

Sample of the *Database of Ugric etymology* (ÖAW APART-GSK project *Hungarian historical phonology reexamined*, University of Vienna, <https://ugric.univie.ac.at>)

Sampsa Holopainen

This is an illustration of etymology entries for the forthcoming *Database of Ugric etymology*. The database will be based on the Wiki-format currently used by the project *Digital etymological dictionary of the oldest vocabulary of Finnish* (led by Santeri Junttila, 2017–2022, University of Helsinki, Kone Foundation). The Wiki-format enables clear presentation and continuous updating easily and gives also other specialists of Uralic/Ugric/Hungarian etymology the possibility to comment the etymologies.

The database will include critical commentary of the Proto-Ugric etymologies suggested in the earlier etymological dictionaries, *Uralisches Etymologisches Wörterbuch* (UEW, Károly Rédei 1988) and *Magyar Szókészlet Finnugor Elemei* (MSzFE, György Lakó 1971–1978), and references to the other relevant sources where Ugric etymologies are discussed (notably WOT = András Róna-tas & Árpád Berta 2011: *West Old Turkic: Turkic loanwords in Hungarian*). The aim is to determine which of the cognates can really be reconstructed to a common proto-language of Hungarian, Khanty and Mansi, and to find alternative explanations to those etymologies that turn out to be irregular and cannot be reconstructed to Proto-Ugric.

In addition to the Proto-Ugric etymologies, Iranian loanwords of Hungarian will also be included in the database. The focus will be on the earliest loanwords borrowed in the separate existence of Hungarian (words borrowed before typical Proto-Hungarian sound-changes such as $*s > \emptyset$, $*VtV > VzV$) as well as the Iranian etymologies considered uncertain in earlier research. As it is disputed whether the Iranian loanwords common to all the Ugric languages were borrowed at Proto-Ugric times or separately by Khanty, Mansi and Hungarian after it had diverged (Holopainen 2019), it is important to address the phonology of the loanwords as well.

The scrutiny of these cognates serves the long-term aim of the project, a comprehensive presentation of Hungarian historical phonology.

--

Proto-Ugric ? $*taltV-$ ~ $*tultV-$ 'magic (?)' (UEW: $*tultz$ 'Zauberei, Zauberkraft'; WOT: $*toltă$)

Hungarian: *táltos* 'magician, shaman; magic horse; Zauberer, Schamane; Zauberpferd'

Old Hungarian: 1211 *Tholtus* (see EWUng :146, s.v. *táltos*)

Khanty: North (Kaz) *tɔʌt* 'Hilfe, Linderung (bei einer Krankheit, in der Armut)', *tɔʌta* 'ohne (große) Mühe, ohne (viel) Lärm (z.B. Beute bekommen); вдруг', (N) *tolt* 'Riese (eigtl. Zauberer)', *toltn* ~ *tolten* 'mit Zauberkraft'; East (Vj) *tolt* 'fever' < Proto-Khanty ? $*tolt$ (Honti: $*tolt/tolt$)

Mansi: North (N) *tūlt*: *tūltèn* 'leicht, einfach' < Proto-Mansi $*tūlt$

Proto-Ob-Ugric: (uncertain) $*t\tilde{V}lt-$ (Honti 1982: 188, no. 637)

Status: Improbable (phonological and semantic problems)

Discussion:

The Ugric etymology (originally stemming from Erdélyi 1960) has been debated in recent research literature: Honti (2017: 62–67) provides a good overview of research history and of arguments for and against the etymology, but he does not really address the problematic sound-correspondences between the Hungarian and Ob-Ugric forms.

The change **u > á* in Hungarian is irregular: if the etymology goes back to Proto-Ugric/Pre-Hungarian, the vowel *á* has to reflect earlier **o* or **a*. It is unclear how Old Hungarian *o* in the form *Tholtus* should be read, but the other Old Hungarian forms from the 15th century show *a* (see EWUng), so *o* in the 13th century form probably does not reflect [o]. An old (Uralic/Pre-Ugric) **o-a* or **a-a* stem would yield long **ū* in Mansi, so a reconstruction **tola-* or **tala-* could account for the Hungarian and Mansi forms, but the Khanty vocalism is anomalous (not reflecting **o* or **a* regularly). Honti (1982) reconstructs Proto-Ob-Ugric **tVlt-*, meaning that the quality of the Proto-Ob-Ugric vowel is uncertain; this makes the whole etymology dubious.

The semantic connection of the Hungarian and Ob-Ugric words is dubious. Also the semantic connection between the Ob-Ugric forms is uncertain. It has to be pointed out that the word is also attested in the northern dialects of Khanty and Mansi; Honti (2017) assumes that the inherited Ugric form has been retained in only these dialects, but Erdélyi (1960) assumes the Mansi word is a loan from Khanty. Due to the irregular vowel-correspondences, the borrowing from Khanty to Mansi cannot be ruled out.

Taken together, the problems of phonology and semantics point to the conclusion that the Ugric etymology should be rejected.

The Turkic etymology suggested by WOT seems the most probable option. The suggested West Old Turkic source form **taltučī* is ‘the one who exercises a loss of consciousness’ is derived from the Turkic root **tal-* ‘to lose strength, to lose consciousness, to faint’. The etymology includes no major problems, but the formation **taltutčī* has not been attested as such, which makes the loan explanation somewhat hypothetical.

The relationship between the Ob-Ugric forms requires additional research.

