

Competing etymologies: Analyzing problems in the origin of some words in Hungarian and other Uralic languages

The purpose of this paper is to discuss some debated etymologies that have competing explanations in Hungarian and in other Uralic languages. It is shown that in most cases quite clearly one of the suggested etymologies is superior and that some phonological details have been neglected in recent discussion of these etymologies. The discussion includes criticism of some recent etymologies that have not yet been commented on elsewhere in detail.

1. Introduction
 2. The etymologies
 - 2.1. Hungarian *sok* ‘much, many’
 - 2.2. Hungarian *orvos* ‘doctor’
 - 2.3. Hungarian *tojik* ‘lay eggs’
 - 2.4. Hungarian *üdül* ‘refresh oneself; recover; rest and relax’
 - 2.5. Hungarian *ing* ‘shirt’
 3. Concluding remarks
- References

1. Introduction¹

The purpose of this paper is to discuss some debated etymologies in Hungarian and in other Uralic languages. The etymologies that are discussed here have competing etymologies: most have both a competing Uralic (Finno-Ugric) etymology and a loan etymology and one has two competing loan etymologies, both of which involve some debated issues. In the case of most etymologies discussed here, problems with both explanations have been noted in the research literature, but there is no consensus on which explanation is correct or more plausible.

1. I am grateful to Luobbal Sámmol Sámmol Ánte (Ante Aikio), Christopher Culver, András Czentnár, Martin Kümmel, and Johanna Laakso for discussions

Even though Uralic etymology and loanword research is a vibrant field, there are etymologies that include unsolved problems. Some of the more problematic Uralic etymologies presented in earlier sources, such as the UEW, MSzFE, or SSA, have not been commented on in detail in more recent etymological dictionaries, such as ÚESz or SES. On the other hand, some problematic Turkic loan etymologies for Hungarian that have been discussed, for example, by Róna-Tas & Berta (2011), have not been analyzed further in more recent research (see, however, Honti 2017 for discussion of some etymologies and Róna-Tas 2017 for answers to criticism). There are also some quite recent etymological proposals that have not yet been, at least to my knowledge, commented on (Aikio 2021 on *üdül*; Róna-Tas & Berta 2011 on *ing*).

The etymologies discussed here form case studies to highlight the problems with competing etymologies. In this article I intend to show that in a few cases, two competing etymologies are equally good. It is also my aim to highlight the fact that discussing problematic Uralic etymologies in more detail can open fruitful aspects for future research. The etymologies discussed in the following are the Turkic etymology of Hungarian *sok* ‘much, many’ as well as the alleged cognates in Mansi, Mari, and Komi; Hungarian *orvos* ‘doctor’ and its alleged cognates in Finnic and Saami; the competing Uralic and Turkic etymologies for Hungarian *tojik* ‘lay eggs’; the Turkic and Iranian loan etymologies of Hungarian *üdül* ‘refresh oneself; recover; rest and relax’ and recently suggested cognates in Komi and Nganasan; and the competing Iranian and Turkic etymologies of Hungarian *ing* ‘shirt’.

on the etymologies discussed here and to two anonymous reviewers for useful comments that have helped to improve this paper. I am solely responsible for the remaining errors. This research has been partly supported by the Research Council of Finland (project 356825) and by an APART-GSK fellowship of the Austrian Academy of Sciences at the University of Vienna.

2. The etymologies

2.1. Hungarian *sok* ‘much, many’

The Uralic etymology: Hungarian *sok* has been assumed to be cognate with the following words.

Mansi: South *šaw*, East *šǝw*, West *šāw*, North *sāw*; Mari (Carevokokšajsk) *ćoka* ‘thick’, (Kozmodemjansk) *čakata* ‘thick, firm, compact’; Komi *ček* ‘thick’ < Proto-Uralic/Proto-Finno-Ugric *čukkV ~ *čokkV (UEW: 62–63)

Loan etymology:

Hungarian *sok* (< ? early Proto-Hungarian *čokV) ← ? West Old Turkic *čok ‘much, many’, cf. Middle Turkic *čoq* id., reconstructed East Old Turkic *čok³ id. < ? Proto-Turkic *taši-ok (Róna-Tas & Berta 2011: 729–735; suggested as an alternative to the Uralic etymology)

Although Hungarian *sok* and its cognates have often been considered reflexes of Proto-Uralic *čukkV ~ *čokkV (MszFE s.v. *sok*; UEW: 62–63; EWUng s.v. *sok*; Honti 2017: 53; ÚESz s.v. *sok*) and even though the etymology is listed among the certain etymologies in the UEW, it is obvious in the light of modern phonological studies that this cannot be correct. Interestingly, in a recent but short commentary on the etymology, Honti (2017: 53) notes that the Finno-Ugric etymology is without a doubt correct (“Minden kétséget kizáróan finnugor eredetű *sok* szavunk”). However, Honti does not comment on the phonological details, which quite clearly point to the opposite conclusion. It is also difficult to understand why the recent ÚESz (s.v. *sok*) claims that the Turkic etymology is incorrect; no arguments are given. The competing Turkic etymology seems probable at least for the Hungarian word; this will be commented on in more detail below.

-
2. The West Old Turkic language is a reconstructed language, without direct descendants, which Róna-Tas & Berta (2011) postulated on the basis of early Turkic loanwords into Hungarian. For criticism and additional notes on the concept of West Old Turkic, see Erdal (2018: 511–512) and Sándor (2021). The problematic issues with this protolanguage cannot be discussed here, so the term and concept of West Old Turkic is accepted here for the sake of this study.
 3. The word is not attested in actual East Old Turkic sources, but Róna-Tas & Berta reconstruct the expected East Old Turkic form as *čok.

