

Dialect materials in the Estonian Etymological Dictionary¹

IRIS METSMÄGI

*Institute of the Estonian Language, Roosikrantsi 6, Tallinn, Estonia,
iris.metsmagi@eki.ee*

SCN III/1 [2010], 196–204

Članek predstavlja estonsko dialektološko gradivo, ki ga je zbral Inštitut za estonski jezik in bo objavljeno v *Estonskem etimološkem slovarju*.

V seznam iztočnic slovarja bo vključeno omejeno število narečnih besed – narečni podatki bodo pomembni tudi za etimologizacijo. V primeru avtohtonih besed in starejših izposojenk so pomembne arhaične fonetične poteze, ki so preživele v narečju; v primeru novejših izposojenk kažejo narečne variante različno stopnjo prilagajanja. Včasih lahko arealna distribucija besede dokaže dragoceno izvorno iztočnico. Etimologizacija lahko temelji tudi na narečnem pomenu besede. Posebno skupino sestavljajo narečne besede, prevzete v knjižno rabo na drugačen način kot strokovno izrazje.

This article is about how Estonian dialect materials are represented in the Estonian Etymological Dictionary being compiled at the Institute of the Estonian Language.

A limited number of dialect words will be included in the headword list of the dictionary. Dialect data may be vital for etymologization as well. In the case of genuine words and older loanwords, the archaic phonological traits that have survived in dialects are essential; in the case of more recent loanwords, the dialectal variants show different degrees of adaptation. Sometimes the areal distribution of a word may prove a valuable cue to its original background. Etymologization may also be based on the dialectal meaning of a word. A specific group consists of dialect words adopted into standard usage in a different sense as technical terms.

Ključne besede: etimološki slovar, estonski jezik, estonska narečja, zgodovinska fonologija, prevzemanje besed, prostorska distribucija besed, jezikovno načrtovanje

Key words: etymological dictionary, Estonian language and dialects, historical phonology, adaptation of loanwords, areal distribution of words, language planning

¹ This article is supported by grant No 7717 of the Estonian Science Foundation.

The Estonian Etymological Dictionary has been a project of the Institute of the Estonian Language since 2003. Preliminary work started in the 1970s already. This is not the first attempt. There are two etymological dictionaries of Estonian available: 1) Julius Mägiste's *Estnisches etymologisches Wörterbuch* (1982–1983, 2. edition 2000; 12 volumes, 4106 pp.; EEW), and 2) Alo Raun's *Eesti keele etimoloogiline teatmik /Etymological reference book of Estonian/* (1982, 2. edition 2000; 222 pp.), both compiled by the authors in exile (they left Estonia during the Second World War) and published abroad. As Mägiste's dictionary is actually a manuscript, published as a facsimile after minor editing (it remained unfinished by the author due to his death) and Raun's dictionary is extremely laconic – just one line per each headword (stem), the need for a new and modern dictionary has been obvious.

Due to the urgent necessity for the dictionary it was decided to start from a short and not too detailed version tailored for a general public. The next step involved concrete decisions on the material to be included. Omitting a thorough survey of the compilation principles worked out for the new dictionary at our Institute, this article will concentrate on how dialect material will be represented in it. The examples presented here will not contain whole dictionary entries as in a shortened form they will much better illustrate what is essential for this article. The example entries are translated into English; the metalanguage of the dictionary is naturally Estonian.

In the dictionary in question dialect material is encountered in 1) headword selection as well as in 2) etymologization.

1 Headword selection

The headword list of our etymological dictionary (over 6000 stem words) is based on the word stems contained in *Eesti õigekeelsussõnaraamat ÕS 2006 /Estonian Orthographical Dictionary/*. Obviously, most of the stems represent standard Estonian, but some dialect words can also be found there, e.g. *kaarel* 'cloudberry', a word spread in the western and central dialect of the North Estonian dialect group (VMS 1: 146), *nugris* 'active, enterprising', known in the same dialects (VMS 2: 104), *raand* '(wooden) pail', known in the insular and western dialect of the North Estonian dialect group (VMS 2: 298) etc. In addition, the etymological dictionary handles such stems that in standard language occur in compound words only, but are known as separate words in dialect use, e.g. the adjective *raag* 'underdone; half-ripe' occurs only in the compound noun *raagnahk* 'alum-tanned leather, greased after drying' in standard Estonian, whereas in the North-Eastern Coastal dialect group and in the central dialect of the North Estonian dialect group it is known as a word in its own right (VMS 2: 297).

