

# Proto-Tocharian Syntax and the Status of Tocharian A<sup>1</sup>

Michaël Peyrot  
*Universiteit Leiden*

Tocharian A must have been a second language for at least some and perhaps all of its users, since only Old Uygur and Tocharian B – not Tocharian A itself – can be shown to have been native languages of the scribes. In view of the considerable differences between Tocharian A and B in the morphological and lexical domains, it is remarkable that they are phonologically and syntactically very similar. This is an indication of substratum influence, probably exerted by Tocharian B, since that seems to have been the dominant language in general. A possible explanation is that speakers of Tocharian A commissioned Tocharian A texts from Tocharian B writers.

## §1 Introduction

In the first half century of Tocharian studies Tocharian A played the leading part, without doubt because it was first deciphered and described. Only in the second half century did Tocharian B gradually gain the position it deserved: its texts outnumber those of Tocharian A by far and they are more diverse, too, both in content and in language, and it has long been recognized that Tocharian B is linguistically more archaic on the whole. The predominance of Tocharian A was certainly one of the many reasons for the relatively slow progress in Tocharian historical grammar. With a bias towards Tocharian A, it was much more difficult to understand the Proto-Tocharian vowel system, for instance, which is closer to Tocharian B.

Another consequence of the focus on Tocharian A was, ironically, that some important facts about it were observed at a rather late stage of the research because these are most salient in comparison with Tocharian B: its language is strikingly uniform, its style formulaic, and the lexicon contains loan-

---

<sup>1</sup>This paper is an adaptation of a lecture with the same title held at the Indogermanische Fachtagung, 23 September 2008, Salzburg. I would like to thank Alexander Lubotsky and Kristin Meier, as well as the two linguistic referees, for comments on an earlier draft.

words from Tocharian B. This has led to the conclusion that unlike Tocharian B, Tocharian A was not a spoken vernacular, but a fossilized liturgical language (Winter 1963: 243-244; Lane 1966: 226-227).

In this paper, I will investigate cases of interference between Tocharian A and B. First I will treat the text level (§2), then the linguistic level (§3), and finally I will consider possible consequences for the reconstruction of Proto-Tocharian and propose a possible explanation (§4).

## §2 Text interference

§2.1 Since Tocharian B was found at all sites where Tocharian A was found, and many manuscripts date from the same period, one could expect to find extensive text interference between the two languages, that is, textual evidence for the contacts between their speakers: for instance, texts translated from one language into the other, passages with citations from the other language, or even explicit mention of the sister language. However, such instances are exceedingly rare: we have no evidence of a Tocharian A text translated from Tocharian B or vice versa. The text parallels we do have are not precise enough to prove that one is translated from the other, or they concern parallel translations from Sanskrit. In my view, there must have been contact between the scribes of Tocharian A and B,<sup>2</sup> but this contact is not manifest in the texts because the languages were so close that it was not necessary to make translations. Nevertheless, some rare but instructive instances of text interference are found.

§2.2 Without doubt the most striking sample of text interference is the much discussed leaf A394 from Turfan in Tocharian A with glosses in Tocharian B and Old Uygur. As has been shown by Winter (1963: 242-243), these glosses are not written by a native speaker of Tocharian B, but rather by a Turk. Apart from the presence of glosses in Old Uygur, this is indicated by the striking syncope of *a* (Winter p. 249) and the calque *šaṅ kekseṅ* (see below). *a*-syncope is found in TB *ynemne* 'going' for *ynemane*, which glosses TA *ymām* a2 'id.'; in TB *yšucats*, gen.pl. of 'beggar', for *yaššucaṃts*, which glosses TA *paṃṣāntās wrasaśsi* a1 'of the begging beings'; and in ...*skemne* (see further below). Perhaps this *a*-syncope can be connected

---

<sup>2</sup>In as far as they were not the same people, see §4.

with the defective writing of especially Old Uygur *a* and *ä* /*æ*/ in the Uygur script, which suggests that these vowels could be very short, but at least it shows that the Tocharian B is peculiar (by contrast, the two Old Uygur words are correct). The TB gloss *ṣaṇ̄ kekseṇ̄* ‘his own body’ must be a calque because it glosses TA *ṣṇ-āñcām* b1 ‘self’, which would rather be *ṣaṇ̄-āñm* in Tocharian B. The notable addition of the word for ‘body’ is strongly reminiscent of Old Uygur *ät’öz*, which means both ‘self’ and ‘body’ (where the latter is of course primary).

