

Revisiting a problematic Uralic and Indo-Iranian word family

Introduction

Much contradicting information can be found in research literature of Uralic etymology about the possible Iranian (or Indo-Iranian) origin of Hungarian *agyar* ‘tusk’ and its possible Permic and Ob-Ugric cognates. Although the idea that these Uralic words are borrowed from an Iranian form that yielded Ossetic *ænsur* / *æssyr* ‘tusk’ was already presented by Munkácsi (1901: 632–633), the Iranian loan etymology has received various judgements over the decades, with most recent etymological sources either rejecting the etymology or at least being cautious about it. In this short paper the problems of the etymology of *agyar* and the competing explanations are investigated; it is argued that the Uralic forms can, after all, be considered Indo-Iranian loanwords. Furthermore, the problems in the reconstruction of the Indo-Iranian donor form will be discussed and some additional ideas will be put forth.

Research problem and its history

The idea that Hungarian *agyar*, Komi Pr¹ *vožir*, Ud *vežir* ‘Eckzahn (Augenzahn), Hauer, Hauzahn’, Udmurt S *važer*, K *važer*, G *wažer* ‘Hauer, Hauzahn’ and Mansi N *ańsər* ‘medvefog’, P *añsər*, KU *añsər*, TJ *əńcər* ‘Hauzahn’ (UEW: 340) are borrowed from Indo-Iranian has a long history, as it was first suggested by Munkácsi (1901: 632–633; see also 1903: 380). The etymology received initial support in Sköld’s (1922: 15) treatment of “Ossetic” (Alanic) loanwords in Hungarian. (Sköld also commented on the Permic and Ob-Ugric words.) This etymology involves two problematic points that have been debated in subsequent research history. One question concerns the Iranian etymology: various arguments have been pre-

¹ A glossary of abbreviations can be found in the source (UEW: XLIII–XLVIII).

sented for and against it. Another problem is the relationship between the Uralic forms, which is not entirely regular.

The etymology has not received wide acceptance in etymological sources since the 1970s: in their treatments of Indo-Iranian loanwords, Korenchy (1972) and Joki (1973: 168) do not accept this Iranian etymology, and the Indo-Iranian etymology is not mentioned at all by UEW, Rédei (1986), or Holopainen (2019). Because of the alleged distribution of the Iranian word limited to Ossetic only, Korenchy (1972: 45) assumes that the similarity of the Uralic and Ossetic words can rather be explained by assuming a loanword from Finno-Ugric² to Ossetic, although she also mentions that it is not certain that the words are connected at all.

² The idea that there are Uralic loans in Ossetic has a long history, but many of the examples are problematic and debated. This matter cannot be discussed here in detail, but an illustrative example that includes similar problems as the etymology of *agyar* can be mentioned: Korenchy (1972) notes that Hungarian *mély* ‘deep’ is not a loan from Ossetic *mal* ‘deep water’, but here the direction of borrowing is vice versa. The idea that Hungarian *mély* is of Iranian origin has not received much support in later research. This is a different case from *agyar* in that the donor form would have to be a late Ossetic-like word, but the Hungarian word has regular cognates in the Ob-Ugric languages: Khanty V *měl*, DN *mět*, O *mǎl* ‘deep’; Mansi TJ *mäl*, KU *mil*, P (irregular) *ńil*, So *mil* ‘deep’ (UEW: 870). The idea of an Iranian loan etymology in Ugric has been supported by Rasmussen (1989: 234–239) in his discussion of the Indo-European word **móri-* ‘sea’. If Rasmussen is right and Ossetic *mal* reflects Iranian **mārya-*, a derivative of Indo-European **móri-* ‘sea’ (see Abaev II: 68–69 s.v. *mal*), the assumed borrowing from (Finno-)Ugric to Ossetic cannot be correct. However, this explanation is far from certain, as there are no reflexes of this PIE stem in any other Iranian language, and all the certain cognates of the word for ‘sea’ are found in the so-called Northwest Indo-European languages (Italo-Celtic, Germanic, and Balto-Slavic). Presuming borrowing in the other direction comes with problems of its own: the idea that the word is a loanword from a late Alanic form with **l* from Proto-Iranian **ry-* to Proto-Ugric looks chronologically suspicious; forms with **l*, such as the ethnonym Ἄλαιοι, do appear rather early in Greek and Latin sources, in the 1st century CE (Alemany 2000: 1; Bielmeier 1989: 24), but the established loanwords shared by the Ugric languages reflect the Proto-Iranian or “Old Iranian” stage, which has to be significantly earlier. It is possible that the word is a separate loan from Alanic into the predecessors of Khanty, Mansi and Hungarian, but the Ugric cognates are regular, and it would be possible to reconstruct a Ugric/Uralic proto-form for these words. This etymology requires further research.

