

LUOBBAL SÁMMOL SÁMMOL ÁNTE (ANTE AIKIO) (Oulu)

STUDIES IN URALIC ETYMOLOGY II: FINNIC ETYMOLOGIES

Abstract. This paper is the second part in a series of studies that present additions to the corpus of etymological comparisons between the Uralic languages, drawing data from all the major branches of the language family. It includes both previously unnoticed cognates that can be added to already established Uralic cognate sets, as well as a few completely new reconstructions of Uralic word roots. In this second part new Uralic etymologies for the following Finnish words are discussed: *aita* 'fence' (< PU **ajta*), *ammottaa* 'gape open' (< PU **ammV-*'yawn'), *kaiho* 'longing, yearning' (< PU **kajšV* 'illness'), *katkera* 'bitter', *katku* 'burnt smell', *katketa* 'break in two' (< PU **kačka-* 'bite'), *korpi* 'dense forest, wildwood' (< PU **korpi*), *ohut* 'thin' (< PU **wokši*), *puhjeta* 'burst; open (of flowers)', *putkahtaa* 'emerge, come up, pop up' (< PU **pučki* 'hollow stalk, tube'), and *sato* 'harvest, crops', *sataa* 'yield harvest' (< PU **čača-* 'grow'). The principles of reconstruction and the citation of lexical material are explained in the first paper of the series (Luobbal Sámmol Sámmol Ánte (Aikio) 2013).

Keywords: Uralic languages, Finnic languages, etymology, historical phonology.

1. Fi *aita* 'fence' ~ KhE *at'* 'fence' < PU **ajta* 'fence'

Fi *aita* 'fence' has cognates in all Finnic languages except for Livonian, and its PFi form can be reconstructed as **aita*. The word has no known etymology beyond Finnic (SSA s.v. *aita*). It has remained overlooked, however, that there is a strikingly similar noun in Khanty with the same meaning: KhE *at'*, *át'*, KhS *ot'*, KhN *os'* 'fence, enclosure' < PKh **āc*.

The vowel correspondence PFi **a* ~ PKh **ā* is quite regular, so to prove the comparison we need to establish that PKh **č* can reflect an earlier cluster **-jt-*. First, one can note that there seem to be no Khanty word roots with a cluster **-jt-*, nor with the sequence **-jət-* which would be the expected result if the cluster had been broken by the addition of an epenthetic schwa. This shows that if the cluster **-jt-* occurred in Pre-PKh, it must have developed into something else in Khanty. That the reflex of Pre-PKh **-jt-* is PKh **č* is demonstrated by the following two Ob-Ugric comparisons that seem to have remained unnoticed by previous research:

- MsN *χujt-*, MsE *kujt-* 'agitate, lure, tempt' (< PMs **kujt-*) ~ KhE *kut'*, KhS *χut'*, KhN *χűś-* 'lure, tempt' (< PKh **kūc-*). — The Khanty verb

has been considered cognate with SaaN *gohččut* 'order, call' and Fi *kutsua* 'call, invite', but apparently erroneously (UEW 192; SSA s.v. *kutsua*). The vowel correspondence between the Saami-Finnic verb and PKh **kūć-* is irregular, and the previously overlooked Mansi cognate shows that medial consonant (**ć* < **jt*) does not match either.

- MsS *wäjtəŋ* 'beautiful' (< PMs **wāj̄təŋ*) ~ KhE *wit'*, KhS *wet'*, KhN *weś* 'beauty' (< PKh **wēc̄*), KhE *wit'əŋ*, KhS *wet'əŋ*, KhN *weśəŋ* 'beautiful' (< PKh **wēc̄əŋ*) — Note also MsN *wēs* 'beauty', *wēsəŋ* 'beautiful', which are Khanty loanwords.

Thus, PKh **āć* 'fence' can be considered a regular cognate of PFi **aita* 'fence', and the PU noun **ajta* 'fence' can be reconstructed. No cognates seem to be found in other branches, but the fully regular sound correspondence and the identical meanings of the compared forms leave little room for doubt regarding the validity of the etymology.

2. Fi *ammottaa* 'gape open' ~ MariW *omeštä-* 'yawn' ~ PSam **ammV-* 'yawn' < PU **ammV-* / **aŋmV-* 'yawn' ← **aŋi* 'mouth, opening'

Fi *ammottaa* 'gape open' has cognates in southern Finnic: Est *ammutama* 'gape open; open one's mouth wide' and Liv *amtlō* 'yawn'. The PFi shape of the verb can be reconstructed as **ammo-tta-*; the Livonian cognate reflects a frequentative derivative **ammo-tt-ele-*. In Finnish there is also a noun **ammo* in expressions such as *olla ammollaan* 'be wide open, be agape'; this may be a back formation based on the verb. SSA (s.v. *ammottaa*) equates the Finnic words with SaaSk *ämmśed* ~ *äämmśed* and MariW *omeštä-* 'yawn', and characterizes the words as "descriptive-colored" (Finn. "desk.-sävyinen"). It is also mentioned that there are similar words in Samoyed languages, but they are not considered cognate with the Finnic word.

The Skolt Saami word must, however, be a loan from Finnic due to its vowel *ä-* (< PSaa **ā-*); in an inherited cognate one would expect a development PU **a-* > PSaa **vuo-*. The sense 'yawn' in Saami seems to show that this meaning was once more widespread in Finnic, even though it is now only found in Livonian. The equation of Finnic **ammo-tta-* and MariW *omeštä-* is, however, phonologically and semantically flawless and one can thus reconstruct a common root **ammV-* for the Finnic and Mari verbs; the derivational suffixes in the items differ. Also some of the Samoyed items mentioned by SSA can be quite naturally included in this etymology: Slk **āmmu-* and Kam *āmoi-* 'yawn' can be straightforwardly equated with **ammV-* reconstructed on the basis of Finnic and Mari. The Selkup and Kamas words have a further cognate in northern Samoyed: Ngan *ŋam'al'əsa* 'yawn'.

Thus, one could reconstruct a PU verb **ammV-* 'yawn'. Ultimately, however, this verb very probably is a derivative of PU **aŋi* 'opening, mouth' (> Komi *vom*, Udm *įm*, KhE *ōŋ*, PSam **aŋ*) (UEW 11–12). According to Alatalo (2004 : 46), Slk **āmmu-* is a derivative of Slk **āŋ* 'mouth', and hence **ammV-* must reflect an earlier form **aŋ-mV-*. Such an analysis is supported by another Samoyed cognate set, NenT *ńaŋkem-*, SlkK *āŋgu-* 'yawn' (< PSam **aŋkV-*), which can be explained as a parallel derivative of the same noun. But importantly, the Finnic and Mari cognates of PSam **ammV-* show that this derivative was already formed in Proto-Uralic. Also the simplification

of the consonant cluster (*-*ɲm-* > *-*mm-*) may have occurred at this stage already, but it may have taken place independently in the daughter branches as well.

3. Fi *kaiho* 'longing, yearning' ~ MdE *kaž* 'bad, poor; accident, misfortune' ~ Komi *kiž* 'stillborn child', Udm *kiž* 'sickness; a spirit that causes sickness' ~ MsN *χojt-* 'be sick' ~ PSam **kājta* 'sickness' < PU **kajšV* 'sickness'

In standard Finnish *kaiho* is mainly known in the meaning 'longing, yearning, wistfulness', but in dialects the word also has meanings such as 'sorrow', 'need, lack', 'worry, fear', 'temper, anger', and 'envy'. The word has the following cognates in other Finnic languages: Kar *kaiho* 'grief, sorrow; poor thing', Est *kahju*, Liv *ka'i* 'damage, injury, loss, harm' (< PFi **kaiho*). No further Uralic cognates have been suggested.

It has been proposed that Fi *kaiho* is of the same origin as Fi *kaihi* 'cataract' (SSA s.v. *kaihi*, *kaiho*), but the semantic difference makes this etymology difficult to substantiate. Despite the skeptical view of SSA and LÄGLOS (Kylstra, Hahmo, Hofstra, Nikkilä, 1991–2012, s.v. *kaihi*), Fi *kaihi* 'cataract' seems best explained as a loan from Proto-Germanic **χaiχaz* (> Gothic *haihs* 'one-eyed'; cognate with Latin *caecus* 'blind'). This etymology presupposes that there was a period of borrowing when PFi **k-* was substituted for Germanic initial **χ-*, while at the same time PFi **-h-* was substituted for Germanic medial **-χ-*. This would be possible if there was some notable allophonic difference between the pronunciation of Germanic initial and medial **χ* at this point (e.g., **[χ]* initially vs. **[h]* medially). In any case, the assumed substitutions have a plausible parallel in Fi *kahle* 'chain' (cf. Old High German *hāhal* 'kettle hook' < **χāχlaz* < Proto-Germanic **χanχlaz*). Fi *haahla* 'chain for hanging a caldron over the fire' is a later loan from the same word.

