
Diachronic Analysis of Word Formation as Basic Studies for Etymologisation

Iris Metsmägi, Vilja Oja

Institute of the Estonian Language, Tallinn
e-mail: iris.metsmagi@eki.ee; vilja.oja@eki.ee

Abstract

The paper focuses on some problems of word formation in Estonian and Finnic, which are etymologically relevant. Diachronic derivational relations may not be ascertained by the synchronic approach, e.g. either due to sound changes in the stem (cf. Estonian *kõõl-us* ‘tendon, string’ and *keel* ‘tongue; language; string’), or because a root does not occur separately or is rare. A suffix may have been dropped from active use or is subject to extensive variation, e.g., many Estonian verb stems containing the component *-ka, -ki, -ku* etc. (*kilka-ma* ‘to scream’) could be interpreted as derivatives with respective suffixes (*kil-ka-ma*, cf. *kil-ise-ma* ‘to clink’). The interpretation of the morphological structure of loanwords is often complicated. The loanwords may contain derivational suffixes of the donor language, sometimes even resembling genuine ones, e.g. Estonian *lusikas*, Finnish *lusikka* ‘spoon’ does not contain historically the Finnic suffix *-k(k)a* but is borrowed from Old Russian *lŭžika*. Sometimes foreign-sounding loanwords are subjected to folk-etymological adaption and the results are similar to a derivative or a compound, e.g. Estonian *puuslik* ‘idol’, looking like a derivative of the stem *puu* ‘tree; timber’, is rather a folk-etymological version of Russian *apostol* ‘apostle’.

Keywords: etymology; derivation; loanwords; Estonian; Finnic languages

1 Introduction

As is well known, the comparison of a word stem with stems of cognate and contact languages is only one step in finding out the word origin. The morphological analysis of the word is also very important. The morphological structure of the derivatives, formed according to active and productive derivational models, is in most cases transparent enough. On the contrary, old derivational patterns fallen out of active use may have become obscured. The detection of diachronic developments, explanation of various exceptions etc. needs thorough research. It is necessary to ascertain: 1) which words share a root, 2) which structural elements can be regarded as old derivational suffixes, 3) which derivational suffixes are genuine and which ones were brought along by loanwords. The solutions of those problems are directly related to practical lexicography, e.g. to the structure of the entry list of a dictionary. Our paper focuses on some problems of word formation in Estonian, which are etymologically relevant. The examples include comparative material from other Finnic languages as well. Those problems were previously addressed e.g. by Valmen Hallap (1958 and 1983), Huno Rätsep (1992/2002), Osmo Nikkilä (1998).

2 Non-transparent Roots

Diachronic derivational relations are not transparent either due to sound changes in the stem, because a root does not occur separately, or is rare. For example, the word *vaatama* ‘to look, to watch’ is diachronically a derivative of the stem *valvama* ‘to guard, to watch’, reconstructed as **valva-tta-ma* (**-tta* is a causative-factitive suffix). In the old literary Estonian (1600), the variant *valatama* is registered, in which an irregular contraction of syllables has taken place (see EEW: 3610). The substantive *kõõlus* ‘tendon, sinew’; dial. ‘lace, string, cord’ is probably derived from the noun *keel* ‘tongue; language; string’. Sporadic velarisation *ee* > *õõ* has taken place in the derivative, cf. Finnish *kielus* ‘string for fastening or tightening a snare’ < *kieli* ‘tongue; language; string’ (see SKES: 188, EEW: 1116, SMS 7: 44). In 19th century Estonian, the variant with regular stem *keelus* is recorded as well (Wiedemann 1973/1893: 271–272).

The adjective *kindel* ‘firm; sure; safe’ (< **kīnte-lä*) is an old derivative of the same stem which occurs in the adverb *kinni* ‘shut, closed; fast, fixed; occupied, engaged, busy’ (< probably an old lative form **kīnti-k*, cf. Estonian dial. *kindi*, Finnish *kiinni*, Lude *kiñd’i*) (see EEW: 831, 835). Assimilation of the consonant cluster has taken place in the adverb. The root does not occur as a separate word.

The Estonian verb *kastma* ‘to water, to wet’ (< **kas-ta-ma*, cf. Finnish *kastaa* id.) is derived with a causative-factitive suffix from the rare dialectal word *kasi* ‘dampness in soil (in spring); dew’ (EEW: 725, SSA 1: 323–324).