2 Etymologization

In etymologization we have to do with dialectal variants, geographic distribution and dialectal meanings.

2.1 Dialectal variants

Dialectal variants are placed in a separate part of an entry. The general idea is to present only those variants that are essential for the etymology, while all possible variants are taken care of in a dialect dictionary.

In the case of genuine words² and older loanwords³ one may have to consider the regular phonological specifics of certain dialects. The Estonian language is divided into at least two dialect groups: North Estonian and South Estonian, while sometimes the North-Eastern Coastal dialect group is pointed out as a third (Pajusalu 2003: 231, 239–260). The literary standard is based on the central dialect of North Estonian. The phonology of the South Estonian dialect group differs considerably from that of the North Estonian one. On the one hand it has retained some old features lacking in North Estonian, on the other hand it bears traces of some sound changes that cannot be found in any other Finnic language (Viitso 2003: 143–144). Some very archaic traits have also survived in the North-Eastern Coastal dialect group. Thus, in many cases South Estonian or North-Eastern Coastal variants can be used to confirm an etymology or to refer the reader to some other important circumstance. Therefore the genuine stems and older loanword stems are provided with their South Estonian or North-Eastern Coastal dialect variant, if the variant has certain phonological traits.

Example 1

kaevama ‘to dig’

In dialects: South Estonian *kaibma*, *kaivma*, North-Eastern Coastal *kaivama*

Cognate equivalents:

Livonian *kouvõ* ‘dig; gore, push with horns’

Vote *kaivaa* ‘dig; hollow out; gore, push with horns; tear out; tear off’

Finnish *kaivaa* ‘dig; grub; hollow out, furrow; bury’

Ingrian *kaivaa* ‘dig; hollow out; drag out (e.g. a horse); remove, take out’

Olonets Karelian *kaivua* ‘dig, grub; hollow out’

Lude *kaivada* ‘dig; bore, hole’

Veps *kaida*, *keida* ‘dig, grub; take out’

Sami *goaivut* ‘scoop; dig, scrape (snow)’

Erzya *kajams* ‘throw; hit; put; pour; take off (clothes); give birth prematurely’

Moksha *kajams* ‘throw, sow; hit; pour; take off (clothes)’

Mari *kuaš* ‘shovel; row’

² I.e. the native Uralic or Finnic words.

³ I.e. the loanwords which have etymological equivalents in other Uralic resp. Finnic languages – the Indo-European, Indo-Aryan, Baltic, Germanic and Slavic loanwords.

Udmurt *kwjanj* ‘throw; leave; expel, fire, overthrow’
 Komi *kojnĵ* ‘scoop, pour out; fill up, slop’
 ? Hungarian *hajít* ‘throw’
 A genuine Finno-Ugric stem.

The South Estonian and North-Eastern Coastal dialects have retained the original diphthong *ai* in the first syllable, still manifested in cognate languages. (In standard Estonian *ai* > *ae*.) The dialectal variants are from EMS 2: 489. The etymology is based on SKES 1: 144, EEW 2: 642, SSA 1: 279, Raun 1982: 26, Häkkinen 2004: 323.

Example 2

nael ‘nail (fastener)’

In dialects: South Estonian *nagõl*, *nakl*, North-Eastern Coastal *naul(a)*
 < Proto-Germanic **naglan-* → German *nagel* ‘nail; fingernail’

Cognate equivalents:

Livonian *na'ggõl* ‘nail’

Votic *naġla* ‘nail; peg; cotter, chock’

Finnish *naula* ‘nail; peg; cotter, chock’

Ingrian *naagla* ‘nail; cotter, chock’

Olonets Karelian *nuaglu* ‘nail; cotter, chock’

Lude *nagl(e)* ‘nail; peg; cotter, chock’

Veps *nagl* ‘nail; stake; cotter, chock’

As we can see, South Estonian has retained the original stop found in the donor language and still in most of the cognate languages. The dialectal variants are from VMS 2: 72, 80; Iva 2002: 260). The etymology is based on SKES 2: 369, EEW 5: 1654–1655, SSA 2: 209, Raun 1982: 100, Häkkinen 2004: 775.

