THE ORIGINS OF STEMS OF STANDARD ESTONIAN – A STATISTICAL OVERVIEW

Sven-Erik Soosaar

Institute of the Estonian Language, Tallinn, and the University of Tartu

Abstract. The lexicon of literary Estonian is analysed by layers of inherited stems and loan words according to the data from the Estonian Etymological Dictionary (2012). The data is compared with previous data presented by Huno Rätsep in the 1980s and main differences are shown and the reasons behind these discussed. The numbers of cognates in related languages is analysed and the data compared with the theories on proto-languages. No affirmation for a separate Finno-Volgaic node is found based on lexical cognates in Estonian. Samoyed languages clearly distinguish from Finno-Ugrian languages with Selkup and Nenets having more cognates than other Samoyed languages.

Keywords: cognate words, lexicostatistics, loan words, Estonian etymology, Estonian lexicon history of Estonian

1. Introduction

The research on the etymology of Estonian lexicon has already lasted for several centuries. Since the last conclusive overview, more than quarter of a century has passed (Rätsep 1983, 1986) and during this time the etymology of many stems has been adjusted and numerous new etymologies have been proposed. The first etymological reference books of Estonian were published in the early 1980s. Julius Mägiste's monumental "Estnisches etymologisches Wörterbuch" remained unfinished because of the death of the author, but Finno-Ugrian Society in Helsinki published the manuscript of it in 1983 in twelve volumes (EEW). One year before that Alo Raun had published a small reference book, that in a very concise style gives in one line the origin of the word and a few cognates (Raun 1982). The most recent attempt to give an overview of the etymology of Estonian lexicon was made in the Estonian Etymological Dictionary (EES), which is the first etymological dictionary of Estonian published in Estonia. Based on

EES, I will give a new overview of the Estonian lexicon and statistical analysis of cognates in related languages. The methods and principles are similar to the ones used in Huno Rätsep's articles (Rätsep 1983, 1986) but there are some differences in the distribution of historical layers of the lexicon. It should also be noted that Rätsep based his research on his own card catalogue, not available to the public, the present research is based on published source in the shape of EES.

First of all it has to be defined what a stem in the case of Estonian is. Based on synchronic linguistics all these lexemes should be considered stems, from which no case endings, number markers, tense markers, derivational suffixes etc. can be separated. But in the case of historical linguistics we also have to find all historical suffixes that have been merged to the stems and are no more considered as derivational suffixes. Plus we have to find all phonological rules, contractions of stems and other possible changes. One lexical morpheme (stem) from the view-point of contemporary language may historically derive from several stems, like e.g. *praegu* 'now' which is composed of two stems or may contain historically a derivational suffix.

EES has made a questionable choice of defining as stems only the lexems that are not derived from other stems either in Estonian or in any proto-language from which Estonian derives. Accordingly the derivations from the Finnic or Finno-Permic period are not considered separate stems. For example *nägu* 'face' and *nägema* 'to see' are considered as one stem, although *nägu* is as Finno-Saamic derivation from the stem in Estonian *nägema* and which derives from Finno-Ugric stem **näke-* (UEW:302). The word *nägu* has cognates in addition to all Finnic also in most Saami languages, e.g. North Saami *niekko* 'dream' (SSA 2:251). Another example is the case of words *valva-(ma)* 'to keep watch' and *vaata-(ma)* 'to watch', because the latter has been derived from the first by means of derivational suffix *ta*. Later the simplification of consonant cluster and contraction of syllables **lv > *l > Ø* has yielded the word *vaatama*. In dubious cases the stems are dealt as separate ones, e.g. in case of *valvama* it has been proposed that this stem is an old derivation from a stem represented by Estonian *vala-(ma)* 'to pour'. The stem *kat-* in the word *kat-(ma)* 'to cover' is in contemporary Estonian undividable lexical morpheme. However, etymologically it derives from the sequence of two morphemes: **kante-* and causative suffix **-tta-*, besides **kante-* is the same morpheme, found in contemporary Estonian *kaas* 'cover, lid'.

According to the existence of cognates in related languages it is possible to divide stems of contemporary language into historical layers. In case of dubious etymologies it has been counted how many stems there are where cognates in some closer language are doubtless or to which other layer of loans the stem in question may belong. In some cases there are several problematic issues: the stem may be a loan and the proposed cognates are not satisfactory in every side. For example *suga* may be an Indo-Iranian loan, but cognates in Mordvin, Mari and Komi are dubious. The stems with several possible etymologies have been counted in several layers. In the case of stems without loan etymology the stems which have dubious cognates in more distantly related languages are counted as certain

stems in the layer where there is clear cognate and as dubious stems in the layer where the most distant dubious cognate is proposed. E.g. the stem *kõrv* 'ear' has clear cognates in Saamic, but dubious cognates up to Samoyedic. This stem is counted as clear Finno-Saamic and as a dubious Uralic stem. In case of every layer the first number refers to stems which belong to the layer under discussion and the other is the number of all possible stems of this layer. E.g. in the case of *suga* there are three possible layers: Finno-Saamic layer, Indo-Iranian loan (dubious), and Finno-Permic layer (dubious). Because of that the sum of all stems in all layers is bigger than the the overall number of stems, since the same stems are counted in different layers.

