmetre.* Adams further claims that we would expect **wtsikai* rather than *wiṭsako* in the manuscript, but this is not correct, as all other nominal forms attested in the formula are inflected as nominative. Therefore, the form to be corrected is *palāṣṣai*, which should have been either *palāṣṣa* if in agreement with *wiṭsako* or possibly *palāṣṣi*, *palāṣṣana* if in agreement with both *wiṭsako* and *enmetre*.

Skt. *rāsnā* (22), Khot. *lagara bāva* (22), TB *rās(n)ā* (22)

As for Skt. *rāsnā* 'groundsel (*Pluchea lanceolata* C. B.)', TB *rāsnā* appears to be a learned non-integrated loanword (for expected *rāsna* or *rasna*, attested in W23 b4 and W29 b2), while Khotanese has *lagara bāva*, which seems to be a calque of Skt. *rāsnā-mūla.*

Skt. *nalada* (23), Khot. *svarṇagūttaryāṃ būśq̄nai* (23), TB *peparṣṣa wi(tsako)* (23)

As mentioned in §3.1.1, the Sanskrit version records the following ingredient as *naṃdaṃ*. Hoernle hesitantly corrected the manuscript reading with *nandana*, an unidentified drug in Su.Ka.2.5 listed among the twelve plants with poisonous fruit and seven plants with poisonous bark, pith or gum. Bailey's restoration *naladaṃ* 'spikenard (*Nardostachys jatamansi* DC.)' is indeed correct.

The Khotanese correspondent is *svarṇagūttaryāṃ būśq̄nai*, literally 'the fragrant (plant) of the *Suvarṇagotrikas'. The substantivised adjective *bŭśānai-* (and variants) 'fragrant, perfumed' translates Skt. *māṃsī* 'nard (*Nardostrachys jatamansi* DC)' in the *Siddhasāra* (Si 21.11, 26.56, 26.57). In the *Jīvakapustaka*, *svarṇagūttaryāṃ bŭśānai* (and variants) corresponds to both *māṃṣī* and *nalada* (cf. JP 12 [= Skt. *mā‹ṃ›ṣī*], 47 [= Skt. *naladaṃ*], 74 [= Skt. °*naladaṃ*], 75 [= Skt. *māṃṣī*], 81 [= Skt. *nalada*]) (see Bailey 1940: 602 and Ṭāme 2014: 206–207). Khot. *svarṇagūttaryāṃ būśq̄nai* must therefore designate a nard native of the Suvarṇagotra land.[65] In fact, Khot. *buśānaa-* is frequently used in collocation with genitive-dative plural forms of ethnonyms (e.g. *cigāṃ būśq̄naa-* 'fragrant (plant) of the Chinese', *ttāgūttāṃ būśq̄naa-* 'fragrant (plant) of the Tibetan, a nard') as referring to specific plants associated with certain ethnic groups (Luzzietti 2023: 132–133). It follows that *svarṇagūttaryāṃ* in JP 11 can be interpreted as a genitive-dative plural of an inferable noun *svarṇagūttaria-* 'inhabitants of the *Suvarṇagotra* land' (← Skt. **Suvarṇagotrika-*).[66] A similar noun phrase is also attested in JP 12(60v1–2). Still, we find additional variants in the text: while *svarṇagūttariṃ būśānį* in JP 81(110v5) can easily be translated with 'fragrant (plant) of the *Suvarṇagotra* (land)', the interpretation of *svarṇagūttaryāṃga būśaunį* (JP 47[91r1]) ~ *svarnagūttaryāṃga būśānį* (JP 74[105v1]) is troublesome. Both Konow and Bailey take *svarṇagūttaryāṃga* as a single word and thus apparently consider *svarṇagūttaryāṃ* in JP 11 and 12 as a defective spelling for *svarṇagūttaryāṃ‹ga›* (cf. Konow 1941: 103, Bailey KT 7.82). Ṭāme (2014: 206) essentially follows Konow's interpretation without attempting an explanation for the alleged compound *svarnagūttaryāṃga.* I propose to divide the sequence as *svarṇa-gūttaryāṃ ga būśaunį* 'the fragrant (plant) of the tribe of the *Suvarṇagotrika*s', where *ga*

[65] See Bailey 1940: 602–603, 1982: 62, 90–91, and KT 7.82–83 with references therein. On the identification and location of the Suvarṇagotra land, see Zeisler 2009–2010: 416–425.

[66] I owe this explanation to Mauro Maggi (p.c.).

is the outcome of the Old Khotanese noun *ggäṣā*ʾ- 'group, crowd, assembly' > LKhot. *ga*ʾ, used here with the meaning of 'tribe' (the subscript hook is often omitted in writing).

The origin of TB *pepar* in the derived adjective *peparṣṣa* is unclear. Typically, Tocharian B uses the loanword *nalat* for 'spikenard', as Maue noted.[67]

Skt. *padmakesara* (24), Khot. *padmakyesärä* (24), TB *upālṣa witsako* (24)

A reverse case to that of (22) is Skt. *padmakesara*, where Khotanese has the adapted loanword LKhot. *padmakyesärä*,[68] while Tocharian B shows the calque *upālṣa witsako*, which literally means 'root of the lotus'. Maue (1990: 164 fn. 48) argued that the Tocharian B version here deviates from the Sanskrit and Khotanese versions, because Skt. *padma* refers to the Indian lotus (*Nelumbo nucifera* Gaertn.), while Skt. *utpala* (the actual source of TB *upāl*) refers to the blue water lily (*Nymphaea stellata* Willd.). Furthermore, the filament (Skt. *kesara*) of the *padma* should be used and not its root (TB *witsako*), as the Tocharian B version seems to imply. However, TB *uppāl* is used to translate a variety of plants belonging to the *Nymphaeaceae* family. Besides rendering Skt *utpala* and *nilotpala*, both referring to the blue water lily, TB *uppāl* also translates Skt. *padma-* in e.g. PK NS 305 + 306 a2.[69] Furthermore, in IOL Toch 305 a4, the noun phrase *uppāläṣṣana witsakaṃmpa* is used to translate the instr pl. *-bisair* 'lotus stalk'.[70] Thus, I take *upālṣa witsako* in THT 500–502 b11 as a close rendering of Skt. *padmakesara* 'lotus filament'.

Ingredients (25)–(29)

The Sanskrit and Khotanese versions preserve the ingredients (25)–(29), while the Tocharian manuscript has a gap after (25) *tālis*(*pattär*) (= Skt. *tālīsapatra*, Khot. *ttālīspattra* 'leaf of silver fir [*Abies webbiana* Lindl.]'). Skt. (26) *mañjiṣṭhā* 'Indian madder (*Rubia cordifolia* Linn.)' is also translated as Khot. *rūṇai* in the *Siddhasāra*.[71] In Tocharian, one may conjecturally reconstruct *mañcaṣṭa* (or *mañcaṣṭä*) in the lacuna (← Skt. *mañjiṣṭhā*), a term frequently found in the Weber manuscript.[72] The Tocharian noun

[67] See Maue 1990: 164 fn. 47, who also corrects Filliozat's interpretation of *nalat* as a type of *uśīra* (1948: 125; = *nadala* ?).

[68] In the Khotanese *Siddhāsara*, the usual way to translate Skt. *padmakesara* is actually *viysä khīysarä* 'lotus filament' (cf. Si § 2.8) or *viysä hīvī khīysarä* 'id.' (cf. Si § 13.33).

[69] Couvreur 1970: 177. See further CEToM s.v. *uppāl*.

[70] See Filliozat 1948: 58 and Del Tomba 2023: 110–111. Parallels in Ca.Sū.26.84(2), AS Sū. 9.2 79, and AH Sū. 7.50.

[71] For the etymology of *rūnaa-*, see Bailey *Dict.* 366.

[72] Adams translates *mañcaṣṭa* 'chay root (*Oldenlandia umbellata* Linn.)'. Conversely, he translates TB *bhant* 'Indian madder (*Rubia cordigolia*)', which corresponds to Skt. *bhaṇḍī* in the *Yogaśataka* manuscript (PK AS 2A b1). This *bhaṇḍī* is regarded as a synonym of *śirīṣa* 'siris tree (*Albizia lebbeck*)'

for (27) ‘grape’ is so far unattested in the Tocharian corpus.[73] Skt. (28) *nīlotpala* is usually translated with either *uppāl* or *nilotpāl* (cf. W 6 a2, W 10 a2, W 11 b4) in Tocharian B. The last ingredient in the formula is only partially preserved as *(w)i(ts)ako* ‘root’, which should be specified by the name of the plant lost in the lacuna. This lost noun likely corresponds to Skt. (29) *balā* ‘sida plant (*Sida cordifolia* Linn.)’ and Khot. *sacha bavā* ‘sida root’. This is usually rendered in Tocharian B by the loanwords *bal*, attested in the *Yogaśataka* (PK AS 2B a6 and b4), or *balā*, attested in W 36 a1 (cf. also *atibala* in W 24 b3). Therefore, Malzahn (*apud* CEToM) suggests restoring *bal* before *(w)i(ts)ako*. The manuscript sequence reads: *·ā* (–) – – *(w)i(ts)ako*. Although a restoration *bal̲* would fit nicely in the space of two akṣaras before *(w)i(ts)ako*, the interpretation of a compound *bal-witsako* (?) from the Tocharian perspective is unclear. A derived adjective *balṣa witsako* is a possibility. However, there may not be enough space for the *ā*-diacritic before this sequence to belong to the preceding *(nilotp)ā(l)* (lost in the lacuna).[74]

3.3. The diseases in the Sanskrit and Khotanese versions

The table below shows the correspondence between the diseases attested in the Sanskrit and Khotanese versions:

Sanskrit	Khotanese
(a) *timira*	(b) *ttamīra*
(b) *kāca*	(a) *pa tcaimañāṃ kāśaʾ*
(c) *paṭala*	(c) *tcaiña paṭarä*
(d) *arbuda*	(d) *pyāṣṭi*
(e) *jātyandhānām akṣīṇīva paśyante*	(e) *ysāna hạna hve . tcịmañī prrahīsīda vījaiṣḍi*

and *mṛdupuṣpaka* ‘id.’ in Aṣṭ.Nigh.1.84.1; in Dhanv.Nigh.1.17.2, *bhaṇḍī* is a synonym of *jiṅgī* which is one of the names of the Indian madder.

[73] See Ching 2017: 414. We only have *kuñi-mot* ‘wine’, a compound of *kuñi*° ‘pertaining to grape’ and °*mot* ‘wine’. Adams 2013: 193 suggests that *kuñi*° may be a loanword from LKhot. *gūräṇaa-* < OKhot. **gūrīṇaa-*. Dragoni 2021: 301–302 reconstructs the following diachronic path: TB *kuñi(-)* < *gūni(-)* < *gūrṇi(-)* < LKh. *gūräṇai*. A problem, however, might be the irregular reduction of *-ai* to *-i* and the palatalisation of the nasal. Dragoni suggests that the reduction might be due to the position of *gūrṇ(a)i*° as first compound member, which has in turn caused a secondary palatalisation of the nasal. Conversely, Ching 2017: 153 connects *kuñi-mot* with Chin. *pi jiu* 皮酒 ‘alcoholic drinks [brewed] in leather-[bag]’ and further links TB *kuñi* to the Khotanese, Gāndhārī, and Sanskrit words for ‘bag’.

[74] Formally, the noun phrases TB *sacchaṣṣa witsäko* (PK AS 9C b5 [archaic]) and *saccaṣṣa witsako* ‘root of *sacc(h)a*’ (PK AS 9B a6, B 497 a9, and W 14 a4) look very similar to Khot. *sacha bavā* ‘sida root’. This *sacc(h)aṣṣa* might be regarded as a loanword from Khot. *sacha* from a formal perspective. Unfortunately, we lack parallels in other languages that may shed light on the meaning of TB *sacc(h)aṣṣe.*

(f) *ardhabheda*	(f) *hālai kamala vī vīni*
(g) *adhimantha*	(g) *pejsa tcaiña rāhaʾ*
(h) *karṇa-śūla*	(h) *gvąña vīna*
(i) *ye ūrdhva-jatrujāḥ rogāḥ*	(i) *uskātta garśä bīsā āchā*
(j) *pratiśyāya*	(j) *haysgye ttajsīda*
(k) *gala-graha*	(k) *garśa rāhiʾ*
(l) *manyāstambha*	(l) *cū paysaṇve strīsīdä*
(m) *śiraḥ-śūla*	(m) *kamala-rāhiʾ*

Skt. (a) *timira*, (b) *kāca*, and (c) *paṭala* refer to three stages of the same disease which affects the pupillary region. The initial stage, *timira*, is documented in Ca.Sū.20.11 among the eighty *nānātmaja* disorders due to *vāta*. According to Su.Utt.7.5–15, illnesses affecting the first, second, and third layer (Skt. *paṭala*) of the pupil (Skt. *dṛṣṭi*) – potentially causing refraction disorders (like hypermetropia, myopia, presbyopia) – collectively fall under the term *timira* (Su.Utt.7.15c *timirākhyaḥ sa vai doṣaḥ*).[75]

The progressive development of the disease, resulting from an affliction of the fourth layer of the pupil, is then called *liṅganāśa*, *nīlīka* or *kāca* (Su.Utt.18ab *sa eva liṅganāśastu nīlikā-kāca-saṃjñitaḥ*). In AH.Utt.12.1–5, *timira* is described as a disorder affecting the first and second layers, potentially evolving into *kāca* when the *doṣa* settles in the third layer. If the disease progresses to the fourth layer, covering the entire pupil, it leads to *liṅganāśa*. AS.Utt.15.1–8ab echoes a similar description. In the Sanskrit version of the *Jīvakapustaka* (cf. also JP 32.5), the third stage of the disease is called paṭala, described as a type of eye disease in Ca.Cik.26.253, 261, 262. The *Aṣṭāṅga-saṅgraha* confirms *paṭala* as a synonym of *liṅganāśa* (AS.Utt.17.2 *liṅganāśe nīlikā paṭalam āndhyam iti paryāyāḥ*). Furthermore, the five varieties of *paṭala* listed in the *Hārītasaṃhitā* (HS 45.20) align with the corresponding varieties of *liṅganāśa* outlined in Su.Utt.87.18–24.

Thus, *timira*, *kāca*, and *paṭala* must describe degenerative stages of a single eye disease, which can be identified as cataract: from disorders of refraction and immature cataract to mature cataract, blurred vision, and near blindness. These three steps are set out in their proper order in the Sanskrit version of JP 11, but they have apparently been confused in the Khotanese version, which first gives *kāśaʾ*, the correspondent of Skt. *kāca*, and then *ttamīra* and *paṭarä*, the correspondents of *timira* and *paṭala*: IOL Khot 92/1 r1–2 *pa tcaimañāṃ kāśaʾ jaidä ttamīra jaida tcaiña paṭarä* ‘it overcomes *kāca* before the eyes, it overcomes *timira* (and) *paṭala* in the eye’.

The interpretation of *pa* in *pa tcaimañāṃ kāśaʾ jaidä* is troublesome. Hoernle (s.d.: 45 fn. 16) suggested that *pa* is short for *paṭala*. This is improbable because the equivalent

[75] Meulenbeld 1999–2002: 1a.302–303 and 1b.410 n. 224.

of Skt. *paṭala* is LKhot. *paṭarä* cited shortly after and mistakenly read by Hoernle as *papari*. Konow (1941: 44–45), followed by Ṭāme (2014: 50), emends *⁺paṃjsa* and explains it as an attribute of *āchā* 'diseases', thus *tta būri āchā jaidä ⁺paṃjsa tcaimañāṃ* 'it overcomes the following five diseases of the eyes' (cf. Ṭāme 2014: 124 این پنج بیماري چشمها را نابود میکند). However, the syntax of the sentence indicates that *pa tcaimañāṃ* must go with the following *kāśaʾ jaidä* and not with the preceding *tta būri āchā jaidä*. A possibility is to interpret *pa* as a defective spelling of nom.sg. *paṃ* 'each' < OKhot. *panä*, used as an attribute of *kāśaʾ*, thus 'it overcomes each *kāca* disease of the eye'. This interpretation implies that *kāśaʾ* was regarded as the general term for the eye disorder leading to mature cataract by the Khotanese. Still, in Sanskrit, the blanket term is generally *timira.*[76] Therefore, I have interpreted *pa* as the preposition *patä* 'before, in front of'. A similar formulation also occurs in JP 75 (106r5) *ṣiʾ cāṇä pa tcịmañāṃ ttāḍā jidä* 'this powder overcomes darkness before the eyes'.

Skt. (d) *arbuda* is 'flesh node, swelling, tumour', which, in the eye region, is said to affect eyelids (cf. Su.Utt.3.24, Su.Utt.15.29–33, AH.Utt.8.24; Meulenbeld 1999–2002: 1b.408 n. 147). In the corresponding Khotanese passage, we find *pyāṣṭi*, which is accordingly translated 'lid tumour' by Konow (1941: 25).[77] However, in Si 26.38 (Ch 149–150) *pyāṣṭāṃ āstaṃna āchā jiṃdä* '(This treatment) removes diseases such as *pyāṣṭāṃ*' corresponds to the Sanskrit compound *pilla-roga-ghnaṃ* (Tib. *mig tshag po sel to* 'it removes eye-trickle'):[78] thus, gen.pl. *pyāṣṭāṃ* is here used to translate Skt. *pilla*, which is not regarded as a synonym of Skt. *arbuda* in Indian medical texts (Emmerick *Studies* 2: 94–95).

A similar list of diseases occurs in JP 32.5ab(77r5–77v1), where Skt. *paṭala-timiraṃ kācaṃ pillaṃ* (or *paillyaṃ*)[79] *naktāndhya-arbudam*[80] is translated with Khot. *paṭala ttamīrị kāśaʾ pyāṣṭyi ca ṣīyi na vījiṣḍi* [77v5–78r1]. Also here, LKhot. *pyāṣṭyi* seems to

[76] In the *Siddhasāra*, *ttamīra-* (and variants) is clearly regarded as the first stage of the disease. The aggravation of the same disease, corresponding to *kāca* in the Sanskrit original, is rendered with *ṣiʾ pyāṃą-haṃgūstä* 'the (disease characterised by being) covered by a covering' (Si 26.45) or *tciṃʾña āchai pyāṃą* 'the "covering" disease in the eye' (Si 26.46). This peculiar rendering of Skt. *kāca* depends on the Tibetan version, which renders *kāca* with *mig nad pris g yogs pa*, translated 'covered with cream (?) eye-disease' by Emmerick 1980–1982: 2.373. As for Khotanese Buddhist texts, *ttämära-* and *kāśa-* are attested together in several passages of the *Book of Zambasta*, where *ttämära-* 'partial blindness, blurred vision' is regarded as a symptom of those men that are affected by *kāśa-* 'cataract' because of ignorance.

[77] Bailey *Dict.* 252 incorrectly takes *tcaiña paṭarä pyāṣṭi* as the Khotanese rendering of Skt. *paṭala-*. See Ṭāme 2014: 207.

[78] Translation by Emmerick 1980–1982: 2. 371.

[79] The term *paillya-* (same as *pilla*) is attested in the Bower Manuscript II.854 and 857.

[80] The reconstruction of the Sanskrit is based on Chen 2005: 345.

translate *pilla*, while *ca ṣīyi na vījiṣḍi* ‘when one does not see at night’ is a rendering of *naktāndhya-* ‘night blindness’. Curiously, the Khotanese rendition of JP 32.5 does not have any clear correspondent of Skt. *arbuda*, which is apparently left untranslated. However, in the Khotanese *Siddhasāra*, the loanword *arbauda* is generally used to translate Skt. *arbuda-* (cf. Si 26.28 *arbaudä* = Skt. *arbudam*; Si 26.51 *arbaudä* = Skt. *arbuda-*). On the basis of these pieces of evidence, Emmerick argues that LKhot. *pyāṣṭa-* is exclusively used to render Skt. *pilla-* and that “*pyāṣṭä* in JP [11] ought to correspond to the following *arbudānāṃ* of the Skt., but the Khotanese is more likely to have been based on a text containing *pilla* since it is hard to see how *pyāṣṭä* could render *arbuda* in JP [11] but *pilla* in JP [32] and Si 26.38” (Emmerick *Studies* 2: 95).

However, given the fact that both *pilla* and *arbuda* are diseases affecting eyelids, the possibility cannot be ruled out that the correspondence between Skt. *arbuda* and Khot. *pyāṣṭi* in JP 11 is the outcome of an interpretation by the Khotanese translator, who used the noun *pyāṣṭa-* as a blanket term for diseases characterised by swelling in the eyelids and left *arbuda-* in JP 32.5 untranslated because this disease was thought to be implied in the preceding *pyāṣṭyi.*[81]

The last eye disease cited in the text is congenital blindness ([e] *jātyandhānām akṣīṇīva paśyante*), which is translated by Khot. *ysāna haṇa hve . tcịmañī prrahīsīda vījaiṣḍi* ‘(As for) a man blind from birth, his eyes will be open (and) he will see’. A similar phrasing also occurs in JP 32(78r1) *cū ysāna haṇa ttyi tcịmiña prrahīsīda* ‘When one (is) blind from birth, his eyes will be open’.

Sanskrit (f) *ardhabheda* ‘hemiplegia’ is a disease of the head described in Su.Utt.25.15cd-16ab (*ardhabhedaka*). In Si 26.100 and 26.114, two types of brain disease are cited, namely *sūryāvarta-* and *ardhabheda-*. In Ca.Siddh.9.71–87, five head diseases are described, where the correspondent of *ardhabheda* is *ardhāvabhedaka*. In Khotanese, it is translated almost literally as *hālai kamala vī vīni* ‘pains (*bheda-*) at the half (*ardha*) of the head’.

The (g) *adhimantha* disease is described in chapters 9, 10, 11, and 12 of Su.Utt. as a disorder of the eye due to *vāta*, *pitta*, *kapha*, or *rakta*. In the *Siddhāsara* (§26.8), it is

[81] Eighteen varieties of the *pilla* disease are listed in AH.Utt.16.44–46 and AS.Utt.20.25. As pointed out by Meulenbeld 1999–2002: 1b.574 n. 405 this list is absent from Ca. and Su. The Tibetan version of the *Siddhāsara* uses *mig tshag* ‘eye-trickle’ as a rendering of both *klinna-vartman-* ‘excess of the lachrymal discharge’ in 26.37 and *pilla-* in 26.38, even though the watering of the eye is not a symptom of every type of *pilla*. In a similar way, Khotanese may have used *pyāṣṭa-* as a blanked term for rendering both *arbuda* and *pilla* in the *Jīvakapustaka*, on account of some similar characteristics of the two diseases. If Khot. *pyāṣṭi* designates a disease characterised by swelling in the eyelids, Mauro Maggi (p.c.) suggests a fitting etymology for this noun: OKhot. *pätālsta-* ‘increased, grown (eye/eyelid)’ (cf. SGS: 82) > LKhot. *pyāṣṭa-* ‘swelling (in the eyelid)’ (cf. the directional suffix OKhot. *-ālsto* > LKhot. *-āṣṭa*).

described as an aggravation of the *abhiṣyanda* disease (i.e. conjunctivitis) and is said to cause severe pain in half of the head and the eyes (*tīvra-mūrdhārdha-netrārtir*). It is translated by Khot. *pejsa tcaiña rāhaˀ* 'severe pains in the eye', which is indeed one of the symptoms of the *adhimantha* disease (AH.Utt.15.3–4).

Sanskrit (h) *karṇa-śūla* is 'earache', rendered in Khotanese with the loan translation *gvąña vīna* 'pain in the ear'.

Sanskrit (i) is *ye ūrdhva-jatrujāḥ rogāḥ* 'diseases which are located above the collar bone', generally said to be caused by disorders of *vāta* (cf. e.g. Ca.Sū.5.62.2; Su.Nid.1.15.2). In Khotanese, it is translated almost literally as *uskātta garśä bīsā āchā* 'diseases (*rogāḥ*) located (*-jāḥ*) above (*ūrdhva-*) the throat (~ *jatru-*)' (see further Ṭāme 2014: 207–208).

As for Sanskrit (j) *pratiśyāya* 'coryza, cold, nasal catarrh', it is described in Ca.Sū.19.4, Ca.Cik.26.104–117, and Su.Utt.24 as having four or five varieties, depending on the number of *doṣa* involved and on the aggravation of the disease. It is translated by Khot. *haysgye ttajsīda*, which literally means 'the nostrils flow'. In the Khotanese *Siddhasāra*, Skt. *pratiśyāya* is always translated with *mālaiga* (Si 2.24, 26.78; Tib. *cham pa*). Khot. *mālaiga* also occurs in JP 45(87r5) as a rendering of Skt. *prattiśū*, which may be regarded as a corrupted form of either *pratiśyāya* or *pratiśyā*, with the same meaning. However, the compound *haysga-ttājsį* 'runny nose (disease)'[82] is also used to translate Skt. *pīnasa* (112v5) in JP 85(113r4). The disease called *pīnasa* is regarded as an aggravation of *pratiśyāya*, as it is a type of cold that appears when all three *doṣa*s are involved (Ca.Cik.26.104–107.1; Su.Utt.25.22–24). Thus, both *mālaiga* and *haysgye ttajsīda ~ haysga-ttājsį* are Khotanese renderings of the *pratiśyāya* disease and its aggravation (*pīnasa*).

Sanskrit (k) *galagrahaṃ* 'compression on the throat' and (m) *śiraḥ-śūla* 'headache' correspond, respectively, to the noun phrase *garśa rāhiˀ* 'pain of the throat' and the compound *kamala-rāhiˀ* 'headache'.

Finally, the Khotanese equivalent of the penultimate disease requires some explanation. Sanskrit (l) *manyāstambha*, a compound of *manyā-* 'back or nape of the neck' (MW: 789) and *stambha-* 'stiffness, rigidity, becoming hard or solid', refers to the torticollis, wry neck, stiffness of the carotid regions (Meulenbeld 1999–2002: 1b.137 n. 789). In the Khotanese translation, it is rendered with the sentence *cū paysaṇve strīsīdä*. Konow (1941: 25) translates it as 'when the cheeks become stiff'. The challenge lies in the interpretation of the Khotanese noun *paysanuā-*, which is not a common word. Its etymology is clear: it is derived through the prefix *pati-* 'beside' from the pre-form of Khot. *ysanuā-* 'jawbone' < **zanukā-* (cf. ManMP *dnwg* [South-Western form], Sogd.

[82] See Degener 1989: 6, who interprets °*ttājsį* as a derivative of *ttajs-* 'to flow' with the *-ǎ-* suffix.

znwq, Balochi *zanūk*, YAv. *zanu-* [f.]).[83] Except for the single occurrence in JP 11, it is attested four times in the entire Khotanese corpus:

- nom.sg. IOL Khot 37/14 r3 (KT 3: 130, *Cat.* 260): *ci syaṃca paysanva spaʾtti bvāṃñä* 'he whose left *paysanva* twitches shall know ...'.
- gen.sg. Si 26.12 (Ch 146r5): *tciṃ vī ā̧na paysaṇu vīrāṣṭä aharṣṭä bịjsā̧ñä* '(The eye lotion) must be poured uninterruptedly from the eye towards the *paysaṇu*'.
- nom.pl. P 2956 l.47 (KT 3: 38) *pą̣ysaṇvakyä* = P 2025 l.68 (KT 3: 48) *paysaṇvakye*.

The last occurrence is from the so-called *Lyrical Poem*, a very difficult text, which is of little help for understanding the meaning of the word.

IOL Khot 37/14 r3 is described by Skjærvø (*Cat*. 260) as giving the interpretation of bodily signs. The progression of the body parts in the text indicates that *paysanva* likely describes a body part located on the face. Its occurrence in the *Siddhasāra* suggests it should either refer to the cheek or to the jaw. Indeed, Khot. *tciṃ vī ā̧na paysaṇu vīrāṣṭä aharṣṭä bịjsā̧ñä* must translate Tib. *mig 'gram logs nas rgyun ma chad par blugs pa ni* 'pours (the liquid) into the eye without interruption from the side of the cheek'.[84] The Tibetan noun *'gram* actually means 'bank, shore, check, jaw' and it is used in the *Siddhasāra* to translate Skt. *gaṇḍa-* 'cheek', *hanu-* 'jaw', *śaṅkhaka-* 'jawbone'. The same range of meanings could be equally applied to Khot. *paysanuā-*, which may mean both 'jaw' and 'cheek'.[85] I am therefore not inclined to translate *paysaṇve* as 'cheeks' in *cū paysaṇve strīsīdä*, since the stiffness of the cheeks is not a pain described in Āyurvedic texts. If *cū paysaṇve strīsīdä* in JP 11 is a close rendering of *manyāstambha*, then *paysaṇve* could be more generally translated 'sides of the neck', corresponding to the sternomastoid region. However, one should bear in mind the peculiar spelling of the form in the manuscript, i.e. *minnyāstaṃbha* (or *mittyāstaṃbha*) with the duplication of the *n*. It can be assumed that the Khotanese copyist or the original translator was unfamiliar with this Sanskrit technical term. Therefore, he not only miswrote the word but also misinterpreted its meaning, likely by conflating it with other conditions with which he was more familiar, such as *ardita* 'facial paresis' or, more probably, *hanugraha* ~ *hanusaṃgraha* 'lockjaw'.[86]

---

[83] Bailey *Dict*. 213. That Khot. *ysanuā-* < **zanukā-* should be regarded as a *uā*-stem is clearly demonstrated by the feminine agreement in Z 20.40, cf. *śśīya śśo ysanuva śśäte* 'there lies one white.F jawbone'.

[84] Edition and translation of the Tibetan by Emmerick 1980–1982: 2.362–363.

[85] Cf. Konow 1941: 94. Bailey 1958 originally translated *paysanva* as 'corner of mouth'.

[86] Note that *manyāstambha*, *śiraḥśūla*, *ardita*, and *hanusaṃgraha* are cited one after the other in Ca.Sū.5.59–60.

4. Conclusion

This article has presented a revised edition, translation, and commentary on the prescription for a medicated ghee known as *Mahāvaidehaghṛta* in *Jīvakapustaka* 11, which is unique in having a Tocharian B parallel. Building on the work of Hoernle and Emmerick, it has proposed a methodology for reconstructing the highly corrupt Sanskrit version by integrating careful observance of its underlying metrical structure, due consideration of the amplified Khotanese rendition, and full account of the distorting effect of the spelling habits of the copyist. The study has shown that the "barbarous" quality of the Sanskrit text likely reflects the influence of a non-native Sanskrit speaker, possibly the Late Khotanese scribe who also authored the Khotanese section. This scribe adapted Sanskrit words to his own writing conventions and phonological system. The article has further provided the first systematic classification of the graphic, phonological, and grammatical errors made by the Late Khotanese scribe, demonstrating their relevance for reconstructing aspects of Late Khotanese phonology. Finally, it has addressed the identification of the Sanskrit mantra at the beginning of the Tocharian manuscript, presented a comprehensive linguistic, textual, and semantic analysis of the ingredients across the three versions, along with an analysis of the diseases in the Sanskrit and Khotanese texts.

Abbreviations

Grammatical and language abbreviations

1, 2, 3	first, second, third person	opt.	optative
acc.	accusative	pass.	passive
act.	active	pl.	plural
caus.	causative	prs.	present
f.	feminine	PT	Proto-Tocharian
gen.dat.	genitive-dative	sg.	singular
instr.abl.	instrumental-ablative	Skt.	Sanskrit
LKhot.	Late Khotanese	Sogd.	Sogdian
loc.	locative	TA	Tocharian A
ms.	manuscript	TB	Tocharian B
nom.	nominative	Tib.	Tibetan
OKhot.	Old Khotanese		

Text and shelfmark abbreviations

AH	*Aṣṭāṅgahṛdayasaṃhitā*, edition by Das and Emmerick 1998
AS	*Aṣṭāṅgasaṅgraha*, edition by Āṭhavale 1980

Aṣṭ.Nigh.	*Aṣṭāṅganighaṇṭu*, edition by Sharma 1973
Ca.	*Carakasaṃhitā*, edition by Sharma 2014 and 1998
Cik.	*Cikitsāsthāna* (section)
Dhanv.Nigh.	*Dhanvantarīyanighaṇṭu*, edition by Puraṃdare 1925
Ka.	*Kalpasthāna* (section)
JP	*Jīvakapustaka*, partial edition of the Skt. version by Hoernle s.d.; complete edition of the Skt. and Khot. versions by Bailey KT 1: 135–195; translation of the Skt. version by Chen 2005; partial edition and translation of the Khot. version by Hoernle s.d.; complete edition and translation of the Khot. version by Konow 1941; complete edition and translation of the Khot. version by Ṭāme 2014
HS	*Hārītasaṃhitā*, edition by Raison 1974
IOL Khot	Khotanese manuscript folios in the India Office Library, London
IOL Toch	Tocharian manuscript folios in the India Office Library, London
Nid.	*Nidanasthāna* (section)
Nigh.Śeṣa	*Nighaṇṭuśeṣa*
PiŚ	*Piṇḍaśāstra*, edition and translation by Luzzietti 2023
PK AS	Pelliot Koutchéen, Ancienne Série, Paris
PK NS	Pelliot Koutchéen, Nouvelle Série, Paris
Rāj.Nigh.	*Rājanighaṇṭu*, edition by Puraṃdare 1925
Si	*Siddhasāra*, edition of the Skt. original by Emmerick 1980–1982, vol. 1; edition and translation of the Tib. version by Emmerick 1980–1982, vol. 2; edition and translation of the Khot. version by Emmerick unpublished
Si.Nigh.	*Siddhasāranighaṇṭu*, edition by Emmerick 1980–1982, vol. 1, pp. 177–194
Siddh.	*Siddhisthāna* (section)
Su.	*Suśrutasaṃhitā*, edition by Sharma 1999–2001
Sū.	*Sūtrasthāna* (section)
Utt.	*Uttaratantra* (section)

Bibliography

ADAMS, Douglas Q., 2013. *A dictionary of Tocharian B: Revised and greatly enlarged*, 2 vols, Amsterdam–New York.

ĀṬHAVALE, Ananta Dāmodara (ed.), 1980. *Śrīmadvṛddhavāgbhaṭaviracitaḥ Aṣṭāṅgasaṅgrahaḥ Induvyākhyāsahitaḥ* [The *Aṣṭāṅgasaṅgraha* by the illustrious Vṛddha Vāgbhaṭa, accompanied by Indu's commentary], repr. Pune (1st ed. 1902).

BAILEY, Harold W., 1937. 'Ttaugara'. *BSOS* 8/4, 883–921.

BAILEY, Harold W., 1938. *Codices Khotanenses: India Office Library Ch. ii 002, Ch. ii 003, Ch. 00274 reproduced in facsimile with an introduction*, Copenhagen.

BAILEY, Harold W., 1940. 'Ttāgutta', *BSOS* 10/3, 599–605.

BAILEY, Harold W., 1958. 'Languages of the Saka', in *Iranistik*, *IV:1 Linguistik* (ed. Karl HOFFMANN *et al*.), Leiden–Köln, 131–154.

BAILEY, Harold W., 1982. *The culture of the Sakas in Ancient Iranian Khotan*, Delmar–New York.

BAUMS, Stefan, 2009. *A Gāndhārī commentary on Early Buddhist verses: British Library Kharoṣṭhī fragments 7, 9, 13 and 18*. Diss. University of Washington.

BHSD = Edgerton, Franklin, 1953. *Buddhist Hybrid Sanskrit grammar and dictionary*. Vol. 2: *Dictionary*, New Haven, CT.

BLAŽEK, Václav & Michal SCHWARZ, 2015. 'Prehistory of Tocharian *yāstaci* 'Juniper' in the perspective of historical phonetics', *Journal of Indo-European Studies* 43/3–4, 423–431.

BOESI, Alessandro, 2014. 'The materia medica of Tibetan medicine according to practitioners from the lower Mustang district in Nepal', *Rivista di Studi Orientali* 87, 73–104.

BURROW, Thomas, 1937. *The language of the Kharoṣṭhi documents from Chinese Turkestan*, Cambridge.

CARLING, Gerd, 2003. 'Fragments bilingues du Yogaśataka. Révision commentée de l'édition de Jean Filliozat', *TIES* 10, 75–79.

CARLING, Gerd, 2007. 'The vocabulary of Tocharian medical manuscripts', *Asian Medicine* 3, 323–333.

CARLING, Gerd & Georges-Jean PINAULT, 2023. *Dictionary and thesaurus of Tocharian A*, Wiesbaden.

Cat. = SKJÆRVØ, Prods Oktor, 2002. *Khotanese manuscripts from Chinese Turkestan in the British Library: A complete catalogue with texts and translations*, with contributions by Ursula SIMS-WILLIAMS, London.

CEToM = *A comprehensive edition of Tocharian manuscripts*, available at cetom.univie.ac.at (accessed in October 2023).

CHEN Ming 陳明, 2005. *Dunhuang chutu Hu yu yi dian "Qipo shu" yanjiu* 敦煌出土胡語醫典《耆婆書》研究 [*A study on Sanskrit text of Jīvaka-pustaka from Dunhuang*], Taipei.

CHING Chao-jung 庆昭蓉, 2017. *Tuhuoluo yu shisu wenxian yu gudai Qiuci lishi* 吐火罗语世俗文献与古代龟兹历史 – *Tocharian secular texts and the history of Ancient Kucha*, Beijing.

CHING Chao-jung 庆昭蓉 & ENAMI Kazuyuki 江南和幸, 2018. 'Tangdai Anxi da duhufu shiqi zhi Qiuci dangdi yong zhi—Riben Longgu Daxue suo cang Kuche chutu Hanwen shu anli yanjiu zhi yi 唐代安西大都护府时期之龟兹当地用纸—日本龙谷大学所藏库车出土汉文书案例研究之一 – Paper used in the 8th-century Anxi Protectorate, Part 1: Some observations on the Chinese document fragments kept in Ryukoku University (Japan) unearthed from Kucha', *XYWS* 12, 157–176.

CHING Chao-jung 慶昭蓉, ENAMI Kazuyuki 江南和幸 & OKADA Yoshihiro 岡田至弘, 2021. 'Paper in eighth-century Kucha: Discovery of cotton fibres within Chinese and Kuchean documents', in *CAJ* 63/1–2 (2020), 71–104.

COUVREUR, Walter, 1970. 'Boeddhistische Sanskritfragmenten in Koetsjische handschriftenverzamelingen', in *ANAMNHCIC. Gedenkboek Prof. Dr. E.A. Leemans*, Brugge, 175–184.

DAS, Rahul Peter & Ronald Eric EMMERICK (ed.), 1998. *Aṣṭāṅgahṛdayasaṃhitā: The romanised text accompanied by line and word indexes*, Groningen.

DE CHIARA, Matteo, 2013. *The Khotanese Sudhanāvadāna*, Wiesbaden.

DEGENER, Almuth, 1989. *Khotanische Suffixe*, Stuttgart.

DEL TOMBA, Alessandro, 2018. 'On the pronominal feminine plural in Tocharian', *IF* 123, 340–356.

DEL TOMBA, Alessandro, 2023. *The Tocharian gender system: A diachronic study in nominal morphology*, Leiden–Boston.

Dict. = BAILEY, Harold Walter, 1979. *Dictionary of Khotan Saka*, Cambridge.

DRAGONI, Federico, 2021. 'Materia medica Tocharo-Hvatanica', *BSOAS* 84/2, 295–319.

DRAGONI, Federico, 2023. Watañi lāntaṃ: *Khotanese and Tumshuqese loanwords in Tocharian*, Wiesbaden.

DRESDEN, Mark J., 1955. *The Jātakastava or 'Praise of the Buddha's former births': Indo-Scythian (Khotanese) text, English translation, grammatical notes and glossaries*, Philadelphia.

EMMERICK, Ronald E., 1970. *The Khotanese Śūraṇgamasamādhisūtra*, London.

EMMERICK, Ronald E., 1979a. 'Contributions to the study of the Jīvaka-pustaka', *BSOAS* 42/2, 235–243.

EMMERICK, Ronald E., 1979b. 'The vowel phonemes of Khotanese', in *Studies in diachronic, synchronic, and typological linguistics: Festschrift for Oswald Szemerényi on the occasion of the 65th birthday* (ed. Bela BROGYANYI), Amsterdam, 239–250.

EMMERICK, Ronald E., 1980. *The Siddhasāra of Ravigupta*. Vol. 1: *The Sanskrit text*, 1980; *The Siddhasāra of Ravigupta*. Vol. 2: *The Tibetan version with facing English translation*, Wiesbaden.

EMMERICK, Ronald E., 1989. 'Khotanese and Tumshuqese', in *Compendium Linguarum Iranicarum* (ed. Rüdiger SCHMITT), Wiesbaden, 204–229.

EMMERICK, Ronald E., 1992a. 'The Svastika antidote', *Journal of the European Āyurvedic Society* 2, 60–81.

EMMERICK, Ronald E., 1992b. *A Guide to the literature of Khotan*, second edition thoroughly revised and enlarged, Tokyo.

EMMERICK, Ronald E., 1994. 'The Mahāsauvarcalādi ghee', in *Memoriae Munusculum: Gedenkband für Annemarie v. Gabain* (ed. Klaus RÖHRBORN & Wolfgang VEENKER), Wiesbaden, 29–42.

EMMERICK, Ronald E., 1997. 'The Mahāsauvarcalādi ghṛta in Hoernle's unpublished edition of the "Jīvakapustaka"', *Journal of the European Āyurvedic Society* 5, 76–81.

EMMERICK, Ronald E., 2009. 'Khotanese and Tumshuqese', in *The Iranian languages* (ed. Gernot WINDFUHR), London, 377–415.

EMMERICK, Ronald E., unpublished. *Edition and translation of the Late Khotanese Siddhasāra*, unpublished manuscript.

EMMERICK, Ronald E. & Edwin G. PULLEYBLANK, 1993. *A Chinese text in Central Asian Brāhmī script*, Roma.

EWAia = MAYRHOFER, Manfred, 1992–2001. *Etymologisches Wörterbuch des Altindoarischen*, 3 vols, Heidelberg.

FILLIOZAT, Jean, 1946–1947. 'Review of Bailey, *Khotanese texts* Vol. 1.', *JA* 235, 134–135.

FILLIOZAT, Jean, 1948. *Fragments de Textes Koutchéens de Médecine et de Magie*, Paris.

FILLIOZAT, Jean, 1979. *Yogaśataka: Texte médical attribué à Nāgārjuna*, Pondichéry.

VON HINÜBER, Oskar, 2001. *Das ältere Mittelindisch im Überblick*, Wien.

HITCH, Doug, 2016. *The Old Khotanese metanalysis*, Diss. Harvard University.

HOERNLE, A. F. Rudolf, 1908. *The Bower manuscript: Facsimile leaves, Nagari transcript, romanised transliteration and English translation with notes. General Sanskrit index*, Calcutta.

HOERNLE, A. F. Rudolf, s.d. *An ancient medical manuscript from Eastern Turkestan. [Ch. ii. 003]. Press-copy*, unpublished manuscript housed at the British Library, London, MSS. EUR.D.723/22.

ISEBAERT, Lambert, 1980. *De Indo-Iraanse bestanddelen in de Tocharische woordenschat. Vraagstukken van fonische productinterferentie, met bijzondere aandacht voor de Indo-Iraanse diafonen a, ā*, Diss. Katholieke Universiteit Leuven.

KBT = BAILEY, Harold W., 1981. *Khotanese Buddhist texts*, 2nd edition, Cambridge.

KHARE, C. P., 2007. *Indian medicinal plants: An Illustrated Dictionary*, Berlin–Heidelberg.

KONOW, Sten, 1941. *A medical text in Khotanese: Ch. ii 003 of the India Office Library with translation and vocabulary*, Oslo.

KT 1–3 = BAILEY, Harold Walter, 1945–1956. *Khotanese texts*, vol. 1, 1945; vol. 2, 1954; vol. 3, 1956 (vols 1–3, 2nd ed. in one vol., 1969; repr. 1980), Cambridge.

KUMAMOTO Hiroshi 熊本裕, 1988. 'Seiiki ryokōsha yō sansukuritto-kōtango kaiwa renshūchō 西域旅行者用サンスクリット=コータン語會話練習帳 – A Sanskrit-Khotanese conversation manual for Central Asian travelers', *SAK* 28, 53–82.

KUMAMOTO Hiroshi, 1990. 'A Sanskrit-Khotanese conversation manual for Central Asian travellers', in *Documents et archives provenant de l'Asie centrale: Actes du colloque franco-japonais de Kyōto 4–8 octobre 1988* (ed. HANEDA Akira), Kyoto, 71–86.

KUMAMOTO Hiroshi, 1995. 'Did Late Khotanese have a three vowels system?', in *Proceedings of the second European conference of Iranian studies held in Bamberg, 30th September to 4th October 1991, by the Societas Iranologica Europaea* (ed. Bert G. FRAGNER *et al.*), Roma, 383–390.

KUMAMOTO Hiroshi, 1996. 'The Khotanese documents from the Khotan area', *Memoirs of the Research Department of the Toyo Bunko* 54, 27–64.

LUZZIETTI, Silvia, 2023. *Piṇḍaśāstra, an Āyurvedic Khotanese text: Critical edition, translation, commentary, glossary, and study of sources*, Diss. Sapienza Università di Roma.

MAGGI, Mauro, 2009. 'Khotanese literature', in *The literature of pre-Islamic Iran. Companion volume 1 to A history of Persian literature* (ed. Ronald E. EMMERICK & Maria MACUCH), London, 330–417.

MAGGI, Mauro, 2018. 'Bits and bites: The Berlin fragment bi 43 and Khotanese **druṣ*-', *ARIRIAB* 22, 247–260, plates 4–5.

MAGGI, Mauro, 2020–2021. 'Sanscrito *hapuṣā*- 'ginepro', un pracritismo gāndhārī di origine iranica nel lessico āyurvedico', *Die Sprache* 54, 1–12.

MAGGI, Mauro, 2021. 'Some remarks on the history of the Khotanese orthography and the Brāhmī script in Khotan', *ARIRIAB* 25, 149–172.

MAGGI, Mauro, 2022. 'Light on the *Siddhasāra* from the *Jīvakapustaka*: The *Hapuṣādyaghṛta* in Khotanese (JP 26) and its Indian sources', *Linguistica e Filologia* 42, 121–136.

MAGGI, Mauro, 2024. 'Chapter 30. The Buddha's preaching of the Law to his father in the *Mañjuśrīnairātmyāvatārasūtra*', in Ronald E. EMMERICK, *A handbook of Khotanese*, edited with supplements by Mauro MAGGI *et al.*, Wiesbaden.

MAUE, Dieter, 1990. 'Das *Mahāvaidehikaṃ ghṛtam* in Tocharisch B', *Historische Sprachforschung* 103/1, 159–165.

MAUE, Dieter, 2008. 'An Uighur version of Vāgbhaṭa's Aṣṭāṅgahṛdayasaṃhitā', *Asian medicine* 4, 113–173.

MAUE, Dieter & Nicholas SIMS-WILLIAMS, 1991. 'Eine Sanskrit-Sogdische Bilingue in Brāhmī', *BSOAS* 54/3, 486–495.

MEULENBELD, Gerrit Jan, 1974. *The Mādhavanidāna and its chief commentary. Chapters 1–10: Introduction, translation and notes*, Leiden.

MEULENBELD, Gerrit Jan, 1999–2002. *A history of Indian medical literature*, 5 vols, Groningen.

MW = MONIER-WILLIAMS, Monier, 1899. *A Sanskrit-English dictionary: Etymologically and philologically arranged with special reference to cognate Indo-European languages*, new edition, greatly enlarged and improved, Oxford.

NADKARNI, Krishnarao Mangeshrao, 1954. *Indian Materia Medica, with Ayurvedic, Unani and home remedies*, revised and enlarged by A. K. Nadkarni, 2 vols, Bombay.

PAN Tao 潘涛 & CHEN Ruixuan 陈瑞翾, 2021. 'The tip of an iceberg: Traces of Chinese Buddhist scrolls in fragments of Tocharian poṭhīs', *ARIRIAB* 24: 187–215.

PEYROT, Michaël, 2008. *Variation and change in Tocharian B*, Amsterdam–New York.

PEYROT, Michaël, 2021. 'On a festive occasion: The inflection of Tocharian B *ekṣalye* and morphologically related words', in *Lyuke wmer ra. Indo-European studies in honour of Georges-Jean Pinault* (ed. Hannes A. FELLNER *et al.*), Ann Arbor, 450–461.

PINAULT, Georges-Jean. 2008. *Chrestomathie tokharienne: textes et grammaire*, Leuven–Paris.

PURAṂDARE (ed.), 1925. *Rājanighaṇṭusahito Dhanvantarīyanighaṇṭuḥ* [The *Dhanvantarīyanighaṇṭu* combined with the *Rājanighaṇṭu*], 2nd ed., Poona.

RAISON, A., 1974. *La Haritasamhita, texte medical sanskrit avec un index de nomenclature ayurvedique*, Pondichéry.

RECK, Christiane & Adam BENKATO, 2018. ‘‘Like a virgin’: A Sogdian recipe for restoring virginity and the Sanskrit background of Sogdian medicine’, *WMO* 8/2, 67–91.

RECK, Christiane & Jens WILKENS, 2015. ‘Soghdische Versionen von Vāgbhaṭas *Aṣṭāṅgahṛdayasaṃhitā*’, in *Kutadgu nom bitig: Festschrift für Jens Peter Laut zum 60. Geburtstag* (ed. Elisabetta RAGAGNIN & Jens WILKENS), Wiesbaden, 431–452.

SGS = EMMERICK, Ronald E., 1968a. *Saka grammatical studies*, London.

SHARMA, Priya Vrat (ed.), 1973. *Aṣṭāṅga Nighaṇṭu of Vāhaṭācārya*, Madras.

SHARMA, Priya Vrat (ed. and trans.), 1998. *Caraka-saṃhitā: Agniveśa’s treatise refined and annotated by Caraka and redacted by Dṛḍhabala: Text with English translation*. Vol. 2: *Cikitsāsthāna to Siddhisthāna*, Varanasi.

SHARMA, Priya Vrat (ed. and trans.), 1999–2001. *Suśruta-Saṃhitā*, *with English translation of text and Ḍalhaṇa’s commentary along with critical notes*, 3 vols, Varanasi.

SHARMA, Priya Vrat (ed. and trans.), 2014. *Caraka-saṃhitā: Agniveśa’s treatise refined and annotated by Caraka and redacted by Dṛḍhabala: Text with English translation*. Vol. 1 (rev. ed.): *Sutrasthāna to Indriyasthāna*, Varanasi.

SKJÆRVØ, Prods Oktor, 2022. ‘Khotan IV: The Khotanese language’, *Encyclopædia Iranica* 17/2, 117–133.

Studies = EMMERICK, Ronald E. & Prods Oktor SKJÆRVØ, 1982–1997. *Studies in the vocabulary of Khotanese*, 3 vols, Wien.

ṬĀME, Majid مجید طامه (ed. and trans.), 2014. *Barresi va tarjome-ye matn-i pezeški be zabān-e xotani (Jivaka-pustaka)* (جیوَکه پوستَکه) بررسی و ترجمهٔ متنی پزشکی به زبان ختنی [= *An examination and translation of a medical text in Khotanese (Jīvaka-Pustaka)*], Diss. Tehrān: Dānešgāh-e Tehrān, 1393 h.š.

ZEISLER, Bettina, 2009–2010. ‘East of the Moon and west of the Sun? Approaches to a land with many names, north of ancient India and south of Khotan’, *The Tibet Journal* 34/3–4 and 35/1–2: 371–463.

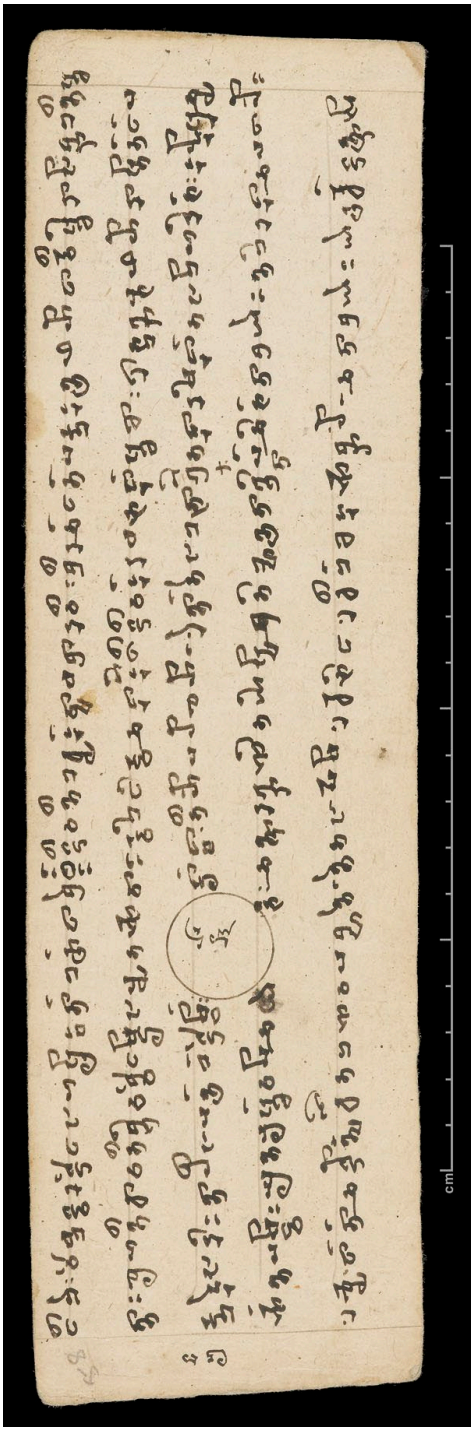


Fig. 1. IOL Khot 91/3 recto.
© British Library Board.

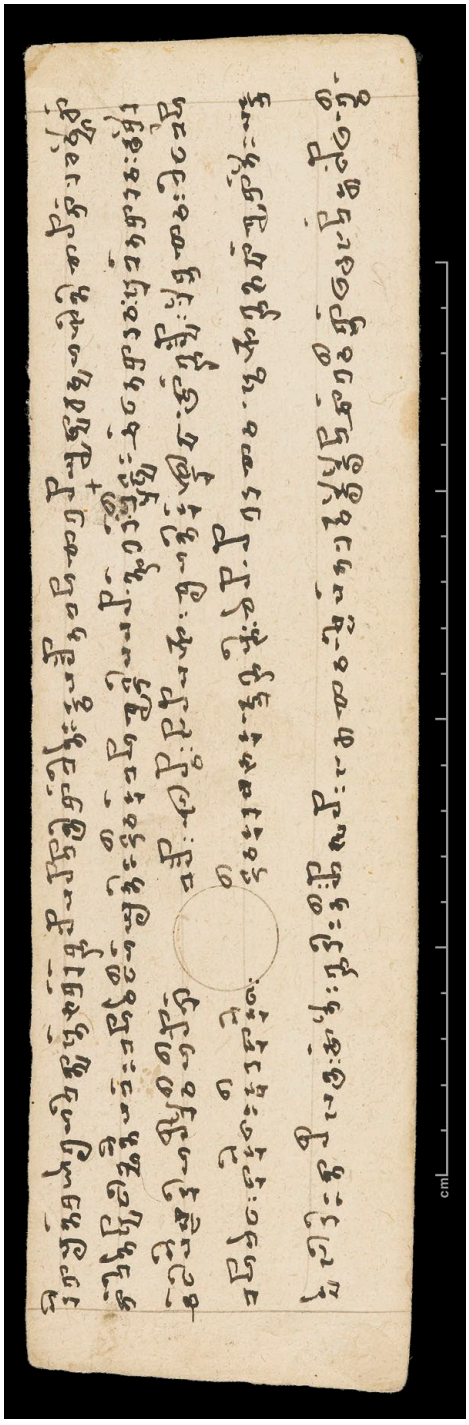


Fig. 2. IOL Khot 91/3 verso.
© British Library Board.

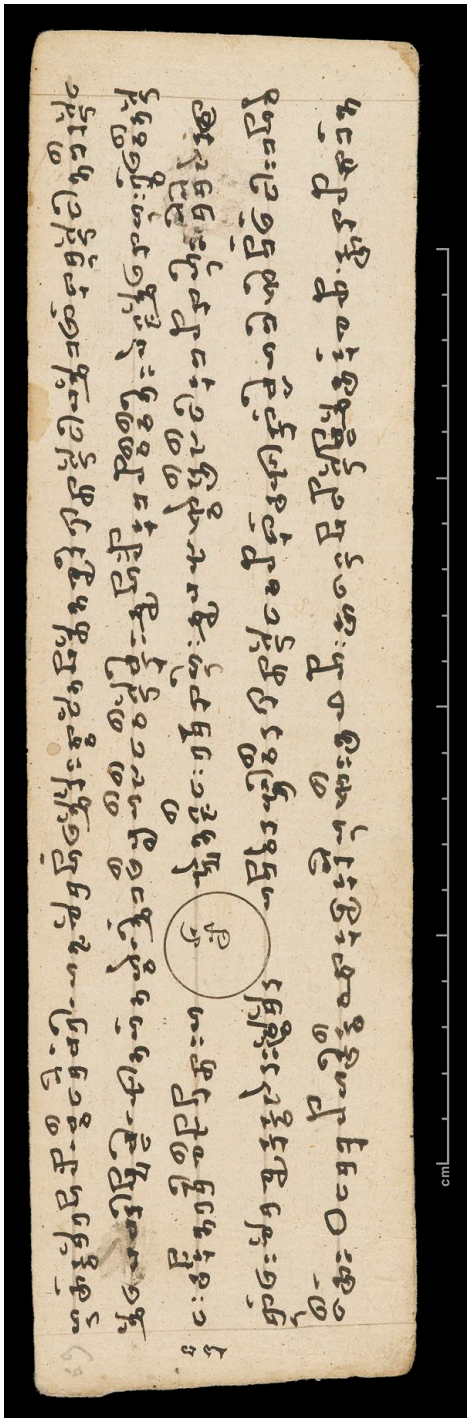


Fig. 3. IOL Khot 92/1 recto.
© British Library Board.

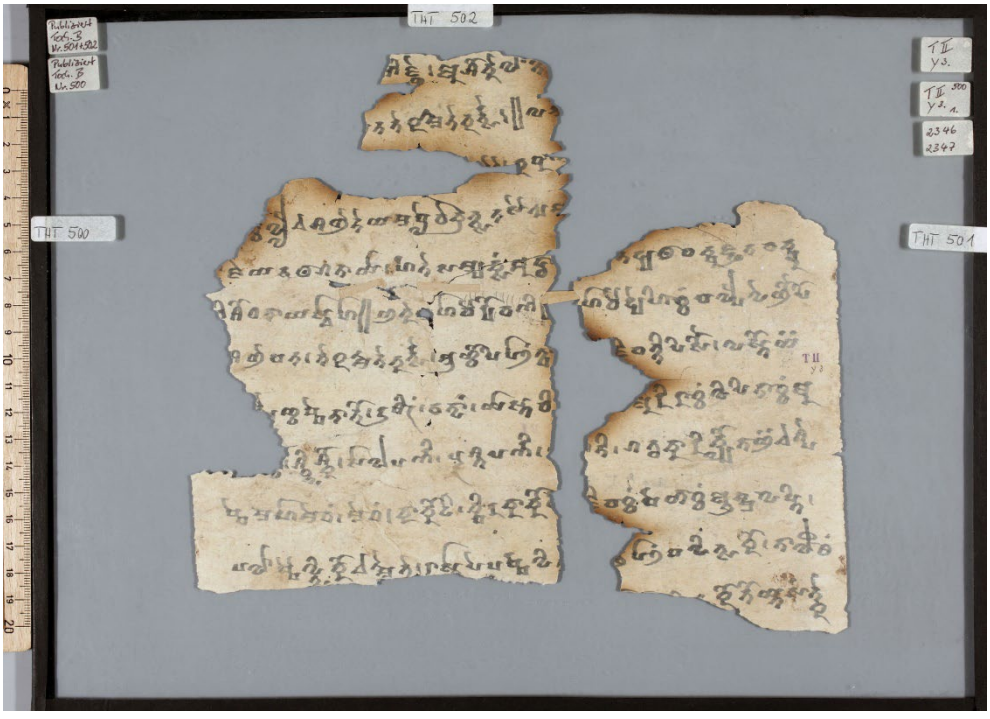


Fig. 4. THT 500+501+502.

© Depositum der Berlin-Brandenburgischen Akademie der Wissenschaften in der Staatsbibliothek zu Berlin – Preussischer Kulturbesitz Orientabteilung.

Was the Khotanese Brāhmī subscript hook borrowed from the Kharoṣṭhī anusvāra?*

Federico DRAGONI

The sign inventory of the Brāhmī script used to write down Khotanese, an eastern Middle Iranian language once spoken in the southwest of the Tarim Basin, includes three additional diacritics. These are the two dots over an akṣara, the St. Andrew's cross, and the subscript hook. The first two diacritics have already been detected in several Kharoṣṭhī documents from the Khotan area. In Khotanese, they represent respectively [ə] and [aə]. The subscript hook, on the other hand, seems to be restricted to Khotanese Brāhmī, but its specific linguistic function, palaeographical developments, and ultimate origin are still a matter of debate among scholars. This article offers a critical overview of the relevant scientific literature and introduces the results of a preliminary palaeographical analysis of the diacritic. The analysis shows that the subscript hook used in Khotanese Brāhmī was borrowed from the Kharoṣṭhī anusvāra. The proposed borrowing trajectory sheds light on the linguistic function of the diacritic as well.

1. The problem

The subscript hook is a diacritic belonging to the sign inventory of the Khotanese Brāhmī script. It owes its name to the fact that it resembles a "hook written below an akṣara" (HK: 281). Its linguistic function(s) and graphic origin are disputed. Emmerick (1992, HK: 281) argues that it indicates breathy voice and derives from a cursive variant of the akṣara *ha*. Skjærvø (2022: 121–122) suggests in passing that it might represent vowel rhotacization and derive from the Tumshuqese "foreign" sign x_5. In this paper, I revisit the whole argument from a palaeographical point of view.

In §2 and §3, I introduce the Khotanese language and its script. In §4, I identify the earliest attested shapes of the subscript hook by testing Maggi's (2022) palaeographical and orthographic classification of the oldest written Khotanese witnesses. §5 discusses

* This research was made possible thanks to the European Research Council (ERC) Consolidator Grant project "The Silk Road Language Web" (Grant agreement ID: 101088902). A preliminary sketch of some of the ideas contained in this paper was presented at the 34th Deutscher Orientalistentag (Freie Universität Berlin, September 16, 2022). For important comments and suggestions, I am grateful to the editors of this volume, Ching Chao-jung and Michaël Peyrot, as well as to Stefan Baums, Alessandro Del Tomba, Mauro Maggi, Dieter Maue, and Niels Schoubben.

10.29091/9783752003635/003

previous scholarship on the subscript hook and introduces the palaeographical analysis of §6. Through an extensive collection of the different shapes of the diacritic covering the whole palaeographical history of Khotanese, I show that both Emmerick's and Skjærvø's graphic derivations cannot stand closer scrutiny (§6). Based on the palaeographical analysis, and after considering the history of the other Khotanese Brāhmī diacritics (§7), I suggest that the subscript hook may have been borrowed from the Kharoṣṭhī anusvāra (§8.1). Furthermore, I consider some linguistic arguments in support of the proposed borrowing trajectory that may contribute to a better comprehension of the phonetic value of the diacritic (§8.2). The results of the study are summarised in §9. An appendix provides information on the sources of the akṣara images, their transliteration and digital processing.

2. The Khotanese language and its stages

Khotanese is an eastern Middle Iranian language once spoken in the southwest of the Tarim Basin. Materials written in Khotanese have been found in the Khotan area (around today's Hetian 和田), and in the Cave of the Thousand Buddhas in Dunhuang 敦煌, where a Khotanese community was active in the 10th cent. CE (Kumamoto 1996). A handful of isolated fragments have also been found in the northwest of the Tarim Basin, in the Tocharian-speaking areas of Kucha and Šorčuq (Dragoni 2023: 233–235).

Scholars traditionally distinguish between two stages of the language, conventionally termed Old and Late Khotanese. Old Khotanese includes the oldest written materials from the Khotan area alone, mainly Buddhist religious texts (approximately from the 5th century onwards) and is therefore considered the "sacred language of Khotanese Buddhism" (Maggi 2009: 333). Late Khotanese is attested from manuscripts of both religious and secular content from the Khotan area (until the 9th cent.) and from Dunhuang (10th cent.). It is likely that the absence of Old Khotanese texts in Dunhuang is due to radical cultural changes in Khotan because of Tibetan presence in the area (8th–9th cent. CE, see Kumamoto 2012: 159).

It is less certain to what extent Late Khotanese texts from Khotan differed (socio)-linguistically from the Dunhuang corpus, and whether the label "Middle Khotanese" for Khotanese texts of 7th and 8th century Khotan is linguistically justified (KMB: lxx, Skjærvø 2022: 120–121; see also Table 1 in this paper).

3. The Brāhmī script used to write down Khotanese

Central Asian variants of the Indian Brāhmī script were employed to write down Khotanese in all its linguistic stages. Palaeographically, one can distinguish between two

main variants, the "book script" and the "documentary script" (Maggi 2022: 157), previously termed "formal" and "cursive". The book script was mainly (but not exclusively) used for Buddhist religious texts, the documentary script for administrative documents and occasional writing, although Late Khotanese literary texts in Dunhuang often make use of several variants of the documentary script. The documentary script did not originate from the book script because it preserves many archaic features that have not survived in the book script. It is reasonable to assume that both scripts were fashioned around the same time, presumably during the 4th–5th cent. CE, and have always been employed side by side (Sander 2009: 140).[1]

Following Sander, the palaeographical periodization of the better researched book script conventionally comprises four stages (Sander 2009, see also Dragoni 2017: 396):

- Early Turkestan Brāhmī (type 2), in both its round and square variants (ETB, 5th cent.)
- Early South Turkestan Brāhmī (ESTB, 6th–7th cent.)
- South Turkestan Brāhmī (STB, 7th–9th cent.)
- Late South Turkestan Brāhmī (LSTB, 10th cent.)

The innovative features of the Central Asian Brāhmī script as used for Khotanese are the following:

1. Three new diacritics
   a. Double dot over an akṣara (<ä> [ə])
   b. St. Andrew's cross (<ei> [aə];)[2]
   c. Subscript hook (<’>)[3]
2. One additional akṣara (<rra> [ɾ])[4]
3. One additional digraph (<ysa> [za])
4. Depending on the period, doubling of <g>, <t>, <ś>, and <ṣ> (<gg> [g], <tt> [t], <śś> [ʃ], <ṣṣ> [ʂ]))[5]

Although their function may be subject to variation, features 1a, 1c, 2, and 3 are found in all extant Khotanese manuscripts. They are the most important distinctive features of the Brāhmī script used for Khotanese. It is noteworthy that feature 1b is only found in manuscripts written in Old Khotanese. On the contrary, the distribution of feature 4 is

[1] For Khotanese numerals written in Tibetan script see Maggi 1995b.

[2] See Emmerick 1998.

[3] For the different phonetic value(s) proposed for the subscript hook, see §5 below.

[4] An alveolar trill, as opposed to <r> [ɾ], according to Skjærvø (2022: 122). Del Tomba (forthc.) convincingly argues that <rr> is in fact a ligature of two <r>s. Thus, it should not be reckoned among the "new akṣaras" *stricto sensu* nor among the innovations limited to the Khotanese system because it also occurs in Tocharian A and B.

[5] See Maggi 2022: 153 table 1.

peculiar, and it provides a solid basis for classifying the different orthographic systems employed to write down Khotanese.

Maggi (2022: 150–151) distinguishes four different orthographic stages:

- Archaic — no doubling
- Old — doubling of <t> alone: <tt> [t] contrasted with <t> [ʔ]
- Classical — doubling of <g, t, ś, ṣ>: <gg> [g] vs. <g> [ɣ], <tt> [t] vs. <t> [ʔ], <śś> [ʃ] vs. <ś> [ʒ], <ṣṣ> [ʂ] vs. <ṣ> [ʐ]
- Late — doubling of *t* alone. Use of the subscript hook (<’>) to distinguish <ś> [ʃ] vs. <ś’> [ʒ], <ṣ> [ʂ] vs. <ṣ’> [ʐ]).

Because of the doubling of *t* alone, it seems reasonable to assume that his “late” and “classical” stages are independently derived from the “old” stage (Maue *apud* Maggi 2022: 151). Table 1 is a preliminary attempt to place on a timeline all the different classification criteria described so far.

Time	5th cent.	6th cent.	7th cent.	8th cent.	9th cent.	10th cent.
Place of finding	Khotan area					Dunhuang
Linguistics	Old Khotanese		Old + Middle Khotanese			LKh.
Palaeography	ETB	ESTB		STB		LSTB
Orthography	Archaic	Old	Classical + Late			Late

Table 1. Preliminary chronological outline of classification criteria.

Having introduced the main features of the Khotanese Brāhmī script and the latest research bearing on its palaeography and orthography, the following section will restrict the field of investigation to the palaeographical development of the Khotanese Brāhmī diacritics, with a special focus on the subscript hook.

4. The earliest shapes of the Khotanese diacritics

Maggi’s (2022) latest research on the archaic orthography allowed the identification of seven items as possibly the earliest extant records of the Khotanese language. They are:

- 90-YKC-040 — a wooden document recently unearthed in Karadong.[6]
- IOL Khot Wood 4 and 5 — two wooden tablets belonging to the same document from the site of Rawak, north of the Dandan Öilik oasis.[7]
- Or. 8211/1474 — a wooden document probably from the Domoko area.[8]

[6] See Maggi 2022: 151–152.

[7] See KMB: 560–561; Maggi 2022: 152–154.

[8] See KMB: 41; Maggi 2022: 154.

IOL Khot 24/9 and 24/10	two fragments of an unidentified literary text on paper from Khadaliq.[9]
SI P 83.2	a fragment of an unidentified literary text on paper.[10]

The identification is based on the convergence of palaeographical, orthographic, and linguistic data. As for the palaeography, Maggi (2022: 162–163) argues that the script of the first three pieces (90-YKC-040, IOL Khot Wood 4 and 5) can be classified as the earliest type of Early Turkestan Brāhmī, corresponding to type 1 in Sander's classification. The other four fragments are written in type 2, the stage immediately following type 1 according to the same classification.[11] They all make use of the archaic orthography.

The palaeographical analysis of these early texts in Maggi's work is limited to the akṣara <ma>, the most significant marker to distinguish the developments of the different varieties of Central Asian Brāhmī. However, the importance of these texts lies also in the fact that they contain the earliest extant shapes of the Khotanese diacritics (1a, 1b, 1c in §3 above). Together with the new akṣara <rra> and the digraph <ys>, they distinguish the Brāhmī script used for Khotanese from its Indian ancestor. The study of their graphic development thus provides important information on the origins and external connections of the system itself. Are their shapes different from the conventional forms found in later texts? If yes, how could they shed light on their origin and subsequent development? The occurrences of the three diacritics in the six items listed above are recorded in Table 2.

The shapes of <ä> and <ei> found in this corpus (see Table 2) do not diverge from their later occurrences. These two diacritics do not show significant palaeographical variation throughout the history of the Khotanese book script.[12] On the other hand, as shown in Table 3 (see below), later manuscripts show a considerable number of different shapes for the subscript hook. Its origin and later developments remain an unsolved issue. As no comprehensive palaeographical analysis of this diacritic has ever been attempted, the data from this early corpus, unmistakably showing six occurrences of a clear semicircle below the akṣara, constitute a solid starting point to historically evaluate later variations. Before proceeding to a palaeographical survey of the occurrences of the

[9] See KMB: 222; Maggi 2022: 154–155.

[10] See Emmerick & Vorob'eva-Desjatovskaja 1995: 82–83; Maggi 2022: 155.

[11] More precisely, IOL Khot 24/9, IOL Khot 24/10, and SI P 83.2 are written in Early Turkestan Brāhmī type 2.1; Or. 8211/1474 in Early Turkestan Brāhmī type 2.2 (Maggi 2022: 162–163).

[12] The situation is slightly more complicated for the less studied documentary script, in which the two dots notoriously develop into a curve above the akṣara, graphically identical to the *i*-diacritic. For a proposal regarding this historical development, see Hitch 1981: 29 with fig. 10. See Peyrot, this volume, for more details on the development and borrowing history of the two dots in the Tarim Basin.

subscript hook throughout the history of Khotanese writing, it is necessary to critically review previous research on its origin and linguistic function.

	90-YKC-040	IOL Khot Wood 4			IOL Khot Wood 5			
<ä>[13]	9 <pä>	1 <mä>	1 <hä>	5 <gä>	1 <sä>	1 <gä>	2 <rä>	2 <ṣä>
<’>		2 <vo’>			1 <vo’>	1 <ga’>		1 <hä’>
<ei>					2 <nei>			

	Or. 8211/ 1474	IOL Khot 24/9		IOL Khot 24/10	SI P 83.2		
<ä>	1 <tä>	5 <nä>[14]	4 <kkrä>	2<nä>[15]	r1 <nä>	r3 <ṣä>	r8 <rä>
<’>					v7 <ṣe’>	v8 <hā’>[16]	

Table 2. Occurrences of the Khotanese diacritics in the earliest Khotanese material. The line number is given before the transliteration of the akṣara.

[13] As the shape of the two dots do not show significant variation, only a selection is recorded in the tables. All of the instances of the subscript hook and of the St. Andrew’s cross are recorded.

[14] Despite KMB: 222 <na>, the two dots are clearly visible.

[15] The right downward stroke of <nä> cannot be read, but Skjærvø’s *hamu* ***nä****jsaḍu* ‘in one and the same manner’ (KMB: 222) appears reasonable, even if these are the only two words of the line that have been preserved.

[16] Following Emmerick & Vorob’eva-Desjatovskaja (1995: 82), I read <hā’> and not <hei’>. The left stroke of the diacritic is clearly rounded – not a straight line as in the St. Andrew’s cross – even if it does not attach directly to the top of the akṣara. Further, I assume that the right stroke is merely an ink stain.

Documentary script				Book script	
Or. 6393/2	Or. 11344/3	Hedin 20	P 2957	IOL Khot 16/10	Z[1]
2 <śa’>	a10 <pa’>	11 <pvī’>	49 <hä’>	v1 <va’>	176v4 <kā’>

Table 3. Variation in the shapes of the subscript hook in later manuscripts.

5. Previous scholarship on the subscript hook

Three major hypotheses on the origin and function of the subscript hook have been put forward during the last century. Leumann (*apud* Konow 1911: 202–203) was the first scholar to recognise in the subscript hook an orthographic device that signals the “absence of some consonant” (Konow 1911: 202). Since he thought that the subscript hook was functionally and graphically derived from the *avagraha* (ऽ) in Indian Brāhmī (Leumann 1912: 1), it is conventionally transliterated with an apostrophe (<’>).

Reviewing Leumann’s hypothesis, Emmerick (1992: 158–159) noted that the Indian *avagraha* is of much later attestation and cannot have been the model for the Khotanese subscript hook. Indeed, it seems that the first occurrences of the *avagraha* are as late as the 9th cent. CE (Bühler 1904: 91; Pandey 1952: 109). Another argument that speaks against this identification is the different function of the two diacritics. Whereas the subscript hook mainly signals the loss of Proto-Iranian **š* between vowels, the Indian *avagraha* only indicates the elision of *a* (Emmerick 1992: 159).

The second hypothesis is due to Skjærvø (1987: 85). He tentatively suggested that the Khotanese subscript hook can be formally compared to the Tumshuqese *Fremdzeichen* x_5, which indicates the Tumshuqese outcome of Proto-Iranian **š* intervocalically (see also Maue & Ogihara 2017: 423). Since the subscript hook can also indicate the loss of Proto-Iranian **š* between vowels, x_5 may indeed provide a perfect graphic and functional *comparandum*. Therefore, it is suggestive to think that both signs have a common origin.

Emmerick (1992: 159) recognised the ingenuity of this proposal but pointed out that x_5 cannot be formally compared with the oldest occurrences of the subscript hook. In fact, x_5 seems to bear more similarities only with the classical form of the subscript hook.

Yet another proposal was put forward by Emmerick (1992: 158–169). In the most extensive investigation of the subject so far, he examined the different functions of the diacritic throughout the whole linguistic history of Khotanese. He concluded that the subscript hook represents most likely a supra-segmental phonetic feature known as

"breathiness" or "breathy voice" (Emmerick 1992: 162–163).[17] As for the palaeography, he proposed that the sign may be derived from the Brāhmī akṣara *ha*. Further, he tentatively argued that the Tibetan *'a-chung* (ཨ) may be formally derived from the subscript hook. Based on the evidence provided by the earliest instances of the subscript hook (see Table 2), Emmerick's derivation of the subscript hook from the Brāhmī akṣara <ha> is no longer defendable.

Recently, Skjærvø (2022: 121, 122) cautiously suggested that the subscript hook might have indicated "rhotacization (retroflexion)" of the preceding vowel, leaving open the possibility that it might also have indicated breathy voice.[18] As for its graphic origin, he describes the sign as 'a bowl-like curve placed below the letter' (l.c.), quoting both his 1987 proposal (cf. *supra*) and Emmerick's without taking a clear stance on the matter.

As no substantial studies on the subscript hook have appeared since Emmerick (1992), a fresh examination of the question is needed to assess the validity of Emmerick's hypotheses. The first task concerns a palaeographical analysis that covers the whole history of Khotanese, starting from the earliest Old Khotanese manuscripts at our disposal.

6. Palaeography

In this section, I gather and interpret the palaeographical data. I have strived to build a representative corpus by selecting manuscripts representative of both the book and the documentary script, and all orthographic systems, from archaic to late. In doing so, I have combined Sander's classification and Maggi's new insights (see §§3, 4). For the book script, I have included materials from all four of Sander's palaeographical stages (from ETB to LSTB). I have arranged the data from the documentary script as attested in the documents from the Khotan area chronologically, according to Zhang's (2018: 60) classification, based on the one initially proposed by Yoshida (2006: 49–88).

The following is a list of the sources of the palaeographical data according to all the criteria mentioned above. Each relevant group of sources has been assigned a letter (in square brackets) that will be used throughout the analysis.

[17] Emmerick (HK: 17) chose to indicate this feature with a small *h* in his phonetic transcription of Old Khotanese. The International Phonetic Association notes this feature with a subscript diaeresis, e.g. [a̤].
[18] Rhotacised vowels have been posited for Sogdian, too (Yoshida 2009: 285).

Documentary script

Khotan area

[a] Early Turkestan Brāhmī (early 5th c.?)
90-YKC-040
IOL Khot Wood 4 and 5
Or. 8211/1474
IOL Khot 24/9 and 24/10
SI P 83.2

[b] Archive 0 (before 755, Domoko)
Or. 9268A, Or. 9268B
IOL Khot Wood 1, Urumqi 1
BH 3-98, 99, […] 131[20]

[c] Archive 1.1 (around 767, Domoko)
IOL Khot Wood 60
Hedin 33, 39, 44, 51, 53, 64, 71, 72

[d] Archive 1.2 (around 773–787, Domoko)
Hedin 4, 26, SI P 103.46

[e] Archive 2 (770–791, Dandan Uilik)
Or. 6392/1, Or. 6392/2
Or. 6393/1, Or. 6393/2
Or. 6394/1, Or. 6394/2
Or. 6395/1, Or. 6395/2
Or. 6396/1, Or. 6396/2
Or. 6397/1, Or. 6397/2
Or. 6398/1, Or. 6398/2
Or. 6399/2.6, Or. 6399/2.7
Or. 6399/2.15
Or. 6400/1.4, Or. 6400/2.1, Or. 6400/2.3
Or. 6401/1.2, Or. 6401/1.3

[f] Archive 3 (798–802, Domoko)
Or. 11252/2, Or. 11252/9, Or. 11252/12
Or. 11252/15, Or. 11252/32
Or. 11252/34, Or. 11344/1
Or. 11344/3, Or. 11344/12
Hedin 1, 3r, 13, 20
Achma 1

[g] Archive 4 (after 791, Mazar Tagh)
IOL Khot 46/2, IOL Khot 46/3
IOL Khot 46/6, IOL Khot 46/7
IOL Khot 50/4, IOL Khot 52/1

[h] Archive 5 (?, Khadalik)
Hedin 7
Or. 9615/1
SI M 1, SI M 15.1, SI M 15.2, SI M 25

Dunhuang

[i] P 3513 (Suv)
Fols. 59v, 60r, 60v, 61r

[j] P 4099 (Mañj)
r9 until r54

[k] P 2957 (Sudh)
ll. 46 to 59

Book script

[l] Early Turkestan Brāhmī, type 2, round ductus
IOL Khot 4/1 (Sgh; Canevascini 1993: 175–176)
IOL Khot 8/1 (KMB: 175–176)
IOL Khot 20/1 (KMB: 206)
IOL Khot 18/16 (Sgh; KMB: 203)
IOL Khot 31/11 (Sgh; KMB: 243)
IOL Khot 155/3 (KMB: 347)
IOL Khot 163/1 (Rk; KMB: 364)
Or. 12637/71.1 (KMB: 156)
Or. 12637/72.10 (Rk; KMB: 160)

[20] I.e. bilingual tallies, see Rong & Wen 2008, Rong & Zhang 2024.

[m] Early Turkestan Brāhmī, type 2, square ductus
- IOL Khot 16/8 (Sgh; KMB: 194–195)
- IOL Khot 16/10 (Sgh; KMB: 194–195)
- IOL Khot 18/1 (Sgh; KMB: 201)
- IOL Khot 26/13 (KMB: 228)
- IOL Khot 114/2–3 (KMB: 311–312)
- IOL Khot 169/6 (Rk; Maggi 2015)
- IOL Khot 141/1 (Sgh; KMB: 322)
- IOL Khot 167/8 (KMB: 375–376)
- bi 33 (Z, see Maggi 2004)

[n] Early South Turkestan Brāhmī
- Manuscript Or. of Suv:
  - Or. 9609A2/1 (Suv 2.16)
  - Or. 9609A2/2 (Suv 2.17)
  - Or. 9609A1/1 (Suv 2.18)
  - Or. 9609B1/1 (Suv 2.19)
  - Or. 9609A1/2 (Suv 2.20)
  - Or. 9609B1/2 (Suv 2.26)
  - Or. 9609B2/1 (Suv 2.28)
  - Or. 9609C1/1 (Suv 2.29)
  - Or. 9609C1/2 (Suv 2.30)
  - Or. 9609C2/1 (Suv 2.30–31)
- VkN (Skjærvø 1986):
  - IOL Khot 153/1, IOL Khot 153/2–3
  - IOL Khot 32/9
- Śgs (Emmerick 1970 and KMB)
  - IOL Khot 10/1, IOL Khot 10/2
  - IOL Khot 10/3, IOL Khot 11/2
  - IOL Khot 11/3, IOL Khot 12/1
  - IOL Khot 12/2, IOL Khot 12/3
  - IOL Khot 13/1, IOL Khot 13/2
  - IOL Khot 13/3, IOL Khot 14/1
  - IOL Khot 14/2, IOL Khot 14/3
- KV (Maggi 1995: 23):
  - IOL Khot 169/3
- Variants of the main manuscript of Z
  - IOL Khot 25/1

[o] South Turkestan Brāhmī
- Main manuscript of Z (Z^1)
  - SI P 6 171r–180v (Z 2.139–3.12)
  - SI P 6 206r– 209v (Z 5.42–89)
- Manuscript C of Suv (Suv 2.49–55)
  - IOL Khot 162/7, IOL Khot 30/9
  - IOL Khot 21/1, IOL Khot 163/3
  - IOL Khot 18/18, IOL Khot 203/16
  - IOL Khot 144/3, IOL Khot 25/2
- Manuscript Q of Suv (Suv 2.64–71)
  - IOL Khot 23/11, IOL Khot 26/4
  - IOL Khot 29/2, IOL Khot 27/1
  - IOL Khot 27/4
- KV (Maggi 1995)
  - IOL Khot 175/1, IOL Khot 176/4

[p] Late South Turkestan Brāhmī
- Si (Ch. ii.002):
  - IOL Khot 116/1–3 (Fols. 1–2)
  - IOL Khot 129/1 (Fol. 130)
- JS (Ch. 00274):
  - IOL Khot 65/1 (Fol. 1v)
  - IOL Khot 70/1 (Fol. 21)
- ApS (Ch. xlvi.0015):
  - IOL Khot 60/1 (Fol. 1)
  - IOL Khot 60/2 (Fol. 2)

Table 4 contains a classification of the shapes of the subscript hook in the documentary script for each of the different source groups.[21] For each different shape, a representative

[21] The press mark references for the selected images in Table 4 are the following:

(a1) SI P 83.2 v7 <ṣe'>
(b1.1) IOL Khot Wood 1 b10 <ṣā'>
(b1.2) IOL Khot Wood 1 a7 <ṣi'>
(b2) Or. 9268A c8 <pa'>
(e1) Or. 6392/2.4 <ḍi'>
(e2) Or. 6392/1.4 <rśaṃ'>
(e3) Or. 6397/1.3 <śä'>
(e4) Or. 6393/1.4 <ha'>
(g1) IOL Khot 52/1.2 <ṣä'>
(g2) IOL Khot 46/2.3 <hau'>
(h1) SI M 25.7 <śi'>
(h2) Or. 9615/1 b3 <śe'>

photographic sample has been provided. A progressive number has been assigned to each one of them, so that e.g. $b_{1.1}$ means 'shape n. 1 of source group b_1'.

[a]	[b]			[c]		[d]		[e]			
a_1	$b_{1.1}$	$b_{1.2}$	b_2	c_1	c_2	d_1	d_2	e_1	e_2	e_3	e_4

[f]				[g]		[h]		[i]	[j]	[k]
f_1	f_2	f_3	f_4	g_1	g_2	h_1	h_2	i_1	j_1	k_1

Table 4. Classification of the different shapes of the subscript hook in the documentary script according to source group (in chronological order).

In the following, I describe each individual shape in the eleven source groups.

[a] Early Turkestan Brāhmī
Only one shape is attested in the earliest examples (a_1), a very neat semicircle under the akṣara.

[b] Archive 0
Two basic shapes are attested: ($b_{1.1}$) is round, ($b_{1.2}$) is squared, and the wavy (b_2) is like ($b_{1.1}$) with the addition of an ornamental coda to the right end of the semicircle.

[c] Archive 1.1
Two shapes: (c_1) = (b_2), (c_2) is like (b_2) with a 90° rightward rotation.

[d] Archive 1.2
Two shapes: (d_1) = (c_2), (d_2) = (b_2)

[e] Archive 2
Four shapes are attested: (e_1) = (c_2), (e_2) is like ($b_{1.1}$) with a 90° rightward rotation or to (c_2) after loss of the coda, (e_3) is a 'mirrored' image of (c_2), (e_4) is like ($b_{1.1}$) with a 90° leftward rotation or to (e_3) after loss of the coda.

[f] Archive 3
Four shapes: (f_1) = (b_2), (f_2) = (e_3), (f_3) = (e_4), (f_4) = (c_2)

[g] Archive 4
Two shapes: (g_1) = (c_2), (g_2) = (e_4), (f_3)

[h] Archive 5
Two shapes: (h_1) = (c_2), (h_2) = (e_4)

[i] P 3513 (Suv)
One shape: (i_1) = (e_4), (f_3)

[j] P 4099 (Mañj)
One shape: (j_1) = (e_4), (f_3)

[k] P 2957 (Sudh)
(k_1) = (e_4), (f_3)

(c1) Hedin 33 a2 <ka'>
(c2) Hedin 64 b1 <ka'>
(d1) Hedin 4.5 <ṣi'>
(d2) SI P 103.46 b2 <śe'>
(f1) Or. 11252/2.10 <ṣa'>
(f2) Or. 11252/12 a2 <pa'>
(f3) Or. 11344/3 b2 <pe'>
(f4) Domoko A4.3 <ma'>
(i1) P 3513 60v4 <ma'>
(j1) P 4099 r30 <ṣa'>
(k1) P 2957.52 <ṣaṃ'>

Table 5 summarises the seven different types of subscript hook in the documentary script that resulted from the palaeographical analysis.

Type 1	Type 2		Type 3	Type 4	Type 5	Type 6	Type 7
	2.1	2.2					

Table 5. The seven types of the subscript hook in the documentary script.