Fi *kaiho* can, however, be equated with MdE *kaž*, *kažo* 'bad, poor; accident, misfortune', MdM *kaž* 'bad; weak, thin, wretched (person or animal)'. The semantic connection to the Finnic word family is obvious, and the equation is straightforward phonologically as well. PFi **kaiho* presupposes a Pre-PFi form **kajšV-* (the final *-o* may be a suffix), and also the Mordvin word can be regularly derived from the same form. The voicing of the sibilant (**š* > *ž*) is regular in a voiced context, and the loss of preconsonantal **j* appears to be regular as well, at least before coronal consonants. There are not many examples of the latter sound change, as clusters of the type **-jC-* seem to have been quite rare in Uralic. However, Aryan loanwords support the assumption of this sound law: cf. MdE *vano-*, MdM *vanə-* 'look, watch' < Pre-PMd **vajna-* < Aryan **vaina-* 'see, look, watch', MdE *sed'*, M *säd'* 'bridge' < Pre-PMd **sājti* < Aryan **seitu-* 'bridge' (Koivulehto 1999 : 230–231). The same development is also attested in some derivatives based on Uralic stems: MdE *vade-*, MdM *vadə-* 'grease, smear' < Pre-PMd **vaj-ta-* (cf. MdE *oj*, MdM *vaj* 'fat, butter, oil' < **vaj* < PU **waji*; UEW 578), MdE *pid'e-*, MdM *pidə-* 'cook (tr.), bake' < Pre-PMd **pej-tä-* (cf. MdE *pije-*, MdM *pijə-* 'cook (intr.)' < PU **peji-*; UEW 368). Hence, MdE *kaž* reflects PU **kajšV* quite regularly.

Further cognates can be found in Mansi and Samoyed. As both branches show a change **š* > **t*, the Finnic-Mordvin root can be quite naturally

compared to MsN *χojt-*, MsE *kojt-*, MsW *k^wojt-* 'be sick' (< PMs **kāj̥t-*) and NenT *χædo* 'epidemic', EnF *kad'e(d)* 'ache', Ngan *kočə* 'sickness', *kočə-* 'be sick', SlkTa *qūt̪i* 'sickness', *qūt̪i-* 'be sick' (< PSam **kāj̥tə-*; Janhunen 1977 : 58–59). PSam **ā* is a regular reflex of PU **a*. According to the Ob-Ugric vowel developments postulated by Sammallahti (1988 : 504) one would, however, expect PMs **ū*, or perhaps PMs **ā* if the word was originally an **i*-stem (**kaj̥ši*). However, the development PU **a* > PMs **ā* is also attested in MsN *oj-* 'run away' (< PMs **āj̥-*), MsN *ojt-*, MsE *āj̥t-*, MsW *ojt-*, MsS *ajt-* 'unbind, unharness; shoot (with an arrow)' (< PMs **āj̥t-*), which derive from PU **aja-* 'drive; escape' (UEW 4–5). Hence, the development PU **a* > PMs **ā* is probably regular before **j*.¹

It appears that also the following Permic words can be derived from PU **kaj̥šV*: Komi (obsolete) *kjž* 'stillborn child', Udm *kjž* 'sickness; an evil spirit that causes sickness'. The consonant correspondences are quite regular: in Permic sibilants became regularly voiced in voiced contexts, and preconsonantal **j* was lost as in PU **aj̥mä* 'needle' > Komi *jem* (UEW 22) and PU **kojra* 'male' > Komi *kj̥r* (UEW 168). As regards vocalism, however, the normal reflexes of PU **a* are Komi *o*, *u* and Udm *u* (Sammallahti 1988 : 530, 533; Reshetnikov, Zhivlov 2011 : 105–107). But there probably was a conditioned development **-aj-* > Komi and Udm *ĩ*; a parallel is provided by PU **kaji* > Komi, Udm *kj̥* 'awn' (Luobbal Sámmol Sámmol Ánte (Aikio) 2012 : 245; 2013 : 166–167). Previously Komi and Udm *kjž* have been compared to Fi *kitu-* 'be in severe pain, linger in pain, languish' and KhE, KhS *kěčə*, KhN *kāši* 'sickness' (< PKh **kič̥i*), but the vowel correspondences are not regular. Moreover, the Khanty word is quite evidently a loan from Permic **kj̥ž*. The vowel substitution PPerm **ỹ* (Komi *ĩ*) > PKh **i* is attested in several dozen Permic borrowings (Toivonen 1956 : 138; PKh **i* = **ə* in Toivonen's notation). Consider the following examples:

- KhE, KhS *kěr* 'mortar' < PKh **kir* < PPerm **g̃yr* (> Komi, Udm *gjr* 'mortar')
- KhE, KhS *měrt*, KhN *märt* 'with difficulty' < PKh **mirt* < PPerm **m̃rd-* (> Komi *mjrden* 'by force', Udm *mjrden* 'hardly; with difficulty')
- KhE, KhS *měčək*, KhN *māšək* 'fist' < PKh **mičək* < PPerm **m̃ž̃yk* (> Komi, Udm *mjž̃yk* 'fist')
- KhE, KhS *pēm*, KhN *pām* 'heat' < PKh **pim* < PPerm **p̃ym* (> Komi *pjm* 'hot, burning')
- KhE *pěrä-*, KhS *pěr-*, KhN *pār-* 'pass (of time); end' < PKh **pirā-* < PPerm **b̃yr-* (> Komi *bj̥r-*, Udm *bj̥rj-* 'run out; end')
- KhE, KhS *sěr*, KhN *sār* 'spade' < PKh **sir* < PPerm **z̃yr* (> Komi *zj̥r* 'spade')

¹ A different vowel development occurred in MsN, MsS *uj-* 'sink' (< PMs **uj-* < PU **wajV-* 'sink'; UEW 551), but this was probably caused by the initial **w-*. Sammallahti (1988 : 500) maintains that in Proto-Ugric **a* was labialized to **o* after an initial labial consonant (**p-* or **w-*); hence the development was probably **wajV-* > **woji-* > PMs **uj-*. The development PU **-oji-* > PMs **-uj-* is regular: cf. PU **koji* 'male' > PMs **kuj* > MsE, MsW *kuj-*, MsN *χuj-* 'male' (UEW 166–167), PU **koji* 'dawn' > PMs **kuj* > MsS *koj*, MsN *χuj* 'dawn' (UEW 167), PU **śoji* 'sound' > PMs **suj* > MsW *suj*, MsE *soj*, MsN *suj* 'voice' (UEW 482–483).

The semantic relationship between 'damage', 'harm', and 'misfortune' found in Finnic and Mordvin and 'sickness' in Udmurt, Mansi and Samoyed is rather obvious, and English *ill* serves as a parallel; in addition to meaning 'sick', it may mean e.g. 'very bad' (*ill weather*), 'unlucky' (*ill fate*), 'harmful' (*ill effect*), 'immoral' (*ill deed*), and 'malevolent' (*ill intent*). It can be assumed that in Finnic and Mordvin a semantic abstraction from 'sickness' to 'misfortune, accident, damage' took place. Hence, there is no semantic obstacle to reconstructing a PU root **kajšV-* 'sickness; be sick'.

Furthermore, Finnic shows a possible trace of the meaning 'sickness'. The root **kajšV-* may have yet another uncertain reflex in Finnic. In addition to PFi **kaiho*, also an adjective **haikeda* with very similar semantics is attested in the Finnic languages: cf. Fi *haikea* 'sad, wistful, melancholy', Est *haige* 'sick'. It appears possible that **haikeda* developed through an irregular metathesis from an earlier form **kaiheda*; the ending **-eda* is an adjective suffix. On the other hand, there are also the similar nouns *haika* 'faint smell; rumor' and *haiku* 'smoke, reek, coal gas; haze, mist; scent, smell', which are cognate with SaaL *suojgge* 'draft' (< Pre-PSaa **šajka*). Hence, it might be that there originally were two distinct words, PFi **haika* 'smoke, smell, haze' (< **šajka*) and PFi **kaiho* 'sickness; sadness; damage, harm' (< **kajšV*), but these became folk-etymologically mixed through a metathesis **kaiheda* >> **haikeda* in Proto-Finnic. This etymology of Fi *haikea* is, of course, speculative, but this has no bearing on the etymology of Fi *kaiho* and its cognates in Mordvin, Permic, Mansi and Samoyed.

4. Fi *katkera* 'bitter', *katku* 'burnt smell', *katketa* 'break in two' < PU **kačka-* 'bite'

In UEW (113) Fi *katkera* 'bitter' is considered cognate with SaaN *guohca* (GEN *guohccaga*) 'rotten, spoiled' (< PSaa **kuocčk*), MariE *kočo*, MariW *kačə* 'bitter' (< PMari **kačə*) and Udm *keš* 'bitter, tasteless, unsalted', possibly also KhE *kjčjm*, KhS *xečem*, KhN *χšəm* 'mold' (< PKh **kjčjm*), MsS *kášək*, MsW *kašša*, MsE *kēsəγ*, MsN *χāssi* 'mold' (< PMs **kjšγj*); the Uralic proto-form is reconstructed as **kačke-*. This etymology is unacceptable for phonological reasons, however: none of the proposed cognates of the Finnic word suggest a cluster **-čk-*. In Saami the word has a geminate affricate **-cc-* (< PU **-čč-*), and the single affricates in Mari and Ob-Ugric can be explained as regular results of geminate shortening. Sammallahti (1988 : 552) equates the Saami, Mari and Udmurt forms and reconstructs the proto-form **kač(č)V-*. There is no reason to leave the Ob-Ugric words for 'mold' out of this cognate set, however, as they can be regularly derived from the form **kjččV-*. The regular reflexes of PU **j* are PMs **j̄* and PKh **ā*. The high unrounded vowel **j̄* in the Khanty cognate is the high ablaut grade of an original **ā* (Helimski 2001; ЖИВЛОВ 2006 : 42); the ablaut was triggered by the vowel in the suffix **-jm*. However, Udm *keš* shows an obscure sibilant as the assumed reflex of an earlier affricate, and also the vowel *-e-* is irregular; for this reason it is best excluded from the cognate set.

