

Uralic evidence for the Indo-European homeland

Jaakko Häkkinen, 13th February 2012

There are linguistic evidence of Proto-Indo-European having contacts with different language families, but none other of these contact partners so far can match the Uralic language family in the accuracy: we can reconstruct several consequent layers of Indo-European loanwords in Proto-Uralic, while the loanwords between Proto-Indo-European and Kartvelian or Semitic are somewhat more indefinite or vague by their stratification.

1. Late Proto-Indo-European / Early Archaic Indo-European loanwords to Proto-Uralic

These loanwords have been traditionally seen to have occurred between Proto-Indo-European and Proto-Uralic, as these stages were considered as contemporaneous (~ 4000 BC). However, new linguistic results of the present millennium show that Proto-Uralic seems to be even 2 000 years younger than was supposed (Kallio 2006; Häkkinen 2009). At this time-depth we cannot speak about Proto-Indo-European anymore, but rather about Archaic Indo-European. Phonologically we can still distinguish at least two separate stages, Early and Late Archaic Indo-European: the one and the same Indo-European word has been borrowed with two different substitutions, and as the distribution of the words does not follow the boundary of the Uralic protodialects, the only possible interpretation seems to be to suppose two temporally differing loanword layers:

- a) IE $*b^heh_1-(ye/o-)$ → U $*pexi-$ ‘to cook’ (Koivulehto 1991)
- b) IE $*b^heh_1-(ye/o-)$ → U $*peša-$ ‘to cook’ (Koivulehto 1991)

The first layer seems to be older, as the Early Proto-Uralic $*x$ (most probably voiceless velar spirant $[x]$) has weakened already before Late Proto-Uralic to $*\gamma$ (voiced velar spirant $[ɣ]$), being later lost in Finnic, Mari, Permian and Samoyed, while remaining as $*\gamma$ in Mansi and Khanty, ν in Hungarian in certain environment, and as j in Mordvin in certain environment, as well as $*k$ in Saami. After the change $*x > *\gamma$ there was a need for more fricative substitute for the Indo-European laryngeals, and at this later stage $*š [ʃ]$ was chosen.

It is for long known that $*x$ was an older substitute for laryngeals than $*š$, but the distinction which was seen between Proto-Uralic and Proto-Finno-Ugric (Koivulehto 1991) can now be seen between Early and Late Proto-Uralic. Phonologically Samoyed seems to derive from the very same East Uralic dialect than the Ugric branches (Hungarian, Mansi and Khanty), which makes Proto-Finno-Ugric obsolete.

From Proto-Uralic to East Uralic (Häkkinen 2007: 71–76)

1. $*š > *s$ (coalescence with original $*s$)
2. $*s > *L$ (both original $*š$ and $*s$ change to voiceless fricopalateral $[ʎ]$)
3. $*ś > *s$ (secondary $*s$ occurs)
4. $*e > *j \sim *e$ (sporadic split; conditions not known)
5. $*k, *w > *\gamma$ (coalescence with original $*\gamma < *x$ between vowels)
6. $*Sk > *\gamma S$ (sibilant metathesis in some obstruent clusters)

There are old Indo-European loanwords proving that the East Uralic sibilants indeed are innovations, while West-Central Uralic (> Saami, Finnic, Mordvin; Mari, Permic) shows the original sibilant (Häkkinen 2009: 21):

Evidence from Proto-Aryan loanwords:

Hungarian *száz* ~ Mansi KM *seēt* ~ Khanty V *sàt* '100'
 < East Uralic **sęta*
 < Proto-Uralic **śęta* (> Mordvin *śado*)
 ← Proto-Aryan **ćata-* / Proto-Indo-Aryan **śata-* '100'

Mansi KM *uutər* 'lord, prince; hero'
 < East Uralic **aLəra*
 < Proto-Uralic **asira* (> Mordvin *azoro*)
 ← Proto-Aryan **asura* > Iranian *ahura* 'lord'

In the Early Archaic Indo-European layer we can list the following words (mainly picked from Johan Schalin's Internet vocabulary):