A more detailed overview of the principles and methodology of EES and a somewhat different viewpoint on lexical build-up of Estonian has been published earlier (Metsmägi, Sedrik, Soosaar 2013) and most of the data on stem layers has been published there as well.

2. On the overall number of stems

The basis for the selection of stems in EES are stems contained in the orthographical dictionary of Estonian (ÕS 2006). Excluded are foreign stems which contain foreign letters (*f, š, z, ž* anywhere or *b, d, g* at the beginning of the word) or which have an unusual phonological structure (e.g. *o* in non-initial syllables like in *verssok* 'former unit of measurement')[1] or that have entered spoken language in the early 20th century. Finnish loanwords are exceptionally included, the bulk of which entered Estonian in the first half of the 20th century.[2] The main drawback is that this kind of selection does not actually reflect Estonian lexicon at any point of time, because some stems that entered the language probably before the 19th century are excluded on the basis of their orthography and stems that were created *ex nihilo* in the first or even second half of the 20th century are included. But the overall number of such stems is not big. There are altogether 6641 entries in EES. Of these 1238 are reference entries, the stem of which is historically the same as some of the main entries. There are 5403 main entries, which is also the biggest possible number of stems. The number of different stems is probably smaller since 580 main entries are considered possible (but not sure) variants, derivations or lexicalised forms of some other stem. Distracting this number from the number of main entries we get the minimal number of stems. Thus there are 4823–5403 stems in EES.

---

[1] Some older loan stems have been excluded because of that or by mistake, e.g. *burlakk, koolera, piiskop, polkovnik.*

[2] About criteria in more detailes see EES:25.

3. Inherited stems

The stems of Estonian have been divided into inherited stems (from some proto-language) and loan stems. Strictly speaking the loan stems which are borrowed into some proto-language are also inherited from this protolanguage but they are discussed separately under loan stems. Inherited stems are further divided between layers depending on in which languages the cognates are found. I have divided inherited stems into 8 groups: Uralic, Finno-Ugric, Finno-Permic, Finno-Maric, Finno-Mordvinian, Finno-Saamic, Finnic and Southern Finnic stems. The stems of unknown origin and stems that have probably emerged in Estonian will be dealt separately.

3.1. Uralic stems

To the oldest layer of inherited stems belong stems that have cognates in one or several Samoyed languages: Nenets (Tundra or Forest Nenets), Enets, Nganasan, Selkup, or already vanished Kamass and Mator and which are not loans from Indo-European.[3] To this layer belong 97 to 149 stems according to EES. In this paper some Uralic etymologies proposed by Eugene Helimski and Ante Aikio have also been taken into account (Helimski 1999, Aikio 2002, Aikio 2006), which unfortunately did not enter into EES.[4] This increases the number of Uralic stems in Estonian to 106–169 or 1.96–3.13% of total number of underived stems.

Among clear Uralic stems are e.g. *ala* 'area', *ela-*(*ma*) 'to live', *ema* 'mother', *ime-*(*ma*) 'to suck', *isa* 'father', *kadu-*(*ma*) 'to disappear', *kaks* 'two', *kala* 'fish', *kand-*(*ma*) 'to carry', *keel* 'tongue', *koer* 'dog', *kool-*(*ma*) 'to die', *kuu* 'moon', *lumi* 'snow', *luu* 'bone', *maa* 'earth, land', *meie* 'we', *mina* 'I', *mine-*(*ma*) 'to go', *murakas* 'cloudberry', *nadu* 'husbands sister', *neela-*(*ma*) 'to swallow', *nool* 'arrow', *pane-*(*ma*) 'to put', *pea* 'head', *peks-*(*ma*) 'to beat', *pesa* 'nest', *pime* 'dark', *pure-*(*ma*), *põlv* 'knee', *päev* 'day', *püü* 'ptarmigan', *see* 'this', *suusk* 'ski', *süda* 'heart', *tema* 'he/she', *tule-*(*ma*) 'to come', *uju-*(*ma*) 'to swim', *väi* 'son-in-law', *õu* 'yard', *üle* 'over'.

Among dubious Uralic stems 6 may be loans. Only Indo-European loans are possible in Proto-Uralic and these are: *aja-(ma)* 'to drive', *pelga-(ma)* 'to be afraid of', *soon* 'sinew', *sõud-*(*ma*) 'to row', *too-*(*ma*) 'to bring', *vesi* 'water' (the loan relation of these stems is not certain for phonetic or semantic reasons).