Winter concluded that the Turkish glossator had only “a modest command of [Tocharian] B” (p. 243) and he was practising Tocharian B with the help of a Tocharian A text. In my view, it is rather the other way round: as a native of Old Uygur, he spoke Tocharian B as a second language, and used it to understand the Tocharian A text. This assumption neatly explains why only parts of words are glossed when Tocharian A and B are almost the same, and why two words are glossed in Old Uygur.

In a number of cases, the glossator noted only parts of Tocharian B words. It seems that these parts systematically concern relevant differences between the Tocharian A and B words, whereas the parts he left out are identical or self-evident. Of TA *kälporā* b3 ‘having attained’ only the last part is rendered with TB *...rsa*: the complete Tocharian B word would have been *kälporsa*, the first two syllables being identical. Of TA *risāt* b1 ‘he gave up’ only the first syllable is rendered with TB *riṃ...*: indeed, the lack of a nasal in Tocharian A is the most striking difference with TB *rintsate*. TA *wrinās* a2, pl.f. of the adj. of ‘water’ is glossed with TB *...yana*, which must be a mistake, since we expect a TB pl.f. *wriyeṣṣana*\* rather than *\*\*wriyana*, but the endings TA *-ās* and TB *-ana* obviously do present the relevant difference between the two languages. TB *...skemne* in turn glosses the most deviating part of TA *wätkäsmām* a3 ‘commanding’: the complete TB word would be *watkäskemane* (for the *a*-syncope, see above).

The idea that the glossator read the Tocharian A with the help of similar words in Tocharian B is further supported by two glosses in Old Uygur to TA words without an etymological TB cognate. We find TA *kātka-ṃ* a3 ‘arose to him’, glossed with OUy. *turti* ‘stood up’, and *mokatsām* a4-b1 ‘strong’ glossed with OUy. *küčlüg* ‘strong’, which would be *tsänkā-ne* and probably *präkre* ‘firm’, respectively, in Tocharian B (on these glosses, cf.

in particular Maue 2009: 14). There are two counterexamples to this explanation: TA *pe b1* ‘also’, glossed with TB *ra* ‘id.’, and TA *tāp* ‘he ate’, glossed with TB *śuwa* ‘id.’. In the first case, it was probably the close similarity in function that made a rendering by Tocharian B instead of Old Uygur more suitable, whereas in the second the TA prs. *śwā-* of the same (suppletive) verb may have been the reason to gloss in Tocharian B.

In short, the highly interesting leaf A394 attests only a very special kind of interference between the two Tocharian languages. It does not prove that speakers of Tocharian B needed to gloss Tocharian A texts, but rather that there were Turks who spoke Tocharian B and studied Tocharian A. In other words, Tocharian B was at a certain point spoken by non-natives, who at the same time held Tocharian A in high esteem.

That Tocharian A was prestigious among a certain group of Turks is well known: in Tocharian A manuscripts we find some names and titles that look Turkish. Most of these manuscripts come from Turfan (e.g. A382, see Maue 2009: 16) but some also from Šorčuq (e.g. A142, see Maue 2009: 15). The Tocharian A leaf from the Musée Guimet edited by Pinault (2007b) is without doubt the most striking example of such a text, as it consists almost completely of Turkish names and titles. We can add a rare sample of Turkish glosses to Tocharian A texts from Xočo, THT1651a1, where Tocharian A *tatämsusā* ‘she has begot’ is glossed with Old Uygur <tuγurγocisi><sup>3</sup> for *tugurgučisi*, agent noun of *tugur-* ‘beget’.

§2.3 Another type of text interference is presented by some instances of Tocharian B additions to Tocharian A texts (not the other way round). In these cases, the Tocharian B is clearly added at a later stage and it does not interfere with the content of the Tocharian A. For instance, in the Sanskrit text THT1524 (Šorčuq), concluded by TA /// [*śa*]*k<sup>ā</sup>* *indrīsyo kaknoṣ<sub>ṣ</sub> tākimassū<sup>4</sup>* ‘may we be equipped with the ten (?) senses’, we find a later addition in Tocharian B: /// *īndrārjune walau*

<sup>3</sup>I owe the correct reading of the syllable <γyo> (instead of <γyā>) to Dieter Maue (p.c.).

