

# Notes on the reflexes of PU sibilants in Hungarian

Sampsa Holopainen

sampsa.holopainen@helsinki.fi

University of Helsinki

CIFU Precongress, 17.8.2020

[correction added 21.8.2020]

The background of the  
problem

# Background

- The Ugric/”Ugro-Samoyed” innovations: PU \*ś > s; \*s, \*š > \*θ

PU \*s, \*š >

Khanty \*λ > λ, l, j, t

Mansi t

Samoyed t

Hung. ø

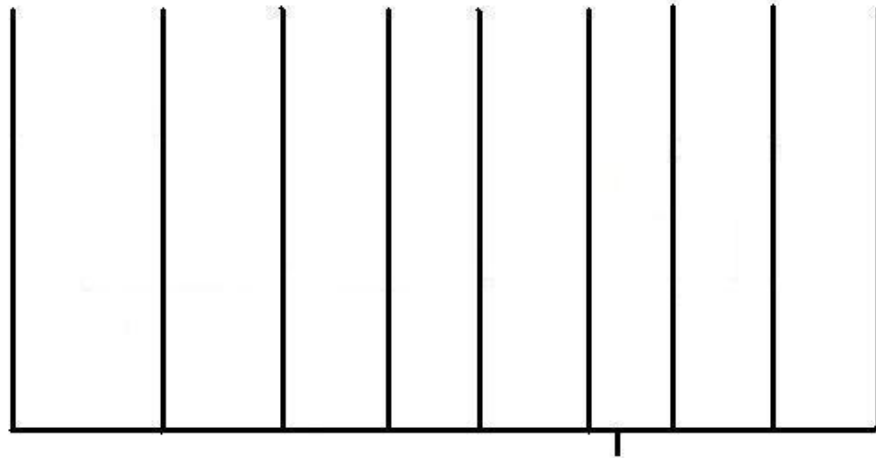
PU \*ś (ć?) >

s

- Areal innovation (Ug+Sam+Yeniseian Pumpokol)? Helimski 1982: 124; Salminen 2002; Blažek 2016: 121
- Shared innovation by Ugric and Samoyed? (Häkkinen 2007, 2009)
  - Recent criticism (Zhivlov 2018a): in Kh and Ms happened at different times, evidence from Ilr. Ilr. Loans (see also Holopainen 2019)
- Connected to the bigger problems of Ugric taxonomy and historical phonology: see Honti 1998, 1999, 2017; Salminen 2002 etc.

# Some family-trees

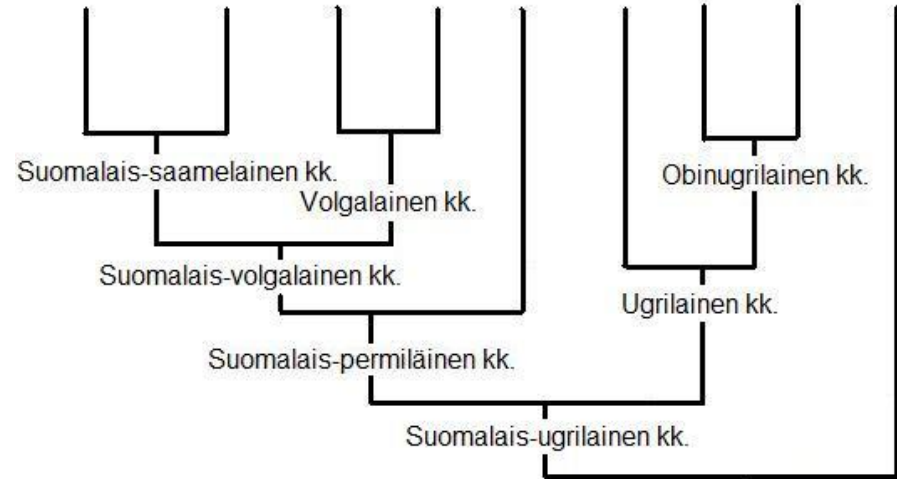
Saame Itämerensuomi Mordva Mari Permi Unkari Mansi Hanti Samoiedi