I assume that type 1 is the original form, because it is attested in the earliest material. Besides, all other types can be derived from type 1:

1 > 2.1 [loss of internal left and right hooks]
1 > 2.2 [loss of internal left and right hooks], [squaring]
1 > 3 [addition of ornamental coda on the right]
3 > 4 [90° rightward rotation]
4 > 5 [mirroring]
2.1 > 6 [90° rightward rotation]
2.1 > 7 [90° leftward rotation]

This allows the reconstruction of the following relative chronology:

a. 1 > 2.1, 1 > 2.2, and 1 >3
b. 3 > 4, 2.1 > 6 and 2.1 > 7
c. 4 > 5

Table 6 contains a classification of the shapes of the subscript hook in the book script for each of the different source groups. As for the book script, a representative photographic sample has been provided, and a progressive number has been assigned to each of them.[22]

[22] The press mark references for the selected images in Table 6 are the following:

(l1) IOL Khot 8/1 r4 <kā'>
(l2) IOL Khot 163/1 r4 <ṣa'>
(l3) IOL Khot 4/1 r5 <ṣo'>
(m1) IOL Khot 114/2-3 v1 <ṣa'>
(m2) IOL Khot 16/8 v4 <hva'>
(m3) IOL Khot 16/10 v1<va'>
(n1) IOL Khot 10/2 r5 <ta'>
(n2) Or. 9609A2/2 v1 <ṣa'>
(n3) IOL Khot 153/1 v2 <po'>
(o1) Z[1] f. 178v4 <bā'>
(o2) Z[1] f. 178v1 <hva'>
(o3) Z[1] f. 209r6 <ta'>
(o4) IOL Khot 203/16 r2 <tce'>
(p1) IOL Khot 70/1 f. 21r4 (JS) <mu'>
(p2) IOL Khot 116/1 f. 1 bis r5 (Si) <ṣi'>

[l]			[m]			[n]		
l_1	l_2	l_3	m_1	m_2	m_3	n_1	n_2	n_3

[o]				[p]	
o_1	o_2	o_3	o_4	p_1	p_2

Table 6. Classification of the different shapes of the subscript hook in the book script according to source group (in chronological order).

A commentary to the different shapes in each source group is provided in the following.

[l] ETB, round ductus	Three shapes are attested: (l_1) = type 2.1 of the documentary script, (l_2) is like (l_1) with the addition of an ornamental coda on the right, (l_3) is like (l_2), but the hook on the left is closed.
[m] ETB, square ductus	Three shapes: (m_1) = (l_1), (m_2) = (l_2), (m_3) = (l_3)
[n] ESTB	Three shapes: (n_1) = (l_1), (n_2) = squared (l_1), cf. type 2.2 of the documentary script, (n_3) = (l_3)
[o] STB	Four shapes: (o_1) = (l_1), (o_2) = (m_2), (o_3) = (l_2), (o_4) = (l_3)
[p] ESTB	Two shapes: (p_1) = (l_1), (p_2) = (n_2)

Table 7 shows the two types and four subtypes of the subscript hook in the book script resulting from the palaeographical analysis.

Type 1		Type 2	
1.1	1.2	2.1	2.2

Table 7. The two types (and four subtypes) of the subscript hook in the documentary script.

Both types occur in the earliest manuscripts in book script, so it is difficult to determine the exact chronology. Based on the evidence of the documentary script, however, it seems reasonable to consider type 2 as an ornamental variant of type 1. Type 2 is very rare in South Turkestan Brāhmī and does not occur in Late South Turkestan Brāhmī. Type 1.2 and 2.2 might be ornamental variants respectively of type 1.1 and 2.1. It is important to note that the documentary script attests a much larger variation of shapes for the subscript hook (seven different types against only two of the book script).

The palaeographical analysis confirms that type 1 of the documentary script, corresponding to type 1.1 of the book script, can be reckoned as the ancestor form of all the other shapes of the subscript hook. I reach this conclusion based on two premises: (a) it occurs in the earliest Khotanese manuscripts, and (b) all the other shapes can be derived from type 1, directly or through intermediate (attested) steps.

In Diagram 1, I summarise the results thus obtained.

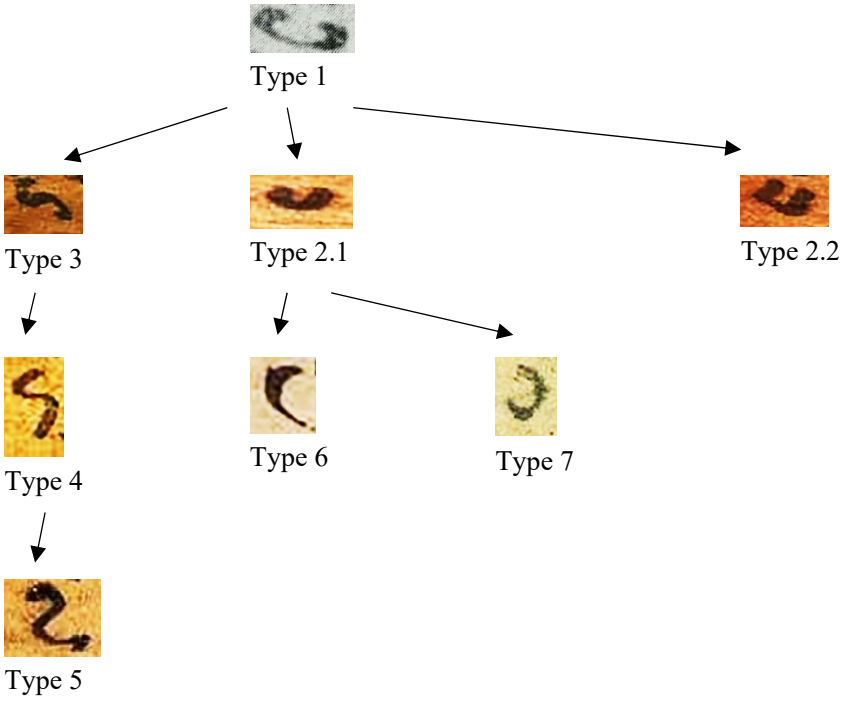


Diagram 1. Reconstruction of the three phases of development of the subscript hook in the documentary script.

In the following sections, I will extend the investigation to the possible graphic origins of the subscript hook. In doing so, I will show how the palaeographical results obtained in §6 may support the hypothesis of a borrowing from another writing system of the area, the Kharoṣṭhī script. Before laying out my argument, I will provide a brief survey of the cultural and linguistic situation of the southern rim of the Tarim Basin in the first centuries of the common era. I will pay special attention to the relationship between Kharoṣṭhī and Brāhmī in this area (§7.1) and the origin of two other Khotanese Brāhmī diacritics, the two dots and the St. Andrew's cross (§7.2).

7. Kharoṣṭḥī and Brāhmī in the south of the Tarim Basin

7.1. Introductory remarks

Before Khotanese was written down for the first time approximately at the beginning of the 5th cent. CE, there is solid evidence that dialects of Gāndhārī, a Middle Indic language, were used as the language of administration and commerce in the southern rim of the Tarim Basin. The transition from Kharoṣṭhī, the writing system of Gāndhārī, to Brāhmī was probably realised over a prolonged period and the two systems might have coexisted for a while, even if the precise dynamics of this process are still largely obscure. On the complex dynamics underlying the growing influence of the Brāhmī script on Kharoṣṭhī, see e.g. Strauch (2012: 162–164), who speaks of a 'Brahmiization' process. Undoubtedly, one should reckon with a network of several concomitant cultural and political factors, such as the transition from Gāndhārī to Sanskrit and the decline of the Kuṣāṇa empire.

As for Khotanese, Middle Indic linguistic influence is evident from the presence of a considerable number of loanwords during its entire linguistic history. No comprehensive study on this loanword corpus is available, but a first look into major Khotanese lexicographical works (e.g. KT 6, Suv 2, KS) shows that borrowing from Middle Indic affected such different semantic fields as Buddhist religious terminology, the administrative and economic language, and the medical language. Buddhist religious terminology of Middle Indic origin in the *Book of Zambasta* has recently been treated by Loukota (2023: 26–27) in a preliminary survey. Specific linguistic influence from Niya Prakrit in the administrative sphere is to be expected, but, also in this case, no comprehensive investigation has ever been attempted. To what extent the Khotanese-speaking administration showed continuity in juridical terminology and practices with the Niya Prakrit-speaking one is still an open problem. In this context, a lexical investigation of the documents from the Khotan area could deliver important historical information on the dynamics of transition from one language to another. Gāndhārī surely exerted a special influence also on the Khotanese medical terminology, as was recently shown e.g. by Luzzietti (2022: 239–235) in the case of Late Khotanese *dūvara-* 'dropsy', borrowed from the unattested Gāndhārī equivalent of Skt. *udakodara-* 'id.'

Maggi (2022: 164) noted that the formal arrangement into columns of early Khotanese documents and the numbers of commercial transactions written both in figures and words recall similar practices in the Niya documents. Similarly, the arrangement of *pāda*s into four columns in the main manuscript of the *Book of Zambasta* (Z^1) continues an earlier formal practice as seen e.g. in the Khotan Dharmapada (Maggi 2004: 187). Besides, the early Khotanese wooden tablets consisting of a covering tablet and an undertablet were a widespread form of writing support in the south of the Tarim Basin before Khotanese was written down (Maggi 2022: 164).

7.2. The borrowing paths of the two dots and the St. Andrew's cross

Given these observations, it only seems natural to expect some degree of Kharoṣṭhī influence during the process of adaptation of the Brāhmī script to write down Khotanese. As a matter of fact, some distinguishing marks of the Khotanese system have already been detected in Kharoṣthī documents. One such case is the presence of the diacritic for <ä> (the two dots above the akṣara) in some Kharoṣṭhī tablets from the southern Tarim Basin. The case of E.6.ii.1. (= CKD 661), a Kharoṣṭhī wooden tablet found in Endere but likely produced in the Khotan area and, according to the dating formula, dated to the 10th year of the King of Khotan, Vijitasiṃha, is the most famous one (Salomon 2012: 182, Dragoni *et al.* 2020: 344).[23] The same diacritic, however, has also been found in other documents, including the newly published Kharoṣṭhī wooden tablet from Khotan BH5-7 (= CKD 843, see Duan 2013). As pointed out by Peyrot (this volume, p. 288), it seems that the graphic origins of this diacritic lie in the Brāhmī visarga, as evidenced by the use of the two dots with this function in Sanskrit texts written in Kharoṣṭhī script (e.g. N.24.viii.9 = CKD 523, see also Hitch 1984: 188–190). Peyrot's contribution in this volume investigates the complex dynamics of adaptation and transfer of this diacritic, which came to represent a fundamental feature of the Tocharian writing system as well.

A hitherto unnoticed detail of BH5-7 is the presence of two instances of a diacritic identical with the later Khotanese Brāhmī St. Andrew's cross on line 7. It is placed on the final akṣara of two distinct lexical items most likely belonging to a list of personal names. It is not to be excluded that Duan's transliteration of the two items as <risavai> and <kṣatarai> is based on a parallel with the Khotanese Brāhmī diacritic, but it cannot be established with certainty. Already in Old Khotanese, <ei> frequently alternates with <ai> and one of the first transliterations of the diacritic was in fact an underlined <u>ai</u> (e.g. Leumann 1912). Possibly, an early shape of the Brāhmī diacritic for *ai*, where the two strokes converged towards the centre and came to resemble an *x*, lies at the origin of the Khotanese Brāhmī (and Kharoṣṭhī) sign, as argued by Leumann (1919: 25).[24]

Another plausible etymology of the sign is due to K.T. Schmidt (*apud* Emmerick 1998: 98), who tentatively compares the Brāhmī *jihvāmūlīya*, which "in Early Gupta (…) has the form of a double axe above an akṣara, and in the Gilgit manuscripts it is

[23] Another unique feature of CKD 661 relevant for the study of the transition from Kharoṣṭhī to Brāhmī in the Khotan area is the presence of two Brāhmī akṣaras at the bottom of the document, together with other unidentified signs. This is a tangible proof of the coexistence and interconnection of the two systems in the Khotan area in the first half of the 1st millennium CE.

[24] Manu Leumann (1934: 8) did not fully endorse his father's etymology: "Daß das liegende Kreuz durch Kreuzung der beiden Keile von ai entstanden sei, leuchtet mir nicht recht ein." See further the discussion of the etymology of the diacritic in Peyrot (this volume, p. 286, fn. 15), cautiously siding with Leumann father.

commonly found in a form resembling the Devanagari *o*-vowel" (l.c.). Even if this option is not completely ruled out by Emmerick on graphic grounds, he nonetheless doubts that the phonetic value of the *jihvāmūlīya* in Sanskrit, something like [x] (Allen 1953: 50), might have justified its most likely value [aə] in Khotanese.

In support of Schmidt's hypothesis, however, I would like to note that in Old Khotanese, the St. Andrew's cross appears most frequently in word-final position, where *-ei* indicates the nominative singular ending of *aa*-stems, mostly derived from Old Iranian *aka*-stems. In other words, the velar feature indicated by the *jihvāmūlīya* might find its counterpart in Pre-Khotanese phonetics, thus providing a linguistic justification (and a possible dating to the 1st half of the 1st millennium CE[25]) for the borrowing and adaptation of the Indian diacritic into the Khotanese system. Another element in favour of Schmidt's hypothesis is the fact that the Brāhmī *jihvāmūlīya* has been suggested as the most likely graphic source of the Tocharian Fremdzeichen for *ka̲* (see Sieg 1908: 186 with fn. 6).[26] Similarly, Maue (2014: 7) notes that the archaic shape of the Tocharian Fremdzeichen *ma̲* resembles an adapted *upadhmānīya*, indicating a voiceless bilabial fricative in the Indian system, noted by Allen (1953: 50) as -ɸ. This shows the high level of "borrowability"[27] of the Brāhmī diacritics used to indicate the allophonic realizations of the word-final aspirate <ḫ>.

Thus, the dynamics of adaptation and transfer of both diacritics, the two dots and the St. Andrew's cross, might have been parallel. Future investigations should aim to establish the exact borrowing trajectories of these graphic items (Indian Brāhmī → Kharoṣṭhī → Khotanese Brāhmī (→?) Tocharian (and Tumshuqese) Brāhmī?) and their significance for the cultural history of the area. As far as the two dots are concerned, this borrowing path is confirmed by Peyrot's extensive investigation found in this volume.

8. The Khotanese Brāhmī subscript hook was borrowed from the Kharoṣṭhī anusvāra

8.1. The borrowing path of the subscript hook

I would like to suggest that the same dynamics of transfer and adaptation from Kharoṣṭhī to Khotanese Brāhmī also applied for the subscript hook and took place during the

[25] See the preliminary dating of the Pre-Khotanese stage in Dragoni 2023: 255–256.

[26] The same shape might have also constituted the graphic base of the Tumshuqese Fremdzeichen *x*$_6$ (see Dragoni 2020: 219 Table 4). If this had the value [j], [χ] would certainly be closer than Hitch's derivation from the akṣara *kha*. However, the precise phonetic value of *x*$_6$ should be investigated more precisely.

[27] I take this term from contact linguistics, see e.g. Haspelmath & Tadmor 2009: 1.

formative phase of the Khotanese system. I propose that the source of the subscript hook lies in an adaptation of the Kharoṣṭhī anusvāra, indicating nasalisation in Gāndhārī.

Table 8 shows the attested shapes of the anusvāra according to Glass' (2000: 134) analysis. He identifies at least three shapes of the diacritic. The first is found in Aśokan inscriptions and in some coins of the Indo-Greek kings and consists of a small *ma* written at the base of the akṣara, i.e. subscript. The shape of the Kharoṣṭhī akṣara *ma* is identical with that of the subscript hook as seen in the earliest examples of documentary scripts (Table 2). In both cases, the sign is subscript and functions as a diacritic. The second and third types consist of 'a hook open to the left' (Glass 2000: 135) which can be either attached (second type) or separated from the base of the akṣara (third type, also known as "floating anusvāra").

Thus, the anusvāra is the only Kharoṣṭhī subscript diacritic that does not require to be attached to the base of the akṣara, thus maintaining a sort of graphic independence. This excludes other Kharoṣṭhī subscript diacritics like the sign for pre-consonantal *r* or the various types of foot marks (see Glass 2000: 121–124, 21–28) as sources of the Khotanese Brāhmī subscript hook and may have facilitated the process of borrowing from one script into another.

Type 1	Type 2			Type 3
Aśokan	BL 9	Niya	Schøyen 1	Schøyen 2
aṃ	aṃ	aṃ	aṃ	kaṃ

Table 8. Shapes of the Kharoṣṭhī anusvāra according to Glass (2000: 134).

Besides, the 90° leftward rotation is much reminiscent of similar developments in the history of the Khotanese subscript hook (see §6). For instance, Glass' type 3 finds its exact counterpart in type 7 of my classification of the subscript hook in the documentary script (see Table 5). These later developments cannot be genealogically connected: they are independent developments in the two systems. Nevertheless, one should note that they unfold along typologically similar pathways.[28]

The proposed borrowing path raises questions of a chronological and geographical nature which cannot be answered in full yet. One may object that the shapes of the Kharoṣṭhī anusvāra palaeographically comparable to the Khotanese Brāhmī subscript hook are mostly found in the Aśokan inscriptions and in some Indo-Greek coins, i.e. not in the Tarim Basin proper. Even if this point needs further research, it cannot be reckoned

[28] As no comprehensive study of the shape of the anusvāra in the Niya material is available, these statements are only tentative and very preliminary. Niels Schoubben (p.c.) informs me of the sporadic occurrence of "floating" anusvāras in the Niya tablets, too (e.g. in CKD 367).

as a serious difficulty against the borrowing hypothesis. The formative phase of the Khotanese writing system in the first half of the 1st millennium CE was not necessarily bound to the Khotan area but benefited from a complex network of cultural connections westwards and southwards. A very well-known instance of this web of relationships is the digraph <ys>, another distinguishing mark of the Khotanese Brāhmī system (see §3) that was already employed in the coins of the western Kṣatrapa dynasty of Ujjain in the 1st–2nd cent. CE (Lüders 1913) and in those of the Pāratarāja dynasty in western Baluchistan in the first half of the 1st millennium CE (Falk 2007; see also Maggi 2022: 156 with further refs.).

8.2. Phonetics and palaeography

If the morphology of the two diacritics can be directly compared, it remains to be investigated whether their linguistic functions have something in common. This has no decisive weight to confirm or reject the hypothesis of a borrowing from Kharoṣṭhī, because one cannot exclude that Khotanese borrowed the sign on a purely graphic basis. Nevertheless, if they had some phonetic feature in common, this could add plausibility to the borrowing hypothesis and contribute to a better understanding of the phonetic value of the Khotanese diacritic.

Although this needs a more thorough investigation that exceeds the limits of this study, I would like to point out that a possible starting point could lie in the feature of nasality indicated by the anusvāra in Indic languages. In the Khotanese Brāhmī system, the subscript hook was probably employed to indicate breathy voice (see §5).[29] The correlation between nasality and breathy voice or some other glottal features is well-known to phoneticians (Ladefoged & Maddieson 1996: 106–118) and was styled as "rhinoglottophilia" by Matisoff (1975). In Avestan, the Proto-Iranian sequence *-*aha*- was nasalised and noted in the script as <aŋha> (see de Vaan 2013, and, recently, Clayton 2020). The sound represented in Old Tibetan script by the *'a-chung*, originally [ɣ] or [ɦ], came to indicate prenasalization in the course of the history of the language (Hill 2009: 131). Moreover, several studies have shown that breathy voice is often perceived as nasality or vice versa (e.g. Garellek *et al.* 2016).

Typologically, the borrowability of the anusvāra accross different scripts to express language-specific features other than nasality is not isolated. In Japanese, the anusvāra borrowed from the *Siddham* script introduced in 9th-century Japan is usually considered among the graphic sources of the *dakuten*, a diacritic employed to distinguish between

[29] Skjærvø's (2022: 122) proposal of rhotacized vowels in Khotanese needs further research. As already noted by Skjærvø himself (l.c.), the suggestion fails to explain some of the occurrences of the subscript hook, e.g. when it signals the loss of an older intervocalic *-*h*-.

sei ‘clear’ and *daku* ‘muddy’ syllables, the distinction that led to the voicing opposition of Modern Japanese (Frellesvig 2010: 163–165).

These phonetic and typological observations strengthen the likelihood of the proposed borrowing path of the Kharoṣṭhī anusvāra into the Khotanese Brāhmī system. Given that the phonetic feature represented by the Khotanese Brāhmī subscript hook was different from that represented by the Khotanese Brāhmī anusvāra (nasalisation, see HK: 9), it could not be used to represent the same phonetic feature. However, it was sufficiently close to it phonetically to inspire the Khotanese to choose the Kharoṣṭhī anusvāra to express it in writing during the formative phase of the script system.

In other words, since the Brāhmī anusvāra had already a well-defined place in the system, the anusvāra of the Kharoṣṭḥī script, the geographically and culturally nearest script available, provided the phonetically closest equivalent of the linguistic feature indicated by the subscript hook. The phonetic similarity triggered the borrowing of the graphic shape of the diacritic by the Khotanese and its adaptation into the Khotanese Brāhmī system.

9. Preliminary results and outlook

The results of this study can be summarised as follows:

1. The earliest shape of the Khotanese subscript hook is a semicircle placed under the akṣara. This shape occurs in the earliest Old Khotanese materials. In addition to this, the preliminary results of a palaeographical analysis covering the whole period of attestation of the language confirm that all other attested shapes can be derived therefrom.
2. This rules out Emmerick’s and Skjærvø’s hypotheses on the graphic origin of the subscript hook. They had made connections respectively with the Brāhmī akṣara *ha* and the Tumshuqese foreign sign x_5.
3. The subscript hook may have been borrowed from the Kharoṣṭhī anusvāra. The earliest attested shape of the Kharoṣṭhī anusvara, a small subscript <ma>, finds a perfect formal match in the earliest occurrences of the subscript hook in Old Khotanese.
4. The borrowing path Kharoṣṭhī anusvāra → Khotanese Brāhmī subscript hook is strengthened by linguistic arguments on the correlation between nasality and breathy voice, one of the most likely phonetic values of the subscript hook.

Whereas the first two points stand on a relatively solid ground, further research is needed to confirm or disprove points 3 and 4. Specifically, a comprehensive palaeographical investigation of the shapes of the anusvāra in the Kharoṣṭhī material could shed much light on the cultural and historical context of the borrowing and on its chronology.

Further research is also necessary to determine the precise linguistic function of the anusvāra in Khotanese Brāhmī and in Kharoṣṭhī. An essential step towards a better comprehension of the phonetic values of the subscript hook cannot avoid an extensive collection and interpretation of the Old and Late Khotanese linguistic forms containing the diacritic.

Abbreviations

Sources of the photos, image processing

IOL Khot, Or., P:	https://idp.bl.uk/ (P also in https://gallica.bnf.fr/)
Domoko A4, Achma 1:	Bailey 1967
Hedin:	https://collections.smvk.se/carlotta-om/web
SI P, SI M:	Emmerick & Vorob'eva-Desjatovskaja 1993
Z[1]:	Vorob'ev-Desjatovskij & Vorob'eva-Desjatovskaja 1965

Image processing for Table 2–7 and Diagram 1: After resolution enhancement to 330 dpi; Adobe Photoshop® automatic adjustment of brightness and contrast has been systematically applied.

Text abbreviations

ApS	*Aparimitāyuḥsūtra*	Si	*Siddhasāra*
JS	*Jātakastāva*	Sudh	*Sudhanāvadāna*
KV	*Karmavibhaṅga*	Suv	*Suvarṇabhāsottamasūtra*
Mañj	*Mañjuśrīnairātmyāvatārasūtra*	Śgs	*Śūraṅgamasamādhisūtra*
Rk	*Ratnakūṭa*	VkN	*Vimalakīrtinirdeśasūtra*
Sgh	*Saṅghāṭasūtra*	Z	*Book of Zambasta*

References

ALLEN, W. Sidney, 1953. *Phonetics in Ancient India*, London.

BAILEY, Harold W., 1967. *Saka documents, 4*, London.

BÜHLER, Georg, 1904. *Indian paleography*, Bombay.

CANEVASCINI, Giotto, 1993. *The Khotanese Saṅghāṭasūtra. A critical edition*, Wiesbaden.

CLAYTON, John, 2020. 'Rhinoglottophilia in Avestan: *h > [h̃] and its orthographic and phonological consequences', in *Proceedings of the 31st Annual UCLA Indo-European Conference (November 8–9, 2019)* (ed. D. M. GOLDSTEIN *et al.*), Hamburg, 79–93.

DEL TOMBA, Alessandro, forthc. 'The palaeography of *rr* and *r* in Khotanese', *JA*.

DRAGONI, Federico, 2017. 'Chotanisch-buddhistische Manuskriptologie', in *Handbuch der Iranistik, Band II* (ed. L. PAUL), Wiesbaden, 394–402.

DRAGONI, Federico, 2020. 'The Tumshuqese year of the Goat and the *Fremdzeichen* x_6', *JA* 308/2, 215–223.

DRAGONI, Federico, 2023. Watañi lāntaṃ. *Khotanese and Tumshuqese loanwords in Tocharian*, Wiesbaden.

DRAGONI, Federico, Niels SCHOUBBEN & Michaël PEYROT, 2020. 'The Formal Kharoṣṭhī script from the Northern Tarim Basin in Northwest China may write an Iranian language', *AOH* 73/3, 335–373.

DUAN Qing 段晴, 2013. 'Neirong buming Quluwen mudu: Guotu BH5-7 内容不明佉卢文木牍:国图 BH5-7', in *Zhongguo Guojia Tushuguan cang Xiyu wenshu: Fanwen, Quluwen juan* 中国国家图书馆藏西域文书:梵文、佉卢文卷 – Xinjiang manuscripts preserved in the National Library of China (ed. DUAN Qing 段晴 & ZHANG Zhiqing 张志清), Shanghai, 203–205.

EMMERICK, Ronald E., 1970. *The Khotanese Śūraṅgamasamādhisūtra,* London.

EMMERICK, Ronald E., 1992. 'The Dunhuang MS. Ch 00120: Its importance for reconstructing the phonological system of Khotanese', in *Turfan and Tun-huang. The texts: Encounter of civilizations on the Silk route* (ed. A. CADONNA), Florence, 145–170.

EMMERICK, Ronald E., 1998. 'Khotanese *ei*', in *Proceedings of the Third European Conference of Iranian Studies held in Cambridge, 11th to 15th September 1995. Part 1: Old and Middle Iranian Studies* (ed. N. SIMS-WILLIAMS), Wiesbaden, 93–97.

EMMERICK, Ronald E. & Margarita I. VOROB'EVA-DESJATOVSKAJA, 1993. *Saka documents. 7: The St. Petersburg collections*, London.

EMMERICK, Ronald E., & Margarita I. VOROB'EVA-DESJATOVSKAJA, 1995. *Saka documents, Text Volume III: The St. Petersburg Collections*, London.

FALK, Harry. 2007. 'The names of the Pāratarājas issuing coins with Kharoṣṭhī legends', *The numismatic chronicle* 167, 171–178.

FRELLESVIG, Bjarke, 2010. *A history of the Japanese language*, Cambridge.

GARELLEK, Marc, Amanda RITCHARD & KUANG Jianjing, 2016. 'Breathy voice during nasality: A cross-linguistic study', *Journal of Phonetics* 59, 110–121.

GLASS, Andrew, 2000. *A preliminary study of Kharoṣṭhī manuscript paleography*, MA Thesis, University of Washington.

HASPELMATH, Martin & Uri TADMOR, 2009. 'The Loanword Typology project and the World Loanword Database', in *Loanwords in the worlds' languages: A comparative handbook* (ed. M. HASPELMATH & U. TADMOR), Berlin, 1–34.

HILL, Nathan, 2009. 'Tibetan <ḥ-> as a plain initial and its place in Old Tibetan phonology', *Linguistics of the Tibeto-Burman Area* 32/1, 115–140.

HITCH, Doug, 1981. *Central Asian Brāhmī palaeography: The relationships among the Tocharian, Khotanese, and Old Turkic Gupta Scripts.* MA Thesis, University of Calgary, Alberta.

HITCH, Doug, 1984. 'Kharoṣṭhī influences on the Saka Brāhmī script', in *Middle Iranian Studies: Proceedings of the international symposium organized by the Katholieke*

Universiteit Leuven from the 17th to the 20th of May 1982 (ed. W. SKALMOWSKI & A. VAN TONGERLOO), Leuven, 187–202.

HK = Ronald E. EMMERICK, 2024. *A handbook of Khotanese*, Wiesbaden.

KMB = Prods O. SKJÆRVØ, 2002. *Khotanese manuscripts from Chinese Turkestan in the British Library*, London.

KONOW, Sten, 1911. 'Note on the "Unknown languages" of Eastern Turkestan', *JRAS* 43, 201–203.

KS = Almuth DEGENER, 1989. *Khotanische Suffixe*, Stuttgart.

KT 6 = Harold W. BAILEY, 1967. *Khotanese texts VI – Prolexis to the Book of Zambasta,* Cambridge.

KUMAMOTO Hiroshi, 1996. 'The Khotanese in Dunhuang', in *Cina e Iran. Da Alessandro Magno alla Dinastia Tang* (ed. A. CADONNA & L. LANCIOTTI), Firenze, 79–101.

KUMAMOTO Hiroshi, 2012. 'Textual sources for Buddhism in Khotan', in *Buddhism across Boundaries* (ed. J. R. MCRAE & J. NATTIER), Philadelphia, 142–149.

LADEFOGED, Peter & Ian MADDIESON, 1996. *The sounds of the world's languages*, Oxford.

LEUMANN, Ernst, 1912. *Zur nordarischen Sprache und Literatur: Vorbemerkungen und vier Aufsätze mit Glossar*, Strassburg.

LEUMANN, Ernst, 1919. *Maitreya-samiti, das Zukunftsideal der Buddhisten*, Straßburg.

LEUMANN, Manu, 1934. *Sakische Handschriftproben*, Zürich.

LOUKOTA, Diego, 2023. '*ne hāḍe vajrropamä vaśärä*: Indic loanwords in the Khotanese Book of Zambasta and the chronology of the spread of Buddhism to Khotan', *BuddhistRoad Paper* 1/7, 1–34.

LÜDERS, Heinrich, 1913. 'Die Śakas und die 'nordarische' Sprache', *SPAW*, Jahrgang 1913, 406–427.

LUZZIETTI, Silvia, 2022. 'Lexical Notes on the Khotanese *Piṇḍaśāstra*', *IIJ* 65, 227–248.

MAGGI, Mauro, 1995a. *The Khotanese Karmavibhaṅgha*, Roma.

MAGGI, Mauro, 1995b. 'Late Khotanese numerals from 1 to 36 in Tibetan transcription', in *Proceedings of the Second European Conference of Iranian Studies held in Bamberg, September 30th – October 4th, 1991 by the Societas Iranologica Europæa* (ed. B. G. FRAGNER *et al.*), Roma, 425–443.

MAGGI, Mauro, 2004. 'The Manuscript T III S 16: Its importance for the history of Khotanese literature', in *Turfan revisited. The first century of research into the arts and culture of the Silk Road* (ed. D. DURKIN-MEISTERERNST *et al.*), Berlin, 184–190.

MAGGI, Mauro, 2009. 'Khotanese Literature', in *The literature of pre-Islamic Iran: companion volume I to A history of Persian literature* (ed. R. E. EMMERICK & M. MACUCH), London, 330–418.

MAGGI, Mauro, 2015. 'A folio of the Ratnakūṭa (Kāśyapaparivarta) in Khotanese', *Fagu Foxue xuebao* 法鼓佛學學報 17, 101–43.

MAGGI, Mauro, 2022. 'Some remarks on the history of the Khotanese orthography and the Brāhmī script in Khotan', *ARIRIAB* 25 (2021[2022]), 149–172.

MATISOFF, James A., 1975. ‘Rhinoglottophilia: The mysterious connection between nasality and glottality’, in *Nasálfest: Papers from a symposium on nasals and nasalization* (ed. Ch. A. FERGUSON), Stanford, 265–287.

MAUE, Dieter, 2014. ‘Simply forgotten? Note on the Tocharian Fremdzeichen la̲’, unpublished paper uploaded on www.academia.edu (last accessed: October 1, 2024).

MAUE, Dieter & OGIHARA Hirotoshi, 2017. ‘Tumschukische Miszellen III. 3. Tumshukese dental affricates’, in *Zur lichten Heimat. Studien zu Manichäismus, Iranistik und Zentralasienkunde im Gedenken an Werner Sundermann* (ed. Team „Turfanforschung“), Wiesbaden, 421–432.

PANDEY, Raj Bali, 1952. *Indian palaeography. Part I*, Banaras.

PEYROT, Michaël, this volume. ‘On the so-called “Fremdvokal” *ä* in Tocharian and Khotanese and its origins’, 275–296.

RONG Xinjiang & WEN Xin, 2008. ‘Newly discovered Chinese-Khotanese bilingual tallies’, *JIAAA* 3, 99–118.

RONG Xinjiang 榮新江 & ZHANG Zhiqing 張志清 (eds.), 2024. *Zhongguo guojia tushuguan cang Xiyu wenshu: Hanwen juan* – 中國國家圖書館藏西域文書・漢文卷, 2 volumes, Beijing.

SALOMON, Richard, 2012, ‘Gāndhārī in the Worlds of India, Iran, and Central Asia’, *BAI* 21 (2007[2012]), 179–192.

SANDER, Lore, 2009. ‘Remarks on the Formal Brāhmī script from the Southern Silk Route’, *BAI* 19 (2005[2009]), 133–144.

SIEG, Emil, 1908. ‘Neue Bruchstücke der Sanskrit-Grammatik aus Chinesisch-Turkistan’, *SPAW*, Jahrgang 1908, 182–206.

SKJÆRVØ, Prods O., 1986. ‘Khotanese fragments of the Vimalakīrtinirdeśasūtra’, in *Kalyāṇa-mitrārāgaṇam. Essays in Honour of Nils Simonsson* (ed. E. KAHRS), Oslo, 229–260.

SKJÆRVØ, Prods O., 1987. ‘On the Tumshuqese Karmavācanā text’, *JRAS* 119/1, 77–90.

SKJÆRVØ, Prods O., 2022. ‘Khotan iv. The Khotanese language’, *Encyclopædia Iranica* 17/2, 117b–133a.

STRAUCH, Ingo, 2012. ‘The character of the Indian Kharoṣṭhi script and the “Sanskrit revolution”: A writing system between identity and assimilation’, in *The idea of writing. Writing across borders* (ed. A. DE VOOGT & J. F. QUACK), Leiden, 132–168.

Suv 1, 2 = Prods O. SKJÆRVØ, 2004. *This most excellent shine of gold, king of kings of sutras: the Khotanese Suvarṇabhāsottamasūtra. Vol. I: The Khotanese text with English translation and the complete Sanskrit text. Vol. II: Manuscripts, commentary, glossary, indexes*, Cambridge, MA.

DE VAAN, Michiel, 2013. ‘On the nasalization of *h* to *ŋh* in Avestan’, in *Le sort des Gâthâs et autres études iraniennes in memoriam Jacques Duchesne-Guillemin* (ed. É. PIRART), Leuven–Paris–Walpole, 49–51.

VOROB’EV-DESJATOVKIJ, V.S. & M.I. VOROB’EVA-DESJATOVSKAJA, 1965. *Skazanie o Bxadre: novye listy sakskoj rukopisi ‘E’*, Moskva.

YOSHIDA Yutaka 吉田豊, 2006. *Kōtan shutsudo 8–9 sēki no kōtango sezoku monjo ni kansuru oboegaki* コータン出土 8–9 世紀のコータン語世俗文書に関する覚え書き – *Notes on the Khotanese documents of the 8th–9th centuries unearthed from Khotan*, Kobe.

YOSHIDA Yutaka, 2009. 'Sogdian', in *The Iranian languages* (ed. G. WINDFUHR), London, 279–336.

ZHANG Zhan, 2018. 'Secular Khotanese documents and the administrative system in Khotan', *BAI* 28 (2014[2018]), 57–98.

The "Unknown script" of Bactria: Unpublished materials and fresh interpretations*

Pavel B. LURJE

The "Unknown script" of Bactria was partially deciphered following the discovery in the Almosi gorge in Tajikistan in summer 2022 of an inscription of the king of kings Wema Taktu in Bactrian as well as in the "Unknown script". In the present paper, the author, after a brief introduction, (1) provides a review of recent studies that fostered the interpretation; (2) publishes previously unknown specimens of this script. Further (3), he clarifies the reading of the diacritics and considers the vocalism of the lan-guage of these inscriptions, where, as he proposes, the reflexes of the Old Iranian *I* and *U* vowels merged. In the next part (4) the author puts forward evidence for the reading of the sign for the initial vowel and revisits the reading of the signs *W* and *H*. In order to do that, he reconsiders the etymology of Bactrian *šaonanošao* 'king of kings' and further (5) analyses the inscriptions on silver items and identifies the word for 'silver' in the inscriptions, *əʳzat*. Finally, (6) he specifies the chronology of the inscriptions (3rd/2nd cent. BCE to 4th cent. CE) and identifies the inscribed bowl from the Issyk kurgan near Almaty as a Bactrian import and (7) locates the position of the "Unknown script" in Central Asia: having originated in Imperial Aramaic it had undergone reforms similar to those of the Indic Kharoṣṭhī script and became one of the indirect sources of the Old Turkic Runic script. As a conclusion (8) he puts forward his considerations on the nature of the language coded by the "Unknown script": it was a dialect close to Bactrian which was perhaps common among the mountaineers of Hindukush or Paropamises, including the silver mines of Panjsher or Ghur.

1. Introduction

The "Unknown script" of Bactria is a writing system of a "Runic" outlook represented in a very limited corpus of inscriptions on hard material which have been discovered by archaeologists in Afghanistan, Tajikistan, Uzbekistan and Kazakhstan during the last half century and which are dated to the first centuries before and after the beginning of the common era.[1]

* This article is the author's translation of the paper "'Неизвестное письмо' Бактрии: неопубликованные материалы и новые интерпретации", published in *Vestnik drevnej istorii* 84/2, 2024, 319–356 (Lurje 2024). It contains some additional data which is marked, if it seems significant to the

10.29091/9783752003635/004

It seems that the first published find was the so-called "*inscription peinte*" from D. Schlumberger's excavations of the Kushan dynastic temple in Surkh-Kotal, northeastern Afghanistan (1954). The inscription, consisting of three lines in "Unknown script" signs interchanging with Greek *alpha*s and *beta*s (perhaps in the function of digits) was published by André Maricq as a photo with a short description (Maricq 1958: 417). The second – and up to now the most voluminous – find is a trilingual inscription (in Bactrian, Gāndhārī and "Unknown script") from Dashti Nawur in central Afghanistan (104 CE) which was found by the Délégation archéologique française en Afghanistan (DAFA) in 1967 and published by Gérard Fussman (1974: 2–50). Later on, short inscriptions on pottery in this lettering appeared among the finds of Southern Uzbekistan and Tajikistan, predominantly in the Buddhist cave monasteries Kara-tepe and Fayaz-tepe in Old Termez (Vertogradova 1995; Fussman 2011). During the excavations of the "golden" late Saka kurgan of Issyk in the Almaty region of Kazakhstan (1969–1970), K. A. Akišev discovered a small silver bowl with a two-line inscription (Akišev 1978: 53–60). A

author. This revised translated version of the article was prepared during my research fellowship at the Paris Institute of Advanced Studies (France), September 2024 – June 2025. It has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 945408, and from the French State programme "Investissements d'avenir", managed by the Agence Nationale de la Recherche (ANR-11-LABX-0027-01 Labex RFIEA+).

[1] The author is grateful to many colleagues for the possibility to study unpublished inscriptions, their considerations on the decipherment, and for various other support. The study benefited greatly from the materials sent by Ja. V. Vassilkov (St. Petersburg), A. N. Gorin (Tashkent), Lauren Morris (Prague), A. N. Poduškin (Chimkent), Kenzo Kawasaki (Tokyo), Claude Rapin (Paris), A. P. Vydrin (St. Petersburg), A. V. Sedov, and M. G. Nikiforov (Moscow). The statistic calculations used in Sections 3 and 4 were done by K. A. Maslinski (Paris). I am also grateful to Ja. V. Vasilkov, O. V. Lundyševa (St. Petersburg), Peter Zieme (Berlin), Ching Chao-jung (Kyoto), Joe Cribb (London/Shijiazhuang), Nicholas Sims-Williams (Cambridge), Ilya Yakubovich (Marburg), Frantz Grenet, Étienne de la Vaissière (both Paris) and L. Morris for their considerations on the decipherment, as well as to A. V. Omel'čenko, A. I. Torgoev, N. N. Nikolaev, S. V. Pankova, P. A. Azbelev (all St. Petersburg), A. V. Sedov, T. V. Udyma and S. V. Bolelov (Moscow) for consultations on archaeological matters. Omel'čenko, Nikolaev and N. A. Vasil'eva provided me with access to the materials kept in the State Hermitage Museum and I. K. Malkiel' made photos of the inscription in X-ray and ultra-violet gamma. Late N. G. Ptitsyna kindly permitted to publish materials from the archive of my teacher, V. A. Livšic, and V. I. Bliznjukova helped with putting the bibliography in order. I would also like to thank the editors of the present volume for valuable suggestions.

In the course of this research, the author often exchanged emails with Svenja Bonmann, Jakob Halfmann and Natalie Korobzow. Some criticism is presented below, but this does not diminish my admiration of their pioneering work. Without the openness of the archaeologists M. G. Bobomulloev, B. S. Bobomulloev and Š. R. Xodžaev (Dushanbe), the author would not have had the possibility to start this work. Its results remain his own responsibility.

fragmented silver ingot with a one-line inscription was found in the treasury of Graeco-Bactrian city Aï Khanoum in northeastern Afghanistan during the last year of the excavations directed by Paul Bernard (1980). A large processed (?) stone with two lines of weakly scratched letters was discovered by A. N. Poduškin (2020) in the fill of a kurgan in the Ugam river valley (Kazakhstan, in the mountains to the North-East of Tashkent).

There has been no common opinion on the interpretation of the script, the encoded language(s), nor of course on the readings of the inscriptions. Fussman supposed that the script is akin to Kharoṣṭhī, and examining the diacritics he reached the conclusion that the script is of the alphasyllabary (abugida) type, in which the consonants are rendered with main signs and vowels with diacritics (or their absence). The direction of writing, as for Kharoṣṭhī, is from right to left, as Fussman has shown. According to the initial suggestion by Fussman, the script coded an ancestor language of Oṛmūrī (he later abandoned this view and considered it possible that different languages had been rendered by it; Fussman 2011: 133). Many scholars, first of all those from Kazakhstan, considered the script and language of the Issyk bowl to be related to the Runic texts of the Ancient Turks.[2] This viewpoint was criticised by V. A. Livšic (1978: 84–87), who assumed an Aramaic prototype for it. Livšic and Bernard, independently, it seems, further added the inscription on the Issyk bowl to the Bactrian epigraphy written in the “Unknown script”.[3]

As concerns the reading of the texts, attempts to read the inscription from the Issyk kurgan in Old Turkic (as well as Old Russian, Proto-Hungarian, etc.) are unceasing. The translations are equally poetic and fantastic.[4] One might consider separately the attempt of János Harmatta (1994) to read them as Khotanese, the script being close to Kharoṣṭhī. His readings have been published with a minimal commentary and apparatus and without figures, and they are not accepted by the scholarly community. My attempt to understand which signs correspond to which phonetic values in Harmatta’s decipherment resulted in a failure. Recently, Orçun Ünal (2019) proposed to read the inscriptions from Issyk and Aï Khanoum in Proto-Mongolic and considered its script as a link between Aramaic

[2] E.g. Sülejmenov 1970; Amanžolov 1971; Akišev 1978; Garkavec 2018.

[3] For the sake of completeness, we mention the works of M. Nasim Khan on the Kohī (Puṣkarasarī) script from Swat in Pakistan and neighboring valleys, which he compares to our “Unknown script” (Nasim Khan 2007, 2020). One should note that the genuineness of the inscriptions of Swat is questionable, and Nasim Khan adds “Unknown script” epigraphy from Southern Uzbekistan as an argument for its authenticity. Altogether one can find similarity in only few signs, and the system of diacritics (if that at all existed in the Kohī script) is totally different. Harry Falk (see below) supports Nasim Khan’s hypothesis (the latter in fact learned about inscriptions of Afghanistan and Central Asia from Falk).

[4] Compare the recent survey in Garkavec 2018.

and Old Turkic runes. Èduard Xuršudjan (2011) considers the language of the Issyk bowl to be "Scythian" Iranian, and the script Aramaic with ideograms (it is worth noting that his interpretation of the sign *T* corresponds to our present-day understanding).

Thus, until recently the scholarly work on the inscriptions in the "Unknown script" was limited to collecting sources, making lists of the signs, to questions of the genesis of the script, and to considerations on the function of the texts based on the archaeological context,[5] while the proposed readings remained at best unproven. A breakthrough was possible only after the discovery of the bilingual inscription in Tajikistan in 2022, and presently we can speak of a partial decipherment of the script.

2. Discovery of the inscriptions from Almosi and consequent progress in the decipherment of the "Unknown script"

After the discovery of inscriptions from Almosi (Tajikistan) in August 2022 a breakthrough in the research of the "Unknown script" occurred. During the two years after the discovery much research was undertaken, in sharp contrast to the usual unhurried pace of ancient philology. In this section I would like to mark the key points in the chronology of the discovery, in which the author took a humble part. I am conscious of the fact that the narrative can be subjective, and it is often a first-person discourse.

On August 13, 2022, while on excavations in Panjakent (Tajikistan), I received a photo of a mysterious rock inscription from the archaeologist Muhsin Bobomulloev (National Museum of Tajikistan). It was found in the north of the Hissar district, in the Fan mountains, in the gorge of Almosi.[6] Thinking that it could have been written in Kharoṣṭhī script or Old Turkic Runic script, I forwarded the photo to Ching Chao-jung and Peter Zieme. Both replies were negative: the inscription did not correspond to either of these writings. A few days later Bobomulloev sent me photos of two more inscriptions found there, as well as the context of the find. One of them was incised in the same script and another one was doubtlessly in Bactrian (in modified Greek script). The latter began with the demonstrative pronoun ειδι 'this' or ειδιηλο 'this very one' and contained the title þαονανοþαε 'king of kings' (with the archaic oblique case ending). Based on this, I compared the other inscriptions with version III of the Dashti Nawur text and made it clear that the script was one and the same.

[5] For example, Vertogradova 2002; Kyzlasov 2020. For more details on the known materials (excluding the Ugam stone, the Tuva inscriptions [see below] and Nasim Khan's hypothesis) and research history, see in Bonmann *et al*. 2023: 295–302.

[6] Tajiki *Almosī* literally means "(made of) diamond".

Having returned on August 29 from the expedition, I sent a note to the Tajik colleagues on August 31, in which I proposed a reading of the Bactrian part (which is by now significantly revised), gave a justification of the attribution of the second part to the “Unknown script”, and provided a preliminary list of correspondences of the signs. I supposed that the sequence of the signs 3–8 of the first line, which can be rendered as *A-B-C*^{1}*-C*^{2}*-A-(B)* corresponds to the rhythm of Bactrian þα-ο-να-νο-þα-ο (this assumption later proved correct). In that note I pointed out that the inscriptions of Dashti Nawur and Almosi are similar both in being incised on natural rocks and in being located next to highland fortresses, and that they somewhat delimitated Kushan territories from South and North.

Obviously, I was not the only philologist who got prompt news about the discovery. On the same day, August 31, I sent my considerations to Nicholas Sims-Williams, and he answered that he had just one hour before received an e-mail from numismatist Joe Cribb concerning these inscriptions (and immediately forwarded me his considerations on the Bactrian version). In a letter dated to September 16 (or a few days earlier) Sims-Williams and Cribb proposed the following reading of the Bactrian version, which was repeated in several later papers: ΕΙΔΙΗΛΟ Υ.(Ο) / ÞΑΟΝΑΝΟÞΑΕ ΟΟΗ/ΜΟ ΤΑΚ/ΤΟΕ ‘this very Y… of king of kings Wima Taktu’. In due course, the media of Tajikistan, and later international media as well, announced the first news about the discovery.[7]

In a collective volume dated to November 22, 2022, B. S. Bobomulloev, Sh. Khojaev and M. G. Bobomulloev, the discoverers of the inscription, published the detailed context of the find with the reading of the Bactrian inscription after Sims-Williams and Cribb (Bobomulloev *et al*. 2022). In December 2022 Gholam Djelani Davary published a German booklet on the find (Davary 2022), in which he proposed his reading of the Bactrian version[8] and justified the relation of the other part of the bilingual to the “Bactrian-Issyk” script (in the terminology of the author).

Meanwhile, I engaged myself in collecting all the specimens of the “Unknown script”, re-tracing and encoding. The inscriptions were transliterated with a code, with the digit indicating the main akṣara (a letter in Indian terminology) and the following letter indicating the diacritic (the letter *a* was conventionally used for the zero diacritic). There turned out to be 59 main akṣaras and 17 variants of diacritics. Consequently, I reduced the number by putting together similar signs from different texts, and as a result got 41

[7] https://avesta.tj/2022/11/11/tadzhikskie-uchenye-rasshifrovyvayut-drevnie-nadpisi-obnaruzhennye-v-ushhele-almosi/ (accessed June 1, 2024); https://uclcaal.org/2022/11/16/discovery-of-inscriptions-in-the-almosi-gorge-tajikistan/ (accessed June 1, 2024).

[8] Davary reads the second word as ΗΑΟΥ(Ν)Ο and compares it to Persian *aywān* ‘veranda’. The latter word, as we know, comes via MP *’’ywn* and Parthian *’pdnky* from OP *āpadāna-*. One can hardly suppose that it could have sounded /ēaun/ or the like in the 1st century CE.

akṣaras and 10 diacritics. Of these, five akṣaras and two diacritics were hapax legomena, i.e. met only once, and five akṣaras were met 2–3 times, but only in one and the same text. Thirty akṣaras (among these, 20 are met five times or more) and eight diacritics could correspond well to the phonology of an Eastern Middle Iranian language. Tables of the compatibility and frequency of the akṣaras and diacritics were drafted. From e-mail exchange I learned that Joe Cribb and Lauren Morris independently saw a similar sequence in Almosi and Dashti Nawur, which can be understood as the name Wema Taktu.

I addressed K. A. Maslinski, a specialist in computational linguistics (INALCO, Paris), who analysed the compatibility of the signs and the occurrence of repetitions (which would mark repeated words) in different inscriptions. Repetitions were noticed relatively rarely: the matches were basically between Dashti Nawur and Almosi, between Aï Khanoum and Issyk, and among a series of inscriptions of Kara-tepe and Fayaz-tepe in Termez. This fact made us suppose that the script encoded different languages. The frequency of the diacritics seemed to indicate that it was considered facultative for some scribes or engravers.

On January 25 K. Maslinskiy, M. Bobomulloev and I presented a paper at the annual Lukonin colloquium in the State Hermitage Museum. We showed the results of our joint work and outlined the major and minor punctuation marks, the sign for initial vowel (*aleph*), and the diacritic for a zero vowel, that is a consonant with no following vowel (equal to Indian *virāma*, Hebrew *shwa* and Arabic *sukūn*). In that paper we tried first of all to provide mathematic arguments for our suggestions.

The National Academy of Sciences of Tajikistan held a colloquium devoted to the inscriptions from Almosi on March 1, 2023. The Tajik archaeologists presented their information about the context of the find, Cribb (online) discussed the history of Wema Taktu based on epigraphic and numismatic data. I presented a revised paper, in which I departed at the end from mathematic rigour, and proposed an interpretation of 11 signs (of which at least seven later proved to be correct) and outlined two deictic pronouns that start with a vowel. I supposed, following the French school, that the script was connected to the movement of the Sakas into India and that the language of the script could be related to Tumshuq Saka (now I consider these conclusions totally wrong).

The colloquium ended with an online lecture of younger generation specialists from Cologne and Würzburg: Svenja Bonmann, Jakob Halfmann and Nathalie Korobzow (hereafter abbreviated BHK). Its first part was quite similar to my report: compiling a corpus of the inscriptions, tracing, digital coding, and calculating the number of signs. Their approach to the diacritics, however, was different: they distinguished only 4 of these, all attached to the left *hasta* of the letter. Then the authors showed a new tracing

and transliteration of the Dashti Nawur inscription based on original photos.[9] They also succeeded to recognise there the sequence “king of kings” and, having identified the regular word-division signs, found some correspondences to the words of the Bactrian version. They then proposed a reading for 14 akṣaras and all the four diacritics. Finally, they provided a text which had not yet been included in the debate: the inscription from cave Hoq on Socotra in three scripts (Strauch 2012: 202). In addition to the Bactrian and Brāhmī versions of the name, there is an “Unknown script” version. Reading the “Unknown script” version based on their decipherment up to that point, all four akṣaras coincided with the Brāhmī and Bactrian versions.

In this way, the researchers from Cologne were several important steps ahead of the author of this paper, although the final decipherment of the script is still a matter of future work.[10]

The author held an online meeting via Zoom with Bonmann, Halfmann and Korobzow on May 7, 2023 and we discussed the results and perspectives of our studies. We agreed to divide forces: the manuscript of the German colleagues was already submitted to the journal; after its publication I was expected to publish addenda to the corpus and my considerations on the inscriptions. The article in *Vestnik drevnej istorii* and its present translation are a fulfilment of this agreement.

Already on March 9, an article of Berlin Indologist Harry Falk about the inscriptions became available on academia.edu (Falk 2023). He also read the sequence “king of kings” and the name of Wema Taktu in inscription I and proposed to explain the sequence ΗΛΟΥΑ[Ο] of the first line of the Bactrian version as a counterpart of Indic *elūka*/*aiḍūka* ‘stone construction for human remains’. The sequence *eleatho* (for *eloatho*, in Falk’s transcription) in inscription I in “Unknown script” he considered as rendering the same term. Falk thinks that the script of Dashti Nawur and Almosi was created by Wema Taktu, that it was closest to Indian Kharoṣṭhī, and that the earlier inscriptions from Issyk and Aï Khanoum should be treated separately. He further compares the inscriptions from Central Asia with the Kohī script (very dubious in our opinion, see above, fn. 3). Even though Falk’s reading of ten letters generally corresponds to that of other scholars, his understanding of the vowels is significantly different.

[9] According to the paper of Fussmann (1974: 3, n. 1), the originals of the photographs were kept in the archive of DAFA (housed in the École française d’Extrême-Orient). My request to the archive remained without answer, and as a result I made tracings based on the photos in the 1974 article (these were quite high in quality for those days, but still insufficient for our needs). The colleagues from Cologne, after a likewise unanswered request to that archive, were able to obtain high quality scans from the Collège de France, where the personal Nachlass of Gérard Fussman is kept.

[10] The recording of the lectures of BHK and myself can be now consulted on the Internet thanks to the efforts of Nurullo Makhmudov (Khujand): www.youtube.com/watch?v=l3spw04yV8c (accessed June 1, 2024); www.youtube.com/watch?v=k-_P-lY1Rms (accessed June 1, 2024).

The voluminous article containing Bonmann, Halfmann and Korobzow's decipherment appeared in the *Transactions of the Philological Society* on June 12, 2023 (Bonmann *et al*. 2023). The publication is also available online, and on the publisher's website one can find three appendices which were not included in the printed version: the archaeological report of the discovery at Almosi, the corpus of all known inscriptions (mostly in photographs and with fresh tracings), and a list of signs and diacritics with proposed readings.[11]

The article provides a new reliable interpretation of the correspondences between Almosi I and Dashti Nawur, with new photos and tracings, in particular: the name of Wema Taktu and the title "king of kings"; parallels to a series of words and expressions from the Bactrian version of Dashti Nawur, namely βαγο, βωγο στοργο, ραβτογο, βαοδανο; and some function words. Further, a reading and translation of Almosi I and lines 3–5 of Dashti Nawur III are given, and the triscript Socotra inscription is provided as an independent confirmation of their readings. In general, the authors identify 15 signs of the script (out of some 25–30), two ligatures and four diacritics (including the zero sign for *ā*, cf. Bonmann *et al*. 2023: 310 and elsewhere).

Bonmann and her co-authors describe the script as a development of Imperial Aramaic which was only indirectly related to Kharoṣṭhī. According to them, an undetermined Middle Iranian language stands behind the script. It is related, but not identical to Bactrian, Sogdian and Khotanese Saka. They consider (following Rtveladze 2005 :146–148 and earlier publications) that it was initially the script of the Da Yuezhi before and after their intrusion into Bactria. In the last footnote, they critically review Falk's attempt at a reading.

The decipherment conducted by the Cologne team was widely commented upon in popular science media[12] (they speak ironically about these posts in another interview),[13] and I did not escape commenting upon their decipherment in the media either.[14]

---

[11] https://onlinelibrary.wiley.com/doi/10.1111/1467-968X.12269 (accessed June 1, 2024); https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2F1467-968X.12269&file=trps12269-sup-0001-Appendix+1.pdf (accessed June 1, 2024); https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2F1467-968X.12269&file=trps12269-sup-0002-Appendix+2.pdf (accessed June 1, 2024); https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2F1467-968X.12269&file=trps12269-sup-0003-Appendix+3.pdf (accessed June 1, 2024).

[12] E.g. https://avesta.tj/2023/03/03/strong-uchenye-blizki-k-polnoj-rasshifrovke-drevnih-nadpisej-ob naruzhennyh-v-almosi-strong/; www.history.com/news/ancient-kushan-script-decoded (accessed June 1, 2024).

[13] https://portal.uni-koeln.de/universitaet/aktuell/koelner-universitaetsmagazin/unimag-einzelansicht/schockierend-positiv (accessed June 1, 2024).

[14] https://tj.sputniknews.ru/20230315/tayny-tajikistan-uchenye-uschele-almosi-1055549757.html (accessed June 1, 2024); Powell 2024.

Exactly two weeks after the publication of that article, Geoffrey Caveney (New York) uploaded a preprint of a paper with an attempt to read the Issyk inscription basing himself on the new decipherment (Caveney 2023). He proposed the identification of another sign, for voiced /j/.

Later, the team of Svenja Bonmann, then also including Indologist Simon Fries, published a substantial article on the Bactrian version of Dashti Nawur. Although there the authors did barely touch the version in the “Unknown script”, it was thanks to this version that the Bactrian term þαvo ‘kingdom’ was detected, since a similar word is attested in the “Unknown Script” version (DN III).[15] Thus, in addition to being the recipient of comparisons with other languages of Central Asia, the “Unknown script” also became a donor.

In December 2023 an article by Joe Cribb on Wema Taktu appeared (Cribb 2023). It barely touches the inscription in the “Unknown script”, but it details the progression in its understanding, with a view perhaps as subjective as mine here. The archaeological context was characterised in another paper by M. G. Bobomulloev (2023).

The fact that, with different approaches, the identification of seven or eight signs of the script is similar in the papers of Falk, BHK, and the unpublished considerations of Joe Cribb, Lauren Morris and the present author, speaks of itself in favour of the right direction of decipherment. Altogether, the Cologne team was significantly ahead of the other specialists involved.

Meanwhile, by winter 2023–24 the Bactrian version of the Almosi inscription and a part of inscription II in the “Unknown script” were removed and transported to Dushanbe, as M. G. Bobomulloev informed me with regrets. The Bactrian version was located on a separate stone, which everybody could remove having loaded their donkey, while inscription II was cut off the rock, and, moreover, split into three parts. I do not know any publications about the process or result of the removal, and I cannot judge the formal and expedient reasons for such a decision. I can only note that the integrity of the unique monument was infringed. By summer 2024 both inscriptions found their way to the exhibition space of the National Museum of Tajikistan, where inscription II was put together and the cracks were carefully treated with putty (Fig. 1–2).

[15] Halfmann *et al*. 2023: 20–21, from OIr. **xšāna-*. Sogdian *š’nwx* ‘excellent’ and the proper names *š’n*, *š’n’kk* (Lurje 2010: 366) can belong here too and may be considered Bactrian loans (Persian *šān*, which was used for its explanation, is a loan from Arabic *ša’n*).

Fig. 1. Inscription II in the “Unknown script”
on exhibition in the National Museum of Tajikistan.
Author’s photo, June 2024.

Fig. 2. The rock with Inscription II in undisturbed condition.
The cut-off part is marked with a dotted line. Photo of M. G. Bobomulloev.

3. Additions to the corpus of inscriptions

The main corpus, with 24 assured inscriptions in the “Unknown script” and 9 doubtful ones, has been published in photographs and tracings.[16] In this section I provide inscriptions which remained unknown to the compilers of the corpus, with photographs, tracings and coded transliterations (with the numbering following that of BHK), as well as with phonetic transliterations, which take into account the suggestions expressed below (see Table 2). Moreover, for two previously known inscriptions I provide higher quality photographs or more exact tracings to which I had access.

[16] Bonmann *et al*. 2023 online version, appendix 2; see above, fn. 11.

Ad-1.

An archaeologically complete oenochoe with red burnish coming from L. I. Al'baum's excavations in the Fayaz-tepe Buddhist monastery in Old Termez (Fig. 3–5). It remains unpublished; the photos are present in the archive of V. A. Livšic (Bactria-9).[17] The vessel itself is kept by heirs of Livšic, where I took the photo in 2019.

A one-line ink inscription, with 6 or 7 signs surviving; a large tamga sign is incised nearby.

Fig. 3, 4, 5. Ad-01. Oenochoe of Fayaz-tepe.
Complete photo and detail, drawing of inscription (not to scale). Author's images.

The inscription has a series of otherwise unattested signs. The first sign is perhaps a variant of sign 16 (or a ligature of signs 23 and 16) with the diacritic ' ($T^{\bar{a}}$ / $\check{S}T^{\bar{a}}$) or, with Halfmann and Bonmann (by e-mail), sign 13 (Γ). The second sign is absent in the list of BHK, and seems not to have analogies in other texts. It is followed by a clear *aleph* (sign 24) with a zero diacritic ($'^{\bar{a}}$). Another thus far unattested sign is the shape of two counterposed "roofs" (Halfmann in his e-mail proposes to see here partially erased *R*). Then follows sign 2 with a zero diacritic ($H^{\bar{a}}$ of BHK, our $W^{\bar{a}}$, see below). The next sign, if it is a single one, is similar to that in the penultimate line of inscription II of Almosi: sign 26; alternatively, one can read it as sign 22 with a diacritic ' ($N^{\bar{a}}$) and a word divider. This is followed perhaps by two parallel diagonal lines, which are characteristic of sign 23 ($\check{S}$).

[17] The archive is kept in the Oriental department of the State Hermitage Museum.

Ad-2.

The shoulder part of a grayware jug with a strip of light-brown slip and horizontal lines below (Fig. 6–7), 9 × 7.5 cm, 0.8 cm wide. Black and white photos are in Livšic’s archive (Bactria-9); the original sherd was rediscovered when the archive was inventorised and presently is in the author’s hands waiting to be handed over to the Hermitage Museum. According to the note accompanying the photographs, the sherd was found during excavations of the facility yard of Fayaz-tepe in 1975 or 76.

The inscription is in ink and has six or seven characters. The beginning has survived, but the end has likely been lost.

Fig. 6, 7. Ad-02. Inscription on a sherd from the facility yard of Fayaz-tepe.
Photo and tracing by the author (not to scale).

In this inscription, one can identify virtually all the signs: 24"-20"-word-divider-21-21-20' (or 20"), which is followed by part of a sign which can be taken to stand for 15 (?). Thus, we read it as $^{\prime\prime i}$-D^{i} | $M^{\bar{a}}$-$M^{\bar{a}}$-$D^{\partial/\prime i}$-(K^{2}). We return to this inscription in the discussion of the underlying language.

Ad-3.

A fragment of the bottom of a storage vessel of light colour baking (Fig. 8) coming from the Phrourion of Kampir-tepe (the fortress on the Amu Darya of the Hellenistic and Kushan periods), from the excavations of 2016, under the brick of register XXVII on the outer side of the Southern gates. The photo was kindly sent by A. N. Gorin.

A deeply incised inscription of three signs.

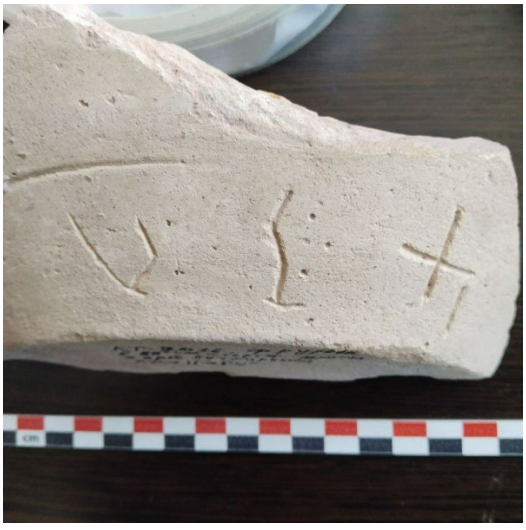


Fig. 8. Ad-03. Inscription from the Phrourion of Kampir-tepe.
Photo by A. N. Gorin.

The first letter resembles most of all sign 21 (although the upper left *hasta* is lacking, $M^{\bar{a}}$); the next sign can be considered to be sign 22 (*N*) with an uncommon diacritic; and the third looks most similar to sign 4 ($B^{\bar{a}}$) or 19. In general, it is not certain that the inscription belongs to our "Unknown script".

Ad-4.

Inscription on the shoulder of a khum (large storage vessel) from Saksanokhur (near Farkhor, Khatlon region of Tajikistan, Fig. 9). I do not know where it is kept presently. Saksanokhur is first of all famous for the Greco-Bactrian period palace, although the craftsmen quarter belongs to the Kushan period, and the sherd in question was found in its second horizon (Litvinskij & Mukhitdinov 1969). The inscription seems to have been deeply incised.

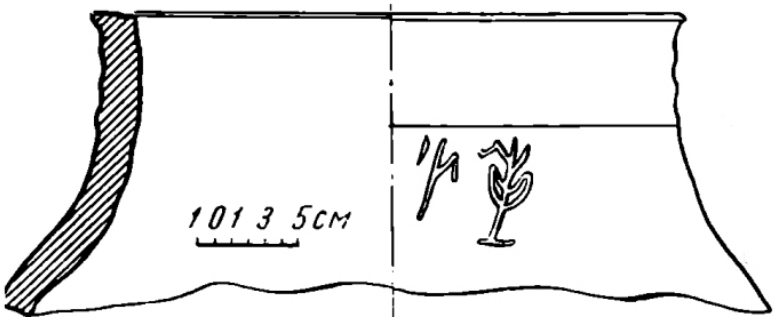
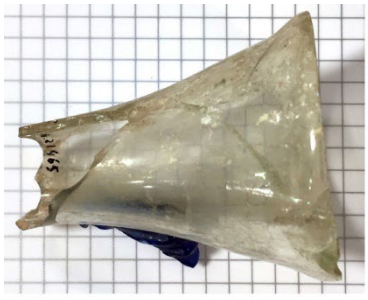


Fig. 9. Ad-04. Inscription from Saksanokhur.
After Litvinskij & Muxitdinov 1969, fig. II.

The first sign corresponds to sign 12 of BHK with an untypical diacritic; the second could be perhaps a variant of 19 ($Y^{\bar{a}}$), 16 ($T^{\bar{a}}$) or 6; then follows a word divider.

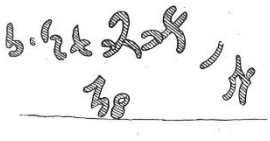
Ad-5.

Inscription on the sprout of an ichtyomorphic silver vessel from the Begram treasury, from excavations by J. Hackin in the framework of DAFA in the late 1930s (Fig. 10). The vessel was made in the Eastern Mediterranean in the late 1st – early 2nd centuries CE. The treasure (room 10) of Begram was left after 260. The vessel and a tracing of the inscription were published by Pierre Hamelin (1954: 180–181), and the inscription was treated anew by Lauren Morris (2021: 380–381), who rightly supposed that it represents the “Unknown script”. It was Morris who courteously informed me of this find and shared the data above. Presently the inscription is almost invisible and one can doubt the exactness of Hamelin’s tracing. The vessel is kept in the National Museum of Asiatic Art in Paris (Musée Guimet).



3. Opening of ichthyomorphic glass flask (4.2.1.12), RAB 358 [214] (Morris, courtesy of Pierre Cambon, MG).

4. Opening of ichthyomorphic glass flask (4.2.1.12), RAB 358 [214] (Morris, courtesy of Pierre Cambon, MG).



5. Hamelin’s illustration of the inscription on RAB 358 [214] (Hamelin 1954, Pl. XLI c).

Fig. 10. Ad-05. Sprout of an ichtyomorphic silver vessel.
After Morris 2021, pl. 50, with the author’s kind permission.
Left to right: profile of the sprout; face with inscription; Hamelin’s tracing.

In the ink inscription one can recognise sign 24 ($^{\bar{a}}$) at 8 o’clock and a part of sign 13 ($I^{\bar{u}}$?) at 4 o’clock. On the tracing one can again see sign 13 in the beginning, then sign 1 ($H^{\bar{a}}$ of BHK, my $W^{\bar{a}}$), then sign 21 with an unusual first diacritic (M^{a}), and then sign 17, also with an unusually placed stroke (K_1^{∂}). The next sign, with two crisscross strokes, does not have clear analogies. Then follow perhaps sign 16 (T?), a word divider, and two more unclear signs before *aleph* 24 with an unusual diacritic.

Ad-6.

An inscription on a rectangular stone measuring 40 × 24 × 11 cm, which is outstanding in colour and size in the fill of a large kurgan near the Ushbastobe site in the Ugam valley (an upper tributary of the Chirchik, one of the rivers of Tashkent) in Kazakhstan (Fig.

11–12). A. N. Poduškin, the leader of the excavations, thinks that the stone was treated from three sides. He dates the kurgan quite widely, between the 3rd century BCE and the 4th century CE, or, in other words, the first centuries before and after the beginning of the common era (Poduškin 2020: 198, 210). Thanks to Poduškin's kindness I obtained the photos of the stone in good resolution, and they allowed me to make a new tracing.

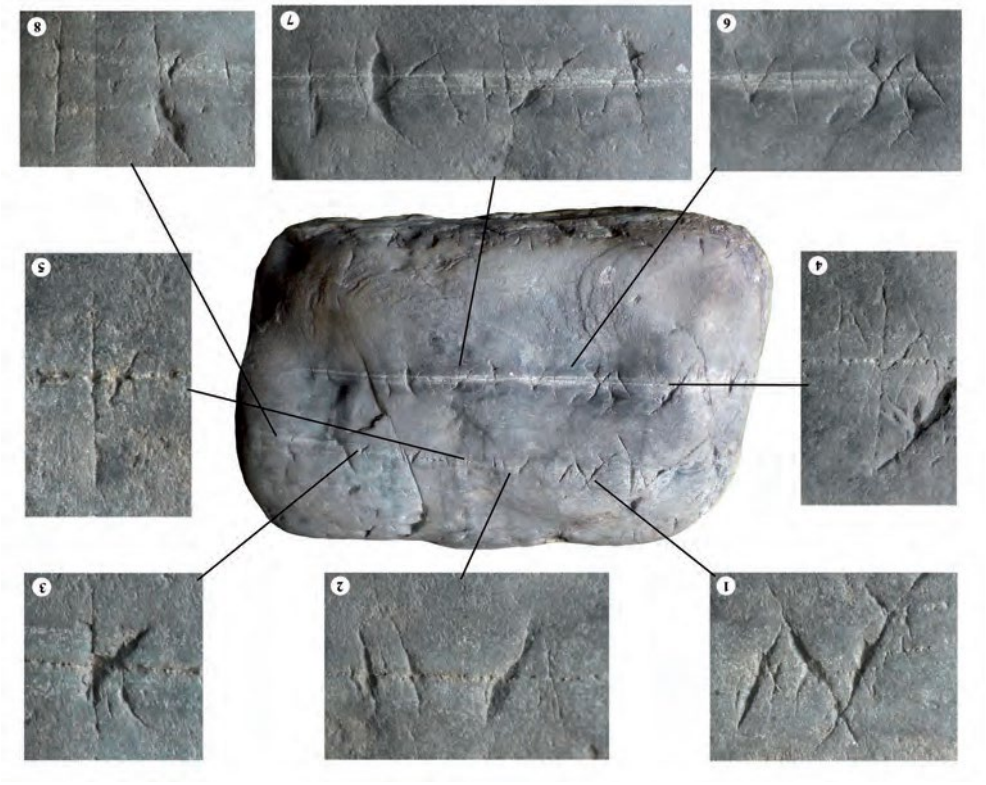


Fig. 11. Ad-05. Ugam stone. After Poduškin 2020, fig. 5.

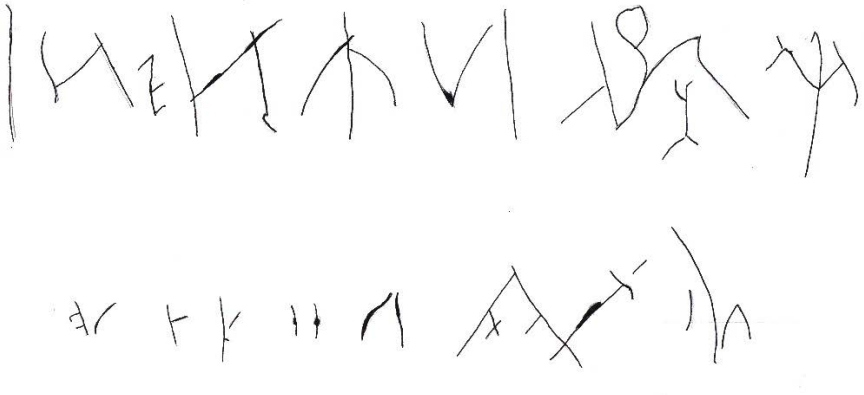


Fig. 12. Ad-06. Ugam stone. Author's tracing (not to scale).

The inscription of 15–25 signs is weakly incised, and the lines follow the white layered geological formations of the stone. It was published upside down by Poduškin.

The first line starts with sign 18 without diacritic and is followed by a ligature where one can see signs 9 ($R^{\tilde{a}}$) and 15 ($K_2^{\tilde{a}}$). The ligature is in turn followed either by a large word divider and sign 4 ($B^{\tilde{a}}$), or by a ligature of signs 3 and 1 (*HW* of BHK) or the longer sign 6. Then come a sign resembling sign 14; a ligature of signs 3 and 1, or sign 6 or 16 ($T^{\tilde{u}}$); perhaps aleph (24) with an untypical diacritic; once again sign 6; and finally, a large word divider.

The signs of the second line are even less clear; closer to the end one can see No. 23 ($\check{S}^{\tilde{a}}$). In general, one has to admit that the question whether this inscription belongs to our “Unknown script” remains open to doubt.[18] Perhaps autopsy of the stone would permit understanding more.

Ad-7.

A photograph without label kept in the archive of V. A. Livšic (Fig. 13). Incised after baking(?) on the bottom(?) of an undiagnostic storage vessel. Thanks to the kindness of A. V. Sedov and T. P. Udyma I learned that this inscription comes from B. A. Litvinski’s excavations at Kalai Kafirnigan (alias Tokkuz-tepe) in Southern Tajikistan in 1976 (Cf. CA-03 below), area VII, room 4, a corridor on the south of a large early medieval household (Litvinskij 1982: 114–121).

Fig. 13. The second inscription from Kalai Kafirnigan.
Photograph from Livšic’s archive (folder Bactria-9).

[18] A sign resembling Russian *И* (sign 6 or a ligature of signs 3 and 1) is attested here three times, and it appears four times on the Issyk inscription (see below, fn. 44). Only one diacritic is encountered here, and it joins the vowel sign, as on the Aï Khanoum specimen, see below.

The first letter is safely read as sign 21' (*M*º), and the second sign, with the same diacritic, is also attested in inscriptions from Surkh-Kotal, Khatyn-Rabat, and Dzhiga-tepe (No. 4, 5 of BHK).

Ad-8.[19]

An untreated relatively large slate stone found at the Karron site, Darvaz district of the Mountainous Badakhshan region of Tajikistan (Fig. 14), kept reportedly at the entrance of the field museum. Although I was at the site several times (and twice conducted small-scale excavations there under the direction of Academician Jusuf Jakubov), I have not noticed this stone. Although the main structures of this unusual highland city belong to the late medieval period, there was some habitation there earlier, as indicated by the finds of Kushan and Chinese coins as well as Kushan period pottery (Jakubov & Sulajmonzoda 2022). M. G. Nikiforov (Moscow) made a photograph of it and kindly sent that to me. Judging from what can be seen on the photos, most of the scratches are vertical lines, long and short, in one longer central line and shorter upper and lower lines. One can recognise signs 3 (*W* of BHK, *H* of Lurje) in the first line and 22'-4 (*N*º-*B*?) in the last line. One needs additional studies to judge whether it is an inscription in the script discussed.

Fig. 14. The stone from Karon.
Photo courtesy M. Nikiforov.

[19] Added after the publication of Lurje 2024.

Additional images

The unpublished data add new material to the interpretations of the inscriptions provided in the appendix to the article by Bonmann, Halfmann, Korobzow and Bobomulloev. The numbering follows their appendix (Bonmann *et al.* 2023, appendix 2; see above, fn. 11).

KT No. 01.

Two of three sherds of one inscribed vessel from Kara-tepe which are kept in the Hermitage Museum (A-1894, A-1895)[20] safely join one another (Fig. 15).

Fig. 15. KT-01. Sherds A-1894, A-1895 joined.
Author’s photograph.

The inscription thus reads 22"/15"-20"-24 (without lower stroke!)-13"-12, or *N*^{i}/*K*$_{1}$^{i}-*D*^{i}-$^{\bar{a}}$-*I*^{i}-*?*.[21]

CA-03.

Inscription from Tokkuz-tepe (Fig. 16). Tokkuz-tepe or Tokkuz-kala is the old name of Kalai Kafirnihan, 80 km southwest of Dushanbe. A clear photograph is present in V. A. Livšic’s archive (Bactria-9) and we publish it below. According to a kind note of A. V. Sedov, who has been directly associated with the find and its field treatment, the sherd was discovered in the mixed layer of the citadel in 1975 (and not 1945, as I thought initially) in the course of the excavations of B. A. Litvinski’s expedition (Litvinskij 1980: 121; see below on its date).

[20] My thanks go to Andrei Omel’čenko, the curator of the collection, for access to the material and for the permission to publish it.

[21] Cf. the beginning of AGII: ^{ri}-*I*^{r}.

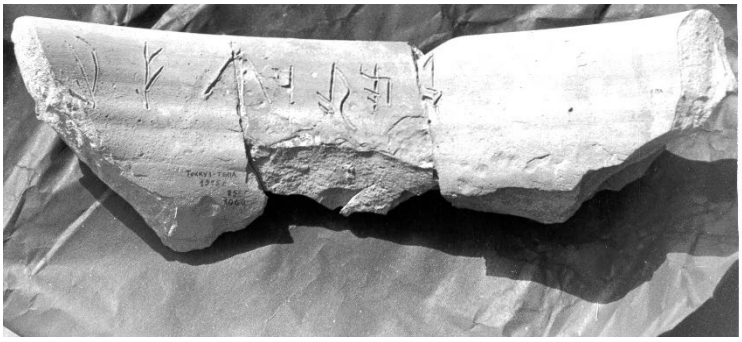


Fig. 16. CA-03. Inscription from Tokkuz-tepe.
Photo from V. A. Livšic's archive (Bactria-9), with kind permission of N. G. Ptitsyna; profile of the rim (from the archive of South-Tajikistan archaeological expedition, Institute of Oriental Studies, Moscow).

This photograph permits clarifying some details of the tracing by Natalie Korobzow. The signs are 24"-21"-16"-6(?)-14(?)-23'-4', i.e. $^{\prime i}$-M^{i}-T^{i}-$Y^{\bar{a}}$-?-$Š^{ə}$-$B^{ə}$.

This corpus of course will surely enlarge in the course of time, and not only with the new finds, but also with archive data on older discoveries. The archaeologists might have sometimes ignored them, considering their paper and ink to be of more worth for other finds from the monuments, rather than for these scribbles, which remained hopeless until very recently.

4. Diacritics

BHK distinguish three diacritics in the "Unknown script": a diagonal line coming from the left lower hasta (marked as *X*′) is read by them as ^{a}; two parallel diagonal lines in the same place (*X*″) are read as ^{u}; and a triangular mark at the center of the letter (*X*<) as ^{e}. The absence of a diacritic is interpreted as $^{\bar{a}}$ after the consonant. In the appendix, they add another probable diacritic, a vertical line which is always placed alongside " (`, i.e. the one only attested with ", see Bonmann *et al.* 2023: appendix 3, p. 2). According to them, the sign ^{a} could render short /a/ as well as a svarabhakti vowel /ə/ and no vowel, while the sign ^{e}, perhaps, served as /i/ too, and ^{u} could also render /o/ (Bonmann *et al.* 2023: 310, 316, appendix 3, p. 2).

The author of this article initially considered that there were many more diacritics, thinking that there were marks at the right-hand side of the lower *hasta*. It seems reasonable now to reject this hypothesis although in some inscriptions (Issyk, Surkh-Kotal, Chim-kurgan, Ad-3 from Kampir-tepe) one can note diacritics on the right-hand side, and in two or three cases one can note *aleph* without lower diagonal stroke (see below). The diacritic of the Begram inscription (Ad-5) looks unusual, but perhaps we are dealing with an inaccurate drawing. There seem to be upper diacritics which we can

record in the inscriptions of Surkh-Kotal, Kara-tepe and especially Aï Khanoum (cf. also Saksanokhur Ad-4 above). Taking Indic analogies into account, one can very cautiously suppose that these signs were used for nasals after vowels.

The considerations on the frequency of these diacritics seem more significant. G. Caveney expressed the opinion that in the Issyk inscription they were optional and they became mandatory in the later texts (Caveney 2023: 7–8). However, the Issyk inscription does not provide a definitive answer to this question due its small volume of 25 signs, out of which 16 are without diacritics. In the inscription from Aï Khanoum of 21 signs, the lower diacritic is attested only two times, and both times on *aleph*,[22] so coincidence is unlikely. With one possible exception, there are no diacritics on the Ugam inscription, which is partially why we doubt its appurtenance to the “Unknown script”.

Let us have a look at the frequency of diacritics in two large mutually related texts: Dashti Nawur[23] and Almosi (Table 1).

Diacritic type	DN III	AG
0 (ā)	44	17
' (a)	44	15
" (u)	43	25
< (e)	2	1
"<	1	1
Unclear, ligatures	23	2

Table 1. Frequency of diacritics in the texts of Dashti Nawur and Almosi.

The first three marks are surprisingly uniformly present in DN III and approximately equally common in AG. The third sign, <, however, is found very rarely: in the name Wema in both inscriptions and once more (near N in DN III). Combined diacritics appear once in DN III (last line, with the sign K_2) and once in Almosi (AG I, 3, with K_1). Consequently, there were three vocalisation marks that formed the basis of the script. Two (or more?) other signs were marginal.

According to the reconstruction of BHK, the system looks very skewed and excentric: the frontal vowels /i, e/ are basically absent there; one can hardly imagine such a system in any Iranian language. Let us look, however, at the Bactrian counterparts of the words with the diacritic " (^{u}). In some cases, as the decipherers rightly observe, it corresponds to *omicron* or *omega*:

[22] The second and perhaps 15th sign have upper diacritics.

[23] According to the tracing of Korobzow.

- K_2^{u}-$\check{S}^{\bar{a}}$-N^{a} – ΚΟÞΑΝΟ
- B^{u}-$\Gamma^{\bar{u}}$-(K_1^{a}) ′ S^{u}-T^{u}-R^{a} – ΒΩΓΟ ΣΤΟΡΓΟ
- $R^{\bar{a}}$-$\check{S}_T^{u}$-K_1^{a} – ΡΑÞΤΟΓΟ

Meanwhile in other cases it corresponds to *epsilon* or *iota*:

- ^{u}-D^{u} – ΕΙΔΙ(ΗΛΟ)
- $(K_1^{u?})$-D^{u} – ΚΙΔΙ

The same is true for the oblique nominal case: the diacritic ^{u} appears in the final of the counterparts of Bactrian ÞΑΟΝΑΝΟÞΑΕ and ΤΑΚΤΟΕ (on the reading of the basic signs see below).

I am unaware of any script where the vowels *U* and *I* would be rendered with one sign, and it seems more likely that these phonemes merged in the language which our script encoded. Parallels can be collected from the living Iranian vernaculars: Iron Ossetic, where the reflexes of OIr. **i*, **ī* and **u*, **ū* merged in ы (*ï*); and Munji, in which the same development happened (with some exceptions). The latter vernacular is geographically and genetically close to Bactrian.[24]

One can object that these vowel shifts are typical for Modern Iranian and could not have happened 2000 years ago, when their ancestors were at a much more archaic stage. I would like to draw attention to the fact that such advanced vocalic systems coexist today next to much more archaic ones in closely related languages and dialects. In Digor Ossetic, for example, **i* and **ī* are by default reflected as /i/, and **u* and **ū* as /u/. The same tendency is shown by the reflexes of long and short Persian *u*, *ū* as /i/ in the Kulyab and Darvaz dialects of Tajiki.[25] These dialects are part of the huge Perso-Tajik dialectal continuum, and its larger part, a number of Afghanistan dialects, and the literary norm of *Darī* based on it, preserve the Middle Iranian eight vowel system *a*, *ā*, *i*, *ī*, *u*, *ū*, *ō*, *ē* almost intact.

Among Middle Iranian languages, this eight-vowel system is reconstructed for Middle Persian, Parthian and Bactrian; as for Sogdian, current scholarship assumes merger of all three short vowels as /ə/ (Yoshida 2009: 284–285); Late Khotanese lost the differentiation in vowel length, and *ā* merged with *ū* (Skjærvø 2002: lxxi; Emmerick 2009: 382–383).

It is obvious that phonetic (and other) changes took place in different languages at different moments, and the more or less "advanced" vocalic systems coexisted in closely related languages or dialects, and one can project this state of affairs from the modern

[24] Èdel'man 1986: 71–72. A somewhat similar development is reconstructed for Wakhi.

[25] Rozenfel'd 1956: 200–201. Being in Darvaz, I heard a sound closer to high-central /ï/.

period into antiquity. The language which was encoded by the “Unknown script” was, as we suppose, phonetically some steps ahead of Bactrian.

Thus, we propose the following system of notation of the main diacritics:

- Zero diacritic – *ā*;
- Mark ' – *ə*;
- Mark " – *ï*;

We will follow this system in the discussion below.[26]

5. Reading of selected signs

In general, BHK proposed the reading of 15 signs (akṣaras in the Indian tradition) and two ligatures. The reading of another sign, sign 11, as *J* was proposed by G. Caveney.[27]

In this section we provide additional arguments in favour of the identification of *aleph* and propose to change the reading of sign 1 from *H* into *W* and sign 3 from *W* into *H* (a hypothetical proposal for the sign *Z* will be given in the next section, and an even more far-fetched one for *x* in the footnote 44, see below).

Aleph

When the author collaborated with K. A. Maslinski in compiling the code of inscriptions and analysed the frequency and compatibility of the signs,[28] we noticed that BHK sign 24 (which we coded as *1*) behaves in an unusual way. We met it 38 times in inscriptions with different diacritical signs and it is represented evenly among almost all groups of texts. Among the 30 inscriptions where the beginning of the text can be identified with some probability, this sign (with or without diacritics) was the first one in 12 cases.[29] Of

[26] Bonmann *et al*. 2023, 323, n. 6 deal with the above hypothesis without reaching a final conclusion. I was happy to realise that in their paper presented at the Fifth Conference of the Hellenistic Central Asia Research Network on March 21, 2024 in Paris Bonmann, Halfmann and Korobzow accepted the notation system used here.

[27] Caveney 2023: 6. The author bases this reading on the last three signs of the first line of the Issyk inscription, which he reads as *J(?)a-ma-ka* “bowl” (MP *yāmak*) with the Khotanese shift /y/ > /dʒ/; we leave this proposal out of consideration as well as the reading of another sign which was proposed by Bonmann *et al.* at the South Asia Languages Analysis Roundtable, 37, Venice, October 4, 2023 and as of the time of writing remains unpublished.

[28] At that time, we did not have access to the originals of the photos of Dashti Nawur, and made a new tracing basing on the published images (Fussman 1974). The photographs and tracings published by BHK of course change the picture significantly. However, we operated here mostly with the minor inscriptions.

[29] Almosi I and II, Aï Khanoum, Issyk, two inscriptions from Kara-tepe and one from Fayaz-tepe, the inscriptions of Tokkuz-tepe, Khatyn Rabat, Zartepa, Kampir-tepa, and Kosh-tepa.

the 18 remaining texts, two start with the signs 1 (*H* or *W*), 20 (*D*), 21 (*M*) and 13 (*Γ*), and the other signs found at the beginning are all met only once. So, sign 24 was the most characteristic of the beginning of a text.

