LINGUISTICA URALICA LXI 2025 3 : 145—184 https://doi.org/10.3176/lu.2025.3.01

JUHA JANHUNEN (St. Petersburg)

SAMOYEDIC VOWEL SEQUENCES AND THE STATUS OF THE CLUSTER **lk* IN URALIC

Abstract. This paper takes up the issue concerning the representation of the Finno-Ugric cluster **lk* in Samoyedic and its possible connection with Proto-Samoyedic vowel sequences. It has been proposed that the cluster **lk* was lost in Pre-Proto-Samoyedic, which, at least in words containing a low vowel in the second syllable, would have led to the origination of a non-canonical type of vowel sequences containing a low vowel as the second component. This proposal presents two challenges: on the one hand, the diachronic evidence for the alleged intervocalic loss of the cluster **lk* needs to be critically examined, while, on the other hand, the presence of the alleged type of non-canonical vowel sequences in Proto-Samoyedic has to be either confirmed or rejected. The conclusion of this paper is that the apparent cases suggesting the loss of the cluster **lk* in Samoyedic have to be explained in a different way, in that the cluster in the items concerned is actually of a secondary derivational origin, that is, **l-k*, in which the derivational segment **k*, as observed in the Finno-Ugric data, is simply originally absent in Samoyedic. The issue concerning the proposed non-canonical vowel sequences is more complicated, but, at least at the level of Pre-Proto-Samoyedic, the second component of all vowel sequences can always be identified with the uniform reduced vowel **ə*. This situation underwent later secondary changes in the individual Samoyedic languages, especially in Nganasan and Enets.[1]

Keywords: Proto-Uralic, Proto-Samoyedic, consonant clusters, vowel sequences, laryngeals, glides.

1. Introduction

It is today generally accepted that vowels occurred in Proto-Samoyedic not only as single segments, but also as sequences of two segmentally separate elements. Although the earliest modern attempts at a Proto-Samoyedic reconstruction (Sammallahti 1975; Janhunen 1976) still operated with only monophthongoid occurrences of vowels, vowel sequences have been present in all

[1] An oral version of this paper was presented (online, in Russian) at the Fourth Conference on Uralic, Altaic, and Palaeo-Siberian languages (Четвертая конференция по уральским, алтайским и палеоазиатским языкам), held at the Institute for Linguistic Studies, Russian Academy of Sciences, St. Petersburg, on 25—27 November, 2024.

Received 3 April 2025, accepted 22 July 2025, available online 10 September 2025.

later reconstructions (starting, apparently, with SW). According to the original formulation, these vowel sequences were assumed to have contained a qualitatively indistinctive but phonetically velar "reduced" vowel (**ə̑*) as the second segment. This vowel was paradigmatically identified with the presumably phonetically very similar reduced vowel (**ə̈*), attested as an independent vowel phoneme in both the initial and any non-initial syllables. It was also assumed that the position of the initial component in vowel sequences could be occupied by any vowel quality of the main paradigm, with, however, the exception of the reduced vowel itself. The absence of sequences of two reduced vowels is probably connected with the origin of this segment both as the main vowel of a syllable and as the second component of vowel sequences.

Phonotactically, as far as independent lexical items are concerned, vowel sequences can be reconstructed only for the initial syllable, meaning that the first component in them was always the first vowel of the word. However, the possibility that they could occur also later in the word at morpheme boundaries is suggested, in particular, by the finite aorist paradigm of vowel-stem verbs, which, as it seems, were formed in Proto-Samoyedic, with the exception of the Nganasan lineage, with the help of a "finite morpheme" consisting of a single vowel segment identical with the reduced vowel. It has been proposed that this finite morpheme could represent a reduction of the sequence **-ŋå-*, as attested in the role of a finite morpheme after consonant stems (Salminen 2024 : 180), but both the mechanism and the possible chronology of this (in itself irregular) development remain unconfirmed. This question aside, it is a matter of terminological interest only whether the vowel sequences should be considered to have been "diphthongs" or "diphthongoids", or, as the term itself suggests, sequences of two separate segments, in which case the second component would also have possessed a syllable-bearing value in its own right.

Later research (Хелимский 1993) has pointed out that there were actually two distinct reduced vowel phonemes in the initial syllable in Pre-Proto-Samoyedic, often, but not always, deriving from Proto-Uralic **u* vs. **i* (> **ŭ* vs. **ĭ*, cf. also Pystynen 2022). These vowels have been written variously as <**ə̑*> vs. <**ə*> (as in Salminen 2012 : 346—347; 2024 : 171, 179 *et passim*) or <**ə̑*> vs. <**ə̈*> (as in Kaheinen 2023 : 34 note 10), but they will in the following be rendered as <**ə*> vs. <**ə̈*>. Functioning as a harmonic pair, they correlated with the choice between the mid-low vowels **å* vs. **ä* in any following syllables, as in **məkå* 'back, spine' (< **muka*) vs. **sə̈ymä* 'eye' (< **çilmä*). The distinction between the two reduced vowels has been lost at the phonemic level in all modern Samoyedic languages (cf. Helimski 1997 : 105 and Salminen 2024 : 183, where an unlikely exception from Mator is mentioned), but it has left occasional traces in the morphophonology, as in Tundra Nenets *tə'* 'summer' (< **təŋ* < Proto-Uralic **suŋ/ə*) : (accusative plural) *təŋ-o* vs. *tər* 'hair' (< **tə̈r* of unknown origin) : (accusative plural) *təry-e*.

The issue concerning the two reduced vowels has relevance to vowel sequences, in that, assuming that vowel harmony functioned consistently in the language, the reduced vowel occuring in the role of the second component of vowel sequences could also be assumed to have represented either the velar segment **ə* (after a back vowel) or the palatal segment **ə̈* (after a front vowel). However, while this possibility cannot be ruled out at the phonetic level, its phonological relevance can be questioned, which

is why it is reasonable to leave the distinction unmarked in vowel sequences (cf. Kaheinen 2024 : 94 notes 1—2). The same principle can, in fact, be applied to all instances of the reduced vowel, for, although the two reduced vowels were originally separate phonemes and later functioned as two separate morphophonemes, they appear to have lost their phonemic and probably also phonetic distinction by the Proto-Samoyedic stage. The harmonically neutralized reduced vowel could therefore be combined with both back and front vowels within a single word, as in **kətå* 'nail' resp. **pətä* (< **pətä*) 'bile', and in Proto-Samoyedic reconstructions it is sufficient to use the single symbol <**ə*> for this vowel (so also Pystynen 2022).

The harmonically neutral status of the reduced vowel in the role of the second component of vowel sequences is also supported by its phonotactic similarity to the glides **w* and **y* [j],[2] suggesting that it itself may originally have been a non-syllabic semivocalic segment of the glide type, perhaps of the velar or laryngeal range (**x*, Janhunen 1998 : 464). The similarity is particularly obvious from the fact that both vowel sequences (*V*ə* < *V*x*) and sequences of a vowel and a glide (*V*w* and *V*y*) in Proto-Samoyedic could be followed by a syllable-final consonant, or also, word-internally, by a consonant cluster, which was phonotactically not possible for other ("true") consonants. It may also be noted that in Nganasan, where the vowel sequences are synchronically represented as sequences of two full (syllabic) vowels, also described in terms of "morae" (Helimski 1984 : 42—43), the labial glide **w* in postvocalic syllable-final position likewise yields a full vowel, as in *kou* 'ear' (< **kåw*) = 'sun' (< **kåyå*). However, for unknown reasons, this behaviour does not extend to the palatal glide **y*, which in most cases (though not in all) is represented as a non-syllabic consonantal segment, as in *ŋoj* 'foot' (< **åy*) : (3rd person singular possessive) *ŋo-cu* (< **åy-tå*) : (nominative plural) *ŋue-'* (< **åy.ə-t*) (SNg 133).

A major change in the reconstructional situation occurred when it was proposed (first, apparently, by Helimski) that there were also other vowel qualities that could form sequences in Proto-Samoyedic. In particular, sequences with either **å* or **a* — though, curiously, not with **ä* — as the second component have been postulated for Proto-Samoyedic. The evidence for these "non-canonical" vowel sequences will have to be examined in more detail (below), but the issue also raises several questions of a more general type. For one thing, irrespective of whether they can be verified

[2] In the present paper, reconstructed forms are quoted in a simplified phonemic transcription. In deviation from the traditional notation, the letters **d*, **j*, **x* stand for the weak obstruents (dental vs. palatal vs. velar/laryngeal, traditionally **δ*, **δ′*, **γ*), the letters **w*, **y* for the glides (traditionally **β*, **j*), the letters **ç*, **ñ* for the palatal sibilant (affricate) and nasal, respectively (traditionally **ś*, **ń*), and the letter **c* for an unspecified affricate with a sibilant release (traditionally **č*). Synchronic data from modern languages are quoted according to the sources in a simplified (quasi-)phonemic notation. For Samoyedic, distinctive palatal consonants are written as *c* (strong or unmarked stop or affricate), *j* (weak stop or affricate), *ç* (sibilant continuant), *ñ* (nasal), and *ĺ* (lateral), while *q* denotes a back velar (uvular) stop (as in Selkup) and *'* (apostrophe) the glottal stop. Nenets and Nganasan are quoted following the phonemic analysis of Salminen (1997) and Kaheinen (forthcoming), but with the glottal stop marked invariably as *'* (also used for the Finnic final "aspiration"). Languages with a Latinized written norm are quoted according to the standard orthography. Unverifiable "readings" from historical or otherwise phonologically inexact sources (as for Mator) are marked by the dagger symbol †.

or not, the "non-canonical" sequences are statistically conspicuously rare, even marginal, in the etymological corpus, as compared with the cases in which the second component may unambiguously be identified as the qualitatively and harmonically neutral reduced vowel, that is, **ə*. Therefore, the "non-canonical" sequences are potentially open to alternative explanations, and, in any case, their diachronic background is likely to be different from that of their "canonical" counterparts.

The issue is closely connected with the reconstruction of vowels in non-initial syllables. It was originally assumed (as in SW) that there were only three vowel qualities in Proto-Samoyedic that could occur in this position: the two harmonically opposed low vowels **å* vs. **ä* and the harmonically neutral reduced vowel **ə* (**ə̑*). These three vowels were supposed to constitute relatively direct reflexes of the corresponding Proto-Uralic segments, which likewise comprised the two harmonically opposed low vowels **a* vs. **ä* and the harmonically neutral reduced or high vowel **ə* or **i* (traditionally **e*), which in consonant stems when following a single consonant could alternate with zero.[3] It is, however, generally accepted today that there were also other vowel qualities that could occur in non-initial syllables in Proto-Samoyedic. The two qualities that may be regarded as fully confirmed are **u* and **i*, as well as, possibly, their non-distinctive harmonic alternants (**ü* resp. **i̮*). Since these vowels are also attested in a few Uralic items, as in **yänti* 'sinew', they could, in principle, go back to Proto-Uralic. Assuming that they are nevertheless secondary, their derivation remains to be clarified (Salminen 2024 : 183).

Several other vowels in non-initial syllables are attested in the individual Samoyedic branches and languages. These vowels are typically secondary, deriving from various combinations of vowels and glides, but it is difficult to establish their relative chronology. For instance, data from Nenets-Enets suggest that the mid-high vowels **o* and **e* could derive from sequences ending in a syllable-final palatal glide **y*, as in Tundra Nenets *xən°* 'sledge' : (accusative plural) *xən-o* (< **kəncə-y*), but we do not know if this situation was valid already in Proto-Samoyedic. In any case, the development **əy* > **o* affected also vowel sequences, as in Tundra Nenets *ya* (< **yåə*) 'earth, land' : (accusative plural) *yo* (< **yao* < **yåə-y*) (cf. Salminen 2024 : 180). Another secondary segment was the low vowel **a* [*a*], attested in a few nominal roots, as in **wota* 'berry', but particularly often in Nganasan secondary stems, as in the genitive plural of nouns and the aorist stem of verbs (Kaheinen 2024). The origin and chronology of this **a*, which in early reconstructions (as in SW) was not distinguished from **ä*, is obscure. It is also not clear whether the **a* of non-initial syllables should automatically be regarded as paradigmatically identical with the **a* (= **ä* in SW) of the initial syllable, which in some cases has been assumed to derive from Proto-Uralic **a* (as also discussed *passim* below).

One of the Proto-Samoyedic items for which a vowel sequence with a low vowel as the second component has been proposed is the noun for

[3] It has also been suggested that the Proto-Uralic non-low vowel in non-initial syllables might have participated in the vowel harmony, in which case it would have had two harmonic variants, **i* [i] and **i̮* [ɨ] ~ [ɯ]. Since, however, there is no trace of this variation in any Uralic language, and since the non-low vowel can alternate with zero, it will here be written **ə* (in agreement with Kallio 2012) or, in cases of alternation with zero (after a single consonant), **/ə*.

'feather, wing', once reconstructed as **tuə̑j* (SW 166) = (in the notation used here:) **tuəy*. Mainly on the basis of the Tundra Enets shape *tua*, this item has later also been reconstructed as **tua* (Helimski, as quoted by Salminen 2024 : 180, apparently with the secondary vowel **a* as the second component) or **tuå* (Kaheinen 2023 : 170, with the more "regular" mid-low vowel **å*). A new perspective to this reconstruction was opened by the claim that the Samoyedic item is a regular reflex of Proto-Uralic **tulka* and a direct cognate of its Finno-Ugric reflexes (cf. Zhivlov 2023 : 164). The underlying idea is that the medial cluster **lk* was lost in Pre-Proto-Samoyedic, leaving only a hiatus between the surrounding vowels, which would not have undergone any substantial qualitative changes, i.e., **tulka* > **tuå*. If correct, this derivation would provide a diachronic motivation for the alleged non-canonical vowel sequence, though it gives no direct motivation for the alternative reconstruction **tua*.

The proposed derivation has several problems. The traceless disappearance of an entire consonant cluster would be exceptional, for, even though the velar stop **k* has been regularly lost in a number of clusters either before or after another consonant, as in **sk*, **ck* (= **čk*), **ks*, **kt* (Salminen 2024 : 178), the other component of the cluster is always segmentally preserved, which means that the development probably took place via either regressive or progressive assimilation, leading initially to geminates, which were later simplified. Moreover, **k* is actually preserved in the clusters **rk* (**sarka* 'branch, division' > **tårkå*) and **dk* > **rk* (**pidkä* 'high, long' > **pirkä*).[4] A medial intervocalic **k* is likewise intact before a second-syllable low vowel (as in **muka* 'back, spine' > **məkå*), as is also the medial lateral **l*, though there are signs of a tendency towards phonetic palatalization of this consonant, as suggested by the change in the quality of the following vowel (**kala* 'fish' > **kålä*). In syllable-final position, **l* has been fully palatalized to **y* [j], at least when following original high vowels (as in **çilmä* > **sə̑ymä* > **səymä* 'eye'). In these cases, Proto-Uralic **u* is preserved as **u* in Samoyedic (**tul/ə* 'fire' > **tul* > **tuy*), while the combination **u—a* would normally be expected to yield **ə—å*, rather than **u—å*.

Disregarding for the time being the potential counterevidence provided by Tundra Enets *tua* and its possible parallels in the etymological corpus, the rest of the facts are better explained by the assumption that the Samoyedic items for 'feather, wing' are not direct cognates of Finno-Ugric **tulka*, but, rather, reflexes of a shorter root corresponding only to the initial syllable of the word, which, at least formally, could correspond to Finnish (and Finnic) *tuuli* 'wind' : *tuule-* 'to blow (of wind)' (Janhunen 1981 : 23 no. 85), with cognates as far east as Permic, and tentatively reconstructable as Proto-Uralic **tuxl/ə*(-). A well-known semantic parallel is offered by Scandinavian, where the words (here quoted from Swedish) *vind* (= English *wind*) and *vinge* (→ English *wing*) are both derived from an Indo-European root originally meaning 'to blow (of wind)' (SEO 1348—1350). Irrespective of this semantic issue, the division of **tulka* as **tul-ka* implies that there was a derivational suffix **-kA*, forming secondary nominal

[4] Note that several etymologies suggesting a correspondence of Finnic **rk* to Samoyedic **r* are mistaken. In these cases it is a question of accidental lookalikes based on, among other things, onomatopoetic resemblance, as in Finnic *kurki* vs. Samoyedic **kərə-* 'crane' (SW 54).

stems from more simple primary roots. Therefore, it is relevant to take a look at other examples of this derivational suffix.

Basically, *-*kA* seems to have been an element deriving deverbal nouns. Transparent examples of this function are rare, but a case in point is Samoyedic **kəmə-* 'to fall' : **kəmə-kå* 'fallen (tree)' (SW 52). Another certain case, deriving from Proto-Uralic, is **pid/ə-* 'to be high, long' : **pid-kä* 'high, long' (erroneously classified as a denominal derivative in Luobbál Sámmol Sámmol Ánte (Aikio) 2022 : 19) > Samoyedic **pir-* : **pir*(-)*kä* (SW 125, as already quoted above). It may be concluded that the suffix *-*kA* could be attached to both vowel stems and consonant stems. Other examples of *-*kA* on consonant stems may include **pår*(-)*kå* 'garment' (SW 116—117),[5] which has either a cognate or an areally transmitted reflex in Finnish *parka* *'poor garment' > 'poor, deplorable person' (Janhunen 1988), as well as **poŋ-kå* 'net' (SW 127), which would seem to be derived from the verbal stem underlying Hungarian *fog* 'to catch' (MSzFE 209—210; Janhunen 2022 : 164) with possible further cognates in Finno-Ugric.[6] In the light of this information, it is reasonable to assume that **tul-ka* would also be derived from the verbal representation of the underlying root **tuxl/ə*(-) 'to blow (of wind)', which itself functions, at least in Finnic, as a *nomen-verbum*.

The Finno-Ugric cluster **lk* is, however, attested also before the reduced (or high) vowel **ə* (or **i*). Somewhat inconsistently, it has been speculated that in this position only the segment **k* would have been lost in Samoyedic, while **l* would yield, at least after a high vowel, and following the loss of the original stem-final vowel, the regular syllable-final reflex **y* [*j*]. This scheme is supported by the comparison of Finnish *kylki* : *kylke-* < Finnic **külki* = **külkə* 'side (of body)' with Samoyedic **kəy* < **kəy* 'side' (SW 57—58, where the item is mistakenly reconstructed as **kåj* = **kåy*). Indeed, the etymological connection (discussed extensively in Salminen 2023 : 377—380) is very probably correct (as elaborated in more detail below, section 3), but the fact that there is no trace of medial **k* in the Samoyedic cognates suggests that this segment was never present in the underlying form. The Samoyedic data would best be derived from a basic root of the type **kül/ə* (or **kil/ə*), from which Finnic **külkə* = **kül-kə* would be a secondary derivative based on the consonant stem of the root. This means that the element *-*kə* was also a suffix deriving secondary nominal stems, though it is less clear whether it was deverbal or denominal.

