

The spread of iron in Central Asia: on the etymology of the word for “iron” in Iranian and Tocharian¹

Michaël Peyrot

Leiden University Centre for Linguistics, Leiden, The Netherlands

m.peyrot@hum.leidenuniv.nl 

Federico Dragoni

Leiden University Centre for Linguistics, Leiden, The Netherlands

f.dragoni@hum.leidenuniv.nl 

Chams Benoît Bernard

Leiden University Centre for Linguistics, Leiden, The Netherlands

c.b.a.s.bernard@hum.leidenuniv.nl 

Abstract

Tocharian B *eñcuwo* “iron” and Tocharian A *añcu** have been connected to the Iranian words for “iron”, notably Khwarezmian *hnčw*. On the basis of insights into the patterns of borrowings from Khotanese into Tocharian, it is argued that the Tocharian words must have been borrowed from a pre-form of Khotanese *hīśšana-* “iron”. Further, a new etymology is proposed for “iron” that accounts for the variation of this word in Iranian. The fact that Tocharian borrowed the word for “iron” from Khotanese, not from the archaic steppe dialect of Iranian that is the source of many other loanwords in Tocharian, suggests that the contacts between this latter dialect and Tocharian took place before iron became widespread in the region.

Keywords: Khotanese, Tocharian, Iron, Historical linguistics, Central Asia, Borrowing

1. Introduction

The word for “iron” in the Iranian languages is notoriously difficult to reconstruct: similar forms are found in many languages, but they show irregular correspondences and cannot be reconstructed to a single protoform. For illustration, we may cite here Pahlavi <’*h(y)n’*>, Manichaean Middle Persian ’’*hwn*, Balochi *āsin*; Avestan *haosafnaēna-* “out of steel”; Khwarezmian *hnčw*; Sogdian

1 We thank the two anonymous reviewers for valuable comments, and Mei Jianjun (Cambridge) for his kind help. For the contents we are alone responsible. This research was supported by the Netherlands Organisation for Scientific Research (NWO, project number 276-70-028).

'*spn**;² Ossetic *æfsæn* “ploughshare; iron (obs.)”; and Khotanese *hīśśana-*. It is possible to distinguish groups of Iranian languages that show more closely related words (Klingenschmitt 2000: 193–4; Adams 2013: 85; Sadovski 2017: 572), and the items that can be reconstructed for these clusters resemble each other, but it seems that no single protoform can account for all of the different clusters together. The Tocharian words for “iron”, Tocharian B *eñcuwo* and Tocharian A *añcu**,³ look similar too, especially if compared to Khwar. *hnčw* (e.g. Adams 1999: 80). It thus seems likely that the Tocharian word is borrowed, but it is difficult to indicate the exact source. The irregular correspondences between the Tocharian and the Iranian forms, as well as those between the Iranian forms themselves are often explained with the assumption of a substrate term or a Wanderwort (e.g. Schwartz 1974: 409).

In this article, we argue that the word for “iron” in Iranian and Tocharian is not a substrate term or a Wanderwort. In sections 2 and 3, we identify the source of Tocharian B *eñcuwo* and Tocharian A *añcu** as a prehistoric form of Khotanese *hīśśana-*. In sections 4 and 5, we revisit the Iranian words for “iron” and we try to show that the irregularity in the correspondences may be reduced considerably. Even though it remains difficult to reconstruct a single proto-form, the remaining variation does not allow the assumption of a substrate term or a Wanderwort. Finally, we draw further conclusions from our revised reconstructions: on the one hand, the Tocharian borrowing is important evidence for the development of the Proto-Indo-Iranian cluster **čy* in Khotanese-Tumšugese (section 6); on the other hand, the fact that the Tocharian word was clearly borrowed from prehistoric Khotanese allows important inferences about the prehistory of Tocharian and its contacts with Iranian (section 7).

2. Iranian and Tocharian terms for “iron”

A cluster of partly reconstructible Iranian terms is formed by Sogdian '*spn**', Ossetic *æfsæn*, and Khotanese *hīśśana-*. The correspondence between Sogd. *sp*, Oss. *fs* and Khot. *śś* points to Proto-Iranian **čw*⁴ > **tsw* as in **áčwā-* “horse” > Sogd. '*sp-*, Oss. Digoron *æfsæ* “mare” (< **áčwā-*), Khot. *asśa-*. Leaving the initial aside for now, this allows a reconstruction **-čwana-* > **-tswana-*. While, at least at first sight, the Sogdian and Ossetic forms can be derived from **tswana-* directly, Khotanese *hīśśana-* shows an initial element *hī*^o that is difficult to explain. The Khwarezmian form *hnčw* and the Tocharian forms B *eñcuwo* and A *añcu**, if related, likewise show an initial element, but lack, in turn, the final element *-n*, *-na*.

The similarity between the Iranian and Tocharian words has been explained in different ways. Schwartz (1974: 409 fn 33; cf. Adams 1999: 80) connects

- 2 We use a following asterisk to indicate that a form is not directly attested but can be deduced synchronically. By contrast, a preceding asterisk is only used for forms that are reconstructed for linguistic pre-stages.
- 3 The Tocharian A noun *añcu**, which is so far unattested, can be posited with certainty on the basis of the derived adjective *añcwāsi* “out of iron”.
- 4 For this notation for Proto-Iranian, see also §6 below.

Tocharian B *eñcuwo* and A *añcu** with Khwarezmian *hnčw*. For Iranian, he reconstructs **ans(u)wan-*, which, together with the Tocharian and Khwarezmian forms, would go back to a “substratic” **anśuwan*.

Based on Klingenschmitt (2000: 193), Tremblay (2005: 424) reconstructs a Proto-Indo-European **h₂ek̑-uon-* “cutting edge”, which developed to Khot. *hīśśana-* through an intermediary “Old Sakan” **ačyan-* due to a change of **ačy* to *īśś* in trisyllabic words. According to him, the Tocharian and Khwarezmian words derive from a nasalized variant **ančyan-* of **ačyan-*. Although we agree that the *ī* in the first syllable of Khot. *hīśśana-* goes back to **a*, we think that the mechanism for this change is different; see §3 below. Tremblay does not explain why the first **a* of *ančyan-* was rendered by Tocharian B *e*, while the second is apparently rendered as *o*.

Pinault (2006: 184–9) accepts Schwartz’s attribution of the Tocharian and Khwarezmian words to a Central Asian substrate source. In addition, he adds Vedic *aṃśú-*, a term referring to the material aspect of the soma plant. According to him, the common semantic dominator is the colour: iron is often called after the reddish-brown colour of its rust, and *aṃśú-* may have referred to the colour of soma juice. In our view, this etymology is in itself not impossible, but it depends on a semantic derivation of “iron” that cannot be verified and has to cope with phonological difficulties.⁵

3. Tocharian B *eñcuwo* is a borrowing from prehistoric Khotanese

In our view, Tocharian B *eñcuwo* and A *añcu** can be derived from an early stage of Khotanese. Two insights are crucial to understand this borrowing:

- Tocharian *ñc* may correspond to Khotanese *ś(ś)* in borrowings from early Khotanese into Tocharian;
- borrowings from early Khotanese into Tocharian regularly show Tocharian **a* for Khotanese **a* word-internally, while the Khotanese stem vowel **-a* is rendered morphologically with Tocharian **-o*.

The correspondence between Tocharian *ñc* and Khotanese *ś(ś)* was established by Chen and Bernard (forthcoming).⁶ They have discovered that Toch.B *śāñcapo* means “mustard” and does not denote a tree, the *Dalbergia Sissoo* or Indian rosewood, as had been supposed since Hoernle (1901: 23; cf. Adams 2013: 681). With this new meaning, they could connect the Tocharian B word with Khot. *śśasvāna-* “mustard”, and the relevant pre-stage of the Khotanese word can be reconstructed as **śaNzapa-* or **śaNzapāna-* (on these options, see below). The nasal in the first syllable is reconstructed on the basis of Tocharian and cannot be reconstructed on the evidence of Khotanese alone, but it is found in variants of this word in other languages, on the basis

5 According to Buyaner (2013: 613), the Tocharian words are borrowed from “the East-Iranian designation of ‘spike’ or ‘blade’, in turn borrowed from proto-Parthian **anjūg* ‘narrow’”. In our view, this etymology is formally and semantically impossible.

