On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic *x

0. Introduction

The paper *Uralilaisen kantakielen sanastosta* (1981), authored by Juha Janhunen over three decades ago, is without doubt one of the most important and influential studies published in the field of Uralic phonological and lexical reconstruction. In this paper Janhunen critically evaluated the stock of proposed Proto-Uralic etymologies and based his analysis of historical phonology on a corpus of only 140 *Gleichsetzungen*; he presented a solidly argued reconstruction of Proto-Uralic phonology and showed that the number of valid Proto-Uralic etymologies is much smaller than what had been earlier assumed. Subsequent research has established new etymologies and amended details in Proto-Uralic phonology, but not changed these main results. In fact, current scholars working on Uralic etymology and historical phonology are in great debt to Janhunen; it is no exaggeration to say that he raised this line of research to a new level of methodological precision.

One of the many novel ideas in Janhunen’s paper is the hypothesis of Proto-Uralic *x* — a phoneme of unknown phonetic quality which Janhunen reconstructed to account for certain correspondences. According to Janhunen there is a regular correspondence between Finnic long vowels and Samoyed bisyllabic vowel sequences which consist of a full vowel followed by a schwa: e.g. PFi *keeli* ~ PSam *kääj* ‘tongue’. In Janhunen’s view, also this correspondence originated in the canonical Proto-Uralic root shape *CVCCV*, but it reflects a consonant cluster with *x* as the first member: PU *käxli* ‘tongue’. In addition, Janhunen assumes an intervocalic *x* for roots where Finnic and Samoyed monosyllabic vowel stems correspond to each other; Saami has retained the consonant as *k*: e.g. PFi *möö- ‘sell’ ~ PSam *mi- ‘give’ ~ P Saam *miekë- ‘sell’ (< PU *mexi*). In such cases a Proto-Uralic velar spirant *γ* had occasionally been suggested by earlier research.

This paper presents a critical evaluation of Janhunen’s hypothesis that Finnic long vowels correspond to Samoyed vowel sequences and reflect a PU vowel followed by preconsonantal *x*. It will be examined whether such a correspondence can indeed be reliably established, and whether PU *x* provides the best account for the genesis of Finnic long vowels and Samoyed vowel sequences.
1. The relationship between Finnic long vowels and Samoyed vowel sequences

Janhunen (1981) presents 14 etymologies for which he reconstructs PU presonantal *x; each etymology is critically evaluated below.

**kaxxi ‘spruce’** > PSam *kåət, PFP *koosi. — This etymology involves three phonological peculiarities: 1) The vowels in Fi *kuusi* and Md *kuz ‘spruce’* do not match the reconstruct *koosi. 2) The Ms cognate shows a consonant cluster *wt (E *kowt, W kowt, kɑ̃rɪt, N *xɔ̃wt < PMs *kawt). 3) The Samoyed cognates do not unambiguously point to a vowel sequence; Kosterkin & al. (2001) cite Ngan *koʔ* (pl. *koʔəʔ*), which suggests PSam *kåt. Strangely, though, the 3SG form is cited as *kuoʔtu*, which does not seem regular (*< *kåətå*). Prokofev cites Ngan *kuoʔ*, and Castrén gives *ku’a* (*= /kuʔə/)*; the historical phonology of these forms is difficult to interpret.

**käxli ‘tongue’** > PSam *käəj (= *keəj in Janhunen’s reconstruction), PFP *keeli. — This is one of the very few clear cases where a Samoyed vowel sequence corresponds to a Finnic long vowel: cf. Ngan *s`iíd’é ~ Est keel ‘tongue’. The reconstruction of PSam *-j is somewhat hypothetical, though; its reflex is encountered in Ngan *s`iíd’ə* (*< *kääjo*) and Mator *käɕ-tä ‘tongue.3sg* (*< *kä(ə)j-tä*), and these might involve a derivational element; this is also the case, for example, with Ngan *n`iéd’é ‘belt’ (*< *niə-jə; see the discussion below). Some other Samoyed languages show different derivational suffixes in this word: cf. EnF *sioro, Kam šikä ‘tongue’.

**d’ixmi ‘bird-cherry’** > PSam *jë(ə)m, PFP *d’oomi. — No vowel sequence can be established in this word; the attested cognates, SlkTa *ɛm and Kam *lem, can reflect PSam *jëm.

**ńixli ‘arrow’** > PSam *ńë(ə)j, PFP *ńooli. — This etymology involves the same problem as the previous one: the Samoyed cognates (NenT *-ni in tu`nii ‘rifle, shotgun’ (cf. tu ‘fire’), Slk *ńejo ‘arrow’, Kam *nie ‘arrow’, *ńa ‘bullet’, Mat *nej ~ *ńej ‘arrow’) do not show evidence of a vowel sequence. One can probably reconstruct PSam *ńejenis, even though the correspondences are somewhat deviant, perhaps because the back vowel *ę was flanked by palatal consonants.

**koxji ‘birch’** > PSam *koəj, PFP *koj-wV. — This cognate set shows no match between a Finnic long vowel and a Samoyed vowel sequence; it is entirely hypothetical that the -o- in Fi *koivu* would go back to a long vowel which was secondarily shortened before the consonant cluster *-jw-. Moreover, this interpretation operates with an unidentified derivational suffix *-wV. It should be added that the reconstruction of the PSam sequence *oəj is not clear; the vowel correspondence between NenT *xo, EnF *kua, Ngan *kua, Slk *qüə and Kam *küjü is unusual and apparently unique. Evidence for *-j in PSam is rather meagre; the only segmental reflex of *j would apparently occur in Kam *küjü, which looks

---

1. I follow the slightly modified reconstruction of Proto-Samoyed vocalism argued by Helimski (2005); see also Aikio (2006: 9–11).
like a derivative. The other suggested Finno-Ugric cognates (MdE kil'ej, M keļu, MsE kēl') are phonologically so obscure that they might be better treated as etymologically unrelated. Thus, the sound correspondence between Fi koivu and PSam *koəj remains unexplained for the time being, and one could even ask whether the etymology is really valid.

*tuxli ‘feather, wing’ > PSam *tua, PFP *tul-ka. — This example is similar to the word for ‘birch’: there is no actual match between a Finnic long vowel and a Samoyed vowel sequence, but only a hypothesis that the expected long vowel would have been shortened before a consonant cluster in the derivative *tul-ka. Here, too, the *-j in the PSam reconstruction can be questioned; none of the attested forms (NenT to, EnT tua, Ngan ěe, SlkTu tu, Ty K tū, Mat tua ~ tuga (toga?)) show direct evidence for it. The expected long vowel would, however, be seen in the PFP noun *tuuli ‘wind’, which was compared to the cognate set for ‘feather’ by Janhunen (1981: 241), but with hesitation. The connection is not entirely clear: even though semantic parallels can be found (e.g. English wind and wing, both ultimately based on PIE *weh₁- ‘blow’), one can ask why the PFP noun *tuuli ‘wind’ would exactly correspond to PSam *tua(j) ‘wing, feather’. The known parallels only show that ‘wing / feather’ and ‘wind’ are related concepts, but do not testify to a semantic change ‘wind’ > ‘wing / feather’ (or vice versa).

*peksi ‘edge, outside’ > PSam *piaj, PFP *peeli. — The PSam reconstruction *piaj appears problematic; it does match the Nenets and Enets forms (NenT pi-, EnF fio-), but Ngan h’ai, (Castrén) feai ‘end, edge’ points rather to *pājä. The Slk form *pō- with its back vowel is in any case irregular. Janhunen (1981: 241) states that the reconstruction *piaj best accounts for the Samoyed vowel correspondences, but as these correspondences are unique, there is no evidence of their regularity. Notably, Janhunen (1981: 242) postulates the same vowel sequence for PSam *tiej ‘pus, rot’, even though the vowel correspondences in this cognate set are different.

*sexji ‘pus’ > PSam *tiej, PFP *seeji. — The Samoyed vowel sequence *iə would account for NenT tīm- ~ Ngan čiim- ‘spoil’ (< PSam *tiə-m-). It is more difficult, however, to find grounds for reconstructing a stem-final *-j in Samoyed; and it should be noted that Slk *tē ‘pus’ shows a deviant vowel (< PSam ? *tā(o)(j)) — the expected reflex of PSam *iə is Slk *ū. As regards Finno-Ugric, the reconstruction of *ee is speculative, as none of the cognates show clear evidence of a long vowel. It is true that the vowel correspondences appear to be somewhat irregular: cf. SaaN siedja, MdE sij, MariW sū ‘pus’, Komi siś, siś, Udm sīś ‘rotten’, KhE lōj, MsE sāj, Hung év, év ‘pus’. Even so, the reconstruction of *ee does not remove these irregularities. The inconsistencies in vowel development may result either from the influence of *j or from the affective semantics, or both.

*nóxma ‘hare’ > PSam *ńāomā, PFP *ńoma-la. — This would be the only example of a cluster *xC in a Uralic *A-stem. The Samoyed forms do not, how-
ever, suggest a vowel sequence, as later admitted by Janhunen (2007: 221); the PSam form can be reconstructed as *ńåmå.

