



Universiteit  
Leiden  
The Netherlands

## **The Formal Kharoṣṭhī script from the Northern Tarim Basin in Northwest China may write an Iranian language**

Dragoni, F.; Schoubben, N.; Peyrot, M.

### **Citation**

Dragoni, F., Schoubben, N., & Peyrot, M. (2020). The Formal Kharoṣṭhī script from the Northern Tarim Basin in Northwest China may write an Iranian language. *Acta Orientalia Academiae Scientiarum Hungaricae*, 73(3), 335-373. doi:10.1556/062.2020.00015

Version: Publisher's Version  
License: [Creative Commons CC BY-NC 4.0 license](https://creativecommons.org/licenses/by-nc/4.0/)  
Downloaded from: <https://hdl.handle.net/1887/139192>

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

# The Formal Kharoṣṭhī script from the Northern Tarim Basin in Northwest China may write an Iranian language<sup>1</sup>

FEDERICO DRAGONI, NIELS SCHOUBBEN and MICHAËL PEYROT\*

Leiden University Centre for Linguistics, Universiteit Leiden, Postbus 9515, 2300 RA Leiden, The Netherlands

E-mail: f.dragoni@hum.leidenuniv.nl; n.schoubben@hum.leidenuniv.nl;

\*Corresponding Author: m.peyrot@hum.leidenuniv.nl

Received: February 13, 2020 • Accepted: May 25, 2020

© 2020 The Authors



---

## ABSTRACT

Building on collaborative work with Stefan Baums, Ching Chao-jung, Hannes Fellner and Georges-Jean Pinault during a workshop at Leiden University in September 2019, tentative readings are presented from a manuscript folio (T II T 48) from the Northern Tarim Basin in Northwest China written in the thus far undeciphered Formal Kharoṣṭhī script. Unlike earlier scholarly proposals, the language of this folio cannot be Tocharian, nor can it be Sanskrit or Middle Indic (Gāndhārī). Instead, it is proposed that the folio is written in an Iranian language of the Khotanese-Tumšūqese type. Several readings are proposed, but a full transcription, let alone a full translation, is not possible at this point, and the results must consequently remain provisional.

---

## KEYWORDS

Kharoṣṭhī, Formal Kharoṣṭhī, Khotanese, Tumšūqese, Iranian, Tarim Basin

<sup>1</sup> We are grateful to Stefan Baums, Chams Bernard, Ching Chao-jung, Doug Hitch, Georges-Jean Pinault and Nicholas Sims-Williams for very helpful discussions and comments on an earlier draft. We also thank the two peer-reviewers of the manuscript. One of them, Richard Salomon, did not wish to remain anonymous, and especially his observation on the possible relevance of Khotan Kharoṣṭhī has proved very useful. An earlier version of this paper was presented on 5 November 2019 at the ‘Hu-manuscripts and the ancient civil tradition’ workshop held at Peking University. In the following, CKD stands for ‘Catalog of Kharoṣṭhī Documents’ and CKI for ‘Catalog of Kharoṣṭhī Inscriptions’ (cf. <https://gandhari.org/>).

## 1 INTRODUCTION

In a posthumously published study, Klaus T. Schmidt (2018: 161–271) presents the decipherment of a number of manuscript fragments written in the so-called Formal Kharoṣṭhī script,<sup>2</sup> a late variant of Kharoṣṭhī from the Northern Tarim Basin in present-day Northwest China as a variety of Tocharian. According to him, they are written in ‘a third Tocharian language’, different from Tocharian A and B, that was originally at home in Lólán, the ancient kingdom in the southeast of the Tarim Basin, so that he terms it ‘Lolanisch’.<sup>3</sup>

Schmidt’s claim was sensational: the discovery of a possible third Tocharian language would have enormous consequences for our understanding of the history and prehistory of Tocharian as well as of the Tarim Basin. Even though his work was published posthumously and he evidently did not consider it ripe for publication during his life, it obviously deserved to be taken into account seriously in view of his earlier work. Yet it was clear that there were problems with his decipherment and therefore an evaluation was urgently needed. To this end, a workshop with the title ‘Schmidt’s Lolanisch Hypothesis’ was organised on 15–16 September 2019 at Leiden University. This workshop was attended by Stefan Baums (Munich), Ching Chao-jung (Kyoto), Hannes Fellner (Vienna), Ogihara Hirotoshi (Kyoto), Georges-Jean Pinault (Paris), and Chams Bernard, Louise Friis, Stefan Norbruis, Abel Warries (all Leiden), as well as by the three authors of this article.

It was soon agreed that Schmidt’s attempt at a decipherment had failed, and that there was no evidence that the language was either Tocharian or originally from Lólán. Since Schmidt’s readings proved to provide no useful basis for further research, it was also clear that new attempts at a decipherment would have to start with a clean slate. Based mainly on the revised initial readings of the Formal Kharoṣṭhī script by Stefan Baums and Ching Chao-jung, it was further found that the most promising perspective to identify the language of the Formal Kharoṣṭhī manuscripts was provided by Khotanese or a language related to Khotanese.

In particular, during the Leiden workshop the following readings have hesitantly been posited:

- A word-final element *-oña* alternating with *-ya*, tentatively compared with Khotanese *-auña* (we have kept this reading, see below § 4.6).
- A sentence-initial element *cu*, compared with Khotanese *cu* (we have kept this reading, see below § 4.3).
- A word *śīrya*, occurring three times, compared with Khotanese *śśāra*- ‘good’ (we doubt that this reading is correct but have so far not found anything better, see below § 4.11).
- Two times a word *mastiya* or *mastiśa*, and once a word *mastena*, compared with Khotanese *māsta*- ‘big’ or *māsti*- ‘month’ (for different options, see below § 4.11).
- A particle *dhi* (we now compare this to Sogdian *-ty* etc., see below § 4.4).
- A word *vagamXgä* (we now read *agadgä* ‘wish’, see below § 4.7).
- A word *vaṣamṭoñä* (we now read *aṣaṣoñä*, see below § 4.6).

<sup>2</sup> The designation ‘Formal Kharoṣṭhī’ goes back to Sander (1999: 72). In an earlier publication (Sander 1986: 169 fn. 10), she still considered calling it ‘Brāhmī-style Kharoṣṭhī’, because ‘the akṣaras are written more upright and square-shaped than in the normal Kharoṣṭhī’. Indeed, the writing style and the resulting appearance of this variety of Kharoṣṭhī seem due to influence from Brāhmī, but the term ‘Formal Kharoṣṭhī’ now has the widest usage and is adopted here, although we agree with Richard Salomon (p.c.) that it is not fully adequate.

<sup>3</sup> Schmidt does not mention Burrow’s hypothesis (1937) that Niya Prakrit (Gāndhāri) contains a Tocharian element and does not refer to Niya Prakrit at all, so that it remains unclear whether he assumed his ‘Lolanisch’ to be related to Burrow’s so-called ‘Tocharian C’ or not.



In this article, we propose selected readings from the best preserved Formal Kharoṣṭhī manuscript fragment known so far, T II T 48. As we will argue, in line with the provisional results of the Leiden workshop, the language is most likely Iranian, probably related to Khotanese and Tumšūqese, and possibly an early form of the latter.<sup>4</sup> We will first briefly list some problems with Schmidt's attempt at a decipherment (§ 2), then we will give an introduction to the corpus (§ 3) and present our tentative readings (§ 4), and finally we will outline why we opt for a language of the Khotanese-Tumšūqese group or an early form of Tumšūqese (§ 5).

## 2 PROBLEMS WITH SCHMIDT'S PROPOSAL

Although it has some internal consistency, there are problems with Schmidt's proposal on all possible levels. Selected points will be raised in the following.<sup>5</sup>

### 2.1 The name 'Lolanisch'

Schmidt's identification of the name of the language as 'Lolanisch' is based not on the find sites of the manuscripts or other contextual information, but only on two of his readings:

- His text number 43: *lolaṃ-kamaṃ* 'die aus Lolaṃ Gekommenen' (2018: 202);
- His text number 1, verso 3: *alōlaṃ* 'nicht aus Lolaṃ stammend' (2018: 180).

Both readings are wrong according to our current understanding of the script, but even if they were right, Schmidt's interpretation would have been impossible. For the putative *lolaṃ-kamaṃ*, it is striking that there would be no case-marking on *lolaṃ*, like the ablative, though it might theoretically have been a compound. Worse is the fact that there is no formation in Tocharian A or Tocharian B that resembles the element *kamaṃ* even remotely: the nom. pl. m. of the pret. ptc. of 'to come', for instance, is TA *kakmuṣ*, TB *kekamoṣ*. On any account, the use of the verb 'to come' (Proto-Tocharian \**kʷəm-*) is unparalleled in expressions denoting provenance. For the supposed *alōlaṃ*, the inflexion is unclear (would this formation contain a suffix or an ending?), and the use of the negative prefix, supposedly *a-*, is without good parallel within Tocharian grammar.<sup>6</sup>

The largest problem, probably, is the fact that the contexts in which these sequences occur by no means suggest this meaning. And even if '*lolaṃ*' referred to a place, it can hardly have been Lólulán, because the name for Lólulán is known from Sogdian as *krʷr'n* and from Niya Prakrit (Gāndhāri) as *kroraīna*, *krorayina*. These forms are compatible with Chin. *lólulán* 樓蘭, but not

<sup>4</sup> In a guest lecture at Heidelberg University entitled 'The Quandaries of an undeciphered script: The Formal Kharoṣṭhī corpus from Kucha' (2 November 2018), Diego Loukota Sanclemente has discussed the script and language of the Formal Kharoṣṭhī corpus. We thank him for sharing his presentation with us in February 2020. He proposes several readings of akṣaras and word boundaries without offering, as yet, interpretations of words. Our readings agree in part with his, but there are many differences at the same time. His conclusions are, amongst others, that the language appears to be Indo-European in view of its inflexional morphology, while the suffix *-oña* and the possible merging of *ā* with *i* point to 'Saka', i.e. Khotanese-Tumšūqese. We fully agree with these findings and hope to have found more evidence pointing in the same direction.

<sup>5</sup> There has been a series of online discussions about Schmidt's decipherment, see <https://linguagelog ldc.upenn.edu/nll/?p=42318>; <https://linguagelog ldc.upenn.edu/nll/?p=42724>; <https://linguagelog ldc.upenn.edu/nll/?p=42828>; <https://linguagelog ldc.upenn.edu/nll/?p=44503>.

<sup>6</sup> All these points were also raised by Georges-Jean Pinault during the workshop.



with *‘lolaṃ’*: several details of the Old Chinese form of this name are unclear, for instance the vocalism, but it is beyond doubt that Modern Mandarin *l* derives from *\*r* in both syllables (Baxter and Sagart 2014: 110). It is implausible that locals from Lólán would have the name of their own kingdom wrong.

Thus, there is no evidence for a connection of the Formal Kharoṣṭhī script with Lólán, and we will therefore avoid the misleading term ‘Lolanisch’.

## 2.2 General problems

Schmidt’s readings have to cope with several problems on a general level.






Most importantly, Schmidt presents his readings and translations without proper scientific report: there is no discussion of uncertainties, alternatives, or of the methods and insights that have led to the final result. For instance, there is no discussion of the script at all. A table of Kharoṣṭhī akṣaras is given, but it does not match the script of the Formal Kharoṣṭhī corpus.

Schmidt’s account of the script as it has to be distilled from his readings is highly questionable. His readings have some internal consistency and many of them are not obviously contradicted by what is known about Kharoṣṭhī, since many akṣaras have no match or no clear match in other varieties of this script. However, some of the akṣaras do have matches elsewhere, and many of Schmidt’s readings are clearly at odds with those identifiable akṣaras. It is also striking that Schmidt assumes a very high degree of connected writing, i.e. akṣaras that contain the final of one word and the initial of the following, a kind of akṣara-style ‘scriptio continua’. This would be highly unusual for Kharoṣṭhī, and also for Tocharian Brāhmī.

Even if Schmidt’s reading of the script is adopted for the sake of the argument, his transcription hardly yields any recognisable Tocharian elements, grammatical or lexical. Finally, the translations resulting from Schmidt’s readings are unconvincing and the content is highly unexpected and does not conform to what is known about the literary and inscriptional genres and usages of the region, as pointed out by Georges-Jean Pinault.

A last point is that bilingual fragments of Sanskrit written in Brāhmī on the one hand and Formal Kharoṣṭhī on the other do exist, and are even included in Schmidt’s corpus, but the possible evidence these bilinguals may provide is completely neglected: Schmidt makes no attempt to give readings of the Sanskrit at all.



## 2.3 Specific problems of the script

While Schmidt’s akṣara <ñā>  looks like <ya> (= *va*)  or <ha>  in other varieties of Kharoṣṭhī,<sup>7</sup> his akṣara <tya>  looks like the usual <ñā> . The reason for this


<sup>7</sup> We now transliterate this akṣara as <ha>, also in view of the *i*-diacritic that is horizontal as can be expected of <ha> (Glass 2000: 40), while it should be vertical for <ya>. Pictures of the Formal Kharoṣṭhī are taken from T II T 48 and the comparanda are from the Niya Kharoṣṭhī table in Boyer *et al.* (1920–1929).



shift in akṣaras is not given, and it is difficult to imagine what it could have been, as the yield is not impressive. Possibly, Schmidt's reason to read apparent <ha> and <ñā> as <ñā> and <tya>


is an akṣara that he reads as <ñca> , which is somewhat similar to his <ñā> .<sup>8</sup> The value <ñca> for this other akṣara gives the apparently very Tocharian-looking *poṃñc* 'all', but it

would be the only obviously Tocharian element, and the reading <ñca> is highly unlikely from the Kharoṣṭhī point of view. Strikingly, this putative Kharoṣṭhī akṣara <ñca> is very similar to the


Brāhmī akṣara <ñca> , (Malzahn 2007: 239), and one suspects that Schmidt was influenced

by the Brāhmī reading, if he did not (without making this explicit) assume influence from Brāhmī on Kharoṣṭhī in this point. Furthermore, Schmidt's unfounded akṣara shift has turned many instances of <ñā> into an unwarranted <tya>, a combination that is rare in Tocharian and needs, in his system, a special work-around, like the assumption of *scriptio continua*, anytime it occurs.

Schmidt assumes several ligatures that show no similarities to the corresponding simple

akṣaras, which is untypical for Kharoṣṭhī. An example is his akṣara <ṣmōṃ> , which shows

no obvious relationship to his simple akṣaras <ṣa> or <ma>. In fact, the unusual ligatures that are known in Kharoṣṭhī are mostly transparent, and formed by 'stacking', or rather simply hanging the

second consonant under the first, e.g. Niya <mṣo> . In Kharoṣṭhī, the only important strate-

gy to write consonant groups if not by stacking is to add a diacritic. Most importantly, there is 1) a preconsonantal *r* diacritic; 2) a postconsonantal *r* diacritic; 3) a postconsonantal *y* diacritic; and 4) a postconsonantal *v* diacritic. In Schmidt's system, the only frequent diacritic is 1), the preceding *r*, e.g. <rya>. Thus, the strategies Schmidt assumes for writing consonant groups are untypical for Kharoṣṭhī, while the strategies that are typical for Kharoṣṭhī he assumes to be hardly used.

Schmidt has one frequent akṣara in virāma: <ç>. This final occurs in Tocharian, but is not particularly frequent,<sup>9</sup> and it is strange that only this consonant would be more frequent in virāma position. More importantly, it does not look like <c>, and it is not attached in the right way: it is not a small akṣara at the lower left, but a circle at the lower right. It seems more likely that this circle is either a consonant diacritic, for instance *v*, or a vowel diacritic, for instance *u*.