Also, the etymological dictionaries of Hungarian and Komi mostly have a negative view of the loan etymology, if it is mentioned at all. The first serious (but unfinished) etymological dictionary of Hungarian by Gombocz & Melich (1914–1944, 32–33 s.v. *agyar*) supports the Finno-Ugric etymology of the Hungarian word but is more cautious regarding the Iranian etymology. Gombocz & Melich mention the Ossetic word and Munkácsi's etymology, but state that the Iranian word cannot be considered the immediate source of the Finno-Ugric words (“... amelyet azonban nem tarthatunk a fgr. szavak közvetlen forrásának”). Of the later etymological dictionaries, MSzFE (73–74 s.v. *agyar*), EWUng (13 s.v. *agyar*) and Zaicz (2006 s.v. *agyar*) make no mention of the Iranian etymology at all, whereas TESz (107 s.v. *agyar*) only states that the Iranian origin of Hungarian *agyar* is unacceptable, without discussing the etymology in any detail or even mentioning the Iranian donor form. KESK (60 s.v. *водзуп*) does not mention the loan etymology. Recently, also WOT (1339) has rejected the Iranian etymology, listing it among unacceptable Iranian loanwords in Hungarian, but no details are given about the reasons. WOT (1302) supports the Uralic etymology found in earlier etymological dictionaries.

On the Iranist side, the loan etymology has received more support, although it has been discussed only in few studies. The only Ossetic etymological dictionary, compiled by Abaev (I: 189 s.v. *æssyr* | *ænsur*(æ)), supports Munkácsi's etymology (Iranian loan into Hungarian). Recently, the etymology has been briefly discussed by Cheung (2002: 164) in his treatment of Ossetic vowel history. Cheung finds the Iranian origin of the Uralic words probable but does not analyse the etymology in more detail. The etymology is also mentioned in the list of Indo-Iranian loans into Uralic by Kümmel (2020: 251), who notes that the Hungarian and Permic words might be parallel loans. The loan etymology has also been supported briefly in a recent presentation on Uralic derivational morphology (Holopainen, Kuokkala, Metsäranta & Rauhala 2020).

A closer look at the Uralic data

Before discussing the Iranian loan etymology in detail, it should be determined whether the Hungarian, Permic, and Mansi words are cognates

or whether they could be possible parallel loanwords from Iranian. The vocalism of the Uralic forms is clearly irregular, and Aikio (UED manuscript) has noted that the comparison includes severe irregularities and that it is not at all certain the words go back to a Proto-Uralic form. However, most traditional sources of Uralic etymology list the word among Proto-Uralic/Proto-Finno-Ugric lexemes.

The Uralic/Finno-Ugric etymology is mentioned and supported in UEW (340), in all the major sources of Hungarian etymology (MSzFE 73–74 s.v. *agyar*; TESz; EWUng; the unfinished dictionary of Gombocz & Melich 1914–1944: 32–33 s.v. *agyar* also support the etymology), and in KESK as well. Zaicz (2006 s.v. *agyar*) mentions only the possible Ob-Ugric cognates of Hungarian *agyar*. In all of these works, the Proto-Finno-Ugric form is reconstructed as **ońća-rV* or **ońćV-rV*. UEW and the Hungarian dictionaries assume that the word is originally a derivative, with a suffix **-rV* (this idea will be discussed below). Sammallahti (1988: 541) includes the etymology in his word list among the Proto-Finno-Ugric words but reconstructs the word with **ę* (**j*) as **ęńsara* ‘canine tooth’. However, neither the reconstruction with **o* or **ę* fits all the Uralic forms.

Hungarian *agyar* can regularly reflect **ońćara* or **ańćara* (or **ońćiri*, **ańćiri*); also **ęńćara* or **ęńćiri* would be possible. There is thus no problem to derive the Hungarian word from the Uralic reconstructions of UEW or Sammallahti.

However, the rest of the words are much more problematic. The Komi and Udmurt words are not regular cognates even within Permic, and the labial consonant makes the assessment of the etymology difficult. KESK, MSzFE and TESz mention that Permic *v-* is secondary, referring to Uotila (1933: 67). In this treatment of the reconstruction of the Permic proto-language, Csúcs (2005: 112) mentions the etymology in the list of Proto-Permic words with **w-* in anlaut. Csúcs lists several words with secondary Proto-Permic **w-*, but in his (and UEW’s) reconstruction they reflect various Proto-Uralic initial vowels. It seems that in words with a labial vowel in anlaut, **w-* developed in Proto-Permic, but details are unclear, and it is not clear at all how the regular development should be described. In any case the words with anlaut **w-* show various first-syllable vowels, and the assessment of these etymologies is very difficult.

Udmurt *važer* alone would not be problematic, as *va-* as the reflex of **o* is found in several inherited words such as Udmurt *war* ‘slave’ < PU **orja* (UEW: 721)³. However, the Komi unrounded central back vowel *e* does not occur as the cognate of Udmurt *a* in any of the other words listed by Csúcs (2005). On the other hand, Metsäranta (personal communication) remarks that the *o* in Komi *vožir* is less problematic as cognate to Udmurt *a*, as other examples of this correspondence can be found (such as Komi *košmj-* ~ Udmurt *kwašmi-* ‘to dry’). Due to the irregularity within Komi dialects, the correspondences within Permic remain problematic.