6 The argument has also been laid out in an online presentation with the title “A spicy etymology. On Tocharian B (and A) *śāñcapo*” by Chen Ruixuan and Chams Bernard on 8 December 2020 at the *Tocharian in Progress* conference of Leiden University.

of which Henning (1965: 44) reconstructed an Iranian preform **sinšapa-*. In these parallel forms, there is no evidence for a palatal stop or fricative after the nasal, and the *-c-* in Tocharian B *šāñcapo* must therefore be a secondary development of **ś* after a nasal; cf. e.g. 3sg.opt. *tañci* “may he stop” from **tānk-’i* vs. *taši* “may he touch” from **tāk-’i* or the frequent epenthesis of *t* in *s*-clusters like *kektseñsa* “with the body” from **kektseñ-sa* (on *šāñcapo* and related matters, see in detail Chen and Bernard forthcoming).⁷

Tocharian borrowings from different stages of Khotanese systematically render word-internal *a* with *a*, and word-final *-a* with *-o* (Dragoni 2022). Both correspondences are found in the newly discovered match between Toch.B *šāñcapo* and Khot. *śśaśvāna-*, but also in loanwords that were already known. Internal *a* is for instance found in Toch.B *tvāñkaro* “ginger” from an earlier form of Late Khot. *ttumgara-*, possibly **tvāmgara-* or **tvamgara-* (Dragoni 2021: 305–7). Final *-o* is further found in e.g. Toch.B *pito* “price” from prehistoric Khot. **pīθa-* (Old Khot. *pīha-*), Toch.B *tono* “garment” from Khotanese *thauna-*, Toch.B *yolo* “bad” from Old Khot. *yaula-* “falsehood”. The correspondence between Khotanese and Tocharian word-internal *a* may not seem surprising, but is relevant because short **a* is rendered with Tocharian **e* in borrowings from the Old Iranian dialect from which words such as Toch.B *perne* “glory” ← **farnah-* were borrowed. Since we suppose that this Old Iranian dialect hails from the steppe, we tentatively term it “Old Steppe Iranian” for the sake of convenience. Tocharian final *-o* is probably a morphological rendering of the Khotanese acc.sg.m. *-u* or the nom.acc. sg.n. *-u*.⁸ Importantly, these correspondences are thus far found with certainty only with Khotanese loanwords and therefore they are crucial clues to identify Khotanese borrowings in Tocharian.

Since the final *-o* of Toch.B *eñcuwo* thus fits this newly established pattern of Khotanese borrowings in Tocharian B, while the *-śś- < *-ćw-* in Khotanese *hīśśana-* may, on the evidence of *šāñcapo* and *śśaśvāna-*, go back to a cluster with a nasal, i.e. **Nćw-*, it is tempting to see if the Tocharian word can be derived from prehistoric Khotanese. Indeed, we argue that this is possible, and much more straightforward than any of the explanations for Tocharian B *eñcuwo* that have been proposed so far.

As noted above, Tremblay assumes a change of **aćw* to *īśś* in trisyllabic words to account for the initial syllable of Khotanese *hīśśana-*. For this change, he gives two parallels: Late Khot. *hīšau’* “loss of appetite (Skt. *aruci*)” and Late Khot. *piśāra-* “disgusting”, which he derives from Plr. **fra-ǰwāba-* (following Bailey 1979: 487) and Plr. **pati-ǰwāra-*, respectively. Both forms would be related to e.g. MSogd. *’zβ’β* “taste”. Whereas the first derivation is difficult from both the phonological and the semantic points of view, the second is impossible because *piśāra-* is only attested in Late Khotanese, where a single <ś> without subscript hook indicates an unvoiced sound, not a voiced one. From **ǰwāra-* one would rather expect Late Khot. **ǰsāra-*. A better etymology for Late Khot. *piśāra-* may be < Old Khot. **paśśāra-* (with Late Khotanese

7 Another possible parallel is Tocharian A *sañce* “doubt”, borrowed from Skt. *saṃśaya* or a related form. However, a form with *mc* is also attested in Gāndhārī *saṃcaya*, from which the Tocharian A word could alternatively have been borrowed.

8 In other borrowings, it may also stand for the acc.sg.f. *-o*.

weakening of unaccented **a* in the first syllable) < **apa-ćwaH-ra-* “disgusting”, a *-ra-* verbal adjective from the root **ćwaH-* “throw”. The verb *paśś-* < **apa-ćwaH-* is attested in Old Khotanese and means “let go, release” (cf. BSogd. *’pspy-* “to reject”).

For the initial syllable of *hīśśana-*, we rather follow Bailey (1979: 352b) in assuming “trajected umlaut”, that is, a fronting effect on an original **a* in the first, apparently accented, syllable due to a **y* in the third syllable. Parallels for “trajected umlaut” are Khot. *ysīrra-* “gold” < Plr. **jāranya-* (Skjærvø 2004: II, 331) and Late Khot. *hījsara*, loc.sg. of *hamjsarā-* “scope, range, place” < **ham-ćarā-* (De Chiara 2014: II, 216).

However, while Bailey reconstructed **aśuanyā-*, the nasal preserved in Tocharian suggests **ham-ćwanyā-*, a reconstruction that accounts at the same time for the initial *h-* of Khotanese, which in the reconstruction of Bailey and others had to be an unexplained secondary development.⁹ According to the correspondences established so far, the Tocharian B vowel *e-* cannot reflect Pre-Khot. **a*; instead, it must render the intermediate stage of **a* after **y*-umlaut and before further raising to *ī*. At this intermediate stage, the Khotanese word seems to have been **hen-śwanyā-*. In Tocharian, the initial **h-* was lost as generally in borrowings from Iranian, while the cluster **nś* developed to **ñc* as in *śāñcapo* “mustard”.

If Toch.B *eñcuwo* (Khot. *hīśśana-*) was borrowed at the same stage of Pre-Khotanese as *pito* “price” (Khot. *pītha-*), the difference in the representation of Khotanese *ī* may be due to the fact that *ī* < **ai* in “price” was always a long vowel, and may have become *ī* earlier than *ī* < umlauted **a* in “iron”, which may at first have become short **e*. Alternatively, *eñcuwo* may have been borrowed earlier than *pito*, at a stage where **henśwanyā-* and **pēθa-* both had not yet shifted their **e* and **ē* to **ī*.¹⁰

The correspondence of the finals is difficult. Although Tocharian B *-o* is compatible with borrowing from Khotanese, at first glance *eñcuwo* (phonologically /eñćáwo/, with ə-epenthesis from **eñćwo*) looks like **henśwa-* rather than the **henśwanyā-* posited above. We have at present no definitive solution for this apparent mismatch, but see two possible scenarios.

The first option we see is that the word was borrowed at the stage **henśwanyā-*, in the form **eñćwañño*, **eñćwañña*. Such forms would have looked like a *ññe*-adjective, like the adjective *eñćwaññe* “out of iron”, which is well attested in Tocharian B. Supposing that the expected **eñćwañño*, **eñćwañña* was indeed interpreted as an adjective, the step to extracting the base noun as *eñcuwo* is very simple: a *ññe*-adjective to a base in *-a* presupposes an oblique singular *-a*, and the most frequent corresponding nominative singular ending is *-o*. In this scenario, it is most likely that the prehistoric Khotanese word was borrowed before Proto-Tocharian split up into Tocharian B and A,

9 The initial *yi-* of Wakhī *yiśn* “iron”, which in close agreement with Khotanese *hīśśana-* shows *ś*, can go back to **ā-*, **ham-* as well as zero (Steblin-Kamensky 1999: 25, 35, 49).

10 This is the solution adopted in Dragoni (2022: 63–4), where two distinct stages of prehistoric Khotanese are distinguished, Proto-Tumshuqese-Khotanese for **henśwanyā-* and Pre-Khotanese for **pīθa-*. In this article, we refer to both stages together with the term “prehistoric Khotanese”.

as it seems quite unlikely that the supposed back-formation took place twice in an identical manner. The only alternative is that the borrowing and the back-formation took place in Pre-Tocharian B, after the dissolution of Proto-Tocharian, and that this Pre-Tocharian B **eñcwo* was borrowed into Pre-Tocharian A early enough to undergo the relevant sound laws to become Tocharian A *añcu**.

The second option is based on the observation that not only Khot. *hīśśana-* shows an extra nasal compared to Tocharian B *eñcuwo*, but, compared to Tocharian B *śāñcapo*, Khot. *śśaśvāna-* as well. It should be considered, therefore, that this unexpected correspondence is in fact regular. In this case, Tocharian **eñcwo* would have been borrowed from a prehistoric Khotanese form in which the **y* in the final syllable had already disappeared: **henśwana-*. This scenario would leave the dating of the borrowing relative to the break-up of Proto-Tocharian open, since an early but independent borrowing into Pre-Tocharian B and Pre-Tocharian A would then be much easier to assume. An obvious advantage of this option is that it accounts at the same time for *śāñcapo*, but it remains to be seen how this correspondence should have come about. Could a form like the Khot. nom.sg. in *-āñā* have been interpreted as an inflected form in *°an°* in Tocharian? Such forms are found in the loc.sg., e.g. Toch.B *eñcuwane*, or the obl.pl., e.g. *eñcuwaṃ*, but it is questionable whether these would form a probable model. Neither “iron” nor “mustard” is likely to have frequently occurred in the plural, and *śāñcapo* “mustard” has an obl.sg. in *-o*, and no form of the paradigm had *-a*.