In this connection, it may be recalled that many Uralic suffixes, especially those deriving deverbal nouns or nominalizations, have two forms, one with the final low vowel **A* and the other without it or, alternatively, with the final reduced (high) vowel **ə* (**i*). Thus, apart from the relatively rarely attested suffix *-*kA*, the vowelless variant *-*k* is also attested in a similar nominalizing function, as in Finnish *lähte-* 'to depart' : *lähde-*' 'spring (of

[5] Tapani Salminen (p.c. 4 July 2024) notes that, because of a possible (though not fully confirmed) difference in the vocalism, the Nganasan verbal root connected (in SW) with **pår*(-)*kå* may actually represent a different etymon. This does not, however, rule out the possibility that Samoyedic **pår*(-)*kå* is a deverbal noun.

[6] According to Christopher Culver (p.c. 16 September 2024 and Culver 2025) the derivation of **poŋkå* from the verbal root **poŋ-* would seem to be confirmed, apart from Hungarian *fog* 'to catch', also by Mari *poŋ.a-m* 'to catch fish by driving it into a net'. Culver further notes that the verbal root might be connected with the noun **poŋə* 'bossom' (> Finnish *povi*), another possible case of a Proto-Uralic *nomen-verbum*.

water)' < **läkte-k* (SSA 2 : 121—122), and this same suffix is also used as the productive marker of the basic imperative and connegative form of verbs. The imperative paradigm shows a synchronic alternation between **-k* and **kA-*, as in Finnish (second person singular) *lähde-'* 'go!' < **läkte-k* : (corresponding plural) *lähte-kä-ä* 'go (you all)!' < **läkte-kä-tä*, a morphophonological feature shared also by Samoyedic. The background of this alternation, which goes back to Proto-Uralic, is unknown.

2. The cluster **lk* in the Uralic etymological corpus

To gain some more clarity on the issue we can take a look at the etymological corpus of items containing the cluster **lk*. Since this stem structure is best preserved in Finnic, a sufficiently large number of relevant examples can be obtained from the standard reverse database of Finnish (SKKS), complemented by a few items from other sources. Apart from nominals ending in Proto-Finnic **lkA*(-) and **lki* : **lke-*, there are verbs in **lkA-* and **lke-*, as well as derived stems in **lkA-tA-*, **lko*(-), **lkU*, **lki*, and **lke-tA*. This gives us the following corpus of Finnic items, listed below without glosses:

- nominals in **lka*: *helka*(-), *ilka*, *jalka*, *malka*, *olka*, *sulka*, *talka*
- nominals in **lkä*: *nälkä*, *selkä*, *ylkä* ,**pelkä* > *peuka-lo*
- verbs in **lka-*: *alka-*, **valka-*
- nominals in **lki* : **lke-*: *elki*, *jälki*, *kylki*, *melki*, *olki*, *nolki*, *solki*, *sylki*, *telki*
- verbs in **lke-*: *kulke-*, *nylke-*, *polke-*, *sulke-*, *sylke-*
- derived verbs in **lkA-tA-*: *hylkää-*, *pelkää-*
- derived nominals in **lko*(-): *halko*, *pelko*, *palko*, *salko*, *selko*, *valko*, *ulko*
- derived nominals in **lkU*: *alku*, *hylky*, *kulku*, *nylky*, *polku*, *sylky*, *ulku*
- derived nominals in **lke-tA*: *ilkeä*, *julkea*, *nilkeä*, *selkeä*, *valkea*
- derived adverbs in **lki*: *halki*, *julki*

It has to be noted that the list contains several cases of derivationally interrelated word forms, notably: *ilka* : *ilkeä*, *alka-* : *alku*, *sylki* : *sylke-* : *sylky*, *kulke-* : *kulku*, *nylke-* : *nylky*, *polke-* : *polku*, *sulke-* : *sulku*, *hylkää-* : *hylky*, *pelkää-* : *pelko*, *halko* : *halki*, *selko* : *selkeä*, *valko* : *valkea*, *julkea* : *julki*. Also, among the independent etyma in the list, there are a few confirmed loanwords: *helka*(-) (from Scandinavian 'holy', SSA 1 : 152 s.v. *helatorstai*), *malka* 'pole' (from Baltic, SSA 2 : 143), *palko* 'pod' (cf. also *palje* : *palkee-* 'bellows', Estonian *pale* : *palge-* 'cheek, face', from Germanic, SSA 2 : 300, 302; EES 349—350), *telki* 'bolt' (from Germanic, SSA 3 : 281). Of a secondary origin are probably also *talka* 'axe butt' (possibly a recent creation with restricted dialectal distribution in Finnish, SSA 3 : 261) and *velka* (from Proto-Finnic, but containing a Post-Proto-Uralic vowel combination, SSA 3 : 424). Several other items also seem to be restricted to Finnic with no established further connections: *elki* 'habitus' (obsolete in the basic form, but used in the derivative *elje'* : (plural) *elkee-t* 'gestures', Karelian *elgo* 'reason', SSA 1 : 103), *hylkä-ä-* : *hylä-t-* 'to abandon' : *hylky* 'wreck' (SSA 1 : 197—198), *julkea* 'brave' : *julki* 'made publicly disclosed' : *julista-* 'to declare' (SSA 1 : 2 v46), *polke-* 'to tread' : *polk-u* 'path' (SSA 2 : 390), and apparently also *alka-* 'to begin' : *alk-u* 'beginning' (borrowed to Saami as *álge-* : *álgu*, SSA 1 : 69; YSS no. 46), with proposed but semantically

and phonologically unacceptable comparanda in Khanty and Mansi, as well as in Samoyedic (UEW 6—7).

Proceeding to the items that have confirmed cognates in the other branches of Finno-Ugric (not counting Samoyedic), we can see that the cluster **lk* is normally preserved as **lk* > *lg* in Saami (quoted below mainly from Northern Saami) and Mordvin (quoted from either Erza of Moksha), but is metathesized to **kl* > *xl* > *γl* > *wl* in Khanty and Mansi (quoted in one sample dialectal form for each), while in Mari (Hill and/or Meadow), Permic (Udmurt and/or Komi), and Hungarian there is no direct trace of the segment **k*, leaving only the lateral **l* > *l* or, in Hungarian, depending on the item, also the geminate *ll*. Etyma that would seem to follow this pattern include the following:

- *jalka* 'foot/leg' (SSA 1 : 234) < *yalka = Saami *juolgi* (YSS no. 292), Moksha *jalga*, Mari (Hill) *jal* = (Meadow) *jol*. Hungarian *gyalog* 'on foot', traditionally linked to this etymon (MSzFE 1 : 230), is an unlikely cognate: the representation of initial **y-* as *gy-* is irregular, and the suffixal *-g* is idiosyncratic (WOT 1318). In view of both the form and the meaning, the word looks much more like a reflex of Turkic **yadag* 'on foot' (EDT 887). The representation of Turkic **y-* > **j-* (conventionally written **ǰ-*) as *gy-* is typical of West Old Turkic loanwords in Hungarian, and the representation of **-g* as *-g* is also well attested (WOT 1082—1083, 1092—1093), while for Turkic **-d-* there is only one uncertain example (*idő*, WOT 437—439), leaving it a likely possibility that this consonant, pronounced *[δ], could yield *l* in Hungarian, although this would have happened later than the earlier change of Uralic **d* to **l*.
- *jälki* : *jälke-* 'trace' (SSA 1 : 256) < **yälkə* = Hungarian *jel* : *jele-* 'sign', also > *jegy* : *jegye-* 'token, ticket' (MSzFE 2 : 338—339), with an apparent cognate in Khanty (Vasyugan) *jäγəl* 'pattern, spot (as on animals)' (DEWOS 341). In this case, the Hungarian and Khanty items fill the phonological and semantic criteria, and the etymology may be considered as sound in spite of the absence of data from the intermediate branches. In Estonian the word has been partly confused with **yärki* : **yärke-* > *järg* : *järje-* 'order' (EES 104, 105).
- *kylki* : *kylke-* 'side' (SSA 1 : 461—462). The only potential Finno-Ugric cognate that has been proposed for this Finnic noun is Hungarian *kül-* : *kül.ö-n* 'outside', which, indeed, would formally fit the regular pattern of **lk* being represented as *l* in Hungarian. However, in spite of its superficial suitability, the Hungarian item is actually a truncation of the longer form *kívül* 'outside (of), apart (from)', itself a primary ablative case form in *-l* of the spatial root *kív-*, reflected also in *ki* : *kint* ~ *künn* 'out(side)', with cognates in Khanty and Mansi (MSzFE 363—365) and with no relation to the Finnic data. Therefore, in the lack of other Finno-Ugric points of comparison, connections for the Finnic item have also been sought in Slavic and/or Baltic, but with no conclusive result (Junttila 2015 : 212).
- *nylke-* 'to skin' : *nylk-y* 'skinning' (SSA 2 : 246—247) < **ñülkə-* = Saami *njalga-* 'to lose hair/feather' (YSS no. 755), Mordvin (Moksha) *ńeĺgə-* 'to deprive (of)', Khanty (Konda) *ñəγət-ma-* 'to molt' (DEWOS 1035).
- *olka* 'shoulder' (SSA 2 : 263) < **w/olka* = Saami *oalgi* (YSS no. 827), Hungarian *váll* : *válla-* (MSzFE 669), for which a connection with Mansi

(Sosva) *wāγlap* 'strap' (WWb 713), including 'shoulder strap', has been proposed (Liimola 1951). The Mansi item has alternatively been assumed to derive from the verb *wāγl-* 'to descend' < **wïlka-* (on which see below), but a connection with 'shoulder' appears more likely, and the etymology may be considered as fairly certain.

- *olki* : *olke-* 'straw' (SSA 2 : 263—264) < **olkə* = Mordvin (Erza) *olgo*, absent elsewhere, apparently also in Moksha. As a Finnish loanword, the item is, however, represented in Saami as *oalga* 'stalk'.
- *nolki* : *nolke-* 'slime, spittle' (SSA 2 : 229—230), Karelian *ńolki* (with a secondary palatal nasal), Estonian *nõlg* : *nõle* (EES 324) < **nolkə* or **ñolkə* = Saami *snuolga* (with a secondary prothetic sibilant, YSS no. 744), Mordvin (Erza) *nolgo*, Hungarian *nyál* (MSzFE 475—476). There is an obvious descriptive component in this etymon, and an association with the verb *nuole-* 'to slick' (SSA 2 : 229) < **ñal/ə-* = Hungarian *nyal* (MSzFE 474—475) is apparent especially in Hungarian. Even so, the Hungarian data fit both formally and semantically into the comparison.
- *salko* 'pole' (SSA 3 : 148) < **çalka-* : **çalka-y* = Saami *čuolgu* (YSS no. 193), Mordvin (Moksha) *śalga* : verbally *śalgo-* 'to stick', Khanty (Vasjugan) *sayəl* 'slat, shingle' (DEWOS 1313—1314), Mansi (Northern) *sāγla* 'pole' (WWb 510), Hungarian *szál* : *szála-* 'stick, fiber, thread', also used as a counter for long and narrow objects (MSzFE 563—564). A distant lookalike that has been quoted from Komi in this connection may be disregarded as a false comparison.
- *selko* 'clarity' : *selkeä* 'clear' : *selke-ne-* ~ *selki-ä-* 'to become clear' (SSA 3 : 167) < **çelkə-* = Saami *čielggas* (YSS no. 151). The Finnish items stand in an irregular relationship with the semantically close item *selvä* 'clear' (SSA 3 : 168), though the variation *lk* vs. *lv* is attested also elsewhere (see examples below).
- *selkä* 'back (of body)' (SSA 3 : 167) < **çelkä* = Saami, with irregular vocalism, *čielgi* (YSS no. 152), and with a plausible but likewise irregular cognate in Mari (Meadow) *šəl*(-)*əž* 'sacrum'.
- *sulke-* 'to close' (SSA 3 : 211) : *sulk-u* 'closure' < **çulkə-* = Mordvin (Erza) *śolgo-* = (Moksha) *śolgə-*, but with no further cognates, and absent also in Saami (unless connected with *solki*, on which see below).
- *ulko*(-) : *ulo-* 'outside' : (lokative) *ulko-na* : (ablative) *ulko-a* : (lative) *ulo-s* (SSA 3 : 370) = Saami *olgu*(-) : (locative) *olgu-n* : (lative) *olggo-s*, Permic (Komi) *ïl-* 'distance' : (locative) *ïl.ï-n* 'far away'. The Finno-Permic comparison (cf. also UEW 803) is phonologically regular and semantically acceptable. Comparisons with Finnish *ulo-tt-u-* 'to stretch, to reach' (SSA 3 : 371) = Saami *olli-* (YSS no. 802) and Finnish *ulappa* 'open space of water' (SSA 3 : 369—370) are, however, far-fetched and would presuppose a low-vowel root of the type **ula-*, which would not be immediately compatible with the cluster **lk*. There is also a problem with the meaning, which is why these comparisons are best rejected.
- *ulku* 'pole' (SSA 3 : 370) < **ulkə-w* : **ulkə* = Saami (dialectally with an irregular prothetic laryngeal initial) *h/olga* (YSS no. 803), Mordvin (Moksha) *olga*, Permic (Komi) **ïl* 'tent pole', Khanty (Kazym) *ŏxəʟ* (DEWOS 39), Mansi (Northern) *åwlä* = *āwla* (WWb 399).
- *valka-* : *valka-ma* 'haven' (SSA 3 : 399) < **wïlka-* (Sammallahti 1988 : 551; Zhivlov 2014: 119) = Saami *vuolgi-* 'to leave' (YSS no. 1434), Mordvin

(Erza) *valgo-* 'to descend', Mari *wale-*, Khanty (Vakh) *waγəl-* (DEWOS 1574—1575), Mansi (Middle Lozva) *wāγl-* (WWb 713), Hungarian *vál-* : *vál-ik* 'to depart; to become' (MSzFE 667—668). Formally, the Saami cognate could also reflect the etymon **alka-* 'to begin', with which it may have become confused.

- *ylkä* (SSA 3 : 490) < **ülkä* 'boy, man, bridegroom', attested almost exclusively in Finnish, but marginally documented also in Estonian as *ülg* = Saami (Aanaar) *alge* (YSS no. 10). In the etymological literature, this item is separated from Finnish *yrkä* with exactly the same meaning, and the latter is connected with Mari (Hill) *ergə* 'boy' and Hungarian *-ér* in *férj* 'husband' (*fi-ér-*) and *férfi* 'man' (< *fi-ér-fi*), as well as with *-ër* in *embër* 'man, human' and *-ar* in *magyar* 'Hungarian' (MSzFE 150—151, 203). Needless to say, these comparisons are multiply mistaken, for there is no doubt that Finnish *yrkä* is simply a secondary variant of *ylkä*. Some of the Hungarian items quoted in this connection (though not *magyar*)[7] are likely to contain a trace of Turkic **er* 'man' (which is, incidentally, mentioned in UEW 84).

For a few items, certain languages show the expected regular representation in the consonantism, while there are unexplained irregularities in other details and/or in other languages. These items do not necessarily as such contradict the general picture, but the cases of irregular representation need to be noted, as they weaken the potential validity of the etymologies. This is the case in:

- *halko* 'split firewood' : *halki* 'split' : *halje-t-* : *halke-a-* 'to split' (SSA 1 : 133—134) < **šalka-* ≈ Erza *čulgo-* 'to peel', Permic (Komi) *šulʹ-al-* 'to split'. In this case, the unexplained irregularities in both the consonants and the vowels, and even in the meanings (in Mordvin), are too numerous to make the comparisons viable, leaving the Finnic item without any confirmed cognates in the other branches.
- *melki* : *melke-* 'breast of animal' (SSA 2 : 158), attested only in the North Bothnian dialects of Finnish and therefore likely to be a borrowing from Saami **mälkə* > *mielga* id. (YSS no. 666) = Erza *melʹke* 'breast', Mari (Meadow) *mel*, Permic (Udmurt) *mïl*, Khanty (Obdorsk) *mewəl* (DEWOS 909—910), Mansi (Northern) *māγl* (WWb 293), Hungarian *mell* : *melle-* 'breast' > 'location beside': (locative) *mell-e.tt* 'beside' : (lative) *mell-é* '(to) beside' (MSzFE 430—431). Here, only Mordvin shows an unexpected representation in the consonantism (perhaps a contracted derivational form, with truncation of the root). However, apart from the Saami borrowing, Finnish has *mälvi* 'breast of bird' (SSA 1 : 133—134), also attested in Estonian as *mälv* : *mälve* (EES 295), suggesting an irregular variation between **lk* and *lv* (as in *selkeä* : *selvä*). Even so, the etymology is sound, with **mälkə* as the original form (so also UEW 267 and Zhivlov 2023 : 167).
- *solki* : *solke-* 'buckle' (SSA 3 : 196) < **çolkə* = Erza *śulga-mo*, Mari (obsolete) *šolka-ma*, also compared with Saami *čulgom* 'side piece of fur shoe'. The consonantal correspondence between Finnic and Mordvin is regular, but the Mari item must be an areal borrowing from Mordvin, which is

[7] The final *-r* in *magyar* (~ *mogyër* ~ *mëgyër*) is most probably a non-Uralic plural marker, widely attested in Inner Asian ethnonyms and this case apparently of a Bulghar Turkic origin (Janhunen 2018).

not surprising since we are dealing with an item of cultural vocabulary (with additional reflexes in Chuvash and Russian dialects). In spite of the difference in the vocalism, there might be a connection with the verb **çulkə-* 'to close', from which the Mordvin (and possibly the Saami) item would be a regular nominalization. A further comparison with Permic (Komi) *śul* 'runner of sledge' (UEW 775) is obviously mistaken.