Many other notes about Schmidt's decipherment can be made, but since his attempt has clearly failed, this is not the place to do so: his interpretation of the script is idiosyncratic and unwarranted, the resulting transcriptions do not yield Tocharian forms, the translations are unconvincing and the identification of the language as 'Lolanisch' is completely unfounded. Nevertheless, we must also here acknowledge that Schmidt's pioneering work has been a decisive stimulus to work on this corpus. Many documents had been known for decades, but only he has proposed a coherent interpretation.

<sup>8</sup> Another reason may have been not to have the sound *h* in his transcription, because this is missing from Tocharian. Obviously, this would have been highly circular.

<sup>9</sup> It is found, for instance, in TA all. -ac, TA/TB nom. pl. -ñc, TB 2sg. obj. -c, and TA 2pl. pres. -c.



### 3 AN OVERVIEW OF THE FORMAL KHAROŠṬHĪ CORPUS

Not all of the available material in Formal Kharoṣṭhī is taken into account by Schmidt (2018). Well aware of the fact that additions will prove to be necessary, we list here the Formal Kharoṣṭhī materials known to us.<sup>10</sup> On this point, too, we have gratefully profited from the input of Ching Chao-jung and Stefan Baums.

#### 3.1 Berlin collection

- T II T 13 / bi 34.<sup>11</sup> A very small fragment, written on recto and verso, which was brought to our attention by Ching Chao-jung. According to the expedition code from Tuyuq. Probably from the same manuscript as T II T 30 / bi 35 and T II T 48. Not included in Schmidt (2018).
- T II T 30 / bi 35.<sup>12</sup> A very small fragment, written on recto and verso, which was brought to our attention by Ching Chao-jung. According to the expedition code from Tuyuq. Probably from the same manuscript as T II T 13 / bi 34 and T II T 48. Not included in Schmidt (2018).
- T II T 48. A complete folio, written on recto and verso (Fig. 1–2). According to the expedition code from Tuyuq. The original as well as negatives of photos of it are missing, but photo prints are preserved as MIK B1928 and MIK B1940. Probably from the same manuscript as T II T 13 / bi 34 and T II T 30 / bi 35. It seems to be the Tuyuq Kharoṣṭhī manuscript that is repeatedly mentioned in von Le Coq's writings (e.g. 1909: 318f.). See also Bernhard (1970; 1976); Bailey (1973: 226); Lin (1996: 199f.); Sander (1999: 71f.) and Salomon (2007: 186). Included in Schmidt as 'Text 1' (2018: 168–181). This text will be discussed below in § 4.
- MIK B1932. Two fragments of which only one side is preserved as a photo. The expedition code of the left fragment is T 36 (or possibly T 35), probably T II T 36; that of the right fragment is not known. Both preserve the right margin, but without leaf number. Included in Schmidt as 'Text 2' (2018: 181–183; the left fragment) and 'Text 3' (2018: 183–186; the right fragment).
- T III Š 88 / bi 36.<sup>13</sup> A larger fragment written on recto and verso, which was brought to our attention by Ching Chao-jung. According to the expedition code from Šorcuq. Many akṣaras have faded considerably. Not included in Schmidt (2018).

The Berlin collection also contains some mural inscriptions in a variety of Formal Kharoṣṭhī. It is not clear whether the language of these inscriptions is the same as that of the manuscripts. They are currently being worked on by Ching Chao-jung and Ogihara Hirotooshi.

- Qizil cave 211. A mural inscription in Kharoṣṭhī script from cave 211 in Qizil (Dreyer *apud* Schmidt 2018: 196<sup>14</sup>). Also known as MIK B1879 (Sander 1999: 73).

<sup>10</sup> For a research history up to 1984, cf. Hitch (1984: 197f.), who also points to the importance of the Kharoṣṭhī materials in the Northern Tarim Basin for Tumšūqese.

<sup>11</sup> Pictures are available online at [http://turfan.bbaw.de/dta/bi/images/bi34\\_seite1.jpg](http://turfan.bbaw.de/dta/bi/images/bi34_seite1.jpg) and [http://turfan.bbaw.de/dta/bi/images/bi34\\_seite2.jpg](http://turfan.bbaw.de/dta/bi/images/bi34_seite2.jpg).

<sup>12</sup> For pictures, see [http://turfan.bbaw.de/dta/bi/images/bi35\\_seite1.jpg](http://turfan.bbaw.de/dta/bi/images/bi35_seite1.jpg) and [http://turfan.bbaw.de/dta/bi/images/bi35\\_seite2.jpg](http://turfan.bbaw.de/dta/bi/images/bi35_seite2.jpg). The picture of the verso side should be turned around 180°.

<sup>13</sup> For the recto, see [http://turfan.bbaw.de/dta/bi/images/bi36\\_seite1.jpg](http://turfan.bbaw.de/dta/bi/images/bi36_seite1.jpg) (which should be turned around 180°) and [http://turfan.bbaw.de/dta/bi/images/bi36\\_seite2.jpg](http://turfan.bbaw.de/dta/bi/images/bi36_seite2.jpg), for the verso.

<sup>14</sup> Dreyer gives the cave number as '213', but Ching Chao-jung has shown in her papers at the Leiden and Peking workshops that the actual location is cave number 211.



- G-Su 43. A mural inscription from Subashi. On the same wall, several inscriptions in Tocharian B are found. The inscription is presently located in the *Museum für Asiatische Kunst* (Sander 1999: 70 fn. 21; Schmidt 2018: 200). See von Le Coq (1928: Tafel 17), Pinault (1987: 157f.; planches LVI-1; LVII) and CKI 171.
- G-Su 44. A second mural inscription from Subashi, with Tocharian inscriptions surrounding (G-Su 12–14). The state of the original, which is at present also in the *Museum für Asiatische Kunst*, has been badly deteriorated since its discovery (Dreyer *apud* Schmidt 2018: 202 fn. 176). See Pinault (1987: 158; planches LXIV–LXV), who gives no transcription.

### 3.2 London collection

According to Bailey (1973: 226), the Formal Kharoṣṭhī materials from the British Library were found ‘in a box of fragments from Kuci which has lain in the India Office Library since about 1910’ and belonged to the Hoernle collection. He published photographs of the fragments that are now classified as IOL Khot 203/2, IOL Khot 203/3 and Or.15002/43 (1973: Plates 3–4). These fragments belong to the subcollections H(oernle) 152 and 156, the first of which arrived in London in August 1908 and the second in June 1911. Or.15009/44, which was unknown to Bailey, belongs to H(oernle) 149, also from Kuča and in London since June 1907 (Ursula Sims-Williams 2009: 2). Ursula Sims-Williams (2009: 3) further confirms that these documents were found in packets 1–6, which came from the region of Kuča, i.e. from Qumtura, Qizil and Yak-Arik (see fn. 13). In Emmerick’s opinion (as quoted in Sander 1986: 169 and Lín 1996: 198), it is most probable that the fragments were specifically found in Qizil, but this seems difficult to confirm.

- IOL Khot 203/2. A small bilingual fragment (5.9 × 4.2 cm) with Sanskrit Brāhmī and Formal Kharoṣṭhī. The Sanskrit appears to be still unpublished. Probably from the same manuscript as IOL Khot 203/3. Earlier classified as Hoernle 152: Kh C5. Included in Schmidt as ‘Text 5’ (2018: 188–191).
- IOL Khot 203/3. A small bilingual fragment (4.2 × 5.3 cm) with both Sanskrit Brāhmī and Formal Kharoṣṭhī. The Sanskrit appears to be still unpublished. Probably from the same manuscript as IOL Khot 203/2. Earlier classified as Hoernle 152: Kh C4. Included in Schmidt as ‘Text 4’ (2018: 186–188; the IOL number is indicated wrongly).
- Hoernle 156: Tb 1 and Hoernle 152: Kh C3. These fragments that can be joined (Schmidt 2018: 192)<sup>15</sup> are missing and have no modern signature. Included in Schmidt as ‘Text 6’ (2018: 192–196).
- Or.15002/43. A bilingual fragment (5.9 × 9 cm), with Sanskrit Brāhmī and Formal Kharoṣṭhī. The Sanskrit is still unpublished. Probably from the same manuscript as Or.15009/44 (Stefan Baums p.c.). Included in Schmidt as ‘Text 56’ (2018: 207).
- Or.15009/44. A bilingual fragment, brought to our attention by Stefan Baums. The Sanskrit text has been identified by Wille (2006: 31 fn. 12) as covering parts from verse 23 and 24 of the *Anaparāddhastotra* by Mātṛceṭa. The Sanskrit has been published in Ye (2009: 120f.). Probably from the same manuscript as Or.15002/43 (Stefan Baums p.c.). Not included in Schmidt (2018).

<sup>15</sup> In Schmidt (2018: 192), the image on the lower left should be put on top of the fragment on the upper left. The fragments on the right can be joined as they are.



### 3.3 Paris collection

Bailey (1973: 226) considers pictures in Filliozat (1958) and Pauly (1967) to contain the same script as the documents he presented. In reality, the fragment in Filliozat (1958: Planche VII) is an example of Kuča Kharoṣṭhī on a wooden tablet, similar to the ones published by Schmidt (2001) and Ching (2013). The Kharoṣṭhī documents which Pauly (1967: 274; 283; Planche IV) refers to were found in Bāmiyān and show the typical late Kharoṣṭhī from the third-fourth century as is also found in the more recent finds, for instance in the Schøyen collection. They are not the type of Formal Kharoṣṭhī that interests us here (cf. also Lín 1996: 196f.). Nevertheless, it seems possible that one or more pieces of Formal Kharoṣṭhī still lie hidden in the Pelliot collection: Hambis *et al.* (1961–64: 111–113) refer to manuscript P413 that was found in Toqquz Saray (Toqqouz-sarai) near Tumšūq according to Pelliot's diary (see also Pinault 2007: 171). Pelliot described the script of this document as a cursive variety of Brāhmī, vaguely resembling Tibetan. However, Hambis *et al.* relate this entry from Pelliot to 'un fragment long de 20 à 21 cm et large d'environ 15 cm écrit en une écriture assez insolite de type *kharoṣṭhī*' (1961–64: 112).<sup>16</sup> Ching Chao-jung has probably identified this fragment in the Bibliothèque nationale de France, where she sent a preliminary note of her research results on 13 June 2014. Since then, she has been doing research on it in collaboration with Stefan Baums and Ingo Strauch.

### 3.4 Dating and origin

With the scanty corpus, partly missing or lost, and difficult to access, it is difficult to get a clear picture of the dating and origin of the Formal Kharoṣṭhī corpus.

For the literary fragments, the find places Tuyuq in Turfan region<sup>17</sup> and Šorčuq in Yānqí region are secured by the German expedition codes, and, judging from the reports about the British Library collection, such fragments were also found in Kuča region, and more specifically in Qizil. It is at this point not possible to say whether the inscriptions from Kuča region are written in the same language; the script, at least, is not identical.

The literary fragments may be from the fifth or perhaps the sixth century. This is suggested by the 'Turkestan Gupta Type' of the Brāhmī in the bilingual fragments, which points to the fifth century (Sander 1986: 169f.; 1999: 72; Salomon 2007: 186; Ye 2009: 120). The Kharoṣṭhī is of a late type, and would appear to be more developed than the documents in Kuča Prakrit (Schmidt 2001), which are written in a related form of Kharoṣṭhī, and have been dated most recently by Ching (2013: 83–87) as belonging to the fifth or sixth century. This would also square with the fact that manuscripts on paper became more common from the fifth century onwards (Sander 1968: 29). Compared to Sanskrit and Tocharian Brāhmī manuscripts, the relatively large space between the lines, the size of the margins and the relatively thick strokes may point to a later period, for instance the sixth rather than the fifth century, but this is certainly not hard evidence.

<sup>16</sup> Cf. also Lín (1996: 192f.) about this, who mentions that more Kharoṣṭhī fragments have been found in Tumšūq by Chinese archaeologists.

<sup>17</sup> According to Salomon (2007: 186), 'Another similar manuscript in this type of Kharoṣṭhī was reportedly found at Murtuq, also near Turfan, but this too has not been published.'



Our impression is that T II T 48 shows a slightly more developed ductus than T III Š 88 / bi 36, where some akṣaras may be a little bit closer to other varieties of Kharoṣṭhī.

## 4 TENTATIVE READINGS FROM T II T 48

The most important Formal Kharoṣṭhī fragments are probably the bilinguals Or.15002/43 and Or.15009/44; the large fragment T III Š 88 / bi 36; and T II T 48. It has turned out to be difficult to map the Formal Kharoṣṭhī of the bilinguals on the Sanskrit text and we have not yet succeeded in making any progress worthy to report. T III Š 88 / bi 36 is clearly a key fragment, but in our view too abraded to be put to good use at this stage. However, T II T 48, the folio labeled as ‘Text 1’ in Schmidt (2018: 168–181), contains the longest continuous so far known sequence of text in Formal Kharoṣṭhī, which makes it an obvious starting point for a first tentative reading. An additional advantage of this folio is that it contains several strings of repetitions. As we will try to show, this allows to tentatively identify word boundaries and to obtain a first impression of the structure of the text.

We must highlight here once more that although our interpretations of individual akṣaras have changed in some cases, we build on the reading sessions during the Leiden workshop in September 2019, which crucially depended on the expertise in the Kharoṣṭhī script of Stefan Baums and Ching Chao-jung.

### 4.1 The external appearance of the folio

T II T 48 is a folio in poṭhī format with text in Formal Kharoṣṭhī on both sides (Figs. 1–2). One side is filled completely with calligraphic writing, and on the other side, the beginning of the first line is calligraphic while the rest is more cursive and the lower half of the folio is left blank. Obviously, the side completely filled with calligraphic writing is the recto, with the text continuing on the verso in calligraphic writing. Probably, this cursive part on the verso was written with a softer brush and, perhaps, by another scribe. As has been suggested to us by Stefan Baums, Ching Chao-jung and Georges-Jean Pinault (p.c.), the cursive part seems to be a list, for instance of personal names. Since in this cursive part we could so far not identify with any certainty elements of names or clues for the reading of akṣaras, we will not discuss this part further in this article.

In contrast to poṭhī folios written in Brāhmī where the string hole is on the left (Sander 1968: 26), T II T 48 has the string hole space on the right, which must be due to the fact that Kharoṣṭhī is written from right to left. The original manuscript has been lost (cf. *supra*), but judging from the photo and comparable fragments on paper, it must have been written on paper, the most frequent writing material for literary texts in the ancient Tarim Basin from the fifth century onwards (Sander 1968: 29).<sup>18</sup>

<sup>18</sup> While Bernhard (1970: 57) first did not express his opinion about the writing material, he later (1976) said that it is written on paper.



The folio was once part of a larger, although relatively small manuscript: in the right margin of the recto (Fig. 3), a Kharoṣṭhī number 16 is found,<sup>19</sup> and the cursive text as well as the large blank space on the verso show that it was almost certainly the last folio of the manuscript. As a consequence, the beginning of the text on the recto may well be in the middle of an original sentence; it is even possible that the recto begins in the middle of a word (although cf. *infra*). Conceivably, the leaf number is written on the recto rather than the verso, as usual, because it is the last leaf of the manuscript. The leaf number is obviously in the right rather than the left margin, as usual for Brāhmī poṭhīs,<sup>20</sup> due to the writing direction from right to left (Sander 1999: 71).<sup>21</sup>

In what follows, we give a tentative transcription of some possibly identifiable words, accompanied by a more detailed philological and linguistic discussion. We will not be able to comment upon the whole text, because many akṣaras still cannot be read with any degree of certainty and differ too much from other varieties of Kharoṣṭhī, as described for instance in Boyer *et al.* (1920–1929) or Glass (2000).