In view of the difficulties with the second option, the first option seems more likely to us. In fact, the parallel between “iron” and “mustard” may be only superficial, since the extra nasal of *śśaśvāna-* must be secondary within Khotanese, so that the Tocharian word may well reflect a shorter form **śaŋzapa-*. One possibility is that *śśaśvāna-* derives from **śśaśva-dānā-* “grain of mustard”, like Sogd. *šyws̄pδn* and Parthian *šyfs̄d'n* (Bailey 1979: 396a). However, the nom.sg. *śśaśvānā* in Z 2.118 (Emmerick 1968b: 31), points to an *a*-stem rather than an *ā*-stem: an *ā*-stem nom.sg. *-ā* is rare in Old Khotanese (Emmerick 1968a: 271). Moreover, enlarged with a **ka*-suffix, *°dāna-* is attested as a second compound member in Late Khot. *pirānaa-* (< **pira-dānaa-*) “worm-seed” and Old Khot. *jūšdānaa-* (< **jūšda-dānaa-*) “musk-grain”, for which see recently Luzziatti (2022: 238). The same suffixed form **dāna-ka-* is reflected also in MPers. *dānag*, NPers. *dāna*. In our view, therefore, the word is not likely to contain a second element **-dānā-*. Another possibility is that *-āna-* is an old adjectival suffix of the type seen in *ysāmāna-* “winter” (Degener 1989: 85). In any case, no matter what the exact origin of the element *-āna-* is, it must be secondary in view of the above-mentioned Sogd. *šyws̄pδn* and Parthian *šyfs̄d'n*, which presuppose the shorter forms *šyws̄p°* and *šyfs̄°*, respectively.

4. Revisiting the Iranian word for “iron”

Now that Khot. *hīśśana-* could be derived from **ham-čwanya-*, it is worthwhile carrying out a survey of the Iranian words for “iron” to see if it is possible to unify the different reconstructions that appear to be needed because of the difficult correspondences.

4.1. Avestan, Ossetic, and Sogdian

As pointed out to us by Alexander Lubotsky (personal communication), Avestan *haosafnaēna-* “out of steel” must be an inner-Iranian borrowing and may be explained as follows. The middle element *-safna-* must be the result of a metathesis of a sequence **-sfana-* in the source dialect: in Avestan, the cluster **-sf-* was not possible, and this metathesis was a strategy to nevertheless keep the **f*. The initial element *hao-* is often interpreted as a *vrddhi* variant of **hu-* (e.g. Abaev 1958: 481), but no variant of “iron” with **hu-* is found anywhere else, and it is unclear what the function of *vrddhi* could have been as *°safnaēna-* already contains an adjective suffix. In view of the reconstruction **ham-ćwana-* on the basis of Khotanese, it is more likely that Av. *hao-* goes back to the same prefix **ham-*, which had apparently become unrecognizable in the source dialect, most probably through a change of **ham-* to **hau-*. The postulated change to **hau-* allowed a simplification of the difficult cluster **-msw-* of **ham-swana-* to **-sw-*; or, if the word was **ham-sfana-* at that point, it allowed a simplification of **-msf-* to **-sf-*.

The source dialect from which **hau-sfana-* was borrowed may be related to Ossetic, since in Ossetic the word is *æfsæn*, whose cluster *fs* < **tsw* may have gone through an intermediate stage **sf*. However, Ossetic appears to show no trace of the prefix **ham-*, which would be expected to have been preserved as *æn-*, seemingly the “normal” development; or as Iron *ys-*, Digoron *in-*, with raising next to a nasal, cf. Iron *yssyn*, Digoron *insun* “sharpen” < **ham-tsai-* (Cheung 2002: 13); or perhaps *w*-umlaut operated. According to the discussion of Cheung (2002: 124–6), *w*- or *u*-umlaut changes a preceding **a* to Iron *y*, Digoron *u* if **a* is found next to a labial and the *u* is apocoped or lost otherwise. If *w*-umlaut operated before the change of **tsw* to **sf* / **sp* > *fs*, it seems that these conditions are met. Still, none of three possible developments of **ham-* would have resulted in loss of the nasal, and it is unclear why a resulting **æn-*, **in-* or **un-* would not have been preserved. It thus seems necessary to assume that the initial *æ-* of *æfsæn* is prothetic, and to derive the Ossetic word from a variant without prefix: **tswana-*. The only alternative is to assume that the nasal was lost before **sf*, as evidently happened in the source dialect of Avestan *haosafnaēna-* too. This is difficult to prove, but not unlikely either. In that case, the initial *æ-* could be the etymological reflex of the original **ham-*.

Even though Avestan *haosafnaēna-* and Ossetic cannot be reconciled so easily, it is tempting to consider the possibility that the Avestan form was borrowed from the “Scythian” dialect assumed by Lubotsky (2002), who attributes to it a change of initial **p-* to **f-*, as in Ossetic, on the basis of **farnah-* “glory” < **parnah-* “abundance, fulness”, borrowed into Avestan as the famous *x^varənah-*. Nothing is known so far about the development of **tsw* in this “Scythian” dialect, and it is not possible to verify whether it became indeed **sf* or something like it, but this dialect at least seems to have had the necessary prestige for borrowing into Avestan. It is tempting to think that the reason for this borrowing was a shift in meaning from the original “iron” to “steel”.

If this “Scythian” dialect is related to Ossetic, as it seems to be because Ossetic also shows the change of **p-* to *f-*, it is worth considering that *u*/*w*-umlaut took place there as well. If so, an alteration of the prefix **ham-* by

w-umlaut may have been a further reason for the adaptation to *hao-* in Avestan. An indication that *u/w*-umlaut occurred in Eastern Iranian steppe dialects is provided by Tocharian B *mot* “wine; alcoholic beverage”. This word is derived from Sogdian *mwδ* “wine” by Pinault (2003: 183; cf. also Tremblay 2005: 438; Peyrot 2015), but this form is attested only in the Christian Sogdian Gospel lectionary E5 (Barbati 2016: 237), too late for borrowing into Tocharian. The normal Sogdian form is *mδw* /*mədu*/ (Brāhmī *mdhu*; see Sims-Williams 1996: 309), which does not fit formally. A putative occurrence of *mwδ* in Ancient Letter IV, l. 5, is rather to be interpreted as “price” (Sims-Williams 1983: 45). A derivation from Sogdian is thus impossible, and the Tocharian word points to a dialect with *u/w*-umlaut of **madu* to **modu* or **mod*. This may well be a dialect from the steppe, possibly Old Steppe Iranian (see §3), since no other known variety qualifies. However, Old Steppe Iranian certainly cannot be equated one-to-one with the source of Avestan *haosafnaēna-*. We will discuss this matter further below.

As with Ossetic, it does not seem possible to derive Sogdian *'spn**, Chr. Sogd. *spn-* from **ham-tswana-*: a form **tswana-* without prefix seems to be needed. In Sogdian, **ham-* regularly becomes *'n-* and this would be expected to be preserved as such. As with Ossetic, the only alternative is to assume a special case of loss of the nasal before **tsw* or the later **sp*. If such loss is not accepted, Sogdian *'spn** points to **spana-* < **tswana-*.

The Sogdian variant *'spyn* reflects **tswana-* (Sims-Williams 2016: 174). This variant is probably to be explained as a secondary, analogical development after “gold” within Sogdian (see also §6). It would otherwise be very difficult to explain the dominant Sogdian variant *'spn** (cf. also Christian Sogdian *spnyq* “iron, of iron”; Sims-Williams 2016: 175), which cannot easily be accounted for if *'spyn* was original.

4.2. Pashto

The Pashto word for “iron” is a feminine noun *óspina*, *óspəna* (also found as *ōspīna* in scientific literature), which is derived from **āćwanyā-* by Morgenstierne (1927: 12), Klingenschmitt (2000: 193) and Cheung (2011: 177). Cheung suggests that the first-syllable accentuation of the Pashto forms is original and reconstructs **áśwan(y)ā-*. The **y* is posited to explain the *-i-* of *óspina*.