The UEW gives two reconstruction variants, *čukkV ~ *čokkV, but it is not explained what the reason for this is. The vocalism of the Hungarian and Mari words is not commented on, but an explanation is offered for the dubious Komi vocalism and the unexpected *-w- in the Mansi word.

To start with the Hungarian word, the *o* in *sok* quite clearly points to earlier *u: Proto-Uralic *o would not be retained in Hungarian except before *čk, for example, Hungarian *močki > Hungarian *mos* ‘wash’. The assumed Turkic original has *o, but there are parallels for Hungarian *o* corresponding to Proto-Turkic *o in loans, and these will be discussed below in more detail.

The alleged Mansi cognate is problematic. The vocalism across the Mansi dialects poses some problems: most of the variants point to Proto-Mansi *šāw (cf. Zhivlov 2006: 69), whereas the Southern dialects show an unexpected short *a*. If the vocalism of the Southern forms is considered to be a result of the shortening of earlier *ā, we can reconstruct Proto-Mansi *šāw, which obviously cannot reflect Proto-Uralic *čokka or *čukka.⁴ Furthermore, Mansi *w cannot reflect Proto-Uralic *kk, as *kk became *k in Proto-Ob-Ugric and this should have been retained as *k in Proto-Mansi (Honti 1999: 139–140; Zhivlov 2023: 142–143). The UEW mentions the parallel of Proto-Finno-Ugric *rakka > East Mansi *row-* ‘let (the game animals) sneak up close [nahe heranschleichen lassen (vom Wild)]’, but this is not a plausible parallel in the light of modern research, as the alleged Finnic cognate *rakas* is a Germanic loan (← Proto-Germanic *frakaz > Old English *fræc* ‘lustful, eager’; LÄGLOS III s.v. *rakas*), meaning that no Proto-Uralic word can be reconstructed. Also ascribing a Proto-Ugric origin to Mansi *row-* and related forms must be regarded as erroneous because of the problem with *w.

Recently the issue with the Mari cognate has been briefly discussed by Bereczki (2013: 3–4), who notes that Mari *čok* ‘very, very much, in a great amount (sehr, sehr viel, in großer Menge)’ (< Proto-Mari *čok) is rather the reflex of Proto-Uralic *čukkV ~ *čokkV, not the dialectal words *čoka* or *čakata* mentioned in the UEW. It is true that the vowel correspondence East Mari *o* ~ West Mari *a*, reflecting Proto-Mari *ā in the reconstruction by Aikio (2014), could not regularly reflect either Proto-Uralic *o or *u. However, the Mari word *čok* could probably also be an archaic loan from Turkic. Proto-Mari *č would be a phonetically suitable substitution of

4. I am grateful to an anonymous reviewer for comments on the vocalism of the Mansi words discussed here.

Turkic *č, but the exact chronology of the possible borrowing from Turkic to Mari would need further research (a reflex of Turkic *čok is attested neither in Chuvash nor in Kazan Tatar, the usual Turkic languages with which Mari has been in contact; Róna-Tas & Berta 2011: 731). It should also be noted that Proto-Mari *ć, not *č, has served as the Mari substitute for Turkic *č in some loans, at least in Eastern Mari *čatan* ‘lame’ (< Proto-Mari *ć) that was borrowed from Middle Bashkir *čatan ‘lame’ (Bereczki 2013: 3–4). The etymology of the unrelated *ćoka* ‘thick’, *čakata* ‘thick, firm, compact’ remains unclear, but these words cannot be explained as borrowings from the same Turkic word, however.

Concerning the Komi vowel *ę*, the UEW notes that a sporadic change **j* > *ó* > *ę* could have taken place in this word. However, Komi *ę* can go back to Proto-Uralic **o* in old **i*-stems; nevertheless, this is not the regular reflex of **u* that we would expect based on Hungarian (Zhivlov 2023: 136).

To sum up, the suggested Uralic cognates involve phonological problems that make the etymology untenable. Now it is time to investigate the phonology of the possible Turkic etymology in more detail.

Hungarian *š* from *č is regular in inherited words, compare, for example, Hungarian *sötét* ‘dark’ < Proto-Ugric *čittV- (UEW: 842), and there are many examples of old Turkic loans borrowed before this change happened, e.g. Hungarian *som* ‘cornel (*Cornus*)’ ← Turkic *čum id., Hungarian *sólyom* ‘falcon’ ← Turkic *čavli id. Regarding the *o* vocalism, parallels can be found among other Turkic loans: for example, *tok* ‘sturgeon (*Acipenser sturio*)’ ← Turkic *toku id., *toklyó* ‘one-year-old lamb’ ← Turkic *toklu id. are found in the material of Róna-Tas & Berta (2011). It is admittedly dubious that these words did not go through the lowering *o* > *a* that probably happened in Old Hungarian, and for Hungarian *o* we would expect Proto-Hungarian **u* as the most probable preform (as is the case with *som* ← Turkic *čum above). However, the fact that we have a number of Turkic loans in Hungarian displaying this vowel correspondence means that the vowel *o* is not a problem, even if not all the details are fully understood.

