

Michaël Peyrot

# Interrogative stems in Hittite and Tocharian

**Abstract:** Hittite and Tocharian share an interrogative pronominal stem in *m-* next to the well known Proto-Indo-European interrogative *\*k<sup>wi-</sup>*, *\*k<sup>w</sup>e-*, *\*k<sup>w</sup>o-*. In Tocharian, the *m-*interrogative is especially frequent as a formative element in several interrogative, relative and indefinite stems. In this paper, these stems are investigated in detail, and it is argued that the Tocharian A interrogative stem *ā-* posited by Sieg, Siegling & Schulze in their *Tocharische Grammatik* is a ghost. Although the reconstruction of the *m-*interrogative for the oldest stage of Proto-Indo-European is beyond any doubt, it is difficult to use this Anatolian-Tocharian isogloss as an argument for the phylogenetic structure of the Indo-European family tree since in the other branches the *m-*interrogative may have been lost independently.

**Keywords:** Hittite, Tocharian, interrogative, pronouns, Proto-Indo-European

## 1 Introduction

It is well known that Anatolian and Tocharian share a number of remarkable isoglosses. Some well known examples are listed in Table 1 (see Pinault 2006: 93).

**Table 1:** Anatolian-Tocharian isoglosses

Indo-European	Hittite	Tocharian
* <i>Kost-</i> ‘hunger’	Hitt. <i>kāšt-</i> / <i>kišt-</i>	TB <i>kest</i> , TA <i>kašt</i>
* <i>h<sub>1</sub>eg<sup>wh</sup></i> - ‘drink’	Hitt. <i>eku<sup>zi</sup></i> / <i>aku-</i>	TA/TB <i>yok-</i>
* <i>h<sub>1</sub>erH-</i> ‘wash’	Hitt. <i>ārr<sup>i</sup></i> / <i>arr-</i>	TA <i>yārā-</i>
* <i>mo-</i> , interrogative stem	Hitt. <i>maši-</i> ‘how many’	TA <i>mānt</i> ‘how’

Such isoglosses are potentially relevant for the phylogenetic tree of Proto-Indo-European. Since there are no indications that Anatolian and Tocharian share innovations that would warrant an Anatolian-Tocharian branch, these isoglosses could reflect an earlier phase of Proto-Indo-European than that reflected by the other branches. This would be in line with the scenario argued for by many scholars, but on the basis of various pieces of evidence, namely that, after Anatolian, Tocharian was the second branch to leave the proto-language (cf. e.g. Carling 2005:

Michaël Peyrot, Leiden University; m.peyrot@hum.leidenuniv.nl

Open Access. © Michaël Peyrot published by De Gruyter. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. DOI 10.1515/if-2018-0003.

48–49; Jasanoff 2003: 204; Kim 2007; Schindler apud Jasanoff 2003: 46; Schmidt 1992; Winter 1997).<sup>1</sup>

It is generally accepted that only shared innovations may be used as evidence for a subbranch, but there is virtually no agreement on which innovations are precisely shared by all branches that were left after Anatolian, and supposedly, Tocharian, split off. Thus, in spite of the fact that the idea that Tocharian split off second is widely found, it is not generally accepted. For instance, Malzahn (2016) criticises the use of lexical evidence, as forwarded most prominently by Winter (1997), arguing that lexical arguments often allow alternative interpretations and should therefore not be used.

Isoglosses between Anatolian and Tocharian can obviously only be used when they are paired with a common innovation of the remaining branches. In practice, it is often difficult to prove that such a common innovation has taken place, and thus most of the isoglosses could simply be only shared archaisms, which are not to be used in subgrouping without a corresponding common innovation. For instance, words for ‘hunger’ are not particularly stable (cf. Buck 1949: 332) and the word *\*Kost-* ‘hunger’, preserved in Hittite and Tocharian, could simply have been replaced independently in the other branches in view of the variation in the terms for ‘hunger’.

Grammatical isoglosses suffer from many of the same problems as lexical isoglosses, but clearly we do not expect to find a large number of equivalent expressions for the same grammatical function, nor a degree of instability comparable to that found with ‘hunger’. I will therefore take a closer look at the interrogative stem *\*mo-* shared by Anatolian and Tocharian. Strikingly, this stem is entirely absent from most introductory handbooks (cf. e.g. Beekes 2011: 227–231; Fortson 2004: 130; Meier-Brügger 2003: 227–229). Its existence can nevertheless not be doubted. In Anatolian, this stem is represented by the following forms (cf. in general Kloekhorst 2008):

- Hitt. *māši-* ‘how many; however many’ < *\*mo-s-i-*;
- Hitt. *mān* ‘when > if; how; like’ (possibly also the modal particle *man*) < *\*mó-n-*;
- Hitt. *mānḥanda* ‘just as’ (Kloekhorst 2010);
- Pal. *maš* ‘as much as’.

It is more difficult to decide whether Hitt. *=ma* ‘and, but’ (Rieken 2000) is related, because the meaning is quite different. A possible path of semantic development

---

<sup>1</sup> Studies in computational cladistics, as for instance Ringe, Warnow & Taylor 2002, also often have Tocharian split off second, but since the method implies that the result is always a binary branching tree, these studies cannot be compared one-to-one to the isolated arguments in the literature listed here.

would be ‘how’ > ‘as, as well as’ > ‘and’ > ‘but’. For this development one might adduce as a parallel PIE *\*k<sup>w</sup>e* ‘and’ < *\*k<sup>w</sup>eh<sub>1</sub>* ‘how’, although the latter derivation is truly hypothetical itself. In Tocharian, the clearest cognates are TA *mānt* ‘how’ and TB *mäksu* ‘which’, but, as I will argue, the *\*mo-* interrogative is found in many more stems and forms.

Below, I will discuss the following Tocharian interrogative, relative and indefinite elements: Tocharian B *mäksu* ‘which’ (Section 2), Tocharian B *intsu* and Tocharian A *äntsam* ‘which’ (Sections 3 and 4), the Tocharian A *ā*-interrogative (Section 4), the newly discovered Tocharian B *mantsu* (Section 5), Tocharian B *miñcek*, of which the meaning is so far unknown (Section 6), Tocharian B *mant* ‘so’ and Tocharian A *mānt* ‘how’ (Section 7), Tocharian B *manta* ‘not any’ (Section 8), and Tocharian A *māñcaṃ*, of which the meaning needs to be corrected (Section 9). I will then come back to the reconstruction of the *m*-interrogative for Proto-Indo-European and its value for the phylogenetic tree (Section 10).

## 2 Tocharian B *mäksu* ‘which’

Alongside the most important and most frequent interrogative and relative *k<sub>u</sub>se*, another interrogative and relative is found in Tocharian B: *mäksu*. While *k<sub>u</sub>se*, which is cognate with Tocharian A *kus*, can be translated with ‘who, what’, *mäksu* means ‘which’. The differences between *k<sub>u</sub>se* and *mäksu* can further be characterised as follows.

*k<sub>u</sub>se* is a pronoun, and normally not used as a nominal modifier. Use as a nominal modifier is found, but seems to result from calquing. An illustrative case is the phrase *k<sub>u</sub>se šamāne* ‘which monk’, which is frequent in the Prātimokṣasūtra, together with its late variant *se šamāne* with reduction of the initial cluster *k<sub>u</sub>s<sup>o</sup>*. As with many other syntactic peculiarities of the Prātimokṣasūtra (cf. e.g. Peyrot 2013: 333, 335), this expression is a calque of the original Sanskrit, i.e. *yaḥ punar bhikṣuḥ*.<sup>2</sup> Modifier use is also found in some short phrases such as *k<sub>u</sub>ce šarmtsa* ‘for what reason’. Otherwise, modifier use does not occur and normally Tocharian B *k<sub>u</sub>se* is used as a pronoun and does not show agreement with noun phrases. Thus it is only inflected for case, and not for gender or number (although in Tocharian A a rare nom. pl. *k<sub>u</sub>ce* is attested; cf. Carling 2009: 156–157). Its inflexion is nom. *k<sub>u</sub>se*, obl. *k<sub>u</sub>ce*,<sup>3</sup> gen. *ket*.

<sup>2</sup> *k<sub>u</sub>se no šamāne*, with the expected rendering of *punar* by *no*, is rare, but found in e.g. B 333 b1.

<sup>3</sup> Forms with secondary case suffixes (formed from the oblique) are also attested.

*mäksu* is much less frequent than *k<sub>u</sub>se*, and predominantly used as a nominal modifier. Accordingly, we do find alongside the usual *k<sub>u</sub>se šamāne* ‘which monk’, the actually expected *mäksu no šamāne*. Predicative use of *mäksu* is rare, and the meaning then seems to be ‘of what kind’, e.g. THT 1335d a6 *wrašše camel mäksu* ‘what is birth in water?’, i.e. ‘what kind of birth is birth in water?’. In agreement with the following noun, *mäksu* is fully inflected for gender, case and number, and it follows the inflexion of the anaphoric demonstrative pronoun *su*.