53 dubious Uralic stems belong to some other inherited layer: 14 Finno-Ugric, 3 Finno-Permic, 2 Finno-Maric, 4 Finno-Mordvinic, 15 Finno-Saamic, and 15 Finnic stems. 2 stems (*oota-*(*ma*) 'to wait', *onu* 'uncle') have clear cognates in more closely

---

[3] There is a new theory that the first breakup of Proto-Uralic was not between Samoyedic and Finno-Ugric but Samoyedic-Ugric and Finno-Permic (Häkkinen 2007, 2012). This theory is based on phonological changes and not on lexicology. Because of that the present article remains on the traditional Finno-Ugric/Samoyed dichotomy.

[4] Of these, clear Uralic stems are the following 9: *aja-(ma), kaeva-(ma), kesi, kura, kätke-(ma), lähe-(dal), lämb, nõlg, ole-(ma)* and I consider dubious following 11: *esi, joo-(ma), kääri-(ma), köis, nisa, pude-(ne-ma), seis-(ma), sõda, söö-(ma), täis, õlg*.

related languages, but these may be old derivations, not originally separate stems. 1 stem (*ta* 'he') can be partly a variant of other Uralic stem (*te-*(*ma*) 'he'). In case of 1 stem (*õeva-*(*ne*) 'excellent, marvellous') all cognates are dubious.

3.2. Finno-Ugric stems

Chronologically next are stems belonging to the Finno-Ugric layer, and have a cognate at least in one Ugric language: Khanty, Mansi or Hungarian and which are not loans from Proto-Indo-European, Proto-Indo-Iranian or Proto-Iranian. Altogether 130 to 202 stems belong to this group or 2.41–3.74% of the total number of underived stems.

Among clear Finno-Ugric stems are e.g. *aht-*(*ma*) 'to put grain on poles, to cram', *ava-*(*ma*) 'to open', *hiir* 'mouse', *ilm* 'weather', *jõuk* 'horde, crowd', *jää* 'ice', *kas-*(*t-ma*) 'to water', *keha* 'body', *kere* 'body, hull', *kivi* 'stone', *koda* 'dwelling, house', *kolm* 'three', *kulm* 'eyebrow', *kuus* 'six', *käsi* 'hand', *kütke* 'trammel, shackle', *leem* 'soup, broth', *leil* 'sauna steam', *luge-*(*ma*) 'to read', *mõist-*(*ma*) 'to understand', *mägi* 'hill', *nai-*(*ne*) 'woman', *neli* 'four', *nülgi-*(*ma*) 'to skin', *pilv* 'cloud', *poeg* 'son', *põu* 'bosom', *savi* 'clay', *seen* 'mushroom', *säga* 'catfish', *sõud-*(*ma*) 'to row', *sügis* 'autumn', *talv* 'winter', *täi* 'louse', *uus* 'new', *veri* 'blood', *õppi-*(*ma*) 'to learn', *öö* 'night', *üks* 'one'.

8 stems with Finno-Ugric cognates may be loans either from Indo-European: *aja-*(*ma*) 'to drive', *sool* 'bowel', *tege-*(*ma*) 'to do', *utt* 'ewe'; from Proto-Indo-Iranian: *osa* 'part', *reba-*(*ne*) 'fox' or from Proto-Iranian: *sündi-*(*ma*) 'to be born', *tõug* 'summer corn'. 1 stem with possible Ugric cognates (*taba-*(*ma*) 'to catch, to hit') may alternatively be a Germanic loan and in this case cognates in Permic and Ugric cannot be valid. 1 stem with dubious cognate in Ugric (*samb* 'sturgeon') may be identical with another Indo-Iranian loan stem (*sammas* 'column'). 58 dubious Finno-Ugric stems belong to some younger layer of inherited stems. Among these are 2 Finno-Permic, 6 Finno-Maric, 7 Finno-Mordvinic, 10 Finno-Saamic, 31 Finnic stems and 1 South-Finnic stem. 2 stems (*aas* 'meadow', *ila* 'drool') have only dubious cognates in related languages. Additionally 2 stems (*parv* 'flock' and *kusi-*(*lane*) 'ant') may be identical to stems belonging to the Uralic layer (*parv* 'raft' and *kusi* 'urine').

3.3. Finno-Permian stems

The stems belonging to the Finno-Permian layer have cognates either in Komi or Udmurt, often in both languages. There are 43–85 such stems or 0.80–1.57% of total number of underived stems.