<sup>4</sup>The 1pl. opt. ending is *-i-mās*: *sū* must be a particle. Evidently, it is to be compared with *päklyossū* A370.4 ‘listen!’, where *sū* was already analysed as a particle by Pinault (2005: 515-518).

*yrāṣaṅṅe yama(ṣṣa)[ly](e) śol<sub>1</sub> prakare tākaṃ {-e -e}<sup>5</sup> ketseṃ snai teki paltskotse anāstaṅṅe tākaṃ* ‘King Indrārjuna is to be respected. [Then] life will be long, the body [will be] without sickness, and there will be *anāstaṅṅe* (?) of the mind.’

The contents of the Tocharian B text do not have a special connection with the Tocharian A text. It is possible, for instance, that the Tocharian B scribe used only the Sanskrit text – not the Tocharian A addition – or just added a wish on the blank space of a manuscript. It must be noted, however, that the Tocharian B is a bit peculiar: the spellings *yrāṣaṅṅe* and *prakare* for *yārṣalṅe* and *pārkare* look “Tocharian-A-like”, just as the inserted *t* in *paltskotse* (for Tocharian B *palskontse*; TA *pālskes*, nom.-obl. *pālsāk*). Although I do not know how to explain *ketseṃ* for *kekteṅe* (TA *kapśaṅi*) as a Tocharian A based error, I have to conclude on the basis of *yrāṣaṅṅe*, *prakare* and *paltskotse* that the scribe was probably more used to writing Tocharian A than Tocharian B.

§2.4 As pointed out by Winter (1963: 243; see also Lane 1966: 228-229), perhaps the most important case of text interference is a mistake in the heading of a Maitreya-samitināṭaka leaf from Šorčuq. Apparently, the mistake was recognized and the leaf was copied again: on the corrected leaf (A252), we read Skt. *prathama* ‘first’ followed by TA *malto* ‘at first’, denoting the first chapter of the text, whereas in the leaf with the mistake (A251), we find *prathama* followed by TB *pārweṣṣe* ‘first’. On another Tocharian A leaf (A372) we find a pious wish in Tocharian B: *(pā)rweṣṣe kartse tāko(y)* ‘may the first be good’. Since in these two cases we have no indications that the scribe was a Turk, the evident conclusion is that he was a speaker of Tocharian B – apparently Tocharian A was written by speakers of Tocharian B, at least some manuscripts as important as this Maitreya-samitināṭaka copy.<sup>6</sup>

§2.5 So far, no speakers of Tocharian A have been traced, which is fully in line with the old idea that it was a book language, not a spoken vernacular (see §1). However, this

<sup>5</sup>Remains of two akṣaras with *ə*-vocalism are visible, but they have been erased; perhaps, it is a false start of the word *kekteṅe*, which follows as *ketseṃ* here.

<sup>6</sup>If the scribe was a Turk, this changes matters a bit, but not radically: in that case he must have been such a fluent speaker of Tocharian B that he could make this type of mistake.

traditional view is not generally accepted, as the two leading scholars of Tocharian disagree. Schmidt claims that glosses in Tocharian A prove that it was in use as a living language (1983: 279), and Pinault argues that a Tocharian A manual on magic and medicine shows that it was not dead (2007a: 180).

In my view, Pinault's argument of text genre is important, but not decisive. It certainly proves that Tocharian A apparently had a wider use than the calligraphic Buddhist manuscripts that make up the vast majority of the collection, but that does not mean that it was still a language with native speakers at the time of the manuscripts. It is clear that Tocharian A had a very high status since it was the holy language of a certain circle of Turkish Buddhists and it was written by speakers of Tocharian B. These non-native writers of Tocharian A must have had a way to pronounce it, and perhaps they spoke it in certain contexts. I think that "monastery language" covers this somewhat wider use. Although its content is not religious, even the manual for magic and medicine can be understood this way, since it is certainly a text that fits into the Indian-Buddhist literary tradition, as is also evident from its verse form.