Although many akṣaras are difficult to identify or, at this point, fully unclear, some can be read with relative confidence. One of these is the consonant <p>, characterised by its long vertical stroke combined with a shorter horizontal stroke (cf. Glass 2000: 82–84). Similarly, vowels are mostly clearly identifiable. The vowel diacritic <i>, a loose vertical line to the left of the akṣara, can be securely identified, and is the same as in Kuča Kharoṣṭhī (cf. the tables in Schmidt 2001: 28f.). The same is true of <e>, marked by a vertical line on top of the akṣara. In principle, the vowel <u> is represented with a closed loop at the bottom of the akṣara and <o> with an oblique stroke going down to the left at the lower part of the akṣara, but the distinction is not fully clear in all cases.

In other varieties of Kharoṣṭhī, anusvāra can either be marked by ‘a hook open to the left attached to the base of the letter’ or ‘a separate stroke floating beneath the radical’ (Glass 2000: 135). The latter form of anusvāra is clearly distinct from the so-called long vowel mark, especially used to write Sanskrit in Kharoṣṭhī (e.g. CKD 523; Boyer, Rapson and Senart 1918), which ‘consists of a short oblique stroke drawn down to the right’ (Glass 2000: 137). Despite Schmidt’s (2001: 11) statement to the contrary, these two vowel diacritics are still distinct in Kuča Kharoṣṭhī, as discovered by Ching (2013: 63 fn. 12). However, at the present stage of our knowledge, we are not sure yet if and how they are distinguished in Formal Kharoṣṭhī, and in our transcription we mark a floating separate stroke with a breve below the vowel, i.e. <ā̆>, and a curl in the stem with a dot below the vowel, i.e. <ā̇>. Importantly, some akṣaras have both diacritics, which seems to suggest that <ā̆> rather indicates vowel length, but this needs further research. A table summarising our readings is given as an appendix, where we also give the corresponding characters from CKD 661 for reference. This Gāndhārī document from Khotan not only deviates in its language, but also shows several peculiarities of the script that seem to have parallels in the Formal Kharoṣṭhī script.

In Late Kharoṣṭhī, it is mostly impossible to tell whether there is a distinction between <śa> and <ya> (e.g. Glass 2000: 94; 99). In Khotan Kharoṣṭhī, however, the two are clearly distinct (cf. the table at the end), as the right downward-going stroke of <ya> is very small when compared

<sup>19</sup> The number is written downward. In his review, Richard Salomon kindly points out that the number 4 is written as a horizontal cross, as in CKD 661, whereas it is normally ‘an X-shaped cross’.


<sup>20</sup> Sander (1968: 158) was aware of only two exceptions among the Sanskrit manuscripts from Xinjiāng.

<sup>21</sup> She also argues that the ductus of the script suggests that the text has been written from right to left.



to other (Central Asian) varieties. Still, we are not sure how to distinguish them exactly in our document and for now, we give <śa> as the first option and <ya> as the second option.

## 4.2 The paragraph sign ‘¶’

The key for a better understanding of T II T 48 is the identification of the sign  as a punctu-

ation mark, a suggestion we owe to Stefan Baums and Ching Chao-jung, who independently of each other arrived at the same conclusion, and whose arguments are summarised below. In the transcription, we represent this sign with a pilcrow, ‘¶’, and we will call it ‘the paragraph sign.’

The paragraph sign was read as <ssū> by Schmidt (2018 *passim*), which is not supported by evidence from other varieties of Kharoṣṭhī: there is no resemblance with the normal akṣara for <s><sup>22</sup> or the vowel mark <u>, which is normally a small ‘leftward stroke at the base of the character’ or ‘a closed loop at the base of the stem’ (Glass 2000: 41).

Indeed, the sign does not look like a Kharoṣṭhī akṣara at all and seems to have a different structure, which suggests that it may be a punctuation mark. This is corroborated by the following observations. First, it is used as the last akṣara of the calligraphic writing on the verso and thus marks the transition from the calligraphic to the more cursively written part. Second, the same sign is used in Or.15002/43 (Schmidt’s Text 56; 2018: 207–208) and Or.15009/44 (cf. *supra*) to mark the transition between the Sanskrit Brāhmī text and the Kharoṣṭhī part. The identification of the paragraph sign ¶ allows to provisionally establish the internal structure of T II T 48, because it is used three times on the recto side, probably indicating a division in text units.

## 4.3 Sentence-initial <cu> and <ca>

After the paragraph sign ¶ (that is, to its left), we read twice an akṣara <cu> and once an akṣara <ca>, both of which are reasonably secure in view of their similarity to the basic Kharoṣṭhī sign for <c> (Glass 2000: 61f.). The same akṣara <cu> is found two more times on the recto, once in line 1 and once in line 2. It is unclear whether <ca> is a variant of the element <cu>; if so, it could perhaps be a contraction of <cu> with a following element, i.e. the *-u* could be lost before a following vowel.

Since this <cu> (and once <ca>), follows the paragraph sign, it must either be the beginning of a word or a word of its own. We opt for the latter, as we are reminded of the Khotanese subordinating conjunction *cu*, whose meanings range from ‘which, what?’ and ‘that’ to ‘since, because’ etc. (Bailey 1979: 104). The same conjunction is found in Tumšūqese (Emmerick 2009: 405). By contrast, *cu* is not a usual beginning of a sentence in either Sanskrit, Middle Indic, or Tocharian A or B.

<sup>22</sup> A related problem in Schmidt’s account (2018) is the consistent reading <z> of a Formal Kharoṣṭhī akṣara that rather looks like the <s> of other Kharoṣṭhī varieties, without commenting on why he does so.



#### 4.4 Second-position clitic <dhi>

In two out of four probable instances of sentence-initial <cu> (not necessarily after a paragraph sign), it is followed by an akṣara that can with relative confidence be transliterated as <dhi>.<sup>23</sup> We assume that this <dhi> is a second-position clitic, a so-called Wackernagel particle, since it seems to be a word of its own, being followed by <aḡadga> in r1 and <paṇāḥā> in r2. The former is certainly a word of its own, see below, and the latter may be a word too.

The tentatively identified particle <dhi> may be compared with the Sogdian particle *-ty* (Gharib 1995: 394), very frequent in second position, and so marking the beginning of a clause, and with Bactrian *-ḍi* (older), *-ḍo*, *-iḍo* (more recent; Sims-Williams 2007: 209), both from Plr. *\*uti*.<sup>24</sup> In Khotanese, this particle is not attested in this form (but see Sims-Williams *forthc.* and § 4.5 below), but an interpretation of <dhi> as Tumšūqese seems possible. After intervocalic lenition (Emmerick 1989: 214), one expects an outcome *\*uḍi*, where the initial /u/ could have coalesced with the preceding <cu> (for other options, cf. *infra*). As pointed out to us by Nicholas Sims-Williams (p.c.), the particle seems to be optional, like in Khotanese, and unlike Sogdian and Bactrian.

This reading would suggest that the <dh> is used for a voiced dental fricative  $\delta$ . This has a parallel in the already mentioned CKD 661, in which  $\delta$  (written <dh>) for Niya Prakrit *d* is due to Iranian influence (Konow 1936: 239). However, in <khvāḡadhi>, possibly /xvarandi/ (see below, § 4.9), <dh> appears to denote a voiced dental stop *d* because it is found after a nasal.

The akṣara <dhi> occurs more often. In some instances, it probably is the 3sg. verbal ending, and in one instance it seems to be the 3pl. verbal ending (see below § 4.8, 4.9 and 4.11).

#### 4.5 Possible sequence <cudhi>

As an alternative to reading sentence-initial <cu> and second-position <dhi>, it is possible to read <cudhi> as one sequence. Sims-Williams (*forthc.*) identifies this sequence in Khotanese *cū* from *cu* and *u*, in Sogdian *cw 'ty*, *'cw ZY* 'what, which, why, whichever, whatever, whether' (Gharib 1995: 130), and in Bactrian (α)σιḍο, σιδι 'what, whatever, which, who, (so) that' (cf. also Sims-Williams 2007: 194f.). On the basis of these forms, an East Iranian *\*cim uti* can probably be reconstructed. For Khotanese, it has to be assumed that the intervocalic *\*t* in *\*uti* has been lost by a special reduction with subsequent contraction of /u/ and /i/, thus yielding *u* (Sims-Williams *forthc.*). Because of the loss of *\*t* in Khotanese, <cudhi> cannot be Khotanese, unless the Formal Kharoṣṭhī language is much more archaic than attested Khotanese. However, in Tumšūqese *\*uti* may have yielded *uḍi* regularly, without the special reduction observed in Khotanese. The initial *u* may either be contracted with the final *-u* of *cu*, or it may have been lost by another type of special reduction, parallel to the developments in Sogdian and Bactrian.

<sup>23</sup> We assume that the difference between <c> and <dh> can be seen from the upper part of the akṣara. The akṣara we transliterate as <c> has a longer horizontal top with a slight dent down, and a knob at the upper left, whereas <dh> has a shorter, straight horizontal top. This is parallel to the distinctions between these akṣaras in other varieties of Kharoṣṭhī (see e.g. Glass 2000: 61f.; 80f. and Fig. 4).

<sup>24</sup> <dhi> was identified as a particle during the Leiden workshop, but not yet etymologised as Iranian *\*uti*.



#### 4.6 Abstract suffix -oña in <aṣ, aṣoña>?

The very beginning of the folio contains a sequence that we provisionally read as <aṣ, aṣoña> (Fig. 5). The same apparent word is found as <aṣ, aṣoña> two other times on the recto (line 2 and line 4). We tentatively transcribe a variant on the verso (line 1) as aṣ, aṣya. This strongly suggests that there is a word boundary after aṣ, aṣoña and that aṣ, aṣya is formed from the same stem. Accordingly, the -oña part is either a derivational or an inflexional suffix. We opt for the former, and compare Khotanese -auña, which derives abstracts from nouns and adjectives (Degener 1989: 158–165).<sup>25</sup>

In taking the first akṣara to be <a>, we follow Schmidt's (2018: 169 *et passim*) reading. We read <a> and not <va>, which would be a possible alternative,<sup>26</sup> because there is a difference between two akṣaras that look superficially the same and could either be <a> or <va> (Fig. 6). We assume that <a> is broader on top and goes slightly down at the upper left, while <va> is shorter on top and does not go down. For the second akṣara our first reading is <ṣ, a>, with the curl in the stem possibly denoting anusvāra, i.e. <ṣ, am>. We mark this akṣara with an index '2' because it has a dent in the top stroke and is different from an akṣara, in our text in <khvaṣa> r3 and <khvaṣo> r4, that is closer to <ṣa> of other Kharoṣṭhī varieties. An alternative reading of this akṣara could be <ka>, but this makes it especially difficult to interpret <ṣ, vidu>, presumably 'milk', which would become <kvidu> (or, less likely, <kridu>). The third akṣara has no clear equivalent in most Kharoṣṭhī varieties, but does have one in CKD 661, transliterated as <ṣa> by Boyer *et al.* (1920–1929: 249). In fact, this akṣara has a horizontal stroke added below, hence our transliteration <ṣ>. In CKD 661, the variant without stroke is found in aṣṭi '8' (end of line 3), before ṭ, and the variant with stroke is found intervocalically and may denote ṛ.<sup>27</sup> <ṣ> is also found in CKI 48, 443, and 564, also in intervocalic position. The lower part of the akṣara may contain an extra tail of which the function is unknown, with an *o*-diacritic next to it; less likely is it that the two strokes denote a vowel such as <au>. The last akṣara can fairly certainly be read as <ñā> in r1 and as <ñā> in r2 and r4 (compare Glass 2000: 67ff.).

Apart from the observation that it probably contains the abstract suffix -oña-, we have at present no interpretation to offer of aṣ, aṣoña.<sup>28</sup> If <k> rather than <ṣ> should be the right reading, the word might be compared with the Old Khotanese adjective *akāṣāñā*- 'unthinkable, unimaginable'. The adjective in its Old Khotanese form occurs once in the instr.-abl. sg. (*akāṣāñāna*, Skjærvø's emendation for manuscript *akāṣāñ[ ]*) in *Suv(arṇabhāsottamasūtra)* 6.3.24 (manuscript Or.). This has the form of a participium necessitatis from the verb *kāt*- (Emmerick 1968: 22–23), to which a negative *a*-prefix has been added. Degener (1989: 56) notes that it appears instead of the more frequent *akāṣṭa*- (formed on the past part.), probably due to the preceding *paphāñāna* 'to be made happy'. Since it occurs again in the later contracted form *akāñā*- in the manuscript Q<sup>29</sup> (Suv 11.27,

<sup>25</sup> This suffix has also been noted by Diego Loukota Sanclemente in his lecture (see fn. 4).

<sup>26</sup> <va> was the initial provisional reading together with Stefan Baums and Ching Chao-jung.

<sup>27</sup> The same akṣara may be attested in bi 36 b4 and b5, in the right part of the lines.

<sup>28</sup> In an earlier draft, we had read this word as <aṣamṭ'oña>, and interpreted it as 'Arhatship', an abstract in -oña from a base related to Khotanese *āṣana*- 'worthy' (Bailey 1979: 26), the equivalent of Sanskrit *arhat*- 'worthy, Buddhist saint'. The problem was the reading of the third akṣara, which we now read as <ṣo>. With this new reading, a connection with *āṣana*- is no longer possible.

<sup>29</sup> Middle Khotanese, according to P.O. Skjærvø's classification.



*akā'ña bvāmatīja śṣadā* 'Unthinkable (is my) faith in knowledge'), the adjective is not a hapax (as Degener 1989: 56 has it). In both cases it translates Skt. *acintya*-.

The final *-ya* of *aṣaṣya* on the verso could perhaps be the locative ending, or, following a suggestion by Sims-Williams (p.c.), it could be a feminine abstract in *-ya*. In the latter case, there is a possibility that the preceding *śa* is a demonstrative (cf. § 4.11): *śa aṣaṣya* 'the/this *a*.-ness'.

#### 4.7 <ḡadga> 'wish'?

The next recurring sequence for which we propose a reading is <ḡadga> (r1) and the variant <ḡadgā> (r2–3 and v1; Fig. 7). Our reading of this sequence is in partial agreement with Schmidt (2018: 169), who read it as *agāldkañ* (his <ñ>, which is certainly wrong, is in fact the beginning of the next word). Again, the repetition strongly suggests the presence of a word boundary before and after this sequence.

Apart from a small knob at the lower right, which is here transliterated with <a> but may be irrelevant, the first akṣara is clearly the same as the first of <aṣaṣoñā> (see above). The second akṣara is fairly certainly <g>, a voiced velar fricative *ɣ*, well known from other varieties of Kharoṣṭhī (Glass 2000: 58),<sup>30</sup> especially from that of Niya (Boyer *et al.* 1920–1929: 404 nr. 24). As discussed above (§ 4.1), it is at present difficult to distinguish between anusvāra and the long vowel mark. Based on the etymology that we propose below, we assume that <a>, the 'floating line below the akṣara', may here indicate a long vowel. The third akṣara is a ligature, probably with <ḡ> as the first element, and certainly with <g> as the second.