3 Obsolete Suffixes

Diachronic derivatives are not ascertained by the synchronic approach, because a suffix has either fallen out of active use or is subject to extensive variation. The fact that the word is an old derivative is indicated by the parallel occurrence of words with a similar or close meaning, but with different components interpretable as a suffix, and a noticeable amount of such word pairs. The relevant material of cognate languages can be used to contribute to suffix identification. The Finnish nouns (and some examples of other Finnic languages) were thoroughly analysed from this point of view by Osmo Nikkilä (1998; note also the examples of earlier treatment of the issues, mentioned by him).

Estonian adjectives with the ending *-eda* in genitive may attract the derivative suffix *-g/-k-* (< **-k-*). Sometimes the stems of the same adjectives may attract the suffix *-k* (< **-kka/-kkä*, *-kk-V*) (Hallap 1983). Similar suffixes are used in Finnish and other cognate languages (see Nikkilä 1998: 87–90), e.g. Estonian *hele* (**hel-edä*) ‘bright, light; high (voice)’, *helge* (**hel-k-edä*) ‘bright, sunny; cheerful, happy’ and *hel-k* (GEN *helga*) ‘high, acute (voice)’, cf. Finnish *heleä* ‘bright, light; high (voice)’ and *helkkä* id.; Estonian *tahe* (**tah-edä*) ‘dry, dried; salty; mealy (potato)’ and *tahke* (**tah-k-edä*) ‘solid; dry, hard’, cf. Finnish dial. *tahea* ‘mealy’, *tahkea* id.; ‘viscous, sticky’ (see also Nikkilä 1998: 83; SSA 3: 251); Estonian dial. *müre* (**mür-edä*) ‘half-sour (milk)’ and *mürk* (**mür-kkü*) ‘poison’, cf. Finnish *myreä* ‘half-sour, strange-tasting’ and *myrkky* ‘poison’. Many Finnic adjectives denoting lightness of tone, originate from the stem *val-* or *valk-* (*valg-*), where *-k-* may be regarded as an old suffix, e.g. Estonian *valge* (**val-k-edä*), Finnish *valkoinen* ‘white’ cf. Estonian dial. *valss* and others (see e.g. Oja 1996: 95), Finnish *vaalea* ‘light’, and others.

Similar phenomenon occurs in the verbs as well. In Estonian, for example, there are a great number of (mainly) onomatopoeic verbs containing the component *-ka*, *-ki*, *-ku*, *-ga* *-gi*, *-gu*. There are also

verb stems that are semantically similar or close to the former, but do not contain the above-mentioned component, e.g. *kil-ka-ma* ‘to scream’ and *kil-ise-ma* ‘to clink’, *kol-ki-ma* ‘to beat (with loud noise)’ and *kol-ise-ma* ‘to rumble’, *nõr-gu-ma* ‘to drip’ and *nõr-ise-ma* ‘to trickle’. Although a synchronic approach points out the verb suffix *-ki*, *-ku*, *-gi*, *-gu*, it classifies the suffix as unproductive that derives verbs only from interjections (with few exceptions), e.g. *mää-gi-ma* ‘to bleat’ < *mää!* ‘baa!’, *oh-ki-ma* ‘groan and moan’ < *oh!*, *näü-gu-ma* ‘to mew’ < *näü!* ‘miaow!’, the exceptions including *kär-ki-ma* ‘shout at sb.’ < *kära* ‘noise’, *kär-ama* ‘to scold’ and even *tõr-ku-ma* ‘to refuse’ < *tõre* ‘curt’. Thus the structural analysis of modern Estonian vocabulary does not make a connection between *kilkama* and *kilisema* or *kolkima* and *kolisema*, and does not discern derivational suffixes in *kilkama* and *kolkima* (see Vare 2012: 579–580, 655). Nevertheless, those words must be derivatives, at least from the historical perspective, including old suffixes *-ka-* (< **-kka-*, **-kkä-*), *-ki-*, *-ku-* (< **-kk-i-*, *-kk-ü-*, *-kk-u-*) or *-ga-* (< **-ka-*, **-kã-*), *-gi-*, *-gu-* (< **-k-i-*, **-k-ü-*, **-k-u-*), once rather productive ones, which have equivalents in other Finnic languages and even in some more distantly cognate languages (see Hallap 1958: 175–183), cf. Est *kilkama* and Finnish *kilkkaa*, *kilkata* ‘clink, clank, tinkle’, North Sami *gilkit* ‘jingle, chink’. The handling of those components as suffixes would considerably reduce the current excess of onomatopoeic-descriptive stems in Estonian (and other Finnic languages). This could also help etymologise some verbs of unknown origin interpreting them as derivatives from verb stems common in Finnic. For example, the verb *kohkuma* ‘to be frightened’ without certain equivalents in cognate languages might be a *ku*-derivative from the Finnic stem **koho-* (in Estonian *kohuma* ‘to boil up, to effervesce’ and Finnish *kohota* ‘to rise’). The etymological connection between these verbs was suggested by Andrus Saareste (he held that *-k-* was “analogical”); as regards the semantic relations, he supposed that the earlier meaning of the verb *kohkuma* was ‘to rise up, to jump up’ (Saareste 1924: 157, 272).