The etymology of *mäksu* has been discussed in detail by Hackstein (2004: 280–283), who derives it from a univerbated phrase *\*mos k<sup>wi</sup>is so + u*, literally ‘which one is this who this is?’. According to Hackstein (2004), in this phrase *\*mos* is the interrogative element, i.e. ‘which one is this?’; *\*k<sup>wi</sup>is* is a “clitic” relative marker, ‘who’; and *\*so + u* is a demonstrative pronoun in the function of a predicate nominal, ‘this is’. Structurally, this phrase would be comparable to French *qu’est-ce que* ‘what’, from earlier ‘what is it that ...’. Pinault (2010: 359) accepts Hackstein’s identification of the three basic elements *\*mos*, *\*k<sup>wi</sup>is* and *\*so + u*, but is not convinced that the first element *\*mos* should really have been fully inflected. He considers it likely that the first element may have been uninflected at the time the phrase was formed and that it was no more than a connective element, since the following *\*k<sup>wi</sup>-* already expressed the interrogative value.

In my view, Pinault is right in noting that the first element, reconstructed as *\*mos* by Hackstein, need not have been an inflected nominative singular. I would extend this reasoning to include also the second element *\*k<sup>wi</sup>i*: that *\*k<sup>wi</sup>i* was inflected even though it was followed by an inflected demonstrative is impossible to prove, and the fact that a demonstrative was added in the first place can be neatly explained under the assumption that *\*k<sup>wi</sup>i* itself was no longer inflected. There cannot be the slightest doubt that if Tocharian inherited a simplified interrogative of the type nom. *\*k<sup>wi</sup>-s*, acc. *\*k<sup>wi</sup>-m*, the inflexion would have disappeared completely through the loss of final *-s* and *-m*, yielding an uninflected *\*k<sup>w</sup>ə*. In fact, this uninflected *\*k<sup>w</sup>ə* is exactly what is needed to derive the attested Tocharian A and B interrogatives.

Apart from the fact that Hackstein’s *\*mos* and *\*k<sup>wi</sup>is* were probably rather uninflected *\*mo* and *\*k<sup>wi</sup>i*, his analysis must be essentially correct. It is true that there are two elements that may be interrogative or relative, both *\*mo* and *\*k<sup>wi</sup>i*, but I see no reason to assume, with Pinault, that *\*k<sup>wi</sup>i* necessarily contributed the interrogative value. After all, relative use of Tocharian B *k<sub>u</sub>se < \*k<sup>wi</sup>-so* is well attested, and Hackstein adduces ample evidence for compound interrogatives that contain an interrogative as the first element, and a relative element as the second. This relative element, then, becomes the connective element in the course of the univerbation. Fronting of interrogatives is frequent cross-linguistically, and the

order of elements assumed by Hackstein fits the rules and tendencies of Proto-Indo-European and Tocharian grammar.

Thus, my reconstruction of Tocharian B *mäksu* is, in Proto-Tocharian terms, *\*mā-k<sup>w</sup>ə-sə-w*. There is no reason to assume that exactly this phrase existed in Proto-Indo-European, but for the sake of clarity a mechanical reconstruction may in this case, as well as in other cases to follow below, be helpful. Such a reconstruction would be *\*mo-k<sup>w</sup>i-so-u*, in which *\*mo* is the interrogative element, *\*k<sup>w</sup>i* is the relative-connective element, *so* is the demonstrative, and *u* is a particle.<sup>4</sup>

As pointed out by Hackstein (2004: 280–281), Tocharian *\*mā-* would in theory allow for several reconstructions:

1. *\*mu-*;
2. *\*me-* and *\*mi-*, if the expected intermediate *\*mā-* did not develop, as otherwise regularly, to Tocharian B *mi-*; and
3. *\*mo-*.<sup>5</sup>

The reconstruction *\*mo-*, which has already been used above, is by far the most likely in view of the Anatolian cognates, but does need the assumption of a special development of *\*o* to *\*ə*, whereas the normal development is PIE *\*o* > PT *\*e* > TB *\*e*, TA *\*a*. This assumption seems unproblematic, since especially in the pronominal system parallels are frequent, for instance in the anaphoric pronoun TA *sām* < *\*sə + m*, TB *su* < *\*sə + w*, both ultimately from PIE *\*so*. Although the exact conditions for this reduction have not been established yet, and reflexes of PT *\*sə* are found next to reflexes of the regular PT *\*se*, e.g. proximal TB *se*, TA *sas*, it is in all probability due to a weak accentual position in the phrase.

### 3 Tocharian B *intsu* and Tocharian A *äntsam*

Next to Tocharian B *mäksu*, a further interrogative *intsu* is found. While *mäksu* has no direct cognate in Tocharian A, *intsu* is obviously related to Tocharian A *äntsam*. Tocharian B *intsu* is less frequent even than *mäksu*, but as far as can be

<sup>4</sup> I will not discuss the etymology or function of *\*u* further here. Suffice it to note that according to Stumpff's analysis (1971), it distinguishes the Tocharian B anaphoric pronoun *su* from the proximal demonstrative *se*.

<sup>5</sup> Hilmarsson (1987: 42) derives *\*mā-* from *\*me-* as in Arm. *merj*, Gk. μέχρι 'until' < *\*mé-ǵ<sup>h</sup>sr-i*. As far as the form is concerned, this derivation is possible, but it lacks any functional or semantic basis.

established, it is parallel to *k<sub>u</sub>se* in being inflected for case only,<sup>6</sup> unlike *mäksu*, which is inflected for case, number and gender (see Table 2).

**Table 2:** Inflexion of TB *intsu*

	nom.	obl.	perl.
TB <i>intsu</i>	<i>intsu</i>	<i>iñcew</i>	<i>iñcewsa</i>
cf. <i>k<sub>u</sub>se</i>	<i>k<sub>u</sub>se</i>	<i>k<sub>u</sub>ce</i>	<i>k<sub>u</sub>cesa</i>

Otherwise, the difference between *intsu* and *mäksu* is difficult to establish. Both mean ‘which, which kind of’, and can be interrogative as well as relative. As noted above, *mäksu* is much more frequent than *intsu*. Further, *intsu* is also found in indefinite use, in which case it is reduplicated or followed by the particles *ra tsa*, while *mäksu* is only very rarely indefinite.

In Tocharian A, the situation is different. Tocharian A *äntsam* is best attested in the masculine singular inflexion, but feminine, neuter and plural forms are found as well (see Table 3).

**Table 3:** Inflexion of TA *äntsam*

TA <i>äntsam</i> , m. sg.:	nom. sg. m.	<i>äntsam</i>	obl. sg. m.	<i>äñcaṃ</i>	gen. sg. m.	<i>äñcanik</i>
other:	obl. sg. f.	<i>äntām</i>	gen. pl. m.	<i>äñcesni</i>		<i>äntanne</i> <sup>7</sup>

Although the paradigm is not complete, the attestation of feminine, plural and neuter forms proves that *äntsam* was inflected for case, gender and number. In this respect, it agrees with Tocharian B *mäksu*, not with *intsu*. However, since Tocharian A *äntsam* does not have a competitor like TB *mäksu*, it is possible that the original Proto-Tocharian situation can be compared to Tocharian B, while in Tocharian A *äntsam* took over the functions that were originally reserved for the equivalent of Tocharian B *mäksu*. The functional differences between the Tochar-

<sup>6</sup> Obviously, as pointed out by an anonymous reviewer, a caveat is needed here: *intsu* is so rare that it is hard to exclude that plural and feminine forms are not attested by chance. This observation is therefore true only as far as the evidence goes at this point.

<sup>7</sup> In origin, this seems to be the basic nom.-obl. neuter *äntam* with an element *-ne*. Though synchronically in Tocharian A *-ne* often functions as a relative marker, *äntanne* is only found with an additional marker *-ne* to mean ‘where’; apparently, the function of the first *-ne* has been bleached.

ian B interrogatives are difficult to establish in detail, but the original distribution may have been:

- *k<sub>u</sub>se*, default ‘who, what’, pronominal, not used as modifier;
- *intsu* ‘which, of which kind’, pronominal, not used as modifier;
- *mäksu* ‘which, of which kind’, modifier.

The initial *ä-* of Tocharian A *äntsam* is without any parallel in Tocharian: apart from this interrogative stem, initial *ä-* is not found in Tocharian A, and in Tocharian B it is not found at all. This situation is reminiscent of the exceptional phonological structure that is more often found in interrogatives, pronouns and occasionally other function words cross-linguistically; compare for instance Turkish *ne* ‘what’, the only inherited stem with initial *n-*; Turkish *hangi* ‘which’, the only inherited word with *h-*; Russ. *эmo* ‘this’, the only inherited stem with initial *э-*; or even Eng. *this* /ðɪs/, together with a small group of mostly demonstratives and personal pronouns unique in having initial *ð-*. Since the initial *ä-* of Tocharian A is unique, no regular reconstruction of this interrogative stem is possible for Proto-Tocharian.

Tocharian B initial *i-* is not without parallels, but nevertheless not as frequent as one might think. In rare cases, initial *y-* before a consonant may alternate with *i-* in archaic Tocharian B, e.g. *yšāmma* ‘among men’, with variant *išāmma*. Regular initial *i-* not from *y-* is rare, and in all certain cases it goes back to a double *i-* or *y-* element, for instance *ike* ‘place’ < \**yäyke* < \**wäyke* < \**ueiiko-*. Even synchronically, a phonological transcription /*yäyke*/ seems to be required. Obviously, \**yäy-* is not a possible reconstruction for Tocharian A *ä-*. By far the most straightforward reconstruction of TA *än-*, TB *in-* is \**ən-*. This \**ən-* must in Proto-Tocharian have been as rare and marked as it is synchronically in Tocharian A, and this must be the explanation of the unique replacement with *in-* in Tocharian B.