At that time, we had all reasons to suppose that the script was an abugida or alphasyllabary of the Indian type, where the vowel after the consonant was written with a diacritic (or absence of it) and an initial vowel was written with a basic sign (akṣara). Consequently, the sign for initial vowel is most characteristic for word-initial position, although it can sometimes appear in other positions.[30] In the inscriptions without word divider the only secure position for the word initial is the beginning of the whole text. For a cross-check we selected Aśoka's edicts from Shahbazgarhi (*ca*. 250 BCE)[31] and the Kara-tepe inscriptions (both in Kharoṣṭhī script). Of the Shahbazgarhi edicts, five start with a vowel-initial akṣara (*aleph*) and eight with different consonant-initial akṣaras. Altogether vowel-initial akṣaras are less than 7% of the total number of akṣaras. Among the Kharoṣṭhī inscriptions from Kara-tepe (Vertogradova 1995: 49–88, 116–119), 14 start with a vowel, 18 with various consonants, and a vowel-initial akṣara not in initial position is met only once! The statistics of the Kharoṣṭhī and "Unknown script" inscriptions are quite similar, also given the fact that most of these Kharoṣṭhī inscriptions start with the deictic pronoun *aya(ṃ)*, and in Iranian languages most of the deictic pronouns also start with vowel.

Note that a similar argumentation was briefly put forward by the Cologne team,[32] and I had initially taken the shape of this vowel sign to be similar to the Arabic number *1*, taking a diagonal stroke at the bottom right as a diacritic, while BHK more reasonably explain it as a part of the base letter. A variant of *aleph* without this lower right stroke is sometimes attested (KT No.1 above, the third sign of the Aï Khanoum inscription, perhaps Fayaz-tepe CA. No. 1 (with an unusual upper diacritic), and Zar-tepe CA No. 6);[33] it is hard to say now whether it is only a graphic feature or has any phonetic relevance. BHK compare its shape with Aramaic *aleph*, and I see a similarity with the vowel-initial akṣara in the Kharoṣṭhī of Aśoka's time and some later inscriptions (Glass 2000: 33–35), as well as with *a/ä* in Turkic Runes (see below).

[30] For instance, the name of Aśoka in the Kharoṣṭhī edicts can be written as *Priyadraśi* or *Priadraśi*.

[31] According to the transliterations in https://gandhari.org/catalog?itemID=1 (accessed June 1, 2024).

[32] "In fact, character 24 is found at the beginning of words (after word dividers or empty spaces) quite frequently (10 times), where we would expect a vowel onset marker to appear most often" (Bonmann *et al.* 2023: 317, here speaking only of the DN inscription). Falk ascribes to this sign (with the diacritic ") the reading *? i*, *ai*, *ei*, following his hypothesis that the "Unknown script" transcribes the Bactrian language (Falk 2023: 11).

[33] The sherd is broken and the lower part of the line is lost. The diacritic of the fifth letter shows that the top of the diagonal stroke might have remained on it.

Reading change: sign 1 & sign 3

The parallel of the Bactrian ϸαονανοϸαο 'king of kings' to the sequence of signs 23-3"-22-22'-23-3" of the first line of Almosi inscription I was recognised independently by BHK, Harry Falk and the author of the present paper. Whereas the reading of signs 23 as *Š* and 22 as *N* is convincing, sign 3 permits different explanations. Falk thinks that it corresponds in a straightforward way to *omicron* in Bactrian. One should note that this last letter has different functions: it can be an *U*-tone vowel, the semi-vowel /w/ or a silent indicator of the end of a word or of a transparent morpheme that does not end in a different vowel.

BHK accept the commonly acknowledged etymology of Bactrian ϸαονανοϸαο and project it onto the "Unknown script". According to this etymology, Bactrian ϸαο was pronounced as /šāw/ and goes back to OIr. **xšāwan-* (nom. **xšāwā*) 'ruler' (as Sogd. *'xšywn'k* 'king' and Khot. *ṣṣau* in dating formulae), and the compound is a genitive plural construction from **xšāwanānām xšāwā*. In the subsequent periods of development of Bactrian (Kushano-Sasanian and later), autochthonous Bactrian /šāw/ was contaminated with MP *šāh* from OIr. **xšāyaθya-* (Livšic 1969: 57, n. 48; Davary 1982: 274–276; Sims-Williams 2007: 283).

This etymology in itself has a weak point: the gen.pl. of an athematic base **xšāwan-* would be **xšāwanām* (as Av. *aṣaunąm* from *aṣauuan-*; Hoffmann & Forssman 2004: 145–146). For the extended genitive **xšāwanānām* one expects thematisation of the base into **xšāwan-a-*. However, we do not find this stem either in the simple ϸαο nor in the end of the compound: a nom.sg. **xšāwā* would regularly result in Bactrian ϸαο(ο), but a derivative **xšāwan-a-(h)* would not, as the required loss of *n* cannot be accounted for. The assumption that only the gen.pl. but not the nom.sg. was thematicised[34] is possible, but does not simplify the picture.

Moreover, we do not see any secure indications of the presence of the semivowel /w/ in the base. As was said above, *omicron* also marks the end of a word or of a (transparent) morpheme unit. It appears also before the plural suffix, for example in the words ειμοανο 'these', οισποανο 'all' (Davary 1982: 183, 248). Consequently, the letters ϸαο in the first and in the second part of the compound can be read as /šā(h)/ as well. The form ϸαοο, which is provided by Davary (1982: 274), and which would most likely have to be read as /šāw/, [35] is a ghost-word: it appears only on Huvishka coins read as ÞAONANOÞAOO-OHÞKIKOÞANO. The name of the ruler on these coins has two omicrons

[34] Livšic 1969: 57, n. 48. N. Sims-Williams (e-mail of March 3, 2024) provides as an important parallel: Khotanese *śve* 'dog', gen. pl. *śvānānu*.

[35] Even with this spelling that reading would perhaps not be necessary. As N. Sims-Williams kindly noted to me (e-mail of February 28, 2024), the Huvishka coins with the Moon deity contain the legend μαοο which is doubtless /māh/ (written μαο, μαυο elsewhere), cf. Davary 1982: 226.

in the beginning (sometimes broken by *upsilon*), i.e. Ο(Υ)ΟΗΡΚΙ (Sims-Williams 2010: 111), and the division of the line into two parts by the king's diadem (noted with a hyphen above) is present on Kushan coins in other, quite unexpected places, which do not correspond to word boundaries: ΚΑ-ΝΗΡΚΙ, ΒΑ-Ζ(Ο)ΔΗΟ, Β-ΑΖΕΡΚΟ.

It is known that the laryngeal fricative /h/ was rendered in Bactrian as *upsilon*, but with many exceptions: in the early inscriptions only the initial group /hi/ was noted as Υ (ΥΝΔΟ /hind-/), and since Kanishka's time the rendering of /h/ through Υ only slowly becomes common (Sims-Williams 2008: 59–60). In the later documents, υ is often dropped, or replaced by a superscript line. Perhaps not all speakers of Bactrian articulated this phoneme (a similar loss of *h* is for example noticeable in many vernaculars of Afghan Dari Persian).

It is even more important that in versions A (line 11) and B (line 3) of the monumental Surkh-Kotal inscription, we find the spelling ΡΑΥΟ, which cannot be read in any way other than /šāh/ (Davary 1982: 53, 57); the same spelling appears in more than half of the cases in the later Bactrian documents (Sims-Williams 2007: 283). The date of the Surkh-Kotal inscription is year 31 of Kanishka's era (around 158 CE according to the commonly accepted dating) and this date definitively excludes any Middle Persian borrowing. Finally, the etymology of the name of *Kanārang* ΡΑΦΑΡΟ of the Rabatak inscription, which was proposed by Sims-Williams to mean '(possessing) the glory of the king' (Sims-Williams 2010: 156–157), is much better explainable as /šāh-far(r)/ than /šāw-far(r)/. Thus, Bactrian ϸα(υ)ο, and likewise (Middle-)Persian *šāh*, is much better traced back to another derivative of the Old Iranian root **xšā-*: **xšāyaθya-*.

In this case, how can one explain the sequence ΝΑΝΟ in the term "king of kings"? It is quite difficult to imagine a frozen gen.pl. that merged with the following base with the change *m(x)š* > *nš* (*xšāyaθiyānām-xšāyaθya-*). For one, the required simplification of *m(x)š* to *nš* (which is attested for example in Sogdian) is not found in the Bactrian theonym Ιαμϸο < **yima-xšāya-*.

It is important to note here the presence of the reduplicated plural ending -(α)νανο in the later Bactrian texts. It is optionally added to the base: πο(υ)ρανανο (also πορ(α)νο), ποσανανο 'sons', φροζινδινανο, φροζινδανανο (also φαρζινδανο) 'descendants' (Sims-Willimas 2007: 258, 274–275). Additionally, one can mention the late Bactrian plural χοηονανο from χο(αδ)ηο 'lord',[36] and μαρηγινανο from μαρηγο 'slave' (Sims-Williams 2025: 6, 117). One can add here that the optional reduplicated plural ending *-å̊nån* is attested for a series of words in the Parachi language on the Afghan-Pakistan border (Efimov 2009: 39).

[36] Sims-Williams 2007: 278. I thank the author of his reminder of this form (e-mail of February 28, 2024).

One can carefully suppose that a similar “reduplicated plural” was used for the term “king of kings”. We leave here the question of the origin of this ending open.

Consequently, we propose to read sign 3 as *H*,[37] and transliterate the term of Shahinshah itself as *Š*^{ā}-*H*^{ï}-*N*^{ā}-*N*^{ə}-*Š*^{ā}-*H*^{ï}. Let us see to what extent this proposal fits the other attestations of the sign.

The name of Wema Taktu is the most important here. According to BHK, it is spelt as *W*^{a}-*H*^{e}-*M*^{a}-*T*^{a}-*K*$_{1}$^{a}-*T*^{a}-*H*^{u} in the inscription of Almosi I, l. 2, and as *W*^{a}-*H*^{e}-*M*^{a}-*T*^{a}-*K*$_{1}$^{a}-*T*^{a}-*H*^{u} in Dashti Nawur, III, l. 4. Importantly, there is a word divider between *W*^{a} and *H*^{e} in the latter, which the scholars consider to be a continuation of a diacritic or unevenness of the rock; however, it is well visible on the photograph.

The name of Wema Taktu seems to have contained phonemes alien to Prakrit and to Bactrian, and was spelt differently (Falk 2009). The presence of the aspirate *h* in the beginning of the name is chiefly based on N. Sims-Williams’ supposition that the name of Wema’s grandson, Huwishka, which is rendered as Bactrian OϷOHÞKO /uhwišk/, Brāhmī and Kharoṣṭhī *Huviṣka*, is a diminutive to OOHMO, so that *h* should be restored in OOHMO (Sims-Williams 2010: 111). In the Kharoṣṭhī script we see the writing *vhema* (and *v́ema*), but the exact phonetic value of the first akṣara has not been established. The records of the second part of his name indicate aspiration of the middle consonant (*Takhtu*-), but not of the final consonant *t*, as suggested by the spelling -*T*^{u}-*H*^{u}.

In view of the above, we prefer to read sign 1 as *W*. Then the name of the king would be rendered as *W*^{e}-*M*^{ə}-*T*^{ə}-*K*$_{1}$^{ə}-*T*^{ə}-*W*^{ï} in Almosi and *W*^{e}-*M*^{ə}-'-*T*^{ə}-*K*$_{1}$^{ə}-*T*^{ə}-*W*^{ï} (with a clear word divider) in Dashti Nawur. In Almosi we can additionally see the rendering of the oblique case which is observed in the Bactrian version OOHMO TAKTOE: /wem taktəwï/.

An additional argument is graphic. Sign 1, which we prefer to render with *W*, is quite close to Imperial Aramaic *wāw* on the one side, to the Kharoṣṭhī akṣara *v* on the other, and ultimately to *o/u* of the Orkhon-Yeniseian Runes as well. Sign 3, which we read as *H*, has a more distant relation to *hē* of Imperial Aramaic, and is somewhat closer to *ḥēth*. In the majority of the Middle Iranian scripts based on Aramaic, *hē* is used for the aspirate /h/, while *ḥēth* is reserved for the velar fricative /x/. In the Manichaean script of the 3rd century CE and later, however, Middle Persian and Parthian /h/ can equally be rendered by the late Aramaic signs *h* and *ḥ* (in Palmyrene and Hatraic ductus), while for /x/ *kaph* with a diacritic is used (Durkin-Meisterenst 2014: 29–39). In New Persian Arabic *ḥā*

[37] Judging from the most recent paper (Halfmann *et al*. 2023: 21, n. 18), the team of Cologne and Würzburg is not far from this explanation either. According to an e-mail of March 14, 2024, Jakob Halfmann agrees with the explanation proposed above (largely on separate ground) and Svenja Bonmann favours the old etymology. In his e-mail of February 28, 2024, Nicholas Sims-Williams informed the author that “I have also been doubting Bonmann *et al*.’s reading of *W* and *H*, for different reasons”.

ḥuṭṭī (*hā-ye Hāfez*) is pronounced the same way as *hā hawwaz* (*hā-ye docheshmeh*), although the latter sign is used in inherited Iranian lexemes.

Ezafe

We are now confronted with another question: what was the function of H^a in front of the king's name in the two inscriptions? As has been noted by BHK, one of the significant differences of the language of our script from Bactrian was the absence of a determiner (*ezafe*, deictic pronoun, article) *I* (Bonmann *et al.* 2023: 324). I do not know examples of the use of a determiner before personal names in early Bactrian texts, although in the Rabatak inscription it often precedes toponyms, e.g.: ... ΟΔΟ Ι ΣΑΓΕΔΟ ΟΔΟ Ι ΚΩΖΑΜΒΟ ΟΔΟ Ι ΠΑΛΑΒΟΤΡΟ ΟΙΔΡΑ ΑΔΑ ΑΒΑ Ι ΖΙΡΙΤΙ-ΑΜΒΟ '(captured)… and Sāketa, and Kauśāmbī and Pāṭaliputra, as far as Śrī Campā' (Sims-Williams 2008: 56, l. 5–6). In the kings' epithets one can find it after titles: ÞΑΟΝΑΝΟ ÞΑΟ Ι ΒΑΓΕΠΟΟ-ΡΟ ΚΑΝΗÞΚΕ 'king of kings, son of gods Kanishka' (Sims-Williams 2008: 57, l. 14–15); ÞΑΟΝΑΝΔΕ ÞΑΟ Ι ΒΩΓΟ Ι ΣΤΟΡΓΟ ΟΟΗΜΟ ΤΑΚΤΟΟ ΚΟÞΑΝΟ 'Of the king of kings, the great salvation, Vema Takhtu, the Kushan' (Halfmann *et al.* 2023: 15, l. 2–4). *Ezafe* appears before the dynastic name Kushan in the Rabatak inscription: ΚΑΝΗÞΚΕ Ι ΚΟÞΑΝΟ.

Ezafe, originally a deictic pronoun which is used for the coordination of nouns, is in Iranian attested for the first time in Old Persian, where it has the form *haya* for the direct and *taya* for the oblique case. Later, as it is well known, it developed into *ī* in Middle Persian, and the suffix *-(y)i* in New Persian, while in Kurdish (as well as in Bactrian and Chorasmian) it preserved gender differentiation (Perry & Sadeghi 1999). The function of *ezafe* is closely related to its Bactrian counterpart, which is commonly derived not from **ha-ya-*, but rather from the deictic pronoun **ya-*, fem. **yā-*, perhaps contaminated with **ayam-* (Sims-Williams 2007: 214; Gholami 2011: 17–21). Altogether, I cannot exclude derivation of Bactrian *ezafe* from **ha-ya-* (similar to Middle Persian) as well: initial **h* in Bactrian tended to be lost in function words, as in αλο, preposition 'with' from **hada*; the preverb αμ- from **ham-*; and αμο, υαμο 'also' from **hamam*.

Based on the above, I propose to understand $H^ə$ in the inscriptions in the "Unknown script" as an *ezafe* or an article which had a different derivational track, namely from **há-(ya)-*, and partially a different function.[38] In Dashti Nawur III the sign $H^ə$ followed by a word-divider is found in front of $Š^{ā(?)}$-H^{i}-$N^{ā}$-$N^{ə}$-$Š^{ā}$-H^{i}', and one can also here suppose an article or ezafe. The last words of line 3 of Almosi II (Bonmann *et al.* 2023: 321) become

[38] I did not encounter other cases, neither in Bactrian, nor in Western Iranian (except most recent optional examples like *Doktor(-e) Malekzādeh*), where the *ezafe*/article would appear between the title and the name (the opposite is quite common: *Gaumāta haya maguš* 'Gaumata (who is) Magian', ΝΟΚΟΝΖΟΚΟ Ι ΚΑΡΑΛΡΑΓΓΟ 'Nokonzok (who is) Margrave').

more transparent when we read *H^{ə}-W^{ē}-M^{ə}-T^{ə}-K$_{1}$^{ə}-T^{ə}-W^{ï}-B^{ï}-H^{ə}-N^{ā}-M^{ï}*. According to the conclusions of BHK, *B^{ï}* corresponds to Bactrian preposition αβο (Bonmann *et al.* 2023: 316), which is often followed by *ezafe* in Bactrian. It is hard not to see in the following word the equivalent of OIr. ***nāman*-**, Bactrian ναμο ‘name’ perhaps as the first member of a compound like ΝΑΜΟΒΑΡΓΟ ‘named, famous’ in the large Surkh Kotal inscription.

As concerns the inscription from Socotra, BHK propose to read a ligature of signs 1 and 3', in their interpretation *H-W^{u}* (Bonmann *et al.* 2023: 322–323), to render the initial /hu/ of the name which is spelt ΟΜΟΙΑΓΟ in Bactrian and *humiyaka(sa)* in Brāhmī. Along the lines of the previous discussion, I see here more likely *(W?)H^{ə}*, which would also suit the rendering of an Old Iranian labialised fricative like *x^{w}*, *h^{w}* (cf. Parthian *wxad* ‘self’, etc.).

6. The inscriptions from Issyk and Aï Khanoum

These two inscriptions are close to one another by a number of parameters. First, both of them are relatively well preserved and clearly written. Second, they are the earliest examples of the script, although the date of both of them is not definitively established (see below). Third, both of the inscriptions are on silver objects: a bowl from Issyk and an ingot from Aï Khanoum, and they might bear information of a similar nature. Virtually all inscriptions on precious objects from pre-Islamic Central Asia known to me indicate the weight of the object in these or those units. The majority also indicate property of the object, and facts of donation or devotion.

It is noteworthy that in all of the numerous approaches to read the Issyk inscription there has been no attempt to find out the recording of its weight. Moreover, its real weight, which could perhaps be a “quasi-bilingual” to the inscription, has not been published. Amir Sakuakas from the National Museum of Kazakhstan kindly shared with me the dimensions of the bowl: weight 78.7 g, height 2.2 cm, diameter 7.7 cm. The initial weight of the bowl might have been somewhat larger: there is a series of losses in the central part. The weight of the Aï Khanoum ingot is 2278 gram, and the dimensions are 165×155×10–22 cm, although hardly half of the original shape and weight of the ingot has survived (its supposed diameter is *ca*. 24 cm; Rapin 1992: 318); the end of the inscription is also lost.

Altogether one can notice three series of similar signs of two or three letters each, which can be taken to be identical words in similar formulae of the inscription. That is why we deal with them together (Fig. 17).[39] As was said above, in these inscriptions,

[39] Ünal (2019: 179–181) worked in a similar way, taking into account the chronological similarity of the two inscriptions only.

especially the Aï Khanoum one, the diacritics seem to be used less, so that we allow ourselves more liberty in the vocalisation.

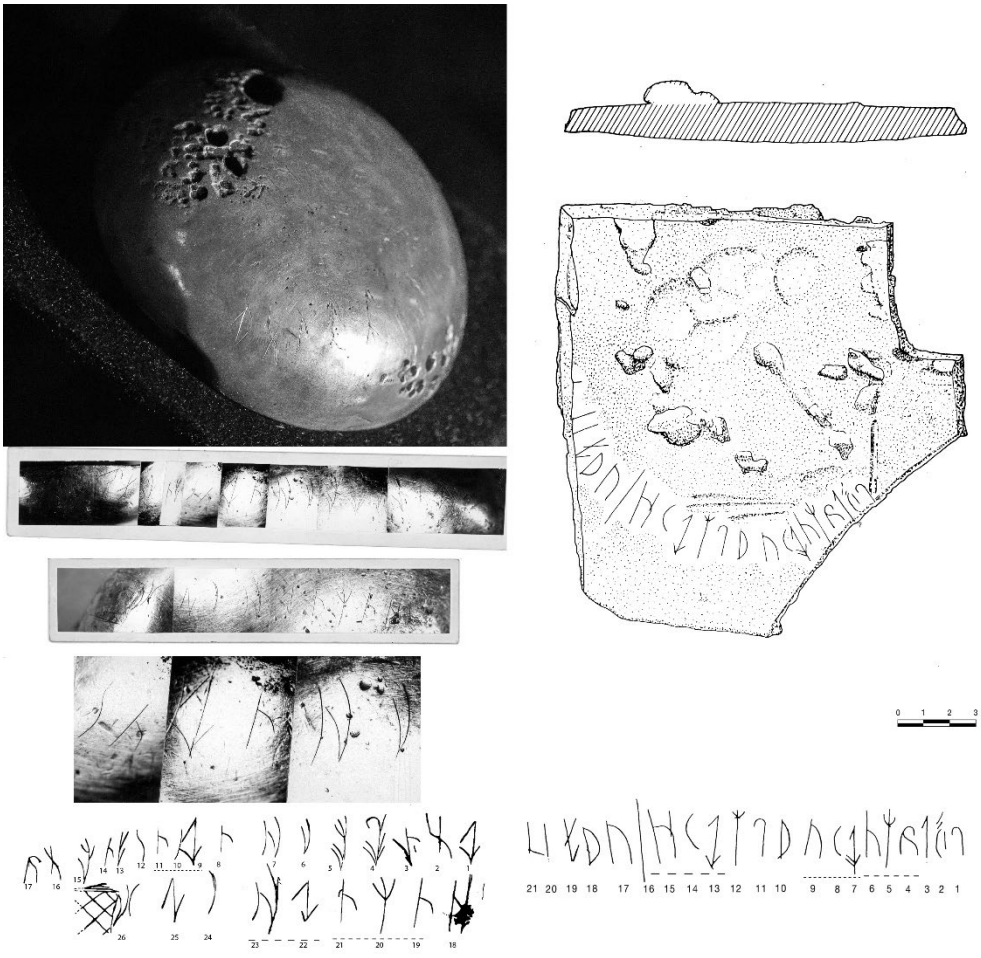


Fig. 17. Inscriptions from the Issyk kurgan and Aï Khanoum.
Top left is a photo of the Issyk bowl.
(https://art16.ru/gallery2/v/kazan_kremlin/manej/shestvie-zolotogo-cheloveka-po-muzeyam-mira/20190315_0151.jpg.html?g2_imageViewsIndex=3; accessed June 1, 2024).
Mid left is a photo montage of the inscription (Livšic's archive, Bactria-9) and an enlarged fragment with signs 6–12; bottom left is the tracing after Bonmann *et al.* (2023: appendix 2, p. 1) with the author's additions.
On the right: cross-section, view, and tracing of the inscription on the Aï Khanoum ingot (with the kind permission of Claude Rapin).
The similar elements of the two inscriptions are underlined.

A similar word can be recognised in signs 9-10-11 of the Issyk inscription[40] and 13-14-15 on the one from Aï Khanoum. The first letter is $^{\vartheta}$, the third is $T^{\bar{a}}$, and the second is a line slightly curved to the left or to the right. Taking into account that the word ‘silver’ is commonly found on such inscriptions one can suppose that the second sign corresponds to /z/ or /ž/, and the word itself is read as /*ə^{r}zata/. It is similar to the Old Iranian term for ‘silver’, **arzata-* or **r̥zata-*: Av. *ərəzata-*, OP *ardata-*, Chorasmian *ž^{y}d*, Ossetic *ærzæt*, Khot. *āljsata-*, coming from the same Indo-European root from which Old Indian *rajatá* and Latin *argentum*, etc. originate.[41] In this case, our language lost the consonantal element of the vocalic *r̥*, as it happened in languages like Sogdian (Gershevitch 1954: 19–22), Chorasmian and sometimes Bactrian.[42] At that period, the post-vocalic /-t/ had perhaps not been voiced to /-d/, which we see in $^{\ddot{\imath}}$-$D^{\ddot{\imath}}$ – ειδο, $K_1^{\ddot{\imath}}$-$D^{\ddot{\imath}}$ – κιδι in the time of Wema Taktu and in later texts.

It is difficult to identify this sign for /z/ or /ž/ on other materials. The letter with the minimalistic shape with the possible phonetic value *z* is not found there; moreover, if it occurred there, it could easily be mistaken for a word divider. Maybe it was “hidden” in the crack in DN III, l. 5, the word BHK read (in their transliteration) as B^{a}-*Γ(?)-[?]-[?]*-N^{u}-K_1^{a}, which would correspond to Bactrian ΒΑΓΟ Ι ΗΖΝΟΓΟ ‘worthy of deity’ (Bonmann *et al.* 2023: 315, 320). The similarity of the discussed sign to Aramaic *zain* is obvious.

My other attempts to read words in the Issyk and Aï Khanoum inscriptions are even more hypothetical and require much more work before publication (cf. however fn. 44 on *baxt(a)* ‘given’?).

7. The chronology and topography of the finds

Svenja Bonmann and co-authors date the monuments of the “Unknown script” to the period from the last centuries BCE to the 7th century CE (Bonmann *et al.* 2023: 301). Geographically, the inscriptions are concentrated in historical Bactria-Tokharestan, both to the south and to the north of the Oxus, but also reaching the southern slopes of the Hindukush (Dashti Nawur, Begram). The inscription from the Issyk kurgan 50 km to the east of Almaty stands aside. Half-way between Bactria and Issyk is located the kurgan of Ugam (Ad-6), but the relation of its inscription to our “Unknown script” remains unproven.

[40] The second letter is lacking on the published tracings but is clearly visible on the photo montage from Livšic’s archive.

[41] The last discussion: Blažek & Schwarz 2016: 123–124.

[42] Cf. αζανο ‘worthy, deserving’ from **arǰyāna-* (Sims-Williams 2007: 188).

The inscription on the small silver bowl from the "golden man" kurgan of Issyk seems to be the earliest attested example. The initial dating of the Issyk kurgan to the 6th–5th centuries BCE (Akišev 1978: 74) is obviously much too early. Now archaeologists propose the 4th–3rd or even the 2nd century BCE (Ünal 2019: 176–177). An inscription in the "Unknown script" in the period before the 3rd century BCE would be extremely problematic. This bowl is significantly different from the other materials of this grave, the outstanding golden artefacts made in the Scythian "Animal style". We see here both the less prestigious silver, and the absence of any decoration. The "royal kurgans" of both the Sakas and later nomads often include importations from the Near East. They have also recently been attested among the materials of the Issyk burial (Torgoev 2024).

Taking into account the geography of the finds, the peculiarity of the cup and the presence of other imports among the grave goods of Issyk, one can suppose that the silver bowl was also an import from the southwest, a result of trade, seasonal migration, raids or a combination of these.

The chronologically second inscription is engraved on the silver ingot from the treasury of the Graeco-Bactrian city of Aï Khanoum. The capture and destruction of Aï Khanoum is dated to around 145 BCE. Many scholars explained the inscribed ingot in the treasury as a result of the Saka invasion (Rapin 2007: 50; Grenet 2015: 203–204; Martinez-Sève 2020: 106). As we proposed above, the Issyk bowl was an import, and so the items with the "Unknown script" in Bactria should not necessarily be linked to the Sakas. There seem to be no other clear indications of the Saka seizure of the city. Laurianne Martinez-Sève argues that the city was perhaps destroyed by the local Bactrian population during the uprisal against Hellenic rule. We think it is more accurate to date the ingot to the first half or the middle of the 2nd century BCE, predating the destruction of Aï Khanoum.

The uppermost date of the monuments of the "Unknown script" is also in need of certain corrections. The limit of the 7th century given by Bonmann *et al.* (2023: 301) is based on the inscription of Tokkuz-tepe, that is Kalai Kafirnigan (cf. CA-03 above). The latter site is known for the decorated Buddhist sanctuary of the Early Middle Ages, although it also contains lower layers of the Kushan period, for example, a bronze coin of Soter Megas (Wema Taktu) and high goblets (for instance, Litvinskij 1983; Bobomulloev & Hasuike 2013: 45, 49). The inscription of Ad-7 comes seemingly from an Early Medieval household. The trench on the excavation plot has shown that the house covers buildings of the Kushan period and, as A. V. Sedov supposes, this fragment could have been redeposited out of this early stratum (Litvinskij 1982: 114–121). Inscription CA-03 was found in the mixed layer of the citadel, the rim is diagnostic and, according to the competent opinion of A. V. Sedov,[43] more likely belongs to the Kushan period.

[43] E-mail dated March 20, 2024.

Thus, with this more nuanced dating of CA-03, the upper limit of the 7th century is no longer valid.

The inscription CA No. 6 from Zar-tepe (Surkhandarya region of Uzbekistan) comes from area I, stratum 18 (Vertogradova 1995: 138), i.e. the monumental building in the center of the site. The excavator of the building, A. Ya. Schetenko, dated it to the Kushan period, but V. A. Zav’jalov corrects this date to Kushano-Sasanian, that is the second half of the 3rd century CE to the 4th century CE (Zav’jalov 2008: 19). A. V. Omel’čenko and S. B. Bolelov informed us, orally, that the oenochoe Ad-01 should be dated to the Kushano-Sasanian period too; it is known that the Fayaz-tepe monastery, where it was found, was inhabited in that period. The shoulder of the vessel with inscription Ad-02 is of the Kushan epoch.

Thus, the reliable life-span of the “Unknown script” should be placed between the 3rd century BCE and the 4th century CE and its area of use should be limited to Bactria-Tokharestan and the southern Hindukush.

8. The position of the “Unknown script” among the ancient scripts of Central Asia

Previously, scholars supposed that the “Unknown script” of Bactria could be related to the Aramaic script of the Achaemenid empire, the Indian Kharoṣṭhī script, or to the Old Turkic runes (cf. above). These hypotheses do not contradict one another, as Kharoṣṭhī ultimately derives from Aramaic and perhaps the Old Turkic Runic script as well. As we shall see below, all these hypotheses are correct to various degrees.

I concur with the opinion of Bonmann and co-authors who, following V. A. Livšic, consider Imperial Aramaic the source of this script. To the eight similar letters provided in the table of BHK (Bonmann *et al.* 2023: 318–319), namely *’*, *D*, *M*, *N*, *S*, *R*, *Š*, and *T*, one can likely add three more: *W*, early *Z* (?), and *H*. If our observation is correct that the last letter derives from Aramaic *ḥēth*,[44] and not *hā*, this suggests that the “Unknown script” derives from Aramaic independently from other Middle Iranian scripts, where *hā* was reserved for /h/. The letters K_1 and K_2 have, according to BHK, only some (loose) similarity to Aramaic *qoph*, and are even less similar to *kaph*; this fact also speaks in favour of independent derivation of the “Unknown script”, as in other old Middle Iranian traditions (unlike later Manichaean or Christian Sogdian) derivatives of *kaph* were used.

[44] The И-shaped letter 6, which is still unexplained, looks even more similar to *ḥēth*. Maybe the sequence in the first line of the Issyk inscription, $B^{ā}$-$6^{ā}$-$T^{ū}$, should be reconstructed as /baxt(a)/ “donated” (PPP), which is well expected in a votive inscription. However, this reading is quite shaky (one also wonders about the absence of diacritics).

Another eye-catching difference of our script to other Middle Iranian scripts based on Aramaic is the complete absence of ideograms. One could suppose that some Arameograms have not been deciphered yet, although, according to the words of V. A. Livšic, as well as according to my own limited experience, ideograms are usually the first to be noticed by scholars in any text.

The connection to Indian writing systems is first of all the way of rendering vowels through diacritics (alphasyllabary or abugida). Diacritics are obviously mandatory in the large inscriptions of the Kushan period, although in the early examples (especially Aï Khanoum) they seem to be written irregularly, and were mostly used with the vowel-initial sign. The position and the shape of the diacritics are quite distant both from early Brāhmī and from Kharoṣṭhī (Salomon 1998: 18, 43). The sound values of the diacritics are also significantly different. Three common diacritics of the "Unknown script" (including the zero diacritic) cannot be compared to the eleven of Brāhmī and the six of Kharoṣṭhī. It is possible that in some texts nasalisation of a syllable was also marked by a diacritic, but by now we lack safe examples. However, in the early Indian texts the *virāma* sign (for zero vowel after consonant) is extremely rare (Falk 2021), while in our script the sign $^{\partial}$ is regularly used for zero vowel. The number of ligatures is different, as these are much more numerous in Indic scripts, and their structure as well: they are oriented from top to bottom in Brāhmī and Kharoṣṭhī, but from right to left in the "Unknown script". The zero diacritic for the long and perhaps short /a, ā/ is similar to the Kharoṣṭhī usage.

The shape of a number of letters is evidently akin to Kharoṣṭhī, especially the early ductus of Aśoka inscriptions (with its pointed vowel-initial akṣara):

aleph	– vowel-initial akṣara	*R*	– ra
K_1	– ka	*W*	– va
T	– ta	*H*	– ha
D	– dha	*S*	– sa
N	– na	and perhaps:	
B	– ba	*Y*	– ja

According to the common opinion, Kharoṣṭhī also derives from the Imperial Aramaic script. The number of persuasive similarities of identified signs of the "Unknown script" to the Aramaic script on the one hand and to Kharoṣṭhī on the other is roughly identical, numbering 9 to 11 letters (7 of them have both Aramaic and Kharoṣṭhī counterparts).[45]

Moreover, it is unlikely that the very idea of the diacritic notation of vowels could emerge among the anonymous creators of the "Unknown script" independently from the

[45] The number of sure similarities between Aramaic and Kharoṣṭhī, namely 12 signs, is also rather close; see Glass 2000: 14 (one should note that we know the reading of *all* signs of these two scripts).

endeavor of the Indian pundits who by then had carefully studied and described the phonetic system of the Vedic hymns (Katenina & Rudoj 1980: 72, 75). Finally, the absence of ideograms also links our script with Kharoṣṭhī.

Thus, the script might have been formed somewhere on the territory of the Achaemenid empire, where Aramaic chancery was common, and in the areas bordering northwestern India, where the Kharoṣṭhī script emerged, and reached its shape in the mid-3rd century BCE, the time of Aśoka (Salomon 1998: 46).

Already the first scholars who were confronted with the Issyk inscription compared its signs to the Orkhon-Yeniseian runes (see footnote 2 above). This viewpoint was criticised by V. A. Livšic, who noticed the similarity of no more than three signs, which could well be random (Livšic 1978: 84–86). Indeed, a few similar-looking signs in two randomly taken scripts will surely be noticed, but when these similarities are seen in the system of the script or at least their phonetic realisation, the possibility of coincidence sharply diminishes.

Now, when we know the likely reading of 16 signs of the "Unknown script", one can re-examine its relation to the Orkhon-Yeniseian runes. Some similarities are doubtlessly coincidental: *M* of the "Unknown script" and *z* of the runes are virtually identical. However, we can find undoubted similarity between our *aleph* and Runic[46] *a/ä*, *W* and *o/u*, *D* and d^1, *Γ* and Yeniseian *γ*, K_2 and a rare Yeniseian variant of *g* (Vasil'ev 1983: table 10, 23–24), *N* and n^1, *T* and t^2. These similarities can hardly be attributed to coincidence.

Although we see a number of similarities in the signs and their phonetic realisation of the "Unknown script" and Old Turkic runes, we do not see any similarity in the system of writing. The "Unknown script" is an alphasyllabary and the Orkhon-Yeniseian Runic script is, with some exceptions, alphabetic,[47] and very well adapted to Old Turkic phonetics with its synharmonism. One can thus state that the "Unknown script" was one of the sources of the Turkic Runic script, but not its direct prototype, since the creation of the latter presupposes a radical reform of its sources. Moreover, many correspondences between the early ductus of Sogdian and the runes which were noticed by Livšic, remain valid (Livšic 1978: 94–95). Furthermore, in two cases (*γ* and *g*) the "Unknown script" demonstrates similarity only with the Western, Yeniseian version of the Runic script, but not the Eastern, Orkhon variant. So, it was apparently used partially as a secondary, local source of the new script of the Ancient Turks.

Finally, one has to take into account the spatial and temporal distance between 3rd century BCE – 4th century CE Bactria and 7th – 8th century Mongolia and Southern Siberia, where and when the Runic script was formed. One would need to find

[46] Following Vasil'ev 1983: 95–147.

[47] Cf. Kyzlasov 1998 on the hypothesis of a syllabic origin of the Old Turkic Runic script.

intermediate stages to be sure of the exact relationship between the two scripts. There are two options: either the script survived in Bactria until the Early Middle Ages, or it was borrowed by ancestors of the Turks during their raids to the southwest. As we have seen above, there are no clear or even likely examples of this script in or around Bactria posterior to the Kushano-Sasanians. It is unlikely that these existed nevertheless, as one can hardly believe in such bad luck of archaeologists.

On the contrary, there seem to be Steppic examples of similar writing prior to the Turks, although very limited in quantity. Signs similar to our "Unknown script" have been found on four birchbark caskets from the excavations of A. M. Mandel'štam on the necropolis of Aymïrlïg in Tuva in 1977–1980,[48] and on a birchbark box from the early Tashtïk burial ground Chernoozernoe II in the Minusinsk Depression, Khakassia.[49] The burial with boxes from Aymïrlïg is dated to not earlier than the 2nd century CE, as I was kindly informed by N. N. Nikolaev; S. V. Pankova dates the box from Chernoozernoe to the 2nd–4th centuries (oral consultations). One can include the inscription without diacritics on the Ugam stone (Ad-8, almost half way to Siberia) in this series. Apart from that, there is a series of illegible rock inscriptions which are specified as special "South Yenisei Runes" by I. L. Kyzlasov (2020: 14–18); they are however dated to a much later period. The signs of all these inscriptions resemble our "Unknown script", while diacritics are virtually absent. The size of this paper does not permit to provide details on them and, moreover, the research is in need of the participation of a qualified specialist in Altaic languages.[50]

The bearers of the South Siberian cultures could have learned the script from Bactria thanks to trade (cf. Bactrian imports in Noin-Ula burials), migration or raids (Chionite invasion?) and have borrowed it. Meanwhile, the diacritics were lost and perhaps the script was adapted to the local Altaic language. This script remained glimmering there until the 7th or 8th century, when the development of the Second Old Turkic Kaghanate led to the need for a new script which was primarily used for proclamation ends, and in the course of its creation, elements of the originally Bactrian script were integrated.

---

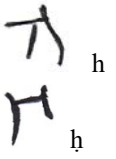
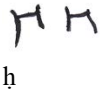
[48] Only sketches of two caskets have been published in small scale, Mandel'štam and Stambul'nik 1992: tables 78, 8; 83, 9. The tracings of two inscriptions of the 1977 excavations are present in Livšic' archive, and two more (from the excavations of 1979 and 1980) were kindly sent to me by Ja. V. Vassilkov. The last-mentioned casket is preserved in the Hermitage, and I am grateful to N. N. Nikolaev for arranging its autopsy. I got acquainted with the birchbark fragment from Černoozernoe thanks to the good service of conservator N. A. Vasil'eva.

[49] Gotlib 2006. Cf. also Kuznecov 2006 where similar (?) signs appear on a Tagar-period spinning whirl.

[50] I use the term "Altaic languages" here as an areal concept: the language may have been Turkic or Mongolic (or perhaps even Palaeoasiatic).

And yet, a number of signs remain orphans. We do not have clear analogies, e.g., to sign 19 (*Y*),[51] or to sign 12, which looks like *phi* with a *háček* (as in Ad-4) etc. It is possible that one should look for a source in the tamga signs of Central Asia (Yatsenko 2019) or in ancient scripts of Asia Minor.[52] However, we do not find a parallel to sign 12 in any of these sources.

The deciphered part of the script can be put together in Table 2.

I	II	III	IV	V	VI	VII
1.		W		va	o/u	*H* acc. to BHK
3.		H	 h ḥ			*W* acc. to BHK, cf. No. 6.
4.		B			b[1]	A rare variant in Old Turkic.
6.		X?	 ḥ	śa		Doubtful, cf. fn. 44.
8.		S				In Kharoṣṭhī in post-Aśokan inscriptions.
9.		R			r[1]	
11.		J?		ca		Doubtful, see Caveney 2023.
13.		Г			γ	

[51] Comparison with Kharoṣṭhī *ja* is possible but would need a phonetic explanation.

[52] This was noted already by Amanžolov 1971. Independently, I. S. Yakubovich proposed me to examine this possibility. The similarity of sign 4 (*B*) with *beta* in Carian is noteworthy (Adiego 2018: 155). A connection between the alphabets of Asia Minor and Bactria cannot be excluded. One can note the town of the Branchides in Bactria, who were punished by Alexander. Allies from Asia Minor were surely among the veterans who were left by Alexander in the cities he founded.

15.		K_2 = G?	q		g	Only in the Yeniseian variant of Old Turkic.
16.		T			t^2	
17.		K_1	q		q	
19.		Y		, ja		Perhaps identical to No. 11.
20.		D		ḍa, dha	d^1	
21.		M				
22.		N	,		, n^1	
23.		Š				
24.		’		a	, a, ä	See p. 154. The Kharoṣṭhī sign is as in Aśoka’s edicts
27.		Z?				In early texts, cf. p. 161.

Table 2. Deciphered signs of the “Unknown script” and their comparanda.

Legend to Table 2:

I. Number according to Bonmann *et al.* 2023, appendix 3.

II. The main shape(s) of the sign.

III. Transliteration (author’s proposals in bold).
IV. Aramaic prototype (Aśoka’s inscriptions, Achaemenid Bactrian documents).
V. Analogue in Kharoṣṭhī (Aśoka’s inscriptions, after Glass 2000).
VI. Old Turkic Runes (after Vasil’ev 1983).
VII. Notes.

9. On the language encoded by the “Unknown script”

The absolute majority of the words read in the inscriptions in the “Unknown script” have safe parallels in Bactrian, especially in the Bactrian version of the bilinguals. This observation led Harry Falk to consider that the Bactrian language was recorded by this script (Falk 2023: 10–11). Bonmann and her team on the contrary notice a series of differences between the two languages (Bonmann *et al.* 2023: 324–326). Seemingly, the close relation between the language of the “Unknown script” and Bactrian could be explained as being due to the early stage of decipherment: scholars tend to find phonetically similar words and proper names, and the abundance of Bactrianisms could be taken as being borrowed from a language with a higher status.

However, the similarity to Early Bactrian in fact is visible also in the function words: the deictic pronouns *ʾï-Dï* – ειδι (AG I, 1, 8INC, Chingiz-tepe, Ad-2, Fayaz-tepe), *ʾï-Mï* – ειμι (CA-03, Tokkuz-tepe, cf. above, also perhaps *ʾə-Mā* in the beginning of the Issyk inscription and in KT-04:*ʾā-Mï*) and the relative pronoun *(K1)ï-Dï* – κιδι (DN III, 6; Bonmann *et al.* 2023: 315). As we have shown on the examples of the formula “king of kings” and the name of Wema Taktu, the direct and oblique cases of this language and early Bactrian were identical. Such structural coincidences are impossible in the case of borrowing and can hardly be taken as the result of language interference.

In some cases, it seems that the “Unknown script” indeed is a variant transcription of Bactrian. In this respect the ostracon Ad-2 from Fayaz-tepe (cf. above) is quite informative. It is read as *ʾï-Dï* | *Mā-Mā-Də/ï-(K2)*, and among the Bactrian epigraphy of Kara-tepe one finds the text Ειδι μα-μαδογο αβο-ι βαο ζουρια ‘this vessel for the king’s libation’ (Livšic 1996). As we see, the beginning of the two inscriptions is very close.

On the other hand, there are also differences between the two languages: the preposition αβο in Bactrian corresponds to *Bï* in the “Unknown script”, and the article or *ezafe* ι/ια in the former perhaps corresponds to *Hə* in the latter,[53] and its syntactic role was perhaps also somewhat different. In the latter language, it seems, the reflex of OIr. **ṛzata* ‘silver’ survived, while in the former it was replaced by σιμ- (ultimately from Greek ἄσημος, MP *asēm*). In the latter language, as in a series of modern Iranian

[53] Both cases can be explained by a different accentuation in the Old Iranian prototype: **ábi*, **hayá* in Bactrian and **abí*, **háya* in the language of the “Unknown script”.

languages and dialects, the reflexes of the *U* and *I* vowels likely merged, and they are perfectly differentiated in Bactrian in the majority of cases. For some reason, the first two letters of inscription III of Dashti Nawur resist reading, although they might contain the date (usually, an easy part of inscriptions).

Consequently, the language of the inscriptions in the "Unknown script" is a dialect which was close to, yet distinct from Bactrian. We observe perhaps differences within the language between the old and new inscriptions, such as possibly the voicing of *-t-* > *-d-*. It was hardly the language of the Sakas or Yuezhi, which is likely to have been more distant from Bactrian. These tribes invaded Bactria from the North and North-East, and their homeland was an unlikely place for the formation of a script based on Aramaic and related to Kharoṣṭhī.

In which part of Bactria or its neighbours should one locate this language? One of the ways to shed light on this is to consider the reasons and means of production of inscriptions (what was called "socioscriptorics" by V. V. Vertogradova). Bactria and its neighbours were a land of a very developed monetary economy. The coin legends are in a variety of scripts: Greek, Aramaic (also in Middle Iranian languages), Kharoṣṭhī, Brāhmī, Bactrian, and Middle Persian. However, we do not know any coin legend in the "Unknown script". One can surely look for various explanations of this *argumentum ex silentio*, but for me a sound hypothesis is that the speakers of this language were astray of any monetary economy. This feature was characteristic (up until the ethnographic past) of the mountain societies of the Hindukush and Pamirs (Lurje 2012: 26–27; 2022: 260). From this observation one could tacitly guess that the speaking area of the language was in the mountain area near Bactria.

Another observation is based on the two earliest inscriptions, from Issyk and Aï Khanoum. Both are incised on silver products, and while the Issyk bowl is a very simple artefact, the ingot from Aï Khanoum is hardly more than a billet, from which the city artisans might strike coins or even create artwork. One can suppose that both items were produced not far from silver mines and inscribed there as well. The main regional silver mines of the Middle Ages were located at the Panjshīr river, modern Panjshēr. These silver mines were exhausted by the Mongol period, but until then they supplied the bulk of the silver coinage of the Samanids and other dynasties. The silver mines of Nuqri-Khāna 'Silver house' were located there and radiocarbon analysis of slags indicates that exploitation started in the early Bronze age (Thomalsky *et al.* 2013: 219–224). Other famous silver mines were located in Ghūr in the western Hindukush.

Immediately to the east of the Panjshēr, beyond the mountain pass, lies Munjān, where a dialect close to Bactrian is still spoken, and, as in the language of the "Unknown script", the reflexes of OIr. **u*, **ū*, **i*, **ī* merge in Munji. To the north of the Panjshēr, downstream Kokcha river, lies Aï Khanoum, wherefrom one of the earliest inscriptions

originates, and downstream the Panjshēr, at its exit to the lowland, is Begram, ancient Kapisa, wherefrom another inscription comes (Ad-5), and further south is Gandhāra, the cradle of the Kharoṣṭhī script. One can thus prudently set up the hypothesis that the “Unknown script” mirrors a vernacular of the ancient mountaineers of the Hindukush, where up to now a variety of relic Indo-Iranian languages is observed, which likely were much more numerous earlier.[54]

Of course, one has to take into consideration that several different languages could have used this script (including, most likely, some Altaic ones), but for reasons of economy we now take for granted that it was one language or a group of closely related dialects. The sequence $^{\partial}$-$M^{\bar{a}}$ in the beginning of the Issyk inscription seems to be a variant of $^{\ddot{\imath}}$-$M^{\ddot{\imath}}$ of the later texts.

10. Conclusions

Thanks to the discoveries of the last few years, the “Unknown script” of Bactria has become less unknown. We know the reading of the majority of the signs and diacritic marks, we know that it coded an Eastern Iranian language, perhaps closely related to Bactrian, and we know a dozen words and some grammatical forms. We know that the script derives from Imperial Aramaic and in its development underwent changes similar to Kharoṣṭhī, and it became one of the sources of the Old Turkic Runic script.

What to call this script? According to the suggestion of Harry Falk, it should be called the Wema Takhtu script, or, abbreviated, “WT script” (Falk 2023: 11), and it was used for the Bactrian language. We know, however, that the script emerged at least three centuries before the coronation of Wema Takhtu, and was at best reformed during the early Kushan period (in fact, it would be more logical to call *Bactrian* the Wema Taktu script since the earliest doubtless examples of it are dated to his rule!). S. Bonmann, J. Halfmann and N. Korobzow propose either “(Issyk)-Kushan script”, or, more hypothetically, “Eteo-Tocharian” (Bonmann *et al*. 2023: 325–326). I do not think that any of these denominations is fortunate: the first (even when removing the brackets) would immediately be confused with Kushan Bactrian; and the second is an anglicised version of the “*étéo-tokharien*” of André Maricq, who after the discovery of Surkh-Kotal coined this term to refer to the Bactrian language in Greek letters. As I tried to show

[54] The so-called Old Vanch language in the northwestern Pamirs became extinct in the 19th century (Laškarbekov 2008); in the early 8th century the pilgrim Huichao/Hyecho noticed the presence of a separate, neither Tokharistanian (= Bactrian), nor Turkic tongue in Khuttal (Kulab region, tr. Yang *et al.* 1985: 54–55). In the Early Islamic period the dialects of Gharjistān (Hindukush) and Banjhīr (Panjshēr) were mentioned (Le Strange 1905: 415, 350).

above, this language and script hardly belonged to the Yuezhi-Tocharians, who captured Bactria and established the Kushan kingdom.

The suggestions presented in the last section are hardly enough for calling it "Hindukush" or "Paropamises" script. In many cases, when we do not know the autochthonous name of the language, it is based on the most significant discovery. Thus, the term "Script of Dashti Nawur" seems to be the most appropriate for the time being.

Abbreviations

Av.	Avestan	Khot.	Khotanese
Bactr.	Bactrian	MP	Middle Persian
DAFA	Délégation Archéologique française en Afghanistan	OIr	Old Iranian
		OP	Old Persian
DN	Dashti Nawur inscription	pl.	plural
gen.	genitive case	Sogd.	Sogdian

References

ADIEGO, I. X., 2018. 'Local adaptations of the alphabet among the non-Greek peoples of Anatolia', in *Paths into script formation in the ancient Mediterranean* (ed. S. FERRARA & M. VALÉRIO), Roma, 145–162.

AKIŠEV, K. A., 1978. *Kurgan Issyk. Iskusstvo sakov Kazaxstana*, Moskva.

AMANŽOLOV, A. S., 1971. 'Runopodobnaya nadpis' iz sakskogo zaxoronenija bliz Alma-Aty', *Vestnik Akademii nauk Kazaxskoy SSR* [*Bulletin of the Academy of Sciences of the Kazakh SSR*] 12 (320), 64–66.

BERNARD, P., 1980. 'Campagne de fouilles 1978 à Aï Khanoum (Afghanistan)', *CRAI* 124, 435–459.

BLAŽEK, V. & M. SCHWARZ, 2016. *The early Indo-Europeans in Central Asia and China. Cultural relations as reflected in Language*, Innsbruck.

BOBOMULLOEV, B. S., Š. XODŽAEV & M. G. BOBOMULLOEV, 2022. 'Naskal'nye nadpisi uščelja Almosi (predvaritel'nye rezul'taty issledovanij)', *Esikskie čtenija* III, 59–65.

BOBOMULLOEV, B. S. & T. HASUIKE, 2013. *Kalai Kofarnigan. Katalog naxodok 2007–2013*, Dušanbe.

BOBOMULLOEV, M. G., 2023. 'Jodgodii Almosī', *Merosi Šarq* 1/3, 30–35.

BONMANN, S., J. HALFMANN, N. KOROBZOW & B. S. Bobomulloev, 2023. 'A partial decipherment of the unknown Kushan script', *TPS* 121/2, 293–329.

CAVENEY, G., 2023. *A new reading of the Issyk Inscription based on the recent decipherment of the Kushan script*. Available at: www.academia.edu/104775728/A_new_reading_of_the_Issyk_inscription_based_on_the_recent_decipherment_of_the_Kushan_script (accessed on: June 1, 2024).

CRIBB, J., 2023. ‘A Kushan royal inscription among the Almosi Rock Inscriptions recently discovered in Tajikistan’, *JRAS* 34/2, 361–375.

DAVARY, G. D., 1982. *Baktrisch, ein Wörterbuch auf Grund der Inschriften, Handschriften, Münzen und Siegelsteine*, Heidelberg.

DAVARY, G. D. 2022. *Drei neu entdeckte Inschriften in Baktrisch und in „Issek-Baktrisch“ aus Tadschikistan*, Tours.

DURKIN-MEISTERENST, D., 2014, *Grammatik des Westmitteliranischen (Parthisch und Mittelpersisch)*, Wien.

ÈDEL’MAN, D. I., 1986. *Sravnitel’naja grammatika vostočnoiranskix jazykov. Fonologija*, Moskva.

EFIMOV, V. A., 2009, *Jazyk parači. Grammatičeskij očerk. Teksty. Slovar’*, Moscow.

EMMERICK, R. E., 2009. ‘Khotanese and Tumshuqese’, in *The Iranian languages* (ed. G. WINDFUHR), London–New York, 377–415.

FALK, H., 2009. ‘The Name of Vema Takhtu’, in *Exegisti Monumenta. Festschrift in honour of Nicholas Sims-Williams* (ed. W. SUNDERMANN *et al.*), Wiesbaden, 105–116.

FALK, H., 2021. ‘Frühstadien des virāma in indischen Schriften’, in *Wort- und Formenvielfalt. Festschrift für Christoph Koch zum 80. Geburtstag* (ed. A. JOURAVEL & A. MATHYS), Berlin, 127–131.

FALK, H., 2023. ‘Wema Takhtu, the graveyard near Almosi, and the end of an “unknown” script’, *ARIRIAB* 26, 253–264.

FUSSMAN, G., 1974. ‘Documents épigraphiques kouchans’, *BEFEO* 61, 1–66.

FUSSMAN, G., 2011. *Monuments buddhiques de Termez I. Catalogue des inscriptions sur poteries. Avec une contribution par Nicholas Sims-Williams et la collaboration d’Éric Ollivier*, Paris.

GARKAVEC, A. A., 2018. ‘Issykskaja nadpis’: novyj variant pročtenija’, *Èpigrafika Vostoka* 33, 25–38.

GERSHEVITCH, I., 1954. *A grammar of Manichaean Sogdian*, Oxford.

GHOLAMI, S. 2011. ‘Definite Articles in Bactrian’, *Topics in Iranian Linguistics* (ed. A. KORN *et al.*), Wiesbaden, 10–22.

GLASS, A., 2000. *A preliminary study of Kharoṣṭhī manuscript paleography*, Washington.

GOTLIB, A. I., 2006. ‘Novye sjužety taštykskogo iskusstva (predvaritel’naja publikacija)’, *Sovremennye problemy arxeologii Rossii. Materialy Vserossijskogo arxeologičeskogo s”ezda (23–28 oktjabrja 2006 g., Novosibirsk)* (ed. A. P. DEREVJANKO & V. I. MOLODIN), Novosibirsk, 78–83.

GRENET, F., 2015. ‘Zoroastrianism among the Kushans’, in *Kushan Histories. Literary sources and selected papers from a Symposium at Berlin, December 5 to 7, 2013* (ed. H. FALK), Bremen, 203–239.

HALFMANN, J., S. BONMANN, S. FRIES & N. KOROBZOW, 2023. ‘A new reading of the Bactrian part of the Dašt-i Nāwur Trilingual (DN I)’, *IIJ* 67/1, 5–51.

HAMELIN, P., 1954. 'Matériaux pour servir à l'étude des verreries de Begram (suite)', *Cahiers de Byrsa* 4, 153–183.

HARMATTA, J., 1994. 'Languages and literature in the Kushan Empire', in *The development of sedentary and nomadic civilizations, 700 B.C. to A.D. 250* (ed. J. HARMATTA), Paris, 407–431.

HOFFMANN, K. & B. FORSSMAN, 2004. *Avestische Laut- und Flexionslehre*, Innsbruck.

JAKUBOV, Ju. Ja. & A. SULAJMONZODA, 2022. *Karon – rezidencija pravitelej drevnego Darvaza*, Dušanbe.

KATENINA, T. E. & V. I. RUDOJ, 1980. 'Lingvističeskie znanija v drevnej Indii', in *Istorija lingvističeskix učenij. Drevnij mir* (ed. A. DESNICKAJA & S. KACNEL'SON), Leningrad, 66–91.

KUZNECOV, N. A., 2006. 'Pamjatniki drevnej èpigrafiki južnoj Sibiri (skifskoe vremja)', in *Sovremennye problemy arxeologii Rossii. Materialy Vserossijskogo arxeologičeskogo s"ezda (23–28 oktjabrja 2006 g., Novosibirsk)* (ed. A. P. DEREVJANKO & V. I. MOLODIN), Novosibirsk, 29–31.

KYZLASOV, I. L., 1998. 'Materialy k rannej istorii tjurkov. II. Drevnejšie svidetel'stva o pis'mennosti (čast' pervaja)', *Rossijskaya arxeologija* 1998/I, 71–83

KYZLASOV, I. L., 2020. 'K poznaniju neraššifrovannogo pis'ma', *Rossijskaja tjurkologija* 28–29/3–4, 9–22.

LAŠKARBEKOV, B. B., 2008. 'Starovančskij jazyk', in *Osnovy iranskogo jazykoznanija. Sredneiranskie i novoiranskie jazyki* (ed. V. EFIMOV), Moskva, 61–109.

LE STRANGE, G., 1905. *The lands of the Eastern Caliphate*, New York.

LITVINSKIJ, B. A., 1980. 'Kalai Kafirnigan (raskopki 1975 g.)', *Arxeologičeskie raboty v Tadžikistane* 15, 120–146.

LITVINSKIJ, B. A., 1982. 'Kalai Kafirnigan (raskopki 1976 g.)', *Arxeologičeskie raboty v Tadžikistane* 16, 106–132.

LITVINSKIJ, B. A., 1983. 'Buddijskij xram Kalai Kafirnigan (Južnyj Tadžikistan) i problemy istorii i kul'tury Central'noj Azii', in *Istorija i kul'tura Central'noj Azii* (ed. G. M. BONGARD-LEVIN *et al.*), Moskva, 282–356.

LITVINSKIJ, B. A. & X. Muxitdinov, 1969. 'Antičnoe gorodišče Saksanoxur (Južnyj Tadžikistan)', *Sovetskaja arxeologija* 2, 160–178.

LIVŠIC, V. A., 1969. 'K otkrytiju baktrijskix nadpisej na Kara-tepe', in *Buddijskie peščery Kara-Tepe v Starom Termeze* (ed. B. Ja. STAVISKIJ), Moskva, 47–81.

LIVŠIC, V. A., 1978. 'O proisxoždenii drevnetjurkskoj runičeskoj pis'mennosti', *Sovetskaja tjurkologija* 4, 84–98.

LIVŠIC, V. A., 1996. 'Baktrijskaja nadpis' na fragmente sosuda iz Kara-tepe (naxodka 1984 g.)', in *Buddijskie peščery Kara-tepe v Starom Termeze* (ed. B. Ja. STAVISKIJ), Moskva, 168–172.

LURJE, P. B., 2010. *Personal names in Sogdian texts*, Wien.

LURJE [LUR'E], P. B., 2012. 'Sogdijskie busy na "Kryše Mira"', *Vostok* [*Oriens*] 1, 23–28.

LURJE, P. B., 2022: ‘Rev.: Richard E. Payne & Rhyne King (ed.), The Limits of Empire in Ancient Afghanistan’, *Orientalistische Literaturzeitung* 117/3, 257–261.

LURJE [LUR’E], P. B., 2024. ‘“Neizvestnoe pis’mo” Baktrii: Neopublikovannye materialy i novye interpretacii’, *VDI* 84/2, 319–356.

MANDEL’ŠTAM, A.M. & E. U. STAMBUL’NIK, 1992. ‘Gunno-sarmatskij period na territorii Tuvy’, in *Stepnaja polosa aziatskoj časti SSSR v skifo-sarmatskoe vremja* (ed. M. G. MOŠKOVA), Moskva, 196–205.

MARICQ, A., 1958. ‘La grande inscription de Kaniṣka et l’étéo-tokharien, l’ancienne langue de la Bactriane’, *JA* 246, 345–440.

MARTINEZ-SÈVE, L., 2020. ‘Greek Power in Hellenistic Bactria: Control and Resistance’, in *The limits of empire in ancient Afghanistan* (ed. R. E. PAYNE & R. KING), Wiesbaden, 81–112.

MORRIS, L., 2021. *The Begram Hoard and its Context*, Munich.

NASIM KHAN, M., 2007. ‘Kohi script from Gandhara’, *Gandhāran Studies* 1, 89–118.

NASIM KHAN, M., 2020. ‘Kohi or Puṣkarasāri. The story of an undeciphered script from Gandhāra’. A preprint is available at: www.researchgate.net/publication/342697750_Kohi_or_Puskarasari_The_Story_of_an_un-deciphered_script_from_Gandhara (accessed on July 21, 2024).

PERRY, J. R. & A. A. SADEGHI, 1999. ‘Eżāfa’, *Encyclopaedia Iranica* IX/2, 127–128.

PODUŠKIN, A. N., 2020. ‘Runopodobnoe pis’mo iz Ugama’, *Margulanovskie čtenija* 2020, 198–212.

POWELL, E. A., 2024. ‘Breaking the Code of the Kushan Kings’, *Archaeology* 77/2, 50–55.

RAPIN, C., 1992. *La trésorerie du palais hellénistique d’Aï Khanoum*, Paris.

RAPIN, C. 2007. ‘Nomads and the shaping of Central Asia: From the Iron Age to the Kushan period’, in *After Alexander. Central Asia before Islam* (ed. J. CRIBB & G. HERRMANN), Oxford, 29–72.

ROZENFEL’D, A. Z., 1956. ‘Darvazskie govory tadžikskogo jazyka’, *Trudy Instituta jazykoznanija AN SSSR* 6, 196–276.

RTVELADZE, È. V. 2005. *Civilizacii, gosudarstva, kul’tury Central’noj Azii*, Tashkent.

SALOMON, R., 1998. *Indian epigraphy. A guide to the study of inscriptions in Sanskrit, Prakrit, and the other Indo-Aryan languages*, New York–Oxford.

SIMS-WILLIAMS, N., 2007. *Bactrian documents from Northern Afghanistan II: Letters and Buddhist texts*, London.

SIMS-WILLIAMS, N., 2008. ‘The Bactrian inscription of Rabatak: A new reading’, *BAI* 18 (2004[2008]), 53–68.

SIMS-WILLIAMS, N., 2010. *Bactrian personal names*, Wien.

SIMS-WILLIAMS, N., 2025. *Bactrian documents* IV: *Documents from the south of the Hindukush,* I, London.

SKJÆRVØ , P. O. 2002. *Khotanese manuscripts from Chinese Turkestan in the British Library*, London.

STRAUCH, I., 2012. *Foreign sailors on Socotra. The inscriptions and drawings from the Cave Hoq*, Bremen.

SÜLEJMENOV, O., 1970. 'Žeti suvdyn köne žazbalary', in *Qazaq ädebieti*, Alma-Ata.

THOMALSKY, J., B. BRÄUTIGAM, M. KARAUCAK & S. KRAUS, 2013. 'Early mining and metal production in Afghanistan: The first year of investigations', *Archäologische Mitteilungen aus Iran und Turan* 45, 199–230.

TORGOEV, A. I., forthcoming. 'Ešče o date kurgana Issyk', in *Gorizonty i paradigmy arxeologii Jevrazii na sovremennom ètape. Sbornik materialov konferencii k 100-letiyu K. A. Akiševa* (ed. A. ONGARULY), Almaty, 100–106.

ÜNAL, O., 2019. 'The Issyk and Ai Khanum inscriptions revisited', *Ural-Altaische Jahrbücher* 27, 172–210.

VASIL'EV, D. D. 1983. *Grafičeskij fond pamjatnikov tjurkskoj runičeskoj pis'mennosti aziatskogo areala (opyt sistematizacii)*, Moskva.

VERTOGRADOVA, V. V., 1995. *Indijskaja èpigrafika iz Kara-Tepe v Starom Termeze. Problemy dešifrovki i interpretacii*, Moscow.

VERTOGRADOVA, V. V., 2002. 'Neizvestnoe pis'mo v mnogoalfavitnom prostranstve Baktrii', *Peterburgskij Rerixovskij sbornik* 5, 129–147.

XURŠUDJAN, È. Š., 2011. 'Nadpis' na čaše iz Issykskogo kurgana. Predvaritel'nye rezul'taty', *Iran-name* 20, 95–105.

YANG Han-sung, Y. JAN, Sh. IIDA & L. W. PRESTON, 1985. *The Hye Ch'o diary: Memoir of the pilgrimage to the Five Regions of India*, Berkeley.

YATSENKO, S. A. (ed.), 2019. *Tamgas of pre-Islamic Central Asia*, Samarkand.

YOSHIDA Yutaka, 2009. 'Sogdian', in *The Iranian languages* (ed. G. WINDFUHR), London–New York, 279–335.

ZAV'JALOV, V. A., 2008. *Kušanšaxr pri Sasanidax. Po materialam raskopok na gorodišče Zartepa*, Sankt-Peterburg.

Postscript

After the present translation was completed, an important paper of Harry Falk, 'Mit Wema Takhtu unterwegs, auch in Almosi und Reh?' appeared in *Asiatische Studien* 79/1 (2025), 39–73; Nicholas Sims-Williams kindly shared with me his manuscript 'Bactrian in two scripts: Greek and Kushan', to appear in *Indo-Iranian Journal* 68/3 (pp. 185–214; online publication: 4 Nov. 2025); furthermore, Jakob Halfmann sent me the draft of his forthcoming paper 'Bactrian in Issyk-Kushan script: Additional readings and decipherments'. These articles are significant contributions to the decipherment of the script and they arrived too late to be taken into account in the present paper; the larger part of its main conclusions, I believe, nevertheless remains valid.

On the origin of Bactrian final -o*

Francesca MICHETTI

This article investigates the origin of the Bactrian use of the letter *omicron* as an orthographic device to mark the end of the word: "final -o rule". While this orthographic rule is systematically applied in documents later than the 4th cent. CE, earlier inscriptions and documents still attest the use of other final vowels (-α, -ε, -ι), interpreted variously as spelling variants or as historical spellings. By collecting and discussing every attestation of final -α, -ε, and -ι in early Bactrian, it will be suggested that the distribution of final -α, -ε, and -ι reflects a vital nominal system, largely comparable to that of Sogdian and Khotanese. It will be argued that final -o in Kushan Bactrian is a historical spelling, pronounced as *-ə* but originally representing the ending of the acc.sg. case *-u* (< OIr. **-am*). Finally, the argument will be made that the Kushan Bactrian inscriptions capture a specific moment in the diachronic evolution of Bactrian phonology that is best understood through the lens of the variationist approach, in which the coexistence of innovative and obsolete standards is viewed as systematic. The innovation consisted in the centralisation of all final vowels to *-ə* (written -o), but the distinction between final vowels (-α, -ε, -ι) was still a possible, however obsolescent, variant.

1. Introduction

Since the very beginning of Bactrian studies, the final *omicron* appearing at the end of most Bactrian words has been interpreted as an orthographic sign with no phonetic value, functioning as a word- and morpheme-divider.[1] This is certainly the case for the bulk of the Bactrian corpus, represented by the manuscript documents published in Sims-Williams' *Bactrian Documents* (henceforth *BD*),[2] where *omicron* is generalised in final position (with few exceptions, cf. *BD* II: 40; *BD* IV: 5, on which see §5.4). This

* This paper is a product of the 2022 PNRR PRIN project (Prot. P2022LWSYY), PI Giancarlo Schirru, University of Naples "L'Orientale". I am grateful to Maria Carmela Benvenuto, Ching Chao-jung, Claudia A. Ciancaglini, Alessandro Del Tomba, Jakob Halfmann, Agnes Korn, Marco Mancini, Michaël Peyrot and Nicholas Sims-Williams for the useful remarks offered at various stages in the process of writing this paper.

[1] Cf. Henning's (1960: 50) well-known remark on the use of final -o in the Surkh Kotal inscription (on which see the next paragraph). Previously, the same interpretation had been suggested by J. de Menasce (*apud* Ghirshman 1953: 124) based on a small inscription on a Hephthtalite seal.

[2] More precisely *BD* I, *BD* I^2, *BD* II, *BD* III, and *BD* IV as given in the references.

10.29091/9783752003635/005

orthographic rule, which may be dubbed "final -ο rule", has the practical effect of helping the reader parse the text, since the letter *omicron* is never linked to the following letter in the cursive script of the documents, hence creating a graphic separation between the words (cf. Sims-Williams 1989: 233, fn. 24).

The transformation of a vowel letter into an orthographic sign to mark the end of the word is a clear indication of the loss of Old Iranian final vowels in the course of the history of the Bactrian language (cf. Sims-Williams 1989: 234). The absence of any counterpart to the final -ο in the only extant Bactrian fragment in Manichaean script from Qočo (fragment M1224 of the Berlin Turfan collection) confirms that final -ο had no phonetic value by the time the Manichaean document was written (7th–8th cent. CE), cf. ManB. *qyrd* = G. κιρδο.[3] However, in the earliest phase of the language, documented by the Kushan inscriptions found in Afghanistan and southern Tajikistan (2nd–3rd cent. CE) and by the earliest manuscript documents (4th cent. CE), the final vowels -α, -ε, and -ι are still attested beside -ο, suggesting that final *omicron* still constituted a phonological reality at that stage (Sims-Williams 1989: 234; Huyse 2003: 60–61). This paper aims at assessing the distribution of the final vowels in the earliest phase of the Bactrian language and establishing their etymological origin. Moreover, it seeks to address the related question of the origin of the final -ο rule in the broader context of the loss of the final vocalism in Bactrian.

2. The debate around the final -ο rule

In his seminal study on the Surkh Kotal monumental inscription (SK4M),[4] the first discovered Bactrian text of significant length, Henning (1960: 50) described the function of final -ο as that of a word-divider. It is worth quoting Henning's remark in full:

> It appears to be the rule that all words must end in a vocalic letter. If a word in fact ended in a consonant, an omicron was added to it […]. It is possible that here and there a closing omicron expressed a vowel pronounced in speech at the time of the inscription; in most cases it functioned virtually as a word-divider. […] the omicron even appears in composition […]. If an inflexional ending or a suffix is added to the word, the final omicron disappears.

Henning (1960: 52) then proceeded to reconstruct a two-case nominal system, where a direct (= nom.-acc.) singular case -Ø represented by the "silent" -ο is opposed to an

[3] The Manichaean fragment was published by Sims-Williams 2009; its dating is discussed in Sims-Williams 2011a.

[4] For the abbreviations of the quoted texts and the relative edition see the list at the end of the article.

oblique singular in -ι (< OIr. *-*ahya*) with both possessive and agentive meaning, while, in the plural, -ε marks the direct case and -ανο the oblique case.

In the subsequent decades, much of the scholarly debate revolved around rejecting Henning's description of final -ο as a word-divider in an attempt to assign to it a phonological value, sometimes overlooking his nuanced position about the possibility of it occasionally representing a vowel still pronounced at the time of the inscription. The main criticism addressed at Henning, especially after the publication of the two other versions of the Surkh Kotal inscription (SK4A, B; Benveniste 1961), regarded his interpretation of the final -ο as a "silent" sign to represent the ending -Ø of the dir.sg. case. Comparing the parallel passages of the three versions, it was clear that Henning's view could not always be upheld, because several nouns that had final -ο in SK4M now presented a final -α, -ε or -ι in the correspondent passages of SK4A or SK4B. Humbach (1960: 18, 47; 1976: 65) also noticed that the application of Henning's model to Kushan coin legends such as ϸαοναvο ϸαο κανηϸκι κοϸανο (NumKan 1304) led to the paradoxical interpretation 'King of Kings, the Kushan *of* Kanishka', with the dynastic name κοϸανο (Henning's dir.sg. case) governing the personal name κανηϸκι (Henning's obl.sg.).

This observation led Humbach to conclude that the final vowels were simply alternating freely as spelling variants, and that no reconstruction of the nominal system should be undertaken on the basis of random final vowels.[5] Altheim (1962: 17–20) also rejected the idea of -ο as a silent sign and regarded the preservation of final vowels in Bactrian as an archaic trait shared with "nomadic" languages such as Ossetic and Khotanese. However, the purpose of this argument was to substantiate the hypothesis of a "nomadic" origin of Bactrian, thereby challenging Henning's (1960: 47) now universally accepted identification as the indigenous language of Bactria. Another critic of the idea of final -ο as a word-divider was Morgenstierne (1970: 126). Observing the alternation between -α-, -ε-, -ι- and -ο- in internal position, e.g. in the word νιϸαλμο/νοϸαλμο 'seat' < OIr. **nišadman-* or νοβιχτο/νιβιχτο/ναβιχτο 'written' < OIr. **ni-pixšta-*, and the use of *omicron* as an epenthetic vowel, he argued that *omicron* represented in those cases "a reduced vowel *ə*, phoneme or allophone".[6] It followed that

[5] Humbach's criticism went far beyond this point and invested much of Henning's interpretation of the Surkh Kotal inscription. Notably, Humbach (1960, 1966) attempted to interpret the inscription as a Mithraic hymn. He later abandoned his own interpretation (Humbach 2003) and adhered to the idea that -α, -ε and -ι are historical spellings that were progressively replaced by the generalised final -ο, cf. §3.

[6] The idea that *omicron* represented "*une voyelle de timbre indéterminé*" both in internal and final position had already been put forward by Maricq (1958: 400, 409) in the first edition of the Surkh Kotal inscription. His interpretation of the text, as is inevitable with any pioneering edition of a newly discovered language, was soon superseded, cf. especially Henning's (1960: 50) remarks on his reconstruction of the nominal system. This caused many of his observations to be somewhat neglected.

also final -ο, as well as final -α, -ε, and -ι, could then be interpreted as representations of a centralised vowel *-ə* resulting from a process of general reduction of vowels in final position. The logical consequence of Morgenstierne's observation (although not explicitly stated by the author) is that the various final vowels simply alternated as free variants, as first proposed by Humbach 1960.

A similar conclusion was reached by Schwartz (1974: 410) in his seminal article about Iranian loanwords in the Tocharian languages. Based on the adaptation of the final vowels of Bactrian loanwords into Tocharian B, he concluded that the OIr. thematic stem vowels were still pronounced in Bactrian at the time of borrowing and that the graphic oscillations in final position were to be interpreted as "alternate representations of schwa" (cf. also Tremblay 2005: 435).

Other scholars set out to revise Henning's grammatical analysis, still rejecting the interpretation of -ο as a silent sign but trying to reconcile its primary vocalic value /u(ː)/ within the Graeco-Bactrian script with the variant spellings -α, -ε, and -ι.[7] Harmatta (1969: 350–352), taking -ο at face value /o/ and comparing it with Av. *-ō* (< OIr. **-ah*), suggested that -ο represented the nom.sg. ending of the thematic masculine nouns. Postulating a complicated development in which most final vowels were first reduced to *ə* and then labialised and confused with -ο, he tried to demonstrate how the ending of the nom.sg. -ο /o/ was generalised to most cases of all the nominal stems. The main problem with this reconstruction is that Harmatta's understanding of the nominal system is based on his own interpretation of a number of Bactrian words, which often relied heavily on conjectures and had no follow-up in subsequent literature.

A few years later, Lazard tried to make sense of the alternation between -ο and -α, -ε, and -ι in the three versions of the Surkh Kotal inscription by assuming that the spelling variants were "*tentatives diverses de noter des articulations qu'on ne savait trop comment rendre en alphabet grec*" (Lazard *et al*. 1984: 220). Accordingly, he proposed the following reconstruction of the nominal system: dir.sg. -Ø or *-u* (because it is most often expressed by -ο); obl.sg. *-i̊* (a transcription by which the author indicates a vowel like [ɨ] or [ʉ], because of the variation -ο/-ι/-ε); dir.pl. *-i* or *-i̊* (variation -ο/-ε); obl.pl. *-ān* or *-ānu* (cf. Lazard *et al*. 1984: 224–226).

3. Historical spellings?

In short, the theories about final -ο diverged on one main point: the significance of the alternative final vowels -α, -ε, and -ι. For one group of scholars, all Bactrian final vowels, including -ο, were functionally equivalent and simply alternated freely as representations of the mid central vowel *-ə*, as they did in internal position. The other group believed

[7] On the various phonological values of *omicron*, cf. Sims-Williams 1989: 233.

that -α, -ε, and -ι in nouns had morphological significance and their oscillating with -ο had a precise phonological meaning.

A compromise between these two opposing views is contained in Sims-Williams' (1989: 234) brief description of the Bactrian language, where he suggested that the vowels -α, -ε, and -ι could represent historical spellings, while "the -ο indicated that the final vowel has been reduced to *ə* or perhaps (as certainly in the late Bactrian of the Manichaean MS) lost altogether". This idea derived from his own observation that, when vowels other than the standard -ο occurred, they consistently reflected their Old Iranian origin, e.g. ιθα 'so' < OIr. **iθā*, οτι < OIr. **uta-uti* etc. (cf. Sims-Williams 1985: 114). Consequently, when -α, -ε, and -ι occurred with nouns and adjectives, they could be interpreted as inflexional endings. Henning's reconstruction of the nominal system could thus be retained,[8] with the addition of the spelling -ε interchanging with -ι for the obl.sg., which Henning could not acknowledge before the discovery of the two other versions of the Surkh Kotal inscription (cf. SK4B, 26 καραλραγγι = SK4M, 23 καραλραγγε). In addition, the ending -α could be traced back to the original *ā*-stem nouns, cf. λιζα 'fortress' < OIr. **dizā-*, φρομανα 'command' < OIr. **framānā-*.

Elaborating on this idea, Huyse (2003: 60–61, 99–100) compared the use of final *-y* in Middle Persian Sasanian inscriptions with that of -ο in Kushan Bactrian inscriptions. In his reconstruction, final IMP *-y*, originally representing *-ĕ^{h}* (a trace of various OP oblique cases), had become a silent sign in early Sasanian inscription and would later undergo a functional change to a purely orthographic sign for the end of the word. Similarly, the Kushan inscriptions would document a phase in which final -ο, originally representing "*une voyelle brève réduite*" (*ə*), was transitioning towards the role of a purely orthographic sign.

While this solution has the merit of reconciling the two opposing views exposed above by simultaneously acknowledging the etymological significance of final -α, -ε, -ι and the process of centralisation of final vowels suggested by Morgenstierne, it has some important implications on the way we understand the relationship between language and script during the Kushan period. First, if vowels in final position had been already completely reduced to *-ə* in Kushan Bactrian, the question arises why *omicron* was chosen among the other short vowels of the Graeco-Bactrian script. As we have seen, the other short vowels of the Greek alphabet (-α, -ε, -ι)[9] could serve the purpose of representing *-ə* just as well, as the spellings of the preverb Bactr. *nə-* in νιβιχτο/νοβιχτο/ναβιχτο clearly show.

[8] Cf. Sims-Williams 1989: 234; Sims-Williams 1988: 348; *BD* II: 40.

[9] Note that the letter υ could not serve this purpose as it was initially only used in the digraph ου to represent long *-ū*. Only later did it come to be used as the sign for the glottal fricative /h/, cf. Sims-Williams 2008: 60.

Moreover, the use of historical spellings implies a long scribal tradition in which words were memorised in a graphic form that reflected an older, obsolete pronunciation. Theoretically, nothing prevents us from assuming that Bactrian had a long writing tradition, even though it is normally held that it was first given a written form with the rise of the Kushan dynasty between the 1st and the 2nd cent. CE, simultaneously with the appearance of the first Bactrian inscriptions.[10] However, writing systems that feature many historical spellings always include a number of hypercorrections and inverse spellings, due to the difficulty of preserving the right etymological form for each word. It suffices to look at the Middle Persian inscriptions to find many examples, e.g. IMP <nydʾk> for *niyāk* 'grandfather', where <d> is used instead of the etymological <y> (OP *niyāka*; Henning 1958: 67; Gignoux 1972: 30). If the Kushan Bactrian final vowels -α, -ε, -ι were historical spellings, i.e. vowels that are no longer pronounced in speech, one would expect to find at least one case where such vowels were misplaced.

In order to properly assess this question, it is necessary to collect all the occurrences of final vowels different from -ο in the available documentation and analyse their distribution. In the survey below, I collected all the words with a final -α, -ε or -ι occurring in the Kushan Bactrian corpus (2nd–3rd cent. CE) and in the earliest manuscript documents (4th cent. CE). In Section 4, adverbs, particles and verbs are listed alongside their OIr. origin. It will be shown that Sims-Williams' (1985: 114) observation about the phonological significance of final vowels still holds true with the much larger amount of data available at present. As for the nouns, all the instances where the final vowels -α, -ε, and -ι represent a case ending have been collected in the Appendix and analysed according to their syntactic context. The result of this analysis will be discussed in Section 5.

4. Etymological final vowels

Adverbs and particles[11]

Final -ι, -ε < OIr. *-*i*

- -δι, -δε (clause-initial particle) < OIr. **uti*: þαοναvδε (DN1, 2), þαοναvδι (Nplate, 1), κιδι (DilV, 10; DN1, 5; Rab *passim*; SK4B, 13), (τα)καλδι (DilV, 12; Nplate, 4, 5; SK4B, 10), ταδι (DilV, 5, 11; Nplate *passim*; Rab *passim*; SK4A, 3; SK4M, 2; SK4B,

[10] Fussman 1974: 35–38; Morgenstierne 1970: 127; Sims-Williams 1988: 345; Sims-Williams 1989: 230.

[11] The origin of the suffix -ηλι (Sogd. -*ʾyδ*) attested in οαρηλι 'thither' (DilV, 6) is unknown. Another final -ι is attested in the preposition ανδ̣ιμαγ̣ι (Rab, 9), cf. Sims-Williams 2008: 61. This is cognate with MP and Parth. *hndym(ʾ)n* 'in front of, in the presence of', which go back to OIr. **ham-* + *daiman-* 'eye, sight' (Rastorgueva & Èdel'man 2003: 293). The final -ι in Bactrian may be a fossilised loc.sg. form from OIr. **handaiman*i 'in the sight (of), in the presence (of)'.

8), σιδι (DilV, 11; Nplate, 4; Rab, 8; SK4B, 2), οτι (Ayr, 5, 6; DilV, 10; DN1, 12; SK4B, 10; *BD* I, **aa**17–18, 19, 22, 34), (τα)σαγωνδι (Rab, 3, 14), ǫǫ̣ηλδι (Rab, 9), ασιδι (SK4B, 6), ατι (SK4B, 22), ειδι (186 KT;[12] 217 KT; Alm3, 1), τιδι[13] (Ayr, 4), πιδοοησατι (*BD* I, **aa**7, < πιδοησαδ° + -δι).

· -σι 'also' < PIr. *-*čit*: αβισσι (Rab, 21), μανδαρ̣σι (Rab, 6).

Final -α < OIr. *-*ā*[14]

· ιθα 'so' < OIr. **iθā* (DilV, 7, 11; SK4A, 25; SK4B, 22; *BD* IV, **km**2, 5).[15]

· μασκα 'hereupon' < OIr. **(i)ma-* + **uskā(t)* (Rab, 11).

· οιδρα αδα 'as far as' < OIr. **witarā(t) atā* (Rab, 5).

---

[12] Fussman's edition has ειδο, but a final -ι seems to me clearly readable here.

[13] On the origin of τιδι cf. Sims-Williams' contribution in this volume, p. 328.

[14] It may be debated whether the negative particle να 'not' (< OIr. **naid* with vocalism influenced by μα < OIr. **mā*, cf. *BD* II: 235) should be considered here. The particle, which is proclitic, is normally prefixed to the following word and is reduced to ν- when attached to an initial vowel (e.g. ν-ηρσιδο). However, in a few instances in relatively early texts (*BD* II, **ed**8: 465 or 475 CE; *BD* IV, **kk**9, **km**11, 15 and perhaps 19: late 4th cent. CE) να is graphically separated from the following word (cf. *BD* II: 115 n. 177; *BD* IV: 5), even before a vowel (*BD* IV, **km**14: να αβιρημο). Moreover, in later texts, both the prohibitive particle μα and the negative να are often written as separate words enlarged by the suffix -γο/-υο (ναγο, ναυο, μαγο, μαυο, cf. *BD* II: 47). The base of these late suffixed forms arguably results from the synchronic reanalysis of sequences like να-λαδο, μα-κιρο, where the negated word begins with a consonant. The same reasoning can be applied to the isolated writing of να in the early texts listed above. Thus, since the preservation of the original final vocalism in να is due to a prosodic feature of the word, it should not be considered an archaism.

On the other hand, the final -α of the 3sg. enclitic pronoun -ηια attested only in Kushan inscriptions (Ayr, 3; DN1, 7, 10, 11; Rab, 3, 10, 12) can be confidently considered an archaic feature, as this is an old spelling of -ηιο found in the *BD* (variant spellings -ηο in Nplate, 1, 2, 3; DilV, 6 and -ιηο in *BD* IV, **ko**, **kp**). Its origin is unclear. Etymologically, -ηια/-ηιο/-ηο is connected to OIr. *-*haį*, but the final diphthong was supposed to result into Bactr. *-ē* > *-e* rather than Bactr. *-ēya*/*-ēyə*. Sims-Williams and Cribb (1995–1996: 91) explained the ending with final -α as the result of the influence of the thematic gen.sg. ending *-*ahya*, while Kreidl (2024: 218) now suggests that it may be a hypercorrect abl.-instr. ending (cf. §5.1). One may perhaps consider the possibility of a univerbation with a postposition cognate to Av. *ā*, which was often suffixed to nouns in the ablative case, cf. Av. *aspāδa* (abl.sg.) < *aspāṯ* + *ā*. If this postposition was inherited by Bactrian and was used after the enclitic pronoun, we may reconstruct the development OIr. *-*haį* + *ā* > *-*aya*, *-*aį̯ya* (with insertion of a glide) > Bactr. *-ēa* (-ηο), *-ēya* (-ηια, -ηιο). A similar case of grammaticalisation of a postposition is attested in Khotanese, where the particle *jsa* < OIr. **hačā* is found both after nouns and pronouns in the abl.-instr. case, cf. OKh. sg. *-ī jsa*, pl. *-n jsa* (cf. Emmerick 1968: 258).

[15] In *BD* IV, **ka**, **ko**, **kv** attested also as ιθο. In later documents, the adverb is often written ιθαο (cf. *BD* II: 218). Final -o is added here as a word-divider, consistently with the general orthography of the documents. The vowel -α is retained also in later documents, as in the case of να (cf. fn. 14), because of the prosodic status of the adverb, which is proclitic.

Verbs

Final -ι, -ε < OIr. *-*i*

- -ημε, -ημι (1sg copula) < OIr. *-*mi*: σṭαδημι, σταδημε ‘I was’ (Nplate, 1, 3), ωσταδημι ‘I was established’ (Nplate, 2).
- -ενδι, -ινδι (3pl copula) < OIr. *-*anti*: νιβιχτιγενδι ‘are inscribed’ (Rab, 11, 17), φροχορτινδι ‘were displaced’ (SK4B, 8), μανο νινδι ρατο ‘I did not call them’ (DilV, 11).
- νιστι (3sg neg. copula), αστι (3sg copula) < OIr. **asti*: νᵉιστι χοτο ‘it did not require(?)’ (SK4B, 5–6), [ασ]ṭι (*BD* IV, **kh**, 13), ασṭι (*BD* IV, **kw**, 10).
- -δι (3sg pres. indic.) < OIr. *-*ti*: ριζδι ‘is called’ (Rab, 8, 10; Ayr, 5).
- -ι (infinitive) < OIr. *-*ayaį*: κιρδι (Rab, 8, 11, 12, 15).[16]

This first survey already shows that OIr. *-*i* is consistently reflected in the Bactrian spellings as -ι or -ε, whereas OIr. *-*ā* is reflected by Bactrian -α. In one case, namely the infinitive κιρδι, the final -ι can be traced back to OIr. *-*ayaį*. In no case the scribes seem to confuse the final vowels, for example by mistakenly writing ιθι instead of ιθα ‘so’ (OIr. **iθā*). Some texts tend to use final -ο more frequently, but this does not seem to depend on the relative chronology of the inscriptions. It suffices to compare the three versions of the Surkh Kotal inscriptions: SK4B has οτι, ατι, ταδι, καλδι, σιδι, ασιδι, κιδι where SK4A and M have οτο, ατο, ταδο, καλδο, σιδο, ασιδο, κιδο. All three versions were written in the same period, version B being the second in chronological order to be engraved,[17] so they can hardly reflect two subsequent stages of the centralisation of final vowels.