The Mansi word (Mansi N *ańśar* ‘medvefog’, P *äńśar* ‘Hauzahn’, KU *äńśar* ‘Hauzahn’, Tj *ǵńćar* ‘Hauzahn’) is likewise problematic, as it cannot be a regular cognate of either the Permic or the Hungarian word. It is unclear what Pre-Mansi vowel the word could reflect: most forms point to Proto-Mansi **ä*, but *ǵ* in South Mansi (Tj) does not support this. The phonology of the Mansi word requires further research, but it is quite clear that the word must be a parallel loan.

Various Khanty forms have been connected to the other Uralic words and it seems that the forms listed by UEW (340), MszFE (73–74 s.v. *agyar*), TESz (s.v. *agyar*), Zaicz (2006 s.v. *agyar*), and Sammallahti (1988: 541) are certainly not cognates of the Mansi words: it is also clear that all the Khanty words listed by UEW are not cognates with each other. MszFE refers to Gulya (1964: 114), who notes that Khanty *ańžar* is found only in the northern dialects and only in the specific meaning ‘Eck-, Reisszahn des Bären’, which points to the conclusion that it is a loan from Mansi. This is a plausible explanation. UEW and DEWOS (133–134) likewise consider the North Khanty words loans from Mansi (UEW only states that the Kazym form is a loan, but obviously the same is true of the Obdorsk form). None of the South Khanty forms mentioned by UEW can reflect an earlier **ńć* cluster. Sammallahti (1988: 541) mentions Proto-Khanty **ańćar* among the list of cognates: although such a form could be postulated on the basis of the North Khanty forms, there is little reason to assume that the word goes back to Proto-Khanty, as it looks like a relatively late Mansi loan.

³ The other alleged Permic cognate included in UEW, Komi *ver* ‘husband’, does not reflect PU **orja* but, contrary to the claim of UEW, is a loan from Iranian **wiHra-* ‘man’ (see Holopainen 2019: 164–165).

The South Khanty forms DT *ǎŋtəl* ‘Reißzahn des Bären’, Kr *ǎŋtəl* ‘Zahn der Bären’, Kam *ǎŋtəl* ‘Backenzahn’, O *ǎńśár* ‘hinten am Gürtel getragener Bärenzahn’, *ǎńśar* ‘Eck-, Reißzahn des Bären’ can hardly be explained as Mansi loans, however. These forms cannot be regular cognates to the other Uralic forms either, and the Khanty *ŋ*-consonantism regularly reflects a Proto-Uralic cluster **ńć*; furthermore, **l* from **r* would be irregular. UEW states that “Die Konsonanten *ŋ* (< **ń*) und *l* (< **r*) in den Formen ostj. DT Kr. Kam. können das Ergebnis einer tabuistischen Deformation sein. Auch *á*, *ǎ* sind unregelmäßig”. The explanation involving a “taboo deformation” is ad hoc if no parallels for such a development are given. The origin of Khanty *ǎŋtəl* remains a mystery, but as it cannot be regularly connected to any of the other Uralic words (also not to the Indo-Iranian original), the word could be a completely unrelated lexeme that only shows accidental similarity.

In UEW, MszFE, and EWUNG, it is assumed that the Uralic word **ońćV-rV* contains a derivational suffix **-rV*, which according to UEW appears in a number of Proto-Uralic, Proto-Finno-Ugric, Proto-Finno-Permian, and Proto-Ugric etymologies (it is argued that a reflex of the same suffix also appears in a number of Proto-Ob-Ugric etymologies, Riese 2001: 52). Such a Proto-Uralic derivational suffix is also reconstructed in other presentations of Uralic historical morphology (see e.g., Rédei 1975: 95–100; Majtyinszkaja 1974: 352; Riese 2001: 52; Maticsák 2015: 154). Although denominal derivational suffixes with *r* do exist in many Uralic languages (for example, Khanty *-r*, Mordvin *-r(V)-*; see Sauer 1967: 159–160, Maticsák 2015: 154), many details in the history of these suffixes are uncertain, and it is far from certain that they all somehow reflect the same Proto-Uralic suffix. In some languages, like Mansi, only petrified inherited forms with a possible **-rV* suffix are found, and no such productive suffix exists (Riese 2001: 57, 107).⁴

As has been noted by Holopainen, Kuokkala, Metsäranta & Rauhala (2020), there is not actually much evidence for the derivational

⁴ It has been argued that a reflex of the Uralic **-rV*- suffix appears in a number of Proto-Ob-Ugric etymologies (Riese 2001: 52), many of which are reflected in both Khanty and Mansi. It is interesting to note that this Ob-Ugric **-r* is not a productive suffix in Mansi anymore (Riese 2001: 107), but according to Sauer (1967: 159–160) it is productive and common in Khanty.