Among clear Finno-Permian stems are e.g. *amb* 'bow, crossbow', *hui* 'netting-needle', *ihu* 'body; flesh', *jaga-(ma)* 'to divide', *jaksa-(ma)* 'to be able to, to be strong enough', *jäse* 'limb', *kaas* 'lid', *kesk* 'center', *koole* 'ford', *kotkas* 'eagle', *kudu-(ma)* 'to weave', *käski-(ma)* 'to order', *külm* 'cold', *lõuna* 'south', *meel* 'mind, sense', *mõni* 'some', *neid* 'young lady, maiden', *nema-(d)* 'they', *niin* 'bass, bast', *okse-(nda-ma)* 'to vomit', *orav* 'squirrel', *ots* 'end, tip', *pedajas*

'pine', *peni* 'dog', *pire* 'restless; grumpy', *pähkel* 'nut', *rehi* 'barn', *räppen* 'vent for hot vapour', *sage* 'frequent', *sari* 'series, cluster', *sarna(-ne)* 'similar', *seitse* 'seven', *säsi* 'pulp', *säär* 'leg, shin', *tee* 'way, road', *tuul* 'wind', *vahe-(ta-ma)* 'to change', *voos* 'annual crop', *äi* 'father-in-law').

8 stems with cognates in Permic languages may be loans, of these *mõni* 'some', *soovi-(ma)* 'to wish', *utt* 'ewe' may be loas from PIE, *ori* 'slave; serf' from PIE or Proto-Indo-Iranian, *reba-(ne)* 'fox', *suga* 'weavers read; rough brush' from Proto-Indo-Iranian, *soe* 'warm', *sündi-(ma)* 'to be born' from Proto-Iranian, some of these have dubious Finno-Ugrian cognates, see 3.2.

33 of dubious Finno-Permian stems belong to some younger layer of inherited stems, of these 4 belong to Finno-Maric, 3 Finno-Mordvinic, 6 Finno-Saamic and 20 to Finnic layer. 1 stem (*piht-* in *pihtaed* 'pole fence') has only dubious cognates in related languages. 1 stem (*oma* 'own') may be an old derivation from the Uralic stem.

3.4. Finno-Maric stems

EES specifies Finno-Volgaic as a separate layer, where the stems have cognates in Mari or Mordvin (either Erzya or Moksha) (p. 10). Although the invalidity of this group is acknowledged, still a separate Finno-Volgaic sub-layer is defined (p. 10 and 16). This is not justified, because the existence of Finno-Volgaic protolanguage was doubted long ago (e.g. Bereczki 1974 and Bereczki 1988:314–315) and later rejected, based on phonology (Häkkinen 2007, 2012), lexical analyses (Michalove 2002) and recently on lexicostatistics (Blažek 2012). The groundlessness of a separate Finno-Volgaic layer is therefore out of question. In the present article Finno-Volgaic and Finno-Maric according to EES have been united into one group. According to EES the number of the so-called Finno-Volgaic stems in Estonian is only 20–28. 2 of these may be derivations of other stems which date back to this period: *kummardama* 'to bow' (derivation of a Uralic stem, the equivalent of which is the Estonian *kumm* 'vault, convexity') and *ootama* 'to wait' (may be Uralic stem or a derivation of a stem, which is represented by Estonian *oda* 'spear') and also *üheksa* '9' and *kaheksa* '8' contain Uralic stems for '1' and '2' respectively. Not taking these 4 into account, only 16 stems remain, which have undisputed equivalents in Estonian, Mordvin and Mari (and not in more distantly related languages). This would be the smallest of all inherited layers, which serves as a further proof that Finno-Volgaic is an invalid node.

Adding up the Finno-Volgaic and Finno-Maric stems of EES we get 38 to 53 stems or 0.7%–1% of the total number of underived stems. Even this is a considerably smaller amount than that of Finno-Permic or Finno-Mordvinic stems. Among undisputed Finno-Maric stems are e.g. *haab* 'aspen', *hapu* 'sour', *ihu(-ma)* 'to hone', *jahva(ta-ma)* 'to grind', *jumal* 'god', *kaheksa* '8', *kevad* 'spring', *kuk(-lane)* 'ant', *kummar(da-ma)* 'to bow', *kärbes* 'fly', *künd(-ma)* 'to plough', *loo(-ma)* 'to create', *mahtu(-ma)* 'to fit', *mari* 'berry', *nood* 'those', *oks* 'twig', *oota(-ma)* 'to wait', *parm* 'horsefly', *püsima* 'to stay', *saar* 'ash tree', *selg* 'back', *tamm* 'oak', *tuum* 'nucleus', *täht* 'star', *uhmer* 'mortar', *vaher* 'maple', *valge*

'white', *üheksa* '9'. Among stems with disputable equivalents 10 belong to younger layers of inherited stems (among these 6 are undisputed Finnic stems: *iga* 'each', *karga-*(*ma*) 'to jump', *kidu* 'fine drifting snow', *kogu* 'whole', *kurvits* 'woodcock', *turd* 'half-dried'; 3 Finno-Saamic stems: *õige* 'right', *pinge* 'tension', *süva* 'deep'; 1 belongs to Finno-Mordvinic: *vast* 'lately') and 2 are possible Baltic loans: *järv* 'lake' and *leht* 'leaf'. The latter may be PIE loan as well as *siduma* 'to bind', *tohtima* 'to be allowed' and *õng* 'fishhook'.