5. Final vowels as case endings (see Appendix)

The distribution of the final vowels in the nouns reveals a remarkably consistent picture of the Bactrian nominal system in this early phase.

5.1. Final -α

Final -α is mostly found with feminine nouns belonging to the OIr. *ā*-stem class, and with adjectives, pronouns and articles showing agreement to a feminine head noun. In (1)–(13) these elements all have syntactic functions typical of the nominative case, in that they are either subjects of nominal sentences, present tense clauses and intransitive past tense clauses, or direct objects of transitive past tense clauses (so-called ergative

[16] Cf. Sims-Williams & Cribb 1995–1996: 92.

[17] As established since its first publication, cf. Benveniste (1961: 139). *Contra* only Göbl (1965: 5–8), but cf. the arguments by Gershevitch (1966: 91) against his reconstruction.

alignment, on which see further). In some examples, the case marker is only present on the noun modifier (article or pronoun) while the noun has the ending -ο, cf. (4) μα λιζο. It is easy to trace back the final -α in these occurrences to the nom.sg.f. ending OIr. *-*ā* of the old *ā*-stems (cf. Sims-Williams 1985: 114). As we shall see (§5.2, 6), the use of the short Greek vowels such as -ε to represent final Bactrian vowels suggests that long vowels were shortened in final position. Thus, final -α probably represented a short central vowel [a].

In (14)–(20), nouns in -α appear governed by the prepositions αβο 'to, in', πιδο 'according to' and ανδ̣ιμαγ̣ι 'before, in front of'. As can be seen in (58)–(74), when old *a*-stem nouns are governed by these prepositions, they are often inflected in the genitive case (-ι/-ε).[18] Instead, *ā*-stem nouns consistently show the same ending -α even when they are governed by a preposition that would normally trigger an oblique case.

The only possible trace of an oblique case in this stem class is the ending -αα attested in the phrase (15) πιδο ια þοδ̣βαα φρομαγ̣α 'according to the *shudva*'s command', where þοδ̣βαα is semantically the possessor, hence a role prototypically marked by the genitive case. The same word occurs in two other places in the same inscription (Ayr), both times unfortunately in ill-preserved contexts where their syntactic function is not easy to determine (in Ayr, 2 as þοδβα or þοδβαο, in Ayr, 4 as þο̣δβα or þο̣δβαα̣, cf. Sims-Williams' contribution in this volume, p. 323). The etymology of this word is also not entirely clear. Sims-Williams (2010: 157) tentatively interpreted it as a title from OIr. **čyuta-pā-* '(official) who protects the expenditures', a root noun in °*pā-* from the IIr. verbal root **paH-* 'to protect' (cf. Kellens 1974: 223). Kreidl (2024: 214–215) cautiously suggested that -αα of (15) þοδ̣βαα may represent the outcome of *-*āyāh*, i.e. the gen.sg. ending of the masculine *ā*-stem inflexion, to which he assumes þοδβα(ο) was transferred (cf. again Sims-Williams' contribution in this volume, p. 326).

Although the possessive meaning of (15) þοδ̣βαα is practically beyond doubt, the identification of -αα as the gen.sg. ending of the *ā*-stem nouns poses some problems. First, the development OIr. *-*āyāh* > Bactr. -αα is admittedly unparalleled.[19] Moreover,

[18] When αβο and πιδο are used in locational sense, old *a*-stem nouns can also bear the ending -α, whose origin is dealt with below in this paragraph.

[19] Even taking -αα as an intermediate stage towards Bactr. -*ā* as suggested by Kreidl (2024: 215, with fn. 19, cf. also Sims-Williams in this volume, p. 326, fn. 8), the parallels invoked for this development are not conclusive. One of them is the word σαγο 'shadow' < OIr. **să̯yā-kā-*, which may point to OIr. *-*ayā-* > Bactr. -*ā-* rather than to OIr. *-*āyā-* > Bactr. -*ā-*, since the exact length of the first vowel cannot be reconstructed with certainty based on the cognates Sogd. *sy'k*, Yidgha *să̯yo* and MP *padisāy* (cf. *BD* II: 261).

The other parallel presents similar problems. Kreidl (2024: 215) compares νανα, the Bactrian name of the Mesopotamian goddess Nana, with the form ναναια attested on Greek coin legends, assuming that this last form represents the older stage of the Bactrian form. Even admitting that ναναια represents

if the etymology from IIr. **paH-* is correct (which, in the words of its proponent, is "no more than a vague possibility", cf. Sims-Williams, this volume, p. 326), þοδβα(ο) would originally be a stem in laryngeal. Laryngeal stems, when preceded by **-a-*, show up in Old Iranian languages as "irregular" (or apophonic) *ā*-stems, whose inflexion differs slightly from that of the regular *ā*-stem.[20] The supposed gen.sg. ending -αα, in particular, would come from OIr. **-āh* < IIr. **-aH-s*. Therefore, the ending -αα may not be representative of the Bactrian treatment of the oblique cases of the regular *ā*-stem nouns, but rather of the apophonic *ā*-stems.[21]

All in all, the present state of our documentation does not allow us to determine whether the Bactrian *ā*-stem nouns retained endings other than that of the nominative in the singular inflexion. The generalised use of final -α even after a preposition suggests that the oblique cases had merged with the nominative in the nouns of *ā*-stem class. It is possible that the acc.sg. ending OIr. **-ām* also resulted in Bactr. **-a* via **-ā(m)* > **-ā* (cf. further in this paragraph on the loss of consonants after long **-ā*). If this were the case, it may have contributed to the generalisation of the ending *-a* to the whole inflexion. However, it is equally conceivable that the final **-m* had a labialising effect on the previous vowel (cf. §7), as is visible in the Khotanese acc.sg.f. ending *-o* (Emmerick 1968: 273). A hypothetical ending Bactr. *-o* could have only been written with *omicron*, making it impossible to recognise due to the orthographic rules of Bactrian.

The form of the nominative seems to be generalised also in a few nouns from the old consonantal stem classes. In the early Bactrian documentation, two kinship nouns belonging to the old *r*-stem class are attested with final -α: (21) ια λογδα 'the daughter' and (22) πιδα 'father'. Due to the apophonic suffix **-tar-*, the inflexion of these nouns in Indo-Iranian languages alternates a nom.sg. form without *r*, e.g. OIr. **pitā* (nom.sg., cf. Skt. *pitā́*, YAv. *pita*), to forms with *r*, e.g. OIr. **pitaram* (acc.sg., cf. Skt. *pitáram*, YAv. *pitarəm*), OIr. **piθrāi̯* (dat.sg., cf. Skt. *pitré*, YAv. *piθre*). In Bactrian, kinship terms generally selected the form without *r*, as can be seen in the later documents, cf. *BD* I, **A**28; *BD* II, **zb**2 μαδο 'mother' (instead of *μαδαρο). This is already visible in the

the older Bactrian name of Nana and not just an independent Greek rendering, we are far from certain that -αια represented *-āyā*, with two long vowels in Bactrian. The only vowel that can be reasonably reconstructed as long is the last one, since Nana was a feminine deity and her name is inflected according to the feminine *ā*-stems, cf. (1), (3) in the Appendix. But we have no indication that the other vowel was long. In fact, the reconstruction of a first short vowel OIr. **-ay-* is suggested by the fact that the name of Nana shows up as νανη°, which points to an older **-ay-* rather than **-āy-* (cf. Gholami 2014: 32), in the theophoric PN νανηβανδο (compare Sogd. *nnyβntk*, an almost identical formation that bears no indication of a long *-ā-*, cf. Sims-Williams 2010: 94).

[20] Hoffmann & Forssman 1996: 123–125; Skjærvø 2009: 78; Cantera & Redard 2023: 61.

[21] Note, additionally, that in Old Persian the gen.sg. of these nouns has the innovative ending *-āha* (cf. Schmitt 2014: 214; Kuiper 1978: 7).

Kushan inscriptions, cf. (22) αβο ι πιδα, where a gen.sg. form would be expected, but the form πιδα can only go back to OIr. nom.sg. **pitā*.[22] This phrase is in apposition with (64) αβο οοημο καδφισε, which shows that the noun governed by αβο was inflected in the genitive case (cf. §5.2).[23] This means that the nom.sg. form πιδα was being generalised throughout the whole singular paradigm. Interestingly, a trace of an old form with *r* is still visible in a Kushan inscription, cf. Nplate, 1: αζο μο νοκονζικο ι πιορο μαρηγο 'I, Nukunzik, (his) father's servant', where πιορο continues the OIr. form **piθrah* (gen.sg.).[24]

A similar phenomenon seems to have happened with the old *n*-stems. In the *BD*, many old neuter nouns of this stem class (specifically *man*-stems) appear to continue the nom.-acc.sg. form, e.g. ναμο 'name' < **nāma* (OIr. **nāman*-), λαμο 'land' < OIr. **dāma* (OIr. **dāman*-), ταχμο 'stream' < OIr. **taxma* (OIr. **taxman*-), instead of *ναμανο, *λαμανο, *ταχμανο. In (23) ασο ια νιβαλμο 'from the seat', we find yet another originally neuter *n*-stem noun, νιβαλμο (OIr. **nišadman*-) preceded by the article ια, which is normally found with feminine nouns. This suggests that, once the nom.-acc.sg. form OIr. **nišadma* was generalised, this old neuter noun was reinterpreted as belonging to the *ā*-stem class.[25] What favoured the reanalysis as an *ā*-stem noun was probably the nom.-acc.sg. ending *-a* (< OIr. **-a* < PIE **-ḗ*) becoming formally identical to the nom.sg. ending *-a* of the *ā*-stem nouns (< OIr. **-ā*) at some stage in the evolution of Bactrian phonological system.

A shift from the neuter to the feminine nouns seems to have occurred also in the case of (24) ρωσο 'day'. This word is attested in one of the earliest *BD* documents (*BD* I, **Aa**, 356 CE) preceded by the article ια in the opening dating formula: [χ]ϸονο ρ΄λ΄δ΄ δηματριγανο μαυο οδο ια ρωσο δινο 'the year 134, the month Dematrigan and the day Din'. As remarked by Sims-Williams & De Blois (2005: 187), this noun, originally a

[22] Similarly, Parthian mostly generalised the forms without *r*, whereas Middle Persian seem to use the forms with and without *r* according to their syntactic function, cf. Sims-Williams 1981a: 167–171. Further details in Jügel 2015: 183–187. Forms with *r* are attested in Bactrian in the plural, cf. βραδαρανο next to βραδανο (cf. *BD* II: 204–205).

[23] Both πιδα and οοημο καδφισε are direct objects governed by the verb κιρδι 'to do'. They are preceded by αβο as is the rule for animate definite objects in Bactrian. About Differential Object Marking (DOM) in Bactrian cf. Sims-Williams 2011b.

[24] In the same inscription, the genitive form also occurs as the first element of the compound πιοριϸτειγανο 'belonging to the father', which is found in later documents with the normalised form πιδοριϸτο 'inheritance, ancestral estate' beside the direct continuation πιριϸτο.

[25] Cf. Sims-Williams & De Blois 2005: 187. Kreidl (2024: 220) carefully considers the possibility of considering ια in the phrase (23) ασο ια νιβαλμο as the abl.-instr. of the article (OIr. **yā*). However, such explanation could not work in the case of (24) ια ρωσο. On the abl.-instr. case see further in this paragraph.

neuter *h*-stem (OIr. **raṷčah-*), was transferred to the feminine nouns after the loss of the neuter gender. This phenomenon can be explained by looking at the inflexion of neuter *h*-stem nouns: the ending of the nom.sg. of these neuter nouns was formally identical to the nom.sg.m. ending of *a*-stem nouns (OIr. **-ah*, cf. Av. *aspō* and *raočō*), whereas the nom.pl.n. ending coincided with that of the *ā*-stems (OIr. **-āh*, cf. Av. *daēnå̊* and *raočå̊*).[26] This situation set the premises for neuter nouns like ρωσο to be reanalysed either as *a*-stem nouns or as *ā*-stem nouns. In the case of (24) ια ρωσο it seems that the noun was shifted to the *ā*-stems and reinterpreted as a feminine.[27]

We can make a few educated guesses about the phonological form of this noun in Bactrian. As we shall see (§5.4), OIr. **-ah* resulted in Bactrian *-i*: the nom.sg. may then be reconstructed as Bactr. *ρωσι < OIr. **raṷčah*. As regards the plural, the final sequence **-āh* probably developed to Bactr. *-e*. This is suggested by the fact that the Bactr. ending -ε (nom.pl.m.) originated from OIr. **-āhah* through **-ā(h)ah* > **-āh* (cf. §5.3); hence OIr. **raṷčāh* probably resulted in Bactr. *ρωσε. The same outcome Bactr. *-e*, of course, may be posited for the homophonous nom.pl.f. ending of the *ā*-stems: IIr. **-āHas* > OIr. **-āh* > Bactr. *-e*. Note that a very similar situation is found in the Khotanese thematic declension, where the confluence of the neuter nouns in the feminine and masculine inflexions is much better documented and has been comprehensively described by A. Del Tomba (2022: 116–127).

One rare case of a direct plural ending in -α is attested in (25) αβισῳι ι̣ π̣αρ̣ηνᾳ ‘many rites’. According to Sims-Williams (1998: 88), π̣αρ̣ηνᾳ may be traced back to a neuter noun OIr. **pari-ayana-* (cf. OP *parīyanam* ‘conduct’) or a feminine noun OIr. **pari-ayanā-* (cf. Av. *ayanā-* ‘way’). Depending on the etymology, the ending -α in (25) π̣αρ̣ηνᾳ could then represent the direct plural form of the *ā*-stem nouns (< OIr. **-āh*) or of the neuter *a*-stem nouns (< OIr. **-ā*). While there is evidence for OIr. **-ā* resulting in Bactr. *-a*, as we have seen throughout this paragraph, the same does not hold true for OIr. **-āh*. As was just mentioned, the development of the nom.pl.m. ending OIr. **-āhah* to Bactr. -ε probably involved a stage **-āh*. Therefore, it may be surmised that the Bactr. nom.pl. ending of the *ā*-stems was -ε, and consequently that the ending -α in π̣αρ̣ηνᾳ̣ represents the regular thematic neuter plural ending Bactr. *-a* < OIr. **-ā*.

---

[26] The nom.pl. ending **-āh* of the *ā*-stems goes back to IIr. **-āHas* (PIE **-eh₂-es*), whereas the nom.-acc.pl. ending of the neuter *h*-stems originates from IIr. **-ās* (PIE **-ōs*), cf. Hoffmann & Forssman (1996: 122, 155); Martínez & de Vaan (2014: 52, 58).

[27] It is likely that, before being reinterpreted as a feminine noun, the inflexion of this and similar old neuter nouns alternated between the *a*-stems and the *ā*-stems for a long time. A similar phenomenon can be observed in Romance languages, where neuter nouns like Lat. *digitum* (sg.) : *digita* (pl.) are continued by It. *dito* (sg.) : *dita* (pl.), respectively masculine (‘il dito’) and feminine (‘le dita’). Cf. Del Tomba 2021 on a similar treatment of old neuter nouns in Khotanese.

In (26)–(33), final -α is attested with singular nouns and adjectives that cannot be traced back to old *ā*-stems, but for the most part belonged to the *a*-stem class. Sims-Williams (2015: 258) proposed to interpret them as instances of a syncretic case resulting from the merger of the OIr. ablative, instrumental and dative endings of the *a*-stem class, respectively *-*āt*, *-*ā* and *-*āi̯*. This entails a revision of the two-case system (direct and oblique) reconstructed by Henning's 1960 article into a three-case system comprising an additional oblique case. The existence of a "third case" has been supported by two recent papers (Kreidl 2024; Halfmann *et al*. 2024). Both refer to it as "ablative-instrumental", leaving the dative out of the picture.[28]

The idea of a syncretic case combining the functions of the old instrumental, ablative and perhaps also dative cases is in line with the attested usages of -α with *a*-stem nouns. In (26)–(29), the ending -α occurs with nouns or adjectives governed by prepositions that typically trigger the ablative or the instrumental case, such as ασο 'from' and πιδο 'in' used in locational sense (cf. Kreidl 2024: 219–222; Halfmann *et al*. 2024: 27). In one case, (26) ασο μο υνδα 'from India', the final -α is attached to a noun originally belonging to the *u*-stem class (cf. OP *hindu*-, YAv. *hiṇdu*-).[29] As pointed out by Kreidl (2024: 219), however, this can hardly be the outcome of OIr. *-*aų̯š*, -*aų̯t* or *-*wā*, -*ū*, respectively the abl.sg. and the instr.sg. endings of the *u*-stems. More likely, *u*-stem nouns had by that time merged with the *a*-stems, and (26) υνδα is simply inflected according to the *a*-stems.

In (30) αβεινα and (31) οβεινα, the case marker -α is present on an element -εινα affixed to the prepositions αβο and οβ° 'in'.[30] Initially, -εινα was compared to MP *ēn*, *īn* 'this here', Sogd. *ʾyny(y)* 'this' (further to Skt. *ena* 'this, that') and interpreted as a demonstrative pronoun from a reconstructed stem **a̯ina*- alongside ειο, ειμο, ειδο and οο.[31] However, the attestation of this pronominal stem appears to be limited to Kushan Bactrian, as it is never used in the *BD* or in later inscriptions. Following a suggestion by

[28] In their brief discussion of the topic, Halfmann *et al*. (2024: 28) do not identify any use of -α compatible with the functions of a dative (see further for a different take), while Kreidl (2024: 212) observes that a possible outcome of the OIr. dat.sg. ending *-*āi̯* in Bactrian is -*ē̆*, which would be homophonous to the gen.sg. ending -ι/-ε (< OIr. *-*ahya*). If this was indeed the case, it may have led to syncretism between the dative and the genitive, as it probably happened in Sogdian and Khotanese (but not in Old Persian, where the syncretism is not a consequence of phonological developments, cf. Benvenuto & Pompeo 2012). However, a development OIr. *-*āi̯* > *-*ā(i̯)* > Bactr. -*ā* is equally possible, and, as we shall see, one of the uses of -α may provide an indication that it also inherited the functions of the dative.

[29] On (28) χϸovo cf. the discussion in the Appendix.

[30] Bactr. οβ° is only attested in Kushan Bactrian and likely goes back to OIr. **upa*. It later merged with αβο, cf. Sims-Williams 2008: 63; Benvenuto & Bichlmeier 2022: 93.

[31] Davary 1982: 256; Mayrhofer 1992: 268; Rastorgueva & Èdel'man 2000: 122–123.

Sims-Williams, Kreidl (2024: 223) argued that -εινα may be a remnant of OIr. **ayanā*, the instr.sg. of the demonstrative pronoun ειο 'this' < OIr. **ayam*. Bactrian would not be isolated in preserving a trace of the instrumental case in the pronominal inflexion: the so-called compound articles of Sogdian *c'wn*, *cnn* 'from the' and *δ'wn*, *δnn* 'with the' derive from the univerbation of the prepositions OIr. **hačā* and **hada* with the instr.sg. forms (OIr. **awanā*, *imanā*) of the demonstrative pronouns (cf. Gershevitch 1954: 232, 234). The instr.sg. form of the pronominal inflexion is also the source for the Khotanese abl.-instr.sg. ending *-äna* of the *a*-stems (cf. Sims-Williams 1990: 277–278). The derivation of -εινα from OIr. **ayanā* offers a good explanation for (30) αβεινα οιαγο 'in this place' where οιαγο, a masculine *a*-stem noun from OIr. **wiyāka-* could hardly be reconciled with the feminine ending -α of a supposed pronominal stem OIr. **aiṇa-*.

Another interesting consideration can be made about (31) οβεινα ζαμ̣ν̣α 'at this time'. The word ζαμ̣ν̣α, cognate to MP *zamān* 'time', Parth. *žamān* 'id.' and Sogd. *žamn-* 'hour' (MSogd. *jmn-*, SSogd. *zmnw*), is attested in the *BD* with the spelling ζαμανο 'time'. As briefly remarked by Sims-Williams (*BD* II: 41, 210–211) and discussed in more detail by Panaino (2017: 167–171), the spellings ζαμανο/ζαμνα seem to preserve a trace of the original ablauting *n*-stem **ǰamān-/ǰaman-/ǰamn-*, among which ζαμ̣ν̣α would be the direct continuation of the weak forms with zero grade of the suffix. The same ablaut variant -ζαμν- appears fossilised in later documents in the compound κιζαμνιδο 'as soon as' < OIr. **kahya ǰamnah* + -δο (*BD* II: 222). The phrase οβεινα ζαμ̣ν̣α can thus be traced back to **up-ayanā ǰamnā(t)*, and represents an interesting relic that further confirms the survival of the ablative-instrumental case in early Bactrian.

Yet another use of the ablative-instrumental case -α is attested in (32) and (33), this time with old *a*-stem nouns that are not preceded by any preposition (cf. Sims-Williams 2015: 258; Halfmann *et al*. 2024: 27–28; Kreidl 2024: 213). In these two examples, final -α marks nouns encoding the agent in transitive past tense clauses. As other Middle Iranian languages, Bactrian shows nominative-accusative alignment in the present and ergative-absolutive alignment in past tense clauses (so-called split-ergativity).[32] In the ergative construction, the transitive verb agrees with the logical object, which is expressed by the direct case, while the logical subject is expressed by an oblique case. In the transitive past tense clauses attested in the early Bactrian documents, the logical subject (agent) is normally expressed by the genitive ending -ι/-ε (cf. (48)–(57)), except in (32) and (33) where we find -α. The clearest example is in the Ayrtam inscription, whose final line, according to a common Bactrian inscriptional practice, consists of a colophon recording the name of the scribe who was entrusted with the task of writing the text: (32) οτι ειμο μιιροζαδα νιβιχτο 'and Mihrzād wrote this'. Here, the PN μιιροζαδα can only be the logical subject of the verb νιβιχτο (3sg. preterite of ναβισ- 'to

[32] On ergativity in Bactrian cf. Gholami 2009 and Jügel 2015.

write'). In (33) τασαγωνδι βαονανο βαο ι βαγεποορα κ̣ανηβ̣κ̣ε φρ̣ο̣μαδο κιρδι 'then, as the king of kings, the son of the gods Kanishka had given orders to do', the compound βαγεποορα 'son of the gods' occurs in a similar context, i.e. as the logical subject of φρομαδο 'ordered' (3sg. preterite of φρομιι- 'to order'). Note that in (33) the agent is expressed by a phrase in which the epithet βαγεποορα is followed by the name of the king κ̣ανηβ̣κ̣ε inflected in the genitive (cf. (51)), both expressing the agent of the ergative construction.

According to Kreidl (2024: 213), the alternation between the ablative-instrumental and the genitive case in the Bactrian ergative construction is comparable to the Avestan use of the genitive, dative, ablative and instrumental cases with agentive function in the so-called pp-construction (past participle + copula), the precursor of the ergative construction. Indeed, the agent of the pp-construction is expressed in Avestan by various cases with slight semantic differences, although from the survey conducted by Jügel (2015: 307–321) a tendency to prefer the dative case emerged, most commonly with personal pronouns. However, this refers to a stage when this construction was not yet fully grammaticalised and there was no codified expression for the logical subject. A fully fledged ergative construction, usually referred to as the *manā kṛtam* construction, is instead already observable in Old Persian, where, as is well known, the agent (logical subject) is always encoded by the gen.-dat. case.

Since Bactrian displays a fully grammaticalised ergative construction, I would rather interpret the oscillation between -α and -ι/-ε to mark the agent as a consequence of the ongoing expansion of the functions of the genitive case. The residual state of the ablative-instrumental case is apparent not only from the fact that the attestations are numerically few, but also from the fact that it is being replaced by the genitive in every usage. As already remarked, the most common way to express the agent in the ergative construction is in fact the genitive case -ι/-ε, and, as we shall see in §5.2, the genitive case is also the most common case used after preposition, even those that originally triggered the ablative or the instrumental cases (πιδο, αβο in locational and instrumental sense and ασο). In both contexts, the domain of the genitive case is encroaching into that of the ablative-instrumental. The juxtaposition between κ̣ανηβ̣κ̣ε and βαγεποορα in (33) (= (51)) is a proof of the ongoing process of syncretism: similar phenomena are attested in Late Greek, where the dative case is being replaced by the accusative, cf. ἀνέστησεν ἐαυτῷ (dat.) καὶ Βαθθιν τὴν ἐαυτοῦ γυναῖκα (acc.) 'he dedicated (a statue) to himself and to his wife Batthis' (cf. Brixhe 1984: 96).

A few more words must be said about the consequences of this reconstruction. If the genitive case is replacing the ablative-instrumental at the time of the Kushan Bactrian inscriptions, it follows that at an earlier stage of the language the agent in the ergative construction was expressed either by the ablative or by the instrumental case. It would

not be unusual for the instrumental case to develop an agentive meaning, as this is exactly what happened in the Indo-Aryan ergative construction.[33] However, a use of the instrumental case in the ergative construction would be at odds with the widespread idea that the genitive case had gained agentive functions with the past participles already in Old Iranian.[34] This view is largely based on the fact that the genitive case has agentive function in the Old Persian *manā kr̥tam* construction, but it overlooks the fact that this case that had fallen together with the dative case (cf. Jügel 2015: 321, fn. 756). Comparing the Old Persian situation with the Avestan data, Jügel (2015: 321, 344–345) argued that the dative case may have had a crucial role in the development of ergativity in Iranian, and pointed rather to this case as the most likely one to have been codified as the case of the agent when the process of grammaticalisation of the ergative construction was completed (cf. fn. 34).

This suggests a possible scenario, though admittedly hypothetical, for the development of the Bactrian case system. Assuming that -α represents the Bactrian outcome of the dat.sg. ending OIr. *-*āi̯*, as well as that of the instr.sg. *-*ā* and the abl.sg. *-*āt*, we may postulate a syncretic case combining not only the functions of the Old Iranian ablative and instrumental cases but also those of the dative case. We have already seen how long final OIr. *-*ā* resulted in Bactr. -*a*; it is likely that both *-*āi̯* and *-*āt* lost their final consonantal element and merged with Bactr. -*a*. Instead of falling together with the genitive case as it did in Old Persian and in the other Middle Iranian languages, the Bactrian dative case could have merged with the ablative and the instrumental cases, due to their endings becoming formally identical. If this was the case, we may interpret (32) μιροζαδα and (33) βαγεποορα as two remnants of a syncretic ablative-

[33] As already remarked by Kreidl 2024: 213 and Halfmann *et al*. 2024: 28. Cf. also Jamison 1979, who maintains that the instrumental was the ordinary case of the agent already in PIE.

[34] Cf. Hettrich 1990: 95–97; Bichlmeier 2011: 329 (with literature). The problem is linked to the much-debated Benveniste's 1952 interpretation of the *manā kr̥tam* construction as a possessive construction (rather than a passive one), which implies that the agentive genitive was originally a possessive genitive. Among the critics of this view cf. especially Cardona 1970; Skjærvø 1985; Ciancaglini 1987. More recently, Haig (2008: 55–81) resumed Benveniste's idea, arguing that the genitive case gained agentive function in OP by extension from the possessive construction, but described the functions of the genitive as encompassing also the roles of the benefactive, which is prototypically encoded by the dative. For a precise critique of his argument cf. Pompeo & Benvenuto 2011: 86–92.

Jügel 2015 gave the question a thorough revision, integrating for the first time the Avestan data into the discussion. According to his reconstruction, the ergative construction originated from the integration of the pp-construction into the verbal system as an analytical perfect with resultative meaning, which was followed by the reinterpretation as the case of the agent of the *dativus commodi*, originally used in this construction to express the benefactive (see especially Jügel 2015: 61–68; 444–445). Jügel (2015: 345) also observed that the use of the dative case to express the subject in the possessive construction may have solidified its position as the case of the agent also in the ergative construction.

instrumental-dative case, which was used, among other functions, to mark the agent of ergative constructions in accordance with the inherited functions of the dative. This reconstruction is of course bound to remain speculative, prompted as it is by the problems posed by a very limited set of attestations, However, if backed by more data in the future, it may show that in Bactrian the Old Iranian case system followed a different path of evolution than other Middle Iranian languages.

5.2. Final -ι/-ε

Final -ι/-ε occurs with nouns originally belonging to the *a*-stem class with two main functions: in (35)–(47) the ending -ι/-ε marks the possessor, whereas in (48)–(57) it marks the agent in ergative constructions. The derivation from OIr. *-*ahya* lies at hand, and it is in fact generally accepted since Henning's 1960 article. The alternative spellings -ι/-ε point to a short front vowel, possibly a close-mid vowel [e] (cf. §6). This may have developed from *-*ahya* through *-*a*(*h*)*ya* > *-*ē*,[35] and finally -*e* (with shortening due to the final position).

The genitive case in -ι/-ε is also attested in prepositional phrases with αβο 'to, in', πιδο 'in, according to' and ασο 'from', cf. (58)–(74). As already mentioned, our documentation shows an ongoing process of extensions of the functions of the genitive at the expense of the residual ablative-instrumental(-dative?) case.[36] Moreover, if we accept the traditional reading in (73) πιδεινι 'by means of' (cf. the discussion in the relative section of the Appendix), the form -εινι may be explained as a recharacterisation with the gen. ending of the pronoun -εινα (cf. (30)–(31)), which, as was discussed in §5.1, is likely to be a remnant of the instr.sg. form of the demonstrative pronoun ειο. A similar process is attested by the 'irregular' genitives in Sogdian *wyspny*, *ʾmyn*, *ʾwyn* formed from the abl.sg. -*na* + the gen.sg. ending -*i* (cf. Sims-Williams 1990: 278).

In (101) and (102), the ending -ι is attested with the article / pronoun OIr. **(i)ma-* 'this': [α]βο μι βαγεᾳ[βο] 'in Bage-ab', μι βαγολαγγο 'in the temple'. According to the most recent interpretation by Sims-Williams (forthcoming), μι goes back to the gen.sg.m. form OIr. **imahya*. Given the shared origin with the gen.sg. of the *a*-stem nouns from OIr. *-*ahya*, it is likely that this ending could also be spelled with -ε.

The same can be said about another use of -ι with the same pronominal stem. In (100) ειμιδβα βαγε 'may these gods', the demonstrative pronoun ειμι (here with its full form) can only have the function of a nom.pl., since it is in the same noun phrase as the nom.pl.

---

[35] Cf. the development OIr. *-*aya-* > Bactr. -η- (-*ē*-) in internal position in αβαχρηγο 'fee' < OIr. **apa-xraya-ka-* (cf. *BD* II: 182) or in the 2pl. ending -ηδο < OIr. *-*aya-ta-*, cf. MP, Parth. -*ēd*.

[36] We can keep track of this process by looking at the relative chronology of the inscriptions, compare (26) ασο μο υνδα 'from India' (2nd cent. CE) with (68) ᾳṣο ι ροζγι 'from the vineyard' (3rd cent. CE).

βαγε, and can thus be traced back to the nom.pl. form of the pronoun OIr. **imaḭ* (Cf. Sims-Williams & Cribb 1995–1996: 91). The proclitic form of ειμι is attested in (99) ασο οισποανο μι βαγανο 'from all the gods'.[37] The development of OIr. **-aḭ* was probably similar to that of OIr. **-ahya*: first it was monophthongised to **-ē*, then it was shortened to Bactr. *-e* (see also §6). Therefore, one can reasonably assume that this ending may have the same variant spelling -ε as the genitive ending (cf. also §5.4 on the development OIr. **-tanaḭ* > -δανι/-δανε).

5.3. Final -ε

In (78)–(82), final -ε marks the direct plural ending of old *a*-stem nouns such as βαγε 'gods'. As observed by Del Tomba (2022: 124, fn. 16), the Bactrian ending -ε can be compared with the Khotanese nom.-acc.pl. ending *-e* that is characteristic of a set of masculine nouns of the *a*-stems. This ending can be traced back to OIr. **-āhah*, an innovative nom.pl. ending attested in Old Iranian languages (cf. Av. *-å̊ŋhō*, OP *-āha*) as well as Old Indian (cf. Ved. *-āsaḥ*). Interestingly, the distribution of the endings *-āsaḥ* in Vedic and *-e* in Khotanese seems to indicate that the secondary ending was used mainly in nouns that have a high degree of agentivity (cf. Del Tomba 2022: 126–127). This is consistent with the Bactrian data, though few in number, since the ending -ε is attested with the animate nouns βαγε 'gods' and ρηδγε 'attendants'.

For Khotanese, a development OIr. **-āhah* > **-ā(h)ah* > **-ā.ah* > **-āh* > *-e* can be confidently reconstructed (cf. Del Tomba 2022: 125). In Bactrian, a similar deletion of intervocalic **-h-* in word-final position is attested in verbal endings such as **-ahi* > -ηιο and in the enclitic copula **ahi* > -ηιο (both 2sg. pres. indicative). Therefore, it is safe to assume that the first glottal fricative of OIr. **-āhah* was deleted in Bactrian as well, resulting in a sequence **-āah* > **-āh*. As for the raising of the central vowel in front of the second glottal fricative, one can compare the outcome Bactr. *-i* < OIr. **-ah* (§5.4). Accordingly, the development of OIr. **-āhah* may be reconstructed as follows: OIr. **-ā(h)ah* > **-ā.ah* > **-āh* > **-ē* > Bactr. *-e*.

5.4. Final -ι

Lastly, a somewhat forgotten function of the ending -ι is that of marking the nominative singular case, as illustrated by examples (83)–(98). This ending was first noticed by

[37] Here one could reasonably have expected the obl.pl. form ειμοανο, since the phrase is governed by the preposition ασο, which in fact triggers the obl. forms οισποανο and βαγανο. The same form (with application of the final -ο rule) is however attested in the same context in SK4M, 9 φαρο οισποανο μο (not **μοανο) οαδοβαργανο 'towards all living creatures'. This may indicate that the article generalised the ending of the nominative case in the plural. Alternatively, this may be a case of group inflexion.

Sims-Williams (1985: 113, fn. 26) in three forms attested in the version B of the Surkh Kotal inscription, namely (84) πιδοριγδι, (85) οζοοαστι, and (86) κιρδι, which he analysed as 3sg. preterites (originally past participles)[38] inflected in the nom.sg. case. Accordingly, he traced back the ending -ι to the nom.sg. ending OIr. *-*ah* of the *a*-stems.[39] However, his observation seems to have gone unnoticed, as the ending -ι is never included in the descriptions of the Kushan Bactrian nominal system.[40]

The development OIr. *-*ah* > Bactr. -ι has clear Middle Iranian parallels, cf. the Sogdian ending -*y* of the nom.sg.m. of the light stems and the Khotanese nom.sg.m. ending -*ä*, -*i* of the *a*-stems.[41] The presence of a close front vowel at the end of the *a*-stem nouns is also indirectly confirmed by an early Bactrian loanword in Nuristani languages recently found by Halfmann 2023. In Katë and Ashkun, two languages of this group located between modern-day Afghanistan and Pakistan, the word *ladír* 'mediator' (*ladér* in northeastern dialects of Katë) can be connected to Bactr. λαδοβαρο 'judge' through a stage **lad(i)vari*. The final -*i*, which is necessary to justify the *i*-umlaut in **vari* > °*vír*~*vér* and is explained by Halfmann (2023: 4, fn. 7) as an adaptation of -*ə*, has an even more straightforward explanation assuming that the borrowed form was the nominative form in -*i* (*λαδοβαρι).[42]

Remarkably, this ending can be detected also in the earliest *BD* documents, cf. (87)–(95), including the recently published 4th-century letters from south of the Hindukush,

[38] In Bactrian, the preterite is formed by suffixing the copula to the past participle in the first two persons. In the 3sg., however, the preterite form coincides with the pp. Due to the split-ergativity (on which see §5.1), with transitive verbs the past participle agrees with the logical object (grammatical subject).

[39] The nom.sg. ending -*i* was noticed also by Harmatta (1969: 351), who postulated the existence of two Bactrian dialects, one characterised by the development of the nom.sg. ending OIr. *-*ah* to /-i/, the other to /-o/ (which he considered to be the pronunciation of final -o, as discussed in §2). This idea found no fortune in the subsequent literature, cf. already Schwartz 1974: 410. In general, apart from a few words with specific phonological features that may be of dialectal origin (for example **w*- > *γu*-, on which cf. Sims-Williams & de la Vaissière 2015: 46–47; Sims-Williams 2023a: 136; Schoubben 2023; *BD* IV: 4), it is hard to find evidence for a dialectal split within the Bactrian corpus. This is in part due to the fact that most of our documentation comes from Northern Bactria, more specifically from the area around the city of Rōb (Sims-Williams 1997). But even the inscriptions and documents originating from south of the Hindukush seem to adopt the same standard language, apart from a few orthographic features (cf. Sims-Williams 2023a: 136; *BD* IV: 5–6). For the possibility of a Bactrian dialect being represented by the so-called Unknown Kushan script see now Lurje 2024 (English translation in this volume).

[40] Cf. Davary 1982: 140; Lazard *et al*. 1984: 226; *BD* II: 40; Gholami 2014: 73.

[41] Cf. Gershevitch 1954: 128, 177; Emmerick 1968: 251–253.

[42] I am grateful to Jakob Halfmann for pointing out this interesting connection to me and for providing me with a draft of his article before it was published.

again mainly (but not exclusively) with participles.[43] The case of (89) αγγιτιδο (< αγγιτι + the usual clause-initial particle -δο) is particularly significant because it shows the ending -ι- preserved in internal position. Together with the gen.sg. in (47) and (69)–(71), these are the only cases of inflexional endings attested in manuscript documents, thus later than the Kushan times. These endings seem to occur only in formulaic expressions such as titles, cf. (47) ι ραβτι βα(ο)βορο βιγανο, (69) τι χοηο, polite clichés, cf. (70)–(71) πιδο ... βιζγι, and dating formulas, cf. (97) σαχτι, and with technical verbs of the administrative and scribal jargon, cf. (88), (91), (95) ναβιχτι 'written', (92)–(94) ναγατι 'heard' (in the common phrase 'I have heard…'), (87) αγγιτι 'received' (said of a payment). Since it is hard to maintain that a nom.sg. ending still survived at the time of these documents (4th cent. CE), these occurrences should probably be regarded as historical spellings, whose old orthography was kept due to their frequent usage in economic writings and letters. Similarly, the other few cases of final -ι and -α in the *BD* are found with high-frequency words (οτι, ιθα, αστι, cf. §4), which are suitable for the creation of traditional spellings.

Furthermore, it is possible that the final -ι in nouns and pronouns with stems ending in *-*w* found in documents until the 7th cent. CE (e.g. **aj**1 γαοι 'cow' for later **U**8 γαο, cf. *BD* II: 40) ultimately goes back to a crystallised form of the word inflected in the nom.sg. (or perhaps even the gen.sg.).[44] The digraph -οι is probably a scribal convention designed to avoid the sequence of two *omicra*.[45]

Finally, a group of nouns that have the syntactic functions of a nom.sg. shows an alternating final vowel -ι/-ε, cf. (75) κιδι … βαοδανε λφαχτο 'who gained the kingship';

[43] It is possible that the rare forms in -αγι analysed by Sims-Williams as future participles contained the same ending (*BD* II: 44). Such a formation, whose origin is unclear, is attested only in document **xp** (and perhaps in **kh** with the spelling -ογο, cf. *BD* IV: 7), where it is employed to express promises or obligations. Formally, it resembles a present participle formed by adding the suffix *-*aka*- to the present stem. However, one would expect the modal meaning to be expressed by an auxiliary verb, cf. Sogd. *β't ptcγšy* 'she should receive' (cf. Gershevitch 1954: 135). N. Sims-Williams (personal communication) sees a connection with the Sogdian 'gerund' in -*kya*/-*kī* (cf. Yoshida 2009: 297). Personally, I prefer a rapprochement with the Khotanese so-called participle of necessity in *-*ya*-*ka*- (e.g. *hvañaa*- 'to be said', cf. Emmerick 1968: 218), as suggested to me by A. Del Tomba. We wish to return to this question in the future.

[44] That -οι is a conventional spelling is further confirmed by the fact that, as any good historical spelling that does not reflect the current pronunciation of the word, it can be overextended by hypercorrection. See for example the 2sg. pronoun το(ο) 'you', etymologically from OIr. **tuwam*, often written τοι or τοοι, with an unetymological final -ι.

[45] Interestingly, as pointed out to me by Sims-Williams (personal communication), final -ι in manuscript documents is attested only after three consonants: -γ-, -τ- and -ο-. There is probably a graphic reason behind this distribution, although, while the use of the digraph -οι may be motivated by the tendency to avoid the sequence -οο, it is hard to find an explanation for -γι and -τι.

(76) [νανα] … αβο βαονανο βαο κανηβκι κοβανο ι χοαδηοδανι λαδο '[Nana] gave the lordship to the King of Kings Kanishka the Kushan'; (77) κιδι ασǫ νανα … ι βαοδανι αβορδο 'who has obtained the kingship from Nana'. One may wonder whether -ε is an alternative spelling for the -ι of the nom.sg. All instances of the alternation -ι/-ε with the nom.sg. occur in nouns ending with a suffix attested in later documents as -δανο, which forms abstract nouns and adjectives (*BD* II: 208). The etymology of this suffix is uncertain, but a comparison with the OP suffix *-tanaį̯* of the infinitive may be suggested.[46] If this suffix formed the present infinitive in an older stage of the Bactrian language, it may have been refunctionalised as an abstract suffix when the past infinitive in -ι (cf. κιρδι < OIr. *kr̥tayaį̯*) was selected as the only form of the infinitive. This etymology is also supported by the cognate Sogdian suffix *-tʾny(h)*, which also forms abstract nouns (cf. Gershevitch 1954: 164). If this reconstruction is correct, the final diphthong **-aį̯* of the suffix **-tanaį̯* is responsible for the alternating spellings -ι/-ε (cf. the development described in §5.2) and there is no need to assume a variant spelling -ε for the nom.sg., otherwise consistently attested as -ι.

6. Notes on the phonetic values

Summing up, the final vowels alternative to -ο can be traced back to various OIr. origins:

-α	OIr. **-ā*, **-āt*, (*-āį̯*?)	generalised sg. ending of the of the *ā*-stems, abl.-instr.(-dat.?) ending of the *a*-stems, adverbs
-ι	OIr. **-ah*	nom.sg. ending of the *a*-stems
-ε	OIr. **-āhah*	nom.pl. ending of the *a*-stems
-ι/-ε	OIr. **-i*, **-ahya*, **-aį̯*, **-ayaį̯*	particles, verbs, gen.sg. ending of the *a*-stems, pronominal nom.pl. ending, infinitive ending

Since the nom.sg. is consistently written with -ι, it can be surmised that it was pronounced as a close front vowel [i]. Although the endings nom.sg. -ι and nom.pl. -ε are graphically identical to the gen.sg. ending -ι/-ε in writing, they may have still been kept separate from each other in speech. Indeed, only the signs <ε> and <ι> were available in the Greek alphabet to represent short front unrounded vowels. The fact that Bactrian scribes alternated -ι and -ε to represent the gen.sg. ending of the *a*-stems may indicate that the pronunciation of this ending was similar to that of both -ι of the nom.sg. and -ε of the nom.pl. At the same time, the alternation itself shows it was identical to neither. In other words, the spelling alternation -ι/-ε suggests that there was a phonological contrast between the gen.sg., nom.sg. and nom.pl. endings. If, as proposed

[46] The same etymology is suggested in J. Kreidl's *Baktrisches Wörterbuch*, still unpublished at the time of writing. I would like to thank the author for kindly sharing a draft of this valuable dictionary.

in §5.2, the alternation between -ι/-ε (typical of the gen.sg. of the *a*-stems, but present also in the infinitive ending -δανι/-δανε) reflects the pronunciation of a close-mid front vowel [e], it is possible that the endings nom.sg. -ι and nom.pl. -ε were pronounced as front vowels of different heights. While -ι most likely represented the high front vowel [i], it may be surmised that -ε was more similar to an open-mid front vowel [ɛ] (cf. §5.3).

The use of the short letter -ε for endings that were originally disyllabic (like *-*ahya* or *-*āhah*) indicates that final vowels were shortened in final position after developing into a long mid vowel. As mentioned above (§5.1), assuming that the shortening of long final vowels was systematic, final -α is likely to represent short [a] rather than [aː].

The fate of the short OIr. vowels in final position is harder to determine with the current data. OIr. *-*i* was presumably pronounced as Bactr. [e], since it appears to be continued by Bactr. -ι/-ε.

7. The Kushan Bactrian nominal inflexion

Overall, the distribution of final vowels shows that Kushan Bactrian retained substantial traces of the Old Iranian nominal system. At least two genders, two numbers and three cases are functionally distinguished. Feminine nouns are characterised by the ending -α, from the OIr. nom.sg. ending *-*ā* of the *ā*-stems, which seems to have been generalised to the whole singular inflexion. The old *a*-stems are continued by a greater variety of endings. In the singular, there are traces of a nom. case -ι < OIr. *-*ah*, a gen. case -ι/-ε < OIr. *-*ahya* and a syncretic abl.-instr.(-dat.?) case -α < OIr. *-*ā*, *āt*, (-*āi̯*?). In the plural, only two endings are distinguished: a dir.pl. ending -ε < OIr. *-*āhah* and an obl.pl. ending -ανο < OIr. *-*ānam*.[47] If (25) παρηνᾳ can be interpreted as a neuter noun, a trace of the nom.pl.n. ending -*a* may also be retained. An expansion of the domain of the genitive can be observed in the singular, caused by the large functional overlap between the genitive and the abl.-instr.(-dat.?) cases. In the plural inflexion, the syncretism between the oblique cases is already complete, as the only surviving oblique ending is the old gen.pl. ending -ανο. The two-case system in the plural inflexion may have enhanced the coalescence of the oblique cases also in the singular.

The Kushan inscriptions (and some of the early *BD*) also preserve traces of other OIr. stem classes, which appear to have been restructured into the *a*- and *ā*-stems. OIr. *u*-stem nouns were likely transferred to the *a*-stems, as is suggested by the case of (26) υνδα (OIr. **hindu*-) which is inflected according to the *a*-stems. Moreover, some neuter nouns seem to have been shifted to the feminine inflexion due to the fact that they shared the nom.pl. ending with the *ā*-stems. It is the case of the *n*-stem noun (23) νιβαλμο and the *h*-stem noun (24) ρωσο, both attested with the feminine article ια. Lastly, the *r*-stem

[47] For this reconstruction of the gen.pl. ending in Eastern Iranian languages cf. Peyrot 2018.

nouns denoting kinship also generalised the nom.sg. form in *-*ā*, cf. (21) and (22). From a comparative point of view, the nominal inflexion preserved in Kushan Bactrian is remarkably close to that of Sogdian and Khotanese, the two other Middle Iranian languages that still have a vital nominal system, as is illustrated by Table 1 (cf. Yoshida 2009: 288; Emmerick 1968: 252):

	OIr.	Kushan Bactrian	Sogd. light stems	Khot. *a*-stems
Masculine				
nom.sg.	*-*ah*	-ι -*i*	<-y> -*i*	-*ä*, -*i*
gen.sg.	*-*ahya*	-ι/-ε -*e*	<-y(y)> -*e*	-*i*
abl.-instr.(-dat?).sg.	*-*āt*, *ā*, (-*āi̯*?)	-α -*a*	<-ʾ> -*a*	(-*äna*)[48]
dir.pl.	*-*āhah*	-ε -*ε*	(<-tʾ> -*ta*)	-*e*
obl.pl.	*-*ānam*	-ανο -*ānə*	<-'nw>[49] -*ānu*	-*ānu*

	OIr.	Kushan Bactrian	Sogd. light stems	Khot. *ā*-stems
Feminine				
nom.sg.	*-*ā*	-α -*a*	<-ʾ> -*a*	-*a*

Table 1. Sogdian and Khotanese parallels for Kushan Bactrian nominal endings.

As can be observed, apart from the innovative Khotanese abl.-instr.sg. -*äna* and Sogdian dir.pl. -*tʾ* (Sims-Williams 1990: 277–278; Gershevitch 1954: 189), the Old Iranian nominal system has a very similar development in Kushan Bactrian, Sogdian and Khotanese.

The large comparability between the nominal endings of these three languages suggests that Bactrian may originally have another form in common with the singular inflexion of the Sogdian light stems and the Khotanese *a*-stems, i.e. an acc.sg. form in -*u* < OIr. *-*am* (Sogd. -*u* <-w>, Khot. -*u*).[50] From the phonological point of view, the labialising effect of the bilabial nasal -*m* on a preceding central vowel -*a*- is paralleled by cases like οδομο = ManB. *'wdwm* 'and by me' (OIr. *-*uta-maį̄*) or ManB. *nyśt qyrdwm* (OIr. *-*kṛta-maį̄*) 'not made for me', where the vowel at the end of a word is labialised by the following 1sg. enclitic pronoun -μο. The same ending OIr. *-*am* is also characteristic of the nom.-acc.sg. of the *a*-stem neuter nouns, which arguably shifted to the masculine inflexion quite early, as their singular inflexion came to be almost identical (cf. Table 2).

[48] Traces of the ending -*a* can be seen in adverbs such as OKh. *uska* 'above', besides *uskäna*, cf. Emmerick 1968: 257.

[49] This form is still attested instead of the innovative ending -*tyʾ* in the 4th-century Ancient Letters, cf. AL 5, r4 *wyšnw βγʾnw* 'to all the gods'.

[50] Note that a Bactr. acc.sg. in -*u* was suggested for comparative reasons already by Humbach (1976: 65) and, more recently, by Kreidl (2021: 160; 2024: 209, 219).

	Masculine		Neuter	
	OIr.	Bactr.	OIr.	Bactr.
nom.	*-*ah*	-*i*	*-*am*	-*u*
acc.	*-*am*	-*u*	*-*am*	-*u*
gen.	*-*ahya*	-*e*	*-*ahya*	-*e*
abl.-instr. (-dat.?)	*-*āt* *-*ā* (*-*āi̯*?)	-*a*	*-*āt* *-*ā* (*-*āi̯*?)	-*a*

Table 2. Singular inflexion of Bactrian *a*-stem nouns

Postulating an acc.sg. case in -*u* also helps to justify the shift of old *u*-stem nouns such as OIr. **hindu-* to the *a*-stems. The expected outcome -*u* of the OIr. nom.sg. and acc.sg. endings, *-*uš* and *-*um* respectively, was formally identical to the acc.sg. -*u* < OIr. *-*am* of the *a*-stems after the deletion of the final consonants (cf. Kreidl 2024: 219).[51]

8. The origin of the final -ο rule

The observation that -ο may originally have represented the acc.sg. -*u* of the *a*-stems suggests a possible solution for the question of the origin of final -ο. If Bactrian had an ending -*u*, it would have certainly been written with -ο, as this is the same vocalic value it has in internal and initial position. It may be surmised that, when Bactrian was first given a written form, the scribes would use final -ο to represent the -*u* of the acc.sg. of the *a*-stem nouns, e.g. writing καδγο to represent Bactr. *kadgu* 'house' (< acc.sg.m. OIr. **kataku*). The same obviously applied to the neuter nouns in the direct cases (þαυρο for Bactr. *ṣahru* < OIr. **xšaθram*), and the oblique plural of both *a*- and *ā*-stems (βαγανο for *βαγānu* < OIr. **bagānam*). [52] In this phase, the nominal endings were still distinguished in speech, and each ending was represented in writing.

Assuming that -*u* was the first vowel to undergo the process of centralisation that will later affect all final vowels, this may explain why the letter *omicron* gained a new value

[51] The nom.sg. and acc.sg. endings of the OIr. *u*-stems could also have full- or long-grade of the suffix (Av. -*āuš*, and -*aōm*, -*āum*) but tend to be levelled analogically to the endings of the proterodynamic inflexion (-*uš* and -*um* respectively) already in Young Avestan, since this was the most frequent type in Indo-Iranian, cf. Kümmel 2018: 1895–1896.

[52] It may be added that other endings were probably written -ο but are indistinguishable from the generalised final -ο. For example, as mentioned in §5.1, the acc.sg. of the *ā*-stem nouns (OIr. *-*ām*) may have resulted in Bactr. -*o*; the acc.sg. of the *i*-stem nouns may also have resulted in Bactr. -*u* (< OIr. *-*im*). A counterargument for the development OIr. *-*im* > -*u* may be the relative pronoun σιδο, from OIr. **čim* + -δο (*BD* II: 194). The -ι- in the stem may be analogical to the other relative pronoun κιδο, or it may also go back to the neuter nom.-acc.sg. form of the pronoun OIr. **čit*.

-ə besides *-u*, *-ū* and *-w*, thus becoming the most obvious choice to represent the end of the word once the process of centralisation of final vowels was completed. But if this is the endpoint of the process, as the later orthography of the *BD* clearly shows, how do we make sense of the Kushan Bactrian evidence?

In Kushan Bactrian, the centralisation of final vowels was evidently still an ongoing process. As the survey conducted above has shown, the distribution of -α, -ε, and -ι in the Kushan Bactrian inscriptions and early Bactrian documents is etymologically motivated and their use reflects the vitality of the nominal system at least during the Kushan period. Among the approximately one hundred nouns examined in the Appendix and the more than sixty attestations of particles, adverbs and verbs listed in §4, there is not a single case of a misplaced final vowel. It is thus hardly credible that the occurrences of -α, -ε, and -ι should be understood as free variants, and they are equally not likely to be historical spellings, since we would expect at least a few cases of hypercorrections or inverse spellings. The lack of errors suggests that the letters written in final position still corresponded to vowels actually pronounced by the Bactrian speakers, which explains why the scribes knew exactly when and where to write them and did not fail to place them in the right position. But if final -α, -ε, and -ι indeed represented vowels that were still pronounced by the speakers, how can this notion be reconciled with the fact that these vowels could be replaced, seemingly at the scribe's whim, by the variant spelling -ο representing the centralised vowel *-ə*?

Perhaps a more fruitful approach to understanding the Bactrian nominal system as reflected by the Kushan inscriptions is to look at the alternation between final vowels as the consequence of an actual situation of variation in the speech of the Bactrian linguistic community at the time. This situation is best understood by adopting a variationist approach, i.e. looking at language change through the categories developed by sociolinguistic studies. [53] In sociolinguistic research, variation within a speech community is viewed as systematic, and language change is understood as the result of the social evaluation of variants. Only the variants that are adopted by the majority of speakers produce a new standard, but before the change takes place, rival forms coexist in a speech community.[54] This process is best described by the following quote from the seminal work by Weinreich *et al*. (1968: 156):

[53] The utility, and the limits, of applying the methodology and the theoretical framework of variation studies to historical linguistics have been discussed by Lass (1976; 1997, esp. pp. 139–145) and, with special regard to ancient languages with limited corpora, by Mancini (2012, 2014). The role of the socio-cultural motivation in linguistic change was already largely outlined by Coseriu (1958, 1992).

[54] This is the phase termed "selection" by Coseriu (1992: 151), which follows the phase of "diffusion", i.e. the "intentional though intuitive" adoption of the innovation by several speakers.

> the transfer takes place when Speaker *A learns* the form or rule used by Speaker *B*, and […] the rule then coexists in *A*'s linguistic competence along with his previous form or rule. Change then takes place *within* the complex linguistic repertoire of *A*: one type is the gradual disfavoring of the original form at the expense of the new one, so that it moves to the status of "archaic" or "obsolete".

Applying this model to Kushan Bactrian, instead of assuming that the final vocalism had already progressed to the stage where *-ə* was generalised everywhere and that the final vowels -α, -ε, -ι were a mere graphic reminiscence known only to the scribes, we can consider the idea that the obsolescent trait (i.e. the distinction between final vowels) was still present in the linguistic repertoire of the Bactrian speakers. This meant that scribes could decide whether to adopt a more conservative orthography, or to lean towards the emerging new standard, i.e. the centralised pronunciation of all final vowels.

To exemplify this situation, one may look at the different orthographic standards adopted in the three versions of the Surkh Kotal inscriptions. As was already remarked, the scribe of SK4B makes use of many more etymological vowels than the two other versions.[55] Rather than representing different stages of the language, the three versions simply reflect different orthographic standards, one adopting a more conservative spelling (SK4B), the other two (SK4A, M) reflecting the emerging tendency to centralise the final vowels towards *-ə* more systematically. Note, however, that the scribes of SK4A and SK4M are perfectly capable of placing the right etymological vowel when it is needed, e.g. to disambiguate between the possessor and the head noun in SK4M, 21–23 κοζγαβκι πουρο 'the son of Kuzgashk' and νοκονζικι καραλραγγε μαρηγο 'the servant of Nukunzik the margrave'.

Interpreted in this way, it is final -ο (rather than -α, -ε, -ι) that best corresponds with the definition of "historical spelling". As discussed above (§3), a historical spelling is a conservative spelling that no longer represents the current pronunciation of the word. After the development **-u* > *-ə*, the vowel <ο> [-u] did not represent the current pronunciation of the final vowel anymore. In all likelihood, however, scribes in the pre-Kushan period kept on writing final -ο as they did before, for example in the accusative of the *a*-stems, although it was no longer pronounced as a close back rounded vowel in that position (much like <gh> in English *light* /laɪt/ is written but not pronounced). Contextually with the centralisation of **-u*, the letter *omicron* gained the additional phonological value *-ə*. Gradually, as the process of centralisation started to affect all final

[55] Cf. Benveniste 1961: 137. This is the reason why Göbl 1965 believes that the version B is chronologically earlier than the other two. Even if this was the case, all three versions were surely written during the lifetime of Nukunzik, who commissioned the inscriptions and is the real protagonist of the texts (cf. Gershevitch 1979: 55), which makes it unrealistic that there is a significant chronological discrepancy between the three versions.

vowels, *omicron* provided the most apt representation for these newly centralised vowels and started to be used in final position even in words that never had a final *-u* (e.g. πιδο < OIr. **pati*) and in the place of nominal endings different from the acc.sg. This is the moment in the historical evolution of the final vocalism of Bactrian that is captured by the Kushan inscriptions. At the time, as we have seen, scribes had the option to stick to the old orthography and represent case endings and final vowels according to their traditional (obsolescent) pronunciation, or to adhere to the new spelling standard that emerged with the innovative centralised variant.

As for the early manuscript documents (4th cent. CE), the traces of final -ι attested there in correspondence with nom.sg. (87)–(95) and gen.sg. nouns (47), (69)–(71) and with particles and verbal endings (cf. §4) are limited to high-frequency words and formulaic expressions of economic documents and letters (cf. §5.4). It is likely that these words preserved the old orthography because of their frequent usage in these documentary types. Therefore, in the *BD* final -ι was hardly still pronounced as [-i], but it rather represented a historical spelling.

9. Conclusions

In the discussion above, I have tried to address the question of the origin of the final -o rule by looking at the distribution of alternative final vowels in Kushan Bactrian and early manuscript documents. The reconstruction of Kushan Bactrian nominal system has prompted the suggestion that final -o was originally the ending of the acc.sg. *-u*, which was soon centralised to *-ə*.

If this reconstruction is correct, final -o represented Bactr. final **-u* (resulting from OIr. **-am*) in an undocumented phase of the Bactrian language and was assigned the additional value *-ə* only after the innovative centralised pronunciation of this vowel came to be adopted widely by the speech community. In this phase, final *omicron* functioned as a historical spelling, as it was not used with the phonological value it originally represented, but rather with the one it acquired after the change from *-u to -ə*. As the process of centralisation of unstressed final vowels progressed, the letter *omicron* gradually became the most common (but not unique) final vowel, as we see in the Kushan Bactrian inscriptions.

In due course, all final vowels came to be pronounced as the central vowel *-ə* and the final *omicron* was generalised at the end of the word. After the process of centralisation of final vowels was completed, these probably started to be lost altogether, and final -o ended up being refunctionalised as a graphic final sign, ultimately becoming a word-divider. Only at this point can we talk about the "final -o rule" as a proper orthographic device. The Kushan Bactrian inscriptions seem to lie in the middle of this process,

capturing a moment in which final -α, -ε, and -ι are still being pronounced, but the innovative centralised pronunciation of final vowels was gradually being accepted by the speech community.

Appendix

The Appendix collects all the passages in which the endings in -α, -ε or -ι are attested in Kushan inscriptions and early Bactrian documents. The syntactic analysis follows the Leipzig glossing rules (see Abbreviations at the end of the article). Relevant endings are highlighted in **bold**. The etymology of the nouns, when known, is given according to the Glossary at the end of *BD* II, if not specified otherwise.

The examples are ordered according to the relative chronology of the texts (cf. Kreidl 2024: 207–208). As for the texts whose dating is uncertain, I adopted the following rough chronological order: KT, Ayr, DilV, Alm, DN1, Nplate, Rab, NumKan, NumHuv, SK1, SK3, SK4, Splate. The place of KT, Ayr, DilV, SK1 and SK3 is based on archaeological data and the presence of archaic traits in the text but remains of course speculative. For the dating of the *BD* texts, cf. Sims-Williams & De Blois 2018.

The following attestations of final vowels could not be included even as 'uncertain cases' because the context in which they are attested is either too fragmentary or too problematic:

219 KT

[ε]ιḍọ [κ]οṇḍ[ι]κ[ο] [α]βọ αρο[γ]αδο αβο ζαο σαγγ**ε** γοṇ**ε**
'This water-pot–may it be for everyone, for life (?)…'.

The last two words of the inscription are unclear, cf. Fussman 2011: 107.

24 KT

διιγζ**ε** ανανδο.

Based on the Khar. and Br. inscription on the same vase, which mention the monk Jīvananda, this should be the Bactrian rendition of the name, but the spelling is puzzling. Note that there is a blank space between the two words.

Ayr, 1

]ọṛọṛ**ę** ι ḅαο[.

The whole line is very damaged.

According to the recent reading by Halfmann *et al*. (2024: 33), lines 9–13 of DN1 contain three words with a final vowel different from -ο (μολ̣**ę**, ṭορριγ**ι**, μιροα̣σ̣α̣ṇ**ę**). I have not included them in the Appendix because they all present interpretative problems, cf. Sims-Williams 2025: 189, fn. 10,11; 190, fn. 12.

As for 20 KT νοβιχτο βυικþο βοδδοζιροσσ**α** ‘written by the monk Buddhaśira’, it is clearly influenced by the (sanskritised) Gandh. inscription *ayo kuḍiya bhikṣusya budhaśirasya pugaligasya* on the same vase. The interesting ending -σσα may reproduce the Gandhari pronunciation of the gen.sg. ending *-sya* according to Fussman 2011: 65–66.

Ending -α used as a feminine nom.sg. ending (< OIr. *-ā)

(1) Rab, 9–10

ǫǫηλδι	ι**α**	αμγ**α**	ναν**α**	οδο	ι**α**
DEM=PART	ART.NOM.SG.F	same-NOM.SG.F	Nana-NOM.SG.F	and	ART.NOM.SG.F

αμγ**α**	ομμ**α**
same-NOM.SG.F	Umma-NOM.SG.F

‘that (is), the above-mentioned Nana and the above-mentioned Umma’

Bactr. Ομμα < OIr. **uxšmā-* or **us-šmā-* (cf. Sims-Williams 2017). It is reasonable to assume that the name of Nana, a feminine goddess of Mesopotamian origin, was added to the *ā*-stems (cf. Falk 2015: 265).

(2) Rab, 22

þαι	μ**α**	λιζγ**α**	αβο	βαγανο	λαδο
king-GEN.SG	ART.NOM.SG.F	fortress-NOM.SG.F	to	god-OBL.PL	give-PRET.3SG

‘the king gave the fortress to the gods’

Bactr. λιζγο < OIr. **dizakā-*, same as (4) with **-ka-* extension.

(3) NumKan, NumHuv (several examples, cf. Alram 1986: 305–309)

ναν**α**
Nana-NOM.SG.F

‘Nana’

On the name of Nana cf. (1).

(4) SK4A, 1 = SK4B, 1 = SK4M, 1

ειδο	μ**α**	λιζο	μο	κανηþκο	οανινδο	βαγολαγγο
DEM	ART.NOM.SG.F	fortress	DEM	Kanishka	victorious	sanctuary

‘this fortress (is) the sanctuary of Kanishka the victorious’

Bactr. λιζο < OIr. **dizā-*, also in (5)–(9).

(5) SK4A, 4–5 = SK4B, 6 = SK4M, 3

μ**α**	λιζ**α**	αβαβγο	σταδο
ART.NOM.SG.F	fortress-NOM.SG.F	waterless	be-PRET.3SG

‘the fortress was without water’

(6) SK4A, 8 = SK4B, 10 = SK4M, 6
μ**α** λιζο πιδοριγδο[56]
ART.NOM.SG.F fortress be.abandoned-PRET.3SG
'the fortress was abandoned'

(7) SK4A, 16 = SK4B, 16 = SK4M, 11
ταδ=ηιο μ**α** λιζο πορ(ο)γατο
then=he-OBL.SG ART.NOM.SG.F fortress survey-PRET.3SG
'then he surveyed the fortress'

(8) SK4A, 23 = SK4B, 20 = SK4M, 16
οτ=ανο μ**α** λιζο μα πιδοριχσηιο
and=HYP ART.NOM.SG.F fortress NEG be.abandoned-PRES.OPT.3SG
'and the fortress might not be abandoned'

(9) SK4A, 26 = SK4B, 23–24 = SK4M, 19–20
υαρουγο μ**α** λιζο χουζο πορooατο
whole ART.NOM.SG.F fortress good stay-PRET.3SG
'the whole fortress fared well'

Uncertain cases:

(10) 217 KT
ειδι μ**α** μαδογο
'this *madugə* (is)'.

The vessel fragment bearing this inscription is broken after the last word, meaning that the sentence might not be complete. Nonetheless, there can be little doubt about the interpretation of the initial syntagm, cf. other vessel inscriptions from the same site such as 218 KT ειδο μο κονδηο̣ 'this water-pot…' (Sims-Williams *apud* Fussman 2011: 106), and, more famously, the beginning of SK4, 1 ειδο μα λιζο 'this fortress (is) …' (Lazard *et al*. 1984: 227; Sims-Williams 2012: 78). The word μαδογο arguably refers to a type of vessel. However, since the etymology of this word is unknown, it is theoretically possible to segment μα μαδογο as μ-αμαδογο, thus eliminating the -α of the feminine article (cf. Livšic 1996: 170). An ostracon from Fayaz Tepe bearing an inscription in the "Unknown script" may support the segmentation μα μαδογο according to the most recent decipherment (cf. Lurje 2024, English translation in this volume, p. 143; Sims-Williams 2025: 205).

(11) Ayr, 2
]ο σταδο þοδβ̣**α** ο̣[
'… was the *shudva*' (?).

[56] SK4B, 10: πιδοριγδι, on which see (84).

On þοδβα cf. §5.1. Sims-Williams' recent interpretation takes it as a nominative (cf. the author's contribution in this volume, p. 326), but the syntactic context cannot be established with certainty because of the poor state of preservation of this line.

(12) Ayr, 4
]δ̣ι λαχþτ̣**α** σ̣ταδ̣ο αβαβσ**α**
'was(?) a desert, without water'.

The word λαχþτ̣α 'desert' is interpreted by Sims-Williams (in this volume, p. 329) as a cognate to CSogd. *δxšt'* 'desert', hence as a feminine *ā*-stem in the nom.sg. case. As for αβαβσα, Sims-Williams (1994: 173) proposed to consider it as the feminine counterpart to αβαβγο 'waterless', cf. (5), with the suffix -σα representing the outcome of the feminine adjectival suffix OIr. **-čī*, recharacterized with the feminine ending **-ā* (cf. Sims-Williams 1981b: 15). If this interpretation is correct, this word displays the final -α of the feminine nom.sg. case, agreeing with the head noun λαχþτα. However, since the whole phrase occurs in a very damaged passage of the Ayrtam inscription, it seems advisable to regard this interpretation as uncertain.

(13) Rab, 4
φροαγδαζο … αγιτ**α**
'there was proclaimed … the capture'.

Sims-Williams (2008: 60–61) tentatively compared this noun to the verb αγισ-: αγιτο 'to take, hold, capture', but the construction of the sentence is not entirely clear.

Ending -α after preposition

(14) Ayr, 3

τι̣δ=ηια	α̣β̣ο̣	μ̣**α**	[λι]ζ**α**	ωσταδο
then=he-OBL.SG	in	ART.SG.F	fortress-SG.F	put-PRET.3SG

'then he placed (it?) in the fortress'

Bactr. λιζο < OIr. **dizā-*.

(15) Ayr, 6

πιδο	ι**α**	þο̣δ̣βαα	φρομαν̣**α**
according to	ART.SG.F	*shudva*-GEN.SG?[57]	command-SG.F

'according to the *shudva*'s command'

Bactr. φρομανο < OIr. **fra-mānā-*.

[57] Although the meaning of the phrase is clear (cf. the colophons of other inscriptions such as SK4B, 27 = SK4M, 23–24 πιδο (ι) χοαδηο φρομανο 'according to the lord's command' and Rab, 21 πιδο þαε φρομανα 'according to the king's command'), it is hard to account for the ending -αα in þο̣δ̣βαα, cf. §5.1.

(16) Rab, 8

φρομαδο	...	βαγολαγγο	κιṛδι	...	αβο	μ**α**	κασιγε
order-PRET.3SG	...	sanctuary	do-INF	...	in	ART.SG.F	Kasig-GEN.SG

ραγ**α**
plain-SG.F
‘(Kanishka) ordered (Shafar) to make … the temple … in the plain of Kasig’

Bactr. ραγο < OIr. **rāgā-*. Sims-Williams (2008: 63) offers two possible translations of the phrase αβο μα κασιγε ραγα: 1) ‘in the plain of Kasig’, interpreting κασιγε as a noun inflected in the gen.sg. and governed by ραγα; 2) ‘in the Kasig plains’, taking κασιγε as an adjective in -σιγο (or -ιγο) inflected in the dir.pl. Sims-Williams 2008 prefers this second interpretation, which implies that ραγα is a dir.pl.f. (< OIr. **rāgāh*). Since pl.f. nouns have not yet been identified with certainty (cf. §5.1 on (25) ϯαṛηνᾳ), Sims-Williams (personal communication) now prefers the interpretation of ραγα as a singular noun, but understands it as an unmarked form qualified by the obl.sg. κασιγε, which he considers to be an adjective. However, if κασιγε is the adjective agreeing with ραγα, we should expect it to take the same ending -α, i.e. the generalised ending of *ā*-stem nouns. Therefore, I preferred the first option presented by Sims-Williams 2008, taking κασιγε as a genitive and μα ραγα as a sg. noun.

(17) Rab, 8–9

βᾳṛᾳνο	κιδ̣ι	μαρο	κιṛδ̣ι	ανδ̣ιμαν̣ι	ǫφαρρο	ομμ**α**
god-OBL.PL	REL	hither	become-PRET.3SG	before	glorious	Umma-SG.F

‘(for these) gods who have come hither into the presence of the glorious Umma’

Bactr. Ομμα < OIr. **uxšmā-* or **us-šmā-* (cf. Sims-Williams 2017).

(18) Rab, 21

π]ιδο	þαε	φρομαν**α**
according to	king-GEN.SG.	command-SG.F

‘according to the king’s command’

Bactr. φρομανο < OIr. **fra-mānā-*.

(19) SK4A, 19 = SK4B, 17–18 = SK4M, 13–14

ατ=ανο	αβο	μ**α**	λιζο	φαρο	καρανο	αβο	μα
so.that=HYP	in	ART.SG.F	fortress	for	people-OBL.PL	water	NEG

γαοηιο
lack-OPT.PRES.3SG
‘so that water should not be lacking to the people in the fortress’

Bactr. λιζο < OIr. **dizā-*.

Uncertain cases:

(20) 217 KT

αβο ι βαο ζουρι**α**

'(this *madugə* is) for the libation(?) of the king'.

Livšic (1996: 171–172) tentatively proposed to identify ζουρια as a cognate of Av. *zaōθrā-* 'libation' (OIr. **zaṷθrā-*) with a **-ya-* extension, but admitted the difficulty of reconciling the use of a Zoroastrian term in a Buddhist context.

Ending -α of the nom.sg. of old *r*-stems

(21) DilV, 8

ι**α** λογδ**α** λαδο

ART.SG.F daughter-SG.F give-PRET.3SG

'the daughter was given/by the daughter was given/(someone) gave the daughter'

Bactr. λογδο < OIr. **dugdā*, nom.sg. of **dugdar-*. This is the only readable part of the sentence, and it is thus hard to establish the syntactic role of ια λογδα.

(22) Rab, 11_14

οτ=ηια φ̣ρ̣ομαδο … ḳ̣ιρδι … αβο ι=πιδ**α**

and=he-OBL.SG order-PRET.3SG … do-INF … to ART=father-SG

'and he gave orders to make (images of these kings): … the father'

Bactr. πιδο < OIr. **pitā*, nom.sg. of **pitar-*.

Ending -α with neuter nouns secondarily incorporated into the *ā*-stems

(23) SK4B, 8, 19–20

ασο ι**α** νιβαλμο

from ART.SG.F seat

'from the seat'

Bactr. νιβαλμο < OIr. **nišadman-*.

(24) *BD* I, **Aa**, 2

ι**α** ρωσο δινο

ART.SG.F day Din

'the day Din'

Bactr. ρωσο < OIr. **raṷčah-*.

Ending -α used as pl. marker, either feminine or neuter (OIr. *-*āh* or *-*ā*)

(25) Rab, 21

αβισ=ọı̣ ı̣=ṗαṛηνα̣ λαδọ

many=ADV ART=rite-DIR.PL.F/N? give-PRET.3SG

'many rites were endowed'

Bactr. παρηνα < OIr. **pari-ayanā* or **pari-ayana-* (Sims-Williams 1998: 88).

Ending -α used as abl.-instr.(-dat.?)sg. (OIr. *-*ā*, -*āt*, -*ā̰*?) after prepositions

(26) Nplate, 4

ασο μο υνδ**α**

from ART India-ABL.-INSTR.SG

'from India'

Bactr. υνδο < OIr. **hindu-* (Sims-Williams & Cribb 1995–1996: 92).

(27) Nplate, 1

π]ı̣δọ [ι=ιωγα] χþον**α**

in [ART=one-ABL.-INSTR.SG] year-ABL.-INSTR.SG

'in the year [one]'

See (28) for the etymology of χþονο and the integration of [ι ιωγα].

(28) Nplate, 2–3

πιδ]ο ι=ιωγ**α** χþον**α**

in ART=one-ABL.-INSTR.SG year-ABL.-INSTR.SG

'in year one'

The origin of Bactr. χþονο, which is most likely the source for the word for 'year' for several languages of the Tarim Basin,[58] is debated. Traditionally, the word is considered an adaptation from Gr. χρόνος 'time'.[59] If so, it probably was included in the masculine *a*-stem class, which indicates that the ending -α here represents the abl.-instr.(-dat.?) ending (cf. Kreidl 2024: 221). The only alternative is to consider it an old *ā*-stem, which

[58] Adams 2013: 261, cf. (Niya) Gandh. *kṣuna*, *kṣana* 'date' (cf. Baums & Glass 2002–a, *s.v.*), TochB *kṣuṃ* 'regnal year', Khot. *kṣuṇa-* 'period of time, regnal period', Tumsh. *xšana-* 'id.'.

[59] Thierfelder *apud* Humbach 1966: 2. Scarborough 2021 recently challenged this view, arguing that the development Bactr. *xr* > *xš* is unparalleled in Bactrian. There are however good reasons to assume that the Greek vibrant may have sounded close to the fricative retroflex Bactr. þ (/ʂ/), cf. the early writings of the dynastic name of the Kushans on the coins of Heraios: κορρανου, κορσανου (Alram 1986: 294). On the retroflex pronunciation of Bactr. þ see Sims-Williams 2011a: 247; Michetti 2024: 186–187.

does not seem supported by early borrowings such as Gandh. *kṣuṇami* (loc.sg.m) and *kṣuṇeṇa* (instr.sg.m) in early Kharoṣṭhī inscriptions (74 CE and 87 CE respectively).[60]

(29) Nplate, 5–6

πιδο ι=λαṣο χþονα
in ART=ten year-ABL.-INSTR.SG
'in the year ten'

See (28) for the etymology of χþονο.

Traces of the instrumental case in the pronominal inflexion and in the *n*-stems

(30) Rab, 7–8

φρομαδο ąβ=εινα οιąγο βαγολαγγο κιṛδι
order-PRET.3SG in=ART.ABL.-INSTR.SG place sanctuary do-INF
'(Kanishka) ordered (Shafar) to make in this place the temple'

Bactr. οιαγο < OIr. **wiyāka-*.

(31) Ayr, 5

οβ=εινα[61] ζαṃṿą
in=ART.ABL.-INSTR.SG time-ABL.-INSTR.SG
'at this time'

Bactr. ζαμανο < OIr. **ǰamă̆n-/ǰamn-*, with the oblique ablauting stem ζαμν- (cf. §5.1).

Ending -α used as abl.-instr.(-dat.?)sg. (OIr. **-ā, -āt, -āi̯*?) with agentive meaning

(32) Ayr, 6

οτι ειμο μιροζαδα νιβιχτο
and DEM Mihrzād-ABL.-INSTR.SG write-PRET.3SG
'and Mihrzād wrote this (inscription)'

Bactr. μιροζαδο < OIr. **miθra-zāta-* (Sims-Williams 2010: 87).

(33) Rab, 14–15

þαονανο þαο ι=βαγεποορą κανηþκε
king-OBL.PL king ART=son.of.gods-ABL.-INSTR.SG Kanishka-GEN.SG

φṛṛṃαδο κιρδι
order-PRET.3SG do-INF
'the king of kings, the son of the gods Kanishka had given orders to do'

Bactr. βαγοπουρο < OIr. **baga-puθra-* (Davary 1982: 173).

[60] CKI 564, CKI 172.

[61] On this preposition cf. Sims-Williams' contribution in this volume, p. 331.

Uncertain cases:

(34) Ayr, 4

τιδι ϧọδϐ**ααᾳ** δι ι[..]νο κανδο / τιδι ϧọδϐ**α** ᾳδι ι[..]νο κανδο
'then the *shudva* dug the (?)'

The writing in this part of the inscription is too effaced to establish a secure reading, and it is unclear whether one should segment ϧọδϐααᾳ or ϧọδϐα. A further difficulty lays in the etymology of this word, §5.1.

Endings -ι, -ε used as gen.sg. markers (OIr. *-*ahya*) with possessive meaning

(35) Alm3, 1

ειδι	ηλου[..]	ϧαονανο	ϧα**ε**	οοημο	τακτο**ε**
DEM=PART	?	king-OBL.PL	king-GEN.SG	Wima	Taktu-GEN.SG

'this (?) (is) of the king of kings Wima Taktu'

The Bactrian word for 'king' ϧαο is usually traced back to OIr. **xšāwan*-, an old *n*-stem (hence the pl. form ϧαονανο). The gen.sg. form ϧαε/ϧαι may accordingly be understood as a defective spelling for *ϧαοε, *ϧαοι. However, from a *n*-stem, we would expect the genitive form to show the nasal of the stem (e.g. *ϧαονι). Since the sign υ for /h/ was only introduced later, I suspect that ϧαε/ϧαι may instead represent a defective spelling for **šāhi*, as suggested by the spelling ϧαυο 'king' attested in SK4A, 11. Bactr. ϧαυο may be the Bactrian outcome of OIr. **xšāyaθiya*- or a MP/Parth. loanword, cf. Halfmann *et al*. (2024: 21, fn. 18), Lurje (2024: 339, English version in this volume, p. 156) and Sims-Williams (2025: 197, fn. 25).

(36) DN1, 2–3

ϧαοναν=δε	ϧᾳ**ι**[62]	ι=ϐωγο	ι=ϛτοργο	οοημο	τακτọο
king-OBL.PL=PART	king-GEN.SG	ART=salvation	ART=great	Wima	Taktu

'[…] of the king of kings, the great salvation, Wima Taktu'

[62] The reading established by Sims-Williams and Cribb (1995–1996: 95) agrees with the corresponding line in the Gandhari version of the inscription (DN4, 2): *rajatirajasa* (gen.sg.) 'of the king of kings', which opens the inscription immediately following the dating formula (Fussman 1974: 21). The two recent editions of the Bactrian text at Dasht-i Nawur diverge on this point. Palunčić *et al*. (2023: 338) agree with the traditional reading and print ϧαι. On the other hand, Halfmann *et al*. (2024: 15) propose a different reading: ϧαονανδε ϧαο ι βωγο ι στοργο οοημο τακτοο, thus changing ϧαι in ϧαο. This would leave us with one less occurrence of the gen.sg. ending -ι. However, they also admit the possibility of an alternative reading ϧαε, but consider it unlikely since a gen.sg. ending would break the sequence of titles with final -ο (cf. Halfmann *et al*. 2024: 17). This does not seem to me a valid reason to exclude the reading ϧαε, since syntagms are often marked on only one element, as demonstrated by several examples in this Appendix and as admitted by the authors. Moreover, the Gandhari version (*rajatirajasa*) provides a strong indication that the Bactrian text opened with a genitive. All in all, I think that a reading ϧαι or ϧαε in the Bactrian text is preferable.

Cf. (35) on the etymology of βαο. The noun governing the genitive is not preserved in the inscription.

(37) Rab, 1

βωγο	ϣτοργο	κανηβκ**ε**	ι=κοβϙνο	
salvation	great	Kanishka-GEN.SG	ART=Kushan	

‘[…] of Kanishka the Kushan, the great salvation’

The noun governing the genitive is lost because the first part of the line is broken.

(38) Rab, 8

αβο	μα	κασιγ**ε**	ραγα
in	ART.SG.F	Kasig-GEN.SG	plain-SG.F

‘in the plain of Kasig’

Cf. (16) on the interpretation of κασιγε.

(39) Rab, 17

βα**ι**	ϕϙομανο
king-GEN.SG	command

‘the king’s command’

Cf. (35) on the etymology of βαο.

(40) NumKan 1310-16, 1319-22

βαο	κανηβκ**ι**
king	Kanishka-GEN.SG

‘(coin) of King Kanishka’

(41) NumKan 1304-08, 1317-18

βαονανο	βαο	κανηβκ**ι**	κοβανο
king-OBL.PL	king	Kanishka-GEN.SG	Kushan

‘(coin) of the king of kings Kanishka the Kushan’

(42) NumHuv 1323-26, 1328, 1331-33, 1335-7, 1341

βαονανο	βαο	οοηβκ**ι**	κοβανο
king-OBL.PL	king	Huvishka-GEN.SG	Kushan

‘(coin) of the king of kings Kanishka the Kushan’

(43) NumHuv 1327

βαονανο	βαο	ουοηβκ**ι**	κοβανο
king-OBL.PL	king	Huvishka-GEN.SG	Kushan

‘(coin) of the king of kings Huvishka the Kushan’

(44) NumHuv 1329, 1338, 1343-45

βαονανο	βαο	οοηβκ**ε**	κοβανο
king-OBL.PL	king	Huvishka-GEN.SG	Kushan

‘(coin) of the king of kings Huvishka the Kushan’

(45) SK4B, 24–26

μο	σαδο	βορζομιορο	κιρδι	κοζγαβκ**ι**	πουρο	νοκονζικ**ι**
DEM	well	Burzmihr	do-PRET.3SG	Kuzgashk-GEN.SG	son	Nukunzik-GEN.SG

καραλραγγ**ι**	μαρηγι
margrave-GEN.SG	servant-GEN.SG

‘this well was made by Burzmihr, the son of Kuzgashk, servant of the margrave Nukunzik’

Bactr. καραλραγγο < OIr. **karāna-dranga-* (Davary 1982: 207–208). Of the four nouns marked by the gen.sg. -ι ending (κοζγαβκι, νοκονζικι καραλραγγι, μαρηγι), μαρηγι expresses agentive meaning in apposition to βορζομιορο (lit. ‘by Burzmihr, the servant’), cf. (56). The other genitives have possessive meaning: κοζγαβκι modifies the noun πουρο (‘Kuzgashk’s son), while νοκονζικι καραλραγγι modifies μαρηγι (‘servant of the margrave Nukunzik’). Cf. Humbach 2003: 162–165, Sims-Williams forthcoming.

(46) SK4M, 21–23

μανο	κιρδο	αμο	βορζομιυρο	αμο	κοζγαβκ**ι**
I-OBL.SG	do-PRET. 3SG	DEM.OBL.SG	Burzmihr	DEM.OBL.SG	Kuzgashk-GEN.SG

πουρο	...	αμο	νοκονζικ**ι**	καραλραγγ**ε**	μαρηγο
son	…	DEM.OBL.SG	Nukunzik-GEN.SG	margrave-GEN.SG	servant

‘(this well and *maštə xirgə*) were made by me, Burzmihr, son of Kusgashk … servant of the margrave Nukunzik’

Bactr. καραλραγγο < OIr. **karāna-dranga-* (Davary 1982: 207–208), on αμο cf. (57).

(47) *BD* IV, **kc**, 1–2 = **kg**, 1 = **kw**, 1, v1

αβο	σανοσιδδο	ι=ραβτ**ι**	βα(ο)βορο	βιγανο
to	Sansidd	ART=just-GEN.SG	Shabur	royal.retainer

‘to Sansidd, the royal retainer of the just Shabur’

Bactr. ραβτο < OIr. **rašta-*.

Endings -ι, -ε used as gen.sg. markers with agentive meaning

(48) Nplate, 1–2

ταδι	ι=βαγεποορ**ε**	…	ωσταδημι
Then	ART=son.of.gods-GEN.SG	…	put-PRET.1SG

‘Then the son of the gods established me’

Bactr. βαγοπουρο < OIr. **baga-puθra-* (Davary 1982: 173). The sentence is extracted from a longer text: ταδι ι βᾰγεποορε πιδο ι χοβι βιζ[α]ᾰ ǫ̣δο πιδο ι μανο ʂπαχτε ταδηο αλο ι πιοριβτειγανο οδο αλο ι νιαγοβτηγανο μαρηγανο αμσασογο ωσταδημι αλο ι φαρδαμγανο 'Then the son of the gods, on account of his own good[ness] and on account of my service—he established me (as) equal(?) with (his) father's and with (his) grandfather's servants, with the foremost (people)' (Sims-Williams 2015: 257). The role of the βαγεποορε as the agent of the verb ωσταδημι is confirmed by the presence of the enclitic pronoun -ηο (lit. 'by him'), attached to the conjunction ταδ- and expressing the agent of the verb ωσταδημι.

(49) Nplate, 5

καλδι κανηβκ**ι** βαο αμο[63] ǫ̣ηβο αγαδο
when Kanisha-GEN.SG king to Wesh bring-PRET.3SG
'when king Kanishka brought (it) to Wesh'

(50) Rab, 7

βα**ι** ᾰ̣ανῃβκ**ε** αβο βαφαρο ᾰ̣αραλραγγο φρομαδο
king-GEN.SG Kanishka-GEN.SG to Shafar margrave order-PRET.3SG
'King Kanishka ordered Shafar the lord of the marches'

Cf. (35) on the etymology of βαο.

(51) Rab, 14–15

βαονανο βαο ι=βαγεποορα κᾰ̣νῃβᾰ̣κ**ε** φρǫ̣μαδο
king-OBL.PL king ART=son.of.gods-ABL.-INSTR.SG Kanishka-GEN.SG order-PRET.3SG
κιρδι
do-INF
'the king of kings, the son of the gods Kanishka had given orders to do'

(52) Rab, 15

βαφαρ**ε** κᾰ̣ραλραγγ**ε** κιρδο ειο βᾰ̣ǫ̣λαγγǫ̣
Shafar-GEN.SG margrave-GEN.SG do-INF this sanctuary
'Shafar the lord of the marches made this sanctuary'

Bactr. καραλραγγο < OIr. **karāna-dranga-* (Davary 1982: 207–208).

(53) Rab, 22

βα**ι** μα λ̣ιζγᾰ αβο βαγανο λαδο
king-GEN.SG ART.SG.F fortress-SG.F to god-OBL.PL give-PRET.3SG
'the king gave the fortress to the gods'

Cf. (35) on the etymology of βαο.

[63] Contracted form of αβο μο, cf. Sims-Williams 2015: 262, fn. 35.

