Universiteit
Leiden
The Netherlands

Traces of language contact in Niya Prakrit: Bactrian and other foreign elements

Schoubben, N.

Citation
Schoubben, N. (2024, November 6). *Traces of language contact in Niya Prakrit: Bactrian and other foreign elements*. Retrieved from https://hdl.handle.net/1887/4108454

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/4108454

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

CHAPTER 6 The "Tocharian C" hypothesis

The data presented in Chapters 3–5 allow for the conclusion that Bactrian and other Iranian languages had a profound influence on Niya Prakrit. Yet, scholars believe that Niya Prakrit has also been influenced by non-Iranian languages, most famously by a hypothetical substrate language known as "Tocharian C". After a brief introduction to the topic (§6.1), this chapter re-examines the lexical, phonological, morphological, and syntactic arguments that have been used to justify the assumption of contact with Tocharian (§§6.2–6.5). This re-examination will enable us to conclude that there is, with the potential exception of one loanword, so far no compelling evidence in favour of Tocharian influence on Niya Prakrit, be it as the result of a substrate or an adstrate type of contact (§6.6).

6.1 Introduction

6.1.1 A research history

Burrow (1935b; LKD: viii–ix) seems to have been the first scholar seriously to consider the possibility of language contact between Tocharian and Niya Prakrit. According to him, a language related to, but different from, Tocharian A and B was the native language of the people inhabiting the Shanshan kingdom in the first centuries of our era (cf. §1.1.2). This hypothetical substrate language, the so-called "Tocharian C", would account for otherwise unexplained nouns and personal names in the Niya documents (§6.2); Niya Prakrit's phonological idiosyncrasies (§6.3); and some grammatical features not easily explained as inherited from OIA (§§6.4–6.5).[892]

If correct, Burrow's hypothesis would have "major implications for … the problem of Tocharian origins" (Mallory 2015: 7). Yet, despite its potential, it has not been the subject of any extensive scrutiny ever since (cf. §§1.1.2–1.1.3). A few scholars have tried to augment the evidence in favour of Burrow's theory (e.g. Harmatta 1967: 11ff.; van Windekens 1971); some authors have accepted it without much reserve (Lin 1996: 196; Carling 2005:

[892] Burrow (LKD: viii) proposed to call the substrate language "Krorainic", after the ancient name of Loulan, the capital of the Shanshan kingdom. Burrow's hypothesis should not be confused with Schmidt's (2018: 161–271) unpersuasive proposal that some fragments in Formal Kharoṣṭhī are written in a third variety of Tocharian, on which see Dragoni, Schoubben & Peyrot (2020).

47; Adams 2017[2]: 453; Kogan 2017: 232); and still others, including me, have taken a more cautious stand in the debate (e.g. Hitch 2009: 5; Schoubben 2021; Peyrot 2022: 86f.).

An unpublished handout by Giorgio Banti (2000) represents an exception to this lack of serious scholarly debate with Burrow's theory. Even though a precious contribution to the topic, Banti's study can still be refined on various levels. Banti did, for example, not discuss every argument in the amount of detail it deserves, and he left out some potentially relevant pieces of evidence, e.g. the feminine marker *-s-* (§6.4.4). Besides, a few of Banti's own suggestions can be further improved upon by paying more attention to philological details in Niya Prakrit and by including pertinent material from Bactrian published in the past two decades. There is thus more than sufficient reason to re-examine Burrow's "Tocharian C" hypothesis.

6.1.2 "Tocharian C" and Niya onomastics

While I have tried to collect all the relevant arguments for "Tocharian C" proposed by Burrow and others,[893] I will tacitly ignore Tocharian etymologies for Niya Prakrit personal names.

§3.4.1 already briefly treated the methodological issues inherent in the study of onomastic material. The problems involved are, however, still more pronounced in the present case, as the principles of Tocharian onomastics are not yet well understood (cf. Krause 1953; Pinault 1987b). Whereas Iranian onomastics has been studied in depth over the past decades based on a considerable body of evidence, onomastic material in Tocharian has long been limited by the lack of publications of secular documents. In Tocharian literary texts, one mainly encounters Buddhist names of Sanskrit or Prakrit origin, and even in the secular documents, one has to carefully distinguish genuine Tocharian names from those of Indian, Iranian, Chinese, Turkic, or unknown origin. Due to these methodological difficulties, etymologising personal names as Tocharian generally does not exceed mechanical comparisons with Tocharian lexical material matching phonetically more or less with the name in question.

[893] Because some of the potential Tocharian elements in Niya Prakrit are more easily explained as borrowings from Bactrian, they have already been discussed more extensively in Chapters 3–5. Therefore, it suffices to summarise the issues with their Tocharian etymologies here.

Besides, even if a couple of Tocharian personal names were attested in the Niya documents, these do not need to point to a Tocharian substrate. Contact between the Shanshan kingdom and the Tocharian-speaking areas in the north of the Tarim Basin is known to have existed (Høisæter 2020: 367–369). Therefore, travellers or even immigrants from these northern areas may be mentioned in the Niya documents, and, if so, the occurrence of some Tocharian names in the Niya corpus is no conclusive proof that the natives of the Shanshan kingdom spoke a form of Tocharian.

6.2 Lexicon

This section critically reviews Burrow's lexical arguments in favour of "Tocharian C" as well as Tocharian etymologies for Niya Prakrit words proposed by other scholars.[894] One of the proposed etymologies, i.e. Burrow's idea to compare the tile *kitsaïtsa* with TB *ktsaitse* 'old' (and TA *ktsets* 'perfect, accomplished'), seems conceivable. All the others are, in my view, too uncertain to base any conclusions on.

1. *aṃklatsa,* a type of camel

Attestations of the base form

DIR.SG. *aṃklatsa*: CKD 163; 195; 383; 561; 590; 592; 897; *aklatsa*: CKD 569; *agiltsa*: CKD 422

DIR.PL. *aklatsa*: CKD 330

The technical term *aṃklatsa*[895] is always used to denote a specific type of camel. The currently available evidence does, however, not permit more precise conclusions on its semantic range (cf. also Lüders 1936b: 647–652).

---

[894] I leave aside TA *ṣoṣṭaṅk** 'finance officer' and Niya *ṣvaṭhaṃgha* / *ṣoṭhaṃga* (LKD: ix), as, in this case, the Tocharian form is most likely borrowed from Niya Prakrit; cf. §3.2.4 s.v.

[895] In early secondary literature, the reading of this word was still *a(ṃ)kra'tsa* (e.g. KD: 329). *kla* is the current transliteration of 𐨒, an uncommon Kharoṣṭhī sign that consists of a default 𐨑 *ka* and an additional stroke crossing the stem of the 𐨑 *ka* horizontally. In CKD 897, the same akṣara occurs in a word which Jiang (2020: 135f.; 140f.) has read as *kulaśukra* and interpreted as a compound of *kula* 'family' and *śukra* 'bright' < Skt. *śukla-*. As the context is about marriage, I think that Jiang's reading should be corrected to *kulaśulka* 'bride-price'. If this interpretation is correct and 𐨒 could also stand for *lka*, one can compare the character 𐨔 *lpa*, which is formed by adding a St. Andrew's cross to the stem of a standard 𐨔 *pa*.

As part of his argumentation on "Tocharian C", Burrow (1935b: 675; LKD: 71 s.v. *aṃkra'tsa*) proposed to equate *aṃklatsa* with TA *āknats* / TB *aknātsa* 'stupid, foolish' (DoT: 3; DTTA: 27). His idea was that *aṃklatsa*, formed with the negative prefix **en*- and the verbal root **kna*- 'to know', would refer to an "untrained camel". As an alternative for the semantics, Banti (2000: 6) hypothesised that 'foolish' would be a way to characterise young camel bulls who tend to act foolishly due to sexual excitement in their first rutting period. Somali *waal-ay* 'the fool one', also used for young male camels, would count as a semantic parallel.