The reconstruction of the initial element *o-* is especially difficult. Indeed, a possible source seems to be **ā-*, as in the reconstructions by Morgenstierne, Klingenschmitt and Cheung, cf. e.g. *oba* “water” < **āp-* (Morgenstierne 1927: 9). According to Cheung (2011: 177), a condition for this sound law is that the **ā* was accented, i.e. **ā̂ > o*. Another possible source of *o-* is probably **hau-*, either by assimilation from **ham-ćwanyā-* or by borrowing from another dialect in which such a dissimilation had taken place (see the Avestan in §4.1 above). Yet a third option is possibly **a-*, which according to Cheung (2011: 198) changes to *o* if it is accented and followed by a labial. Indeed, a labial follows in *óspəna*. Also in this case, the **a-* probably needed to be accented, because it would otherwise have been lost, as in *bán* “co-wife” < **ha-páθnī-* (Cheung 2010: 118).

To conclude, Pashto *óspina* could reflect **á-ćwanyā-*, **hau-ćwanyā-*, or **há-ćwanyā-*.¹¹ The first form could be compared with the forms reconstructed below for Persian, Parthian and Kurdish, but these languages point to **ć* instead of **ćw*, and so Pashto would be the only language requiring **ā-* together with **ćw*. The other two forms, **hau-ćwanyā-* and **há-ćwanyā-*, would derive from **ham-ćwanyā-* through assimilation of **mćw* to **ućw* or simplification of **mćw* to *ćw*, respectively. Importantly, the prefix must have been accented, as it was in the **hám-ćwanyā-* leading to Khotanese *hīśśana-*.

4.3. Khwarezmian

In Khwarezmian, the following words connected to the notion of “iron” can be found in Benzing (1983): *spny* m. “iron (Ar. *ḥadīd* ‘iron’)” (var. *’spny*, *sbny*), *spnynk* and *spynykn* “made of iron”, *hnčw* “spearhead” and *hnjw* “having an iron tip”, *θ:hnjw* “iron-pointed staff” (with the Khwarezmian prefix *θ-* “with”). Both *<hnjw>* and *<hnčw>* can be read /*hančwa/*, /*hanjwa/*, /*hanjuwa/*, or /*hanju/*, etc. If they are to be read /*hanjwa/*, /*hančwa/* or /*hanjuwa/*, /*hančuwa/*, the final *-a* could be read as a feminine nom.-acc. If the reading is /*hanju/*, /*hanču/*, it would be a masculine noun.

Of these forms, *spny* seems to be the genuine Khwarezmian form, with *sp* from PIr. **ćw-* > *sp-* as in *sptyr*, *spdyr* “mule” < **aćwa-tara-*. There is no trace of a prefix, and the final *-y* suggests a preform **tswanya-* < **ćwanyā-*.¹² On the other hand, *hnčw* cannot be inherited: apart from *-ćw*, which cannot go back to **ćw*, *h-* points to borrowing, since original **h-* was dropped in pre-vocalic position as in *’pn* “co-wife”, from Old Iranian **hapaθnī-* (cf. Durkin-Meisterernst 2009: 341).¹³ There is no exception to this, and all Khwarezmian words that start with *h-* are loanwords.¹⁴ The prefix **ham-* is also reduced to *an-* in inherited Khwarezmian words, as in *’nbnc* “the (middle of the) way” < **ham-panti-*, *’nbθ* “companion” < **ham-paθ-* (cf. Emmerick 1970: 68), *’nbnc(y)* “to intend” < **ham-bandaya-* (cf. Henning 1971: 29), etc.

Thus, because of the initial *h-* and the medial *č*, *ǰ*, it can safely be assumed that Khwarezmian *hnčw*, *hnjw* is a loanword, borrowed after the loss of original word-initial **h-* in Khwarezmian. As a possible source language, one could think of prehistoric Khotanese **hen-śwanyā-*, since this would account for the initial *h-* and the palatal cluster *nčw*, *nǰw*. However, this presupposes loss of the nasal final of Khotanese, which is difficult.¹⁵ Furthermore, we assume that the prehistoric Khotanese form borrowed into Tocharian had **nśw*, not **nćw*, although the latter option is difficult to exclude. Another obvious problem is that Khotan and Chorasmia are geographically far apart from each other. Perhaps the

11 **á-ćwanyā-* is also possible, but has no parallels in the reconstructions we propose.

12 The assumption of borrowing from or into Sogdian *’spn-*, *’spyn-*, etc. (Gharib 1995: 64) is unnecessary.

13 Except **hy-* and *hu-* which become *xu-* in Khwarezmian. The word *hwny* “blood” shows that **wahūnī* “blood” (Av. *vohunī-*) became *hūnī* only after *hu-* became *xu-*.

14 Khwarezmian *hrs* “bear (Arabic *dubb* ‘bear’)” remains problematic. In this word the *h-* is to be compared with the *x-* of Persian *xirs*. Possibly, therefore, *hrs* “bear” is a borrowing in Khwarezmian.

15 Tocharian is lacking this nasal final, too, but can be excluded as a possible source of borrowing because the language does not have **h* at all.

Khwarezmian word was borrowed at a very early stage, when prehistoric Khotanese had not yet moved into the Tarim Basin and still had **nčw*, but it is questionable whether **h-* was lost so early in Khwarezmian. In view of all these problems, it seems likely that if *hnčw*, *hnjw* is to be explained from prehistoric Khotanese, it was not borrowed directly, but mediated by a third language.

4.4. Persian, Parthian, and Kurdish

Although the Achaemenids must have known iron,¹⁶ there is no trace of a word meaning “iron” in Old Persian. In Zoroastrian Middle Persian (Pahlavi) the word for “iron” is <’hyn’>, <’syn’>, which we think should be read as /āhen/ or /āhin/ (cf. also Skjærvø 1994: 271);¹⁷ in Manichaean Middle Persian it is ’’*hwn*, transcribed as /āhun/ by Durkin-Meisterernst (2004: 35); and in Parthian it is (’)*’swn*. Durkin-Meisterernst’s transcription of ’’*hwn* with short *u* must be based on etymological considerations. As far as the spelling is concerned, both ’’*hwn* and (’)*’swn* could also have long *ū*, i.e. *āhun*, *āsun* or *āhūn*, *āsūn*. The *ū*-vocalism of Manichaean Middle Persian ’’*hwn* must have been taken over from Parthian (’)*’swn*.¹⁸ New Persian has *āhan* for “iron”, which cannot be derived from either *āhen*, *āhin* or *āhūn*.

The New Persian and Parthian forms are difficult to reconstruct. Klingenschmitt (2000: 193) derived the Manichaean Middle Persian and Parthian forms from **āsuna-*, in our notation **ācuna-*.¹⁹ Indeed, while a reconstruction with **čw* is not possible, a variant with **č* but without the following **w* would indeed yield the right correspondences for the medial consonant. However, the correspondences of the vowels remain unresolved. While Parthian might indeed reflect **ā-cun(y)a-* directly, New Persian *āhan* suggests **ā-čan(y)a-*. The only option for Middle Persian *’h(y)n*, /āhen/ or /āhin/, seems to be **ā-čanya-*, but the vowel *e* or *i* is, nevertheless, not regular and should originally have been a dialectal feature. A close parallel is offered by the alternation between *Ahriman* ~ *Ahri/amen* and *dušman* ~ *dušmen* in Middle and New Persian varieties, which reflects final **ny* too, < **-manyu-*.

In Kurdish we find Sanandajî *āsin*, Kurmanjî (*h*)*āsin*, *hesin*, and in Zazaki *asi*, *āsin*. These words cannot derive from a preform with **čw* either, since the cluster **čw* becomes *-sp-* as in Kurdish *hesp* “horse”, *spî*, *spah* “louse”.

16 We know of iron knives in an Achaemenid context (cf. Moorey 1982: 96); a Neo-Elamite text from the Acropole Archive (s 49) mentions iron explicitly (cf. Henkelman 2008: 361); and iron was used in the construction of Cyrus’ tomb at Pasargadae (cf. Stronach 1985: 840).

17 Pace Klingenschmitt (2000: 193), the variants with *y* do not reflect /āhēn/, since *y* would always be written if the vowel was long. Here, it is not always written and therefore it is rather a *mater lectionis*, representing a short *e* or *i*. This short *e* or *i* is also found in other Iranian languages, such as Balochi *āsin* (Korn 2005: 87), or in the closely related Bakhtiari language, where one finds the form *āhen* (Anonby and Asadi 2014: 201).

18 A parallel can be observed in the Balochi form *āhin* “iron”, which can be interpreted as Persian *āhan* with the *i* taken over from genuine Balochi *āsin* (so Korn 2005: 193) or, conversely, as Balochi *āsin* with the *h* taken over from Persian *āhan*.

19 He derives Pahlavi <’*h(y)n*’> from **āsyan(i)ā-*. This is not possible because **čw* (his notation *su*) does not become *h* in Pahlavi.

As in Persian, these forms may derive instead from a preform with **ć*, not **ćw*. The vowel *i* in the second syllable seems to point to **ā-ćanya-* with **y*.