*pu/oxli ‘knee’ > PSam *puej, Finnic-Saami *pol-wi. — This etymology is similar to the words for ‘birch’ and ‘feather’ discussed above: there is no actual correspondence between a Finnic long vowel and a Samoyed vowel sequence, but it is merely assumed instead that Finnic-Saami *o derives from a long vowel that secondarily shortened in the derived form *pol-wi. Moreover, the element *-wi is no known derivational suffix. Here, too, the Samoyed forms lack clear evidence of a stem-final *-j, so PSam *pue can probably be reconstructed instead. This root is only attested in derivatives and obscured compounds: 1) NenT pu$lië, NenF punLa¨ ‘knee’, Slk *pūlə (Ta puli), Mat hulu < PSam *puu(-n)-IV, apparently compounded with PSam *lē ‘bone’; 2) EnF juu<se (? < PSam *sajmå ‘eye’); 3) Mat ?hoŋoaj ~ ?hoŋoj ‘knee, shinbone’ (attested forms: <hongoi>, <hooi>, <ñoi>, <ñoœi[-]>, where -oj perhaps < PSam *âj ‘foot’; 4) Ngan hüagaj ‘knee’, with an obscure element -gaj. This gives reason to ask whether the Finnic-Saami form *polwi should actually be analyzed as cognate with the first subset deriving from PSam *pue(-n)-IV. The sound correspondence is of course not regular, but if this hypothesis is correct, a PU compound ? *po/u(x/wi)(-n)-luwi (or the like) can be reconstructed, consisting of ? *po/u(x/wi) ‘knee’ and *luwi ‘bone’. Also Janhunen (2000: 72) suggests that already in PU there would have been a compound *puxi-lixi ‘knee-bone’; he does not, however, consider Finnish polvi the exact cognate of PSam *pue(-n)-IV.

? *üxji ‘belt’ > PSam *ńioj ~ *jioj, PFi *vōō. — The reconstruction *üxji cannot be justified, as SaaN avvi, Fi vyō and Hung öv ‘belt’ regularly reflect PU *iwä. Contra Janhunen (1981: 260), SaaN avvi cannot be a loanword; borrowing from Finnic would predict a Saami reflex *vivēv or *viejē. The PSam root is best reconstructed as *niø ~ *jio, where both *n- and *j- are prothetic consonants. The phonological development was probably *iwä > *iwa > *iâ > *iœ > PSam *nio ~ *jia. EnT niiojo and Ngan niad?q point to PSam *njia, where *-jœ may be a suffix; cf. Ngan muod?q (< *moajœ) vs. NenT mo (< *moa) ‘branch’.

? *uxt northern ‘track, way’ > PSam *uut, Ugric *uC(V)t(t)IV. — The reconstruction of a cluster *uxt is not justified, as there is no evidence of a vowel sequence: Ngan ṣuta, NenT ŋu?, EnF u? < PSam *u(t)ə. At best one could reconstruct PU *uktı, but it is questionable whether this is a valid cognate set at all; the vowel correspondence between KhE ɣar ‘isthmus’, MsE êkt ‘outflow’, Hung ŋt ‘way’ and PSam *u(t)ə is not regular. Saarikivi (2004: 349) suggests that an otherwise lost Finnic cognate is preserved in Karelian and Vepsian place-names of the shape Uht-; if this hypothesis is correct, PU *ukti can be reconstructed, but the vocalism of the Ugric cognates must be regarded as irregular.

? *KVeV(x)(l)V- ‘die’ > PSam *kåœ-, PFi *koole-

? *nVeV(x)(l)V- ‘lick’ > PSam *ńåœ-, PFi *noole-

These two etymologies are considered irregular by Janhunen (1981: 263–264), but the cognate sets display the same vowel correspondence in most languages, and this can hardly be attributed to chance:
On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic *x

○ SaaN goallu- ‘freeze, feel cold’, Fi kuole-, MdM kulɔ-, MariW kọle-, Komi kul-, Udm kuli-, KhE kāla-, Ms kōl-, Hung hal, PSam *kāo- ‘die’
○ SaaN njoallu-, Fi nuole-, MdM nola-, MariW nôle-, Komi nul-, Udm nūli-, KhE nāla-, MsE nālant-, Hung nyal, PSam *ńąo- ‘lick’

In our current framework of Uralic historical phonology, however, this vowel correspondence lacks a satisfactory explanation. The two examples suggest that this is a result of some kind of regular development, but how the PU forms should be reconstructed remains unclear.

The fourteen etymologies discussed above involve nine roots for which a vowel sequence can be reliably reconstructed in Samoyed, and seven roots which show a long vowel in Finnic (this figure does not include PFi *vöö ‘belt’, as it is well-known that Finnic monosyllabic vocalic stems are secondary and go back to roots of the shape *(C)VCV-). In only four instances does a Finnic long vowel actually correspond to a Samoyed vowel sequence.

PFI *keeli ~ PSam *kāo(j) ‘tongue’
PFI *peeli ‘outer edge; post’ ~ PSam *piə ‘outside’
PFI *koole- ~ PSam *kāo- ‘die’
PFI *noole- ~ PSam *ńąo- ‘lick’

It is notable that in all cases the intervocalic consonant is -l- in Finnic. This already suggests the correspondence should be explained in some other way than by postulating a PU cluster *xl.

Let us hypothetize that Finnic long vowels and Samoyed vowel sequences are diachronically unrelated phenomena after all, and that their correspondence in the four word roots above is only apparent. This hypothesis predicts that there are also word roots where a Finnic word with a long vowel has a Samoyed cognate that demonstrably lacks a vowel sequence. In practice such examples are somewhat difficult to find because vowel sequences were only preserved in Nganasan and partially in Enets; in other Samoyed languages only indirect traces of vowel sequences are occasionally found. Hence, in most cases only a Nganasan cognate can conclusively prove the absence of a vowel sequence. Despite this limitation two examples can apparently be found:

PSam *jämVjo (> Ngan demə ‘porridge made of blood, meat, flour and water’, Nent jewej ‘soup, broth’) ~ Fi liemi, SaaN liepma ‘broth’, MdM lâm, MariW lem, Udm lìm ‘soup, broth’, MsE lōäm ‘soup, thin porridge’, Hung lē ‘liquid, juice’. This is a new etymology; the development *l- > PSam *j- is regular.

PSam *cëen (?) *tën (> Ngan taj, Nent te?, EnF ti?, Slk *čën, Kam ten, Mat ten) ~ Fi suoni, SaaN suotna, MdM san, MariW šiun, Komi, Udm sen, KhE lan, jan, MsE tèn, Hung in ‘sinew’. — The problem with this well-known equation, which has been rejected by Janhunen (1981) and Sammallahti (1988), is the unexpected affricate in the Slk cognate; one would expect Slk *tën. But as the correspondence is otherwise precise and the word is in any case attested in every Finno-Ugric language, it would be quite unusual if the Samoyed word were of
different origin after all. A similar unexpected affricate as the reflex of PU *t is found in SlkK čumžu ‘riddle’ (~ Ngan tumtə ‘riddle’, tumtə- ‘guess’, NenT tumtə- ‘know’, En tudda- ‘guess, find out’, Kam təmna- ‘know’, Mat tumdo- ‘recognize’ < PU *tumti-). If the Samoyed cognates of Fi suoni are rejected due to the Selkup affricate, the demand of consistency requires the rejection of the Samoyed cognates of Fi tunte- ‘feel, know’ as well.

2. The origin of Finnic *ee and *oo

The material discussed in the previous section indicates that separate historical accounts are needed for the origin of Finnic long vowels and Samoyed vowel sequences. It is well-known that the occurrence of long vowels in Finnic stems of Uralic origin is subject to strict phonological restrictions: they only occur in stems of the type *CVV- and *(C)VVCe-. In the latter type (the so-called “primary long vowels”) the quality of the vowel is restricted to e, i, o and u (a, ä, ö and ü do not occur). Monosyllabic stems of type *CVV- have, in turn, developed secondary long vowels from earlier bisyllabic stems of the type *CVCV- due to loss of intervocalic *w, *j, *x and *ŋ.

Let us first examine the background of Finnic *ee and *oo. Lehtinen (1967) has suggested that *ee and *oo in stems of the type *(C)VVCe- have arisen through a process of secondary lengthening. According to her, non-high vowels would have been lengthened before single voiced consonants in Pre-Finnic *e-stems, after which long low vowels merged with long mid vowels. She postulates the following development:

*ńale- > *ńaale- > PFi *noole- ‘arrow’
*kole- > *koole- > PFi *koole- ‘die’
*käle- > *kääle- > PFi *keele- ‘tongue’
*ńele- > *ńeele- > PFi *neele- ‘swallow’

Lehtinen’s study has largely overlooked by later research. Only recently Reshetnikov and Zhivlov (2011: 97) have followed some of her ideas, and suggest that “Proto-Balto-Finnic *ō is [...] the result of regular lengthening of *a in Balto-Finnic e-stems before intervocalic *r, *l, *m, *n and *ð.” They do not, however, attempt to substantiate this hypothesis with more detailed analysis of data.