Our reading <ḡadgā> reminds one of Tocharian B *akālk* 'wish' (Adams 2013: 2), borrowed from Bactrian *αγαλγο* /aɣalg/ <PIr. \**āgādaka*- (Schwartz 1974; Sims-Williams 2007: 187).<sup>31</sup> If this interpretation is right and the language of our texts is Tumšūqese or related to it, then we must be dealing with a loanword. In Tumšūqese as well as in Khotanese, the \**k* in PIr. \**āgādaka*- (Sims-Williams 2007: 187) would not have been preserved as *g*, but would have been fully lost, yielding an *-aa* stem. The assumption of a loanword would at the same time explain the spelling <ḡ> for expected *l* (as in Bactrian) or *d*, *ḍ* (as in Sogdian):<sup>32</sup> presumably, *l* did not occur in this position, and had to be replaced with a similar sound. Whereas Niya Prakrit <ḡ> seems to stand for some rhotic sound and Gāndhāri intervocalic retroflexes are borrowed with <r> in both Tocharian and Sogdian, they are borrowed with <l> in Khotanese (Burrow 1937: 7), so that <ḡ> seems to be at least a possible way to write or represent *l*.<sup>33</sup>

<sup>30</sup> In the older literature (and still in Glass 2000), it was transcribed as <ḡ>. It is sometimes difficult to distinguish from <gr>.

<sup>31</sup> As noted by Nicholas Sims-Williams (p.c.), Bactrian *αγαλγο* may, on the basis of its etymology, be expected to have had originally a long vowel in the first and perhaps also in the second syllable: /āɣalg/ or /āyālg/. In view of the uncertainties with the readings of the Formal Kharoṣṭhī, we think it is too early to decide whether the initial *a*- of <ḡadgā> denotes a long or a short vowel, and, in the latter case, whether it was not written long because of a script convention, or because the vowel was effectively short.

<sup>32</sup> A borrowing from Tocharian B *akālk* would need the assumption of an adaptation of the first *k* to *g* = *ɣ* and of the second to *g*. It would therefore seem more likely that it was borrowed from Bactrian *αγαλγο* /aɣalg/ directly.

<sup>33</sup> Note, furthermore, such cases as the spelling <Keraḍa> in CKI 2 (Rock Edict 2 from Shāhbāzgarhi) for /Kerala/ (which is spelled as such in the parallel CKI 16 (from Mansehra).



With due caution, one might hypothesise that *agadgä* is the nom. sg., as it would be expected for Khotanese. According to this Khotanese pattern, *agadga* could not be the acc. sg., which ends in *-u*. Instead, it might be the plural (nominative or accusative).<sup>34</sup>

#### 4.8 <habheryadhi> ‘fulfils’?

The next sequence to be discussed is <habheryadhi> (r1; Fig. 8). This must be a word, because it directly follows *agadgä* and precedes the paragraph sign. In addition, in the first line of the verso, there is a similar sequence, which we provisionally read as <habharyä>.

The reading of <ha> and <dhi> is relatively secure, but the two akṣaras in the middle are more difficult. The second sign, <bhe> in our tentative reading, shows some vague resemblance to <bha> as we know it from other forms of Kharoṣṭhī (see Glass 2009 [2013]). In addition, no other akṣara in the text can be securely interpreted as <bha>, which allows the possibility that the akṣara we have here has to be read so. None of these arguments are conclusive, however. In reading the akṣara after it as <rya>, we follow Schmidt’s (2018: 169 *et passim*) interpretation, although other possibilities exist, for instance a reading <rda>.

Following Georges-Jean Pinault’s (p.c.) suggestion to search for a verb at the end of a sentence, before the paragraph sign, we conjecture that <dhi> is the 3sg. verbal ending. This would be *-ḍi* in Tumšūqese (Emmerick 2009: 395), and the <dh> could represent a voiced dental fricative *ḍ*. However, <khvaṛadhi>, discussed below, seems to be rather a 3pl. pres., and would have <dh> after a nasal, where it more likely represents a voiced dental stop *d*. The first syllable <ha> may be a verbal prefix. As far as Khotanese is concerned, this could be both from PIr. *\*fra-* ‘for’ and *\*ham-* ‘together’, but if we are dealing with Tumšūqese (or a language close to it), it can only be from *\*ham*, since in Tumšūqese the verbal prefix *\*fra-* develops into *ra-* (Konow 1935: 787; Emmerick 1989: 213).

We suggest to derive <habheryadhi> from PIr. intransitive *\*ham-parya-ti* ‘is filled, is fulfilled’ (Khot. *hambīr-*, 3sg. pres. *hambiḍā*), or, alternatively, from PIr. transitive *\*ham-pāraya-ti* ‘fills, fulfils’ (Khot. *hamber-*, 3sg. pres. *hamberāte*; Emmerick 1968: 143; Emmerick 1989: 215). In Tumšūqese, there are examples of a non-etymological nasal in front of a voiced occlusive, e.g. Tum. *pandam-* ‘to construct’, corresponding to Khot. *padam-* (Konow 1935: 820).<sup>35</sup> Apparently the intervocalic *-d-* of *\*pa-dam-* < *\*pad-dam-* < *\*pati-dam-*, rather than lenited *-ḍ-*, was felt as equivalent to *-nd-*, the only other context with medial unlenited *d*. We assume that the opposite happened here, i.e. that the etymological nasal *m* has dropped in <habheryadhi>. However, the question is whether <bh> may stand for *b*. Since the postconsonantal *y* could not have been preserved in either Khotanese or Tumšūqese, it seems to point to a very archaic stage of either language. Finally, the vowel *e* is difficult to account for: it could be due to a palatalisation effect, typical of Khotanese rather than Tumšūqese.

If the verb is correctly identified, we would have an instance of a typical Buddhist phrase that is also found in other Tarim Basin languages, i.e. ‘a wish is fulfilled’. A meaning ‘to fulfil’ is attested for Khotanese *hambīr-* (Bailey 1979: 462–464) and this phrase is also found in e.g. Tocharian B

<sup>34</sup> The latter interpretation is due to Nicholas Sims-Williams (p.c.).

<sup>35</sup> A similar phenomenon is attested in Modern Greek, cf. the use of <μπ> to express either *mb* or (more frequently) simply *b*.



*akālk kən*- ‘a wish is fulfilled’ (base verb), or ‘to fulfil a wish’ (causative; Adams 2013: 169f.). Were it not for the many uncertainties in our analysis and the series of akṣaras in between that we cannot at present interpret at all, one could even tentatively translate the first part of T II T 48 as ‘fulfils the wishes.’<sup>36</sup>

Nicholas Sims-Williams (p.c.) has suggested another possible interpretation of the sequence <habheryadhi>. In case the Formal Kharoṣṭhī language would not be a form of Tumšūqese, but more closely related to Khotanese, or if it is a third language of this branch of Iranian, then the prefix <ha> could also derive from PIr. \**fra*. If so, <habheryadhi> could alternatively go back to PIr. \**fra-bārayati* literally ‘brings forth’, hence e.g. ‘gives’. In this context, it could mean ‘grants’. One could, for instance, compare this with the typical Old Persian formula *Auramazdā xšačam manā frābara* ‘Ahuramazda granted me the lordship’ (e.g. DB §5.D; Schmitt 2009: 38). In this interpretation, the spelling <bh> would then stand for β. Obviously, this interpretation is not possible if the Formal Kharoṣṭhī language is a form of Tumšūqese, and if it is a form of Khotanese, it requires an early prestage in which *-ry-* is still preserved, and the 3sg. pres. ending was apparently *-ḍi* instead of the historically attested *-tā /-də/* or */-dī/*, later */-ṛə/*.

The related form *habharyä* in v1 could be a 3sg. opt. from \**-yāt*, or a ptc. nec. from \**-ya-*.<sup>37</sup> The final *-ä* would seem to fit the latter option better; from \**-yāt* one would rather expect *-ya*. At present, we do not dare to speculate on the difference in root vocalism between *habheryadhi* and *habharyä*.

#### 4.9 <khvaṛadhi> ‘they eat’? and preceding sequences

In this section, we discuss the first half of line three on the recto. Comparing Khotan Kharoṣṭhī, we identify the sixth and eighth akṣara after the § sign as <khva>.

With the identification of the akṣara <khva>, we get a relatively straightforward reading <khvaṛadhi> in r3 (Fig. 9). The first akṣara of this presumed word has a floating stroke below, i.e. <khvā>, but for etymological reasons we think that the *a* is short here. As in *purra* (discussed below), we read the second akṣara as <ra>, with two diacritics: a floating stroke below, and an extra curl on the stem, i.e. <raṣ>. Our etymological interpretation suggests that this extra curl denotes a nasal, i.e. <raṃ>. The third akṣara is <dhi>, which seems to stand here for *di*, not *ḍi*, because of the preceding nasal.

*khvaṛadhi* may be compared with the Khotanese verb *hvar*- ‘to eat’ (Emmerick 1968: 156). The exact phonetic reason for the spelling *khva*- for initial \**hwa*- is unclear, but it is paralleled in Khotan Kharoṣṭhī.<sup>38</sup> As argued by Burrow (1935: 789, correcting his earlier proposal), the name Khvarnarse in CKD 661 probably contains as its first part Iranian *xvar* ‘sun’. The curl in the <r> seems to denote anusvāra, i.e. <khvaramḍhi>, so that the form must be a 3pl. pres. from \**hwaran-di* (without the palatalisation attested in OKh. *hvarindä*, e.g. *Z(ambasta)* 3.59). Apart from the anusvāra-like curl on <ra>, an interpretation as 3pl. is necessary because the 3sg. should have

<sup>36</sup> This transitive reading of *habheryadhi* (as if from \**ham-pāraya-ti*) is needed if it is indeed 3sg., and if *agadga* is indeed nom.-acc. pl. This interpretation additionally avoids the problem of a possible syncope of the stem vowel in the intransitive form \**ham-parya-ti*, as occurred in Khot. intrans. *hambiḍā*, but not in trans. *hamberāte*.

<sup>37</sup> This was suggested to us also by Nicholas Sims-Williams (p.c.).

<sup>38</sup> It is conceivable that \**hw*- was spelled as <khv> in view of the rarity of \**khw*- < \**xw*-.



been disyllabic, with syncope of the middle syllable: \**hwara-ti* > \**hwar-ti* (Khot. *hviḍā*). In view of this interpretation, the floating strokes in <khvāṛa> do not seem to indicate length here.

The words preceding *khvāṛadhī* are possibly to be interpreted as direct objects. The clearest is <ṣ<sub>2</sub>vādu>, of which a variant <ṣ<sub>2</sub>vidu> occurs also in r1 and r4 (Fig. 10). The observed alternation between <i> and <ä> is characteristic of Khotanese, but only from the *Book of Zambasta* onwards, and not yet in the *Śūraṅgamasamādhisūtra* (Śgs) (Emmerick 1979: 240).<sup>39</sup> We tentatively read the top element of the first akṣara as <ṣ<sub>2</sub>> (see above). At the bottom, an upward-going line is attached, which we interpret as a postconsonantal <v> (Glass 2000: 152f.). In theory, it is also possible that it is postconsonantal <r> (as we interpreted it previously, together with Stefan Baums and Ching Chao-jung), but we opt now for <v>, because the diacritic goes rather high up. In addition, there is not one word in Bailey's dictionary of Khotanese (1979) that starts with /šr/,<sup>40</sup> whereas there are words starting with /šv/ (Bailey 1979: 415f.). The diacritic cannot be the vowel <u>, because we already have another vowel: <i> or <ä>. The reading of the second akṣara is uncertain, but we opt for <du>, an interpretation that seems to be supported by the context, but is not obvious palaeographically.

ṣ<sub>2</sub>vidu, ṣ<sub>2</sub>vādu seems related to OKh. *ṣvīda*- 'milk' (Bailey 1979: 415–416, already OKh. in Z 15.93); the final -u would be the ending of the acc. sg. Deriving from Proto-Iranian \*xšwifta- (cf. Av. xšuuīpta° Bartholomae 1904: 562, Pa. šyft Durkin-Meisterernst 2004: 320), ṣ<sub>2</sub>vidu would share with Khotanese the development of \*-ift- to -īd-. It is likely, however, that the same development has taken place in Tumšūqese in view of *hoda*- '7' < \**hafta*. The exact value of ṣ<sub>2</sub> is unclear, but it seems likely that it represents ṣ (ṣ̣), and the original cluster \*xšw- was thus simplified to šv-, like in Khotanese. Since the upper edges of <kṣa> are in other Kharoṣṭhī varieties bent upward instead of downward (Glass 2000: 115f.), as here, it seems less likely to us that ṣ<sub>2</sub> represents xš. For this particular word a reading xš would have been a possibility, but the akṣara is quite frequent in this text and it is unlikely that they would all stand for xš. For ä ~ i, see above.

The word preceding ṣ<sub>2</sub>vādu in r3 ends in -u too, and may modify it. The first akṣara can be read as <stri>, and the second seems to be either <śu> or <yu>. We opt for <yu>, because this allows to interpret the word as the acc. sg. m. of an adjective *striya*-, possibly related to Khotanese *striya*- 'stiffened' < \**straxta*-, the past part. of *strīs*- 'to become stiff' (Emmerick 1968: 135). Perhaps, *striyu* ṣ<sub>2</sub>vādu 'stiffened milk' refers to churned milk or a similar dairy product.

Between *striyu* ṣ<sub>2</sub>vādu and *khvāṛadhī* in r3 we read *khvāṣa*, which seems to be the same word as *khva*<sub>x</sub>so in r4, despite the fact that the initial akṣara has an extra stroke at the lower right, here provisionally marked with  $\overset{\cdot}{x}$  as well as the fact that *khvāṣa* in r3 has a floating stroke under the first akṣara, unlike *khva*<sub>x</sub>so in r4. This would be the only word in this text with a 'regular' <ṣa>. If the reading is correct, it could be compared with OKh. *hvāṣsa*- 'plant, herb' (e.g. Z 2.14, cf. Bailey 1979: 506). As with *khvāṛadhī*, the initial *khv*- would represent \**hw*-. The form *khvāṣa* could be the plural, here acc. pl. as the object of *khvāṛadhī*. Perhaps *khva*<sub>x</sub>so is the corresponding acc. sg., but it would have -o for -u, if it is not an -aa stem, where an acc. sg. in -o could be a contraction of earlier -au (cf. Emmerick 1968: 297).<sup>41</sup>

<sup>39</sup> Sander (1986: 169; 1999: 71f.), following a suggestion by Lín Meicūn, took the double dot as an indication that the language should either be Gāndhārī or Tocharian, but the double dot is also known in Sanskrit Kharoṣṭhī (e.g. CKD 523) and in Khotanese and Tumšūqese Brāhmī. This has independently also been noted by Diego Loukota.

<sup>40</sup> Old Khotanese ṣṣ- goes back to \**sr*- (\**ʃr*-), so that *ṣr*- could only be the result of syncope from \**ṣVr*-.

<sup>41</sup> It may be noted that a -ka-derivative *hvāṣṣaka*- is also attested in OKh. (Z 22.117, cf. Degener 1989: 199).



As the basic meaning of *\*hwar-* is ‘consume’, denoting often ‘eat’, but sometimes also ‘drink’ (Cheung 2007: 147f.), it is possible that the whole string *striyu ś<sub>2</sub>vādu khvāṣa khvāṣṛadhi* in r3 means ‘they consume stiffened milk [and] herbs’.

