

Substrate in Finno-Permic languages

Niklas Metsäranta

University of Helsinki (Helsingin yliopisto)

A considerable number of vocabulary items that have been traditionally reconstructed for the now defunct “Finno-Volgaic” or “Finno-Permic” proto-languages involve irregular sound correspondences and phonological features either foreign or atypical of Uralic languages (Aikio 2015: 44–47). Especially the rare phoneme *š (singleton and different clusters e.g. *kšt, *kšn, *šm, *šn) is clearly overrepresented in these words whose distribution is limited to Finno-Permic languages (Saami, Finnic, Mordvin, Mari and Permic). Aikio offers a list of 21 words that contain the phoneme *š in various combinations. Words in this list mostly pertain to animal husbandry (e.g. Finnish *lehmä* ‘cow’, Erzya *l’išme* ‘horse’), agriculture (e.g. Mari *šož* ‘barley’, Komi *čuž*, Udmurt *čužjem* ‘malt’, Fi *riihi*, Komi *riñiš* ‘drying barn’), nature (e.g. Fi *pähkinä*, Mari *pükš*, Udmurt *pašpu* ‘nut’, Fi *vaahtera*, Erzya *ukštor*, Mari *waštar* ‘maple’) and primitive technology (e.g. Fi *huhmar*, Mordvin *čovar* ‘masher’). Given the irregular sound correspondences, the foreign phonological structure, the limited distribution, and the clustering semantics, Aikio puts forth the idea that these types of words in Finno-Permic originate from an extinct substrate source and calls for the topic to be explored further. It seems that Aikio’s call has thus far mostly fallen on deaf ears.

In this presentation, I will explore the topic of Finno-Permic š-substrate or shubstrate further. Firstly, I aim to expand the potential substrate corpus beyond the 21 words mentioned by Aikio by gathering all words that could potentially reflect a cluster containing an š but are distributionally more limited, i.e. are found only in one or two branches of Finno-Permic. Even a cursory look into this type of vocabulary shows that there is potential there. For example, there are a number of words limited to Saami and Finnic that seem to reflect an earlier *šŋ-cluster, e.g. North Saami *vuoskku* ‘perch’ (< Proto-Saami **vuosŋōn*) and Finnish *ahven* (< Early Proto-Finnic ?**ašŋənə*). The idea of Finnish *ahven* being a substrate word is not new (Ariste 1971: 257) but to my knowledge it has not been previously suggested that it together with the Saami word could belong to the Finno-Permic š-substrate. If ‘perch’ turns out to be part of the š-substrate – like its phonological structure and meaning would suggest – then from this observation it can be extrapolated that other words reconstructable to Proto-Saami with the

cluster **sɲ*, e.g. **kēsɲes* ‘juniper’ > South Saami *gasnges*, North Saami *gaskkas* (Lehtiranta 2001: 42–43), are also potentially part of the same substrate.

Secondly, after collecting the potential words reflecting clusters containing an *š*, I will examine their co-occurring structural properties. For instance, most *ahven*-type words in Finnic, e.g. Finnish *ien* ‘gum’, *haiven* ‘single hair’, *huomen* ‘morrow, tomorrow’, *höyhen* ‘feather’, *joutsen* ‘swan’, *jäsen* ‘member, limb’ (Hahmo 1994: 122–123) have comparanda that is for the most part limited to Finno-Permic languages. Thus, even words that do not bear the explicit phonological hallmarks of the *š*-substrate could potentially be identified as such if some of their co-occurring structural properties can be determined to be indicative of substrate origin.

There are words that bear many of the same phonological and structural characteristics as words belonging to the Finno-Permic *š*-substrate but have not been listed as such, e.g. North Saami *lasta*, Finnish *lehti* ‘leaf’, Luleå Saami *sassne* ‘tanned reindeer leather’, Finnish *hihna* ‘leather strap’, Finnish *tuhat*, East Mari *tüžem* ‘thousand’. Due to the fact that there are similar looking words particularly in Baltic, cf. Lithuanian *šikšnà* ‘untanned leather; leather strap or belt’, *tūkstantis* ‘1000’, the Finno-Permic words have often been considered Indo-European loanwords. Recently, given that the Baltic words themselves often lack an Indo-European etymology, this idea has been challenged and it has been proposed that rather than being loanwords from Baltic to Finnic or vice versa, words such as Finnish *hihna* and Lithuanian *šikšnà* instead represent parallel loans from a substrate that has been tentatively identified to be the same substrate as the Finno-Permic *š*-substrate (Jacob 2023: 188, 478). The idea of Baltic and Finno-Permic languages sharing the same substrate raises interesting questions about the distribution and geography of the *š*-substrate that I will also be exploring in my presentation.

References

- Aikio 2015 = Luobbal Sámmol Sámmol Ánte (Ante Aikio). The Finnic ‘secondary e-stems’ and Proto-Uralic vocalism. *Journal de la Société Finno-Ougrienne* 95. 25–66.
- Ariste, Paul. 1971. Die ältesten Substrate in den ostseefinnischen Sprachen. *Linguistica Uralica* 7(4). 251–258. <https://doi.org/10.3176/lu.1971.4.02>.
- Hahmo, Sirkka-Liisa. 1994. *Grundlexem oder Ableitung? Die finnischen Typen kämmen und pähkinä und ihre Geschichte* (Studia Fennica Linguistica 5). Helsinki: Suomalaisen Kirjallisuuden Seura.
- Jacob, Anthony. 2023. *A History of East Baltic through Language Contact*. Leiden: University of Leiden. (Doctoral dissertation.) <https://hdl.handle.net/1887/3630131>.

Lehtiranta, Juhani. 2001. *Yhteissaamelainen sanasto* (Mémoires de La Société Finno-Ougrienne 200). Helsinki: Suomalais-Ugrilainen Seura.