1 Estonian stem, which has a cognate in Mari (*pistma* 'to prick'), is an old derivation from another Finno-Maric stem, represented by Estonian *püsima*[5]).

3.5. Finno-Mordvinic stems

Stems belonging to Finno-Mordvinic stems have cognates in Mordvin (Erzya or Moksha) languages, but not in Mari. 67–101 stems belong to this layer (or 1.24%–1.87% of all stems).

Clear Finno-Mordvinic stems are e.g. *aher* 'barren', *hüva (hea)* 'good', *istu-*(*ma*) 'to sit', *jahe* 'cool', *juur* 'root', *kõrbe-*(*ma*) 'to burn', *käpp* 'paw', *kümme* 'ten', *küsi-*(*ma*) 'to ask', *lehm* 'cow', *lisa* 'addition', *muhk* 'bump', *murd-*(*ma*) 'to break', *pett* 'buttermilk', *pisar* 'tear', *poo-*(*ma*) 'to hang', *sattu-*(*ma*) 'to happen', *sõrm* 'finger', *sääsk* 'mosquito', *töö* 'work', *vaim* 'spirit', *õlg* 'straw'.

Among dubious stems 8 stems have an alternative loan etymology, of these 5 may be loans from PIE: *nidu-(ma)* 'to connect', *pese-(ma)* 'to wash', *sang* 'handle', *sidu-(ma)* 'to bind', *sõre* 'of large grains', 1 from Proto-Indo-Iranian: *sinine* 'blue', 1 from Proto-Baltic: *lepp* 'alder', 1 from Proto-Germanic: *vähe* 'few'. Of the remaining stems in this group 2 (*kõnts* 'filth', *meigas* 'ringdove') have only dubious cognates, 18 are Finnic stems: *eda-(si)* 'forward', *jama* 'rubbish', *jära-(ma)* 'to gnaw', *kang* 'crowbar', *kõver* 'crooked', *laam* 'larg piece', *matsa-(kas)* 'thump', *nali* 'joke' *palu-(ma)* 'to beg', *pool* 'half', *sirge* 'straight', *sise-(mine)* 'inner', *sõre* 'thin, coarse', *sõstar* 'currant', *tõrju-(ma)* 'to ward off', *tädi* 'aunt', *urise-(ma)* 'to snarl', 1 Finno-Saamic stem: *lüps-(ma)* 'to milk' and 1 has cognate only in Votic: *kokuta-(ma)* 'to stammer'. 2 are old derivations: *känd* 'stump', *künnis* 'threshold', and 1 variant of older stem: *veis* 'neat'.

3.6. Finno-Saamic stems

The stems belonging to this layer are common to one or more Saami language and Finnic languages. They derive from the Proto-Finno-Saamic (Early Pre-Finnic) period before the split of Finno-Saamic protolanguage about 1000 BC (Kallio 2006:14).

Stems belonging to Finno-Saamic layer have cognates besides Finnic only in Saami languages. 110–174 stems belong to this layer or 2.04%–3.22% of all stems.

---

[5] Accordingly this stem is not separately accounted among Mari-Finnic stems.

Clear Finno-Saamic stems are e.g. *ahven* 'perch', *haige* 'ill', *huul* 'lipp', *hõbe* 'silver', *jätka-(ma)* 'to continue', *katsu-(ma)* 'to touch', *kaudu* 'through', *kindel* 'sure', *kiru-(ma)* 'to curse', *kole* 'horrible, ugly', *küla* 'village', *luule-(ta-ma)* 'to compose poetry', *nina* 'nose', *noor* 'young', *nälg* 'hunger', *org* 'valley', *põde-(ma)* 'to be ailing', *põõsas* 'bush', *rebi-(ma)* 'to tear', *rida* 'row', *rind* 'breast', *salva-(ma)* 'to bite', *sigi-(ma)* 'to breed', *sile* 'smooth', *sitke* 'tough', *viga* 'error, mistake', *vihm* 'rain', *ämm* 'mother-in-law'.