Likewise, Schmidt's argument based on the Tocharian A glosses does not prove that it was a spoken vernacular: in a religious context, it is perfectly possible to use a holy language for glosses – all are in fact found in Sanskrit Buddhist texts. Most Tocharian A glosses are extremely difficult to read and decipher, and many have not yet been published (cf. Malzahn 2007a). I will pick out only a few that might be especially interesting because they are well understood and display linguistic peculiarities at the same time (SHT5-1098 = THT4083, Murtuq): *ets*<sub>v</sub> (e.a4), which glosses *irsyako*<sup>7</sup> 'jealous', for *emts* 'jealousy'; *yisli* (e.a4), which glosses *matsari* 'envious, jealous', for *yäslyi* 'envy'; *pakar<sub>v</sub>ka<sub>v</sub>* (e.a5), which glosses *badhnata*, 2pl.ipv. of 'bind', for *pākärkäs*. The shape of the first two of these glosses may not seem very striking (*ets* is well attested elsewhere, for instance), but it should be noted that the sound changes attested are perfectly parallel to those in late Tocharian B, namely *nts* > *ts*, *yä* > *yi*, whereas *li* for *lyi* may be hypercorrect (Peyrot 2008: 55, 69, 109). However, in the third gloss the scribe clearly mixed up the vowels, which might point to Tocharian B influence, since in this language the

---

<sup>7</sup>For *irsyakah*.

vocalism would be exactly the reverse: *pākarkäs\** /pākǎrkəs/.

§2.6 Since it could be used for the claim that Tocharian A was spoken, it is necessary to discuss a small text that is – exceptionally – written on a piece of cloth (THT1559, from Šorčuq): *tāš\ pak\ turkšāñč<sup>8</sup>\ y(ā)mtsāt\ put(i)šparsim<sup>8</sup> akālyo k<sub>a</sub>lyi maškaṃ tākim<sub>a</sub>* ‘Turkšāñč has made this *pāk*;<sup>9</sup> may I through the wish for the Buddha-worth no longer be a woman!’ We are dealing with a personal note by a woman who wishes to be reborn as a man, a necessary step towards enlightenment. In my view, this sample does not prove that Tocharian A was spoken, because the language is correct, and the text must have been written by a well-trained and religious scribe. Since the subjects of the two clauses are probably the same, the name of the scribe was *Turkšāñč*, which does not look very ‘Tocharian’ (the suffix is probably to be compared with the Sogdian feminine suffix *-āñč*, see Gershevitch 1954: 158-159). Probably, *Turkšāñč*’ mother-tongue was not Tocharian A, but for instance Old Uygur.

§2.7 As far as text interference is concerned, evidence that Tocharian A was more than a monastery language is scanty indeed. It is very difficult to prove that no scribe was a native speaker of the language, but the fact that the only scribes that we can trace through the texts turn out to be native speakers of Tocharian B and Old Uygur is a strong indication nevertheless – in any event, it is a distinct possibility that Tocharian A was not a normal spoken language (on alleged morphological influence of Tocharian A on Tocharian B, see §3.3).

### §3 Linguistic interference

§3.1 On the linguistic level, the most obvious type of interference is without doubt found in the lexicon. Although

<sup>8</sup>The *t* of *put(i)* is a bit below the line, but a reading *pu(t)t(i)*, the correct form, seems to be excluded.

<sup>9</sup>Perhaps the word *pāk* (f.) is related to Tocharian B *pakai* /pǎkay/ (obl.) ‘chowry’ (a kind of fan). For the formation, cf. obl. TB *kolmai* ‘boat’ vs. TA *kolām* (f.) and TB *salyai* ‘border’ vs. TA *slyi* (f.). If *pāk* means ‘chowry’, this piece of cloth could be a part of that chowry. Alternatively, *pāk* might be a loan-word from Chinese, as was suggested to me by Georges-Jean Pinault (p.c.). Unfortunately, neither of these hypotheses can be verified as long as the meaning of the word is uncertain.

the lexical differences between the two languages are large, as shown by Lane (1966: 221-223), a considerable number of words are actually “too similar” in the two languages: they do not conform to the systematic differences between cognate inherited lexical items. As shown by Winter in his fundamental study on “lexical interchange” (1962<sup>10</sup>), most of these similarities have to be explained with borrowing from Tocharian B into Tocharian A. If these loan-words are scrutinized in the light of the chronological stratification of Tocharian B, they appear to be of classical shape (Šorčuq type): they do not generally display the typical cluster simplifications of the late language, and they reflect a triple *a* vowel system ( $\bar{a} : a : \ddot{a}$ ), just like the written forms of classical Tocharian B (cf. in general Peyrot 2008).