4 Complicated Interpretation of Loanwords

In some words, the apparent derivational relationship cannot be explained only on the basis of the Estonian language, because the relevant context includes loanwords.

4.1 Loanword groups with the same stem

In numerous cases, several words with the same stem have been borrowed, one of them being a stem word and the other(s) containing a derivational suffix of the donor language. For example, there are Middle Low German loanwords *teenima* ‘to serve’ (< MLGm *dēnen* id.) and *teener* ‘servant’ (< MLGm *dēner* id.) in Estonian (see, e.g. EEW: 3108, 3109). Although the stem is clearly the same, the noun *teener* cannot be regarded as an Estonian derivative, because *-er* is not an Estonian suffix. Sometimes the loan source may be obscured, because the original suffix has changed in Estonian. The Middle Low German agent suffix *-er*, for example, has often developed into *-el* due to dissimilation. For example, *röövel* ‘robber’ (< MLGm *rover* id.) is a separate loanword besides the noun *rööv* ‘robbery’ (< MLGm *rōf* id.) and the verb *röövima* ‘to rob’ (< MLGm *roven* id.) (see, e.g. EEW: 2621, 2622; SKES: 907–908; SSA 3: 122).

4.2 Loanwords with suffixes resembling genuine ones

Sometimes a loanword contains a derivational suffix of the donor language, which phonetically resembles a genuine (Estonian or Finnic) suffix. The Estonian *lusikas* (GEN *lusika*), Finnish *lusikka*

‘spoon’, e.g., seems to contain the suffix *-k(k)a* (cf. Estonian *pihla-kas* (GEN *pihlaka*) ‘rowan’ < *pihl* id.; Estonian *mustikas* (GEN *mustika*), Finnish *mustikka* ‘blueberry’ < *must-* ‘black’), but there is no such stem as **lusi* in Finnic languages. The word is indeed neither Estonian nor Finnic derivative, but a loanword (Finnic **lusikka* < Old Russian *лъжка* (*lužika*)) probably containing a Slavic suffix resembling a Finnic one (see SKES: 314; EEW: 1392; SSA 2: 113; Blokland 2009: 161). The Estonian *kleenuke* ‘very lean; thin, slim’ is sometimes interpreted as a derivative with two diminutive suffixes *-u-ke* (cf., e.g. *pal-u-ke* ‘small crumb, bit’ < *pala* ‘piece’) (see, e.g. EEW: 864). As the stem **kleen* is not borrowed into Estonian separately, the word *kleenuke* is sooner borrowed from Low German (or Swedish) along with derivational suffixes (later adapted to phonetically similar Estonian ones), cf. LMGm *klēn(e)* ‘thin; fine, frail, graceful; small, slight, unimportant’, *klenik*, *klenlik*, *kleinlik* ‘thin, fine, frail, graceful’ (SchL: 478; cf. also Raun 1982: 42), Swedish *klen* ‘small, little (obs.); thin, fine; weak; scanty, tenuous; sick, sickly’, obs. *klenlig* id.; ‘somebody or something small, thin, weak etc.’ (SAOB s.v. *klen*, *klenlig*).