Example 3

kahv ‘a bag net attached to a long pole for catching crayfish and fish’

In dialects: (part of) South Estonian *kauh*

< Baltic **kauša* → Lithuanian *kaušas* ‘large dipper; wooden drinking vessel’, Latvian *kauss* ‘cup; goblet; dipper’

Cognate languages:

Votic *kavi* ‘dipper’

Finnish *kauha* ‘ladle; dipper; piggin’

Ingrian *kaavi* ‘ladle; dipper, liquid scoop’

Olonets Karelian *kawhu* ‘dipper, liquid scoop’

Lude *kauh* ‘dipper, liquid scoop; bag net for scooping fish’

Veps *kauh*, *kouh* ‘dipper, liquid scoop’

Julius Mägiste suggested Estonian *kahv* as metathetical (*kahva-* ← **kauha-*), belonging, etymologically, with the above-mentioned Baltic word loaned to Finnic languages, although he did not mention the South Estonian variant *kauh* (EEW 2: 654). The metathesis of *h* after a diphthong is regular in the rest of the Estonian dialects. The form *kauh*, surviving in a very narrow strip of the South Estonian dialect area (EMS 2: 532), proves that Mägiste was right. The

Baltic etymology of Finnic **kauha* resp. Finnish *kauha* is based on SKES 1: 172, SSA 1: 330, Häkkinen 2004: 381; in none of them is the Estonian *kahv* handled as a possible equivalent of that stem.

Irregular stem variation in dialects is important, first and foremost, for more recent loans,⁴ showing different degrees of adaptation. A dialect word may actually be closer to the source than the standard variant. Therefore in the dictionary more recent loans are provided with dialect variants.

Example 4

kapsas ‘cabbage’

In dialects: *kapust(as)*, *kapst(ass)*, *kapus*, *kaapsus*, *kaapsas*

< Russian *kapústa*

The dialectal variants are from EMS 2: 712, 715, 717; the etymology is based on EEW 3: 700; Raun 1982: 31; Must 2000: 79–80; Blokland 2005: 143–144.

Example 5

kartul ‘potato’

In dialects: *kartuhvel*, *karduhbel*, *kardulas*, *kardules*, *kardulis*, *kardul*, *kartles*, *tuhlis*, *tuhl(i)*, *tohris*, *toh(v)el*, *tih(v)el*

< German *Kartoffel*

The dialectal variants are from EMS 2: 768, 771, 772 and VMS 2: 553; the etymology is based on EEW 3: 719.

2.2 Geographic distribution

The geographic distribution of a word may be another important aspect in etymologization. Some Estonian loanwords come from mutually closely related languages. For instance, there are loans from different Germanic languages such as Low German, (High) German and Swedish; in addition to these, loans from a German dialect spoken once in Estonia and Latvia (Baltic German) and from a Swedish dialect, once spoken by the Swedish population of the islands and the coast of Estonia (Estonian Swedish), are treated as separate loanword groups. Moreover, the possible donor languages have been borrowing from each other. Finnish, for example, has a lot of Swedish loans and so some originally Swedish words may have reached Estonian via Finnish, not directly from Swedish. In cases like these, dialectal (areal) distribution may be an important or even the only criterion to decide the source of borrowing. If the dialectal distribution of a word is etymologically important, it is discussed in the commentary to the entry.

⁴ I.e. loans from Low German, (High) German, Russian, Swedish, Latvian and Finnish.

Example 6

traksid ‘braces’

← Baltic German *tragsen* ‘brace, strap’

or

← Swedish (dialectal) *dragsa, drakksu* ‘tug’

As the word is known all over the Estonian territory the Baltic German version looks more likely.

The word is etymologized as a Swedish loan in SKES 3: 726, EEW 10: 3261, SSA 3: 44; the Baltic German etymology is preferred by Paul Ariste (1933: 11–12) and repeated in Raun 1982: 181.

Example 7

risla ‘dunlin (a coastal bird), *Calidris*’

← Finnish (dialectal) *risla* ‘black guillemot, *Cepphus grylle*’

or

← Swedish *grissla* ‘guillemot, *Uria*’

Finnish *risla* is a Swedish loan (← *grissla*). As *risla* is a word form characteristic only of the Estonian North-Eastern Coastal dialect group and of some Finnish dialects spoken on the coast of the Gulf of Finland it is likely that *risla* has been borrowed from Finnish dialects rather than directly from Swedish.

The Swedish etymology is given for the word in SKES 3: 811–812, while a loan from Finnish is supposed in EEW 8: 2500–2501; this should be preferred, moreover as the same word is borrowed into Estonian from Estonian Swedish dialects in the form *kriüsel* ‘*Cepphus grylle*’ (EEW 3: 1010).