Uralilainen kantakieli

*Uralilainen sukupuu Tapani Salmisen mukaan (1999).*

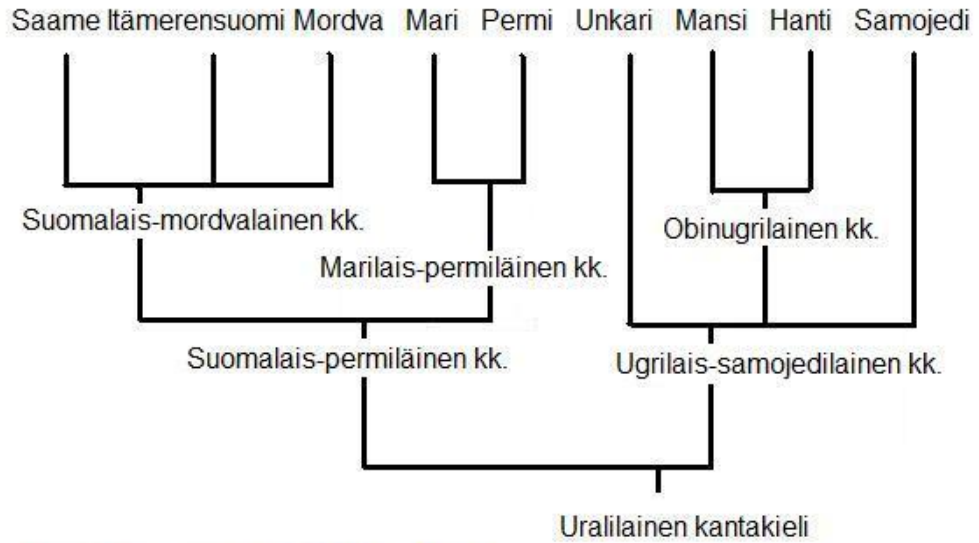
Saame Itämerensuomi Mordva Mari Permi Unkari Mansi Hanti Samoiedi



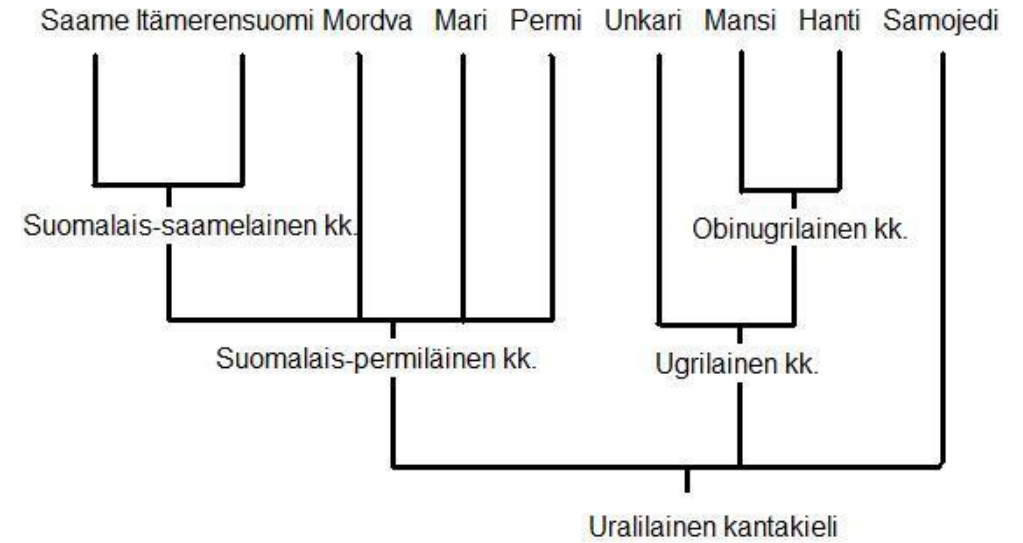
*Uralilainen sukupuu Otto Donnerin mukaan (1879; ilman samojedia)*