Of possible loans 5 may be Baltic: *allikas* 'source', *kägu* 'cuckoo', *lava* 'stage', *põud* 'drought', *tõsi* 'truth', 11 may be Germanic: *kare* 'rough', *kilp* 'shield', *paljas* 'bare', *põhi* 'north', *püha* 'holy', *rahu* 'peace', *raiu-(ma)* 'to cut, to chop', *saun* 'sauna', *tupp* 'sheath', *tõsi* 'truth', *täis* 'full', 1 Indo-European (*õng* 'fishhook'*)* and 4 Indo-Iranian loans: *osa* 'part', *põhi* north', *suga* 'coarse brush', *tõota-(ma)* 'to promise'. Some of these may be derivations of older inherited stems, e.g. *jooks-(ma)* 'to run', *keela-(ma)* 'to forbid', *mõnu* 'pleasure', *oja* 'brook', *oma* 'own', *onu* 'uncle', *tähn* 'spot' and three may be variants of old loan words: *kinner* 'hollow behind the knee', *laast* 'chip', *vait* 'silent'. 35 are clear Finnic stems, e.g. *hiilga-(ma)* 'to shine', *haige* 'ill', *hool* 'care', *ilu* 'beauty', *kibe* 'bitter', *koht* 'place', *kõhn* 'thin', *laul-(ma)* 'to sing', *lible* 'leaflet', *mahl* 'juice', *paka-(ne)* 'frost', *rüsi-(ma)* 'to crowd', *saarmas* 'otter', *suits* 'smoke', *sära-(ma)* 'to shine', *taht-(ma)* 'to want', *tuim* 'dull', *õmb-(le-ma)* 'to sew'.

3.7. Finnic stems

To the Finnic layer belong stems that have cognates in Northern Finnic languages (Finnish, Ingrian, Carelian or Veps). This is the most numerous layer with 960-1367 stems or 17.77%-25.31% of total amount of stems.

Clear Finnic stems are e.g. *ahm* 'wolverine', *eile* 'yesterday', *hais* 'stink', *halb* 'bad', *higi* 'sweat', *homme* 'tomorrow', *hõõru-(ma)* 'to rub', *hüppa-(ma)* 'to jump', *ilu* 'beauty', *isu* 'appetite', *janu* 'thirst', *jänes* 'hare', *kallis* 'expensive', *kaval* 'cunning', *kerge* 'easy', *kiri* 'writing', *kõht* 'belly', *kõri* 'throat', *külva-(ma)* 'to sow', *lenda-(ma)* 'to fly', *linn* 'town', *loksu-(ma)* 'to slop', *lõke* 'bonfire', *must* 'black', *naer-(ma)* 'to laugh', *neem* 'cape', *nut-(ma)* 'to cry', *nõges* 'nettle', *näri-(ma)* 'to chew', *püga-(ma)* 'to clip', *roie* 'rib', *soo* 'swamp', *sõrg* 'cloven hoof', *toru* 'pipe', *täna-(ma)* 'to thank', *udu* 'fog', *valmis* 'ready', *võlg* 'debt', *õhtu* 'evening'.

There are 407 dubious stems, among them 38 do not have clear cognates outside Estonian, e.g. *jumi-(kas)* 'knapweed', *kahk* 'sligth rustle', *koksi-(ma)* 'to knock', *lõpus* 'gill', *mõhn* 'bulge', *natu-(ke)* 'a little', *pupe-(rda-ma)* 'to tremble, to throb', *ritsi-(kas)* 'grass-hopper', *rämps* 'trash', *sorakas* 'straggly', *tatsa-(ma)* 'to toddle', *tola* 'fool', *tuhla-(ma)* 'to rummage', *tuust* 'wisp', *töga-(ma)* 'to chaff', *tömp* 'blunt', *õgi-(ma)* 'to devour' and another 5 have clear cognates only in Votic or Livonian.

135 Finnic stems can be loans, of these 43 are possibly Baltic loans, e.g. *aina* 'only; always', *ais* 'thill', *ale* 'sart', *aur* 'steam', *habe* 'beard', *haukuma* 'to bark', *kaur* 'loon', *kinnas* 'glove', *kints* 'thigh', *kiur* 'pipit', *kõik* 'all', *lang* 'relative by

marriage', *laud* 'table', *lein* 'mourning', *liiv* 'sand', *narts* 'rag', *oder* 'barley', *pard* 'beard', *peig* 'groom', *petma* 'to deceive', *piim* 'milk', *põder* 'elk', *rahn* 'lump, chunk', *rangid* 'horse collar', *sammal* 'moss', *sober* 'friend', *teder* 'black grouse', *tolm* 'dust', *toores* 'raw', *tume* 'dark', *vikat* 'scythe'.

