

# Speculative Iranian loanwords in Hungarian and Turkic

Sampsa Holopainen, Universität Wien & ÖAW

András Czentnár, ELTE

Deutscher Orientalistentag, 12.-17.9.2022

[sampsa.petteri.holopainen@univie.ac.at](mailto:sampsa.petteri.holopainen@univie.ac.at)

[zentnar.andras@btk.elte.hu](mailto:zentnar.andras@btk.elte.hu)

# Aims of the talk

- Criticize some untenable/speculative Iranian loan etymologies in Hungarian and Turkic suggested by Harmatta (1997), Szemerényi (1981)
  - These suggestions have received little criticism (some words criticized in WOT), and occasionally the ideas pop up in current studies on the history of Hungarian; a critical scrutiny is needed
- Present methodological problems in operating with “speculative” languages: problematic to assume unattested variants of some language solely on the basis of loanwords
  - Sound-substitutions
- Compare the problematic loanword layers with what we really know of Iranian influences in Hungarian/Uralic and Turkic > show how these loans can potentially be useful for Iranists if properly investigated

# Why?

- Interesting case studies on “speculative” suggestions
  - Harmatta (1997) operates with ad hoc reconstructions of East Iranian
  - Szemerényi assumes a Persian-type language (showing innovations such as  $*\theta < *ć$ ,  $*d < *j$ ) spoken in the steppe (= no evidence apart from loanwords)
- No comprehensive criticism of these ideas so far: Klima (1999) is cautious of Harmatta’s etymologies; some etymologies criticized in WOT, some alleged loans related to equestrian culture analysed by Holopainen (2022)
  - The ideas occasionally pop up in research on the history of Hungarian
- Relevance for cultural history: Ugric and Turkic equestrian nomadism and Iranian influences
- > important to show the scholars (of Iranian, Hungarian and Turkic) that not all the loan evidence is equally convincing

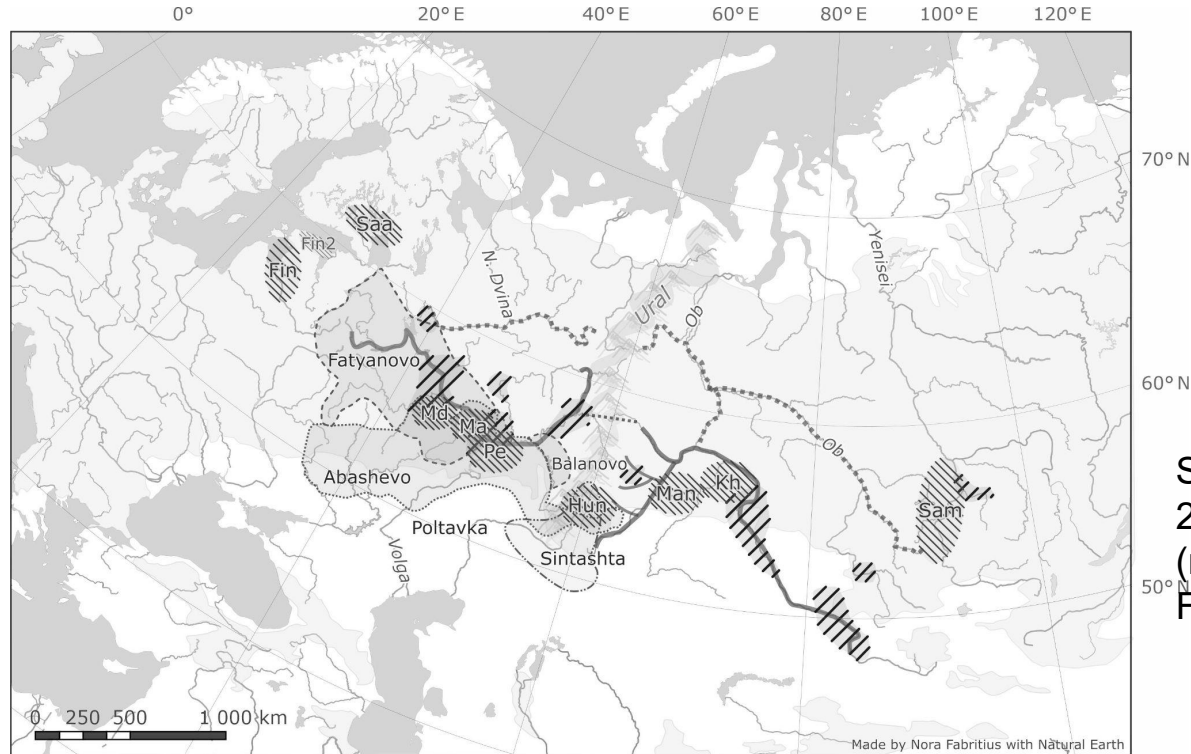
# Iranian loanwords in Hungarian: what do we know?

