Recurring irregularities in West Uralic 1: Para-Slavic loanwords

Jaakko Häkkinen

hakkinenjaakko@gmail.com

1. West Uralic

West Uralic (WU) consists of the westernmost Uralic branches: Saami, Finnic, Mordvin, and probably some extinct branches like Meryanic and West Chudic. Proto-West Uralic (PWU) is the stage, when the language was still uniform and spoken in a narrow area.

Common West Uralic (CWU) is the stage, when the language dispersed into a wider area and in which the earliest dialectal differences developed, even though those loanwords lacking the diagnostic sounds still show identical reconstructions between the branches. This stage can be associated with the pre-stages of the branch-specific protolanguages.

Diverging West Uralic (DWU) is the stage, when we begin to see clearly different reconstructions between the branches; this can be associated with the early and middle stages of the branch-specific protolanguages. After DWU, we rarely see loanwords shared by the WU branches any longer, indicating even greater level of areal dispersal.

There is no space to dwell deeper in the PWU phonology here, so I only shortly describe the two new vowels * and * , sketched already in Häkkinen (2007). The need for these new vowels arises from the fact, that there are two different cognate sets for both PFi *e-i and PFi *i-A, earlier left without satisfying explanation

Furthermore, we can find even wider Uralic support for the new vowels. Considering the *i-A combinations, compare the following examples, which only coalesce in the westernmost branches:

PSa * *elme* 'eye' ~ PFi *silmä* ~ PMd * *elm*   ~ PMr *šin- ä* ~ PPe * *in(m-)* ~ PMn *šam* ~ PKh *sem*

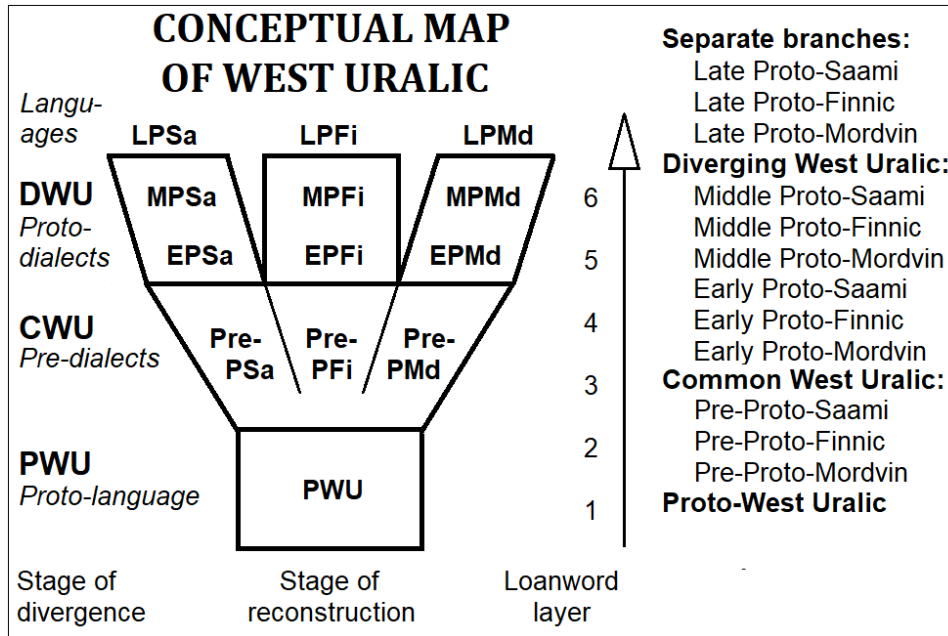PSa * *ele* 'inner bark' ~ PFi *nila* ~ PMd *nola* ~ PMr *nol*   ~ PPe * *ol* ~ PMn * *ol-* ~ PKh * *al-*

Figure 1. Conceptual map of West Uralic.

Proto-West Uralic	Proto-Finnic	Proto-Saami	Proto-Mordvin	Earlier reconstruction
*e-ə	*e-i	*ie-e̮	*i̮-ə	*e-i
*ə-ə	*e-i	*e̮-e̮	*e-ə	*e-i
*i-a	*i-ä	*e-ē	*e-a	*i-ä
*ə̂-a	*i-a	*e̮-ē	*o-a	*i-a

Figure 2. West Uralic cognate sets for Proto-Finnic *e-i and *i-A.

As pointed out by Mikko Korhonen (1988), it is very unlikely, that the PU *i* would have been the only vowel, after which there existed the opposition *ä* vs. *a* in the second syllable. Therefore the opposition should be shifted onto the vowel of the first syllable, leading to an update of the PU and PWU reconstructions followingly: *i-a* * *-a* and *i-ä*   *i-a*.

Considering the two cognate sets for PFi *e-i*, the strongest support comes from Nganasan: the distinction between Nganasan *i* and *i*, argued to have been preserved from Proto-Uralic through Proto-Samoyedic by Eugen Helimski (2005), seems to agree with the different PWU cognate sets:

Recurring irregularities in West Uralic 1: Para-Slavic loanwords

A) Ngan i ~ PSa *e ~ PFi *e ~ PMd *e

PU *m n - 'to go' > Ngan *min-* ~ PSa *mene- ~ PFi *mene- ~ PMd Ø || PU *p l - 'to fear' > Ngan *hil-* ~ PSa *pele- ~ PFi *pel-kä- ~ PMd *pel - || PU *w t 'water' > Ngan *bi"* ~ PSa Ø ~ PFi *vete- ~ PMd *ved || PU *k t 'shape, figure' > Ngan *si"* ~ PSa *kete ~ PFi *kete- ~ PMd *ked.

B) Ngan i ~ PSa *ie ~ PFi *e ~ PMd *i

PU *mex - 'to give' > Ngan *mi-* ~ PSa *mieke- ~ PFi *möö- ~ PMd *mij - || PU *pex - 'to cook' > Ngan *hi-* ~ PSa Ø ~ PFi Ø ~ PMd *pij -.

There are more words in these groups to confirm the PWU cognate sets, but unfortunately they lack cognates in Nganasan. However, there is no visible sound environmental reason for the difference between the vowels, if we look at the larger list. Consequently, I split the earlier PU and PWU reconstructions followingly: A) *e- * - vs. B) *e- . In both of these cases, the Finnic vowel now seems to be of dual origin (PFi *i < *i, * and PFi *e < *e, *), misleading scholars who earlier often had a predetermined conception of the archaicness of the Finnic vowels.