Uralilainen kantakieli

# Some family-trees



*Uralilainen sukupuu Jaakko Häkkinen mukaan (2007).*



*Uralilainen sukupuu Ulla-Maija Kulosen mukaan (2002).*

# Background

- Irregular developments in all the branches of Ugric, as well as Samoyed (?)
  - Honti (1983, 1986, 1999: 124–33, 2017: 174–, 2019), Tálós (1982)
- Ob-Ugric:
  - Mansi (see Tálós 1982: 90):
  - MsN *sāt* < PMs \**sātə* '7' (Zhivlov 2006) < ? PUg \**ṡäptä* ← (l)Ir \**sapta* (Honti 1999); parallel loans (Zhivlov 2018a; Holopainen 2019)
  - MsS *säj* 'pus' < ? PU \**säji*, > ? Hu *ev*, *év* (Honti 1999, 2019: irregular vocalism according to Aikio (2012: 229) - > perhaps not real cognates.
  - Khanty: \**s* retained if \**ś* follows, numerous examples:
  - PU *sükśi* 'autumn' > PKh \**süwəs* (Zhivlov 2006) > KhV *söŷəs*; PUg (?) \**ṡasV-* 'dry' > Kh O *säsǎ-*, V *sös-* etc.
  - Has been explained as a sound-law (Honti 2019: 84–85; see also Honti 2017: 171–173): first \**s* > *ṡ*, then assimilation by word-internal \**s*
  - The notion of regularity is convincing, but **more probable that it happened early**

# Background

- Conclusions regarding the Ob-Ugric situation:
- Khanty  $*s\text{--}\acute{s} > *s\text{--}s$  is a sound-law, but it is more natural that this happened early in Khanty, *before*  $*s$  became  $\vartheta$
- The irregular examples in Mansi are so few that it is difficult to doubt the regularity of the changes  $*s > *\vartheta > t$ ,  $*\acute{s} > s$

# Background

- Irregularities in Samoyed?
  - In general, the changes  $*ś > s$ ,  $*s > t$  seem to have been regular in Samoyed
- Mikola 2004: 29–36 deems some cases irregular and offers alternative explanations (a third PU affricate  $*č$ ):
  - TN *śer* < PU ?  $*śerV$ ; TN *śexə* < ? PU  $*śäkerV-$ , vrt. SaN *čiegar*; TN *śij* < PU  $*śüδ'i$  'coal'; Selk *šara* 'hard' ~ SaN *čares* 'coarse' <  $*čarV$ ; En *śuso* 'breast' < ? PU  $*śünši$ ; Kamas *šē* < ? PU  $*če$
  - These exceptions are not very convincing, as all of the etymologies include various phonological problems (see Holopainen 2019: 25, footnote 3)
- Recently Zhivlov 2018b and Kümmel 2019 have noted that  $*č$  was still retained as an affricate in Proto-Samoyed -> evidence for later change of  $*č > s$  in Samoyed (= no Ugro-Samoyed change)



# Background

- Uralic/Finno-Ugric \*ś vs. \*ć
- Two different phonemes in traditional reconstructions (MSzFE, UEW)
- Sammallahti 1988: \*ć appears only in Proto-Finno-Ugric, no Proto-Uralic examples (with PFU-Samoyed distribution); in PU, only one affricate \*c (= \*č)
- Zhivlov 2014: PU (= PFU) \*ś = \*ć; no minimal pairs?
- Hungarian is one of the languages that show different reflexes of \*ś (> sz) and \*ć (> cs, s) according to traditional reconstructions
- Besides PU/PFU, \*ć > cs, s occurs also in Ugric vocabulary (UEW)

# Background

- The tradition of *sporadic* sound-changes in Uralic etymological research has allowed exceptions to sound-laws
  - This has been heavily criticized lately: Ponaryadov 2012, Aikio 2013, Zhivlov 2014
- Some Hungarian words that show aberrant sibilants/affricates have been simply accepted as examples of sporadic sound-change
- In this presentation, these issues are approached through more strict approach