In loan-words into Tocharian A, the Tocharian B phoneme /a/ (accented < $\bar{a}$ >, unaccented <a> in the classical language) is represented as / $\bar{a}$ /, irrespective of the Tocharian B accent:

Tocharian A	meaning	TB (class.)	TB (arch.)	TB phonological
<i>āyāto</i>	‘suitable’	<i>ayāto</i>	* <i>āyāto</i>	/ayāto/
<i>āsāwe</i>	‘gross, rough’	<i>aśāwe</i>	* <i>āsāwe</i>	/aśāwe/
<i>lālamṣkā</i>	‘tender’ (f.)	<i>lalamṣka</i>	* <i>lālāmṣkā</i>	/lalónṣka/

On the other hand, the Tocharian B phoneme /ə/ (accented <a>, unaccented < $\bar{a}$ > in the classical language) is represented as / $\bar{a}$ / or /a/ in Tocharian A, according to the accent rules of classical Tocharian B:<sup>11</sup>

Tocharian A	meaning	TB (class.)	TB (arch.)	TB phonological
<i>enkāl</i>	‘passion’	<i>enkāl</i>	<i>enkāl</i>	/énkəl/
<i>enkalsu</i>	‘passionate’	<i>enkalsu</i>	* <i>enkālsu</i>	/enkól̄su/
<i>ṣñāṣse</i>	‘relative’	<i>ṣñāṣse</i>	* <i>ṣñāṣse</i>	/ṣñól̄ṣse/
<i>ymassu</i>	‘thoughtful’	<i>ymassu</i>	<i>ymāssu</i>	/ymól̄ssu/
<i>cāmpamo</i>	‘able’	<i>cāmpamo</i>	* <i>cāmpāmo</i>	/cāmpól̄mo/

<sup>10</sup>Couvreur (1947: 77 and passim) gives the first systematic explanation of irregular word-final correspondences in terms of borrowing, although the concept clearly goes back to Pedersen (1941, e.g. p. 71, 96).

<sup>11</sup>The different treatment of TB /a/ and /ə/ may indicate that the difference between the phonetic values behind the graphemes <a> (TB /a/ and / $\acute{a}$ /) and < $\bar{a}$ > (TB /ə/) was larger than that between those for < $\bar{a}$ > (TB / $\acute{a}$ /) and <a>. In other words, the writers of Tocharian A apparently found it more important to distinguish the Tocharian B allophones of /ə/ than those of /a/.

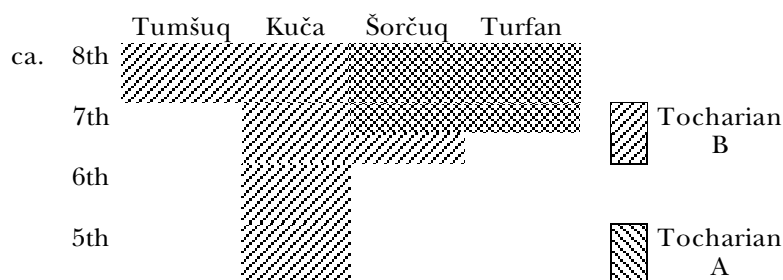


We do not find late features like for instance a colored shwa in *\*\*cāmpamo* for *cāmpamo*. In three cases we find cluster simplification, but the simplified Tocharian B forms are well attested in the classical material (Malzahn 2007b: 239-240; Peyrot 2008: 63-64, 71).

Tocharian A	meaning	TB (class.)	TB (arch.)
<i>ñātse</i>	‘need’	<i>ñātse, ñyātse</i>	<i>ñyātse</i>
<i>ñās</i>	‘desire’	<i>ñās, ñyās</i>	<i>ñyās</i>
<i>šmoññe</i>	‘place’	<i>šmoññai</i> (obl.), <i>šmoññai</i> (obl.)	<i>šmoññai</i> (obl.)

The fact that the Tocharian B phoneme /a/ is always rendered by Tocharian A /ā/ could suggest an early date of the borrowings, namely in the archaic period of Tocharian B, but the double representation of /ə/ clearly rules out this possibility, since the accent rules affected /a/ first and /ə/ only afterwards (Peyrot 2008: 33-41).

The dating of these loan-words fits very well into our picture of the spread of Tocharian B to the east as it is attested by the manuscripts (Peyrot 2008: 191-196). A later date could still be reconciled with what we know about the dates of the manuscripts, but an earlier date would leave us with considerable problems concerning the prehistoric movements of especially the speakers of Tocharian A. In the scheme below, the period of the contacts between Tocharian B and Tocharian A is represented vertically, and the contact area horizontally. During its development from approximately the 5th to the 8th centuries, Tocharian B spread west and east from its original home Kuča. When it reached Šorčuq and Turfan in the 7th century, it came into contact with Tocharian A.