- *sylki* : *sylke-* 'spit' : *sylke-* 'to spit' : *sylk-y* 'spitting' (SSA 3 : 229—230) < **çülkə*(-) = Saami *čolga* : *čolga-* (YSS no. 161), Mordvin (Erza) *śeľge* : *śeľge-*, Mari (Meadow) *šüwəl*(-), Permic (Udmurt) *śal* = (Komi) *śël*, Khanty (Vasyugan) *sŏ̈jəγ-* (DEWOS 1298), Mansi (Northern) *saľγ-* (WWb 520—521). In this case the data from the West Uralic branches (Finnic-Saami-Mordvin) stand in a regular relationship with each other and represent a *nomen-verbum*, which itself may have more ancient roots in Uralic. However, the other proposed cognates show irregularities in the vowels (Permic) or consonants (in the other branches). It is particularly noteworthy that Khanty and Mansi do not indicate the regular metathesis in this item. This suggests random descriptive variation (as also recognized in UEW 479—480), which does not allow a uniform reconstruction beyond West Uralic.

Of greater relevance for the current topic are the items in which the cluster **lk* can be shown to be divided by a derivational morpheme boundary. In these cases, the root is a consonant stem ending in **lə*(-) : **l-*, to which the suffixes **-kA*(-) or **-kə*(-) have been added. The basic root is preserved either as such or in the composition of other (correlative) derivatives which lack the elements **-kA*(-) or **-kə*(-). In some cases, the original morpheme boundary is revealed by simple internal reconstruction based on a single language or branch (especially Finnic), while in other cases the derivational origin is confirmed by comparative evidence from other Finno-Ugric branches. In this context, because of the apparent loss of the stop segment of the cluster **lk* in Mari, Permic, and Hungarian, information from these branches is inconclusive. The derivation of the cluster **lk* from the complex structure **l-k* may be considered as fully confirmed in, at least, the following two cases, both of which involve deverbal nouns derived by the suffix **-kA*:

- *nälkä* 'hunger' < **ñäl-kä*, a nominal derivative from the verb **ñäl/ə-* > Finnic **neel/i-* > Finnish *niel/e-* 'to swallow' with the regular development **ä—ə* > Finnic **ee—e-* : **ee—i* (Luobbál Sámmol Sámmol Ánte (Aikio) 2012).[8] The Finnic derived stem has a cognate in Saami *njálgge-s* 'sweet' (YSS no. 762), possibly also in Ter Saami *ńielge* 'meagre fish' (SSA 2 : 251), but it is also present in Saami in the form *nealgi*, borrowed from Finnish (YSS no. 721, cf. also Aikio 2002 : 53 note 6, 2015: 40). A related derivative is **ñäl-mä* 'tongue, language' (Janhunen 1992 : 239—240; Luobbál Sámmol Sámmol Ánte (Aikio) 2015 : 36; O'Rourke 2016 : 244), which yields Saami *njalbmi* 'mouth' (YSS no. 763), Mari (Hill) *jəlmə*, Khanty (Vakh-Vasyugan) *ńäləm* (DEWOS 1049—1050), Mansi (Northern) *ńēləm* (WWb 360), Hungarian *nyelv* (MSzFE 480—481), probably also Finnish *nälvä* 'slime' : *nälvi-* 'to mock' (cf. SSA 2 : 251, where a "descriptive"

[8] As has been pointed out by Pystynen (2018 : 47—50 *et passim*), the idea concerning the developments **ä* > **ee* and **a* > **oo* in Pre-Proto-Finnic goes back to Meri Lehtinen (1967), who first identified this regularity, also known as "Lehtinen's Law".

origin is proposed). The verbal base **ñäl/ə-* is attested in all branches of Finno-Ugric: Saami *njiella-* (YSS no. 776), Mordvin (Erza) *ńilʹe-*, Mari (Hill) *nel.ä-*, Permic (Komi) *ńӧl-*, Khanty (Vakh-Vasyugan) *ńel-t-* : *ńelə-ŋ* 'hungry' (DEWOS 1042—1043), Mansi (Northern) *ńāl-t-* (WWb 352), Hungarian *nyel* (MSzFE 479).

- *pelkä-ä-* : *pelä-t-* 'to fear, to be afraid' : *pelko* 'fear' (SSA 2 : 335) < **pel-kä-* 'fear'. In this case, the form with **lk* is present only in Finnic, while all the other Finno-Ugric languages point to the basic verbal root **pel/ə-* 'to fear', represented as Saami *balla-* (YSS no. 855), Mordvin (Erza) *pelʹe-*, Permic (Komi) *pol-*, Khanty (Vakh-Vasyugan) *pəl-* (DEWOS 1142—1143), Mansi (Northern) *pil-* (WWb 440), Hungarian *fél* (MSzFE 198). There is no doubt that the Finnic forms are based on the deverbal noun **pel-kä* 'fear', from which both the secondary noun *pelko* and the verb *pelkä-ä-* < **pelkä-tä-* are further derivatives.

It is, then, obvious that the word for 'feather, wing' fits the pattern of **ñäl-kä* and **pel-kä* and may be reconstructed as **tul-ka* at the Finno-Ugric level. Another item that must be discussed in this context is Finnish *peukalo* 'thumb' and its cognates.

- *sulka* 'feather' (SSA 3 : 211) < **sulka* < **tulka* 'feather, wing', with secondary palatalization of the initial dental stop **t-* > **ç-* > *s-* in Proto-Finnic[9] = Saami *dolgi* (YSS no. 1267), Mordvin (Erza = Moksha) *tolga*, Permic (Komi) *tӧl*, Khanty (Vakh-Vasyugan) *tŏɣəl* (DEWOS 1412), Mansi (Northern) *towəl* (WWb 668), Hungarian *toll* : *tolla-* (MSzFE 637). A minor aberration is present in the Estonian cognate, which has secondarily been transferred to a different stem type, yielding **sulki* : **sulke-* > *sulg* : *sule-* (EES 487). If we accept that this word is a derivative of the *nomen-verbum* attested in Finnish as *tuule-* 'to blow (of wind)' : *tuuli* : *tuule-* 'wind', there remains the issue concerning the difference in vowel quantity. It may be noted that the cognates of *tuuli* (SSA 3 : 340—341) in other Uralic languages, including Mari (Hill) *tul* < **tul* and Permic (Komi) *tӧl*, are different from those of *sulka*, and also different from those of the minimal pair *tuli* 'fire' (SSA 3 : 323—324) < **tul/ə* = Saami *dolla* (YSS no. 1266), Mordvin (Erza = Moksha) *tol*, Mari (Hill) *tə̑l* < **tŭl*, Permic (Komi) *tӧl* (so also Itkonen 1969 : 91—92 for Mari, though not for Permic). This means that the synchronic vowel length in Finnic must in this case derive from an additional feature that was originally present in the root, and the best candidate for this feature is the "laryngeal" consonant **x*. We do not know at what stage this consonant was segmentally lost, but there is no direct trace of it in the derivative **tul-ka*, suggesting that there may have been a phonotactic process deleting it before a syllable-final consonant, that is: **tuxl-ka* > **tul-ka*. Whatever the chronology of this development, it should be

[9] The development of initial **t-* to **ç-* (traditional **ś-*) to **s-* in Finnic is exceptional and apparently irregular, but it has a well-known parallel in the correspondence of Finnic **soo* > Finnish *suo* 'bog, marshland' (SSA 213—214) to Hungarian *tó* : *tava-* 'lake' (MSzFE 634), with cognates in Mansi, Khanty, Permic, and Samoyedic (Janhunen 1981 : 39 no.109). In these cases we have to accept the factor of sporadic sound change, and there is no need to complicate the obstruent system of the reconstructed protolanguage with an additional segment, as has been proposed by Honti (2002; 2013 : 18—24).

obvious that the basic root for 'feather, wing', which underlies the derivative **tul-ka*, cannot have been of the simple form **tul/ə* (or **tuli*, as suggested by O'Rourke 2016 : 244), for this would not explain its systematic difference with regard to **tul/ə* 'fire'.

- *peukalo* 'thumb' (SSA 2 : 346) < **pelka-lo* : **pelkä-lä* < **pelkä*(-) = Saami *bealgi* (YSS no. 909), Mordvin (Erza) *pelʹka* = (Moksha) *päʟʹkä*, Permic (Udmurt) *pëlï* = (Komi) *pel*. The Finnic word is somewhat variable, as exemplified by Estonian (older and dialectal) *peial* : *peigla* (< **peükä-l/ä*) ~ (modern standard) *pöial* : *pöidla* (EES 406). Mordvin suggests a secondary derivative of the same type as in the reflex of **mälkə* 'breast' (see above), which means that it shows no unambiguous trace of the cluster **lk*, while Permic is inherently unambiguous for this detail. However, there is reason to assume that **pelkä* = **pel-kä* is actually a derivative of the etymon represented in Finnish as *pieli* : *piele-* '(outer) side, edge, jamb (of door)' (SSA 2 : 347—348) = Mordvin (Erza) *pelʹ* = (Moksha) *päl′*, Mari (Hill) *pel*, Hungarian (*ajtó-*)*fél* : (-)*fele-* 'jamb (of door)' (MSzFE 197—198), implying that the thumb is the "outermost" finger (Janhunen 1981 : 23 no. 86). In this case, the base root, Proto-Finnic **peeli* : **peele-*, seems to be a noun with a structure analogous to that of **tuuli*(-) : **tuule-*, and the explanation of vowel length in both items must be the same and different from the case of **neele-* < **ñäl/ə-* (on which see above). In the etymological literature, **peeli* has often been confused with **pälä* 'half, one of two, relative, friend' (UEW 362—363), and the two etyma have collapsed in Hungarian (MSzFE 194—196), but are distinct elsewhere, including Finnic, where the latter yields ultimately **pooli* > Finnish *puoli* : *puole-* 'half' (Luobbál Sámmol Sámmol Ánte (Aikio) 2015 : 47—48, 63). A separate item that may actually be another derivative of **peeli* is **pel*(-)*yä* 'ear', with reflexes in all Finno-Ugric branches except Finnic, including Saami *bealji* (YSS no. 908) and Hungarian *fül* 'ear' : *figy-el* 'to listen' (MSzFE 224—225).

While the four items discussed above — **ñäl-kä* 'hunger', **pel-kä* 'fear', **tul-ka* 'feather, wing', and **pel-kä* 'thumb' — are all nouns derived with the suffix **-kA*, there are also several items that suggest a verbal derivative in **-kə-*:

- *kulke-* 'to go, to wander, to pass, to proceed' : *kulku* 'passage' (SSA 1 : 429—430) < **kulkə-* = Saami *golga-*'to flow, to leak' (YSS no. 442), Mordvin (Erza) *kolʹge-* 'to drop, to leak', Permic (Komi) *kïl-al-* 'to float downstream', Khanty (Vakh-Vasyugan) *kɔɣəl-* 'to walk' (DEWOS 457—458, suggesting **kukla-* < **kulka-*, Zhivlov 2023 : 164), Hungarian *hala-d* ~ *halla-d* 'to move foreward, to proceed'. It should be obvious that this verb is a correlative derivative of the Finnish reflexive stem *kul-u-* 'to pass (of time), to be worn' (SSA 1 : 433), which derives from the basic root **kul/ə-* = Saami *golla-* 'to pass (of time), to be wasted' (if not a borrowing from Finnish), with a number of phonologically and/or semantically unlikely cognates in Permic, Khanty, and Mansi, including Hungarian *hull* 'to fall (leaves, hair), to flow (blood, tears)' (MSzFE 308—309). Another possible derivative is the causative **kul-ta-* 'to fish with a dragnet' (Luobbál Sámmol Sámmol Ánte (Aikio) 2022 : 9) > Saami *goldi-* (YSS no. 445) = Finnish *kulta-* (if not a borrowing from Saami, SSA 1 : 432—433).

- *valke-* : *valke-a* 'white, bright' (< **walke-ta*) : *valke-ne-* 'to become white, bright' (< **walke-m/e-*) : *valko* 'white colour' : *valko-inen* 'white' (SSA 3 : 399—400) vs. *vaale-a* 'light-coloured' (SSA 3 : 384) vs. *valo* 'light' (SSA 3 : 401). In this case, the Finnic synchronic data suggest a verbal root of the type **wal/ə-* 'to be bright', from which the derivative **wal-kə-* 'to be bright' was formed. There is some descriptive variation in the form of this etymon, due to which the stem consonant can be geminated and/or the vowels can be fronted, as in Finnish *välkeä* 'light' : *välkky-* 'to glitter' (SSA 481), Estonian *välk* : *välgu-* 'lightning' (< **wälkko*, EES 620). The only cognate beyond Finnic which shows an unambiguous trace of the cluster **lk* is Saami *vielga-d* 'white' (< **wälkə-*, YSS no. 1399), while Mordvin (Erza) *val-do* 'light' would rather seem to be based on the simple root. A somewhat enigmatic case is offered by Mari (Hill) *walgədə* = (Meadow) *wolgədə* 'brightness', which certainly contains the same root, but in which the synchronic cluster *lg* can hardly reflect an original **l-k*, as this would be the only example of this kind in Mari. More likely, the Mari item has a different derivational structure, unless it is a borrowing, though the exact source would in that case be difficult to identify. In addition, apart from unlikely lookalikes in Permic, Khanty, and Mansi, the reconstructions **wal/ə- ~ *wal-kə-* have a true reflex in Hungarian *világ* : *világo-* 'light' (> 'world') : *villám* : *villamo-* 'lightning' (MSzFE 691—693), which confirm that the original shape of the root was **wïl/ə-*, with the later regular merger of **ï* with **a* in West Uralic.
- *ilke-* : *ilke-ä* : *ilki-* 'bad, unpleasant' : *ilke-ä-* ~ *ilke-ne-* 'to dare' (SSA 1 : 223—234). This etymon is restricted to Finnic, where it has also forms with a velar vocalism, including *ilka* ~ *ilko*(-) 'trick'. It may also be connected with *ilve* : *ilvee-* 'trick' (SSA 1 : 226, cf. *selkeä* : *selvä*), but, more importantly, and in spite of the superficial difference in the meaning, it may be related to *ilo* 'joy, merriment, uproar, noise' (in Southern Finnic also > 'beauty', SSA 1 : 225; EES 91). If so, the original root might have been **il/ə-* 'to be tricky' (?) with front vocalism and with a basically verbal morphology. However, since this item has no confirmed cognates in the other branches of Finno-Ugric, it is only relevant as a possible additional case of the derivational relationship between **l* and **lk* (= **l-k*).
- *nilke-* : *nilke-ä* 'slimy' : *nilki* 'hairless (skin)', also with back-vocalic forms. including *nilka-* : *nilko-a-* 'to remove skin' (SSA 2 : 221—222) and *nilja* ~ *nilva* ~ *nilpa* (SSA 2 : 222) 'slimy'. This etymon shows a lot of descriptive variation and is to some extent confused with *nylke-* 'to skin' < **ñülkə-* (as discussed above), but it is also clearly associated with *nila* 'slime, phloem, skin (of tree)' (SSA 2 : 221) = Saami *njalli* 'phloem' : *njalla-* : (Aanaar) *njal-de-* 'to remove skin (of tree)' (YSS nos. 754, 756), Permic (Komi) *ñilʹïd* 'slimy', with further lookalikes in Khanty and Mansi. A formally separate item is Finnish *nuli* 'hairless' (SSA 2 : 237), and likewise separate are Mordvin (Erza) *nola* 'phloem', Mari *nolo* 'sap', Khanty (Obdorsk) *ńel* 'sap' (DEWOS 1043), Mansi (Northern) *ńalʹ* 'slime, sap' (WWb 352), all of which also resemble *nolki* 'slime, spittle' (as discussed above).

To take a comprehensive look at the corpus, we may once more list the Fin(nish) etyma with **lk* with their attestations in the other Finno-Ugric

branches, including Saa(mi), Mor(dvin), Mar(i), Per(mic), Kha(nty), Man(si), and Hun(garian), as presented in Table 1 below. The list contains 27 separate etyma, quoted here in the same order as discussed above. Non-cognates, such as confirmed inter-branch borrowings or false etymologies based on unacceptable comparisons are excluded from the list. From the point of view of the present topic, the acceptable cognates may be divided into two categories: those which show a trace of the cluster **lk* (marked +) and those which do not show such a trace (marked o), which means that they only show a segmental trace of **l*. The absence of a trace of **k* in the latter category can depend on several circumstances:

(a) the cognates in the non-Finnic languages may altogether represent a different (more basic) stem, as in the case of *pelkä-*;
(b) the situation may have been obscured by secondary derivational processes or truncation, as in the case of the Mordvin cognates of *melki*, *peuka-*, and *valke-*; or
(c) the cognates of items with **lk* are systematically indistinguishable from the cognates of those with **l*, as is the case in Mari, Permic, and Hungarian.

Table 1

Cognates of Finno-Ugric etyma with **lk* in Finnish/Finnic

Fin	Saa	Mor	Mar	Per	Kha	Man	Hun	T
jalka	+	+	o					2/3
jälki					+		o	1/2
kylki								0/0
nylke-	+	+			+			3/3
olka	+					+	o	2/3
olki		+						1/1
nolki	+	+					o	2/3
salko	+	+			+	+	o	4/5
selko	+							1/1
selkä	+		o					1/2
sulke-		+						1/1
ulku	+	+		o	+	+		4/5
ulko	+			o				1/2
valka-	+	+	o		+	+	o	4/6
ylkä	+							1/1
halko								0/0
melki	+	o	o	o	+	+	o	3/7
solki	+	+						2/2
sylki	+	+						2/2
nälkä	+							1/1
pelkä-	o	o		o	o	o	o	0/6
sulka	+	+		o	+	+	o	4/6
peuka-	+	o		o				1/3
kulke-	+	+		o	+		o	3/5
valke-	+	o	o				o	1/4
ilke-								0/0
nilke-	o			o				0/2
T	19/21	12/16	0/5	0/8	8/9	6/7	0/10	

The calculations based on the data (marked T), indicate, in the vertical columns, the total number of cognates found in each branch for the Finnish etyma, and, on the horizontal lines, the number of cognates that each particular etymon has in the non-Finnic branches of Finno-Ugric (with the exclusion of Samoyedic). In each case, the number of positive matches for **lk* (+) is related to the total number of etymological matches, which also include the cases in which no unambiguous trace of **lk* is present (o).

The table reveals several interesting facts. First, though not unexpectedly, the largest overall number of cognates of the Finnish items is present in Saami, where a cognate is found for as many as 21 of the 27 items, with 19 of them showing an unambiguous trace of the cluster **lk* (19/21). Saami is followed by Mordvin (12/16), suggesting the relevance of the West Uralic branch (Finnic-Saami-Mordvin). By contrast, conspicuously few cognates are present in Mari (0/5) and Permic (0/8), two branches in which the reflexes of **lk* are indistinguishable from those of the single consonant **l*. This is in accordance with the central position and generally "worn" state of the languages of these two branches (which may ultimately represent sub-branches of a single more primary branch). Rather few cognates are also present in Khanty (8/9) and Mansi (6/7), while, surprisingly, Hungarian appears to have slightly more, though with no trace of the cluster **lk* (0/10).