Fi *katkera* cannot reflect PU **kjččV-* due to its consonant cluster *-tk-*, and it must thus have a different etymology. Curiously, even though it has been hypothesized that *katkera* goes back to a proto-form **kačke-*, it has apparently never been connected with the reconstructed verb **kačka-* (>

SaaN *gáskit* 'bite', MariE *kočka-* 'eat', Komi *gečki-* 'ruminant, chew the cud'; UEW 641). The derivation of an adjective meaning 'bitter' from a verb meaning 'bite' is semantically straightforward, and an exact parallel is provided by English *bitter* (< Proto-Germanic **bitra-*), an obscured derivative of *bite* (< Proto-Germanic **bīta-*). The only problem in this etymology is the second-syllable vowel *-e-* instead of expected *-a-* (**katkara*). This slight irregularity need not invalidate an otherwise straightforward etymology, however. The ending *-ra* can be analyzed as an adjective suffix; even though Finnish adjectives ending in *-ra* are rare, there are a couple of adjectives derived from Uralic roots with this suffix: Fi *kumara* 'stooping, bent' < PU **kuma-* 'fall over; knock over' (UEW 201–202), Fi *avara* 'spacious, wide, open' (< PU **ana-* 'open, take off'; UEW 11). A noteworthy case is Fi (dial.) *viherä* ~ Veps, Vote *viher* 'green' (< **viherä*), derived from PU **wiša-* 'green' (UEW 823–824), which serves as a parallel for the irregular change **a > *e* in the second syllable of an adjective derived with **-ra*. The standard Finnish adjective for 'green', *vihreä*, is a later irregular development of **viherä*; note also dialectal *viheriä* 'green'.

In addition to Fi *katkera* 'bitter' there is a similar noun *katku* 'burnt smell, fumes, stink'. This noun, in turn, has been considered cognate with MdE *kačamo* 'smoke' and Komi *kočjs* 'burnt smell', and the proto-form is reconstructed as **kačkV-* in UEW (641). This etymology is not satisfactory because the Mordvin and Komi forms do not support an original cluster **-čk-*. Instead, it is tempting to analyze the Mordvin form as a derivative of the PU root **kiččV-* 'rotten, moldy' discussed above. While it might seem far-fetched to include a noun meaning 'smoke' in this cognate set, there is also a parallel verbal derivative with a meaning closer to the Uralic word family: MdE *kačado-*, MdM *kačadə-* 'smoke (intr.); smell (intr.)'. A parallel for the connection of words for 'bad smell' and 'smoke' can be found in Germanic: cf. English *reek* (earlier 'smoke, vapor, mist') ~ German *Rauch*, Swedish *rök* 'smoke'. The concept that connects the Mordvin words to PU **kiččV-* 'rotten; moldy' is some kind of unpleasant or bad smell; SaaN *guohcistit* 'smell rotten or bad' and SaaL *guohtsa* 'foul smell' come semantically particularly close. Furthermore, there is a previously unnoticed Samoyed cognate that exhibits the meaning of 'bad smell': SlkTa *qētj* 'bad smell, stink', SlkK *qēči-* 'smell, stink'. These words reflect PSam **keca-*, which suits phonologically exactly as the reflex of PU **kiččV-*.

As regards Komi *kočjs*, both SSA and UEW present it as an uncertain cognate. As a reason for this uncertainty it is mentioned that the word had Proto-Komi **o* in the first syllable, which is the regular reflex of PU **e(-ä)* (Sammallahti 1988 : 530): cf., e.g., PU **pesä* 'nest' > PKomi **poz(j-)* (UEW 375), PU **elä-* 'live' > PKomi **ol-* (UEW 73); PU **čečä* 'uncle' > PKomi **čož* (UEW 34–35). UEW states the irregularity could be explained by the influence of the following affricate **č*, but this is an *ad hoc* explanation, as there appear to be no parallels for the assumed sound change. Due to the irregular vowel it remains unclear whether Komi *kočjs* has any relation to PU **kiččV-* 'rotten, moldy'.

Let us return to Fi *katku* 'burnt smell, fumes, stink'. This word can have nothing to do with the PU **kiččV-* 'rotten, moldy', as it shows an unmatching consonant cluster *-tk-*. Fi *katku* also has a south Finnic cognate which demonstrates that the word originally had an affricate: Est *katk*, Võro *katsk*

'plague, epidemic' < **kacku*.² Semantically it is not at all difficult to connect a noun meaning 'burnt smell' and 'fumes' to a verb meaning 'bite'. One can note that both the Saami and Mari cognates are contextually used in reference to the unpleasant effects of smoke: cf. SaaN *suovva gáská čalm-miid* 'the smoke irritates ("bites") the eyes', MariE *šičam šikš kočkeš* 'the smoke irritates ("eats") my eyes'. Hence, also Fi *katku* 'burnt smell, fumes, stink' can be included in the same word family.

Also the Finnish verbs *katketa* (*katkea-*) 'break in two (intr.)' and *katkaista* (*katkaise-*) 'break in two (tr.)' need to be considered in this connection. These also reflect a PFi root **kacke-*; the affricate is preserved in Vöro *katški* 'broken' (~ Fi *katki* 'broken in two'). UEW (641) derives these words from yet another reconstruct **kačka-*, this time supposedly reflected also in MariE *kuškeða-*, MariW *kâškeðä-* 'tear off, tear in two' and Udm *kwačka-* 'tear (intr.)'. At least the Mari verb can have nothing to do with the others, as Mari *-šk-* does not reflect earlier **-čk-*, and also the vowel (MariE *u*, MariW *ê* < PMari **ũ*) is irregular. Instead, the Mari verb must be a reflex of PU **kiško-* 'tear' (> SaaN *gaikut*, Fi *kiskoa* 'tear', Komi *koš-*, Udm *kešj-* 'rip', KhE *kös-* 'tear down, break up'; cf. UEW 162; SSA s.v. *kiskoa*). The development PU **i* > PMari **ũ* occurred in disharmonic stems, cf. PU **wiša-ra* 'green' > PMari **ũžar* > MariE *užar*, MariW *êžar* ~ *žar* (this is cognate with the Finnish formation *viherä* 'green' discussed above).

Whether Udm *kwačka-* can be analyzed as a cognate of Fi *katketa* is not clear. The vowel correspondence is unusual, as in other cases Udm *kwa-* corresponds to Fi *ko-*: cf. Udm *kwaś* 'shallow' ~ Fi *koski* 'rapids' (SSA s.v. *koski*; cf. UEW 674), Udm *kwala* 'summer cabin' ~ Fi *kota* '(Saami) tent' (UEW 190), Udm *kwaldj-* 'split' ~ Fi *kolo* 'hole, hollow' (UEW 174). Hence, Udm *kwačka-* would be a regular reflex of the form **kočkV-*, not **kačka-*. Regardless of the origin of Udm *kwačka-*, however, it can hardly be denied that also Fi *katketa* and *katkaista* must be related to PU **kačka-* 'bite'. Here we seem to have yet a different semantic development in Finnic from 'bite loose' to 'break in two'.

Finally, it can be added that the PU verb **kačka-* 'bite' also has previously unnoticed Ob-Ugric reflexes: KhE *kjč-*, KhS *χeč-*, *χeš-*, KhN *χiš-* 'hurt, ache; sting (of a nettle)' and KhE *kjč*, KhS *χeš* 'nettle' (< PKh **kjč-*) and MsS *kōš-*, MsW *kuš-*, MsN *χūs-* 'sting (of a nettle)' (< PMs **kūš-*). The semantics of these verbs comes very close to the secondary meanings of the Saami reflexes of PU **kačka-*: SaaN *gáskit*, I *käskid* 'bite; sting, smart, burn'. It is also noteworthy that SaaN *gáskálas* 'nettle' is a semantic development of the adjective *gáskálas* 'apt to bite', a derivative of the verb *gáskit*. As regards the historical phonology of the Ob-Ugric forms, the vowel correspondence is regular: the regular reflexes of PU **a*(-*a*) are PKh **ā* and PMs **ū*, and

² The semantic development 'burnt smell, fumes, stink' > 'plague, epidemic' can be understood in context of the so-called miasmatic theory, i.e. the belief that diseases and epidemics are caused by bad air emanating from rotting organic matter. This belief was prevalent in Europe, India and China since ancient times, and was only gradually displaced in the 19th century by the germ theory of disease. As a parallel to the semantic development of Finnic **katku* one can mention *malaria* from medieval Italian *mala aria* 'bad air'. It is also noteworthy that in some dialects Est *katk* means 'rotten spot in a swamp, deep puddle of mud'. It was commonly held that disease-causing bad air emanated from swamps, and also malaria has been formerly called *marsh fever* due to its association with swamps and marshlands.

the attested PKh **ĭ* is the high ablaut grade of **ā* (Helimski 2001; Живлов 2006 : 42), implying that there has been a lost suffix in the stem that triggered the ablaut. Also the loss of PU **k* in the cluster **čk* may be regular, but no other examples of the development of this cluster Ob-Ugric seem to be known that would allow this to be verified.