Loan etymology: 1. Hu ← West Old Turkic **taltučī* ‘the one who exercises a loss of consciousness’ (WOT: 841–846); 2. ← Turkic **tal-* ‘schlagen, prügeln’ (see UEW)

Status: 1. probable (see above); 2. rejected (UEW)

References:

Erdélyi 1960: Proto-Ugric (Hu + Kh; Mansi borrowed from Khanty)

UEW: Proto-Ugric http://www.uralonet.nytud.hu/eintrag.cgi?locale=en_GB&id_eintrag=1862

EWUng: 1475, s.v. *táltos*: Proto-Ugric

Honti 1982: 188, no. 637, POUG **tVlt-*

Honti 2017: 62–67: Proto-Ugric

Vásáry 2012: 37: Proto-Ugric or Hu ← Turkic

WOT: 841–846: Proto-Ugric or Hu ← Turkic

...

Proto-Ugric ? *ǰora or *ǰara ‘lake (?)’ (UEW: *ǰarʒ ‘während des Hochwassers entstandener See’)

Hungarian: *ár* ‘stream; flood; to stream, flood’

Old Hungarian: 1193 ? *aruod* [verb]; 1266 ? *Wisarahel* [noun; place name] (see EWUng: 43, s.v. *ár^l*)

Khanty: East (V) *lar* ‘während des Hochwassers an einem Wiesenufer entstandener See’, (Vj) ‘tiefliegendes, baumloses Wiesenufer od. Wiesengelände, das im Frühling überschwemmt

wird‘; South (DN) *tor* ‘See‘; North (O) *lar* ‘bei Hochwasser überschwemmtes Ufergebiet; See
‘ < Proto-Khanty **lār* (Honti 1982: **lar*)

Mansi: East (KU) *tūr*, West (P) *tūr*, (LO) *tor*, North (So) *tūr* ‘See’ < Proto-Mansi **tūrə* (Honti:
tūrə)

Proto-Ob-Ugric: **θēra* (Zhivlov 2006: 163), **θūrə* (Honti 1982: 139, No. 158)

Status: Unproblematic

Loan etymology: ← Indo-Iranian **sáras-* ‘lake’, cf. Vedic *sáras-* ‘See, Teich’ (Harmatta 1977: 171; Koivulehto 1999: 215; Katz [1985: 119–120] 2003: 102; for the source form, see EWAia II: 707)

Status: Accepted (see Holopainen 2019: 217–218)

Discussion:

The Ugric etymology features no problems and it is universally accepted (TESz; MSzFE; UEW; Zhivlov 2014). Katz’s (2003) idea to remove the Hungarian cognate from the Ob-Ugric forms is based on obsolete views of Uralic historical phonology. No cognates outside the Ugric branch are found: the Permic words Komi *šor*, Udmurt *šur* have been connected here by earlier research (see UEW for references) as well as more recently by Harmatta (1977) and Häkkinen (2009), but these reflect a different Uralic stem **šerä* (UEW; Holopainen 2019: 218). The Proto-Ugric form can formally reflect an earlier Uralic form with **s* or **š* (Zhivlov 2014: 127). The reconstruction of both **o-a* and **a-a* stem is possible.

EWUng assumes that the Ugric word is originally a nomen verbum, but it is very difficult to substantiate this claim; no traces of a verb are found in Ob-Ugric.

An Indo-Iranian etymology for the Ugric word has been suggested independently by several researchers (Harmatta 1977: 171; Koivulehto 1999: 215; Katz [1985: 119–120] 2003: 102), and the idea that the Ugric word reflects a loan from Proto-Indo-Iranian (or Proto-Iranian) **sáras-* is formally unproblematic (the Indo-Iranian word is an *s*-stem, but all the early Indo-Iranian loans in Ugric/Uralic are borrowed as vocalic stems). While it is true that a word for ‘lake’ had to belong to core vocabulary of the Ugric speakers, the word **sara* is likely to be a loan from somewhere, as it has no Uralic etymology.

If the loan etymology is correct, the word has to be borrowed before **s*, **š* > **ʒ* took place in Ugric. The same is true of most other Iranian loans as well, such as **θeraña* ‘gold’. More precise dating of the loan is difficult (cf. the problems with the dating of **s* > *h* in Iranian, see Hintze 1998).

Due to the meanings referring to ‘flood’ and ‘stream’ attested in Hungarian and Khanty, it would also be tempting to connect the Ugric words with Indo-Iranian forms like Vedic *sarít-* ‘stream’, derived from the root *sar-* ‘to run’. However, as the etymology of such forms is not quite clear and as no meaning ‘to flow’ can be reconstructed for the Indo-Iranian root **sar-* (see EWAia II: 707), this possibility remains quite speculative.

Alternative etymologies: --

References:

EWUng: 43, s.v. Proto-Ugric; nomen verbum?

Harmatta 1977: Proto-Ugric + Permic ← Indo-Iranian

Holopainen 2019: 217–218: Proto-Ugric ← Indo-Iranian

Honti 1982: 139 Proto-Ob-Ugric **θūrə*

Häkkinen 2009: 22: Proto-Ugric + Permic ← Indo-Iranian

Katz 2003: 102 Ob-Ugric (Hungarian not here) ← Indo-Iranian

Koivulehto 1999: Proto-Ugric ← Indo-Iranian

UEW: Proto-Ugric http://www.uralonet.nytud.hu/eintrag.cgi?locale=en_GB&id_eintrag=1747

Zhivlov 2006: 163 Proto-Ob-Ugric **θ̄era*

Zhivlov 2014: 127: Proto-Ugric/Uralic