(54) SK1

]ṿοβιχτο	μο	μαβτο	ουβ=εμο	παγδο	ιωλεσαγωγ**ι**
write-PRET.3SG	ART	*maštə*	on=DEM	*pagdə*	Yōlesagōg-GEN.SG

‘Yōlesagōg wrote the *maštə* on this *pagdə*’

A few points of this inscription are obscure, mainly because the meanings of μαβτο and παγδο are unknown. According to the interpretation provided above, μο is the usual definite article governed by μαβτο. This word, arguably a noun, is attested in a similar context in the SK4M colophon: οτο ειιο μο σαδο οδο μαβτο χιργο μανο κιρδο ‘and this well and *maštə xirgə* were made by me’ (Humbach 2003: 162; Sims-Williams, forthcoming). In both occurrences, μαβτο can also be segmented as μ-αβτο ‘the *aštə*’, thus taking μ- as the univerbated article (Sims-Williams 1975: 95, fn. 1). The preceding μο could be interpreted as the 1sg. enclitic pronoun: ‘I, Yōlesagōg, wrote the *aštə* on this *pagdə*’ (Sims-Williams, forthcoming). In either case, the role of the PN ιωλεσαγωγι as an agent marked by the genitive case does not change.

(55) SK4A, 1–3 = SK4B, 2–4 = SK4M, 1–2

… βαγολαγγο σιδο[64]	ι=βαγο	þαο	κανηþκ**ι**[65]	ναμοβαργο	κιρδο
… sanctuary REL	ART=lord	king	Kanishka-GEN.SG	name.bearer	do-PRET.3SG

‘(this fortress is Kanishka the victorious’) sanctuary, which the lord, king Kanishka made name-bearing’[66]

(56) SK4B, 25–26

μο	σαδο	βορζομιορο	κιρδι	…	υαστιλογανζειγο	…	μαρηγ**ι**
DEM	well	Burzmihr	do-PRET. 3SG	…	native.of.Hastilgan	…	servant-GEN.SG

‘this well was made by Burzmihr…, native of Hastilgan, the servant…’

Bactr. μαρηγο < OIr. **mariyaka-*.

(57) SK4M, 20–23

ειιο	μο	σαδο	…	μανο	κιρδο	…	αμο	αστιλογανσειγ**ι**
DEM	ART	well	…	me-OBL.SG	do-PRET.3SG	…	DEM.OBL.SG	native.of.Hastilgan

‘this well was made by me … the native of Hastilgan’

Bactr. (υ)αστιλογανσειγο < *Hastilgan* + OIr. **-čiya-ka-* (cf. Sims-Williams 2010: 139). As convincingly argued by Sims-Williams (forthcoming), the repeated element αμο is

[64] SK4B: σιδι.

[65] SK4B: κανηþηþκι, clearly a mistake.

[66] This is the generally accepted translation of this passage (cf. Gershevitch 1966: 106; Gershevitch 1979: 64; Sims-Williams 2012: 78), which was first elucidated by Henning (1960: 52, *contra* Humbach 1976: 66; Davary 1982: 235). According to an alternative translation provided by the same scholar, κανηþκι can be interpreted as a possessive genitive: ‘sanctuary, which was made bearing the name of the lord, king Kanishka’ (cf. Henning 1960: 52).

best interpreted as an oblique form of the article and demonstrative μο ‘the, this’. If this interpretation is correct, each phrase or word preceded by αμο (βορζομιυρο, κοζγαþκι πουρο, αστιλογανσειγι, νοκονζικι καραλραγγε μαρηγο) is marked by it as an apposition to the agent, which is expressed by the oblique pronoun μανο ‘by me’. Of all these elements, only αστιλογανσειγι also shows the ending of the genitive case with agentive function. Conversely, in the shorter variant of this passage contained in SK4B, the ending of the genitive case is present on μαρηγι but not on υαστιλογανζειγο, cf. (56).

Ending -ι, -ε (< OIr. *-*ahya*) used after prepositions

(58) DN1, 5–6

πιδο	ι=χοβ**ε**	þαν**ε**
according to	ART=own-GEN.SG	authority-GEN.SG

‘according to his own will’

Bactr. þανο < OIr. **xšāna-* (cf. Halfmann *et al*. 2024: 20–21).

(59) Nplate, 1

αβο	þαονανο	þαο	κανηþκ**ι**	κοþανο	ι=χοαδηοδανι
to	king-OBL.PL	king	Kanishka-GEN.SG	Kushan	ART=lordship-NOM.SG

λαδο
give-PRET.3SG
‘[Nana] gave the lordship to the king of kings, Kanishka the Kushan’

(60) Nplate, 2

πιδο	ι=χοβ**ι**	þιζ[δ]**ε**[67]
on account of	ART=own-GEN.SG	goodness-GEN.SG

‘on account of his own goodness’

Bactr. þιζδο < OIr. **xšiǐa-tāt-*, cf. *BD* II: 228.

(61) Nplate, 2

πιδο	ι=μανο	ϙπαχτ**ε**
on account of	ART=I-OBL.SG	service-GEN.SG

‘on account of my service’

Bactr. σπαχτο < OIr. **spaxšti-*.

[67] I prefer this reading, presented as an alternative in a footnote, to the one given in the main text by Sims-Williams (2015: 257, with fn. 4). The expected gen.sg. form of þιζαο (< OIr. **xšiǐa-āwa-*) would be *þιζαοε, not þιζ[α]ε.

(62) Nplate, 3

οτ=ηο αζο αβο ι=κηρ**ι** μαρηγο αλοοαδγο σταδημε
and=he-OBL.SG I in ART=work-GEN.SG servant trusted? be-PRET.3SG
'and I was a trusted (?) servant in his work'

Bactr. κηρο < OIr. **kārya-* (Sims-Williams 2015: 260).

(63) Nplate, 5–6

πιδο ι=λαϙο χβοṇα ṿ̣ɛ̣ıṣ̣ą̣ν**ɛ** μα**ɛ**
in ART=ten year-ABL.-INSTR.SG Nisan-GEN.SG month-GEN.SG
'in the year ten, (in) the month Nisan'

Cf. (29). OIr. μαο < **māhV-*. Bactr. νεισανο comes from the Babylonian month-name *Nisannu*, apparently adapted to the *a*-stems (cf. Sims-Williams & De Blois 2018: 25). It seems that πιδο governs two nouns with different case marking (χβονα and μαε), cf. the discussion in (72).

(64) Rab, 11–13

οτ=ηια φṛṛομαδο … ḳιṛδι … αβο οοημο καδφισ**ɛ̣**
and=he-OBL.SG order-PRET.3SG … do-INF … to Wima Kadphises-GEN.SG
'and he gave orders to make (images of these kings): … Wima Kadphises'

Examples (64)–(66) are cases of Differential Object Marking (cf. Sims-Williams 2011b): the direct object is marked with the preposition αβο governing the genitive.

(65) Rab, 11–14

οτ=ηια φṛṛομαδο … ḳιṛδι … αβο ι=χοβισαρο κανηϸκ**ɛ**
and=he-OBL.SG order-PRET.3SG … do-INF … to himself Kanishka-GEN.SG
'and he gave orders to make (images of these kings): … himself, Kanishka'

(66) Rab, 17–18

ειμι=δ=βα βαγε … αβṛ κανηϸκ**ɛ** κṛϸανο …
DEM.DIR.PL=PART=HORT god-DIR.PL … to Kanishka-GEN.SG Kushan …

[….]ιṿδι
[keep]-PRES.OPT.3PL
'may these gods … [keep] Kanishka the Kushan (for ever healthy, fortunate and victorious)'

(67) Rab, 18

αβο ιαοηδαν**ι** ζορριγ**ι**
to eternal-GEN.SG time-GEN.SG
'for ever'

Bactr. ιαοηδανο < OIr. **yāwaṯ-tāna-*, ζορ(ρ)ιγο < OIr. **zrunaka-*.

(68) Splate
ąṣo ι=ροζγ**ι**
from ART=vineyard-GEN.SG
'from the vineyard'

Bactr. ροζγο < OIr. **razu-ka-*.

(69) *BD* II, **cc**, 7–8
αβο **τι** χοηο
to DEM.GEN.SG lord
'to your lordship'

Bactr. το < OIr. **ta-*.

(70) *BD* IV, **kc**, 11–12
π[ιδο ι=χοβο þι]ζγ**ι**
in ART=own goodness-GEN.SG
'in your own goodness'

Bactr. þιζγο < OIr. **xšiǰa-ka-*. The same phrase is attested in *BD* IV, **kw**, 7 (πιδο χοβọ þιζγ[ι), where final -ι is reconstructed based on the usage of the document, cf. (71).

(71) *BD* IV, **kw**, 8–9
πιδο το χοαδηο þιζγ**ι**
in your lordship goodness-GEN.SG
'in your lordship's goodness'

Bactr. þιζγο < OIr. **xšiǰa-ka-*.

Uncertain cases:

(72) Nplate, 5–6
πιδο … λασσο σαχτ**ε**
'on the tenth (day)'

The use of the word Bactr. σαχτο 'elapsed' < OIr. **saxta-* (pp. of **sač-* 'to pass') in dating formulas has parallels in other Middle Iranian languages and is reminiscent of the Old Persian way of expressing dates with the same verb (e.g. *DB* 1.42: *IX raučabiš θakatā āha avaθā*… '9 days were gone past, then …', cf. MacKenzie 1994). The full dating formula is as follows: πιδο ι λαṣο χþονα ṿẹịṣąνε μαε λασσο σαχτε. In (63) above, I took χþονα 'year' and μαε 'month' to be both governed by πιδο 'in'. Based on parallels in other Middle Iranian languages, such as MParth. *pd sxt cwhrm* 'on the fourth (day) passed' (MMiii, d, 57 = M5 R ii 27; cf. Sims-Williams 2015: 261), it seems likely that πιδο here governs also λασσο σαχτε (*scil*. ρωσο 'day'), in which case the -ε should be interpreted as a gen.sg. ending and the whole formula may be translated 'in the year ten, (in) the

month Nisan, (in) the tenth (day)'. However, in the recently published birchbark letters (late 4th cent. CE) a dating formula consisting only of the word σαχτι + cardinal number is attested at least three times: *BD* IV, **kc**, 19: σαχτι β΄ 'day 2' (lit.'(day) 2 elapsed'); **kh**, 48: σα[χτ]ι ΄ 'day 10'; **kn**, 25: νισανο μᾱ̄ō̄ σᾱ̄χ̄τι ϰ΄ 'month Nisan, day 20'; **kq**, 10: σαχτ]ι 'day…' (the number here may have preceded σαχτι and get lost in the lacuna, cf. *BD* IV: 95). Sims-Williams (personal communication) points out to me that, if one separates the phrase λασσο σαχτε from the preceding πιδο, the date of Nplate may represent another example of the same dating formula. He further observes that the form σαχτι/σαχτε may be locative in origin, as in the version of the formula often used in Sogdian texts: cardinal + loc.sg. *saγdyā* 'on the X(th day) passed' (cf. Henning 1937: 134). It is not impossible that an old locative form was retained in fixed expressions (cf. fn. 11). However, I would not exclude that σαχτι used independently in the birchbark letters represents a nom.sg. form with ending -ι, cf. (83)–(95), perhaps as a shortening of a longer formula like 'day X had elapsed' (cf. the Old Persian formula with *θakatā*). Whatever the origin of σαχτι, I consider it more likely that the formulation in Nplate 5–6 is closer to that of the Sogdian Ancient Letters, cf. AL 3, r26 *pr ʾtδrtyk YRḤʾ pr 10 sxth* 'in the third month on the 10th day', with πιδο governing all the following words.

(73) SK4B, 22–23
πιδειν**ι** σαδο πιδειν**ι** αχβτριγο
'by means of this well and this winch(?)'.

The traditional reading πιδεινι has been recently challenged by Kreidl (2024: 210, fn. 10), who prefers the reading πιδειμι, interpreting it as a remnant of the locative form of the demonstrative pronoun ειμο, cf. Av. *ahmǐ*. Cf. §5.2.

(74) SK4A, 18 = SK4B, 16–17 = SK4M, 13
πιδο ασαγγ**ε**
'with stones'

The etymology is clear (OIr. **asanga*-), but the ending -ε may be interpreted as a dir.pl. form ('with stones'; Maricq 1958: 365; Henning 1960: 53), or as an oblique form of a collective singular, as proposed by Lazard *et al*. 1984: 225.

Ending -ι/-ε with infinitives, probably from OIr. **-aį̇*

(75) DN1, 5–7
κιδι … þαοδαν**ε** ḷφαχτο
REL … kingship-INF gain-PRET.3SG
'who gained the kingship'

From þαο 'king' + infinitive suffix OIr. **-tanaį̇* ? (cf. §5.4)

(76) Nplate, 1

αβο βαονανο βαο κανηβκι κοβανο ι=χοαδηοδανι
to king-OBL.PL king Kanishka-GEN.SG Kushan ART=lordship-INF

λαδο
give-PRET.3SG
'[Nana] gave the lordship to the king of kings, Kanishka the Kushan'

From χοαδηο 'lord' + infinitive suffix OIr. *-*tanaį* ? (cf. §5.4)

(77) Rab, 2

κιδι ασǫ νανα … ι=βαοδανι αβορδο
REL from Nana-SG.F … ART=kingship-INF obtain-PRET.3SG
'who has obtained the kingship from Nana'

From βαο 'king' + infinitive suffix OIr. *-*tanaį* ? (cf. §5.4)

Ending -ε used as dir.pl. marker

(78) Rab, 17

ειμι=δ=βα βαγε κιδι μαρο νιβιχτενδι
DEM.DIR.PL=PART=HORT god-DIR.PL REL here write-PRET.3PL
'may these gods who are inscribed here'

Bactr. βαγο < OIr. **baga*-.

(79) Rab, 21

αβισσι ρηδγε λαδο
many attendant-DIR.PL give-PRET.3SG
'many attendants were endowed'

Bactr. ρηδγο < OIr. **raįtaka*- (cf. Sims-Williams 1998: 88). The 3sg. form of the verb is used as an unmarked form also with plural subjects, cf. (96) and Sims-Williams 2008: 64.

(80) SK4A 6–7 = SK4B 8 = SK4M 4–5

ταδο ι=βαγε ασο ι=νοβαλμο[68] φροχορτινδο[69]
then ART=god-DIR.PL from ART=seat be.displaced-PRET.3PL
'then the gods were displaced from (their) seat'

Bactr. βαγο < OIr. **baga*-.

[68] SK4B: ια νοβαλμο, on which see (23).
[69] SK4B: φροχορτινδι.

(81) SK4B 19–20 = SK4M 15–16

ταδ=ανο ι=βαγ**ε** ασο ι=νοþαλμο[70] μα φροχοαþονδηιο
then=HYP ART=god-DIR.PL from ART=seat NEG be.displaced-PRS.OPT.3SG
'then the gods might not be displaced from (their) seat'

Bactr. βαγο < OIr. **baga-*.

Uncertain cases:

(82) Rab, 4

ąβǫ ι=υṇδο φροαγδαζο αβο þατριαγγ**ε** þαορ**ε**
'there was proclaimed to India, to the cities/city of the *kṣatriya*s(?)'.

Bactr. þατριαγγο < Skt./Gandh. *kṣatriya* + -αγγε < OIr. **-ānaka-* (Sims-Williams & Cribb 1995–1996: 83; Bactr. þαρο < OIr. **xšaθra-*. Since the preposition αβο often governs the gen.sg. rather than a direct case, the ending -ε may be interpreted both as a gen.sg. or a dir.pl.

Ending -ι as a marker for the nom.sg. (< OIr. *-*ah*)

(83) Nplate, 3

ȩιμο κηρ**ι** μανο οαρο κιρδ[ο
DEM work-NOM.SG I-OBL.SG there do-PRET.3SG
'I performed there this work'

Bactr. κηρο < OIr. **kārya-*.[71]

(84) SK4B, 10

οτι μα λιζο πιδοριγδ**ι**
and DEM.NOM.SG.F fortress be.abandoned-PRET.3SG
'and this fortress was abandoned'

Bactr. πιδοριγδο < OIr. **pati-rixta-*, pp. of πιδοριχσ- (Davary 1982: 257). It is hard to tell why the preterite form here is πιδοριγδι and not a form agreeing in gender and number with the feminine noun μα λιζο (cf. (4)–(9)) such as πιδοριγδα.[72] Perhaps in Bactrian the nom.sg.m. form of the pp. was generalised with the grammaticalisation of the ergative

[70] SK4B: ια νοþαλμο, on which see (23).

[71] Theoretically, the ending -ι may be a trace of the **-y-* in the suffix. However, *-y-* is responsible for the *i*-umlaut of the stem, causing the OIr. **-ā-* to be raised to Bactr. -η-. Since no traces of the *-y-* suffix are preserved in other Kushan Bactrian words that present the same phenomenon, cf. SK4M, 14 λρουμινο < OIr. **duš-manyu-*, one may surmise that the vowel triggering the umlaut had already been deleted.

[72] The participial form still agrees with the logical object in the 4th-century Sogdian Ancient Letters, cf. AL 3, r23 *ʾYKZYm ZK twʾxky <sxwn> krty* 'when I did your bidding', AL 4, r4 *ʾḤRZY(m) ZK δy(k)h wyth* 'I saw the letter' (cf. Sims-Williams 2023b: 49).

construction. Such generalisation of the masculine can be observed in the Sogdian intransitive preterite, which is formally identical to a past participle inflected in the nom.sg.m. ending *-y* in the 3sg. (cf. Gershevitch 1954: 861).

(85) SK4, 16

οτηιο αβο οζοοαστ**ι**
and=him-OBL.SG water bring.out-PRET.3SG
'and he brought out water'

Bactr. οζοοαστο < OIr. **uz-wāsta-*, pp. of αζοηλ-. Bactr. αβο 'water' (OIr. **āp-*) is originally a feminine noun, cf. (84) for the agreement.

(86) SK4, 24–25

οτ[ι] μο σαδο βορζομιορο κιρδ**ι**
and DEM well Burzmihr do-PRET.3SG
'and Burzmihr made this well'

Bactr. κιρδο < OIr. **kr̥ta-*, pp. of κιρ-.

(87) *BD* I, **aa**, 19

οτι αγγιṭ**ι** μανο μo̤ λ̣ą[δογο]ζγo̤
and receive-PRET.3SG I-OBL.SG DEM Ladguzg
'and there has been received by me, Ladguzg'

Bactr. αγγιτο < OIr. **ham-kašta-*, pp. of *αγγισ-.

(88) *BD* I, **aa**, 35

o̤ṭ̣ι μ̣[ο πω]στογο μαγo̤ ṿ̣ιβo̤χṭ̣**ι**
and DEM letter I-OBL.SG write-PRET.3SG
'and this contract was written by me'

Bactr. ναβιχτο < OIr. **ni-pixšta-*, pp. of ναβισ-.

(89) *BD* I, **A**, 15–16

αγγιτ**ι**=δο αμαχο μανο βαβο οδο πιδοκο αβο ραλικο ολο
receive-PRET.3SG=PART us I-OBL.SG Bab and Piduk to Ralik wife
'we have received Ralik – I, Bab, and (I), Piduk – (as) a wife'

Bactr. αγγιτο < OIr. **ham-kašta-*, pp. of *αγγισ-. Here αβο marks the direct object.

(90) *BD* II, **cc**, 4, 9

τ**ι** χοηο
DEM.NOM.SG lord
'your lordship'

Bactr. τι, το < OIr. **ta-*.

(91) *BD* I, **C**, 1'
ειδο μο λαþνοβω[στ]ǫγο μαλο ναβιχτ**ι**
DEM ART deed.of.gift here write-PRET.3SG
'this deed of gift was written here'

Bactr. ναβιχτο < OIr. **ni-pixšta-*, pp. of ναβισ-.

(92) *BD* IV, **kw**, 3
οτο=μο ναγατ**ι**
and=I-OBL.SG hear-PRET.3SG
'and I have heard'

Bactr. ναγατο < OIr. **ni-kašta-*, pp. of ναγαυ- (cf. Michetti 2024), also in (93)–(94).

(93) *BD* IV, **kw**, 5
ταδ=μο ναγ[α]τ**ι**
then=I-OBL.SG hear-PRET.3SG
'so I have heard'

(94) *BD* IV, **kw**, 11
αβαþιδο=μο ǫαγατ**ι**
next=I-OBL.SG hear-PRET.3SG
'next: I have heard'

(95) *BD* IV, **kw**, 9
αβ[ο] ṭ[ο χ]οαδηο νοβιχτ**ι**
to your lordship write-PRET.3SG
'[I have] written to your lordship'

Bactr. ναβιχτο < OIr. **ni-pixšta-*, pp. of ναβισ-.

Uncertain cases:

(96) Rab, 8–9
βǫṇǫο κιḍι μαρο κιṛδ**ị** ανḍιμαṇị ǫφαρρο ομμα
'(for these) gods who have come hither into the presence of the glorious Umma'.

This passage has proven difficult to read and interpret. According to the most recent interpretation by Sims-Williams (2008: 64), κιṛδị would be the 3sg. preterite form of βο- 'to become'. The plural subject (βαγανο) would have theoretically required a form *κιρδενδι (later κιρδινδο), but cf. SK4M, 21–22 οτο ειιο μο σαδο οδο μαþτο χιργο μανο κιρδο 'this well and *maštə xirgə* were made by me', where there is no 3pl. copula -ινδο attached to the pp. κιρδο ("unmarked" form, cf. Sims-Williams 2008: 64). If the reading κιρδι is correct, final -ι is thus arguably a nom.sg. ending. This may corroborate the idea of a generalisation of this specific form of the past participle for the third person (both sg. and pl.) of the preterite inflexion, cf. (84).

(97a) *BD* IV, **kc**, 19
σαχτ**ι** β΄
'day 2'

(97b) *BD* IV, **kh**, 48
σα[χτ]**ι** ι΄
'day 10'

(97c) *BD* IV, **kn**, 25
νισανο μᾱ̄ο̄ σᾱ̄χ̄τ**ι** ϰ΄
'month Nisan, day 20'

(97d) *BD* IV, **kq**, 10
σαχτ]**ι**
'day…'.

Cf. the discussion in (72).

(98) *BD* IV, **kn**, 5
αμβιτ**ι**
'?'

The context in which this unknown word occurs is too fragmentary to establish its syntactic role.

Ending -ι with pronouns

(99) Rab, 2

ασο	οισποανο	μ**ι**	βαγανο
from	all-OBL.PL	DEM.DIR.PL?	god-OBL.PL

'from all the gods'

Bactr. μι, μο < OIr. nom.pl. **imaį̇*. See the discussion in §5.2.

(100) Rab, 17

ειμ**ι**=δ=βα	βαγε
DEM.DIR.PL=PART=HORT	god-DIR.PL.

'may these gods'

Bactr. ειμι, ειμο < OIr. nom.pl. **imaį̇*. See the discussion in §5.2.

(101) Rab, 22

[α]βο	μ**ι**	βαγεα[βο]
in	ART.OBL.SG	Bage-ab

'in Bage-ab'

Bactr. μι, μο < OIr. **imahya*. See the discussion in §5.2.

(102) SK3, 2

μι	βαγολαγγο
ART.OBL.SG	temple

‘in the temple’

Bactr. μι, μο < OIr. **imahya*. See the discussion in §5.2.

Abbreviations

1,2,3	first, second, third person	Lat.	Latin
abl.	ablative	m.	masculine
adv	adverb	ManB.	Manichaean Bactrian
art	article	MParth.	Manichaean Parthian
AL	Ancient Letters, see Sims-Williams 2023b	MP	Middle Persian
		MSogd.	Manichaean Sogdian
Av.	Avestan	n.	neuter
CKI	Catalog number of Gāndhārī texts: Inscriptions, see Baums & Glass, 2002–b	nom.	nominative
		NP	New Persian
		obl.	oblique
CSogd.	Christian Sogdian	OIr.	Old Iranian
DB	Bisitun Inscriptions of Darius, see Schmitt 1991	OP	Old Persian
		part	particle
dem	demonstrative	Parth.	Parthian
dir.	direct	pl.	plural
EIr.	Eastern Iranian	PN	personal name
f.	feminine	pp.	past participle
IMP	Inscriptional Middle Persian	pret	preterite
G.	Graeco-Bactrian script	rel	relative
Gandh.	Gandhari	sg.	singular
gen.	genitive	Skt.	Sanskrit
GN	geographic name	Sogd.	Sogdian
Gr.	Greek	SSogd.	Sogdian in cursive script
hort	hortative particle	TochB	Tocharian B
instr.	instrumental	Tumsh.	Tumshuqese
It.	Italian	Ved.	Vedic
Khot.	Khotanese	YAv.	Young Avestan

Kushan Bactrian inscriptions

Alm	Almosi gorge (Tajikistan). Ed. by Bobomulloev *et al*. 2022: 64.
Ayr	Ayrtam (Uzbekistan). Ed. by Turgunov *et al*. 1981 + Sims-Williams 2008: 63 (line 5) + Sims-Williams 1994: 173 (line 4) + Sims-Williams 2010: 157 (line 2, 4, 6). New edition by Sims-Williams in this volume (p. 323).
DilI–V	Dilberjin (Afghanistan). Ed. by Livšic & Kruglikova 1979 + Sims-Williams 2008: 64 (line 11) + Sims-Williams 2010: 56 (line 7) + *BD* II: 242 (line 7), 228 (line 8).
DN1	Dasht-i Nawur 1 (Afghanistan). Ed. by Sims-Williams & Cribb 1995–1996: 95 + Davary & Humbach 1976: 6. New reading by Halfmann *et al*. 2024, cf. also Sims-Williams 2025: 188–190.
DN4	Dasht-i Nawur 4 (Afghanistan). Ed. by Fussman 1974.
KT	Kara Tepe (Uzbekistan). Ed. by Fussman 2011.
Nplate	Nukunzik plate (unknown provenance). Ed. by Sims-Williams 2015.
NumKan	Coins of Kanishka (several findspots). Alram 1986, nos. 1304–1308, 1310–1322.
NumHuv	Coins of Huvishka (several findspots). Alram 1986, nos. 1323–1346.
Rab	Rabatak (Afghanistan). Ed. by Sims-Williams 2008.
SK1	Surkh Kotal 1 (Afghanistan). Ed. by Benveniste 1961: 147.
SK3	Surkh Kotal 3 (Afghanistan). Ed. by Benveniste 1961: 151.
SK4A, B, M	Surkh Kotal 4 A, B, M (Afghanistan). Synoptic edition by Davary 1982: 53–64. Translation by Sims-Williams 2012 and Sims-Williams forthcoming (colophon).
Splate	Sengul plate (private collection). Sims-Williams 2013: 194–195.

References

ADAMS, Douglas, 2013. *A dictionary of Tocharian B: Revised and greatly enlarged*, Amsterdam–New York.

ALRAM, Michael, 1986. *Nomina propria iranica in nummis*, Wien.

ALTHEIM, Franz, 1962. *Geschichte der Hunnen.* Band 5: *Niedergang und Nachfolge*, Berlin.

BAUMS, Stefan & Andrew GLASS, 2002–a. *A dictionary of Gāndhārī.* Available at: https://gandhari.org/dictionary.

BAUMS, Stefan & Andrew GLASS, 2002–b. *Catalog of Gāndhārī Texts.* Available at: https://gandhari.org/catalog.

BD I = SIMS-WILLIAMS, Nicholas, 2000. *Bactrian documents from Northern Afghanistan, I: Legal and economic documents*, London.

BD I^2 = SIMS-WILLIAMS, Nicholas, 2012. *Bactrian documents from Northern Afghanistan, I: Legal and economic documents*, revised edition, London.

BD II = SIMS-WILLIAMS, Nicholas, 2007. *Bactrian documents from Northern Afghanistan, II: Letters and Buddhist texts*, London.

BD III = SIMS-WILLIAMS, Nicholas, 2012. *Bactrian documents from Northern Afghanistan, III: Plates*. London.

BD IV = SIMS-WILLIAMS, Nicholas, 2025. *Bactrian documents* IV. *Documents from South of the Hindukush, I, with a contribution from Frantz Grenet*, London.

BENVENISTE, Émile, 1952. 'La construction passive du parfait transitif', *Bulletin de la Société Linguistique* 48/1, 52–62.

BENVENISTE, Émile, 1961. 'Inscriptions de Bactriane', *JA* 249, 113–152.

BENVENUTO, Maria Carmela & Harald BICHLMEIER, 2022. 'On Bactrian umlaut', *Studi e saggi linguistici* 60, 81–108.

BENVENUTO, Maria Carmela & Flavia POMPEO 2012. 'Il sincretismo di genitivo e dativo in persiano antico', *Vicino Oriente* 16, 151–165.

BICHLMEIER, Harald, 2011. *Ablativ, Lokativ und Instrumental im Jungavestischen*, Hamburg.

BOBOMULLOEV, Bobomullo S., Šerali XODŽAEV & Muhsin G. BOBOMULLOEV, 2022. 'Naskal'nye nadpisi uščel'ja Almosi (predvaritel'nye rezul'taty issledovanij)', in *III Esikie čtenija. Sbornik materialov. Meždunarodnoj naučno-praktičeskoj onlajn-konferencii na temu: "Pis'mennye Pamjatniki Central'nyj Azii (problemy perevoda i interpretacii)"* (ed. G. R. MUXTAROVA), Almaty, 59–65.

BRIXHE, Claude, 1984. *Essai sur le grec anatolien au début de notre ère*, Nancy.

CANTERA, Alberto & Céline REDARD, 2023. *An introduction to Young Avestan: A manual for teaching and learning*, Wiesbaden.

CARDONA, George, 1970. 'The Indo-Iranian construction *mana* (*mama*) *kr̥tam*', *Language* 46, 1–12.

CIANCAGLINI, Claudia A., 1987. *Morfologia e funzione del tipo «manā kartam» nel persiano antico*, Roma.

COSERIU, Eugenio, 1958. *Sincronía, diacronía e historia. El problema del cambio linguistico*, Montevideo.

COSERIU, Eugenio, 1992. '«Linguistic change does not exist»', in *Charisteria Victori Pisani oblata* (ed. G. BOLOGNESI & C. SANTORO), vol. II, Galatina, 167–179.

CRIBB, Joe, 1998. 'The end of Greek coinage in Bactria and India and its evidence for the Kushan coinage system', in *Studies in Greek Numismatics in memory of Martin Jessop Price* (ed. R. ASHTON & S. HURTER), London, 83–98.

DAVARY, Gholam Djelani, 1982. *Baktrisch. Ein Wörterbuch auf Grund der Inschriften, Handschriften, Münzen und Siegelsteine*, Heidelberg.

DAVARY, Gholam Djelani & Helmut HUMBACH, 1976. *Die baktrische Inschrift IDN1 von Dasht-e Nāwūr (Afghanistan)*, Mainz–Wiesbaden.

DEL TOMBA, Alessandro, 2021. 'Genus alternans in Khotanese? A synchronic analysis of the agreement classes', *Archivio Glottologico Italiano* 106, 164–194.

DEL TOMBA, Alessandro, 2022. 'Avestan *-å̊ŋhō*, Young Avestan *-å̊*, Old Khotanese *-e* and the development of the Old Iranian *i*- and *u*-stems in Khotanese', *Studi e Saggi Linguistici* 60, 111–171.

EMMERICK, Ronald E., 1968. *Saka grammatical studies*, London.

FALK, Harry, 2015. 'Kushan rule granted by Nana: The background of a heavenly legitimation', in *Kushan Histories. Literary sources and selected papers from a Symposium at Berlin, December 5 to 7, 2013* (ed. H. FALK), Bremen, 265–299.

FUSSMAN, Gérard, 1974. 'Documents épigraphiques kouchans', *BEFEO* 61, 1–76.

FUSSMAN, Gérard, 2011. *Monuments bouddhiques de Termez, 1. Catalogue des inscriptions sur poteries*, Paris.

GERSHEVITCH, Ilya, 1954. *A grammar of Manichaean Sogdian*, Oxford.

GERSHEVITCH, Ilya, 1966. 'The well of Baghlan', *Asia Major*, n.s. 12, 90–109.

GERSHEVITCH, Ilya, 1979. 'Nokonzok's Well', *Afghan Studies* 2, 55–73.

GHIRSHMAN, Roman, 1953. 'Une intaille hephtalite de la collection de Mrs. Newell', *The Numismatic Chronicle and Journal of the Royal Numismatic Society* 13, 123–124.

GHOLAMI, Saloumeh, 2009. 'Ergativity in Bactrian', *Orientalia Suecana* 58, 132–141.

GHOLAMI, Saloumeh, 2014. *Selected features of Bactrian grammar*, Wiesbaden.

GIGNOUX, Philippe, 1972. *Glossaire des Inscriptions Pehlevies et Parthes*, London.

GÖBL, Robert, 1965. *Die drei Versionen der Kaniška-Inschrift von Surkh Kotal*, Wien.

HAIG, Geoffrey, 2008. *Alignment change in Iranian languages: A construction grammar approach*, Berlin–Boston.

HALFMANN, Jakob, 2023. '*Lād* "law": A Bactrian loanword in the Nuristani languages', *BSOAS* 86, 1–6.

HALFMANN, Jakob, Svenja BONMANN & Natalie KOROBZOW, 2024. 'A new reading of the Bactrian part of the Dasht-i Nawur Trilingual (DN 1)', *IIJ* 67/1, 5–51.

HARMATTA, János, 1969. 'Late Bactrian inscriptions', *Acta Antiqua Academiae Scientiarum Hungaricae* 17, 297–432.

HARMATTA, János, 1986. 'The Bactrian inscription of Ayrtam', in *Studia grammatica iranica. Festschrift für Helmut Humbach* (ed. R. SCHMITT & P. O. SKJÆRVØ), München, 131–146.

HENNING, Walter B., 1958. 'Mitteliranisch', in *Handbuch der Orientalistik I.41.1* (ed. A. SPULER), Leiden–Köln, 20–130.

HENNING, Walter B., 1960. 'The Bactrian inscription', *BSOAS* 23, 47–55.

HETTRICH, Heinrich, 1990. 'Der Agens in passivischen Sätzen altindogermanischer Sprachen', *Nachrichten der Akademie der Wissenschaften zu Göttingen, Philologisch-Historische Klasse*, Jahrgang 1990/2, 57–108.

HOFFMANN, Karl & Bernhard FORSSMAN, 1996. *Avestische Laut- und Flexionslehre*, Innsbruck.

HUMBACH, Helmut, 1960. *Die Kaniška-Inschrift von Surkh-Kotal: ein Zeugnis des jüngeren Mithraismus aus Iran*, Wiesbaden.

HUMBACH, Helmut, 1966. *Baktrische Sprachdenkmäler, Teil 1*. Mit Beiträgen von Adolf Grohmann, Wiesbaden.

HUMBACH, Helmut, 1976. 'Zu den baktrischen Materialien aus Kara-Tepe', *MSS* 35, 63–72.

HUMBACH, Helmut, 2003. 'The great Surkh Kotal inscription', in *Religious themes and texts of pre-Islamic Iran and Central Asia: Studies in honour of Professor Gherardo Gnoli on the occasion of his 65th birthday on 6th December 2002* (ed. M. MAGGI *et al*.), Wiesbaden, 157–166.

HUYSE, Philip, 2003. *Le -y final dans les inscriptions moyen-perses et la "loi rythmique" proto-moyen-perse*, Paris.

JAMISON, Stephanie W., 1979. 'The case of the agent in Indo-European', *Die Sprache* 25, 129–143.

JÜGEL, Thomas, 2015. *Die Entwicklung der Ergativkonstruktion im Alt- und Mittel-iranischen. Eine korpusbasierte Untersuchung zu Kasus, Kongruenz und Satzbau*, Wiesbaden.

KELLENS, Jean, 1974. *Les noms-racines de l'Avesta*, Wiesbaden.

KREIDL, Julian, 2021. 'The ablative case in Pashto', in *Proceedings of the 4th Conference on Central Asian Languages and Linguistics* (ed. Ö. ÖZÇELIK & A. K. KENT), Bloomington, 151–162.

KREIDL, Julian, 2024. 'On the Kushan Bactrian ablative-instrumental case', *Iran and the Caucasus*, 28/2, 206–228.

KUIPER, Franciscus B. J., 1978. *On Zarathustra's language*, Amsterdam–London–New York.

KÜMMEL, Martin, 2018. 'The morphology of Indo-Iranian', in *Handbook of Comparative and Historical Indo-European Linguistics*, *3* (ed. J. KLEIN), Berlin–Boston, 1888–1924.

LASS, Roger, 1976. 'Variation studies and historical linguistics', *Language in Society* 5, 219–229.

LAZARD, Gilbert, Frantz GRENET & Charles DE LAMBERTERIE, 1984. 'Notes bactriennes', *StIr* 13, 199–232.

LIVŠIC, Vladimir A., 1996. 'Baktrijskaja nadpis' na fragmente sosuda iz Kara-tepe (naxodka 1984 g.), in *Buddijskie kompleksy Kara-tepe v Starom Termeze. Osnovnye itogi rabot 1978–1989 gg* (ed. B. Ja. STAVISKIJ), Moskva, 168–172.

LIVŠIC, Vladimir A. & Irina T. KRUGLIKOVA, 1979. 'Fragmenty baktrijskoj monumental'noj nadpisi iz Dil'berdžina', in *Drevnjaja Baktrija, 2* (ed. I. T. KRUGLIKOVA), Moskva, 98–112.

LURJE [LUR'E], Pavel B., 2024. ''Neizvestnoe pis'mo' Baktrii: neopublikovannye materialy i novye interpretacii', *VDI* 84/2, 319–356.

LURJE, Pavel B., this volume. 'The "Unknown script" of Bactria: Unpublished materials and fresh interpretations', 131–176.

MACKENZIE, David N., 1994. 'Dates and dating', *Encyclopædia Iranica, Vol. VII, Fasc. 2*, 124–126.

MANCINI, Marco, 2012. 'Su alcune questioni di metodo in sociolinguistica storica: le *defixiones* sannite', in *Per Roberto Gusmani 1. Linguaggi, culture, letterature 2. Linguistica storica e teorica. Studi in ricordo* (ed. by G. BORGHELLO & V. ORIOLES), Udine, 239–271.

MANCINI, Marco, 2014. 'Testi epigrafici e sociolinguistica storica: le '*defixiones*' sannite', in *Le lingue dell'Italia antica oltre il latino: lasciamo parlare i testi (Atti del Convegno, 29 maggio 2007)* (ed. R. GIACOMELLI & A. ROBBIATI BIANCHI), Milano, 29–61.

MARICQ, André, 1958. 'La grande inscription de Kanishka et l'étéo-tokharien, l'ancienne langue de la Bactriane', *JA* 246, 345–440.

MARTINEZ, Javier & Michiel DE VAAN, 2014. *Introduction to Avestan*, Leiden–Boston.

MAYRHOFER, Manfred, 1992. *Etymologisches Wörterbuch des Altindoarischen*, I. Band, Heidelberg.

MICHETTI, Francesca, 2024. 'A Bactrian sound change (OIr. **š* > Bactr. *h*)', *ZDMG* 174/1, 173–190.

MMiii = ANDREAS, Friedrich C. & Walter B. HENNING, 1934. 'Mitteliranische Manichaica aus Chinesisch-Turkestan III', *SPAW*, Jahrgang 1934, 848–912.

MORGENSTIERNE, Georg, 1970. 'Notes on Bactrian phonology', *BSOAS* 33, 125–131.

PALUNČIĆ, Filip, Daniella PALUNČIĆ & B.T. MAHARAJ, 2023. 'Revisiting the Bactrian and Gāndhārī bilingual inscriptions from Dasht-e Nāwūr', *IIJ* 66, 333–383.

PANAINO, Antonio, 2017. 'The origins of Middle Persian *Zamān* and related words: A controversial etymological history', *Iran and the Caucasus* 21, 150–195.

PEYROT, Michaël, 2018. 'On the East Iranian genitive plural ending', *IIJ* 61,118–130.

POMPEO, Flavia & Maria Carmela BENVENUTO, 2011. 'Il genitivo in persiano antico. Un caso esemplare di categoria polisemica', *Studi e saggi linguistici* 49, 75–123.

RASTORGUEVA, Vera S. & Džoj I. ÈDEL'MAN, 2000. *Ètimologičeskij slovar' iranskix jazykov, tom 1. a–ā*, Moskva.

RASTORGUEVA, Vera S. & Džoj I. ÈDEL'MAN, 2003. *Ètimologičeskij slovar' iranskix jazykov, tom 2. b–d*, Moskva.

RTVELADZE, Èdvard V., 2018. 'Ajrtam – dinastijnyj xram carja Xuviški', *VDI* 78, 365–388.

SCARBOROUGH, Matthew J. C., 2021. 'Bactrian χþovo '(calendar) year, (regnal) year'', *JRAS* 31/3, 599–607.

SCHMITT, Rüdiger, 1991. *The Bisitun Inscriptions of Darius the Great: Old Persian text*, London.

SCHMITT, Rüdiger, 2014. *Wörterbuch der altpersischen Königsinschriften*, Wiesbaden.

SCHOUBBEN, Niels, 2023. 'The Iranian sound change **w-* > **γw-* in the Indo-Iranian borderlands and a new etymology for Gāndhārī and Sanskrit *guśura(ka)-*', *Iran and the Caucasus* 27, 285–298.

SCHWARTZ, Martin, 1974. 'Irano-Tocharica', in *Mémorial Jean de Menasce* (ed. Ph. GIGNOUX & A. TAFAZZOLI), Louvain, 399–411.

SIMS-WILLIAMS, Nicholas, 1975. 'A note on Bactrian syntax', *IF* 78, 95–99.

SIMS-WILLIAMS, Nicholas, 1981a. 'Notes on Manichaean Middle Persian Morphology', *StIr* 10, 165–176.

SIMS-WILLIAMS, Nicholas, 1981b. 'Some Sogdian Denominal Abstract Suffixes', *Acta Orientalia* 42, 11–19.

SIMS-WILLIAMS, Nicholas, 1985. 'A note on Bactrian phonology', *BSOAS* 48, 111–116.

SIMS-WILLIAMS, Nicholas, 1988. 'Bactrian language', in *Encyclopædia Iranica, Vol. III, Fasc. 4*, 344–349.

SIMS-WILLIAMS, Nicholas, 1989. 'Bactrian', in *Compendium Linguarum Iranicarum* (ed. R. SCHMITT), Wiesbaden, 230–235.

SIMS-WILLIAMS, Nicholas, 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* (ed. G. GNOLI & A. PANAINO), Roma, 275–296.

SIMS-WILLIAMS, Nicholas, 1994. 'Bactrian ownership inscriptions', *BAI* 7 (1993[1994]), 173–179.

SIMS-WILLIAMS, Nicholas, 1997. *New light on ancient Afghanistan. The decipherment of Bactrian. An inaugural lecture delivered on 1 February 1996*, London.

SIMS-WILLIAMS, Nicholas 1998. 'Further notes on the Bactrian inscription of Rabatak, with an appendix on the names of Kujula Kadphises and Vima Taktu in Chinese', in *Proceedings of the third European Conference of Iranian Studies held in Cambridge, 11th to 15th September 1995. Part 1: Old and Middle Iranian Studies* (ed. N. SIMS-WILLIAMS), Wiesbaden, 79–92.

SIMS-WILLIAMS, Nicholas, 2008. 'The Bactrian inscription of Rabatak: A new Reading', *BAI* 18 (2004[2008]), 53–68.

SIMS-WILLIAMS, Nicholas, 2009. 'The Bactrian fragment in Manichaean script (1224)', in *Literarische Stoffe und ihre Gestaltung in Mitteliranischer Zeit* (ed. D. DURKIN-MEISTERERNST *et al.*), Wiesbaden, 245–268.

SIMS-WILLIAMS, Nicholas, 2010. *Bactrian personal names*, Wien.

SIMS-WILLIAMS, Nicholas, 2011a. 'Remarks on the phonology of the Manichaean Bactrian fragment (M 1224)', in *Leksika, ètimologija, jazykovye kontakty, k jubileju Doktora filologičeskix nauk, Professora Džoj Iosifovny Èdel'man* (ed. E. K. MOLČANOVA), Moskva, 244–251.

SIMS-WILLIAMS, Nicholas, 2011b. 'Differential object marking in Bactrian', in *Topics in Iranian linguistics* (ed. A. KORN *et al.*), Wiesbaden, 23–38.

SIMS-WILLIAMS, Nicholas, 2012. 'Bactrian historical inscriptions of the Kushan period', *The Silk Road* 10, 76–80.

SIMS-WILLIAMS, Nicholas, 2013. 'Some Bactrian inscriptions on silver vessels', *BAI* 23 (2009[2013]), 191–198.

SIMS-WILLIAMS, Nicholas, 2015. 'A new Bactrian inscription from the time of Kanishka', in *Kushan Histories. Literary sources and selected papers from a Symposium at Berlin, December 5 to 7, 2013* (ed. H. FALK), Bremen, 255–264.

SIMS-WILLIAMS, Nicholas, 2017. 'The name of Kushan goddess Ομμα', in *Studia Philologica Iranica. Gherardo Gnoli Memorial Volume* (ed. E. MORANO *et al*.), Roma, 449–453.

SIMS-WILLIAMS, Nicholas, 2023a. 'A new collection of Bactrian letters on birchbark with an additional note by Frantz Grenet', *BAI* 31 (2022–2023[2023]), 135–144.

SIMS-WILLIAMS, Nicholas, 2023b. *The Ancient Letters and other Early Sogdian documents and inscriptions*, London [= AL].

SIMS-WILLIAMS, Nicholas, 2025. 'Bactrian in two scripts: Greek and Bactrian', *IIJ* 68/3, 185–214.

SIMS-WILLIAMS, Nicholas, forthcoming. 'Names and ghost-names in the colophon of the great inscription of Surkh Kotal'.

SIMS-WILLIAMS, Nicholas, this volume. 'The Bactrian inscription of Ayrtam: A minimal reading', 321–338.

SIMS-WILLIAMS, Nicholas & Joe CRIBB, 1995–1996. 'A new Bactrian inscription of Kanishka the Great', *SRAA* 4, 75–142.

SIMS-WILLIAMS, Nicholas & François DE BLOIS, 2005. 'The Bactrian calendar: New material and new suggestions', in *Languages of Iran: Past and present. Iranian Studies in memoriam David Niel MacKenzie* (ed. D. WEBER), Wiesbaden, 185–196.

SIMS-WILLIAMS, Nicholas & François DE BLOIS, 2018. *Studies in the chronology of the Bactrian documents from Northern Afghanistan*, Wien.

SIMS-WILLIAMS, Nicholas & Étienne DE LA VAISSIERE, 2015. 'A Bactrian document from Southern Afghanistan?', *BAI* 25 (2011[2015]), 39–53.

SKJÆRVØ, Prods O., 2009. 'Old Iranian', in *The Iranian languages* (ed. G. WINDFUHR), London, 43–195.

SKJÆRVØ, Prods O., 1985. 'Remarks on the Old Persian Verbal System', *MSS* 45, 211–227.

TREMBLAY, Xavier, 2005. 'Irano-Tocharica et Tocharo-Iranica', *BSOAS* 68/3, 421–449.

TURGUNOV, Baxadir A., Vladimir A. LIVŠIC & Èdvard V. RTVELADZE, 1981. 'Otkrytie baktrijskoj monumental'noj nadpisi v Ajrtame', *Obščestvennye nauki v Uzbekistane* 3, 38–48.

WEINREICH, Uriel, William LABOV & Marvin I. HERZOG, 1968. 'Empirical foundations for a theory of language change', in *Directions for Historical Linguistics* (ed. W. P. LEHMANN & Y. MALKIEL), Austin, 95–195.

YOSHIDA Yutaka, 2009. 'Sogdian', in *The Iranian languages* (ed. G. WINDFUHR), London, 279–335.

Notes on Wakhsh and Rām-sēt in the Bactrian documents*

MIYAMOTO Ryoichi 宮本亮一

In the Bactrian documents, there are eleven contracts in which the god Wakhsh or the god Rām-sēt is mentioned as one of the witnesses. The former is a well established deity in Tukhāristān since ancient times, but the latter is a Sogdian god that is attested in the local documents only as late as the second half of the 7th century CE. Strikingly, among the numerous contracts for worldly transactions, besides a few exceptions, solely the Bactrian ones mention a deity as a witness. In this paper, I will collate the information on the two gods and examine some aspects related to their worship in the pre-Islamic period.

1. Introduction

Several world religions that originated in India or Western Asia, such as Buddhism, Manichaeism, and Christianity, spread to Eastern Asia one after another, but the situation of the local religions and cults in Central Asia, one of their important transit areas, is not yet fully unraveled. The situation is especially obscure in its western part, namely Tukhāristān (in Bactrian τοχαραστανο/τοχοαραστανο)[1] and Sogdiana. However, the discovery of more than 150 Bactrian documents since the mid-1990s has allowed us to improve our understanding of the situation in Tukhāristān. This paper organizes the information concerning the two deities mentioned in the Bactrian contracts so far published, i.e. Wakhsh and Rām-sēt, and discusses related topics in the socio-cultural history of Central Asia.

* This paper is a slightly revised and shortened version of my previous Japanese article (Miyamoto 2023b). I would like to express my deepest gratitude to Dr. Ching Chao-jung for giving me a great number of valuable comments on this paper and for improving my English. It should be noted that I am only responsible for all errors remaining in this paper. This research was supported by JSPS KAKENHI Grant Number 24K04275.

[1] The toponym, Tukhāristān, roughly corresponds to northern Afghanistan, southern Uzbekistan, and southern Tajikistan. The area was called Bactria, but that term was already obsolete in BCE, and the area came to be called “Tukhāristān” after the name of a nomadic tribe, the Tokharoi, who invaded it. The oldest known mention of Tukhāristān in the Bactrian sources is found in an inscription from the time of Kanishka (Sims-Williams 2015b).

10.29091/9783752003635/006

2. Worship of Wakhsh in Central Asia: An overview

Deifications of important rivers are well known all over the world.[2] In Central Asia, the Amu Darya, or the Wakhsh (Oxus) River, may have been worshipped during the Achaemenid period (mid-6th century – early 4th century BCE).[3] Among the collected findings, this possibility is not only suggested by a gold finger ring in the Oxus Treasure, which bears an inscription in Aramaic script, but also by seven personal names from the newly discovered Aramaic documents in the late Achaemenid period that contain the element *wḫšw*.[4] As to the inscriptions that have a clearer archaeological context, those dedicated to the Oxus River, inscribed on several objects excavated from the temple site at Takht-i Sangīn (early 3rd cent. BCE – 4th cent. CE, southern Tajikistan; see Map 1), demonstrate that this river was a deified object of worship there.[5] At Aï Khanoum (early 3rd cent. BCE – mid-2nd cent. BCE), Greek inscriptions on potsherds attest personal names with an element derived from the name of the river Oxus.[6] Moreover, there is one specimen of a Huvishka coin with an image of Wakhsh (as a deity) on the reverse.[7] It

[2] A typical example in the Indo-Iranian context is the Sarasvatī, an extremely important river for the Indo-Aryans, who gradually advanced from present-day Afghanistan and its surrounding regions to the Indian subcontinent (Yamada 2011). A connection between the goddess Anāhitā in the Avesta and the Wakhsh River has also been pointed out (Oettinger 2001).

[3] The correspondence between the ancient Wakhsh and present-day rivers is an unsolved question. While Frantz Grenet identifies the ancient river with the present-day Wakhsh River (Grenet & Rapin 2001: 80–81; Grenet 2006: 327–328), some scholars doubt this identification. See, for example, Kuwayama 2016: 126 n. 58; Falk 2018: 10–13.

[4] Images of the ring and the possible readings of the inscription can be seen on the British Museum website (https://www.britishmuseum.org/collection/object/W_1897-1231-105, accessed on 22 October 2024). As to the Aramaic documents, see Naveh & Shaked 2012: 57–60.

[5] At Takht-i Sangīn, the temple area has been excavated since the 1970s (Litvinskij & Pičikjan 2000; Litvinskij 2001; Litvinskij 2010). After an interruption due to the civil strife in Tajikistan, the excavation of the urban area extending north and south of the temple began in 1998, and the temple area has been excavated again from 2004 onward (Drujinina & Boroffka 2006; Drujinina *et al.* 2009; Druzhinina *et al.* 2010; Druzhinina *et al.* 2011; Družinina *et al.* 2019; Druzhinina 2016; see also Bernard 1994; Litvinskii & Pichikian 1996; Grenet 2005; Shenkar 2011: 120b–123b; Shenkar 2012). The structure of this temple, in which the cloister encloses the inner sanctuary, was later introduced into Buddhist temples in Central Asia and Northwest India (Iwai 2019). For the inscriptions from the temple, see Rougemont 2012: 196–199, 274–276. Concerning Wakhsh, see also Boyce & Grenet 1991: 179–181; for its iconography, see Abdullaev 2013; and for Iranian personal names including the name of the deity, see Schmitt 2017.

[6] Grenet 1983: 376–378, who gives the same interpretation for some personal names in the biography of Alexander the Great. In this regard, see also Schmitt 2017.

[7] Rosenfield 1967: 92. For the image of the coin, see https://www.britishmuseum.org/collection/object/C_1893-0506-21 (accessed on 22 October 2024).

should be added, of course, that several theophoric names with Wakhsh, such as Wakhsh-burd (οαχβοβορδο) and Wakhsh-gul (οαχβογολο), are attested in the Bactrian documents (Sims-Williams 2010: nos. 321–325, 547).

It seems very likely that the cult of Wakhsh crossed the boundary of Tukhāristān. It has been suggested that a temple dedicated to this deity may have existed near Samarkand and that a dragon-crowned divinity depicted in the Penjikent murals may be identified with Wakhsh.[8] Recently, a sealing with an inscription of Wakhsh (*(w)xwšw*) was found at the site of Kafir-kala near Samarkand (Begmatov 2022). In Sogdian texts, a significant number of Sogdian personal names containing the element *(w)xwšw* have been identified.[9] There is also sporadic onomastic evidence with regard to the situation in Khwarazm.[10] In addition, there was a local festival named after *Wakhsh* according to al-Bīrūnī (Sachau 1878: 225).

Map 1. Important sites and regions in Tukhāristān.

[8] Lurje 2004: 209; Marshak & Raspopova 1991: 194, pl. LXXIV.

[9] Lurje 2010: nos. 211(?), 219, 1307(?), 1355, 1356, 1363–1372, 1373(?), 1454(?); Lurje 2022: 451b–452a.

[10] Livšic 1984: 258; Livšic 2004: 190b–191a.

3. Wakhsh as a deity in the Bactrian documents

As introduced above, the worship of Wakhsh was evident in several parts of Central Asia. However, due to the scarcity of relevant materials, little is known about its actual practice. The information provided by the Bactrian documents, therefore, is valuable, albeit limited. As an example, let us look at the opening of the oldest document mentioning Wakhsh, i.e. the land purchase contract (L) written in Warnu.[11]

Document L, lines 1–4

> '(It was) the year 379, the month Ab,[12] when (this) sealed document, this purchase contract, was written here in the province[13] of Warnu, **with the cognizance of the god Wakhsh whom we worship here in Warnu** (αζδδηβιδο βαγο ι οαχþο κιδο μαλαβο οαρνο ασπισαμο), and in the presence of Torman Aspandagan, and in the presence of Sisan Sibukan, and in the presence of Samb Abkharagan, and also in the presence of the other citizens, ...' (Sims-Williams 2012a: 58–59)

In most of the eight cases (Doc. L, Nn, O, Tt, U, Uu, V, W, see Table 1), including this document, the god Wakhsh (βαγο ι οαχþο) is mentioned at the beginning of the list of witnesses. The locations where these contracts were written down show that this deity was worshipped in several places in southern Tukhāristān, such as Warnu, Gūzgān, and Rōb, in the period from the early 7th century to the mid-8th century. It is noteworthy that it is referred to with various epithets, although they are not seen in the example above. An interesting example of such epithets is in Document Nn, because the appellation there has common elements with the epithet in Document P and Q applied to the Sogdian god Rām-sēt, on which see below in §4.

---

[11] Hereafter, the underlining and Greek letters in the translations are added by the author.

[12] This date in the Bactrian era corresponds to 602 CE January/February. For the details of the so-called Bactrian era, see Sims-Williams & de Blois 2018. While they considered the first year of the era to be 223 CE, i.e. the first regnal year of Ardashir I, Nicholaus Schindel and Étienne de la Vaissière proposed that 227 CE, the first year of the second century in the Kanishka era, was the first year of the Bactrian era (Schindel 2011; de la Vaissière 2018). I keep to the former conversion in this paper, because it seems to me that the accession year of a ruler is more appropriate for the first year of a new era.

[13] Here I have made a slight modification to Sims-Williams' translation of þαρο/þαυρο, which he mostly translates as 'city' (Sims-Williams 2007: 284a). According to a detailed study by Rika Gyselen, the Sasanian administrative divisions were categorised from largest to smallest as 'région', 'province', and 'canton' (for further information, see Gyselen 2019). As the Middle Persian term *šahr* corresponds to the category of 'province', it appears to me that the Bactrian word þαρο/þαυρο should better be translated as 'province' (Miyamoto 2019: 166 n. 7).

Doc.	Date	Type	Location	Deity	Epithet of deities
L	Jan./Feb. 602	contract for the purchase of an estate	Warnu (Kadagstan)	Wakhsh	the wonderful, the granter of favours (and) granter of wishers, whose renown have reached the whole earth
Nn	27 Jan. 659	contract for the purchase of an estate	Lizg (Guzgan)	Wakhsh	the king of gods
O	Aug./Sept. 662	undertaking to keep the peace	Stof (Guzgan)	Wakhsh	the granter of favours (and) granter of wishers, the wonderful
P	4 Jan. 669	contract for the purchase of a male slave	Samingan	Rām-sēt	the granter of favours (and) granter of wishers, the wonderful
Q	30 July 671	contract for a loan of money	Samingan	Rām-sēt	
Ss	31 May 698	contract for a loan of 100 drachmas	Amber (Guzgan)	Rām-sēt	the king of gods, the granter of favours (and) granter of wishes
Tt	23 July 705	deed of gift	Lizg (Guzgan)	Wakhsh	
U	712/713	contract for the lease of a vineyard	Madr (Rob)	Wakhsh	the king of gods, woderful (and) renowned
Uu	9 Sept. 722	undertaking to keep the peace	Lizag (Guzgan)	Wakhsh	the granter of favours (and) fulfiller of wishes, great (and) wonderful
V	May/June 729	contract for the purchase of an estate	Rizm (Rob)	Wakhsh	the granter of favours (and) fulfiller of wishes, great (and) wonderful
W	Aug./Sept. 747	contract for the purchase of an estate	Gandar (Rob)	Wakhsh	the wonderful, the granter of favours (and) granter of wishers, whose renown have reached the whole earth

Table 1. List of Bactrian contracts with a mention of Wakhsh or Rām-sēt as a witness.

Document Nn, lines 1–7

> '(It was) the year 436, the month Ab,[14] the day Wad, when (this) purchase contract was written here in the province which they call Lizg, with the cognizance of the god Wakhsh, **the wonderful, the granter of favours (and) granter of wishes, whose renown have (sic!)**[15] **reached the whole earth** (οαρσοχοανδδιγο λαδοιανο λαδοαγαλγο ακιδδηιο ναμο αβ(ο) υαρογο ζαμιγο βοοαδογινδο), and with the cognizance of Kulzan, the leader of the people of Lizg, and also in the presence of the other freemen (and) witnesses who have written (their names as) witnesses (to this) sealed document, (this) purchase contract, on the other signed document[16] and (who) have placed (their) seals hereupon. [...]' (Sims-Williams 2012a: 74–75)

In Document Tt, Wakhsh is also referred to as βαγανο þαυο 'king of the gods', although it is unclear whether this epithet is sufficient to testify the supreme status of this god in the local pantheon in Tukhāristān:

Document Tt, lines 1–4

> '(It was) the year 483, the month Hurezhn and the day Risht,[17] when (this) sealed document, (this) familial[18] contract (concerning) a trust(?) (in return) for service rendered, was written here in Lizg, at the court of the fortress, in the presence of the god Wakhsh, the king of gods, the granter of favours (and) granter of wishes, who has (his) seat in Lizg (κιδαβο λιζγο νιþλμινδο) [...]' (Sims-Williams 2012a: 104–105)

The phrase 'who has (his) seat at Lizg' suggests the existence of a place of worship dedicated to the deity in the early 8th century.

4. Rām-sēt: A Sogdian god in Tukhāristān

Besides Wakhsh, another deity appears as a witness in the Bactrian contracts, namely Rām-sēt. At present, this deity is mentioned in three documents: P, Q, and Ss. The first two were written at Samingān to the north of Rōb, and the last one at Amber (present-day Sar-i Pul) in Gūzgān (see Table 1 and Map 1). Let us look at Document Q, a well-preserved one for a loan of silver drachmas:

[14] 27 January 659 (Sims-Williams & de Blois 2018: 84).

[15] The "sic!" is in Sims-Williams' original translation, rendering the number disagreement between ναμο (sg.) and βοοαδογινδο (pl.).

[16] For the word πιδοναμο 'signed document', see also Yoshida 2013b: 158, who prefers to translate it as 'copy'.

[17] 23 July 705 (Sims-Williams & de Blois 2018: 84).

[18] For the word παλοβωστιγο, see Sims-Williams 2015a: 45.

Document Q, lines 1–5

> '(It was) the year 449, the month Second New-year, the day Din,[19] when (this) loan-contract was written here in the district of Samingan, at Marogan, the market of the khars of Rob, **with the cognizance of the god Ram-set, the granter of favours (and) granter of wishes, the wonderful, whose worship (and) seat here at Marogan, at the market** (εζδδηβιδο βαγο ραμοσητο λαδοιανο λαδοαγαλγο οαρσοχοανδο ακιδδηιο σπασο νιϸαλμο μαλαβο μαρωγανο αβο οασαρινδο), with the cognizance of Zhun-lad Shaburan, the tapaghlïgh iltäbir of the qaghan prosperous in glory, khar of Rob, (and) in the presence of the tarkhan, the son of Khusaru the tarkhan, and in the presence of Deb-raz, senior overseer of the market of the khars of Rob, and also in the presence of the other freemen who were present there amongst (them) and (who) bear witness concerning this matter. [...]' (Sims-Williams 2012a: 88–89)

Rām/Rāman (< **rāman* 'peace, tranquillity') is a Zoroastrian deity known as the god of peace and as the patron of the twenty-first day.[20] The Bactrian form ραμοσητο (< ραμο + σητο 'spirit') is probably borrowed from Sogdian and the deity was widely worshipped among the Sogdians.[21] A good example of its popularity in the Sogdian community is the story of 'the martyrdoms of St. George' in the Christian Sogdian text E23 (9th–10th cent., found in Turfan, but presumably translated in Semirech'e), in which Heracles of the Athenian Greek version is rendered as Rām-sēt.[22] Another example is Or. 8210/S. 367 *Shazhou Yizhou dizhi canjuan* 沙州伊州地志殘卷, a Chinese geographical text written in the late 9th century. This mutilated scroll records a Zoroastrian shrine dedicated to Rām (*Alan* 阿覽, Middle Chinese *â-lâm* following Karlgren 1957) near the oasis of Yizhou (Hami, to the east of Turfan).[23] In Dunhuang, there was a town or settlement of Sogdian immigrants which was called An 安, and a Zoroastrian temple was located there. It is well known that in Early Mediaeval China, An 安 was the typical Chinese surname attributed to the Sogdians from Bukhara and their descendants, but, according to Ikeda (1965: 50–52), with regard to this town's name, one might consider if it is named after the deity, because the Chinese character literally means 'peaceful, tranquil'.

At any rate, it seems likely that the worship of Rām was brought to Tukhāristān as a result of the migration of the Sogdians, and places of its worship were established in the

[19] 30 July 671 (Sims-Williams & de Blois 2018: 84).

[20] There is a Yašt named after Rām in the Avesta, but its hymns are dedicated to Vāyu. The reason for the discrepancy between its title and content is not clear (Malandra 2014).

[21] Sims-Williams 2010: nos. 388, 436; Lurje 2010: nos. 998, 1003.

[22] In the same text, Apollo is rendered as Mahākāla. These examples indicate that the Sogdians replaced the original divinities with those familiar to them in their translations. For the details on the possible area where E23 was translated and the rendering of the divine names, see Yoshida 2022a.

[23] Yoshida 1994: 392b; cf. Tang & Lu 1986: 41.

areas where their settlements were developed. As in the case of Wakhsh, the phrase 'whose worship (and) seat here at Marongan' in Document Q implies the existence of a place of worship dedicated to Rām-sēt.

Activities of Sogdian immigrants or their descendants were indeed observable in Tukhāristān. Among the three Bactrian documents referring to Rām-sēt, Document Ss (written in Amber) mentions that some merchants there, including a man from Bukhara named [...]-tis, son of Berd. He borrowed 200 silver drachmas from another merchant (Sims-Williams 2012a: 96–97). Berd (βηρδο) is considered to be a Sogdian, thus the document suggests that there was a Sogdian trading post at Amber (Sims-Williams 2010: nos. 81, 581). It is noteworthy that the place is described by the 10th-century Persian geographical source *Ḥudūd al-'Ālam* as "a good and prosperous town, the residence of merchants and the emporium of Balkh" (Minorsky 1970: xxii).

Another document from the end of the 7th century (S), although not mentioning Rām-sēt, indicates that some Sogdians had certainly settled in southern Tukhāristān by this period. Written in Gūzgān in 693, it mentions Farn-agad's sons Potan and Miyar (πωτανο οδο ... μιαρο φαρνοαγαδο πορανο, cf. Sims-Williams 2012a: 94–95). The personal names of this father and his sons are considered to be Sogdian, and the place of their origin, Bunsuglig (βονοσογολιγο), most likely means 'land of the Sogdians'.[24] Furthermore, one of the witnesses named Sit (σιτο) is probably a Sogdian as well (Sims-Williams 2010: no. 436).

The aforementioned documents all date from the 7th century, but presumably a considerable number of Sogdians had settled in Tukhāristān before then. Their footprints were particularly clear in Gūzgān. Document ag also sheds a light on this process (Sims-Williams 2012a: 156–157). This document, which appears to be a list of wine production, mentions 30 personal and family names. Those associated with the Sogdians include Sugdukan (σογδοκανο), Yawar (ιαοαρο), and Sēt (σητο).[25] The document is undated, and the place of writing is unknown, but based on various criteria it is considered to have been written after the early 8th century (Sims-Williams & de Blois 2018: 66). Pottery analysis has already shown that Sogdian culture extended into the northern part of Tukhāristān during the 7th and 8th centuries, and it has been suggested that exchanges between the Balkh region and Sogdiana may have flourished at this time (Iwai 2004: 14 n. 21).

As previously mentioned, a Sogdian man from Bukhara appears in Ss, and it is intriguing that coins with the inscription *r'mcytk βγγ* 'the god Rām-čētē' were issued in this oasis (Lurje 2010: no. 1003). According to al-Bīrūnī, a festival was held on the 28th

[24] As to the personal names, see Sims-Williams 2010: nos. 490, 384, 249. On the interpretation of Bunsuglig, see Lurje *apud* Sims-Williams 2010: no. 249.

[25] Sims-Williams 2010: nos. 441, 175; Sims-Williams & de Blois 2018: 131.

day of the first month in a fire temple in Bukhara dedicated to Rām-sēt.[26] Considering such information, it is probable that some Sogdians migrated to Tukhāristān via the route between Bukhara and Merv, or the route from Bukhara via Nakhshab and other places, crossing the Amu Darya at Kalif or Zamm. Furthermore, there must have been other travellers who used the route via the Iron Gate. In any case, the appearance of Rām-sēt in southern Tukhāristān in the second half of the 7th century was undoubtedly related to the Sogdian migration and settlement.[27]

According to Yoshida, a merchant named *Tuhuoluo Moseduo* 吐火羅磨色多 recorded in a Chinese travel document found in Turfan (64TAM29: 17a, 95a, 108a, 107, 24, 25) was of Sogdian origin, possibly coming from a Sogdian settlement in Tukhāristān.[28] In the west, a pottery sherd with Sogdian inscriptions has been found in Merv, and later Arabic sources record that the Sogdians were engaged in moneylending and other activities in the city.[29] Furthermore, Sogdian inscriptions also began to appear on coins issued in Khwarazm in the late 8th century.[30]

5. Wakhsh and Rām-sēt: Witnesses of transactions

To find eight contracts having Wakhsh or Rām-sēt as one of the witnesses among the dozens of published ones is statistically a very significant ratio. We do not know in detail in what places they were worshipped, but in two of the above-mentioned documents, Tt and Q, it is stated that there were "seats" of Wakhsh and Rām-sēt in Warnu and in Lizg, respectively, and the same word also appears in another document (Doc. Uu, cf. Sims-Williams 2012a: 112–113). The word νιϸαλμο 'seat' also appears in the Surkh Kotal inscription (SK4), where it means 'seat' of the gods.[31] Thus, the evidence from the Bactrian documents shows that there were indeed places of worship for both deities in different parts of Tukhāristān, which served as a space to conclude a contract.

Regarding the relationship between the place of worship and the conclusion of a contract, one may evoke a similar practice in Sogdiana. Two Sogdian marriage contracts

[26] Sachau 1878: 221; Shenkar 2017: 206a. The *Tārīkh-i Bukhārā* mentions a village named Rāmush where there was a fire temple constructed by Kay Khusraw (Frye 1954: 17).

[27] Another notable aspect of influence from Sogdiana to Tukhāristān is the use of indentation as a graphic means to express modesty. This practice, introduced from China to Sogdiana, appears in a Bactrian document (Doc. Y) written in the late 8th century (Yoshida 2013a: 48; Yoshida 2019: 49 n. 52). With regard to the possible characteristics of this document, see Miyamoto 2023a: 94.

[28] Yoshida 2013a: 61. See also Yoshida's remark on the author of the Buddhist manuscripts written in Bactrian and unearthed from the Turfan region (Yoshida 2022b: 51–53).

[29] Livshits 2015: 54–55; Lurje 2010: no. 1255; cf. de la Vaissière 2016: 245–248.

[30] Henning 1957: 57; Lurje 2010: no. 1160; cf. de la Vaissière 2016: 229–231.

[31] Gershevitch 1979; Sims-Williams 2012c: 78b–79a.

(Nov. 3 and Nov. 4) from the early 8th century state that these documents were 'made in the Place of Foundations in the presence of its chief Wakhushukān, son of Varkhumān (*'krty ZNH βwnty-n'k 'st'ny pt'yc ZKn xxy-št wxwšwk'n ZKn βrxm'n BRY*)'.[32] If *xwy-št* 'chief' refers to a (Zoroastrian?) priest as Livshits (2015: 34) supposed, *βwntyn'k 'st'n* 'Place of Foundation' would have been some kind of place of worship and the situation in Sogdiana would be similar to that in Tukhāristān.[33]

The sealing of documents is also important with regard to this issue. In principle, the Bactrian contracts were sealed, and sometimes the names and titles of the persons who sealed them as witnesses are written on the reverse side of the document and correspond with the inscriptions of the stamp impression. It is remarkable that the names Wakhsh (L, Uu) and Rām-sēt (P, Q) appear among these signatures on the reverse side (Sims-Williams 2012b: pls. 34a, 84b, 54a, 58b). Unfortunately, among the impressions that are decipherable from the sealing of Bactrian documents, no one shows the names of these deities. However, there is one seal in Bactrian (i.e. Sig 9 in Bivar 1968: pl. II/2) bearing the name of Wakhsh.[34] Judging from what we have seen so far, it is conceivable that some contracts were concluded at the religious spaces dedicated to Wakhsh or Rām-sēt, and that the priests serving there affixed the seals with the name of the deity.[35] It is worth noting here that in other regions of Central Asia, the act of concluding a contract with divine witnesses was extremely rare. The exceptions include a few contracts written in Sogdian and in Old Uyghur.[36] The Sogdian one is the above-mentioned marriage contract (Nov. 4), in which the bridegroom swears an oath to the bride's guardians "in the presence of the god Mithra" (*ZKn βγy ZY ZKn myδr' nβ'nty*);[37] as to the ones written in Old Uyghur, *tört maxarač tngri-lär* 'Four Mahārāja Gods' and *yiti äkä balṭïz tngrim-lär* 'Seven Sister Goddesses' are mentioned as witnesses (cf. EM01 and WP02 in

[32] In this part, Nov. 3 and Nov. 4 have almost identical texts. The text from Nov. 4 and its translation quoted here are from Livshits 2015: 25, 27. For the texts and translations of these two documents, see also Yakubovich 2006.

[33] With regard to the Sogdian phrase *βwntyn'k 'st'n*, see Yakubovich 2006: 322–323; Livshits 2015: 34.

[34] Cf. Sims-Williams 2010: no. 475; Lerner & Sims-Williams 2011: 56.

[35] Divine seals may have something to do with the validity of documents. On this point, see the relationship between the type of seal and the validity of the document in Middle Persian sources (Macuch 2007: 202). I would like to express my sincere gratitude to Dr. Ogihara Hirotoshi for informing me about this point.

[36] As far as I have been able to ascertain, it appears that this practice is not to be observed in Gāndhārī, Khotanese, Tumshuqese, Kuchean, and Old Tibetan contracts (see Burrow 1940 and Baums & Glass, 2002– ; Skjærvø 2002; Ogihara & Ching 2014; Ching & Ogihara 2012; Takeuchi 1995, respectively). Nor is it found in a Parthian contract drawn up in Western Asia (Haruta 2001). However, I could only search the translations and I may perhaps have overlooked some examples.

[37] Yakubovich 2006: 314, 324; Livshits 2015: 28, 35–36.

Yamada 1993: 130–131, 136–137, respectively). However, it is likely that the gods mentioned in the Old Uyghur contracts resulted from East Asian or Buddhist influences and are therefore not directly relevant to the issues discussed here.[38]

The common custom of naming divine witness in Tukhāristān and Sogdiana suggests that it might be a western tradition introduced in the two regions in the Achaemenid period, although there is only one example in the latter region. The background is explainable: during the Achaemenid period, Sogdiana was under the jurisdiction of the satrap of Bactria, and two regions were considered to be politically united (Naveh & Shaked 2012). Given that Mesopotamian traditions dating to the second millennium BCE have been preserved in the formulae of Sogdian and Bactrian letters, and that the letter format has also inherited traditional Achaemenid forms, the transmission of the practice of divine witness may be placed in the same historical context.[39]

Regarding the socio-cultural context in which Wakhsh and Rām-sēt are mentioned as witnesses in the Bactrian contracts, the characteristics of these divinities must be taken into consideration. First, regarding Wakhsh, it should be noted that this deity is a river god. A recent study of Vedic texts shows that in ancient India, there was a notion that water, as a divine being circulating perpetually through the universe, guaranteed oaths and contracts (Sakamoto [Gotō] 2008: 94–93). Furthermore, it is known that in ancient Indian rituals, people swore an oath in front of water in a jar or by touching the water (Gotō 2008: 91b–92b). Although the Bactrian documents provide no information about the notion of water in the local society of Tukhāristān, it would not be surprising if similar notions and customs existed in this region, as customs similar to Indian rituals were widespread in the Indo-European-speaking areas in the ancient times. The personal name Wakhshukān (*wxwšwk'n*) in the Sogdian marriage contracts mentioned above contains an element of Wakhsh. In addition, the name of the scribe's father in the contracts, Akhushfarn (*'xwšprn*), also has the same element.[40] The facts that the chief of a place of worship where the documents were drawn up and a family member of the

[38] For the (Seven) Sister Goddesses, see Fen & Tenišev 1960: 148; Matsui 2011: 40. I am deeply grateful to Prof. Matsui Dai for sharing information on the Old Uyghur contracts and sending me these papers.

[39] On the formulae of the Bactrian letters, see Sims-Williams 1996; Sims-Williams 2025: 55–86, and on the format, see Yoshida 2013a: 47–48. Concerning the known instances of deities being listed as witnesses in ancient Mesopotamian legal documents, see Sheikh 2023: 45–48. It should be noted, however, that not only Rām-sēt but also Wakhsh appears in Bactrian contracts only from the 7th century onwards, although the fact that there are only about 30 specimens of extant Bactrian contracts, despite their dating from the early 4th century to the mid-8th century, suggests that this phenomenon may be accidental.

[40] For this name, see Yakubovich 2006: 323.

scribe have such names might suggest that a connection between water and contract also existed in Tukhāristān and Sogdiana.

On the other hand, while it is easy to imagine Sogdian merchants praying to the god Rām-sēt for their safe journeys, there appears to be no direct link between the characteristic of this god of "peace" and the making of a contract. However, it should be noted that in the Avestan sources, Rām (Rāman) has a close connection with Mithra, the deity symbolising contract, and that there may be a few onomastic examples of the compound of the two deities in Middle Persian and Bactrian.[41]

Not being an art-historian, I am not able to comprehensively collect the iconographical information on the two deities. As far as I know, no image has so far been successfully identified with Rām-sēt. On the other hand, Wakhsh is depicted on the above-mentioned coin of Huvishka as well as on the seal from Kafir-kala; in both cases, the god holds a long branch in one hand and an underwater creature or cornucopia in the other. It should be noted, however, that a male deity may be represented on the coin and a female one on the seal. In addition to the aforementioned dragon-crowned divinity depicted in the Penjikent wall paintings, the association of this divinity with the horse has also been suggested. Michael Shenkar points to the possibility of Wakhsh being depicted in the Penjikent and Bunjikat wall paintings, referring to the horse represented on the stone panels preserved in the MIHO Museum and the description in the 9th-century Chinese text *Youyang zazu* 酉陽雜俎, etc.[42] Incidentally, although images representing the two deities must have been worshipped, the word πιδογιρβο 'image' in the Rābatak inscription is not so far attested in the Bactrian documents from the fourth century onwards.[43]

6. Conclusion

This paper treats the worship of Wakhsh and Rām-sēt, the only divinities mentioned in the Bactrian documents with a special function of witnessing or granting a contract. If new contracts mentioning other deities are discovered in the future, their role will need

[41] Peschl 2022. I would like to express my deep gratitude to Dr. Benedikt Peschl for sending me his paper. Theophoric names with Mithra are popular in Tukhāristān (Sims-Williams 2010: 98, 229, 253–255, 257, 259–263, 266, 267).

[42] Shenkar 2014: 128–131. He also suggests that the deity depicted in the wall painting of Dokhtar-i Nōshirwān (northern Afghanistan) might also be Wakhsh. For the Chinese source, see Drège & Grenet 1987. With regard to the horse without rider depicted on the stone panels in the MIHO Museum, Frantz Grenet prefers to identify it with the horse of the god Tishtrya (Grenet 2020: 324). The stone panels can be seen at http://www.miho.or.jp/booth/html/artcon/00000432.htm (accessed on 22 October 2024).

[43] For this Bactrian word, see Sims-Williams & Cribb 1996: 93; Sims-Williams 2008: 64b.

to be reconsidered together with new data. I hope that this paper may serve as a stepping stone to a deeper understanding of the religious practice in Tukhāristān.

Finally, one should not ignore that two Bactrian contracts (V and W) for the purchase of estates provide insight into the religious situation in Tukhāristān, where various religions coexisted. These contracts, written in 729 and 749, respectively, state that the purchaser of land can establish a βαυαρο 'monastery', βαγολαγγο 'temple', λαχμιγο 'place of burial', or λαχβατανιγο 'place of cremation'.[44]

References

ABDULLAEV, Kazim, 2013. 'Images d'un dieu-fleuve en Asie centrale: l'Oxus', *CRAI* 157/1, 173–192.

BAUMS, Stefan & Andrew GLASS, 2002–. *Catalog of Gāndhārī texts.* https://gandhari.org/catalog (accessed on 22 October 2024).

BEGMATOV, Alisher, 2022. 'Cross cultural exchange across Eurasia as reflected in the sealings from Kafir-kala in Samarkand', in *Cultures in contact: Central Asia as focus of trade, cultural exchange and knowledge transmission* (ed. Ch. BAUMER & M. NOVAK), Wiesbaden, 193–203.

BERNARD, Paul, 1994. 'Le temple du dieu Oxus à Takht-i Sangin en Bactriane: temple du feu ou pas ?', *StIr* 23/1, 81–121.

BIVAR, A. D. H., 1968. *Kushan and Kushano-Sasanian seals and Kushano-Sasanian coins: Sasanian seals in the British Museum*, London.

BOYCE, Mary & Frantz GRENET, 1991. *A history of Zoroastrianism III. Zoroastrianism under Macedonian and Roman rule*, Leiden.

BURROW, Thomas, 1940. *A translation of the Kharoṣṭhī documents from Chinese Turkestan*, London.

CHING Chao-jung & OGIHARA Hirotoshi, 2012. 'A Tocharian B sale contract on a wooden tablet', *JIAAA* 5 (2010[2012]), 101–127.

DREGE, Jean-Pierre & Frantz GRENET 1987. 'Un temple de l'Oxus près de Takht-i Sangin, d'après un témoignage chinois du VIIIe siècle', *StIr* 16/1, 117–121.

DRUJININA [DRUŽININA], A. P. & N. R. BOROFFKA, 2006. 'First preliminary report on the excavation at Takht-i Sangin 2004', *BMM* 6, 57–69.

DRUJININA [DRUŽININA], A. P., H. INAGAKI, T. HUDJAGELDIEV & F. ROTT, 2009. 'Excavations of Takht-i Sangin city, territory of Oxus temple, in 2006', *BMM* 9, 59–84.

DRUZHININA [DRUŽININA], A. P., 2016. 'The structure of the urban site Takht-i Sangin and its vicinity', *BMM* 16, 53–101.

[44] Sims-Williams 2012a: 120–121, 130–131. For the burial culture in Tukhāristān, see Iwai 2017.

DRUZHININA [DRUŽININA], A. P., T. U. KHUDZHAGELDIYEV & H. INAGAKI, 2010. 'Report of the excavations of the Oxus temple in Takht-i Sangin settlement site in 2007', *BMM* 10, 63–82.

DRUZHININA [DRUŽININA], A. P., T. U. KHUDZHAGELDIYEV & H. INAGAKI, 2011. 'The result of the archeological excavations at Takht-i Sangin in 2008', *BMM* 11, 13–30.

DRUŽININA, A. P., T. U. XUDŽAGELDIEV & X. INAGAKI, 2019. 'Otčet o raskopkax na gorodišče Taxti-Sangin v 2009–2010 gg. Raboty na ploščadi xrama Oksa v raskopax N^{o} 17 i N^{o} 18', *BMM* 19, 87–115.

FALK, Harry, 2018. 'The Five Yabghus of the Yuezhi', *BAI* 28 (2014[2018]), 1–43.

FEN Czja-šen [FENG Jiasheng] & È. TENIŠEV, 1960. 'Tri novyx ujgurskix dokumenta iz Turfana', *Problemy Vostokovedenija* 1960/3, 141–149

FRYE, Richard N., 1954. *The history of Bukhara. Translated from a Persian abridgment of the Arabic original by Narshakhī*, Cambridge (Massachusetts).

GERSHEVITCH, Ilya 1979. 'Nokonzok's well', *Afghan Studies* 2, 55–73.

GOTŌ Toshifumi 後藤敏文, 2008. 'Kodai indo no saisiki gaikan: keisiki, kōsei, genri 古代インドの祭式概観: 形式・構成・原理 – A survey of the ancient Indian Rituals: Form, structure and logic', *Sōgōningengaku sōsyo* 総合人間学叢書 – *Generalized science of humanity series* 3, 57–102.

GRENET, Frantz, 1983. 'L'onomastique iranienne à Aï Khanoum', *Bulletin de correspondance hellénique* 107/1, 373–381.

GRENET, Frantz, 2005. 'The cult of the Oxus: a reconsideration', in: *Central Asia from the Achaemenids to the Timurids: Archaeology, History, Ethnology, Culture. Materials of an international scientific conference dedicated to the Centenary of Aleksandr Markovich Belenitsky, St. Petersburg, November 2–5, 2004* (ed. V. P. NIKONOROV), St. Petersburg, 377–378.

GRENET, Frantz, 2006. 'Nouvelles données sur la localisation des cinq Yabghus des Yuezhi: l'arrière plan politique de l'itinéraire des marchands de Maès Titianos', *JA* 294, 325–341.

GRENET, Frantz, 2020. 'Histoire et cultures de l'Asie centrale préislamique', *Annuaire du Collège de France 2017–2018*, 311–333.

GRENET, Frantz & Claude RAPIN, 2001. 'Alexander, Aï Khanum, Termez: Remarks on the Spring Campaign of 328', *BAI* 12 (1998[2001]), 79–89.

GYSELEN, Rika, 2019. *La géographie administrative de l'empire Sassanide. Les témoignages épigraphiques en moyen-perse*, Bures-sur-Yvette.

HARUTA Seiro 春田晴郎, 2001. 'Auroman monjo No. 3 (daiei toshokan Or. 8115) no shin'yaku アウロマン文書 No. 3(大英図書館 Or. 8115)の新訳 – A new translation of the Avroman Parchment No. 3 (British Library Or. 8115)', *Oriento* オリエント – *Orient* 44/2, 125–134.

HENNING, Walter B., 1957. 'Mitteliranisch', *Handbuch der Orientalistik* I/IV/1 (ed. B. SPULER), Leiden–Köln, 20–130.

IKEDA On 池田温, 1965. '8 seiki chūyō ni okeru tonkō no sogudojin shūraku 8 世紀中葉における敦煌のソグド人聚落 – A Sogdian colony at Dunhuang in the mid-eighth century', *Yūrashia bunka kenkyū* ユーラシア文化研究 – *Bulletin of the Institute of Eurasian Cultural Studies, Hokkaido University* 1, 46–92.

IWAI Shumpei 岩井俊平, 2004. 'Tohārisutān ni okeru chiiki kan kankei no kōkogakuteki kentō トハーリスターンにおける地域間関係の考古学的検討 – Archaeological study for the regional relationships in Tokharistan', *SAK* 60, 1–18.

IWAI Shumpei 岩井俊平, 2017. 'Bakutoria no sōsō bunka to shinkō バクトリアの葬送文化と信仰 – Funerary culture and beliefs in Bactria', in: *Ajia bukkyō bijutsu ronshū 1: chūō ajia~gandāra* – アジア仏教美術論集 1: 中央アジア～ガンダーラ (ed. MIYAJI Akira 宮治昭), Tokyo, 283–307.

IWAI Shumpei 岩井俊平, 2019. 'Chūō ajia ni okeru bukkyō jiin no garan haichi no hensen 中央アジアにおける仏教寺院の伽藍配置の変遷 – Transition of the layout of Buddhist Temples in Central Asia', *Teikyō daigaku bunkazai kenkyūsho kenkyū hōkoku* 帝京大学文化財研究所研究報告 – *Bulletin of Research Institute of Cultural Properties, Teikyo University* 18, 79–97.

KARLGREN, Bernhard, 1957. *Grammata Serica Recensa*, Stockholm.

KUWAYAMA Shoshin 桑山正進, 2017. 'Kisō Kyūshūkyaku no botsunen 貴霜丘就却の歿年 – The rise of the Kushan empire and the year of Qiujiu Que's demise', *Tōhō gakuhō* 東方學報 – *Journal of Oriental Studies* (Kyotō) 92, 77–134.

DE LA VAISSIÈRE, Étienne, 2016. *Histoire des marchands Sogdiens*, 3ème édition, Paris.

DE LA VAISSIÈRE, Étienne, 2018. 'L'ère kouchane des documents bactriens', *JA* 306/2, 281–284.

LITVINSKIJ, Boris A., 2001. *Xram Oksa v Baktrii (Južnyj Tadžikistan)*, tom 2 (Baktrijskoe vooruženie v drevnevostočnom i grečeskom kontekste), Moskva.

LITVINSKIJ, Boris A., 2010., *Xram Oksa v Baktrii (Južnyj Tadzhikistan)*, tom 3 (Iskusstvo, xudožestvennoe remeslo, muzykal'nye instrumenty), Moskva.

LITVINSKII [LITVINSKIJ], Boris A. & Igor R. PICHIKIAN, 1996. 'The Hellenistic Architecture and Art of the temple of the Oxus', *BAI* 8 (1994[1996]), 47–66.

LITVINSKIJ, Boris. A. & Igor R. PIČIKJAN, 2000. *Èllinističeskij xram Oksa v Baktrii (Južnyj Tadžikistan)*, tom 1 (Raskopki, arxitektura, religioznaja žizn'), Moskva.

LIVŠIC, Vladimir A., 1984. 'Dokumenty', in *Toprak-kala. Dvorec* (ed. Ju. A. RAPOPORT & E. E. NERAZIK), Moskva, 251–286.