To conclude, the Uralic etymology of Hungarian *sok* in its traditional form is phonologically impossible and has to be rejected, whereas the Turkic etymology does not involve any major problems and is clearly the more probable alternative. The loan etymology is also semantically more probable. Based on the material of Róna-Tas & Berta (2011), the lexical evidence of Turkic–Hungarian contacts points to an intensive period of contact, where verbs and central concepts such as body-part terms have also

been borrowed, so the borrowing of a function word ‘many’ would not be surprising. The alleged Mari cognate *ćok* is probably also a Turkic loan, although the exact periodization of this borrowing needs additional research. Mari also has a high number of Turkic loanwords as well as Turkic structural influence (Róna-Tas 1988: 768–774), so the borrowing of a function word is not problematic also in this case. The problem of the origin of the Komi and Mansi words discussed above requires further research. Formally, Hungarian *sok* and Mari *ćok* might reflect a common protoform *čukka, but the borrowings from Turkic are more probable.

2.2. Hungarian *orvos* ‘doctor’

(dial. *olvos*, *óros*, *orvas*; see ÚMTSz IV: 212, s.v. *orvos*)

Uralic (Finno-Ugric) etymology: *orvos* has been assumed to be cognate with the following words.

Finnish *arpa* ‘Wahrsagungsgerät, Schlagrute, Los (Loskugel, Würfel); Schlagrute um verborgene Dinge zuentdecken; budskap, uppbud avfolket (genom kringskickad budkavle); Botschaft, Aufgebot des Volks (durch umhergeschickten Aufgebotsstab)’ (cognates in other Finnic languages), North Saami *vuorbi* ‘each of the two or more pieces of wood, stones, etc., used by persons who are going to cast lots something; lot; destiny’ (cognates in many other Saami languages) < Proto-Uralic *arpa ‘soothsaying instrument, magic instrument, potion’ (UEW: 16: uncertain Finno-Ugric etymology)

Loan etymology:

Hungarian *orvos* ← West Old Turkic *orwuči, cf. East Old Turkic *arvišči* ‘spellbinder, sorcerer’, derived from Proto-Turkic *arva- ‘make magic, cast spells’ (MSzFE: 505–506; Róna-Tas & Berta 2011: 656–659)

Already MSzFE mentions the Turkic origin of Hungarian *orvos* as an alternative to the Uralic etymology. Also in the UEW and in many sources published after that, both the Uralic and Turkic origin is mentioned as an alternative (see, for example, EWUng: 1071–1072, s.v. *orvos*; Róna-Tas & Berta 2011: 656–659 and Honti 2017: 48–49). It is, however, not noted by these Hungarian sources that Koivulehto (1972: 249) had suggested a Germanic origin for the alleged Finnic and Saami words; Koivulehto’s suggestion has been since accepted by LÄGLOS (s.v. *arpa*). Furthermore, the Hungarian word involves phonological problems, something that was noted already

by MSzFE: Hungarian *o* does not correspond regularly to *a* in Finnic and *vuo-* in North Saami, both of which regularly reflect Proto-Uralic **a* in **a*-stems. Furthermore, the *-s* element in the Hungarian word is difficult to explain on the basis of the Uralic word.

Honti (2017: 48–49) considers both etymological alternatives possible, noting that perhaps the Hungarian word is of Finno-Ugric origin, but it was later influenced by the Turkic word. This would not be impossible if the Finnic and Saami words did not have a Germanic etymology, but in the current situation it is much more likely that the words have nothing to do with each other. Regarding the phonological problem of the Hungarian vocalism, Honti assumes that affectiveness might have played a role here. This is a difficult explanation, however. Róna-Tas (2017: 51–52), in a commentary on Honti’s criticism, considers both the Uralic and Turkic explanations possible.

Helimski (manuscript) has also defended the traditional etymology, arguing that Hungarian *o* is due to the assimilating influence of the following syllable.⁵ However, his defense is ad hoc, as there seem to be no parallel examples of Uralic **a* being reflected as *o* in Hungarian due to the following syllable.

The recent ŪESz gives a rather confusing explanation, mentioning that the Hungarian word is of uncertain origin, perhaps a derivative of a stem of Finno-Ugric origin. It is also mentioned that the Finno-Ugric words might have Turkic connections, but it is difficult to understand what is meant here exactly.

The Turkic etymology is more plausible from a phonological point of view. Hungarian *s* can easily be explained from Turkic **ç*, as numerous parallels can be found; see the discussion of Hungarian *sok* above. There are also many parallels for the substitution of Turkic **a* with Hungarian *o* (Dybo 2009: 77–78; Róna-Tas & Berta 2011: 1115). The semantic connection is also plausible.

Regarding the vocalism of the Hungarian word, Róna-Tas & Berta (2011) assume West Old Turkic **o*, because Old Hungarian /*u*/ in the early attestations of the word, such as *Vruus* (personal name) and *vruufok* (ŪESz s.v. *orvos*) would be difficult to explain in the case of West Old Turkic **a*.

5. Helimski’s original formulation was as follows: “Das unregelmäßige wortanlautende *o* im ung. *orvos* ‚Arzt‘ ist vermutlich der Entwicklung *a* > *o* vor einem *o* in der zweiten Silbe zu verdanken.”

However, in Old Hungarian, the grapheme <u> can also denote the phoneme /o/ (Benkő 1980: 89–94), so the argument is not compelling. Perhaps the Turkic donor form had *a and Hungarian substituted it with *o, possibly because of the labial in the following syllable.