Although the normal reflex of Proto-Indo-European initial \**h<sub>2</sub>-* is Proto-Tocharian \**en-*, a syllabic nasal seems to be the only possible source of Proto-Tocharian \**ən-*. For the reconstruction with a syllabic nasal, compare in particular the regular variant in TB *ente* ‘where, when’, which seems to have *en-* as a variant of *in-*.<sup>8</sup>

For the origin of an interrogative syllabic nasal one might think of a zero grade \**h<sub>2</sub>-* of the negation \**ne*. The semantics might be interpreted as being parallel to phrases like Eng. *isn't it?* or French *n'est-ce pas?* However, the initial position of the interrogative element would be very surprising, and is not in line with expected syntax. Moreover, the syllabic nasal certainly does not need to be dental \**h<sub>2</sub>-*: since

<sup>8</sup> Note that it must be an old variant, not a secondary development within Tocharian B: there are no parallels for a development *i-* > *e-*.

\*ən is found always before *t* (~ *c*) or *s*, original \*ṇ̥ is possible as well. For \**mt* > *nt*, compare for instance TB *kante* ‘100’ < \**dḱmtóm*. The cluster \**ms* eventually developed into \**s*, probably through an intermediary \**ns* (cf. e.g. Peyrot 2013: 423), but obviously the nasal, \**n* at that stage, may have been reintroduced on the basis of forms with \**t* or \**c* at any time.

I would like to propose that the interrogative element \**Ṇ* is to be identified as a reduced form of interrogative \**mo-*. The most important part of this reduction, of \**mo-* to \**mə-*, is assured for Tocharian by Tocharian B *mäksu* with \**mə-* < \**mo-* (see above). Once this reduced variant \**mə-* had come about, it may, perhaps in pretonic position, have been further shortened to \**m-*, which developed into \**n-*, \**ən-* > TA *än-*, TB *in-*.<sup>9</sup> Thus, I think that Tocharian B *intsu* derives from Proto-Tocharian \**ən-sə-w*, and Tocharian A *äntsam* from \**ən-se-nə*; in both languages, an epenthetic *t* developed in the cluster *-ns-*. A mechanical reconstruction for Proto-Indo-European would be (with the particle \**nə* possibly from PIE \**nu*) as shown in Table 4.

**Table 4:** Reconstruction of TB *intsu* and TA *äntsam*

TB <i>intsu</i>	<	PT * <i>ən-sə-w</i>	<	* <i>mo-so-u</i>
TA <i>äntsam</i>	<	PT * <i>ən-se-nə</i>	<	* <i>mo-so-nu</i>

As with *mäksu*, the Proto-Indo-European reconstruction \**mo-* is tentative; another vowel would also be possible. For the structure of Proto-Tocharian \**ən-sə-w* and \**ən-se-nə*, compare TB *k<sub>u</sub>se*, TA *kus* < PT \**k<sup>w</sup>ə-se*.<sup>10</sup>

<sup>9</sup> The exact mechanism behind the split of \**mo-* > \**mə-* into \**mə-* on the one hand and \**Ṇ* on the other is unclear. Obviously, one thinks of an effect of the accent or the phonological structure of the initial of the following element.

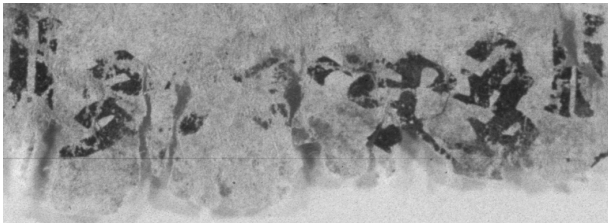
<sup>10</sup> A similar derivation of \**ən-* from \**ṇ-* is proposed by Hilmarsson (1987: 42–43). The difference is that he identifies this \**ṇ-* as the zero grade of \**me*, comparing Gr. μέχρι ‘until’ < \**me-* and ἄχρι ‘until, to the utmost’ < \**ṇ-*. This explanation fails on the semantics, see fn. 5. Also, he reconstructs PT \**e-* with multiple independent rounds of weakening to TA *ä-*, TB *i-* etc., which is improbable. Functionally weak is Kortlandt’s proposal to derive the initial of TB *ente*, TA *äntāne* ‘when’ from PIE relative \**io-* (1983: 321); these words are clearly interrogative in origin. Pedersen (1941: 126) unconvincingly derives *nt-* < \**mnt-* < \**mənt* (as in TA *mänt* ‘how’).

## 4 The Tocharian A $\bar{a}$ -interrogative

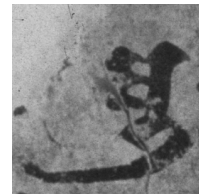
According to Sieg, Siegling & Schulze (1931: 182) there exists a Tocharian A interrogative stem  $\bar{a}$ -, attested two times: A 217 b7  $\bar{a}\eta$  ‘who (obl.)’ and A 463 b5  $\bar{a}\acute{s}\acute{s}i$  ‘where then’, as the translation of Skt. *kuṭra nu* ‘where then’. The existence of this interrogative stem is generally accepted, and not contested as far as I know; cf. e.g. Carling 2009: 23a, 35a; TEB: 166. The second attestation  $\bar{a}\acute{s}\acute{s}i$  clearly contains  $\bar{a}\acute{s}\acute{s}i$  as the second element, so that it needs to be analysed as  $\bar{a} \acute{s}\acute{s}i$ : here  $\bar{a}$  corresponds to Skt. *kuṭra*, so that it must mean ‘where’, and  $\acute{s}\acute{s}i$  corresponds to Skt. *nu*, meaning ‘then’. The morphological relationship between the obl. sg.  $\bar{a}\eta$  ‘who’ and the local interrogative  $\bar{a}$  ‘where’ is unclear and cannot be compared to any known pattern. Obviously, the isolation of this stem in Tocharian A, which has no cognate in Tocharian B and no obvious Indo-European etymology, is puzzling on any account.

In my view, A 463 b5  $\bar{a}\acute{s}\acute{s}i$  ‘where then’ has to be read differently. I also think that A 217 b7  $\bar{a}\eta$  ‘who’, which is then fully isolated, is not read correctly.

A 463 b5<sup>11</sup>  $\bar{a}\acute{s}\acute{s}i$  is attested in a bilingual manuscript with selected key terms from the *Dirghāgama*. The problem with the reading is that the long  $\bar{a}$  is not formed in the usual way, so that it simply cannot be a normal  $\bar{a}$ , cf. Figure 1 (a). A normal  $\bar{a}$  from the same manuscript is given for comparison in Figure 1 (b).



(a) A 463 b5 || *ku[tra nu] ♦  $[\bar{a}]\acute{s}\acute{s}i$*  ||



(b) A 465 a3  $\bar{a}$

**Fig. 1:** Comparison of alleged  $\langle[\bar{a}]\acute{s}\acute{s}i\rangle$  with regular  $\langle\bar{a}\rangle$

The normal  $\langle\bar{a}\rangle$ , to the right, has a small nob on the upper right corner, and the length stroke is attached at the bottom: a long bend a little to the right and then back to the left. Although the akṣara in A 463 b5 is a little damaged, it is clear that the body is identical, but there is a marked difference on the upper right corner,

<sup>11</sup> After Couvreur 1969: 163 recto and verso are reversed compared to Sieg & Siegling 1921: 249.

where no small nob is found, but a length stroke. A length stroke on top of initial ⟨a⟩ instead of at the bottom is simply not possible; there is no such variation with the akṣara ⟨ā⟩. Moreover, the bottom of the akṣara in A 463 b5 is also peculiar. In the lower right corner, the diagonal line is at least two times longer than with the normal ⟨ā⟩. These traces are perfectly compatible with the consonantal characters ⟨nt⟩, but not with the regular length stroke.

Since the normal word for ‘where’ is *āntā* in Tocharian A, I suppose that the peculiar character here is a special way of writing *āntā*. More specifically, I think that this akṣara combines in the typical Tocharian way “two akṣaras in one”. Examples of this spelling peculiarity have been known since the decipherment of Tocharian. Sieg & Siegling (1908: 921) give as examples: ⟨<sub>u</sub>pā⟩, ⟨<sub>u</sub>śā⟩, ⟨<sub>u</sub>pre⟩, ⟨<sub>n</sub>na⟩, ⟨<sub>w</sub>īnā⟩, ⟨<sub>u</sub>le⟩, ⟨<sub>u</sub>ya⟩, ⟨<sub>u</sub>pā⟩, ⟨<sub>u</sub>re⟩. This list is not exhaustive, but it is representative. Perhaps the only important addition should be A 371 a2 ⟨<sub>o</sub>ka⟩ = *okā*. In this case, it was the initial *ā*- of *āntā* ‘where’ that became “nonsyllabic”: the above observations about this akṣara point to ⟨<sub>a</sub>ntā⟩ (since the sequence *nt* cannot be read with certainty, the transliteration is ⟨<sub>a</sub>[nt]ā⟩). Apparently, the expected vowel ⟨ä⟩ is not indicated. As a parallel one could adduce A 371 b5 *antāne* ‘where (relative)’ for *āntāne*, but it seems more likely that there was no need to distinguish nonsyllabic *ā* from *a*; if the spelling ⟨<sub>a</sub>ntā⟩ has any phonetic reality, it is probably simply *ntā*, with a heavily reduced vowel before it, i.e. [ǎnta], [ʔǎnta] or [ʔnta].<sup>12</sup>

The reading proposed here is based on the assumption of a so far unattested nonsyllabic *a*, but, as pointed out above, it has parallels within the system, since nonsyllabic *u*, *i* and *o* were already known. It should be stressed that the akṣara in A 463 b5 cannot possibly be read as normal ⟨ā⟩, and that the manuscript is otherwise completely regular, and written in careful calligraphy. Any reading that does justice to the actually attested akṣara shape needs the assumption of a special spelling technique.