There are some obvious shortcomings in Lehtinen’s explanation, and perhaps because of this it has not received the attention it deserves. Exceptions such as PFi *mene- ‘go’, *pel-ko ‘fear’ (← *pele-) and *vere- ‘blood’ Lehtinen attributes to a change *i > *e after labial consonants; the proto-forms of these words would thus have been *mine-, *pile- and *wire-. Even more hypothetical explanations are offered for the lack of lengthening in PFi *ole- ‘be’ and *mone- ‘many’; Lehtinen suggests the possibility of analogy (cf. on ‘be’ 3sg, ovat 3pl,
which do not fulfill the conditions for lengthening), and points to the Finnish dialectal form *munta ‘many’ (~ standard *monta), which would reflect the supposedly original form *mune-. Lehtinen also leaves unmentioned some etymologies which do not fit her scheme of vowel development: cf. Fi keri ~ SaaN garra ~ MdM ker `bark, rind’ and Fi ääni ~ SaaN jietna ‘sound, voice’. As Lehtinen argues her explanation only on the basis of Finnish, Saami and Mordvin data, word roots of Uralic origin which lack Saami and Mordvin cognates (such as PFi *voote- ‘year’ and *pane- ‘put’) are left without explanation.

Despite these problems it appears that Lehtinen’s attempt to explain Finnic long vowels as results of secondary lengthening was on the right track after all, and that plausible explanations can be offered for apparent exceptions. First, her hypothesis must be modified so that lengthening is assumed to have affected only the low vowels *ä and *a, and that no lengthening of *e and *o ever took place in Pre-Finnic. The lengthening is conditioned by a postvocalic single voiced consonant and the e-stem, as Lehtinen suggests. At least the following examples are known:

*ad'i- ‘bed’ ? > Fi (derivative) vuode : vuotee- ‘bed’ (~ Komi val ‘reindeer hide (for sleeping on it)’, Udm wal’i- ‘spread; make the bed’, MsS álát, Hung ágy ‘bed’. — The etymology is uncertain, as vuode could also be a derivative of the Baltic loanword vuota ‘hide’.

*jani ‘path’ > Fi juoni ‘plot; row’ (~ SaaS joene ‘way’, MdM jan ‘path’). — A loan from Proto-Aryan *yāna- > Sanskrit yāna- ‘going, travel; vehicle; way’ (Koivulehto 2009: 81).

*kari ‘skin?’ > Fi kuori ‘bark, crust’ (~ NenT šar ‘skin (under the hair); surface’ < PSam *kar). — This etymology was, with hesitation, briefly suggested in Aikio (2002: 50).

*wari ‘forest / hill’ > Fi vuori ‘mountain, rocky hill’ (~ Komi ver ‘forest’, Udm vir ‘hill, highlands’, KhE wör ‘ridge along a river’, MsE wör ‘evergreen forest’). — Koivulehto (1999: 218) has proposed that the word is a loan from Proto-Aryan *aras- (~ Greek óros ‘mountain’; the word is not attested in Aryan). The suggestion is phonologically problematic, however; Koivulehto postulates the Uralic proto-form *orV, which accounts neither for the long vowel in Finnic nor for the initial *w- attested in all the cognates. — It is interesting that one can also reconstruct the very similar PU noun *wara ‘mountain, hill’ (> SaaN várri ‘mountain, hill’, KhE ur ‘wooded ridge’, MsE wör ‘mountain, peak’, Ngan bért ‘mountain, cliff’, Kam bør ‘mountain, ridge’; Aikio 2006: 27–28). Obviously there is some kind of etymological connection between PU *wara and *wari, but its nature remains unexplained; the two words may have been in an obscured derivational relationship already in Proto-Uralic.

*idi ‘year / autumn’ > *adi > Fi vuosi ‘year’ (~ Komi vo ‘year’, Udmwapum ‘time, moment’, KhE al ‘year’). — A new Samoyed member can be added to this cognate set: PSam *érö > NenT ērō, SlkTy ara, Kam ere, Mat órōh ‘autumn’ (as for the semantics, cf. Sanskrit sārāḥ ‘year; autumn’). Koivulehto (1999: 218–219) has suggested a borrowing from Proto-Aryan *vatas- / *vatah-
‘year’; the loan etymology must be rejected for phonological reasons, though. Koivulehto posits the Uralic form *wotV, which is in several ways problematic: 1) Fi v- and Udm w- are secondary prothetic consonants, and the original vocalic initium can be seen in Khanty and Samoyed; 2) Khanty and Samoyed demonstrate that the medial consonant was *d and not *t; 3) The reconstruction *wotV accounts neither for the long vowel in Finnic nor for the prothetic w- in Udmurt, as PU *w- is reflected as Udmurt v- and not w.-2

*ńili ‘arrow’ > *ńali > Fi nuoli ‘arrow’ (~ SaaN njuolla, MdM nal, MariE nölo-pikš, Komi, Udm niel, KhE nal, MsE něl, Hung nyil, Kam nie, Mat ňej ‘arrow’)

*ńiri ‘moist, wet (?)’ > *ńari > Fi nuori ‘young’ (~ SaaN njuoras ‘weak (e.g. of an infant)’, MariW nörö ‘flexible, pliable’, nöre- ‘get wet’, Komi nür ‘bog’, Udm nür ‘moist; moistness; bog’, Hung nyirkos ‘moist, humid, damp, raw’, Tenn nër ‘sap; white of an egg’, SlkK nör ‘semen’; Aikio 2006: 20–21). — The semantic development may have been approximately ‘wet, moist’ > ‘soft’ > ‘weak’ > ‘young’. Previously SaaN nuorra ‘young’ has been cited as a cognate of Fi nuori, but it must be a Finnic loanword.

*sïni ‘vein, sinew’ > *sani > Fi suoni (~ SaaN suohtna, MdM san, MariW šùn, Komi, Udm sen, KhE lan, jan, MsE țen, Hung in, Tenn te?, Ngan ţan)

*šimi ‘scale’ > *sami > *soome- > Fi (der.) suoma (~ SaaN ċuoppma ‘fish skin’, MdM šav ‘money’, MariW šùm ‘scale; tree bark’, Komi šem ‘scale; money’, Udm šem, KhE sam, MsE sêm ‘scale’)

*käli ‘tongue’ > Fi kieli ‘tongue, language’ (~ SaaN giella ‘language’, MdM käl’, Komi, Udm kil ‘tongue, language’, KhE köl ‘word, speech, language’, MsE köälmt- ‘inquire, find out’, Tenn sé, EnF sioro, Ngan šıad’ọ, SlkTa še, Kam šiká ‘tongue’)
The sound law outlined above has apparently no such exceptions where the vowel *a or *ä would have been lengthened in Finnic even though no voiced consonant followed. Fi *kuusi ‘spruce’, which according to Janhunen (1981: 240) derives from PU *kaxsi, does not serve as a counterexample due to the multiple irregularities in this cognate set (see the discussion in section 1 above). The examples above reveal that at least PU *m, *n, *l, *r, as well as *d and *d’ the reflex of which was probably a dental spirant (*δ) in Pre-Finnic, have triggered the lengthening of a preceding low vowel in *e-stems. Hence, the result validates Reshetnikov and Zhivlov’s (2011: 97) suggestion regarding the conditions of the shift *a > *oo in Pre-Finnic.

The effects of dorsal voiced consonants on Pre-Finnic *a and *ä require further examination. At least before *ŋ vowel raising did not take place, as indicated by the following examples:

- *jäŋi ‘ice’ > Fi jää (~ SaaN jiekŋa, MdM äj, jäį, MariW i, Komi ji, Udm je, KhE jönk, MsE jöändk, Hung jég ‘ice’)
- *päŋi ‘head, end’ > Fi pää ‘head; end’ (~ MdM pe, Komi pom, Udm pum ‘end’, MsE pöänk, Hung fő ‘head’)
- *säŋi ‘weather, sky’ > Fi sää ‘weather’ (~ Hung ég ‘sky’)
- *kaŋiri ‘boat rib’ > Fi kaari ‘curve; boat rib’ (~ Saal kuonār ‘boat rib’)

The role of the glide *j in the change is ambiguous. The interpretation of the material is complicated by the small number of etymologies with the sequence *-ji-, as well as by the fact that in stems of ending in *-ji-, a long vowel would in any case have been secondarily shortened in Finnic as the glide was vocalized into -i-. The following three cases are known:

- *täji ‘louse’ > Fi tä. — This etymology suggests that at least vowel raising did not take place before the sequence *-ji-; otherwise the expected development would have been *täji > *tääji > *teeji > Fi *tei. On the other hand, the word shows phonological peculiarities in other branches: the vowels in the Permic cognates (Komi toj, Udm tej) are entirely irregular, and the Ugric forms (KhE toýtóm, MsE t̥öötm, Hung tetű) reflect a derived form tā(ji)ktVmV, without a trace of the sequence *-ji-.³

- *ka/oji ‘dawn’ > Fi koi. — The vocalism in this cognate set is difficult to reconstruct; the vowel correspondences in the cognates have been partly obscured by derivational elements attached to the root: 1) Komi kía ‘the red of dawn; sunset’; 2) KhE kuńól, kuńol ‘red sky (at dawn or dusk)’, Hung hajnal ‘dawn, daybreak’; 3) KhE, MsE kotól ‘day’; 4) MsN _VM- ‘shine, shimmer, flash’. If one postulates *kaj as the original form, one can assume an obscured derivational relationship to PU *kaja ‘sun, dawn’ (> SaaL guojiđ- ‘rise (of

---

³ The putative Saami cognate, SaaN dihkki ‘louse’ (< PSaa *tikkē), is much better explained as a loan from Germanic *tikkōn- > English tick, German Zecke ‘tick’. A semantic shift ‘tick’ > ‘louse’ is quite easy to understand. The Uralic etymology of dihkki presupposes a unique irregular development: PSaa *ti- would be a truncated reflex of PU *täji and the part *-kkē a derivational suffix.
moon); dawn (of day), Fi kajasta- ‘dawn, shimmer’, MariW kaja- ‘come in sight’, Ngan kou, Kam kuja ‘sun’, SlkTa qeči ‘heat’).