I nevertheless agree with Peyrot (2022: 86 with fn. 3) that Burrow's Tocharian etymology is "not convincing". For his derivation to work, Burrow has to resort to the ad hoc assumption that the source of *aṃklatsa* would have undergone a dissimilation from *-*nkn*- to *-*nkl*-. In addition, the by-form *agiltsa* is not easily explained from the supposed Tocharian source.[896] While being attested only once, this alternative form requires due attention because it occurs in CKD 422, the earliest Niya document at our disposal and one displaying a number of archaic linguistic features (cf. Brough 1965: 594 = 1996: 289). It therefore seems quite possible that *agiltsa* represents a more original form of the word than *aṃklatsa*, which would make the comparison with TA *āknats* / TB *aknātsa* formally even less attractive.

For historical reasons, one may want to consider that *aṃklatsa* is instead a Bactrian word. The association of camels with Bactria reflected in the term "Bactrian camel" goes back all the way to Aristotle (*Historia animalium* 498b.9), and camels were even used as a symbol for Bactria on Indian coins (Falk 2015–19: 18).[897] Yet, no convincing etymology has presented itself so far.

2. *o̱gana*, an agricultural product

Attestations of the base form

DIR.SG. *o̱gana*: CKD 154; 207; 713 (2x); 714

[896] Thomas' (1944: 65) alternative proposal that *agiltsa* would instead be the opposite of *a(ṃ)klatsa* and connected to TA *āklye* 'learning' (TB *akalye*) has little to recommend itself.

[897] Note also that another unexplained term for camels, i.e. *put̨getsa*, contains the typically Bactrian cluster -*t̨g*- (← -δγ- /dg/), perhaps also the same element -*tsa* as *aṃklatsa*.

Attestations as a PN

DIR.SG. *Ogana*: CKD 286; 701
GEN.SG. *Oganaṣa*: CKD 562 (2x)

Although the exact semantics of *ogana* remain unclear, there is sufficient reason to believe that it refers to an agricultural product. Because of its meaning, Burrow (1935b: 673; LKD: 81 s.v.) connected *ogana* to TB (→ TA) *oko* 'fruit' (DoT: 115; DTTA: 78f.) and TA *okar* 'herb, plant' (DTTA: 76). Yet, this etymology is unconvincing, as it leaves the element *-ana* in *ogana* without an explanation.[898]

According to Bailey (1946a: 794f.), *ogana* would be a hypercorrect spelling of **oyana* < Skt. *odana-*, which he translates as 'rice'. *odana-* means, however, more precisely 'porridge, boiled rice', whereas *ogana* seems to be an unprocessed food product.[899] In addition, there are no good parallels for the phonological development assumed by Bailey.[900] Consequently, also his etymology is not persuasive.[901]

3. *kitsaïtsa*, a title

Attestations of the basic form

DIR.SG. *kitsaïtsa*: CKD 187 (2x); 245; 327; 377; 437 (2x); 495 (2x); 549; 571; 572; 579; 580; 581; 582; 586; 587; 589; 590; 606; 640; 648 (2x); 654; 655; 656; 666; 715; 719; 730; 797 (2x); 809; 849; *kitsayitsa*: CKD 436; 561 (2x); 859; 879; *kiṃtsayitsa*: CKD 415 (2x); *kitsatsa*: CKD 574 (2x); 652; 704; 808
GEN.SG. *kitsaïtsaṣa*: CKD 579

---

[898] *oko* is mainly used metaphorically as the Tocharian equivalent of BSkt. *phala-* 'fruit of an action' (DTTA: 78f.; DoT: 115). Yet, Ching (2016: 56) points to a few passages in administrative texts where *oko* refers to actual cereals.

[899] The normal Sanskrit word for 'rice' is *vrīhi-*. Although the Chinese Annals and Xuanzang's travelogue record the existence of local rice cultivation, references to rice in administrative texts from the Tarim Basin are scarce to non-existent (Ching 2010: 67f.). TA/TB *klu* 'rice', for instance, never occurs in administrative texts (Ching 2016: 55 fn. 55; Peyrot 2018a: 254f.).

[900] Bailey cites Gāndhārī loanwords into Khotanese that show a loss of intervocalic **d*. As **d* is, however, also lost in inherited words, as in Khot. *pāa-* 'foot' < **pāda-*, Bailey's examples are not necessarily informative on Gāndhārī.

[901] Formally, it would be attractive to derive *ogana* from a cognate of Oss. D. *(w)ogæn* 'secret hiding-place, stock' < **awa-kana-* (Cheung 2002: 243). Yet, I have not found evidence for **awa-kana-* in the required agricultural sense.

GEN.PL. *kitsaïtsaṃna*:[902] CKD 671

Being an official title from the judicial sphere, *kitsaïtsa* has been compared since Burrow (1935b: 673; LKD: 82 s.v.) with TB *ktsaitse* 'old' and TA *ktsets* 'perfect, accomplished' (DoT: 263; DTTA: 175). As this etymology is semantically and formally conceivable, various scholars have accepted it, be it with more or less hesitation.[903]

Also approving of Burrow's etymology, Peyrot (2022: 86) has recently discussed what the exact source form of *kitsaïtsa* may have looked like. He sees three options. First, *kitsaïtsa* may reflect an archaic form *$kət^{s}ait^{s}t^{s}e$ with a rendering of **ə* as *i* and adoption of the word into the productive *a*-class declension in Niya Prakrit. Second, a source comparable to TB *ktsaitse* may have been borrowed with *i*-epenthesis in the first syllable (similarly Banti 2000: 6). Third, depending on the chronology of monophthongisation in Tocharian A, the Niya title may also come from an earlier form of *ktsets* where **ai* was still preserved.

Irrespective of which scenario one opts for, *kitsaïtsa* need not be the result of substrate borrowing from an unattested "Tocharian C": it could instead be due to some form of adstrate borrowing.

4. *kilme* 'family estate'

According to Burrow (1935b: 674f.; LKD: 83 s.v.), *kilme* 'family estate' is to be compared to TA *kälyme* and TB *kälymiye* 'region, direction' (DoT: 186f.; DTTA: 131f.). While this etymology poses no major problems on the formal side, *kilme* is no decisive evidence in favour of Tocharian influence on Niya Prakrit because of the existence of competing etymologies. See §3.2.4 s.v. for a more detailed discussion.

5. *tasuca*, a title

Attestations of the base form

DIR.SG. *tasuca*: CKI 170; CKD 26; 78; 103; 126; 143 (2x); 147; 157; 160; 187; 259; 375 (2x); 390; 422 (2x); 436; 494 (2x); 495; 506; 507; 517; 525; 531 (2x); 533; 552 (2x); 554; 558; 580; 588; 637 (3x); 639 (2x); 648; 676 (2x); 709; 749; 774; 788; 805; 808 (2x)

---

[902] This plural form proves that Atwood (1991: 176 fn. 11) was mistaken in claiming there to have been only one *kitsaïtsa* at a time.

[903] Cf. e.g. Thomas (1944: 66), Banti (2000: 6), Padwa (2007: viii; 271; 280), and DoT (263).