Balochi *āsin*, in turn, could have *-s-* as a regular result of PIr. **ćw-* (Korn 2005: 90–91) as well as **ć-*. In theory, one might try to derive the Kurdish forms from Balochi, since they could then be derived from a preform with **ćw*. Indeed, Balochi is likely to have originally been spoken on the border of the Caspian Sea, and borrowing between Kurdish and Balochi occurred (cf. Korn 2005: 51, and especially Korn 2006). However, it is unlikely that a Balochi form was borrowed so widely, and it would not account for the form of the Persian and Parthian words. It is thus better to group Balochi with Kurdish and Persian and derive it from **ā-ćanya-*.

Thus, the most likely reconstruction for Persian, Balochi, Kurdish, and Zazaki, etc., is **ā-ćanya-*. For Parthian, which influenced the vocalism of Manichaean Middle Persian *ʿhwn*, a source form **ā-ćūn(y)a-* seems needed.

5. The Iranian word for “iron”: reconstructions and etymology

In the preceding section, we have made reconstructions for several groups of Iranian words for “iron”. We now attempt to reduce the number of reconstructions further. In this way, we think we can come closer to a unified reconstruction and an etymology. Even though full unification appears to be impossible, the variation in preforms can be reduced to such an extent that the assumption of a substrate term or a Wanderwort is no longer warranted. Only Khwarezmian *hnjw*, *hnčw* cannot be inherited and must have been borrowed if it is related.

The reconstructions for the Iranian words for “iron” that we have made thus far are the following:

- 1) Khot. *hīśśana-* < **ham-ćwanya-*.
- 2) Av. *haosafnaēna-*, borrowed from a form going back to **ham-tswana-* < **ham-ćwana-*.
- 3) Sogdian *ʿspn** and Ossetic *æfsæn* < **tswana-* < **ćwana-* (only to be derived from **ham-ćwana-* with the assumption of special loss of **m* in the cluster **mćw* or its later stages **mtsw* or **m̄sp*).
- 4) Khwarezmian *spny* < **tswana-* < **ćwanya-*.
- 5) Pashto *óspina*, *óspəna* < **-tswanā-* or **-tswana-* < **-ćwan(y)ā-* (the prefix can be reconstructed in various ways).
- 6) Parthian *ʿswn*, *ʿswn* < **ā-ćūna-* or **ā-ćūnya-* (the *ū*-vocalism has probably influenced Manichaean MPers. *ʿhwn*).
- 7) Kurdish *āsin*, *hāsin*, *hesin* < **ā-ćanya-*.
- 8) Balochi *āsin* < **ā-ćanya-* or **ā-ćwanya-*.
- 9) Pahlavi <ʿhyn>, <ʿhn> probably stands for /āhen/ or /āhin/, apparently from **ā-ćanya-* with a dialectal reflex of *e* or *i* because of the final **-ya*.

In these reconstructions, there is variation in the prefix, the root, and the final, which is *-a* or *-ya*.²⁰ In spite of this large variation, it seems that two main

20 As noted above in §4.2, Pashto requires **-ā* or **-yā*. Most other languages allow both **-a* and **-ā*, except for Ossetic and Khotanese, which require **-a* and **-ya*, respectively.

groups can be distinguished: an eastern group with *-ćw- and a western group with *-ć-. The variation in the prefix conforms to these main groups as well: the prefix is *ā- in the western group, and in the eastern group it is *ham- (as in Khot. *hīśśana-*) or it is – at least superficially – lacking (as in Oss. *æfsæn*). We may thus tentatively assume that Pashto *o-* in *óspina*, *óspəna* reflects *ham-, because the *-sp-* goes back to *-tsw- < *-ćw-. Likewise, Balochi *āsin*, which could reflect both *āćanya- or *āćwanya-, is now best derived from *āćanya-. We propose a tentative solution to connect the reconstruction of the western and eastern groups below, but since this is better understood in light of our etymology, we will discuss the etymology first.

Abaev (1958: 481) derives the Iranian word for “iron” from the root set up as *suan- (< *tswan- < *ćwan-) “benefit, bless” by Cheung (2007: 370). Indeed, this could formally work for the eastern group, but the semantic connection to “iron” is far from compelling. The semantic range attested for this root includes “useful” on the evidence of Sogdian *ptspyn-* “be useful, beneficial”, and obviously iron may be beneficial and useful, but so are many other things.

A better match semantically is offered by the root set up as *suaH- “move, throw, erect” (< *tsuaH- < *ćwaH-) by Cheung (2007: 369–70).²¹ For the semantic connection between “throw” and “iron”, it is tempting to compare English *cast*, which means “throw” (amongst many other things) and is used in the term *cast iron*. A problem with this comparison is that there is no indication that the Iranian words referred specifically to cast iron. As a second option, it may be suggested that *ćwaH- with the preverb *ham- “together” meant “put together, mix”, a shade of meaning of *ćwaH- that is attested in Late Khot. *niśś-: niśśāta-* (< *ni-ćwaH-; Emmerick 1968a: 56), used in the Siddhasāra to describe the action of throwing together different ingredients in a vessel to prepare a medicament or a special drug (see e.g. Si §2.34, §2.35). *ham-ćwaH- could then have referred to mixing iron with e.g. carbon to make steel. Indeed, this would fit the Avestan meaning “out of steel” for *haosafnaēna-*. However, many variants of the word for iron have no prefix, or the prefix *ā- instead, and Avestan is the only language showing the specific meaning “steel”. A third, and probably the best option is to take *ćwaH- to refer to striking iron in smithing. French *cingler* provides a neat semantic parallel, since it refers exactly to this, and further means “whip, lash”. An even better parallel is P’urhepecha (Mexico) *tayacata* “silver”, which derives from the root *taya-* “to give blows” (Bellamy 2018: 8). From Iranian we may cite Ossetic Iron *caēgdyn*, Digoron *caēgdun*, which means (amongst others) “strike iron” and derives from P’Ir. *ćak/g- “hit, strike” (Cheung 2007: 31–2). This third option also accounts best for the variation in the preverb: the basic form may have been *ćwaH-* “strike”, while *ham-ćwaH- and *ā-ćwaH- meant “strike together” and “strike at”, respectively.

For the eastern group, the form of the word can be explained with the assumption of a suffix *-ana- to the zero grade of the root: *ćuH-ana-. As noted by LIV (p. 339), this requires a secondary replacement of the expected

21 On the basis of Iranian only, LIV (p. 339) sets up *kueH- “werfen” with a question mark, separating it from *kueh- “anschwellen”.

**ćuw* by the attested **ćw*. In other words, the formation **ćuH-ana-* > **ćuwana-* was analogically adapted to the verb, which had **ćw* in the full grade variant **ćwaH-*. Indeed, the verb is attested, with reflexes of the cluster **ćw*, in several eastern languages: Av. *spā-* < **ćwaH-*, Sogd. *'pspy-* < **apa-ćwaH-ya-* and Khot. *paśś-* < **apa-ćwaH-* and *niśś-* < **ni-ćwaH-*.

The protoform **ćuH-ana-* offers perspectives to connect the reconstructions for the western group, since in this reconstruction **ć* is found before **u* rather than **w*, and would regularly develop into MPers. *h* and Parthian and Kurdish *s*, etc. However, it is by no means obvious how the vocalism of the respective reconstructed forms is to be explained from **uHa*. We make a number of suggestions below, but note that these remain tentative.

The only forms with *ũ*-vocalism are Parthian *'swn*, *'swn* and Manichaean MPers. *'hwn*, which we believe was influenced by Parthian. As noted above, Durkin-Meisterernst (2004: 35) interprets the *ũ*-vocalism as short *u*, but it could also be long *ū*. If this latter interpretation is correct, Parthian *'swn* /*āsūn*/ could continue **ā-ćuwana-* or **ā-ćuwanya-* directly. We would then interpret Manichaean MPers. *'hwn* as /*āhūn*/ too.

The Persian, Kurdish and Balochi forms together point to **ā-ćanya-*, but this is not so easily derived from **ā-ćuwanya-*. We tentatively propose that these forms are to be explained from a secondary *w*-cluster, which arose from **uwa* after primary **ćw* had developed further, so that it was treated differently. For Persian, the development of primary **ćw* may have been as follows: **ćw* > **tsw* > **sw* > *s*. That is, the following **w* caused a change of **ć* > **ts* to *s*, so that **ts* here did not develop to *θ* as elsewhere (see also n. 24). We propose that secondary “**ćw*” arose after the simplification of **tsw* to **sw*, probably in the form **tsw*. However, at this stage the same cluster was simplified differently: instead of the first element **t*, it was this time rather the last element **w* that was lost.

The proposed development is schematized in Table 1.