? *waji ‘fat’ > Fi voi ‘butter’. — The reconstruction *waji is based on the vowel correspondence between Ugric, Mordvin and Saami: SaaN vuodja ‘butter, grease’, MdM vaj, KhE woj ‘fat, butter, oil’, MsE waj ‘fat’, Hung vaj ‘butter’. The Permic and Mari forms do not reveal much: the front vowel in MariW ü ‘fat, butter, oil’ is in any case irregular, and Komi vij and Udm vej ‘butter, oil’ are not even mutually in regular correspondence, so the reconstruction of the Proto-Permic vowel remains uncertain. If the reconstruction *waji is correct, one could assume a development *waji > *waaji > *wooji > voi in Finnic. On the other hand, the labial vowel could also be a result of irregular rounding caused by the initial labial glide (*vai > voi). — Koivulehto (1999: 217–218) has proposed that the Uralic word is a loan from Aryan *ag`ya- / *a$g`ya- > Old Indic ā́jva- ‘melted butter used for oblations’. This explanation is problematic for the same reason as the Aryan etymology suggested for Fi vuori ‘mountain, rocky hill’ which was discussed above: Koivulehto postulates the reconstruction *ojV, which does not explain the initial *w- attested everywhere in Uralic.

On the other hand, the labial vowel could also be a result of irregular rounding caused by the initial labial glide (*vai > voi). — Koivulehto (1999: 217–218) has proposed that the Uralic word is a loan from Aryan *ag`ya- / *a$g`ya- > Old Indic ā́jva- ‘melted butter used for oblations’. This explanation is problematic for the same reason as the Aryan etymology suggested for Fi vuori ‘mountain, rocky hill’ which was discussed above: Koivulehto postulates the reconstruction *ojV, which does not explain the initial *w- attested everywhere in Uralic.

On the basis of the material above it seems clear that in Pre-Finnic *e-stems a sound change *a > *oo and *ä > *ee took place before single non-dorsal voiced consonants (at least *m, *n, *l, *r, *ð). Next, I shall discuss potential counterexamples for this development; for the most part they turn out to be only apparent.

Fi ääni ‘sound, voice’. — The word is cognate with SaaN jietna ‘sound, voice’ and Hung ének ‘song’ (< PU *äni). The Finnic cognate shows the vowel lengthening (*äni > *ääni) but lacks the expected raising (*ääni > *eeni). However, no Finnic language has word roots for which Proto-Finnic initial *ee- could be reconstructed. Finnish words with initial ie- do not reflect PFi *ee-, but instead represent the weak grade of a stem of the type *ikeC-: ien ‘gum’ (< *iken), ies ‘yoke’ (< *ikes). Hence, it can be assumed that the raising *ää > *ee did not take place in word-initial position.

Fi kääri- ‘wrap’. — The Finnic word is apparently cognate with SaaU giarastahtee- ‘catch with a lasso; tie, bind’, MdM kärks ‘string (e.g. of pearls)’, kärma ‘bunch’, MariW kerä- ‘put in, stick into’, Komi kert-., Udm kertti- ‘tie, tie up’, KhE keč, MsN kwarok ‘bunch, bundle’, NenT sør-, SlkTa sør-, Kam sør-, Mat ker- ‘put on (clothes), dress’ (Aikio 2002: 18–20). The raising of the vowel may have been blocked by the derivational suffix -i-; the presupposedly primary *e-stem (*kääre-) is not attested in Finnic. As another possibility one can think of avoidance of homonymy with PFi *keeri- ‘roll’ (> Fi kieri-), which has no cognates outside Finnic.

Fi sali- ‘chop wood shingles’. — SSA cites SaaN čūolla- ‘chop (esp. wood)’ (< *šali-) as the only certain cognate of the Finnish verb. In principle one could attribute the lack of vowel lengthening to the derivational suffix -i-. Probably, however, this is a false etymology. The distribution of Fi sali- is limited to

4. SaaL guojijd- has been incorrectly cited as a cognate of Fi kuu ‘moon’ (SSA s.v. kuu; UEW: 211).
On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic *x

Tavastia, Satakunta and Southern Ostrobothnia dialects, and no cognates are found elsewhere in Finnic. As this dialectal verb bears a very close resemblance to the more widespread Fi säli- ‘chop (e.g. wood shingles)’, it may simply be a recently developed, irregular back vocalic variant. Fi säli-, in turn, ultimately derives from PU *śälä- ‘cut’, with widely attested cognates: SaaN čálli- ‘cut; write’, KhE sül- ‘cut open’, MsE silt- ‘cut’, Hung szel ‘cuts, slits, carves’, Ngan šelj ‘sharpness’, šel’i- ‘sharpen’.

Fi sälyttä- ‘load, put a burden on’. — According to Sammallahti (1988: 548), this verb is a derived from PU *säla-, and cognate with Komi sel-, KhE lel-, jel-, MsE töäl- ‘mount (a horse), board (a boat or sledge)’, archaic Hung ellik ‘mounts’; SSA doubts the etymology, but without good reason. The reconstruction *säli- would predict a development *säli- > *sääle- > *seele- in Finnic. However, the reconstruction of the original stem vowel seems to be guesswork; one could equally well posit the form *sälä- as the starting point. Moreover, in Finnic the word is only attested in derivatives such as Fi sälyttä- ‘load’ and Est sältis (< *sälüttüs) ‘burden, load’, and the derivational suffixes attached may have blocked the vowel lengthening and raising which only took place in *e-stems.

Fi pane- ‘put’. — This verb seems to be the only genuine counterexample, as it really is a Finnic *e-stem in which a short low vowel occurs before a single non-dorsal voiced consonant. But as noted by Reshetnikov and Zhivlov (2011: 97), even this example is somewhat unclear, as also the proposed Permic and Ob-Ugric cognates (Komi pen- ‘fuck’, Udm poní- ‘put; fuck’, KhE pân-, MsE pon- ‘put’) show irregularity in vowels. Moreover, paradigmatic analogy provides a potential explanation. The predictable regular development would have been blocked in Pre-Finnic forms based on the consonant stem (*pan-), resulting in PFi paradigmatic alterations such as *pan-dak INF : *poone-n 1SG (< Pre-Finnic *pan-tak : *pane-n). In Finnish this would yield the hypothetical alteration panna : *puonen. Such alterations, of course, do not occur in Finnish or in any other Finnic language; hence, it must be assumed that they were levelled via paradigmatic analogy before the breakup of Proto-Finnic.

Thus, the Pre-Finnic conditioned sound changes *ä > *ee and *a > *oo have only one counterexample, Fi pane- ‘put’. As even this exception has a possible, albeit hypothetical explanation, there is no obstacle to treating the changes as regular.

There is one more important piece of evidence which supports the hypothesis of Pre-Finnic sound changes *ä > *ee and *a > *oo. As is well-known, Finnic languages possess a group of word roots which secondarily developed into *e-stems in Pre-Finnic; the extra-Finnic cognates of these roots namely seem to point to an original *a- or *ä-stem. This change of stem type is nearly always accompanied by a change in either the quality or the quantity of the first-syllable vowel. Three subtypes of this process have been traditionally distinguished:

5. I am obliged to Mikhail Zhivlov (personal communication) for this suggestion.
1) PU *CoCCa- > PFi *CaCCe-, e.g.:  
PU *komta > Fi kansi : kante- ‘lid’ (~ SaaN goavdi ‘canopy’,  
MdM kunda, MariW komōš ‘lid’, Komi šin-kud ‘eyebrow’)  
PU *šorwa ‘antler, horn’ > Fi sarvi : sarve- (~ SaaN čoarvi,  
MdM šura, MariW šur, Komi, Udm šur, Hung szarv; a loan from  
Proto-Aryan *šrwa-)  

2) PU *CoCa- > PFi *CooCe-, e.g.:  
PU *śorwa ‘antler, horn’ > Fi sarvi: sarve- (~ SaaN čoarvi,  
MdM s`ura, MariW s=ur, Komi, Udm s`ur, Hung szarv; a loan from  
Proto-Aryan *śrwa-)  