To sum up, it is not justified to reconstruct the four near-homonymous roots **kačka-* 'bite', **kačka-* 'tear in two', **kačke-* 'bitter', and **kačkV-* 'smoke, burnt smell' (cf. UEW 113, 641–642). Instead, the following two cognate sets can be postulated:

- PU **kačka-* 'bite': SaaN *gáskit* 'bite; sting, smart, burn', Fi *katkera* 'bitter', *katku* 'burnt smell, fumes', *katketa* 'break in two', MariE *kočka-* 'eat', Komi *ggčki-* 'ruminant', KhE *kjč-* 'hurt, ache; sting (of a nettle)', MsN *χūs-* 'sting (of a nettle)'
- PU **kjččV-* 'smelly, rotten, moldy': SaaN *guohca* 'rotten', Mde *kačamo* 'smoke', *kačado-* 'smoke; smell', MariE *kočo* 'bitter', KhE *kjčjm*, MsN *χāssi* 'mold', SlkK *qēči-* 'smell, stink'

**5. Fi *korpi* 'dense forest, wildwood' ~ Mde *kuro* 'bush, shrub' ~ MsW *kōrp* 'forest, woods'
 < PU **korpi* 'woods'**

Fi *korpi* has cognates in most Finnic languages, e.g. Veps *kořb*, Est *kõrb* 'woodland, wildwood' (< PFi **korpi*), but no generally accepted cognates outside Finnic. SSA (s.v. *korpi*) maintains that the cognates proposed from more distantly related languages are quite uncertain. UEW (217) notes that Finnish *korpi* has been compared to Mde *kuro*, Mdm *kur(a)* 'bush, shrub; bushes', KhS, KhN *χār* 'forest (esp. as a hunting ground)', MsW *kōrp* 'forest' and NenF *kur^ə* 'dense forest on the shore of a river', but rejects the inclusion of the Finnish word in this etymological set. Also the equation of the Mordvin, Khanty, Mansi and Nenets words is considered uncertain by UEW.

The equation between the Finnic and Mansi words was first presented, with hesitation, by Liimola (1956 : 243–244). This comparison can be rehabilitated; the sound correspondence between the items is fully regular. On the basis of MsN *χōrp*, MsW *kōrp*, Mse *kōrəp*, MsS *kərp*, *kərəp* a Proto-Mansi form **kārp* can be reconstructed. As regards vowels, Proto-Uralic **o* has developed into PMs **ā* in stems of the type *(C)*o*CCi- (Sammallahti 1988 : 504), so the match between PMs **kārp* and PFi **korpi* is fully regular. The consonants in the two forms are identical, and the correspondences **k* ~ **k*, **r* ~ **r* and **p* ~ **p* are of course regular. Even so, UEW states the comparison between Finnic and Mansi is rejectable due to the Finnic consonant cluster **rp*. This statement is not logical, as the Mansi word has the same cluster. UEW maintains that Mansi *-p* is a derivational suffix, but this suggestion is not substantiated in any way. The meanings of the words are nearly identical: the Finnic words generally mean 'dense forest, wildwood', and in Mansi meanings such as 'birch forest' and 'fir forest' have been attested; the MsS form *kərp*, *kərəp* means 'grove'.

On the basis of Finnic and Mansi a Proto-Uralic noun **korpi* 'forest' can be reconstructed. Also Mde *kuro*, Mdm *kur(a)* 'bush, shrub; bushes' (< PMd **kurə*), which was already mentioned above, can be considered a reflex of this Uralic word. Semantically the comparison is quite natural;

cf., e.g., MariE *čodəra* 'forest, woods' ~ MariW *caδra* 'twigs, withered branches'. The vowel correspondence requires more detailed argumentation, however, because according to Itkonen (1946 : 205) the regular reflex of PU *o(–i) is PMd *o, not *u. Itkonen gives eight examples of the correspondence PFi *o(–i/e) ~ PMd *o, on which he bases his assumption of regular development:

- MdM *Jov* (name of the river Moksha) (< PMd *jov) ~ Fi *joki* 'river'
- MdE, MdM *jonks* (< PMd *jonks) ~ Fi *jousi* 'bow'
- MdE *kolmo*, MdM *kolma* (< PMd *kolmə) ~ Fi *kolme* 'three'
- MdE *lokšej*, MdM *lokšt'i* (< PMd *lokštəŋ) ~ Fi *joutsen* 'swan'
- MdE *nolgo*, MdM *nolga* (< PMd *nolgə) ~ Fi *nolki* 'snot'
- MdE *ovto*, MdM *ofta* (< PMd *ovtə) ~ Võro *ots* : GEN *ote* 'bear'
- MdE *olgo* (< PMd *olgə) ~ Fi *olki* 'straw'
- MdE *pongo*, MdM *pov* (< PMd *poŋə) ~ Fi *povi* 'bosom'

Not all of these examples are convincing. As regards Fi *joki* 'river' and *kolme* 'three', their Saami cognates point to an original *u instead (SaaN *johka* 'river' < *juki, *golbma* 'three' < *kulmi). As the cognates of these words in more eastern Uralic languages also show irregular vocalism (cf. UEW 99, 174; Sammallahti 1988 : 537, 543), the reconstructions *joki 'river' and *kolmi 'three' cannot be substantiated; it is possible that the Mordvin forms derive from *juki and *kulmi instead and thus reflect the regular change PU *u > PMd *o. The words for 'swan' are phonologically even more obscure; Fi *j- ~ PMd *l- is a completely irregular correspondence, and also the other suggested cognates show strange deviations (e.g., SaaN *njukča* 'swan' < PSaa *nukčə has an initial nasal and *u < Pre-PSaa *ū), so this etymology cannot serve as an example of regular vowel development. Two of the proposed examples, 'bear' and 'straw', are only attested in Finnic and Mordvin, and they may have been borrowed between Pre-Proto-Finnic and Pre-Proto-Mordvin.

The only widespread Uralic words not following the sound law *o(–i) > PMd *u are PMd *jonks 'bow' (< PU *jon̄si), PMd *poŋə 'bosom' (< PU *poŋi), and PMd *nolgə 'snot' (< PU *nolki). In contrast, there are at least twelve cases showing the vowel development *o(–i) > PMd *u:

- PU *oksinta- 'vomit' > PMd *uksəndə- > MdE *uksno-*, MdM *uksəndə-* (UEW 716)
- PU *omti 'cavity, hollow' > PMd *undə > MdE *undo*, MdM *unda* (UEW 338)
- PU *colmi 'knot' > PMd *sulmə > MdE *šulmo*, MdM *šulma* (UEW 38)
- PU *korpi- 'blaze' > PMd *kurva- > MdE *kurva-* (UEW 186)
- PU *moški- 'wash' > PMd *muškə- > MdE *muške-*, MdM *muškə-* (UEW 289)
- PU *soksi 'worm, maggot' > PMd *suks > MdE, MdM *suks* (UEW 764)
- PU *sormi 'finger' > PMd *sur > MdE, MdM *sur* (UEW 765)
- PU *soski- 'chew' > PMd *suske- > MdE *susko-*, MdM *suske-* 'bite' (UEW 448–449)
- PU *solki 'clasp' > PMd *sulgamə > MdE *šulgamo*, MdM *šulgam* (UEW 774–775)
- PU *šodi- 'leak, flow' > PMd *čudə- > MdE *čude-*, MdM *šudə-* (UEW 786)
- PU *totki 'tench' > PMd *tutke > MdM *tutka* (UEW 532)
- PU *woli- 'be' > PMd *ulə- > MdE *ule-*, MdM *ulə-* (UEW 580–581)

Thus, the regular development is clearly PU $*o(-i) > \text{PMd } *u$, contra Itkonen (1946). The lack of the change $*o > *u$ in the words $*jõnks$ 'bow' (< PU $*jõnsi$) and $*põnõ$ 'bosom' (< PU $*põni$) seems to be conditioned by the following velar nasal. There is also a third etymology that serves as an example of this rule: Mde *onkšt'* (PL), Mdm *ovs ~ ovks* 'bit (in bridle)' (< PMd $*õns$) ~ SaaN *vuõnas* 'muzzle of a dog' (< PSaa $*vuõngs$), the proto-form of which can be reconstructed as PU $*õnis$. UEW (11) and Sammal-lahti (1988 : 542) further equate these words with PU $*õji$ 'opening, mouth', but this is not supported by the vocalism of the Mordvin form: the development PU $*a > \text{PMd } *o$ would be irregular. This leaves only PMd $*nolgõ$ 'snot' as an unexplained exception; the word is affective, which might account for its irregularity.

We can now return to the case of Mde *kuro*, Mdm *kur(a)* 'bush, shrub; bushes'. The regular vowel correspondence established above supports deriving this word from PU $*kõrpi$ 'woods'. As for the consonant correspondence, PU $*-rp-$ > Md $-r-$ is apparently not fully regular, as the reflex of this cluster is Md $-rv-$ in at least two cases: cf. PU $*kõrpi-$ 'scorch' > Mde *kurva-* (UEW 186), PU $*turpa$ 'lip' > Mde *turva*, Mdm *tõrva* (UEW 801). However, the complete loss of $*p$ in the cluster $*-rp-$ is attested in one already well-established etymology: PU $*õrpas$ 'orphan' > Mde *uros*, Mdm *urõs* (UEW 343). Thus, there is no obstacle to analyzing Mde *kuro*, Mdm *kur(a)* as a reflex of PU $*kõrpi$. However, the remaining proposed cognates (KhE $-kãrj$ 'place' (in compounds), KhS $\check{\chi}ãr$, KhN $\check{\chi}ãr$, $\check{\chi}ãri$ 'forest (esp. as a hunting ground); clearing' (< PKh $*kjrj$) and NenF *kur^a*) show no regular correspondence to PU $*kõrpi$; hence, they must be excluded from this etymology.

**6. Fi *ohut* 'thin' ~ KhE *woγal'*, MsN *wõwta* 'thin'
< PU $*wokši-$ 'thin'**

The standard Finnish word for 'thin' is *ohut*. In dialects also parallel forms with different derivational suffixes are attested, such as *ohea* (< $*oheda$), *ohukainen* and its syncopated variant *ohkainen* (< $*ohukkainen$). All three forms have cognates in other Finnic languages as well: cf. Votic *ghud* (< $*ohut$) and Est *õhe* (< $*oheda$), *õhukene* (< $*ohukkainen$) 'thin'. A Proto-Finnic root $*ohe-$ can be reconstructed on the basis of these forms; a similar pattern of parallel derivatives is attested in some other Finnic adjectives as well, e.g. Fi *kevyt* (< $*kebü$) ~ *kepeä* (< $*kepedä$) ~ *köykäinen* (< $*kebükkäinen$) 'light'. The root $*ohe-$ has had no further etymology so far (SSA s.v. *ohut*). However, it is of Uralic origin: it has fully regular cognates in the Ob-Ugric languages, which have so far remained unnoticed due to the rather non-transparent sound correspondences involved.