GEN.PL. *tasucana*: CKD 884

Compounds

DIR.SG. *rajyatasuca* '*tasuca* of the kingdom': CKD 883 (2x)

Despite a wealth of attestations of this official title, there is still no good evidence for the exact meaning of *tasuca* (cf. LKD: 94 s.v.). Loukota (2020: 110) tentatively proposed to explain *tasuca* as an agent noun in *-uca* derived from the Tocharian B root *tās-* 'to put, to place'.[904]

There are a number of linguistic problems with this suggestion. First, 'placer' is a semantically not entirely satisfactory denomination of an official title. In addition, the root *tās-* is not known to form an *uca*-agent noun in Tocharian, the agent noun of this verb instead being construed with the synonymous suffix *-ñca*, i.e. *taṣ(ṣ)eñca.* As a rule, agent nouns in *-uca* are moreover derived from the subjunctive V stem, which is *tättā-* for the root *tās-* (Malzahn 2010: 641f.; Peyrot 2013: 757).[905] In the unlikely case that an agent noun in *-uca* to the root *tās-* existed, the required form would thus have been **tättauca* or the like, which is clearly incompatible with *tasuca*.

6. *paḱe* 'allotment, share'

Burrow identified *paḱe* 'allotment, share' with TB *pāke* 'parcel, share' (LKD: 102 s.v.). This etymology does, however, not account for the consistent spelling of *paḱe* with 𐨎 *ḱa*, and *paḱe* can be more easily derived from Bactr. *παþκο /pəškə/ 'share' < **pišta-ka-* 'that which is cut'. Compare Schoubben (2021: 52f.) and §3.2.3 s.v.

7. *maḱa* 'bean'

Burrow (1935b: 673; LKD: 110 s.v.) connected *maḱa* to TA *malke* 'milk' (DTTA: 332). Yet, this proposal suffers from both phonological and semantic problems, discussed in more detail in Schoubben (2021: 50–52) and §3.2.2 s.v. A more likely source is provided by Bactr. *μαþκο /māškə/ 'bean' < **māša-kă̆-*.

---

[904] For the Tocharian B agent nouns in *-uca*, see now Friis (2024: 271–276).

[905] Michaël Peyrot (p.c.) kindly points out that the subjunctive V stem an *uca*-participle is derived from usually also forms a pair with a nasal present, a formation likewise unattested for the root *tās-*.

8. *śuka*, a type of wine

Attestations of the base form

ADJ. *śuka*: CKD 272 (2x); 283 (?); *śuki*: CKD 169; 349; 431 = 432 (2x); *śukha*: CKD 387[906]

In all its attestations, *śuka* modifies the noun *masu* 'wine' (for which cf. §3.7.1 s.v.). On this basis, we may presume that *śuka* was an adjective specifying a particular type of wine (LKD: 125f. s.v.).

According to Burrow, *śuka* would be an inherited word, namely from OIA *śuṣka-* 'dry' (LKD: 125f. s.v.).[907] The variant *śukha* could then be the phonologically regular outcome and *śuka* a deaspirated form (cf. *nikhal-* next to *nikal-* 'to take out, provide' < *niṣkālaya-*).[908]

There is, however, a problem with Burrow's idea, in that there is no parallel for the concept of 'dry wine' (as in French *un vin sec*) in the Tarim Basin languages of the first millennium. Aware of this issue, Burrow points to the mentioning of dry grapes in Tibetan documents from the region (citing Thomas 1934b: 475), to which may be added the Khotanese collocation *hauṣka gūra* 'dry grapes' (DKS: 87; 494). *Pace* Burrow, it would, however, be unwarranted to assume on this basis that *masu* means 'grape' when combined with *śuka*, given that in all other instances the meaning of *masu* is clearly 'wine' (as admitted by Burrow).

An additional drawback to Burrow's idea is that other adjectives describing types of 'wine', such as *potga* (*potgoña*) and *ṣamiyo*, are clearly not inherited (§3.7.2 and §6.2 s.vv.). It therefore seems worth considering that *śuka* is a loanword too.

The possibility of a loanword was also entertained by van Windekens (1971; 1976: 538), who proposed that *śuka* was borrowed from a "Tocharian C" cognate of TA *śuk* and TB *śuke* 'taste, juice, nectar' (~ Skt. *rasa-*, cf. DoT:

---

[906] The compound *śukamuli* 'dues payment' (CKD 59; 211; 309) should be put aside. In this case, *śuka-* may instead come from *śulka-* 'price' (TKD: 13; 40; 56).

[907] Burrow mentions two more possibilities, i.e. OIA **śukna-* 'fermented' next to *śukta-*, like Pkt. *mukka-* 'released' < **mukna-* next to Skt. *mukta-*, and *śukla-* 'bright'. The derivation from **śukna-* would be semantically understandable, as *śukta-* can also mean 'vinegar' (cf. also Romani *šuk* 'idem'); but because of the absence of other MIA and/or NIA derivatives of the variant form **śukna-,* I find this etymology not fully persuasive. With *śukla-*, there is the problem that in other varieties of Gāndhārī this word is continued as *śukra* rather than as ***śuka* /śukka/ (cf. Baums & Glass 2002a– s.v.).

[908] Burrow does not account for the variant *śuki.*

692f.; DTTA: 472); *śuka masu* would thus literally mean 'tasty wine', i.e. 'wine of a high quality'. Formally, this etymology could work (it is accepted by Lin 1996: 196); but it can hardly be called compelling seeing that the meaning attributed to *śuka* is solely based on etymology, and *śuka* is an adjective, while the Tocharian words are nouns. Moreover, even if van Windekens' idea would turn out to be correct, this etymology can still not be used as a conclusive indication of a "Tocharian C" substrate, for a word of this kind could easily have been borrowed as the result of trade.

9. *ṣamiyo*, a type of wine/corn

Attestations of the base form
DIR.SG. *ṣamiyo*: CKD 225
INS.SG. *ṣamiyena*: CKD 637; 703 (2x); 796; 797[909]

ṣamiyo (ins.sg. *ṣamiyena*) is always combined with either *aṃna* 'corn' (so in CKD 703) or *masu* 'wine' (so elsewhere). Therefore, this adjective presumably indicates a specific type of corn and wine, but it is a "completely obscure term" otherwise (LKD: 126f. s.v. *ṣamiyena*).

According to Harmatta (1967: 13f.), *ṣamiyo* would come from **ṣmeyo*, a derivative of TA *ṣme** 'summer' (DTTA: 502) with a suffix *-o*.[910] The variant *ṣamiyena*, by contrast, would be borrowed from **ṣmeyeṃ*, an adjective derived from the same word for 'summer' with the suffix *-eṃ*, as in TA *oṅkalmeṃ* 'of the elephant' ← *oṅkaläm* 'elephant' (DTTA: 80). These etymologies require two phonetic substitutions to have happened. The un-Indic initial cluster **ṣm-* must have been broken up with a prop vowel *-a-*, so as to result in Niya Prakrit *ṣam-*, and the vowel *-e-* must have been raised to *-i-*, as we do not have Niya ***ṣameyo* or ***ṣameyena.*

Harmatta's etymology is to be rejected for several reasons. To begin with, it is unwarranted to posit both **ṣmeyo* and **ṣmeyeṃ* for the donor language, none of which is attested. In fact, it seems more likely that *ṣamiyena* is an ins.sg. in *-ena*, and *ṣamiyo* the basic form (cf. §3.7.2 s.v. *potg̲e*; *potg̲a* / *potg̲oña*). Moreover, the resulting meaning 'summer wine' or 'summer corn'

[909] *ṣamiya* in CKD 770 is probably unconnected.
[910] Although Harmatta is not very explicit, it seems he is referring with this suffix *-o* to Tocharian A adjectives like *ekro* 'poor' ← *ekär* 'empty, deprived' (DTTA: 68), *mkälto* 'little, small' (DTTA: 355), and *tālo* (cf. TB *tallāu*) 'miserable' (DTTA: 208).

does not receive any support from the contexts (cf. supra). I therefore conclude that a convincing etymology for *ṣamiyo* remains to be found.