The second attestation is found in A 217, where a strophe from the Udānavarga is cited, cf. ex. (1).

(1) [ma]tt(a)k nū yn(e)ś [karso](rās) āṃ ṣarpñim

A 217 b7 (Sieg & Siegling 1933: 170)

<sup>12</sup> Although the data allow no definitive interpretation, the first option has my preference, since Tocharian tolerates no sequences of heterosyllabic vowels, which are always contracted. Also, initial *y*- alternates with *i*- in archaic Tocharian B texts, and in verse *o*- can be reduced to *w*-. It seems to me that Tocharian had no glottal stop, [ʔ]. Koller (2016) has argued that Tocharian A has an automatic initial glottal stop. I agree with many of his observations, but I think that the assumption of an initial glottal stop is not necessary, and unlikely in view of the arguments just given.

*sayam abhiññāya kam uddiseyyam* Pāli Dhṛp 353d (= Vin. I.8.19–20)  
*svayam hy abhijñāya kam uddīṣeyam*  
 ‘having learned [it] myself, to whom should I point [as teacher]?’

Skt. Uv. 21.1d

The correspondence is clear: *ṣārpñim* ‘I should point out’ is the equivalent of Pāli *uddiseyyam*, Skt. *uddīṣeyam*, and what has been read as *ām* corresponds to Pāli *kam*, Skt. *kam* ‘who’. Again I compare the relevant akṣara with a regular ⟨ā⟩ from the same manuscript, cf. Figure 2.



(a) A 217 b7 *ām*



(b) A 218 b6 *ā*

**Fig. 2:** Comparison of alleged ⟨*ām*⟩ with regular ⟨*ā*-⟩

Although the case is not as clear as for A 463 b5, here, too, the reading is not evident: compared to the regular ⟨*ā*⟩ to the right,

1. the upper part of the length stroke is too far down and not fully straight;
2. the white space in the length stroke is too large;
3. the point to the right of the length stroke is different.

The normal word for ‘who’ is *kus*, obl. sg. *kuc* or *āntsam*, obl. sg. *āñcam*. Although the problems with the reading are subtle in this case, I propose to read ⟨<sub>a</sub>[ñ](c)am⟩ or perhaps ⟨<sub>a</sub>{ṃ}[c]am⟩ with “double use” of ⟨*m*⟩. The first option would seem to make more sense a priori, while the second appears to be more likely in view of the black and white photo (the original manuscript is missing), although that photo hardly allows verification of the reading in such detail. As with ⟨<sub>a</sub>ntā⟩ above, I suppose that <sub>a</sub> can stand for a reduced *ā*-, i.e. ⟨<sub>a</sub>[ñ](c)am⟩ for *āñcam*.<sup>13</sup>

In order to illustrate my interpretation of the two damaged akṣaras in question, I have made reconstructions of how they could have looked originally (see Figure 3, p. 76).

Thus, since A 463 b5 *ā ṣṣi* simply is not read correctly, the assumption of a new type of “Tocharian” spelling with initial <sub>a</sub> is unavoidable. This initial <sub>a</sub> is perhaps a very short shwa, [ə], or otherwise perhaps [ʔ] or [ʔə]; i.e. *a<sub>n</sub>tā* [ənta] etc. Unlike for

<sup>13</sup> There is a second occurrence of the phrase *a<sub>n</sub>cam ṣārpñim* in the line (Sieg & Siegling 1933: 170), but the manuscript is so damaged before *ṣārp(ñ)i(m)* there that it is of no use.

(a) ⟨<sub>a</sub>ntā⟩(b) ⟨<sub>a</sub>{ṁ}caṁ⟩(c) ⟨<sub>a</sub>ñcaṁ⟩**Fig. 3:** Reconstructed akṣaras for illustration

*i* and *u*, the position of the tongue was probably not relevant for reduced *ä* vs. *a*, so that reduced *ä* could be written *a*. The new reading of A 463 b5 *ā =śśi* as *a<sub>a</sub>ntā =śśi* makes the difficult reading of A 217 b7 *āṁ* completely isolated. Although my new proposal *a<sub>a</sub>{ṁ}[c]aṁ* or *a<sub>a</sub>[ñ](c)aṁ* is less obvious than the reading *a<sub>a</sub>ntā*, it is backed up by A 463 b5 *a<sub>a</sub>ntā* and morphologically much simpler than *āṁ*. My conclusion is that there was no interrogative stem *ā-* in Tocharian A.

## 5 Tocharian B *mantsu*

In the Udānavarga bilinguals, which contain the entire Sanskrit Udānavarga interrupted by fully literal Tocharian translations of each half pāda, I have identified a pronominal stem of Tocharian B that was previously entirely unknown: *mantsu*. As it happens, the passage where *mantsu* (written *maṁtsu*) is found has a variant in the Sanskrit text as well as in the Tocharian B rendering, cf. ex. (2). Variants in the Sanskrit text are frequent, but usually the Tocharian B text conforms to the bold-print “vulgata” of Bernhard (1965), that is, Schmithausen’s recension I (1970: 58–59). The variant of *mantsu* is *k<sub>u</sub>se ksa* ‘whoever’, and the Sanskrit variants are *kaś cit* ‘anyone’ and *kiṁ cit* ‘anything’.

- (2) */// sa mā treñcitra maṁtsu ///* IOL Toch 838 a2  
*/// (k<sub>u</sub>se) ks(a) :* SHT 7354 b2  
*yayā nābhiṣajet kaś cit*  
 ‘through which nobody is hurt’ Skt. Uv. 33.17c; cf. Hahn 2007: 148  
*yayā nābhiṣajet kiñ cit*  
 ‘through which nothing is hurt’ Skt. Uv. 33.17c (variant)

Since *k<sub>u</sub>se ksa* is known to correspond to *kaś cit*, I have suggested that *k<sub>u</sub>se ksa* translates *kaś cit* while *mantsu* translates *kiṁ cit* (Peyrot 2016: 189). However, this interpretation of the variants is only tentative, and it must be kept in mind that *k<sub>u</sub>se* has no neuter counterpart and can be used for persons and things alike.

Consequently, it is very likely that the same is true of *mantsu*. It is conceivable that *mantsu* means ‘anyone; anything’, as different from *k<sub>u</sub>se ksa*, which is known to mean ‘whoever’. However, the exact meaning of *mantsu* cannot be established on the basis of this bilingual; the only thing that is clear is that *mantsu* was compatible here with the negative indefinite meaning shared by both *kaś cit* and *kiṃ cit* in this passage.

If the newly discovered, but apparently exceedingly rare pronoun *mantsu* was inflected for case like *k<sub>u</sub>se*, we would expect an obl. *mañce<sub>u</sub>* and a gen. *mañcpi* (cf. *mākcpī* to *mäksu*).<sup>14</sup> If it was inflected for case, gender and number, like *mäksu*, we would in addition expect a nom. sg. f. *mantsau* (with initial accent), and a nom. pl. m. *mañcai* (cf. *mākcai* to *mäksu*),<sup>15</sup> etc. As can be expected with a pronoun stem that has been overlooked so far, the yield of possible further attestations is meagre. An instance may be the one in ex. (3).

(3) /// *krent pelaiykne ce mañce<sub>u</sub>* 27

B 138 b6

For this passage, the editors Sieg & Siegling (1953: 71) suggested reading *mañce<sub>u</sub>* as a sandhi variant of *mānt ce<sub>u</sub>*: “wohl für *ce mānt ce<sub>u</sub>*”. The problem with this suggestion is that the sentence would then have two direct objects in the oblique case. This is not impossible a priori, but it is difficult to see what the preceding part of the sentence might have been in this case: it is likely that *krent pelaiykne ce* is one noun phrase, ‘this good law’, for *ce krent pelaiykne* with standard word order. Following Sieg & Siegling, Pinault (2008: 387) suggests that *mañce<sub>u</sub>* stands for *mānt ce<sub>u</sub>*, i.e. ‘so him’, ‘so it’, and is to be interpreted as “tel, de cette sorte”. I see three problems with this analysis:

1. the syntax of *mānt* would be without parallel;
2. on the evidence of the Udānavarga bilingual cited above, the corresponding nom. sg. m. *mantsu* exists and means *kaś cit* ‘somebody’, ‘anybody’, not ‘so he’ or ‘such’;
3. sandhi of the type assumed for *mañce<sub>u</sub>* from *mānt ce<sub>u</sub>* is difficult to exclude but would be exceptional, and it should be kept in mind that B 138 is an archaic text.