3) PU *CäCCä- > PFi *CaCCe-, e.g.:  
PU *säppä ‘bile’ > Fi sappi: sappe- (~ SaaN sáhppi, MdM  
säpä, Komi se1p, Udm sep, MsE töäp, Hung epe)  
PU *tälwä ‘winter’ > Fi talvi: talve- (~ SaaN dálvi, MdM t`ala,  
MariW tel, Komi tel, Udm tol, KhE těľγ, MsE töäl, Hung tël)  

Two other subtypes have, however, remained previously unnoticed:  

4) PU *CäCä- > PFi *CaCe-  
 PU *käsä > Fi kasi : kase- ‘cold mist; steam; smoke, coal gas’  
(~ SaaS gaasoe ‘frost mist rising from a lake or a river’, KhE kelo,  
MsE kṑlʾəl’-wit’ ‘dew’, NenT sėdəʔ, SlkK qa$ttaŋ ‘hoarfrost’; Aikio  
2009: 72–73)  

5) PU *CäCä- > PFi *CooCe-  
PU *pälä ‘side, half’ > Fi puoli : puole- (~ SaaN bealli, MdM  
päl’, pälä, MariW pel, Komi pel, Udm pal, KhE pelək, MsE pōäl,  
Hung fël, NenT pēlə, Ngan heli). — Fi puoli has not usually been  
considered a reflex of PU *pälä, and Koivulehto (1987: 202–204)  
proposes that it instead goes back to Pre-Finnic *pola and is a loan  
from Balto-Slavic *palu- > Russian пол ‘gender; half’. This etymology  
is formally flawless, but it would be quite surprising if Finnic  
*pooli ‘side, half’ had no connection to PU *pälä ‘side, half’, refle-

xes of which are attested in every other Uralic language.  

PU *lämä ‘rash, scab’ > Fi luomi : luome- ‘birthmark; eyelid’  
(~ MdE leme ‘rash, scab’, MariW lim, Komi lem, Udm lom ‘scab’).  
— Fi luomi has not earlier been included in this cognate set. As  
regards semantics, it is important that Karelian luomi has the mean-

ing ‘some skin disease; chap’.  

PU *kälä- ‘wade’ > Est koole : gen koolme ‘ford’ (~ SaaN gálli-,  
MdE kāl’, MariW kela-, Komi kel-, Udm kol- ‘wade’, KhE kūl-,  
MsE kōal- ‘rise, get up; go ashore’, Hung kel ‘rises, gets up’). — This  
is a new etymology for the Est word. The verb root has not been pre-
served in Finnic as such, perhaps because it became homonymous  
with PFi *koole- ‘die’.6  

6. In many references Fi kaalaa- ~ kahlaa- ‘wade’ is cited as a possible reflex of PU *kälä-, but due to  
phonological reasons it must be a loan from Saami (Wickman 1968; Aikio 2009: 74–75).
On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic *x

The overall picture of the development of secondary *e-stems becomes much simpler if one assumes that originally there were only two different subtypes to this development: *Co(C)Ca- > *Ca(C)Ce- and *Cä(C)Cä- > *Ca(C)Ce-. After this the regular change *CaCe- > *CooCe- took place, if the medial consonant was a voiced non-dorsal consonant. The development would have been as follows:

<table>
<thead>
<tr>
<th>PU</th>
<th>Pre-Finnic</th>
<th>PFi</th>
</tr>
</thead>
<tbody>
<tr>
<td>*o–a &gt; *a–e:</td>
<td>*šorva</td>
<td>*sarve-</td>
</tr>
<tr>
<td>*šola</td>
<td>*sale-</td>
<td>*soole-</td>
</tr>
<tr>
<td>*ä–ä &gt; *a–e:</td>
<td>*tälwä</td>
<td>*talve-</td>
</tr>
<tr>
<td>*käšä</td>
<td>*kase-</td>
<td>*kase-</td>
</tr>
<tr>
<td>*pälä</td>
<td>*pale-</td>
<td>*poole-</td>
</tr>
</tbody>
</table>

Thus, the sound laws proposed above significantly simplify the picture of the development of the so-called secondary *e-stems in Finnic. Even though some parts of this process remain to be accounted for, the examples discussed above still serve as evidence for the hypothesis that PFi *ee and *oo developed from earlier *ä and *a through lengthening and raising under specific phonological conditions.

Next, the correspondents of Finnic *ä, *ee, *a and *oo in other Finno-Ugric languages need to be examined. One can immediately note that at least Saami, Mordvin, Khanty and Mansi offer no evidence in support of the archaism of the Proto-Finnic oppositions *ä : *ee and *a : *oo. As the table below shows, the correspondents of the two vowels are identical in each case:

<table>
<thead>
<tr>
<th>Fi</th>
<th>SaaN</th>
<th>MdM</th>
<th>KhE</th>
<th>MsE</th>
<th>PFU (Sammallahti 1988)</th>
</tr>
</thead>
<tbody>
<tr>
<td>*kieli</td>
<td>giella</td>
<td>käl'</td>
<td>kół</td>
<td>kōálmt-</td>
<td>*keeli ‘tongue’</td>
</tr>
<tr>
<td>*liemi</td>
<td>liepma</td>
<td>l’äm</td>
<td>lōäm</td>
<td>*leemi ‘broth’</td>
<td></td>
</tr>
<tr>
<td>*käsi</td>
<td>giehta</td>
<td>käd’</td>
<td>köt</td>
<td>kōät</td>
<td>*käti ‘hand’</td>
</tr>
<tr>
<td>*jää</td>
<td>jiekņa</td>
<td>āj, jāj</td>
<td>jōŋk</td>
<td>jōŋk</td>
<td>*jāņi ‘ice’</td>
</tr>
<tr>
<td>*nuoli</td>
<td>njuolla</td>
<td>nal</td>
<td>ńal</td>
<td>ńēl</td>
<td>*ńīlī ‘arrow’</td>
</tr>
<tr>
<td>*suoni</td>
<td>suotna</td>
<td>san</td>
<td>lan, jan</td>
<td>tēn</td>
<td>*sīīni ‘sinew’</td>
</tr>
<tr>
<td>*hapsi</td>
<td>vuokta</td>
<td>–</td>
<td>awōt</td>
<td>ēt</td>
<td>*īpši ‘hair’</td>
</tr>
<tr>
<td>*kanto</td>
<td>guottu</td>
<td>kanda</td>
<td>kant</td>
<td>kēnt</td>
<td>*kīnta ‘tree stump’</td>
</tr>
</tbody>
</table>

7. There are also two unique cases among the secondary e-stems: Fi järvi : *järve- ‘lake’ < PU *jäwrä and Fi sieni : siene- ‘mushroom’ < *säni < *śānā ‘shelf fungus’. These seem to lack the expected backing of the vowel (*ä > *a). The Finnic reflexes of *jäwrä, however, also include back-vocalic forms: Votic jarvi, Liv jōra < *jarvi. The oldest PFi form may have been *jarvi, in which case other languages underwent an irregular fronting *a > *ā due to the influence of initial *j-. The front vowel in sieni is more difficult to account for. It may be that the word-initial palatalized sibilant *ś- prevented the otherwise expected backing of the vowel. On the other hand, as *s and *ś merged in Pre-Finnic, there may have been avoidance of homonymy with Pre-Finnic *sāni ‘blood vessel’ (< PU *sāni). Such speculations are, of course, impossible to prove, but in any case one must assume the phonological development *śānā > *śāne- > *scene-, so sieni serves as yet another piece of evidence for the regularity of the change *ā > *ee.
A potential counterargument is, however, provided by data from Permic. As noted by E. Itkonen (1954: 322–327), Finnic short vowels and long vowels have in certain cases different correspondents in Permic languages. As the following table shows, Finnic stems of the type *ä–e- correspond to Permic stems with either the vowel correspondence Komi i ~ Udm i (in stems of the type CV-) or Komi e ~ Udm o (in other stem types), whereas Finnic stems of the type *ee–e- have Permic cognates showing the correspondence Komi -i- ~ Udm -i-:

<table>
<thead>
<tr>
<th>Fi</th>
<th>SaaN</th>
<th>Komi</th>
<th>Udm</th>
<th>PFP (Sammallahti 1988)</th>
</tr>
</thead>
<tbody>
<tr>
<td>jäsäń</td>
<td>jëz</td>
<td>jëz</td>
<td>jëz</td>
<td>*jäsi[n] ‘joint’</td>
</tr>
<tr>
<td>käske-</td>
<td>kësjë-</td>
<td>kösji-</td>
<td>kösji-</td>
<td>*käski- ‘order’</td>
</tr>
<tr>
<td>käsi</td>
<td>ki</td>
<td>ki</td>
<td>kës</td>
<td>*käti ‘hand’</td>
</tr>
<tr>
<td>väki</td>
<td>viehka-</td>
<td>vi</td>
<td>vi</td>
<td>*wëki ‘strength’</td>
</tr>
</tbody>
</table>

| kie  | giella | kël | kël | *këeli ‘tongue’       |
| mieli | miella | mil | mil | *meeli ‘mind’         |
| niele-| njiella-| nëlal-| nëlili-| *néeli- ‘swallow’    |
| liemi | liepma-| lêm | lêm | *leemi ‘broth, soup’*8 |
| –     | nierra | nir | nir | *neeri ‘cheek / nose’ |