Concerning the Germanic etymology of the Finnic and Saami words, it involves no major problems: Koivulehto (1972) argues that these words were borrowed from Proto-Germanic *arþa- or *arþa-z, as reflected by e.g. Old Swedish *arf*, *arver* ‘possession of land, inheritance’. Phonologically the explanation is plausible, and Koivulehto also argues that semantic parallels exist, such as Gothic *hlaus* ‘lot, inheritance, the share of inheritance’, noting that in past eras inheritance was often determined by casting lots. LÄGLOS also notes that reflexes of the related Germanic form *arþija- (> Old Norse *arþija* ‘wake (Begräbnismahl)’) come closer to the semantics of the Finnic and Saami words (see also Bjorvand & Lindeman 2000: 46–47, s.v. *arv*; Seebold 2019; Koch 2020: 113 and Zimmer 2020: 61 for more detailed discussions of the semantics and the possible background of the Germanic word family). Whether the words in Saami and Finnic are parallel borrowings or cognates (as they regularly could be) should be discussed elsewhere in more detail, but the Germanic etymology for these words is plausible.

2.3. Hungarian *tojik* ‘lay eggs’

Uralic etymology: *tojik* has been considered cognate with the following words.

Finnish *tuo-* ‘bring’ (cognates in other Finnic languages); South Saami *doeke-* ‘sell’ (cognates in Ume and Pite Saami); Mordvin (Erzya and Moksha) *tuje-* ‘bring’, Khanty (East, South, North) *tu-* ‘bring’, ? Mansi *tōl-* (South), *tūl-* (West, East, North) ‘bring’, Tundra Nenets *tā-* ‘give, bring’ (cognates in other Samoyed languages) < Proto-Uralic *toyi- ‘bring, carry, give’ (MSzFE: 635–636; UEW: 529–530, where Hungarian *tojik* is mentioned with a question mark; Sammallahti 1988: 550⁶)

Loan etymology:

Hungarian *tojik* ← West Old Turkic *tuγ- or *toγ-, cf. East Old Turkic *tug-*, *tog-* ‘be born’ (Róna-Tas & Berta 2011: 911–914; first suggested by Vambéry 1870)

6. The Uralic word is reconstructed as *toxi- by Sammallahti (1988). The differ-

This etymology is a rather similar case to the two etymologies discussed above: the Hungarian word has a possible Uralic etymology and a possible Turkic one. However, unlike in the two etymologies above, in this case we are dealing with a plausible Uralic etymology as such, but the cognacy of Hungarian *tojik* with the rest of the forms is debated. The UEW and also SSA (s.v. *toivoa*) are uncertain of the inclusion of the Hungarian word in this cognate set and problems are mentioned already by MSzFE. The Turkic etymology is mentioned as a possibility by both the UEW and MSzFE. ŪESz mentions both possibilities, noting that the Turkic origin of Hungarian *tyúk* ‘chicken’ makes the Turkic borrowing likely. Among the reflexes of Uralic **toyi-*, Hungarian *tojik* is a semantic outlier, as it shows a specific meaning ‘lay eggs’ divergent from the rest of the meanings that quite transparently continue Proto-Uralic ‘bring’ or ‘give’. It has been mentioned in the UEW that the semantic development ‘bring’ > ‘lay eggs’ is plausible, but here it should be noted that the meaning attested in East Old Turkic is even closer. Honti (2017: 68–69) also discusses the etymology but does not state clearly which explanation he supports.

Bigger problems are posed, once again, by the phonology. As has been noted above, Proto-Uralic **o* should not have been retained as *o* in Hungarian: **o* > *a* or *á* would be the expected development. On the other hand, it was already noted above that the correspondence Hungarian *o* ~ West Old Turkic **o* has parallels, and furthermore, in the case of this etymology, forms with both *o* and *u* are attested in East Old Turkic according to Róna-Tas & Berta (2011). Whichever vowel the donor form had, the Hungarian vocalism is easier to derive from that than from Proto-Uralic **o*.

The consonantism also involves problems: already MSzFE noted that Hungarian *j* is unexpected as the hiatus filler after the loss of Uralic **γ*; Hungarian usually has *v*, cf. *tavat* (accusative of *tó* ‘lake’) < Proto-Uralic **tuγi* ‘lake’. The matter of *j* from Turkic **γ* requires further research, but there seem to be few examples of **γ* in intervocalic position among loans (Róna-Tas & Berta 2011: 1081). Even though the matter of *j* is not completely clear, the phonology and semantics otherwise point to a Turkic donor rather than Uralic inheritance.

ence in notation is due to the uncertainty regarding the nature of the Proto-Uralic phoneme **x*/**γ*. This phoneme is reconstructed as a velar spirant by some (for example, in the UEW) but its phonetic realization and its status in the phonological system of Proto-Uralic remains insufficiently understood (see Aikio 2022: 7–8 for a recent discussion).

Lastly, it should be noted that the Uralic word has been considered a loan from Proto-Indo-European *doh₃- ‘give’ (see e.g. Koivulehto 1991: 8–9); the Indo-European etymology is also mentioned in the UEW, but it is mentioned with a question mark in SSA and SES. This has little impact on the origin of the Hungarian word as it would be quite unlikely that a very old Indo-European loanword would have been borrowed separately into Hungarian, but as this is theoretically possible, this issue should be mentioned. The existence of Proto-Indo-European loanwords in Uralic languages has been doubted by Simon (2020), who notes that this particular etymology is semantically problematic as the meaning ‘give’, prevalent in Indo-European, can hardly be reconstructed for the Proto-Uralic word as it is found only in couple of cognates where it might reflect an areal semantic innovation (Simon 2020: 252). It could be added here that the Indo-Iranian branch of Indo-European, a known contact language of early Uralic (see Holopainen 2019) would not be a very probable donor of either the Proto-Uralic word or Hungarian *tojik*, as no plausible donor form can be found among the Indo-Iranian conjugated forms; the present stem is a reduplicated stem (for example, Sanskrit *dādāti* ‘gives’) and it is unclear how this kind of stem would have been borrowed into Uralic (see EWAia I: 713–715 for more information on the Indo-Iranian forms).