§3.2 In phonetics and phonology, another domain of possible linguistic interference, we cannot make much progress,

unfortunately. The manuscript material as we have it does not allow us to draw conclusions about whether Tocharian A was spoken with a Tocharian B accent, or the other way round, or whether there was no interference in this domain at all. The only observation we can make is that on the synchronic level the phonological systems are almost exactly the same, and the spelling is identical. In fact, the script is also identical, if we leave archaic Tocharian B out of consideration. The script type of Tocharian A suggests that it was based on a Tocharian B script contemporaneous with classical Tocharian B, but probably not with the earliest stage of it, that is to say, a script type from the second half of the seventh century or later.

§3.3 In the domain of morphology, we again find a relatively large divergence between the two languages. We can quote George Lane, who stated that “It is indeed in morphology that the two dialects go most widely asunder” (1966: 215). I cannot repeat all his examples here, but he adduces all kinds of differences in nominal and verbal endings, and divergences in the formation patterns of the verb.

Recently, Malzahn and Pinault have argued that Tocharian A influenced Tocharian B in its morphology, which is surprising in view of the general direction of borrowing from Tocharian B to Tocharian A. For Malzahn’s idea that the aberrant *au*-vocalism of the eastern and late form *prautkar* B108b6 ‘they (were) filled’ is due to Tocharian A influence (2007c: 290), I refer to my suggestion that it is a causative form (2008: 133, 209). Another instance was brought up by Pinault (2008: 643), who assumed that the late 3pl. copula *stare* was formed from the suffixed singular *star-*. According to him, this means that the singular and plural copula forms had merged functionally, which in turn would point to influence from Tocharian A, where we find the same functional overlap (Sieg, Siegling and Schulze 1931: 167). However, the intermediate form that is indifferent to number is not attested and *stare* is the free rather than the suffixed form (it is attested with a suffixed pronoun only once). Moreover, *stare* is not an eastern form, but a late one (it is well attested in the Kuča region), and it can be explained without recourse to the singular *star-* (Peyrot 2008: 141-142). The second instance Pinault adduced (2008: 185) concerns a case mistake in an eastern text, where we find the nom.sg. *oski[ye]* B108a9

‘dwelling’ in oblique function as proven by the preceding adjective *paiyneṣṣai*. Since the corresponding Tocharian A word *oṣke* has no difference between nominative and oblique, Pinault explains *oskiye* as a hypercorrection, assuming that the Tocharian A writer has mixed up the Tocharian B case forms. It is difficult to find an alternative explanation for *oskiye*, as no other interpretations of the passage seem possible,<sup>12</sup> but a confusion of nominative and oblique may evidently have other causes, too. In any event, if these phenomena should indeed be due to influence from Tocharian A, it does not yet prove that Tocharian A was a spoken vernacular. As it is plausible that the two languages were in part written by the same scribes, a little bit of confusion may have come about.

§3.4 Unfortunately, there is no systematic description of the syntax of either Tocharian language, let alone a comparative syntax of both. However, we are in the lucky position to have many case studies: apart from Thomas’ publications (which can be traced via Peyrot 2007), we may mention for instance Schmidt (1974) and Carling (2000). Admittedly, it is difficult to generalize from all these case studies, as the results vary, naturally, and the scope, size and depth of the inquiries are very different as well. Still, a tentative conclusion can be that if the formal inventories of the two languages do not match, we do obviously find differences in their usage, but if the formal inventories of the two languages are the same, the differences in usage are slight at most.

The method for detecting loan-words in the two languages was an evaluation of the systematic differences between Tocharian A and B which yielded some words that do not conform to the system because they are too similar (§3.1). Although we cannot apply exactly the same method to syntax, the lack of differences in this domain is striking, so that it is possible that at least some of the syntactic matches do not reflect the Proto-Tocharian state of affairs. Although we will

---

<sup>12</sup>The only solution I can think of is that we should perhaps read *osk[aine]* (loc.sg.), which would resolve the double accusative reconstruction, translating ‘in the dwelling of your feet we make our refuge’. However, this text correction cannot be verified since the manuscript is lost; its only support is that Sieg and Siegling’s transliteration “*oski[ye]*” (1953: 44) suggests that the passage was difficult to read. In any case, Thomas’ analogy with *prašciye* ‘rain’ (1964: 46) is inadequate: although this word may indeed lack a difference between nominative and oblique, it is masculine instead.

never be able to tell on the basis of similarities between the daughter languages *that* the proto-language was different, caution is called for: a superficial match between Tocharian A and B does not necessarily mean that Proto-Tocharian was identical.