IE **b^heh₁-* → U **pexi-* 'to cook' (Koivulehto 1991)

IE **keh₂w-* → U **kaxi-* 'to fell' (Koivulehto 2009)

IE **h₂meiH-* → U **mexi-* 'to give, sell' (Koivulehto 1991)

IE **g^wneh₂-* → U **näxi* 'female' (Koivulehto 1991)

IE **bhuH-* → U **puxi* 'tree' (Koivulehto 1991)

IE **deh₃-* → U **toxi-* 'to bring' (Koivulehto 1991)

IE **weg^h-* → U **wixi-/wêjxi-* 'to carry' (Koivulehto 1991)

When we cannot apply the laryngeal criterion, it is difficult to date a loanword. Temporal range for such Archaic Indo-European loanwords varies from the late 4th millennium BC (Late Proto-Indo-European) to the early 2nd millennium BC (Late Northwest Indo-European) (see Mallory 2001; Kallio 2006).

2. Early and Middle Proto-Aryan loanwords to Proto-Uralic

This layer is about as early as the previous one, as the Early Proto-Uralic **x* was still in use, although not for a substitute for a laryngeal but aspirated palatalized velar consonant. Nevertheless we can hardly take the Early Proto-Aryan layer earlier than the early 3rd millennium BC, but we could in theory take the Archaic Indo-European layer to the late 4th millennium BC.

EPA **g^ug^hew-* → EPU **juxi-* 'to drink' (Koivulehto 1991)
 < IE **g^hew-* 'to pour'

EPA **ǵeǵ^hē-* → EPU **jüxi-* ‘to stay’ (Koivulehto 1999)
< IE **ǵ^heh₁-* ‘to leave’

EPA **kekro-* → EPU **kekra/i* ‘annual cycle’ (Carpelan & Parpola 2001)
< IE **k^wek^wlo-* ‘wheel’

These words show distinctively Aryan features like reduplication or the change **l* > **r*, although they precede the vowel changes (**e*, **o* and syllabic resonant > **a*) and the palatalization of velars before *i* and *e* (**k^w*/**k* > **č*, **g^w*/**g* > **dž*). Additionally a meaning of a word may be closest to the Aryan branch. At this stage Proto-Aryan was already dividing into two entities, Proto-Iranian and Proto-Indo-Aryan, as can be seen from the words which show already the Iranian depalatalization (**ǵ* >) **dž* > **dz* but not yet the vowel changes. Aryan palatalized affricate would have been substituted differently in Proto-Uralic: **dž* → **ś* word-initially.

MPA = EPIr. **dzen-* → MPU **sen-ti-* ‘to be born’ (Koivulehto 2001)
< IE **ǵenh₁-* ‘to be born’

MPA = EPIr. **dziwie-* → MPU **sewi-* ‘to eat’ (Koivulehto 2003)
< IE **ǵiewH-* ‘to chew’

As Early Proto-Uralic seems to have been contemporaneous with Early Proto-Aryan, and Late Proto-Uralic seems to have been contemporaneous with Late Proto-Aryan, we can call the language stages here Middle Proto-Uralic and Middle Proto-Aryan.

3. Late Archaic / Northwest Indo-European loanwords to Proto-Uralic

Here I only present a few words showing the substitute for laryngeal: **š* word-internally and **k* word-initially.

IE **b^heh₁-* → U **peša-* ‘to cook’ (Koivulehto 1991)

IE **ǵnh₁-e/o-* → LPU **inši-* ‘human’ (Koivulehto 1991)

IE **ǵnh₃-me-* → LPU **inš-mi* → **išmi* ‘wonder, sign’ (Koivulehto 1991)

IE **h₂eleuio-* → LPU **kali-ja* ‘(weak) beer’ (Kallio 1998)

IE **h₂esg-* → LPU **kaski* ‘burnt-over clearing’ (Koivulehto 1991)

IE **h₁es-en-* → LPU **kesa* ‘summer’ (Koivulehto 1991)

4. Late Proto-Aryan loanwords to Proto-Uralic

This layer is contemporaneous with the Late Archaic Indo-European layer, because here are no traces of the Early Proto-Uralic *x, even if there are phonetically suitable sounds for such a substitution: instead we see Late Proto-Uralic *w.