- Very early Indo-Iranian loanwords in Uralic (Korenchy 1972, Joki 1973, Holopainen 2019, Grünthal et al 2022), contacts spanning millenia
  - Early Indo-Iranian loans: PFU/PU (?) \**ćęta* ‘hundred’ ← PIIr \**ćatám*, Hu *ostor* (< \**oćtVrV*) ← PIIr \**aćtrā* ‘whip’
- Later contacts between the Iranian branch and the various Uralic languages: especially many loans in the Permic languages, Hungarian, Ob-Ugric
- Several layers of Iranian loanwords (Alanic loans)

# Iranian loanwords in Hungarian: what do we know?

- Early loans (in Hungarian and other branches of Uralic) point to an archaic donor (Proto-Iranian/Old Iranian) > no specific dialectal features (East Iranian or the like)
  - Hu *tehén* ← Plr \**daHinu-* ‘cow’; Hu *vásár* ‘market’ ← (M)Ir \**wačar*; Hu *arany* < PUg \**θəraña* ← OIr \**zaranya-* ‘gold’, Hu *hét* < PUg (?) \**säptä* ← Plr \**sapta-* ‘seven’
- Some disputed issues: “Andronovo Aryan” loans (Helimski 1997, Zhivlov 2013)
- Later layer(s) of Alanic loanwords in Hungarian (also in the Ob-Ugric and Permic languages)
  - Examples: Hu *kard* ← Alanic \**kard*, Hu *egész* ← Alanic \**agas*, Hu *ezüst*, ← Alanic \**æzvist*
  - Some issues with vocalism: Harmatta assumes “Chwarezmian” Alanic loans (*méreg* ‘poison’ ← *marg* etc.)

# Map: early spread of the Uralic languages



Source: Grünthal et al.  
2022, *Diachronica* 39  
(map by Nora  
Fabritius)

# Turkic–Iranian contacts: what do we know?

- Earliest contacts most probably date back to the distant past; Iranian nomads dominated the Eurasian steppe until the 4th century AD (some Eastern Iranian groups may have reached Mongolia and South Siberia – nomadism brought to the Turkic peoples by them?)
- Turkic tribes may have been present among the Scythians
- Turkic dominance from the middle of the 1st millennium in the whole steppe zone
- Long-lasting symbiosis between Iranian and Turkic peoples (even today), Iranian elements in the ethnogenesis of some groups
- More layers can be assumed, great number of Persian loans in most Turkic languages

# The material of this talk

- East Iranian loans suggested by Harmatta (1997)
- 24 loan etymologies in Hungarian
- The earliest loans are borrowed into an already disintegrated “Common Ugric” according to Harmatta (phonological irregularities)
- South-West Iranian loans suggested by Szemerényi: parallel loanwords into early phases of Hungarian and Turkic in the steppe, presumably from a Persian-type language
  - Single “Persian-type” etymologies for certain Ob-Ugric words suggested also by Honti (1997), criticized by Holopainen (2019: 267–268)