2.4. Hungarian *üdül* ‘refresh oneself; recover; rest and relax’

Uralic etymology: Hungarian *üdül* ‘refresh oneself; recover; rest and relax’, *üdvözöl* ‘greet, salute’, *üdvöz* (obsolete) ‘greeted, saluted’ is cognate to the following words.

Finnish *synty-* ‘be born’ (cognates in other Finnic languages), Komi *sod-* ‘grow’, Nganasan *tjntud’i* ‘come back to life; wake up’ < Proto-Uralic *sen-tiw- ‘be born’ (Aikio 2021: 171)

Loan etymology:

Hungarian *üdül* is related to *ül* ‘feast (verb)’ (< *üdl-), (obsolete) *ëgy* ‘holy’ (in the compound *ëgyház* ‘church (*holy house)’) and *ünnep* ‘festival’ (< *üd-nap ‘holy day’); the word family is borrowed from West Old Turkic *edü ‘good; holy (?)’, cf. East Old Turkic *ädgü* (Róna-Tas & Berta 2011: 307–310).

The idea that Hungarian *üdü* and its apparent cognates are Turkic loans is considered plausible by Róna-Tas & Berta (2011) and the issue has a very long research history (see also Ligeti 1977: 7–9; 1986: 33, 195, 273 for a thorough discussion of this etymology and some phonological problems). Recently, Aikio (2021) has noted that *üd-* could reflect Proto-Uralic **sen-tiw-*; he mentions that the Hungarian word lacks an etymology, but this is incorrect, as the Turkic origin has been widely discussed. However, it is true that EWUng (s.v. *üdü*) does not consider the idea of a Turkic etymology correct. As Aikio's new idea would be a plausible possibility as such, it is important to discuss it in more detail.

Aikio's etymology is semantically plausible, and the Uralic cognates could be derived from a Proto-Uralic word that would have had the meaning 'be born' or 'refresh' or 'come back to life'. However, phonologically the issue is less clear. While Hungarian *üd-* could reflect **sün-tiw-* (> Finnish *synty-*) without problems, Komi and Nganasan point to Proto-Uralic **e*, and the labial **ü* in Finnic is due to regressive assimilation. Aikio mentions parallels to this Finnic development, but for a similar development in Hungarian no parallels are known. One possibility would be to assume that **sen-tiw-* > **sün-tiw-* is a Proto-Uralic development and the vocalism in Komi and Nganasan was influenced analogically by other forms of the verbal stem **sen(i)-* where the original **e* vocalism was retained.

The matter of the vocalism of this Proto-Uralic word requires further research. However, it is also clear that problems with the vocalism are involved with the Turkic etymology. Again, semantically it is plausible that all the Hungarian forms mentioned above belong together etymologically, but the alternation of *ë* and *ü* would be difficult to explain. Issues with *gy-* in *ëgy* and *d* in *üd-* have also raised problems in the past (see Ligeti 1977: 7–9), meaning that there are various problems with the Turkic etymology that would require a more detailed treatment. Furthermore, the relationship of *üdü* and *ül* (< ? **üdl-*), which seem to reflect similar formations, requires further research.

It seems clear that the *ëgy-* in *ëgyház* is probably a Turkic loan, as a similar compound 'holy house > temple' is attested in early Turkic (such as medieval Cuman *yix-ön* 'church (literally holy house)' in the Codex Cumanicus), as noted by EWUng, Ligeti (1986), and Róna-Tas & Berta (2011). However, the phonological details in deriving *ëgy* (< ? Old Hungarian *igy*) from a reconstructed Turkic **edü* or **ädgü* remain obscure, regarding both the phonology (the vocalism and the relationship of Hungarian *gy*

to Turkic *d or *dg) and the semantics, and further research is needed. One can be quite certain that *ëgy* ‘holy’ does not have, in any case, anything to do with the Uralic verb *sen-tiw-.

Although no satisfactory solution to the problem of *üdüil* and related forms is available, this treatment has hopefully shown that the existing etymologies involve problems that need addressing, and that both the new suggestion by Aikio and the earlier etymologies circulating in the literature should not be taken for granted without a more thorough scrutiny. However, the possibility that we are dealing with unrelated forms, one inherited from Uralic and the other borrowed from Turkic, is a plausible option based on the scrutiny above.

Lastly, it should be mentioned that Szemerényi (1981: 241–242) had also suggested an Iranian etymology for this Hungarian word family, assuming a borrowing from a Persian-type *yad- (< Proto-Iranian *yadz- ‘holy’), but this idea is implausible: there is a lack of parallels for borrowings from such a Persian-type language and the etymology also involves problems with the vocalism (Holopainen & Czentnár 2022).