Admittedly, the evidence for the complementary vowel combinations * - and * - a is less conclusive (see Häkkinen 2007). Theoretically it could be possible, that there was only one reduced vowel in PrePU, showing the back reflex * before the open back vowel *a, and the front reflex * before the close front vowel * . Such an allophony could be caused by regressive assimilation, produced by the second syllable vowel on the weaker "inner" vowel of the first syllable. Still, based on the different realizations in the WU branches and more widely in the Uralic language family, I reconstruct the following ten vowel system for PWU:

PWU vowel system	
1. syllable	2.
i ü i̵ u e ə ə̂ o ä a	ə (o) a

Figure 3. Proto-West Uralic vowel system.

For the second syllable vowels I follow Petri Kallio (2012). Towards PWU, the PU *e split into two: *e-a > *a-a vs. *e- > *i- (Aikio 2015: 39). The combination * - , which still in PU was more frequent than *e- , now became unproductive, as nearly all of the new words with the WU distribution were adopted into the combination *e- (the cognate set PSa *ie ~ PFi *e ~ PMd *i). After the PWU stage, the vowels * (-) and * (-a) coalesced with other vowels, but differently in different branches:

PWU *e, * , *i
 > PrePFi *e, *e, *i > PFi *e, *e, *i
 > PrePMd *e, * , * > PMd *i, *e, *e
 > PrePSa *e, * , * > PSa *ie, *e, *e

PWU *o, * , *u
 > PrePFi *o, *i, *u > PFi *o, *i, *u
 > PrePMd *o, * , * > PMd *u, *o, *o
 > PrePSa *o, * , *u > PSa *oa, *e, *o

2. Balto-Slavic

Here I call the shared ancestor of Baltic and Slavic languages Proto-Balto-Slavic (PBSl), although some scholars call the very same stage Proto-Baltic (PB). The former practice appears preferrable now, when I shall in this article propose two new early Balto-Slavic branches in addition to Slavic and Baltic (or in the tripartite division: Slavic, West Baltic and East Baltic).

PBSl is defined here by the result of satemization: *k > * and *g, *gʰ > *z. In East Baltic these changed into *š and *ǯ, while in West Baltic and Slavic the ultimate results were *s and *z (Kim 2018: 1975–1976). I follow the PBSl reconstructions by Rick Derksen (2008; 2015), although updated according to Petit (2018) and Kim (2018). Derksen still has the short *o (< PIE *o) in PBSl, but as it early on coalesced into *a, my reconstructions have *a instead.

Here I follow the definition and periodization of Proto-Slavic (PSl) by Georg Holzer (e.g. 2002) and Emanuel Klotz (2017: 16–17, 47): PSl is a more archaic stage than in some of the earlier models, ending ca. 600 AD. Common Slavic (CSl) is the label used here for the later stage, which still Derksen's dictionary (2008) was marked as "late stage of Proto-Slavic". Most of the PSl and CSl words are from the dictionary of Klotz (2017), but PBSl, PB, and some of the CSl forms are from the dictionaries of Rick Derksen (2008; 2015).

The WU words have been collected from the following dictionaries: for Finnic, SSA and SMS; for Saami, Lehtiranta (1989/2015); for Mordvin, Paasonen (1990, 1992, 1994, 1996, 1998, 1999). The reconstructions for PFi, PSa, and PMd are my own, based on the recent studies on Uralic sound history, and for Proto-Mordvin on my own ongoing research.

3. Para-Slavic loanwords in West Uralic

Here I propose a whole new loanword layer for the Diverging West Uralic stage. I label this layer Para-Slavic, because it shows some shared developments with Slavic proper, although it also shows independent developments. Because the contacts with Para-Slavic seem to have begun earlier than with Slavic, the former cannot descend from Proto-Slavic, but it must be a parallel lineage.

The clue for this new layer is the sibilant correspondence MPFi *š ~ MPMd * , which is impossible in the inherited vocabulary. Curiously, this correspondence set seems to relate to the Proto-Slavic *s. In old Germanic loanwords into Finnic, we observe a substitution G *s Fi *š, but this is considered possible only because there

Recurring irregularities in West Uralic 1: Para-Slavic loanwords

was no *š in Germanic, leaving wider space for the realizations of *s (LÄGLOS I: XX). Here that condition is not fulfilled, as there were both *s and *š already in Proto-Balto-Slavic and Proto-Slavic.

Moreover, the Proto-Uralic *s, *š, and * remained distinct in Proto-Mordvin as *s, *š, and * , while in Proto-Finnic *s , *š > *h and * > *c (then coalesced into *s outside South Estonian). Therefore, the correspondence set Fi *š ~ Md * cannot be derived from the common ancestral stage, but it must represent parallel borrowing into Finnic and Mordvin. Furthermore, because both branches opposed the plain *s with other sibilants, it seems inevitable, that the source sound was not *s like in PBSl and PSl, but instead this sibilant had changed into *š, * , or something similar.

We see the DWU *š ~ * reflecting similarly all the PBSl sibilant (*s, * , *), which points to the direction, that all the sibilants were coalesced into * in Para-Slavic. Quality of the resulting sound is based on Finnic sound history: PU * changed into *c already very early (EPFi depalatalization), while *š remained until very late stage, before it changed into *h (Kallio 2007). Consequently, there was a stage around Middle Proto-Finnic, when * of the source language would have been substituted with *š due to the lack of * /* (so already Kallio 2008: 267–268), while in Mordvin it would have been substituted with * in all stages.

1. MPFi *šera > PFi *hera 'milk serum' (> Ing heravesi 'cheese liquid')
    MPMd * ür > PMd * er > MdM er -ved 'milk serum' ("cheese-water")
       ParaSl * u/iro 'cheese' < PBSl *sú r(i)os > PSl *su̱ ˑru > CSl *syru

2. MPMd *mü a > MdE M me as 'mole (Talpa), bat (Chiroptera)'
       ParaSl *mu/i e 'mouse (Mus/Muridae)' < PBSl *mu - > PSl *mu̱ ˑši > CSl *myši

In these two words Mordvin points to the original *ü, which can be derived from ParaSl *u, *u, while Finnic points to *e, which rather points to ParaSl *i. Possibly PMd *e could also derive from the latter sound via MPmd * (more about the Para-Slavic sound in chapter 3.2).