<sup>14</sup> As pointed out to me by an anonymous reviewer, the gen. sg. variant *mākcepi* recorded by Adams (2013: 485) belongs rather to *makte* ‘self’. In any case, the attestation of *mākcepi* in B 78 b4 given by Adams certainly is the gen. sg. of *makte* ‘self’, not of *mäksu* ‘which’.

<sup>15</sup> Again, the anonymous reviewer points out to me that the nom. pl. variant *mākci* recorded by Adams (2013: 485) rather belongs to *makte* ‘self’.

I therefore tentatively propose translating *māñce<sub>u</sub>* in ex. (3) with ‘somehow’:<sup>16</sup>

(3) */// krent pelaiykne ce māñce<sub>u</sub> 27*  
 ‘... this good law somehow. 27’

B 138 b6

In my view, the formation of *mantsu* is relatively straightforward (Peyrot 2016: 204). It shares its first element with *mäksu* ‘which’, and the second element is identical to *intsu* (except for the *i-*, which must be secondary, see above). Thus, *mantsu* reflects Proto-Tocharian *\*mā-ən-sə-w*, and is parallel to *mäksu* < *\*mā-k<sup>w</sup>ə-sə-w*, the difference being the second element *\*ən* in the preform of *mantsu* vs. the second element *\*k<sup>w</sup>ə* in the preform of *mäksu*.<sup>17</sup>

## 6 Tocharian B *mīñcek*

A form that might be related to *mantsu* is Tocharian B *mīñcek*, which is so far limited to late texts, where it occurs twice.<sup>18</sup> The first attestation shown in ex. (4) is from a letter from the French collection.

(4) *kokale(ci m)iñcek skāya-ñ kārñā(ts)i*  
 ‘Les charretiers feront l’effort pour moi d’acheter justement cette sorte de chose.’  
 PK LC 10 a13–14; Pinault 2008: 392<sup>19</sup>

*(m)iñcek* is here rendered as ‘justement cette sorte de chose’, and the reason for this translation has been given above: Pinault (2008: 387) analyses *mīñcek* as *mant* ‘so’ + *ce* ‘this’ + *k*, emphatic particle, with accent shift onto the second syllable because of the particle *k* < *\*-kə*, i.e. *mānt cék* > *māñcék*, and with subsequent colouring of the shwa by the neighbouring palatal *ñ*, which is characteristic of the late language. Although this translation works in the passage just cited, it is difficult to use for the second occurrence shown in ex. (5).

(5) *mīñcek šeme preke palsko yolmuwa-š* 59BTA7–50 D7046 1.3

<sup>16</sup> A further instance seems to be found in THT 2377t a3, where *māñce* can be read with certainty, possibly followed by [w]. Unfortunately the rest of the line is too obscure to allow an interpretation at this point.

<sup>17</sup> Obviously, as pointed out by an anonymous reviewer, there is also a difference in accent: *māntsu* vs. *mäksú*. The meaning of this accentual difference is difficult to interpret.

<sup>18</sup> Close in form is B 147 fgm4 b1 *māñc ce*, according to Sieg & Siegling (1953: 77) “für *mānt ce*”. Unfortunately the relevant passage is too fragmentary to be of any use in the following.

<sup>19</sup> Cf. also Adams’ translation, which follows the one by Pinault closely: ‘The wagon-drivers will try to buy for me just this sort of thing’ (2012: 9).

Schmidt (1997: 236), who first published this text, had translated ‘[Wenn] ich so nur dieses eine Mal dein Herz erreicht habe, ...’, where *miñcek* corresponds to ‘so nur’. Assuming that *miñcek* is a shortened form of late *nemñcek* for *nemcek* ‘certainly’, I have suggested translating ‘Certainly, I have got the reasoning [you] once [had] ...’ (Peyrot 2013: 707). Although the passage is difficult to interpret, Pinault’s interpretation ‘justement cette sorte de chose’ is not easy to fit in. Again, I must stress that Pinault’s interpretation is based on a use of *mant* ‘so’ that is without parallel, and I suppose that Schmidt has done the same in his translation ‘nur so’, where ‘so’ seems to correspond to *mant*, and ‘nur’ apparently to *cek*, literally ‘precisely this’.

In light of the discovery of *mantsu* ‘some; any’, I would now like to propose that *miñcek* is an indefinite as well. While in the first passage, ex. (4’), ‘something’ would be possible, this is difficult for the second, ex. (5’), so that I opt for ‘somehow’, as an adverbialised original oblique (accusative):

- (4’) *kokale(cī m) iñcek skāya-ñ kārñā(ts)i*  
 ‘The wagon-drivers will somehow try to buy it for me’ PK LC 10 a13–14
- (5’) *miñcek šeme preke palsko yolmuwa-ś*  
 ‘I have somehow got the reasoning [you] once [had] ...’ 59BTA7–50 D7046 l.3

Although ‘somehow’ certainly is compatible with the two contexts in question, I must note as a caveat that ‘somehow’ is quite easy to fit in any translation, so that my proposal is unfortunately difficult to falsify.

Formally, a connection between *miñcek* and *mantsu* obviously requires that the element *u* of *mantsu* was added later, or became obligatory only later. As mentioned above, late *miñcek* presupposes a class. *māñcek\**, which apparently had final accent because of the emphatic particle *-k < \*-kə*, i.e. *māñcek\** /māncék/ < \*māncé-kə. The form *miñcek* < *māñcek* with final *-k* in turn presupposes a form without *-k*: nom. sg. m. *mantse\**, obl. sg. m. *mañce\**. Thus, it seems that *mantsu* and *mantse\** differ only in the presence or absence of the marker *-u < \*-w*.

Alternations between forms with and without *-u* are also found elsewhere. For instance, as explained above, TB *intsu* ‘which’ with *-u* corresponds to TA *äntsam* without *-u*. Since there is no functional match between anaphoric TB *su* and distal TA *sam*, it is not likely that either of the two was replaced with the other in one of the languages, and therefore we should probably reconstruct Proto-Tocharian *\*ən-se*, which could receive either *\*-w* or *\*-nə*. That the use of markers such as *-u < \*-w* and *-n < \*nə* was not yet obligatory with *\*ən-se* in Proto-Tocharian is also suggested by forms without marker such as TA *äntā*, *tā* ‘where’ (interrogative), *tāne* ‘where’ (relative), and perhaps *änta-ne* ‘where’ (if this is not for *äntan-ne*), and TB *inte*, *ente* ‘where, when’. Finally, one may compare the difference between

TB *mäfte* ‘how’ without *-u* and *mäksu* ‘which’ with *-u*, as well as between TB *k<sub>u</sub>se* ‘who’ without *-u* and *mäksu* ‘which’ with *-u*.

Nevertheless, I must stress that synchronically the use of *-u* is NOT facultative in Tocharian B. On the other hand, diachronically *-u* may have been more transparent, so that doublets with and without *-u* could occur. The elements *-u* and *-kə* may occur together in the simple demonstrative, e.g. *suk* ‘precisely he’, *tuk* ‘precisely this’, but they may have been in complementary distribution in the derived pronouns *mantsu* and *mantse\**. It is also possible that with pronoun stems that were so rare as these two, the use of *-u* was not regularised as rigorously as with the more frequent pronouns *su* and *se*.

## 7 Tocharian B *mant* ‘so’ and Tocharian A *mänt* ‘how’

On the basis of TA *mänt* ‘how’ (relative *mäntne*) and TB *mant* ‘so’, we can reconstruct Proto-Tocharian *\*mäntə* ‘how’. For the development of the meaning in Tocharian B, compare German *wie* ‘how, as’. The original function of *\*mäntə* in Tocharian B has been taken over by *mäfte* ‘how’ < PT *\*mä-k<sup>w</sup>ə-te*.<sup>20</sup> Since the original meaning of *\*mäntə* is interrogative, it must contain the interrogative stem *\*mä-* as its first element. The explanation of the *nt*-element is so far unclear, and it has no parallel within Tocharian.

A connection of the *nt*-element of *\*mäntə* with the *nt*-element in Lat. *quantus*, *tantus* seems tempting, and Hackstein (2004: 287) has in effect reconstructed *\*mäntə* as *\*meh<sub>2</sub>-nt* (see also Pedersen 1941: 124). Unfortunately, this reconstruction is excluded because of the vocalism: instead of *\*mänt-*, this preform would have yielded *\*\*mont-* or perhaps *\*\*mant-*.

As an alternative, I would like to propose analysing *\*mäntə* fully within the framework of the other Tocharian pronoun stems: it could simply be the neuter of *mantsu* < *\*mä-ən-sə-w* and *miñcek* < *\*mä-ən-ce-kə*.<sup>21</sup> As the neuter to this stem, one would at first have expected *\*mä-ən-te*, with the full form of the neuter demonstrative *\*te* < PIE *\*tod*. However, alongside this full form, the reduced form *\*tə* is well attested, cf.:

<sup>20</sup> Pace Hilmarsson (1987: 42), *mäfte* is most probably not in origin identical to *makte* ‘self’ (see in detail Pinault 2010).

<sup>21</sup> Koller (2015: 150) also connects the element *-n-* of TA *mänt* with the *-n-* of *äntä* ‘when, where’, which he analyses as an interrogative particle.