10. *ṣilýoga* 'document, written tablet'

Attestations of the base form

DIR.SG. *ṣilýoga*: CKD 470
INS.SG. *ṣilýogena*: CKD 492
LOC.SG. *ṣilýogaṃmi*: CKD 140; 359

Compounds

DIR.SG. *ṣilýogalihitaga* 'document, written tablet': CKD 470; *ṣilýogalihidaga*: CKD 471; *ṣilýokalihitaga*: CKD 359; *ṣilýokalihidaga*: CKD 312; 561 (2x); *ṣulgalihidaga*:[911] CKD 582
DIR.SG. *ṣilýogahastalekha* 'idem': CKD 729

The available contexts leave no doubt that the general meaning of *ṣilýoga* is 'document, written tablet'. According to Lüders (1936b: 652–654; so also LKD: ix; 127 s.v. *ṣilṕoga*), *ṣilýoga* needs to be compared with TA *ṣlyok* 'strophe', next to *ślok* ← Skt. *śloka*- (DTTA: 484; 503).[912] As *śloka*- would have resulted in Niya **śiloga*,[913] Lüders proposed that Niya Prakrit reborrowed this word from Tocharian.

However, not only is such a reborrowing *a priori* implausible (Tremblay 2001: 35 fn. 54), there is the additional problem that neither TA *ṣlyok* nor Skt. *śloka*- or one of its MIA/NIA derivatives ever means 'official document' or the like (CDIAL: no. 12748). It thus seems advisable to disconnect TA *ṣlyok* and Niya *ṣilýoga*, the etymology of which remains to be clarified.

6.3 Phonology

Burrow (1935b: 667–671) started his article about potential Tocharian elements in Niya Prakrit with an overview of eight phonological features

[911] The variant *ṣulgalihidaga* is most likely a lapse of the pen. The last few lines of CKD 582 contain two other scribal mistakes, i.e. loc.sg. *Caḍotaṃtimi* instead of *Caḍotaṃmi* 'in Caḍota' and the personal name *Ramaṣtso* (also gen.sg. *Ramaṣtsoaṣa* instead of *Ramṣotsa* (also gen.sg. *Ramṣotsaṣa*).
[912] There is no immediately obvious distinction between TA *ślok* and *ṣlyok*, and both forms occur in parallel sentences in A 353. Tocharian B invariably has *ślok* 'stanza, strophe'.
[913] Cf. the SA Gandh. ins.sg. *śilogano* in Av^{L6} recto 25.

which the supposedly non-Indian and non-Iranian nouns and personal names in the Niya documents would seem to share with Tocharian.[914]

At first sight, Burrow's comparisons with Tocharian phonology seem solid evidence in favour of his substrate theory. Yet, most of the features he listed are not exclusive to Tocharian: they are also found in other potentially relevant languages, notably Bactrian. Another methodological problem with Burrow's approach is that the unexplained nouns and personal names in Niya Prakrit probably do not all derive from the same source. For instance, Burrow's supposedly non-Iranian nouns also include words now explainable as Bactrian, making it difficult to decide which features are representative of the phonology of the supposed substrate language and which are not.

In what follows, I shall present and comment on each of Burrow's arguments individually. The conclusion will be that none of them is compelling evidence for a Tocharian substrate.

6.3.1 The devoicing of voiced stops

Burrow's first phonological argument for "Tocharian C" concerns the devoicing of voiced stops in Niya Prakrit (1935b: 667f.). In §2.1.4 and §4.3.28, I pointed out that, both in inherited lexemes and in loanwords, voiced and unvoiced stops can sporadically interchange in Niya Prakrit. Burrow used this phonological peculiarity as evidence that the language of the Shanshan people lacked voiced stops, if correct, a remarkable isogloss with Tocharian, as the latter too has only unvoiced stops in its phonemic inventory.

Devoicing initial consonants is probably the strongest phonological argument in favour of a Tocharian substrate in Niya Prakrit.[915] Yet, it can hardly bear the burden of proof on its own, especially since there are

[914] Burrow's ninth item, i.e. "no attempt is made to represent the reduced vowel *ä*, which is common in Tokharian" (1935b: 671), is only an observation and not really an argument in support of "Tocharian C". I therefore leave it aside.

[915] It would be difficult to argue that contact with Bactrian caused this type of devoicing, seeing that, in the case of Sanskrit and Gāndhārī loanwords into Bactrian, original voice distinctions are usually well preserved (cf. e.g. βοδδο / βουδο ← *buddha*- or δηβο 'divinity' ← *deva*-). In non-initial position, there are some peculiar forms in Bactrian, e.g. βοτο (also MBactr. *bwt*) ← *buddha*- or the name βατρο ← *Bhadra*- (BPN: no. 75). I doubt, however, that these are connected to the devoicing in Niya Prakrit. Loanwords from other Iranian languages into Bactrian also preserve the original voicing; note e.g. Δηβαυρο 'arrest' ← Parth. *dēbahr* or MBactr. *b'gyg* 'having a share' ← Parth. (?) **bāgīg* (cf. Sims-Williams 2009a: 258).

sporadic instances of initial devoicing in South Asian Gāndhārī, e.g. *caṇa* 'meditation' next to *jaṇa* / *jāṇa* / *jhaṇa* < OIA *dhyāna-* (cf. Silverlock 2015: 166). Because of such examples, it is possible that devoicing was not an exclusive feature of Central Asian Gāndhārī, which makes the assumption of a Tocharian substrate origin less attractive.

6.3.2 The loss of aspiration

A second, equally well-known, characteristic of Tocharian is its lack of aspirated consonants, including simple *h*. These sounds are also rare to non-existent in the corpus of foreign nouns and personal names in Niya Prakrit, which Burrow (1935b: 669) considered as another argument in favour of a Tocharian substrate.[916] The loss of aspiration is, however, already found in South Asian Gāndhārī, conceivably due to contact with Iranian languages (cf. §5.1), so Tocharian does not need to have played a role here.

As concerns the loss of simple *h*, it bears mentioning that this sound also frequently drops in Bactrian, e.g. in αζαρο next to υαζαρο 'thousand' < **hazahra-*. If one wants to attribute the irregular treatment of *h* in Niya Prakrit to foreign influence, Bactrian is therefore as good a source for it as Tocharian.

6.3.3 The rarity of the fricatives *-x-*, *-θ-*, and *-f-*

Other sounds absent from Tocharian phonology are the fricatives *-x-*, *-θ-*, and *-f-*, an absence compared by Burrow (1935b: 669) to the rarity of such fricatives in Niya Prakrit. In Iranian names, *-f-* is rendered as 𐨯𐨪 *vha*, but Burrow noticed that *-x-* and *-θ-* are largely absent from the onomastic corpus. Yet, once again, the rarity of *-x-* and *-θ-* does not necessarily imply a Tocharian substrate. First, OIr. **-θ-* by default developed to *-h-* in Bactrian (cf. SFBG: 38), so this could be an alternative explanation for the general absence of *-θ-* in Niya loanwords and foreign names. Second, the rarity of *-x-* in Niya Prakrit could possibly result from its being substituted by a sibilant, a sound substitution I discussed in more detail in §4.3.27.

[916] Aspiration is sometimes dropped in inherited words too, resulting in forms such as *buma* 'grounds, land' next to *bhuma*, or *asta* 'hand' for *hasta* (see also LKD: §28).