3. MPFi *šarakka 'magpie (Pica)' > PFi *harakka
    ? MPMd * ärakka > PMd * ä aka > MdE e aka, e gan, M ä gan
       ParaSl * arka 'magpie' < PBSl * ár ka > PSl *sa̱r ka˙ > CSl *sorka

4. MPFi *ša ka > PFi *ha ka 'fork etc.', *ha ko 'hayfork'
    MPMd, PMd * a ka > MdE a go, M a ga 'hayfork'
       ? ParaSl * aka 'forked stick' < PBSl * ak a > PSl *saxa˙ > CSl *soxa
       ? NB *šanka 'hook etc.' / *šankas 'bough, twig' (Junttila 2020: *hanka)

5. MPFi *šalma > PFi *halm-eh (derivative) 'swidden (growing cereal)'
       ? ParaSl * alma 'straw' < PBSl *sál m- > PSl *sa̱l ma˙ > CSl *sòlma
       ? NB *ḟalm- 'sprout; growing field' (Junttila 2020: *halmëh)
       ? PSc *halma- 'stem, stalk; harvest' (Junttila 2020: *halmëh)

In the word 3. 'magpie', the Finnic word could also come from the PB *šarka, but the Mordvin word could not, because it would show *š as the reflex of the *š. ParaSl origin for both branches can be proposed. Still, both *ä* and   in the Mordvin word for 'magpie' are unexpected, but the similarity of this peculiar three-syllable word with the PFi and PSl words can hardly be coincidental.

There occurs some unexplained variation between PBSl *a and *e in the modern Baltic and Slavic languages (Kim 2018: 1977). This might also be connected to a variation seen in some Baltic loanwords in Finnic: Fi *rastas* ~ *rästäs* 'thrush (*Turdus*)', *harma* ~ *härmä* 'hoarfrost'. The same BSl variation could be behind the PMd *ä instead of *a, but this might also be secondary development in Mordvin (cf. the word 6). Sibilant assimilation could explain the change * -r > * -   > * -   (between vowels), although I cannot present any further examples at the moment.

In the word 4. 'hayfork' we see an unexplained *   in the Uralic side. An "extra" homorganic nasal in front of a consonant occurs in Balto-Slavic in certain verbal forms (Villanueva Svensson 2016), but no traces of such phenomenon are presented concerning nouns. However, if the West Uralic speakers understood such an exotic "nasal infix" as a marked property of Balto-Slavic, they might have hypercorrectly added it to some loanwords in which it did not belong. Similar markedness is seen in a tendency of modern Finnish speakers to adopt "exotic" voiced stops hypercorrectly in words, which have voiceless stops in a donor language.

North Baltic etymology would agree better with the Proto-Finnic cognate, but it cannot explain the Mordvin cognate with * . Semantically the Slavic word is also closer than the Baltic candidate.

The word 5. for 'swidden' is only found in Finnic, and it has been connected to Scandinavian and Baltic words. In the former case the PFi *h would come directly from the PSc *h. Without a cognate in Mordvin, the sibilant criterion cannot be applied here, and therefore Baltic or Scandinavian etymologies seem also possible.

Some words show a more peculiar vowel correspondence DWU *a ~ PSl *i. The close PSl front vowel and the open WU back vowel could be bridged by assuming a mid-close central vowel *   in ParaSl. The following three words also show the WU *š ~ *   corresponding to PSl *j, pointing to a strongly fricative or even affricative pronunciation *j in ParaSl. The WU reflexes for the ParaSl *   and *j were therefore exactly the same.

6. ? MPFi *šala a ~ *jala a ~ * ala a > PFi *halaga 'bay willow (*Salix pentandra*)', *jalaga 'elm (*Ulmus*)', *calaga > *salaga 'crack willow (*Salix fragilis*)'
   MPMd * äla a > MdE *ele* , M *äli* 'elm'
   PrePMr * ala- > PMr *šol 'elm'
      ParaSl *j l mo 'elm' < BSl * il-mo- > PSl *jilimu > CSl *jilimu

7. MPFi *šamu- > PFi *hamu- 'to collect, assemble'
   ? MPMd * ama- > PMd * ava- > MdE *avado-*, *ijado-*, M *avad -*, *ud -* 'to envy'
      ParaSl *j ma- 'to take, gather' < BSl *im- > PSl *jima > CSl * -, *jimà-

8. MPFi *šaru > PFi *haru 'cold spring wind; freeze'
      ParaSl *j ro 'northern wind?' < BSl *iu r- 'water, sea' > PSl *jiru > CSl *jiru
   strong wind; 'warm lands to which birds migrate'

Recurring irregularities in West Uralic 1: Para-Slavic loanwords

In the word 6, the Finnic *salaga and the Mari cognate are regular outcomes from * ala a, although Mordvin again points to *ä, reminiscing the word 3. * ärakka. Still, the existence of the variant PFi *halaga connects this to the other words of this layer (although arguments for the original PFi *halaba have been presented; see discussion in Junttila 2020: hala-). The third syllable is not a good match, but it can be explained by replacement of the third syllable by the WU "tree-formant" *- a (> PFi *-ga), which is seen also in the words Fi pihlaja, pihlava ~ Est pihlakas, pihelgas 'rowan (Sorbus)' and Fi kataja, katava ~ Est kadakas 'juniper (Juniperus)'.

Moreover, the Finnic variation *j ~ (* >) *s could be explained phonetically from ParaSl *j: the *j-substitution emphasizes the voicedness, while the * -substitution emphasizes the affrication of the source sound. Admittedly, more examples are needed to support the existence of such parallel substitution patterns. There are still a lot of unexplained questions in these similar-looking tree names, but I see it economical to find one and the same explanation for them all.

In the word 7, PMd * ava- would regularly reflect earlier * ama-, but the meaning 'to envy' is not so close to the meanings in Finnic and Slavic. It is still possible: envy occurs, when one gathers less than somebody else.

Considering the word 8, *šaru, the PSl meanings are quite far from the PBSl meaning, but I leave that for the Indo-Europeanists to decide. Still, both of the PSl meanings could be derived from the meaning 'northern wind', which triggers the migration of the birds to the south, and therefore 'northern wind' could have been the meaning of this word in Para-Slavic, too.

The next words with the same vowel correspondence have a narrow distribution, but otherwise they fit in the phonological framework of the above presented words:

9.   MPFi *pašna > PFi *pahna 'straw (under a pig or another animal)'
       ParaSl *p eno 'corn, millet (Panicum, Setaria)' < PBSl *pi - > PSl *pišena > CSl *pišeno

10. MPMd * arta > MdE ardo, M arda 'elk, deer'
       ParaSl * r na 'roe deer' < PBSl * ir na > PSl *sir˙na˙ > CSl *sirna

11. PrePSa *š/särta (*š/sarta?) > PSa *sarte 'heart of a reindeer; meat chunk'
       ParaSl * rd- 'heart' < PBSl * ird- > PSl *sirdika > CSl *sirdice

The word 9. *pašna is quite clear: straws or stalks of any cereal-like plant were used under animals. Adopting a three-syllable word as a two-syllable word in WU with syncope is seen also in the word 22. *ri ma (chapter 4).

The word 10. * arta shows *Ct as a substitute for ParaSl *C n, similar to the ABSl word 19. *palt (chapter 4). This indicates very early loanword, as in the other ParaSl loanwords there are no traces of * .