Obviously, it is possible to assume that the hypothesized **sw* < **ćw* and **tsw* < **ćuw* were simplified to **s* and **ts* at the same time in Persian, since no reflex of **w* is found in either case. In Kurdish, however, a reflex of **w* is found in *sp* < **ćw*, so that we need to assume that in primary **ćw* > **tsw* the **w* was not lost completely. In other words, primary **ćw* > **tsw* developed to *sp*, while secondary “**ćw*” developed as follows: **ćuw* > **tsuw* > **tsw* > **ts* > *s*. As remarked above, Balochi points to **āćanya-* or **āćwanya-*, but assuming it is to be grouped with the other western languages, the merger of primary **ćw* with secondary “**ćw*” could have occurred at any of several stages.

Table 1. The proposed development of **ćw* and **ćuw* in Persian.

Primary * <i>ćw</i>	Secondary “* <i>ćw</i> ” < * <i>ćuw</i>
* <i>ćw</i>	* <i>ćuw</i>
* <i>tsw</i>	* <i>tsuw</i>
* <i>sw</i>	* <i>tsuw</i>
* <i>sw</i>	* <i>tsw</i>
* <i>sw</i>	* <i>ts</i>
<i>s</i>	<i>θ</i>

We are aware that the proposed different outcomes of secondary “*ćw” are hypothetical, but point out that secondary “*ćw” from *ćuw must have been a very rare cluster, so that a special explanation is warranted. Nevertheless, it is clear that variation in the terms for iron remains: the western group needs *ā-ćuHan(y)a-, while the eastern group needs *ham-ćuHan(y)a-, and perhaps in addition *ćuHan(y)a- without prefix. Further, some forms require a preform *-ćuHana-, while there are many forms that presuppose *-ćuHanya-.²² At present, we have no definitive explanation for this latter variation: the *y was probably taken over analogically from *jaryana- > *dzaryana- “gold”, but we cannot at present say at which stage this adaptation occurred or how often it occurred in parallel, but independently.

Keeping in mind the many uncertainties discussed above, our proposal to reduce the variation in the preforms needed for the Iranian words for “iron” may be summarized as follows:

- 1) *ham-ćuHana-: with replacement of *ćuH by *ćw, probably analogical after the verb *ćwaH-, this developed to *ham-ćwana-. This form is the source of (in many cases the final is -ya):
 - a) Khot. *hīśšana-* < *ham-ćwanya-;
 - b) Av. *haosafnaēna-*, borrowed from a form going back to *ham-ćwana-;
 - c) Pashto *óspina*, *óspəna* < *ham-ćwan(y)ā-.
- 2) *ćuHana-: as above, with replacement of *ćuH by *ćw, and sometimes with the final -ya. This *ćwana- can only be derived from *ham-ćwana- if special loss of the nasal in the cluster *mćw is assumed. *ćwana- is the source of:
 - a) Sogdian *’spn**;
 - b) Ossetic *æfsæn*;
 - c) Khwarezmian *spny*.
- 3) *ā-ćuHanya-: this form is the source of (some forms need -ya, the other forms may have had -ya too):
 - a) Parthian *’swn*, *’swn*, if /āsūn/, < *čūnya- < *ā-ćuwanya- < *ā-ćuHanya-;
 - b) Kurdish *āsin*, *hāsin*, *hesin*, Balochi *āsin*, Pahlavi <’hyn>, <’hn> < *ā-ćanya-, with loss of *w in secondary *ā-ćwanya- < *ā-ćuwanya- < *ā-ćuHanya-.

The variation between *ā-ćuHana-, *ham-ćuHana- and *ćuHana-, as well as possibly that between the forms with final *-ya and those without, is probably due to the fact that iron spread throughout the Iranian world around the turn of the second to the first millennium BCE (Moorey 1982; Pigott 2004), long after the break-up of Proto-Iranian.

However, in spite of this remaining variation,²³ our reconstructions account for most of the correspondences that were always seen as irregular. It is in

22 So many forms require *y that one may even consider that it was lost in some cases, like for instance in Av. *haosafnaēna-*, where simplification of an earlier **hau-safnyaina-* would not be surprising.

23 It is conceivable, but in our view impossible to prove currently, that *ā-ćuHana- is adapted from a variant of *ham-ćuHana- that arose through a development of the cluster

our view unlikely that the word for “iron” is a substrate term or a Wanderwort. The formation is genuinely Iranian and shows no marks of borrowing.

6. The Proto-Indo-Iranian cluster *ćy̅ in Khotanese-Tumšūqese

As a corollary, our derivation of Tocharian B *eñcuwo* from Khotanese provides important evidence for the development of the Proto-Indo-Iranian cluster *ćy̅ in the Khotanese-Tumšūqese branch. It is well known that the Proto-Indo-Iranian cluster *ćy̅ as in Skt. *ásva-* “horse”, Av. *aspa-* “id.” < **ac̣ua-* is represented in Khotanese as *ś(ś)* (Emmerick 1989: 216). This strongly suggests that this cluster remained palatal in Khotanese, and was never depalatalized to **sw* or **tsw* as in all other Iranian languages except Tumšūqese and Wakhī (see e.g. Sims-Williams 1998: 136; Peyrot 2018). This special archaism of Khotanese, Tumšūqese and Wakhī gives these languages a special status within the Iranian branch: apparently, these languages form a branch that left the Proto-Iranian speech unity at an early date, after which the remaining unitary variety, or close-knit group of dialects, underwent the well-known further development to **tsw* at first, and then to **sw* > Av. *sp*, Old Persian *s*,²⁴ etc.

Although it was fairly improbable, it was in fact theoretically possible that Khotanese *ś(ś)* does not continue Proto-Indo-Iranian *ćy̅ directly. Since **sy* < Proto-Indo-Iranian *ć̣i̅ is known to become Khot. *ś(ś)* too, there was still a minor option that Proto-Indo-Iranian *ćy̅ developed to **tsw* > **sw* as in the other Iranian languages, after which a change of **sw* to **sy* > *ś(ś)* took place. We are not aware of any evidence in favour of this complicated alternative course of events, but it was difficult to exclude with certainty.

With our explanation of Tocharian B *eñcuwo*, there is now for the first time clear evidence that in prehistoric Khotanese the cluster was palatal and contained **w* at the same time. On the basis of *śāñcapo* (see above), we suppose the prehistoric Khotanese cluster was **św* at the time of borrowing, not *ćw, but if it was in fact *ćw, this would not affect our argument. Since **ś* (or perhaps *ć) was found together with **w*, it is impossible that the **ś* is due to a change of **w* to **y*.

Another alternative explanation for Khot. *ś(ś)* has been offered by Kümmel (2007: 234; see also Novák 2013: 121–2), who attributes the palatal character of Khot. *ś(ś)* to rounding of **s* by the following **w*. In other words, **s*^w < **sw* < *ćw should have merged with **ś* < **sy* < *ćy when the **w* was lost because **ś* was phonetically [ś^w]. Although there are no indications that Khot. *ś(ś)* was rounded, this is difficult to exclude. However, it is obvious that if it was rounded, this must have been a secondary feature, and it is in our view very unlikely that a rounded dental *s* was shifted to become a rounded palatal

**mć*. In our view, it is more likely that these two basic variants reflect a difference in morphological formation rather than a phonological development.

24 It is difficult to rule out that Old Persian *s* has developed from *ćw directly. However, if an intermediate stage **tsw* is assumed, the change to *s* can be conceived of as a simplification of **ts* to **s* before **w*: **tsw* > **sw* > *s* (see also §5 above). This simplification is paralleled by the development of PIr. *ćm > **tsm* to OP *sm*.

ś: in Khotanese, ś clearly is the palatal counterpart of *s*. The evidence we here adduce for an earlier stage *św (or possibly *św) of Khot. ś(ś) is a further counter-argument against Kümmel's explanation. Since it shows that *ś and *w were found together in the same cluster, it follows that ś cannot have arisen secondarily through the merger of *ś < *śy with *ś^w < *św after the loss of *w.

Thus, our explanation of *eñcuwo* further confirms that Khotanese (and with it, Tumšūqese and Wakhī) preserves an archaism in the palatal outcome of the Proto-Indo-Iranian cluster *śy and therefore split off from Proto-Iranian earlier than any other known dialect or branch of Iranian.

7. Iron and bronze in the Tarim Basin

Another interesting consequence of our finding that the Tocharian word for “iron” derives from Pre-Khotanese is the fact that it allows important inferences about the prehistory of the Tarim Basin. Most importantly, Tocharian B *eñcuwo* “iron” cannot have been borrowed from Old Steppe Iranian, the dialect that is the source of an important layer of Iranian borrowings (see above, §3). Although this prehistoric dialect is only indirectly attested, the number of borrowings from this dialect into Tocharian is sufficiently large to establish regular sound correspondences that allow us to accept or reject further etymologies (Bernard, ongoing research). These correspondences may even be used to formulate exact hypotheses about the shape of thus far unattested borrowings from this dialect into Tocharian.