# The Hungarian situation

# Uralic/Finno-Ugric material of UEW/MSzFE

- Contains many dubious etymologies with \*ć (many marked with ? in UEW)
- **Most can be explained as false due to other irregularities involved:**
  - \*ćappV > *csap* 'hit' (irregular vocalism in cognates, esp. Permic, Ko *ćapkj*) UEW: 29
  - \*ćiŋkä > *seg, ség* 'Hügel' (irregular vocalism both within and outside Ug) UEW: 37
  - \*ćorV- > *csorog* (irregular, Hu probably onomatopoeitic) UEW: 40
  - \*ćukkv- > *csukorod* 'sich zusammenziehen' (completely irregular vocalism) UEW: 42
  - \*ćukkv-rV- > *csokor* 'bunch' (only in Hu and P, Ko *ćukęr*, perhaps a loan from P to Hu?) UEW: 43
  - \*ćumpV- > *csobolyó, csoborló* etc. 'Handlägel' (the variation within Hu makes PU unlikely) UEW: 43
  - \*ćuppV > *csúp* 'top' (the Fi *suippu*, Saa L *tjuhppa* cognates irregular) UEW: 44
  - \*ćuppV > *csupor* 'Töpfchen' (only in Hu and P) UEW: 45
  - \*ćuttV > *csont* 'bone' (Hu nasal can't be explained from PU) UEW: 45
  - \*ćVŋV > *csüd, csüg* 'limb' (Hu variation unexplained, P cognate means 'finger') UEW: 49
  - \*ćVppv > *csëpp* 'drop' (even UEW considers possibly onomatopoeitic) UEW: 49
  - \*ćVćV > *csúcs* 'point' (only in Md and Hu, the vocalism cannot be reconstructed)

## PU \*ć > Hu z

- PU ? \**pećV* > *fűz-fa* 'willow', Ko *paća*, Ud *pućj* (UEW: 367)
- An **irregular etymology** with limited distribution
- The relationship between the Hu and P words can be borrowing at best
- PU \**ipsi* 'smell' > *íz* 'taste', Mari *üpš*, Ms *āt*, *at* 'smell' etc. (UEW: 83–4)  
4)
- The Hungarian word might be an **Iranian (Alanic) loan** from a form akin to Ossetic *ad / adæ* 'taste' (Sköld 1925: 19; Helimski 2002), although this explanation has its problems (should be very early loan)
- Note that the cognates clearly point to \**s*, although UEW reconstructs a variant with \**ś*, too

# PU \*ć > Hu z

- ?? PU \**äsä-wäškä* > Hu *ezüst* 'silver', Ud *azveś* 'silver' (Honti 2017: 81–95, 169–170; regarding Permic, see Häkkinen also 2009: 25–28)
- Hu more probably borrowed from Alanic, Oss *æzvīst / æzvæstæ* (Helimski 2002; see Viredaz 2017 for discussion of the Ossetic word)
- PU \**sićV* > *szűz* 'virgin; pure', Ud *śuzal-* 'to clean' (UEW: 441)
- Turkic etymology for Hu has been suggested: West Old Turkic \**süzök* 'clarified' (WOT 833–836: Turkic uncertain, rather from Uralic?)
- Even if the Turkic etymology remains uncertain, the Uralic set of cognates is is not very convincing
- Indo-Iranian etymology (← PIIr \**ćwaitá-* > OI *śvetá-* 'white') by Katz (2003) is likewise problematic regarding Hu vocalism

PU \*ć > Hu s

- PU \*śänki > Hu *segg* 'ass', Ms P *sāṅk* 'loin', Mari *šenyel* 'back-side' (UEW: 472)
- The Mari word has dubious vocalism: *e* in all varieties except *ə* in Hill Mari > very few parallels, none convincing (Aikio 2014)
  - (Note that Aikio UED draft explains otherwise)
- In the light of this, the Hu-Ms word family could be a Ugric item > in Ugric vocabulary, \*ć > Hu s occurs more widely

# PU \*ć > Hu s

- PU \**ejćä* (?) > *õs* 'ancestor', Fi *isä* 'father' etc. (UEW:78 )
  - Irregularities in practically all branches
  - (Note the possible loan-etymology and other problems: Holopainen 2019: 97–100; Aikio UED draft)
- PU \**ec(k)ä*- > *es-ik* 'fall', Ko *uš*- etc. (UEW: 71)
- Sammallahti 1988: \**ec̣kä*
  - This would explain Hu s regularly
  - No trace of *k* in Permic (?)
- PU \**kaća* > *hõs* 'hero', Selkup (T) *kêêčə* 'worker, servant' etc. (UEW: 110 )
- The Selkup cognates irregular, Hu unlikely to be a reflex of a PU word (Ante Aikio, comment in Academia.edu, 2020)