76 may actually be Germanic loans, e.g. *aed* 'fence; garden', *heit-(ma)* 'to throw', *hüljes* 'seal', *ilge* 'loathsome', *kaeba-(ma)* 'to complain', *kangur* 'weaver', *kehv* 'poor, bad', *lait-(ma)* 'to blame', *laud* 'table', *leba-(ma)* 'to lay', *lõika-(ma)* 'to cut', *madu* 'snake', *maga-(ma)* 'to sleep', *mees* 'man', *mäger* 'badger', *otsi-(ma)* 'to search', *paha* 'bad', *palju* 'many', *pea* 'head', *peen* 'fine', *peie(-d)* 'funeral feast', *peit-(ma)* 'to hide', *pilku-(ma)* 'to twinkle', *pühki-(ma)* 'to wipe', *rahe* 'hail', *rahvas* 'people', *rohi* 'grass', *rohke* 'plenty', *sadam* 'port', *telg* 'axle', *tuhk* 'ash', *õlu* 'beer', *ümber* 'around'.

14 are dubious Scandinavian loans: *andur* 'keel', *ent* 'but', *haldjas* 'fairy', *kangur* 'weaver', *karikas* 'beaker', *lõvi* 'lion', *rait* 'colossal', *rangi-(d)* 'horse collar', *rohke* 'plenty', *rõuge-(d)* 'smallpox', *rüüpa-(ma)* 'to sip', *siig* 'whitefish', *tila* 'spout', *täbar* 'awkward'.

45 dubious Finnic stems are otherwise of unknown origin: e.g. *asuma* 'to dwell', *anum* 'vessel', *hatt* 'villus', *hõlst* 'linen cloth', *pugu* 'crop', *toigas* 'stick' etc.

135 stems may be early derivations of other stems, e.g. *abi-(elu)* 'marriage' , *astu-(ma)* 'to step', *auk* 'hole', *ebel* 'vain', *ehk* 'or; perhaps' etc. and 63 stems are possible variants of other stems: e.g. *eha* 'afterglow', *jupp* 'stump', *järg* 'succession', *kagu* 'southeast', *kippu-(ma)* 'to strive', *koba-(ma)* 'to grope', *kääbus* 'dwarf', *lahti* 'open', *lõhke-(ma)* 'to blast', *lääs* 'west', *metsis* 'capercaillie', *mõlema* 'both', *pull* 'bull', *rähn* 'woodpecker', *song* 'hernia', *till* 'penis', *togi-(ma)* 'to poke', *tõmmu* 'dark', *vaja* 'needed', *virk* 'diligent', *vale* 'fastä, *välk* 'lightning'.

3.8. Southern Finnic stems

The stems belonging to this layer have cognates in either Votic or Livonian or in both, but not in other Finnic languages. 90–130 stems or 1.67%–2.41% of all stems belong here.

There are 86 possible Estonian-Livonian stems, among these 58 are clear and 28 dubious, of these 8 may be variants of other stems: *kaal* 'weight', *kõrk* 'haughty', *käkk* 'clod, chunk', *laasima* 'to lop off branches', *lodu* 'soggy ground', *sagar* 'cluster', *suga* 'strip of bast', *tsipa* 'a little'; 8 may be derivations of other stems: *kest* 'husk, hull', *kihv* 'fang', *kikas* 'rooster', *käi* 'grindtone', *laima-(ma)* 'to slander', *ometi* 'yet, still', *raas* 'bit', *sai* 'white bread'; 9 stems have no other etymology except a dubious Livonian or Votic cognate: *mäke-(rda-ma)* 'to rumple', *pau* 'large bead', *pint* 'flail', *posi-(ma)* 'to cure by magic', *pusa* 'bunch', *pants* 'large wet lump', *rong* 'procession; train', *süga-(ma)* 'to scratch', *võigas* 'sinister'; 3 are possibe Baltic loans: *nakkama* 'to be infectious', *peig* 'groom', *tõlv* 'club' and 1 possible Germanic loan: *peig*.

There are 23 to 33 Estonian-Votic stems, e.g. *lõhn* 'smell', *mokk* 'lip', *õun* 'apple'.

There are only 11 stems in all 3 languages (Estonian, Livonian and Votic), of these 9 are clear (*hääl* 'voice' (may be a Finno-Ugric stem)*, kasima* 'to clean'*, kudrutama* 'to coo'*, niiske* 'wet'*, peerg* 'splinter'*, puts* 'vagina'*, põrkama* 'to bump'*, vänge* 'rank, smelly'*, õel* 'evil') and 2 are dubious (*lina* – may be Old Russian or Baltic loan, *part* – may be a variant of another stem).

4. Loan stems

Loan stems are borrowed from other languages into Estonian or into some proto-language of Estonian. In some cases the loan source has been a derivation or a compound word. In such cases the loan word is a more appropriate term instead of a loan stem. In this article the number of stems is taken into account and hence loan stems and loan words are not distinguished terminologically. The derivations and compounds of the same stem are counted as separate loans.

Loan stems are grouped according to the loan period and loan source. According to loan period the loans are grouped into older and newer loans.