LPA *argha- → LPU *arwa ‘value, price’ (Joki 1973)

Proto-Uralic was earlier considered older than Late Proto-Aryan, but now that a word does not require a cognate in Samoyed to be counted as Proto-Uralic, we can re-interpret a few Late Proto-Aryan loanwords in Late Proto-Uralic. Also early loanwords from Aryan dialects (Proto-Iranian and Proto-Indic) are included, because their split is already a Middle Proto-Aryan phenomenon and therefore even distinctly dialectal features are not to be judged automatically post-Proto-Aryan.

Late Proto-Aryan loanwords are posterior to the change of *e, *o and syllabic resonants > *a. The following examples are from Häkkinen (2009); here are listed only words with wide enough regular distribution so that they can be counted as Proto-Uralic:

LPA *ćata- / IA *śata- → LPU *śeta ‘100’ (Joki 1973)
< IE *k̑ntó-m ‘100’

LPA *asura → LPU *asira ‘lord, prince; hero’ (Koivulehto 2001)
< IE *ḡsu- / *hensu- ‘god’

LPA *sáras → LPU *sara(s) ‘lake’ (Koivulehto 1999)
< IE *selos / *seles ‘lake, marsh’

LPA *ārā → LPU *ora ‘thorn’ (Koivulehto 2001)
< IE *ēlā ‘thorn’

LPA *aras- / *arah → LPU *woxri ‘mountain’ (Koivulehto 1999)
< IE *h₁er(H)os/es ‘mountain’

PIr. *zaranya → LPU *sērña ‘gold’ (Joki 1973)
< IE *ǵ^h(o)l(H)- ‘yellow; gold’

5. Conclusion

I stage (*x) 2800 BC–	Early Archaic (Proto- or Northwest?) IE borrowings to Early Proto-Uralic: IE * <i>b^heh₁-(ye/o-)</i> → EPU * <i>pexi-</i> ‘to cook’	Early Proto-Aryan borrowings to Early Proto-Uralic: EPA * <i>ǵuǵ^hew-</i> → EPU * <i>juxi-</i> ‘to drink’
II stage (no *x) 2300 BC–	Late Archaic (Northwest) IE borrowings to Late Proto-Uralic: IE * <i>b^heh₁-(ye/o-)</i> → LPU * <i>peša-</i> ‘to cook’	Late Proto-Aryan borrowings to Late Proto-Uralic: LPA * <i>argha-</i> → LPU * <i>arwa</i> ‘value, price’
III stage (no *x) 1800 BC–	Northwest IE borrowings to Uralic dialects: NwIE * <i>b^hleh₁-tó-</i> → WU * <i>lešti</i> ‘leaf, blade’	Proto-Iranian borrowings to Uralic dialects: (No examples of similar source consonants)

Proto-Uralic substitutes for the Indo-European laryngeal and velar aspirate: *x*-strata (Early Archaic Indo-European and Early Proto-Aryan), *š*-stratum (Late Archaic Indo-European) and *w*-stratum (Late Proto-Aryan).

What is the conclusion about the Uralic evidence on the location of Proto-Indo-European? If the oldest Indo-European loanwords are already Late Proto-Indo-European – which now seems to be the less probable option – the case would be clear: Proto-Uralic cannot in any case be located south from the taiga zone, and therefore Late Proto-Indo-European could not be very far. Southeast European steppe would then be the obvious solution (the Kurgan theory).

