Notes on the reflexes of PU sibilants in Hungarian

Samps Holopainen
samps.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

<table>
<thead>
<tr>
<th>PU *ś (ć?) &gt;</th>
<th>Khanty</th>
<th>*λ &gt; λ, l, j, t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mansi</td>
<td>t</td>
</tr>
<tr>
<td></td>
<td>Samoyed</td>
<td>t</td>
</tr>
<tr>
<td></td>
<td>Hung.</td>
<td>∅</td>
</tr>
</tbody>
</table>

• 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

Uralilainen sukupuu Tapani Salmisen mukaan (1999).
Some family-trees

Urallainen sukupuu Jaakko Hakkisen mukaan (2007).

Background

• Irregular developments in all the branches of Ugric, as well as Samoyed (?)
  • Ob-Ugric:
    • Mansi (see Tálos 1982: 90):
    • MsN sāt < PMs *sātə ‘7’ (Zhivlov 2006) < ? P Ug *θäptä ← (I)r *sapta (Honti 1999); parallel loans (Zhivlov 2018a; Holopainen 2019)
    • Khanty: *s retained if *ś follows, numerous examples:
      • PU sūkśi ‘autumn’ > P Kh *sūwas (Zhivlov 2006) > KhV sōyas; P Ug (?) *θasV- ‘dry’ > Kh O sāsā-, V sōs- etc.
      • Has been explained as a sound-law (Honti 2019: 84–85; see also Honti 2017: 171–173): first *s > θ, 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–ś > *s–s is a sound-law, but it is more natural that this happened early in Khanty, before *s became θ
  • The irregular examples in Mansi are so few that it is difficult to doubt the regularity of the changes *s > *θ > t, *ś > 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 *ć):
  • TN śer” < PU ? *šerV; TN šexe < ? PU *šäkerV-, vrt. SaN čiegar; TN šij < PU *šüđ’i ‘coal’; Selk šara ‘hard’ ~ SaN čares ‘coarse’ < *ćarV; En šuso ’breast’ < ? PU *šünši; 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 *ć was still retained as an affricate in Proto-Samoyed -> evidence for later change of *ć > s in Samoyed (= no Ugro-Samoyed change)
Background

• Uralic/Finnno-Ugric *ś vs. *ć
• Two different phonemes in traditional reconstructions (MSzFE, UEW)
• Sammallahti 1988: *ć appears only in Proto-Finnno-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
  - *ćünkä > seg, ség ’Hügel’ (irregular vocalism both within and outside Ug) UEW: 37
  - *ćorV- > csorog (irregular, Hu probably onomatopoietic) UEW: 40
  - *ćukkV- > csukorod ’sich zusammendziehen’ (completely irregular vocalism) UEW: 42
  - *ćukkV-rV- > csokor ’bunch’ (only in Hu and P, Ko čukger, 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 suiippu, 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 onomatopoietic) 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 pućő (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 *ś, 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 æzvišt / æzvæstæ (Helimski 2002; see Viredaz 2017 for discussion of the Ossetic word)
- PU *sičV > szűz ‘virgin; pure’, Ud śuʒal- ‘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 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 *śäŋki > Hu segg 'ass', Ms P sāŋk 'loin', Mari šenyle 'back-side' (UEW: 472)

- The Mari word has dubious vocalism: e in all varieties except ǝ 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 (?) > ős ‘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 *eć(k)ä- > es-ik ’fall’, Ko uś- etc. (UEW: 71)
- Sammallahti 1988: *ećkä
  - This would explain Hu s regularly
  - No trace of k in Permic (?)
- PU *kaća > hős ‘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 *č > Hu s

• *čiklă > sūly ’scurvy’, Fi syylä etc. (UEW: 36)
• The Hu word is rather from West Old Turkic *šiyūl < 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: *mućV > 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’, *lănći > légy ’fly’ etc.
  - Also in loanwords: agyar ’tusk’ ← IIr *ancūra- (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:
  - *moški > mos ’wash’ (UEW: 289); *čaćkV (?) > sas ’eagle’ (UEW: 51); *paćka > fos (UEW: 396–7)
  - Few exceptions: fëcske (< fëcs-ke) < *päćkV: note that this etymology is criticized by EWUng; dial. fës-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 metathezis (Zhivlov ibid.)
• How to explain the Hu and Ms forms?
  • In Hu, *mć > gy (compare *ńć > gy)? However, m remains (and why not > n?)
  • 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

• *weŋći/wäŋći > vés ’carve’, Fi veitsi ’knife’ etc. (UEW: 565)
• Seems convincing: however, Hu a verb, the rest of the cognates nouns
• Perhaps *ŋć > s regular?
  • No parallels
PU *s > Hu s

• *wäšV- > 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 *∅ resort back to s
  • OK, but why *∅ > *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, *m̥ksa> 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-wäškä > 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 $s < ^*\acute{c}$, whereas Khanty or Mansi show $s$ that points to “normal” $^*\acute{s}$
Ugric cases of *ć > s

• ? *kaćV- > hasad ’split’, MsKu elčaśl- etc. (UEW: 854)
  • More convincing explanation: Metsäranta 2017: 220; PU *kačka- ’bite’ > Fi katke-
• *ke(ń)ćä > ? keshed ’to become narrow’, keskëny ’narrow’
  • The only example of *ńć > s; *ńć has to be reconstructed for the ObUg forms
• *kećä- > kísér ’follow’, KhV kőt’ (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 wäşwy 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 *č?

• *? *kVśV > küzd, küszöd
  • Assimilation due to d? küszöd could manifest the regular outcome
• ? *kućV > -hoz (UEW: 857–58)
• *IVćV > laza ’loose’
• *mVćV > ? mez-telen ’naked’
  • Forms with sz attested in dialects
• ? *đjćV > ázik ’become wet’, KhV lästə- (!)
• *kVIVćV > köles ’Hirse’, MsSo kolas ’a kind of flour’
  • UEW considers possibble 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 *š/*ć
Conclusions

• 1) Most of the PU/PFU etymologies (of UEW, MSzFE) that allegedly manifest PU *ś > Hu cs, s, z turn out to be false cognates or later loans
• 2) There are many words with Ugric distribution that show *ć > cs in Hungarian, also some possible examples of *ć > s and *ć > 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 *ŋ > *ŋk or > (?) *γ
   • 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 < *wæŋći

- 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.
- Aikio UED draft = Uralic etymological dictionary (draft version of entries A–Ć: https://www.academia.edu/41659514/URALICETYMOLOGICAL_DICTIONARY_draft_version_of_entries_A_%C4%86)
- 1999: Az obi-ugor konszonantizmus történte.
- 2017: A magyar és a nyugati ótörök szókészleti kapcsolatairól.
- Kümmel, Martin 2019: Sound change typology and the reconstruction of sound changes in Uralic [presentation, University of Helsinki, 16.12.2020]
References

- UEW = Rédei, Karoly 1988: Uralisches Etymologisches Wörterbuch
- WOT = Róna-Tas, András & Árpád Berta 2011: West Old Turkic
- 2018a: Историческая фонетика и внутренняя классификация уральских языков. Presentation, XIII традиционные чтения памяти С. А. Старостина
Thank you and see you (live) next year!