- TB *tu* ‘this’ < \**tə* + *w*;
- TB *tamp* ‘that over there’ < \**tə* + *mp*;
- TB *taka* ‘then, thus’ < \**tə* + *ka*;
- TA *täm* ‘this’ < \**tə* + *m*;
- TA *täş* ‘this’ < \**tə* + *ş*.

Moreover, this short form of the neuter demonstrative is not found only before particles, as proved by TB *tane* ‘here’ < \**tə-ne* ‘in this’, where *-ne* is the locative suffix.

Thus, I analyse Proto-Tocharian \**məntə* as \**mə-ən-tə*, as if from PIE \**mo-m(o)-tod*. The original meaning of \**məntə* may therefore be assumed to have been ‘what’. As an adverbialised original oblique (accusative) it then came to mean ‘how’, as in TB *mäkte* ‘how’ < PT \**mə-kʷə-te* < \**mo-kʷi-tod*, which also originally must have meant ‘what’.

## 8 Tocharian B *manta* ‘not any’

With Pedersen (1941: 126), I take TA *tā* ‘where, when’ to be a shortened form of the more transparent *äntā* ‘where, when’; see also Koller (2015: 145–148), who shows that the two forms cover the same range of meanings. *äntā* is apparently the obl. sg. f. of a pronominal stem closely related to *äntsam* ‘which’, but without the marker *-n*, i.e., as if from a stem \**äntsa*.<sup>22</sup> As suggested by Pedersen, *äntā* may originally have been combined with a feminine noun such as *ytār* ‘way’. Since, as mentioned above, it lacks the marker *-n* of *äntsam* ‘which’, the specialisation of *äntā* in local function must be relatively old.

As the Tocharian B pendant of Tocharian A *äntā* one might have expected *intā* < \**ənta*. Instead of the form with *i-*, we find *nta*, an element that has an otherwise impossible onset *nt-*, which can go back virtually only to \**ənta*. The fact that the initial vowel was lost can be explained from its being consistently unaccented, shown by the spelling ⟨*nta*⟩ for /*nta*/ (/ntā/) would have been spelled ⟨*ntā*⟩, and this at the same time accounts for the exceptional onset, since it is always enclitic. The meaning differs from that found in Tocharian A, but fits its enclitic behaviour very well: *nta* is an indefinite particle found in negated contexts that is best rendered

<sup>22</sup> It is questionable whether the \**äntsa* presupposed by *äntā* is a well-formed word. Apart from the peculiar initial *ä-* discussed above, it has a final *-a* that would be unique. For Tocharian A, this gives a good motivation for the replacement of an original \**äntsa* with the actually attested *äntsam*: the final *-n* (ṃ) prevented the apocope of *-a*.

with English ‘ever’. On the assumption that the original meaning was ‘where’, as in Tocharian A, one may think of the following path of semantic development:<sup>23</sup>

TB *nta* ‘ever’ (INDEFINITE) < \*‘wherever’ < PT \**ənta* ‘where’ (TA *tā*, *äntā* ‘where, when’).

As has long been recognised, the Tocharian B adverb *manta* ‘never, not any’ most probably contains *nta* ‘ever’ as the second element (e.g. Adams 2013: 473; Hackstein 2004: 289). However, the analysis of the first element is more difficult. Adams and Hackstein opt for a derivation from the negation *mā* ‘not’, i.e. *manta* < \**mā nta*. Semantically, this is attractive, because *manta* ‘never, not any’ seems indeed to combine the meanings of *mā* ‘not’ and *nta* ‘ever’. Formally, however, this analysis is only possible if *manta* is unaccented and stands for /manta/, in which case one would expect archaic spellings of the type *mantā*, *mānta*, *māntā*, etc. This is not what we find. Archaic spellings are attested, but instead show that the vowel of the first syllable was /ə/: arch. B 284 b7, B 295 a7, IOL Toch 302 b3 *mānta*. Consequently, the classical form *manta* is to be analysed as /mənta/.

Since there are no parallels for weakening of \**a* to \**ə*, as would be needed to derive *manta* from \**mā nta*, an alternative solution is needed. I propose that *manta* is an old indefinite related to *mant* ‘so’ and *mantsu* ‘some’. Thus, the second element *nta* is indeed related to *nta* ‘ever’, but the first part *ma*<sup>o</sup> is from Proto-Tocharian \**mə-*. Originally *manta* < \**mə-ən-ta* was probably the obl. sg. f. of the same pronominal stem that is seen in TB *mant* ‘so’, TA *mānt* ‘how’ < \**mə-ən-tə*. The meaning may be derived from an earlier indefinite, ‘any, anyhow’, which became negative when an earlier negation was left out or lost; cf. Gk. *οὐ*, Arm. *օ՛ճ*, Fr. *pas*, all negations from elements that were earlier used to enforce a negation, but were not originally negative themselves. Note, in any case, that the original Proto-Indo-European negation \**ne* has been lost in Tocharian, and compare further TB *mapi* ‘isn’t it?’, which is mostly positive, but sometimes negative (Peyrot 2013: 363–367). The phonological form of *mapi* is /məpi/, and one may therefore speculate that the first element is, again, the indefinite stem \**mə-*, while the second element is to be identified with the particle *pi* ‘please’ (Peyrot 2013: 360–363). However, in the case of *mapi* the morphological structure is not fully clear, since \**mə-* otherwise does not occur alone, but only as \**mə-ən-se*, \**mə-kʷə-sə-w*, etc.

<sup>23</sup> Adams (2013: 373) rather connects TB *nta* with TA *ontam*, which is also an indefinite particle. I do agree that a connection between *nta* and *ontam* seems possible, but I think that this connection is limited to the second part of the word: *tam* may be identical to the nom.-obl. sg. n. *tam* ‘this’, and *ntam* may be the nom.-obl. sg. n. of *äntsam*. The first part *o*<sup>o</sup> remains unclear (see also Koller 2015: 133). Formally the closest cognate in Tocharian B seems to be *ente* ‘where’, but it is unclear what could have caused rounding of \**e* to *o* in this word, and *o*<sup>o</sup> may have a completely different origin.

## 9 Tocharian A *māñcaṃ*

In Tocharian A a word *māñcaṃ* is found that is relevant to the discussion because it could be related to Tocharian B *miñcek*. Unfortunately, *māñcaṃ* is rare, found only in A 461 b3 and YQ II.13 a4 (with *mānt caṃ* in the parallel A 261 b6), and its meaning has not yet been able to be established. The attestation in A 461 b3 had been read differently in the edition of Sieg & Siegling (1921), so that at the time no connection could be made with A 261 b6 *mānt caṃ*.

In the Dirghāgama bilingual A 460–466 it translates Sanskrit *apy eva* as shown in ex. (6).

- (6) /// (a)py evāhaṃ ♦ māñcaṃ nāṣ ||  
 [SKT:] even so I [TA:] *māñcaṃ* I A 461 b3

It was this attestation that was overlooked, because Sieg & Siegling (1921: 248; cf. also Couvreur 1969: 160) had read *mañcā*. However, even though the passage is damaged, it is clear that there is no trace of a virāma stroke, and the dot over the last akṣara is not a virāma dot (which could alternatively have been transliterated as *mañcā* ♦), but an anusvāra dot. Moreover, there is now a second attestation of *māñcaṃ*, while no word *māñc* is otherwise known. Although this passage gives an important first indication for the meaning of *māñcaṃ*, it must be noted that the meaning of Skt. *apy eva* is relatively weak. It occurs with the optative, mostly in wishes, but also in unreal constructions. It can have adversative value, ‘even so’, but may also just emphasise a wish, ‘perhaps’, ‘hopefully’ (SWTF: 1, 100a, “vielleicht, hoffentlich; wenn doch nur”). Obviously, we cannot simply reverse the correspondence from this bilingual by saying that the meaning of *māñcaṃ* is ‘*apy eva*’.

The second attestation shown in ex. (7) is from the Maitreyasamitināṭaka.

- (7) *māñcaṃ klyom wañi te napęsaṃ*  
 ‘Was, o Edler, ist Vergnügen unter den Menschen?’  
 YQ II.13 a4 (parallel in A 261 b6); Geng Shimin, Laut & Pinault 2004: 364

In my view, it is unlikely that *māñcaṃ* in this passage is to be translated as ‘what’, as in the translation of Geng Shimin, Laut & Pinault. The Old Uyghur rendering is shown in ex. (8).

- (8) *nāč[ük tözün] maytri bo yer suvda [ögrünč ol(?)]*  
 ‘Wa[s, edler] Maitreya, [ist] auf dieser Welt [Vergnügen?]’  
 Hami 2.14 b27–28; Geng Shimin, Laut & Pinault 2004: 355–356

In this Old Uyghur translation, *nāč[ük]* evidently corresponds to *māñcaṃ* of the Tocharian A original. Again Geng Shimin, Laut & Pinault translate *nāč[ük]* as

‘what’, in spite of the fact that *näčük* actually means ‘how’. Interestingly, Sieg, Siegling & Schulze (1931: 190) had noted on the parallel A 261 b6, which reads *mänt caṃ* instead of *mäñcaṃ*, “Unklar ist das Nebeneinander von *mänt* und *te*”. Indeed, *te* is a question particle that marks polar questions, i.e. questions that invite a yes/no answer, and it is not used together with interrogatives. The conclusion must therefore be that YQ II.13 a4 *mäñcaṃ* and the parallel A 261 b6 *mänt caṃ* are not interrogative. In view of the variant *mänt caṃ*, the simplest solution is to take *caṃ* as the indefinite pronoun, here in the function of an indefinite particle, and translate *mänt* ‘how’ + indefinite *caṃ* as ‘somehow’, cf. ex. (7’).