origin of the Proto-Uralic words allegedly involving this derivational suffix. The etymologies listed as derivatives with an **rV*-suffix include opaque words like **šijiri* ‘mouse’ and also the well-known Indo-Iranian loan **oćtVrV* ‘whip’, which certainly is not a Uralic derivative (see Korenchy 1972: 56–57; Zhivlov 2013: 220; Holopainen 2019: 156–158). Many other words that are reconstructed with the same suffix in UEW are also uncertain or unclear (see the list of these etymologies in Riese 2001: 38–39), such as P(F)U **ćukkv-rV* ‘Haufen, Schar’ that is an irregular etymology limited to Permic and Hungarian, and P(F)U **ćiyVrV* ‘grave’ that has been refuted by Bereczki (2013: 256). There are good etymologies involving a suffix **-rA*, such as **kojra* ‘male animal’ (from **koji* ‘male’) or **kűńärä* ‘elbow’ (from **kűńä* id.), but the alleged derivatives with **-ri* or **-rV* are much more problematic.

A thorough analysis of all the relevant etymologies reconstructed with **-r*-suffixes into Proto-Uralic would be an important task, which cannot be done here, but it is clear that there is no reason to consider Uralic words grouped under **ońćV-rV* as derivatives. Cheung (2002), too, has criticized the idea that **ońćV-rV* ‘tusk’ is derived from **ońća* ‘forehead’, although he does not explain his view in detail.

A possible connection to Uralic/Finno-Permic **ońća* ‘front part of something’ mentioned in UEW is very unlikely, especially as the reconstruction of such an item is debated (Fi *otsa* ‘forehead’ is probably a loan from Germanic according to Koivulehto 1979: 290). Finnish *otsa* would also rather reflect Early Proto-Finnic **ońćća*, as there is another Uralic item **ońća* ‘part’ (UEW: 339–340) that gives Finnic *osa* ‘id.’ regularly.

To sum up the arguments regarding the Uralic forms, it is obvious that only Hungarian *agyar* can reflect the reconstructions **ońća-rV*, **ońćV-rV* or **ęńsara* given in earlier sources. The Permic and Mansi words cannot be regularly derived from the same form as the Hungarian word. It is thus possible that if the Indo-Iranian etymology is correct, the words are parallel loanwords. It is also clear that there is no need to consider the Uralic words derivatives, as no underived base is found anywhere and the existence of the **-rV* nominal suffix is also uncertain.

A closer look at the Iranian data

In addition to the troubling relationship between the Uralic cognates, also the Iranian etymology of Ossetic *æssyr* ~ *ænsur(æ)* has been disputed. Munkácsi initially mentioned only the Ossetic forms in connection with the Iranian loan etymology, but Sköld (1925: 15) made the important observation that the Ossetic word has a cognate in (Young) Avestan *asūra-* ‘tusk’. The distribution of the word within Iranian is rather crucial, as a lexeme appearing only in Ossetic but missing from all the other Iranian languages with earlier attestations would be suspicious. The view that Ossetic *æssyr* ~ *ænsur(æ)* is related to Avestan *asūra-* was followed also by Abaev (I: 189 s.v. *æssyr* | *ænsur(æ)*); in the 1995 index to his dictionary Abaev also added the Sogdian cognate *’ns’wr*, *’swr* ‘tusk’ as well as Tocharian A *ānkar*, B *ānkār* ‘tusk’. Since then, cognates in various Iranian languages have been suggested, and the Avestan cognate has also been doubted. It is interesting that despite Sköld’s and Abaev’s views, Korenchy still argued that the Hungarian word cannot be an Iranian loan as the assumed source is found only in Ossetic.

As a side note, it can be mentioned here that Sköld also argued that the relationship Ossetic *ns* ~ *ss* = Av *s* is regular, also occurring in the word *fissin* [= *fyssyn*] ~ *finsun* = *paēs-* ‘schreiben’. However, this is not a real parallel, as *finsun* is a verb and the *n* here reflects the nasal infix that appeared in the present forms of the Proto-Indo-European verbs, whereas the Avestan root is only attested in the participle forms (pres.) *aṅku.paēsəmna-* (Yt 17.10) and perf. pass. YAv. *frapixšta-* (Abaev I 501–502; Cheung 2007: 291–292; LIV: 465), where no reflex of the nasal would be possible.