#### 4.10 <paṇasāma> ‘we take’? and neighbouring sequences

A sequence whose reading seems quite straightforward is <paṇasāma> in line 4 of the recto (Fig. 11). This can be interpreted as the 1pl. pres. or subj. act. of the verb attested in OKh. as *nās-* ‘to take’ (Emmerick 1968: 52), to which the preverb *pa-* (perhaps, but not necessarily, from *\*pati*, cf. Emmerick 1968: 265) has been added. *nās-* is well-attested in Old Khotanese (Śgs, Sgh, Suv, Z), where it is always middle, however, and not active. It is only in Late Khotanese that the verb occurs both in the active and in the middle (Emmerick 1968: 52). Possibly, Late Khotanese has preserved an old state of affairs. Alternatively, the difference in diathesis might be due to the added preverb *pa-*, as shown by such pairs as *thaṃj-* (act.) : *pathaṃj-* (mid.) ‘to pull’, and *dajs-* (act./mid.) : *padajs-* (mid.) ‘to burn’.<sup>42</sup> Reading a 1pl. mid. and not act. would need an additional syllable *-ne* after the ending *-ama* to obtain *\*-amane* (cf. the Khotanese ending *-āmane* Emmerick 1968: 198, 201). However, this is not possible on palaeographic grounds, as the unclear akṣara after <ma> does not resemble <na> and bears no trace of the diacritic for <e>.

If the word is <paṇasā>, with a word boundary before <ma>, one could see in it, bearing in mind the Khotanese endings, a 2sg. imper. act., a 3sg. subj. act., or a 2sg. subj. mid. One might also envisage the possibility that the subst. *nasa-* ‘share’ is involved, to which the prefix *pa-* has been added. If not from *nās-* ‘to take’ (Proto-Iranian *\*Hnas-*, cf. Cheung 2007: 183), it may be from *nās-* ‘to be hungry, starving’ (attested in Khotanese with preverb *bi-*, if from *\*nād*, with Cheung 2007: 277) or from *nās-* ‘to quiver, shake’ (Kh. *va-nās-* < *\*nas-* ‘to perish’, Cheung 2007: 282). However, as the overall meaning of the line still escapes us, the most satisfactory solution for the time being seems to be *nās-* ‘to take’, which is the most frequent and well-attested verb in Khotanese among the three above.

If read as a verb, <paṇasāma> might have as its object the preceding <aṣ<sub>2</sub>ṣoṇa> and perhaps also <ś/yirya>, which might then both be interpreted as acc. pl. in *-a*. If correctly analysed as a 1pl., one should note that in the same line another verb in the 1pl. might occur. This may be hidden behind the sequence <siḷama>, which reminds us of the Khotanese verb *śāj-* ‘to succeed’ (from a Prakrit form of Skt. *śidhyati* ‘to succeed’, cf. Emmerick 1968: 133).<sup>43</sup>

#### 4.11 Further sequences

There are five further recurring sequences that we present in the order of their first occurrence.

First, immediately after <aṣ<sub>2</sub>ṣoṇa> (r1), there is a sequence, provisionally read as <patija>, which is repeated in line 3 of the recto (Fig. 12). An *-a* derivation from the verb *patāts-* ‘to give up,

<sup>42</sup> Needless to say, there are also many cases where the prefixed verb takes the same diathesis as the simplex; cf. *jsañ-* (mid.) : *pajsañ-* (mid.) ‘to be struck’, and *bañ-* (act.) : *pabañ-* (act.) ‘to bind’.

<sup>43</sup> Alternatively, one might read *ja ma ś/yirya* and interpret *ja ma* as a sequence of relative pron. (a similar form is attested in Tumšūqese, cf. Konow 1935: 818) + demonstrative (cf. again Konow 1935: 819) referring to *ś/yirya*.



abandon' (Emmerick 1968: 67) is difficult because of the single *-t-* of Khotanese, which we expect to be represented by *-dh-* (cf. <d<sub>1</sub>>, probably δ, of Tumšūqese). Moreover, one should also reckon with a correspondence *ts ~ j*, which is remarkable. A derivation from the verb *pattamj-* 'to produce' (Emmerick 1968: 66–67, *a*-derivative from the present stem) accounts for the *t < \*tt* (cf. Khotan Prakrit intervocalic <t> for *tt*; Konow 1936: 234f.), but it would leave the vocalism unexplained.<sup>44</sup>

Second, a sequence of two akṣaras is repeated 5 times (Fig. 13). Schmidt (2018: 169) had read this in three instances as <poṃñc>, Tocharian for 'all', but this is to be rejected (see above § 2.3). During the Leiden workshop, Stefan Baums and Ching Chao-jung suggested to read it as <pu-tra>. With this reading, we have not been able to find any useful interpretation, and in view of the akṣara shapes in CKD 661 we now read it as <purra>. While Kharoṣṭhī normally does not write geminate consonants, <rra> has also been tentatively read by Duàn Qíng (2013: 204) in the newly discovered Khotan Prakrit document (CKD 843) and a combination <rra> is, of course, well-known from Khotanese Brāhmī. The following occurrences are found: <purra> (r1), followed by <ṣ<sub>2</sub>ä>; <purra> (r2), followed by <ṣ<sub>2</sub>a>; <purra> (r2), followed by <na>; <purre> (r1–2), followed by <jā>; and <purr[e]> (v1), followed by <jä>. Although we cannot fit them in clauses, it is possible that all five instances are forms of *purra-* 'full', cf. Khot. *purra-* 'full'. Alternatively, some or all of the occurrences could also be compared with Khot. *purrā-* '(full) moon' (Bailey 1979: 244).

As suggested to us by Nicholas Sims-Williams (p.c.), the elements *ṣ<sub>2</sub>ä* and *ṣ<sub>2</sub>a* could perhaps be forms of the demonstrative (cf. OKh. *ṣā* masc. and *ṣa* fem., from *\*aiśa-*), although the position after *purra-* suggests that they rather belong with the following, i.e. *ṣ<sub>2</sub>ä cu dhi* 'this who' in r1 and *ṣ<sub>2</sub>a ś/yiryā* 'this goodness (?)' in r2 (see below). One may think of the following *jā* and *jä* as forms of the relative, from *\*ya-*, but the context is too unclear to decide. Should the akṣara now read as <ṣ<sub>2</sub>a> rather have to be read as <ka>, then we would have <purrakā> in r1 and <purrika> in r2. This word could be compared with OKh. *purrāka-* 'overcomer, conqueror',<sup>45</sup> an *-āka*-derivative (Degener 1989: 49) from the root *purr-* 'to overcome' (Emmerick 1968: 84).

A fourth repeating sequence is <mast->, found in two different variants, i.e. <mastena> (r3) and <mastiśa/mastiya> (r1 and r4) (Fig. 14). The exact interpretation of this word is still unclear, but there are different options to interpret it in the vein of Khotanese or Tumšūqese. A first option would be to connect it to the word for 'moon' (Khot. *māsti-*; Bailey 1979: 331), although one would expect to find it next to a month name, which we have been unable to identify. However, if some of the occurrences of <purra> refer to the full moon, this option would still be conceivable. At first sight, it does not seem plausible to connect it to the Iranian verbal root *\*mad-* 'to be intoxicated' either, even though it could somehow fit the consummation of milk and herbs (§4.9). As a final option, one could relate it to Khotanese *māsta-* 'big, great' <PIr. *\*masita-* (Bailey 1979: 333), although we could not find a noun with which this adjective could be in a noun phrase.

Finally, the sequence <ś/yiryā> (r2) is repeated twice (Fig. 15). A variant without the floating line below, i.e. <ś/yiryā> is found in r4 and v1. As pointed out by Georges-Jean Pinault (p.c.), if the first akṣara is read as <śi>, the word is similar to Khotanese *śśāra-* 'good' (Bailey 1979: 400f.; Emmerick and Skjærvø 1982: 117f.). As suggested to us by Nicholas Sims-Williams (p.c.), it might be an abstract in *-ya*, i.e. *śiryā* 'goodness'. This option seems better than a locative in *-ya*.

<sup>44</sup> *i*-vocalism is expected in the past part., cf. Khot. *pattiya-*, but it is unlikely that the *-y-* would be written or realised as *-j-*.

<sup>45</sup> In the *Book of Zambasta* it is used with reference to the Buddha, who is 'conqueror of Māra' (*mārāṇā purrāki* Z 13.85) and as gloss to the 'scatterers of the army of Māra' (*[māri]ñi hīñi tcalaljāka* Z 24.643), who are 'overcomers' (*purrāka*). In Suv 6.1.65 it refers to the the *sūtra* itself, which is 'conqueror of enemies' (*sāñānu ... purrākā*).



#### 4.12 Some final proposals

We end our discussion of T II T 48 with some final proposals, even more tentative than what we have presented thus far, as in these cases the word boundaries are more uncertain.

At the end of r1, there is a sequence <śadhi>, which may tentatively be compared with Old Khotanese *śśāta-*, the past part. of *śś-* ‘to lie down’ (Emmerick 1968: 127).

Further, it may be possible to read a sequence <paṇaṭhā> in the second half of r2, after another <cu dhi>. While the reading <pa> cannot be doubted, this is not true of the two other signs. The overall ductus of the second akṣara is reminiscent of retroflex nasal, but one has to assume that the characteristic long tail has been adopted to give it a more formal appearance. Similarly, the last akṣara certainly resembles <ṭhā>, but other options are not to be excluded a priori. Assuming that as in Gāndhāri the retroflex and dental nasals are interchangeable and <ṭha> represents a cluster *ṣṭ(h)a*,<sup>46</sup> <paṇaṭhā> reminds of the Khotanese past participle *paṇaṣṭā* ‘destroyed’ (Emmerick 1968: 70).

Second, we find a sequence <hajza> in r2. The reading of the second akṣara is very tentative. The upper part looks like <ja>, and the middle stroke ends in a fork that might derive from the upper part of <za>. The same akṣara occurs Or.15009/44 v3, where the <z> is clearer. As suggested by Chams Bernard (p.c.), this reading reminds of the Khotanese postposition *hamtsa* ‘together’, constructed with the instr.-abl. case (Emmerick 1965: 32). Conceivably, we could have the same case preceding it here, i.e. *purraṇa hajza* (for *purra* itself, see above). This would mean that, at least in this particular case, *purra-* would not be feminine (the instr.-abl. sg. of the *ā*-declension ends in <sup>(i)</sup>*e* or <sup>(i)</sup>*ä*; Emmerick 1968: 271) unless the masculine ending would have spread analogically. Less likely, *-ana* could stand here for the gen.-dat. pl. in *-ānu*. If the *na* stands rather on its own, it could alternatively be the negation. An open question is whether this *hajza* has any connection with the recently identified *adzo* ‘together with’ in Tumšūqese (Maue and Ogihara 2017). In spite of being a fairly good semantic match of Khotanese *hamtsa*, Tumšūqese *adzo* cannot be cognate with it according to Maue and Ogihara (2017: 426). Naturally, then, *hajza* would be cognate with Khotanese *hamtsa* but not with Tumšūqese *adzo*. If our very tentative reading turned out to be correct, a special combination <jz> would have been used to represent *dz*, somewhat similar to Khotanese <jś>, which has the same value (e.g. Emmerick 1989: 208; 2009: 381).

Between the possible word <hajza> and <ḡaḡdḡä> (r2–3), we read <vagyādhi>, which could in turn also be a word. The <dhi> might be the 3sg. pres. verbal ending, but we cannot offer an interpretation of this word. An alternative reading would be <vargādhi>.

Finally, at the end of the calligraphic part in v1, a sequence <hima> can be read. This has to be a word of its own, because it is immediately preceded by <ḡaḡdḡä> and followed by the paragraph sign. In theory, this could be the 2sg. mid. of the Khotanese root *hām-* ‘to be(come)’. However, this interpretation is not possible if the language is rather Tumšūqese, as the corresponding verb there is *rām-* (Konow 1935: 821).

<sup>46</sup> For Gāndhāri, cf. e.g. Konow (1929: cii–civ) as far as the nasal is concerned and (1929: cx) for the phonetic cluster.



#### 4.13 Transliteration of selected sequences and akṣaras

To hopefully further the scholarly discussion on Formal Kharoṣṭhī we give here a tentative transliteration and transcription of the whole folio except the cursive part. Akṣara elements for which no reasonable suggestion can be made are denoted with ‘C’ for the consonantal part and ‘V’ for the vowel. However, whenever we do suggest a reading, this has always to be treated with due caution, as other options can mostly not be excluded.

##### *Tentative transliteration*

- [r1] a ṣ<sub>2</sub>a ṣo ñā pa ti ja pu rra ṣ<sub>2</sub>ä cu dhi a ḡga ḡga ha bhe rya dhi ¶ cu su va rya Ca<sup>47</sup> Ca<sup>48</sup> ta ṣ/ya ṣ<sub>2</sub>vi du ma sti ṣ/ya dhi pu
- [r2] rre ja ḡ ṣ/ya pu rra ṣ<sub>2</sub>a ṣ/yi ryā CV CV ṣ/ya a ṣ<sub>2</sub>a ṣo ñā Ci te cā cu dhi pa ṇa ṭhā pu rra na<sup>49</sup> ḡ ja za va gyā<sup>50</sup> dhi a ḡ
- [r3] ḡḡ ¶ ca stri yu ṣ<sub>2</sub>vā o du khvā ṣa khvā ṣa dhi ḡ te va pu te na dhu Ca tu gu CV za pa ti ja Cu Ca ñā ma ste na
- [r4] dhi khvā<sub>x</sub> ṣo dhu CV ja ¶ cu pho ṣ<sub>2</sub>vi du Ci Ca dhi ṣ<sub>2</sub>a<sub>x</sub><sup>51</sup> ṣa ṣi ja ma ṣ/yi rya a ṣ<sub>2</sub>a ṣo ñā pa na ṣa ma Ca ma sti ṣ/ya ṣa<sub>x</sub>
- [v1] pu rr[e] jā ṣ/yi rya ṣ/ye za he CV ṣ<sub>2</sub>a a ṣ<sub>2</sub>a ṣya ha bha ryā a ḡ ḡḡ hi ma ¶ (*beginning of the cursive part*)

##### *Transcription with provisional word boundaries*

- [r1] aṣ<sub>2</sub>aṣoṇā patija purra ṣ<sub>2</sub>ä cu dhi ḡḡḡḡ habheryadhi ¶ cu su va rya Ca Ca ta ṣ/ya ṣ<sub>2</sub>vidu mastiṣ/ya dhi pu-
- [r2] rre ja ḡ ṣ/ya purra ṣ<sub>2</sub>a ṣ/yiryā CV CV ṣ/ya aṣ<sub>2</sub>aṣoṇā Citēcā cu dhi paṇaṭhā purra na hajza vāgyādhi ḡḡ-
- [r3] ḡḡ ¶ ca striyu ṣ<sub>2</sub>vādu khvāṣa khvāṣadhi ḡ te va pu te na dhu Ca tu gu CV za patija Cu Ca ñā mastena
- [r4] dhi khvā<sub>x</sub> ṣo dhu CV ja ¶ cu pho ṣ<sub>2</sub>vidu CiCaḡḡ<sup>52</sup> ṣ<sub>2</sub>a<sub>x</sub> ṣa ṣijaṃa ṣ/yiryā aṣ<sub>2</sub>aṣoṇā paṇasāma Ca mastiṣ/ya ṣa<sub>x</sub>
- [v1] purr[e] jā ṣ/yiryā ṣ/ye za he CV ṣ<sub>2</sub>a aṣ<sub>2</sub>aṣya habharyā ḡḡḡḡ hima ¶ (*beginning of the cursive part*)

<sup>47</sup> A tentative possibility would be <la>. The same akṣara (but with anusvāra) occurs in the right fragment of MIK B1932, line 2.

<sup>48</sup> This akṣara is close to our <ṣo> in <aṣ<sub>2</sub>aṣoṇā>, but the top is somewhat different. See also the akṣara table.