# Methodological problems

- Sound substitution: laissez-faire approach
- Chronological problems (stratigraphy of Uralic–Iranian contacts)
- Geographical context (problems with “homelands”)
  - Szemerényi’s “South-West Iranian” in the steppe
- Outdated reconstructions
- Small amount of data (relatively few, uncertain etymologies; few languages)
- Competing etymologies

# Methodological problems

- Examples of other postulated donor languages (languages/branches assumed on the basis of loanwords):
- “Andronovo Aryan” (Helimski 1997, Zhivlov 2013)
  - An unattested branch of Indo-Iranian, based on loanwords in Ob-Ugric and Permic
  - Much of the evidence can be explained otherwise (Holopainen 2019)
- West Old Turkic (WOT)
  - Archaic variety of Turkic reconstructed on the basis of loans in Hungarian (also in Ossetic, Slavic, and Mongolian); relatively well accepted but some criticism
- North Baltic
  - Donor language of the Baltic loanwords in Finnic (Kallio 2008; Junttila 2015; different opinion on the Baltic source language expressed by Jakob 2022)
- “Frühurarisches” loanwords (Katz 2003)
  - Peculiar Indo-Iranian reconstructions
- West Baltic loans in Chuvash, Mari (Agyagási 2019)
  - Problematic Baltic reconstructions (Holopainen & Metsäranta 2020)

# Historical and geographical context

- Harmatta 1994, 1997 claims that part of the Iranian loans in Hungarian, Ob-Ugric and Turkic is from the languages of the Xiongnu, which he assumes to be Iranian.
- However, this assumption is most likely incorrect, as most researchers do not agree with Harmatta (Vovin 2000), and recently it has been suggested that the Xiongnu were probably speakers of Turkic (Savelyev & Jeong 2020)
- Also the Yuezhi were Iranian speakers according to Harmatta (1997) > far from certain
- Harmatta also mentions a Saka-type donor language, but in practice the loans bear little resemblance to Khotanese Saka

# Harmatta's reconstructions and the taxonomy of Iranian

- East Iranian as a genealogical node rejected in modern research (Korn 2019: 248-251)
- Areal grouping rather than a real branch
- Does not mean that there could be no loans from eastern varieties of Iranian
  - > most loans in Hungarian and the other Uralic languages almost certainly “eastern” and not “western” Iranian
- Examples of East Iranian sound changes according to Harmatta:
- $*w > *γ$
- $*r > *l$  in certain contexts ( $*baraka-$  >  $*vlayə$  ‘horse’, cf. Saka  $bāraa-$  ‘vehicle; riding animal’)
- Syncope ( $*baraka-$  >  $*vlayə$  ‘horse’, cf. Saka  $bāraa-$ )
- Various kinds of assimilations ( $*paxštaka-$  >  $*pekkə$  ‘reins’)
- $*sr > s$  ( $*sravas$  >  $*savä$  ‘word, fame’, cf. Saka  $ššuva$ )
- $*a > *i?$  ( $*manyu-$  >  $*minyi-$  ‘heaven (??)’)
- Important: a full list of changes is not listed anywhere > the criteria of “East Iranian” loans ambiguous

# Analysis of the East Iranian etymologies in Hungarian

- Harmatta (1997) suggests 24 “East Iranian” loanwords, part of which have cognates in other Uralic/Finno-Ugric/Ugric languages (parallel loanwords?)
- 8 of these loans are borrowed into “Middle Proto-Hungarian (közép ősmagyar)” according to Harmatta
- In addition, Harmatta assumes late loans into Hungarian from “Chwarezmian Alanic”, a variety of Alanic he reconstructs (rather loosely connected to actual attested Chwarezmian); loans show vocalism divergent from most Alanic loans in Hungarian (not discussed here in more detail)
  - WOT accepts some of the “Chwarezmian” etymologies
  - NB the (Old Hungarian) ethnonym *káлиз* ‘Chwarezmian (?)’