The most comprehensive and up-to-date discussion of the Iranian cognates is done by Cheung (2002: 164). He assumes that Ossetic *æssyr* / *ænsur(æ)* and Sogdian *’ns’wr*, *’swr* ‘tusk’ are exact cognates and reconstructs a Proto-Iranian form **ansura-/*ansūra-*. The Avestan cognate *asūra-* is considered problematic by Cheung, as it lacks a reflex of the nasal **n*. However, as the Avestan word is a *hapax* (attested only in the compound *tiži.asūra-*), one could assume that the Avestan word does not have much value here. On the other hand, due to the irregularity on the Iranian side, this word might be considered a substrate word. Also, some

other Indo-Iranian names related to animals (such as **kapauta-* ‘pigeon’ > Old Indo-Aryan *kapota-* id., **pusća-* ‘tail’ > Old Indo-Aryan *púccha-*, **warāfha-* ‘boar’ > Old Indo-Aryan *varāha-* id.) originate in a substrate source (Lubotsky 2001: 311–312).

Cheung also mentions the idea that *asūra-* might reflect the zero-grade form with **ŋ*, but he deems this unlikely, as a rare word like ‘tusk’ would not likely have ablaut variants. Cheung also assumes that analogy from the *a-* in the word *aršnō* ‘male’, preceding *tiži.asūra-* in the Avestan text, could have played a role.

The argument about the lack of ablaut is not very strong, however, especially if the word is of Indo-European origin, as has been assumed. Cheung refers to Hackstein’s (1998: 227) etymology of Tocharian A *ānkar*, B *ānkār* ‘tusk’, which he derives from PIE proterokinetic **h₂énk-r-*, **h₂nk-én-s*; Latin *uncus* ‘curved’ stems from the same root according to Hackstein. He does not mention any Iranian words in this connection but states that *uncus* is often an epithet for teeth in Latin, which reinforces the etymological connection between the Latin word and the Tocharian and Indo-Iranian forms. Cheung considers it possible but uncertain that the Iranian words for ‘tusk’ are related to the aforementioned Tocharian and Latin words, noting that the Iranian etymology of the Ossetic word is “somewhat puzzling”. The Proto-Iranian stem does not correspond to the Tocharian or Latin words and might have been modelled after **bāzura-* ‘wing’.

If Ossetic *æssyr* indeed reflects Indo-European **h₂énk-r-*, **h₂nk-én-s*, there certainly has been ablaut in the Indo-Iranian word originally, and it cannot be ruled out completely that the Avestan word would not have generalized the zero-grade here. However, this kind of argumentation is speculative as we have too little Indo-Iranian data here, and admittedly it would be a bit strange if the zero-grade root would have been generalized in the Avestan derivative but not in the predecessors of the Sogdian and Ossetic words. On the other hand, if Cheung’s idea of analogy from **bāzura-* [= Proto-Iranian **bādzura-*, if we employ the Proto-Iranian phonology reconstructed by Mayrhofer (1989)] ‘wing’ is correct, we could assume that the Avestan lack of nasal is also due to this remodeling. It is not immediately clear why ‘tusk’ should be modelled after ‘wing’, however.

Also, another possible explanation for the aberrant Avestan stem could be mentioned: analogy to the Indo-Iranian word **ankura-* ‘hook’ could have played a role here. If ‘tusk’ was originally ‘curved’, the word for ‘hook’ would be semantically and formally rather close. This might be a better option than remodelling after the word for ‘wing’ and might explain the aberrant stem but does not solve the problem of the absence of nasal in Avestan, however.

Further connections of the Iranian word are not central to the Uralic loan etymology, as several of the well-established loanwords reflect roots that are attested in Indo-Iranian only (Holopainen 2019), but it can still be noted here that the Indo-European etymology mentioned cautiously by Cheung is not accepted in all sources.

De Vaan (2008: 640) makes no mention of the Tocharian words but gives the traditional etymology according to which *uncus* derives from PIE **h₂onko-*, which yields also Indo-Iranian **Hanka-* ‘hook’ (a possible source of Finnish *onki* ‘fishing rod’, North Saami *vuogga* id. and Mari *oŋgo* ‘hoop’, see Joki 1973: 295–296; Holopainen 2019: 161–162). The same root also gives Latin *ancus*; de Vaan refers to LIV (268), where no mention of the Tocharian or Iranian nouns is made. The root is reconstructed with plain **k* in LIV. (De Vaan gives palatovelar **k̑* but this must be a mistake, as the Indic and Iranian forms he lists obviously cannot reflect that.)

Adams (2013: 42) notes that the Tocharian A and B forms are somehow related, but the second-syllable vowels do not match. He considers the possible Indo-European connections uncertain, referring to Pokorny who connected the Tocharian words with Greek *ἄγκυρα* ‘anchor; pruning hook,’ Sogdian *’nswr* ‘fang,’ Ossetic *ænsuræ* ‘fang,’ deriving them from **h₂enk-*. Adams considers the Tocharian word for ‘elephant’ possibly related, but that does not help much in the assessment of the Iranian etymology. However, in regard of the Greek word, Beekes (2010: 12–13) does not mention the Tocharian or Indo-Iranian words. He derives the Greek word from the IE root **h₂enk-*, reconstructed with a plain velar.