PU \*ć > Hu s

- \*ćiklä > süly 'scurvy', Fi syylä etc. (UEW: 36)
- The Hu word is rather from West Old Turkic \*šiyül < PT \*sigöl 'wart' (WOT: 747–49)

PU \*ć > Hu cs

- \**ćolmi* > *csomó* 'knot', Fi *solmu*, SaN *čuolbma* etc. (UEW: 38)
- Listed by Zhivlov (2014), who does not comment the irregular (?) *cs*
- Otherwise a promising etymology; difficult to deny the PU status of this word

# PU \*ć > Hu cs

- \*pućV-rV-t- (?) > *facsar* 'squeeze', Fi *puserta-* etc. (UEW: 397; Aikio 2013)
- The example seems convincing, but has its problems:
  - Hu *a* < \**u* is very rare (perhaps regular after labials and \**k*, though: \**mučV* > *hagy-máz*; \**kulki-* > *halad* 'wander'?), as many alleged examples can be reconstructed without \**u*: \**pawi* rather than \**puwi* > *fa* 'tree' etc.
  - UEW (407) also gives a different root with similar meaning: \**puńćV-* ~ \**pućV-* > Ko *pićki-*; could Hu cs reflect \**ćk* here too? (But why not *s*?)
- Last resort: affective vocabulary?

# Clusters with \*ć in Hungarian

- \*ńć > *gy* is regular in Hungarian (note that \*ś/ć remained \*ć in PKhanty and PMansi too – no depalatalization in this environment; see Zhivlov 2018a)
  - Examples: \**lońća* > *lágy* 'soft', \**läńći* > *légy* 'fly' etc.
  - Also in loanwords: *agyar* 'tusk' ← Ilr \**anćūra*- (Cheung 2002)
  - The apparent examples to this rule are not very strong: *hárs*, *hangya*
- PU \*ćk (? śk) cluster must have remained in Proto-Ugric (?) too:
  - \**mośki* > *mos* 'wash' (UEW: 289); \**čaćkV* (?) > *sas* 'eagle' (UEW: 51); \**paćka* > *fos* (UEW: 396–7)
  - Few exceptions: *fěcske* (< *fěcs-ke*) < \**päćkV*: note that this etymology is criticized by EWUng; dial. *fěs-ke* shows the regular development
    - The Uralic cognate set is not entirely regular, however

# Clusters with \*ć in Hungarian

- A troubling example:
- *hagyma* < \**kaćmi* 'onion'
  - PP \**komiž*, PMs \**kāćam* (UEW: 164–65, Zhivlov 2014: 130)
- The Permic form shows metathesis (Zhivlov *ibid.*)
- How to explain the Hu and Ms forms?
  - In Hu, \**mć* > *gy* (compare \**ńć* > *gy*)? However, *m* remains (and why not > *v*?)
  - Voicing in the cluster \**ćm*? The cluster explains the retained *m*?
  - Mansi irregular? Note that \**ć* occurs in many inherited PUg words in Mansi
- Last resort: a central Uralic substrate word? A plant name with very limited distribution...

# Clusters with \*ć in Hungarian

- \**wenći/wänći* > *vés* 'carve', Fi *veitsi* 'knife' etc. (UEW: 565)
- Seems convincing: however, Hu a verb, the rest of the cognates nouns
- Perhaps \**ηć* > *s* regular?
  - No parallels

PU \**s* > Hu *s*

- \**wäsV-* > *vész* 'lose', Ud *vez-* (UEW: 566; Honti 2017)
- Due to the completely irregular vocalism, this cognate should be rejected

# PU \*s > Hu s

- \**pesä* > *fészek* 'nest'
- Honti's explanation: \**f* strident, causes \**ʃ* resort back to *s*
  - OK, but why \**ʃ* > \**s* first? Could \**s* not have been simply retained here?
- \**s* regular reflex of \**s* in word-internal position?
- No counter-examples? (but clusters with \**s* show the regular loss, \**miksa* > *má-j* 'liver')
  - \**s* lost in the reflexes of PU 3p. possessive suffixes (\**sA* etc.)
  - Helimski 1999: Hu *le-* (verbal prefix) and *leg-* (prefix in superlatives) from PU \**läs* 'near'; "Auslaut" according to Helimski, but the PU stem **must** have been \**läsi*!