LIVŠIC, Vladimir A., 2004. 'Nadpisi i dokumenty', in *Kalaly-gyr 2. Kul'tovyj centr v Drevnem Xorezme IV-II vv. do n.è.* (ed. B. I. VAJNBERG), Moskva, 188–212.

LIVSHITS [LIVŠIC], Vladimir A., 2015. *Sogdian epigraphy of Central Asia and Semirech'e*, London.

LERNER, Judith & Nicholas SIMS-WILLIAMS, 2011. *Seals, sealings and tokens from Bactria to Gandhara (4th to 8th century CE)*, Wien.

LURJE [LUR'E], Pavel B., 2004. *Istoriko-lingvističeskij analiz sogdijskoj toponimii*, PhD thesis, Saint Petersburg State University. Available at (accessed on 22 October 2024): https://www.orientalstudies.ru/rus/images/pdf/dis_lurje_2004.pdf.

LURJE, Pavel B., 2010. *Personal names in Sogdian texts*, Wien.

LURJE, Pavel B., 2022. 'The pantheon of Sogdians from the onomastic viewpoint', *Sichou zhi lu yanjiu jikan* 丝绸之路研究集刊 – *Journal of the Silk Road Studies* 9, 445–462.

MACUCH, Maria, 2007. 'The Pahlavi model marriage contract in the light of Sasanian family law', in *Iranian languages and texts from Iran and Turan* (ed. M. MACUCH *et al.*), Wiesbaden, 183–204.

MALANDRA, William W., 2014. 'VĀYU', *Encyclopædia Iranica*, https://www.iranicaonline.org/articles/vayu (accessed on 22 October 2024).

MARSHAK, Boris I. & Valentina I. RASPOPOVA, 1991. 'Cultes communautaires et cultes privés en Sogdiane', in *Histoire et cultes de l'Asie centrale préislamique. Source écrites et documents archéologiques* (ed. P. BERNARD & F. GRENET), Paris, 187–195, pls. LXXIII-LXXVIII.

MATSUI Dai 松井太, 2011. 'Tonkō shutsudo no uigurugo rekisen monjo: tsūsho gyokukōki tono kanren o chūshin ni 敦煌出土のウイグル語暦占文書: 通書『玉匣記』との關連を中心に – Uigur almanac divination fragments from Dunhuang: The relation with the Chinese almanac *Yuxiaji*', *Jinbun shakai ronsō jinbun kagakuhen* 人文社会論叢・人文科学篇 – *Studies in the Humanities. Cultural sciences* 26, 25–48.

MINORSKY, Vladimir, 1970. *Ḥudūd al-'Ālam 'The regions of the world', a Persian geography 372 A.H. –982 A.D.*, Cambridge.

MIYAMOTO Ryoichi, 2019. 'Étude préliminaire sur la géographie administrative du Tukhāristān', *StIr* 48/2, 163–186.

MIYAMOTO Ryoichi, 2023a. 'Letters from Kadagstān', *BAI* 31 (2022–2023[2023]), 87–104.

MIYAMOTO Ryoichi 宮本亮一, 2023b. 'Wahushu shin to Rāmusēto shin: bakutoriago monjo kara mita tohārisutān ni okeru shūkyō jijō no ichi sokumen ワフシュ神とラームセート神: バクトリア語文書から見たトハーリスターンにおける宗教事情の一側面 – God Wakhsh and God Rāmsēt: A religious aspect in Tukhāristān viewed from the Bactrian documents', in *Bukkyō tōzen no michi: Indo-chūōajia hen* 仏教東漸の道: インド・中央アジア篇 – *The eastward transmission of Buddhism: India and Central Asia* (ed. The Institute of Oriental Philosophy), Tokyo, 361–390.

NAVEH, Joseph & Shaul SHAKED, 2012. *Aramaic documents from Ancient Bactria (fourth century BCE.) from The Khalili Collections*, London.

OETTINGER, Norbert, 2001. 'Das Benennungsmotiv der iranischen Göttin *Anāhitā* (mit einer Bemerkung zu vedisch *Aditi*)', *MSS* 61, 163–167.

OGIHARA Hirotoshi 荻原裕敏 & CHING Chao-jung 慶昭蓉, 2014. 'Shinshutsu tōmushukugo keiyaku monjo ni tsuite 新出トゥムシュク語契約文書について – On the Tumshuqese contracts newly published in China', *SIAL* 29, 7–55.

PESCHL, Benedikt, 2022. 'The god Rāman and the question of dual deities in Middle Iranian onomastics' (paper read at Iranistik-Kolloquium at Freie Universität Berlin, 29 June 2022).

ROSENFIELD, John M., 1967. *The dynastic arts of the Kushans*, Los Angeles.

ROUGEMONT, Georges, 2012. *Inscriptions grecques d'Iran et d'Asie Centrale*, London.

SACHAU, Carl E., 1878. *The chronology of ancient nations*, London.

SCHINDEL, Nicholaus 2011. 'The era of the Bactrian documents: a reassessment', *Gandharan Studies* 5, 1–10.

SCHMITT, Rüdiger, 2017. 'Der Flußgott Oxos in der iranischen Anthroponymie', in *Studia Philologica Iranica. Gherardo Gnoli Memorial Volume* (ed. E. MORANO *et al.*), Roma, 413–426.

SAKAMOTO (GOTŌ), Junko 阪本(後藤)純子, 2008. 'Mizu tachi *ā́pas* to shin *śraddhā́* kodai indo shūkyō ni okeru sekaikan「水たち」ā́pas と「信」śraddhā́: 古代インド宗教における世界観 – *ā́pas* 'the waters' and *śraddhā́* 'trust, belief'', *Ronshū* 論集 35, 110–89.

SHEIKH, Hossein, 2023. *Studies of Bactrian legal documents*, Leiden–Boston.

SHENKAR, Michael, 2011. 'The temple architecture in the Iranian world in the Hellenistic period', in *From Pella to Gandhara: Hybridisation and identity in the Art and Architecture of the Hellenistic East* (ed. A. KOUREMENOS *et al.*), Oxford, 117–140.

SHENKAR, Michael, 2012. 'On the temple of Oxus in Bactria III (Review article)', *StIr* 41/1, 135–142.

SHENKAR, Michael, 2014. *Intangible spirits and graven images: The iconography of deities in the pre-Islamic Iranian world*, Leiden–Boston.

SHENKAR, Michael, 2017. 'The religion and the pantheon of the Sogdians (5th–8th centuries CE) in light of their sociopolitical structures', *JA* 305/2, 191–209.

SIMS-WILLIAMS, Nicholas, 1996. 'From Babylon to China: astrological and epistolary formula across two millennia', in *La Persia e l'Asia Centrale da alessandro al X secolo*, Roma, 77–84.

SIMS-WILLIAMS, Nicholas, 2007. *Bactrian documents from Northern Afghanistan II: letters and Buddhist texts*, London.

SIMS-WILLIAMS, Nicholas, 2008. 'The Bactrian inscription of Rabatak: a new reading', *BAI* 18 (2004[2008]), 53–68.

SIMS-WILLIAMS, Nicholas, 2010. *Bactrian personal names*, Wien.

SIMS-WILLIAMS, Nicholas, 2012a. *Bactrian documents from Northern Afghanistan I: Legal and economic documents*, revised edition, London.

SIMS-WILLIAMS, Nicholas, 2012b. *Bactrian documents from Northern Afghanistan III: Plates*, London.

SIMS-WILLIAMS, Nicholas, 2012c. 'Bactrian historical inscriptions of the Kushan Period', *The Silk Road* 10, 76–80.

SIMS-WILLIAMS, Nicholas, 2015a. 'A Bactrian document from Southern Afghanistan?' (with an appendix by Étienne DE LA VAISSIÈRE), *BAI* 25 (2011[2015]), 39–53.

SIMS-WILLIAMS, Nicholas, 2015b. ‘A new Batrian inscription from the time of Kanishka’, in *Kushan Histories. Literary sources and selected papers from a Symposium at Berlin, December 5 to 7, 2013* (ed. H. FALK), Bremen, 255–264.

SIMS-WILLIAMS, Nicholas, 2025. *Bactrian documents IV: Documents from South of the Hindukush, I*, London.

SIMS-WILLIAMS, Nicholas & François de BLOIS, 2018. *Studies in the chronology of the Bactrian documents from Northern Afghanistan*, Wien.

SIMS-WILLIAMS, Nicholas & Joe CRIBB, 1996. ‘A new Bactrian inscription of Kanishka the Great’, *SRAA* 4, 75–142.

SKJÆRVØ, Prods O., 2002. *Khotanese manuscripts from Chinese Turkestan in the British Library. A complete catalogue with texts and translations*, London.

TAKEUCHI Tsuguhito, 1995. *Old Tibetan contracts from Central Asia*, Kyoto.

TANG Geng’ou 唐耕耦 & LU Hongji 陸宏基 1986. *Dunhuang shehui jingji wenxian zhenji shilu* 敦煌社會經濟文獻真迹釋録, Beijing.

YAKUBOVICH, Ilya, 2006. ‘Mariage Sogdian style’, in *Iranistik in Europa – Gestern, Heute, Morgen* (ed. H. EICHNER *et al.*), Wien, 307–359.

YAMADA Nobuo 山田信夫, 1993. *Uigurubun keiyaku monjo shūsei* ウイグル文契約文書集成 – *Sammlung Uigurischer Kontrakte* (ed. ŌDA Juten 小田寿典 *et al.*), Osaka.

YAMADA Tomoki, 2011. ‘The image of Ṛgvedic Sarasvatī: RV II 95, 96’, *IBK* 59/3, 1103–1108.

YOSHIDA Yutaka, 1994. ‘Review: Sims-Williams, N. *Sogdian and other Iranian inscriptions of the Upper Indus* 1’, *BSOAS* 57/2, 391–392.

YOSHIDA Yutaka 吉田豊, 2013a. ‘Bakutoriago monjo kenkyū no kinkyō to kadai バクトリア語文書研究の近況と課題 – Current issues of the Bactrian Studies’, *SIAL* 28, 39–65.

YOSHIDA Yutaka, 2013b. ‘Review of Sims-Williams 2012a & 2012b’, *BSOAS* 76/1, 156–159.

YOSHIDA Yutaka, 2019. *Three Manichaean Sogdian letters unearthed in Bäzäklik, Turfan*, Kyoto.

YOSHIDA Yutaka, 2022a. ‘Sogdian Christians in China, Turfan and Sogdiana’, in *Byzantium to China: Religion, History and Culture on the Silk Road* (ed. G. B. MIKKELSEN & K. PARRY), Leiden–Boston, 596–617.

YOSHIDA Yutaka 吉田豊, 2022b. ‘Fayazutepa hekiga ni soerareta bakutoriago meibun no shotai niyoru nendai hantei oyobi kanren suru mondai nitsuite ファヤズテパ壁画に添えられたバクトリア語銘文の書体による年代判定および関連する問題について Palaeographical dating of the Bactrian inscription accompanying the wall painting of Fayaztepa and related problems’, in *Uzubekisutan nanbu fayazutepa iseki shutsudo shoki bukkyō hekiga no hozonshūhuku to kenkyū* 2 ウズベキスタン南部ファヤズテパ遺跡出土初期仏教壁画の保存修復と研究 2 – *Conservation and research of early Buddhist wall paintings from the Fayaztepa site, Southern Uzbekistan 2* (ed. KAGEYAMA Etsuko 影山悦子 *et al.*), Nagoya, 37–71.

A new look at ownership clauses in Tumshukese sale contracts*

OGIHARA Hirotoshi 荻原裕敏

As demonstrated by Hitch (1988) and other philologists, documents written in various pre-Islamic Central Asian languages share common terms and expressions as a result of cultural and social contact. Tumshukese, a Middle Iranian language from the North-western Tarim Basin, is one of the languages in this Central Asian cultural sphere, and Tumshukese contracts share a number of traits with contracts in other languages. Ownership clauses in Tumshukese contracts are not yet properly understood, and in this paper, I will take a new look at these. After explaining the difficulty in distinguishing <g> and <z> in ligatures of the "North Turkestan" Brāhmī cursive script, I propose to modify my earlier reading *grap^{h}ā* in the ownership clauses of TUMXUQ 001a+b, 003a+b, and 004 to *zrap^{h}ā*. Based on a comparison of parallel clauses in Niya-Prakrit, Bactrian, and Sogdian sale contracts, I assume that the newly identified word *zrap^{h}ā* is derived from a Tumshukese verb *zrap^{h}-*, going back to OIr. **uz-raf-* 'to hit, hurt, exploit(?)'.

1. Introduction

Since Sten Konow founded Tumshukese studies in 1935, much effort has been made to interpret the grammatical structure of this Middle Iranian language. Tumshukese, a language closely related to Khotanese, was the native language of the Tumshuk–Maralbashi region between Kucha and Kashgar. Its corpus, approximately dating to the 6th–8th centuries CE, is the smallest of the seven Middle Iranian languages, and it is entirely written in "North Turkestan" Brāhmī (cf. Sander 2009). The Tumshukese socio-economic documents, the best-preserved materials in this language, have contributed

* This paper is an adaptation of a part of my joint presentation with Ching Chao-jung at the *100. Deutscher Orientalistentag*, Berlin, 14 September, 2022, entitled "An attempt to elucidate Tumshuqese secular documents" and my communication "Reconsidering the ownership clause in the Tumshuqese sale contracts" (with the collaboration of Ching Chao-jung) at the ECIS 10, Leiden, 28 August, 2023. The conference trip to Leiden was financially aided by the Murata Science and Education Foundation and the ERC project "The Tocharian Trek". I am deeply grateful for valuable comments from Ching Chao-jung, from Dieter Maue, from Michaël Peyrot, and from members of the research group "Socio-economic base of ancient traffic in Inner Asia in the light of Chinese and non-Chinese sources" of the Institute for Research in Humanities, Kyoto University. All the mistakes that remain are mine.

10.29091/9783752003635/007

most to the elucidation of its linguistic features through comparative analysis with other Central Asian documents (cf. Hitch 1988).

Recently, ownership clauses have been identified in Tumshukese documents archived in the Bureau of Cultural Relics of Tumshuk City (Ogihara & Ching 2012; TPR; Ogihara & Ching 2014), but they could so far not be understood in all detail. The main difficulty lies in the phrase *añix$_6$ā grap^{h}ā u parād$_1$ane* where the meaning of *parād$_1$ane* 'to sell' (inf.) is assured, but the word thus far read as *grap^{h}ā* remains obscure. I repeat here the sentences in which the relevant ownership clause occurs, in our previous transcription and with our previous translation, but for convenience, I have improved the word segmentation and punctuation:[1]

TUMXUQ 001a, line 8–10

bid$_1$i ***añix$_6$ā grap^{h}ā u parād$_1$ane*** *bimex$_6$id$_1$ane ax$_8$o puryo dud$_1$aryo hmalo Lekyo saṅgarme bula rendu.* (cf. Ogihara & Ching 2012: 46)

'Moreover, **to sell [him] to another person and to …**, [to treat him as] *bimex$_6$a* together with his (= *Wag$_1$ereźa*'s) sons [and] daughters, must be a business(?) [to be treated] in the monastery.' (Ogihara & Ching 2017: 463–464)

TUMXUQ 003a, line 11–12

bid$_1$i ***añix$_6$ā grap^{h}ā u parād$_1$ne*** *rād$_1$a paṣto rordane, bśi ci Kusācā pad$_1$itu.* (cf. TPR: 68)

'Moreover, **to sell [it] to another person and to …**, to give [as] a gift(?) …, all [right] must yield(?) to the *Kusāca*-official.' (Ogihara & Ching 2017: 466)

TUMXUQ 004, line 7–8

bid$_1$i ***añix$_6$ā grap^{h}ā u parād$_1$ane*** *rād$_1$a paṣto rordane, bśi rāzi Kusācā pad$_1$itu.* (cf. TPR: 71)

'Moreover, **to sell [it] to another person and to …**, to give [as] a gift(?) …, all right must yield (?) to the *Kusāca*-official and his sons.' (Ogihara & Ching 2017: 467)

At first glance, Tumsh. "*grap^{h}ā*" might be supposed to go back to OIr. **grabH-* 'to grab, take, seize' (EDIV: 119–121).[2] However, this is in fact difficult, since OIr. **bH* would be expected to yield Tumsh. <v/w> rather than <p^h> in intervocalic position.[3] So we had

[1] Many sincere thanks are due to Desmond Durkin-Meinsterernst for his improvement of the English translation of TUMXUQ 001–004 in Ogihara & Ching 2017. However, the word segmentation and punctuation in TPR: 85–86 are outdated. A fully revised edition of the four documents will be published in the "Tumshukese volume" of the *Corpus Inscriptionum Iranicarum* by Ching, Maue, and Ogihara.

[2] Independently, this idea was suggested by Yoshida Yutaka on 6 May, 2014, in his (non-anonymous) peer review of Ogihara & Ching 2014.

[3] To my knowledge, no Tumshukese word reflects OIr. **bH* in intervocalic position, but it may be assumed to have resulted in <v/w> based on the outcome /w/ of intervocalic **bH* in Khotanese (cf. Khot. *staura-* 'severe, cruel', see EDIV: 362, **stamb(H)-* 'to withstand, be stiff'). OIr. **b* in intervocalic position is reflected in Tumsh. *usänavara* (Pelliot Fragments divers D.A. Fragment G, line 31) and

left "*grap*^{h}*ā*" untranslated in the above passages, and marked it with "…" in the translation. In view of the difficulties with the interpretation of "*grap*^{h}*ā*", we had considered *zrap*^{h}*ā* as an alternative reading (Ogihara & Ching 2014: 29–30, fn. 44), but we had nevertheless opted for the reading *grap*^{h}*ā* as an interim solution, because no akṣara <zra> had been read previously by anyone else in the Tumshukese texts available until then.[4]

Among the nine Tumshukese contracts thus far identified (see the list in Ogihara 2019: 298), "*grap*^{h}*ā*" is attested in three. Two of these, TUMXUQ 003a+b and 004, are sale contracts, and the third, TUMXUQ 001a+b, is very probably an agreement between a family or clan and a monastery concerning the change of identity of a servant (Tumsh. *bimex*$_{6}$*a-*, see §4). For my new look at the ownership clauses, I first need to make a palaeographical analysis of the akṣaras <gra> and <zra> to examine the plausibility of the alternative reading *zrap*^{h}*ā*.

2. Palaeographical analysis of <zra> and <gra>

Although no clear attestation of <zra> has been recognized in the formal writing of "North Turkestan" Brāhmī used for Tumshukese as well as for Kuchean (Tocharian B), Agnean (Tocharian A), Sogdian, and Old Uyghur, in this section I will show that for the word "*grap*^{h}*ā*" the readings <zra> and <gra> are equally possible.

It is known that <gra> and <zu> are clearly distinct in formal Brāhmī, see Kuchean THT 337, side a, line 2 *rājagrīne* and Tumshukese TS 14, side a, line 2 *azu* (Fig. 1–2).[5]

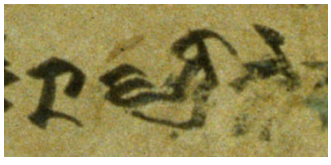


Fig. 1. <rā ja grī ne> in THT 337.

Fig. 2. <a zu> in TS14.

© Depositum der Berlin-Brandenburgischen Akademie der Wissenschaften in der Staatsbibliothek zu Berlin – Preussischer Kulturbesitz Orientabteilung.

uznawaranu (TS 37–40, line 4), the nominative-accusative and genitive-dative plural of *uzanavara-* 'living being' (cf. Khot. *uysnaura-*) < OIr. **uz-anā-bara-* 'bearing breath' (Sims-Williams 2025: 108), respectively. See also Emmerick 1989: 214 for the development of the intervocalic **b* > Khot. /w/.

[4] Nevertheless, it is interesting that Skjærvø once transcribed *zrä* (i.e. <z̲ra̲> in transliteration) in his unpublished reading of TS 1+6+21. See below, §2.

[5] See Sieg and Siegling 1953: 219 and Maue 2009c: HL No. 18e, respectively. The formal writing of <za>, recognized as a Fremdzeichen, is attested in TS 9, side a, line 2 *uzanayya* and IOL Toch 185, side a, line 3 *azaränu*. The akṣara <ga> is seen in the Sanskrit fragment SHT 9.3, side b, line 2 (*bhagavān*, *bhagavāṃ*) found in Maralbashi–Tumshuk region. As for the gloss written in Tumshukese in the same line, see Ogihara 2015: 88 fn. 16.

However, in the cursive variant of the same script, the two akṣaras are not so easily distinguished, because the vowel diacritic *u* attached below the radical <za> is similar to the postconsonantal diacritic *r* attached to the radical <ga> in several documents. For instance, see <gra> and <zu> in TUMXUQ 003a, line 9 *granda* (Fig. 3) and line 7 *azu* (Fig. 4), respectively.

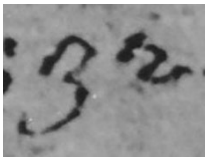


Fig. 3. <gra nda> in TUMXUQ 003a.

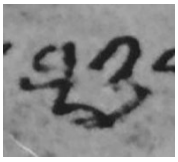


Fig. 4. <a zu> in TUMXUQ 003a.

© Bureau of Cultural Relics of Tumshuk City.

The problem of distinguishing <gra> and <zu> is parallel to the problem under discussion, namely how <gra> and <zra> can be distinguished. More precisely, in contrast to the vowel diacritic *u*, which has three variants in "North Turkestan" Brāhmī (e.g. in <ku>, <khu> and <gu>, cf. Sander 1968: Tafel 33), the diacritic for post-consonantal *r* shows no such variants (cf. Sander 1968: Tafel 30, 34, 38). Moreover, in ligature, the akṣara <za> only retains the upper part and looks like <ga>; see <zño> in IOL Khot 203(9), side b, line 3 *azño* (Fig. 5), cf. Maue 2009c: HL No. 17b. Because of this, the first akṣara in "*grap*^{h}*ā*" in TUMXUQ 001a, line 8–9 (Fig. 6), 003a, line 11 (Fig. 7), and 004, line 7 (Fig. 8) is indeed very similar to cursive <gra>.[6]

Fig. 5. <a zño> in IOL Khot 203(9).

© British Library.

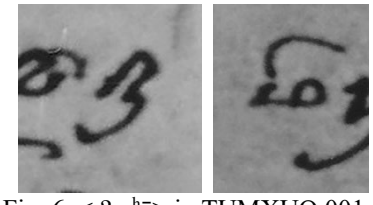


Fig. 6. < ? p^{h}ā> in TUMXUQ 001a. (end of line 8 and beginning of line 9)

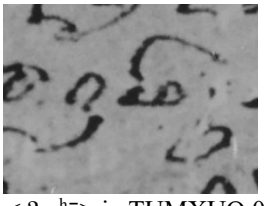


Fig. 7. < ? p^{h}ā> in TUMXUQ 003a.

© Bureau of Cultural Relics of Tumshuk City.

[6] TPR: 98, 102, 104. See also <gri> in TUMXUQ 004, line 5 *gridıu* 'bought'. It can be read <gri> or <zri> from palaeographical perspective, but the reading of <gri> is guaranteed by the context and by its etymology.

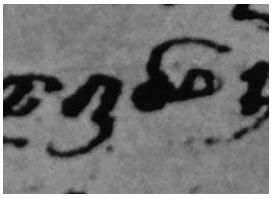


Fig. 8. < ? $p^hā$> in TUMXUQ 004.
(with "?" designating the akṣara under discussion)
© Bureau of Cultural Relics of Tumshuk City.

In order to clarify the situation, clear instances of the ligatures <gr> and <zu> in TS 25+28 (= HL No. 1 in Maue 2009b and c) are compared here, namely in line 23 *haṅgre* '?' (Fig. 9) and in line 5 *azu* 'I' (Fig. 10). Although the meaning of *haṅgre* is obscure,[7] Konow's reading <g> (1935: 804) is safe as it follows <ṅ>. Interestingly, a very similar akṣara is attested in line 12 of the same document, which was read by Konow as *zup^{h}i* (1935: 803) and *gup^{h}i* (1947: 163) and is no doubt related to our word "*grap^{h}ā*" (Fig. 11).

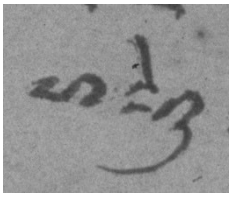


Fig. 9. <ha ṅgre> in TS 25+28.

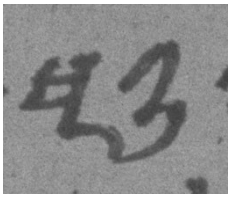


Fig. 10. <a zu> in TS 25+28.

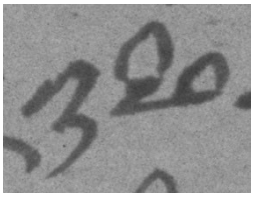


Fig. 11. <? p^hi> in TS 25+28.
© Depositum der Berlin-Brandenburgischen Akademie der Wissenschaften in der Staatsbibliothek zu Berlin – Preussischer Kulturbesitz Orientabteilung.

The word "*zup^{h}i*" or "*gup^{h}i*" has so far not received a convincing interpretation or etymology, and it has been read variously by different scholars (see the full sentences in §3):

Konow 1935	*zup^{h}i*
Konow 1947	*gup^{h}i*
Bailey 1958	*gup^{h}i*
Hitch 1985, 2009	*grap^{h}i*
Maue 2009c	*grap^{h}i* (with alternative suggestions <gu> or <zu> in his fn. 9)

[7] Tumsh. *haṅgre* is interpreted as the 1sg. mid. of a verb **haṅgar-* 'to agree, promise' (< **haṃ-gar-*) by Hitch (1990: 4), according to him cognate to Skt. *saṅgara-* 'promise, agreement' and Skt. *saṅgīrṇa-* 'agreed, promised'.

I further suppose that another form, attested in TS 1+6+21, side b, line 2 (= HL no. 8 in Maue 2009c, see Fig. 12), is related, too:

Konow 1935, 1947	*śṛp^{h}a*
Hitch 1985	*gṛap^{h}a*[8]
Maue 2009c	*grp^{h}a*
Skjærvø (apud Maue 2009c: fn. 142)	*zräp^{h}a*

Fig. 12. <? p^{h}a> in TS 1+6+21. (mode of Auto Tone by Adobe Photoshop 2025)
© Depositum der Berlin-Brandenburgischen Akademie der Wissenschaften in der Staatsbibliothek zu Berlin – Preussischer Kulturbesitz Orientabteilung.

Unfortunately, I cannot decide on the basis of the palaeography whether the readings with *gra*, *grä* or with *zra*, *zrä* are correct. I will therefore base my further argument on etymology, and propose the reading *zrap^{h}i* for TS 25+28 and *zräp^{h}a* for TS 1+6+21.

3. The context of *zrap^{h}i*, *zräp^{h}a*, and *zrap^{h}ā*

In my opinion, the obscure words are very likely to be read *zrap^{h}i*, *zräp^{h}a*, and *zrap^{h}ā*, derived from a Tumshukese verb to be set up as *zrap^{h}-*. In this section, I will examine the relevant passages to establish the most likely meaning of *zrap^{h}i*, *zräp^{h}a*, *zrap^{h}ā*, and the verb *zrap^{h}-*. Based on this meaning, I will then propose an etymology that supports the reading with *zra-*, *zrä-*.

zrap^{h}i

The word is found in a penalty clause, which Konow read and translated as follows:

TS 25+28, line 11–13

Kwa hve hmaḍa jānu wa pura bi ywo druhvam^{e}ne, ji nu ***zuphi*** *daṇḍi dzaḍu gyāźdi riḍe ḳe śa bārre roro patsasu bandi ta χšerā χšiṣta*. (Konow 1935: 803)

'Wenn aber wir Brüder von selbst, oder deren Söhne, ins Ohr falsch aussagen werden, dann soll, was die **angemessene(?)** Strafe ist, dem Gyāźdi-König zugehen, wovon die

[8] In fact, *gṛap^{h}a* is a transliteration rather than a transcription. A transcription of his reading would be *gräp^{h}a*.

einzelnen Beträge: fünfzig für die Reichsbehörde, so (und) sechzig den Distrikten.' (Konow 1935: 804)

Konow, assigning the meaning 'angemessen, gebührend' to his *zuphi* in the glossary (1935: 823), also briefly mentioned *guphi* as a possibile reading there as well as in a footnote (1935: 803 fn. 2). Afterwards, he modified his reading and provided the following English translation, with 'combined' for *guphi*:

Kwa hve hmaḍa jā nu wa pura bi gwo druhwamnai,[9] *ji nu* ***guphi*** *daṇḍi dzaḍu gyāźdi riḍe ḳe śa bārre roro patsasu bandi ta χšerā χšiṣta*. (Konow 1947: 163–164)
'Or when we ourselves together, or those who are our sons, make a false statement to the ear (*i.e.* clandestinely), then the **combined** fine shall go hundred coins to His Majesty the King, fifty to the treasury, and sixty to the province.' (Konow 1947: 164–165)

As to the etymology, he explained *guphi* 'combined' as a loan word from Skt. *guṃphita-* 'tied, strung together' in his updated glossary (Konow 1947: 182; cf. MW: 359).

According to Bailey (1958: 154), *guphi* means 'appropriate' and may derive from OIr. **guf-ya-* 'to be stated', etymologically "connected with Old Persian *gaub-* 'to assert'" and Khot. *gguph-* 'to abuse'. He translated the passage as follows:

kwa hve hmaḍa jānu wa pura bi (?)o druhvamnai ji nu ***guphi*** *daṇḍi dzaḍu.*
'and if anyone my own or theirs shall dispute … for the future, the **appropriate** penalty shall apply.' (Bailey 1958: 152)

However, it is unclear how he derives the meaning "appropriate" from OIr. **guf-ya-* 'to be stated', or how it relates to the meanings of the adduced Iranian cognates.

A different interpretation has been provided by Hitch, who reads *graphi* and assumes that it is a certain type of penalty, although he does not give a more exact translation:

kwa hve = hmaḍa jānuwa pura bišo druhvamnai, ji nu ***graphi*** *daṇḍi dzaḍu: gūẓdi = riḍe theśa bārre roro patsasu, bandina xšerā xšiṣta*.
'Or if we ourselves likewise *jānuwa pura* should quarrel, then *nu* (sic!) the **graphi** penalty must apply: to the Gūẓdi king I must give fifty *theśa bārre*, the fine for the state is sixty.' (Hitch 2009: 24)

Hitch's interpretation of the word in its context is in my view correct. I will return to this point in §4.

zräp^{h}a-

The word is only attested in TS 1+6+21, a Buddhist text written in formal calligraphy, which has not yet been identified. Konow's initial decipherment is:

[9] According to his system of transcription, the reading should be *druhvamnai* (cf. Konow 1947: 184). *druhwamnai* seems to be a typographical error. See further discussion of this word in §4.

TS 1+6+21, side b, line 2

amace pur-sickari mare hvaźandi tāri ***gṛphalāźa*** *la …* (Konow 1935: 813)
'Des Ministers Söhne werden die hiesigen .. essen, die anderen sind **.. Fruchtesser** ..' (Konow 1935: 814)

For *gṛphalāźa*, he gives "Skt. .. *phalāśa-* die .. Früchte essend" in his glossary (1935: 817). In other words, he saw in °*phalāźa* a loan word composed of Skt. *phala-* 'fruit' (MW: 716) and *aś-* 'to eat' (MW: 112), but he was unable to solve *gṛ°*. Later, he modified his reading to:

amace pursickari mare hvaźyandi tāri ***śṛphalāźa*** *la …* (Konow 1947: 172)
'The ministers feed here the judges, the others are a little **bilva-eaters**.' (Konow 1947: 173)

Accordingly, Konow gave the meaning of *śṛphalāźa* as "eating bilva" in his updated glossary (Konow 1947: 188), now assuming that the first part is a borrowing from Skt. *śrīphala-* 'the Bilva-tree, *Aegle Marmelos*' (MW: 1099c), and maintaining his explanation of the second part as borrowed from Skt. *aś-* 'to eat'. In my view both of Konow's intepretations of this word, as well as the etymologies proposed, are not very convincing in the context. Also, as I will argue below, the reading should, following Skjærvø, rather be *zräp^{h}a-*.

zrap^{h}ā

As explained in §2, all three attestations of *zrap^{h}ā* (as I will argue it is to be read, cf. Fig. 6–8) are in the clause concerning new ownership, where *añix$_{6}$ā* was translated as "to another person" in the previous editions on the assumption that it is the genitive-dative of *anix$_{6}$a-* < OIr. **aniya-ka-* 'other' (Ogihara & Ching 2014: 29 fn. 43; 2017: 457). This is rightly rejected by Dragoni (2020: 220), who points out that it cannot be the genitive-dative singular of an *aka*-stem. However, *añi*/*ani*, *añix$_{6}$ā*, and *añix$_{6}$e*, by Dragoni (2020: 221) taken as case forms of the same proto-form **ania-*, can be analyzed differently. In my view, *añix$_{6}$e* is more likely a variant form of *añix$_{6}$ā*, and I see *añipre*/*anipre* as one word, not as two words *añi pre*/*ani pre* by Dragoni's analysis.[10] Based on this interpretation, I would like to propose that *añix$_{6}$ā* derives from OIr. **anyaθā* 'in a different way, otherwise', cf. Skt. *anyáthā* 'otherwise, falsely', MP *'yny'* /ēnyā/ 'otherwise', Gath. *anyāθā* (Nyberg 1974: 71a; Bartholomae 1904: 138). Given the three attestations in the same formula, I assume that the scribe erroneously omitted this expression in the relevant part of TUMXUQ 002a+b (cf. TPR: 66) and restore the passage as follows:

[10] I have discussed the phonetic value of <x$_{6}$> in the joint presentation with Ching at the *100. Deutscher Orientalistentag* (see above, the headnote). It will be published in a future work.

TUMXUQ 002a, line 10–11

bid$_{1}$i ***anix$_{6}$ā {zrap^{h}ā u} parād$_{1}$āne*** *ruru rordane, bśi kusācā pad$_{1}$itu.*

In my view, the key to determining the meaning of *zrap^{h}ā* is to be sought in a comparison with contracts in other Central Asian languages. Among these, the Niya-Prakrit contracts often clearly define the rights of a new owner in transactions of properties such as slaves, farming lands, and vineyards. For example:

CKD 590 uo, line 4–6

*saṃna saṃna taya striya lẏipaae vaṃti ajuy̲adae atra tivira Ramaṣotsaṣa eśvari huda taḍaṃnae baṃnaṃnae vikranaṃnae aṃñano y̲a prahu<*ḍa> deyaṃnae namaṃniya deyaṃnae badho deyaṃnae ṣarva boga kikama karaṃni siyati* (Baums & Glass 2002–)

'From now on the scribe Ramṣotsa has ownership of that woman, to beat her, to bind her, to sell her, to give her to others as a present, to exchange her, to pledge her, to do whatever he likes with her.' (TKD: 126)

CKD 580 uo, line 6–8

taha te miṣiya bhuma avi akriya bhumaṣa vaṃti ajuvadae atra ṣoṭhaṃgha Ramṣotsaṣa eśvarya huda, kiṣaṃnae, vavaṃnae, aṃñisya prahuḍa deyaṃnae, namaṃniya deyaṃnaye, sarva boga kikama karaṃniya siyati. (Baums & Glass 2002–)

'Therefore in that *miṣiya* land and also in the *akriya*-land, from now on the *ṣoṭhaṃgha* Ramṣotsa has ownership, to plough, to sow, to give to another as a present, to exchange, to enjoy in all ways whatever is wanted to be done.' (TKD: 119)

CKD 581 cr, line 1–3

eda masuśaḍaṃmi tivira Ramṣotsaṣa eśvari huda, aṃña prahuḍa deyaṃnae, namaṃniya deyaṃnae, ṣarva boga paribhuchaṃnae kikama karaṃni siyati. (Baums & Glass 2002–)

'In this vineyard the scribe Ramṣotsa has ownership to present it to others, to exchange it, and to enjoy it in every way whatever is wanted to be done.' (TKD: 120)

All the examples show that a new owner has the full right to do whatever he or she would like to do with the property after the transaction by listing various behaviours. A similar legal concept is observable in Bactrian contracts, with regard to slaves, vineyards and other possessions:

Document P', line 15–17 (669 CE, cf. Sims-Williams & de Blois: 84)

οτο μισο παδοαχβανηδο ασιδδηιο υαρσο κιρδο αριμαδο χοαδο δδριγδο παραλαδο ναβαγο ωσταδο λαβνο αμβαγδο αχρινο οιχρινο κιρδο πιδο γαοοανο πιδοκιβτο πιδο σπασο αζαδο υιρτο (BD I^{2}: 87)

'And also you (pl.) have the right (to do) whatever it may suit you (sg.!) to do, to keep (him) yourself, to sell (him), to pawn (him), to give (him as) a gift, to put (him for) purchase (or for) hire, to detain (him) for a misdeed, (or) to let (him) free (in return) for service.' (BD I^{2}: 86)

Document L', line 20–22 (602 CE, cf. Sims-Williams & de Blois: 84)

ṭοι βαζανọκο αλο ι χ̣οβο καδγιγαγο αβο ι μαϙκọ ναβιχṭιγο αγγαργο παραλαδο οδο ναβαγο αβο ωσταδο οδọ ị ọιχρηγανο αβο κιρδο ọ̣δο λαβ̣νọ αβο αμβαγδο οδο χοαδọ αβο δδριγδο σαγονδο υαρσο χοα̣δο ριμαδ̣ọ (BD I^2: 63)

‘And you, Bazanuk, together with the members of your household, have authority (over) it for the future, to sell the property described herein, and to pawn (it), and to put (it for) hire, and to give (it) away (as) a gift, and to hold (it) yourself, just as may suit yourselves.’ (BD I^2: 62)

Likewise, in a Sogdian contract concluded between a Sogdian and a Chinese in Turfan in 639 CE, the new ownership is explicitly defined, even though such a detailed statement is unseen in the many Chinese sale contracts made in the same oasis during the 5th–8th centuries, nor in the Chinese, Old Tibetan, Khotanese, Old Uyghur or any other medieval contracts discovered in Northwest China (Yoshida *et al.* 1989: 21–22).[11]

69 TAM 135:1, line 8–12

KZNH ’PZYšn wβyw šmny y’nsy’n xwty ’PZYšy ’xw BRY ’PZY ’xw npyšn ’PZY ’xw pδy ’PZY ZKh ’wzyh pr k’m’kw x’w’t rnp’tw βynt’t pr’yδ’t np’kw ’wsty’t r’tw βxš’t prβxš’t wn’tw wyspy ’cwtyšn ’’δprmw k’m’t ’krty (Yoshida *et al*. 1989: 5)

‘それで沙門 y’nsy’n 自身と彼の息子, 孫, 一族及び子孫は, 彼女を好きなように打ち, 虐待し, 縛り, 売り, 人質とし, 贈り物として贈与し, 何でもしたいことをしてもよい.’ (Yoshida *et al.* 1989: 8)

‘Accordingly, the monk Yansyan himself and his sons, grandsons, family, and descendants may at will hit her, abuse her, bind her, sell her off, pledge her, give and offer her as a gift, and do whatever they want.’ (tr. mine)

In the above Turfan contract, Sogd. *rnp’tw* gives an important clue to the meaning of Tumsh. *zrap^{h}ā*. This verbal form, here translated as “may abuse”, is the subj. 3sg. of *rnβ-ranβ/p* ‘to attack, fight’ (Yoshida *et al.* 1989: 41; cf. Gharib 1995: 342b, No. 8499 ‘to attack, violate’; Gharib 2004: 342b, No. 8503). The same verb has been thought to appear in the *Vessantara-jātaka*, cf. VJ 1093–1094 *rty βn wβyw xw’w’nt ZY rnp’nt* ‘On vous maltraite et on vous brutalise à la fois’ (Benveniste 1946: 66), but this latter form probably rather belongs to another verb: *ranp-* ‘to insult, humiliate’ (cf. Sims-Williams 2018: 367–368).

It is worthy of mention that a noun derived from the verb is attested in a marriage contract belonging to the Mugh Documents.

[11] See also Arakawa 1989. Incidentally, such a precise definition of absolute ownership is not found in the Kuchean contract THT 4001 either, cf. Ching & Ogihara 2012.

Nov. 4. Recto, line 12–15 (711 CE, cf. Goyibov *et al.* 2022: 112)

rtšw ms ky c'm'k ZY MN s'n'n ky-r'n s'r ''s't ZY pcxw'y't rtšw 'zw y-wn pw γyδrpH ZY pw ry-βyH xwy-ckH w'c'nk'm (Yakubovich 2006: 313)

'And if someone, from my [side] or from the enemies' side takes her and detains her, I shall have her immediately released without damage or injury.' (Yakubovich 2006: 314)

According to Yakubovich's commentary, both the noun *ryβy(H)* in this sentence and *rypH* in recto line 18 of the same contract mean 'injury, defect' and go back to OIr. **rafya-*. Yakubovich claims that this noun "provides a semantic link with Sogd. *rnβ-* 'to attack, fight' (also 'to beat, abuse' in a contract for the purchase of a slave girl, Yoshida-Moriyasu 1988)[12] and Sogd. *r'β(H)* 'illness'" (Yakubovich 2006: 326). However, in my view, these Sogdian forms are rather derived from the Old Iranian verb **Hrab/f-* 'to attack, fight', cf. Parth. *rf-* 'to attack' (EDIV: 185).[13] It is noteworthy that in Cheung's dictionary, a Khotanese nominal form *rraphai jsa* 'in fight' (Sid. 104v3, cf. Bailey 1945: 40)[14] is cited as a cognate of **Hrab/f-*, following the etymology proposed by Bailey (1979: 358, Khot. *rraph-* 'to attack'). If my assumption about these Sogdian words is correct, then Tumsh. *zrap^{h}i-*, *zräp^{h}a-*, and *zrap^{h}ā* are also derived from **Hrab/f-*. More precisely, they might come from OIr. **uz-raf-* 'to hit, hurt, exploit(?)'. If so, the preverb **uz-* was reflected differently in Tumshukese and Khotanese: Tumsh. *z-* ≠ Khot. *uz-* < OIr. **uz-*. Interestingly, some modern Eastern Iranian languages in the Pamirs seem to have preserved more cognates: cf. Shugh. *cirafc-*/*cirūvd-* '[intrans.] to hurt (саднить)' and Yazgh. *cərafs-*/*cərafšt-* '*idem*'.[15]

From the morphological perspective, Tumsh. *zrap^{h}i*, *zräp^{h}a-*, and *zrap^{h}ā* can be assigned to two different nouns:

(1) Tumsh. *zrap^{h}a-* < **zrap^{h}-a-*

- nom. sg. *zrap^{h}i* (TS 25+28, line 12), also attested as *zräp^{h}a-* in compound (TS 1+6+21, side b, line 2).

Cf. Khot. instr.-abl. sg. *rraphai jsa* < *raphaa-* < **raph-aka-* < **raph-* (without the preverb **uz-*).[16]

[12] Sic! Actually, Yoshida *et al.* 1989 is meant.

[13] Cheung supposes that the nasal infix in the Sogdian verb *rnβ-* results from a contamination with **ran(H)-?* 'to fight, enjoy (to fight)' (EDIV: 185, 313).

[14] Bailey (1979: 358) wrongly gives the source of the quotation as Sid. 104v2.

[15] Èdel'man 2020: 383–384, who reconstructs the etymology as **us-raf-sa-*. In my opinion, the expected form of the prefix **us-* in this position, before a voiced sound, is rather **uz-*.

[16] Normally the instr.-abl. sg. ending is *-aina*, but the Late Khotanese form *-ai jsa* is also attested for *āchai jsa* (Emmerick 1968: 298).

(2) Tumsh. *zrap^{h}aā-* < **zrap^{h}-akā-*

- nom. sg. *zrap^{h}ā* (TUMXUQ 001a, line 8–9; 003a, line 11; 004, line 7).

In Khotanese, both the suffixes *-aa-* (< OIr. **-aka-*) and *-aā-* (< OIr. **-akā-*) derive action nouns from the present stem of verbs (Degener 1989: 13–16). The fact that Tumsh. *aa*-stem nouns show a nom. sg. in *-a* instead of *-ā* (cf. Tumsh. *pesa* < *pesaa-* 'teacher', Khot. *pīsaa-* '*idem*'), suggests that Tumsh. *zrap^{h}ā* is an *aā*-stem noun, cf. Khot. *haysānaā-* 'perception' < *haysān-* (cf. OIr. **fra-zan-*) = Skt. *saṃjñā-* (Canevascini 1990: 19–20).[17] The above general comparison of the Tumshukese, Niya-Prakrit, Bactrian, and Sogdian contracts shows that Tumsh. *zrap^{h}aā-* is very probably an expression of exploitation (i.e. the act of using something or someone unfairly for one's own advantage) or maltreatment (i.e. the act of treating someone cruelly or violently).

4. Revised interpretation on Tumsh. *zrap^{h}-*

Based on the above considerations, and with *zrap^{h}a-* and *zrap^{h}aā-* meaning '[a kind of?] penalty' and '± exploitation', respectively, the translation of the aforementioned passages can now be revised to a significant extent:

TS 25+28, line 11–12

kwa hve hmad$_{I}$a jā nu wa pura biźo druhvamnai, ji nu ***zrap^{h}i*** *daṇḍi v$_{I}$ad$_{I}$u.*[18]

'If [we] both accuse each other (lit. 'ourselves'?)[19] falsely(?) … in the future(?),[20] then there shall be a **penalty** [and] a punishment (*zrap^{h}i daṇḍi*) to us.' (tr. Ogihara & Ching)

The verbal phrase *biźo druhvamnai* is by Hitch (2009: 24–27) translated as 'to quarrel', which is probably correct. He sees the former word (= *bišo* in his system)[21] as the acc. sg. f. of the Tumshukese word for 'tongue' and the latter as the subj. 1pl. mid. of *druhv-* 'to misbehave'. Initially, the Tumshukese verb was interpreted by Konow as 'falsch aussagen' (1935: 817) and 'to lie' (1947: 184), and it was linked to Khot. *drūja-*.[22] The problem with Konow's interpretation is that *druhv-* cannot be derived OIr. **drauǰ-* 'to

[17] See Ogihara 2020: 21 for the declension of Tumsh. *aa*-stem nouns.

[18] Apart from my reading *zra°* rather than *gra°*, the transliteration is based on Maue 2009c, with my word segmentation added.

[19] Our translation is based on the assumption that Tumsh. *hmad$_{I}$a* is the acc. pl. m. of the adjective *hmad$_{I}$a-*, which was intitially translated by Konow (1935: 818) as 'von selbst', cf. Khot. *hamata-* 'in itself, of itself' (Bailey 1979: 458).

[20] We follow Bailey's grammatical note on *pura* 'in the future' (1958: 154).

[21] Hitch's <š> has the same phonetic value as Maue's <ź>, that is, both indicate a voiced palatal sibilant (Hitch 2009, 25 fn. 81). For clarity, I transcribe this akṣara as <ź>.

[22] I.e. Khot. *drrūja-*/*drūja-* 'falsehood' in Bailey 1979: 168. At present, the stem of this noun is interpreted as *drūgyā-*/*drūjā-* in Khotanese studies (Emmerick 2024: 226 and Sims-Williams 2025: 119).

lie, deceive' (EDIV: 80–81), the root from which Khot. *drūja-* is derived, because OIr. **ǰ* corresponds to /dz/ in Tumshukese; cf. Tumsh. *tsata-*,[23] Khot. *jsata-*, past stem of *jsan-* 'to strike' (Emmerick 1968: 37), from OIr. **ǰan* 'to kill, strike' (EDIV: 224). Bailey (1958: 154) translated it as 'to dispute' and traced it back to *drah-*, a "derivative from *drang-* 'to confirm' with *-u-* due to the following *-v-*". This etymology is difficult in view of the fact that a present stem with suffix *-u-* is not attested for PIE **dreg^{h}-* 'to hold, fasten', from which for instance Avestan *drəṇj-* (*draž-*) 'to fix, fasten, hold' derives (LIV^{2}: 126; EDIV: 76). Hitch (2009: 25) only mentions Bactr. δρουο- of uncertain meaning as a possible phonological match without explaining his semantic suggestion 'to misbehave'. In other words, his interpretation of *druhv-* seems to be based on the context where it appears.

In my view, Tumsh. *druhv-* may derive from Skt. *druhu-* 'injury, harm' (MW: 502b): it can be interpreted as a denominative verb of Class A formed within Tumshukese on the basis of the Sanskrit noun (Ogihara 2019: 307 fn. 13). In Tumshukese as well as in Khotanese, most of the denominative verbs assigned to Class A go back to OIr. **-aya-* stems (Emmerick 1968: 178, 188–189). If this is correct, *druhv-* could literally mean 'to injure, harm'.[24]

As for Tumsh. *biźo*, Hitch compares it with Khot. *biśāa-* 'tongue' (Emmerick 1968: 307). Although he does not give a stem or class, his grammatical analysis of this form as acc. sg. f. implies that it must be an *ā*-stem, unlike its Khotanese counterpart, which is an *āa*-stem.[25] However, if *biźo*, the only form of this word so far attested, is indeed an *ā*-stem form, then it is more probably the instrumental-ablative plural (cf. TUMXUQ 002a, line 19 *diźo* < *diźā-* 'direction' borrowed from Skt. *diśā-* '*idem*'; see also Khot. *diśā-* '*idem*'). Morphologically, *biźo* is indeed possibly the acc. sg. of a feminine *ā*-stem

[23] Restored from the infinitive *tsatänayyā* in the Tumshukese Karmavācanā (Pelliot Fragments divers D.A. Fragment G, line 31, cf. Bailey 1950: 662). Only in this Buddhist text, the phoneme /dz/ is uniquely written <ts>, whereas in other documents it is written <x$_3$> (namely <g$_2$> in Maue 2009c), cf. Maue 2022; Maue & Sims-Williams 2024: 32–35; Ogihara & Ching 2024: 308 fn. 14, 317 fn. 57. As for the etymology of Khot. *drūgyā-*/*drūjā-*, Sims-Williams (2025: 32–35) reconstructs it as **drauǰyā-*, derived from **drauǰ-* 'to lie, deceive' (EDIV: 80–81). It is noteworthy that Tumsh. *biśe drox$_3$e* appears in Pelliot Fragments divers D.A. Fragment S+R+T, side b, line 4. If *drox$_3$e* in this phrase is cognate to Khot. *drūgyā-*/*drūjā-*, it could be the nom.-acc. pl. f. form of **drox$_3$ā-*. The discrepancy between Khot. <gy/j> and Tumsh. <x$_3$> may be due to a phonetic change of OIr. **-ǰy-* to Tumsh. /dz/ <x$_3$>, or to some morphological process. If *drox$_3$e* is indeed cognate to Khot. *drūgyā-*/*drūjā-*, Tumsh. *biśe drox$_3$e* would denote 'all the falsehood'.

[24] Incidentally, the first component of Bactr. λρουομινανο 'enemies(?)' (cf. Lazard *et al.* 1984: 228) in the Surkh-Kotal inscription and δδρουμινο, δρουμινο, and λρουμινανο 'enemy' in documents (cf. BD II: 228–229) may be also have been borrowed from a Sanskrit derivative of *druh(u)-*, perhaps **druh-min-* 'harmful, with harm' with the rare suffix *-min-*.

[25] On the declension of the Khotanese *āa*-stems, see Emmerick (2024: 297–298).

noun, but one cannot exclude the possibility that the word in the verbal phrase *biźo druhvamnai* in fact indicates a means or manner and therefore would rather require the instr.-abl. case.[26] In this case, *biźo druhv-* could literally mean 'to injure/break [the deal] by [one's] words (lit. tongues)', and, as a legal term, it could mean 'to falsely accuse' (Ching Chao-jung, p.c. 15 August, 2024). It is noteworthy that in contracts *biźo* so far occurs only in the collocation with *druhv-* (TS 25+28, line 12, 25; TS 31, line 10; TUMXUQ 002a, line 14), while *bi-druhv-* with a preverb *bi-* (< OIr. **vi-*) is used in case that *druhv-* does appear without *biźo* (TS 25+28, line 10; TUMXUQ 004, line 9; H 111, line 9).[27] Nevertheless, both *biźo druhv-* and *bi-druhv-* appear in TS 25+28, and the difference in usage is to be solved.

Hitch further reads *dzaḍu* instead of Maue's *vɪadɪu*, adopting Konow's reading (1935) of <x11> as /dz/. However, Maue's solution /v1/ (1996: 215–219) is convincing not only on etymological grounds, but it is also supported by the variants *udɪu* and *ūdɪu* in parallel passages, as the same verb appears in the penalty clause of another contract:

TS 24, line 5–6 (= HL no. 3)

ki wa marā kruṣtā ñiwāne ardi, ki ṣu palaci hampā pandamidɪi, ji nu ***daṇḍi vɪadɪi***.[28]

'If someone [of the transacted party] makes complaints(?) to this contract, or someone makes a legal dispute to the [other] party(?),[29] then **there shall be a punishment** to them.' (tr. Ogihara & Ching)

As argued in another paper (Ogihara & Ching 2017: 470), Tumsh. *vɪadɪu* and *vɪadɪi* are the impv. 3sg. act. and prs. 3sg. act. respectively of *u-/vɪa-* 'to be(come)' < OIr. **bauH-* 'to be(come)' (EDIV: 16–18). The difference between the two clauses in TS 25+28 and TS 24 lies in the presence of *zrap^{h}i*. This leads me to suppose that *zrap^{h}i* and *daṇḍi* most probably belong to the same semantic field. Incidentally, the context suggests that the hapax legomenon *ñiwāne* is an action noun related to *ñiwānā* '± plaintiff, claimant'. Both forms can be regarded as derivatives from the putative verb *ñi-wān-/ni-wān-* (causative of **ni-ban-*, cf. OIr. **ban-* 'to afflict', cf. EDIV: 4). In my view, Tumsh. *ñiwāne* would be the nom.-acc. pl. of *ñiwānaā-* 'complaint'.[30]

---

[26] See the usage classified as "instrumental of means" by Emmerick (1965: 31).

[27] Maue kindly reminded me (p.c. 27 August 2021) that H 111.9 *biṣtru hviḍte* so read by Duan (2021: 5) is an infinitive form *bidruhvidɪne*, which can be further supported by the variant *bidruhvidɪane* in TUMXUQ 004.9–10. I totally agree with his suggestion.

[28] Hitch 2009: 27 reads *dzadi*. Hitch's sentence segmentation is signifcantly different from mine. In his view, the very contract clause should begin with *ji* 'then' (an introductory particle of a new sentence?) and is a subordinate clause, which is followed by the main clause (not quoted here).

[29] On Tumsh. *hampā*, see Ogihara & Ching 2017: 459.

[30] See Ogihara & Ching 2017: 459 fn. 28 for the alternative etymology of *ñiwānā* suggested by D. Durkin-Meisterernst.

TS 1+6+21, side b, line 2

amace pursickari mare hvax$_5$andi nāri ***zräp^{h}a-lāẓa*** *la* [...][31]

'Ministers, judge, these people,[32] wife, one who **punishes (?) (lit. gives a penalty)** […]' (tr. mine)

Recently, Dieter Maue has assigned the meaning 'to give' to Tumsh. *lāẓ-*, assuming that it is a loan from Niya-Pkt. *laṣi* 'gift' (p.c. 27 August, 2021).[33] Therefore, I would now read *zräp^{h}a-lāẓa*, the nom. sg. m. of a compound of *zrap^{h}a-*° and Tumsh. **lāẓaa-*, a noun derived from the verb posited by Maue.

TUMXUQ 001a, line 8–10

bid$_1$i añix$_6$ā ***zrap^{h}ā*** *u parād$_1$ane bimex$_6$id$_1$ane ax$_8$o puryo dud$_1$aryo hmalo Lekyo saṅgarme bula rendu.*

'Moreover, [to do] **exploitation** [of him = the transacted person] in [any] other way and to sell [him], [and] to treat him [fully] as *bimex$_6$a* together with his sons [and] daughters in all, shall be the affair in the monastery (founded?) by the (clan of?) Leka.' (tr. Ogihara & Ching)[34]

TUMXUQ 003a, line 11–12

bid$_1$i añix$_6$ā ***zrap^{h}ā*** *u parād$_1$ne rād$_1$a paṣto rordane, bśi ci Kusācā pad$_1$itu.*

'Moreover, [to do] **exploitation** [of this *hampara* 'storehouse(?)'][35] in [any] other way and to sell [it], [and] to give [it as] a gift under a pact(?), all [ownership] must be transferred (lit. 'must fall') to the *Kusāca*.' (tr. Ogihara & Ching)

In the above sentence, Tumsh. *paṣto* 'under a contract/pact(?)' may go back either to:

(A) OIr. **paštakam*, cf. MP *pašt* 'a binding promise' (Nyberg 1974: 153b), 'a verbal and written contract' (Macuch 2022a: 153; 2022b: 313 fn. 55), Bactr. παβτο 'agreement, pact' < **pašti-* (BD II: 253b), Sogd. *pšt* 'contract, pact' (Gharib 1995

[31] Apart from my reading *zrä*° rather than *gṛ*°, the transliteration is based on Maue 2009c, with my word segmentation added. See Maue 2004 on the value /ʒ/ of <x$_{10}$>, which was read <ź> by Konow.

[32] See Skjærvø 1987: 84 on the interpretation of Tumsh. *hvax$_5$andi* and the phonetic value of <x$_5$>, which Konow had transcribed as <ź>/<źy>.

[33] Interestingly, it has been noted that Niya-Pkt. *laṣi* is very likely an Iranian loan (LKD 115; BD II: 226 λαβνο 'gift').

[34] Taking the clauses in TUMXUQ 003a+b and 004 into consideration, one cannot exclude an alternative solution "[…] together with his sons [and] daughters, all shall be the affair in the monastery (founded?) by the (clan of?) Leka." In this case, *hmalo* (adv.) is to be regarded as the beginning of the final phrase.

[35] The exact meaning of Tumsh. *hampara-* remains unclear. The translation of *hampara* 'storehouse' was proposed by Bailey (1958: 152). We assume that it is something bringing yearly profit (Ogihara & Ching 2014: 24; 2017: 471), which is not necessarily a storehouse but possibly another kind of lucrative real estate.

and 2004: 302a, No. 7518), Parth. *paštag* ‘bound, fettered’ (DMMP: 285a), MP *pašn* ‘bond, agreement’ (MacKenzie 1971: 66), or

(B) **pati-stav-*, cf. Khot. **paṣtu-* ‘to promise’, Parth. *pdyst’w-* ‘to promise’ (Emmerick 1968: 78; EDIV: 366).

TUMXUQ 004, line 7–8

bid$_1$i añix$_6$ā ***zrap^{h}ā*** *u parād$_1$ane rād$_1$a paṣto rordane, bśi rāzi Kusācā pad$_1$itu.*

‘Moreover, [to do] **exploitation** [of this *histanaka* ‘a kind of estate(?)’][36] in [any] other way and to sell [it], to give [it as] a gift under a pact(?), all ownership/right must be transferred (lit. ‘must fall’) to the *Kusāca* as well as his sons.’ (tr. Ogihara & Ching)

TUMXUQ 002a, line 10–11

*bid$_1$i anix$_6$ā {**zrap^{h}ā** u} parād$_1$äne ruru rordane, bśi Kusācā pad$_1$itu.*

‘Moreover, {[to do] **exploitation**} [of this *kuẓda* ‘a kind of estate(?)’][37] in [any] other way and to sell [it], to give [it as] a gift, all [ownership] must be transferred (lit. ‘must fall’) to the *Kusāca*.’ (tr. Ogihara & Ching)

Dragoni (2020: 218–220) has recently provided an etymology of *bimex$_6$a-* and translates it as ‘lodger’ or ‘dweller’. While his connection with PIr. **maiθH-* ‘to dwell’ is very likely correct, his suggestion “lodger” or “dweller” for the meaning is not fully satisfactory if one follows the explanation in the Cambridge Dictionary, i.e. “someone who pays for a place to sleep, and usually for meals, in someone else’s house” and “a person who lives in a city, town, cave, etc.”.[38] Based on the context and nature of the transaction, Tumsh. *bimex$_6$a-* most probably designates a servant or a person of inferior social status who is counted as a tradable property or transferable member of a clan or household, cf. TUMXUQ 001a, line 8–10: *bid$_1$i añix$_6$ā* ***zrap^{h}ā*** *u parād$_1$ane bimex$_6$id$_1$ane ax$_8$o puryo dud$_1$aryo bid$_1$i añix$_6$ā zrap^{h}ā u parād$_1$ane* ***bimex$_6$id$_1$ane*** *ax$_8$o puryo dud$_1$aryo hmalo* ‘Moreover, [to do] **exploitation** [of him = the transacted person] in another way, and to sell [him], to **treat** [him fully] **as *bimex$_6$a*** together with his sons [and] daughters all together’. The interpretation of *bimex$_6$a-* as ‘servant’ could also be supported by Ching’s earlier proposal that Tumsh. *ñizeja-*/*nizeja-* mentioned in the document as a status of this *bimex$_6$a-* could be equated to Bactr. νιζαδαγο ‘born (in one’s own house)’ (BD II: 239)

[36] See Ogihara & Ching 2017: 462 on Tumsh. *histanaka-*. It is very probably a borrowing with prothetic *h* from Sogd. *’st’ny* (or *’stny*) *(ə)stǎnē* ‘place’ (Gharib 1995 and 2004: 66b No. 1683 and 67b No. 1703, respectively), which functions as a suffix indicating ‘place of’ < **-stāna(ka)-* (Gershevitch 1961: 171 No. 1118).

[37] See Ogihara & Ching 2017: 456 fn. 14, 462 on Tumsh. *kuẓda-*. Etymologically, it is comparable to Khot. *kūṣḍa-* ‘mansion, palace’.

[38] See https://dictionary.cambridge.org/dictionary/english/lodger and https://dictionary.cambridge.org/ dictionary/english/dweller (last access 2 August 2025).

and Sogd. *nyz'tcw* 'born at home (f.)' (Ogihara & Ching 2014: 14 fn. 18; 2017: 463 fn. 47), so that such a *bimex$_6$a-* qualified by the word *ñizeja-*/*nizeja-* may indicate a domestic servant.

5. Concluding remarks

In this paper, I have reexamined the obscure phrase *añix$_6$ā grap^{h}ā u parād$_1$ne* 'to sell something to another person and to ...' by retranscribing and translating it as *añix$_6$ā zrap^{h}ā u parād$_1$ne* 'to [do] exploitation in [any] other way and to sell [the transacted property]', where the fixed expression *añix$_6$ā zrap^{h}ā* is composed of a noun derived from the newly identified verb *zrap^h-* 'to hit, hurt, exploit(?)' and an adverb *añix$_6$ā* 'in another way, otherwise, etc.', on the assumption that the legal concept of a rather absolute ownership is shared among Tumshukese, Sogdian and other Iranian societies. This paper further touches etymological, morphological, and syntatic issues concerning Tumsh. *añix$_6$ā*/*anix$_6$ā* 'in another way', *ñiwāne* 'complaints (?)', *druhv-* 'to injure', *paṣto* 'under a contract/pact', *bimex$_6$a-* 'servant (?)', *biźo druhv-* 'to falsely accuse', *lāẓa* 'giver', *histanaka-* 'place, estate (?)', *zrap^{h}ā* '± exploitation, maltreatment', and *zrap^{h}i*/*zräp^{h}a* '[a kind of?] penalty'.

With the use of *zrap^{h}ā* '± exploitation, maltreatment', Tumshukese contracts became relatively concise, without the need of further listing various actions that a new owner has the right to do as we see in the contracts from Afghanistan, Turfan and Niya. A wider investigation of the contracts in the ancient world is to be carried out to pursue the question whether the parallels observed in socio-cultural practices may be traced back to a Kushan or pre-Kushan tradition (cf. Sims-Williams 1991, Sims-Williams 1996; Yoshida 2013). This question is not only important to historians but also to linguists.

Abbreviations

acc.	accusative	H	Huang Wenbi Document, see Duan 2021
act.	active	HL	Handlist
Bactr.	Bactrian	impv.	imperative
CKD	Catalog number of Gāndhārī texts: Documents, see Baums & Glass 2002–	instr.-abl.	instrumental-ablative
		intrans.	intransitive
cr	cover-tablet reverse	Khot.	Khotanese
D.A.	Douldour-âqour	m.	masculine
f.	feminine	MP	Middle Persian
Gath.	*Gāθā*s	nom.	nominative

o	obverse	Shugh.	Shughni
OIr.	Old Iranian	Sid.	*Siddhasāra*, see Bailey 1945
OP	Old Persian	Skt.	Sanskrit
Parth.	Parthian	Sogd.	Sogdian
p.c.	personal communication	subj.	subjunctive
PIE	Proto Indo-European	Tumsh.	Tumshukese
Pkt.	Prakrit	uo	under-tablet obverse
pl.	plural	VJ	*Vessantara-jātaka*, see Benveniste 1946
prs.	present		
sg.	singular	Yazgh.	Yazghulami

References

ARAKAWA Masaharu 荒川正晴, 1989. 'Turufan shutsudo "kikushi kōshōkoku jidai no sogu-dobun onna dorei baibai monjo" no rikai wo megutte トゥルファン出土「麹氏高昌国時代のソグド文女奴隷売買文書」の理解をめぐって – A consideration on the Sogdian sale contract of a female slave recently discovered from Turfan', *SIAL* 5, 137–153.

BAILEY, Harold. W., 1945. *Khotanese texts I*, Cambridge.

BAILEY, Harold. W., 1950. 'The Tumshuq Karmavācanā', *BSOAS* 13/3, 649–670.

BAILEY, Harold. W., 1958. 'Languages of the Saka', in *Handbuch der Orientalistik: 1. Abteilung. Iranistik. Linguistik* (ed. K. HOFFMANN *et al.*), Leiden, 131–154.

BAILEY, Harold. W., 1979. *Dictionary of Khotan Saka*, London.

BARTHOLOMAE, Christian, 1904. *Altiranisches Wörterbuch, zusammen mit den Nacharbeiten und Vorarbeiten*, Berlin–New York.

BAUMS, Stefan & Andrew GLASS, 2002–. *Catalog of Gāndhārī texts.* See https://gandhari.org/catalog (Last access: 16 August 2024).

BD I^2 = SIMS-WILLIAMS, Nicholas, 2012. *Bactrian documents from Northern Afghanistan, I: Legal and economic documents* (revised edition), London.

BD II = SIMS-WILLIAMS, Nicholas, 2007. *Bactrian documents from Northern Afghanistan, II: Letters and Buddhist texts*, London.

BENVENISTE, Émile, 1946. *Vessantara Jātaka: Texte sogdien*, Paris.

CANEVASCINI, Giotto, 1990. 'New readings and new words in the Khotanese *Saṃghāṭa-sūtra*', *StIr* 19, 13–20.

CHING Chao-jung & OGIHARA Hirotoshi 2012. 'A Tocharian B sale contract on a wooden tablet', *JIAAA* 5 (2010[2012]), 101–128.

DEGENER, Almut, 1989. *Khotanische Suffixe*, Stuttgart.

DMMP = DURKIN-MEISTERERNST, Desmond, 2004. *Dictionary of Manichaean Middle Persian and Parthian,* Turnhout.

DRAGONI, Federico, 2020. 'The Tumshuqese year of the goat and the Fremdzeichen x_6', *JA* 308/2, 215–223.

DUAN Qing 段晴, 2021. ‘Huang Wenbi faxian de liangjian Jushideyu wenshu 黄文弼发现的两件据史德语文书 – Two Gūẓdiya documents discovered by Huang Wenbi’, *XYWS* 15, 1–18.

ÈDEL’MAN, D. I., 2020. *Ètimologičeskij slovar’ iranskix jazykov*, t. 6 (p–ṛ), Moscow.

EDIV = CHEUNG, Johnny, 2007. *Etymological dictionary of the Iranian verb*, Leiden–Boston.

EMMERICK, R. E., 1965. ‘Syntax of the cases in Khotanese’, *BSOAS* 28/1, 24–33.

EMMERICK, R. E., 1968. *Saka grammatical studies*, London.

EMMERICK, R. E., 1989. ‘Khotanese and Tumshuqese’, in *Compendium Linguarum Iranicarum* (ed. R. SCHMITT), Wiesbaden, 204–229.

EMMERICK, R. E., 2024. *A handbook of Khotanese* (completed and edited by M. MAGGI *et al.*), Wiesbaden.

GERSHEVITCH, Ilya, 1961. *A grammar of Manichean Sogdian*, Oxford.

GHARIB, Badresaman, 1995. *Sogdian dictionary: Sogdian–Persian–English*, Tehran.

GHARIB, Badresaman, 2004. *Sogdian dictionary: Sogdian–Persian–English*, second edition, Tehran.

GOYIBOV, Bobir, Pavel LURJE, Alisher BEGMATOV, Valentina RASPOPOVA & Gaybulla BOBOYOROV, 2022. *Catalogue of Sogdian writings in Central Asia*, Samarkand.

HITCH, Doug, 1985. ‘Tumshuqese transcriptions’, PDF created on December 31, 2008, see https://www.researchgate.net/publication/329030821_Tumshuqese_Transcriptions (Last access: 16 August 2024).

HITCH, Doug, 1988. ‘Penalty clauses in Tumshuqese, Khotanese and the Shanshan Prakrit’, *StIr* 17, 147–152.

HITCH, Doug, 1990. ‘Prakrit administration at Tumshuq’, in *Eastern Asia: Literature and humanities.* Vol. 3 of *Contacts between cultures (selected papers from the 33rd International Congress of Asian and North African Studies, Toronto, August 15–25,1990)* (ed. A. HARRAK *et al.*). New York, 323–326. Available at https://www.researchgate.net/publication/320291220_Prakrit_Administration_at_Tumshuq (Last access: 2 July 2025).

HITCH, Doug, 2009. *The special status of Turfan*. See https://sino-platonic. org/complete/spp186_turfan.pdf (Last access: 16 August 2024).

KONOW, Sten, 1935. ‘Ein neuer Saka-Dialekt’, *SPAW* 20, 770–823.

KONOW, Sten, 1947. ‘The oldest dialect of Khotanese Saka’, *Norsk tidsskrift for sprogvidenskap* 14, 156–190.

LAZARD, Gilbert, Frantz GRENET & Charles de LAMBERTERIE, 1984. ‘Notes bactriennes’, *StIr* 13, 199–232.

LIV^2 = RIX, Helmut, 2001. *Lexikon der Indogermanischen Verben: Die Wurzeln und ihre Primärstammbildungen. Zweite, erweiterte und verbesserte Auflage bearbeitet von Martin Kümmel und Helmut Rix*, Wiesbaden.

LKD = BURROW, Thomas, 1937. *The language of the Kharoṣṭhi documents from Chinese Turkestan*, Cambridge.

MACKENZIE, D. N., 1971. *A concise Pahlavi dictionary*, London.

MACUCH, Maria, 2022a. 'Pahlavi legal documents from Tabarestān: The juristic context of TAB. 16, 19, 20, 22BIS and 25', *Res orientales* 29, 131–160.

MACUCH, Maria, 2022b. 'A legal transaction in the Zoroastrian community and its social consequences', in *The reward of the righteous: Festschrift in honor of Almut Hintze* (ed. A. CANTERA *et al.*), Wiesbaden, 131–160.

MAUE, Dieter, 1996. 'Zu den uigurischen und iranischen Brāhmī-Handschriften der Berliner Turfanfunde', *Turfan, Khotan und Dunhuang* (ed. R. E. EMMERICK *et al.*), Berlin, 211–220.

MAUE, Dieter, 2004. 'Konows Zeichen Nr. 10', in *Turfan revisited. The first century of research into the arts and culture of the Silk Road* (ed. D. DURKIN-MEISTERENST *et al.*), Berlin, 208–212.

MAUE, Dieter, 2009a. 'Corpus of Tumshuqese fragments: introduction', see https://titus.uni-frankfurt.de/texte/iranica/tumshuq/introduction.pdf (Last access: 11 February 2022).

MAUE, Dieter, 2009b. 'Corpus of Tumshuqese fragments: handlist', see https://titus.uni-frankfurt.de/texte/iranica/tumshuq/handlist.pdf (Last access: 11 February 2022).

MAUE, Dieter, 2009c. 'Corpus of Tumshuqese fragments: transliteration', see https://titus.unifrankfurt.de/texte/iranica/tumshuq/transliteration.pdf (Last access: 11 February 2022).

MAUE, Dieter, 2022. 'Konow's sign no. 3', paper read at the workshop "Tocharian and Iranian in the Tarim Basin and beyond", Leiden, 23/06/2022.

MAUE, Dieter & Nicholas SIMS-WILLIAMS, 2024. *Sogdian documents in Brāhmī script*, London.

MW = MONIER-WILLIAMS, Monier, 1899. *A Sanskrit-English dictionary: Etymologically and philologically arranged with special reference to cognate Indo-European languages* (new edition, greatly enlarged and improved), Oxford.

NYBERG, H. S., 1974. *A manual of Pahlavi, part II: Glossary*, Wiesbaden.

OGIHARA, Hirotoshi, 2015. 'Kuchean secular documents in the Pelliot Koutchéen Nouvelle Série', *TIES* 16, 81–105.

OGIHARA, Hirotoshi, 2019. 'Tumshuqese imperfect and its related forms', in *Proceedings of the Eighth European Conference of Iranian Studies*. Vol. 1 (ed. P. B. LURJE), Saint Petersburg, 297–310.

OGIHARA, Hirotoshi, 2020. 'Miscellany on the Tumshuqese documents', *CAJ* 63/1–2, 11–24.

OGIHARA Hirotoshi 荻原裕敏 & CHING Chao-jung 庆昭蓉, 2012. 'Xinchu Jushideyu siyuan qiyue wenshu ji qita 新出据史德语寺院契约文书及其他 – A newly found Tumshuqese monastic contract and other related findings', *XYWS* 7, 43–61.

OGIHARA Hirotoshi 荻原裕敏 & CHING Chao-jung 慶昭蓉, 2014. 'Shinshutsu tumushukugo keiyaku monjo ni tsuite 新出トゥムシュク語契約文書について – On the Tumshuqese contracts newly published in China', *SIAL* 29, 7–55.

OGIHARA Hirotoshi & CHING Chao-jung, 2017. 'Some observations on the Tumshuqese documents newly published in China', in *Zur lichten Heimat: Studien zu Manichäismus, Iranistik und Zentralasienkunde im Gedenken an Werner Sundermann* (ed. 'Turfan-forschung'), Wiesbaden, 453–482.

OGIHARA Hirotoshi & CHING Chao-jung 2024. 'Miscellany on the Tumshuqese documents (III): Words relating to cereals', in *Siddham. Studies in Iranian philology in honour of Mauro Maggi* (ed. G. BARBERA *et al.*), Wiesbaden, 305–323.

SANDER, Lore, 1968. *Paläographisches zu den Sanskrithandschriften der Berliner Turfan-sammlung*, Wiesbaden.

SANDER, Lore, 2009. 'Remarks on the formal Brāhmī script from the Southern Silk Route', *BAI* 19 (2005[2009]), 133–144.

SIEG, Emil & Wilhelm SIEGLING, 1953. *Tocharische Sprachreste. Sprache B, Heft II. Fragmente Nr. 71-633*, aus dem Nachlaß herausgegeben von Werner Thomas, Göttingen.

SIMS-WILLIAMS, Nicholas, 1991. 'A Sogdian greeting', in *Corolla Iranica. Papers in honor of Prof. Dr. David Neil MacKenzie on the occasion of his 65th birthday on April 8th, 1991* (ed. R. E. EMMERICK *et. al.*), Frankfurt am Main–New York, 176–187.

SIMS-WILLIAMS, Nicholas, 1996. 'From Babylon to China: Astrological and epistolary formulae across millenia', in *La persia e l'Asia centrale da Alessandro al X secolo* (ed. Academia Nazionale dei Lincei), Roma, 77–84.

SIMS-WILLIAMS, Nicholas, 2018. 'The wisdom of Aḥiqar and the wisdom of Ādurbād: A Manichaean parallel', in *A thousand judgements. Festschrift for Maria Macuch* (ed. A. HINTZE *et al.*), Wiesbaden, 363–372.

SIMS-WILLIAMS, Nicholas, 2025. *An Old Khotanese reader: The tale of Bhadra* (with contributions by Jonathan A. SILK), Wiesbaden.

SIMS-WILLIAMS, Nicholas & François de BLOIS, 2018. *Studies in the chronology of the Bactrian documents from northern Afghanistan, with contributions by Harry Falk and Dieter Weber*, Wien.

SKJÆRVØ, P. O., 1987. 'On the Tumshuqese *Karmavācanā* text', *JRAS* 1987/1, 77–90.

TKD = BURROW, Thomas, 1940. *A translation of the Kharoṣṭhi documents from Chinese Turkestan*, London.

TPR = Tumushukeshi wenwuju 圖木舒克市文物局 Bureau of Cultural Relics of Tumshuk City, Beijing daxue zhongguo gudaishi yanjiu zhongxin 北京大學中國古代史研究中心 Center for Research on Ancient Chinese History, Peking University & Zhongguo renmin daxue guoxueyuan xiyu lishi yuyan yanjiusuo 中國人民大學國學院西域歷史語言研究所 Institute for Historical and Philological Studies of China's Western Regions, Renmin University of China, 2014. 'Xinchu sanjian Jushideyu qiyue 新出三件據史德語契約 – Three newly found Tumshuqese sale contracts', *Xiyu lishi yuyan yanjiu jikan* 西域歷史語言研究集刊 – *Historical and Philological Studies of China's Western Regions* 7, 63–105.

YAKUBOVICH, Ilya, 2006. 'Marriage Sogdian style', in *Iranistik in Europa – Gestern, Heute, Morgen* (ed. H. EICHNER *et al.*), Wien, 307–344.

YOSHIDA Yutaka, 2013. 'When did Sogdians begin to write vertically?', *Tokyo University Linguistic Papers* 33, 375–394.

YOSHIDA Yutaka 吉田豊, MORIYASU Takao 森安孝夫 & Xinjiang Uighur Autonomous Museum [sic!] 新疆ウィグル自治区博物館, 1989. 'Kikushi Kōshōkoku jidai sogudobun onna dorei baibai monjo 麹氏高昌国時代ソグド文女奴隷売買文書 – A Sogdian sale contract of a female slave from the period of the Gaochang 高昌 kingdom under the rule of Qu 麹 clan', *SIAL* 4, 1–50.

On the so-called "Fremdvokal" *ä* in Tocharian and Khotanese and its origins*

Michaël PEYROT

The Tocharian languages, attested in manuscripts from the 2nd half of the 1st millennium CE from the Tarim Basin in Northwest China, have a so-called "Fremdvokal", a weak vowel close to shwa that is transcribed with *ä*. In the Tocharian Brāhmī script, *ä* is spelled in part with special consonant signs, the so-called "Fremdzeichen", and in part with a vowel diacritic consisting of two dots. Khotanese, spoken in the southern Tarim Basin, and attested in manuscripts from approximately the same period, also has a vowel diacritic, originally consisting of two dots, that is transcribed with *ä*. It is argued that the sound value of Tocharian *ä* was close enough to that of Khotanese to be derived from it. Khotanese *ä* in turn seems to derive from a special use of visarga, *ḥ*, in Khotan Prakrit, the local variant of Gāndhārī written in the Kharoṣṭhī script.

1. Introduction

In the Tarim Basin of the first millennium of the Common Era, the local languages Tocharian and Khotanese were written down in especially adapted varieties of the Brāhmī script originally developed for Sanskrit. Although Tocharian and Khotanese are both Indo-European, they are not closely related, so that their phonological systems were quite different, and the Brāhmī script thus needed to be adapted for each in a specific way. Indeed, many script and spelling conventions as well as the palaeography of Tocharian and Khotanese Brāhmī are so different that they have clearly been developed independently.

Yet, the introduction of the Brāhmī script to write the local languages occurred approximately at the same time, probably in the early fifth century of the Common Era,[1]

* This article is an adaptation of a lecture held in the *Collegium Turfanicum* series (number 78) of the *Turfanforschung* of the Berlin-Brandenburgische Akademie der Wissenschaften, Berlin, on 10 December 2015. The research was first supported by a Marie Curie Intra European Fellowship (P7-PEOPLE-2013-IEF 626656) and then by the European Research Council (ERC-2022-COG 101088902). I thank Ching Chao-jung (Kyoto), Federico Dragoni (Leiden) and Niels Schoubben (Leiden) for valuable comments on an earlier draft.

[1] See notably Maggi (2004), Dragoni (this volume), Malzahn (2007b: 277) and Peyrot (2008: 204–206). These datings are no longer controversial, but remain approximate. In any case, we should most probably allow some time between the development of the script and the earliest preserved manuscripts.

10.29091/9783752003635/008

and there are also commonalities between the two Brāhmī varieties that set them apart from Sanskrit. In the following, I will investigate what is probably the most striking match of this kind between the two Brāhmī varieties: the vowel diacritic conventionally transcribed as *ä*. Comparing its use and sound value, I will study the origin, development and spread of this vowel diacritic, so shedding further light on the interrelationship between the different Brāhmī varieties of the Tarim Basin as well as their relationship to Kharoṣṭhī, the other large Indic script.

2. The Tocharian "Fremdvokal" *ä*

Tocharian Brāhmī famously has a "Fremdvokal" that is transcribed as *ä*, a vowel that does not occur in Sanskrit. The Fremdvokal is sometimes written with a special vowel diacritic, transliterated <ä>, and sometimes it is the inherent vowel of special consonant signs, the so-called "Fremdzeichen" (e.g. Sieg & Siegling 1908: 919; Malzahn 2007a: 223–225). Fremdzeichen are transliterated with an underline, so that, for instance, <k̲a̲> stands for *kä*, etc.

In Brāhmī syllabaries from the Tocharian areas, the Fremdzeichen are sometimes found listed as an addition to the simple akṣaras, for example in IOL Toch 200 (Fig. 1, cf. also e.g. SHT 794), where the second line of the recto reads /// *m̲a ś̲a ṣ̲a s̲a wa r̲a* **:** /// (Couvreur 1965: 119):[2]

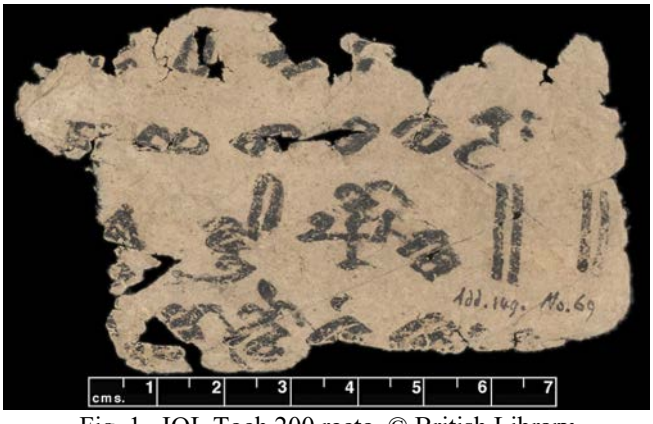


Fig. 1. IOL Toch 200 recto. © British Library.

Thus, if the earliest manuscripts are from the early 5th century, the development of the scripts may have to be dated in the late 4th century.

[2] Cf. also Filliozat (1948: 25), where it is cited as "India Office, Add. 149 (?), nº 69". In this syllabary, <wa>, another Tocharian addition to the Sanskrit Brāhmī inventory, is included among the Fremdzeichen. Current scholarship varies as to whether <wa> should be called a Fremdzeichen: it is, likewise, an addition to the original Sanskrit akṣara inventory, but unlike <k̲a̲> etc., it does not have an inherent vowel *ä*.

Akṣaras with a corresponding Fremdzeichen are *k*, *t*, *n*, *p*, *m*, *r*, *l*, *ś*, *ṣ*, *s* (the digraph <tsa> has a Fremdzeichen variant <t̲s̲a̲>). No Fremdzeichen counterparts exist for akṣaras that are only used in Sanskrit loanwords, such as *kh*, *g*, etc. However, the following akṣaras used in genuinely Tocharian words have no Fremdzeichen counterpart either: *ṅ*, *c*, *ñ*, *y*, *w*. In such cases, the vowel diacritic <ä> must be used.

In spelling, the Fremdzeichen are obligatory. For instance, TB *kätkare* ‘deep’ is always spelled <k̲a̲tkare> with Fremdzeichen <k̲a̲>, never <kätkare> with the vowel diacritic <ä>. Exceptions to these spelling conventions are limited to certain consonants only, e.g. <nä> is often found for <n̲a̲> in Tocharian B and <śä> for <ś̲a̲> in Tocharian A.

In consonant clusters, it is the last consonantal element of the akṣara that determines the vowel. If the last consonantal element is a Fremdzeichen, the preceding element may be a Fremdzeichen or a non-Fremdzeichen, for instance: TB <post̲a̲ṃ> *postäṃ* ‘after’ with Fremdzeichen <t̲> preceded by non-Fremdzeichen <s>; but TB <akāl̲k̲anta> *akālkänta* ‘wishes’ with Fremdzeichen <k̲> preceded by Fremdzeichen <l̲>.[3]

For word-final consonants, Fremdzeichen are obligatory if available, e.g. TB <eṅkaskemt̲a̲r̲ˎ> *eṅkaskemtär* ‘we take’ with Fremdzeichen <r̲ˎ> in virāma position, or TB <akāl̲k̲ˎ> *akālk* ‘wish’ with Fremdzeichen <k̲ˎ> in virāma position. If no Fremdzeichen is available, the *ä* diacritic is used, but it remains silent and does not denote a vowel, and is conventionally transcribed with a superscript ^{ä}, e.g. TB <lac^{ä}ˎ> *lac* ‘s/he went out’.

As is clear from its behaviour in phonology, *ä* denotes a reduced vowel in Tocharian. For instance, in Tocharian B *ä* may disappear through syncope, e.g. TB *mäsketrä* ‘is’ ~ *msketrä*,[4] and it may appear through apparent epenthesis, e.g. TB *srukalñe* ‘death’ ~ *srukalläññe*.[5] In Tocharian A, syncope and epenthesis are very frequent: *ä* regularly disappears from any potentially open syllable and many clusters regularly receive epenthetic *ä*:

mnu ‘wish’, perl. *mänwā* (initial *mn-* is permitted, but *mnw-* is not permitted)
rtär ‘red’, f. *rätri* (initial *rt-* is permitted, but initial *rtr-* and final *-tr* are not permitted)
okät ‘eight’ (final *-kt* is not permitted)

[3] A rare exception appears to be the spelling of *cä* in Archaic texts, where <ś̲ca> and <ṣ̲ca> likely stand for <ś̲ca̲> *ścä* and <ṣ̲ca̲> *ṣcä*, respectively (Peyrot 2008: 179–180).

[4] This particular syncope in a closed syllable is only found in verse.

[5] These forms are verbal abstracts. Verbal abstracts of the first, shorter type are much more frequent than forms of the second, longer type. However, morphologically the longer form seems to be primary, so that this apparent epenthesis is probably rather an archaism, while the shorter forms are the result of syncope.

In view of the high predictability of *ä* in Tocharian A it has been argued that this vowel is not phonemic (Jasanoff 1987: 110; Kim 2007: 2). The above forms may thus be analysed as *mnu* /mnw/, *mänwā* /mnwā/, *rtär* /rtr/, *rätri* /rtry/, *okät* /okt/: the appearance of *ä* is automatic and predictable in all cases.

Another well-known pattern showing that *ä* is a reduced vowel is provided by the accent alternations of Tocharian B, in which accented *ā* is weakened to *a* /a/ when unaccented, and accented *a*, here for convenience phonologically noted as /á/, is weakened to *ä* /ə/ or zero when unaccented, e.g.

palsko /pálsko/ 'thought', pl. *pälskonta* /pəlskónta/
cake /cáke/ 'river', pl. *ckenta* /ckénta/

Finally, alternations between *i* or *u* and zero are best explained as instances of syncope of *ä* /ə/ coloured to *i* or to *u*, e.g. TB 1sg.pret. *prekwa* /prekwa/ ~ *prekuwa* /prekáwa/ 'asked' (cf. 2sg.pret. *prekasta* /prekásta/) or TB *yokye* /yokye/ 'thirst' ~ *yokiye* /yokáye/. The fact that *ä* /ə/ was susceptible to colouring to *i* or *u* further shows that it was a weak vowel.

Even though the exact value of the vowel prototypically represented by <ä> remains difficult to determine, it was weak, susceptible to colouring, and therefore probably close to [ə]. An alternative value [ɨ] is also found in the literature (Jasanoff 1978: 30; Pinault 1989: 37–38), but the arguments in favour of this higher, closed value are the same as those given above. If the vowel was weak and susceptible to colouring, it was very likely to some extent variable in phonetic realisation, also if not coloured.