The word 11. 'heart' would regularly go back to PrePSa *šärta or *särta, showing a similar *ä-variant like in some Mordvin words earlier (3. * äzakka and 6. * äla a). Still, there are few occasions, in which PU *a-a has also yielded PSa * - (Aikio 2015 appendix: *ka ka-, *wan a-, *wara), so the *a-substitution is also possible. Sibilant-wise

Saami agrees with Finnic, showing the PrePSa *š as a reflex of the ParaSl * . This is expected, as in PrePSa there was no * , because the PU * remained as an affricate.

There are two possible words corresponding to the ParaSl denasalization of nasal before a stop:

12. ? PrePSa *š/sak / *š/soka > PSa *soake 'birch (*Betula*)'
       ParaSl * ako 'bough, knot (in wood)' < PBSl * ank- > PSl *sanˈku > CSl *s ku

13. MPFi *ša(t)tara > PFi *hattara, *hatara 'small cloud; big snowflake'
       ParaSl * adarga 'fine hail' < PBSl *sam-dar ga > PSl *samˈdarga > CSl *s dorga

The PSa vowel combination *oa-e can be regularly derived either from PrePSa *o-a or from *a- (see Aikio 2015: 39), so the forms PrePSa *šak and MPFi *šattara could show similar reflex: short *a with no traces of the following nasal consonant.
     In the word 13. *šat(t)ara, single *t would be an expected reflex of the ParaSl *d, and *hatara* 'cloud, foggy' is indeed seen in the Finnish dialects alongside the more common *hattara* 'small cloud; big snowflake' (SMS: *hatara*).

14. MPMd *śära > MdM śära 'swine (*Sus*)'
     ← ParaSl *śɛr- 'hair, bristle (of swine?)' < PBSl *śér-; cf. Lit *šerys* 'bristle' → *šernas* 'wild boar (*Sus scrofa*)'

15. MPFi *ärä > PFi *äre-dä 'grumpy, cranky', *ärä-htä- 'to snarl, snap'
     ← (Para)Sl *ɛ̄ro 'furious' < PBSl *eɔr- > PSl *ēru > CSl *ĕrŭ

These two words show that the PBSl *e and *e developed into more open vowel in ParaSl, just like they did in PSl (Kim 2018: 1979). It seems that bristle was primarily associated with swine in BSl, although the word 14. was not preserved in Slavic proper. The word 15. could also be a later Slavic borrowing, as there is no sibilant to confirm the old age.

16. MPFi *šura > PFi *hura 'unpleasant, sad'
     E/MPMd *śurə̂ > PMd *śorə̂ > MdE śor-ks 'angry'
     ← ParaSl *źūre- 'to rage' < PBSl *źeuɔr- > PSl *zūri- > CSl *zuri-, *žuri-

This is so far the only example of WU *u, but the word shows the typical sibilant correspondence and can therefore be derived from ParaSl. The existence of long *uu in Finnic seems to be secondary, possibly from multiple sources (Pystynen 2018: 50–53), so the ParaSl long *u would still have been substituted by the short *u before the second syllable *a.
     Here Mordvin shows the primary *u, which coalesced into * and later changed into *o towards PMd, which makes this loanword very old.

Recurring irregularities in West Uralic 1: Para-Slavic loanwords

3.1. Para-Slavic consonants

1. Para-Slavic does not differ from Proto-Slavic or Proto-Balto-Slavic:
    [No examples, because the aim is to find distinctly Para-Slavic loanwords]

2. Para-Slavic shares an innovation with Proto-Slavic:
    DWU *š ~ *   ParaSl *ǰ < *j (prosthetic) < PBSl 0 > PSl *j > CSl *j (words 6, 7, 8)

3. Para-Slavic shows an independent innovation:
    DWU *š ~ *   ParaSl * < PBSl *s > PSl *s > CSl *s (words 1, 5, 13)
    DWU *š ~ *   ParaSl * < PBSl * > PSl *š > CSl *š (words 2, 9)
    DWU *š ~ *   ParaSl * < PBSl * > PSl *s > CSl *s (words 3, 4, 10, 11, 12, 14, 16: *)
    DWU *š ~ *   ParaSl *ǰ < PBSl 0 > PSl *j > CSl *j (words 6, 7, 8)

Para-Slavic shares with Slavic the appearing of the prosthetic initial *j, but only in Para-Slavic it developed further into *ǰ (strong fricative or affricate).

3.2. Para-Slavic vowels

1. Para-Slavic does not differ from Proto-Slavic or Proto-Balto-Slavic:
    DWU *a   ParaSl *a < PBSl *a > PSl *a > CSl *o (words 3, 4, 5)

2. Para-Slavic shares an innovation with Proto- or Common Slavic:
    DWU *e ~ *ü   ParaSl *i/*u < PBSl *u > PSl *u > CSl *y (words 1, 2)
    DWU *ä   ParaSl * , * < PBSl *e( ) > PSl * , * > CSl *e, *e (words 14, 15)
    DWU *u   ParaSl *u < PBSl *eu > PSl *u > CSl *u (word 16)

3. Para-Slavic shows an independent innovation:
    DWU *a   ParaSl *a < PBSl *aN > PSl *aN > CSl * (words 12, 13)
    DWU *a   ParaSl * < PBSl *i > PSl *i > CSl *i (words 6, 7, 8, 9, 10, 11)

The change *š > *h is considered the latest Proto-Finnic innovation, occurring around 200 AD (Kallio 2007; 2014). This agrees well with the archaic vowels of type 1, and also with the innovations of type 2 shared with Proto-Slavic. Finnic dating agrees also with the independent Para-Slavic innovations of the type 3, because Proto-Balto-Slavic is seen to have diverged very early, already around 1000 BC. There was more than a millennium for the distinct Para-Slavic developments to appear before the Proto-Finnic change *š > *h.

However, the Common Slavic delabialization *u > *y [*i] in the type 2 is traditionally seen centuries too late (Collins 2018: 1461) to agree with the *š-borrowing in Finnic. However, this problem only concerns the MPFi *šera (the word 1), as the PMd *e in words 1 and 2 can be explained from earlier *ü, which sometimes derives from BSl *u: the substitution BSl *u *ü is seen in the PFi words *kürsä 'hard or dry)

bread' (Koivulehto 2001: 241), *tüttäre- 'daughter', and *tühjä 'empty' (Kalima 1936: 71).

Still, it would not seem totally impossible, that similar delabialization and centralization could have occurred independently in Para-Slavic, too. After all, also the other close vowel, the PBSl *i, must have been centralized and lowered into * , because the WU *a-borrowings are phonetically impossible to derive straight from the PBSl/PSl *i. The central mid-close * bridges the gap between the close front vowel and the open back vowel.