2.5. Hungarian *ing* ‘shirt’

(dial. *imëg*, *ümëg*; ÚMTSz s.v. *ing*)

Iranian loan etymology:

Hungarian *ing* < *imëg < (? *jämVgV) ← Middle Iranian *yāmag ‘shirt’, cf. Middle Persian *jāmag* ‘shirt’ (Joki 1973: 264; Szemerényi 1981: 242; Har-matta 1997: 81; Katz 2003: 308)

Turkic loan etymology:

Hungarian *ing* < *ümeğ < *ümmeg ← (?) West Old Turkic *öñmāk, derivative of *öñ ‘front’; cf. Eat Old Turkic *öñ* ‘front’

The Turkic etymology is a new idea of Róna-Tas & Berta (2011: 457–459). As this etymology has never been, to my knowledge, discussed in subsequent research, it is important to analyze it here. ÚESz (s.v. *ing*) mentions the Iranian etymology as an uncertain possibility but does not mention the new suggestion by Róna-Tas & Berta. To start with, this etymology involves two problems: the speculative donor form and certain phonological obstacles. The form *öñmāk is postulated by Róna-Tas and Berta to account for the Hungarian word, but it does not have any actual Turkic basis. The Turkic derivational suffix *-māk is a real and productive suffix according to those

authors, so the etymology is not impossible, but it is still dubious to reconstruct unattested forms simply on the basis of loanwords. The semantic side of the Turkic etymology is also more problematic compared to the Iranian one.

Regarding the phonological problems, it is not clear how Hungarian *i* and *ü* in dialectal forms like *umëg* could be explained from Turkic *ö*. Róna-Tas & Berta assume a substitution Hungarian **ü* ← Turkic **ö*, but this would need good parallels before the idea can be accepted. Furthermore, the idea of a Proto-Hungarian geminate **mm* in the reconstruction **ümmeg* seems to be ad hoc. The voiced *g* in Hungarian would not be a problem as such, as there are some loanwords where Hungarian word-final *-g* corresponds to Proto-Turkic word-final **-k* although in most cases Turkic **-k* is reflected either as Hungarian *-k* or as zero (< Proto-Hungarian **γ*; Róna-Tas & Berta 2011: 1076).

The Iranian etymology, which is an older idea, also involves some problems, but it is overall much more likely than the Turkic etymology. EWUng (I: 614) considers the Iranian etymology possible but mentions problems with the vocalism. However, the idea that Hungarian *i* can correspond to Iranian **yā-* is not that problematic: if it is assumed that a reconstructed Middle Iranian **yāmag* (< Proto-Iranian (?) **yāmaka-*⁷) was borrowed as **jāmVgV* into Proto-Hungarian, *i* can be explained from **jä* in a similar way as *ideg* ‘bowsting’ from Proto-Uralic **jänti(η)* (> Finnish *jänne* etc.). EWUng also refers to other etymologies where Hungarian *i* alternates with a sequence of *j* and a vowel, such as *juhár* ~ *ihár* ‘maple’. It is not completely clear that the modern alternation in the Hungarian dialects is connected to the change *i* < **jä* that seem to have produced *i* already quite early; it is also not clear under which conditions this change happens, and also the chronology is not completely clear, but this Iranian loanword can be helpful in the discussion of these issues.

7. The Middle Iranian word can be reconstructed on the basis of Middle Persian *jāmag* [jʸmkʰ], New Persian *jāma* ‘clothing, garment’ as well as related forms in some modern Iranian languages (Rastorgueva & Edelʹman 2007: 76; Horn 1893: 93; MacKenzie 1971: 46). According to Rastorgueva & Edelʹman (2007: 76) and Bailey (1979: 108) the Middle Persian word is derived from the Proto-Iranian verbal root **yam-* ‘hold, bear’, reconstructed by Cheung (2007: 211–212) with the meanings ‘hold, stretch, reach out’ (see also AiWb: 1262–1263; *yam* ‘halten, fassen’; Kellens 1995: 46). The semantic side of this idea is not entirely satisfactory, but that does not affect the loan into Hungarian.

However, even if $i < *j\ddot{a}$ is regular, the origin of dialectal forms with labial \ddot{u} remains somewhat problematic. If *ing* goes back to earlier $*j\ddot{a}$ -, the forms with \ddot{u} have to reflect secondary labialization.

The Iranian word is not widely attested in the Middle Iranian period, and it is doubtful whether a Middle Persian or early modern Persian word could have been directly borrowed into Hungarian. However, we cannot exclude the possibility that the word would have been found in other Middle Iranian varieties, and a Persian cultural term could have found its way into Proto-Hungarian through trade contacts or through some intermediary language (see Holopainen 2024: 32, 39, 42, 56 for discussion of problems with some allegedly Persian loanwords in Hungarian).

3. Concluding remarks

The etymological analysis of this paper has shown that in several cases where two alternative etymologies for a Hungarian word have been suggested, in the end one of the alternatives is clearly inferior. The Uralic etymologies of *sok* ‘much, many’ and *orvos* ‘doctor’ involve phonological problems that make them extremely unlikely, whereas the Turkic etymologies show established patterns of sound substitution; some of the alleged Uralic cognates of *sok* can also be explained as Turkic loans, and the alleged Finnic and Saamic cognates of *orvos* have a better alternative etymology. Concerning the etymology of *tojik* ‘lay eggs’, the Turkic etymology is also less complicated and, in this case, also semantically more plausible than the Uralic etymology. The situation of *üdül* ‘refresh oneself; recover; rest and relax’ and its relationship to *ëgy* ‘holy’ is more complicated and requires further research, but the idea presented here that the verb is of Uralic origin and the noun from Turkic can be supported with at least some arguments. The Turkic etymology of *ing* ‘shirt’ involves a speculative donor form that cannot be supported by the actual data.