# Analysis of the East Iranian etymologies in Hungarian

- Phonological problems on the Iranian side
  - Hu *kés* ‘knife’ < PFU/PU \**keči* ← East Ir \**keči*, cf. Wakhi *kəz*
  - Hu *köles* ← East Iranian \**goliči*
- Phonological problems on the Hungarian/Uralic side
  - Hu *ló* ‘horse’ ← East Ir \**loyə*, *nyerég* ← \**nəyer* ‘saddle (?)’, Hu *fék* ‘rein’ ← \**pekkə*, *imád* ← \**wi-mand-*, Hu *rés* ← East Ir, cf. Avestan *raēša-* ‘hole’
- Problems with non-existing forms
  - Hu *imád* ‘pray’ ← East Ir \**wi-mand-* ‘say liturgical text (?)’, Hu *kedv* ‘mood’ ← East Ir ? \**känti-*
- Semantic problems
  - Hu *nyerég* ‘saddle’ ← East Ir \**nəyer* ‘saddle (?)’
- Problems with relative chronology
  - Hu *ég* ‘sky’, Fi *sää* ‘weather’, etc. < PU/PFU \**säŋi* (\**säŋe*) ‘sky’ ← East Ir \**sangi-* ‘stone’
  - Hu *kés* ‘knife’ < PFU/PU \**keči* ← East Ir \**keči*, cf. Wakhi *kəz*

# Example etymologies

- Hu *ló* ‘horse’ + Kh, Ms < ? Ugric *\*loyɜ* ← East Iranian *\*loyə* < *\*vlayə* < *\*vəlayə* < Old Iranian (= Proto-Iranian) *\*bāraka-* ‘horse’
- Turkic *\*ulay* ← East Iranian *\*loyə*
  - The East Iranian reconstruction is ad hoc
  - The Ugric words probably loans from somewhere but clearly not from Iranian
  - The Turkic word can be a derivative of the verb *ula-* ‘to join sg to sg’ (Clauson 136)
- Hu *nyerэг* ‘saddle’ + Kh, Ms < ? Ugric *\*newrɜ* ~ *\*neyrɜ* ← East Iranian *\*neyər* < *\*newər* < Old Iranian (= Proto-Iranian) *\*niwarn-* ‘harness’
- Turkic ? *\*neyer* ‘caddle’ ← East Iranian *\*neyər*
  - Recently suggested also by Ponaryadov (2022)
  - Phonological problems (very vague resemblance)
  - The Khotanese word does not denote ‘saddle’
  - Zhivlov (2016) notes that the Hungarian and Ob-Ugric words have to be loans from somewhere
  - The reconstruction of the Turkic etymon is problematic. Clauson (Clauson 63) does not separate *\*ädär* and *\*ingir* (both ‘saddle’), but WOT (WOT 1210-1213) does and dismisses the possibility to reconstruct a Proto-Turkic form with an initial *\*ń* or even *\*y*. Very common in the Turkic family.

# Example etymologies

- Hu *imád* ‘pray’ < ? \**wimantV-* ← East Iranian \**vimand-* ‘recite a liturgical text’ < Proto-Iranian (?) \**wi-mand-* (cf. Sanskrit *mand-*, Avestan *mand-*)
  - Problems with the Iranian word (no such root); the formation \**wimand-* in any case unattested
  - Problems with the Hungarian development \**wi-* > *i-* (cf. *özvegy* ‘widow’)
- Hu *köles* ‘millet’, Mansi *kolás* id. ← East Iranian \**goliči* < (?) Proto-Iranian \**gauvriji* ‘millet’
  - Phonological problems
  - Accidental similarity
  - Hu and Ms cognates formally regular (a possible Proto-Ugric etymology; East Ir loan unlikely)
- Hu *kéj* ‘lust’ ← East Iranian \**gayi-* (\**gaHu-* ‘need, want, desire’), Ossetic *qæwyn*, *ǰæwun* ‘be in need of something’
  - Rather < PU/PFU \**keji* ‘rut’
  - Hu *kíván*, *kéván* (if not from Turkic) might be connected with the Iranian words listed here, but further research is needed