2.3 Dialectal meanings

In many cases word meaning differs widely between standard and dialect use. In the standard language a dialect word may come into use in some peripheral meaning; also, its meaning may change over time either spontaneously or due to language planning. In the early decades of the 20th century, standard Estonian vocabulary was consciously enriched with dialect material. Many dialect words were adopted into use to refer to concepts felt to be lacking proper designation in the standard language, while their meanings were handled rather liberally in terms of enlargement or even change. A number of dialect words were used to build up scientific terminology in Estonian. Dialect stems, in a changed meaning, have been adopted into the standard language later as well. Such stems can be etymologized only on the basis of their dialectal meaning – hence the dire necessity to represent dialectal senses for quite a number of words. Word definitions, together with the relevant dialectal meanings are, in general, given at the head of the entry.

Example 8

nugima ‘to live on/in an organism as a parasite, to parasitize; (in dialect) to attend weddings uninvited’

← Estonian Swedish *snwg* ‘to beg, to sponge off others’

The dialect word is adopted into the standard language, in a changed meaning as a biological term.

The etymology of the stem is based on EEW 6: 1736, Raun 1982: 105.

Example 9

tävi- (first component in compound words) ‘inflatable, to be filled up with gas or liquid’

Cognate equivalents:

Livonian *tä'u(v)*, *tä'ugõz* ‘lung’

Votic *tävü* id.

Finnish (dialect) *tävy*, *täy*, *täty* id.

Veps *tävu* id.

Erzya *t'ev'elav* id.

Moksha *t'evlav* id.

Udmurt *tj* id.

Komi *tj* id.

Nenets *t'iwāk* id.

Enets *t'ji* id.

Nganasan *t'ieind'a* id.

Selkup *tūmäktä* id.

Kamassian *tu* id.

(The Erzya, Moksha, Nenets, Enets and Nganasan equivalents contain derivational affixes; the Selkup equivalent is a compound word.)

A Uralic stem, preserved in dialects in the forms *täü*, *tähi*, *tähü*, *täht* ‘lung; spleen’.

The stem (in the form *tävi* of older records) was adopted into standard Estonian in the 1970s, giving it a new meaning. The verb *tävima* ‘to inflate or pump up, to fill with gas or liquid’ has been derived from the same stem.

This is a special case, because the stem is adopted into standard language in the form *tävi*, which, although registered in Wiedemann 1893, seems to be an artificially restored North Estonian shape, actually not found in dialects (Saari 1987: 53–55). The etymology of the stem is based on SKES 5: 1486, EEW 11: 3464, Raun 1982: 189, SSA 3: 358.

The presented overview suggests that the inclusion of dialect material in the Estonian Etymological Dictionary is undoubtedly necessary for two reasons: firstly, there is a number of dialect words, the etymology of which cannot be omitted, secondly, dialect material (dialectal phonetic variants (2.1), areal distribution of words (2.2) and dialectal meanings (2.3)) give a valuable complementary information for the etymologization.