4. Archaic Balto-Slavic loanwords in West Uralic

There seem to exist also Archaic Balto-Slavic loanwords in West Uralic, distinguished by the reflex of the PBSl glottal stop * (from PIE laryngeal). Here I propose that the effect of this sound on the nearby consonant results as a voiceless stop in the loanwords borrowed into West Uralic.

17. PrePMd *aks > PMd *uks > MdE *ukso* 'elm (*Ulmus*)'
     MrE *oško* 'poplar (*Populus*)'
         ABSl *a s-en-; *a s-i-o- 'ash (*Fraxinus*)' > PSl *asi > CSl *osi

18. ? PrePFi *mükša > PFi *mükrä 'mole (*Talpa*); small rodent (*Arvicolinae*)'
         ABSl *mu - 'mouse (*Mus/Muridae*)' > PSl *mu ši > CSl *myši

There are two possible examples of the substitution * S   *kS. The word 17. is quite clear: borrowed names for trees often denote different trees than in the donor language, and even different trees than in the related languages. PWU *a- regularly yields PMd *u- (Aikio 2015: 39). The consonant metathesis *kS   *Sk is a well-known phenomenon in Mari vocabulary, although it is considered regular only in Permic, and therefore Niklas Metsäranta considers such Mari words as borrowings from Permic (Metsäranta 2020: 234).

The word 18. *mükša requires some irregular developments and therefore remains uncertain. As mentioned earlier, there are some examples of the substitution BSl *u   WU *ü. The sporadic change *š > *r is more difficult to support, but there are some possible parallels with PFi *r corresponding to hushing sibilants elsewhere — interestingly, in the domain of animal names:

PFi *orih, *oreh 'stallion' ~ PMr *oʃ ~ PPe *uʃ (< EPPe *a a; Metsäranta 2020: 213, 214)
PFi *poro 'reindeer' ~ PSa *poac j ~ PMr *pu   ~ PPe *puʃ < PU *po av (Sammallahti 1998: 232)

It could be theoretically possible to derive PFi *mükrä from *mükša and ultimately from ABSl *mu -. Admittedly, the Finnic *r-words could also reflect intra-Uralic borrowing substitution *ʃ   *r (especially 'stallion' based on its irregular vowel correspondences), but there is no trace of the word *mükša in other Uralic branches (although somewhat later ParaSl borrowing is seen in the chapter 3: word 2).

Recurring irregularities in West Uralic 1: Para-Slavic loanwords

19. PrePFi *palt > PFi *palci 'hard clay/soil layer, especially in the bottom of a lake or river; shoal, shallows' > Fi *palsi-kko* 'the part of the shore which is flooded'
    PrePSa *palt > PSa *puolte 'slope of hill/mountain'
        ABSl *bal ni- 'low flooded place; pasture' > PSl *baľ ni- > CSl *bolnije

SSA presents the word *palci as a Germanic borrowing, but its meaning 'fold, tuck, hem' is a poor match (SSA: *palsi*), so the ABSl etymology seems more credible. For the substitution BSl *R n WU *Rt, there is another example in the chapter 3: word 10.

Luobbal Sámmol Sámmol Ánte has argued, that the original PU *a- > PSa *oa-e, thus coalescing with the PU *o-a (Aikio 2015). We can assume, that after some early stage of change in the primary *a- combination, there occurred the secondary *a- in PrePSa, which then developed symmetrically with the original *a-a combination (*a > *uo in both of them). Such development is confirmed by several Germanic loanwords: PG *hamen PrePSa *am n > PSa *vuomen 'a kind of hunting fence' || NwG *haba-z PrePSa *ap /as > PSa *vuopse 'depth of a fishing net' || PG *skallV- PrePSa *kal > PSa *kuole 'testicle' || PG *skapa- PrePSa *kap- s/š > PSa *kuopes 'witch' (Aikio 2006: 10–11).

20. PrePFi *l ppa 'alder (*Alnus*)' > PFi *leppä
    PrePMd *l ppa > PMd *lep > MdE lepe
    MPSa *lejpa > PSa *leajpe
        ABSl *léi pa-/*lé ipa- 'linden (*Tilia*)' > PB *leipa

Proto-Finnic *leppä has usually been considered a Baltic loanword. Still, only the Saami cognate *leajpe can be regularly derived from the Baltic original (PrePSa *lejpa PB *leipa), while Finnic and Mordvin point to the original *pp, which could be explained from the ABSl cluster * p. At the early WU stages, the structure *VjCC was not allowed, so the result of *éi p would have been * pp/*epp. Also Santeri Junttila has considered the possibility that the PBSl glottal stop or the earlier laryngeal could explain the PFi geminate stop in some of the Baltic loanwords (Junttila 2017: 140–144).

The Finnic front combination *e-ä points to a very old loanword, as almost always even the Baltic loanwords show the back combination *e-a (or the secondary *ei-ä from PFi *ai-a; see Kallio 2014: 159). In Mordvin there is the secondary *e, which is often considered as a sign of a young word, but it goes regularly back to the PWU * and can therefore be very old (cf. chapter 1).

21. PrePFi *me a > PFi *meccä 'forest'
    PrePSa *me a > PSa *mea e 'far, distant' (or via Finnic)
        ABSl *medj-o/a- > PB *medias 'forest'

This word is usually seen as a Baltic loanword, but the old *e-ä combination in Finnic suggests it could be earlier than the Baltic loanwords, which are borrowed already into the younger *e-a-combination — the only exceptions being this and the previous word. It would be better to consider these two *e-ä words already ABSl borrowings.

In South Estonian there is apparently a younger borrowing with the back combination, which goes back to PFi *mecca. This could be treated as an independent Baltic borrowing, as there are plenty of those in South Estonian. Another possibility could be a secondary fronting of the vowels caused by palatalized consonant, as pointed out by a referee: *me a > *meccä. However, the Baltic loanwords are usually dated from MPFi onwards, and at that point the depalatalization had already occurred (Kallio 2007: 233). Only around the same time the *e appeared in the first syllable of MPFi (Häkkinen 2019: 34), so there probably was no stage at which the *e and * coexisted.

22. (? PrePFi *rjćma 'string, rope' >) MPFi *rišma > PFi *rihma
    MPMd *ri m > MdE ri me
    ? PSa *reaćmē 'rope (of fishing net)'
        ABSl *ri ima- 'binding' > PB *rišima- (Grünthal 2012: 328)
        ? PrePII *Hre mi- 'rope' (Holopainen 2019: 249)

The *šm in Finnic and Saami can now be seen as the regular outcome of PWU * m (see Aikio 2015: 44), but the Mordvin * cannot come from *š, so here the Baltic origin seems unlikely. Riho Grünthal proposed, that the Baltic *š could have been strongly palatalized in this environment, but there is no further evidence of such phenomenon. The sibilants cannot help to distinguish between ABSl and ParaSl origin, because of the conditioned change * m > *šm in Finnic, but the preserved *i points towards ABSl: in ParaSl it would have been developed into * , substituted with the WU *a (cf. the chapter 3.2).