The origin of Avestan *asūra-* and related forms has also been recently discussed by Kümmel (2021) in an unpublished presentation. He assumes that Iranian **(H)antsura-* ~ **(H)atsura-* might be derived from

Proto-Indo-Iranian **Hanć-* ‘spitz’, a nasalized variant of Proto-Indo-European **h₂aḱ-* ‘scharf’.

The riddle of the Iranian forms cannot be solved completely, but in any case, the fact that the Sogdian and Ossetic forms can reflect a Proto-Iranian word whereas the Uralic cognates show irregularities pointing to parallel borrowing means that it is more likely that the Uralic forms are loans from Indo-Iranian.

Cheung (2002: 164 footnote 20) also notes that Khotanese *haska-* ‘tusk’, New Persian *ašk, yašk* id. and Buddhist Sogdian *’sk* ‘thorn’, reflecting Proto-Iranian **aska-* (**atska-*) might share the same root as the Ossetic and Sogdian words, although he considers this uncertain. Bailey (1979: 11 s.v. *aše*) lists Khotanese Saka *aše* ‘plant name’ as a cognate of Avestan *asūra-* ‘point’, Ossetic *æssyr* and Sogdian *’nswr*, deriving them from the Iranian root **as-* (= **Hats-*) ‘to sharpen’. The exact proto-form of Saka *aše* is reconstructed as **as-ya-* ‘sharp, pointed’ by Bailey. It is not clear what kind of formations the other Iranian forms would reflect in Bailey’s view, and the nasal of Ossetic and Sogdian cannot be derived from this root without problems. Interestingly, also Khotanese *haskä* ‘(elephant) tusk’ (< **aska-*) is derived from the same root **as-* ‘be sharp’ by Bailey (1979: 474 s.v. *haskä*). This is semantically closer to the Ossetic word, but it looks unlikely that the words could be cognates.

Revisiting the loan etymology

It seems obvious that the Uralic words are not regular cognates with each other, whereas based on Sogdian and Ossetic, a Proto-Iranian or Proto-East-Iranian form can be regularly reconstructed. Cheung’s defence of the loan etymology is thus convincing, although it must be tested with a detailed investigation of the sound-substitutions involved.

All of the substitutions are well-established: both the **o* or **a* vocalism required by Hungarian and the **ä* to which Mansi points are frequent substitutions of Indo-Iranian **a*. The medial syllables have generally little value in the discussion of Indo-Iranian loans, but at least the Hungarian and the Mansi words can be old *A*-stems as most of the early Indo-Iranian loanwords are. The consonant cluster **ńć* from Indo-Iranian **ńć* has a parallel in the loan **ońća* ‘part’ ← Indo-Iranian **(H)anća-*.

The loan into Hungarian must be early in any case, as Hungarian *gy* points to earlier **ńć*, and this can result only from Indo-Iranian **nć*, not a later Iranian cluster **nts* or **ns*. Also, the palatal affricate or sibilant in the Permic forms points rather to an early Indo-Iranian donor, although this is not as clear: a palatal affricate appears also in some loans in Permic that otherwise seem to be quite late (for example Proto-Permic **beriz̥* ‘linden’ ← Iranian, cf. Ossetic *bærz/bærzæ*), and thus they cannot be used as definitive arguments, as noted by Palmér et al. (2021: 14–15).

Conclusions

Despite the reluctance to accept the etymology, the idea that the Uralic words (with the exception of Khanty) originate in Indo-Iranian is unproblematic: the meanings are similar and there are no phonological problems. The Uralic words are not regular cognates but look rather like parallel borrowings. The most logical explanation is to assume that they have been borrowed from Indo-Iranian. The **ć* indicates an early borrowing. It is also interesting to note that many other terms relating to animals have been borrowed from Indo-Iranian into Uralic languages, such as West Uralic **wočara* ‘boar’ ← Proto-Indo-Iranian **warājha-* ‘id.’, and some terms for body parts might have been borrowed in a similar context (referring to animals), such as **wärkkä* ‘kidney’ ← Proto-Indo-Iranian **wṛtka-* id. Like the word for ‘tusk’, many of these Indo-Iranian words lack a proper Indo-European etymology, as noted above (see Lubotsky 2001 for more details). It is also clear that fangs and tusks of animals had a notable significance in the cultures of some Uralic peoples (for example, the Mansi carried a bear’s tooth on the back of their belt⁵), which might explain why a name for ‘tusk’ has been borrowed.

The origin of the Iranian word remains uncertain, but it is clear that a Proto-Iranian form *(*H*)*antsura-* can be reconstructed, and technically nothing prevents us from reconstructing a Proto-Indo-Iranian word *(*H*)*anćura-* that would fit well as the origin of the Uralic words.

⁵ I am grateful to an anonymous referee for pointing this out to me.

Acknowledgments

I am indebted to two anonymous reviewers and Niklas Metsäranta for useful comments and remarks. The remaining errors are my own responsibility. I also would like to thank Martin Kümmel for giving me access to his presentation.