#### §4 Discussion

In conclusion, there is an imbalance in the degree of divergence between Tocharian A and B in the different linguistic domains. The differences in morphology and lexicon are relatively large, whereas those in phonology and syntax are small. In order to explain this imbalance, I have three suggestions to make, which are not necessarily mutually exclusive.

First of all, certain syntactic similarities may be due to parallel copying from Sanskrit patterns, as many Tocharian texts are translated from Indic.<sup>13</sup> Even if a text is not a translation from Sanskrit, but for instance a Tocharian (re)composition, we can safely assume that Indian models had a strong influence. Although this suggestion does not work for syntactic agreements not matched by Sanskrit, it is very attractive because influence from Sanskrit is guaranteed, not only by numerous loan-words, but also in the form of a large number of lexical and syntactic calques in certain texts, for instance the Udānavarga and Udānālaṅkāra in Tocharian B. Candidates for Indian influence may be the usage of participles, the usage of certain cases, and probably some word order patterns.

Second, agreements between Tocharian A and Tocharian B in syntax and phonology could be due to a Tocharian B substrate in Tocharian A. Although there are many types of language contact, depending on the social situation and the types of the languages, and despite the fact that there is no ready solution for what a Tocharian B substrate in Tocharian A should look like, it is plausible that the phonetics and syntax of the substrate language were the domains that influenced the ultimate outcome most. Admittedly, it is difficult to offer hard proof for this suggestion, especially because we have little information about the precise nature of the contacts between the two languages, and their predecessor Proto-Tocharian is only reconstructed. Nevertheless, influence of Tocharian B on

---

<sup>13</sup>Mostly at a rather late stage, that is, from Buddhist Sanskrit.

Tocharian A is ascertained by the loan-words and the similarities in the script types. As an additional advantage, the substrate theory accounts for the high mutual translatability: it is usually possible to make very precise translations from one language into the other. Candidates for Tocharian B influence may be the usage of tenses and moods, and possibly adverbs and particles.

The question is how the substrate could have come about. In my view, the most plausible scenario is that the Tocharian A texts were written by professional Tocharian B writers, commissioned by Tocharian A nobles. This explains the spread of Tocharian B to the east as instrumental in the development of Tocharian A writing. It also explains why the influence of Tocharian B on Tocharian A took place before the oldest attested Tocharian A text was composed, in spite of the demonstrably late date of the earliest contacts. In addition, it allows for a largely independent Tocharian A tradition as far as the content of the texts is concerned, which in turn could be the reason why there are hardly any parallel texts in the two languages, nor translations from one into the other.

Third, it cannot be excluded that there is a common foreign element which makes Tocharian A and B look similar. For instance, it has often been suggested that the agglutinative features in the Tocharian nominal system are of the Altaic type. Although we do not know where and when contacts with agglutinative languages should have taken place, it is possible that these features were borrowed only after Proto-Tocharian broke up. Independent borrowing of features from a common foreign source would account for differences in the secondary cases, for instance, case suffixes that have the same function but are not etymologically related.

### **§5 Conclusion**

To sum up, strong influence of Tocharian B on Tocharian A is certain and no speakers of Tocharian A can be traced through the manuscripts. The two languages must have converged to a certain degree and it is even possible that Tocharian A was actually written by speakers of Tocharian B. Consequently, Proto-Tocharian may have differed more from its daughter languages than is often suggested by superficial similarities between them. Therefore we should be very cautious when trying to reconstruct Proto-Tocharian syntax, especially when Tocharian A and B agree perfectly.