6.3.4 The absence of retroflex consonants

The fourth phonological feature Burrow observed is the non-occurrence of retroflex consonants in the unexplained vocabulary in Niya Prakrit (1935b: 669f.). Retroflexes are, however, not only absent in Tocharian, but also in many Iranian languages, including Bactrian.[917] Their absence in loanwords into Niya Prakrit is hence no forceful argument in favour of a Tocharian substrate.

6.3.5 The semivowel *-w-*

In §2.1.6, it was stated that the Niya akṣara 𐨆 *y̲a*, phonologically /w/, occurs in non-Indic personal names, e.g. *Y̲arpa*, while at the same time also interchanging with 𐨅 *va* in inherited words like *viṃñaveti* (also *viṃñay̲eti*) '(s)he orders' < Skt. *vijñāpayati*. As this akṣara is seldomly used in South Asian Gāndhārī, Burrow (1935b: 670; cf. also LKD: §29) argued that its more frequent use in Niya Prakrit resulted from contact with Tocharian, given the existence of a phoneme /w/ in the latter. Contact with Bactrian is, however, an at least equally likely explanation, given that there are Niya Prakrit loanwords in which 𐨆 *y̲a* can be shown to render Bactrian -o- /w/ (see §4.3.22).

6.3.6 A palatalised lateral

Burrow's sixth phonological argument has to do with the Niya akṣara *lýa* (1935b: 670f.). As I discuss in more detail in §2.1.6, this akṣara can be taken as evidence that in Niya Prakrit, *l* could be palatalised in front of *ĭ*, a type of palatalisation connected by Burrow to the existence of a phoneme /Ĺ/ (written <ly>) in Tocharian. Palatalisation is, however, a typologically so widespread phenomenon (cf. Kümmel 2007: 250–266) that it is no cogent argument for contact between Niya Prakrit and Tocharian. One should moreover note that the phonemic inventory of Pāli, a MIA language that has certainly never been in contact with Tocharian, may also have included a palatalised /Ĺ/ (so Oberlies[2] 2019: 133 who speaks about "a *l mouillé*").[918]

---

[917] The Bactrian sibilant þ /š/ may have been a retroflex (cf. §2.2.4.1 and §4.3.18), but the same is true about Tocharian <ṣ>.

[918] The Graeco-Bactrian and Manichaean scripts are regrettably uninformative whether there was a (phonematic) palatalised lateral in Bactrian. In any case, the Niya PN *Lýaka* ← Bactr. λιακο /l(i)yākə/ < **dahyu-k(k)a*- (SA Gandh. *Liaka*, cf.

6.3.7 The affricate *-ts-*

Burrow (1935b: 671) further remarked on the presence of an affricate *-ts-*, a common phoneme in Tocharian, in Niya Prakrit loanwords and foreign names. Two Niya examples are *tsaṃgina*, a type of corn, and the PN *Tsina*. As /ts/ is, however, one of the commonly accepted values of Bactrian σ (cf. §2.2.4.2 and §4.3.15), its occurrence in Niya Prakrit cannot be conclusively attributed to Tocharian.

6.3.8 Syncope and heavy consonant clusters

Finally, there is, due to syncope, a variety of heavy consonant clusters in Niya Prakrit, e.g. in the word *crorma* (also *curorma* / *cirorma*), an agricultural product, or in the inherited PN *Butsena* (cf. also *Butsenga*) < Skt. *Buddhasena-*. Burrow (1935b: 671) ascribed such syncopes to Tocharian. His argument is, however, not compelling, as syncope and heavy clusters are as much a feature of Bactrian phonology as they are of Tocharian.

6.4 Morphology

As already briefly indicated in §6.1, Burrow based his "Tocharian C" hypothesis not only on phonological and lexical arguments; he also thought to recognise some traces of Tocharian derivation and inflection in Niya Prakrit. In my view, none of his proposals stands closer scrutiny. Some of the topics to be discussed still lack a convincing analysis, e.g. the plural form *paḱeyu* 'allotments, shares' (§6.4.5). Yet, at least some features are more likely Bactrian in origin than Tocharian, i.e. the adjective suffix *-ci* (§6.4.2) and the feminine marker *-s-* (§6.4.4). Still others need not be due to language contact in the first place, e.g. the adjective suffix *-i* (§6.4.3).

6.4.1 The suffix *-ina*

A few loanwords into Niya Prakrit contain a suffix *-ina* (sometimes *-iṃ*), most likely a marker of (substantivised) adjectives. Examples include the title *cuṿala(y)ina*, *cilaṃḍhina* 'shared', and *koyimaṃḍhina* 'an official in charge of *koyimaṃḍha* corn'.

According to Burrow (1935b: 672; LKD: §77), *-ina* (*-iṃ*) would correspond to a Tocharian A adjective suffix *-iṃ*, for which he refers his

§3.4.1) shows that *lýa* could also occur in Bactrian names occurring in the Niya documents.

readers to TG (21f.). However, none of the words where *-ina* occurs is easily analysable as a borrowing from Tocharian. In addition, the Tocharian words ending in *-iṃ* cited in TG (21f.) are adjectives in *-i* in the oblique case, e.g. *āriñciṃ* (A 338 b7; THT 1646.c a6), the obl.sg. of *āriñci** 'belonging to the heart' ← *āriñc* 'heart', and *ṣulyiṃ* (A 77 b3), the obl.sg. of *ṣuli** 'mountainous' ← *ṣul* 'mountain'.[919] In other words, a suffix *-iṃ* does not exist in Tocharian, and it would be far from plausible that *-ina* represents a borrowing from the obl.sg. of the adjective suffix *-i*.

Incidentally, the suffix *-ina* (*-iṃ*) occurs in two of the potential Bactrian loanwords in Niya Prakrit discussed in §§3.2.3–3.2.4. These are *acoviṃna* from *aco* 'frontier, border post' ← *αδοþαοο /adšāwə/ < **ati-čyāwa-* and *vaṣḍhiga̱ïṃ* from *vaṣḍhiga̱*, festival on the mountain ← *βαγδιγο /vaγdīgə/ 'tribute' from βαγδο 'distributed' < **baxta-* + -ιγο < **-iya-ka-*.

The question arises whether the suffix *-ina* could also be Bactrian in origin. This option seems indeed conceivable, given the existence of an adjective suffix -ηνο/-ινο in Bactrian.[920] This suffix is attested in some personal names, e.g. ραþτηνο < **rāšt-aina-* 'related to justice', in the adjective σιμινο 'made from silver' based on *σιμο 'silver' ← Gr. ἄσημος, and perhaps in the place name σαγγινο (Korn & Olsen 2012: 214).[921] Nevertheless, there is so far only little evidence that -ηνο/-ινο was an inherited suffix in Bactrian. Sims-Williams (BPN: no. 125) noted that all of the personal names ending in -ηνο/-ινο possibly come from Middle Persian or Parthian, and in these languages *-ēn* < **-aina-* is a frequent suffix (GW: §325), whereas in Bactrian the suffix -ηγγο/-ιγγο < **-aina-ka-* occurs more often (BD II: 213).[922] In order

[919] Of *āriñci**, there are no other forms attested (cf. DTTA: 42f.). For *ṣuli**, cf. additionally the nom.pl. *ṣuliñi* (A 75 a1; A 318 + 319 b4) and the obl.pl. *ṣulinäs* (A 200 a3).

[920] The raising of -η- to -ι- is probably due to the following nasal (so also Korn & Olsen 2012: 215). Cf. the derivation of Bactr. πριγγο* < **pir-aina-ka-* discussed in §3.2.1 s.v. *prigha*.