References

- Abaev = В.И. Абаев. 1958–1989. *Историко-этимологический словарь осетинского языка*. I–V. Москва & Ленинград: Академия Наук СССР.
- Adams, D. Q. 2013. *A Dictionary of Tocharian B*. Leiden Studies in Indo-European 10. Second edition, revised and greatly enlarged edition. Leiden: Brill.
- Alemany, Agustí. 2000. *Sources on the Alans: a critical compilation*. Handbook of Oriental Studies. Section 8 Uralic & Central Asia Studies 5. Leiden: Brill.
- Bailey, Harold. 1979. *Dictionary of Khotan Saka*. Cambridge: Cambridge University Press.
- Beekes, Robert Stephen Paul. 2010. *Etymological dictionary of Greek*. I–II. Leiden Indo-European Etymological Dictionary Series 10. Leiden / Boston: Brill.
- Bereczki, Gábor. 2013. *Etymologisches Wörterbuch des Tscheremissischen (Mari)*. *Der einheimische Wortschatz*. Nach dem Tode des Verfassers herausgegeben von Klára Agyagási und Eberhart Winkler. Veröffentlichungen des Societas Uralo-Altica 86. Wiesbaden: Harrassowitz.
- Bielmeier, Roland. 1989. Sarmatisch, Alanisch, Jassisch und Altossetisch. In Rüdiger Schmitt (eds.), *Compendium linguarum iranicarum*, 236–245. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Cheung, Johnny. 2002. *Studies in the historical development of the Ossetic vocalism*. Beiträge zur Iranistik 21. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Cheung, Johnny. 2007. *Etymological dictionary of the Iranian verb*. Leiden Indo-European Etymological Dictionary Series 2. Leiden / Boston: Brill.
- Csúcs, Sándor. 2005. *Die Rekonstruktion der permischen Grundsprache*. Budapest: Akadémiai Kiadó.

- DEWOS = Wolfgang Steinitz. 1966–1993. *Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*. Berlin: Akademie Verlag.
- EWUng = Loránd Benkő (ed.). 1992–1997. *Etymologisches Wörterbuch des Ungarischen Sprache*. Budapest: Akadémiai Kiadó.
- Gombocz, Zoltán & Melich János. 1914–1944. *Magyar etimológiai szótár*. 1–2. köt. I–XVII. füzet [A-Geburnus]. Budapest: Magyar Tudományos Akadémia.
- Gulya, János. 1964. Etimológiai jegyzetek. *Nyelvtudományi Közlemények* 66. 111–115.
- Hackstein, Olav. 1998. Tocharisch und Westindogermanisch, Strukturell uneinheitliche laryngalreflexe im Tocharischen (Uridg. *-Uh₁C- versus *-Uh_{2,3}(C-) und *-h₁RC- versus *-h_{2,3}RC-) In Wolfgang Meid (ed.), *Sprache und Kultur der Indogermanen: Akten der X. Fachtagung der Indogermanischen Gesellschaft Innsbruck, 22.-28. September 1996* (Innsbrucker Beiträge zur Sprachwissenschaft 93), 217–236. Innsbruck: Universität Innsbruck.
- Holopainen, Sampsa, Juha Kuokkala, Niklas Metsäranta & Ilona Rauhala. 2020. Challenges and opportunities in reconstructing Proto-Uralic nominal derivation (Handout of a presentation at the workshop *Word-formation in diachrony*, 7 February 2020, Vienna). www.academia.edu/42017894/.
- Holopainen, Sampsa. 2019. *Indo-Iranian borrowings in Uralic: Critical overview of sound substitutions and distribution criterion*. Helsinki: University of Helsinki PhD thesis.
- Joki, Aulis J. 1973. *Uralier und Indogermaner*. Suomalais-Ugrilaisen Seuran Toimituksia 151. Helsinki: Suomalais-Ugrilainen Seura.
- KESK = V.I. Lytkin & E.S. Gulyaev. 1999. *Краткий этимологический словарь коми языка*. Syktyvkar: Коми книжное издательство.
- Koivulehto, Jorma. 1979. Lainoja ja lainakerrostumia. *Virittäjä* 83. 267–301.
- Korenchy, Éva. 1972. *Iranische Lehnwörter in den obugrischen Sprachen*. Budapest: Akadémiai Kiadó.
- Kümmel, Martin. 2020. Substrata of Indo-Iranic and related questions. In Romain Garnier (ed.), *Loanwords and substrata. Proceedings of the Colloquium held in Limoges (5th–7th June, 2018)* (Innsbrucker Beiträge zur Sprachwissenschaft, Band 164), 237–277. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Kümmel, Martin. 2021. Neuere Forschungen zum (Indo-)Iranischen. Talk at the workshop *Current issues in historical linguistics*, 46. Österreichische Linguistik-Tagung, University of Vienna, 10.12.2021.