<sup>49</sup> The reading <na> is not fully certain and is largely based on the long tail below the akṣara which one also finds in other Kharoṣṭhī varieties.

<sup>50</sup> Alternatively, <rga>, cf. § 4.11.

<sup>51</sup> This akṣara has a small circle added to it at the right, which is unknown so far in Kharoṣṭhī (cf. also § 2.3). It is indicated here with the ‘<sub>x</sub>’ after our transliteration of the akṣara.

<sup>52</sup> The same word occurs in the right fragment of MIK B1932, line 3 (Schmidt’s ‘Text 3’) with *u* in the last akṣara: *CiCaḡḡhu*. The form in *-i* may be the nom. sg. of a masculine *a*-stem, in which case *CiCaḡḡhu* would be the corresponding acc. sg. Otherwise, *CiCaḡḡhi* may be a 3sg. or 3pl. present form, and *CiCaḡḡhu* the corresponding 3sg. or 3pl. imperative. The first akṣara may be <khi>.



## 5 ARGUMENTS FOR THE LANGUAGE OF THE FORMAL KHAROṢṬHĪ MANUSCRIPTS

As is clear from the above, our hypothesis is that the language of the Formal Kharoṣṭhī manuscripts is related to Khotanese and Tumšūqese. We are aware of the fact that much more secure evidence is needed to prove (or disprove) this hypothesis, but we like to point out that many other options are unlikely or can be excluded:

- Sanskrit can be discarded as highly unlikely, in view of the bilinguals with Sanskrit written in Brāhmī.
- Old Turkic and Tibetan are unlikely on chronological grounds: although the corpus cannot be dated exactly, it may be from the fifth or perhaps sixth century, which seems too early for Turkic and Tibetan. No elements pointing to these languages have been found.
- Chinese is unlikely because this language has a well-established script tradition and no elements pointing to Chinese have been found.
- Gāndhārī would seem an obvious option to consider at least, since Gāndhārī is so strongly associated with the Kharoṣṭhī script (Salomon 1999: 112). Yet, this option in turn is very unlikely in view of the fact that all specialists of Kharoṣṭhī know Gāndhārī and they should have been able to read it.<sup>53</sup>
- In our view, Schmidt's identification of the Formal Kharoṣṭhī language as Tocharian has definitely failed, but this does not mean that it could not be Tocharian under another interpretation.<sup>54</sup> However, the initial readings by Stefan Baums and Ching Chao-jung during the Leiden workshop have not yielded any indication that the language could be Tocharian, and the large number of different consonants, which include voiced stops and fricatives, makes this very unlikely.

The evidence that we think to have identified points definitely to an Iranian language,<sup>55</sup> probably related to Khotanese and Tumšūqese but seemingly more archaic than either. We summarise here the linguistic clues for a more exact identification we see so far. More generally, it should be borne in mind that recent research (Salomon 1998; Strauch 2012) has emphasised the flexibility of the Kharoṣṭhī script to write Sanskrit, so that it should in principle also be possible to adapt it for a Middle Iranian language.<sup>56</sup>

- *cu* and *cudhi*, or *dhi* or *udhi*: *cu* < \**čim* is identical to Khotanese *cu*, but also to Tumšūqese *cu*. Sogdian shows the same development of \**čim* to *cw*, but Bactrian σιδι etc. preserves the

<sup>53</sup> Salomon (2007: 186) remarks about the bilingual fragments: 'it might be supposed that the Kharoṣṭhī text represents some version of the same text [viz. as the Sanskrit Brāhmī text], but apparently in a language other than Gāndhārī. But it has not yet proven possible to confirm this, and thus this peculiar manuscript [i.e. Or.15009/44] remains a mystery' (clarification in rectangular brackets ours). Franz Bernhard (1970: 57; cf. also Lín 1996: 200) thought them to be written in Gāndhārī.

<sup>54</sup> Lín (1996: 200) suggested that it would be a 'kind of Tocharo-Gāndhārī mixed language' (cf. also Sander 1999: 73), but without any detailed argumentation. Also, he declares to have found Bactrian loanwords in the Formal Kharoṣṭhī, but he does not cite them, nor does he give exact references. Sander (1986: 169; 1999: 72) considered both Tocharian and Gāndhārī as possibilities, but made no definitive choice.

<sup>55</sup> This was also the opinion of Wille upon the identification of the bilingual fragment Or.15009/44 as part of the *Anaparāddhastotra* (2006: 31 fn. 12; also mentioned in Salomon 2007: 186).

<sup>56</sup> With many thanks to Georges-Jean Pinault for drawing our attention to these publications.



old *\*i*. On the other hand, *dhi* or *udhi* < *\*uti* is clearly different from Khotanese *u* (Sims-Williams forthc.), where the element *\*ti* is completely lost, and possibly matched by a hypothetical Tumšūqese *uđi\** or *đi\** only.

- *-oña*: The suffix *-oña* matches Khotanese closely, but might as well have been Tumšūqese.
- *agadga* ‘wish’: This word can hardly be genuine Khotanese or Tumšūqese because of the *g* in the third syllable and is probably a loanword.
- *habheryadhi* ‘fulfils’: If from *\*ham-parya-ti* or *\*ham-pāraya-ti*, the loss of *\*m* fits Tumšūqese better than Khotanese. The preservation of *ry* would need an archaic prestage for both. If the ending *-dhi* stands for *-đi* < *\*-ti*, this fits Tumšūqese better than Khotanese, because in Khotanese, *\*-ti* becomes *-tä*, probably /-də/ or /-dɪ/ at first, and later /-ʔə/. However, in *khvaṛadhi* ‘they eat’ <dh> seems to occur after a nasal and thus more likely stands for *d*. One would have to assume that <dh> could be used for *d* in this position.
- *khvaṛadhi* ‘they eat’: The absence of palatalisation in the ending fits Tumšūqese better than Khotanese (which has *-indä*), or this text would show very archaic Khotanese. As noted above, <dh> presumably stands here for /dɪ/ after a nasal. *ṣ<sub>2</sub>vādu/ṣ<sub>2</sub>vidu* ‘milk’ seems to fit both Khotanese and Tumšūqese, although the exact phonetics of <ṣ<sub>2</sub>v> are not clear.

Thus, as far as the scanty linguistic evidence goes, the language of the Formal Kharoṣṭhī fragments is more likely to be an early stage of Tumšūqese if <dh> in the 3sg. ending represents *δ* from old intervocalic *\*t*, and the particle *udhi* cannot be from Khotanese because it has reduced the particle to *u*. Most other readings are too uncertain to be used or would fit early Khotanese as well as early Tumšūqese.

Obviously, as pointed out to us by Nicholas Sims-Williams and Sasha Lubotsky (p.c.), the option that the Formal Kharoṣṭhī language is not ancestral to either Khotanese or Tumšūqese cannot be fully excluded. This option would allow to take *agadga* as inherited, and to take *habheryadhi* alternatively from *\*fra-bāraya-ti*. We find this option unlikely in view of the evidence so far: in this hypothetical language *\*-aka* would have been preserved as *-ga* (with syncope), which would need a fairly early separation from Khotanese-Tumšūqese. At the same time, this language would share with Tumšūqese the development of intervocalic *\*t* to *δ* and possibly the development of *\*fra-* > *\*hra-* > *ha-* with Khotanese. It would also be difficult to understand within the linguistic landscape of the ancient Tarim Basin.

Indeed, there are a number of nonlinguistic reasons to believe that the Formal Kharoṣṭhī language may be an earlier form of Tumšūqese.

First of all, the find sites of the Formal Kharoṣṭhī manuscripts in the Kuča, Yānqí and Turfan regions, that is, in the whole area of Tocharian B, fits Tumšūqese best: Tumšūqese manuscripts are found in Tumšūq, but also in Turfan region (Maue 2009) and the language was apparently spread throughout the sphere of influence of Kuča. Although a Khotanese folio was found in Šorcuq in Yānqí region (Maggi 2004), the distribution of the Formal Kharoṣṭhī corpus otherwise does not fit Khotanese at all.

Second, we know from archaeological evidence that Tumšūq had already a flourishing Buddhist culture, at the end of the fourth century, long before the adoption of the Brāhmī script from the Tocharians in the 7th century or even later (Hambis *et al.* 1961–1964: 115–118).<sup>57</sup> It would not be surprising if they had had a written literature before, for instance in Formal Kharoṣṭhī.

<sup>57</sup> Hambis *et al.* (1961–1964: 115) refer to Pelliot, according to whom Tumšūq ‘est l’un des plus anciens sites bouddhiques du Turkestan chinois, et plus ancien tout au moins que l’on a trouvée au Nord du Tarim’.



Third, although there is no doubt that the Tumšūqese Brāhmī script has been adopted from the Tocharian B script, it is not in all details easily explained from that source alone. For instance, the use of the *ā* diacritic is different in Tumšūqese, and the typical Tocharian Fremdzeichen are hardly used. More importantly, 12 extra akṣaras have been added to represent sounds missing from the standard Sanskrit Brāhmī inventory (Konow 1935: 776). Some of these could have been taken from Formal Kharoṣṭhī (Ching Chao-jung p.c.). Hitch (1984: 198–202) also suggested that Fremdzeichen nr. 12 /xš/ could derive from Kharoṣṭhī <kṣ>.<sup>58</sup>

Thus, it is conceivable that in the earlier period of local literature, when Khotanese and Tocharian B were first written down, approximately from the fifth century onwards, Tumšūqese was written down too, but in the Formal Kharoṣṭhī script. When Kuchean influence became stronger in Tumšūq, Formal Kharoṣṭhī was gradually replaced by the Brāhmī script, and thanks to Kuchean influence towards the east, Tumšūqese could spread as far as Turfan too.

FIGURES

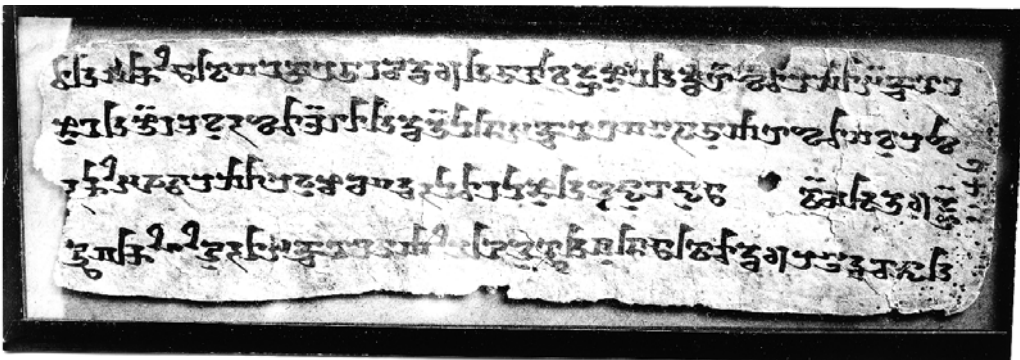


Figure 1: T II T 48 recto side (© Museum für Asiatische Kunst Berlin)

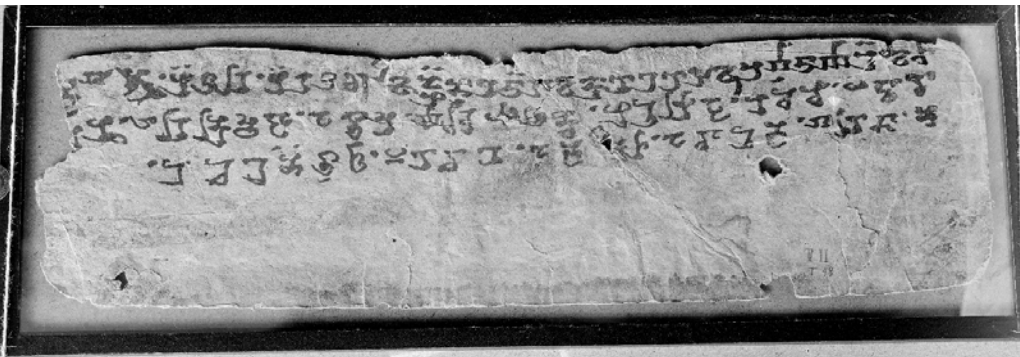


Figure 2: T II T 48 verso side (© Museum für Asiatische Kunst Berlin)

<sup>58</sup> See also Hitch (2009: 19ff.) on the Gandharan cultural influence on Tumšūq, including Kharoṣṭhī.



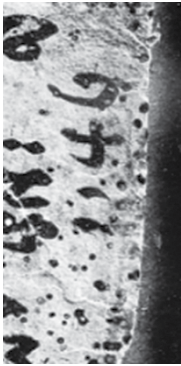


Figure 3: close-up of the leaf number of T II T 48 recto side. The sign on top stands for <10>, the second sign for <4> and the two signs below each for <1>, which makes <16>.



Aśokan	BL 1	BL 9	BL 21	KDhp	Niya	Schøyen 1	Schøyen 2
𑀓, 𑀔	𑀕	𑀖	𑀗	𑀘	𑀙	𑀚	𑀛
Aśokan	BL 1	BL 9	BL 21	KDhp	Niya	Schøyen 1	Schøyen 2
𑀜	𑀝	𑀞, 𑀟	𑀠	𑀡	𑀢	𑀣	𑀤

Figure 4: difference between <c> (here <cu>) and <dh> (here <dhi>); attached are different forms of <ca> and <dha> from other varieties of Kharoṣṭhī (Glass 2000: 61; 80).



Figure 5a: <aṣ₂aṣoṅā> (r1).



Figure 5b: <aṣ₂aṣoṅā> (r2).

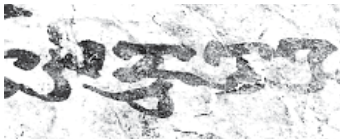


Figure 5c: <aṣ₂aṣoṅā> (r4).

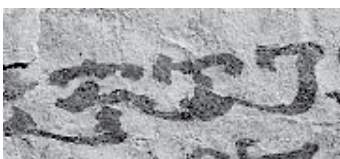


Figure 5d: <aṣ₂aṣya> (v1).





Aśokan	BL 1	BL 9	BL 21	KDhp	Niya	Schøyen 1	Schøyen 2
𑀅	𑀅, 𑀅	𑀅, 𑀅	𑀅	𑀅, 𑀅	𑀅	𑀅	𑀅



Aśokan	BL 1	BL 9	BL 21	KDhp	Niya	Schøyen 1	Schøyen 2
𑀅	𑀅	𑀅	𑀅	𑀅	𑀅	𑀅	𑀅

Figure 6: difference between akṣaras provisionally read as <a> and <va> with other Kharoṣṭhī shapes from Glass (2000: 33; 97) for comparison.

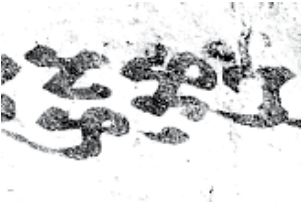


Figure 7a: <agaḍga> (r1).



Figure 7b: <agaḍgā> (end r2; beginning r3).



Figure 7c: <agaḍgā> (v1).

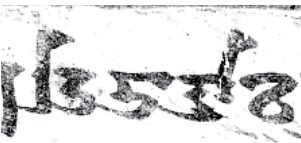


Figure 8a: <habheryadhi> (r1).



Figure 8b: <habharyā> (v1).





Figure 9:  $\langle khvar\grave{a}dhi \rangle$  (r3).



Figure 10a:  $\langle \grave{s}_2vidu \rangle$  (r1).

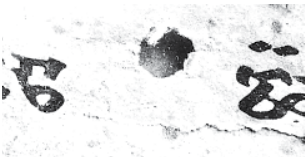


Figure 10b:  $\langle \grave{s}_2v\acute{a}du \rangle$  (r3).