Based on the sound correspondences of Old Steppe Iranian borrowings established thus far, the following hypotheses may be set up for the various possible forms of the word for “iron”:

- **tswana-* (< *śwana-): Toch.B ***tswene*, Toch.A ***tswaṃ*. The vowels **a* *a* would certainly have been represented with Toch.B *e* *e*, and the cluster **tsw* is represented as *tsw* in Toch.B *etswe* ← **atswa-* “horse” (< **acwa-*; Peyrot 2018).
- **ham-tswana-* (< **ham-śwana-*): Toch.B ***entswene*, Toch.A ***ontsaṃ* (?). The **h-* of **ham-* would certainly have been lost, and the **m* would probably have been assimilated to **n* before **ts*. In Tocharian A, *w*-umlaut could have taken place, which would have given ***ontsaṃ*; without *w*-umlaut, ***antswaṃ*. In words with the vowels *e* *e* *e*, the middle syllable is syncopated. If that happened here too, Toch.B ***entswne*, ***entsune* or ***entsne* would be expected.
- **hom-tswana-* (< **ham-tswana-* with *w*-umlaut): Toch.B ***ontswene*, Toch.A ***ontswaṃ*.

Importantly, in the stage of this dialect that yielded so many borrowings in Tocharian, the cluster **tsw* had not developed to the **sf* or **tsf* needed for Avestan *haosafnaēna-* (Peyrot 2018). If Avestan has borrowed that form from the same or a closely related dialect, it must be from a later stage.

None of the hypothetical forms postulated above is attested in Tocharian, and all these forms are clearly different from Toch.B *eñcuwo*, Toch.A *añcu**, the terms that are actually found. At the same time, it has been noted before that many borrowings from Old Steppe Iranian are military terms and terms for

weapons,²⁵ like Toch.B *retke* “army” (cf. MPers. *radag* “line, row”), *tsain* “arrow” (cf. Av. *zaēnu-* “baldric”), *kertte* “sword” (cf. Av. *karāta-* “knife”), *peret* “axe” (cf. Oss. *færæt* “axe”). The fact that Tocharian borrowed so many words for weapons from this dialect, but not the word for iron, which was instead borrowed from prehistoric Khotanese, strongly suggests that at the time these weapon terms were borrowed, iron was not yet in use, or its use was not widespread. Although it cannot be proved in the strict sense, it thus appears that the *tsain* “weapon”, *kertte* “sword” and *peret* “axe” that the Tocharians took over from the Iranians were made of bronze. Only later, when the Tocharians came into contact with prehistoric Khotanese, they must have become well acquainted with iron, when they borrowed the word for it from this other archaic Iranian dialect.

Thus, the fact that the Tocharian word for iron has been borrowed from prehistoric Khotanese and not from Old Steppe Iranian, from which otherwise many terms for weapons were borrowed, suggests that contacts with Old Steppe Iranian took place at an early period: in the Bronze Age, or at the latest at a time when the use of iron was not widespread. It appears that the Andronovo presence in northern Xīnjiāng, dated to the thirteenth to ninth centuries BCE by Kuz'mina (2008: 98–107) provides a neat match for the contacts between Tocharian and Old Steppe Iranian. Only for the early first millennium BCE are there iron finds from Xīnjiāng (Guo 2009).

On the other hand, prehistoric Khotanese is another archaic Iranian dialect that does not descend from Old Steppe Iranian. While Old Steppe Iranian was probably spoken in northern Xīnjiāng, prehistoric Khotanese (or Khotanese-Tumšūqese) probably entered the Tarim Basin from the west, which fits the historical distribution of Khotanese in the southwest Tarim Basin and the closely related Tumšūqese in the northwest. The best archaeological match for the earliest stage of this Iranian branch in the Tarim Basin seems to be the Aqtala Culture, which spread in the western Tarim Basin from the early first millennium BCE (Peyrot 2018).

It has been suggested that iron was introduced into the Tarim Basin from the west too, since iron is attested relatively early in the Chust Culture of Fergana, and there are cultural links between the Chust Culture and Xīnjiāng (Mei and Shell 2002: 229–30; Wagner 2008: 97; Mallory 2015: 25). If this is correct, it is possible that prehistoric Khotanese speakers introduced iron in the Tarim Basin. However, iron spread quickly in the early first millennium BCE, as shown also by ninth-century finds from Arzhan in Tuva (Guo 2009: 108), and no conclusive routes for its spread could thus far be established on the basis of the relatively small number of finds (see Mei et al. 2015: 226 for iron in Central China). Therefore, it is very well possible that iron was introduced in parallel in northern Xīnjiāng. For our purposes, however, it suffices to say that speakers of prehistoric Khotanese made iron known to the early Tocharians.

25 Winter (1971) makes this semantic argument but misidentifies the language as Bactrian. See further Schmidt 1985.

References

- Abaev, Vasilij I. 1958. *Istoriko-ëtimologičeskij slovar' osetinskogo jazyka*. Tom 1. A–K. Moscow: Akademija Nauk SSSR.
- Adams, Douglas Q. 1999. *A Dictionary of Tocharian B*. Amsterdam: Rodopi.
- Adams, Douglas Q. 2013. *A Dictionary of Tocharian B*. Second edition, revised and greatly enlarged. Amsterdam: Rodopi.
- Anonby, Erik and Ashraf Asadi. 2014. *Bakhtiari Studies. Phonology, Text, Lexicon*. Uppsala: Uppsala Universitet.
- Bailey, Harold W. 1979. *Dictionary of Khotan Saka*. Cambridge: Cambridge University Press.
- Barbati, Chiara. 2016. *The Christian Sogdian Lectionary E5 in Context*. Vienna: Österreichische Akademie der Wissenschaften.
- Bellamy, Kate. 2018. “Investigating interaction between South America and West Mexico through the lexicon of metallurgy”, in Guus Kroonen and Rune Iversen (eds), *Digging for Words*. Oxford: BAR, 1–19.
- Benzing, Johannes. 1983. *Chwaresmischer Wortindex*. Wiesbaden: Harrassowitz.
- Buyaner, David B. 2013. “K ètimologii toxarskogo nazvanija železa”, in Sergej R. Toxtas'ev and Pavel B. Lurje (eds), *Commentationes Iranicae. Sbornik statej k 90-letiju V.A. Livšica*. Saint Petersburg: Nestor-Istorija, 608–13.
- Chen, Ruixuan and Chams Benoît Bernard. Forthcoming. “A spicy etymology. Remarks on Tocharian AB *sāncapo*”.
- Cheung, Johnny. 2002. *Studies in the Historical Development of the Ossetic Vocalism*. Wiesbaden: Reichert.
- Cheung, Johnny. 2007. *Etymological Dictionary of the Iranian Verb*. Leiden: Brill.
- Cheung, Johnny. 2010. “Selected Pashto problems I: The accent in Pashto”, *Persica* 23, 109–21.
- Cheung, Johnny. 2011. “Selected Pashto problems II. Historical phonology 1: On vocalism and etyma”, *Iran and the Caucasus* 15, 169–205.
- De Chiara, Matteo. 2014. *The Khotanese Sudhanāvādāna. Commentary*. Wiesbaden: Harrassowitz.
- Degener, Almuth. 1989. *Khotanische Suffixe*. Stuttgart: Steiner.
- Dragoni, Federico. 2021. “Materia medica Tocharico-Hvatanica”, *Bulletin of the School of Oriental and African Studies* 84/2, 295–319.
- Dragoni, Federico. 2022. *Watañi lāntaṃ*. Khotanese and Tumshuqese loanwords in Tocharian. PhD dissertation, Leiden University.
- Durkin-Meisterernst, Desmond. 2004. *Dictionary of Manichaean Middle Persian and Parthian*. Turnhout: Brepols.
- Durkin-Meisterernst, Desmond. 2009. “Khwarezmian”, in Gernot Windfuhr (ed.), *The Iranian Languages*. London: Routledge, 336–76.
- Emmerick, Ronald E. 1968a. *Saka Grammatical Studies*. London: Oxford University Press.
- Emmerick, Ronald E. 1968b. *The Book of Zambasta. A Khotanese poem on Buddhism*. London: Oxford University.
- Emmerick, Ronald E. 1970. “Some Chorasmian and Khotanese etymologies”, *Journal of the Royal Asiatic Society of Great Britain and Ireland* 1, 67–70.
- Emmerick, Ronald E. 1989. “Khotanese and Tumshuqese”, in Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert, 204–29.
- Gharib, Badr-az-Zamān. 1995. *Sogdian Dictionary: Sogdian–Persian–English*. Tehran: Farhang.