[921] It seems conceivable that σαγγινο is a derivative from σαγγο* 'stone', indirectly attested in personal names (cf. BPN: no. 401 and note also dir.pl. ασαγγε 'stones' in SK_4M 13; SK_4A 18; SK_4B 17).

[922] The suffix **-aina-* is frequently used in Iranian languages as a "material suffix", forming adjectives from nouns denoting substances. -ηγγο/-ιγγο and its Khotanese equivalent *-īnaa-* also have a more general meaning 'affected with, related to' (cf. Degener 1989: 138–152; Korn & Olsen 2012: 205f.). Therefore, if -ηνο/-ινο was also a native Bactrian suffix, its general meaning may also have been 'related to', fitting with *-ina* in Niya Prakrit.

to confirm -ηνο/-ινο as the origin of Niya -*ina*, one would need not only more evidence for this suffix in Bactrian, but also further Bactrian etymologies for Niya Prakrit words containing this suffix -*ina*.[923]

Whatever its correct etymology, one can safely conclude that the suffix -*ina* (-*iṃ*) is unlikely to be Tocharian in origin, and this conclusion is sufficient for the purposes of this chapter.

6.4.2 The suffix -*ci*

The appurtenance suffix -*ci* has been treated in §5.2.2. There, I argued against Burrow's comparison with the TA adjective suffix -*ñci* and proposed an alternative derivation from Bactr. -σ(ε)ιγο /tśīgə/.

6.4.3 The suffix -*i*

In Niya Prakrit, a common way to derive adjectives from nouns is by deleting the stem vowel -*a* and replacing it with the suffix -*i* < OIA -*ika*- (cf. LKD: §75). A good example is *goṭhi karma* 'work related to the farm', *goṭhi* being a derivation from *goṭha* 'farm, household'.

In his article about "Tocharian C" (1935b: 672), Burrow noted that the use of the suffix -*i* got "much extended compared with ordinary Sanskrit and Prakrit, at the expense of the Tatpuruṣa compound". This observation, in turn, induced Burrow to compare the default Tocharian translation of Indo-Aryan tatpuruṣa compounds by the combination of an adjective and a noun, e.g. TB *pelaikneṣṣe cākkär* 'wheel of the dharma' ~ Skt. *dharmacakra*- 'idem'.[924]

It seems, however, unnecessary to assume Tocharian influence on the *i*-suffix. First of all, the more extensive use of compounds in Sanskrit probably had more to do with literary taste than with the actual linguistic situation, as Aśokan Prakrit and the modern languages likewise have a more restricted use of compounds (Bloch 1965: 107). Moreover, if the spread of the *i*-suffix was due to contact influence, there is a more likely source than Tocharian in the form of the Bactrian adjective suffix -ιγο /īgə/ < *-*iya-ka*-. As the usage of -ιγο closely resembles that of -*i* in Niya Prakrit (cf. BD II: 217

[923] In case the Bactrian etymology turns out to be correct, it is still not to be included under the heading of "grammatical borrowings" from Bactrian, as the suffix -*ina* never attaches to native words (cf. §5.2.3 for this criterion).

[924] In fact, the translation of Sanskrit compounds into Tocharian is more complicated than Burrow presents it. See recently Meunier (2015).

on -ιγο), it may have contributed to its spread, although there is still no absolute necessity to maintain that the *i*-suffix underwent foreign influence in the first place.

6.4.4 The feminine marker *-s-*

Some Niya female PNs of non-Indic origin end in *-(i)sa* (*-(i)ṣa*). Examples include *Smaga̱sa* (CKD 71), *Apiṣa* next to masc. *Apika* / *Apita* (CKD 11), and *Kuñiṣa* next to masc. *Kuñita* (CKD 110). Outside personal names, another instance of a feminine marker *-s-* may occur in the otherwise unexplained adjective *tirsa*.[925] This form is twice combined with *vaḍ̱avi* 'mare' (CKD 39; 771), whereas the similarly looking *tirṣa* by default modifies *aśpa* 'horse' (CKD 39; 45; 434; 573). Put differently, *tirsa* is probably the feminine equivalent of masculine *tirṣa* (cf. LKD: 95 s.v.).[926]

Burrow (1935b: 672) included the feminine marker *-s-* in his list of possible Tocharian elements in Niya Prakrit. Although he made no explicit comparison, Burrow was presumably thinking of the feminine preterite participles in Tocharian. Whereas, for instance, the nom.masc.sg. of the preterite participle of *yām-* 'to do' is TA/TB *yāmu* 'done', its nom.fem.sg. is TA *yāmus* / TB *yāmusa*.[927] Yet, preterite participles are the only place in Tocharian grammar where *-s-* marks feminines: names borne by females are not characterised in this way, which renders Burrow's idea unconvincing.

Sims-Williams (1996: 52) observed that "[a] more plausible source can be found in Bactrian, where *-(i)s(a)* is the expected outcome of the Old Iranian feminine suffix **-čī-*" (cf. also Tremblay 2001: 35 fn. 54).[928] The Bactrian

[925] Sylvain Lévi's (*apud* Bloch 1937: 52) suggestion that *tirsa* means 'three years old' does not receive strong support from the available contexts.

[926] In a handwritten note added to his copy of LKD (106), Burrow suggested the adjective *pursa(ka)*, like *tirṣa* / *tirsa* used for animals, to be another feminine form containing the same suffix *-sa*. That *pursa(ka)* could be used for female animals is shown by CKD 383, where the gen.fem.sg. *taya* < *tasyāḥ* 'her' is co-referential with an aforementioned *pursaka* camel.

[927] On the formation of the Tocharian preterite participles, including the feminine, see Peyrot (2010).

[928] Not having spread to the inherited lexicon, the suffix *-(i)sa* (*-(i)ṣa*) is, however, not to be included among the "morphological loans" from Bactrian into Niya Prakrit studied in Chapter 5 (for this criterion, cf. §5.2.3).

suffix referred to typically occurs as -σο (-ζο after nasals).[929] Yet, the Kuṣāṇa Bactrian adjective αβαβσα 'waterless' (Ayrtam 4, cited in fn. 374), a feminine counterpart of masc. αβαβγο 'idem' < **apāpa-ka-*, additionally evidences a more archaic variant -σα. The latter may be the direct source of Niya *-(i)sa* (*-(i)ṣa*), as some Bactrian loanwords into Niya Prakrit were observably borrowed at a time when (certain) final vowels were still preserved (see §4.2.1).

-σα is derived by Sims-Williams (1994: 173) from **-čī+ā-* based on a comparison with Sogd. *-č* / *-čī* (heavy stems) and *-čā* / *-čyā* (light stems) and Khot. *-kyā-* / *-cā-* (*-gyā-* / *-jā-*) (cf. Sims-Williams 1982: 15 with fn. 15; Degener 1989: 199–208). Alternatively, **-čī-* could have been remade into **-čā-* in the ancestor of Bactrian by analogy with the feminines in **-ā-.*

In either case, -σα presumably stands for /tsă/. Unless /ts/ could occasionally be simplified to /s/ in Bactrian (e.g. in certain clusters),[930] one consequently has to presume a simplification (and subsequent voicing) of Bactr. /ts/ to Niya *-s-*/*-ṣ-* /z/. If one derives it out of Bactr. παρασο /parātsə/ 'away' < **parāčā,* Niya *paraṣa* 'abduction, plundering' (also 1x with *-s-*) would be a parallel for this sound substitution (on which see §3.3.2 s.v. *paraṣa* and §4.3.15).