3. *ä* in Old Uyghur Brāhmī

Old Uyghur is written in a variety of scripts, the most important being the Uyghur script, which is derived from the Sogdian script. A minor, but still considerable part of the Old Uyghur corpus is written in a variety of Brāhmī that is palaeographically very similar to Tocharian Brāhmī. Old Uyghur Brāhmī has its own spelling conventions, which are needed especially to render the front-rounded vowels *ö* and *ü* and the open front vowel *ä* (not to be confused with Tocharian <ä> /ə/),[6] but there are also parallels with the conventions of Tocharian.

[6] The development of Brāhmī strategies to write Old Uyghur *ä*, *ö* and *ü* are not the topic of this paper. The spelling with an additional <y> after consonants is easy enough to understand, in particular in the case of the velars, which were probably slightly palatalised, e.g. <kya> *kä*, <kyo> *kö*, etc. From there, the more complicated <yyo> *yö* may easily have been generalised. The spelling of initial front vowels is peculiar, with combinations such as <eya> *ä-*, <oyo> *ö-*, <oya> *ö-*, etc., with two vowels in one akṣara. However, the combination of two vowels in one akṣara finds a model in Tocharian akṣaras of the type TochA <$k_u$pre> /kwpre/ 'whether', <$k_u$li> /kwli/ 'woman', <$_a$ntā> /ə̃nta/? 'when' (Sieg &

In accordance with the Tocharian spelling rules, Old Uyghur Brāhmī uses Fremdzeichen with the accompanying virāma stroke in word-final position instead of the regular non-Fremdzeichen consonants. Thus, we find word-finally <k̲, t̲, p̲, m̲, r̲, ś̲, ṣ̲, s̲> instead of <k, t, p, m, r, ś, ṣ, s>. Gabain calls the Fremdzeichen in this use "Vokallose Endzeichen" (1974: 34). She further notes that in Old Uyghur Brāhmī, as in Tocharian, "wird in manchen Hss. auch ein Doppelpunkt gesetzt, um mit dem Verbindungsstrich eine Vokallosigkeit anzuzeigen" (1974: 38). An example of the latter use is *adırtlamak* 'distinguishing', written <atirtlāmāhkh^{ä}˴> in Tocharian based transliteration, and transliterated by Gabain as <a ti rtlā mā hk̈h>.

Apart from the systematic use of Fremdzeichen for word-final consonants, there is limited use to denote a vowel. While the regular spelling of *i* and *ı* [ɨ] is with <i, ī>, Maue (1996: xxiii) notes that spellings with the "Fremdvokal" are found in the suffix *-lıg*, *-lig*, where *lı* or *li* is written as <l̲a, l̲ä, l̄a̲ l̄ä̲>,[7] and he writes, "Es liegt nahe, anzunehmen, daß diese Grapheme *l* mit einem inhärierenden *i*-ähnlichen, vielleicht reduzierten Vokal darstellen. Das steht im Einklang mit den Verhältnissen des Tocharischen und der Ableitung aus den silbischen Liquiden 'l̥, l̥̄' des Indischen."

Written with Fremdzeichen, *ä* is further found to stand for /ı/ in e.g. 24 v6 *sarıg* <sār̲aγ˴> 'yellow' (Maue 1996: 129), *arıg* <ār̲aγ˴> 'pure' (Gabain 1954: 58 / I.15), and *tonlarıg* <toṃlār̲aγ˴> 'clothes' (Gabain 1954: 59 / I.20). It stands for zero in 3c r4 *savıklagalı* <syāwehk̲älāγālī> 'to talk incoherently' (Maue 1996: 42). Gabain assumes that *ä* denotes a "Murmelvokal" in e.g. *burhannıŋ* <būrhāṃnäṅ˴> (her transliteration is <būrhāṃn̈əṅ>) 'of the Buddha' and *kapıg* <qaphäg˴> (her transliteration is <qapḧəg>) 'gate' (1954: 4).

In my view, it is questionable whether Old Uyghur Brāhmī <ä> really denotes an *i*-like vowel, as Maue suggests. Importantly, most examples of the spelling <l̲aγ˴> are found in back-vocalic words, where the shape of the suffix is expected to be *-lıg* with the back vowel *ı* [ɨ]. The front-vowel variant *-lig* is almost always written <liγ˴>, <līγ˴>. The only exceptions seem to be 26a r7 <·iliγ˴l̲aγ˴> *biliglig*; 26a v2, v4 <pyutyurmyāk̲˴l̄ä̲γ˴> *bütürmäklig*; and 26a r2 <yyormyāk̲˴l̄ä̲γ˴> *yörmäklig* (all these examples Maue 1996: 136). Since <ä> mostly stands for /ı/ and only rarely for /i/, there is no reason why <ä> would be closer to /i/ than to /ı/.

Siegling 1908: 921; Peyrot 2018: 74; Wilkens & Peyrot 2024). The Old Uyghur initial vowels would according to the conventions for Tocharian Brāhmī have to be transliterated as e.g. <ₒya>, <ₒyo>, <ₒya>.

[7] Maue's transliteration <l̄a̲> with a "long" Fremdzeichen is based on the fact that <l̲a> derives from Sanskrit <l̥> (initial vocalic *l*) and <l̄a̲> from Sanskrit <l̥̄> (initial long vocalic *l*). It is not likely that this rare akṣara really denoted a long vowel. Gabain had transliterated <l̄a̲> as <lr̄>, which is clearly wrong; however, it is understandable since <l̄> looks like a combination of <l̲> and <r̲>.

Also, the assumption that <ä> is an "*i*-like vowel" does not fit the evidence from Tocharian. On the contrary, a value /ɪ/ would fit Tocharian <ä> quite well, and if Old Uyghur Brāhmī <ä> denoted a reduced variant of /ɪ/, this would fit the situation in Tocharian perfectly.

The spelling conventions of Old Uyghur Brāhmī are far from strict and there is much variation, but even so a special case is U 6889 (catalogue Nr. 89), a manuscript that does not follow the common patterns of spelling to be noted elsewhere (Maue 2015: 33–38). In this manuscript, <ä, a̲> are used more often. In most cases, <ä, a̲> render Old Uyghur *ä*, for instance: 89 r2 <k̲a̲siγ$_{[\diagdown]}$> *käzig* 'sequence' (Maue 2015: 41, <k̲a̲ si g$_{1}$>); 89 r2 <k̲a̲smakiṃtaki> *käsmäkindäki* 'while cutting' (Maue 2015: 42); 89 r3 <1 2 ndiśakä> *bir ikindiškä*[8] 'together' (Maue 2015: 43); 89 r4 <yāṃa̲> *yemä* 'and' (Maue 2015: 45); 89 v1 <irak̲a̲ga> *erkäk* (or *irkäk*) 'man' (Maue 2015: 49–50); 89 v4 <āmak̲a̲tamak̲a̲> *ämgätmäk* 'torment' (Maue 2015: 52, the first <k̲a̲>). In a smaller number of cases, <ä, a̲> is "silent", i.e. the virāma stroke has been left out, for instance: 89 v2 <diśa til̲a̲k$_{u}$o> *tiši tilkü* 'vixen, she-jackal' (Maue 2015: 51–52, <di śa ti l̲a̲ kuo>); 89 v3 <$_{u}$yumak̲a̲dhiṃ> *ümäkdin* '?' (Maue 2015: 52, <ʾyu ma k̲a̲ d^{h}iṃ>); 89 v4 <āmak̲a̲tamak̲a̲> *ämgätmäk* 'torment' (Maue 2015: 52, the second <k̲a̲>). There is no easy explanation for the use of <ä, a̲> for *ä* in this manuscript, other than that the spelling on the whole is highly unusual and unsystematic.

4. *ä* in Sogdian Brāhmī

There are some Sogdian manuscripts in Brāhmī too, but the number of fragments preserved is very small. Compared to Old Uyghur Brāhmī, which makes up only a small part of the Old Uyghur corpus, Sogdian Brāhmī is markedly less frequent.

In Sogdian Brāhmī, final *k*, *t*, *p*, *m*, *r*, *l*, *ś*, *ṣ* and *s* are, as in Old Uyghur and Tocharian, written with the corresponding Fremdzeichen: <k̲$_{\diagdown}$, t̲$_{\diagdown}$, p̲$_{\diagdown}$, m̲$_{\diagdown}$, r̲$_{\diagdown}$, l̲$_{\diagdown}$, ś$_{\diagdown}$, s̲$_{\diagdown}$, s$_{\diagdown}$> (Maue & Sims-Williams 1991; Sims-Williams 1996: 309, 311; Maue & Sims-Williams 2024: 30, 36). We find, for instance, SogdBr 1 r6 <piṃ zirk̲$_{\diagdown}$> /pīn zirk/ 'yellow cream' (earlier read /pēn zerk/), and SogdBr 2 v.e <āp̲$_{\diagdown}$> /āp/ 'water'. Only for final *t* is the normal non-Fremdzeichen akṣara used, in a clearly smaller number of cases (Maue & Sims-Williams 2024: 30, fn. 45). In a few such cases, the two dots are found in the same function as Fremdzeichen <t̲> (l.c., fn. 42), for instance in SogdBr 5 B4 <pcāyt^{ä}$_{\diagdown}$> /p(ə)čāit/ 'is useful'. As in Tocharian and Old Uyghur, final *-c* does not have a Fremdzeichen and is written <c$_{\diagdown}$>, e.g. SogdBr 5 A4 <pizāc$_{\diagdown}$> *pīzāč* 'crushed'; and final *-n* is written <ṃ> (Maue & Sims-Williams 2024: 27).

[8] The "1" and "2" in the original text are to be read out in full as *bir iki* in order to arrive at a coherent interpretation.

In a few cases, the vowel <ä, a̲> occurs, and it then most probably stands for shwa (Maue & Sims-Williams 2024: 28–29). Examples in native words are SogdBr 11 B2 <pa̲cmr[]o> *pəčmr°* 'count', SogdBr 11 A4 <γärwi> *γərvi* 'much' (Maue & Sims-Williams <g$_1$ä *r*wi>) and SogdBr 11 B2 <phäthkuṃ v$_1$eñe> *pətkūn-wēne* 'having heretical views'. It is further found in the loanwords SogdBr 5 B2 <durāla̲m꞉> *durāləm* 'camel thorn' and Sogd Br 4 Bc <tr̥phāl̲꞉> *trəpāl* 'the three myrobalans'.[9] Maue & Sims-Williams attribute the use of <ä, a̲> in these loanwords to the model of Tocharian. Indeed, <tr̥phāl̲꞉> *tr̥phāl* (or *träphāl*) is well attested in Tocharian B. However, even though it is not decisive that no Tocharian (B?) *durāläm* is attested (Maue & Sims-Willimas), it is problematic that a Tocharian borrowing from Skt. *durālambhā* would rather be expected to be approximately **duralāmbh*. Since the *r̥* in <tr̥phāl̲꞉> *trəpāl* may equally be based on Sanskrit spellings, it is perhaps better to see, for the time being, in <durāla̲m꞉> *durāləm* a genuinely Sogdian treatment of the Sanskrit word.

A special case is SogdBr 2 v.c <rust̲^{ä}꞉>, in which Sims-Williams earlier already noted, "a final <ä> may represent a vowel in *ru-st̲ä* (2), if this stands for [rusti] "has grown" " (1996: 309, 311). The problem is that the virāma is very clear (Fig. 2).

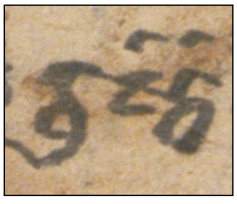


Fig. 2. SogdBr 2 v.c <rust̲^{ä}꞉>.
© Depositum der Berlin-Brandenburgischen Akademie der Wissenschaften in der Staatsbibliothek zu Berlin – Preussischer Kulturbesitz Orientabteilung.

For this word, Maue & Sims-Williams (2024: 39, 55, fn. 177) therefore consider the option that the virāma here is a mistake, and that <st̲ä> stands for [sti]. However, as they have made clear, and as summarised above, <ä> is very rare, and in the few cases it is found, it does not stand for /i/.

Perhaps the most striking feature of Sogdian Brāhmī is that the most frequent spellings for shwas earlier posited on the basis of the Sogdian and Manichaean scripts are <a> and zero. Examples of spellings with <a> are (Maue & Sims-Williams 2024: 29, fn. 29, p. 40): SogdBr 1 r3 <γarwak̲꞉> *γə^{r}vāk* 'wise' (Maue & Sims-Williams <g$_1$a rwa-k̲>), SogdBr 11 A3 <γarwi> *γə^{r}vī*, *γə^{r}vi* (p. 89) 'much (?)' (Maue & Sims-Williams <*g$_1$a r*wi>, next to SogdBr A4 <γärwi>), SogdBr 3 r3 <pat̲꞉pδi> *pətpδi* 'long pepper' (Maue & Sims-Williams <pa-t̲ pd$_1$i>, SogdBr 3 r1 <patyāp̲꞉> *pətyāp* 'amount'. In initial

[9] Maue & Sims-Williams (2024: 29) further note SogdBr 5 B3 <ma̲t̲꞉> *meθ* 'with', which they call an "unexpected spelling".

open syllables, shwa and other short vowels indicated by the Sogdian and Manichaean scripts are frequently not written. Sims-Williams called this "the frequent non-writing of short vowels" and noted a shwa in brackets (1996: 311). In the most recent edition of Maue & Sims-Williams (2024), these shwas have been entirely deleted from the transcription because it is impossible to know precisely when a shwa was there or not (p. 41): SogdBr 2 v.d, v.g <knā> earlier [k(ə)nā], now /knā/ 'make!'; SogdBr 2 v.g <mdhu> earlier [m(ə)δu], now /mδu/ 'wine'; SogdBr 3 r1, 3 r5 <mγoṃ> earlier [m(ə)γōn], now /mγōn/ 'equal'; SogdBr 5 b3 <pcāyt^{ä}ꞈ> (next to other spellings, but all with initial *pc*-) earlier [p(ə)čāit], now /pčāit/ 'is useful'.

5. *ä* in Tumšuqese Brāhmī

The diacritic <ä> also occurs in Tumšuqese Brāhmī. Fremdzeichen in principle do not occur, and <ä> is found on all akṣaras, e.g. (Maue handlist; Skjærvø 2002: 446–447, 576): <ä, jä, tä, δä, nä, pä, mä, yä, wä, ṣä, sä, ẓä, hä>. The Fremdzeichen <t̲a, n̲a, p̲a, m̲a, ṣ̲a, s̲a> are not attested at all. The only clear exception is <r̲a>, which is found frequently, while <rä> is rare if not completely absent (Maue & Sims-Williams 2024: 37).[10] There are also a few instances of <k̲a> (Maue & Sims-Williams 2024: 37), e.g. TS 17 b3. A further instance is IOL Khot 203/9a4 (Fig. 3), where Skjærvø reads (2002: 447): */ khisti pä o -i [x]* for what clearly seems to be */// khi sti pä k̲a bi ///*.[11]

Fig. 3. IOL Khot 203/9a4 (detail). © British Library.

However, also for <k> the normal spelling is with <ä>, cf. for instance <kä> in Pelliot Div. 410, TS 16 a1. Finally, a rare case of <ṭ̈> is attested in TS 7c b1.

According to Maue & Sims-Williams (2024: 37), the value of <ä> in Tumšuqese is *ə*, but no detailed argument is given, and the precise phonetics are difficult to establish. According to Emmerick (2024: 284), it notes *ə* for original unstressed **a*.[12]

[10] Possible instances of <rä> are TS 16 a4 and <ndr*ä*> in DA 6 b3.a1. Cf. further IOL Toch 185 a4, where <ẓä> may have to be read.

[11] Maue (handlist) reads "[...] k^{h}i si-ś pä *j*a bi [...]", but I think that the character under <si> is too small for <ś>; in my view, Skjærvø is right that the lower part is <t> and the whole akṣara <sti>.

[12] It is further suggested that "Other vowels may have had an allophone [ə] in unstressed syllables" (Emmerick 2024: 284). Unfortunately this remains unclear to me in the detail.

6. The Tocharian origin of word-final Fremdzeichen and “silent” -*ä*

As laid out above, in Tocharian as well as in Old Uyghur and Sogdian, Fremdzeichen are in word-final position preferred to their non-Fremdzeichen counterparts, or a “silent” -*ä* may be added. This phenomenon must be connected to the occurrence of word-final vowel -*ä* in early Tocharian B, which I will now discuss in more detail.

In Tocharian A, no final vowel -*ä* regularly occurs: words end either in another vowel, or in a consonant, but not in the vowel -*ä*. In Classical Tocharian B, the situation is very similar, but there is more variation in general, and there are more exceptions to this rule too. Final -*ä* is found occasionally, especially in verse. Moreover, in verse many words normally ending in a consonant sometimes show an additional final -*o*, called “mobile *o*” (Malzahn 2012), for instance *kektseñ* (obl.sg.) ‘body’ has a verse variant *kektseño*.

This “mobile *o*” is real because it appears precisely there where the verse requires an extra syllable, and also there where the accent requires one according to the analysis of Marggraf (1970). For instance, in the preterite of *nes-* ‘to be’, formed from the suppletive stem *taka-*, all trisyllabic forms have their accent on the second syllable, while the disyllabic 3sg. has it on the first syllable:

takāwa	/takáwa/	‘I was’
takāsta	/takásta/	‘you were’
tāka	/táka/	‘he was’

However, if a pronoun suffix is added, the accent shifts to the second syllable in the 3sg. as well, and this also happens with the pronoun suffix of the 2sg., which normally is only -*c*, but has a longer variant -*co*, where the necessary extra syllable can still be seen:

takā-ne	/takáne/	‘it was to him’
takā-c	/takác/	‘it was to you’, next to *takā-co*

Thus, it is generally assumed that “mobile -*o*” corresponds to an earlier final shwa that was otherwise lost in Classical Tocharian B, but may make an extra syllable in verse, and still shows its effect on the placement of the accent.

In the oldest stage of Tocharian B, Archaic Tocharian B, the situation is different. Where Classical Tocharian B has mobile -*o*, Archaic Tocharian B usually has final -*ä* instead, e.g. *kektseñä*, *takā-cä*. In addition, there are more forms ending in -*ä* in general, also in prose. In a very old manuscript that may be the oldest preserved for Tocharian B (Malzahn 2007b: 267, on THT 2668 etc.), we even find that almost all words that would have ended in a consonant in Classical Tocharian B rather end in -*ä*, for example:

THT 4122 a3

/// (vira)[lo]nä • ṣe pāke 1 se curn̲a naracak̲a ñem̲a • /// (Peyrot 2014: 141)

'... *viḍalavaṇa* [salt]: one part 1. This powder [has] the name *naracaka*.'

In this small extract, *(vira)lonä* would be *viraloṃ* in Classical Tocharian B, *curnä* would be *curn* or *curm* (the change of *-n* to *-m* does not matter to the argument made here), *naracakä* would be *naracak*, and *ñemä* would be *ñem*.

Thus, in the oldest Tocharian B, almost all words that later would end in a consonant still ended in *-ä*. This final *-ä* was written with the vowel diacritic <ä> or with the corresponding Fremdzeichen. In the course of the development of Archaic to Classical Tocharian B, the final *-ä* was lost and virāma introduced, but the spelling with Fremdzeichen or *-ä* was kept, so making *-ä* both "silent" and redundant. It is difficult to be fully certain about the phonetic value of each and every instance of final *-ä* in Archaic Tocharian B prose, as it is theoretically possible that virāma was simply not written, but for *-ä* in verse the syllable counting shows that *-ä* was indeed there in Archaic Tocharian B, and lost onto Classical Tocharian B.

As an intermediate summary, the following points can be highlighted. In Tocharian, there is an extra vowel <ä>, probably [ə], which is written in part with a vowel diacritic, and in part with separate akṣaras, the so-called Fremdzeichen. Word-finally, the vowel diacritic and the Fremdzeichen are used even when there seems to be no need, but in older Tocharian B, many words ended in *-ä*, which explains word-final *-ä*.[13]

Based on the spelling conventions and approximate sound value of *ä*, the other non-Sanskrit Brāhmī varieties of the Northern Tarim Basin appear to derive in essence from the Brāhmī variety of Tocharian B:

- Tocharian A Brāhmī has the same conventions as Classical Tocharian B.
- Old Uyghur Brāhmī has adopted word-final *-ä* as well as word-final Fremdzeichen. The marginal use of *ä* in Old Uyghur points to /ɪ/ or a reduced form of it. Old Uyghur Brāhmī cannot be derived from Tocharian B Brāhmī only, as for instance Old Uyghur <l̲a> is not used for Tocharian B, and it must have been taken from Sanskrit Brāhmī.
- Sogdian Brāhmī uses word-final Fremdzeichen and sometimes word-final *-ä*. As a vowel, *ä* is rare, and mostly stands for shwa, although shwa where it was earlier supposed to occur is in Sogdian Brāhmī mostly written with *a* or zero.
- Tumšuqese Brāhmī uses the vowel diacritic <ä>, but only marginally Fremdzeichen. The extended use of the vowel diacritic in Tumšuqese next to marginal Fremdzeichen is difficult to understand, but Tocharian influence on Tumšuqese is otherwise assured.

[13] Thus, the use of Fremdzeichen for word-final consonants is a nontrivial feature that can be used to trace the development and spread of the Brāhmī varieties of the Northern Tarim Basin. There is no reason to consider it "pretentious" as per Maue and Sims-Williams (2024: 36).

7. *ä* in Khotanese Brāhmī

In the southern Tarim Basin, in Khotanese Brāhmī, there is also a diacritic <ä>. In the case of Khotanese, this diacritic is not in competition with any Fremdzeichen, and it is consistently used on all akṣaras. The problem is rather what the exact phonetic value of <ä> is.[14]

The current consensus seems to be that <ä> has the value [ə], e.g. Emmerick (2024: 8, 287; and cf. already Emmerick 1998: 93). The best argument in favour of this value is that *ä* often notes an unaccented vowel resulting from some sort of vowel reduction. Probative in particular is the reduction of unstressed *a* to *ä* as in *hvatäna-* 'Khotanese' < *hvatana-*, which is most straightforwardly interpreted as a reduction of *a* to *ə*. In my view, the phonetic value *ə* is probably correct for most phases of Khotanese, but it is not the most likely option for the oldest stage. Therefore, an explanation of the origin of the diacritic <ä> should not be based on the value [ə].

Originally, Emmerick had argued for a quite different value of *ä*: a closed [ẹ]. According to him, this value was likely because the two dots "are used to indicate a vowel sound that arose as the result of palatalisation of Old Iranian [a]: Khotanese *mästa-* "great" from **masita-* (cf. Av. *masita-*). The resultant vowel sound merged early with [i]: *mista-* "great" (Emmerick 1989: 207–208). In my view, it is certainly possible that a fronted and raised *a* develops to *ẹ*, as would have happened here with *i*-affection or "*i*-Umlaut". However, there are certainly more possible outcomes of such an affection, and the value *ẹ* does not provide the most obvious phonetic interpretation of *ä* for the other sound changes resulting in it:

- *ä* may go back to final **i*, after causing palatalisation of a preceding consonant or *i*-affection in the preceding syllable, e.g. *vīrä* 'upon' < **upari*;
- nom.sg.m. *-ä* goes back to **-ah*;
- *är*, *äl* are possible outcomes of **r̥* (next to *arr*, *ir*, *il*, *ur*, *urr*, *ul*, *ri*), e.g. **kr̥ta-* > *yäḍa-* 'done'; **mr̥ya-* > *mär-* 'die'.

A further argument for the phonetics of *ä* comes from another well-known early sound change in historical Khotanese, by which *ä* and *i* merged. Among other things, this caused confusion between nom.sg. *-ä* < **-ah* and gen.sg. *-i* < **-ahya*.

[14] In Leumann's table of Khotanese Brāhmī (1934: 17), the <ä> diacritic consists of two dots over the akṣara, similar to the Tocharian diacritic. However, it may also have the shape of an arc (e.g., Dragoni, this volume, p. 111). The two dots must represent the original shape of the diacritic, since that is the variant found in the oldest material identified by Maggi (2022), e.g. SI P 83.2 on p. 156. See in particular the discussion and the selected akṣaras in Dragoni (this volume, p. 109–111).

A piece of evidence that is not straightforward is the development of *-*akah* to a sound rendered with a diacritic transcribed as <ei>. This transcription is conventional, and it should not be taken for granted that it stands for a diphthong *ei*, but nevertheless this or a similar diphthong ending in -*i* is a likely original phonetic value. First of all, *ei* becomes confused with *ai*, so that the two were probably relatively close. Second, the "St. Andrews cross" diacritic used in the Brāhmī appears to derive historically from the *ai*-diacritic (Leumann 1919: 25; see also Dragoni, this volume, p. 120[15]). These two facts are not easily explained if we assume, with Emmerick's earlier value *ẹ* for <ä>, that <ei> somehow stands for or derives from -*aẹ* < *-*akah*. Nor are they any easier to understand if we assume Emmerick's later value *ə*, i.e. -*aə* < *-*akah* (Emmerick 2024: 287).

Thus far, the following clues for the value of *ä* have been found:

- it was often a reduced vowel, resulting from the reduction of word-final *-*i* and *-*ah*;
- it was close to *i*, since one of its sources is **i* and it merges with *i* in Old Khotanese;
- *-*ah* probably yielded an *i*-like vowel in Khotanese because *-*akah* probably yielded an *i*-diphthong, transcribed as <ei>;
- it was a relatively front vowel because it represents the outcome of *i*-affection of **a*.

I think that the best interpretation of the different pieces of evidence together is that <ä> originally stood for a reduced *i*. In other terms, it may have been a more centralised vowel than *i*, approximately [ɪ], like in E. *pit*. Later, this vowel may easily have developed to a shwa, compare the further centralisation of the vowel in *pit* to [ə] in "broad" South African English. This neutral central vowel *ə* accounts best for the reduction of *a* to *ä* in unstressed syllables; for this change, an outcome [ɪ] is not plausible.

8. Khotan Prakrit *ȧ* and *ḥ*

As I will try to show, following Hitch (1984: 188–190), there is yet another script variety that is relevant for the understanding of the Brāhmī diacritic <ä>: the variant of the Kharoṣṭhī script used to write Khotan Prakrit. As is well known, there is next to the well attested Niya Prakrit, the Gāndhārī variety of Niya in the Shànshàn 鄯善 kingdom, a variety of Gāndhārī from Khotan, a "Khotan Prakrit". The existence of this variety is known because of the important document CKD 661 (Boyer *et al.* 1920−29: 249; Baums & Glass 2002−), of which the dating formula shows that it is from Khotan:

[15] Dragoni has a slight preference for Schmidt's derivation of <ei> from the Sanskrit *jihvāmūlīya* [x]. As Dragoni argues, the phonetic rationale of this derivation would be that the velar that is lost in the frequent final -*ei* < *-*akah* was still preserved as a velar element at the earliest stage of writing. In my view, Leumann's derivation is more probable, especially since <ei> is also found word-internally where no velar element was ever there.

saṃvatsare 10 maṣe[16] *3 dhivaza 18 ija kṣunami khotanamaharaya*
'On the 18th day of the 10th month of the 3rd year, at this time in the reign of the king of Khotan ...'

Document CKD 661 is remarkable because of palaeographic peculiarities, Brāhmī in the closing formula, "Runic" characters (apparently too many to be "signa"), and because of linguistic pecularities (Burrow 1936: 430–434; Schoubben 2024: 523–526). The most important linguistic peculiarities are:

- the systematic spelling of *dh* for *d* initially, for example *dhivaza* for *divaṣe* 'day (loc.)', *dhitu* for *dita* 'given', *dhaḍu* for *daṃḍa* 'penalty';
- the systematic spelling of *d* for *t*, for example *haradi* for *harati* 'carries', *grahidu* for *grahita* 'received', *dhinadi* for *denati* 'gives' (in *dhitu* 'given' the intervocalic *t* possibly stands for *tt*, as in Skt. *datta* 'given');
- *u* for *o*, for example *cudiyadi* for *codeyati* 'would complain';
- *i* for *e*, for example *dhinadi* for *denati* 'gives', *vidiyadi* for *veteyati* 'would inform';
- no *ṃ*, for example *madrädi* for *maṃtreti* 'says'.[17]

Of these, the changes of *o* to *u* and of *e* to *i* are strongly reminiscent of the derivations of Khotanese *ī* < **ai* and *ū* < **au*, as in Khot. *śśīta-* 'white' < IIr. **ćuaiṭa-* (Av. *spaēta-*) or *ggūna-* 'colour' < **gauna-* (Av. *gaona-*). These vowel changes of Khotanese most probably went through an intermediate stage with monophthongised long mid-vowels **ē* and **ō*, i.e. Khot. *ī* < **ē* < **ai* and *ū* < **ō* < **au*. The parallelism between Khotan Prakrit and Khotanese may be explained in two ways: Either the changes of **ē* to *ī* and of **ō* to *ū* were not yet completed, and as this change occurred in Khotanese it also affected the local variety of Prakrit; or the sound changes were completed, but speakers of Khotanese still could not pronounce *ē* and *ō* and replaced these in their variety of Prakrit with *ī* and *ū*.

The Khotan Prakrit changes *d* > *dh* and *t* > *d* are not so easily explained. If *dh* should stand for *δ*, both changes would be suggestively parallel to the values that were long assumed for Khotanese <d> and <t>, namely /δ/ and /d/ respectively (Emmerick 1981: 187–188, 203). However, this interpretation of Khotanese <d> and <t> has been challenged and is no longer supported. Rather, <d> is better interpreted as /d/, while <t> stands for /t/, or for a glottal stop intervocalically (Emmerick 2024: 8, 9).

[16] KI read *maṣ.e*; Baums & Glass read *ma[ṣ].e*. However, the curl at the bottom of <ṣa> is probably just ornamental, and the most straightforward reading is *maṣe* (p.c. Niels Schoubben and Francesco Barchi).

[17] It should be noted that loss, or more probably omission in spelling of *ṃ* is not exceptional for Gāndhārī generally (Niels Schoubben, p.c.).

A further parallel between Khotan Prakrit and Khotanese is found in the inflexion of *a*-stem nouns: Unlike Niya Prakrit, Khotan Prakrit seems to have a difference between m. nom.sg. *-aḥ* or *-ȧ*, acc.sg. *-u* or *-o* and neutre *-u*. No rule could as yet be established for the alternation between the two nominative endings *-aḥ* and *-ȧ*, nor for the alternation between the accusative endings *-u* and *-o*, but the distinction between nominative and accusative in these endings is clear. This distinction is exactly parallel to Khotanese, where the same contrast is made in nom.sg. *-ä* and acc.sg. *-u*, for instance in nom.sg. *balysä* 'Buddha' vs. acc.sg. *balysu.*

Although Khotan Prakrit "ä" and "aḥ" look quite different if they are so transcribed, they are in Kharoṣṭhī in fact very similar: *ȧ* represents one dot over an akṣara, and *aḥ* represents two dots over an akṣara (Fig. 4 & 5).

Fig. 4. CKD 661 o6 *tȧna* for *tena* 'therefore'.

Fig. 5. CKD 661 o5 *uṭaḥ* for *uṭa* 'camel'.

(images from Baums & Glass 2002−, and ultimately Boyer *et al.* 1920−29)

While *ȧ* is a unique diacritic, *aḥ* can be interpreted in two ways. As the transcription suggests, *ḥ* could be taken to stand for the consonant visarga, as in the Sanskrit Kharoṣṭhī document CKD 523. In that text, it is clearly not a vowel, but a consonant *ḥ*, because it is combined with the vowel *i* in the word CKD 523 r4 *budhiḥ*. However, it is implausible that *ḥ* would still be in use as a consonant to mark the nom.sg. in Khotan Prakrit, and it would in many places occur where a different sandhi would be regular. Since *aḥ* interchanges with *ȧ*, and *ȧ* also occurs internally, where it clearly stands for a vowel, *aḥ* is probably best interpreted as a vowel as well. In that case, a better transcription would be *ä* (as already done by Konow 1936).

Probably, the symbol that in Kharoṣṭhī appears to be used as visarga *ḥ* for Sanskrit as in CKD 523, and as a vowel *ä* for Khotan Prakrit as in CKD 661, is a borrowing from the Brāhmī script, in which visarga is also noted with two dots (see also Dragoni, this volume, p. 120). Since in Brāhmī visarga is written as a colon with the two dots aligned vertically as a separate akṣara, this shape seems close enough to the Kharoṣṭhī way of writing visarga (or the vowel *ä*) as two dots over the akṣara.

A partial distribution of the two vowel diacritics *ȧ* and *ä* can be established: *ȧ* is used word-finally and internally, but *ä* only word-finally. If the original distribution was that *ȧ* was used internally and *ä* finally, and *ä* derives from word-final visarga *ḥ*, this would

suggest that *ä* was adapted to *ȧ* for internal use especially, and that subsequently *ä* and *ȧ* became partly confused so that *ȧ* could also be used word-finally.

The values of *ȧ* and *ä* must have been close to one another, since the diacritics are in part confused. Since *ȧ* derives from *e* (see above), it is likely to have been relatively close to *e*. It is therefore likely that *ä* was close to *-e* as well; in any case, this ending of Khotan Prakrit can hardly have been close to *-o*, because *-o* and *-u* are found as systematic spellings of the accusative already.[18] However, neiter *ȧ* nor *ä* can obviously have been identical to *e* or *i*, because then they would in all probability have been written with either <e> or <i>. A compromise that does justice to all these observations seems to be the assumption of a reduced vowel close to *e* and *i*, for instance [ɪ]. In parallel, the confusion between *-o* and *-u* at the end of the word may point to a reduced vowel close to *o* and *u*, for instance [ʊ] (as in Gm. *muss* or in E. *foot*).

The Khotan Prakrit vowel *ȧ* derives from Sanskrit *e*, but it is in competition with Khotan Prakrit *i*, which is also a reflex of Sanskrit *e*. It seems that the distribution between *i* and *ȧ* from Skt. *e* is related to the position in the word. Word-finally, Sanskrit *e* develops into Khotan Prakrit *i* regularly:

- *kali* 'at that time' < *kāle*;
- *sa̲ga̲ji* 'before' < *sakāśe*;
- *uvadayi* for Niya Prakrit *uvadaye* 'starting from' < *upādāya*.

Internally, *ȧ* is found when on the basis of the etymology it is expected to be the only long vowel in the word, or, in the case of *mulyȧna*, the rightmost long vowel:

- *tȧna* for Niya Prakrit *tena* 'therefore' < *tena*;
- *madrȧdi* for Niya Prakrit *maṃtreti* 'says' < *mantrayati*;
- *mulyȧna* for Niya Prakrit *muliyena* 'in price' < *mūlyena*;

Instead, *i* is found internally when on the basis of the etymology a long vowel is expected to follow:

- *vidiyadi* for Niya Prakrit opt. *veteyati* 'would inform';
- *uthaviyadi* for Niya Prakrit opt. *uthaveyati* 'would bring up';
- *cudiyadi* for Niya Prakrit opt. *codeyati* 'would complain';
- *dhinadi* for Niya Prakrit *denati* 'gives', as if from *denāti*.

[18] Schoubben (2024: 402) also assumes that *ȧ* and *ä* are "some *e*-sound". He notes, however, that the Khotan Prakrit dialect may in origin rather be an *o*-dialect because of nom.sg.m. *so* in CKI 661 o2, o5. He argues that this contradicts Burrow's idea that there are no original dialect differences between Niya and Khotan Prakrit (Burrow 1936) because Niya Prakrit is in origin an *e*-nominative dialect as shown by e.g. *se* 'he' < *saḥ* and *tade* 'then' < *tataḥ* (Burrow 1937: 4).

The length of the *a* of the optative forms *vidiyadi*, *uthaviyadi* and *cudiyadi* was still preserved in the Niya dialect, since it is occasionally written (Burrow 1937: 46): CKD 320 r3 *gṝheyāti* 'would take', CKD 582 cr1 *coteyāti* 'would complain', CKD 678 uo5 *bhaveyāti* 'would be', and CKD 437 cr7 *deyāṃti* 'would give'.

Since the distribution of *ȧ* and *i* < *e* is dependent on the position in the word and the position relative to other long vowels, I suppose that it is due to a difference in accent. The accent rule for words containing long vowels appears to be that the accent falls on the last long vowel, but not on a final syllable.[19] If this rule is correct, *e* developed to *i* in unaccented position, and to *ȧ* in accented position:

- after the accent: *káli* < *kāle*, *sa̱ǵa̱ji* < *sakāśe*, *uvadáyi* for Niya Prakrit *uvadaye*;
- before the accent: *vidiyádi* for Niya Prakrit *veteyati*, *uthaviyádi* for Niya Prakrit *uthaveyati*, *cudiyádi* for Niya Prakrit *codeyati*, *dhinádi* for Niya Prakrit *denati*;
- under the accent: *tȧ́na* < *tena*, *madrȧ́di* < *mantrayati*, *mulyȧ́na* < *mūlyena*.

9. Another Khotan Prakrit related document?

A further relevant document, which has not yet been successfully interpreted or translated, is BH5-7 (CKD 843). This document shows several features that are similar to Khotan Prakrit document CKD 661 (Fig. 6 & 7). Based on the initial decipherment of Duan, the following features linking this document to Khotan Prakrit may be listed (Duan 2013: 203–205):

- Loss of *ṃ* and voicing of *c*: *pajamu* for *paṃcama* 'fifth'[20]
- *i* for *e*: *ika̱ṣa* for *eka̱ṣa*, gen. of 'one'?
- *dh* for *d*: *dhivaza* for *diva̱ṣa* 'day'
- Many names, most end in *-aḥ*, or for that matter, *-ä*
- The name *kharṣakaḥ* / *kharṣakä* is reminiscent of Khot. *Kharsali* and *Kharsa* in WBH1 (Duan 2008: 12–14).
- *ustamu*, *ustama* = Khot. *ustama-* 'utmost'?; *maśta* = Khot. *mästa-* 'great' (among other options)?
- *vharri* = Khot. *phārra-* 'fortune' (OIr. **farnah-*)?

[19] This concords with the rules of the accentuation of Classical Sanskrit as given by e.g. Bühler, who explicitly says that this accentuation is that of the "indischen Brahmanen" and thus not just a Western Latin-based system (1927, verso of the "Schrifttafel"). Obviously, this accent is not noted in the texts, which is presumably why Renou (1996: 25) holds that "rien ne permet d'indiquer avec précision la nature de l'accent dans la période post-védique".

[20] The loss of anusvāra is also found in Khotan Prakrit, but as noted above, this phenomenon is widespread in Gāndhārī generally and therefore not a good diagnostic. As noted by Niels Schoubben (p.c.), the voicing of *c* to *j* is parallel to the spelling peculiarities of the Khotan Dharmapada.

In addition, it shows a diacritic that looks like Khotanese <ei>, and is accordingly transcribed with <ai> by Duan (Fig. 8; cf. also Dragoni, this volume, p. 120).

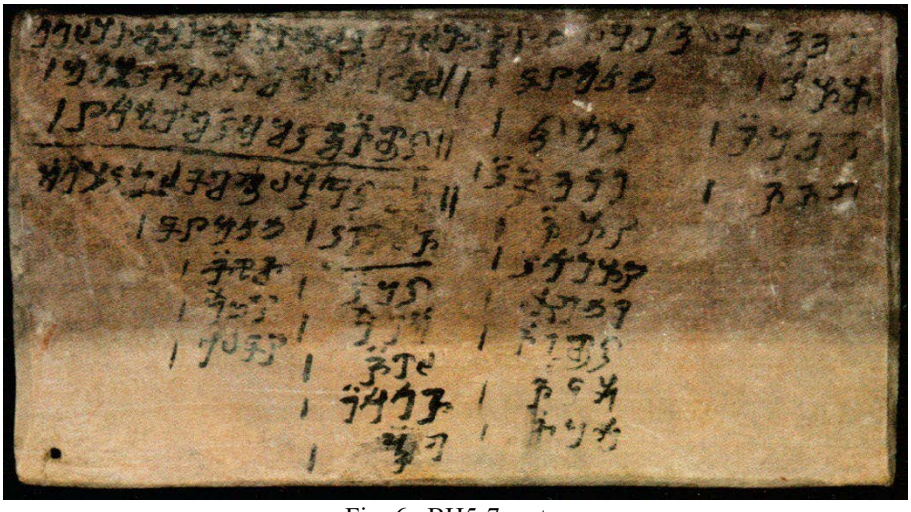


Fig. 6. BH5-7 recto.

Fig. 7. BH5-7 verso.

Fig. 8. BH5-7 r7 *risavai* | *kṣatarai* |.

© National Library of China.

Although BH5-7 / CKD 843 is important for a better understanding of Khotan Prakrit, otherwise attested only in CKD 661, it does not as of yet offer good evidence for the development, value, or spread of the diacritic <ä>. It does nevertheless seem to show that Khotan Brāhmī and Khotan Kharoṣṭhī depend on each other and show several rather specific shared features that set both apart from other varieties of Brāhmī and Kharoṣṭhī.

10. Further Kharoṣṭhī evidence: Kuča Prakrit

Further evidence for the development and spread of the diacritic <ä> may come from the Prakrit variety of Kuča. This variety is barely known, but some of the available documents have been studied by Ching (2013, 2014). In her editions, a diacritic consisting of two dots does occur, but it is found in one instance under instead of above the akṣara (Fig. 9). Another possible instance has been read as *ai* by her (Ching 2013: 62, see Fig. 10). For this occurrence, an alternative reading *nanaṃkäñeme* might be considered in view of the possible Tocharian B equivalent *na[nnak̲]ñ(e)m(eṃ)* in a1 of the Brāhmī side (cf. the suggestion by Georges-Jean Pinault *apud* Ching 2013: 62).

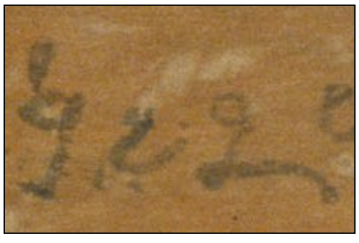


Fig. 9. Kha4.2 (Ching 2014: 62) *E[s̲]ạle*, proper name.

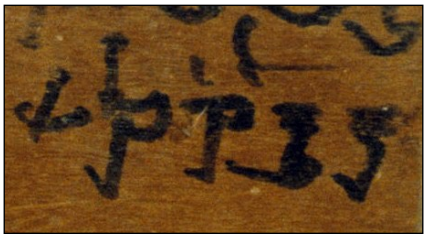


Fig. 10. THT 4059 b2 (= Kha6.2): *nanaṃk[ai]ñeme.*
© Depositum der Berlin-Brandenburgischen Akademie der Wissenschaften in der Staatsbibliothek zu Berlin – Preussischer Kulturbesitz Orientabteilung.

Unfortunately, this Prakrit variety is too scarcely attested to date to draw any firm conclusion. However, as far as the evidence goes so far, <ä> seems only to be used in proper names, not in genuinely Prakrit words, and it may therefore render a sound taken from Tocharian B, and foreign to the Prakrit dialect itself.

11. Further Kharoṣṭhī evidence: Formal Kharoṣṭhī

A further script variety may be relevant: the so-called Formal Kharoṣṭhī script. In this script, there is a vowel diacritic with two dots, but the decipherment is not advanced enough to claim anything useful on the sound value of this diacritic. It may nevertheless be observed that *-ä* occurs mostly if not always word-finally and may there represent the outcome of **-ah* if the language is Iranian as argued by Dragoni *et al.* (2020).

12. Conclusions

To conclude, I summarise my main results. The diacritic <ä> of the Kharoṣṭhī variety that is used for Khotan Prakrit appears to derive from the visarga of Brāhmī. Accordingly, it may be transcribed <ä> or <aḥ>. This diacritic probably denotes a vowel that was close to *i* and *e*, but different, e.g. a centralised *i*, i.e. [ɪ]. Word-finally, the range of use of <ä> or <aḥ> coincides exactly with Khotanese *-ä*. The Khotanese diacritic <ä> clearly denoted a vowel, but it is not entirely clear which. In any case, it was different from *i* and *e*, but seems to have been close to *i* originally, so that it may have been [ɪ], just like Khotan Prakrit <ä>. Later, <ä> seems to have developed to shwa, especially when it denotes a reduced vowel derived from *a* as in *hvatana-* > *hvatäna-* 'of Khotan'.

Since Khotan Prakrit Kharoṣṭhī <ä> or <aḥ> and Khotanese Brāhmī <ä> are used with the same approximate sound value, and partly in the same function, namely word-finally in the nominative singular of masculine nouns, and both scripts are from the same region, it is likely that these two diacritics are related. The most likely scenario is that the diacritic was borrowed from Khotan Prakrit Kharoṣṭhī into Khotanese Brāhmī when Khotanese replaced Prakrit as the literary language of Khotan.

The sound value of the Tocharian vowel diacritic <ä> appears to be different from the original value [ɪ] postulated for Khotanese <ä>, but would be close or identical to the later value [ə] of Khotanese. Since Khotanese <ä> has a plausible origin in Khotan Prakrit Kharoṣṭhī, while there is no obvious source for Tocharian <ä> otherwise, it is attractive to assume that the Tocharian diacritic was borrowed from Khotanese Brāhmī.

A further aspect relevant to the scenario that Tocharian <ä> derives from Khotanese <ä> is the chronology. Obviously, the Khotanese use must be older than that of Tocharian. As far as the evidence goes, the earliest manuscripts of both languages are from the early 5th century CE, so that this does not need to be a problem. The fact that Brāhmī was adapted to write the local languages Tocharian B and Khotanese around the same time points to contact by itself. In addition, we may cautiously connect the Khotanese influence on Tocharian Brāhmī with the borrowings from Old Khotanese into Tocharian B (Dragoni 2023).

At this point, the option that Tocharian <ä> may alternatively derive from a local variant of Kharoṣṭhī directly does not appear to be attractive. Kuča Prakrit would be an obvious candidate. However, although Kuča Prakrit has a diacritic <ä>, it does not have the same distribution as in Khotan Prakrit: it rather seems to be used to write Tocharian names, and thus may represent a foreign sound for which the corresponding diacritic was borrowed from Tocharian B Brāhmī.

Finally, the question is why Tocharian Brāhmī has a mixed system to denote shwa, making use of Fremdzeichen next to the diacritic <ä>. As it is improbable that a regular system with <ä> for all occurrences of shwa would be changed to the attested mixed system, the Fremdzeichen will represent an older system to write shwa, which was (perhaps only slightly) later expanded with <ä> to write shwa consistently in all positions. This relative chronology of the development of Tocharian B Brāhmī is supported by the fact that <ä> is in the oldest Tocharian B texts not consistently written on the akṣaras <c>, <ñ>, <y>, <w> (Peyrot 2008: 35), which are the only akṣaras for genuinely Tocharian consonants that are lacking a Fremdzeichen counterpart.

References

BAUMS, Stefan & Andrew GLASS, 2002–. *Catalog of Gāndhārī texts.* See: https://gandhari.org/catalog.

BOYER, Auguste M., Edward J. RAPSON, Emile SENART and Peter S. NOBLE, 1920−29. *Kharoṣṭhī inscriptions discovered by Sir Aurel Stein in Chinese Turkestan*, Oxford.

BÜHLER, Georg, 1927. *Leitfaden für den Elementarkursus des Sanskrit*, Wien.

BURROW, Thomas, 1936. 'The dialectical position of the Niya Prakrit', *BSOS* 8, 419–435.

BURROW, Thomas, 1937. *The language of the Kharoṣṭhi documents from Chinese Turkestan*, Cambridge.

CHING Chao-jung, 2013. 'Reanalyzing the Kuchean-Prākrit tablets THT4059, THT4062 and SI P/141', *TIES* 14, 55–94.

CHING Chao-jung 庆昭蓉, 2014. 'Kezier chutu De cang Quluwen Qiuciwang zhaoyu yu qiyue wenshu yanjiu 克孜尔出土德藏佉卢文龟兹王诏谕与契约文书研究 – On the decree of Kuchean King and two pieces of contracts written in Kharoṣṭhī among the findings from the Kizil Grottoes and kept in Berlin', *XYWS* 9, 51–73.

COUVREUR, Walter, 1965. 'Koetsjische schrifttabellen in Slanting Gupta', *Orientalia Gandensia* 2, 111–143.

DRAGONI, Federico, 2023. Watañi lāntaṃ: *Khotanese and Tumshuqese Loanwords in Tocharian*, Wiesbaden.

DRAGONI, Federico, this volume. 'Was the Khotanese Brāhmī subscript hook borrowed from the Kharoṣṭhī anusvāra?', 105–129.

DRAGONI, Federico, Niels SCHOUBBEN & Michaël PEYROT, 2020. 'The Formal Kharoṣṭhī script from the Northern Tarim Basin in Northwest China may write an Iranian language', *AOH* 73, 335–373.

DUAN Qing 段晴, 2008. 'Yutianyu gaoseng mainu qiyue 于闐語高僧買奴契約', *Dunhuang Tulufan yanjiu* 敦煌吐魯番研究 – *Journal of the Dunhuang and Turfan Studies* 11, 11–27.

DUAN Qing 段晴, 2013. 'Neirong buming Quluwen mudu: Guotu BH5-7 内容不明佉卢文木牍:国图 BH5-7', in *Zhongguo Guojia Tushuguan cang Xiyu wenshu: Fanwen, Quluwen juan* 中国国家图书馆藏西域文书:梵文、佉卢文卷 – *Xinjiang manuscripts preserved in the National Library of China: Sanskrit fragments and Kharoṣṭhī documents* (ed. DUAN Qing 段晴 & ZHANG Zhiqing 张志清), Shanghai, 203–205.

EMMERICK, Ronald E., 1981. 'The consonant phonemes of Khotanese', in *Monumentum Georg Morgenstierne* I, Leiden, 185–209.

EMMERICK, Ronald E., 1989. 'Khotanese and Tumshuqese', in *Compendium linguarum Iranicarum* (ed. Rüdiger SCHMITT), Wiesbaden, 204−229.

EMMERICK, Ronald E., 1998. 'Khotanese *ei*'. In *Proceedings of the Third European Conference of Iranian Studies held in Cambridge, 11th to 15th September 1995. Part 1: Old and Middle Iranian studies* (ed. Nicholas SIMS-WILLIAMS), Wiesbaden, 93–97.

EMMERICK, Ronald E., 2024. *A handbook of Khotanese*. Completed and edited by Mauro MAGGI, John S. SHELDON and Nicholas SIMS-WILLIAMS, Wiesbaden.

FILLIOZAT, Jean, 1948. *Fragments de textes koutchéens de médecine et de magie, Texte, parallèles sanskrits et tibétains, traduction et glossaire*, Paris.

GABAIN, Annemarie von, 1954. *Türkische Turfan-Texte VIII*, Berlin.

GABAIN, Annemarie von, 1974. *Alttürkische Grammatik.* 3rd ed., Wiesbaden.

HITCH, Doug, 1984. 'Kharoṣṭhī influences on the Saka Brāhmī scripts', in *Middle Iranian Studies. Proceedings of the International Symposium organized by the Katholieke Universiteit Leuven from the 17th to the 20th of May 1982* (ed. Wojciech Skalmowski & Alois Van Tongerloo), Leuven, 187–202.

JASANOFF, Jay H., 1978. *Stative and middle in Indo-European*, Innsbruck.

JASANOFF, Jay H., 1987. 'Some irregular imperatives in Tocharian', in *Studies in Memory of Warren Cowgill* (ed. Calvert WATKINS), Berlin–New York, 92–112.

KIM, Ronald I., 2007. 'Vowel weakening in Tocharian A preterite participles and abstract nouns', *Kyoto University Linguistic Research* 26, 1–30.

KONOW, Sten, 1936. 'Note on Khotanī Saka and the Central Asian Prakrit', *Acta Orientalia* 14, 231–240.

LEUMANN, Ernst, 1919. *Maitreya-samiti, das Zukunftsideal der Buddhisten*, Straßburg.

LEUMANN, Manu, 1934. *Sakische Handschriftproben*, Zürich.

MAGGI, Mauro, 2004. 'The manuscript T III S 16: Its importance for the history of Khotanese literature', in *Turfan Revisited – The first century of research into the arts and cultures of the Silk Road* (ed. Desmond DURKIN-MEISTERERNST *et al.*), Berlin, 184–190.

MAGGI, Mauro, 2022. ‘Some remarks on the history of the Khotanese orthography and the Brāhmī script in Khotan’, *ARIRIAB* 25, 149–172.

MALZAHN, Melanie, 2007a. ‘A Tocharian Brahmi chart’, in *Instrumenta Tocharica* (ed. Melanie MALZAHN), Heidelberg, 223–254.

MALZAHN, Melanie, 2007b. ‘The most archaic manuscripts of Tocharian B and the varieties of the Tocharian B language’, in *Instrumenta Tocharica* (ed. Melanie MALZAHN), Heidelberg, 255–297.

MALZAHN, Melanie, 2012. ‘Now you see it, now you don’t — Bewegliches *-o* in Tocharisch B’, in *Linguistic developments along the Silk Road: Archaism and Innovation in Tocharian* (ed. Olav HACKSTEIN & Ronald I. KIM), Wien, 33–82.

MARGGRAF, Jürgen, 1970. *Untersuchungen zum Akzent in Tocharisch B.* Diss. Kiel.

MAUE, Dieter, 1996. *Alttürkische Handschriften Teil 1: Dokumente in Brāhmī und tibetischer Schrift*, Stuttgart.

MAUE, Dieter, handlist. *Tumshuqese manuscripts and inscriptions.* [unpublished]

MAUE, Dieter, 2015. *Alttürkische Handschriften Teil 19: Dokumente in Brāhmī und tibetischer Schrift Teil 2*, Stuttgart.

MAUE, Dieter & Nicholas SIMS-WILLIAMS, 1991. ‘Eine Sanskrit-Sogdische Bilingue in Brāhmī’, *BSOAS* 54, 486–495.

MAUE, Dieter & Nicholas SIMS-WILLIAMS, 2024. *Sogdian documents in Brāhmī script*, London.

PEYROT, Michaël, 2008. *Variation and change in Tocharian B*, Amsterdam–New York.

PEYROT, Michaël, 2014. ‘La relation entre la chronologie du tokharien B et la paléographie’, in *Lecteurs et copistes dans les traditions manuscrites iraniennes, indiennes et centrasiatiques* (ed. Nalini BALBIR & Maria SZUPPE), Roma, 121–147.

PEYROT, Michaël, 2018. ‘Interrogative stems in Hittite and Tocharian’, *IF* 123, 65–90.

RENOU, Louis, 1996. *Grammaire sanscrite*. 3rd ed., Paris.

SCHOUBBEN, Niels, 2024. *Traces of language contact in Niya Prakrit. Bactrian and other foreign elements.* PhD thesis Leiden.

SIEG, Emil & Wilhelm SIEGLING, 1908. ‘Tocharisch, die Sprache der Indoskythen. Vorläufige Bemerkungen über eine bisher unbekannte indogermanische Literatursprache’, *SPAW*, Jahrgang 1908, 915–932.

SIMS-WILLIAMS, Nicholas, 1996. ‘The Sogdian manuscripts in Brāhmī script as evidence for Sogdian phonology’, in *Turfan, Khotan und Dunhuang. Vorträge der Tagung „Annemarie v. Gabain und die Turfanforschung“* (ed. Ronald E. EMMERICK *et al.*), Berlin, 307–315.

SKJÆRVØ, Prods O., 2002. *Khotanese manuscripts from Chinese Turkestan in the British Library*, London.

WILKENS, Jens & Michaël PEYROT, 2024. ‘Tokharian, Turkic and’, in *Encyclopedia of Turkic languages and linguistics online* (ed. Lars JOHANSON), Leiden–Boston.

Gāndhārī light on Eastern Middle Iranian and vice versa: Three new examples*

Niels SCHOUBBEN

This paper aims to illustrate with three new examples that crossing the disciplinary boundaries between Indology and Iranian studies benefits both fields, especially where Gāndhārī and Eastern Middle Iranian languages are concerned. The following proposals will be made: (i) MSogd. *krmšwhn* 'absolution' is a loan from Gandh. **karma-śohana-* 'purification/removal of (bad) *karma*' < OIA **karma-śodhana-*, and MBactr. *šwwh-* 'avoid, escape' a loan from Gandh. **śoh(e)-* < OIA *śodhaya-* 'purify, remove (something noxious), clear off, exculpate'; (ii) the sequence πιδοβισταρο in Doc kd, a Bactrian birchbark letter, must be read as two words, i.e. the preposition πιδο 'in, on, with, etc.' and βισταρο 'detail', the latter borrowed from Gandh. (less likely Skt.) *vistara-* 'idem'; (iii) the sequence *divinajanami* in the Gāndhārī silk fragment CKM 430 conceals an administrative formula *divi na janami* 'I do not witness the document', cognate with Bactr. λιβο να ζανινδο 'they do not witness the document'. An annotated translation of CKM 430 is appended to the article.

1. Introduction

"The content of Indian and Iranian studies" – Harold W. Bailey remarked in his Cambridge inaugural lecture (1938: 5–6) – "is the study of a culture which has become two but was originally one, and of which the later developments came once again into close contact". In the first millennium CE, the contacts between Indian and Iranian languages mentioned by Bailey took place in the Indo-Iranian borderlands in Pakistan and Afghanistan, but also in the Tarim Basin (NW China), where Gāndhārī functioned as a language of law and religion, Khotanese and Tumshuqese were spoken by locals, and Bactrian and Sogdian were introduced by socio-politically influential migrants. The scholarly study of these languages thus requires an integrated approach in which the disciplinary boundaries between Indian and Iranian studies are regularly crossed. Each text in an Eastern Middle Iranian language may contain hitherto unnoticed loanwords or

* I wish to thank Ching Chao-jung for inviting me to contribute to this volume and for highly useful suggestions, Sasha Lubotsky and Michaël Peyrot for their very helpful feedback on an earlier draft, and Nicholas Sims-Williams for discussion on the ideas presented in Sections 2 and 3. This research was funded by the European Research Council (ERC-2022-COG 101088902).

10.29091/9783752003635/009

calques from an Indian language, while texts in Gāndhārī, including its Tarim Basin varieties (Niya Prakrit, Khotan Prakrit, Kuča Prakrit), can be expected to contain Iranian loanwords and idioms. Sometimes, parallel passages in Eastern Iranian and Indian texts have to be identified and compared before either can be elucidated in full.

A recent example from the literature nicely illustrates the utility of a comparative approach to Eastern Middle Iranian and Gāndhārī. In the royal titulature of two Kuṣāṇa inscriptions (both 2nd century CE), the reigning monarch is introduced as βαγοιηζνογο (DN1, line 5) or βαγοηζνογο (Rab., lines 1–2). Assuming a word boundary behind βαγο, Sims-Williams (2008: 58) interpreted this Bactrian sequence as 'god worthy of worship' < **baga-* + **yazn-iya-ka-*. Recently, though, Halfmann *et al.* (2024: 18–20) observed that βαγοιηζνογο / βαγοηζνογο corresponds to *devavrata-* 'devoted to the gods' in Gāndhārī versions of the Kuṣāṇa titulature. Based on this observation, they convincingly suggest that the Bactrian title is a compound, too, meaning 'worshipping the gods' < **baga-yazna-ka-*.[1]

The aim of this paper is to present three more examples where a closer look at Gāndhārī (and Indian languages in general) casts new light on Eastern Middle Iranian or vice versa. Sections 2 and 3 concern unrecognised Gāndhārī loanwords into Bactrian and/or Sogdian. In Section 4, the Bactrian legal formula λιβο να ζανινδο 'they do not witness (literally: 'know') the document' (Doc A line 4 in BD I^2: 26–27) is adduced to elucidate the so far obscure sequence *divinajanami* in the Gāndhārī land-sale contract CKM 430. A revised translation of this Gāndhārī document, a silk fragment recently published by Falk (2021), is included as an appendix. The notes added to this translation are meant as a further illustration of the benefits obtainable from integrating Indian with Iranian studies, as they contain additional cases where comparisons with Bactrian idioms and formulae throw light on the interpretation of the Gāndhārī text.

2. MBactr. *śwwh-* 'avoid, escape' and MSogd. *krmšwhn* 'absolution'

The sole text in Manichaean Bactrian surviving is the leather fragment M 1224 (Boyce 1960: 69, 150). In view of the expedition code T I α written on the recto, this fragment was probably discovered in the SW corner of the Qočo ruins during the first Turfan expedition (Nov. 1902–March 1903; cf. Boyce 1960: X–XIII). It is presently housed in

[1] As duly admitted by Halfmann *et al.* (2024: 20), there is still a minor phonological issue with the etymology in that there is so far no exact parallel for the required development of **-ya-* to Bactr. -(ι)η-/(y)ē/.

Berlin,[2] and it has been dated on linguistic grounds to the late 7th, early 8th century CE (Sims-Williams 2011: 169).

On its recto, M 1224 contains an otherwise unattested 3sg. verb form *śwwhyyd* governing a direct object *wy(ś)p nrḫ* 'all hells'. The context in which this verb occurs suggests that it approximately means 'avoid' or 'escape' (cf. also DMSB[2]: 285 s.v. *śwh-*); see (1).

(1) M 1224 recto 11–16

t'd 'ww q[d](y)βryyg 'qydyy(ḫ) lhyyd [ṭ]'(d pyd) wyśp pwwn b'gy(g)ynd 'wd ywg py(d) hz'r pwnyng βrg 'βyryyd 'wd wy(ś)p nrḫ śwwhyyd 'wd pwwn pydrw(βy)d y'wyd'nźyg

'[T]hus, that layman who gives it—he becomes a sharer in all the merits, and obtains merit-fruit a thousandfold and escapes(?) all hells [*wy(ś)p nrḫ śwwhyyd*] and receives merits eternally.' (Sims-Williams 2009: 249, 252)

In unpublished notes cited by Sims-Williams (2009: 258), Henning compared MBactr. *śwwh-* to the second element of MSogd. *krmšwhn* 'absolution' (also spelled *qrmšwhn*). One of the occurrences of this Sogdian noun is in the so-called "additional text f" appended to modern editions of the *Bet- und Beichtbuch*;[3] see (2).

(2) *Bet- und Beichtbuch* f5–12 (= M 139 i. recto 5–12)

'rty nwkr (xw) γw'nw'cyy pr δyn I p't (x)[c](y)y ˚˚ wy(δp')ṭ c('nw) 'ty cn spyn'w' zwrtyy ˚ 'ty cn δs(') z(n)g'n 'kt'ny(y) 'ty cn wtšnyḫ γnd'k 'kty' 'pstwyy 'rt(y)w' qrmšwhn 'ty 'xšn'm w(y)δ p'ty βyrṭ

'Now, in the religion forgiveness of sins occurs (just) once; at that time, if one turns from rebelliousness and renounces the ten kinds of misdeed and (one's) former wicked actions, one then obtains absolution [*qrmšwhn*] and pardon.' (Sims-Williams *et al.* 2022: 189)

Since the first element of *krmšwhn* 'absolution' is obviously Gandh./Skt. *karma-* '(good/bad) deed, consequences of actions',[4] *-šwhn* must roughly mean 'removal' or 'remission'. Like the first element *krm-*, *-šwhn* must be a loan because, isolated dissimilations left aside, *h* is only found in loanwords in Sogdian (Sims-Williams 1989: 179).

According to Henning (*apud* Sims-Williams 2009: 258), MSogd. *-šwhn* and MBactr. *śwwh-* continue the Iranian verbal root **xšaud-* 'wash' (cf. MP *šōy-*, Parth. *šōδ-*; EDIV: 455–456).[5] This would be formally possible under the assumption that both are

[2] Depositum der Berlin-Brandenburgischen Akademie der Wissenschaften in der Staatsbibliothek zu Berlin.

[3] For further information on the "additional text f", see Sims-Williams *et al.* 2022: xvii.

[4] *krm* (*qrm*) 'deed, consequences of actions' is also found as a loanword in Manichaean Parthian, Sogdian, and Bactrian (DMMP: 208; DMSB[2]: 116, 284).

[5] MSogd. *-šwhn* would be from a verbal noun **-xšaudana-* 'washing (away)'.

borrowings from Western Middle Iranian (cf. DMSB[2]: 116 s.v. *krmšwhn* 'Indian-WestMIr. hybrid'). A small number of loanwords into Armenian and Middle Persian, e.g. Arm. *spah* 'army' and MP *(i)spāh* 'idem' ← **spāh* < **spāda-*, require postulating an otherwise unattested Western Middle Iranian language in which intervocalic **d* developed to **h* and not to *y* or *δ* as in Middle Persian and Parthian, respectively (cf. e.g. Meyer 2023: 24). One could therefore conjecture that *-šwhn* and *śwwh-* find their origin in this hypothetical Western Middle Iranian language (though see below).

Sims-Williams (2009: 258) objected to Henning's comparison between MBactr. *wy(ś)p nrẖ śwwhyyd* 'he escapes (?) all hells' and MSogd. *krmšwhn* 'washing away of (bad) *karma* > absolution' that "the metaphor "he washes away all hells" seems rather strained". More recently, he noted that the metaphor would be understandable "if one assumes that the usage has its basis in the noun *krmšwhn* "the washing away of (bad) *karma*"" (Sims-Williams, p.c., 9 October 2023). Indeed, MBactr. *śwwh-* and MSogd. *-šwhn* in *krmšwhn* match formally well, and both words are used in contexts that deal with the good and/or bad consequences of one's karmic actions. Therefore, Henning's suggestion that MBactr. *śwwh-* and MSogd. *-šwhn* are related seems sound.

In my view, it is Henning's etymological proposal which is doubtful. For how plausible is it that an unattested Western Middle Iranian language only known by way of a few loanwords into Armenian and Middle Persian contributed a technical term to the religious vocabulary of Manichaean Bactrian and Sogdian – a term which moreover does not occur in Manichaean Parthian or Middle Persian? Since *śwwh-* governs the Indian loanword *nrẖ* 'hell'[6] and the first element of *krmšwhn* is borrowed from Indian too (cf. *supra*), the first option to explore is whether *śwwh-* and *-šwhn* could also be explained as Indian loans. After all, we know that many Indian terms (especially Buddhist ones) were incorporated into the religious terminology of Manichaean communities speaking Iranian languages (cf. e.g. Sims-Williams 1983).[7]

Because of the presence of a palatal sibilant, two donor languages come into question if *śwwh-* and *-šwhn* are Indian loans, i.e. Sanskrit and Gāndhārī.[8] Of these, Sanskrit can probably be ruled out, as there is no root or stem of the type ***śuh-*, ***śoh-*, or ***śauh-* in this language. With Gāndhārī, there are more possibilities because *-h-* can, first of all, continue OIA velar and labial aspirated stops whose occlusive element was lost

[6] On *nrẖ* 'hell', which is also attested in Parthian, see Sims-Williams 1983: 134.

[7] In Buddhist Sogdian, Indian loanwords are even more common (cf. now Lurje 2021). A few such loanwords also occur in Christian Sogdian (cf. now Sims-Williams 2021).

[8] Except for Gāndhārī and far away Māgadhī, Middle Indian languages do not have a palatal sibilant in their phoneme inventory due to the merger of the OIA sibilants into one dental sibilant (von Hinüber 2001: §219). According to Sims-Williams (2011), MBactr. <ś> could write either [sʲ] or [ʃ]. I favour the latter, but either option would be fine as a rendering of Skt./Gandh. *ś-*.

intervocalically, e.g. *suha-* 'happiness, bliss' < OIA *sukha-*.[9] In many forms of Gāndhārī, the dental aspirates *-th-* and *-dh-* did not partake in this deocclusion, as they developed to /z/ (probably via **δ*), e.g. dir.sg. *purvapraniṣi* /pu̯rvapranizi/ 'former resolution' < *pūrva-praṇidhi-* (CKM 18, verso 41). Yet, there are clear examples of *-th-* and/or *-dh-* becoming *-h-* in Niya Prakrit, the Gāndhārī of the *Khotan Dharmapada*, and a Gāndhārī dialect contributing loanwords to Khotanese (LKD: §27; Brough 1962: §42; Loukota 2023: 10). Thus, if *śwwh-* and *-šwhn* are indeed loans from Gāndhārī, the *-h-* could also go back to *-th-* or *-dh-*.

With this in mind, I propose (i) that MSogd. *krmšwhn* 'absolution' is a loan from a Gāndhārī compound **karma-śohana-* < OIA **karma-śodhana-* 'purification/removal of (bad) *karma*'; and (ii) that MBactr. *śwwh-* 'avoid, escape' was borrowed from Gandh. **śoh(e)-* < OIA *śodhaya-* 'purify, remove (something noxious), clear off, exculpate', the causative of the root *śudh-* 'become pure' (PW vol. 7: 255–257; MW: 1082).[10]

The compound **karma-śodhana-* does not seem to be attested in Indian texts, but there is the synonymous Skt. *karma-viśuddhi-*, Pāli *kamma-visuddhi-*, and Ardh. *kamma-visohi-* 'purity/purification of action(s)' (PTSD s.v. *kamma*; Mylius 2003: 204). Pāli *kamma-visuddhi-* occurs in a pair of verses in the *Dhammapada* (see (3)), which, like the Manichaean text passages quoted in (1) and (2), talk about the results of *karma* and the need for *puṇya* 'merit' (MBactr. *pwwn* ~ Pāli *puñña-*):

(3) *Dhammapada,* vv. 15–16

idha socati pecca socati pāpakārī ubhayattha socati, so socati so vihaññati disvā kamma kiliṭṭham attano | idha modati pecca modati katapuñño ubhayattha modati, so modati so pamodati disvā kammavisuddhim attano (von Hinüber & Norman 1994: 4)

'Here he grieves, having passed away he grieves; the evil-doer grieves in both places. He grieves, he is tormented, seeing his own defiled action. Here he rejoices, having passed away he rejoices; the one who has done merit rejoices in both places. He rejoices, he exults seeing the purity of his own action [*kammavisuddhim*].' (tr. Norman 1997: 2, accompanying notes on pp. 64–65)[11]

---

[9] Gandh. *suha-* is, for instance, attested in Bajaur fragments 4 and 11 (tentatively datable to 50–150 CE); see Schlosser (2022: 316–317).

[10] The Gāndhārī outcome of *śodhaya-* is attested in the Niya documents in the meaning 'pay (off)' (e.g. fut.3pl. *śodheṣyaṃdi* 'they will pay' in CKD 272 o9), but only in a Sanskritic spelling with *-dh-* preserved (LKD: 126 s.v. *śudha*). Therefore, I have put an asterisk in front of Gandh. **śoh(e)-*. As Federico Dragoni (p.c.) has kindly reminded me, Chen & Loukota (2018: 164 fn. 86) also reconstruct a Gāndhārī ("Middle Indic") form **śohe-* < *śodhaya-* 'purify' + *śobhaya-* 'beautify', which they cautiously propose to be the source of Khot. *śūh-* (*śśūh-*), pret. *śūsta*, 'recite, adorn, purify, prepare'.

[11] These verses are parallelled in *Udānavarga* 28.34–35 (Bernhard 1965: 363–364, with *karma viśuddham* instead of a compound ***karmaviśuddhim*), *Khotan Dharmapada* 205–206 (Brough 1962: 151, with *kama višudhu*, as also in v. 204d), and *Saindhavī Dharmapada* = *Patna Dharmapada* 3–4

In other words, a defiled action (*kamma kiliṭṭham*) committed by an evil-doer (*pāpakārī*) causes him to grief (*socati*) in and after this life (*idha*, *pecca*), while the pure actions (*kammavisuddhim*) of a meritorious person (*katapuññho*) cause him to rejoice (*modati*).

In addition, it is noteworthy that Skt. *karma-* and its Middle Indian cognates can form a phraseological pair with the root *śudh-* 'become pure' and its derivatives.[12] A particularly telling example of this phraseology occurs in the Pāli *Vatthūpamasutta*, the 'discourse on the simile of the cloth'. Near the end of this discourse, the Buddha recites a handful of verses to the brahmin Sundarika Bharadvāja, partially quoted in (4). These verses teach the brahmin (and the reader) that contrary to a popular belief, bathing in rivers will not purify, i.e. absolve, evil-doers from the bad deeds they committed. Importantly, they contain three instances of the stem *kamma-* < *karma-* (*kaṇhakammo* 'evil deed', *pāpakamminaṃ* 'intent on bad deeds', *sucikammassa* 'whose actions are pure') and five of the root *sudh-* < *śudh-* (*sujjhati* 'becomes pure', *sodhaye* 'shall purify', *suddhassa* 'pure' (3×)):

(4) *Majjhima-nikāya* 7: *Vatthūpamasutta*[13]

Bāhukaṃ Adhikakkañ ca, Gayaṃ Sundarikām api, Sarassatiṃ Payāgañ ca, atho Bāhumatiṃ nadiṃ | niccam pi bālo pakkhanno kaṇhakammo na sujjhati. kiṃ Sundarikā karissati? kim Payāgo? kim Bāhukā nadī? | veriṃ katakibbisaṃ naraṃ na hi naṃ sodhaye pāpakamminaṃ; | suddhassa ve sadā phaggu, suddhass' uposatho sadā, suddhassa sucikammassa sadā sampajjate vataṃ. (Trenckner 1888: 39, lines 13–20)[14]

'Bāhukā and Adhikakkā, Gayā as well as Sundarikā, Sarassatī and Payāga, and the river Bāhumatī—even when a fool constantly jumps (into these rivers), (his) evil deed does not become pure [*sujjhati*]. What shall the Sundarikā do? What the Payāga? What the river Bāhukā? For certainly (a river) shall not purify an inimical (and) sinful man intent on bad deeds; for a pure (man) (there is) always the spring festival, for a pure (man) (there is) always a holy day, for a pure (man) whose actions are pure there is always proper conduct.' (tr. mine)

The content of (4) also resembles that of the Sogdian passage quoted in (2): there is only one way to obtain absolution, namely by deliberately abstaining from impure deeds. The

(Cone 1989: 104–105, with the compound *kammaviśuddhim*). For Chinese parallels to the *Vatthūpamasutta*, see https://suttacentral.net/mn7?view=normal&lang=en (accessed 20 February 2025).

[12] Similar phraseology is also found in Khotanese, where one can use *kīra-* 'act, deed' (~ Skt. *karma-*) + *vasūj-* 'purify' (~ Skt. *śodhaya-*) to refer to the cleansing of one's karmic actions (cf. DKS: 60 s.v. *kīra-*, 380 s.v. *vasūj-*).

[13] I have partially adapted Trenckner's punctuation of the Pāli text.

[14] The final verse (*suddhassa … vataṃ*) has a parallel in verse 327 of the *Khotan Dharmapada* (Brough 1962: 170, 273) and was thus also known in Buddhist circles in Central Asia.

phraseological parallelism between *kaṇhakammo* 'evil deed' and *pāpakamminaṃ* 'intent on bad deeds' in the Pāli text and *'kt'ny(y)* 'misdeed' and *γnd'k 'kty'* 'wicked action' in the Sogdian text lends still further support to my proposal that MSogd. *krmšwhn* 'absolution' was borrowed from a Gāndhārī outcome of **karma-śodhana-* 'purification/removal of (bad) *karma*'.

In (late forms of) Gāndhārī, intervocalic *-h-* can also go back to OIA *-bh-* (via *-vh-*), e.g. *śuha-* 'pure' < OIA *śubha-*.[15] Therefore, an alternative option could be that *krmšwhn* 'absolution' was borrowed from Gandh. **karma-śohana-* < OIA **karma-śobhana-* to the root *śubh-* 'beautify'. From 'beautification of (bad) *karma*', a derived meaning *'forgiveness for (bad) *karma*' may have developed, comparable to (archaic) Dutch *om verschoning vragen* 'ask for forgiveness', which is based on *schoon* 'beautiful, clean'. For three reasons, I nonetheless deem this option less likely than **karma-śodhana-*: (i) the semantic development that has to be assumed; (ii) compounds ending in *-śobhana-* typically mean 'beautiful by reason of X' rather than 'beautification of X' (PW vol. 7: 313–314; MW: 1092; Norman 1971: 131 ad v. 331); and (iii) I could not find much relevant evidence for a phraseological pairing of *karma-* and its cognates with the root *śubh-* 'beautify'.[16]

Let us now return to MBactr. *śwwh-* as attested in the phrase *wy(ś)p nrḫ śwwhyyd* and my proposal that this Bactrian verb was borrowed from Gandh. **śoh(e)-* < OIA caus. *śodhaya-*. The intended meaning of *wy(ś)p nrḫ śwwhyyd* is obviously that a meritorious person is saved or redeemed from hell[17] – hence Sims-Williams' glossing *śwwh-* as 'avoid' or 'escape' (cf. *supra*). Some of the specialised meanings of *śodhaya-* 'purify' are 'remove (impurity of anything noxious)', 'clear off, pay (debts)', and 'acquit, exculpate' (PW vol. 7: 255–257; MW: 1082). These clearly belong to the same semantic sphere as MBactr. *śwwh-* 'avoid, escape'. For instance, it would not be a big step to derive a meaning 'avoid' from 'remove (anything noxious)', comparable to how Tocharian A *wäyk-* and B *wəyk-* can mean both 'remove' and 'avoid' (Adams 2013: 652–653).[18] Alternatively, one could imagine the phrase 'purify/remove one's (bad) *karma*'

[15] Gandh. *śuha-* is, for instance, attested in Bajaur fragments 4 and 11 (Schlosser 2022: 313).

[16] The most relevant passage I found is a fragmentary Gāndhārī verse in which *karma* 'deed' and the adjective *śohaṇa* 'beautiful, excellent' < *śobhana-* are paired in a context referring to karmic deeds done in previous lifetimes: *suvutam=eḍa agileṇa taḍiṇo | eva ta[ḍ]. ///* + + + + + + + + | + + + + + + + + *(*taṣa)gaḍe | karma kiḍe śohaṇa parvagaḍiṣu* 'This was well said by the Savior, [who is] free of mental harshness. Thus, the Savior … (*the Tathā)gata. An excellent [*śohaṇa*] deed [*karma*] was done in previous lifetimes' (*Anavatapta-gāthā* on Senior scroll 14, verse 18 = CKM 246 recto 23–24; perhaps from Haḍḍa and datable to ca. 140 CE; Salomon 2008: 399).

[17] Cf. e.g. MParth. *'c hw nrḫ 'w hwyn … bw(x)[tn]* 'to save them from that hell' in *Huyadagmān* V 12b (Boyce 1954: 88–89).

[18] I owe the comparison with the Tocharian verb to Louise Friis (Leiden).

to have formed the basis of a derived metaphor 'purify/remove hell', i.e. 'avoid (rebirth in) hell', given that purifying bad *karma* is a prerequisite for not ending up in hell.[19]

Phraseological support for my etymology of *šwwh-* comes from a passage in the Sanskrit *Sarvadurgatipariśodhana*, the 'Purification of all evil destinies',[20] in which *naraka-* 'hell' (cf. MBactr. *nrḫ*) and *śodhitā* 'purified' (cf. MBactr. *šwwhyyd*) co-occur:

(5) *Sarvadurgatipariśodhana Tantra* (version B) 5b–6a

... idaṃ sarvadurgatipariśodhanarājanāmatathāgatahṛdayaṃ niścārayām āsa

oṃ śodhane śodhane sarvapāpaviśodhani śuddhe viśuddhe sarvakarmāvaraṇaviśuddhe svāhā

asyā vidyāyā bhāṣaṇānantaram eva sarvasattvānāṃ durgatir vinipātitā sarvanarakatiryakpretagatiḥ śodhitā tīvraduḥkhāni praśāntāni bahavaś ca jātāḥ sukhīmukhībhūtāḥ (Skorupski 1983: 126, l. 10–16)

'He (the Buddha) uttered the secret spell of the Tathāgata called *Sarvadurgatipariśodhanarāja* [King-Purifier of All Evil Destinies]:

oṃ purifier, purifier, purifier of all sins, pure, most pure, most pure (with regard) of obstruction due to past actions *svāhā*

Immediately after uttering this formula, the evil destiny of all living beings was destroyed; every rebirth (in) hell, (as) an animal, or a hungry ghost was eliminated [literally: purified]; severe sufferings were removed; and many living beings became happy.' (tr. after Skorupski 1983: 6–7)

Hinting once again at the necessity of purifying one's karmic actions (cf. *sarvapāpaviśodhani*, *sarvakarmāvaraṇaviśuddhe*), the tantric formula uttered by the Buddha in (5) makes that living beings are no longer reborn in hell (*naraka*). The word used to express the elimination of rebirth in hell is *śodhitā* 'purified', the verbal adjective (nom.sg.fem.) of *śodhaya-* 'purify'. This supports my idea that MBactr. *šwwh-* in *wy(š)p nrḫ šwwhyyd* 'he avoids/escapes all hells' is a loan from Gandh. **śoh(e)-* < *śodhaya-*.

To summarise, the element *-šwhn* in MSogd. *krmšwhn* 'absolution' and MBactr. *šwwh-* 'escape, avoid' do not seem to be loanwords from an unattested Western Middle Iranian language in which PIr. **xšaud-* 'wash' became **šōh-*. More plausibly, both are Indian loanwords: *krmšwhn* can be traced back to Gandh. **karma-śohana-* 'purification/

[19] In the Buddhist tradition, any remaining bad *karma* would have to be purified in hell. Cf. the formulaic sequence *tan narakavedanīyaṃ karma kṣapayitvā devamanuṣyeṣu pratisandhiṃ gṛhṇanti* 'having thrown of *karma* (still) to be experienced in hell, they are reborn among gods and men' in *Avadānaśataka* 1 *et passim*.

[20] This Buddhist Sanskrit text is an anonymous tantric work of which one version (the so-called "Version A") was translated into Old Tibetan in the 8th century (cf. Lindsay 2024: 29–32). I learn from Allon (2008: 171) that a further witness to the Sanskrit text has appeared in a birchbark manuscript written in Gilgit/Bamiyan Type I Brāhmī (6th/7th century CE). Gregory Schopen is preparing an edition of this manuscript, which is housed in the J. Paul Getty Museum in Los Angeles.

removal of (bad) *karma*' < OIA **karma-śodhana-*, while MBactr. *śwwh-* can be a borrowing from Gandh. **śoh(e)-* < OIA *śodhaya-* 'purify, remove (something noxious), clear off, exculpate'.

3. Bactr. πιδο βισταρο 'in detail'

My second example concerns the sequence πιδοβισταρο in Doc kd, a Bactrian birchbark letter recently published by Sims-Williams (2023a: 138).[21] Like others in the same collection, this letter was probably written in the late 4th century CE, in or close to SE Afghanistan.[22] It was sent by a certain Razz-shamsh (ραζζοþαμþο).[23]

The relevant passage of letter kd runs as follows:

(6) Doc kd, lines 9–11
οδο σαγωνδο πιδοβισταρο πιδοβανδο ταδο βαγο þιγανο ασο τιρογολο πηδανο βοοαδο
'And in as much as (there is any) more extensive [πιδοβισταρο] answer, sir, (your) excellency will be informed by Tir-gul.' (Sims-Williams 2023a: 138)

In his 2023 edition, Sims-Williams takes πιδοβισταρο to be one word, translating it as 'more extensive'. He further explains (2023a: 140) that he analyses πιδοβισταρο as a comparative in -ταρο < **-tara-* from an otherwise unattested adjective *πιδοβιστο 'connected, extended' < **pati-basta-* 'bound on/to'. Under this analysis, πιδοβισταρο would form a *figura etymologica* with the immediately following πιδοβανδο 'answer' < **pati-banda-*. Yet, I find Sims-Williams' idea not entirely persuasive because (i) the required semantic development 'bound on/to' > 'connected, extended' seems somewhat forced, and (ii) **basta-* 'bound' is known to yield Bactr. βαστο (BD II: 202a), not **βιστο (unlike λιστο 'hand' < **dasta-*).

In my opinion, the sequence πιδοβισταρο conceals a prepositional phrase πιδο βισταρο 'in detail, detailed' consisting of (i) the preposition πιδο 'in, on, with, etc.' < **pati* (BD II: 254), and (ii) βισταρο 'detail' as a loanword from Gandh. (less likely Skt.)

[21] Meanwhile, Doc kd has been re-edited by Sims-Williams in BD IV (19f.), where he adopts my interpretation of the sequence πιδοβισταρο as argued for below (cf. also BD IV: 3, 88, 109).

[22] See Sims-Williams 2023a: 135–136, 141–142 (note by Frantz Grenet).

[23] Sims-Williams (2023a: 140, also BD IV: 123) notes that "[t]he name ραζζοþαμþο perhaps represents Skt. *Rāja-śaṃsa-*, with assimilation of the two sibilants". If so, the source language will probably be Gāndhārī, where *s* regularly becomes *ś* in the vicinity of another palatal sibilant, also in the root *śaṃs-* 'to praise' (cf. Baums 2009: 187–188). The required substitution of *ś* with Bactr. þ remains nonetheless remarkable, given that the default way to render an Indian *ś* is with Bactr. σ, while þ normally renders the retroflex *ṣ* (cf. Michetti 2024: 186–187 fn. 84).

vistara- ‘detail, expansion’.[24] Analysed thus, Bactr. πιδο βισταρο nicely corresponds to the adverbial ins.sg. *vistarena* ‘in detail, detailed’, which is used in the Niya documents (3rd–4th centuries CE) in similar contexts as πιδο βισταρο.[25] Note in particular the parallelism between πηδανο βοοαδο ‘he will be informed’ in the Bactrian sentence (6) and *ñadarthama*/*ñadartha bhavidavya* ‘we have/you should become informed’ in the Niya Prakrit passages (7) and (8):[26]

(7) CKD 357 o3
avi lekhaharagaṣa hastaṃmi viñatilekha prahideṣi taha vistarena ñadarthama
‘Also you have sent a letter of information in the hand of a letter-carrier, so we have become informed in detail [*vistarena*].’ (TKD: 69)

(8) CKD 376 o6–7
Opgeya pratilekha yajita tahi karana lekha gaṃdavya tena karana ahu varita pratilekha na hastaṃmi kiḍa eda me lekhade vistarena ñadartha bhavidavya
‘Opgeya demanded a letter in reply. On your account a letter is to go. For that reason I prevented him. A letter of reply was not taken in hand. From this letter of mine you can become informed in detail [*vistarena*].’ (TKD: 76)

The instrumental case ending in *vistarena* can be seen as a semantic correspondent of the preposition πιδο in πιδο βισταρο, as it occurs more often that the same concept is expressed in Gāndhārī with a noun in the instrumental case, and in Bactrian with a prepositional phrase introduced by πιδο. I list some examples of this kind below (again taking the Gāndhārī examples from the Niya documents):

- πιδο (...) οαυαγο ‘for a price’ (e.g. Doc ed line 7) = *muliyena* ‘idem’ (e.g. CKD 327 o4)
- πιδο οαχο ‘with interest’ (Doc Q line 14, 17) = *ayogena* ‘idem’ (e.g. CKD 14 cr3)

[24] For the use of Bactr. β- to render Indo-Aryan *v-*, cf. e.g. βιυαρο (βαυαρο) ‘monastery’ ← *vihāra-* ‘idem’ (BD II: 203b). The Bactrian birchbark letters contain more Indian loanwords; see Sims-Williams (2023a: 136; BD IV: 3).

[25] It is geographically more likely that βισταρο was borrowed from South Asian Gāndhārī (or Sanskrit) rather than from Niya Prakrit, but I focus on Niya Prakrit evidence because we have very few administrative documents in South Asian Gāndhārī whose phraseology can be compared with the Bactrian documents. *vistara-* is also attested in South Asian Gāndhārī, e.g. in the 1st-century *avadāna*s from the British Library collection (Lenz 2010: 176). The ins.sg. *vistarena* was also borrowed into Khotanese as *västarna* (*västārna*); see Skjærvø (2004 II: 342).

[26] Compare also the Niya Prakrit phrase *livivistarena anatilekha* ‘a letter of command [*anatilekha*] with a detailed [*vistarena*] statement in writing [*livi*]’, which is attested in several orthographic variants: *livivistarena anatilekha* (CKD 4 uo1–2), *levistarena anatilekha* (CKD 160 o7; 788 uo5–6), *livistarena anadilekha* (CKD 310 uo3, 2×).

- πιδο ... φρομανο ‘at the command of ...’ (Doc ab line 20) = *anatiyena* (*anadiyena*) ‘idem’ (e.g. CKD 436 cr5)
- πιδο χαγγαρο ‘with a sword’ (Doc O line 10; Doc O’ line 7, 11) = *taravarena* ‘idem’ (CKD 585 co5; cf. also *śastrena* ‘with a knife’ in the same line)
- πιδο <α>στωρο ‘with the animals’ (Doc xb line 6) ~ *storena* ‘with/on a beast’ (e.g. CKD 135 uo5)[27]
- πιδο λοοι μολρο ‘with two seals’ (Doc ci line 12) ~ *muṃtrena* ‘with a seal’ (CKD 247 o4)

In conclusion, the translation of passage (6) can be modified to ‘And in as much as (there is any) detailed [πιδο βισταρο] answer, sir, (your) excellency will be informed by Tirgul’.