The number of cognates on the horizontal lines is also a measure of the credibility of the etymology: the fewer cognates an item has, the less credible the etymology is. At the same time, assuming that linguistic distances correlate, at least to some extent, with geographical distances, items that are represented in mutually distant branches are likely to be older than those present in two adjacent branches, in which they can be shared either on a genetic or an areal basis.[10] As it can be argued that Hungarian is the branch linguistically most distant to Finnic (which is the branch of reference here), it is important to note that all items shared by Hungarian and Finnic are also attested in at least one other branch. However, considering that Hungarian is conventionally grouped together with Khanty and Mansi in the so-called "Ugric" branch, only five Hungarian items are shared by both Mansi and Khanty (the cognates of *salko*, *valka-*, *melki*, *sulka*, as well as the basic root of *pel-kä-*), two by Khanty without Mansi (the cognates *jälki*, *kulke-*), and one by Mansi without Khanty (the cognate of *olka*). Also, considering that Khanty and Mansi are conventionally grouped in the so-called "Ob-Ugric" sub-branch, there is only one additional item that is shared by both Khanty and Mansi (the cognate of *ulku*). One would expect even more coherence between these languages, if the conventional taxonomy were correct.

In this connection, Hungarian deserves some more attention. While Khanty and Mansi, due to an apparently areal innovation shared by these two

[10] In the currently popular "comb" or "rake" model (on which see Salminen 2002), in which all the main branches of Uralic (including even Samoyedic) are supposed to be linguistically equidistant to each other, there can be no *a priori* differences in the numbers of inherited lexicon they share, though, of course, some languages can have more innovations and less retentions than others due to a variety of extra-linguistic contextual factors. Obviously, this new model, which presupposes a relatively recent no-time linear expansion of Uralic in the east-to-west direction (as implied in, for instance, Grünthal et al. 2022), accounts poorly for the empirically observed distribution of the shared lexicon.

branches, preserve the cluster **lk* in the metathesized form **kl*, Hungarian, like Mari and Permic, shows no sign of a similar metathesis, nor any other indication of the segment **k* in the etyma under discussion. In principle, in some etyma, this could mean that Hungarian never had the cluster **kl*, but only the basic underived stem with **l*, as in the case of *fél* 'to fear'. Even so, it has been suggested that the presence of a geminate *ll* in several items, including *váll*, *mell*, and *toll*, and the alternation of *ll* with *l* in the items *halad* ~ *hallad* and *villám* : *világ* would mean that the synchronic geminate actually reflects the cluster **lk*. The development would, then, have been something like **lk* > **lγ* > *ll* (cf. e.g. UEW 267 for *mell*, MSzFE 637 and WOT 1308 for *toll*). There are, indeed, cases in which the Hungarian geminate *ll* can be thought to stand for a former consonant cluster, as in *hall* 'to hear' < **hadl*V- < **kuntə-l/ə-* (MSzFE 253—255, more recently Holopainen 2023a : 122—123), a minimal pair to *hal* 'fish' < **kala* (MSzFE 250). However, this explanation does not seem to be possible in the case of **lk*, for there remain several other items which show a non-alternating single *l*, as in *jel*, *nyál*, *szál*, *vál-ik*. Unless these etymologies are wrong — which is, of course, possible — the only logical conclusion is that the variation *l* ~ *ll* in Hungarian is a secondary and random phenomenon, which gives no clue as to whether the original form had **l* or **lk*. Consequently, in a case like *villám* : *világ* there is no way to tell whether these items correspond to the simple root **wïl/ə*(-) (> Finnish *valo*) or the derived stem **wïl-kə-* (> Finnish *valkea*).

3. The non-canonical vowel sequences in Samoyedic

Several of the Finno-Ugric items discussed above have either assumed or confirmed cognates in Samoyedic. Some of them would seem to contain a vowel sequence on the Samoyedic side, while others contain a single vowel. The question is what this difference means and to what extent it is relevant for the reconstruction of the cluster **lk* as opposed to the single consonant **l*. This issue, like vowel sequences, in general, remains one of the most disputed topics in Samoyedic historical phonology, and a variety of opinions have been presented on it (most recently Зайцев 2024). However, in one case, at least, Samoyedic, like all the non-Finnic branches of Finno-Ugric, reflects a basic stem with **l*:

- **pelkä*(-) 'fear' : **pel/ə-* 'to fear' > Early Proto-Samoyedic **pey-* 'to fear' > Late Proto-Samoyedic **piy-* (SW 124—125, reconstruction updated).[11] In this item, the lateral **l*, after coming to stand in syllable-final position after the loss of the final vowel, is regularly represented as the palatal glide **y* [*j*] in Samoyedic. The basic root **pey-* > **piy-* is not attested as

[11] It is here assumed that Proto-Samoyedic involved two chronological stages: Early Proto-Samoyedic, which includes Nganasan, and Late Proto-Samoyedic, which excludes Nganasan but includes all the other Samoyedic languages: Enets (Tundra and Forest), Nenets (Tundra and Forest), Selkup (with all varieties), Kamas, and Mator. After the separation of Nganasan, the lineage of Late Proto-Samoyedic underwent several innovations absent in Nganasan (Янхунен 1991). Among them is the merger of the Early Proto-Samoyedic vowels **e* (< Uralic **e*) and **i* (< Uralic **i* & **ü*), as first noted by Helimski (2005). For Samoyedic items absent in Nganasan, the distinction between **i* and **e* can be verified only if a cognate is present outside of Samoyedic, that is, either in Finno-Ugric or in some non-Uralic external contact language.

such in any Samoyedic language, but it is present in several derivatives, including the inchoative Early Proto-Samoyedic *pey-m- 'to become frightened' > Nganasan *xï ïm-* = Late Proto-Samoyedic *piy-m- > Enets Tundra *piim-* = Forest *pim-*, Kamas *pim-*, Mator (augmentative) †*xim-er-*, the causative *pey-tä- > *piy-tä- 'to frighten, to chase away' > Enets Tundra *pii-ze-* = Forest *pi-ze-* (with *z* [*δ*] < *t*), Nenets Tundra *pyí-dye-* = Forest *pyi-tye-*, Selkup (Northern) < *pity-* > †*piitə-*, and the imperfective *piyntə- 'to be scared' > Enets Tundra *pii-do-* (EH ms.) = Forest *pi-do-*, Nenets Tundra *pyí-nə-* = Forest *pyi-nă-*. It may be noted that in Nganasan *xï ïm-*, the sequence *ey > *ïy is represented as the long monophthong *ï ï*, in which the glide *y is vocalized to a full vowel segment. This seems to be the regular representation of the sequence *ey > *ïy and is paralleled by the representation of *iy as *ii*, as in Nganasan *ciimi* 'fathom' < *tiy-mä (SW 163), although syllable-final *y normally remains asyllabic. Importantly, the original root-final *l is also preserved in Nganasan *xïle* < *pelə 'frightening, terrible'. This must be either the nominally used generalized vowel stem of the word, supposing that it originally functioned as a *nomen-verbum*, or a derivational form with an obscured and segmentally lost suffix after the root. A further derivative from this stem is apparently *xili'-* : *xilije-* 'to fear' (SNg 192), with regressive assimilation *e > *ï > *i* in the initial syllable. It may be concluded that Proto-Samoyedic had residual morphophonological alternations of the type *pey- (consonant stem) vs. *pelə- (vowel stem), conditioned by the circumstance that the development *l > *y was active only in syllable-final position, while *l was preserved before a vowel. A similar case seems to be present in Nganasan *xale* [hᵘalə] 'stone' (SW 112), which corresponds to Tundra Nenets *pae* < *pəy or *påy = Selkup *püü* (SkWb no. 385). In this case, the vowel stem with *l is preserved also in Mator *xilä* (Helimski 1997 : 245 no. 298).[12]

There are three other cases in which the Samoyedic cognates of Finno-Ugric items with *lk would seem to contain only a syllable-final *y, suggesting that the Finno-Ugric cluster is of a derivational origin, while Samoyedic preserves the original basic root:

- *ulku* 'pole' < *ulkə- : *ul/ə > Samoyedic *uy (SW 29—30) > Nganasan *ŋüj*, Enets Tundra *ŋu* (EH ms.) = Forest *ŋu-zu* [ŋuðu] (ES 304), Nenets Tundra *ŋú* = Forest *ŋu*, Selkup *üü* 'stanchion (of sledge)' (SkWb no. 8). The assumption that the Samoyedic data here reflect an original with the cluster *lk (as implied by Zhivlov 2023 : 166, so also Зайцев 2024) would mean that Samoyedic somehow lost the segment *k. This could have happened, for instance, by way of metathesis and the subsequent

[12] Without going into the details, it may be recalled that the reflexes of the sequences *əy and *åy are difficult to tell apart in the etymological material. Nganasan *xa* [hᵘa] normally derives from *pa, while *på and *pë yield *xo* [ho] and *pə yields *xe* [hə], but there are exceptions from this, cf. e.g. *xotï* 'bile' < *pətä vs. *xote-* 'to decorate, to write' < *påtə- (SNg 199; on the rules governing the occurrence of the vowel qualities in combination with consonants, see Kaheinen, forthcoming). There are also cases of variation between *ə and *a in Nganasan, as in *kəyŋ ~ *kay(--)ŋ 'thunder' (Kaheinen 2023 : 86). In Selkup, *åy would normally seem to yield *üü*, while *əy would yield *aa*, but, again, the representation is not systematic. Even so, the cognateship of Nganasan *xale* with Nenets *pae* and Selkup *püü* may be regarded as safe.

simplification of the new cluster (**ulkə* > **uklə** > **ul/ə*). However, one would expect this development to have affected also the cases in which the cluster **lk* was followed by a low vowel, of which there are no examples. It is therefore more likely that Samoyedic **uy* reflects the primary root **ul/ə*, while Finno-Ugric **ul-kə*(-) is a secondary derivative (as already suggested in Janhunen 1981 : 15 no. 53). Another issue concerns the Nganasan form of this word. It seems that the development **uy* > *üj* in Nganasan, as observed here, is regular, while **uy* > *uj*, as in **tul/ə* 'fire' > Samoyedic **tuy* (SW 166) > Nganasan *tuj* is irregular and suggests an intermediate stage with **oy*, which regularly yields *uj* in Nganasan, as in **toymå* 'larchtree' (SW 164) > Nganasan *tujmu-*. The frontening of **u* to *ü* is probably the rule in Nganasan, connected with the general vowel shift, which also raised **o* and **ö* to *u* (cf. Kaheinen 62—66). This means that any items in which Nganasan *u* stands for Proto-Samoyedic **u* must be due to an irregular merger of **u* with **o* > *u*.

- *kulke-* 'to move' < **kulkə-* : **kul/ə-* > Samoyedic **kuy-*. A link of the Finno-Ugric item (on which see above) with Samoyedic data has often been proposed in the etymological literature (cf. e.g. UEW 198), but the connection is problematic. In view of the items **tul/ə* > **tuy* 'fire' and **ul/ə-* > **uy* 'pole', the Samoyedic reflex of **kul/ə-* must be **kuy-* and not, for instance, **ku-* (SW 76) or **kuə-* (Zhivlov 2023 : 164; Зайцев 2024). The only Samoyedic data that would seem to fit this reconstruction are Nenets Tundra *xú-w°* = Forest *ku-w°* 'driftwood', as well Tundra Nenets *xú-'-la-* : *xú-'-lyo-* 'to float downriver', which are derived from the root *xú-* = *ku-* with a semantic profile reminiscent of the presumed Permic cognate. The Nenets data are, however, formally ambiguous, as they do not directly confirm the presence of **y* in the root. Within Samoyedic, the Nenets items have been compared with Selkup *quu*(-)*rə-* 'to float' (SkWb 330—331 no. 2284), but on the analogy of **uy* > *üü* and **tuy* > *tüü* (SkWb 139—140 no. 954) one would expect the vowel *üü* here, if the original is **kuy-*. A potentially more suitable point of comparison is Nganasan *kujü-* 'to float' (SNg 71). It seems that the intervocalic *j* can in this case only stand for original **y*, whose exceptional retention after the initial syllable may signal a complex origin of the second-syllable vowel (cf. Kaheinen 2023 : 94—95). If this is so, the only irregularity in the Nganasan item is the representation of **u* as **o* > **u* (cf. above). It may be concluded that the connection between Finno-Ugric **kulkə-* : **kul/ə-* with Samoyedic data can be neither fully confirmed nor fully rejected.
- *kylki* : *kylke-* 'side' < **külkə* : **kül/ə* > Samoyedic **käy* > **kəy* (as already preliminarily discussed above). Since the Finnic item in this case has no known cognates on the Finno-Ugric side, the Uralic etymology depends on the Samoyedic cognate, if accepted. The cognateship is formally correct, assuming that **ü* underwent the regular reduction after having merged with **i*, that is, **ü* > **i* > **ə* > **ə*. The situation is, however, not fully unambiguous, for, although the sequence **il* yields **äy* > **əy* in **çilmä* > **səymä* > **səymä* 'eye' (SW 132), the sequence **ül* > **il* yields **iy* in **sül/ə* 'fathom' > **tiy-* : **tiy-mä* (SW 163); note also that the Late Proto-Samoyedic **i* from **e*, as in **pel-* > **piy-* 'to fear',

did not undergo reduction. It is therefore possible that the original root was actually **kil/ə*, from which Finnic got the primary derivative **kil-kə* and only secondarily (by sporadic rounding) **kül-kə* > **külki*. However this may be, the root **kəy* > **kəy* is not attested as such in Samoyedic, but only in two parallel derivational forms which may be reconstructed as **kəy-ü* (Salminen 2023 : 378) > Nganasan *kei* vs. **kəy-ɯə* > Enets Tundra *koo* (EH ms.) = Forest *keo* ~ *kio* (ES 169), Nenets Tundra *xaew°* = Forest *kaew°*, Selkup *qöö* (SkWb no. 1768). The Nganasan form is bimoraic, as is obvious from its declension: (3rd person singular possessive) *kei-di* ~ *kei-du* (SNg 79, with secondary variation in the vocalism, see Salminen l.c.). In spite of the superficial difference, the forms **kəy-ɯə* and **kəy-ü* are actually full cognates, for they represent two different stems of a single derivative: the consonant stem **kəy-(ə)ɯ* > **kəyü* and the vowel stem **kəy-ɯə(-)*. The two parallel stems are also present in the derivative **kəyü-tə-yə* ~ **kəy-ɯə-tə-yə* 'rib' (SW 57—58), which yields Nganasan *kei-de-e* vs. Enets Tundra *koo-zi* = Forest *kiu-zi* ~ *kiu-ze*, Nenets Tundra *xaew°-di* = Forest *kaew°-ti*, Selkup *qöö-tə-ɦ*, Kamas *koo-t* (Klumpp 2022 : 821), Mator †*kajbə-dɜ-* (Helimski 1997 : 261 no. 392).[13] The suffix **ɯə* in **kəy-ɯə* is rather exceptional (though cf. Nenets *xú-w°* = *ku-w°* 'driftwood', as discussed above), and we do not know when it appeared in the word. Assuming that it was there already before the developments **ü* > **i*, **i* > **ə̈* > **ə*, and **l* > *y*, we would have had Pre-Proto-Samoyedic **kül-ɯə* (or **kilɯə*) vs. Pre-Finnic **külkə* (or **kil-kə*). It is, however, unlikely that the suffixes **kə* and **ɯə* are cognates, for this correspondence has no parallels. The most likely explanation is, then, that the forms attested in Finnic and Samoyedic are separate derivatives from the Uralic root **kül/ə* (or **kil/ə*). It may be added that the locative coaffix **-kə-* in Samoyedic nominal declension is very probably based on this same spatially used noun (so also Honti 2022 : 219—225, unnecessarily rejected in Holopainen 2023b : 242—243), in which case its source may have been the basic underived root + **kəy*, rather than the more complex secondary derivative.

Compared with the above items, the Samoyedic cognates of **pel-kä* 'thumb' and **tulka* 'feather, wing' are clearly of a different type in that they contain unmistakable traces of vowel sequences. The question is only what these sequences exactly were, and how they relate to the structure of the corresponding Finno-Ugric items. In the following it will be assumed that they should be viewed in relation to the Finno-Ugric basic stems **pexl/ə* 'side' and **tuxl/ə* 'wind'. At the level of Proto-Uralic these stems stand in a clear contrast with their minimal pairs **pel/ə-* 'to fear' and **tul/ə* 'fire', respectively. This difference is here explained as being due to the presence of the segment **x* in them.

[13] From Forest Enets, the form *ke* 'сторона' is also listed (ES 167), but rather than a direct reflex of the basic root **kəy* it is probably a false abstraction from inflected forms like (locative) *keo-xon* > *ke-xon* 'beside' (ES 170) < **kəy-ɯə-kə-nå*. The same is true of Selkup *qöö*, in which the rounded vowel *öö* probably contains a trace of the labial glide in the derived stem **kəy-ɯə*. Note also that the representation of **kəy-(ə)ɯ* as **kəy-ü* in Nganasan is idiosyncratic and unique, yet the connection with the form **kəy-ɯə* is obvious and cannot be explained without assuming a direct correspondence between these two stem variants.