In brief, Sims-Williams' idea to derive Niya *-(i)sa* (*-(i)ṣa*) out of the Bactrian feminine suffix -σα (later -σο) is to be preferred over Burrow's comparison with Tocharian.

6.4.5 The plural/collective in *-u* in *paḱeyu* 'allotments, shares'

Compare §3.2.3 s.v. *paḱe* for a rejection of Burrow's comparison between the plural/collective *paḱeyu* 'allotments, shares' and the rare Tocharian A plural in *-u*, as in *cmolu* 'births' ← sg. *cmol*.

---

[929] Cf. e.g. the female PN βαγοαζιιασο < **baga-āzāyata-čī-* 'dear to god' (BPN: no. 45) and the adjective φρομανζο 'commanding, possessing authority', seemingly a feminine to *φρομαγγο < **fra-māna-ka-*.

[930] A comparison with MBactr. *'s* /as/ 'from' < **hačā* should probably be resisted since the occasional spelling of ασο with <σσ> in the newly discovered birchbark letters suggests that this word still had /ts/ in earlier forms of Bactrian. Cf. §2.2.4.2 with fn. 116.

6.4.6 The plurals in *-aṃca*

As first observed by Thomas (1927), in Niya Prakrit, a few nouns can form their dir.pl. in *-aṃca* (12 types; 22 tokens; see below), even when in other contexts, they are attested with the default dir.pl. ending *-a*.

Attestations of this plural formation[931]

aśpaṃca 'horses': CKD 387; 681
uṭaṃca 'camels': CKD 132; 387;[932] 665; 681; 696
jaṃdunaṃca 'worms': CKD 565
dajhaṃca 'slaves': CKD 133
paṭaṃca 'silk rolls': CKD 660 (2x)
paśunaṃca 'sheep': CKD 683
bhumaṃca 'lands': CKD 366 (2x); 713 (2x)
makaḍaṃca 'monkeys': CKD 565
mahatvaṃca 'magistrates': CKD 696
śaḍaṃca 'enclosures, vineyards': CKD 85
storaṃca 'horses': CKD 13
striyaṃca 'women': CKD 552 (2x)

Forms in *-aṃca* combine with (i) dir.pl. *ede* 'these';[933] (ii) numerals higher than one;[934] and (iii) verbs with unambiguous plural morphology.[935] There is consequently no doubt that they are indeed plurals. Leaving aside that *-aṃca* is much less frequently attested than *-a*, it has, however, proved impossible to discern a difference in usage or distribution between the two plurals.[936]

---

[931] I have excluded *valaḡaṃ ca* (CKD 367, 2x) from this list of attestations, as this probably means 'and a guard' rather than 'guards'. In CKD 797, Lin's (1999: 231) reading *aẏanaṃ ca*, theoretically interpretable as a dir.pl. **aẏanaṃca* 'villages', is clearly mistaken. Based on the hand copy of CKD 797 in Hasuike (1996: 307), I would read and translate the relevant sentence as *tuo aṭa naṃdha goṭhaṃmi thaviteṣi* 'you deposited a load of flour in the farm' (cf. also §3.3.2 s.v. *aṭa*).
[932] Boyer et al.'s (KD: 139) reading *uṭa ca* can be corrected to *uṭa[ṃ]ca*; on the photograph, one can see a clear trace of a floating anusvāra below the 𐨟 *ṭa*.
[933] CKD 366: *ede bhumaṃca* 'these lands'; 387: *ede uṭa[ṃ]ca* 'these camels'.
[934] CKD 660: *4 3 paṭaṃca* 'seven silk rolls'; *paṭaṃca 2* 'two silk rolls'; 665: *uṭaṃca 10* 'ten camels'. Cf. also CKD 85: *śaḍaṃca sarva* 'all enclosures, vineyards'.
[935] CKD 681: *aśpaṃca … khayitaṃti* 'horses … ate'; 696: *mahatvaṃca … kareṃti* 'magistrates … do'.
[936] As is the case with *-a*, *-aṃca* can be used for both the subject and object of a sentence; plurals in *-aṃca* can be definite or indefinite, animate or inanimate; they can occur in both main and sub-clauses; they can be negated or not; they occur both with and without numerals; and so forth.

Another unsolved problem is the etymological derivation of *-aṃca*. Thomas' (1927: 546) proposal to etymologically compare *-aṃca* to the pluralising suffix *ca* / *can* found in later Tibeto-Burman languages is unattractive, as there are no secure parallels for lexical or grammatical borrowings from Tibeto-Burman into Niya Prakrit.[937]

Burrow (1935b: 673) made an alternative comparison with the Tocharian A nom.pl. in *-āñ*. To account for the evident mismatch between this ending and *-aṃca*, Burrow had, however, to hypothesise that "a palatal consonant might have dropped off at the end [of *-āñ*]". As a parallel for the supposed consonant loss, he pointed to TA *päñ* 'five' (~ TB *piś*) < PToch. **pə́ńśə* < PIE **penkʷe* (DTTA: 275). By now, it is nevertheless commonly agreed that *-āñ* comes from the PIE nom.pl. ending **-ōn-es* (so e.g. Pinault 2008: 499). Burrow's proposal is thus phonologically impossible.

More recently, Banti (2000: 3f.) has argued that the plural in *-aṃca* originated in Prakrit-internal developments. He maintained that these plurals resulted from a reanalysis of coordinated acc.pl. forms of masculine *a*-stems of the type *dāsān aśvāṃśca* 'slaves and horses' (ex. mine). First, *dāsān aśvāṃśca* underwent a phonological development to **daṣaṃ aśpaṃ ca*. In an unprotected environment, i.e. when no clitic was following, *-aṃ* further developed to *-a*, resulting in **daṣa aśpaṃ ca*. Once nominative and accusative merged to become the direct case, **daṣa aśpaṃ ca* would, according to Banti, be reanalysed to **daṣa aśpaṃca*. A potential motivation for such a reanalysis would be the Niya Prakrit tendency to avoid using *ca* 'and' when coordinating two or more nouns (§6.5.2). Once this process resulted in the creation of a new ending *-aṃca*, it could have spread further to other syntactic environments.

In my view, also Banti's scenario is not compelling. First, one wonders how likely it is that these plurals came about in sequences of nouns coordinated by *ca* 'and'. After all, as Banti writes, Niya Prakrit rarely uses *ca* to conjoin noun phrases. There are moreover only two instances where a plural in *-aṃca* occurs in such a coordinated sequence.[938]

Besides, if Banti is correct that *-aṃca* originates in the acc.pl. of masculine *a*-stems, this implies that the *u*-stem forms *jaṃdunaṃca* 'worms' (CKD 565) and *paśunaṃca* 'sheep' (CKD 683) are later analogical formations. Yet, if so, it is not entirely clear how one can account for their morphology. According

[937] See Barnett (1927) and Thomas (1928) for further debate on this proposal.
[938] CKD 13: *vaḍavi storaṃca* 'mares and horses'; 387: *khulaüṭa aśpaṃca* 'camels from the herd and horses'.

to Banti, final nasals should have dropped before the creation of the new ending *-aṃca* (cf. supra). So, to explain the *-n-* in between *-u* and *-aṃca* in the forms just quoted, one would probably need to assume that *-aṃca* was secondarily added to the old acc.pl. *jantūn* and *paśūn*. This assumption contradicts, however, Banti's own chronology.[939]

Finally, Banti's preference for a language-internal derivation is primarily based on the fact that *-aṃca* most frequently attaches to inherited nouns. A borrowed morpheme, by contrast, is expected first to appear in loanwords before spreading to inherited nouns. Nevertheless, at least one Bactrian loanword is attested with a plural in *-aṃca*, i.e. *stora* (see §3.2.1 s.v.), and the Indo-Aryan etymology of some of the other words combined with *-aṃca* is less secure than Banti seems to believe.[940] In other words, we should perhaps not too quickly dismiss the possibility that *-aṃca* is a borrowed morpheme after all.