At first sight this data would appear to support Itkonen’s solution to reconstruct distinct long vowels (in this case, *ä vs. *ee) in Proto-Finno-Permic. Nevertheless, one must note that there is no real obstacle to reconstructing PU *ä in all these words, as the different Permic reflexes can be straightforwardly explained by a conditioning factor: the development PU *ä(–i) > Komi, Udm i occurred whenever the vowel was followed by a voiced consonant (*l, *r or *m). What is more, the conditions of this development are not identical to those of the Finnic change *ä > *ee: Komi and Udm i occurs as the reflex of *ä also before consonant clusters where the first member was voiced (*l or *r), as shown by the following examples:

<table>
<thead>
<tr>
<th>Fi</th>
<th>SaaN</th>
<th>Komi</th>
<th>Udm</th>
<th>PFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>mälvä?</td>
<td>mielga-</td>
<td>mil</td>
<td>*mälki ‘breast’</td>
<td></td>
</tr>
<tr>
<td>(palo-)kärki</td>
<td>kir</td>
<td>kir</td>
<td>*käri ‘woodpecker’</td>
<td></td>
</tr>
</tbody>
</table>

There is a similar, superficial difference in the Permic correspondents of Fi a (< PU *ī) and Fi uo (< *oo < *a < *ī). On these grounds, Sammallahti (1988) reconstructs an opposition between short *ī and long *īī into Proto-Finno-Ugric. The former vowel would be reflected as Komi and Udm u, the latter as Komi and Udm e.

---

*Komi l‘em ‘jelly’ has often been included in this cognate set (UEW: 245; SSA s.v. liemi), but it actually seems to be the same word as Komi l‘em ‘glue’ (< PU *d‘imä, cf. Fi tymä ‘glue’).*
On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic *x

241

Fi SaaN Komi Udm PFU (Sammallahti 1988)

maa – mu mu *miixi ‘earth’
saa- – su- su-ti- *siixi- ‘come, arrive’
asu- – uz- izi-, izi- [*izi-w ‘camp’]
– – sus- susi- *siixi ‘Siberian pine’

nuoli njuolla nel nel *niiili ‘arrow’
suomi cuopma sem sem *siimi ‘scale’
suoni suotna sen sen *siini ‘sinew’
tuomi duopma lem lem *diiimi ‘bird-cherry’

Nevertheless, a similar conditioning can be seen here: u occurs in Permic CV-stems as well as before PU unvoiced consonants, whereas e occurs before PU voiced consonants (*l, *m or *n). Hence, one can postulate proto-forms with *i for all these words. In conclusion, the Permic correspondents of Finnic long vowels do not provide evidence for reconstructing a series of distinct long vowels into Proto-Finno-Permian.

3. The origin of Finnic *ii and *uu

Next, we have to deal with the Finnic long vowels *ii and *uu. Lehtinen derives Finnic *ii and *uu from earlier sequences of a short vowel and a glide. The vocalization of the glide would be a shared Finnic-Saami innovation, as also the Saami reflexes of the words with *ii and *uu lack the glide:

*pijre > PFi *piire- ‘circle’ (cf. SaaN birra ‘around’)
*kuwle- > PFi *kuule- ‘hear’ (cf. SaaN gulla- ‘hear’)

Finnic words containing these vowels only very rarely have cognates elsewhere in the family. Leaving aside monosyllabic stems with secondary long vowels (e.g. Fi pii ‘tooth’ < PU *piŋi), the vowel *ii is found in three words with cognates further than Saami:

Fi piiri ‘circle’ ~ SaaN birra ‘around’ (postp., adv.), MdE piře, M perā ‘fence, enclosure, garden’. — Note that the correspondence MdE i ~ M e is irregular.
Fi niini ~ MariW nī, Komi, Udm nìn ‘bast’
Fi visi (; viite-) ~ SaaN vihtta, MdM vetā, MariW wēc, Komi vit, Udm vīt’, KhE wet, MsE āt, Hung āt ‘five’, NenT ju?, EnF biu?, Ngan bīi?, SlkTa kōt, Kam bōʔ, Mat čuʔ ‘ten’

9. The illabial vowel in Udm results from a secondary change *u > *i; see Lytkin (1964: 21–23) for discussion.
On the basis of such limited material it is difficult to say much about the Uralic background of Fi ii, but at any rate it is clear that such limited examples can hardly support the reconstruction of a special long vowel phoneme *ii back to the Finno-Permic or Finno-Ugric level, as multiple other explanations are possible. The first two words could be simply reconstructed as *pijri and *ńijni instead. Furthermore, as the word for ‘bast’ does not have Saami or Mordvin cognates, one cannot rule out some yet more complex trisyllabic structure such as *ńijini, *ńınji, *ńixini; an intervocalic *j, *ŋ or *x would in any case have been regularly lost in such an environment in all the cognates. This leaves us with the numeral for ‘five’ (‘ten’ in Samoyed), which shows well-known difficulties of reconstruction; Janhunen (1981: 261) postulates the alternative proto-forms *witi / *witti / *wiixti. In the absence of any parallels for the sound correspondences in this lexical item, the postulation of a unique cluster *xt, and the assumption of a further development of a long vowel (*wixti > *wiiti), cannot be supported by the principle of regular sound change.

The vowel *uu is somewhat more frequent. Excluding monosyllabic stems (such as *luu ‘bone’ < PU *luwi), eight examples can be found:

Fi huali ‘lip’ ~ ? SaaN sulla ‘approximation, resemblance’, KhE lul, jul ‘mouth’
Fi kuule- ~ SaaN gulla-, MdM kuľə-, MariW kola-, Komi kįl-, Udm kili-, KhE kől-, MsE kől- ‘hear’
Fi kuusi ~ SaaN guhtta, MdM kotə, MariW kut, Komi kvaŋt ~ kvat’, Udm kwat’, KhE kut, MsE kőt, Hung hat ‘six’
Fi kuusi ‘spruce’ ~ SaaN guossa, MdM kuž, MariW kož, Komi koz, kež, Udm kįz, KhE kol, MsE kowt ‘spruce’, PSam *kā(ə)t
Fi tuuli ‘wind’, tuule- ‘blow (of wind)’ ~ MariW tul ‘storm’, Komi Udm tel ‘wind’, MsE tol ‘cloud’ (not previously included in this correspondence set; its semantic development may have been ‘wind’ > ‘storm’ > ‘storm cloud, rain cloud’ > ‘cloud’), ? PSam *tuə ‘feather, wing’
Fi uuddin ‘bed curtain’ ~ MariW amaš ‘shelter, hut’, Komi von, en, Udm in ‘bed curtain’, KhE olɔw ‘sleeping tent’, MsE āmɔl ‘cradle cover which protects the child from mosquitos’
Fi uuhi ‘sheep’ ~ MdM ucə ‘sheep’, MariW əţra ‘sheep-skin’, Komi, Udm ıţ, KhE ač, MsE ăş ‘sheep’
Fi uuusi ‘new’ ~ SaaN oδas, MdM od, MariW u, Komi, Udm  нагрузкens, Hung új ‘new’

Regardless of the higher frequency of Fi uu in inherited vocabulary, the picture is not at all clearer than in the case of ii. The correspondents of Fi uu are so heterogeneous that no single correspondence set matches any other, save for those of huali ‘lip’ and kuusi ‘six’, where the superficial match may simply be due to the scarcity of cognates of the former word. Compare the correspondence sets in the table below:
On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic *x

Fi | SaaN | MdM | MariW | Komi | Udm | KhE | MsE | Hung | PSam
---|------|-----|-------|------|-----|-----|-----|------|------
*huul*ī ‘lip’  | ? u – – – – u – – –
*kuule- ‘hear’  | u u o i i o õ – –
*kuusi ‘six’  | u o u va wa u õ a –
*kuusi ‘spruce’  | uo u o o, ê i o ow – *å(ə)
*tuuli ‘wind’  | – – u ê ê – o – *ua?
*uusi ‘new’  | o o u i i, i – – ü –
*uu*udin ‘bed curtain’  | – – a vo, ê i o å – –
*uuhi ‘sheep’  | – u õ i i a õ – –

These chaotic correspondences offer few obvious conclusions regarding the background of Fi *uu*. One thing is nonetheless clear: due to the fundamental irregularity of the correspondences there is no reason to assume that Fi *uu* would have one single Proto-Uralic source. For example, it is *ad hoc* to reconstruct the stems *kuuli- ‘hear’ and *tuuli ‘wind’, as the vowels in the cognates of these two items match in none of the languages where both are attested. This being the case, there is little evidence for the idea that the latter word ultimately reflects PU *tuxli and is thus cognate with PSam *tua ‘wing, feather’.