4. Gandh. *divi na janami* ‘I do not witness the document’

The subject of my final case study is the sequence *divinajanami* in the Gāndhārī land-sale contract CKM 430, a silk fragment recently published by Falk (2021: 10–15).[28] This document is of unknown provenance, but since the Palola kings are mentioned in line 2, we should probably locate it somewhere in the Gilgit area. According to Falk (2021: 11), it is to be dated to the 4th century CE.

Leaving *divinajanami* unanalysed for now, the last line of CKM 430 can be read and translated as follows:[29]

(9) CKM 430 line 4

likhita mea Ṣaghamitreṇa Pigeṣa ajeṣaṇaï ya karaṇu divinajanami atra sakṣia upari likhiṣati

‘This contract [*karaṇu*] has been written by me, Ṣaghamitra, at the request of Pige. *divinajanami*. The witnesses in this matter will write (their names) above.’ (tr. mine)

Observing the same sequence (with retroflex nasals) to occur in a parallel context in the so far unpublished Bajaur fragment 15,[30] Falk (2021: 14–15) proposed to segment

[27] In the Niya documents, *stora* usually refers to a beast of burden or to a mount (mainly horses and camels); cf. Schoubben 2022: 345–346.

[28] On the 15th of April 2021, Stefan Baums corrected some of Falk’s readings in the online catalogue of Gāndhārī texts (Baums & Glass 2002– *ad locum*), e.g. changing Falk’s reading *diviṇa-* (with a retroflex nasal) into *divina-* (with a dental nasal). In what follows, I cite from Baums’ online edition (see also the appendix, p. 311).

[29] For a full translation of CKM 430 with additional comments, see the appendix, p. 311.

[30] Gudrun Melzer (Munich) is preparing an edition of Bajaur fragment 15 (a loan contract).

divinajanami as *divina janami*,[31] which he glossed as *dīvinaḥ jānāmi* in his Sanskrit *chāyā* and translated as 'the players I know (personally)'. Yet, this proposal seems unlikely for a number of reasons: (i) **dīvin*- 'player' is not a known word in Sanskrit, forcing Falk to hypothesise that it is 'a local variant of the regular *devin*' (*ibid*., p. 15); (ii) it seems unparallelled that a scribe has to express "that the main agents, called the "players", are known to him personally" (*ibid*., pp. 14–15); and (iii) the verbal root *dīv*- typically means 'gamble', 'play with the dice', not **'play a part in a legal transaction' (PW vol. 3: 616; MW: 478).

I propose to segment *divinajanami* as *divi na janami* 'I do not (act as a) witness (of) the document', literally 'I do not know the document' (Skt. *dipiṃ na jānāmi*).[32] When read that way, the Gāndhārī phrase has an exact cognate in Bactr. λιβο (να) ζανινδο 'they do (not) witness (literally: 'know') the document', which occurs at the beginning of Doc A (13 October 332 CE, Rob), in a context bearing close resemblance to passage (9):[33] both texts mention a contract being written (*likhita … ya karaṇu* ~ νοβιχτο μο ολοβωστογο) and witnesses writing their signatures (*sakṣia … likhiṣati* ~ οιγαλφανο … πιδοναμγο νοβιχτιγινδο).

(10) Doc A, lines 1–5

> … καλδο νοβιχτο μο ολοβωστογο … αζδηβδο ειμουανο αστηβιγανο αζαβοργανο οιγαλφανο μολροδηβγανο κιδο ι ωνδο λιβο ζανινδο ταδο αβο μασκο πιδοναμγο νοβιχτιγινδο οδο κιδο ι ωνδο λιβο ναζανινδο ταδοιηνο αβο μο νιγο λιστοβαρο κιρδο …
>
> '… when this marriage contract was written … with the cognizance of these freemen of Asteb, the witnesses (and) those who have impressed (their) seals (upon it), who witness the present document and (whose) signatures are written hereupon—and those who do not witness the present document have put (their signatures) on the other manuscript …'. (BD I^2: 26–27)

By adding the phrase *divi na janami* 'I do not witness (literally: 'know') the document', the scribe may have wanted to imply that he only wrote down the text, and, therefore, is not to be held legally responsible for its content: that responsibility lies with the witnesses who have officially signed the contract and know the exact conditions in which the transaction was completed.

---

[31] As noted in fn. 28, Falk reads *diviṇa* with a retroflex nasal, but I adopt Baums' reading with a dental nasal.

[32] As discovered by Ching (2014: 59; cf. also Ching 2021), Gandh. *divi* 'document' is also attested in some Kuča Prakrit documents, e.g. in CKD 869 = TS 42 (dir.sg. *divi*, ins.sg. *diviyena*).

[33] See further Doc aa1–6 (BD I^2: 146–147).

More tentatively, I would like to argue that 'to (not) witness a document' is an administrative formula whose roots go back to the Official Aramaic writing tradition in Bactria and Gandhāra in Achaemenid times (6th–4th centuries BCE).[34] A first argument in favour of this hypothesis is the etymology of Gandh. *divi-* < *dipi-* 'document' and Bactr. λιβο 'idem': since these words are independent borrowings from OP *dipi-* (Sims-Williams 2002: 227–228; Falk 2010: 208–209), they entered Pre-Gāndhārī and Pre-Bactrian in the Achaemenid period. Secondly, 'to (not) witness (literally: know) a document' is reminiscent of the Official Aramaic formula *PN yd' ṭ'm' znh* 'PN is cognisant of (literally: knows) this order', which, like the Gāndhārī formula, is used in colophons mentioning the scribe of a particular document (Aram. *spr*).[35] In the Aramaic formula, *yd'* 'know' is used in a technical sense 'be cognisant of', 'be aware of', 'be in charge of'.[36] Although similar uses of verbs meaning 'know' are attested in other administrative traditions,[37] this polysemy of Aram. *yd'* 'know' invites a comparison with the specialised sense '(act as a) witness' of Gandh. *jan-* 'know' and Bactr. ζαν- 'idem'[38] – especially so when, as I suggested above, the person 'witnessing' a document is the one to be held responsible for what is stipulated in the document.

[34] The Aramaic texts from Achaemenid Bactria have been published by Naveh & Shaked (2012). In Gandhāra, Aramaic documents have not yet been found. Nevertheless, they must have existed because Kharoṣṭhī is a reworking of the Aramaic script, and Aramaic writing habits and phraseology were continued in Gāndhārī (cf. e.g. Sims-Williams 1996: 81; Yakubovich 2006: 338; Baums 2014, 2022).

[35] In the Aramaic documents from Egypt, the chancellor (*b'l ṭ'm*, literally 'master of the order') is the person "knowing the order", while the scribe is mentioned in the next sentence. In the documents from Bactria, usually one person is chancellor and scribe at the same time, with the result that the scribe becomes the person "knowing the order". For detailed discussions of the Aramaic formula, see Folmer (2017: 424–432) and Tavernier (2017: 355–378, 2020: 87–94); the latter also discusses Elamite and Demotic versions of the same formula.

[36] The glosses are taken from the translations of Aramaic letters in Driver (1957) and Naveh & Shaked (2012).

[37] See e.g. Ching & Ogihara (2010: 85) on similar uses of Tocharian B *kərsa-* 'know, understand, recognise' and Chinese *zhī* 知 'know'.

[38] This specialised meaning of Gandh. *jan-* and Bactr. ζαν- is not restricted to the administrative formula under discussion. For the same meaning in different contexts, cf. e.g. *aṃñe sakṣi janati* 'other witnesses bear witness (of it)' (CKD 415 cr2, tr. mine) or οιγαλφανο κιδο ... οτασιμο ιερο ζανινδο 'witnesses who … and bear witness concerning this matter' (Doc J' line 3, BD I^2: 48–49). Gandh. *jan-* can additionally be used to express that one takes responsibility over other people; cf. e.g. *ka ci puna iśa agamiṣyati emeva aho teṣa tanu saṃña janiṣyami* 'whichever of your people comes here in the future I will look upon them as my own' (CKD 161 cr3–4; TKD: 30).

5. Conclusion

This paper has aimed to show how evidence from Gāndhārī (and other Indian languages) can shed light on problems in Eastern Middle Iranian and vice versa. This has been done on the basis of three case studies in which we have examined texts in various languages and from different periods of time.

In Section 2, we were concerned with MBactr. *šwwh-* 'escape, avoid' and MSogd. *krmšwhn* 'absolution', words attested in Manichaean texts discovered in the Tarim Basin and datable to the later part of the 1st millennium CE. I have proposed that these words are loans from Gāndhārī: *šwwh-* ← Gandh. **śoh(e)-* < OIA *śodhaya-* 'purify, remove (something noxious), clear off, exculpate', and *krmšwhn* ← Gandh. **karma-śohana-* 'purification/removal of (bad) *karma*' < OIA **karma-śodhana-*. To support this proposal, I have adduced phraseological evidence drawn from texts in better-attested Indian languages, viz. Pāli and Buddhist Sanskrit.

Section 3 has treated a further Indian loanword into Bactrian, this one attested in a birchbark letter (Doc kd) from SE Afghanistan (4th century CE). I have argued that the sequence πιδοβισταρο in this letter should be divided into two words, i.e. (i) the common preposition πιδο 'in, on, with, etc.' < **pati*, and (ii) a so far unattested noun βισταρο 'detail' ← Gandh. (less likely Skt.) *vistara-* 'detail, expansion'. Interpreted that way, πιδο βισταρο 'in detail' nicely parallels the phraseology of ins.sg. *vistarena* 'idem' as used in the Niya documents (3rd–4th centuries CE), although βισταρο must be a loan from a Gāndhārī dialect spoken in the Indo-Iranian borderlands rather than from a Tarim Basin variety like Niya Prakrit.

In Section 4, finally, we dealt with a legal formula from the Gāndhārī silk fragment CKM 430 (presumably from the Gilgit area, perhaps 4th century CE). I have suggested that the sequence *divinajanami* should be understood as *divi na janami* 'I do not witness (literally: know) the document', thus being cognate with the Bactrian formula λιβο να ζανινδο 'they do not witness (literally: know) the document' in Doc A (332 CE). In addition, I have hypothesised that this piece of legalese has its roots in Achaemenid Official Aramaic (6th–4th centuries BCE). If this is true, it shows that the disciplinary boundaries between Indology and Iranian studies are not the only ones worth crossing: Indo-Iranianists may find an occasional glance at scholarship on Semitic languages also to be worth their while.[39]

[39] For a comparative approach to Indo-Iranian studies and Semitics, see e.g. Schwartz (2013). He connects Skt. *sārtha-* 'caravan' with Arabic *sayyāra(t)-* and Aram. *šayyārtā-*, both 'caravan', and argues that *sārtha-* is a loan from South Arabian adopted as a result of the Indian Ocean trade.

Appendix

An annotated translation of CKM 430, a sale contract written on silk

Since the analysis of the sequence *divinajanami* is not the only point in which my interpretation of CKM 430 differs from Falk's, I present in this appendix a revised translation of the entire document. The Gāndhārī text is cited from the online catalogue (Baums & Glass 2002–) and accompanied by a brief apparatus mentioning the more substantial differences between Baums' online edition and Falk's printed edition (cf. fn. 28).[40] The notes justify certain aspects of my translation, they comment on linguistically interesting forms, and they point to parallel passages and/or idioms in Bactrian that help with the interpretation of the Gāndhārī.

Text

1 /// *[ñ]uputra kṣetru yatra gachati*[a] *biyaṣa droṇa 10 1 1 vikriṇami Pigeaṣa Hastika-poravastavaṣa*

2 /// *[kṣ](*e)traṣa karanena japeati tena Palolarayaṣa daḍu dadavi siati su sadera 2*[b] *Anegaṣara kṣatravaṣa*

3 /// *(*kṣe)tru*[c] *vikriṇami ya karaṇu rayakaüaku dadavi vikritu mea Jeṭhavarmena kritu mea Piena Kaporavastavena*

4 /// *likhita mea Ṣaghamitreṇa Pigeṣa ajeṣaṇaï ya*[d] *karaṇu divi na*[e] *janami atra sakṣia upari likhiṣati*

[a] *gachati* (Baums); *garjati* (Falk). [b] *su sadera 2* (Baums); *susadera ya* (Falk).
[c] *(*kṣe)tru* supplied by Falk in his Sanskrit *chāyā*. [d] *ajeṣaṇaï ya* (Baums, supported by ins.sg. *ajiṣanayi* in CKD 661 o7); *ajeṣaṇa iya* (Falk). [e] *divi na* (Schoubben, cf. §4); *divina* (Baums); *diviṇa* (Falk).

Translation

I, son of …, sell a field, in which twelve *droṇa*s of seed are spent … of Pige, an inhabitant of Hastikapora … (Whoever) might make a claim concerning this field, he must pay a penalty to the Palola king, (namely) two golden staters, (and) … to the *kṣatrapa* Anegaṣara … I sell this field … must give this contract [*karaṇu*] stored in the royal (archive) (?). I Jeṭhavarma sold (the field), I Pia, an inhabitant of Kapora, bought (it) … This contract [*karaṇu*] has been written by me, Ṣaghamitra, at the request of Pige. I (Ṣaghamitra) do not witness [i.e. am not responsible for] the document. The witnesses in this matter will write (their names) above.

[40] The right margin of the silk fragment is not so easily readable on the plates provided in Falk's article. Therefore, the reading of the first word in lines 1, 2, and 3 is tentative.

Notes

Line 1

My translation of pres.3pl. *gachati*,[41] literally ‘go’, with ‘are spent’ is supported by the fact that Bactr. þαο-, pret. þοδο, ‘go’ can also mean ‘be spent’ in this type of context (see BD II: 283). The same phraseology is moreover attested in the Tarim Basin languages: in Tocharian B, pres. *y*-, pret. *m*-, ‘go’ can have a derived meaning ‘be spent’ (Ching & Ogihara 2010: 116);[42] Khot. *tsu*- ‘go’ and its Tumshuqese cognate *ch*- ‘idem’ can likewise be used for expenses (Ogihara & Ching 2024: 317–318 with fn. 61); and the same is true of Niya Pkt. *nikas*- < OIA *niṣ-kas*- ‘to go out’ (e.g. *vyaga nikasta* ‘expenditure was incurred’ in CKD 637 o2) and *nikram*- < OIA *niṣ-kram*- ‘to go out’ (e.g. *aṃnapačevara nikraṃta* ‘the expenditure of corn for their provisions’ in CKD 637 o6).

Line 2

Falk (2021: 12) comments that ‘[t]he king is involved in that he will punish the seller in case the latter would accept the payment but later on declare that neither transaction nor payment ever took place.’ He then adds that the reasons for inflicting a punishment are ‘not expressed verbally’. On the latter point, I disagree. Because Falk takes *karanena* to mean ‘through this document’ and he analyses *japeati* as an indicative present (cf. Skt. *jalpayati* in his *chāyā*), he assumes that *[kṣ](*e)traṣa karanena japeati* and the part starting with *tena* ‘by him’ are co-ordinated main clauses. But, if we take *karanena* to mean ‘regarding’[43] and analyse *japeati* as an optative (thus also Baums in the online edition), *[kṣ](*e)traṣa karanena japeati* can be the final part of a preposed subclause stipulating in which cases a fine will be imposed. Similar structures in Bactrian sale contracts confirm that this syntactic analysis is the correct one; compare, for instance, οδαλδο ασταδο μαπαρσο ζαμανο ... κιδο ... σιγισο χοηναδο οισηραδο ταδιιο ι χοανδο οδο οισαρο ναοαραδο οτο λαυαδο αβινδ{α}αμο φαρο υασταμογαρανο δινα ρο οιστο οδο φρηλο φαρο ιπαδαρλιγανο ‘And if in the future there should be (anyone) … who

[41] I assume *gachati* to be 3pl. because its subject is the dir.pl. *droṇa 10 1 1* ‘twelve *droṇa*s’: *-ati* can stand for /anti/ given that the scribe of CKM 430 does not make use of anusvāra (see also below on *likhiṣati* ‘will write’). For more examples of 3pl. forms in *-ati* in a text in which anusvāra is mostly not written, see e.g. Cox 2025: 250, 298 on British Library fragment 28. Compare also the Niya Prakrit instances of pret.3pl. forms in *-dati* (next to *-taṃti*), on which see LKD: §105.

[42] One such example is *śiktālyeś ysāre masa* ‘wheat went/has been spent for sowing’ (PK DA M 507.14 a1; Ching 2013: 364).

[43] Cf. the Niya Prakrit combination of *karaṃna* ‘regarding’ and *jalp*- ‘to speak’ in *avi vasaṃtaṃmi uṭaṣa karaṃna tahi vaṃti jalpita* ‘also in spring he spoke to you on the matter of the camels’ (CKD 594 o2–3; TKD: 127).

might claim or argue … (over) anything, then his claim and argument shall be invalid, and he shall pay a penalty to (their) excellencies of twenty dinars, and the same to the opponents' (Doc ab, lines 10–18; BD I[2]: 148–149). The opt.3sg. *japeati* corresponds to the Bactrian subj.3sg. χοηναδο 'will claim', literally 'will say' (cf. BD II: 280b).

For *daṇḍa-* 'punishment, fine' + *dā-* 'give' (*daḍu dadavi siati*) in the meaning 'pay a fine', Falk (2021: 12) cites *varam ahaṃ daṇḍaṃ dadyām* 'I better pay that fine' in *Avadānaśataka* 80. One may further compare *Manusmṛti* 8.274 (*dāpyo daṇḍaṃ* 'must be compelled to pay a fine'), 8.341 (*na daṇḍaṃ dātum arhati* 'he ought not to pay a fine'), 8.383 (*daṇḍaṃ dāpyo*), 9.229 (*daṇḍaṃ dātum*);[44] and various instances of the same phraseology in the Niya documents (e.g. *eta daṃḍa daditva* 'having paid this fine' in CKD 345 cr4). *daṇḍa-* + *dā-* parallels Bactr. αβινδαμο λαυ- 'pay (literally: give) a fine' (cf. λαυαδο αβινδ{α}αμο in the quotation above and λα̅η̅ιο … αβινδαμο in Doc C line 18 = Doc C' line 20; BD I[2]: 34–35).

I suspect that a second fine was mentioned in the missing text after *Anegaṣara kṣatravaṣa* 'to the *kṣatrapa* Anegaṣara'; compare again the Bactrian documents where there are usually two fines mentioned in this type of sentence, including the passage from Doc ab cited above.

Line 3

The semantics of *rayakaüaku* are not fully clear, which, combined with the missing portion of text at the beginning of the line, complicates the syntactic analysis and overall interpretation of this part of the document. Though somewhat reluctantly because of the required lenition of *-pt-* to *-Ø-*, I follow Falk (2021: 13) in assuming that *rayakaüaku* continues OIA *rājaka-guptaka-*. He translated this compound as 'king's [*rājaka*] permanent loan [*guptaka*]', hypothesising that 'every land-owner possesses a written permanent loan issued by the king, which must be handed over to the new proprietor'. I opt for 'stored [*guptaka*] in the royal (archive) [*rājaka*]' because (i) we know from the Bactrian documents that duplicates of contracts were stored in royal archives (cf. Sheikh 2023: 23–24), and (ii) *rayaka* by default means 'royal' rather than 'king' in the Niya documents.[45]

What precisely is meant by 'this contract must be given' (*ya karaṇu … dadavi*) and in which circumstances this ought to be done is admittedly not clear to me. For the

[44] Cited from Olivelle (2005: 714, 727, 736, 788).

[45] See e.g. *rayakade aṃnatilekha kṛdama* 'we wrote a letter of command from the royal (office) [abl.sg. *rayakade*]' (CKD 399 o4, tr. mine). Note in passing that the apparent usage of *kṛdama*, literally 'we made', in the sense 'we wrote' may be due to calquing from an Iranian language since Bactr. κιρδο 'made' and its Sogdian cognate *krt* 'idem' are more often used in the sense 'wrote' (cf. BD II: 223 and Sims-Williams 2023b: 39, who also compares MP *nāmag kardan* 'to write (literally: make) a letter').

phraseology, compare Bactr. παροβωστογο ... λαδο ‘to give a loan contract’ in Doc Q line 27 (BD I^2: 90–91).

Line 4

Here, as in line 3, I consider *ya* to be a variant of *iya(ṃ)* ‘this’.[46] Interpreted thus, *ya karaṇu* equals μο πωστογο ‘this contract’ in Bactrian colophons of the same type, e.g. οτο μο πωστογο μανο νιβιχτο αζαδοφαρδαρο πιδο οηþολαδο φρομανο ‘and this contract was written by me, Azad-fardar, at the command of Wesh-lad’ (Doc ab lines 19–20; BD I^2: 148–149).

Falk (2021: 14) glosses *atra sakṣia* as a nom.sg. *atrasākṣikaḥ* in his Sanskrit *chāyā* and he translates with “the witness in this case”. Yet, elsewhere in his paper (2021: 12, 15), he renders *sakṣia* with the plural ‘witnesses’. I favour the latter translation and analyse *sakṣia* as a dir.pl. in *-a* to the *i*-stem *sakṣi* < OIA *sākṣin-* (analogical after the *a*-stems).

Consequently, I take *likhiṣati* ‘will write’ as a 3pl. verb form. This requires that *-ati* stands for /anti/, which seems possible given that the scribe of CKM 430 does not write preconsonantal nasals (cf. e.g. *daḍu* < *daṇḍam* in line 2, and see fn. 41). For my assumption that *likhiṣati* governs an unexpressed complement ‘their names’, some instances of Bactr. ναβισ-, pret. ναβιχτο, ‘write’ provide relevant parallels, e.g. οιγαλφανο ακιδδηνο μοδδραγο χιρσοβωστιγο αβο μανιυο πιδοναμο οιγαλφο **ναβιχτιγινδο** ‘witnesses who **have written (their names** as) witnesses (to this) sealed document, (this) purchase contract, on the other signed document’ (Doc Nn lines 5–7, BD I^2: 74–75). Note, finally, that *likhiṣati* is reminiscent of the futures in *-iṣa-* < OIA *-iṣya-* in the *Khotan Dharmapada*, on which see Caillat (1977–1978).[47] Since in other Gāndhārī texts the OIA future tense morpheme *-iṣya-* is continued as *-iśa-*, one wonders whether these futures in *-iṣa-* could be dialectal.[48]

Abbreviations

abl.	ablative	Ardh.	Ardhamāgadhī
Aram.	Aramaic	Arm.	Armenian

[46] In both instances, *ya* is preceded by a word ending in *-i* (*vikriṇami*, *ajeṣaṇai*), which could be a reason why the initial *i-* in *iya(ṃ)* disappeared (or was left unwritten).

[47] The attested forms in the *Khotan Dharmapada* are: *kariṣadi* ‘will make’ (v. 125; Brough 1962: 137), *bhaviṣadi* ‘will be(come)’ (v. 188; Brough 1962: 149), and *akamiṣadi* ‘will come’ (vv. 209, 210; Brough 1962: 152).

[48] Two further instances of Gāndhārī futures in *-iṣa-*, i.e. *nivartiṣadi* ‘will occur’ and *upa(*j)iṣadi* ‘will arise’, have turned up in the recently published Abhidharma text contained in British Library fragment 28 (Cox 2025: 256, 273, 300, 403).

Bactr.	Bactrian	OIA	Old Indo-Aryan
caus.	causative	OP	Old Persian
CKD	see Baums & Glass 2002–a	opt.	optative
CKM	see Baums & Glass 2002–b	Parth.	Parthian
co	cover-tablet obverse	p.c.	personal communication
cr	cover-tablet reverse	PIr.	Proto-Iranian
dir.	direct (case)	PK DA	Pelliot Koutchéen Duldur Aqur
DN1	Dašt-i Nāwur inscription 1	Pkt.	Prakrit
fem.	feminine	pl.	plural
fut.	future	PN	personal name
Gandh.	Gāndhārī	pres.	present
ins.	instrumental	pret.	preterite
Khot.	Khotanese	Rab.	Rabatak inscription
MBactr.	Manichaean Bactrian	sg.	singular
MP	Middle Persian	Skt.	Sanskrit
MParth.	Manichaean Parthian	subj.	subjunctive
MSogd.	Manichaean Sogdian	uo	under-tablet obverse
nom.	nominative	v(v).	verse(s)
o	obverse		

Symbols

←	borrowed from	|	verse boundary
>	developed into	~	corresponds to
<	developed from	< >	orthography
**x*	reconstructed form (diachronic)	[ ]	phonetics
***x*	expected, but wrong form	{ }	editorial deletion

References

ADAMS, Douglas Q., 2013. *A dictionary of Tocharian B: Revised and greatly enlarged*, Amsterdam–New York.

ALLON, Mark, 2008. 'Recent discoveries of Buddhist manuscripts from Afghanistan and Pakistan and their significance', in *Art, architecture and religion along the Silk Roads* (ed. Ken PARRY), Turnhout, 153–178.

BAILEY, Harold W., 1938. *The content of Indian & Iranian studies: An inaugural lecture delivered on 2 May 1938*, Cambridge.

BAUMS, Stefan, 2009. *A Gāndhārī commentary on early Buddhist verses: British Library Kharoṣṭhī fragments 7, 9, 13 and 18*, PhD Dissertation University of Washington, Seattle.

BAUMS, Stefan, 2014. ‘Gandhāran scrolls: rediscovering an ancient manuscript type’, in *Manuscript cultures: Mapping the field* (ed. Jörg B. QUENZER *et al.*), Berlin–Boston, 183–225.

BAUMS, Stefan, 2022. ‘The earliest colophons in the Buddhist northwest’, in *The syntax of colophons: A comparative study across pothi manuscripts* (ed. Nalini BALBIR & Giovanni CIOTTI), Berlin–Boston, 15–41.

BAUMS, Stefan & Andrew GLASS, 2002–a. *Catalog of Gāndhārī texts: Documents*. https://gandhari.org/catalog?section=documents.

BAUMS, Stefan & Andrew GLASS, 2002–b. *Catalog of Gāndhārī texts: Manuscripts.* https://gandhari.org/catalog?section=manuscripts.

BD I^2 = SIMS-WILLIAMS, Nicholas, 2012. *Bactrian documents from Northern Afghanistan, I: Legal and economic documents* (revised edition), London.

BD II = SIMS-WILLIAMS, Nicholas, 2007. *Bactrian documents from Northern Afghanistan, II: Letters and Buddhist texts*, London.

BD IV = SIMS-WILLIAMS, Nicholas, 2025. *Bactrian documents IV: Documents from south of the Hindukush, I* (with a contribution by Frantz GRENET), London.

BERNHARD, Franz, 1965. *Udānavarga. Band I: Einleitung, Beschreibung der Handschriften, Textausgabe, Bibliographie*, Göttingen.

BOYCE, Mary N. E., 1954. *The Manichaean hymn-cycles in Parthian*, London.

BOYCE, Mary N. E., 1960. *Catalogue of the Iranian manuscripts in Manichean script in the German Turfan collection*, Berlin.

BROUGH, John, 1962. *The Gāndhārī Dharmapada*, London.

CAILLAT, Colette, 1977–1978. ‘Forms of the future in the Gāndhārī Dharmapada’, *Annals of the Bhandarkar Oriental Research Institute* 58–59, 101–106.

CHEN, Ruixuan & Diego LOUKOTA, 2018. ‘Mahāyāna Sūtras in Khotan: Quotations in chapter 6 of the *Book of Zambasta* (I)’, *IIJ* 61/2, 131–175.

CHING, Chao-jung, 2013. ‘The activities of Sogdian Buddhists in Kucha as observed in the Tocharian B secular documents’, in *Buddhism among the Iranian peoples of Central Asia* (ed. Matteo DE CHIARA *et al.*), Wien, 333–384.

CHING Chao-jung 庆昭蓉, 2014. ‘Kezier chutu De cang Quluwen Qiuciwang zhaoyu yu qiyue wenshu yanjiu 克孜尔出土德藏佉卢文龟兹王诏谕与契约文书研究 – On the decree of Kuchean King and two pieces of contracts written in Kharoṣṭhī among the findings from the Kizil Grottoes and kept in Berlin’, *XYWS* 9, 51–73.

CHING, Chao-jung, 2021. ‘Kucha’, in *Encyclopaedia Iranica Online.* https://referenceworks.brill.com/display/entries/EIRO/COM-337721.xml.

CHING, Chao-jung & Hirotoshi OGIHARA, 2010. ‘Internal relationships and dating of the Tocharian B monastic accounts in the Berlin collection’, *SIAL* 25, 75–135.

CONE, Margaret, 1989. ‘Patna Dharmapada. Part I: text’, *Journal of the Pāli Text Society* 13, 101–217.

COX, Collett, 2025. *A Gāndhārī Abhidharma text: British Library Kharoṣṭhī fragment 28*, Seattle.

DKS = BAILEY, Harold W., 1979. *A dictionary of Khotan Saka*, Cambridge.

DMMP = DURKIN-MEISTERERNST, Desmond, 2004. *Dictionary of Manichaean Middle Persian and Parthian,* Turnhout.

DMSB^2 = SIMS-WILLIAMS, Nicholas & Desmond DURKIN-MEISTERERNST, 2022. *Dictionary of Manichaean Sogdian and Bactrian* (revised and enlarged by Nicholas SIMS-WILLIAMS), Turnhout.

DRIVER, Godfrey R., 1957. *Aramaic documents of the fifth century* (revised and abridged), Oxford.

EDIV = CHEUNG, Johnny, 2007. *Etymological dictionary of the Iranian verb*, Leiden–Boston.

FALK, Harry, 2010. 'Foreign terms in Sanskrit pertaining to writing', in *The idea of writing: Play and complexity* (ed. Alex DE VOOGT & Irving FINKEL), Leiden–Boston, 207–217.

FALK, Harry, 2021. 'Two fragmentary Kharoṣṭhī inscriptions', *ARIRIAB* 24, 7–16.

FOLMER, Margaretha L., 2017. 'Bactria and Egypt. Administration as mirrored by the Aramaic sources', in *Administration in the Achaemenid Empire: Tracing the imperial signature* (ed. Bruno JACOBS *et al.*), Wiesbaden, 413–454.

HALFMANN, Jakob, Svenja BONMANN, Simon FRIES & Natalie KOROBZOW, 2024. 'A new reading of the Bactrian part of the Dašt-i Nāwur trilingual (DN I)', *IIJ* 67/1, 5–51.

HINÜBER, Oskar VON, 2001. *Das ältere Mittelindisch im Überblick*, Wien.

HINÜBER, Oskar VON & Kenneth R. NORMAN, 1994. *Dhammapada*, Oxford.

LENZ, Timothy, 2010. *Gandhāran avadānas: British Library Kharoṣṭhī fragments 1–3 and 21 and supplementary fragments A–C*, Seattle.

LINDSAY, Rory, 2024. *Saving the dead. Tibetan funerary rituals in the tradition of the Sarvadurgatipariśodhana Tantra*, Wien.

LKD = BURROW, Thomas, 1937. *The language of the Kharoṣṭhi documents from Chinese Turkestan*, Cambridge.

LOUKOTA, Diego, 2023. '*ne hāḍe vajrropamä vaśärä*: Indic loanwords in the Khotanese Book of Zambasta and the chronology of the spread of Buddhism to Khotan', *BuddhistRoad Paper* 1/7, 1–34.

LURJE, Pavel B., 2021. 'Buddhist Indian loanwords in Sogdian and the development of Sogdian Buddhism', *Entangled Religions* 11.6, 1–42.

MEYER, Robin, 2023. *Iranian syntax in Classical Armenian: The Armenian perfect and other cases of pattern replication*, Oxford.

MICHETTI, Francesca, 2024. 'A Bactrian sound change (OIr. **š* > Bactr. *h*)', *ZDMG* 174/1, 173–190.

MW = MONIER-WILLIAMS, Monier, 1899. *A Sanskrit-English dictionary: Etymologically and philologically arranged with special reference to cognate Indo-European languages* (new edition, greatly enlarged and improved), Oxford.

MYLIUS, Klaus, 2003. *Wörterbuch Ardhamāgadhī-Deutsch*, Wichtrach.

NAVEH, Yoseph & Shaul SHAKED, 2012. *Aramaic documents from ancient Bactria (fourth century BCE.): From the Khalili Collections*, London.

NORMAN, Kenneth R., 1971. *The elders' verses II (Therīgāthā)*, London.

NORMAN, Kenneth R., 1997. *The word of the doctrine (Dhammapada)*, Oxford.

OGIHARA Hirotoshi & CHING Chao-jung 2024. 'Miscellany on the Tumshuqese documents (III): Words relating to cereals', in *Siddham. Studies in Iranian philology in honour of Mauro Maggi* (ed. Gerardo BARBERA *et al.*), Wiesbaden, 305–323.

OLIVELLE, Patrick, 2005. *Manu's code of law: A critical edition and translation of the Mānava-dharmaśāstra*, Oxford.

PTSD = RHYS DAVIDS, Thomas W. & William STEDE, 1921–1925. *The Pali Text Society's Pali-English dictionary*, Chipstead.

PW = BÖHTLINGK, Otto & Rudolph ROTH, 1855–1875. *Sanskrit Wörterbuch*, St. Petersburg.

SALOMON, Richard, 2008. *Two Gāndhārī manuscripts of the songs of lake Anavatapta (Anavatapta-gāthā): British Library Kharoṣṭhī fragment 1 and Senior scroll 14*, Seattle.

SCHLOSSER, Andrea, 2022. *Three early Mahāyāna treatises from Gandhāra: Bajaur Kharoṣṭhī fragments 4, 6, and 11*, Seattle.

SCHOUBBEN, Niels, 2022. 'Linguistic evidence for Kuṣāṇa trade routes: Bactrian *λιρτο 'load, cargo' and Sanskrit *lardayati* 'to load', *IF* 127, 343–357.

SCHWARTZ, Martin, 2013. '*Sārtha-* and other caravan words', *BAI* 23 (2009[2013]), 179–180.

SHEIKH, Hossein, 2023. *Studies of Bactrian legal documents*, Leiden–Boston.

SIMS-WILLIAMS, Nicholas, 1983. 'Indian elements in Parthian and Sogdian', in *Sprachen des Buddhismus in Zentralasien: Vorträge des Hamburger Symposions vom 2. Juli bis 5. Juli 1981* (ed. Klaus RÖHRBORN & Wolfgang VEENKER), Wiesbaden, 132–141.

SIMS-WILLIAMS, Nicholas, 1989. 'Sogdian', in *Compendium Linguarum Iranicarum* (ed. Rüdiger SCHMITT), Wiesbaden, 173–192.

SIMS-WILLIAMS, Nicholas, 1996. 'From Babylon to China: Astrological and epistolary formulae across two millennia', in *La Persia e l'Asia Centrale da Alessandro al X secolo* (ed. Academia Nazionale dei Lincei), Roma, 77–84.

SIMS-WILLIAMS, Nicholas, 2002. 'Ancient Afghanistan and its invaders: Linguistic evidence from the Bactrian documents and inscriptions', in *Indo-Iranian languages and peoples* (ed. Nicholas SIMS-WILLIAMS), Oxford, 225–242.

SIMS-WILLIAMS, Nicholas, 2008. 'The Bactrian inscription of Rabatak: A new reading', *BAI* 18 (2004[2008]), 53–68.

SIMS-WILLIAMS, Nicholas, 2009. 'The Bactrian fragment in Manichaean script (M 1224)', in *Literarische Stoffe und ihre Gestaltung in mitteliranischer Zeit* (ed. Desmond D. MEISTERERNST *et al.*), Wiesbaden, 245–268.

SIMS-WILLIAMS, Nicholas, 2011. 'Remarks on the phonology of the Manichaean Bactrian fragment (M 1224).' in *Leksika, ètimologija, jazykovye kontakty: K jubileju doktora filologičeskix nauk, professora Džoj Iosifovny Èdel'man* (ed. Elena K. MOLČANOVA *et al.*), Moskva, 244–251.

SIMS-WILLIAMS, Nicholas, 2021. 'A Buddhist technical term in Christian Sogdian', *Pis'mennye pamjatniki Vostoka* 18, 32–36.

SIMS-WILLIAMS, Nicholas, 2023a. 'A new collection of Bactrian letters on birchbark with an additional note by Frantz GRENET', *BAI* 31 (2022–2023[2023]), 135–144.

SIMS-WILLIAMS, Nicholas, 2023b. *The "ancient letters" and other early Sogdian documents and inscriptions* (with contributions by Frantz GRENET), London.

SIMS-WILLIAMS, Nicholas, John S. SHELDON & Zsuzsanna GULÁCSI, 2022. *A Manichaean prayer and confession book*, Turnhout.

SKJÆRVØ, Prods O., 2004. *This most excellent shine of gold, king of kings of sutras: The Khotanese Suvarṇabhāsottamasūtra*, Harvard.

SKORUPSKI, Tadeusz, 1983. *The Sarvadurgatipariśodhana Tantra (elimination of all evil destinies): Sanskrit and Tibetan texts with introduction, English translation and notes*, Delhi–Varanasi–Patna.

TAVERNIER, Jan, 2017. 'The use of languages on the various levels of administration in the Achaemenid Empire', in *Administration in the Achaemenid Empire: Tracing the imperial signature* (ed. Bruno JACOBS *et al.*), Wiesbaden, 337–412.

TAVERNIER, Jan, 2020. 'Persian in official documents and the processes of multilingual administration', in *Aršāma and his world. The Bodleian letters in context, volume iii: Aršāma's world* (ed. Christopher J. TUPLIN & John MA), Oxford, 75–96.

TKD = BURROW, Thomas, 1940. *A translation of the Kharoṣṭhi documents from Chinese Turkestan*, London.

TRENCKNER, Vilhelm, 1888. *The Majjhima-nikāya Vol. I*, London.

YAKUBOVICH, Ilya, 2006. 'Marriage Sogdian style', in *Iranistik in Europa – Gestern, Heute, Morgen* (ed. Heiner EICHNER *et al.*), Wien, 307–344.

The Bactrian inscription of Ayrtam: A minimal reading

Nicholas SIMS-WILLIAMS

This paper contains a new edition of an early Kushan Bactrian inscription from Ayrtam in southern Uzbekistan. The author proposes a "minimal reading" of the poorly-preserved text, avoiding the speculative readings and restorations found in earlier editions.

1. Introduction

The Bactrian inscription of Ayrtam is carved on the pedestal of a limestone block containing the lower part of two standing figures, one male and one female. The sculpture and inscription were discovered in 1979 by B. A. Turgunov in his excavation of the northeastern part of the monumental building of the cult complex at Ayrtam, a site situated 18 km east of Termez in the Surkhan Darya Province of Uzbekistan, and is now preserved in Tashkent, in the Museum of the History of Uzbekistan (Inventory no. A13936, Collection no. 199/8). The surface of the pedestal is about 78 cm wide and 23 cm high.[1] The text, consisting of six lines, is inscribed on the upper part of this surface, occupying about 11 cm at the left and about 14 cm towards the right, where the writing tends to straggle downwards. Only traces are visible at the beginnings of the lines, where the surface is eroded, and some letters are lost at the ends of most lines and in the middle of lines 1–2, where part of the stone has broken away completely. The letters are of the "monumental" type characteristic of the early Kushan period (2nd cent. CE). Their size and spacing is somewhat variable, though the individual letters are generally formed in a consistent way.

A partial reading and interpretation of the inscription was published in 1981 by V. A. Livšic and È. V. Rtveladze in collaboration with the excavator, B. A. Turgunov. According to their preliminary reading the inscription was dated in the fourth year of Huvishka and recorded construction works by a certain Shodia (ϸοδια), while the last line of the inscription states that it was written by Mihir-zad (μιροζαδα) at Shodia's command (Turgunov *et al*. 1981). An attempt at a complete reading and interpretation was offered in 1986 by J. Harmatta, who unfortunately had access only to the quite

[1] Most of these details are taken from the chapter 'The Airtam inscription: text, translation, interpretation' in Rtveladze 2021: 148–167.

10.29091/9783752003635/010

inadequate published photos. According to Harmatta, the inscription dates from the thirtieth year of Huvishka (reading λ′ '30' rather than δ′ '4'), and the name of the principal figure is Shodila (ϸοδιλα) rather than Shodia. Otherwise, he accepted almost all of his predecessors' readings, filling in all the gaps which they had left with adventurous new proposals of his own.[2]

Meanwhile, Livšic and Rtveladze continued their work on the inscription, partly independently and partly in collaboration. Rtveladze published his own revised edition and translation at least twice, in Russian and English.[3] So far as I am aware, Livšic never published a new reading of the text, though his translation is cited by Rtveladze in two books.[4] In addition to photos of the inscription (including those reproduced here as Fig. 1–4), many drafts, in the form of drawings and typed or handwritten notes by both authors, are preserved in the Oriental Department of the State Hermitage Museum, St. Petersburg (Livšic archive, folder Bactria-42: Ayrtam). These drafts have kindly been made available to me by Dr Pavel Lurje, but I have decided that it would not be fair or useful to cite them, as most of them clearly represent early ideas which the authors would hardly have wished to publish. However, I also have access to a handwritten text, sent to me many years ago by È. Rtveladze for inclusion in a putative volume of the Corpus Inscriptionum Iranicarum, which he described as the joint work of Livšic and himself. This will be cited in the commentary below (as "LR"), since it corresponds closely to Livšic's published translation and helps to explain the readings and restorations on which the latter is based.

Unlike Harmatta's work, which is too fanciful to contribute anything useful to the reading or interpretation, that of Livšic and Rtveladze must certainly be taken seriously, being based on careful (and in Rtveladze's case no doubt frequent) autopsy of the inscription. Nevertheless, it seems to me that the readings of Livšic and Rtveladze, as well as those of Harmatta, show a strong tendency to over-interpret vague traces of letters and to seek words and expressions with close parallels in better-preserved Kushan inscriptions, even against the epigraphic evidence. Unfortunately, such over-interpretation leads to false deductions regarding the historical content of the inscription. It is hardly an exaggeration to say that almost everything which has been written about the content of the inscription is without foundation (see the short discussion at the end of this paper). In my view, large parts of the inscription are illegible and will probably

[2] Harmatta 1986. Virtually the same text, apart from a few misprints, is included in Harmatta 1994: 432–433.

[3] Rtveladze 2018; Rtveladze 2021: 154–160. Each contains a few misprints, which can be corrected by comparison with the other.

[4] Rtveladze *et al.* 2000: 112–113; Rtveladze 2008: 155. There are only insignificant differences between the two translations. I am grateful for Frantz Grenet for providing me with copies of both.

always remain so. Moreover, many speculative readings and reconstructions offered in the editions are incompatible both with what can be seen on the stone and with what we now know of the Bactrian language.

There is of course an element of subjectivity in any reading of such a poorly-preserved inscription, and I do not claim that my reading represents the last word on the matter. I have to admit that each time I have had a chance to examine the original – in Tashkent in 1990 and 2004,[5] through the courtesy of the late Academician È. V. Rtveladze, and in the exhibition 'Archaeological Treasures of Uzbekistan: From Alexander the Great to the Kushan Empire' in the James-Simon-Galerie, Berlin, in 2023, thanks to the kindness of the curator, Dr Manfred Nawroth – I have been tempted to interpret some less legible parts of the text in different ways.[6] Similarly, it is clear from their publications, and even more from their unpublished drafts, that Rtveladze and Livšic changed their minds many times. In this paper, therefore, I propose what I would like to call a "minimal reading" of the inscription, limited to what I found legible or at least plausibly reconstructible on each occasion when I was able to study the original. Although the resulting interpretation may be less exciting than those previously proposed, it is less likely to mislead.

2. "Minimal" text and translation

1 [c. 15]κᾳ[c. 8]ǫρε̣ ι̣ βαο[c. 13]αǫ ѡ[c. 4]μι[•••]γ̣[c. 10]οογᾳ[
… the king(?) …

2 [c. 22]ο σταδο βοδβ̣α ǫ[c. 8]ιγᾳν[c. 5]β̣[••]ᾳ[•••]γ̣ο ζιδο α[
… was the *shudva* … struck(?) …

3 [c. 20]δο τι̣δηια ᾳβ̣ǫ μᾳ [λι]ζα ωσταδο[c. 12]ο λβαρǫ [
… Then he placed (it?) in the citadel … the court …

4 [ ]ᾳ[c. 18]δ̣ι λαχβ̣τᾳ σ̣ταδ̣ǫ αβαβσα τιδι β̣ǫδ̣βαᾳ δ̣ι ι̣••νο κανδο α[
… was(?) a desert, without water, then the *shudva* dug the …

5 [c. 14 σι]δ̣[ι]ᾳλ̣βογανδο ριζδι οτι οβεινα ζαμ̣νᾳ •[•]τομ̣[••]οβ̣[••]δο βα[
… which is called Alv-gand. And at this time …

6 [ ]ᾳ[c. 11]ο οτι ειμο μιιροζαδα νιβιχτο πιδο ια β̣οδ̣βαα φρομαν̣α *rest blank*
… And Mihir-zad wrote this at the *shudva*'s command.

[5] Not in 2006 as inadvertently stated by Rtveladze (2021: 153)!

[6] I am also grateful to Stefan Härtel for alerting me to the fact that the inscription was included in this exhibition.

3. Commentary

For each line, or part of a line, I first present the readings of the available editions (T = Turgunov *et al*. 1981; H = Harmatta 1986; LR = Livšic and Rtveladze [unpublished, see above]; R = Rtveladze 2018/ Rtveladze 2021), followed by my own, all arranged so that the equivalent letters stand one above the other as far as possible, and the translations by Livšic, Rtveladze and myself (L = Livšic in Rtveladze *et al*. 2000/Rtveladze 2008, translated by me from Russian into English; R = Rtveladze 2021; SW = Sims-Williams). The text of each edition is reproduced as faithfully as possible, but transcribed into a consistent system of notation: subscript points = letters not fully visible; [ ] = letters wholly lost or illegible; ••• = approximate number of missing or illegible characters; < > letters assumed to have been omitted by the mason. In the case of the partial edition T, the gaps between the parts for which a reading is proposed are indicated by the ellipsis (…).

Line 1

T	þαοοηοþκο̣/ι̣	ι	[χ]þονο	δ	κα[λδ]ο	…	þαο …
H	[þα]ο̣ ο̣ο̣ηþ̣κο̣	ι	χ̣þ̣ο̣νο̣	λ	κα[λδ]ι̣	ι	α̣[ρδο]χþο φαρρ[ο πιδογαρα]
LR	ι þαο̣ ο̣ηο̣þκο̣	ι	χþον[ο]	δ	κα[λ]δο	[••••••]ο̣ρ[a]ε	ι þα̣ρ̣[ο •••••••
R	þαο οηοþκο̣/ι̣	ι	[χ]þονο	δ	κα[λδ]ο	ο••ωρε	ι[b] þαορ[ο •••••••
SW	[				]κα[		]ο̣ρ̣ε̣ ι̣ þα̣ο[

T			βαγδ[c]ο	…					
H	μ̣α̣λ̣[ι	ι]	βα̣[γ]ο	þαο	βαγδο	οδο	ωσταδ[ο]		
LR		þ]α̣ο	βαγδ[c]ο̣	μι	σ̣αγγο̣	α[d]	α[d] þ[e]α[d]αρο	•σο̣	ογα
R	••••]α̣/λ̣ο		βαγαομ[f]ιαδγγογο			μαο	<ο>ρμιζδοογα		
SW		]αο̣	ω̣[	]μι̣[•••]γ[f][				]οογ̣α̣[	

[a] Or þ. [b] ι is missing in Rtveladze's printed text, but included in his drawing (Rtveladze 2021: 153). [c] Or α. [d] Or δ. [e] Or ρ. [f] Or ν.

L Of King Huvishka (it was) the fourth regnal year, when … city (or: country) … the king endowed(?) (or: donated?), gave as a gift(?) this *saṅgha* and(?) city …

R King [is] Huvishka. The Era-year [is] 4. When (earthquake?) … country/city … in the month Bagaom/niadgogo [on the day of] Ohrmazd

SW … the king(?) …

All previous editors claim that the inscription begins with the name of Huvishka and an indication of his regnal year. According to T, the first letters are to be read þαοοηοþκο or þαοοηοþκι, presumably a mason's error for *þαο οηþκο/ι or *þαο οοηþκο/ι 'King Huvishka', which H 'corrects' to [þα]ο̣ ο̣ο̣ηþ̣κο̣. In my opinion neither these words, nor

the supposed dating formula which follows, can be read, even in outline, and consequently there is no reason to assume that the inscription belongs to the reign of Huvishka.[7] Some traces are visible, possibly including the ι which has been interpreted as an article (before the supposed χβονο), but the first reasonably clear letters are]κα̣[, so far interpreted as belonging to κα[λδ]ο or κα[λδ]ι 'when'. A little further on R reads and restores ο••ωρε ι βαορ[ο], while LR has]o̰ρε ι β̣ao̰ρ[ο. I independently interpreted the traces as]o̰ρε̣ ι̣ β̣αο[(the ι being particularly uncertain). The restoration of a form of βαορο 'city' or 'realm' is certainly possible, though a form or derivative of βαο 'king' seems equally likely.

After a substantial lacuna, further text is partially visible, beginning with an almost certain α (either α or λ according to R) followed by ο or σ. The restoration þ]α̣o̰ 'king' in LR is possible but quite hypothetical. Then comes what T interpreted as βαγδο 'gave' (though βαγαο is offered as an alternative); R reads βαγαομ/νιαδγγογο, which he understands as an otherwise unknown name of a month (μαο). In my opinion, none of these readings corresponds to what can be seen on the stone. The only clear letter in this sequence is the sixth(?), read by R as μ or ν, which seems to me a certain μ. For R's °μ/νιαδγγο° and °μαο <ο>ρμ°, LR has μι σ̣αγγo̰, which could be understood as 'this *saṅgha*', and β̣α̣αρο, evidently taken as a variant of βαορο or βαυρο, but hardly anything is clearly visible here on the stone. At the end of the line R has <ο>ρμιζδοογα, supposedly naming a day dedicated to Ohrmazd. Of this I see only the last few letters, which may well be]οογα̣[in agreement with R.

Line 2

T			…	σταδο	þοδια ο[
H	[τα κ]αλ[δι]	φρογι[ρδ]ο	μ̣α̣λ̣ι̣ζα	σταδο	þοδιλα̣ [
LR	α̣[••]α̣α[a]•ο þ[b]ιπ[c]σ[d]ι[••••]		ειδο	σταδο	þοδια ο[τι
R	α̣••α̣α̣• κ̣ρ̣ιπν••••		ειδο	σταδο	þ[b]οδιαο [
SW	[		]ο	σταδο	þοδβ̣α ο̰[

T	…]α … γανια …		βαγαλαγγο	ζιδο …
H	] ι γανζαβ̣α̣ρ̣α̣	α̣βο	βαγολαγγο	ζιδο ατι
LR	••••]α̣••γα νι̣α̣ κ̣ιδ̣ι		βαγγαλαγγο[e]	ζιδο
R	[••••••• α]δ̣ι̣ιγα μαλια•		βαγαλαγγο	ζιδοα
SW	]ιγα̣ν[		]β̣[•••]α̣[•••]γο	ζιδο α[

[a] Or δ. [b] Or ρ. [c] Or ν. [d] Or γ. [e] Sic.

[7] Thus already Sims-Williams 2010: 111 (#354).

L … this (i.e. *saṅgha*) Shodiya founded (established) and … who put in order (or: decorated) the temple

R … this became/happened. Sh/Rodia came/arrived(?). [He had] a beautiful temple

SW … was the *shudva* … struck(?) …

At the beginning of the line, for which H and R offer partial (and incompatible) readings, I see no recognizable letters. The verb σταδο 'was', already recognized in T, is preceded by a rather clear ο. It does not seem possible to read *ωσταδο 'placed, established' as L's translation seems to suggest.

The following word is a name or title, which occurs again in lines 4 and 6. In T it is read þοδια in the first two occurrences and þοδ[ι]αα in line 6. H prefers þοδιλα. R reads þοδιαο here and þοδ[ι]αα in line 6, in both cases offering ρ as an alternative reading of the initial. In line 4, however, he reads the relevant sequence as °ρο δβαα, though he does not regard this as representing the same word. In my opinion, the fourth letter is clearly β in line 4, and more likely β than ι in the other two instances. The initial is clearly þ in line 2, though the occurrences in lines 4 and 6 would allow the reading ρ.

In Sims-Williams 2010: 157 (#557) I have suggested that þοδβα may derive from **čyuta-pā-* '(official) who looks after the expenditure', but this is of course no more than a vague possibility. If it is indeed a title, the word-order σταδο þοδβ̣α '[So-and-so](?) was (the) *shudva*', with the verb preceding its complement, may be compared with a statement in the silver plate of Nukunzik (Sims-Williams 2015: 257, line 1 of the text): αζο μο νοκονζικο … στ̣αδημι ą̣μβουκαο 'I, Nukunzik …was (the) *amboukao*'.

The declension of this word is problematic. Here, where it presumably functions as a nominative, it has the form þοδβ̣α or perhaps þοδβ̣ą̣, depending on the word-division. A similar ambiguity affects þ̣ǫ̣δβα or þ̣ǫ̣δβαą̣ in line 4, where the word seems to be the agent of a past tense verb 'dug'. In the final instance, in line 6, the form þ̣οδ̣β̣αα is clearly genitive in function: πιδο ια þ̣οδ̣β̣αα φρομαν̣α 'at (the) *shudva*'s command'. I see no plausible alternative to the supposition that the ending -αα derives from **-āyāh*, the genitive singular of the feminine *ā*-stems, here transferred to a masculine *ā*-stem.[8] In that case, one can probably expect the same form to have been used in line 4, whether the agent function was expressed by the old genitive or by an instrumental-ablative form (cf. below on μιιροζαδα in line 6).

After a lacuna the sequence]ιγą̣ν[seems fairly certain. Towards the end of the line one can read]χ̣ο ζιδο α[. Here T and R read βαγαλαγγο 'temple' (altered by H to the normal spelling βαγολαγγο), interpreting the following ζιδο as the past stem of a verb

[8] See Kreidl 2024, where σαγο 'shadow' < **sāyākā-* (cf. Peyrot 2018: 123 n. 7) and the divine name νανα < ναναια are cited as possible parallels for a phonological development of **āyā* to *ā*. The form -αα, attested only here, may represent an intermediate stage in this development.

meaning ‘to adorn’, cf. Khotanese *āysän-* ‘id.’ (Emmerick 1968: 10). This is also the interpretation implied by L’s translation, while H translates ζιδο as ‘was sent’. However, ζιδο is actually attested in Bactrian as the past stem of two different verbs, [1]ζιν- ‘to strike’ and [2]ζιν- ‘to exact, seize, take by force, etc.’(Sims-Williams 2007: 212). It is tempting to interpret the traces of the preceding letters as ασαγγο ‘stone’ (corresponding to °αλαγγο in T and R), giving with ζιδο an expression meaning ‘to break stones’, but I prefer not to include such a speculative reading in my text. At any rate, βαγολαγγο is excluded by the rather clear traces of α (where ο would be required), and the following letter can hardly be λ, which rules out even a variant such as βαγαλαγγο or βαγγαλαγγο.

Line 3

T	… φ	… ϰ … ρα	[þο]δια	κιρδο	<ο>τι	ανια	
H	[ειδο	πιδ]ογαρα	þọδịλα	κιρδο	<ο>τι	ανι <ι>α	
LR	[•••] ι	þ[••]ọịρ[a]α[•	ν]ąμα	κιρδο	[ο]τι	ανια	
R	αμ[ο]	κ[οþαν]ο[b]	þ[c]αọ	ναμα	κιρδο	<ο>τι	ανια
SW	[		]δο	τịδηια			

T	αβο	μαλιζα	ωσταδο	…		ολ/σβαρο
H	αβο	μαλιζα	ωσταδο	αṭị	ϰąλδι ι	[α]βο φρολβαρδọ
LR	ąβα[d]	μαλị̣ζα	ωσταδο	[••]•	ι[α]ζαδε̣ọ[e]	μαζι ολβαρο ị̣
R	αβο	μαλιζα	ωσταδο	••ι	ζαδε̣	•μαζιο δβαρο •••• (?)
SW	ąβọ	μą̣	[λι]ζα	ωσταδο[		]ο λβαρọ [

[a] Or þ. [b] Or þ[αοναν]ο. [c] Misprinted ρ. [d] Or ο. [e] Or ą.

L … (the sanctuary?) which the king named after (made in the name of) Kanishka(?) and besides Shodiya(?) built a big gate(?) for the gods in the acropolis

R made/built here in the name of the Kushan king (or: king of kings) and restored the other [buildings] in the fortress. (And) … by image(?) … the entrance …

SW … Then he placed (it?) in the citadel … the court …

Just after the beginning of the line, R offers two options, κ[οþαν]ο þαọ ‘Kushan king’ and þ[αοναν]ο þαọ ‘king of kings’, while L’s translation suggests a third alternative *κ[ανηþκ]ο þαọ ‘king Kanishka’.[9] In reality, nothing is clearly legible before]δο. As the last word of a clause, this is very likely the ending of a past tense form, but κιρ]δο ‘made’ as in all editions is only one among many possibilities.

After this, all previous editions assume the omission by haplography of the first letter of οτι ‘and’. Rather than <ο>τι ανια, however, one can read (without emendation) τịδηια

[9] Not in LR, but cf. Livšic 1996: 171 n. 12, where an earlier reading by Rtveladze is cited: [κ]α[ν]ηþκαα ναμα κιρδο ‘made in the name of Kanishka’.

'then he', cf. τιδι 'then' in line 4. Both τιδι and τιδηια have now been recognized in the Dasht-i Nawur inscription (DN1), in lines 7 and 10 respectively.[10] Possibly τιδι, which no doubt contains the sentence-initial particle -δι, is cognate with Khotanese *ttīttä*, *ttīyä* 'then'.[11] The reading of the following words αβ̣ο μ̣α̣ [λι]ζα ωσταδο 'placed (it?) in the citadel' is plausible, though they are not quite so well preserved as would appear from the published editions.

At the end of the line one can clearly read λβαρ̣ο, a word attested in the later documents as αλβαρο, with prothetic vowel, always in the sense 'court' (Sims-Williams 2007: 189b). It is possible that it also had the more basic meaning 'door, gate' like the underlying Old Iranian **ɗwar-*, as implied by the translations of L and R (ворота, 'gate', 'entrance'); however, the first letter is almost certainly λ, as expected in Bactrian, rather than R's δ. Of the preceding words, read as ••ι ζαδ̣ε •μαζιο by R,[12] only the final -ο is clear. The reading of LR is similar (though not entirely clear) and includes the sequences ι[α]ζαδ̣ε and μαζι, evidently understood as 'gods' and 'big' respectively, cf. Avestan *yazata-* and *maz-*, *mazaṇt-* etc. Neither word is attested elsewhere in Bactrian (if one excludes ιαζαδο transcribing Middle Persian *yazad* on Kushano-Sasanian coins and ιεζιδ°, ειζιδο 'god' as a New Persian loanword in the phrase 'in the name of God, *bismillāh*' in the very latest Bactrian texts).[13] In any case, there seems to be no justification for the readings on which these interpretations are based.

Line 4

T	…
H	[τ]αδι [ι ιαζαδε] οαᵩ̣ṭ̣ι̣νδο α̣[σ]ο [ι] μ̣[α]λ[ιζ]α
LR	αλ̣α [•••••••]α̣[••]α̣ πι̣δ̣ι λα̣χβι βᵃοα̣βα̣[•]ο̣
R	αλα ••••• α••α [ε]ιδι γαχ β[αο]ρο αβα[ν]οᵇ
SW	[ ]α̣[ ]δ̣ι λαχβṭ̣α σṭαδ̣ο

[10] See Halfmann *et al*. 2024: 24, 36, where the authors prefer a different interpretation of τιδι as equivalent to the later demonstrative τιδο. (Contrary to Sims-Williams 2007: 269b, I now think that this form is to be understood as a compound of the pronouns το and ειδο. Since it does not begin a new clause, there is no reason to think that it contains the particle -δο.)

[11] In Sims-Williams 1997: 319 n. 5 I suggested that *ttītä* 'then' was formed on the analogy of *cītä* 'when', and that the latter derives from **čiyat-*.

[12] R apparently took °μαζ° as part of a word meaning 'image', but his explanation (Rtveladze 2021: 157, 159) is incomprehensible.

[13] See Humbach 1966: 52; Sims-Williams 2007: 217b.

T	αβαβο	αϙιδι	þοδι̣α	ǫ̣δι[	]νοκανδο …
H	αβαβο	ατι <ι> δι	þοδιλα	σα̣δ̣ι	νιγανδο ατι
LR	αβαβο	οτιδι	ροαβα	α̣λιρζι̣	νοκανδο αϙ[c
R	αβαβοα	τιδιρο	δβαα	λιιμενο	κανδο αγ•
SW	αβαβσα	τιδι	þǫ̣δ̣βαα̣	δ̣ι ι̣••νο	κανδο α[

[a] Or ρ. [b] Or αβα[γ]ο. [c] Or λ̣.

L … the flowing water was dry, so Shodiya dug (in?)

R Because of … this part of the city became waterless (or: remained without water). [He had] two reservoirs (or: ponds)(?) dug …

SW … was(?) a desert, without water, then the *shudva* dug the …

Apart from a likely α (as its 2nd or 3rd letter), the beginning of the line is illegible. The first word which is fairly clear is λαχþτ̣α̣, which I take to be a fem. noun for 'desert' cognate with Sogdian *δxšt-* (feminine), Persian *dašt* (hence the feminine adjective αβαβσα, on which see below). A later form of the word is attested as λ̣αχþτο in a Buddhist text in the phrase ζ̣αδομ[ο]ρδδι[γο]λ̣αχþτο, probably meaning 'desert of birth-death (i.e. *saṃsāra*)'.[14] Here LR read λα̣χþι, R γαχ þ[. My own reading λαχþτ̣α̣ was reached independently in 2004. The first letter is certainly λ, while the final α is virtually certain in view of its distinctive 'tail'. The fifth letter is not clear, but its width makes τ more likely than ι.

Before λαχþτ̣α̣, one can discern traces which can perhaps be read as]δ̣ι. If so, τι]δ̣ι 'then' (a word found again later in this line) seems a more likely restoration than R's [ε]ιδι 'this'(?) or πι̣δ̣ι (a variant of the preposition πιδο?) as in LR.

After λαχþτ̣α̣, I read (with some hesitation) στ̣αδ̣ǫ̣ 'was'. The σ and α seem to me clear, and the width of the letter between them suggests τ, but the last two letters of the word are quite uncertain. R reads αβα[ν]ο or αβα[γ]ο and LR, more cautiously, °α̣βα̣[•]ǫ̣, but I do not see the supposed β. L's translation may imply that he understood this word as a form of αβο 'water'.

Although all editions read αβαβο, I already pointed out, after my first examination of the inscription in 1990, that the fifth letter is a clear σ, and that αβαβσα can be understood as a regular feminine form of the adjective αβαβγο 'waterless' attested at Surkh Kotal.[15]

[14] Turfan fragment h8, line 11, see Sims-Williams 2004: 331. I am grateful to Jonathan Silk and Peter Zieme for drawing my attention to a number of parallels, e.g. Uygur *sansarlıg uzun öŋ körtük* 'die endlose Wüste$_2$ des *saṃsāra*' (Wilkens 2016: 470–471), probably based on Sanskrit *saṃsāra-kāntāra-* 'wilderness of *saṃsāra*' (Kaygusuz 2021: 211).

[15] Sims-Williams 1994: 173. The form αβαβο, without the suffix *-*ka-*, does occur in the compound αβαβοδαχμο 'dried up stream' (Sims-Williams 2007: 181a).

There I also suggested the possibility that the following τιδι means ‘then’, as is now confirmed by the Dasht-i Nawur inscription (see above on τιδηια in line 3).

Regarding the next word, the name or title þοδβα, see above on line 2.[16] ‘Shudva’ or ‘the *shudva*’ is presumably the agent of the verb at the end of the line, which may be κανδο (as at Surkh Kotal) or νοκανδο (cf. later νακανδο, Sims-Williams 2007: 237b), both meaning ‘dug’. As argued above, the form þοδβαα is to be expected in agent function, in which case the sequence following it, which presumably contains a noun denoting the logical object of the verb, may be δ̣ι̣••νο (followed by κανδο) or δ̣ι̣•• (followed by νοκανδο). The initial is almost certainly δ, as it was read in T; the horizontal base stroke seems to be visible, excluding a reading λ.[17] Since the only native Bactrian words beginning with δ are derivatives of the 2nd person demonstrative **aita*- (δαγογγο, δανομανο, δαρο, δαφρηλο, see Sims-Williams 2007: 208), it seems likely that δ̣ι is a separate word, a proclitic form which would be related to ειδο, εδο ‘this (Latin *iste*)’ < **aita*- in the same way as μι, μο to ειμο, εμο ‘this (Latin *hic*)’. However, it cannot of course be excluded that δ̣ι̣••νο or δ̣ι̣•• is a loanword.

Line 5

T			…	α̣βογανδο	ριζδι
H	þοδιλα	αβο	μαλιζα	αβογανδο	ριζδι
LR	[•••]ν̣[••••	þ]ο̣[a]δ̣ια		αβο γανδο	ριζδι
R	•••ν•••••	[κ]ιδια[b]		αβογανδο	ριζδι
SW		σι]δ̣[ι	]α̣λ̣βογανδο		ριζδι

T	… οβει …				
H	οτι οβει ι	ιαζαδε	μα̣λ̣ι	αβ̣ο βα[γ]ολα̣γ[γ]ο	αβ[α]
LR	οτι οβει•	ιαζαδοε̣	μα̣[λι][c]	α̣γ[d]οδ[ο] οβ[••]δο	þ[e]α
R	οτι οβει	μαζα[ν]ι <ι>ιμα[α]		οτο ν̣α̣[ν]ο β••δο	þ[e]α••
SW	οτι οβεινα	ζαμ̣να̣		•[•]τομ̣[••]οβ[••]δο	þα[

[a] Or ι̣. [b] Or [σ]ιδια. [c] Or μα̣[λα]. [d] Or τ. [e] Or ρ.

L … Shodiya dug and both(?) deities were brought (or: came) here

R … which [he] named Abogando and both images(?) of Yima and Nana … the city(?)

SW … which is called Alv-gand. And at this time …

[16] Differently R, but his interpretation of δβαα as ‘two’ is impossible. The Bactrian form of this word is now well-attested as λοο, λοοι, λοι, see Sims-Williams 2007: 228a.

[17] LR reads α̣λιρζι, while R has λιμενο, which he translates as ‘reservoirs’ or ‘ponds’, with a surprising derivation from Greek λιμήν ‘harbour’.

The meaning of ριζδι ‘is called, is named’ was established from its occurrence in the Rabatak inscription (Sims-Williams & Cribb 1996: 94). The preceding word should therefore be the name of whatever was dug by the *shudva*, perhaps most likely a canal. Although all editors have read αβογανδο, the letter before β seems to be λ rather than α. With the preceding α, as read by almost all editors, we gain the name α̣λβογανδο, possibly a compound of *αλβο ‘way’ (Av. *aduuan-*, *aδβan-*) + κανδο ‘dug’. Before this R reads [κ]ιδι° ‘who’ or [σ]ιδι° ‘which’; the latter seems plausible, although only the δ is more or less clear.

The reading οβεινα ζαμ̣ν̣α̣ ‘at this time’ seems to me almost certain. The last three letters are only partially preserved, but the surviving traces are very distinctive. All earlier editions have equated οβει° with ουβε in the “*inscription pariétale*” of Surkh Kotal, generally understood as ‘both, also’ < OIran. *uba-*:](ν)οβιχτο μο μαβτο ουβε μο παγδο ‘wrote the μαβτο and also the παγδο’. However, as I have proposed elsewhere (Sims-Williams forthcoming), a more likely interpretation of these words is ‘wrote the μαβτο (or: αβτο) on the παγδο’, with ουβεμο < **upa imam*. Here, οβεινα may be taken as a similar combination of preposition + demonstrative, < **upa ayanā*, the ending being a pronominal instrumental singular in *-*anā*.[18] In Sims-Williams 2008: 62–63 the expression οβεινα ζαμ̣ν̣α̣ was compared with α̣βεινα οιαγο ‘in this place’ (< **abi ayanā*) in the Rabatak inscription,[19] to which one may add πιδεινο σαδο (< **pati ayanā*), variant πιδειμι σαδο (< **pati ahmi(yā)*, locative, or **pati imahya*, genitive as generalized oblique?),[20] in the great Surkh Kotal inscription, a later spelling with -ο for -α as in many other words. The variety of case-forms attested with some of these prepositions need not be regarded as problematic in view of the fact that **abi*, **upa*, **pati* and other prepositions are used with more than one case in Avestan and other Iranian languages. Moreover, the system of nominal morphology was evidently in a state of flux during the early Kushan period, so that some variation in usage is to be expected.

At the end of the line I see •[•]τομ̣/ν̣[••]οβ[••]δοþα[(almost entirely in agreement with R). A possible, but quite hypothetical, restoration of this passage might be α̣[σ]τομ̣[ο ν]οβ[αν]δο þα[‘there is with me …’, though the contrast between the writing of the prefix **ni-* here as *νο- and in line 6 (νιβιχτο) as νι- would be a little surprising.

[18] Cf. Khotanese *-äna*, Sogdian *-nʾ* (Sims-Williams 1990: 276–278). On the forms in -να and -νο discussed here see now the detailed discussion in Kreidl 2024.

[19] Slight traces of the initial α seem to be visible in some photos, but I do not exclude the possibility (suggested by Halfmann *et al*. 2024: 27) of reading ǫ̣βεινα as at Ayrtam.

[20] For a likely survival of locative **ahmi(yā)* cf. Ossetic *am*, Digoron *ami* ‘here’; for *imahya* cf. Sims-Williams forthcoming on the forms αμο and μι.

Line 6

T	…	οτι ειμο μιροζαδα	νιβιχτο πιδο ια ϸοδ[ι]αα	φρομανα
H	[σ]ọ ỿịọạṭινδο	οτι ειμο μιροζαδα	νιβιχτο πιδο ια ϸοδιλα	φρομανα
LR	α α[a]ọạ[b]•ṿδο	οτι ειμο μιροζαδο[c]	νιβιχτο πιδο ια ϸοδ[ι]αα	φρομανα
R	•••••δο	οτι ειμο μιροζαδα	νιβιχτο πιδο ια ϸ[d]οδ[ι]αα	φρομαṿα
SW	[ ]ạ[ ]ο	οτι ειμο μιροζαδα	νιβιχτο πιδο ια ϸọδ̣βαα	φρομαṿα

[a] Or δ. [b] Or δ̣. [c] Sic, probably a copying error. [d] Or ρ.

L … and this was written by Mirozado on Shodiya's orders.

R … And this was written by Miirozada (Mirzad) by the order of Sh/Rodiya.

SW … And Mihir-zad wrote this at the *shudva*'s command.

This is by far the best preserved line of the inscription, almost all of which has been correctly read and understood ever since its discovery. The final -α of μιροζαδα is now recognized as an ablative-instrumental form marking the agent of the verb (see Sims-Williams 2015: 258; Kreidl 2024). On ϸοδ̣βαα see the note to line 2.

* * *

In conclusion, what can we say about the content and significance of the Ayrtam inscription? That lines 4–5 commemorate the construction of a canal or some other structure intended to mitigate a lack of water, and that line 6 names the writer of the inscription, has always been clear and is confirmed by the present edition. It is possible that some other building works were mentioned in the poorly-preserved lines 1–3, but the reference to a Buddhist *saṅgha* (Livšic) or a "dynastic temple" (Rtveladze) cannot be confirmed, any more than the supposed mention of the gods Ardukhsh and Farr (Harmatta) or Yima and Nana (Rtveladze). Above all, the dating to the reign of Huvishka, which has been taken for granted in all previous editions, appears to be baseless. In fact, as already suggested by Kreidl (2024), the morphological and orthographic features of the text tend to suggest (though they cannot prove) a somewhat earlier date. It is to be hoped that a competent archaeologist will re-examine Rtveladze's dating of the layer in which the inscription was found to the mid to late 2nd century CE, which seems to be almost entirely based on his reading of the name of Huvishka (Rtveladze 2021: 160–167).

References

EMMERICK, Ronald E., 1968. *Saka grammatical studies*, London.

HALFMANN, Jakob, Svenja BONMANN, Simon FRIES & Natalie KOROBZOW, 2024. 'A new reading of the Bactrian part of the Dašt-i Nāwur trilingual (DN I)', *IIJ* 67/1, 5–51.

HARMATTA, János, 1986. 'The Bactrian inscription of Ayrtam', in *Studia grammatica iranica. Festschrift für Helmut Humbach* (ed. Rüdiger SCHMITT & Prods Oktor SKJÆRVØ), München, 131–146.

HARMATTA, János, 1994. 'Languages and scripts in Graeco-Bactria and the Saka kingdoms', in *History of civilizations of Central Asia* II, Paris, 397–416.

HUMBACH, Helmut, 1966. *Baktrische Sprachdenkmäler*, Teil I, Wiesbaden.

KAYGUSUZ, Arzu, 2021. *Buddhistische Bildersprache in alttürkischen Texten. Eine literaturwissenschaftliche und philologische Analyse*, Wiesbaden.

KREIDL, Julian, 2024. 'On the Kushan Bactrian ablative-instrumental case', *Iran and the Caucasus*, 28/2, 206–228.

LIVŠIC, Vladimir A., 1996. 'Baktrijskaja nadpis' na fragmente sosuda iz Kara-tepe (naxodka 1984 g.)', in *Buddijskie kompleksy Kara-tepe v Starom Termeze. Osnovnye itogi rabot 1978–1989 gg.* (ed. Boris Ja. STAVISKIJ), Moskva, 168–172.

PEYROT, Michaël, 2018. 'On the East Iranian genitive plural ending', *IIJ* 61/2, 118–130.

RTVELADZE, Èdvard V., 2008. *Civilizacii, gosudarstva, kul'tury Central'noj Azii* (2nd ed.), Tashkent.

RTVELADZE, Èdvard V., 2018. 'Ajrtam – Dinastijnyj xram carja Xuviški', *VDI* 2018/2, 365–388.

RTVELADZE, Èdvard V., 2021. *The Kushan kingdom: Dynasties, state, people, language, writing systems, beliefs*, Tashkent.

RTVELADZE, Èdvard V., A. X. SAIDOV & E. V. ABDULLAEV, 2000. *Očerki po istorii civilizacii drevnego Uzbekistana: Gosudarstvennost' i pravo*, Tashkent.

SIMS-WILLIAMS, Nicholas, 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 held in Turin, September 7th–11th, 1987 by the Societas Iranologica Europaea* I (ed. Gherardo GNOLI & Antonio PANAINO), Roma, 275–296.

SIMS-WILLIAMS, Nicholas, 1994. 'Bactrian ownership inscriptions', *BAI* 7 (1993[1994]), 173–179.

SIMS-WILLIAMS, Nicholas, 1997. 'The denominal suffix *-ant-* and the formation of the Khotanese transitive perfect', in *Sound Law and Analogy. Papers in honor of Robert S. P. Beekes on the occasion of his 60th birthday* (ed. Alexander LUBOTSKY), Amsterdam–Atlanta, 317–325.

SIMS-WILLIAMS, Nicholas, 2004. 'Two Bactrian fragments from Yar-khoto', in *Turfan revisited. The first century of research into the arts and culture of the Silk Road* (ed. Desmond DURKIN-MEISTERERNST *et al.*), Berlin, 325–332.

SIMS-WILLIAMS, Nicholas, 2007. *Bactrian documents from Northern Afghanistan II*, London.

SIMS-WILLIAMS, Nicholas, 2008. 'The Bactrian inscription of Rabatak: A new reading', *BAI* 18 (2004[2008]), 53–68.

SIMS-WILLIAMS, Nicholas, 2010. *Bactrian personal names*, Wien.

SIMS-WILLIAMS, Nicholas, 2015. 'A new Bactrian inscription from the time of Kanishka', in *Kushan Histories. Literary sources and selected papers from a Symposium at Berlin, December 5 to 7, 2013* (ed. Harry FALK), Bremen, 255–264.

SIMS-WILLIAMS, Nicholas, forthcoming. 'Names and ghost-names in the colophon of the great inscription of Surkh Kotal'.

SIMS-WILLIAMS, Nicholas & Joe CRIBB 1996. 'A new Bactrian inscription of Kanishka the Great', *SRAA* 4, 75–142.

TURGUNOV, Baxodir A., Vladimir A. LIVŠIC & Èdvard V. RTVELADZE, 1981. 'Otkrytie baktrijskoj monumental'noj nadpisi v Ajrtame', *Obščestvennye nauki v Uzbekistane* 1981/3, 38–48.

WILKENS, Jens, 2016. *Buddhistische Erzählungen aus dem alten Zentralasien. Edition der altuigurischen Daśakarmapathāvadānamālā*, Teil 1, Turnhout.

Fig. 1. Ayrtam, statue base with inscription.
© Oriental Department of the State Hermitage Museum, St. Petersburg.

Fig. 2. The Ayrtam inscription, left.
© Oriental Department of the State Hermitage Museum, St. Petersburg.

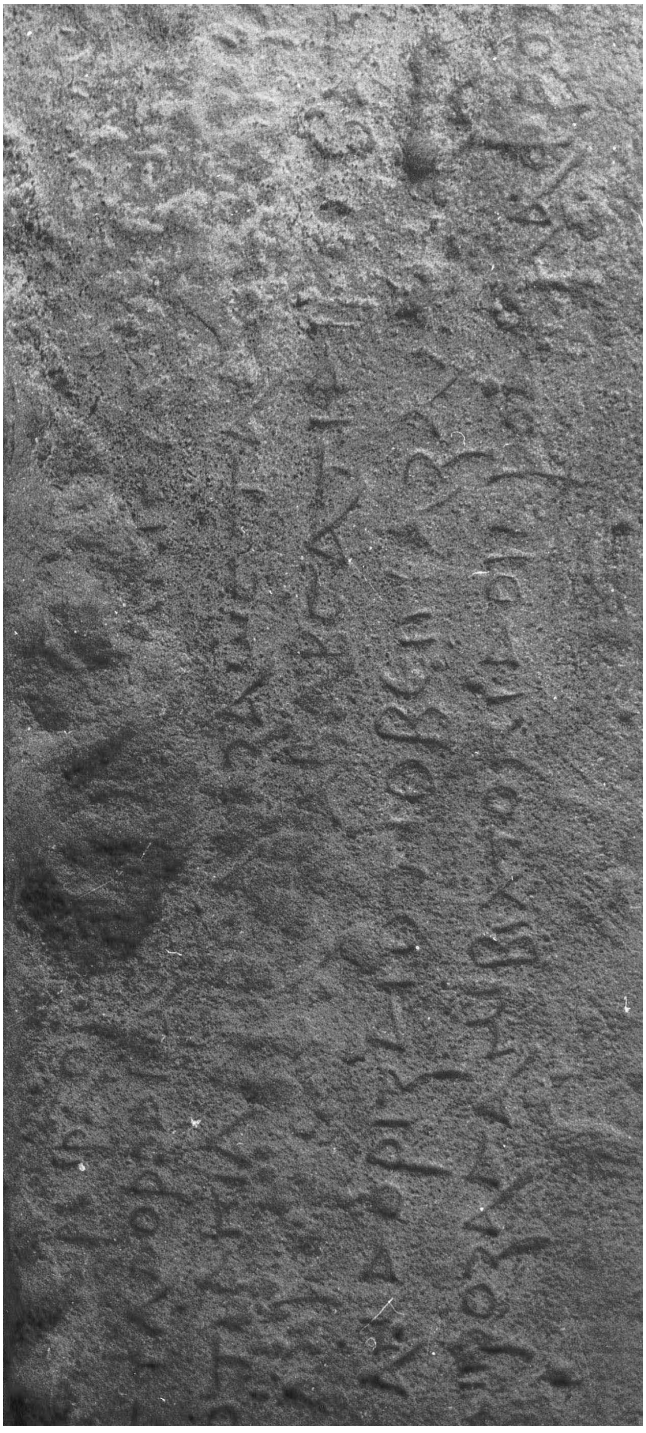


Fig. 3. The Ayrtam inscription, middle.

© Oriental Department of the State Hermitage Museum, St. Petersburg.

Fig. 4. The Ayrtam inscription, right.

© Oriental Department of the State Hermitage Museum, St. Petersburg.

List of contributors

CHING Chao-jung 慶昭蓉 • ching.chaojung.5j@kyoto-u.ac.jp
Kyoto University

Alessandro DEL TOMBA • alessandro.deltomba@uniroma1.it
Università degli Studi di Roma "La Sapienza"

Federico DRAGONI • f.dragoni@hum.leidenuniv.nl
Leiden University

Pavel B. LURJE • lurje@panchkath.ru
State Hermitage Museum

Francesca MICHETTI • francesca.michetti@unior.it
Università di Napoli "L'Orientale"

MIYAMOTO Ryoichi 宮本亮一 • miyamotor@daibutsu.nara-u.ac.jp
Nara University

OGIHARA Hirotoshi 荻原裕敏 • diyuanyumin@gmail.com
Ryukoku University

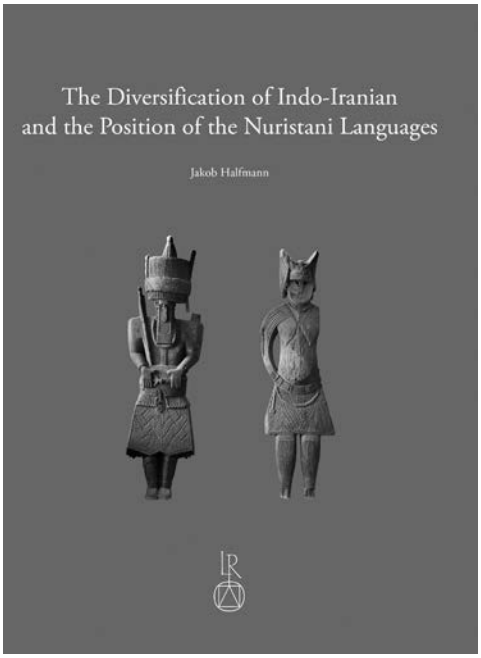
Michaël PEYROT • m.peyrot@hum.leidenuniv.nl
Leiden University

Niels SCHOUBBEN • n.schoubben@hum.leidenuniv.nl
Leiden University

Nicholas SIMS-WILLIAMS • ns5@soas.ac.uk
SOAS University of London

The Diversification of Indo-Iranian and the Position of the Nuristani Languages

Jakob Martin Halfmann
Beiträge zur Iranistik Vol. 54
2025. 8°. 160 pp., hc.
Print: 978-3-7520-0912-5
eBook: 978-3-7520-0354-3



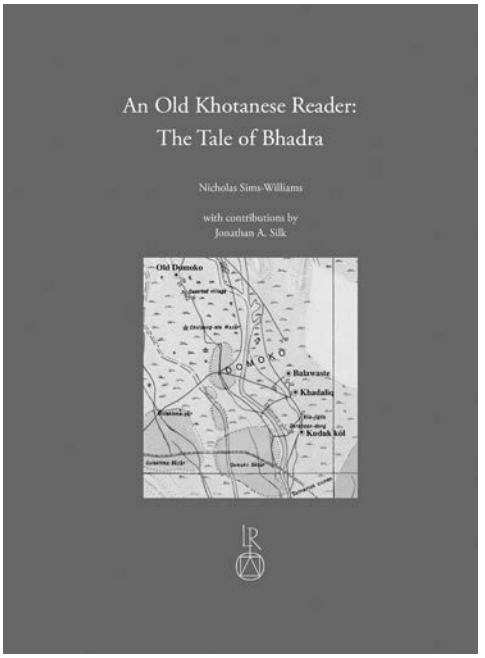
This book offers a new approach to the long-standing problem of the genealogical affiliation of the Nuristani languages, a small group of closely related languages spoken in the Eastern Hindu Kush, within the Indo-Iranian subgroup of Indo-European. This topic is approached via a step-by-step examination of the crucial isoglosses, while taking into account more sample data than was available to previous researchers. The author concludes that the Nuristani languages were likely historically more closely affiliated with the Iranian than the Indo-Aryan subgroup, though they must have been isolated from the Iranian continuum early on and subsequently have come under intense contact influence from Indo-Aryan languages.

An Old Khotanese Reader

The Tale of Bhadra

Nicholas Sims-Williams
With contributions by Jonathan A. Silk

Beiträge zur Iranistik Vol. 53
2024. 8°. 144 pp., 2 ill. (colour) hc.
Print: 978-3-7520-0917-0
eBook: 978-3-7520-0357-4



This reader contains the complete text of the Buddhist 'Tale of Bhadra', the second chapter of the Old Khotanese Book of Zambasta, accompanied by a translation, commentary and glossary. All morphological forms occurring in the text are identified in the glossary and in the introduction, which also includes a survey by Jonathan A. Silk of sources and parallels in other languages. The volume thus provides everything required to make this text accessible either to a student who has already worked through the 'Introduction to Khotanese' which forms the first part of R. E. Emmerick's Handbook of Khotanese (BzI 51, 2024) or even to a complete beginner. It also contains substantial original material, particularly in the commentary and the etymological notes in the glossary, which will be of interest to specialists in Khotanese and in Iranian and Indo-European languages in general.

Beiträge zur Iranistik

Siddham
Studies in Iranian philology in honour of Mauro Maggi
Ed. by Gerardo Barbera, Matteo De Chiara, Alessandro Del Tomba, Bhikkhunī Dhammadinnā, Federico Dragoni, Paola Orsatti
Beiträge zur Iranistik Vol. 52
2024. 8°. 480 pp., 10 b/w- and 6 ill. (colour), hc.
Print: 978-3-7520-0774-9
eBook: 978-3-7520-0312-3

A Handbook of Khotanese
By Ronald Eric Emmerick (†)
Ed. by Mauro Maggi, John S. Sheldon and Nicholas Sims-Williams
Beiträge zur Iranistik Vol. 51
2024. 8°. 360 pp., hc.
Print: 978-3-89500-444-5
eBook: 978-3-7520-0289-8

Watañi lāntaṃ
Khotanese and Tumshuqese Loanwords in Tocharian
By Federico Dragoni
Beiträge zur Iranistik Vol. 50
Vol. 50. 2023. 8°. 332 pp., hc.
Print: 978-3-7520-0690-2
eBook: 978-3-7520-0243-0

The Book of Zambasta
Metre and Stress in Old Khotanese
By Nicholas Sims-Williams
Vol. 49. 2022. 8°. 310 pp., hc.,
Print: 978-3-7520-0688-9
eBook: 978-3-7520-0248-5

Oral Narration in Iranian Cultures
Ed. by Maryam Nourzaei, Carina Jahani and Agnes Korn
Vol. 48. 2022. 8°. 198 pp., 4 b/w- and 1 ill. (color), hc.,
Print: 978-3-7520-0640-7
eBook: 978-3-7520-0153-2

Le verbe simple en pashto
État des lieux
By Matteo De Chiara and Daniel Septfonds
Vol. 47. 2022. 8°. 416 pp., cl.,
Print: 978-3-7520-0638-4
eBook: 978-3-7520-0134-1

A Grammar of Şirvan Tat
By Murad Suleymanov
Vol. 46. 2020. 8°. 396 pp., 1 b/w- ill., cl.,
Print: 978-3-7520-0011-5
eBook: 978-3-7520-0572-1

A Turquoise Coronet
Studies in Persian language and literature in honour of Paola Orsatti
Ed. by Mauro Maggi and Mohsen Ashtiany
Vol. 45. 2020. 8°. 336 pp., 4 b/w- ill., cl.,
Print: 978-3-95490-510-2
eBook: 978-3-7520-0508-0

The Lantern of Spirits
Miṣbāḥ ul-arvāḥ
By Shams ud-dîn Muhammad b. Îl-Tughân
Ed. by Bo Utas
Vol. 44. 2019. 8°. 160 pp., cl.,
Print: 978-3-95490-417-4
eBook: 978-3-95490-710-6

Le verbe pashto
By Matteo De Chiara and Daniel Septfonds
Vol. 43. 2019. 8°. 176 pp., 1 b/w- ill., cl.,
Print: 978-3-95490-375-7
eBook: 978-3-95490-708-3

Azandname
An edition and literary-critical study of the Manichaean-Sogdian Parable-Book
By Adam Benkato
Vol. 42. 2017. 8°. 216 pp., 42 b/w- ill., cl.,
Print: 978-3-95490-236-1
eBook: 978-3-7520-0537-0