- **pelkä* 'thumb' : **peeli* < **pexl/ə* '(outer) side'. The Samoyedic cognate here would seem to be Early Proto-Samoyedic **peəy* > Late Proto-Samoyedic **piəy* 'outside space' (SW 124, where the reconstruction is mistakenly given as **piə̑* = **piə*) : (dative) **piəy-ntə-ŋ* : (locative) **piəy-kə-nä*, which yields Enets Tundra *pio* : (dative) *pio-do'* : (locative) *pio-xone* (EH ms.) = Forest *pie* > *pe* : (dative) *pie-do* > *pe-d* : (locative): *pe-xon* (ES 330—331), Nenets Tundra *pyí* : (dative) *pyí-n°'* : (locative) *pyí-x°nya* ~ (secondary) *pyí-x°na* = Forest *pyi* : (dative) *pyi-n°*, Selkup *poo-* : (locative) *poo-qən* 'outside' : (adverb) *poo-nä* 'out' (SkWb no. 376), Mator (dative) <*индамъ*> †*xi-ndəŋ* : (locative) <*едынъ*> †*xe-gən* : (Helimski 1997 : 242 no. 276 and 253 no. 344). The simple root **peəy* seems also to be the source of Nganasan *xïaje* 'thumb' (SNg 208), which reflects the generalized vowel stem **peəyə* (SW 123, where the reconstruction is mistakenly given as **pi-*). Elsewhere in Samoyedic, 'thumb' is expressed by derivatives, which are different in each major language: Enets Tundra *pii-cu* (EH ms.) = Forest *pi-cu* ~ *pi-ju* (ES 337—338), Nenets Tundra *pyí-k°cya* = Forest *pyi-'k°sya*, Kamas *pii-di*, Mator <*егебтu*> †*xe-gəbti* (Helimski 1997: 242 no. 275). As may be seen, Early Proto-Samoyedic **peəy* corresponds segment by segment to the reconstructed Finno-Ugric stem **pexl/ə*, with the regular development of syllable-final **l* to **y* and with the vocalisation of preconsonantal **x* to **ə*. The modern reflexes show, however, a number of irregularities. The *oo* in Selkup *poo* points to an earlier vowel sequence, and this sequence can unambiguously be identified as **eə* > **iə* on the basis of Enets *pio-* = *pie-* > *pe-* 'outside', but its representation as *ii* > *i* in Enets 'thumb' is unexpected. The synchronic sequence *ïa* in Nganasan *xïaje* 'thumb' is unique and remains without a definitive explanation (Kaheinen 2023 : 174). Since, however, *a* in non-initial syllables is a secondary vowel of a complex origin, it can here hardly be anything else but an aberrant reflex of **ə* (on which see further below). Theoretically, it might even be the regular representation of **ə* in the sequence **eə* before a syllable-final **y*, for, although there are cases of the synchronic sequence *ïe* in Nganasan, they seem to be due to the secondary loss of intervocalic **y*, as in *sïe-* 'to drown, to sink' = Tundra Nenets *seyə-* < **sëyə-* (SW 138, an example of the sporadic representation of **ë* as *ï* in Nganasan, cf. Kaheinen 2023 : 59—60).
- **tulka* 'feather, wing' : **tuuli* < **tuxl/ə* 'wind' ~ 'feather, wing'. Assuming that the development of this item was similar to that observed in **pexl/ə*, the Proto-Samoyedic reconstruction would have to be **tuəy* (SW 166). Synchronic forms that are compatible with this reconstruction include Enets Forest *to* (ES 438), Nenets Tundra *to* : (accusative plural) *tú* = Forest *to* : *tu*, and Selkup *tuu* (SkWb 136—137 no. 940). In particular, the Nenets form *to* may be explained as being a contraction of earlier **tuo*, a regular reflex of **tuəy* (as in *ya* : *yo* < **yåə* : **yåə-y*, as mentioned above). Several other languages show, however, forms that do not quite fit the pattern of **pexl/ə*: Nganasan *cüe* (*cue*) 'wing' : *cüe-ru* 'fin' (SNg 217),[14] Tundra Enets *tua* ~ *tue* (EH ms.), Mator (Karagas) <*тóга*> ~

[14] Note that the distinction between the vowels *ü* and *u* is neutralized in Nganasan in the position following the palatal stops *c* [c] and *j* [ɟ], orthographically Cyrillic <*чу*> and <*дю*>. In spite of this, for systemic reasons, the notations *ü* and *üe* will be used here.

(Mator proper) <ry>, which may stand for †*tua* ~ †*tu* (Helimski 1997 : 367 no. 1070). While these do not necessarily allow a uniform Proto-Samoyedic original to be reconstructed, it is also clear that they do not unambiguously point to a segment other than *ə as the second component of the vowel sequence. Nganasan has synchronically the sequence *üe* (*ue*) < *uə, while Tundra Enets has both *ua* and *ue*, and Mator has both *ua* and *u*. Apart from the inherently problematic data from Mator (and its tribal variants), the only item that would seem to point to *a* < *å as the second component of the sequence is Tundra Enets *tua*, which, however, is also attested as *tue*. There is, therefore, no obvious single alternative to the reconstruction **tuəy*. Even so, the issue requires a more comprehensive survey of the etymological material showing similar or related features.

Starting with Nganasan, which is the language that most consistently preserves traces of vowel sequences, and excluding the standard cases in which vowel sequences contain synchronic *e* < *ə as the second component, and focusing on the cases in which the vowel sequence may be assumed to have been followed by a syllable-final palatal glide **y*, we may distinguish between three types of representation: (1) the type in which the glide has been moved to a prevocalic position due to the generalization of the vowel stem of the word, as in the case of *xïaje* 'thumb', (2) the type in which there is no segmental trace of the glide, as in the case of *cüe* 'wing', as well as (3) the type in which a synchronic syllable-final glide is still present in Nganasan. Of relevance in this connection is also (4) the type which does not point to a syllable-final glide in any language, but which in Tundra Enets shows a vowel sequence with synchronic *a* as the second component.

Type (1) is exemplified by four additional etyma, all of which exhibit in Nganasan vowel sequences with *ə as the second component, followed by the syllable *je* [ɟə], which apparently contains the original root-final palatal glide **y* and the vowel *ə, generalized from the vowel stem of the word (Salminen 2024 : 217). For some of these items, a similar generalization of the vowel stem is also observed in Enets:[15]

- Nganasan *ŋieje* (SNg 133) ~ *nieje* [ñ-] (Castrén) 'belt', Enets Tundra *ñiojo* (EH ms.) = Forest (Castrén) *ñiejo* > *ñejo* > *ñej* (ES 266), Nenets Tundra *nyí* = Forest *nyi*, Selkup *cüü* (SkWb 208 no.1433), Kamas *cii*, Mator †*ni* or †*ñi* (Helimski 1997 : 319 no. 750). In this case, the generalized vowel stem is also present in Enets, which confirms the presence of the glide **y* in the word. Most of the data could be explained by assuming Proto-Samoyedic **iəy* : **iəyə*(-) (instead of **ni* or **niə* in SW 102) > Kamas **yiəy*, but the rounded vowel *üü* in Selkup could correlate with the Nganasan initial *ŋi*, suggesting Proto-Samoyedic **üəy* > Selkup **yüəy*. The prothetic nasal is regular in Nganasan, Enets, and Nenets,

[15] The generalization of the vowel stem to the status of the basic form of nouns is a phenomenon relatively frequently observed in the Samoyedic etymological corpus, and the number of affected languages varies. For instance, in the Uralic etymon **suŋ/ə* 'summer' (UEW 451—452) > Proto-Samoyedic **təŋ* (SW 148), only Nenets Forest *tăŋ* = Tundra *tə'* and Enets *to'* > *to* (ES 438—439) preserve the consonant stem, while Nganasan *teŋe* (SNg 187), Selkup *taŋə* (SkWb 178 no. 1211), Kamas *taŋa*, and Mator *taŋa* (Helimski 1997 : 352 no. 960) are based on the generalized vowel stem.

but irregular in Mator, while the prothetic **y* > *c* in Selkup and Kamas is likewise irregular but might reflect a shared Selkup-Kamas innovation. The word has well-known potential cognates in Finno-Ugric, including Finnish *vyö* (SSA 3 : 477), Saami *avvi* (YSS no. 27), and Hungarian *öv* (MSzFE 520—521), but the correspondences are chaotic (cf. also Aikio 2002 : 53 note 5), one problem being that Uralic **ü* would normally yield Samoyedic **i* and not **ü* (cf. most recently Zhivlov 2023 : 170, where the conventional Uralic reconstruction **üwä*, based on Saami, is supported; note that this reconstruction does not account for the Samoyedic final **y*, which is quite certainly present in the word). In the etymological literature (e.g. UEW 575), there is also a frequent confusion with Samoyedic **wenä* > **winä* 'strap' (SW 175), which is a different etymon.

- Nganasan *ciije* 'sleeve' (SNg 214), Enets Tundra <*tiojo*> †*ciojo* (EH ms.) = Forest (Castrén) †*ciejo* > *cej* (ES 475—476), Nenets Tundra *tyú* = Forest *tyu*, Selkup *tüü-naŋ* ~ *tü-naŋ* (SkWb 169 no. 1129), Kamas *tü* (SW 20). In this item, which features the generalized vowel stem in both Nganasan and Enets, the vowel sequence **üə* > **iə* has been homogenized to *ii* in Nganasan, apparently because of the palatal initial, which itself is due to the impact of the following *i* < **i* & **ü* (for several similar cases cf. SNg 214—215). The Selkup form *tüü-naŋ* ~ *tü-naŋ* may be analysed as a compound with genitival linking: **tüü-n+aŋ* 'mouth = opening of the sleeve', with *aŋ* < **aŋ* 'mouth' (SW 20, cf. Finnish *hiha-n-suu* 'mouth of the sleeve'). The Proto-Samoyedic reconstruction is unambiguously **tüəy* (SW 167, so also Kaheinen 2023 : 95; Salminen 2024 : 176). Tundra Enets <*tiojo*> (quoted from Castrén) is unexpected and must be an error for †*ciojo*, for the sequence *ti* would normally presuppose original **tï*, as in *tira* 'fist' < **tïrå* (SW 160). The word has conventionally been compared with distantly similar Finno-Ugric items, including Hungarian *ujj* (MSzFE 652—653) and Saami *soadji* (YSS no. 1154), which may presuppose an original form of the type **soya*. Needless to say, the comparison is wrong.
- Nganasan *mueje* 'branch (of tree), twig' (SNg 100), Enets Tundra *moo* (EH ms.) = Forest *moe* ~ *moj* : (diminutive) *moj-ku* (ES 239—240), Nenets Tundra *mo* = Forest *mo*, Selkup *moo* (SkWb 103 no. 728), Kamas *moo* ~ *mu* (also 'arrow'), Mator <*moh*> : <*moho*-> †*moo* ~ <*muhu*> †*muu* 'arrow' (Helimski 1997 : 307 no. 683 and 309 no. 697). The Proto-Samoyedic form of this item may fairly confidently be reconstructed as **moəy* (instead of **mo* or **moə* in SW 95). A minor issue is connected with Forest Enets *moj*, Cyrillic <*моӥ*>, which could theoretically be a primary form based on the generalized vowel stem, as in the items *ñej* (< **iəyə*) and *cej* (< **tüəyə*) above, but which more likely is a late secondary variant of *moe* (> **moi* > *moj*). Formally it could perhaps also be a derivative in **-yə*. In the latter case, one could speculate that Nganasan *mueje* could also be a similar derivative, and that the basic root is, after all, just **moə* (Kaheinen 2023 : 94). This is, however, unlikely in view of the parallels offered by the other items of this same type, and also since no similar derivative is known in this case from Tundra Enets.
- Nganasan *sieje* 'tongue, language' (SNg 154), Enets Tundra *çioro* (Castrén) ~ *çi'oro* (EH ms., apparently with an inetymological intervocalic hiatus-

filling glottal stop) = Forest (Castrén) *çioro*, Nenets Forest *sye*, Selkup *çee* (SkWb 346 no. 2420), Kamas *çe-kə* ~ *çi-kä*, and Mator †*kej* ~ †*ki* : (3rd person singular possessive) <*kischtä*> †*kis-tä* ~ <*kauıte*> †*käs-te* (Helimski 1997 : 217 no. 450).[16] Most of these data suggest the Late Proto-Samoyedic form **keəy*, which means that the Early Proto-Samoyedic form underlying Nganasan *sieje* must have been **käəy* : **käəyə*(-), although this is not immediately visible from the Nganasan synchronic form. More importantly, the Enets form *çioro* must represent the generalized vowel stem of the Proto-Samoyedic alternation pattern **käəy* : **käələ-* > **keəy* : **keələ-*, which shows that the root originally ended in the lateral **l* (Salminen 2024 : 217). This root itself has well-known cognates in Finno-Ugric, including Finnish *kieli* (SSA 1 : 353). Because of the presence of both a vowel sequence and the consonant **l* > **y* in the Samoyedic cognates, the Proto-Uralic reconstruction cannot be simply **käli-* (Zhivlov 2023 : 166), but is more likely **käxl/ə*, from which Finnic gets the long vowel either directly (> **kääli* > **keeli*) or indirectly (> **käli* > **keeli*).

Type (2) involves two additional etyma containing the Nganasan sequence *üe*, which, unlike *ue* from **oə* in *mueje* < **moəyə*, presupposes original **uə*. It is important to note that the frontening of **u* to *ü*, as also the palatalization of initial **t* to *c* before *ü* in *cüe* (followed by the neutralization of the distinction between *ü* and *u* after the palatal stop *c*), are processes that do not point to the presence of a syllable-final **y*, for they have taken place also in items with no palatal glide, as in Nganasan *cü'* (*cu'*) 'reindeer fat' (SNg 217) < **tut* (in SW mistakenly reconstructed as **tujt* = **tuyt*, so also Kaheinen 2023 : 41, but corrected in Helimski 1997 : 372 no. 1106).

- *küe* 'birch' (SNg 77), Enets Tundra *kua* ~ *kue* (EH ms.) = Forest *ko* ~ *koe* (ES 175, 197), Nenets Tundra *xo* = Forest *ko*, Selkup *qwë* ~ *qwää* (SkWb 257—258 no. 1772), Kamas *koju* ~ *kuju* ~ *küjü*, Mator <*kuga*> †*kua* ~ *kuo* ~ *ku* (Helimski 1997 : 289 no. 565). Assuming that this item follows the pattern of *cüe*, we may reconstruct it tentatively as **kuəy* (instead of **koə̑j* = **koəy* in SW 73). It has to be stressed that the Nganasan form shows no trace of a palatal glide, as is also suggested by declensional patterns, cf., for instance, *küe* : (3rd person possessive form) *küe-du* vs. *küj* 'spoon' : *kü-cü* (SNg l.c.) < **kuy* : **kuy-tå* (SW 76). Even so, there remains the possibility that the sequence *üe* was once followed by a palatal glide, which was lost in this specific context. Among the cognates of *küe* in the other Samoyedic languages, Kamas *koju* ~ *kuju* ~ *küjü*, if segmentable as *koj-u* ~ *kuj-u* ~ *küj-ü*, may actually preserve the root-final **y* [*j*], while Selkup *qwää* and the Enets variants, Tundra *kue* = Forest *koe*, might also point to the former presence of a palatal element in the word structure. In Selkup, however, the devel-

[16] Note that the initial *k* in Mator *kej* etc. cannot represent a retention of the original **k* but, instead, stands for an earlier **s* [ɛ], which itself in this case goes back to **k* before a front vowel (as in the other Samoyedic languages), while the medial *s* [ɛ] is a regular reflex of **y* before **t* (cf. Helimski 1997 : 79—80, 89). Thus, it has to be assumed that Mator initially had the development of **k* to **s* before front vowels and then the opposite (and typologically rare) development of **s* to **k* in a similar context. The latter development affected also the original (dental) sibilant **s*.

opment of the vowel quality is also connected with the origination of the labiovelar initial *qw*, whose most typical source is **kå* > *qwë*, which is also attested in this item as a variant form. In the Uralic context, the Samoyedic items for 'birch' have been compared with Finnish *koivu* (SSA 1 : 386). Since, however, the sequence **yw* is regularly preserved in Samoyedic as **yw*, as in **kaywa*(-) 'to dig' > Samoyedic **kaywå* 'spade' (Aikio 2002 : 41—42), the comparison is acceptable only if we assume that the Finnish (Finnic) item (with a cognate in Mordvin and several more distant lookalikes elsewhere in Finno-Ugric) is based on a derivative of the type **koy-wa* (UEW 169—170), of which the root **koy-* can be identified with Samoyedic **kuəy* only if we postulate a Proto-Uralic form of the type **kuxy/ə* or **koxy/ə* (Janhunen 1981 : 23 no. 84). There still remains an unexplained difference in the root vowel, but if correct, the Uralic comparison could serve as external support for the reconstruction of a palatal glide in the Samoyedic item.

- Nganasan *xüe* 'year, age', Enets Tundra *pua* ~ *pue* ~ *pe* (EH ms.) = Forest *po* ~ *pea* (ES 344), Nenets Tundra *po* = Forest *po*, Selkup *po* ~ *poo* (SkWb 58 no. 371), Kamas *pje* ~ *-pi*, Mator *xa* ~ *xaa* (Helimski 1997 : 239 no. 252). Assuming that this item goes back to **puəy* (instead of **poə̑j* = **poəy* in SW 127), we can see that a possible indirect trace of the presumed root-final palatal glide is present in the Tundra Enets variant *pue* (cf. *kue* above), as well as in Kamas *pje* ~ *-pi*. As in other similar cases, the postconsonantal element *j* [j] (synchronically probably to be understood as a manifestation of consonant palatalization) in Kamas *pje* may indicate the former presence of a syllable-final **y* in the word structure (Klumpp 2022 : 822—823).

Type (3) comprises two rather problematic cases which among themselves are different, in that only one of them contains a synchronically "regular" vowel sequence with *e* (< **ə*) as the second component (*üe*), while the other shows an "irregular" sequence with *a* (*üa*), reminiscent of the case of *xïaje*.

- Nganasan (Castrén) †*jüej* 'dam, closure (in a river, for the purposes of fishing)', Enets Tundra *juu* = Forest *juu* (EH ms.), Nenets Tundra *yú* = Forest *ju*, Mator (Karagas) <*dschui*> †*cuj* (Helimski 1997 : 236 no. 234). It should be noted that the Nganasan word is apparently not attested in any modern sources, and the historical recording, written (Castrén) <*jụai*>, does not actually confirm the quality of the second component of the vowel sequence. The vowel may nevertheless with some certainty be identified as *e* on the basis of what seems to be a related derivative, *jüe-lir* 'closure' (SNg 50), which, in turn, would seem to be based on the verb (Castrén) †*jüe-*, from **yuə-* 'to dam, to block (a river)' (SW 48; Kaheinen 2023 : 70). The verb has cognates in Enets Tundra *ju-a-* (EH ms.), Nenets Tundra *yú-* : *yo-* = Forest *jo-*, Selkup *cuu-* (SkWb 207 no. 1430). However, the Nenets variant stem with *o*, which is also attested in derivatives such as *yo-bta-* 'to dam', is unexpected and might go back to **yuo-* < **yuəy-*. Formally, Nganasan †*jüej* seems to be a deverbal derivative of †*jüe-*, formed by the nominalizing suffix **-y*, that is, **yuə-* : **yuə-y*, which would, then, be the origin of all the cognate forms, as well.