REFERENCES

- Paul ARISTE, 1933: *Eesti-rootsi laensõnad eesti keeles. /Estonian Swedish loanwords in Estonian./* Tartu. (Acta et Commentationes Universitatis Tartuensis B 29, 3).
- Rogier BLOKLAND, 2005: *The Russian loanwords in literary Estonian.* Groningen. *Eesti õigekeelsussõnaraamat ÕS 2006. /Estonian orthographical dictionary./* Ed. Tiiu Erelt. Tallinn: Eesti Keele Instituut, Eesti Keele Sihtasutus. <http://www.eki.ee/dict/qs2006/>
- EEW = Julius MÄGISTE, 1982–1983: *Estnisches etymologisches Wörterbuch. 1–12.* Helsinki: Finnisch-Ugrische Gesellschaft.
- EMS = *Eesti murrete sõnaraamat. /Dictionary of Estonian dialects./ 1–4.* Tallinn: Eesti Teaduste Akadeemia. Eesti Keele Instituut. 1994–2008.
- Kaisa HÄKKINEN, 2004: *Nykysuomen etymologinen sanakirja. /Etymological dictionary of modern Finnish./* Helsinki: WSOY.
- Sulev IVA, 2002: *Võru-eesti sõnaraamat. /Võru [an Estonian dialect] – Estonian dictionary./* Tartu–Võru: Võru Instituut. (Võru Instituudi Toimetised 12).
- Mari MUST, 2000: *Vene laensõnad eesti murretes. /Russian loanwords in Estonian dialects./* Tallinn: Eesti Keele Sihtasutus.
- Karl PAJUSALU, 2003: Estonian Dialects. *Estonian Language.* Ed. Mati Erelt. Tallinn: Estonian Academy Publishers. (Linguistica Uralica supplementary series, 1). 231–272.
- Alo RAUN, 1982: *Eesti keele etimoloogiline teatmik. /Etymological reference book of Estonian./* Rooma–Toronto: Maarjamaa.
- Henn SAARI, 1987: *Murre tuleb kirjakeelde. /Dialect enriching the standard language./ Kirjakeel 1985.* Tallinn: Valgus. 9–65.
- SKES = Y. H. TOIVONEN et al., 1955–1981: *Suomen kielen etymologinen sanakirja. /Etymological dictionary of Finnish./ 1–7.* Helsinki: Suomalais-Ugrilainen Seura. (Lexica Societatis Fenno-Ugricae 12).
- SSA = *Suomen sanojen alkuperä. Etymologinen sanakirja. /The origin of Finnish words. An etymological dictionary./ 1–3.* Helsinki: Suomalaisen Kirjallisuuden Seura, Kotimaisten kielten tutkimuskeskus. (Suomalaisen Kirjallisuuden Seuran toimituksia 556. Kotimaisten kielten tutkimuskeskuksen julkaisuja 62). 1992–2000.
- Tiit-Rein VIITSO, 2003: Rise and Development of the Estonian Language. *Estonian Language.* Ed. Mati Erelt. Tallinn: Estonian Academy Publishers. (Linguistica Uralica supplementary series, 1). 130–230.
- VMS = *Väike murdesõnastik. /Concise dictionary of dialects./ 1–2.* Ed. Valdek Pall. Tallinn: Valgus. 1982–1989.
- Ferdinand Johann WIEDEMANN, 1893. *Ehstnisch-Deutsches Wörterbuch.* St.-Petersburg.

DIALEKTOLOŠKO GRADIVO V ESTONSKEM ETIMOLOŠKEM SLOVARJU

Seznam besed za *Estoniski etimološki slovar (Estonian Etymological Dictionary)*, ki ga je zbral Inštitut za estonski jezik, temelji na izvorihi besed, vsebovanih v *Eesti õigekeelsussõnaraamat/Estonian Orthographical Dictionary* (2006). Očitno večina besed izvorno pripada knjižni estonsčini, vendar pa je za etimološki slovar pomembno tudi narečno gradivo. (1) V slovar bo vključeno omejeno število narečnih besed, ki jih najdemo v ortološkem slovarju, in tudi tistih korenskih besed, ki se v knjižni rabi pojavljajo le v zloženkah. (2) V nekaterih primerih so lahko narečni podatki (narečne inačice, geografska distribucija in narečni pomeni) ključni za etimologizacijo. (2.1) Knjižni standard estonsčine temelji na severnoestonskem narečju. V primeru izvornih korenskih besed in starejših izposojenk lahko nekatere fonetične poteze, preživele v južnoestonskem narečju in v obalnih narečjih, uporabimo za potrditev etimologije. Zato so ti koreni opremljeni s svojo južnoestonsko ali obalno narečno inačico, npr. *nael* 'noht' (< proto-germansko **naglan-*): južnoestonsko *nagõl*, *nakl*, obalno *naul(a)*. Novejše izposojenke so opremljene z nepravilnimi narečnimi inačicami in tako kažejo različne stopnje adaptacije. Narečna beseda je lahko dejansko bližje izvoru kot knjižna ustreznica, npr. samostalnik *kapsas* 'zelje' < rus. *kapústa*, ki ima narečne inačice *kapust(as)*, *kapst(ass)*. (2.2) Včasih lahko arealna distribucija besede dokaže dragoceno izvorno iztočnico. Npr. za besedo *traksid* 'braces' so domnevali, da ima izvor bodisi v švedskih narečjih (< *dragsa*, *drakksu* 'tug') ali v baltsko-germanskem narečju (< *tragsen* 'brace, strap/zatega, trak'), glede na to, da je beseda baltsko-germanska inačica. (2.3) Etimologizacija lahko temelji tudi na narečnem pomenu besede. Posebna skupina je sestavljena iz narečnih besed, v knjižno rabo prevzetih z drugačnim pomenom (npr.: izvorno je sodobni biološki izraz *nugima* 'parazitirati' pomenil 'pritti na poroko nepovabljen'; koren je bil izposojen iz estonsko-švedskega *snõg* 'beračiti, izžemati drugega').