The vowel cognate set PFi *i-a ~ PMd *i- shows the secondary *i in Mordvin (which appeared after the primary *i > *). Moreover, in Finnic the word shows the younger back combination *i-a instead of the older front combination *i-ä. Therefore we are dealing with a somewhat later borrowing compared to the other ABSl loanwords.

The Saami word points to the original combination *e-a, and also the Mordvin word could be derived from that (if an older borrowing before the primary *e > secondary *i), but the Finnic word could not. Moreover, in the old words we see the change * C > *šC towards Proto-Saami, so the preserved * C here points to a much later borrowing, not until Proto-Saami. Therefore we must reject the Pre-Proto-Indo-Iranian etymology. At the moment, the origin of the Saami word remains unknown.

5. Dating and locating the contacts

Afore I have proposed two possible new loanword layers in West Uralic: Para-Slavic and Archaic Balto-Slavic layers. Even though the latter stage on phonological ground could in theory equal Proto-Balto-Slavic and therefore be ancestral to the Para-Slavic layer, there are reasons to assume, that here we are dealing rather with temporally overlapping sister dialects.

The first criterion is based on the Mordvin secondary close vowels: after the PrePMd primary *i and *u had coalesced into the * and * (which later changed into the PMd secondary *e and *o, respectively), there appeared the new, secondary *i and *u (> PMd *i and *u, into which the primary *e and *o coalesced, respectively; cf.

Recurring irregularities in West Uralic 1: Para-Slavic loanwords

chapter 1). The ABSl loanword 22. *ri ma shows the secondary PMd *i, while the ParaSl loanword 16. * ur shows the primary PrePMd *u > PMd *o, thus appearing as older one of the two.

Another criterion is the ABSl glottal stop * , which is still seen also in the ParaSl loanword 10. * arta, thus appearing to represent the same chronological stage with the ABSl loanwords. Still, the ABSl loanwords in general require pre- or early proto-stage, while most of the ParaSl loanwords only require middle proto-stage, so the temporal overlapping is only partial.

Based on the sound correspondences in the Baltic loanwords in Finnic, there seem to be at least two Baltic loanword layers. Petri Kallio calls the latter North Baltic, but its status either within West or East Baltic or as an independent branch is still unclear (Kallio 2008). The older layer is traditionally seen as Proto-Baltic, but it is uncertain, if there ever was a stage connecting only West and East Baltic after the separation of Slavic (Junttila 2016).

The "Early Baltic" loanwords in Finnic are here seen to represent Archaic Balto-Slavic and/or Early East Baltic, before the "Late Baltic" contacts were established with North Baltic (these could have overlapped, if represented different dialects). We know, that the Baltic-speaking area was earlier much wider, so any innovation shared by Latvian and Lithuanian might actually represent only a small part of the original East Baltic diversity. Proto-East Baltic should therefore be considered an earlier stage than "Proto-Latvo-Lithuanian", although no other East Baltic languages have survived.

Already Proto-Balto-Slavic loanwords in West Uralic can be supported (Kallio, forthcoming). These are included in the column of Archaic Balto-Slavic, in its earlier end. The scenario presented here makes geographical sense, because Finnic spread from the east to the west.
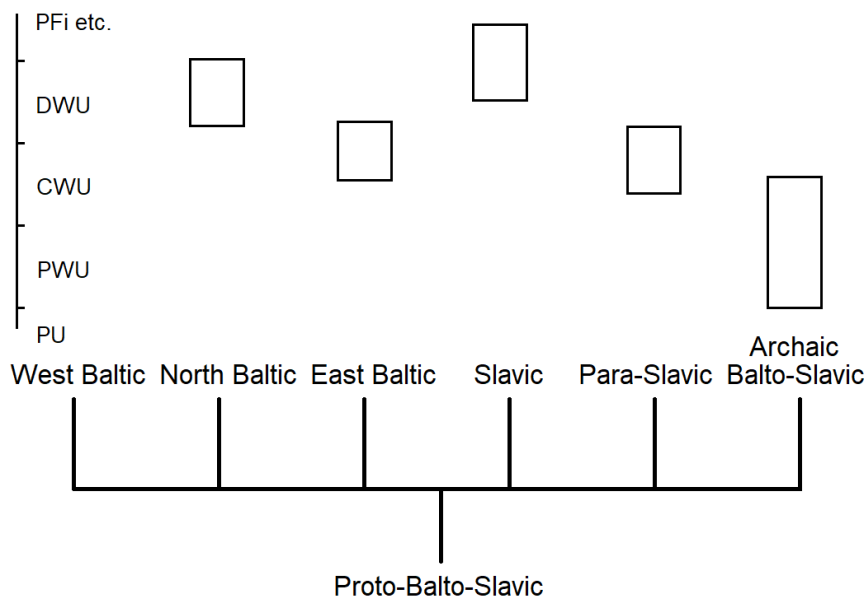

Figure 4. Early Balto-Slavic dialects
and the approximate chronology of the loanword strata in West Uralic.

Consequently, roughly during the first half of the 1st millennium BC, there probably was still/already at least six Balto-Slavic dialects: West Baltic, North Baltic, East Baltic, Slavic, Para-Slavic, and Archaic Balto-Slavic. The order of the branches roughly indicates their location in the west–east axle, although Slavic proper was the southernmost branch for a long time, before the first Slavic wave expanded northwards.