- LIV = Helmut Rix. 2001. *Lexikon der indogermanischen Verben. Die Wurzeln und ihre Primärstambildungen*. Bearbeitet von Martin Kümmel, Thomas Zehnder, Reiner Lipp und Brigitte Schirmer. Zweite, erweiterte und verbesserte Auflage, bearbeitet von Martin Kümmel und Helmut Rix. Wiesbaden: Dr. Ludwig Reichelt Verlag.
- Lubotsky, Alexander. 2001. The Indo-Iranian substratum. In Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: linguistic and archaeological considerations*, 301–317. Suomalais-Ugrilaisen Seuran Toimituksia 242. Helsinki: Suomalais-Ugrilainen Seura.
- Majtinskaja. 1974. = К.Е. Майтинская 1974. Сравнительная морфология финно-угорских языков. In В.И. Лыткин, К.Е. Майтинская & Карой Редей (eds.), *Основы финно-угорского языкознания*, 214–382. Москва: Наука.
- Maticsák, Sándor. 2015. *Az erza-mordvin névszóképzők nyelvtörténeti vizsgálata*. Akadémiai doktori értekezés. Debrecen: Debreceni Egyetem.
- Mayrhofer, Manfred. 1989. Vorgeschichte der iranischen Sprachen; Uriranisch. In Rüdiger Schmitt (ed.), *Compendium linguarum iranicarum*, 4–24. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Munkácsi, Bernát. 1901. *Árja és kaukázusi elemek a finn-magyar nyelvekben*. Budapest: Magyar Tudományos Akadémia.
- Palmér, Axel, Anthony Jacob, Rasmus Thorsø, Paulus van Sluis, Cid Swanenvleugel & Guus Kroonen. 2021. Proto-Indo-European ‘fox’ and the reconstruction of an athematic *k̄*-stem. *Indo-European Linguistics* 201. 21–30.
- Rasmussen, Jens Elmegård. 1989. *Studien zur Morphophonemik der indogermanischen Grundsprache*. Innsbrucker Beiträge zur Sprachwissenschaft 55. Innsbruck: Universität Innsbruck.
- Rédei, Károly. 1986. *Zu den indogermanisch-uralischen Sprachkontakten*. Veröffentlichungen der Kommission für Linguistik und Kommunikationsforschung 16. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Rédei, Károly. 1975. Wortbildung in der PU/PFU-Grundsprache. *Nyelvtudományi Közlemények* 77. 93–102.
- Riese, Timothy. 2001. *Historische Nominalderivation des Wogulischen*. Studia Uralica Band 10. Wiesbaden: Harrassowitz.
- Sammallahti, Pekka. 1988. Historical phonology of the Uralic languages with special reference to Samoyed, Ugric and Permic. In Denis Sinor (ed.), *The Uralic languages: description, history and foreign influences* (Handbuch der Orientalistik 8: Handbook of Uralic studies 1), 478–554. Leiden: Brill.

- Sauer, Gert. 1967. *Die Nominalbildung im Ostjakischen* (Finnisch-Ugrische Studien V). Berlin: Akademie-Verlag.
- Sköld, Hannes. 1925. *Die ossetischen Lehnwörter im Ungarischen*. Lund: Gleerup / Leipzig: Otto Harrassowitz.
- TESz = Benkő Loránd et al. (eds.). 1967–1976. *A magyar nyelv történeti-etimológiai szótára*. I–III. Budapest: Akadémiai Kiadó.
- UED manuscript = Ante Aikio. *Uralic Etymological Dictionary*. Unpublished manuscript.
- UEW = Károlyy Redei. 1986–1991. *Uralisches Etymologisches Wörterbuch*. Band I-II: Unter Mitarbeit von Marianne Bakró-Nagy, Sándor Csúcs, István Erdélyi †, László Honti, Éva Korenchy †, Éva K. Sal und Edit Vértes, Band III: Register Zusammengestellt von Attila Dobó und Éva Fancsaly. Wiesbaden: Harrassowitz.
- Uotila, T. E. 1933. *Zur Geschichte des Konsonantismus in den permischen Sprachen*. Suomalais-Ugrilainen Seuran Toimituksia 55. Helsinki: Suomalais-Ugrilainen Seura.
- De Vaan, Michiel. 2008. *Etymological dictionary of Latin and the other Italic languages* (Leiden Indo-European Etymological Dictionary Series 7). Leiden / Boston: Brill.
- WOT = Róna-Tas, András & Árpád Berta (with the assistance of László Károly). 2011. *West Old Turkic. Turkic loanwords in Hungarian*. Wiesbaden: Harrassowitz.
- Zaicz, Gábor. 2006. *Etimológiai szótár. Magyar szavak és toldalékok eredete*. Budapest: Tinta Könyvkiadó.
- Zhivlov, Mikhail. 2013. Андроновский арийский язык. In Ю.Б. Коряков & А.А. Кибрик (eds.), *Реликтовые индоевропейские языки Передней и Центральной Азии*, 217–220. Языки мира. Москва: Academia.