Lehtinen (1967) suggested that Fi *uu* could reflect a sequence *-uw-, but due to its multiple correspondents elsewhere in Uralic the reality is obviously more complex. It is possible, for instance, that combinations of various different vowels and the glide *w lie behind it. For example, the Mari and Permic vowels in the cognates of *tuuli* match those in the cognates of Fi *oksenta- ‘vomit’ (cf. MariW *uksënza-*, Komi *es- (~ vos-)*, KomiJ *ursot-*, Udm *es- ‘vomit’ < *oksi-*), so perhaps one should reconstruct *towli and assume a change *ow > *uw > *uu in Finnic. As regards *kuule- ‘hear’, it is not impossible that this verb is ultimately in a derivational relationship with PSam *kaw ‘ear’, even though at the moment this remains a matter of speculation. In the case of *kuusi ‘spruce’ the glide *w is even synchronically found in Mansi (E *kowt, W *kowt, kawt, N *kawt < PMs *kawt), but the original shape of the word is difficult to reconstruct. In Fi *uusi ‘new’ the initial long vowel might reflect PU *wu- (*wud'i ‘new’), as suggested by Itkonen (1969: 102).

The exact background of Fi *ii* and *uu* must be left for future research to solve. However, it seems unnecessary to invoke PU *x to explain the emergence of these long vowels in Finnic, as they can represent a development of various combinations of vowels and the glides *w and *j, perhaps also other sounds.

4. The origin of Samoyed vowel sequences

If Finnic long vowels turn out not to correspond to Samoyed vowel sequences in a consistent manner, the origin of the latter must be treated separately. The Uralic etymologies for Samoyed words with vowel sequences can apparently be divided into at least three distinct types of cases. In the first type of case the vowel sequence derives from a PU sequence *-VwV-:
Luobbal Sámmol Sámmol Ánte (Ante Aikio)

PSam *puə- ‘blow’ (> NenT püʔ-, En fuasa-, Ngan hūolə-, Slk *pū-, Kam puʔ-). — Contrary to Janhunen (1981) and Sammallahti (1988: 547), the Samoyed verb can hardly be treated as unrelated to PU *puwa- (> MdE puva-, MariW pue-, KhE pōɣ-, MsE pow-, Hung fūj ‘blow’). The verb is no doubt onomatopoietic in origin, but as the sound correspondences appear to be quite regular, it can be reconstructed to Proto-Uralic already.

PSam *pu(o)I- ‘swell’ (> Slk *pūl-: K pūliŋ 3sg ‘swells’). — The Slk word has not been previously etymologized, but it is no doubt cognate with MariW puala-, KhE pōɣol- ‘swell’ (< *puwVlV-). The PSam vowel sequence cannot, of course, be verified on the basis of the Slk cognate only, but the reconstruction *pu(o)- is supported by the fact that the verb appears to be a derivative of PU *puwa- ‘blow’ (> PSam *puə-); as for the semantic connection, cf. e.g. SaaK pōssa- ‘blow; swell’, German blasen ‘blow’ ~ Blase ‘blister, bubble’, Old Norse blása ‘blow’ → blástr ‘blister, swelling’, Lithuanian pušti ‘blow’ → pušlė ‘blister, bubble, bladder’. The original trisyllabic structure of the verb can clearly be seen in the Mari cognate puala-; note the homonymous puala- ‘blow (once)’, a derivative of Mari pue- ‘blow’ < PU *puwa-. The original trisyllabic structure also explains the preservation of the stem-final lateral *-l- in Slk; a parallel is provided by SlkTa ašil- ‘step over’ (< PSam *asəl- < PU *askili-). Otherwise one would expect a change *l > PSam *j in stem-final position.

PSam *tiə ‘year ring’ (> NenT tį, Ngan (Castrén) tį, Slk *tů). — The vowel correspondence between Nenet and Selkup indicates that the PSam vowel sequence *iə must be reconstructed for this word. As is well-known, the Samoyed word is cognate with Fi syy, MariW sōj and Udm si ‘year ring’, to which one can probably add SaaS sīve ‘line, streak’ and MsE tāw ‘fiber, streak, layer’ as new cognates. Janhunen (1981: 245) and Sammallahti (1988: 540) reconstruct PU *sūxi, but this form explains neither the Mari and Permic illabial vowel nor the vowel sequence in Samoyed. It seems more appropriate to postulate PU *siwV or the like; the labial vowel in Fi syy would have been caused by a development *iw > *ūū. Note, however, that SaaS sīve is deviant in regard to vocalism: it would presuppose a proto-form *sāwi.

PSam *ńiə ~ jia ‘belt’ (> NenT nį, Slk *cū (Ta čū, čū), Kam ji, Mat Nį), *ńiə-jə (> EnT niojo, Ngan C niéja). — This word derives from PU *ūwā ‘belt’; see the discussion in section 1.

Ngan kuogunu ‘long ago’ < PSam *kuwkā-nā. — This word appears to have no cognates elsewhere in Samoyed, and no etymology has been proposed for it. However, *kuwakā- can be derived from PU *kuwa-kka ‘long’ (> SaaN guhκki, MdM kowaka ‘long’). The underived root is found in KhE koɣ ‘long’ (< *kuwa), and MdM kavor ‘a long time’ reflects a parallel derivative. On the etymology, see Aikio (2000).

Ngan buo ‘current’ < PSam *wuə. — The word has no cognates elsewhere in Samoyed, but it is apparently a previously unnoticed reflex of PU *(w)uwa- ‘current; flow’ (> SaaN avvi- ‘leak (of boats)’, Fi vuo, KhE oɣ ‘current’, oɣa- ‘flow’, MsW ow ‘current’, ow- ‘flow’).
In the second type of case the vowel sequence corresponds to a stem of the shape *CVli- in Finno-Ugric or at least Finnic; all of these etymologies are discussed in more detail in section 1 above:

**PSam** *kää- ‘die’ (> NenT χa-, EnF kā-, Ngan kuo-, SlkTa qu-, Kam ku-, Mat kā-) ~ Fi kuole- (< *kali-)

**PSam** *ńåa- ‘lick’ (> SlkTa nü-, Kam nē-) ~ Fi nuole- (< *ńali-)

As mentioned in section 1, these two roots pose an unresolved reconstructional problem: some Finno-Ugric cognates suggest that Fi kuole- and nuole-might belong to the group of ‘secondary e-stems’, and their proto-forms should be reconstructed as *kola- and *ńola-. Samoyed, however, does not support such a conclusion: the predictable reflexes of PU *kola- and *ńola- would be PSam *kålā- and *ńålā-

**PSam** ? *piə- ‘outside’ ~ Fi pieli ‘outer edge; post’ (?) < *pāli).

**PSam** *kää(j) ‘tongue’ ~ SaaN giella, Fi kieli, MdM käl’, Komi, Udm kil, KhE kōl (< *kāli).

**PSam** *tuə ‘feather, wing’ ~ Finno-Ugric (derivative) *tul-ka (> SaaN dolgi, MdM tolga, Komi til-bord, Udm tili, KhE tōgel, MsE towl, Hung toll).

Third, on the basis of the following two examples it appears that a vowel sequence may correspond to a Finno-Ugric stem type *CVji:

**PSam** *kää > Slk *qū ‘slender object (?)’ (Ta qu: nüti qu ‘stalk of grass’, opti qu ‘a single hair (on the head), pōl qu ‘tree trunk’, tūl pot qu ‘conifer needle’, etc.). — This word apparently belongs to a widespread but previously unnoticed Uralic cognate set; its cognates include Saal kuojā ‘sedge’, Komi, Udm kī ‘awn’, MsE kōj ‘hair, tuft, mane’, Hung haj ‘hair (on the head)’ (< PU *ko/aji). Despite the lack of cognates in other Samoyed languages the vowel sequence *åə is securely reconstructed. An identical development is attested in Slk *qū- ‘die’ (< PSam *kāо-), and there appear to be no alternative proto-forms that could account for Slk *qū. Because the change PSam *k- > Slk *q- took place before PSam *a, *ə, *å and *o, in this word Slk *qū must go back to a non-high back vowel; PSam *ku- is reflected as Slk *ku-. The proto-form cannot have been PSam *kā or *ko, as both would predictably have yielded Slk *qō (cf. PSam *sā- ‘sharpen’ > Slk sō-, PSam *ko- ‘find, see’ > Slk *qō-). The stem cannot have contained the glide *j in PSam, as the reflex of PSam sequences *åj, *oəj, *uj and *aj is Slk *ū.

**PSam** ? *tiə- ‘pus, rot’ < PU ? *sājī. — See the discussion in section 1.

Of the three types of cases discussed above, the first type is the easiest to account for. One can simply assume that a vowel sequence developed through a regular loss of intervocalic *w; there do not seem to be any plausible examples of retention of intervocalic *w in Samoyed (Aikio 2002: 35–36 contra Janhunen 1981: 253).