- Nganasan *ŋüaj* 'island', Enets Tundra *ñue* (EH ms.) = Forest *nuj* (ES 284), Nenets Tundra *ŋo* = Forest *ŋo*, Selkup *ko* ~ *kuu* (SkWb 247 no. 1743), Mator *o* (Helimski 1997 : 326 no. 793). Except for the Nganasan form with the sequence *üa*, all these items are potentially derivable from **ɯuəy* (instead of **ɯoə̑y* = *ɯoəy* in SW 177). Other reconstructions that have been proposed include **ɯoåj* = **ɯoåy* (Kaheinen 2023 : 70), and **ɯuaj* = **ɯuay* (Salminen 2024 : 180, 217), of which only the latter is congruent with the Nganasan data without, however, going deeper into the origin of the otherwise problematic vowel *a*, while **ɯoə* (Helimski l.c.) leaves unexplained the final **y* that is physically present in the Nganasan cognate. Forest Enets *nuj*, Cyrillic <*ныӱ*>, is probably a secondary variant of *nue*, as attested in Tundra Enets, rather than a separate form (cf. *muj* above).

Type (4), which here serves as a point of comparison for some of the items discussed above, comprises two etyma:

- Nganasan *taa* 'domestic reindeer' (SNg 168), Enets Tundra *tia* (EH ms.) = Forest *te* (ES 464), Nenets Tundra *ti* : *te-* = Forest *ti* : *te-*, Kamas *to*, Mator (Karagas) <*dége*> †*te-ge* ~ (Mator proper) <*tíggä*> †*ti-gä* (Helimski 1997 : 357 no. 1001). While it is obvious that this item involves an original vowel sequence, suggesting an original form like **tëə* (SW 155), the presence of *a* as the second component of the sequence in Nganasan and Tundra Enets has stimulated alternative reconstructions of the type **tëå* (Kaheinen 2023 : 70) and **tëa* (Salminen 2024 : 216). It has to be noted, however, that Nganasan *taa* can also be a regular reflex of **tëə*, for the sequence **ëə* > *ae* would automatically become homogenized to *aa*,[17] leaving only the Tundra Enets data to be explained as potentially irregular.
- Nganasan (Castrén) *ŋoa* ~ *ŋua* 'door' (SNg 136), Enets Tundra *ŋia* ~ *ñia* (EH ms.) = Forest *no* (ES 273), Nenets Tundra *nyo* = Forest *nyo*, Selkup *-a* (SkWb 109 no. 742), Kamas *aa-je*, Mator (Karagas) *no* ~ (Mator proper) *ño* (Helimski 1997: 321 no. 765). In this case Nganasan shows an unambiguous *a* as the second component of the vowel sequence, which has again stimulated reconstructions of the type **öå* (Kaheinen l.c.) or **öa* (Salminen, l.c.) instead of **öə* (SW 29). For this item there is a potential Finno-Ugric cognate in Finnish *ovi* : *ove-* 'door' (SSA 2 : 277; UEW 344), but the correspondences, also within Finno-Ugric, are too irregular to allow the etymology to be confirmed or any unambiguous Uralic reconstruction to be proposed.

Gathering the data discussed above, Table 2 below shows the correspondences of the four types of Nganasan (Ng) items (1—4) in the other Samoyedic languages, accompanied by preliminary reconstructions. It is tentatively assumed here that all the vowel sequences in these items contain the reduced vowel **ə* as the second component, but whether this is so or not has to be decided on the basis of the actual correspondences.

[17] Note that Proto-Samoyedic **lë* 'bone' does not contain a vowel sequence, as is obvious from its derivative *la-tee* 'bone' (SNg 87) = Enets Tundra *lizi* = Forest *lizi*, Nenets Tundra *ledi* = Forest (Western) *riti* 'spine' < **lë-tə-yə* (SW 82, so also Kaheinen 2023 : 88). Nganasan *laa* 'ring', on the basis of which a reconstruction of the type **lëə* has been suggested (Salminen 2024 : 216), must also be a derivative, i.e., **lë-yə*.

Table 2

Types of correspondences involving Proto-Samoyedic vowel sequences

		Ng	ET	EF	NT	NF	Sk	Km	Mt
1	**peəy*	*xïaje*	*pio*	*pie* *pe*	*pyí*	*pyi*	*poo*	*pii-*	*xi-* *xe-*
	**y/iəy* **y/üəy*	*ŋieje* *nieje*	*ñiojo*	*ñiejo* *ñejo* *ñej*	*nyí*	*nyi*	*cüü*	*cii*	*ni* *ñi*
	**tüəy*	*ciije*	*ciojo*	*ciejo* *cej*	*tyú*	*tyu*	*tüü-* *tü-*	*tü*	
	**moəy*	*mueje*	*moo*	*moe* *moj*	*mo*	*mo*	*moo*	*moo* *mu*	*moo* *muu*
	**käəy*	*sieje*	*çioro*	*çioro*		*sye*	*çee*	*çe-* *çi-*	*kej* *kis-* *käs-* *ki*
2	**tuəy*	*cüe*	*tua* *tue*	*to*	*to*	*to*	*tuu*		*tua* *tu*
	**kuəy*	*küe*	*kua* *kue*	*ko* *koe*	*xo*	*ko*	*qwë* *qwää*	*koj-* *kuj-* *küj-*	*kua* *kuo* *ku*
	**puəy*	*xüe*	*pua*	*po* *pea*	*po*	*po*	*poo* *po*	*pje*	*xaa* *xa*
3	**yuəy*	*jüej*	*juu-*	*juu-*	*yú* *yo-*	*ju* *jo-*	*cuu-*		*cuj*
	**wuəy*	*ŋüaj*	*ñue*	*nuj*	*ŋo*	*ŋo*	*kuu* *ko-*		*o*
4	**tëə*	*taa*	*tia*	*te*	*ti* *te-*	*ti* *te-*		*to*	*te-* *ti-*
	**öə*	*ŋoa* *ŋua*	*ŋia* *ñia*	*no*	*nyo*	*nyo*	*-a*	*aa-*	*no* *ño*

As may be seen, the two Nenets languages, Tundra (TN) and Forest (FN), are mutually congruent and show the middle vowel *o* for both presumed **oəy* (*mo*) and presumed **uəy* (*to*, *xo*, *po*, *ŋo*), with the exception that **yuəy* yields both *yo-* and *yu(-)*, which could be connected with the impact of the palatal initial **y*. Similarly, the sequence **äəy* > **eəy* yields the middle vowel *e* (*sye*).[18] The other front-vocalic sequences **eəy* > **iəy* and **iəy* have merged and yield the high stretched unrounded vowel **í* combined with the palatalization of the preceding consonant (*pyí* = *pyi*, *nyí* = *nyi*), while **üəy* yields the high stretched rounded vowel **ú*, also with palatalization (*tyú* = *tyu*). The high stretched vowels themselves indicate a complex origin of the segment, though they are ambiguous as to whether they reflect a trace of a vowel sequence or the former presence of the glide **y* in the root structure, or both (Salminen 1993 : 348—349).

[18] According to Salminen (2007 : 367) Forest Nenets vowels underwent a secondary process of "monosyllabic shortening", which is why an item like *sye* is actually represented as *syĕ* in the language. This recurrent detail is not incorporated in the notation here.

Selkup (Sk) exhibits a somewhat less systematic picture, and there is both qualitative and quantitative variation in the dialectal data. The dominant representation is, however, a long middle vowel for the Late Proto-Samoyedic sequences *eəy (çee) and *oəy (moo) and a long high vowel for the sequences *uəy (tuu, cuu, kuu), *üəy (tüü-) and *y/iəy ~ *y/üəy (cüü). However, both *peəy > *piəy and *puəy yield poo, apparently due to the impact of the labial initial *p. The only item that might point to the former presence of a palatal element in the root structure is qwää, if derived from *kuəy.

Kamas (Km) has a few lacunae in the data and shows also random variation in both quantity and quality, but offers a picture similar to Selkup for the items with *äəy > *eəy (çe- ~ çi-), *oəy (moo), and *üəy (tü), and is also congruent with Nenets in the case of *eəy > *iəy (pii-). For the item *y/iəy ~ *y/üəy, Kamas suggests *yiəy (cii), with the vowel congruent with Nenets and with the consonant congruent with Selkup (as mentioned above). A physical trace of *y may be preserved in koj- ~ kuj- ~ küj-, if this is a direct reflex of *kuəy-. Also, pje, if from *puəy, points to a palatal element in the original.

Mator (Mt) offers a more challenging picture. Mator has preserved the syllable-final glide *y quite well, also in monosyllabic roots (Helimski 1997 : 90—91). In this respect, the Mator reflexes of *käəy > *keəy (kej : kes-) and *yuəy (cuj) are exactly what can be expected. The reflexes of *peəy > *piəy (xe- ~ xi-) do not necessarily contradict this, since they only concern medial position. The lack of any trace of *y in all the other items, including *muəy (moo ~ muu) and *iəy (ni ~ ñi), is, however, surprising, a circumstance which has stimulated reconstructions without a final *y in these cases. Typically, the data show doublets of "short" forms with a single vowel (tu, ku, xa, o) and "long" forms ending in a or o (tua, kua ~ kuo, xaa), but most of the "long" forms come from older, less reliable, and dialectally more diffuse sources, which reduces their informative value.

The two Enets languages, Tundra (TE) and Forest (FE), show much fewer mutual similarities than the two Nenets languages. Both Enets languages agree on the form *keələ (çioro), which corresponds to *käəyə in Nganasan and *käəy > *keəy in the other Samoyedic languages, as well as on the reflexes of *iəy : *iəyə (ñiojo = ñiejo) and *tüəy : *tüəyə (†ciojo = ciejo) of type 1. On the reflex of *yuəy (juu) of type 3, both Enets languages agree with Nenets. For *moəy (moe ~ moj), *kuəy (koe), and *ɯuəy (nuj), as well as for *peəy > *piəy (pie ~ pe), Forest Enets suggests a palatal final of the type e or j, but there are also forms following the pattern of Nenets for the items of type 2 (to, ko, po). A similar final e is present in the Tundra Enets reflexes of *uəy (tue, kue, ñue), paralleled by variants with a final a (tua, kua, pua), while the items *peəy > *piəy and *moəy have o (pio, moo). In Forest Enets a final a is attested in a variant form of *puəy (pea), provided that this form really represents Forest Enets. Finally, Tundra Enets differs from Forest Enets on the point that the items *tëə and *öə contain a vowel sequence ending in a (tia, ŋia ~ ñia), while Forest Enets has simple monosyllables (te, no) of the same type as Nenets. Altogether, the picture offered by Enets appears internally rather incoherent, which may indicate some mixing between the two Enets languages, or a confusion in the sources, but which is also conditioned by the chronological heterogeneity of the database. It is, in any case, obvious that for several items Tundra Enets

shows a clear parallelism with Nganasan, while Forest Enets more often follows the pattern exhibited by (both varieties of) Nenets.

Since we are looking for information that would allow us, on the one hand, to confirm the presence or absence of a final palatal glide **y* in the items under discussion, and, on the other hand, to identify the types of vowel sequences originally present in them, we may now summarize the evidence for both of these points. As far as a physical glide segment is concerned, it is present in all items of types 1 and 3 in Nganasan, in the reflexes of **iəy* and **tüəy* of type 1 in Enets, as well as in the reflexes of **käəy* > **keəy* of type 1 and **yuəy* of type 3 in Mator (*kej*, *cuj*), possibly also in the reflex of **kuəy* of type 2 in Kamas (*koj-* ~ *kuj-* ~ *küj-*). An indirect trace of **y* may be present in the reflexes of **kuəy* in Selkup (*qwää*) and **puəy* in Kamas (*pje*), both of type 2, as well as in the reflexes of **tuəy* and **kuəy* of type 2 (*tue*, *kue*) and **wuəy* of type 3 (*ñue*) in Tundra Enets, which have analogies in the reflexes of **peəy* > **piəy* and **moəy* of type 1 and **kuəy* of type 2 in Forest Enets (*pie*, *moe*, *koe*). A secondary asyllabic pendant of this representation seems to be present in the Forest Enets reflexes of **moəy* of type 1 (*moj*) and **wuəy* of type 3 (*nuj*), which may, however, also be compared with the maximally "worn" Forest Enets reflexes of **tüəyə* (*cej*) and **iəyə* (*ñej*). When reviewed in this way, it turns out that for each of the items of types 1 to 3 there is at least one language that preserves either a direct or an indirect trace of a palatal element, presumably **y*, in the root composition.

We are, of course, most concerned with the items of type 2, which show no physical glide segment in Nganasan, and for which no final glide is registered in Mator. However, each of the items of this type has possible traces of a palatal segment in one or several languages: for **tuəy* in Tundra Enets (*tue*), for **kuəy* in both Tundra (*kue*) and Forest Enets (*koe*), Selkup (*qwää*), and Kamas (*koj-* ~ *kuj-* ~ *küj-*), and for **puəy* in Kamas (*pje*). The fact that none of these items shows a segmental glide in Mator is potentially significant, but it has to be noted that the Mator items *ni* ~ *ñi*, *moo* ~ *muu*, and *o* also lack any trace of the glide that is physically present in their Nganasan counterparts *ŋieje* ~ *nieje*, *mueje*, and *ŋüaj*. If we assume that these words are full cognates, we have to conclude that the glide must have been present in the Proto-Samoyedic forms, as well.

Assuming, then, that all of the items of types 1 to 3 had originally a root-final palatal glide **y*, which was only secondarily lost in the Nganasan reflexes of the items of type 2, there remains the issue of the preceding vowel sequences and their divergent representations in the modern languages. This concerns also the items of type 4, in which no palatal glide was ever present. Primary vowel sequences are synchronically attested only in Nganasan and Tundra Enets, with a few, possibly residual or contact-induced analogies also in Forest Enets. For this reason, the Mator items suggesting vowel sequences ending in *a* or *o* in the items of type 2 (*tua*, *kua* ~ *kuo*, *xaa*) are hardly of any relevance, especially since they are in less obscure sources attested as simple monosyllables (*tu*, *ku*, *xa*). It may nevertheless be noted that exactly these same items appear in Tundra Enets with vowel sequences ending in *a* (*tua*, *kua*, *pua*), though they also have variants in *e* (*tue*, *kue*). The problem is that, although the vowels represented as *a* in Mator and Enets would be regular counterparts (from **å*), both Tundra Enets *e* (from

ä) and Nganasan *e* (from **ə*) in the cognates of these same items require a different reconstruction. Since, consequently, the vowels do not match, the correspondences are by definition irregular, and at least some of the synchronically attested forms must be results of secondary developments.

Tundra Enets has vowel sequences ending in *a* also in the two items of type 4 (*tia*, *ŋia ~ ñia*). In these cases Nganasan has analogous sequences ending in *a*, though only one of the items has diagnostic value (*ŋoa ~ ŋua*), while Mator has simple monosyllables (*te- ~ ti-*, *no ~ ño*). However, there is again a mismatch in the vowels, since Enets *a* (from **å*) does not correspond to Nganasan *a* (from **a*). Nganasan has *a* also in the reflexes of **peəy* (*xïaje*) and **wuəy* (*ŋüaj*), but in these items Tundra Enets has either *o* (*pio*) or *e* (*ñue*). This means that the two languages do not show congruent correspondences, which is why it is reasonable to assume that the cases with a synchronic [*a*] as the second component of a vowel sequence are somehow secondary. As far as Tundra Enets is concerned, the situation appears to be rather straightforward, for all the cases with *a* involve items with a high root vowel, which is either *u* (*tua*, *kua*, *pua*) or *i* (*tia*, *ŋia ~ ñia*). The fact that in these same contexts both *o* (*pio*) and *e* (*tue*, *kue*, *ñue*) are also attested, suggests that there has been an uncompleted tendency in Tundra Enets to polarize these vowel sequences by lowering the second component to [*a*]. Whether this is also true of the Nganasan items concerned (*xïaje*, *ŋüaj*, *ŋoa ~ ŋua*) is more difficult to tell. However, any reconstructions of the type **öå* (Kaheinen 2023 : 63) or **öa* (Salminen 2024 : 216) are bound to be inconclusive, since they can at most account for the synchronic representation in one language at a time.

The vowel sequences in Nganasan *xïaje* and *ŋüaj* call for special attention, as almost all the other items of types 1 (*ŋieje ~ nieje*, *mueje*, *sieje*) and 3 (*jüej*), as well as those of type 2 (*cüe*, *küe*, *xüe*), with the single exception of the reflex of **tüəy* : **tüəyə* (*ciije*) of type 1, which has a secondary monophthong, contain sequences ending in Nganasan *e* [ə], as could be expected if Proto-Samoyedic really had only sequences with the generic reduced vowel (**ə*) as the second component. The vowel *a*, at least in positions other than the initial syllable of a word, is (as already pointed out above) a diachronically secondary segment, which represents an innovation conditioned by a sequence of so far not fully understood processes of suffixation, as is obvious, in particular, from its role in the morphophonology of Nganasan paradigmatic root-final alternations. Evidence for lexical items ending in *a* are few and controversial: for istance, Nganasan *basa* 'iron' (SNg 20) suggests Proto-Samoyedic **wasa* or **wäsa* (Kaheinen 2023 : 39), while its cognates in all the other Samoyedic languages, including Mator †*bese ~ †base* 'iron' (Helimski 1997 : 216—217 no. 115), suggest **wäsä* > **wesä* (SW 175), which is congruent with its cognates in West Uralic, including Finnish *vaski* : *vaske-* < **wäçkä* (SSA 3 : 416), meaning that the Nganasan form may well involve an innovation — perhaps a generalization of a secondary stem.[19] From the diachronic point of view it is even more difficult to motivate the presence of *a* in the Nganasan items *xïaje* and *ŋüaj*, for

[19] Aikio (Luobbál Sámmol Sámmol Ánte (Aikio) 2015 : 42—43) argues that the reconstruction of **a* (and not **ä*) in the second syllable of this word is also supported by the initial-syllable vowel in Selkup *kwësə* (SkWb 318 no. 2185) and the second-syllable vowel in Kamas [*båzå*] *baza* (KWb 8), but in the absence of exact

the processes responsible for *a* in root-final position can hardly have been active root-internally. The only possible conclusion is that the sequences *ïa* and *üa* in *xïaje* and *ŋüaj* represent secondary and probably irregular developments of **ïə* (< **eə*) and **üə* (< **uə*), respectively. In the case of *xïaje* the development may have been contextually conditioned, but in the case of *ŋüaj* this is less likely in view of *jüej*, which, as it seems, has not undergone a similar irregular development (assuming that this item is correctly quoted in the sources, which may not be taken as fully confirmed).