Considering the semantic fields of the loanwords, the layers considered in this article include words for plants, animals, terrain, weather, and tools and products of local livelihood – semantic fields, which are typical in substrate loanwords. This supports the model, in which West Uralic superseded Balto-Slavic dialects while spreading to the south and west.
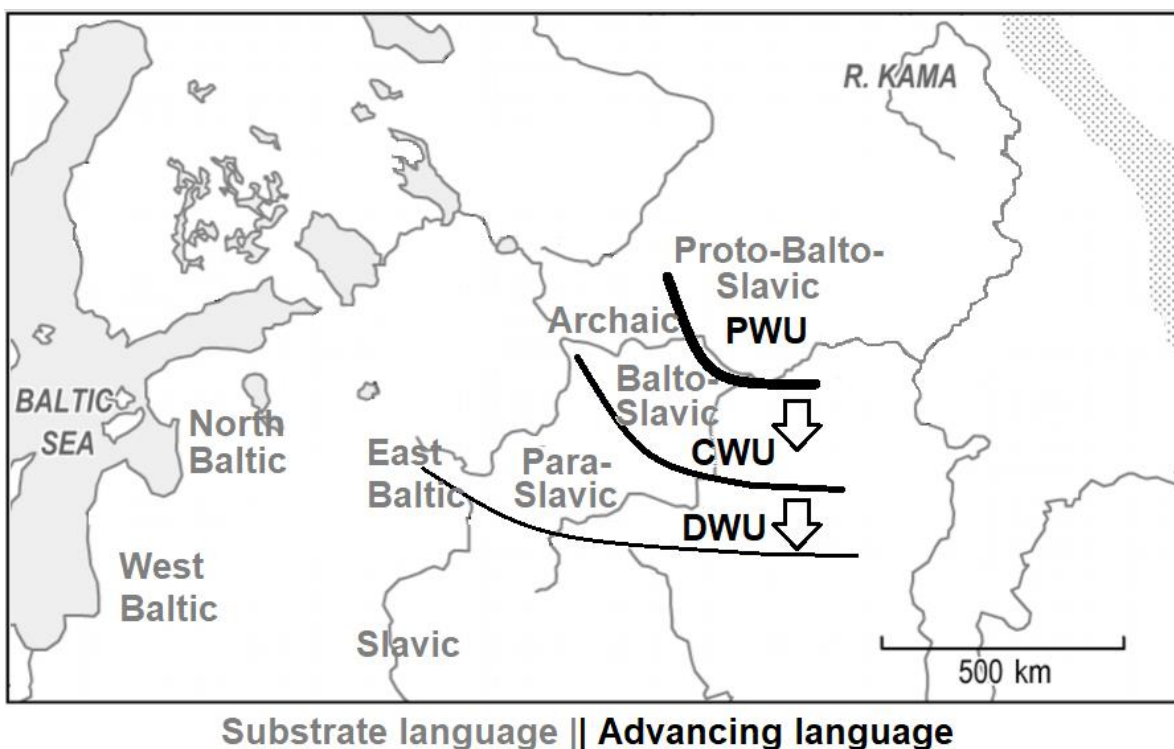

Figure 5. Stages of the West Uralic primary dispersal.

The southern borderline of the DWU stage in the map (Figure 5) is based on the toponymic studies by Pauli Rahkonen. In the eastern end it follows the earlier Mordvin area all the way to the Upper Oka in the west, and from there towards the Volkhov River it follows the West Chudic area (Rahkonen 2013: 241).

At this point, the absolute chronology can only be estimated very roughly. Proto-West Uralic can be located around the Upper Volga–Oka fork in the mid-second millennium BC (Häkkinen 2009: 37–40; Kallio, forthcoming), to where it arrived from the east. Proto-Balto-Slavic was there earlier, descending from the Northwest Indo-European dialect continuum, the spread of which is generally connected to the spread of the Corded Ware Cultures from the Daugava–Dniepr Region to the Middle Volga Region during the early 3rd millennium BC (the Fatyanovo Culture).

Recurring irregularities in West Uralic 1: Para-Slavic loanwords

Common West Uralic probably dates in the centuries around the beginning of the first millennium BC, spreading to the south and west to the Upper Volga–Oka Region. Here the contact language was Archaic Balto-Slavic and towards the end also Para-Slavic and the early stage of East Baltic.

Diverging West Uralic probably dates around the mid-first millennium BC. At this stage the contact languages were first Para-Slavic and the Early East Baltic, and later contacts with North Baltic were established, as Middle Proto-Finnic spread nortwards from the Daugava River.

The chronology presented here is considerably later than that of Valter Lang, according to whom Finnic and Saami spread to all the way to Finland quite early, already in the end of the second millennium BC (Lang 2020). However, the linguistic results do not require so early dispersal, but several centuries later datings would also be acceptable. As we still see Early Baltic loanwords shared by Finnic and Saami, and roughly contemporaneous Para-Slavic loanwords shared by Finnic, Mordvin, and even Saami, it is difficult to assume the final areal dispersal of the WU branches occurring much earlier than during the mid-first millennium BC.

Furthermore, the archaeologically perceivable expansions — even if they were connected to certain language speakers — cannot be taken as unambiguous signs of linguistic expansions. A group of migrating language speakers might spread to a new area (1) their language; (2) some loanwords; (3) no linguistic influence at all. All the subsequent migrations have all the same options, so it is impossible to see afterwards, which wave actually brought the language to the area. Lang's proposition is the earliest possible, but not the latest possible correlate for the Saami and Finnic linguistic expansions to Finland.

LITERATURE

Luobbal Sámmol Sámmol Ánte (Aikio) 2006: On Germanic-Saami contacts and Saami prehistory. — *Journal de la Société Finno-Ougrienne* 91: 9–55.
<http://www.sgr.fi/susa/91/aikio.pdf>

Luobbal Sámmol Sámmol Ánte (Aikio) 2015: The Finnic 'secondary e-stems' and Proto-Uralic vocalism. – *Journal de la Société Finno-Ougrienne* 95: 25–66.
<https://journal.fi/susa/article/download/82642/41883/>

Derksen, Rick 2008: *Etymological Dictionary of Slavic Inherited Lexicon.* Edited by Alexander Lubotsky. – *Leiden Indo-European Etymological Series*, Volume 4. Leiden/Boston: Brill.

Derksen, Rick 2015: *Etymological Dictionary of Baltic Inherited Lexicon.* Edited by Alexander Lubotsky. – *Leiden Indo-European Etymological Series*, Volume 13. Leiden/Boston: Brill.

Grünthal, Riho 2012: *Baltic loanwords in Mordvin. – A Linguistic Map of Prehistoric Northern Europe. Mémoires de la Société Finno-Ougrienne* 266. 297–343.
<https://www.sgr.fi/sust/sust266/sust266_grunthal.pdf>

Helimski, Eugen 2005: *The 13th Proto-Samoyedic Vowel. – Mikola-konferencia 2004*, Szeged: SzTE Department of Finnougristics.

Jaakko Häkkinen

Holopainen, Sampsa 2019: *Indo-Iranian Borrowings in Uralic. Critical Overview of the Sound Substitutions and Distribution Criterion.* Dissertation, University of Helsinki. <https://helda.helsinki.fi/handle/10138/307582>

Holzer, Georg 2003: *Urslavische Fonologie. – Wiener Slavistisches Jahrbuch*, Vol. 49 (2003).

Häkkinen, Jaakko 2007: *Kantauralin murteutuminen vokaalivastaavuuksien valossa.* Master's thesis, University of Helsinki. <https://helda.helsinki.fi/handle/10138/19684>

Häkkinen, Jaakko 2009: Kantauralin ajoitus ja paikannus: perustelut puntarissa. – *Journal de la Société Finno-Ougrienne* 92. 9–56. <http://www.sgr.fi/susa/92/hakkinen.pdf>

Häkkinen, Jaakko 2019: *Kantasuomen keskivokaalit: paluu.* Santeri Junttila & Juha Kuokkala (editors): *Petri Kallio rocks. Liber semisaecularis 7.2.2019.* Kallion etymologiseura.