## References

- Carling, Gerd  
 2000 *Die Funktionen der lokalen Kasus im Tocharischen*. Berlin – New York: Mouton de Gruyter.
- Couvreur, Walter  
 1947 *Hoofdzaken van de Tochaarse klank- en vormleer*. (Katholieke Universiteit te Leuven, Philologische Studiën, Teksten en Verhandelingen II, 4) Leuven: Beheer van Philologische Studiën.
- Gershevitch, Ilya  
 1954 *A grammar of Manichean Sogdian*. Oxford: Blackwell.
- Lane, George S.  
 1966 On the interrelationships of the Tocharian dialects. In: Henrik Birnbaum and Jaan Puhvel (eds.) *Ancient Indo-European dialects, Proceedings of the conference on Indo-European linguistics, held at the University of California, Los Angeles, April 25-27, 1963*, 213-233. Berkeley and Los Angeles: University of California.
- Malzahn, Melanie  
 2007a A preliminary survey of the Tocharian glosses in the Berlin Turfan collection. In: Melanie Malzahn (ed.) *Instrumenta Tocharica*, 301-319. Heidelberg: Winter.  
 2007b Tocharian desire. In: Alan J. Nussbaum (ed.) *Verba Docenti, Studies in historical and Indo-European linguistics presented to Jay H. Jasanoff by students, colleagues, and friends*, 237-249. Ann Arbor – New York: Beech Stave.  
 2007c The most archaic manuscripts of Tocharian B and the varieties of the Tocharian B language. In: Melanie Malzahn (ed.) *Instrumenta Tocharica*, 255-297. Heidelberg: Winter.
- Maue, Dieter  
 2009 Uigurisches in Brāhmī in nicht-ugurischen Brāhmī-Handschriften. *Acta Orientalia Academiae Scientiarum Hungaricae* 62: 1–36.
- Pedersen, Holger  
 1941 *Tocharisch vom Gesichtspunkt der indoeuropäischen Sprachvergleichung*. (Det Kongelige Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser 28,1) København: Munksgaard.
- Peyrot, Michaël  
 2007 Index to the translations by Werner Thomas. In: Melanie Malzahn (ed.) *Instrumenta Tocharica*, 129-161. Heidelberg: Winter.  
 2008 *Variation and change in Tocharian B*. (Leiden Studies in Indo-European 15) Amsterdam – New York: Rodopi.

Pinault, Georges-Jean

- 2005 Impératif et exhortation en tokharien. In: Gerhard Meiser and Olav Hackstein (ed.) *Sprachkontakt und Sprachwandel, Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17.-23. September 2000, Halle an der Saale*, 495-523. Wiesbaden: Reichert.
- 2007a Concordance des manuscrits tokhariens du fonds Pelliot. In: Melanie Malzahn (ed.) *Instrumenta Tocharica*, 163-219. Heidelberg: Winter.
- 2007b Le tokharien pratiqué par les Ouïgours: à propos d'un fragment en tokharien A du Musée Guimet. In: Jean-Pierre Drège and Olivier Venture (eds.) *Études de Dunhuang et Turfan*, 327-366. Genève: Droz.
- 2008 *Chrestomathie tokharienne, textes et grammaire*. (Collection linguistique publiée par la Société de Linguistique de Paris 95) Leuven – Paris: Peeters.

Schmidt, Klaus T.

- 1974 *Die Gebrauchsweisen des Mediums im Tocharischen*. Diss. Göttingen.
- 1983 Vorläufige Bemerkungen zu den in der Staatsbibliothek Preußischer Kulturbesitz in Berlin neu gefundenen tocharischen Handschriftenfragmenten. *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, Supplement V: 271-279.

Sieg, Emil and Wilhelm Siegling

- 1953 *Tocharische Sprachreste. Sprache B, Heft 2, Fragmente Nr. 71-633, Aus dem Nachlaß herausgegeben von Werner Thomas*. Göttingen: Vandenhoeck & Ruprecht.

Sieg, Emil, Wilhelm Siegling and Wilhelm Schulze

- 1931 *Tocharische Grammatik*. Göttingen: Vandenhoeck & Ruprecht.

Thomas, Werner

- 1964 *Tocharisches Elementarbuch, Band II, Texte und Glossar*. Heidelberg: Winter.

Winter, Werner

- 1962 Lexical interchange between 'Tocharian' A and B. *Journal of the American Oriental Society* 81: 271-280.
- 1963 Tocharians and Turks. In: Denis Sinor (ed.) *Aspects of Altaic civilization, Proceedings of the Fifth Meeting of the Permanent International Altaistic Conference Held at Indiana University, June 4-9, 1962*, 239-251. (Uralic and Altaic Series 23) Bloomington – The Hague: Indiana University.