(7’) *mänt caṃ klyom wañi te napem*

‘O noble one, is there somehow pleasure among men?’

A 261 b6 (parallel in YQ II.13 a4)

This is perfectly compatible with the Old Uyghur rendering, since in Old Uyghur interrogatives can be used as indefinites without special marking (cf. Erdal 2004: 217–218). In accordance with the Tocharian A original, we may therefore also restore a question particle for Old Uyghur, here *mu*, cf. ex. (8’).

(8’) *näč[ük tözün] maytri bo yer suvda [ögrünč bar mu]*

‘O noble Maitreya, is there somehow pleasure in this world?’

Hami 2.14 b27–28

This would mean that *mäñcaṃ* is a recent univerbation of *mänt* and *caṃ* after all, and not of direct relevance to the discussion of Tocharian B *miñcek* and *mantsu* above.

## 10 Tocharian interrogative, relative and indefinite stems: an overview

For the sake of clarity, an overview of the above reconstructions of the Tocharian interrogative, relative and indefinite stems is given in Table 5. It must again be noted that the purpose of the mechanical Proto-Indo-European reconstructions is only to show the structure of the Tocharian formations: most of these combinations are obviously not of Proto-Indo-European date.

**Table 5:** Tocharian interrogative, relative and indefinite stems

Tocharian B	Tocharian A	Proto-Tocharian	Proto-Indo-European
<i>k<sub>u</sub>se</i>	<i>kus</i>	<i>k<sup>w</sup>ə-se</i>	<i>k<sup>w</sup>i-so</i>
interrogative and relative; pronominal, inflected for case			
<i>ksa</i>		<i>k<sup>w</sup>ə-se</i>	<i>k<sup>w</sup>i-so</i>
indefinite; inflected for case			
	<i>sam̄</i>	<i>se-nə</i>	<i>so-nu</i>
indefinite; inflected for case			
	<i>āntā</i>	<i>ən-ta</i> , stem <i>ən-se</i>	<i>mo-teh<sub>2</sub>m</i> , stem <i>mo-so</i>
'where'			
	<i>tā</i>	<i>ən-ta</i> , stem <i>ən-se</i>	<i>mo-teh<sub>2</sub>m</i> , stem <i>mo-so</i>
'where', reduced form of <i>āntā</i>			
<i>intsu</i>		<i>ən-sə-w</i>	<i>mo-so-u</i>
'which'; interrogative and indefinite; pronominal, inflected for case			
	<i>āntsam̄</i>	<i>ən-se-nə</i>	<i>mo-so-nu</i>
'which'; interrogative; modifier, inflected for case, gender and number			
<i>mākte</i>		<i>mə-k<sup>w</sup>ə-te</i> , stem <i>mə-k<sup>w</sup>ə-se</i>	<i>mo-k<sup>w</sup>i-tod</i> , stem <i>mo-k<sup>w</sup>i-so</i>
'how' < *'what'			
<i>māksu</i>		<i>mə-k<sup>w</sup>ə-sə-w</i>	<i>mo-k<sup>w</sup>i-so-u</i>
'which'; interrogative and relative; modifier, inflected for case, gender and number			
<i>mant</i>	<i>mānt</i>	<i>mə-ən-tə</i> , stem <i>mə-ən-se</i>	<i>mo-mo-tod</i> , stem <i>mo-mo-so</i>
'how' < *'what'			
<i>manta</i>		<i>mə-ən-ta</i> , stem <i>mə-ən-se</i>	<i>mo-mo-teh<sub>2</sub>m</i> , stem <i>mo-mo-so</i>
'never'			
<i>miñcek</i>		<i>mə-ən-ce-kə</i> , stem <i>mə-ən-se</i>	<i>mo-mo-tom</i> , stem <i>mo-mo-so</i>
'somehow'			
<i>mantsu</i>		<i>mə-ən-sə-w</i>	<i>mo-mo-so-u</i>
'some'			
	<i>**ā, ām̄</i>	—	—
ghost stem, nonexistent			

## 11 The *m*-interrogative: reconstruction and the phylogenetic tree of Indo-European

On the basis of Hitt. *māši*- ‘how many’ and Pal. *maš* ‘as much as’, Hackstein (2004: 281–282) suggests that *\*mo-* was a quantifying interrogative ‘how many’. In my view, this is not attractive for the more basic meanings attested in Tocharian: the evidence of the various interrogative, relative and indefinite stems of Tocharian rather suggests that the meaning of the *m*-interrogative was more basic.

One might think of a more specific interrogative. According to Hackstein (2004: 281), “TB *māksu* is best described as an adjectival interrogative for restricting reference, ‘which one of a given class or group’”. This would also account for the function of TB *intsu* ‘which’, while the meaning of the other *m*-formations in Tocharian, such as the indefinites, may well result from bleaching. The distinction between a more general interrogative *\*k<sup>w</sup>i-*, *\*k<sup>w</sup>e-*, *\*k<sup>w</sup>o-* and a more specific *\*mo-* is perhaps reflected by the tendency in PIE to use *\*mo-* for derived interrogatives and the promotion of *\*k<sup>w</sup>i-*, *\*k<sup>w</sup>e-*, *\*k<sup>w</sup>o-* to the main and basic interrogative.<sup>24</sup>

Although altogether the evidence for a Proto-Indo-European interrogative stem *\*mo-* is fairly strong, it is difficult to use this Anatolian-Tocharian isogloss as an argument for the phylogenetic tree of Indo-European. The basic problem was already voiced by Pedersen (1938: 72):

Es genügt hervorzuheben, dass Verlust des Alten (in diesem Falle des *m*-Pronomens), der allmählich in jedem Sprachzweige für sich eingetreten sein kann, nicht als eine gemeinsame Neuerung gewertet werden darf.

As pointed out in the introduction above, I think that the loss of a grammatical category such as an entire interrogative stem is more significant than the loss or replacement of a lexical item. Also, it is in my view easier to assume that loss of an element occurred once than that it occurred eight times independently in the remaining branches of Indo-European. However, this is an argument of probability, and other, more solid evidence is needed before weaker arguments of this type can be considered in the discussion.

An additional problem in the case of *\*mo-* is that there are further possible traces of this stem in Celtic, where OIr. *má* ‘if’ and related forms are found (Dunkel

---

<sup>24</sup> As pointed out to me by Martin Kümmel (p.c.), one might expect, if there are two basic interrogatives, that one refers to persons, i.e. ‘who’ and the other to things, i.e. ‘what’. I find it attractive to search for a difference of this kind between the stems *\*mo-* and *\*k<sup>w</sup>i-*, *\*k<sup>w</sup>e-*, but I see so far no evidence for such a distribution.

2014: 518–523).<sup>25</sup> These forms have no interrogative value, and an origin of this stem as a local conjunction is only suggested by Breton and Cornish (Pedersen 1913: 230). All in all the semantics would have been seriously bleached out, but a local conjunction would in turn make a derivation from an interrogative likely. Even if the appurtenance of the Celtic forms remains only a possibility, it further weakens the evidence of the *m*-interrogative for the phylogenetic tree of Indo-European.

**Acknowledgment:** This article is an adaptation of a lecture held at Arbeitstagung *100 Jahre Entzifferung des Hethitischen – Morphosyntaktische Kategorien in Sprachgeschichte und Forschung* of the Indogermanische Gesellschaft (Marburg, 21 September 2015). This research was supported by the Netherlands Organisation for Scientific Research (NWO, project number 276-70-028).

## Abbreviations

- SWTF *Sanskrit-Wörterbuch der buddhistischen Texte aus den Turfan-Funden* [und der kanonischen Literatur der Sarvāstivāda-Schule] (1973–). Begonnen von Ernst Waldschmidt. Im Auftrage der Akademie der Wissenschaften zu Göttingen hrsg. von Heinz Bechert et al. Göttingen: Vandenhoeck & Ruprecht.
- TEB Wolfgang Krause & Werner Thomas (1960). *Tocharisches Elementarbuch*. Vol. 1: *Grammatik*. Heidelberg: Winter.

## Bibliography

- Adams, Douglas Q. (2012). “Another look at three Kuci-Prākṛit–Tocharian B bilinguals”. In: *Tocharian and Indo-European Studies* 13, 7–55.
- Adams, Douglas Q. (2013). *A Dictionary of Tocharian B*. Revised and greatly enlarged. 2 vols. Amsterdam & Atlanta: Rodopi.
- Beekes, Robert S. P. (2011). *Comparative Indo-European Linguistics. An Introduction*. 2nd ed. Amsterdam & Philadelphia: Benjamins.
- Bernhard, Franz (1965). *Udānavarga*. Vol. 1: *Einleitung, Beschreibung der Handschriften, Textausgabe, Bibliographie*. Göttingen: Vandenhoeck & Ruprecht.