# Analysis of the East Iranian etymologies in Hungarian

- Some etymologies listed by Harmatta (based on earlier suggestions) are indeed correct but not necessarily from “East Iranian”
  - Hu *arany* ‘gold’ (< PUg \**θerańa*) ← Old Iranian/Proto-Iranian \**zaranya-* (< \**dzrHanya-*)
  - Hu *hány-* ‘throw’ ← (?) Proto-Indo-Iranian \**kanH-* ‘dig, throw’
  - **Hu *hús* ‘meat’** points to some kind of “eastern” source language (< **PIr \**gauštrā-*** [Morgenstierne et al. 2003] or \****gaušti-*** [Gerschevich 1975])
  - Hu *rés* ‘hole’ ← ? Ir \**raiša-*
  - Hu *réz* ‘copper’ ← (?) Ir \**rawḏa-* is generally accepted to be an Iranian loan (probably an old loan showing \**z* from \**VtV*)
  - Hu *vászon* ‘cotton’ is clearly ultimately of Indo-Iranian origin, but the loan has to be late (because of *sz*); Wanderwort?

# Analysis of the East Iranian etymologies in Turkic

- *\*at* ‘horse’ ← *\*atya-*
  - Word-initial *\*h-* in Proto-Turkic (cf. Doerfer–Tezcan 127); not an Iranian word (only in Old Indic).
- *\*sab* ‘word’ ← *\*savā* < *\*sravah-*
  - Ad hoc East Ir reconstruction. It seems that in OT monuments, *sav* appears to be a pair to *söz*, the former meaning ‘a (full-length) speech; a narrative or story, a message’, while the former is ‘a single word, or short utterance’. Common until the 14th c, rare after that (Clauson 782–783).
- *\*buzayu* ‘calf’ ← East Ir (?) *\*busay* < Plr *\*watsa-ka-* > Saka *basaka*
  - If the Turkic word is from Iranian, then it should have been a very early loan, since it can be found in every branch of the Turkic family. Furthermore, it shows rhotacism in Chuvash *păru* (and also in Mongolian). Replacement of initial *\*w* with *\*b* would be possible. Vocalism also remains unexplained.
- *\*bar-* ‘to go’ ← East Ir *\*bar-* ‘to ride’ > Saka *bar-* ‘to bear; (middle) to ride’
  - Improbable etymology because of the meaning, which is much wider in Turkic with the basic movement ‘to go’ in the focus. Attested and common in every branch of the Turkic family (Clauson 354).
- *\*tez-* ‘to run away, to fly’ ← East Ir *\*täz-* (cf. New Persian *täz-*, *taxstän* ‘to gallop, to turn around, to attack’)
  - The Turkic word has a deverbal derivation (*täz-* ‘to run away > *\*täzk* > *tärk* ‘quickly’), which shows rhotacism, and can be related to Mongolian *derkire-* ‘to run, to race’ (WOT 1110), hence, in case of borrowing, an explanation of chronology would be necessary.

# Problems with the “South-West Iranian” loans in Hungarian

- Hu *tanú* ‘witness’, Old Turkic *tanu-* ‘to recommend’ ← ? Ir \**dan-*
  - Turkic *tanuq* ‘witness’ (WOT 848–849) can be analysed morphologically: *tanu-* ‘to recommend, to suggest’ and deverbil suffix *-q*.
  - Also, the Ir verb would mean ‘to know’.
- Hu *tegnap* ‘yesterday’ ← Ir \**dēg* < \**dyiaka-*
  - Problems with chronology (voiced *g*, voiceless *t*)
- Hu *szó* ‘word’ ← Middle Iranian \**sav* < \**srawah-* ‘fame’
  - Competing etymologies (Ugric \**sawV*); semantic difference
- *egy-* in *egyház* ‘church’ (also *ünnep* ‘feast’) ← Ir \**yad-* ‘verehren, anbeten’ (Avestan *yazatē*)
  - Usually considered a Turkic loan: OT *edgü* ‘good’
  - Turkic has parallel compounds with the same meaning: *yïx-öv* ‘église, lit. holy house’, *yïh-kun* ‘dimanche, lit. holy day’. The Iranian form \**yad-* is back-vocalic, however, *nap* ‘day, Sun’ became front-vocalic due to vowel harmony in the fused compound *ünnep* ‘holiday’.