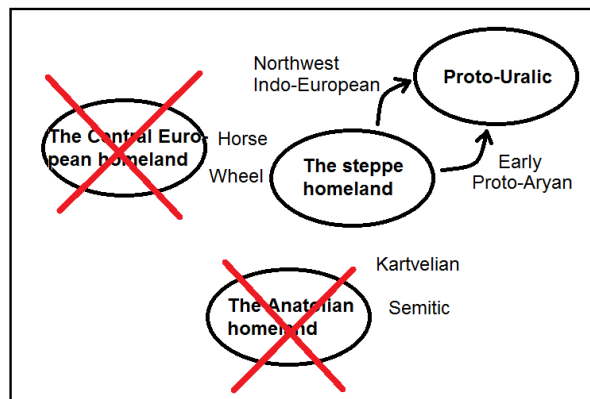
If, however, the oldest Archaic Indo-European loanwords are only contemporaneous with the Early Proto-Aryan loanwords – which now seems the most probable option – things get more complicated. First we must identify the donor language. On geographical reason it can hardly be the Graeco-Armenian or any other remote dialect, so we are left with either Tocharian or Northwest Indo-European. Tocharian is generally seen to split off very early from the Indo-European stock, and it is connected to the Afanasyevo Culture. The main problem with the Tocharian explanation is that it could not have been present in the Volga region anymore at the Early Proto-Aryan stage. On the other hand, it is possible that Pre-Proto-Uralic was spoken in Asia and met Tocharian there; but the Archaic Indo-European loanword layer was contemporaneous with the Early Proto-Aryan layer, and Early Proto-Aryan was not spoken in Asia. Tocharian explanation thus seems to be a dead end.

Northwest Indo-European, which is connected to the Corded Ware Culture, matches better both the temporal and spatial closeness to Early Proto-Aryan: in the first half of the 3rd millennium BC the Corded Ware Fatyanovo-Balanovo Culture and the steppe Poltavka Culture reached each other in the Mid-Volga area (Carpelan & Parpola 2001), which happens to be the Proto-Uralic homeland (Kallio 2006; Häkkinen 2009). This solution explains best the oldest Indo-European loanword layers in Uralic, the *x*-strata. In the same area we can locate the next oldest loanword layers, the Northwest Indo-European *š*-stratum and the Late Proto-Aryan *w*-stratum.

Consequently, we have a situation where we have two very close dialects of Proto-Indo-European spoken in adjacent areas in the easternmost Europe near the great Volga bend at the 3rd millennium BC. The farther in time and space we go from there, the more implausible is the solution concerning the Proto-Indo-European homeland. No language remains unchanged for millennia, least of all when spreading thousands of kilometers to new areas. It is most credible to derive these dialects from the homeland from less than 1 000 kilometers southwest and one millennium back in time (the Copper/Bronze Age steppe homeland). It would be very improbable indeed to derive them from more than 2 000 kilometers southwest, behind the Black Sea, and up to 4 000 years back in time (the Neolithic Anatolian homeland). So much we get from the Uralic anchor: *the Kurgan theory seems to be the only credible one.*

Caucasian and Semitic contacts – even if they were as accurately stratified as the Uralic contacts – do not require the Anatolian homeland, either: it does not matter whether the Proto-Indo-European was spoken north (Ukraine) or west (Anatolia) of these contact languages. However, they exclude the Central European homeland, as does the Uralic argument as well.

Neither can the fact, that Anatolian is seen as the first entity to split off from the Indo-European unity, prove that the homeland of Early Proto-Indo-European was in Anatolia – it could equally well have been in Ukraine, because no branch is stronger than another concerning the location. And indeed the Anatolian branch has been included in the palaeolinguistic analysis of Proto-Indo-European, and words denoting to concepts like ‘horse’ and ‘wheel’ (which are also found in Anatolian; see Mallory & Adams 2006: 139, 248) clearly point to the European origin: the oldest wheels are found in Europe, and the word for ‘wheel’ in Sumerian and Semitic seems to be an Indo-European loanword (see Parpola 2008). Admittedly, such words do not exclude the Central European homeland, but the contacts with Uralic, Kartvelian and Semitic do exclude it.