References

- AIKIO, ANTE (LUOBBAL SÁMMOL SÁMMOL ÁNTE). 2014. On the reconstruction of Proto-Mari vocalism. *Studies in Language Relationship* 11. 125–157.
- AIKIO, ANTE (LUOBBAL SÁMMOL SÁMMOL ÁNTE). 2021. Studies in Uralic etymology V: Permic etymologies. *Linguistica Uralica* 57(3). 161–179.
- AIKIO, ANTE (LUOBBAL SÁMMOL SÁMMOL ÁNTE). 2022. Proto-Uralic. In Bakro-Nagy, Marianne & Laakso, Johanna & Skribnik, Elena (eds.), *The Oxford guide to the Uralic languages*, 3–27. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198767664.003.0001>
- AiWb = BARTHOLOMAE, CHRISTIAN. 1961. *Altiranisches Wörterbuch*. 2., unveränderte Auflage. Berlin: De Gruyter.
- BAILEY, HAROLD. 1979. *Dictionary of Khotan Saka*. Cambridge: Cambridge University Press.
- BENKŐ, LORÁND. 1980. *Az Árpád-kor magyar nyelvű szövegelemélei*. Budapest: Akadémiai Kiadó.
- 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-Altaica 86). Wiesbaden: Harrassowitz Verlag.
- BJORVAND, HARALD & LINDEMAN, FREDRIK OTTO. 2000. *Våre arveord: Etymologisk ordbok*. Oslo: Novus Forlag & Instituttet for sammenlignende kulturforskning.
- CHEUNG, JOHNNY. 2007. *Etymological dictionary of the Iranian verb* (Leiden Indo-European Etymological Dictionary Series 2). Leiden & Boston: Brill.
- ДУБО, ANNA. 2009. Вокализм раннетюркских заимствований в венгерском. *Finnisch-Ugrische Mitteilungen* 32/33. 83–132.
- ERDAL, MARCEL. 2018. [Review of] Róna-Tas, András & Berta, Árpád †: *West Old Turkic: Turkic loanwords in Hungarian*. *Orientalistische Literaturzeitung, Zentralasien* 113-6. 510–513. <https://doi.org/10.1515/olzg-2018-0172>
- EWÁia = MAYRHOFER, MANFRED. 1986–2001. *Etymologisches Wörterbuch des Altindoarischen I–III*. Heidelberg: Winter.
- EWUng = BENKŐ, LORÁND. 1992–1997. *Etymologisches Wörterbuch des Ungarischen*. Budapest: Akadémiai Kiadó.
- HARMATTA, JÁNOS. 1997. Iráni nyelvek hatása az ősmagyar nyelvre. In Kovács, László & Veszprémi, László (eds.), *Honfoglalás és nyelvészet*, 71–83. Budapest: Balassi.
- HELIMSKI, EUGENE. (Manuscript.) *Marginalia ad UEW*. Available online: <https://www.slm.uni-hamburg.de/ifuu/download/marginalia-ad-uew.pdf>
- HOLOPAINEN, SAMPSA & CZENTNÁR, ANDRÁS. 2022. Speculative Iranian loanwords in Hungarian and Turkic. Presentation at the Deutscher Orientalistentag, Freie Universität Berlin, 16.09.2022. https://ugric.univie.ac.at/fileadmin/user_upload/p_ugric/Berlin_DOT_Holopainen_Czentnar_updated.pdf (Accessed 2022-09-26.)

- HOLOPAINEN, SAMPSA. 2019. *Indo-Iranian borrowings in Uralic: Critical overview of sound-substitutions and distribution criterion*. Helsinki: University of Helsinki. (Doctoral dissertation.)
- HOLOPAINEN, SAMPSA. 2024. Notes on the phonology of Alanic loanwords in Hungarian. *Die Sprache: Zeitschrift für Sprachwissenschaft* 56. 20–64.
- HONTI, LÁSZLÓ. 1999. *Az obiugor konzonantizmus története* (Studia Uralo-Altaiica Supplementa 9). Szeged: Szegedi Tudományegyetem.
- HONTI, LÁSZLÓ. 2017. *A magyar és a nyugati ótörök szókészleti kapcsolatairól*. Budapest: Tinta Könyvkiadó.
- HORN, PAUL. 1893. *Grundriss der neupersischen Etymologie*. Strassburg: K. J. Trübner.
- JOKI, AULIS J. 1973. *Uralier und Indogermaner: Die älteren Berührungen zwischen den uralischen und indogermanischen Sprachen* (Mémoires de la Société Finno-Ougrienne 151). Helsinki: Société Finno-Ougrienne.
- KATZ, HARTMUT. 2003. *Studien zu den älteren indoiranischen Lehnwörtern in den uralischen Sprachen*. Aus dem Nachlass herausgegeben von Paul Widmer, Anna Widmer und Gerson Klumpp (Indogermanische Bibliothek, 3. Reihe). Heidelberg: Winter.
- KELLENS, JEAN. 1995. *Liste du verbe avestique*. Wiesbaden: Reichert.
- KOCH, JOHN T. 2020. *Celto-Germanic: Later prehistory and post-Proto-Indo-European vocabulary in the North and West*. Aberystwyth: Canolfan Uwchefrydiau Cymreig a Cheltaidd Prifysgol Cymru.
- KOIVULEHTO, JORMA. 1972. Germanisch-finnische Lehnbeziehungen II. *Neuphilologische Mitteilungen* 73. 575–628.
- KOIVULEHTO, JORMA. 1991. *Uralische Evidenz für die Laryngalthorie* (Veröffentlichungen der Kommission für Linguistik und Kommunikationsforschung, Heft 24). Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- LÄGLOS = KYLSTRA, A.D. & HAHMO, SIRKKA-LIISA & HOFSTRA, TETTE. 1991–2012. *Lexikon der älteren germanischen Lehnwörter in den Ostseefinnischen Sprachen I–III*. Amsterdam: Rodopi.
- LIGETI, LAJOS. 1977. *A magyar nyelv török kapcsolatai és ami körülötte van I* (Budapest Oriental Reprints, Series A 1–2). Budapest: Magyar Tudományos Akadémia & Körösi Csoma Társaság.
- LIGETI, LAJOS. 1986. *A magyar nyelv török kapcsolatai a honfoglalás előtt és az Árpád-korban*. Budapest: Akadémiai Kiadó.
- MACKENZIE, DAVID. 1971. *A concise Pahlavi dictionary*. London: Oxford University Press.
- MSZFE = LAKÓ, GYÖRGY. 1971–1978. *A magyar szókészlet finnugor elemei*. Budapest: Akadémiai Kiadó.
- RASTORGUEVA & EDEL'MAN 2007 = Расторгуева, В. С. & Эдельман, Д. И. 2007. *Этимологический словарь иранских языков, Том. III*. Москва: Восточная литература.
- RÓNA-TAS, ANDRÁS & BERTA, ÁRPÁD (with the assistance of LÁSZLÓ KÁROLY). 2011. *West Old Turkic: Turkic loanwords in Hungarian* (Turcologica 84). Wiesbaden: Harrassowitz.