Such cases are not always easy to identify. For example, in addition to the words *teenima* and *teener* discussed above, there is a third Middle Low German loanword with the same stem, *teenistus* ‘service’ < MLGm *dēnst*, *denest* id. (+ Estonian *-us*). At the first sight, it could be interpreted as an Estonian derivative *teeni-st-us* from *teeni-ma* ‘to serve’ (cf. *jooni-st-us* ‘drawing’), but this is not the case. The noun *joonistus* is derived from the verb *jooni-sta-ma* ‘to draw’ (including a causative-factitive suffix *-sta-*), but there is no such verb as **teeni-sta-ma*. In old literary Estonian we find an earlier form of the word, *teenst*, which is closer to the Low German source, proving that the word is a separate loan. Some Estonian verbs seem to contain genuine suffixes *-lda-*, *-n(-)da-*, *-r(-)da-*, but stem verbs without the mentioned suffixes are lacking. Probably those verbs were borrowed from German (including a German dialect once spoken in the Baltics, called Baltic German), along with the components *-l-*, *-n-*, *-r-*, and only causative-factitive suffix *-da-* was added in Estonian, e.g. *traageldama* ‘to baste’ (< BGm *trakeln* id.) (EEW: 3256), *rehkendama* ‘to reckon’ (< Gm *rechnen* id.) (EEW: 2445), *klimberdama* ‘to strum (on the piano)’ (< Gm *klimpern* id.) (EEW: 866), *sahkerdama* ‘to traffic, to speculate’ (< Gm *schachern* id.) (EEW: 2673).

4.3 Derivative- or compound-like loanwords as the results of folk etymology

Often foreign-sounding loanwords are subjected to folk-etymological adaption and the results are similar to a derivative or a compound. The word *puuslik* ‘idol’, for example, sounds like a derivative from *puu* ‘tree; timber’ with an exceptional suffix *-slik*. It is more likely, though, that *puuslik* is a folk-etymological version of the Estonian dialect word *puus(t)li* ‘icon; idol; guardian spirit’, which is borrowed from Russian *апостол* ‘apostle’ (Blokland 2009: 226–227). The animal name *tuhkur* ‘polecat’ could be interpreted as a derivative from the stem *tuhk* ‘ashes’, containing the suffix *-ur* (c.f. colour word *tuhkur* ‘ashen’ and *tuhkur (hobune)* ‘pale grey horse’), but probably the Estonian word for polecat is borrowed from Old Russian *дѣхорѣ* (*dūhorī*) id. (see e.g., Blokland 2009: 292–293). Sometimes the stem of a loanword is similar to an Estonian word, and also the suffix fits Estonian derivational system. For example, the Võru dialect word *villan* ~ *villañ* ‘women’s blanket, a large wool kerchief’ seems to be an Estonian derivative with *ne-* suffix from the noun *vill* ‘wool’, that is an old Baltic loanword. Actually, the word for ‘wool kerchief’ - *villan* is a younger loanword from the Latvian *villane*, *vilnaine* id. (Vaba 1997: 242).

In Estonian dialects most of the foreign colour terms display phonetic similarity to one or another familiar word and structural similarity to a comparison-based expression (Oja 2007: 204–205). For example, the German loanword *Anilin* is *aniliin* in standard Estonian, but in dialect speech it is

interpreted as a compound from two words *hani* ‘goose’ + *lilla* ‘purple’, and the colour is called *hanililla*.

5 Conclusion

In the etymological studies of the vocabulary of Estonian and cognate languages, several problems are related to word formation. On the one hand, diachronic derivational relations may not be ascertained by the synchronic approach due to sound changes in the stem or in the suffix, or because the suffix has fallen out of active use. Often the data from the dialects, old literary language and cognate languages help explain the diachronic changes. Many Estonian verbs with the component *-kV-*, *-gV-* are usually classified as different onomatopoeic-descriptive stems. Probably those words are derivatives with suffixes containing consonant *-k-* or *-g-*.

On the other hand, sometimes loanwords may be treated as derivatives erroneously. Loanwords with the same stem, one of them being a stem word and the other containing a derivational suffix of the donor language, are problematic. Misinterpretations could be avoided by means of comparing the loanword with the derivatives of the donor language. A loanword containing a component, similar to an Estonian/Finnic suffix, may have been interpreted as a derivative. If the separate stem is lacking, the word is most probably borrowed together with the suffix-like component. In dialect speech many incomprehensible loanwords (including derivatives and compounds) were adapted by folk-etymology, and some of these words subsequently entered the standard language as well. Sometimes the original shape of such distorted words is easier found with the help of their meaning.