The second type of case, a PSam vowel sequence corresponding to Finno-Ugric *(C)Vli-, obviously requires another explanation. Janhunen’s hypothesis of PU preconsonontal *x was based on the following patterning of sound correspondences:
But as Finnic long vowels can be explained as a regular product of secondary lengthening, the second mora of the long vowel cannot be matched with the schwa component of the vowel sequence. Instead, the correspondence can be interpreted in two alternative ways. The schwa may be either of a reflex of PU *l or the remnant of an original stem vowel, implying that *l was either vocalized (*l > *ə) or lost (*l > Ø) in some contexts in Pre-Proto-Samoyed:

<table>
<thead>
<tr>
<th>alternative 1:</th>
<th>alternative 2:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proto-Finnic</td>
<td>k oo l e</td>
</tr>
<tr>
<td>Proto-Samoyed</td>
<td>k å ø Ø</td>
</tr>
</tbody>
</table>

At this point it is interesting to consider the third type of case where a vowel sequence appears to correspond to a PU stem of the type *CVji. As convincingly argued by Janhunen (1981: 250), PU *l and *j seem to have merged in PSam in nearly all contexts due to a change *l > *j. Thus, the correspondences PSam *

<table>
<thead>
<tr>
<th>'come'</th>
<th>'die'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Proto-Samoyed</td>
<td>*tul- : *tuli-</td>
</tr>
<tr>
<td>*tuj- : *tujo-</td>
<td>*kåj- : *kåjo-</td>
</tr>
<tr>
<td>*tuj- : *tujo-</td>
<td>*kåj- : *kåjo-</td>
</tr>
<tr>
<td>*tuj- : X</td>
<td>X : *kåo-</td>
</tr>
<tr>
<td>Proto-Samoyed</td>
<td>*tuj- : *tuj-</td>
</tr>
</tbody>
</table>

Such an explanation is of course hypothetical, but not far-fetched. As a partial parallel, alterations similar to those hypothesized for Pre-Proto-Samoyed later
On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic *x
devolved in Nganasan through loss of intervocalic (but not syllable-final) *j:
e.g. Ngan ŋoj ‘foot’ : PL.NOM ŋuo? (< *āj : *āo-t < PSam *āj : *āo-t). The dif-
ference is that in Nganasan the process is recent enough for the alterations to
remain a regular part of the language’s morphophonological system.

Other speculations, however, could also be pursued here. As known ex-
amples of the development *CVli- > *CVj- seem to be limited to stems with
non-high vowels, one could surmise that the development *CVli- > *CVj- >
*CVə- took place if the first-syllable vowel was low (as in *kāv- ‘die’, *kāa(j)
‘tongue’). However, this explanation would appear to be weaker than the sug-
gestion of analogical generalization, as it does not account for PSam *tuə ‘wing,
feather’ (? < *tuli-, cf. Finno-Ugric *tul-ka ‘feather’).

Finally, it must be noted that Samoyed vowel sequences occur in a couple of
word roots which show some kind of unique Uralic correspondence. While some
examples might simply involve wrong etymologies, not all such cases seem to be
easy to dismiss. Cases in point are PSam *puə-, *puə(-n)-IV ‘knee’ (~ Fi polvi)
and PSam ? *koaj ‘birch’ (~ Fi koivu), which were discussed in section 1; the
phonological problems in these etymologies remain to be accounted for. The
following case is also interesting:

PSam *puə- ‘behind’ (> NenT pū, EnF fuo-, Ngan huo, SlkTa pu) ~ Fi
poo ‘anus’, KhE pūj, MsE pəj ‘ass’. — This etymology is rejected by Janhunen
(1981) and Sammallahti (1988: 547), even though the match seems both phono-
logically and semantically highly plausible. The exact PU reconstruction of the
word remains obscure, however. Ob-Ugric forms suggest intervocalic *j (PU
*puji?), but *j is not known to have disappeared in Finnic after a back vowel.

5. Conclusion

The main results of the study above can be summarized as follows:
a) There is no consistent correspondence between Proto-Samoyed vowel se-
quences and Finnic long vowels. The only environment where a correspondence
can be established is before PU *l, and even here it seems to be coincidental.
b) The Finnic “primary” long vowels *ee and *oo — which have been recon-
structed as long vowels as far back as Proto-Finno-Permic (Itkonen 1954) or
Proto-Finno-Ugric (Sammallahti 1988) and which Janhunen (1981) derives from
a Proto-Uralic sequence *Vx — can be explained as secondary Finnic innova-
tions. They arose through the changes *ä > *ee and *a > *oo, which regularly
took place before single voiced non-dorsal consonants in Pre-Finnic *e-stems.
c) The background of the Finnic “primary” long vowels *ii and *uu is less
clear. Both of these show multiple correspondents outside Finnic, suggesting
they have no single source. Probably *ii and *uu have developed from various
combinations of vowels and the glides *j and *w, perhaps also other sound se-
quences; the issue requires further study.
d) The vowel sequences in Samoyed have several sources: they occur in the re-
flexes of the PU stem types *CVwV-.* CVli- and *CVji-. Apparently, they arose
through the loss of a Pre-Proto-Samoyed intervocalic glide *w or *j (after the shift *l > *j); the details of the process require further study. In individual lexical items vowel sequences may have yet some other, more complex background, as suggested by correspondences such as PSam *pua(-n)-IV ~ Fi polvi ‘knee’.

e) Since there are satisfactory alternative accounts for the development of both Finnic “primary” long vowels and Samoyed vowel sequences, and because there is no regular correspondence between the two in the first place, there is no need to reconstruct the consonant *x in preconsonantal position into Proto-Uralic (contra Janhunen 1981).

Even though we have reached a negative conclusion in our review of Janhunen’s hypothesis regarding *x, this should not be seen as diminishing the value of his contribution to the reconstruction of Proto-Uralic in any way. It goes without saying that we would not even be able to address such details of Uralic reconstruction, were it not that Janhunen set a whole new methodological standard for this line of research with his groundbreaking paper. That new details offer themselves to scrutiny once a breakthrough has been made is only an expected turn in the progress of science.

Moreover, the results presented in this paper do not so greatly contradict Janhunen 1981 as might seem at first, as he initially formulated his hypotheses regarding *x in a very cautious manner: “it must be noted that all the presented assumptions regarding the independent phonemic status and the distribution of *x are highly tentative” (Janhunen 1981: 28; translated from Finnish). Only later, it seems, he became fully convinced by his own hypothesis and described PU *x as a ‘laryngeal’ segment (Janhunen 2007). No new evidence in favor of *x seems to have surfaced after Janhunen’s initial study, however, so here one could with some justification criticize him for momentarily neglecting his own, remarkably high methodological standards.

Finally, it must be noted that the arguments presented here do not affect the reconstruction of PU *x as a distinct phoneme in intervocalic position. As argued by Janhunen (1981), a PU stem of the shape *CVxi- must be reconstructed for cognate sets which show the correspondence Saami *CVkę- ~ Finnic *CVV- ~ Samoyed *CV-, such as SaaSk miökka- ~ Fi myy- ‘sell’ ~ PSam *mi- ‘give’ (< PU *mexi-). In such cases *x can be distinguished from *k on the basis of Finnic; the latter is attested in Saal kuohä- ‘check (nets, traps)’ ~ Fi koke- ‘experience; check (nets, traps)’ ~ PSam *ko- ‘see, find’ (< PU *koki-). In Saami one can often indirectly distinguish *x from *k on the basis of the quality of the first syllable vowel: cf. SaaSk suukka- ‘row’ (< PU *suxi-) vs. lookka- ‘read, count’ (< PU *luki-), miökka- ‘sell’ (< PU *mexi-) vs. töökkä- ‘fuck’ (< PU *teki- ‘put, do’). Janhunen (2007: 217–218) and Koivulehto (1991: 17–19) are probably right in their suggestion that intervocalic *x derives from Pre-Proto-Uralic *k through phonemic split, but nevertheless the two sounds were distinct in PU already.

As *x can be established as an independent member of the Proto-Uralic phonemic inventory, one cannot completely exclude the possibility that it occurred in some consonant clusters after all. Who knows if clusters with *x are hiding behind the strange sound correspondences exhibited by the cognates of
Fi koivu ‘birch’, polvi ‘knee’ or viisi ‘five’? Thanks to Janhunen (1981) we have a clear basic picture of the Proto-Uralic sound system, and Uralic historical phonology can move on to deal with such mysteries.

Abbreviations

| EnF  | Forest Enets | Ngan | Nganasan  |
| EnT  | Tundra Enets | PFi  | Proto-Finnic |
| Est  | Estonian     | PFP  | Proto-Finno-Permic |
| Fi   | Finnish      | PFU  | Proto-Finno-Ugric |
| Hung | Hungarian    | PSaa | Proto-Saami   |
| Kam  | Kamas        | PSam | Proto-Samoyed |
| KhE  | East Khanty  | PU   | Proto-Uralic |
| Liv  | Livonian     | SaaI | Inari Saami   |
| MariE| East Mari    | SaaK | Kildin Saami |
| MariW| West Mari    | SaaL | Lule Saami |
| Mat  | Mator        | SaaN | North Saami |
| MdE  | Erzya Mordvin| SaaSk| Skolt Saami |
| MdM  | Moksha Mordvin| SaaU| Ume Saami |
| MsE  | East Mansi   | Slk  | Selkup |
| MsN  | North Mansi  | SlkK | Ket Selkup |
| MsS  | South Mansi  | SlkTa| Taz Selkup |
| NenF | Forest Nenets| SlkTy| Tym Selkup |
| NenT | Tundra Nenets| Udm | Udmurt |

References