Everything taken together, Banti's proposal needs to be treated with caution too, and a convincing explanation of the Niya Prakrit plurals in *-aṃca* remains to be found. Anyhow, *-aṃca* is unlikely to be Tocharian in origin, so also this morpheme cannot be used as evidence for a Tocharian substrate in Niya Prakrit.

6.5 Syntax

Syntactic arguments were lacking in the original formulation of the "Tocharian C" hypothesis by Burrow, but Banti (2000: 5f.) adduced two syntactic similarities between Tocharian and Niya Prakrit. In this section, I argue that neither of them is necessarily the result of language contact.

---

[939] Irrespective of whether this is diachronically also the correct analysis, *jaṃdunaṃca* and *paśunaṃca* look more like a gen.pl. followed by *-(ṃ)ca* (cf. gen.pl. *paśuna* < Skt. *paśūnām* in CKD 725).

[940] First, *uṭa* is probably not the phonologically regular outcome of *uṣṭra-*. Given its semantics (cf. §6.2 s.v. *aṃklatsa*), it could be a Bactrian loanword, perhaps from *οτρο /utrə/ < **uštra-*, with **-št-* to Bactr. -τ- /t/ (SFBG: 47). As there is, however, no parallel for the development of **-štr-* in Bactrian, this cannot be proved for now. Second, *śaḍa* 'enclosure, vineyard' is commonly derived from Skt. *śālā-* 'enclosure' (see e.g. LKD: 111 s.v. *masu śaḍa*). Yet, this etymology is not very attractive as *śaḍa* would be the only instance of *-ḍ- /ɽ/ corresponding to Skt. *-l-*. Third, the etymology of *paṭa* 'silk roll' and its IA cognates remains uncertain (EWAia III: 298), so this may also be a recent loanword.

6.5.1 Group inflection in noun phrases

In Niya Prakrit, case endings only need to be attached to the last element in a noun phrase. This noun phrase may consist of the combination of (i) an adjective or a pronoun with a noun; (ii) a noun in apposition with another noun; or (iii) two or more coordinated nouns (LKD: §136; Jamison 2000: 68 fn. 22; Barchi & Peschl 2022: 413 fn. 21).

Banti (2000: 5) compared this phenomenon to Tocharian group inflection citing as an example *kokleṃ oṅkolmaṃnpā* 'with chariots and elephants' (THT 362 b9).[941] In this example, the comitative ending *-mpā*, here written as *-ṃnpā*, is only visible in *oṅkolmaṃnpā*, whereas *kokleṃ* is obl.pl. Such group inflection occurs in various languages from the region, including Sogdian and Tibetan. It may hence be an areal feature.

Nevertheless, Niya Prakrit group inflection does not need to result from contact-induced change. Though it has not always been described as such, there are instances of the same phenomenon in various forms of MIA, including South Asian Gāndhārī and Buddhist Hybrid Sanskrit. [942] Incidentally, group inflection occurs relatively frequently in Apabhraṃśa (de Vreese 1959), which, like Niya Prakrit, represents a more vernacular and grammatically more advanced type of MIA. In addition, some Prakrit grammarians, e.g. Rāmaśarman in his *Prākṛtakalpataru* (1.6.2), prescribe group inflection as a characteristic feature of the language. This counts as another argument that Niya Prakrit's group inflection may be due to language-internal developments.

6.5.2 Asyndetic coordination of noun phrases

As mentioned in §6.4.6, Niya Prakrit prefers asyndetic coordination of noun phrases. The conjunction *ca* 'and', so frequently found in other forms of early Indo-Aryan, including South Asian Gāndhārī, has largely fallen out of use.

According to Banti (2000: 5f.), one should compare this feature of Niya syntax to the use of asyndetic constructions in Tocharian, e.g. in the TB example THT 31 b6 loc.pl. *riṃne k_uṣaiṃne ostwane* 'in/to towns, villages, and houses'. Yet, once again, there is no need to invoke external influence. The conjunction *ca* has hardly left a trace in NIA (cf. Bloch 1965: 310; CDIAL: no.

---

941 Sieg, Siegling & Schulze (TG: 205ff.) discuss Tocharian A group inflection in detail. The situation in Tocharian B is largely comparable.

942 See Caillat (1990: 11ff.; 23 n. 18) and von Hinüber (2001^2: §295) with the references listed there.

4533), and its infrequency in Niya Prakrit may merely be an early sign of its general disappearance.[943] The more colloquial nature of some of the Niya Prakrit documents may have acted as a contributing factor in this regard.

6.6 Conclusions and open questions

To conclude, hardly any argument so far adduced in favour of "Tocharian C" stands closer scrutiny. Some of the supposedly Tocharian features in Niya Prakrit are more easily explained as language-internal developments; others are probably due to contact with Bactrian and thus in line with the conclusions of Chapters 3–5. Of the evidence cited so far, the official title *kitsaïtsa* has the best chance of being of Tocharian origin. This one word can, however, not bear the burden of proving a Tocharian substrate. As official titles can be borrowed quite easily, *kitsaïtsa* may merely be a linguistic vestige of the contacts between the southern and northern parts of the Tarim Basin which have to be assumed on non-linguistic grounds in any case (Høisæter 2020: 367–369).

The question remains which language the native people of the Shanshan kingdom spoke. As it would be historically unlikely that there were more than a few speakers of Bactrian in the region before Kuṣāṇa times, the Bactrian elements in Niya Prakrit should not make us believe that Bactrian was the substrate language spoken by the Shanshan people.[944] A more sensible, though admittedly not quite exciting, conclusion would be that the available evidence does not permit us to deduce with an acceptable degree of certainty which language was spoken. The onomastic record in Niya Prakrit possibly contains names that find their etymology in the language of the Shanshan people. Yet, building large-scale conclusions on such inherently ambiguous evidence would be methodologically too hazardous.

Recently, Peyrot (2018b: 275–277) has hypothesised that the Aqtala archaeological culture, which spread over large parts of the Tarim Basin from about 1000 BC to 650/550 BC, can be identified with the arrival of speakers of Proto-Tumshuqese-Khotanese to the region.[945] Based on this

[943] Cf. the disappearance in Niya Prakrit of the particle *hi* 'for', which is still used in some other forms of Gāndhārī (Ollett 2022: 274). Another parallel is the loss of enclitic personal pronouns discussed in §5.4.10.2.

[944] Even during Kuṣāṇa times, it seems unlikely that a large part of the population had shifted their language to Bactrian.

[945] Dragoni's examination of Proto-Tumshuqese-Khotanese loanwords in Tocharian has added linguistic evidence to this hypothesis (2023b: 254f.).

identification, he suggests that the area east of Khotan and the Keriya river "was Iranian-speaking before Middle Indian was introduced there", although "the details and consequences of this hypothesis need further study" (Peyrot 2018b: 276).[946] In other words, the Shanshan language may have been closely related to Tumshuqese and Khotanese. However, also this theory, while attractive in itself, has yet to find linguistic support in Niya Prakrit or elsewhere. Given our still imperfect knowledge of the linguistic prehistory of the Tarim Basin, caution is therefore called for.

[946] Peyrot's theory should not be confused with Barchi and Peschl's hypothesis of a Khotanese substrate in (an earlier form of) Niya Prakrit, against which I argued in §5.4.