6 References

- Blokland, R. (2009). *The Russian Loanwords in Literary Estonian*. Veröffentlichungen der Societas Uralo-Altaica 78. Wiesbaden: Harrassowitz.
- EEW = Mägiste, J. (1982–1983). *Estnisches etymologisches Wörterbuch*. Helsinki: Finnisch-ugrische Gesellschaft.
- Hallap, V. (1958). Materialy po istorii glagolobrazovatel'nyh suffiksov v mordovskih jazykah. In *Keele ja Kirjanduse Instituudi uurimused. Trudy instituta jazyka i literatury II*. Eesti NSV Teaduste Akadeemia. Akademiija nauk Èstonskoj SSR. Tallinn: Eesti Riiklik Kirjastus, pp. 175–256.
- Hallap, V. (1983). Eesti ja sugulaskeelte adjektiivitüüpe. In *Keel ja Kirjandus*, pp. 422–432.
- Nikkilä, O. (1998). *nop-ea* → *nop-sa*, *sel-keä* → *sel-vä*: suffiksinvaihto ja suomen sanojen etymologiointi. In *Kirjoituksia muoto- ja merkitysopista*. Folia Fennistica & Linguistica 21. Tampereen yliopiston suomen kielen ja yleisen kielitieteen laitos. Tampere: Tampereen yliopisto, pp. 77–101.
- Oja, V. (1996). Compound adjectives used in the Estonian dialects to denote light and pale shades of colour. In *Linguistica Uralica*, XXXII (2), pp. 92–100.
- Oja, V. (2007). Colour naming in Estonian and cognate languages. In R. E. MacLaury, G. V. Paramei, D. Dedrick (eds.). *Anthropology of Color. Interdisciplinary multilevel modeling*. Amsterdam, Philadelphia: John Benjamins, pp. 189–209.
- Rätsep, H. (2002). Uute lihttüvede saamise viisidest eesti sõnavaras. In *Sõnaloo raamat*. Tartu: Ilmamaa, pp. 78–83. [First published in *Festschrift für Károly Rédei zum 60. Geburtstag. Emlékkönyv Rédei Károly 60. születésnapjára*. Wien-Budapest, 1992, pp. 371–375.]

- Raun, A. (1982). *Eesti keele etimoloogiline teatmik*. Rooma, Toronto: Maarjamaa.
- Saareste, A. (1924). *Leksikaalseist vahekordadest eesti murretes. I. Analüüs 60 kaardi ja 1 skeemiga*.
Résumé: Du sectionnement lexicologique dans les patois estoniens. I. Analyse avec 60 cartes et 1 esquisse schématique. Acta et Commentationes Universitatis Dorpatensis B VI, 1. Tartu.
- SAOB = *Svenska Akademiens ordbok*. Accessed at: <http://www.saob.se> [28/04/2016].
- SchL = Schiller, K., Lübben, A. (1876). *Mittelniederdeutsches Wörterbuch. Zweiter Band (G–L)*.
Bremen: Verlag von J. Kühtmann's Buchhandlung. Accessed at: <http://www.rzuser.uni-heidelberg.de/~cd2/drw/s/Sa-schm.htm#Schiller-Lubben> [28/04/2016].
- SKES = Itkonen, E., Toivonen, Y. H., Joki, A. J. (1955–1981). *Suomen kielen etymologinen sanakirja. I–VII*. Lexica Societatis Fenno-Ugricae XII. Helsinki: Suomalais-Ugrilainen Seura.
- SMS = *Suomen murteiden sanakirja*, 7. Head ed. M. Vilppula. Suomalaisen Kirjallisuuden Seuran Toimituksia 556, Kotimaisten kielten tutkimuskeskuksen julkaisuja 36. Helsinki: Kotimaisten kielten tutkimuskeskus, Suomalaisen Kirjallisuuden Seura, 2003.
- SSA = *Suomen sanojen alkuperä. Etymologinen sanakirja*. Eds. E. Itkonen, U.-M. Kulonen. Suomalaisen Kirjallisuuden Seuran Toimituksia 556, Kotimaisten kielten tutkimuskeskuksen julkaisuja 62. Helsinki: Suomalaisen Kirjallisuuden Seura, Kotimaisten kielten tutkimuskeskus, 1992–2000.
- Vaba, L. (1997). *Uurimusi läti-eesti keelesuhetest*. Tallinn – Tampere: Eesti Keele Instituut, Tampereen yliopiston suomen kielen ja yleisen kielitieteen laitos.
- Vare, S. (2012). *Eesti keele sõnapered. Tänapäeva eesti keele sõnavara struktuurianalüüs*. Tallinn: Eesti Keele Instituut, Eesti Keele Sihtasutus.
- Wiedemann, F. J. (1973/1893). *Eesti-saksa sõnaraamat. Estnisch-deutsches Wörterbuch*. Vierter unveränderter Druck nach der von Jakob Hurt redigierten Auflage. Tallinn: Valgus.

Acknowledgements

Our participation in the XVII Euralex Congress is supported under the Governmental basic funding of the Institute of the Estonian Language by the Estonian Ministry of Education and Research.