- Guo, Wu. 2009. "From western Asia to the Tianshan Mountains: on the early iron artefacts found in Xinjiang", in Jianjun Mei and Th. Rehren (eds), *Metallurgy and Civilisation: Eurasia and Beyond*. London: Archetype, 107–15.
- Henkelman, Wouter F.M. 2008. *The Other Gods Who Are. Studies in Elamite–Iranian Acculturation Based on the Persépolis Fortification Texts*. Leiden: Nederlands Instituut voor het Nabije Oosten.
- Henning, Walter B. 1965. "A grain of mustard", *Annali dell'Istituto Universitario Orientale di Napoli*, Sezione linguistica 6, 29–47.
- Henning, Walter B. 1971. *A Fragment of a Khwarezmian Dictionary*. Edited by N. MacKenzie. London: Lund Humphries.
- Hoernle, A.F. Rudolf. 1901. "A report on the British Collection of antiquities from Central Asia", *Journal of the Asiatic Society of Bengal*, Extra Number 1.
- Klingenschmitt, Gert. 2000. "Mittelpersisch", in Bernhard Forssman and Robert Plath (eds), *Indoarisch, Iranisch und die Indogermanistik: Arbeitstagung der Indogermanischen Gesellschaft vom 2. bis 5. Oktober 1997 in Erlangen*. Wiesbaden: Reichert, 191–229.
- Korn, Agnes. 2005. *Towards a Historical Grammar of Balochi. Studies in Balochi Historical Phonology and Vocabulary*. Wiesbaden: Reichert.
- Korn, Agnes. 2006. "Counting sheep and camels in Balochi", in Mixail N. Bogoljubov et al. (eds), *Indoiranskoe jazykoznanie i topologija jazykovyx situacij. Sbornik statej k 75-letiju professora A.L. Grjunberga (1930–1995)*. Saint Petersburg: Nauka, 201–12.
- Kümmel, Martin J. 2007. *Konsonantenwandel. Bausteine zu einer Typologie des Lautwandels und ihre Konsequenzen für die vergleichende Rekonstruktion*. Wiesbaden: Reichert.
- Kuz'mina, Elena E. 2008. *The Prehistory of the Silk Road*. Philadelphia: University of Pennsylvania.
- LIV = Helmut Rix et al. (eds). 2001. *Lexikon der indogermanischen Verben. Die Wurzeln und ihre Primärstammbildungen*. Zweite, erweiterte und verbesserte Auflage. Wiesbaden: Reichert.
- Lubotsky, Alexander M. 2002. "Scythian elements in Old Iranian", in Nicholas Sims-Williams (ed.), *Indo-Iranian Languages and Peoples*. Oxford: Oxford University, 189–202.
- Luzzietti, Silvia. 2022. "Lexical notes on the Khotanese *Piṇḍasāstra*", *Indo-Iranian Journal* 65, 227–48.
- Mallory, James P. 2015. *The Problem of Tocharian Origins: An Archaeological Perspective*. Philadelphia: University of Pennsylvania.
- Mei, Jianjun and Colin Shell. 2002. "The Iron Age cultures in Xinjiang and their steppe connections", in K. Boyle et al. (eds), *Ancient Interactions: East and West in Eurasia*. Cambridge: McDonald Institute for Archaeological Research, 213–34.
- Mei, Jianjun et al. 2015. "Archaeometallurgical studies in China: some recent developments and challenging issues", *Journal of Archaeological Science* 56, 221–32.
- Moorey, Peter R. Stuart. 1982. "Archaeology and pre-Achaemenid metalworking in Iran: A fifteen year retrospective", *Iran* 20, 81–101.
- Morgenstierne, Georg. 1927. *An Etymological Vocabulary of Pashto*. Oslo: Dybwad.
- Novák, Lubomír. 2013. Problem of Archaism and Innovation in the Eastern Iranian Languages. PhD dissertation. Prague.
- Peyrot, Michaël. 2015. "Tocharian language", *Encyclopædia Iranica*. www.iranicaonline.org/articles/tocharian-language
- Peyrot, Michaël. 2018. "Tocharian B *etswe* 'mule' and Eastern East Iranian", in *Farnah. Indo-Iranian and Indo-European Studies in Honor of Sasha Lubotsky*. Ann Arbor: Beech Stave, 270–83.

- Pigott, Vincent C. 2004. "Hasanlu und das Auftreten des Eisens in Westiran im frühen 1. Jahrtausend v. Chr.", in Thomas Stöllner et al. (eds), *Persiens antike Pracht. Bergbau, Handwerk, Archäologie*. Bochum: Deutsches Bergbau-Museum, 350–57.
- Pinault, Georges-Jean. 2003. "Sur les thèmes indo-européens en *-u-: dérivation et étymologie", in Eva Tichy, Dagmar S. Wodtke and Britta Irslinger (eds), *Indogermanisches Nomen. Derivation, Flexion und Ablaut. Akten der Arbeitstagung der Indogermanischen Gesellschaft. Freiburg, 19. bis 22. September 2001*. Bremen: Hempen, 153–88.
- Pinault, Georges-Jean. 2006. "Further links between the Indo-Iranian substratum and BMAC", in Bertil Tikkänen and Heinrich Hettrich (eds), *Themes and Tasks in Old Middle Indo-Aryan Linguistics*, vol. 5. Delhi: Motilal Banarsidass, 167–96.
- Sadovski, Velizar. 2017. "The lexicon of Iranian", in Jared Klein et al. (eds), *Handbook of Historical and Comparative Indo-European Linguistics*, vol. 1. Berlin: De Gruyter Mouton, 566–99.
- Schmidt, Klaus T. 1985. "Zu einigen der ältesten iranischen Lehnwörter im Tocharischen", in Ursula Pieper and Gerhard Stickel (eds), *Studia Linguistica, Diachronica et Synchronica. Werner Winter sexagenario anno MCMLXXIII gratis animis ab eius collegis, amicis discipulisque oblata*. Berlin: de Gruyter, 757–67.
- Schwartz, Martin. 1974. "Irano-Tocharica", in Philippe Gignoux and Ahmad Taffazzoli (eds), *Mémorial Jean de Menasce*. Louvain: Imprimerie orientaliste, 399–411.
- Sims-Williams, Nicholas. 1983. "Chotano-Sogdica", *Bulletin of the School of Oriental and African Studies* 46, 40–51.
- Sims-Williams, Nicholas. 1996. "The Sogdian manuscripts in Brāhmī script as evidence for Sogdian phonology", in Ronald E. Emmerick et al. (eds), *Turfan, Khotan and Dunhuang: Vorträge der Tagung "Annemarie von Gabain und die Turfanforschung"*. Berlin: Akademie Verlag, 307–15.
- Sims-Williams, Nicholas. 1998. "The Iranian languages", in Anna Giacalone Ramat and Paolo Ramat (eds), *The Indo-European Languages*. London: Routledge, 125–53.
- Sims-Williams, Nicholas. 2016. *A Dictionary: Christian Sogdian, Syriac, and English*. Wiesbaden: Reichert.
- Skjærvø, Prods Oktor. 1994. "Of lice and men and the Manichean anthropology", in Renate Söhnen-Thieme and Oskar von Hinüber (eds), *Festschrift Georg Budruss zur Vollendung des 65. Lebensjahres und zu seiner Emeritierung dargebracht von Schülern, Freunden und Kollegen*. Reinbek: Dr. Inge Wezler, 269–86.
- Skjærvø, Prods Oktor. 2004. *This Most Excellent Shine of Gold, King of Kings of Sutras: the Khotanese Suvarṇabhāsottamasūtra*. Vol. I: *The Khotanese Text with English Translation and the Complete Sanskrit Text*. Vol. II: *Manuscripts, Commentary, Glossary, Indexes*. Cambridge, MA: Harvard University.
- Steblin-Kamensky, Ivan M. 1999. *Ėtimologičeskij slovar' vaxanskogo jazyka*. St. Petersburg: St. Petersburg State University.
- Stronach, David. 1985. "Pasargadae", in *The Cambridge History of Iran. Volume 2: The Median and Achaemenian Periods*. Cambridge: Cambridge University Press, 838–55.
- Tremblay, Xavier. 2005. "Irano-Tocharica et Tocharo-Iranica", *Bulletin of the School of Oriental and African Studies* 68/3, 421–49.
- Wagner, Donald B. 2008. *Ferrous Metallurgy*. (Civilisation in China Volume 5: Chemistry and Chemical Technology, Part 11.) Cambridge: Cambridge University Press.
- Winter, Werner. 1971. "Baktrische Lehnwörter im Tocharischen", in Robert Schmitt-Brandt (ed.), *Donum Indogermanicum. Festgabe für Anton Scherer zum 70. Geburtstag*. Heidelberg: Winter, 217–23.