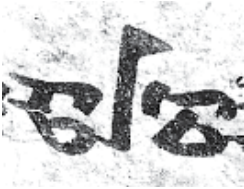


Figure 10c:  $\langle \grave{s}_2vidu \rangle$  (r4).

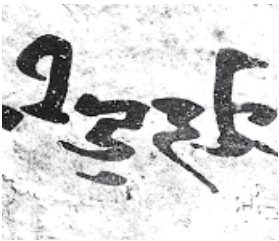


Figure 11:  $\langle \grave{p}anas\grave{a}ma \rangle$  (r4).

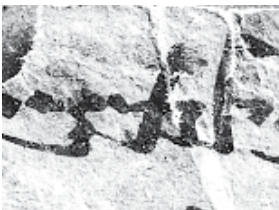


Figure 12a:  $\langle patij\grave{a} \rangle$  (r1).



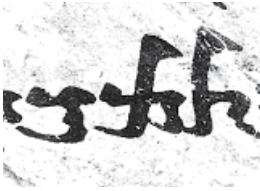


Figure 12b : &lt;patija&gt; (r3).



Figure 13a: &lt;purra&gt; (r1).



Figure 13b: &lt;purre&gt; (end r1, beginning r2).

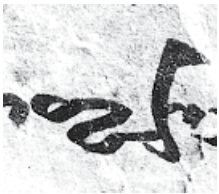


Figure 13c: &lt;purra&gt; (r2).



Figure 13d : &lt;purra&gt; (r2, second time).



Figure 13e: &lt;purr[e]&gt; (v1).





Figure 14a: <mastiś/ya> (r1).



Figure 14b: <mastena> (r3).

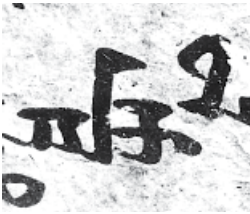


Figure 14c: <mastiś/ya> (r4).

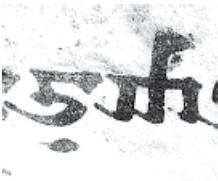


Figure 15a: <ś/yirya> (r2).



Figure 15b: <ś/yirya> (r4).



Figure 15c: <ś/yirya> (v1).



## BIBLIOGRAPHY

- ADAMS, Douglas Q. 2013. *Dictionary of Tocharian B: Revised and Greatly Enlarged*. [Leiden Studies in Indo-European 10]. Amsterdam: Rodopi.
- BAILEY, Harold W. 1973. 'Taklamakan Miscellany.' *BSOAS* 36: 224–227.
- BAILEY, Harold W. 1979. *Dictionary of Khotan Saka*. Cambridge: Cambridge University.
- BARTHOLOMAE, Christian 1904. *Altiranisches Wörterbuch*. Strassburg: K.J. Trübner.
- BAXTER, William H. and Laurent SAGART 2014. *Old Chinese. A new reconstruction*. Oxford: Oxford University.
- BERNHARD, Franz 1970. 'Gāndhāri and the Buddhist Mission in Central Asia.' In: JayadevaTILAKASIRI (ed.) *Añjali: Papers on Indology and Buddhism: A Felicitation Volume Presented to Oliver Hector de Alwis Wijesekera on His Sixtieth Birthday*. Peradeniya: University of Ceylon, 55–62.
- BERNHARD, Franz 1976. 'Bilingual (Gāndhāri and Kuchean) Documents Written in Kharoṣṭhī and Brāhmī from Chinese Turkestan.' In: Albert R. DAVIS (ed.) *Proceedings of the 28<sup>th</sup> International Congress of Orientalists, Canberra 6–12 January 1971*. Wiesbaden: Harassowitz, 274.
- BOYER, Augustin M., Edward James RAPSON and Émile SENART 1918. 'Une tablette kharoṣṭhī-sanskrite de la collection de Sir Aurel Stein.' *JA* [11<sup>ème</sup> série] 12: 319–327.
- BOYER, Augustin M. et al. 1920–1929. *Kharoṣṭhī Inscriptions Discovered by Sir Aurel Stein in Chinese Turkestan*. Oxford: Clarendon.
- BURROW, Thomas 1935. 'Iranian Words in the Kharoṣṭhī Documents from Chinese Turkestan – II.' *BSOAS* 7: 779–790.
- BURROW, Thomas 1937. *The Language of the Kharoṣṭhī Documents from Chinese Turkestan*. Cambridge: Cambridge University.
- CHEUNG, Johnny 2007. *Etymological Dictionary of the Iranian Verb*. [Leiden Indo-European Etymological Dictionary Series 2.] Leiden: Brill.
- CHING, Chao-jung 2013. 'Reanalyzing the Kuchean-Prākṛit Tablets THT 4059, THT 4062 and SI P/141.' *Tocharian and Indo-European Studies* 14: 55–94.
- DEGENER, Almuth 1989. *Khotanische Suffixe*. [Alt- und neu-indische Studien 39.] Stuttgart: Steiner.
- DUÀN Qíng 段晴 2013. 'Nèiróng bùmíng Qūlúwén mùdú: Guótú BH5-7. 内容不明佉卢文本牍: 国图 BH5-7 [A Kharoṣṭhī wooden tablet with unidentified content: BH5-7 from the National Library of China].' In: DUÀN Qíng 段晴 and ZHĀNG Zhìqīng 张志清 (eds.) *Zhōngguó Guójiā Túshūguǎn cáng Xīyù wénshù: Fànwén, Qūlúwén juàn* 中国国家图书馆藏西域文书: 梵文、佉卢文卷 [Fànwén bèiyèjīng yǔ Fójiào wénxiàn xiliè cóngshū 梵文贝叶经与佛教文献系列丛书 3.] Shànghǎi: Zhōngxī Shūjú, 203–205.
- DURKIN-MEISTERERNST, Desmond 2004. *Dictionary of Manichaean Middle Persian and Parthian*. Turnhout: Brepols.
- EMMERICK, Ronald 1965. 'Syntax of the Cases in Khotanese.' *BSOAS* 28: 24–33.
- EMMERICK, Ronald 1968. *Saka Grammatical Studies*. [London Oriental Series vol. 20.] London: Oxford University.
- EMMERICK, Ronald 1979. 'The Vowel Phonemes of Khotanese.' In: Bela BROGANYI (ed.) *Studies in Diachronic, Synchronic, and Typological Linguistics: Festschrift for Oswald Szemerényi*. Amsterdam: Benjamins, 239–250.
- EMMERICK, Ronald 1989. 'Khotanese and Tumshuqese.' In: Rüdiger SCHMITT (ed.) *Compendium linguarum Iranicarum*. Wiesbaden: Reichert, 204–229.
- EMMERICK, Ronald 2009. 'Khotanese and Tumshuqese.' In: Gernot WINDFUHR (ed.) *The Iranian Languages*. London: Routledge, 377–415.



- EMMERICK, Ronald and Prods O. SKJÆRVØ 1982. *Studies in the Vocabulary of Khotanese I*. Vienna: Österreichische Akademie der Wissenschaften.
- FILLIOZAT, Jean 1958. 'L'agalloche et les manuscrits sur bois dans l'Inde et les pays de civilisation indienne.' *JA* 246: 85–93.
- GHARIB, Badr al-Zamān 1995. *Sogdian dictionary: Sogdian-Persian-English*. Tehran: Farhangan.
- GLASS, Andrew 2000. *A Preliminary Study of Kharoṣṭhī Manuscript Paleography*. (MA Thesis. University of Washington)
- GLASS, Andrew 2009 [2013]. 'Bha.' *Bulletin of the Asia Institute* 23: 79–86.
- HAMBIS, Paul et al. 1961–1964. *Toumchouq*. Paris: Adrien-Maisonneuve.
- HITCH, Douglas 1984. 'Kharoṣṭhī Influences on the Saka Brāhmi Scripts.' In: Wojciech SKALMOWSKI and Alois VAN TONGERLOO (eds.) *Middle Iranian Studies: Proceedings of the International Symposium Organized by the Katholieke Universiteit Leuven from the 17th to the 20th of May 1982*. Leuven: Peeters, 187–202.
- HITCH, Douglas 2009. *The Special Status of Turfan*. [Sino-Platonic Papers 186.] Philadelphia: University of Pennsylvania, 1–61.
- KONOW, Sten 1929. *Kharoṣṭhī Inscriptions with the Exception of Those of Aśoka*. [Corpus Inscriptionum Indicarum, Vol. II, Part I.] Calcutta: Government of India Central Publication Branch.
- KONOW, Sten 1935. 'Ein neuer Saka-Dialekt.' *Sitzungsberichte der Preußischen Akademie der Wissenschaften, philosophisch-historische Klasse* 1935: 772–823.
- KONOW, Sten 1936. 'Note on Khotanī Saka and the Central Asian Prakrit.' *AO* 14: 231–240.
- LE COQ, Albert von 1909. 'A Short Account of the Origin, Journey and Results of the First Royal Prussian (Second German) Expedition to Turfan in Chinese Turkistan.' *JRAS* 1909: 299–322.
- LE COQ, Albert von 1928. *Von Land und Leuten in Ostturkistan: Berichte und Abenteuer der 4. Deutschen Turfanexpedition*. Leipzig: Verlag der J. C. Hinrichs'schen Buchhandlung.
- LÍN, Méicūn 1996. 'Kharoṣṭhī Bibliography: The Collections from China (1897–1993).' *CAJ* 40: 188–220.
- MAGGI, Mauro 2004. 'The Manuscript T III Š 16: Its Importance for the History of Khotanese Literature.' In: Desmond DURKIN-MEISTERERNST et al. (ed.) *Turfan Revisited – The First Century of Research into the Arts and Cultures of the Silk Road*. [Monographien zur indischen Archäologie, Kunst und Philologie, Band 17] Berlin: Dietrich Reimer, 184–190.
- MALZAHN, Melanie 2007. 'A Tocharian Brahmi Chart.' In: Melanie MALZAHN (ed.) *Instrumenta Tocharica*. Heidelberg: Winter, 223–254.
- MAUE, Dieter 2009. *Tumshuqese Manuscripts. A Provisional Handlist*. <http://titus.uni-frankfurt.de/texte/iranica/tumshuq/handlist.pdf>
- MAUE, Dieter and OGIHARA Hirotooshi 2017. 'Tumshukische Miscellen III. 3. Tumshukese dental affricates.' In: Team TURFANFORSCHUNG (eds.) *Zur lichten Heimat: Studien zu Manichäismus, Iranistik und Zentralasienkunde im Gedenken an Werner Sundermann*. [Iranica Band 25.] Wiesbaden: Harrassowitz, 421–432.
- PAULY, Bernard 1967. 'Fragments sanskrits d'Afghanistan (fouilles de la D.A.F.A.).' *JA* 255: 273–283.
- PINAULT, Georges-Jean 1987. *Sites divers de la région de Koutcha. Épigraphie Koutchéenne*. Paris: Collège de France.
- PINAULT, Georges-Jean 2007. 'Concordance des manuscrits tokhariens du fonds Pelliot.' In: Melanie MALZAHN (ed.) *Instrumenta Tocharica*. Heidelberg: Winter, 163–219.
- SALOMON, Richard 1998. 'Kharoṣṭhī Manuscript Fragments in the Pelliot Collection, Bibliothèque nationale de France.' *Bulletin d'études indiennes* 16: 123–60.
- SALOMON, Richard 1999. *Ancient Buddhist Scrolls from Gandhāra: The British Library Kharoṣṭhī Fragments*. Seattle: University of Washington.



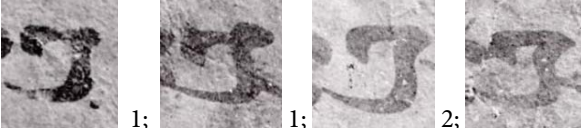
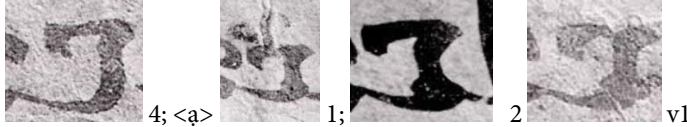

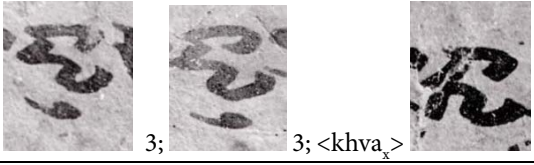

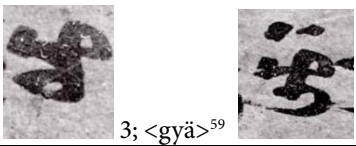
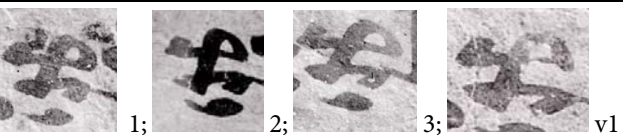

- SALOMON, Richard 2007. 'Gāndhārī in the Worlds of India, Iran, and Central Asia.' *Bulletin of the Asia Institute* 21: 179–192.
- SANDER, Lore 1968. *Paläographisches zu den Sanskrithandschriften der Berliner Turfansammlung*. Wiesbaden: Franz Steiner.
- SANDER, Lore 1986. 'Brāhmī Scripts on the Eastern Silk Roads.' *Studien zur Indologie und Iranistik* 11-12: 159–192.
- SANDER, Lore 1999. 'Early Prakrit and Sanskrit Manuscripts from Xinjiang (Second to Fifth / Sixth Centuries C. E.): Paleography, Literary Evidence, and Their Relation to Buddhist Schools.' In: Jan NATTIER and John R. McRAE (eds.) *Collection of Essays 1993: Buddhism across Boundaries: Chinese Buddhism and the Western Regions*. Sanchung: Fo Guang Shan Foundation for Buddhist & Culture Education, 61–106.
- SCHMIDT, Klaus T. 2001. 'Entzifferung verschollener Schriften und Sprachen: dargestellt am Beispiel der Kučā-Kharoṣṭhī Typ B und des Kučā-Prākritis.' *Göttinger Beiträge zur Asienforschung* 1: 7–35.
- SCHMIDT, Klaus T. 2018. *Nachgelassene Schriften: 1. Ein westtocharisches Ordinationsritual, 2. Eine dritte tocharische Sprache: Lolanisch*. (Bearbeitet und herausgegeben von Stefan Zimmer.) [Monographien zur indischen Archäologie, Kunst und Philologie Band 24.] Bremen: Hempen.
- SCHMITT, Rüdiger 2009. *Die altpersischen Inschriften der Achaimeniden. Editio minor mit deutscher Übersetzung*. Wiesbaden: Reichert.
- SCHWARTZ, Martin 1974. 'Irano-Tocharica.' In: Philippe GIGNOUX and Ahmad TAFAZOLLI (eds.) *Mémorial Jean de Menasce*. [Fondation culturelle iranienne 185.] Louvain: Imprimerie orientaliste, 399–411.
- SIMS-WILLIAMS, Nicholas 2007. *Bactrian Documents from Northern Afghanistan, II: Letters and Buddhist Texts*. [Studies in the Khalili Collection, Volume III.] London: The Nour Foundation.
- SIMS-WILLIAMS, Nicholas forthc. 'Chotano-Sogdica III: Old Khotanese *u*, *-u*, *-ū*.'
- SIMS-WILLIAMS, Ursula 2009. 'The British Library Hoernle Collection, Part 1.' In: Klaus WILLE and Seishi KARASHIMA (eds.) *The British Library Sanskrit Fragments Volume II.1 Texts*. Tokyo: The International Research Institute for Advanced Buddhology Soka University, 1–24.
- STRAUCH, Ingo 2012. 'The Character of the Indian Kharoṣṭhī Script and the 'Sanskrit Revolution': A Writing System between Identity and Assimilation.' In: Alex DE VOOGT and Joachim Friedrich QUACK (eds.) *The Idea of Writing: Writing across Borders*. Leiden: Brill, 131–168.
- WILLE, Klaus 2006. 'Some Recently Identified Sanskrit Fragments from the Stein and Hoernle Collections in the British Library, London (2).' In: Klaus WILLE and Seishi KARASHIMA (eds.) *The British Library Sanskrit Fragments Volume I*. Tokyo: The International Research Institute for Advanced Buddhology Soka University, 27–64.
- WILLE, Klaus 2009. 'Buddhist Sanskrit Sources from Khotan.' In: Seishi KARASHIMA and Klaus WILLE (eds.) *The British Library Sanskrit fragments II.1 Texts*. Tokyo: The International Research Institute for Advanced Buddhology, 25–72.
- YE, Shaoyong 2009. 'The Sanskrit Fragments Or.15009 in the Hoernle Collection: Or.15009.1–50.' In: Seishi KARASHIMA and Klaus WILLE (eds.) *The British Library Sanskrit fragments II.1 Texts*. Tokyo: The International Research Institute for Advanced Buddhology, 105–127.