The Indian homeland is no longer maintained by anybody else than some fanatic indocentrists. It suffices to say that the continuum of consequent Aryan loanword layers (Early Proto-Aryan, Middle Proto-Aryan, Late Proto-Aryan) in Proto-Uralic make it practically impossible that the Aryan languages could be native in India. The development of the Aryan branch must be located in the Southeast European steppes, south of the Volga-Ural taiga zone (Carpelan & Parpola 2001).

Literature

Carpelan, Christian & Parpola, Asko 2001: Emergence, contacts and dispersal of Proto-Indo-European, Proto-Uralic and Proto-Aryan in archaeological perspective. Carpelan et al. (edited): *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations*, p. 55–150. Suomalais-Ugrilaisen Seuran Toimituksia 242. Helsinki.

Häkkinen, Jaakko 2007: *Kantaauralin murteutumisen vokaalivastaavuuksien valossa*. Pro gradu -työ, Helsingin yliopiston Suomalais-ugrilainen laitos. <https://oa.doria.fi/handle/10024/7044>

Häkkinen, Jaakko 2009: Kantaauralin ajoitus ja paikannus: perustelut puntarissa. – *Suomalais-Ugrilaisen Seuran Aikakauskirja* 92, p. 9–56. <http://www.sgr.fi/susa/92/hakkinen.pdf>

Joki, Aulis J. 1973: *Uralier und Indogermanen. Die älteren Berührungen zwischen den Uralischen und Indogermanischen Sprachen*. Suomalais-Ugrilaisen Seuran Toimituksia 151. Helsinki: Suomalais-Ugrilainen Seura.

Kallio, Petri 1998: Kalja. – *Sananjalka* 40, p. 87–94.

Kallio, Petri 2006: Suomen kantakielten absoluuttista kronologiaa. – *Virittäjä* 1 / 2006, p. 2–25. http://www.kotikielenseura.fi/virittaja/hakemistot/jutut/2006_2.pdf

Koivulehto, Jorma 1991: *Uralische Evidenz für die Laryngaltheorie. Österreichische Akademie der Wissenschaften. Philosophisch-Historische Klasse. Sitzungsberichte*, 566. Band. Wien.

Koivulehto, Jorma 1999: Varhaiset indoeurooppalaiskontaktit: aika ja paikka lainasanojen valossa. – Paul Fogelberg (ed.): *Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan*, p. 249–280. Bidrag till kännedom av Finlands natur och folk, 153. Helsinki: Societas Scientiarum Fennica.

Koivulehto, Jorma 2001: The Earliest Contacts between Indo-European and Uralic Speakers in the Light of Lexical Loans. Christian Carpelan, Asko Parpola & Petteri Koskikallio (editors): *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations*, p. 235–264. Mémoires de la Société Finno-Ougrienne 242. Helsinki: Finno-Ugrian Society.

Koivulehto, Jorma 2003: Frühe Kontakte zwischen Uralisch und Indogermanisch im nordwestindogermanischen Raum. Alfred Bammesberger & Theo Vennemann (editors): *Languages in Prehistoric Europe*. Heidelberg.

Koivulehto, Jorma 2009: Etymologisesti hämäriä *-(is)tA*-johdosverbejä, lainoja ja omapohjaisia. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 92, p. 79–102. <http://www.sgr.fi/susa/92/koivulehto.pdf>

Mallory, J. P. 2001: Uralics and Indo-Europeans: Problems of time and space. Carpelan et al. (edited): *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations*, p. 345–366. Suomalais-Ugrilaisen Seuran Toimituksia 242.

Mallory, J. P. & Adams, D. Q. 2006: *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Oxford / New York: Oxford University Press.

Parpola, Asko 2008: Proto-Indo-European speakers of the Late Tripolye culture as the inventors of wheeled vehicles: Linguistic and archaeological considerations of the PIE homeland problem. *Proceedings from the 19th Annual UCLA Indo-European Conference* (JIES Monograph Series). Karlene Jones-Bley & al. (editors.). Washington, DC: Institute for the Study of Man.

Schalin, Johan: Lexicon of Early Indo-European Loanwords Preserved in Finnish. <http://kotisivu.lumonetti.fi/js749/lexicon.htm>