---

<sup>25</sup> Dunkel further adds, among many other forms, Arm. *imn* ‘something’, *omn* ‘somebody’, and Ved. *néma-* ‘the one’, because the *m*-pronoun is to account for the indefinite semantics of these forms. I find this highly doubtful, since there are many different sources possible for indefinite formations, cf. in particular the thorough typological study by Haspelmath (1997). There is, in other words, no “need” to explain indefinite semantics with a specifically indefinite formative element.

- Buck, Carl D. (1949). *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*. A contribution to the history of ideas. Chicago: University of Chicago.
- Carling, Gerd (2005). "Proto-Tocharian, Common Tocharian, and Tocharian. On the value of linguistic connections in a reconstructed language". In: *Proceedings of the Sixteenth Annual UCLA Indo-European Conference*. Los Angeles, November 5–6, 2004. Ed. by Karlene Jones-Bley et al. Washington, DC: Institute for the Study of Man, 47–70.
- Carling, Gerd (2009). *Dictionary and Thesaurus of Tocharian A*. Vol. 1: a–j. Wiesbaden: Harrassowitz.
- Couvreur, Walter (1969). "Sanskrit-Tochaarse en Sanskrit-Koetsjische trefwoordenlijsten van de Dīrghāgama (Dīghanikāya)". In: *Orientalia Gandensia* 4, 151–165.
- Dunkel, George E. (2014). *Lexikon der indogermanischen Partikeln und Pronominalstämme*. Vol. 2: *Lexikon*. Heidelberg: Winter.
- Erdal, Marcel (2004). *A Grammar of Old Turkic*. Leiden & Boston: Brill.
- Fortson, Benjamin W. IV (2004). *Indo-European Language and Culture. An Introduction*. 1st ed. Oxford: Wiley-Blackwell.
- Geng Shimin, Jens Peter Laut & Georges-Jean Pinault (2004). "Neue Ergebnisse der Maitrisimit-Forschung". In: *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 154, 347–369.
- Hackstein, Olav (2004). "From discourse to syntax. The case of compound interrogatives in Indo-European and beyond". In: *Proceedings of the Fifteenth Annual UCLA Indo-European Conference*. Ed. by Karlene Jones-Bley, Martin E. Huld, Angela Della Volpe, et al. Washington, DC: Institute for the Study of Man, 257–298.
- Hahn, Michael (2007). *Vom rechten Leben. Buddhistische Lehren aus Indien und Tibet*. Frankfurt am Main & Leipzig: Verlag der Weltreligionen.
- Haspelmath, Martin (1997). *Indefinite Pronouns*. Oxford: Clarendon.
- Hilmarsson, Jörundur G. (1987). "Stray notes on the interrogative pronominal stems in Tocharian". In: *Tocharian and Indo-European Studies* 1, 40–48.
- Jasanoff, Jay H. (2003). *Hittite and the Indo-European Verb*. Oxford & New York: Oxford University Press.
- Kim, Ronald I. (2007). "The Tocharian subjunctive in light of the  $h_2e$ -conjugation model". In: *Verba docenti*. Studies in historical and Indo-European linguistics presented to Jay H. Jasanoff by students, colleagues, and friends. Ed. by Alan J. Nussbaum. Ann Arbor: Beech Stave, 185–200.
- Kloekhorst, Alwin (2008). *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden & Boston: Brill.
- Kloekhorst, Alwin (2010). "Hittite *mān*, *maḥḥan*, *māḥḥan*, *māḥḥanda* and *mānḥanda*". In: *Ex Anatolia Lux*. Anatolian and Indo-European studies in honor of H. Craig Melchert on the occasion of his sixty-fifth birthday. Ed. by Ronald I. Kim et al. Ann Arbor: Beech Stave, 217–226.
- Koller, Bernhard (2015). *Studies in Tocharian Phonology above the Word-Level*. PhD thesis. Los Angeles: University of California. URL: <https://escholarship.org/uc/item/3dk2b17w>.
- Koller, Bernhard (2016). "Virāma spelling and Tocharian A prosody". In: *Tocharian and Indo-European Studies* 17, 77–108.
- Kortlandt, Frederik H. H. (1983). "Demonstrative pronouns in Balto-Slavic, Armenian, and Tocharian". In: *Dutch Contributions to the Ninth International Congress of Slavists*. Kiev, September 6–14, 1983. Linguistics. Ed. by André G. F. van Holk. Amsterdam: Rodopi, 311–322.

- Malzahn, Melanie (2016). "The second one to branch off? The Tocharian lexicon revisited". In: *Etymology and the European Lexicon*. Proceedings of the 14<sup>th</sup> Fachtagung der Indogermanischen Gesellschaft, 17–22 September 2012, Copenhagen. Ed. by Bjarne Simmelkjær Sandgaard Hansen et al. Wiesbaden: Reichert, 281–292.
- Meier-Brügger, Michael (2003). *Indo-European Linguistics*. In cooperation with Matthias Fritz and Manfred Mayrhofer. Berlin: de Gruyter.
- Pedersen, Holger (1913). *Vergleichende Grammatik der keltischen Sprachen*. Vol. 2. Göttingen: Vandenhoeck & Ruprecht.
- Pedersen, Holger (1938). *Hittitisch und die anderen indoeuropäischen Sprachen*. København: Levin & Munksgaard.
- Pedersen, Holger (1941). *Tocharisch vom Gesichtspunkt der indoeuropäischen Sprachverglei-  
chung*. København: Munksgaard.
- Peyrot, Michaël (2013). *The Tocharian Subjunctive*. A Study in Syntax and Verbal Stem Formation. Leiden & Boston: Brill.
- Peyrot, Michaël (2016). "Further Sanskrit-Tocharian bilingual Udānavarga fragments". In: *Tocharian and Indo-European Studies* 17, 153–211.
- Pinault, Georges-Jean (2006). "Retour sur le numéral *un* en tokharien". In: *Indogermanische Forschungen* 111, 71–97.
- Pinault, Georges-Jean (2008). *Chrestomathie tokharienne. Textes et grammaire*. Leuven & Paris: Peeters.
- Pinault, Georges-Jean (2010). "Le pronom d'ipséité en tokharien". In: *Typologie et comparatisme. Hommages offerts à Alain Lemaréchal*. Ed. by Injoo Choi-Jonin, Marc Duval & Olivier Soutet. Leuven: Peeters, 351–365.
- Rieken, Elisabeth (2000). "Die Partikeln *-a*, *-ja*, *-ma* im Althethischen und das Akkadogram *Ū*". In: *125 Jahre Indogermanistik in Graz*. Ed. by Michaela Ofitsch & Christian Zinko. Graz: Leykam, 411–419.
- Ringe, Don, Tandy Warnow & Ann Taylor (2002). "Indo-European and computational cladistics". In: *Transactions of the Philological Society* 100.1, 59–129.
- Schmidt, Klaus T. (1992). "Archaismen des Tocharischen und ihre Bedeutung für Fragen der Rekonstruktion und der Ausgliederung". In: *Rekonstruktion und relative Chronologie*. Akten der VIII. Fachtagung der Indogermanischen Gesellschaft, Leiden, 31. August – 4. September 1987. Ed. by Robert Beekes, Alexander Lubotsky & Jos Weitenberg. Innsbruck: Institut für Sprachwissenschaft, 101–114.
- Schmidt, Klaus T. (1997). "Liebe und Sexualität im Spiegel der tocharischen Sprachzeugnisse". In: *Eros, Liebe und Zuneigung in der Indogermania*. Ed. by Michaela Ofitsch. Graz: Leykam, 227–262.
- Schmithausen, Lambert (1970). "Zu den Rezensionen des Udānavargaḥ". In: *Wiener Zeitschrift für die Kunde Südasiens* 14, 47–124.
- Sieg, Emil & Wilhelm Siegling (1908). "Tocharisch, die Sprache der Indoskythen. Vorläufige Bemerkungen über eine bisher unbekannt indogermanische Literatursprache". In: *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*. Jahrgang 1908, 915–932.
- Sieg, Emil & Wilhelm Siegling (1921). *Tocharische Sprachreste*. Vol. 1: *Die Texte. A. Transcription*. Berlin: de Gruyter.
- Sieg, Emil & Wilhelm Siegling (1933). "Bruchstück eines Udānavarga-Kommentars (Udānālamkāra?) im Tocharischen". In: *Festschrift Moriz Winternitz 1863 – 23. December 1933*. Ed. by Otto Stein & Wilhelm Gampert. Leipzig: Harrassowitz, 167–173.

- Sieg, Emil & Wilhelm Siegling (1953). *Tocharische Sprachreste*. Sprache B, Heft 2. Fragmente Nr. 71–633. Göttingen: Vandenhoeck & Ruprecht.
- Sieg, Emil, Wilhelm Siegling & Wolfgang Schulze (1931). *Tocharische Grammatik*. Göttingen: Vandenhoeck & Ruprecht.
- Stumpf, Peter (1971). *Der Gebrauch der Demonstrativ-Pronomina im Tocharischen*. Wiesbaden: Harrassowitz.
- Winter, Werner (1997). “Lexical archaisms in the Tocharian languages”. In: *Historical, Indo-European, and Lexicographical Studies*. A festschrift for Ladislav Zgusta on the occasion of his 70th birthday. Ed. by Hans H. Hock. Berlin: Mouton de Gruyter, 183–193.