PU \**s* > Hu *z*

- ?? PU \**äsä-wäskä* > Hu *ezüst* 'silver', Ud *azveś* 'silver' (see Häkkinen 2009: 25–28)
- More probably borrowed from Alanic, Oss *æzvīst* / *æzvæstæ* (most recently Viredaz 2017; see also Helimski 2002)

# Notes on the Ugric sibilants in Hungarian

# Ugric material of UEW/MSzFE

- The Ugric material features several irregular etymologies: however, their analysis shows that most are irregular words that can be loans or substrate, some rather accidental similarities
- In some cases, Hu shows  $s < *ć$ , whereas Khanty or Mansi show  $s$  that points to "normal"  $*ś$

# Ugric cases of \*ć > s

- ? \*kaćV- > *hasad* 'split', MsKu *elχaśl-* etc. (UEW: 854)
  - More convincing explanation: Metsäranta 2017: 220; PU \*kačka- 'bite' > Fi *katke-*
- \*ke(ń)ćä > ? *keshed* 'to become narrow', *kesküny* 'narrow'
  - The only example of \*ńć > s; \*ńć has to be reconstructed for the ObUg forms
- \*kećä- > *kiser* 'follow', KhV *köt'* (UEW: 856)
- ? \*kVćV 'bitter' > ? *keserű*, MsP *kwaśert-*; irregular vocalism? (UEW: 861)
- \*läćV > *les* 'lauern, nachstellen', MsTJ *läć-*, Kh N *lāši* (UEW: 863)
- \*muśV > *mosolyog* (UEW: 872)
  - Mansi cognates P *mus-* etc. point regularly to \*ś (not \*ć)
- \*wäćä 'penis' > *vese* 'kidney', MsP *wäśəγ* etc. (UEW: 899–90)
  - Phonologically unproblematic etymology

# Ugric cases of \*ć > cs?

- \*ćukkv- > *csókol* 'to kiss' (UEW: 838)
- The only convincing example?
- Affection?

# Other reflexes of Ugric \*ć?

- \*? \**kVśV* > *küzd, küszöd*
  - Assimilation due to *d*? *küszöd* could manifest the regular outcome
- ? \**kućV* > *-hoz* (UEW: 857–58)
- \**lVćV* > *laza* 'loose'
- \**mVćV* > ? *mez-telen* 'naked'
  - Forms with *sz* attested in dialects
- ? \**ʃjćV* > *ázik* 'become wet', KhV *lǎstǎ-* (!)
- \**kVlVćV* > *köles* 'Hirse', MsSo *kolas* 'a kind of flour'
  - UEW considers possible loan
  - MsSo sibilant unclear

# Notes on the Ugric data

- Much more research needed to the poorly understood vowel-developments
- However, it seems that \*ć can probably be reconstructed to a pre-form for some Hu words that have Ugric connections
- Many of the alleged Proto-Ugric words are, however, probably not real Ugric cognates but parallel loans/substrate words
- The three Ugric languages/branches sometimes offer contradictory evidence regarding \*ś/\*ć

# Conclusions

- 1) Most of the PU/PFU etymologies (of UEW, MSzFE) that allegedly manifest PU  $*\acute{s} > \text{Hu } cs, s, z$  turn out to **be false cognates or later loans**
- 2) There are many words with **Ugric** distribution that show  $*\acute{c} > cs$  in Hungarian, also some possible examples of  $*\acute{c} > s$  and  $*\acute{c} > z$ 
  - **However, many of these words show other irregularities (esp. in the vocalism) > the words not real cognates; parallel loans (some from Ir., like *ostor*), or unknown conditioned changes?**
  - (Nobody likes loans from unknown [substrate] languages, but...)
  - The situation is rather **similar** as with **other Ugric sound-changes**, like  $*\eta > * \eta k$  or  $> (?) * \gamma$
  - More research on **Ugric** vocabulary needed!