Junttila, Santeri 2016: Die baltisch-slawische frage im lichte der alten baltischen lehnwörter des ostseefinnischen. – *Baltistica* Li(2) 2016, p. 217–238.

Junttila, Santeri 2017: Lähtökielen sanansisäisten soinnittomien klusiilien edustus kantasuomen balttilaislainoissa: Osa I. Edustus kohdekielen lyhyen vokaalin jäljessä. – *Journal de la Société Finno-Ougrienne* 96 (2017), p. 127–148. <https://journal.fi/susa/article/view/70223>

Junttila, Santeri 2020: *Kantasuomen balttilaislainat -käsikirjoituksen artikkeliluonnokset.* <https://sanat.csc.fi/wiki/Luokka:Kantasuomen_balttilaislainat_-k%C3%A4sikirjoituksen_artikkeliluonnokset>

Kalima, Jalo 1936: *Itämerensuomalaisten kielten balttilaiset lainasanat. – Suomalaisen Kirjallisuuden Seuran toimituksia* 202.

Kallio, Petri 2007: *Kantasuomen konsonanttihistoriaa. –* Jussi Ylikoski and Ante Aikio (eds.): *Sámit, sánit, sátnehámit. Riepmo ála Pekka Sammallahtii miessemánu 21. beaivve 2007. Mémoires de la Société Finno-Ougrienne* 253. 229–249. <http://www.sgr.fi/sust/sust253/sust253_kallio.pdf>

Kallio, Petri 2008: On the 'Early Baltic' Loanwords in Common Finnic. Alexander Lubotsky, Jos Schaeken & Jeroen Wiedenhof (editors): *Evidence and Counter-Evidence: Essays in Honour of Frederik Kortlandt 1. Balto-Slavic and Indo-European Linguistics*, p. 265–277. Studies in Slavic and General Linguistics 32. Amsterdam – New York.

Kallio, Petri 2012: The Non-initial-syllable Vowel Reductions from Proto-Uralic to Proto-Finnic. In: Tiina Hyytiäinen, Lotta Jalava, Janne Saarikivi, and Erika Sandman (eds.), *Per Urales ad Orientem: Iter polyphonicum multilingue. Festskrift tillägnad Juha Janhunen på hans sextioårsdag den 12 februari 2012*, pp. 163–175. Mémoires de la Société Finno-Ougrienne 264. <Helsinki 2012. https://www.sgr.fi/sust/sust264/sust264_kallio.pdf>

Kallio, Petri 2014: The Diversification of Proto-Finnic. – Joonas Ahola and Frog (eds.): *Fibula, Fabula, Fact: The Viking Age in Finland. Studia Fennica Historica* 18. 155-168. <https://www.academia.edu/10172149/The_Diversification_of_Proto_Finnic?auto=download>

Kallio, Petri (forthcoming): *The Proto-Balto-Slavic Loanword Stratum in the Uralic Languages.* Manuscript.

Kim, Ronald I. 2018: Balto-Slavic Phonology. – Herbert Ernst Wiegand (ed.): *Handbook of Comparative and Historical Indo-European Linguistics*, Volume 3. 1974–1985. Berlin/Boston: Walter De Gruyter.

Klotz, Emanuel 2017: Urslavisches Wörterbuch. Facultas Verlags- und Buchhandels AG.

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.

Korhonen, Mikko 1988: Uralilaisten kielten jälkitavujen vokaaliston historiaa. – *Virittäjä* 1988. 8–25. <https://journal.fi/virittaja/article/download/38186/46502/>

Recurring irregularities in West Uralic 1: Para-Slavic loanwords

Lang, Valter 2020: *Homo Fennicus – Itämerensuomalaisten etnohistoria.* Helsinki: Suomalaisen Kirjallisuuden Seura.

LÄGLOS = *Lexicon der Älteren Germanischen Lehnwörter in den Ostseefinnischen Sprachen.* Bd. I (1991), II (1996), III (2012). A. D. Kylstra, Sirkka-Liisa Hahmo, Tette Hofstra, Osmo Nikkilä. Amsterdam–Atlanta: Rodopi.

Lehtiranta, Juhani 1989/2015: *Yhteissaamelainen sanasto. – Mémoires de la Société Finno-Ougrienne* 200. <https://www.sgr.fi/sust/sust200.pdf>

Metsäranta, Niklas 2020: *Periytyminen ja lainautuminen. Marin ja permiläisten kielten sanaston-tutkimusta.* Dissertation, University of Helsinki. <https://helda.helsinki.fi/handle/10138/321695>

Paasonen, Heikki 1990, 1992, 1994, 1996, 1998, 1999: *Mordwinisches Wörterbuch.* Zusammen-gestellt von Kaino Heikkilä. Bearbeitet und herausgegeben von Martti Kahla. – *Lexica Societatis Fenno-Ugricae* XXIII: 1–6 / *Kotimaisten kielten tutkimuskeskuksen julkai-suja* 59. <https://www.mv.helsinki.fi/home/rueter/PaasonenMW.shtml>

Petit, Daniel 2018: *Balto-Slavic.* – Herbert Ernst Wiegand (ed.): *Handbook of Comparative and Historical Indo-European Linguistics*, Volume 3. 1960–1973. Berlin/Boston: Walter De Gruyter.

Pystynen, Juho 2018: *Itämerensuomen pitkien vokaalien alkuperä.* Master's thesis, University of Helsinki. <https://helda.helsinki.fi/bitstream/handle/10138/236213/Pystynen_Juho_Pro_gr adu_2018.pdf?sequence=2&isAllowed=y>

Rahkonen, Pauli 2013: *Southeastern Contact Area of Finnic Languages in the Light of Onomastics.* Dissertation, University of Helsinki. <https://helda.helsinki.fi/handle/10138/38908>

Sammallahti, Pekka 1998: *The Saami Languages. An Introduction.* Kárášjohka: Davvi Girji OS.

SMS = *Suomen murteiden sanakirja (a–lööveri). –* Kotimaisten kielten keskuksen verkkojul-kaisuja 30. <https://kaino.kotus.fi/sms/>

SSA = *Suomen Sanojen Alkuperä: Etymologinen sanakirja*, volumes 1–3 (1992, 1995, 2000). Erkki Itkonen & Ulla-Maija Kulonen (eds.). Helsinki: Suomalaisen Kirjallisuuden Seura.

Villanueva Svensson, Miguel 2016: Zero-grade denominative nasal and sta-presents in Baltic. — *Baltistica* Li(1) 2016, p. 37–59.