# Problems with the “South-West Iranian” loans in Hungarian

- Hu *ölyv* ‘a kind of falcon’ ← Persian-type Iranian *\*āluh-*
  - The Hu word has a hypothetical Turkic etymology, and it only has Mongolian counterparts (*eliye* < *\*elige*) that are assumed to be borrowed from Turkic; Turkic *\*ālig* is thus based on Hu and Mo; Hu cannot be directly from Mo; the Iranian loan criticized by WOT (667–668), as Róna-Tas considers the OIr source *\*āluf* > *\*aluh* > *\*ālih* phonetically unacceptable. Two competing hypothetical forms, *\*ǎ* > *ö* and *\*g* > *v* are more plausible than *\*ǎ* > *ö* and *\*h* > *v*.

# Conclusions

- The evidence for the loanword layers suggested by Harmatta and Szemerényi is very limited, and most etymologies have to be rejected
- While there certainly are loans from *eastern* Iranian varieties in Hungarian, the etymologies of Harmatta contribute little to our knowledge of such contacts
- Based on Turkic, there is no need to assume an East Iranian donor in the cases mentioned as all words can be analysed to be native or internal development in Turkic, however, further study of the early Iranian loans in Proto-Turkic would be useful

# References

Agyagási, Klára 2019: Chuvash historical phonetics. Wiesbaden: Harrassowitz.

AiWb = Bartholomae, Christian 1961: Altiranisches Wörterbuch. 2., unveränderte Auflage. Berlin: De Gruyter.

Bailey, Harold 1979: Dictionary of Khotan Saka. Cambridge: Cambridge University Press.

Cheung, Johnny 2002: Studies in the historical development of the Ossetic vocalism. Beiträge zur Iranistik, Band 21. Wiesbaden: Dr. Ludwig Reichert Verlag.

Clauson, Gerard 1972: An Etymological Dictionary of Pre-Thirteenth Century Turkish. Oxford: Clarendon Press.

Doerfer, Gerhard-Tezcan, Semih 1980: Wörterbuch des Chaladsch. Budapest: Akadémiai Kiadó.

EWAia = Mayrhofer, Manfred 1986–2001: Etymologisches Wörterbuch des Altindoarischen I–III. Heidelberg: Winter.

EWUng = Etymologisches Wörterbuch des Ungarischen Sprache. 1992–1997. Hrsg. von Loránd Benkő. Budapest: Akadémiai Kiadó.

Gerschevitch, Ilya 1976: More meat in Iranian. Morpurgo Davies, Anna, and Meid, Wolfgang (eds.), Studies in Greek, Italic, and Indo-European Linguistics offered to Leonard R. Palmer on the occasion of his seventieth birthday. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. 63–64

Harmatta, János 1997: Iráni nyelvek hatása az ősmagyar nyelvre. – László Kovács & László Veszprémi (eds.): Honfoglalás és nyelvészet. 71–83. Budapest: Balassi.

Helimski, Eugene 1997: The southern neighbours of Finno-Ugrians: Iranians or an extinct branch of Aryans („Andronovo Aryans“)? – Sirikka-Liisa Hahmo, Tette Hofstra, László Honti, Paul Van Linde & Osmo Nikkilä (hrsg.), *Finnisch-ugrische Sprachen in Kontakt*. Maastricht: Shaker. 117–25.