# Conclusions

- 3) Some otherwise regular (?) PU cases with suspicious sibilants/affricates remain:
  - *csomó* < \**ćolmi*
  - *fészek* < \**pesä*
  - *hagyma* < \**kaćmi* (note that also PMansi has \**ć*)
  - *vés* < \**wäŋći*
- 4) Due to the very **limited evidence**, it is a matter of **interpretation** whether the situation in Hungarian gives evidence for

# References

- Aikio 2012: On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic \*x... *Per Urales ad Orientem (FS Janhunen)*. 227–250.
- 2013: Uralilaisen kantakielen vokaalistosta. Etymologia ja kielihistoria: Erkki Itkosen ja Aulis J. Joen 100-vuotisjuhlaseminaari 19.4.2013 (handout). Helsinki.
- 2014: On the reconstruction of Proto-Mari vocalism. – *Journal of language relationship* 11: 125–157.
- Aikio UED draft = *Uralic etymological dictionary* (draft version of entries A–Ć: [https://www.academia.edu/41659514/URALIC\\_ETYMOLOGICAL\\_DICTIONARY\\_draft\\_version\\_of\\_entries\\_A\\_%C4%86\\_](https://www.academia.edu/41659514/URALIC_ETYMOLOGICAL_DICTIONARY_draft_version_of_entries_A_%C4%86_) )
- Blažek, Václav 2016: On the classification of the Samoyedic languages. *FUF* 63: 79–125.
- Helimski, Eugen 1999: PU \*lās 'bei, in der Nähe'im Samojedischen (und im ungarischen?). *FUM* 21/22: 77–81.
- 2002: Zu den iranischen Lehnwörtern im Ungarischen. *Finno-Ugrians and Indo-Europeans: Linguistic and Literary contacts*. 105–111.
- Holopainen, Sampsa 2019: *Indo-Iranian borrowings in Uralic*.
- Honti, László 1982: *Geschichte des obugrischen Vokalismus der ersten Silbe*.
- 1999: *Az obi-ugor konzonantizmus történte*.
- 2017: *A magyar és a nyugati ótörök szókészleti kapcsolatairól*.
- 2019: Zur Frage nach den dentalen Spiranten in den Vorgängern der uralischen Sprachen. *Folia Uralica Debreceniensia* 26.
- Katz, Hartmut 2003: Studien zu den älteren indoiranischen Lehnwörtern in den uralischen Sprachen.
- Kümmel, Martin 2019: Sound change typology and the reconstruction of sound changes in Uralic [presentation, University of Helsinki, 16.12.2020]

# References

- MSZFE = Lakó, György 1971–1978: *A magyar szókészlet finnugor elemei*.
- Ponaryadov, Vadim 2012: Насколько достоверны современные реконструкции прафинно-пермского вокализма? – *Вопросы финно-угорской филологии* 3 (8): 78–91.
- Salminen, Tapani 2002: Problems in the taxonomy of the Uralic languages in the light of modern comparative studies. *Лингвистический беспредел: сборник статей к 70-летию А. И. Кузнецовой*. Москва: Издательство Московского университета. 44–55.
- Sammallahti, Pekka 1988: Historical phonology of the Uralic languages. In Sinor (ed.): *Uralic languages*. 478–554.
- Sköld, Hannes 1925: *Die ossetische Lehnwörter in Ungarischen*. Lund: Gleerup. Leipzig: Otto Harrassowitz.
- Tálos, Endre 1984: Vogul + Osztyák/2. *Nyelvtudományi Közlemények* 86. 89–99.
- UEW = Rédei, Karoly 1988: *Uralisches Etymologisches Wörterbuch*
- Viredaz, Remy 2017: Trois étymologies ossètes: rasyg, ævzist, ærxy. *Studia Iranica* 46: 31–48.
- WOT = Róna-Tas, András & Árpád Berta 2011: *West Old Turkic*
- Zhivlov, Mikhail 2006: *Реконструкция праобско-угорского вокализма*. Unpublished dissertation.
- 2014: Studies in Uralic vocalism III. *Journal of Language relationship* 12: 113–148.
- 2018a: Историческая фонетика и внутренняя классификация уральских языков. Presentation, *XIII традиционные чтения памяти С. А. Старостина*
- 2018b: Прасамодийский \*s - фонология и фонетика. Presentation, Семинар по сравнительно-исторической фонетике самодийских языков

Thank you and see you (live)  
next year!