Open Access statement. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium for non-commercial purposes, provided the original author and source are credited, a link to the CC License is provided, and changes – if any – are indicated.



**APPENDIX: TABLE WITH THE TENTATIVELY READ AKṢARAS**

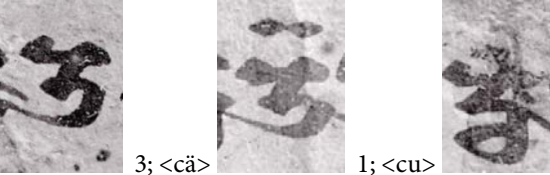
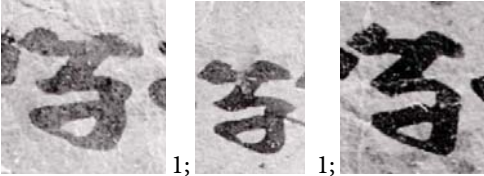


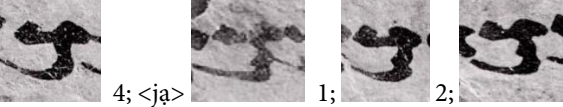
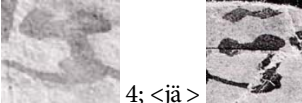


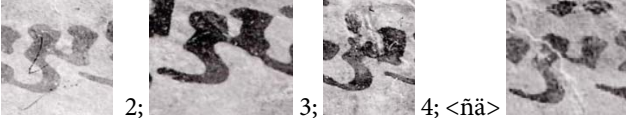



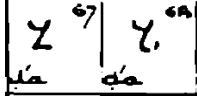

This table contains only the akṣaras that have been tentatively assigned a phonetic value; too uncertain examples are not included. The akṣaras are listed in the *varṇamālā* sequence, with the addition of *z* after the sibilant series. To the right of the image, a reference to the line is given. The paragraph sign (¶) is added at the very end of the table and, whenever available, we add a comparandum from CKD 661, the Khotan Prakrit document, also with a line reference to the right, once supplemented by CKD 843, the new Khotan Prakrit source.

	Attestations	Cf. Khotan Kharoṣṭhī
a	 	 2
khva		 2
ga		
ḡa		 2

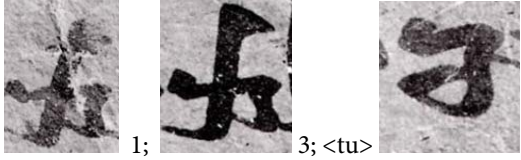
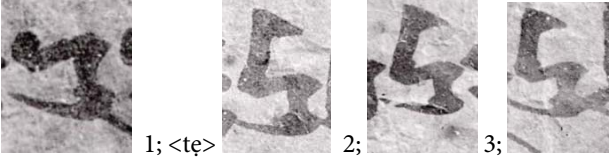




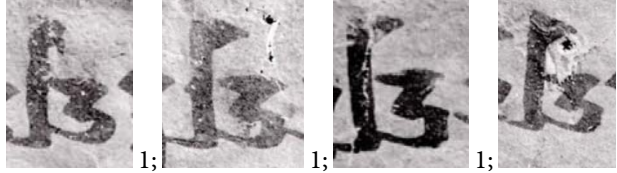

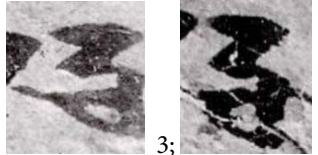


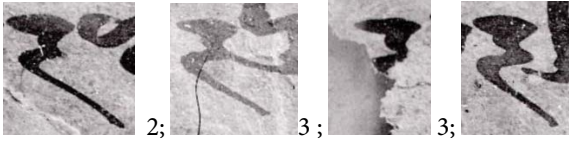

59 The extra stroke at the lower left is here provisionally noted as “<sub>x</sub>”. One might perhaps see in it an *o*-diacritic.

60 Other possibility <rgā>.

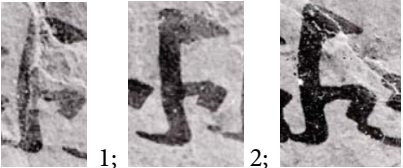
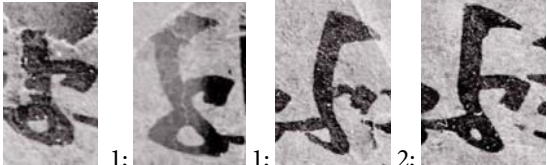



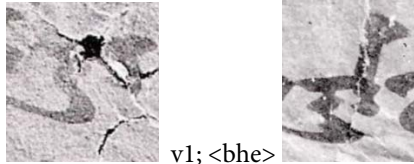
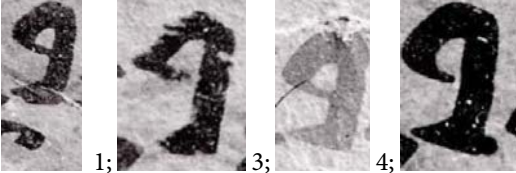
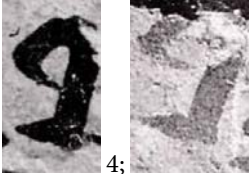



<p>ca</p>	 	 
<p>ja</p>	 	
<p>jza</p>		
<p>ña</p>		
<p>ṭha</p>		
<p>ḍga</p>		 <p>(Boyer et al. 1929)</p>
<p>ṇa</p>		




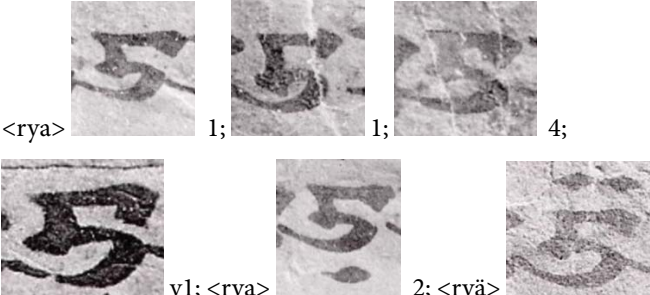
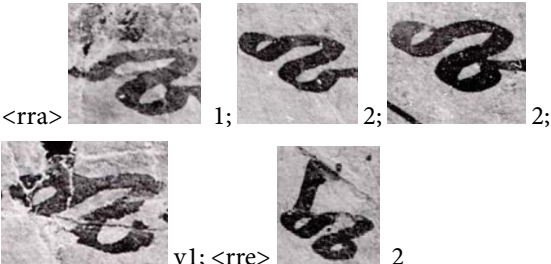
<p>ta</p>	 <p>&lt;ti&gt; 1; 3; &lt;tu&gt; 3; &lt;tə&gt;</p>  <p>1; &lt;tɛ&gt; 2; 3; 3</p>	 <p>&lt;ta&gt; 1;</p>  <p>&lt;ti&gt; 4</p>
<p>da</p>	 <p>&lt;du&gt; 1; 3; 4</p>	 <p>&lt;du&gt; 5</p>
<p>dha</p>	 <p>&lt;dhi&gt; 1; 1; 1; 1;</p>  <p>2; 2; 3; 4; 4;</p>  <p>&lt;dhu&gt; 3; 4</p>	 <p>&lt;dha&gt; 4;</p>  <p>&lt;dhi&gt; 1</p>
<p>na</p>	 <p>&lt;na&gt; 2; 3; 3; 4</p>	 <p>1</p>



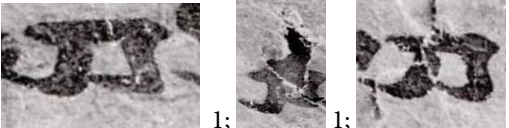





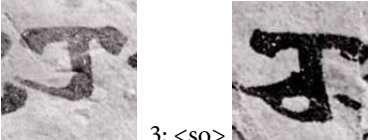

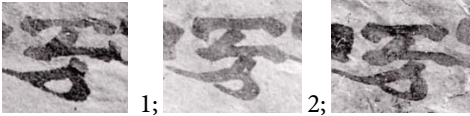
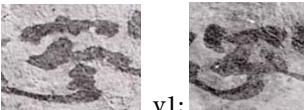

<p>pa</p>	 <p>&lt;pa&gt; 1; 2; 3;</p>  <p>&lt;pu&gt;<sup>60</sup> 1; 1; 2; 2;</p>  <p>3; v1; &lt;pa&gt; 4</p>	 <p>6</p>
<p>pha</p>	 <p>&lt;pho&gt; 4</p>	
<p>bha</p>	 <p>&lt;bha&gt; v1; &lt;bhe&gt; 1</p>	
<p>ma</p>	 <p>&lt;ma&gt; 1; 3; 4; 4;</p>  <p>4; v1</p>	 <p>1</p>

<sup>60</sup> Cf. the discussion in the main article about the difficulty to distinguish between <pu> and <po>.

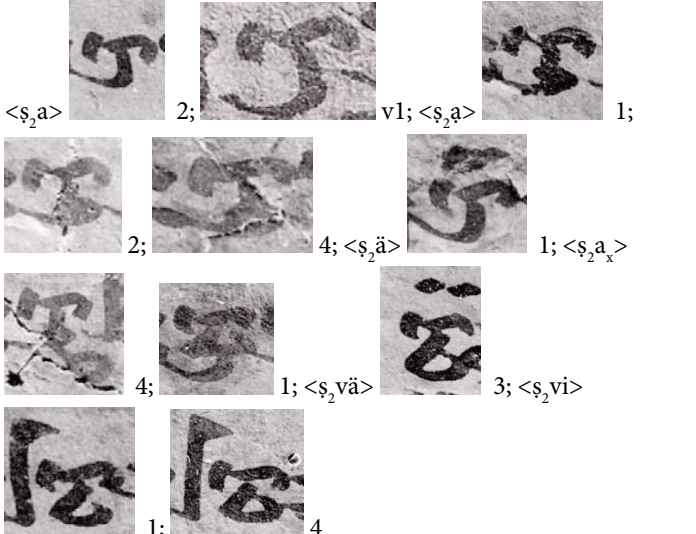

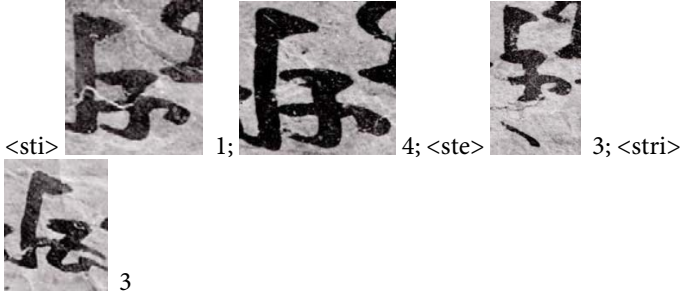

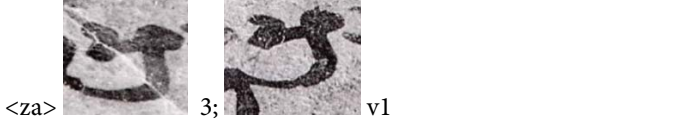

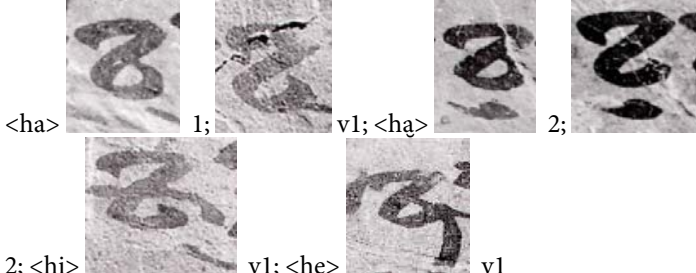




ya	 <p>&lt;yu&gt; 3</p>	
ra	 <p>&lt;ra&gt; 3; 4</p>	 <p>3</p>
rya	 <p>&lt;rya&gt; 1; 1; 4; v1; &lt;rya&gt; 2; &lt;ryä&gt; v1</p>	
rra	 <p>&lt;rra&gt; 1; 2; 2; v1; &lt;rre&gt; 2</p>	
va	 <p>&lt;va&gt; 1; 2; 3</p>	 <p>1</p>



<p>śa (or ya)</p>	 <p>&lt;śa&gt; 1; 1; 2;</p>  <p>2; 4; &lt;śi&gt; 2;</p>  <p>4; v1; &lt;śe&gt; v1</p>	 <p>&lt;śa&gt; 1;</p>  <p>&lt;śi&gt; 9;</p>  <p>&lt;ya&gt; 1</p>
<p>ṣa</p>	 <p>&lt;ṣa&gt; 3; &lt;ṣo&gt; 4</p>	 <p>&lt;ṣti&gt; 3</p>
<p>ṣa</p>	 <p>&lt;ṣo&gt; 1; 2; 4;</p>  <p>&lt;ṣya&gt; v1; 1<sup>61</sup></p>	 <p>3</p>

<sup>61</sup> See the tentative transcription (§ 4.13) for the problems with this akṣara; see also under <ṣ<sub>2</sub>a> below.

<p>ś<sub>2</sub>a<sup>62</sup></p>	 <p>&lt;ś<sub>2</sub>a&gt; 2; v1; &lt;ś<sub>2</sub>a&gt; 1;          2; 4; &lt;ś<sub>2</sub>ä&gt; 1; &lt;ś<sub>2</sub>a<sub>x</sub>&gt;          4; 1; &lt;ś<sub>2</sub>vä&gt; 3; &lt;ś<sub>2</sub>vi&gt;          1; 4</p>	 <p>&lt;ka&gt; 2</p>
<p>sta</p>	 <p>&lt;sti&gt; 1; 4; &lt;ste&gt; 3; &lt;stri&gt;          3</p>	 <p>&lt;sti&gt; 2</p>
<p>za</p>	 <p>&lt;za&gt; 3; v1</p>	 <p>&lt;za&gt; 1</p>
<p>ha</p>	 <p>&lt;ha&gt; 1; v1; &lt;hä&gt; 2;          2; &lt;hi&gt; v1; &lt;he&gt; v1</p>	 <p>&lt;ha&gt; 1</p>
<p>¶</p>	 <p>1 3 4 v1</p>	

<sup>62</sup> As far as the shape is concerned, <ka> would seem a possible alternative.