- RÓNA-TAS, ANDRÁS. 1988. Turkic influence on the Uralic languages. In Sinor, Denis (ed.), *The Uralic languages: Description, history and foreign influences* (Handbuch der Orientalistik 8: Handbook of Uralic studies 1), 742–780. Leiden: E. J. Brill.
- RÓNA-TAS, ANDRÁS. 2017. Vítás magyar etimológiák: Válasz Honti Lászlónak. *Nyelvtudományi Közlemények* 113. 37–84.
- SAMMALLAHTI, PEKKA. 1988. Historical phonology of the Uralic languages with special reference to Samoyed, Ugric and Permic. In Sinor, Denis (ed.), *The Uralic languages: Description, history and foreign influences* (Handbuch der Orientalistik 8: Handbook of Uralic Studies 1), 478–554. Leiden: E. J. Brill.
- SÁNDOR, KLÁRA. 2021. A török–magyar nyelvi kapcsolatok újraértelmezésének lehetőségei. In Klima, László & Türk, Attila (eds.), *Párhuzamos történetek: Interdiszciplináris östörténeti konferencia a PPKE Régészettudományi Intézetének szervezésében Budapest, 2020. november 11–13.* (Studia ad Archaeologiam Pazmaniensia 23), 77–101. Budapest: Pázmány Péter Katolikus Egyetem & Martin Optiz Kiadó & Bölcsészettudományi Kutatóközpont, Magyar Östörténeti Kutatócsoport.
- SEEBOLD, ELMAR. 2019. Arme Waise mit reichem Erbe. *Amsterdamer Beiträge zur älteren Germanistik* 79. 151–155.
- SES = *Suomen etymologinen sanakirja* (Kotimaisten kielten keskuksen verkkojulkaisuja 72). Helsinki: Kotimaisten kielten keskus. URN:NBN:fi:kotus-202259. Updated 2023-12-05. <https://kaino.kotus.fi/suomenetymologinensanakirja> (Accessed 2024-06-27).
- SIMON, ZSOLT. 2020. Urindogermanische Lehnwörter in den uralischen und finno-ugrischen Grundsprachen: Eine Fata Morgana? *Indogermanische Forschungen* 125. 239–265.
- SSA = ITKONEN, ERKKI & KULONEN, ULLA-MAIJA (eds.). 1992–2000. *Suomen sanojen alkuperä: Etymologinen sanakirja 1–3*. Helsinki: Suomalaisen Kirjallisuuden Seura & Kotimaisten kielten tutkimuskeskus.
- SZEMERÉNYI, OSWALD. 1981. „Südwestiranische“ Lehnwörter in Ungarischen und Türkischen. *Études Finno-Ougriennes* 13. 229–246.
- ÚESZ = GERSTNER, KÁROLY. 2022. *Új magyar etimológiai szótár* (digital publication, under construction). Budapest: Nyelvtudományi kutatóközpont. uesz.nytud.hu
- UEW = RÉDEI, KÁROLY. 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.
- ÚMTSz = B. LÓRINCZY, ÉVA & HOSSZÚ, FERENC. 1979–2010. *Új magyar tájszótár I–V*. Budapest: Akadémiai Kiadó.
- VAMBÉRY, ÁRMIN. 1870. Magyar és török–tatár szóegyezések. *Nyelvtudományi Közlemények* 8. 109–189.

- ZHIVLOV 2006 = Живлов, Михаил. 2006. *Реконструкция праобско-угорского вокализма*. Moscow: Russian State University for the Humanities. (Unpublished doctoral dissertation.)
- ZHIVLOV, MIKHAIL. 2023. Reconstruction of Proto-Uralic. In Abondolo, Daniel & Valijärvi, Riitta-Liisa (eds.), *The Uralic languages*, 117–175. 2nd edn. London & New York: Routledge.
- ZIMMER, STEFAN. 2020. Celtic, Germanic and HARIGASTI TEIWA. In Markey, Thomas L. & Repanšek, Luka (eds.), *Revisiting dispersions: Celtic and Germanic ca. 400 BC – ca. 400 AD: Proceedings of the International Interdisciplinary Conference Held at Dolenjski Muzej, Novo Mesto, Slovenia, October 12–14, 2018* (Journal of Indo-European Studies Monograph Series 67), 48–80. Washington, D.C.: Institute of the Study of Man.