Honti (1982 : 193) equates the following Ob-Ugric adjectives: KhE (Vj) *woγal'*, (Trj) *wõγ.əλ*, KhS *woχət*, KhN (Ni) *uχət*, (Kaz) *oχəλ*, (O) *oχəλ* 'thin (of flat things)' (< PKh $*wayəl$), MsW *wayta*, MsN *wõwta* 'thin' (< PMs $*waytā$). These words show no superficial resemblance to PFi $*ohe-$, but the sound correspondence is actually fully regular. The initial $*w-$ presupposed by Ob-Ugric is lost in Finnic before the vowel $*o$, as in PU $*woli-$ 'be' > Fi *ole-* (UEW 580), PU $*wolka$ 'shoulder' > Fi *olka* (UEW 581) and PU $*wosta-$ 'buy' > Fi *osta-* (UEW 585). As for the medial consonants, the corre-

spondence PKh **-ɣəl-* ~ PMs **-ɣt-* implies Ob-Ugric **-ɣθ-*, which can reflect four different PU clusters: **sk*, **ks*, **šk* and **kš*. The sibilants **s* and **š* have merged in Ugric languages, and a regular metathesis also took place in clusters of the type **sk* and **šk* in Ob-Ugric. The last one of the four alternatives, PU **kš*, accounts for the correspondence between Finnic and Ob-Ugric, as the regular reflex of this cluster is **h* in Finnic: cf. PU **mekši* 'bee' > Fi *mehi-läinen* (UEW 271), PU **makša-* 'rotten (wood)' > Fi *mahi* (UEW 698). Hence, a Proto-Uralic root **wokši-* 'thin' can be reconstructed on the basis of Finnic and Ob-Ugric.

As regards vowels, PFi **ohe-* presupposes a Uralic vowel combination **o—i*. PU **o* in stems of the type **(C)oCCi-* is usually reflected as PKh **a* and PMs **ā*. Examples include:

PU <i>*omti</i> 'cavity'	> PKh <i>*ant</i> , PMs <i>*āntər</i> (UEW 338)
PU <i>*ponči</i> 'tail'	> PKh <i>*pač</i> , PMs <i>*pānš</i> (UEW 353)
PU <i>*sorši</i> 'span'	> PKh <i>*sarəs</i> , PMs <i>*tārās</i> (UEW 448)
PU <i>*woŋki</i> 'den'	> PKh <i>*waŋk</i> , PMs <i>*wāŋkā</i> (UEW 583)
PU <i>*korpi</i> 'woods'	> PMs <i>*kār̥p</i> (see etymology 5)

The reflexes of **wokši-* 'thin', however, display the correspondence PKh **a* ~ PMs **ā*. The difference in Mansi vowel length results from a conditioning factor: if the following consonant was PMs **ɣ*, a shortening **ā* > **a* has taken place. Compare the following two parallels:

PU <i>*joŋsi</i> 'bow'	> PKh <i>*jaɣəl</i> , PMs <i>*jaɣt</i> (UEW 101—102)
PU <i>*soski-</i> 'chew'	> PKh <i>*Laɣəl-</i> , PMs <i>*taɣt-</i> (UEW 448—449)

The otherwise very common vowel correspondence PKh **a* ~ PMs **ā* apparently never occurs in words where the following consonant is PMs **ɣ*; no such case can be found in the list of Ob-Ugric cognate sets presented by Honti (1982 : 123—198). This offers further support for the hypothesis that a secondary vowel shortening **ā* > **a* took place before PMs **ɣ*. In addition to the well-established Mansi reflexes of PU **joŋsi* 'bow' and **soski-* 'chew', the word **wokši* 'thin' > PMs **waytā* provides a third example of an Uralic item that follows this sound law.

**7. Fi *puhjeta* (*puhkea-*) 'burst; open (of flowers), come out (of leaves)' and *putkahtaa* 'emerge, come up, appear (suddenly), pop up'
< PU **pučki* 'tube; stalk (?)'**

The Fi verb *puhjeta* (*puhkea-*) has cognates in all Finnic languages, while the similar word *putkahtaa* is only attested in Finnish and Karelian (SSA s.v. *puhjeta*, *putkahtaa*). SSA hesitatingly equates these words with MdE *počkode-* 'burst', MdM *počkod'ə-* 'burst; open (of sprouts)', Komi and Udm *bički-* 'sting, pierce'. In addition, the dictionary describes the verbs *puhjeta* and *putkahtaa* as 'descriptive'. This characterization appears misleading, as it remains unclear what sound symbolic conventions might be manifested in these words.

In any case, the equation of the Finnic, Mordvin and Permic words is quite convincing, and the underlying Uralic root can be reconstructed as **pučki-*. The dual representation of the consonant cluster (*-hk-* ~ *-tk-*) in Finnic appears to result from an irregular split of the preconsonantal affricate

*č in Pre-Finnic. The cluster *-tk-* is the regular reflex of PU **-čk-*; cf., e.g., Fi *kotka* 'eagle' (< PU **kočka*; UEW 668) and Fi *notko* 'depression' (< **nočko*; UEW 714). The form *puhkea-* can be explained as the outcome of an irregular assibilation in Pre-PFi (**pučki-ta* >> **puški-ta* > PFi **puhke-da-*). Exactly the same phonological split has occurred in another Finnic word-family as well: Fi *potka* 'shank' (< **počka*) and Fi *pohje* (GEN *pohkeen*), dial. *pohkea* 'calf (of the leg)' (< PFi **pohkeda* < Pre-PFi **poškita* << **počkita*); the word is cognate with SaaN *boaski* 'ankle (of an animal)', SaaS *bâetskie* 'heel' (< PSaa **poackē*) (SSA s.v. *potka*).

The initial *b-* in Komi and Udm *bički-* is not regular, but there are several other, well-known examples of sporadic voicing of initial stops in Permic: e.g., Komi *bež*, Udm *biž* 'tail' (< PU **ponči* 'tail'; UEW 353), Komi *dor*, Udm *dur* 'edge' (< PU **terä* 'edge, blade'; cf. UEW 522, 795),³ Komi *gžž*, Udm *gžž* 'nail' (< PU **künči* 'nail'; UEW 157). In the verb *bički-* also the affective semantics might have played a role in the voicing.

It is notable that the reconstructed root **pučki-* 'burst' is homonymous with another well-established PU root **pučki* 'tube; stalk; Angelica (a plant with a hollow stalk)'. This word has widely attested reflexes, e.g., SaaN *boska* 'garden Angelica (Angelica Archangelica)', Fi *putki* 'tube; Angelica', MdE *počko* 'tube; hollow stalk; Angelica', MariE *puč* 'stalk; tube', NenT *pud^o* 'small metal tube', Slk **pūčō* 'tube; soft inner part of plants; inside; middle', Kam *pūt* 'marrow; inside, inner part; groove'. As the reconstructed stems are homonymous, it is in order to examine whether they could also be etymologically identical.

It is noteworthy that the Finnish verb *puhjeta* also has meanings related to the growth of plants: cf. *lehdet puhkeavat* 'the leaves come out', *puhjeta kukkaan* 'blossom, open its flowers' — literally 'the leaves burst', 'to burst into flowers'. A similar usage is attested in MdM *počkədə-* 'open (of sprouts)'. This already brings the comparison closer to the word **pučki* 'tube; stalk; Angelica'. One can thus reconstruct a derived verb **pučki-ta-* 'open (of sprouts, flowers), blossom', which already at an early stage developed the secondary meaning 'burst'; as parallels, compare Lithuanian *sprógti* 'burst; blossom' and Hungarian *feslik* 'rip (intr.), get torn (e.g., of clothes); blossom'. The meaning 'burst', then, gave rise to its transitive equivalent 'pierce' in Permic.

Notably, also the nominal reflexes of PU **pučki* 'tube; stalk' have developed a wide range of derived and more abstract meanings. In the Samoyed reflexes one encounters, in addition to 'tube' and the like, also more abstract meanings connected to 'inside, inner part of something'. This is evident in the Selkup reflexes, for instance; Bykonya's Selkup dialect dictionary gives the forms ObSh, ObCh, Ty *puž*, ObS, Vas *pūžə*, Tur, El *pūtj*, Ket *puča* (Быконя

³ UEW (522, 795) distinguishes the cognate sets for 'blade' and 'edge': SaaN *dearri*, Fi *terä*, MariE *tür*, MariW *tər* 'blade', Udm *tir* 'ax', Hung *tór* 'dagger' (< PU **terä*) and MariE *tür*, MariW *tər*, Komi *dor*, Udm *dur* 'edge; shore' (< PU **terV*). This analysis is clearly erroneous; in Mari the words are homonymous, and since there are parallels for the semantic relationship 'edge' ~ 'blade' (such as English *edge* < Germanic **agjō-* '(edge of a) blade'), it is reasonable to postulate only one underlying etymon, PU **terä* 'edge/blade'. The distinction between Udm *tir* 'axe' and *dur* 'edge; shore' does not support the reconstruction of two originally distinct roots, as Udm *u* is a regular reflex of PU **e* but Udm *i* is not. Therefore, Udm *tir* 'axe' most probably has another etymology, and is not related to Fi *terä* 'blade' and its cognates.