It is more difficult to explain why the final palatal glide **y* was lost in Nganasan in the items of type 2, when it was not lost in the two items of type 3. The other Samoyedic languages do not suggest anything abnormal in the cognates of *jüej* and *ŋüaj*. In the case of *jüej* < **yuəy* = **yuə-y* one could think of the possibility that the morpheme boundary and the synchronically preserved connection with the verb *jüe-* < **yuə-* has prevented the loss of the final **y* and also the development **ə* > *a*, but this is not relevant for *ŋüaj*. More likely, both items share a parallel background, which might be connected with the frequently observed variation of consonant and vowel stems in monosyllabic nouns, including roots containing a vowel sequence. Since Nganasan, like also Enets, shows several cases of a vowel stem being generalized to the status of the basic form of a word instead of the phonologically regular consonant stem (*xïaje*, *ŋieje ~ nieje*, *ciije*, *mueje*, *sieje*), it is thinkable that the opposite trend was also active, meaning that the original consonant stem can have been restored on the basis of the generalized vowel stem. If this is so, the forms *jüej* and *ŋüaj* would be secondary consonant stems formed from the corresponding vowel stems *jüeje-* and *ŋüaje-*, respectively. Sychronically, consonant stems in *j* and vowel stems in *je* do not differ in inflected forms requiring a connective or suffix-initial vowel, as in (plural nominative and genitive) *ŋüaj* : *ŋüaj-e'* : *ŋüaj-i'* ~ *ŋüaj-u'* (SNg 141) vs. *xïaje* : *xïaje-'* : *xïaj-i*['] (SNg 208). There is, however, a difference when the suffix contains a syllable added directly to the root, as in (3rd person singular possessive) *ŋüaj* : *ŋüaj-cu* vs. *xïaje* : *xïaje-tï*.

Summarizing the diachronic situation for the items under discussion, it seems likely that there existed originally only one structural type, containing a vowel sequence with the reduced vowel **ə* as the second component and followed by a root-final **y*, which could also represent the lateral **l*. This structure underwent diversification in Nganasan and Enets, in that the final **y* was either regularly deleted in at least three Nganasan items of type 2 (*cüe*, *küe*, *xüe*) and possibly their Tundra Enets counterparts (*tue*, *kue*, **pue*), which later, in Tundra Enets, participated in a tendency to lower the second component to *a* (*tua*, *kua*, *pua*). A different path of development was followed by items for which the vowel stem was generalized as the basic form of the word, which happened before the loss of the final **y* in type 2. This type comprises at least five Nganasan items of type 1 (*xïaje*, *ŋieje ~ nieje*, *ciije*, *mueje*, *sieje*) and the Enets cognates of two of them (*ñiojo* = *ñiejo*, †*ciojo* = *ciejo*), as well as one earlier case of a similar general-

parallels the evidence is inconclusive. The sequence *kwë ~ qwë* appears variously as a reflex of both **wä* > **we* and **wå* (as well as **wə*) in Selkup (SW 168—175), and there are several other examples of second-syllable **ä* being represented as *a* in Kamas, as in **yälä* 'sun, day' > *cala*, **yapä* 'leaf' > *caba*, **yëkcä* 'female (animal)' > *căkta* (~ *căktə*), **pətä* 'bile' > *pada*, and even **səymä* 'eye' > *sima* (SW 40—41, 115, 132).

ization in Enets (*çioro*). The two items of type 3 (*ŋüaj*, *jüej*) also participated in this generalization, but a later development restored the consonant stem for them. Also, in two Nganasan items, the second component of the vowel sequence is not *e* but *a* (*xïaje*, *ŋüaj*). This detail remains without a definitive explanation, but it can hardly represent an original feature inherited from Proto-Samoyedic.[20]

4. Conclusion

The main result of the above argumentation is that the derivation of the Samoyedic items for 'feather, wing' from the form **tulka*, which is the origin of the Finno-Ugric words with the same meaning, is untenable. The idea that the cluster **lk* would have totally disappeared in Samoyedic is premature, as Samoyedic reconstructions like **tuå* or **tua* cannot explain the comparative data. The Proto-Samoyedic form is best reconstructed as **tuəy*, which is an exact reflex of Proto-Uralic **tuɤl/ə* > Finnish *tuuli* : *tuule-* 'wind'. An analogous relationship exists between Samoyedic **peəy* > **piəy* 'outside space; thumb' and Finnish *pieli* : *piele-* 'side' < Proto-Uralic **peɤl/ə*, which also yields the derivative **pel-kä*, as attested in Saami *bealgi* 'thumb'. The relationship between the Samoyedic and Finno-Ugric data can only be explained by assuming the presence of the "primary laryngeal **ɤ* (Janhunen 2007) in the Proto-Uralic source forms. It is the presence of this segment that distinguishes these lexemes from their minimal pairs **tul/ə* 'fire' and **pel/ə*(-) 'to fear; fear(some)'. Although many cases of Finnic long vowels are now known to be secondary and can be explained by a regular sound change with no additional segment (Luobbál Sámmol Sámmol Ánte (Aikio) 2012), this does not apply to the items whose Samoyedic cognates contain vowel sequences, in which the "laryngeal" **ɤ* is typically reflected as the postvocalic reduced vowel segment **ə*.

In Samoyedic **tuəy* and **peəy* > **piəy* the final palatal glide **y* represents the original Uralic lateral **l*, which is preserved on the Finno-Ugric side. The change **l* > **y* seems to have taken place very late in Pre-Proto-Samoyedic, as is evident from the occasional preservation of **l* in individual lexemes in some Samoyedic languages, for istance, in the Enets reflexes of **käəy* > **keəy* 'tongue, language'.[21] This change is observed also in a couple of other items, Proto-Samoyedic **pel-* > **pey-* > **piy-* 'to fear', **kil-/*kül-* > *käy* > **kəy* 'side', and possibly **kul-* > **kuy-* 'to float downriver', whose Finno-Ugric cognates contain items with the cluster **lk*, which, however, derives from

[20] Other examples of the sequence *üa* are *xüa-le-* : *xüa-rkuj-* 'to blow' (SNg 207), as mentioned by Salminen (2012 : 350) and Kaheinen (2023 : 164—165), and <*xɵaɾaü*> †*xüa-gaj* 'knee' (SNg 199), both of which would at first glance seem to require a reconstruction of the type **puə-* without a following glide (instead of **pu-* or **puə̑j-* = **puəy-* in SW 128—129, 130), though the presence of a glide cannot be excluded. The history of 'knee' is particularly complicated and would call for a separate study.

[21] Note that the development **l* > **y* was active also word-initially, and in this position it was likewise not fully completed at the Proto-Samoyedic stage, leaving doublets like **lëpsə* (in Nganasan and Enets) > **yëpsə* 'cradle' (in the other languages) (SW 41) < Proto-Uralic **lïpsə* > Finnish *lapsi* : *lapse-* 'child' (SSA 2 : 48—49). Other items, like **lë* 'bone' (SW 82) = Finnish *luu* (SSA 2 : 114) remained altogether untouched by the development. There are no obvious contextual factors that could explain the different behaviour of **l* in such cases.

the combination of the root-final lateral **l* with a secondary suffix-initial **k*. This means that there is no single etymon with a Proto-Uralic **lk* that would have a full cognate in Samoyedic. This may appear unexpected, but it is probably simply due to the small size of the Proto-Uralic etymological corpus. In principle, we would expect that the cluster **lk* would develop along lines similar to **rk* and **dk* > **rk*, except that the syllable-final **l* would be represented as **y*. We may nevertheless briefly discuss an often-quoted counterexample, which would superficially suggest that the cluster **lk* is preserved in Samoyedic:

- Selkup *ñulqə* 'fir' (SkWb 245 no. 1718) < **ñulkå* (SW 112) > Kamas *nolgo* (KWb 45), possibly also reflected in the Forest Nenets (Western) derivative *nyŭrki* 'cambium' : *nyŭrki pya* 'cambium tree' = 'fir', but potentially confused with other dendronyms, including Nenets Tundra *nyurka* 'aspen' = Forest Nenets *nyŭłka* (SLD 86, apparently a borrowing from Tundra Nenets, Tapani Salminen p.c.), and *nyerka* 'willow' < **ñër-* : **ñër-kå* ~ **nër-kå* (SW 108) > Nganasan *nirku* 'aspen' (SNg 117), Enets Tundra *niga* (EH ms.) = Forest <ныга> *niga* 'willow bush' (ES 285), Selkup *ñarqə* 'willow' (SkWb 242 no. 1694), Kamas †*nerga* 'willow' (KWb 45), Mator †*nerga* 'willow' ~ <нарге> 'fir' for ?'willow' (Helimski 1997 : 314 no. 723 and 319 no. 748). Of these, the items registered as meaning 'fir' have conventionally been compared with similar data on the Finno-Ugric side, reconstructed as **ñulkɜ* 'fir (*Abies*)' (UEW 327), which has also been used as a palaeolinguistic reference for the Uralic homeland (Saarikivi 2022 : 57, with the reconstruction **ñulki*). However, the Finno-Ugric data are mutually incongruent: Mari (Eastern) *nulgo*, Permic (Udmurt = Komi) *ñïl*, Khanty (Vakh-Vasyugan) *ñălkï* (DEWOS 1049), Mansi (Northern) *ñuli* (WWb 374). The fact that the internal consonantism in these items cannot stand for an original Finno-Ugric cluster **lk* is clearly shown by the representation in Mari (*lg* instead *l*), Khanty (*lk* instead of *γl*), and Mansi (*l* instead of *wl* or the like). It may be concluded that all these items are reflexes of secondary areal transmission between the eastern branches of Uralic (so also Kaheinen 2023 : 41, 179).

In this connection we have to reject also a couple of other comparisons that have been made between Finno-Ugric and Samoyedic and involving words that on the Finno-Ugric side contain the cluster **lk*:

- Finno-Ugric **w/olka* 'shoulder' (as discussed earlier). This word has been compared with Samoyedic **wayk* : **wayk-kə* 'neck' (SW 173) > Nganasan *bake-* : *bake-dee* (SNg 19), Enets Tundra *beko* (EH ms.) = Forest *bäk* (ES 65), Nenets Tundra *yík°* = Forest *wye'k°*, Selkup *qwëq* (SkWb 308 no. 2106) ~ *qwët/ə* (SkWb 276 no. 1902) 'shoulder', Kamas *baj'gə*, Mator †*böjkö* (Helimski 1997 : 215 no. 107). Although this would seem to make a case for an item with a Proto-Uralic cluster **lk* represented in Samoyedic as **yk* (Aikio 2002 : 54; Zhivlov 2014 : 139, 2023: 162), the presumed loss of the final vowel is a major irregularity that makes the comparison unacceptable. Moreover, since the development **l* > **y* is relatively late in Samoyedic, it would have to be assumed that the word underwent initially an irregular change in the vocalism, that is, **wolka*

> **wålkå* > **walkå*, then the regular change **l* > *y*, **walkå* > **waykå*, and finally the irregular loss of the final vowel, **waykå* > **wayk*, followed by an unexpected pattern of suffixation or gemination. Although Finno-Ugric **ay* is represented as Samoyedic **ay* in the item **kaywa-* 'to dig' > Samoyedic **kaywå* 'spade', the development **a* > **å* > **a* would hardly have taken place before a syllable-final **l*.[22] Also, the regular representation of Finno-Ugric **oy* in Samoyedic is **åy*, as in **oywa* 'head' > Finnish *oiva* (SSA 2 : 261) = Samoyedic **åywå* (SW 17). All of this means that Samoyedic **waykkə* 'neck' remains without a satisfactory Uralic etymology. An additional complication is that the otherwise unexplained Selkup doublet *qwëq* ~ *qwët/ə* might imply that the original form of the word actually contained the cluster **tk*, i.e. **wayt-kə*.

- Finnic and Saami **ülkä* 'boy' (as discussed earlier). This word has been tentatively compared (UED) with Selkup *ii* 'son' (SkWb 3 no. 13). Apart from the obvious problem connected with the very limited distribution of these items in both Finno-Ugric and Samoyedic, the phonemic correspondence is imperfect. Formally, Selkup *ii* could go back to **iy*, which could theoretically reflect the first syllable of **ül-kä*, supposing that this word is another derivative in **-kA*. However, it is more likely that the Selkup item is connected with the data for 'son, child' in the other Samoyedic languages, including Nganasan *ñüe* (*ñue*) (SNg 124), Enets Tundra *ñio* (EH ms.) = Forest *ñie* ~ *ñe* (ES 266), Nenets Tundra *nyú* = Forest *nyu*, Kamas *ñii*, Mator *ñu* ~ *nu* (Helimski 1997 : 323 no. 779). Although the reconstruction **ñuə* has been favoured for this set (most recently Salminen 2024 : 216), there is a clear parallelism with the items of the type **tuəy* 'wing, feather' (type 2 above), which means that the Proto-Samoyedic form may actually have been **ñuəy*. It is probably relevant to note that Nenets has also a stem variant with *í*, as in Tundra Nenets (accusative plural) *nyí* : (captative verb) *nyí-s-* : *nyí-cy°*, paralleled by Forest Enets *ñi-c* 'to calve (of reindeer)' (ES 271), while Kamas has stem variant with *ee* in the derivatives *ñee-* 'to give birth' and (diminutive) *ñee-kə* 'young (of animals)'. Since **ñuəy* is in a complementary distribution with Selkup *ii*, the likelihood is great that we are actually dealing with a single etymon. If so, Selkup *ii* would have to be an innovative form that has lost the initial palatal nasal which is present in all the other Samoyedic languages. In spite of occasional synchronic variation in the harmonic status of the word, as in Tundra Nenets (3rd person singular possessive) *nyú-dya* (JSWb 330) vs. Nganasan *ñüe-du* (SNg l.c.), the root is originally back-vocalic, which means that a connection with Finnic and Saami **ülkä* can be ruled out.

[22] The representation of Finno-Ugric **a* as Samoyedic **a* is often considered to involve a primary phonemic split, due to which Uralic **a* had from the beginning a dual reflex in Samoyedic: either **å* (regular) or **a* (irregular) (cf. e.g. Aikio 2002 : 50). However, it is more likely that the representation of **a* was initially always **å*, which only later, possibly under certain contextual conditions, but often without any observable reason, was divided between **å* (conservative) and **a* (innovative), of which the latter was a secondary phoneme in Samoyedic. The correspondence of Finno-Ugric **å* to Samoyedic **a* is always a factor that weakens an etymology. Even so, there are enough examples to corroborate the reality of the change **a* > **å* > **a* in Pre-Proto-Samoyedic. Of course, it would also be possible to postulate both **å* and **a* for Proto-Uralic and assume that their distinction was lost in Finno-Ugric and retained only in Samoyedic.

In the present paper the focus has been on items in which Samoyedic vowel sequences occur before an original root-final **l*, represented in Samoyedic as **y*. The second component of the vowel sequence in these cases goes back to a Proto-Uralic "laryngeal", which is one of the sources of long vowels in Finnic, but which is also indirectly reflected in the data from the other branches of Finno-Ugric, as in the minimal pair **tul/ə* 'fire' vs. **tuxl/ə* 'wind; wing, feather'. Vowel sequences were, of course, allowed also before other consonants, including obstruents, as in **wüət* 'ten' (instead of **wüt* in SW) < **wixt/ə* > Finnish *viisi* : *viide-* (SSA 3 : 444—445),[23] and, in particular, root-final vowel sequences could be followed by several types of consonants, as in **yåə* 'earth, land' : (genitive) **yåə-n* : (accusative) **yåə-m* : (nominative plural) **yåə-t* : (accusative plural) **yåə-y*, as well as by entire syllables, as in (3rd person singular possessive) **yåə-tå*. This situation is synchronically still preserved in Nganasan, but with the difference that the system of vowel sequences has been multiplied by a number of new combinations whose origin is mainly connected with the loss of intervocalic **y* in the position after a stressed (odd-numbered) syllable (Kaheinen 2023 : 91—92). In the present paper it has been argued that the phonotactic situation was different in Proto-Samoyedic, where the only segment that could occupy the position of the second component of a vowel sequence was the qualitatively neutralized reduced vowel **ə*, whose phonotactic status resembled in some respects that of the glides **w* and **y*. Therefore, forms of the type **tuå* or **tua*, as have been proposed on the basis of Finno-Ugric comparisons, would have been a systemic impossibility in Proto-Samoyedic.

Proto-Samoyedic was certainly not alone in having this type of vowel sequences. A well-known parallel is offered by English, especially in its British standard pronunciation, in which all vowel qualities can be followed by the reduced vowel [ə] or its allophones, as in *peer* [pʰɪə] */piə/*, *poor* [pʰʊə] */puə/*, *pear* [pʰɛə] */peə/*, *pour* [pʰɔː] */poə/*, *par* [pʰɑː] */paə/*, *purr* [pʰɜː] */pəə/*. In these cases, the second component of the vowel sequence represents a trace of the liquid *r*, which is still present in the morphophonological (deep) form of the lexeme and can appear in the speech if a vowel follows, as conditioned by the sandhi rules active at morpheme boundaries. However, there are also examples of sequences in which no *r* is involved, as in *vehicle* [vɪəkəl] */viəkəl/*. Without going any deeper into the discussion concerning English segmental phonology — which can certainly be analysed in many different ways — the English vowel sequences, in the interpretation favoured here, resemble sequences of vowels and the glides *w* and *y*, which means that the segment */ə/* in the vowel sequences is also functionally equal to a glide. This is, consequently, an obvious typological parallel between British English and Proto-Samoyedic. Certainly, it would be possible to locate still other languages with a similar system.

[23] Another item for which a vowel sequence has been reconstructed before an obstruent is **kåət* 'spruce (*Picea*)' (SW 61). In this case, however, modern Nganasan shows two mutually contradictory stems: *ko'* : (oblique) *kode-* and *kue* : (3rd person singular possessive) *kue-du*, of which the former suggests original **kåt* (with no vowel sequence but with a final glottal stop ' < **t*), while the latter suggests **kåə* (with a vowel sequence but with no final consonant). Even so, the evidence favours the reconstruction of both a vowel sequence and a final consonant (Salminen 2024 : 188), i.e., **kåət* < Proto-Uralic **kaxs/ə*, which may also explain the exceptional vocalism of the Finnic cognate, Finnish *kuusi* : *kuuse-* (SSA 1 : 460).