# References

Holopainen, Sampsa 2019: Indo-Iranian borrowings in Uralic. Critical overview of sound-substitutions and distribution criterion. Doctoral thesis. University of Helsinki.

Holopainen, Sampsa 2022: Lingvistisen paleontologian ongelmista. Mitä sanasto voi kertoa kulttuurista? Kaisla Kaheinen, Larisa Leisiö, Riku Erkkilä & Toivo Qiu (eds.), Hämeenmaalta Jamalille. Kirja Tapani Salmiselle. 7.4.2022. Helsinki: Helda Open Books. 101-114.

Jakob, Anthony 2022: Symmetry, sound change, and the Baltic loanwords in Finnic. Presentation at the workshop Areal effects in prehistoric contacts between Uralic and Indo-European, CIFU XIII, University of Vienna, 24.8.2022.

Junttila, Santeri 2015: Tiedon kumuloituminen ja trendit. Kantasuomen balttilaislainojen tutkimushistoria. Doctoral thesis. University of Helsinki.

Kallio, Petri 2008: On the "Early Baltic" Loanwords in Common Finnic. – Alexander Lubotsky, Jos Schaecken & Jeroen Wiedenhof (eds.), Evidence and counter-evidence: Essays in honour of Frederik Kortlandt 1, Balto-Slavic and Indo-European linguistics. Studies in Slavic and General Linguistics 32; Amsterdam – New York: Rodopi. 265–277.

Katz, Hartmut 2003: 2003: Studien zu den älteren indoiranischen Lehnwörtern in den uralischen Sprachen. Wiesbaden: Reichelt.

Klima, László 1999: Honfoglalás és nyelvészet. *Magyar nyelv* XCV.

Korn, Agnes 2019: Isoglosses and subdivisions of Iranian. Journal of Historical Linguistics 9:2. 239-281. <https://benjamins.com/catalog/jhl.17010.kor>

Morgenstierne, Georg 2003: A new etymological vocabulary of Pashto. Compiled and edited by J. Elfenbein, D. N. MacKenzie and Nicholas Sims-Williams. Wiesbaden: Reichert.

# References

Parpola, Asko 2020: Royal “Chariot” Burials of Sanauli near Delhi and archaeological correlates of prehistoric Indo-Iranian languages. *Studia Orientalia Electronica* 8. 75–198.

Sammallahti, Pekka 1988: Historical phonology of the Uralic languages with special reference to Samoyed, Ugric and Permic. – Denis Sinor (ed.), *The Uralic languages: description, history and foreign influences*. *Handbuch der Orientalistik 8: Handbook of Uralic studies* 1; Leiden: E. J. Brill. 478–554.

Savelyev, Alexander & Choongwon Jeong 2020: Early nomads of the Eastern Steppe and their tentative connections in the West. *Evolutionary Human Sciences* 2. 1-35.

Sköld, Hannes 1925: *Die ossetischen Lehnwörter in Ungarischen*. Lund: Gleerup. Leipzig: Otto Harrassowitz.

Szemerényi, Oswald 1981: Südwestiranische Lehnwörter in Ungarischen und Türkischen. *Études Finno-Ougriennes* XIII. 229-246.

UEW = Károly Redei 1988: *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 Zusammenge stellt von Attila Dobó und Éva Fancsaly. Wiesbaden: Harrassowitz.

Vovin, Alexander 2000: Did the Xiong-nu Speak a Yeniseian Language? *Central Asiatic Journal* 44. 87-104.

WOT = Róna-Tas, András & Árpád Berta (with the assistance of László Károly) 2011: *West Old Turkic. Turkic loanwords in Hungarian*. Wiesbaden: Harrassowitz.

Zhivlov, Mikhail 2013: Андроновский арийский язык. – Ю.Б. Коряков, А.А. Кибри (редакционная коллегия) *Реликтовые индоевропейские языки Передней и Центральной Азии. Языки мира*. Москва: Academia. 217–220.



Thank you for your attention!