2005 : 197), with the diverse meanings 'inside, interior; soul; stomach; navel; seed; core; pipe (instrument); barrel (of a gun)'. Also the meaning of a type of plant with a hollow stalk is found in SlkEl *pūt* 'Angelica (Russ. *дудка*)' (Быконя 2005 : 198) and SlkTa *pūtj* 'Siberian hogweed (*Heracleum sibiricum*)' (Хелимский 2007).⁴

Semantic abstraction similar to Selkup is also found in the other Samoyed cognates: Kamas *pūt* 'groove; inside; marrow', *bün pūt* 'riverbed' (*bün* 'water, river' GEN), *kūzan pūt* 'human innards' (*kūzan* 'person' GEN), Ngan *hüüdəä* 'core; middle of a river' (Helinski 1997 : 249). According to Lehtisalo (1956 : 369), the NenT cognate *pud^o* means 'small metal tube (used as an ornament in a girl's cap)', but a more abstract meaning of 'core' is found in expressions such as *paṃ pud^o* 'core of a tree' (*pa* 'tree, wood'), *nuw^om pud^o* 'the Milky Way' (*num* 'sky, heaven'), and *jam pud^o* 'the core of the earth' (*ja* 'earth'). Tereshchenko (Терещенко 1968 s.v. *nyd*) gives only an abstract meaning 'that which keeps something stable or balanced (e.g. spine, core of a tree)'. The derivative *pudo* means 'spinal cord', as does its Forest Enets cognate *puduj*. A further degree of abstraction is found Ngan *hüütəä* 'body, figure' and EnF *pudodo* 'body; the person himself'; from the meaning 'body' the word was finally grammaticalized as a reflexive pronoun in Mator (Helinski 1997 : 249); apparently, the Tundra Nenets pronoun root *pid-* (< **piṭə-*) is of the same origin, even though it shows an unexpected illabial vowel. Thus, the long path of semantic divergence has created a semantically absurd cognate relationship between the westernmost and easternmost Uralic languages: South Saami *batske* 'flower stalk of Angelica' turns out to be etymologically identical with the Mator reflexive pronoun *hudu!*

The semantic abstraction '(hollow) stalk, stem (of a plant)' > 'inside, core' has apparently begun already in Proto-Uralic, as the meaning 'inside' is also attested in Komi *pič* (*pičk-*) and Udm *puč* (*pučk-*) ~ *puš* (*pušk-*), which can be included in this cognate set; local case forms of these Permic nouns also function as adverbs and postpositions in the senses 'in', 'into'. In Udmurt also the more concrete meanings 'insides, intestines', 'womb' and 'stomach' are attested. The meaning 'stomach' is found in Selkup, too, and also the Kamas expression *kūzan pūt* 'human innards' is notable in this connection. The Permic forms had been included in the cognate set by Sammallahti (1979 : 35), but for some reason they are no longer cited in his later paper (Sammallahti 1988 : 539). UEW (397–398) makes a distinction between two both phonologically and semantically distinct etymological sets: on the one hand, the Permic words meaning 'inside' and the Samoyed items with similar meanings are derived from a PU word **pučkV* 'inside', and on the other, the words with concrete meanings such as 'tube', 'stalk' and 'Angelica' are derived from another reconstruct **pučke*. The latter reconstruction is, however, clearly incorrect because none of the cognates show any evidence for a palatalized affricate **č*; the cognate set quite unambiguously points to the PU cluster **-čk-*.

As the two correspondence sets yield an identical Proto-Uralic reconstruction, and their assumed reflexes in Samoyed languages also show mean-

⁴ It is not altogether clear whether the gloss of some variants in Bykonja's dictionary (Быконя 2005) should be read as 'Angelica' instead of 'pipe', as Russian *дудка* means both. SlkEl *pūt* is unambiguously glossed as 'дудка (растение)', however.

ings that bridge the semantic gap between the two groups, there are no grounds for treating them as etymologically separate sets. The remaining question is how exactly the semantic heterogeneity of the attested cognates is to be accounted for, or more precisely, how the duality between the meanings 'hollow stalk; Angelica, plant with a hollow stalk' and 'inside, interior, core' originally developed. As the semantic bifurcation seems to date to a very early phase — likely to Proto-Uralic already — the details of the process can hardly be reliably reconstructed any longer. One possibility, however, is offered by the meaning 'soft inner part of plants' attested in Selkup dialects. One could surmise a development 'plant stalk' > 'the soft inner part of a plant stalk' and further > 'some kind of soft inner part in general (e.g. marrow, intestines)' > 'inside, core'. On the other hand, one could also think of a development 'hollow stalk; tube' > 'marrow bone', and a further metonymic shift yielding the sense of 'marrow'.

Regardless of the how exactly each semantic shift in this word family has taken place, all the attested meanings of the cognates can be rather straightforwardly derived from a relatively narrow axis of basic meanings, as shown in Figure 1.

8. Fi *sato* 'harvest, crops', *sataa* 'yield harvest' < PU *čaća- 'grow'

According to SSA (s.v. *sataa*, *sato*), Fi *sato* 'harvest, crops' is etymologically the same word as Finnic **sato* 'falling; rain' (> Fi comp. *sulka-sato* 'molt-ing' (*sulka* 'feather'), Vot *sato* 'heavy rain', Est *sadu*, Liv *sa'd* 'rain'), and a derivative of PFi **sata-* 'rain; fall' (> Fi *sataa* 'rain', Est *sadama* 'rain, fall', etc.). In Finnish dialects there is also a verb *sataa* 'yield harvest (subj.: field, corn)'. The verb **sata-* 'rain' derives from PU **šada-* and is cognate with Proto-Samoyed **sârâ-* 'rain' (Janhunen 1981 : 221). From a semantic perspective it is far from obvious, however, that Fi *sato* 'harvest' and *sata-* 'yield harvest' are reflexes of PFi **sato* 'falling; rain' and **sata-* 'fall; rain', and not merely coincidentally homonymous with them.

Setälä (1902 : 222) was probably the first to connect *sato* 'harvest' with *sataa* 'rain', and he suggested that the Finnic word is cognate with SaaN *čuohcit* 'get stuck (of fish in a net); have an effect on', MdE *čačo-*, MariE *šoča-* 'be born; grow, yield harvest', Komi *čuž-* 'be born; grow', Udm *čijžj-vjžj* 'relative' and KhN *šošə* 'local, native'. Toivonen (1928 : 87–88), however, treated the two Finnic words as etymologically distinct: he considered Fi *sataa* 'rain' the cognate of SaaS *tjuetsedh* 'snow' and SaaI *čuoccâd* 'molt', whereas *sataa* 'yield harvest' and *sato* 'harvest' were in his view cognate with the Mordvin, Mari, Permic and Khanty words mentioned above, as well as MsW *šošiγ*, MsE *sāsə*, MsN *sossa* 'homestead, homeland'. Ravila (1938 : 19), in turn, considered Fi *sataa* 'yield harvest' and *sato* 'harvest' etymologically identical with Fi *sataa* 'rain', and maintained that they are in no connection with the Mordvin and Mari verbs meaning 'be born; grow'. UEW (52) reconstructs the Proto-Uralic form *čaćV- ~ *čančV-, and states that Fi *sato* 'harvest' cannot be included in this cognate set due to its meaning and its initial *s-*. In addition to the words mentioned above, the dictionary also presents Samoyed cognates: NenT *tencə* 'kind, sort; tribe', EnF *tjz*, Ngan *tansə* 'tribe, clan', and Slk *čāči (SlkTa *tātj*, SlkTy *čāž*, SlkK

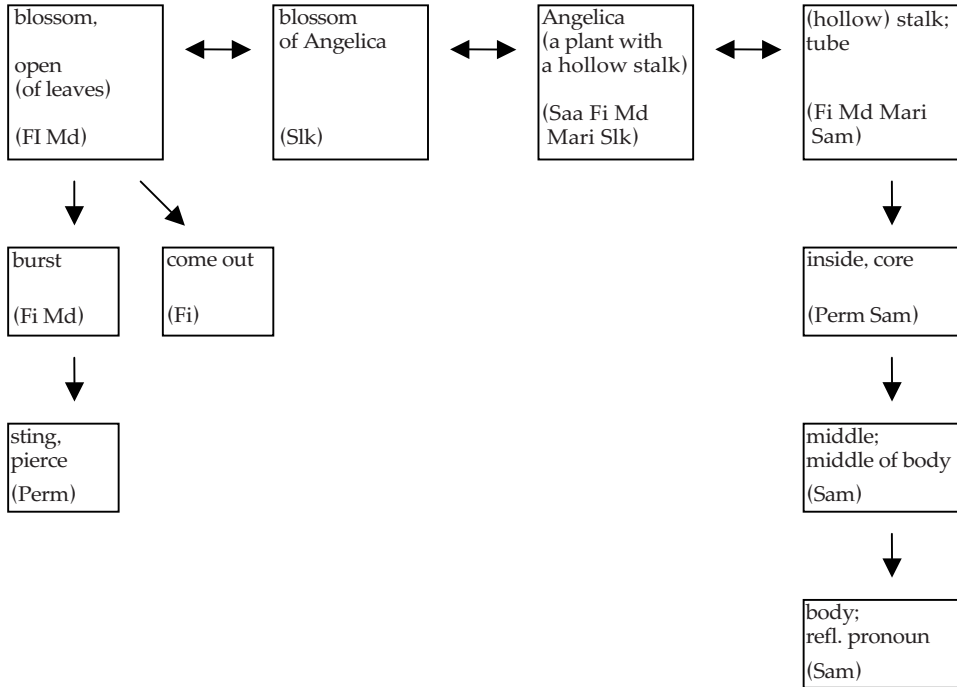


Figure 1. A rough scheme showing the semantic divergence of the words deriving from PU **pučki*. The subbranches and languages where each meaning is attested are given in parentheses.