Since the Finno-Ugric cognates of Samoyedic **tuəy* 'feather, wing' and **peəy* > **piəy* 'outer space; thumb' are of the complex form **tul-ka* and **pel-kä*, respectively, the addition of the suffix **-kA* may be seen as a Finno-Ugric innovation. It has often been claimed that there are no innovations that would characterize the Finno-Ugric languages as a coherent primary branch of Uralic, but these items now go against this claim. Other similar cases are involved in the comparisons of Finno-Ugric **ul-kə-* 'pole', **kul-kə-* 'to move, to float', and **kül-kə-* 'side' with Samoyedic **uy*, **kuy-*, and **kəy-* > **kəy-*, respectively, supposing that these etymologies are accepted. In fact, there are also other examples of derivational forms that unite the Finno-Ugric languages against Samoyedic, a well-known case being Samoyedic **ñåmå* 'hare' (SW 105—106) < Uralic **ñoma*, probably derived from **ño-* 'to chase' (SW 111) and therefore originally **nox/ə-* : **ñox-ma*, whose cognate on the Finno-Ugric side is **ñoma-la* > Saami *njoammil* (YSS no. 780), Hungarian *nyúl* : *nyula-* (MSzFE 491). The conclusion, corroborated by the generally large lexical discrepancy of Finno-Ugric and Samoyedic, is that Samoyedic was, as has conventionally been assumed, the first branch to separate from the context of Proto-Uralic. It may be noted that this conclusion is in no contradiction with the general framework of Uralic language spread in Eurasia, which by all criteria may be assumed to have taken place from east to west.

Acknowledgements. The publication costs of this article were covered by the Estonian Academy of Sciences.

The author thanks Kaisla Kaheinen and Tapani Salminen for valuable comments and advice, as well as for digital sources that have been essential for this paper. Thanks are also due to László Honti for providing digital access to his publications, as well as to the members of the "Etymological workshop" (*Etymologiatyöpaja*) at the University of Helsinki for inspiring discussions concerning a number of individual etyma. Finally, the author gratefully acknowledges the impact of Gerson Klumpp and two anonymous reviewers whose observations and suggestions have helped to clarify a number of important technical issues.

Address

Juha Janhunen
Institute for Linguistic Studies, Russian Academy of Sciences, St. Petersburg
E-mail: asiemajeure@yahoo.com

Abbreviations

EDT — Sir Gerard Clauson, An Etymological Dictionary of Pre-Thirteenth-Century Turkish, Oxford: At the Clarendon Press, 1972; **EH ms.** — Eugene Helimski, Dictionary of Tundra Enets. Digitalized manuscript; **ES** — И. П. Сорокина & Д. С. Болина, Энецкий словарь с кратким грамматическим очерком, Санкт-Петербург: ИЛИ РАН, 2009; **JSWb** — T[oivo] Lehtisalo, Juraksamojedisches Wörterbuch, Helsinki: Suomalais-Ugrilainen Seura, 1956 (LSFU XIII); **KWb** — Kai Donner, Kamassisches Wörterbuch nebst Sprachproben und Hauptzügen der Grammatik. Bearbeitet und herausgegeben von A. J. Joki, Helsinki: Suomalais-Ugrilainen Seura, 1944 (LSFU VIII); **SEO** — Elof Hellquist, Svensk etymologisk ordbok. Ny omarbetad och utvidgad upplaga I—II, Lund: C. W. G. Gleerups förlag, 1939; **SkWb** — Jarmo Alatalo, Sölkupisches Wörterbuch aus Aufzeichnungen von Kai Donner, U. T. Sirelius und Jarmo Alatalo, Helsinki: Suomalais-Ugrilainen Seura, 2004 (LSFU XXX); **SLD** — М. Я. Бармич &

И. А. В э л л о, Словарь ненецко-русский и русско-ненецкий (лесной диалект). 2-ое издание, доработанное, Санкт-Петербург: Просвещение, 2002; **SNg —** Н. Т. К о с т е р к и н а & А. Ч. М о м д е & Т. Ю. Ж д а н о в а, Словарь нганасанско-русский и русско-нганасанский. Санкт-Петербург: Филиал издательства «Просвещение», 2001; **SW —** Juha J a n h u n e n, Samojedischer Wortschatz. Gemeinsamojedische Etymologien, Helsinki 1977 (Castrenianumin toimitteita 17); **SKKS —** Tuomo T u o m i, Suomen kielen käänteissanakirja. Reverse Dictionary of Modern Standard Finnish. 2. korj. painos, Helsinki: Suomalaisen Kirjallisuuden Seura, 1980 (SKST 274); **UED —** Luobbal Sámmol Sámmol Ánte (Ante A i k i o), Uralic Etymological Dictionary (draft version of entries A—Ć). https://www.academia.edu/41659514; **WOT —** András R ó n a-T a s & Árpád B e r t a, West Old Turkic: Turkic Loanwords in Hungarian 1—2, Wiesbaden: Harrassowitz Verlag 2011 (Turcologica 84); **WWb —** Wogulisches Wörterbuch. Gesammelt von Bernát Munkácsi, geordnet, bearbeit und herausgegeben von Béla Kálmán, Budapest: Akadémiai Kiadó, 1986; **YSS —** Juhani L e h t i r a n t a, Yhteissaamelainen sanasto-Helsinki: Suomalais-Ugrilainen Seura, 1989 [2.painos 2001; 3. painos 2015] (MSFOu 200).

REFERENCES

A i k i o, Ante 2002, New and old Samoyed etymologies. — FUF 57, 9—57.

L u o b b á l Sámmol Sámmol Ánte [A i k i o, Ante] 2012, On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic *x. — Lotta Jalava & Janne Saarikivi & Erika Sandman & Tiina Hyytiäinen (eds.), Per Urales ad Orientem. Iter Polyphonicum Multilinguae, Helsinki (MSFOu 264), 227—250.

—— 2015, The Finnic 'secondary *e*-stems' and Proto-Uralic vocalism. — JSFOu 95, 25—66.

—— 2022, Proto-Uralic. — Marianne Bakró-Nagy & Johanna Laakso & Elena Skribnik (eds.), The Oxford Guide to the Uralic Languages, Oxford: Oxford University Press (Oxford Guides to the World's Languages), 3—27.

C u l v e r, Christopher 2025, Proto-Uralic **poŋi* 'fishing seine'. Handout at Etymologiatyöpaja, 27 February 2025. University of Helsinki.

G r ü n t h a l, Riho & H e y d, Volker & H o l o p a i n e n, Sampsa & J a n h u n e n, Juha & K h a n i n a, Olesya & M i e s t a m o, Matti & N i c h o l s, Johanna & S a a r i k i v i, Janne & S i n n e m ä k i, Kaius 2022, Drastic demographic events triggered the Uralic spread. — Diachronica 39 (4), 490—524.

H e l i m s k i, Eugene 1984, Phonological and morphophonological properties of quantity in Samoyed. — Péter Hajdú & László Honti (eds.), Studien zur phonologischen Beschreibung uralischer Sprachen, Budapest: Akadémiai Kiadó, 13—17 [quoted according to the republication in Хелимский 2000 : 41—44].

—— 2005, The 13th Proto-Samoyedic vowel. — Beáta Wagner-Nagy (ed.), Mikola-konferencia 2004, Szeged: SzTE Department of Finnougristics, 27—39.

H e l i m s k i, Eugen (unter Mitarbeit von Beáta Nagy) 1997, Die matorische Sprache: Wörterverzeichnis, Grundzüge der Grammatik, Sprachgeschichte, Szeged (Studia Uralo-Altaica 41).

H o l o p a i n e n, Sampsa 2023a, Notes on an old problem on Hungarian historical vocalism: the sporadic (?) change of Uralic **u* > Hungarian *a*, *á*. — FUF 68, 101—140.

—— 2023b, Hunting for the Uralic accusative(s). — FUF 68, 237—244.

H o n t i László 2002, "The missing link", avagy rokon-e a magyar *toll* és a finn *sulka*? — Mikola-konferencia, Szeged: SZTE Finnugor Tanszék. 75—84. [Republished in: H o n t i László, Magyar nyelvtörténeti tanulmányok, Budapest: Károli Gáspár Református Egyetem — L'Harmattan Kiadó, 2013, 139—151].

—— 2013, Comments on Uralic historical phonology. — ALHung 60 (1), 1—68.

—— 2022, Az ősi uráli tárgyragok története és vesszőfutása. Accusatum et expulsum, Budapest: Tinta Könyvkiadó (Segédkönyvek a nyelvészet tanulmányozásához 226).

Itkonen, Erkki 1969, Zur Wertung der finnisch-ugrischen Lautforschung. — UAJb. 41 (1—4), 76—111.
Janhunen, Juha 1976, Adalékok az északi-szamojéd hangtörténethez: Az első szótagi magánhangzók. — Néprajz és Nyelvtudomány 19—20 [1974—1975], Szeged, 165—188.
—— 1981, Uralilaisen kantakielen sanastosta. — JSFOu 77 (9), 1—56.
—— 1988. Lisiä samojedilaisen vaatetustermin vaiheisiin. — Vir., 88—91.
—— 1992, Petrified verbal nouns in Uralic. — Rédei Festschrift, Wien—Budapest (Studia Uralica 6), 239—244.
—— 1998, Samoyedic. — Daniel Abondolo (ed.), The Uralic Languages, London: Routledge (Routledge Language Family Descriptions), 457—479.
—— 2007, The primary laryngeal in Uralic and beyond. — Jussi Ylikoski & Ante Aikio (eds.), Sámit. sánit, sátnehámit, Helsinki (MSFOu 253), 203—227.
—— 2018, From Tatar to Magyar: Notes on Central Asian Ethnonyms in *-r*. — Ákos Bertalan Apatóczky & Christopher Atwood (eds.), Philology of the Grasslands. Essays in Mongolic, Turkic, and Tungusic Studies, Leiden: Brill (The Languages of Asia Series 17), 138—146.
—— 2022, Two Samoyedic words in Yakut. — Kaisla Kaheinen & Larisa Leisiö & Riku Erkkilä & Toivo E. H. Qiu (eds.), Hämeenmaalta Jamalille. Kirja Tapani Salmiselle, Helsinki: Helda Open Books, 159—166.
Junttila, Santeri 2015, Tiedon kumuloituminen ja trendit lainasanatutkimuksessa. Kantasuomen balttilaislainojen tutkimushistoria, Helsinki. https://helda.helsinki.fi/server/api/core/bitstreams/ce19b2ea-2070-4280-bcdd-b3f4185d019e/content.
Kaheinen, Kaisla 2023, Etymologia ex silentio. Nganasanin äännehistoria ja kielikontaktit, Helsinki (Dissertationes Universitatis Helsingiensis 111/2023).
—— 2024, The origin and development of the Nganasan indicative aorist perfect. — FUF 69, 93—134.
—— [Forthcoming], Remarks on Nganasan phonology: a systematic phonemic account.
Kallio, Petri 2012, The non-initial-syllable vowel reductions from Proto-Uralic to Proto-Finnic. — Lotta Jalava & Janne Saarikivi & Erika Sandman & Tiina Hyytiäinen (eds.), Per Urales ad Orientem. Iter Polyphonicum Multilinguae, Helsinki (MSFOu 264), 163—175.
Klumpp, Gerson 2022, Kamas. — Marianne Bakró-Nagy & Johanna Laakso & Elena Skribnik (eds.), The Oxford Guide to the Uralic Languages, Oxford: Oxford University Press (Oxford Guides to the World's Languages), 817—843.
Lehtinen, Meri 1967, On the origin of the Balto-Finnic long vowels. — UAJb. 39, 147—152.
Liimola, Matti 1951, Etymologisia lisiä. — Vir., 366—370.
O'Rourke, Patrick 2016, Comments on Proto-Uralic Etymology: Derivations and Lexemes. — LU LII, 241—246.
Pystynen, Juho 2018, Itämerensuomen pitkien vokaalien alkuperä. Pro gradu -tutkielma, Helsingin yliopisto. https://www.academia.edu/37417639.
—— 2022, The sources of Proto-Samoyedic **ə*. Powerpoint presentation at the Workshop on Samoyedic Languages and Cultures. Universität Hamburg, 8 September, 2022. https://www.academia.edu/90134140.
Saarikivi, Janne 2022, The divergence of Proto-Uralic and its offspring: A descendent reconstruction. — Marianne Bakró-Nagy & Johanna Laakso & Elena Skribnik (eds.), The Oxford Guide to the Uralic Languages, Oxford: Oxford University Press (Oxford Guides to the World's Languages), 28—58
Salminen, Tapani 1993, A phonemization of Tundra Nenets long vowels. — Marianne Sz. Bakró-Nagy & Enikő Szíj (eds.), Hajdú Péter 70 éves, Budapest: MTA Nyelvtudományi Intézet, 346—352.
—— 1997, Tundra Nenets Inflection, Helsinki (MSFOu 227).
—— 2002, Problems in the taxonomy of the Uralic languages in the light of modern comparative studies. — Т. Б. Агранат & О. А. Казакевич (eds.), Лингвистический беспредел. Сборник статей к 70-летию А. И. Кузнецовой, Москва: Издательство Московского университета, 44—55.

—— 2007, Notes on Forest Nenets phonology. — Jussi Ylikoski & Ante Aikio (eds.), Sámit, sánit, sátnehámit, Helsinki (MSFOu 253), 349—372.
—— 2012, Traces of Proto-Samoyedic vowel contrasts in Nenets. — Lotta Jalava & Janne Saarikivi & Erika Sandman & Tiina Hyytiäinen (eds.), Per Urales ad Orientem. Iter Polyphonicum Multilinguae, Helsinki (MSFOu 264), 339—358.
—— 2023, Suomalais-samojedilaisia sanavertailuja. — Sampsa Holopainen & Jeongdo Kim & Niklas Metsäranta (eds.), Elämä ja etymologia. Janne Saarikiven 50-vuotisjuhlakirja, Helsinki: Helda Open Books, 377—399.
—— 2024, The Samoyed languages. — Edward Vajda (ed.), The Languages and Linguistics of Northern Asia. Vol. 1, Berlin: De Gruyter Mouton, 167—251.
Sammallahti, Pekka 1975, Über das Vokalsystem des Urnordsamojedischen. — FUF 41, 86—112.
—— 1988, Historical phonology of the Uralic languages, with special reference to Samoyed, Ugric, and Permic. — Denis Sinor (ed.), The Uralic Languages: Description, History and Foreign Influences, Leiden: E. J. Brill (Handbuch der Orientalistik VIII.I), 478—554.
Zhivlov, Mikhail 2014, Studies in Uralic vocalism III. — Journal of Language Relationship 12, 113—148.
—— 2023, Reconstruction of Proto-Uralic. — Daniel Abondolo & Riitta-Liisa Valijärvi (eds.), The Uralic Languages, London: Routledge (Routledge Language Family Series), 117—175.
Зайцев Л. А. 2024, Развитие прасамодийских дифтонгов: место в истории самодийского языка-предка. Powerpoint. — Четвертая конференция по уральским, алтайским и палеоазиатским языкам, Санкт-Петербург: Институт лингвистических исследований РАН.
Хелимский Е. А. 1993, Прасамодийские *ə̑ и *ə̈: прауральские источники и нганасанские рефлексы. — Marianne Sz. Bakró-Nagy & Enikő Szíj (eds.), Hajdú Péter 70 éves, Budapest, 125—133 [quoted according to the republication in Хелимский 2000 : 196—201].
—— 2000, Компаративистика, уралистика. Лекции и статьи, Москва: Языки русской культуры.
Янхунен Юха 1991, Нганасаны и распад прасамодийской языковой общности. Семинар "Проблемы происхождения народов уральской языковой семьи". (Тезисы докладов), Ижевск, 16—19.

Ю. А. ЯНХУНЕН (Санкт-Петербург)

САМОДИЙСКИЕ СОЧЕТАНИЯ ГЛАСНЫХ И СТАТУС СОЧЕТАНИЯ **lk* В УРАЛЬСКИХ ЯЗЫКАХ

В данной статье рассматривается проблема представления финно-угорского сочетания **lk* в самодийских языках и его возможная связь с прасамодийскими сочетаниями гласных. В последнее время было выдвинуто предположение, что сочетание **lk* было утрачено на этапе препрасамодийского языка, что, по крайней мере в словах с гласным нижнего подъема во втором слоге, могло привести к формированию нетипичных сочетаний гласных, в которых вторым компонентом выступает гласный нижнего подъема. Эта гипотеза вызывает две проблемы: с одной стороны, необходимо критически пересмотреть диахронические данные, предполагающие утрату сочетания **lk* в интервокальной позиции; с другой — подтвердить или опровергнуть наличие предполагаемых нетипичных сочетаний гласных в прасамодийском языке. В статье делается вывод, что предполагаемые случаи утраты сочетания **lk* в самодийских языках требуют иного объяснения, поскольку данное сочетание является вторичным и производным (**l-k*), причем в рассматриваемых примерах словообразовательный сегмент **k*, присутствующий в финно-угорских языках, изначально

отсутствовал в самодийских языках. Проблема предполагаемых нетипичных сочетаний гласных представляется более сложной, однако, по крайней мере на уровне препрасамодийского языка, вторым компонентом всех таких сочетаний всегда можно считать редуцированный гласный *ə. Это состояние в дальнейшем претерпело вторичные изменения в отдельных самодийских языках, особенно в нганасанском и энецком.

JUHA JANHUNEN (Peterburi)

SAMOJEEDI VOKAALIJÄRJENDID JA UURALI KONSONANTÜHENDI *lk STAATUS

Artiklis käsitletakse soome-ugri konsonantühendi *lk samojeedi esinemust ning selle konsonantühendi ja samojeedi algkeele vokaalijärjendite võimalikku seost. On eeldatud, et *lk on samojeedi algkeele eelsel ajal kadunud ja et vähemalt neis sõnades, mille teises silbis on olnud madal vokaal, on see kaasa toonud ebatüüpilisi vokaalijärjendeid, mille teine osis on madal vokaal. Sellisel juhul tuleb esiteks kriitiliselt uurida vokaalidevahelise konsonantühendi *lk väidetava kao diakroonilisi tõendeid ning teiseks ebatüüpiliste vokaalijärjendite algsamojeedi olemasolu kas tõestada või ümber lükata. Siinses artiklis jõutakse järeldusele, et näiteid, mis osutavad, nagu oleks konsonantühend *lk samojeedi keeltest kadunud, peab selgitama teistmoodi: vaadeldav ühend on sekundaarne (tegelikult *l-k), s.o. *l*-lõpulistele tüvedele on lisatud soome-ugri tuletusliide *k, mida samojeedi keeltes algselt ei olnudki. Ebatüüpiliste vokaalijärjendite probleem on keerulisem, aga vähemalt algsamojeedi eelsel ajal on selle teine osis tagasiviidav redutseeritud vokaalile *ə. Edaspidi on samojeedi eri keeltes toimunud sekundaarseid muutusi, iseäranis nganassaani ja eenetsi keeles.