čāž(i) 'family, tribe'. Sammallahti (1988 : 552) includes only the Mordvin, Mari and Permic words in the cognate set.

As the semantic connection of Fi *sataa* 'yield harvest' and *sato* 'harvest' to PFi **sata-* 'rain; fall' is not obvious, the alternative connection of these words to PU **čac̣V-* needs to be reconsidered. Semantically the comparison is flawless, as MdM *šac̣ə-* means not only 'be born' but also 'grow' and 'yield harvest'; the derived noun *šac̣əma* means both 'birth' and 'harvest'. Also MariE *šoča-*, MariW *šača-* mean both 'be born' and 'grow (of plants), yield harvest'. Hence, the reference to semantic problems made in UEW (52) remains incomprehensible. The only problem in the etymology is phonological: Fi *s-* is not a normal reflex of PU **č-*.

However, the development **č-* > Fi *s-* is accounted for by the fact that the verb originally contained two identical affricates. A dissimilation of the affricates first took place (**čac̣a-* > **cača-*), after which there was a regular change **c* > Fi *s*. This dissimilation has one well-established parallel, Fi *setä* 'paternal uncle' < **cečä* < PU **čečä* (Janhunen 1981 : 225; Sammallahti 1988 : 536; cf. UEW 34). In this case there is also a Saami cognate exhibiting the same dissimilation: SaaN *čeahci* 'paternal uncle (younger than father)' < **cečä*; note the regular changes PU **č* > PSaa **c* and PU **c* > PSaa **č*. That the word for 'uncle' originally contained two identical affricates is evident from cognates in other branches: e.g., MdE *čiče* 'brother-in-law (one's sister's husband)', MariW *čəčə* 'maternal uncle', Udm *čuž-murt* 'uncle', MsE *šüş* 'uncle' (Ms *š* < PU **č*), SlkK *čiččə* 'uncle'. In Komi, however, a dissimilation

identical to Finnic and Saami took place: Komi *čož* 'maternal uncle'. This is a regular development found in several other words as well, including *čuž-* 'grow; be born'. Other examples include Komi *čužji-* ~ Udm *čijži-* 'kick', Komi *čuž* ~ Udm *čužjem* 'malt', and Komi *čež* ~ Udm *čež* 'wild duck'.

Moreover, in Saami yet a third example of the same dissimilation has been discovered: SaaN *čoska* 'block of wood' < **čučki* < **čučki*, cognate with MdE *čočko* 'timber, log' (Luobbal Sámmol Sámmol Ánte (Aikio) 2013 : 164–165). The dissimilation **č—č* > **ć—č* can be considered a regular sound change in Finnic and Saami, which removes any problem in connecting Fi *sataa* 'yield harvest' and *sato* 'harvest' to PU **čača-* 'grow'. From a semantic perspective it is obviously much more plausible to analyze these words as cognate with MdM *šača-*, MariW *šača-* 'be born; grow, yield harvest' and MdM *šačama* 'birth; harvest' instead of connecting them with a verb meaning 'rain' and 'fall'.

The suggested Ob-Ugric and Samoyed cognates, which are not cited by Sammallahti (1988 : 552), are also worth closer examination. The inclusion of NenT *tenc^a* 'kind, sort; tribe', EnF *tjz* and Ngan *tansə* 'tribe, clan' (UEW 52) in the cognate set is obviously incorrect, as they presuppose a PSam form **tensə* or **censa*; the cluster **-ns-* does not match PU **-č-*, but would instead presuppose PU **-nś-* (cf. PU **kunśi-* 'urine' > PSam **kunsə*; Janhunen 1981 : 236). However, Slk **čāči* would continue a PSam form **cacV-*, which matches PU **čača-* as far as consonantism is concerned. The same is true of the KhS *čačə*, KhN *šoša*, *šoši*, *sasi* 'local, native' (< PKh **čičəγ*) and MsW *šošiγ*, MsE *sāsə*, MsN *sossa* 'homestead, homeland' (< PMs **šəšγā*); in Mansi there was a regular sound change **č* > **š*. The vowel correspondences require closer scrutiny, however.

Sammallahti (1988 : 552) reconstructs the verb as Finno-Permic **čęčV-*, which corresponds to PU **čičV-* in the present notation. The reason for not reconstructing the vowel **i* is the Komi cognate *čuž-*. According to Sammallahti's theory of Permic historical vocalism PU **a* developed to PPerm **u* (> Komi *o*, Udm *u*), whereas PU **i* is reflected as PPerm **ü* (> Komi and Udm *u*). In his framework the reconstruction **čičV-* is incompatible with the Khanty, Mansi and Selkup forms, as PU **i* should yield PMs **i*, PKh **ā* and PSam **i* or **e* (Sammallahti 1988 : 484, 504). However, recently Reshetnikov and Zhivlov (2011) have suggested that **a* and **i* in Janhunen's (1981) and Sammallahti's (1988) system of PU vocalism can actually be identified as a single phoneme (**a*), and its varying reflexes in Permic, Mansi and Samoyed can be explained by conditioned sound changes. As for Permic, they assume that there was vowel a shift PU **a*(*-a*) > PPerm **u* before palatalized consonants and the cluster **-rj-*, whereas a shift **a*(*-a*) > PPerm **ü* occurred before other types of consonants. Reshetnikov's and Zhivlov's hypotheses regarding Mansi and Samoyed historical vocalism cannot be assessed here, but their sound laws appear to account for the Permic data, and thus Permic does not seem to offer evidence for reconstructing the opposition **a* : **i*. Hence, the verb root can be reconstructed as PU **čača-* instead of **čiča-*, which allows also Slk **čāči* 'family' to be included in the cognate set. Semantically the Selkup form comes close to the Udm cognate *čijži-vjži* 'relative'.

The vocalism in the Ob-Ugric forms (PKh **čičəγ*, PMs **šəšγā*) is less clear. As PKh **i* is the high ablaut grade of original **a* (Helimski 2001; ЖИВЛОВ 2006 : 42), the underlying Ob-Ugric vowel correspondence is PKh

**a* ~ PMs **ɔ*. According to Sammallahti (1988 : 504) the Uralic source of this correspondence is roots of the shape *(C)*uCa*- and *(C)*iCa*-, whereas the reflexes of PU *(C)*aCa*-roots display the vowel correspondence PKh **ā* ~ PMs **ū*. Hence, these forms are irregular, and they must be excluded from the etymology because there is another Ob-Ugric word family which shows both phonologically and semantically a better match with the reconstructed Proto-Uralic verb **čača*-: MsN *sūs*- ~ *sūns*- 'grow, increase (intr.); have cubs (of a bear)' (< PMs **šū(n)š*-), MsN *sūsm*-, MsW *šūsm*- 'grow, increase (intr.)' (< PMs **šūšm*-), MsN *sūst*-, MsW *šūšt*-, MsS *sōst*- 'grow, increase (tr.)' (< PMs **šūšt*-), MsN *sūsam*, MsE *šūšam* 'one-year old bear cub' (< PMs **šūšām*), KhE *čjčj̄m*, MsN *ššam* 'bear cub in its first year' (< PKh **čjčj̄m*). The PKh vowel **ī* is the high ablaut grade of **ā*, so the underlying vowel correspondence is PKh **ā* ~ PMs **ū*, exactly as expected in the reflex of PU **čača*-. The only irregularity is the unexpected nasal in MsN *sūs*- ~ *sūns*-, but this is in all likelihood a secondary development: the verb has analogically acquired the morphophonological alteration *-s* : *-ns*-, which developed in roots with an original cluster **-nč*- (Honti 1999 : 49–51). The secondary origin of *n* is also evident from the lack of the nasal in all the derived forms as well as in the Khanty cognate.

Address

Ante Aikio
University of Oulu
E-mail: ante.aikio@oulu.fi

Abbreviations

EnF — Forest Enets; **Fi** — Finnish; **Hung** — Hungarian; **Kam** — Kamas; **KhE** — East Khanty; **KhN** — North Khanty; **KhS** — South Khanty; **MariE** — East Mari; **MariW** — West Mari; **Mat** — Mator; **Md** — Mordvin; **MdE** — Erzya Mordvin; **MdM** — Moksha Mordvin; **Ms** — Mansi; **MsE** — East Mansi; **MsN** — North Mansi; **MsS** — South Mansi; **MsW** — West Mansi; **NenF** — Forest Nenets; **NenT** — Tundra Nenets; **Ngan** — Nganasan; **Pfi** — Proto-Finnic; **PKh** — Proto-Khanty; **PMari** — Proto-Mari; **Pmd** — Proto-Mordvin; **Pms** — Proto-Mansi; **PPerm** — Proto-Permic; **Pre-Pfi** — Pre-Proto-Finnic; **Pre-Pmd** — Pre-Proto-Mordvin; **Pre-PSaa** — Pre-Proto-Saami; **PSaa** — Proto-Saami; **PSam** — Proto-Samoyed; **PSik** — Proto-Selkup; **PU** — Proto-Uralic; **SaaI** — Inari Saami; **SaaL** — Lule Saami; **SaaN** — North Saami; **SaaS** — South Saami; **SaaSk** — Skolt Saami; **SaaT** — Ter Saami; **Sik** — (Proto-)Selkup; **SikK** — Ket Selkup; **SikTa** — Taz Selkup; **SikTy** — Tym Selkup; **Udm** — Udmurt.

SOURCES OF LEXICAL DATA

- Alatalo, J. 2004, Sölkupisches Wörterbuch aus Aufzeichnungen von Kai Donner, U. T. Sirelius und Jarmo Alatalo, Helsinki (LSFU XXX).
- Benkő, L. 1992–1997, Etymologisches Wörterbuch des Ungarischen, Budapest.
- Bergsland, K., Mattsson Magga, L. 1993, Åarjelsaemien-daaroen baakoegærja — Sydsamisk-norsk ordbok, [Lakselv].
- Fokósfuchs, D. R. 1959, Syrjänisches Wörterbuch, Budapest.
- Grundström, H. 1946–1954, Lulelappsk ordbok, Uppsala (Skrifter utgivna genom dialekt- och folkminnesarkivet i Uppsala. Ser. C:1).
- H. Paasonen Mordwinisches Wörterbuch. Zusammengestellt von Kaino Heikkilä. Bearbeitet und herausgegeben von Martti Kahla, Helsinki 1990–1999 (LSFU XXIII. Kotimaisten kielten tutkimuskeskuksen julkaisuja 59).

- H e l i m s k i, E. 1997, Die matorische Sprache. Wörterverzeichnis, Grundzüge der Grammatik, Sprachgeschichte. Unter Mitarbeit von Beáta Nagy, Szeged (Studia Uralo-Altaica 41).
- H e r r a l a, E., F e o k t i s t o v, A. 1998, Mokšalais-suomalainen sanakirja, Turku (Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 58).
- I t k o n e n, E. 1986–1991, Inarilappisches Wörterbuch, Helsinki (LSFU XX).
- I t k o n e n, T. I. 1958, Koltan- ja kuolanlapin sanakirja, Helsinki (LSFU XV).
- Kai Donners Kamassisches Wörterbuch nebst Sprachproben und Hauptzügen der Grammatik. Bearbeitet und herausgegeben von A. J. Joki, Helsinki 1944 (LSFU VIII).
- L e h t i s a l o, T. 1956, Juraksamojedisches Wörterbuch, Helsinki (LSFU XIII).
- M a k s i m o v, S., D a n i l o v, V., S a a r i n e n, S. 2008, Udmurtilais-suomalainen sanakirja, Turku (Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 79).
- M i k o l a, T. 1985, Morphologisches Wörterbuch des Enzischen, Szeged (Studia Uralo-Altaica 36).
- M o i s i o, A., S a a r i n e n, S. 2008, Tscheremissisches Wörterbuch, Helsinki (LSFU 32).
- N i e m i, J., M o s i n, M. 1995, Ersäläis-suomalainen sanakirja, Turku (Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 48).
- O r s z á g h, L., M a g a y, T. 2006, Angol-magyar nagyszótár, Budapest.
- S a m m a l l a h t i, P., M o r o t t a j a, M. 1993, Säämi-suomâ sänikirje — Inarin-saamelais-suomalainen sanakirja, Ohcejohka.
- S a m m a l l a h t i, P. 1989, Sámi-suoma sátnegirji — Saamelais-suomalainen sanakirja, Ohcejohka.
- S t e i n i t z, W. 1966–1993, Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache 1–15, Berlin (= DEWOS).
- Suomen sanojen alkuperä. Etymologinen sanakirja 1–3, Helsinki 1992–2000 (SKST 556; Kotimaisten kielten tutkimuskeskuksen julkaisuja 62) (= SSA).
- U o t i l a, T. E. 1942, Syrjänischer Wortschatz nebst Hauptzügen der Formenlehre. Aufgezeichnet von Yrjö Wichmann, Helsinki (LSFU VII).
- Wogulisches Wörterbuch. Gesammelt von Bernát Munkácsi, geordnet, bearbeitet und herausgegeben von Béla Kálmán, Budapest 1986.
- Wotjakischer Wortschatz. Aufgezeichnet von Yrjö Wichmann. Bearbeitet von T. E. Uotila & Mikko Korhonen. Herausgegeben von Mikko Korhonen, Helsinki 1987 (LSFU 21).
- Б ы к о н я В. В. 2005, Селькупско-русский диалектный словарь, Томск.
- К о с т е р к и н Н. Т., М о м д е А. Ч., Ж д а н о в а Т. Ю. 2001, Нгансанско-русский и русско-нгансанский словарь, Санкт-Петербург.
- С о р о к и н а И. П., Б о л и н а Д. С. 2001, Энецко-русский и русско-энецкий словарь, Санкт-Петербург.
- Т е р е щ е н к о Н. М. 1965, Ненецко-русский словарь, Москва.
- Х е л и м с к и й Е. 2007 [unpublished manuscript], Северноселькупский словарь.

R E F E R E N C E S

- H e l i m s k i, E. 2001, Ablaut als Umlaut im Ostjakischen: Prinzipien und Grundzüge der lautgeschichtlichen Betrachtung. — Fremd und eigen. Untersuchungen zu Grammatik und Wortschatz des Uralischen und Indogermanischen in memoriam Hartmut Katz, Wien, 55–76.
- H o n t i, L. 1982, Geschichte des obugrischen Vokalismus der ersten Silbe, Budapest.
— — 1999, Az obi-ugor konzsonantizmus története, Szeged (Studia Uralo-Altaica Supplementum 9).
- I t k o n e n, E. 1946, Zur Frage nach der Entwicklung des Vokalismus der ersten Silbe in den finnisch-ugrischen Sprachen, insbesondere im Mordwinischen. — FUF XXIX, 222–337.
- J a n h u n e n, J. 1977, Samojedischer Wortschatz. Gemeinsamojedische Etymologien, Helsinki (Castrenianumin toimitteita 17).
— — 1981, Uralilaisen kantakielen sanastosta. — JSFOu 77, 219–274.

- Koivulehto, J. 1999, Varhaiset indoeurooppalaiskontaktit: aika ja paikka lainasanojen valossa. — Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan, Helsinki (Bidrag till kännedom av Finlands natur och folk 153), 207–236.
- Kylstra, A. D., Hahmo, S.-L., Hofstra, T., Nikkilä, O. 1991–2012, Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen, Amsterdam—Atlanta (= LÄGLOS).
- Liimola, M. 1956, Etymologische bemerkungen. — FUF XXXII, 226–264.
- Luobbal Sámmol Sámmol Ánte (Aikio, A.) 2012, On Finnic Long Vowels, Samoyed Vowel Sequences, and Proto-Uralic *x. — Per Urales ad Orientem. Iter polyphonicum multilingue. Festschrift tillägnad Juha Janhunen på hans sextioårsdag den 12 februari 2012, Helsinki (MSFOu 264), 227–250.
- 2013, Studies in Uralic Etymology I: Saami Etymologies. — LU XLIX, 161–174.
- Ravila, P. 1938, Über die entstehung des tscheremissischen konjugationssystems. — FUF XXV, 1–25.
- Reshetnikov, K., Zhivlov, M. 2011, Studies in Uralic Vocalism II. Reflexes of Proto-Uralic *a in Samoyed, Mansi and Permic. — Journal of Language Relationship / Вопросы языкового родства 5, 96–109.
- Sammallahti, P. 1979, Über die Laut- und Morphemstruktur der uralischen Grundsprache. — FUF XLIII, 22–66.
- 1988, Historical Phonology of the Uralic Languages with Special Reference to Samoyed, Ugric and Permic. — The Uralic languages. Description, History and Foreign Influences, Leiden—New York—København—Köln, 478–554.
- Setälä, E. N 1902, Zur finnisch-ugrischen lautlehre. — FUF II, 219–280.
- Toivonen, Y. H. 1928, Zur geschichte der finnisch-ugrischen inlautenden affrikaten. — FUF XIX, 1–270.
- 1956, Über die syrjänischen Lehnwörter im Ostjakischen. — FUF XXXII, 1–169.
- Живлов М. 2006 [unpublished], Реконструкция праобско-угорского вокализма. Диссертация на соискание ученой степени кандидата филологических наук, Москва.

ЛУОББАЛ САММОЛ САММОЛ АНТЕ (АНТЕ АЙКИО) (Оулу)

ИССЛЕДОВАНИЯ УРАЛЬСКИХ ЭТИМОЛОГИЙ II. ПРИБАЛТИЙСКО-ФИНСКИЕ ЭТИМОЛОГИИ

Статья представляет собой вторую часть из серии исследований, в которых сопоставляются новые этимологии слов из разных уральских языков. В статьях этой серии рассматривается лексика языков и диалектов уральской языковой семьи, предлагаются новые соответствия уже известных этимологий, а также совершенно новые этимологии. В этой части исследования автор знакомит с новыми уральскими этимологиями для следующих финских слов: *aita* 'изгородь' (< праур. **ajta*), *ammottaa* 'быть настежь открытым' (< праур. **ammV-* 'зевать'), *kaiho* 'тоска' (< праур. **kajšV* 'болезнь, беда'), *katkera* 'горький', *katku* 'угар, запах горелого, запах паленого', *katketa* 'сломаться, прерваться' (< праур. **kačka-* 'кусать'), *korpi* 'бор, дремучий лес' (< праур. **korpi*), *ohut* 'тонкий' (< праур. **wokši*), *puhjeta* 'разразиться; раскрыться (о бутонах цветов)', *putkahtaa* 'объявиться, выйти, обнаружиться' (< праур. **pučki* 'полюй стебель, труба') и *sato* 'урожай', (диал.) *sataa* 'давать урожай' (< праур. **čāca-* 'расти'). Принципы реконструкции и выбора лексического материала изложены в первой статье данной серии (Luobbal Sámmol Sámmol Ánte (Ánte Aikio) 2013).