4.1. Older loans

Older loans originate from the time, when today's attested Uralic and Indo-European languages were not yet formed as separate languages. These have been borrowed into some stage of proto-language and have nearly always equivalents in related languages. These stems have been part of the development from the proto-language to present language and have been subject to same phonological changes that the inherited stems have. Among older loans are Indo-European, Indo-Iranian, Iranian, Baltic, Germanic, Scandinavian, and Slavic loans.

4.1.1. Indo-European loans

Indo-European loans have come from the time when probably neither PIE nor PU had started to break up. The latest possible time of the start of contacts is considered to be the 4th millennium BC (Koivulehto 1999:231). Among PIE loans are stems, that have equivalents in all Uralic languages including Samoyed, but also stems which are found only in western languages or even only in Finnic (e.g. *puhas* 'clean'*, maja* 'house'). It can be explained by the disappearance of corresponding stems from more distantly related languages or because the area of Uralic-Indo-European contacts was located further in the west and all loans did not spread to the eastern part of the Uralic language area (Koivulehto 1999:231). The high age of narrowly distributed loans is attested by phonological changes affected (Koivulehto 1999:209). The amount of Indo-European loan stems is still rather small: 16–40.

Clear Indo-European loan stems are e.g. *iva* 'grain', *mesi* 'honey', *mõsk-*(*ma*) 'to wash', *müü-*(*ma*) 'to sell', *nimi* 'name', *os-*(*t-ma*) 'to buy', *puhas* 'clear, tidy', *punu-*(*ma*) 'to braid, to plait', *pura* 'icicle-like thing', *sool* 'salt', *sõlg* 'brooch', *veda-*(*ma*) 'to draw', *vii-*(*ma*) 'to take, to lead', *vili* 'grain'.

Among dubious Indo-European loans 3 may be later loans, they may be Indo-Iranian (*ori* 'slave'), Germanic (*rõht* 'horizontal') or Baltic loans (*leht* 'leaf'). The rest of altogether 20 stems are dubious loans for phonological, semantic or other reasons and these stems may actually be inherited: *pelga-*(*ma*) 'to fear', *soon* 'vein; sinew', *sõud-*(*ma*) 'to row', *too-*(*ma*) 'to bring', *vesi* 'water', *aja-*(*ma*) 'to drive', *sool* 'bowel', *tege-*(*ma*) 'to do', *utt* 'ewe', *mõni* 'some', *soovi-*(*ma*) 'to wish', *pese-*(*ma*) 'to wash', *sidu-*(*ma*) 'to bind', *nidu-*(*ma*) 'to connect', *sang* 'handle', *sõre* 'of large grains', *tohti-*(*ma*) 'to be allowed', *õng* 'fishhook', *maja* 'house', *susi* 'wolf'. 1 stem may be Finnic derivation from another stem (*roht* 'grass').

4.1.2. Proto-Indo-Iranian and Proto-Iranian loans

Indo-Iranian loans originate from the time when both PIE and PU had broken up. The oldest of these may derive from the beginning of the 3rd millennium BC (Koivulehto 1999:215). Proto-Indo-Iranian was the common ancestor of Indic, Iranian and Nuristani languages. A few stems are considered later, Proto-Iranian loans, from the period when Proto-Indo-Iranian had broken up. Indo-Iranian loans may have equivalents in every branch of Finno-Ugric, but there are several which appear only in the western branches. It has been explained by the relatively western situation of the contact area (Koivulehto 1999:232–233). They have no equivalents in Samoyed languages. There are 20–33 Indo-Iranian and 2–6 Iranian loans.

Clear Indo-Iranian loans are *abi* 'help', *aru* 'reason', *iha* 'lust', *keder* 'disk, whorl', *marrask* 'scarfskin', *ora* 'spike', *paks* 'fat', *paras* 'proper, fit', *petkel* 'pestle', *põrsas* 'young pig', *saa-*(*ma*) 'to get', *sada* 'hundred', *sarv* 'horn', *tarn* 'sedge', *udar* 'udder', *vang* handle; crook', *varss* 'foal', *vasar* 'hammer', *vasi-*(*kas*) 'calf', *viha* 'anger'. Proto-Iranian loans are *era* 'private' and *maks-*(*ma*) 'to pay'.

Of dubious Indo-Iranian loans 4 stems have an alternative loan etymology: Indo-European (*ori* 'slave'), Germanic (*peie-*(*d*) 'funeral feast', *põhi* 'north'), Germanic or Baltic (*tala* 'scaffolding'). The rest (*osa* 'part', *reba-*(*ne*) 'fox', *sammas* 'column', *sini-*(*ne*) 'blue', *suga* 'coarse brush', *ternes* 'beest', *tõo-*(*ta-ma*) 'to promise') have no alternative loan etymology and Indo-Iranian source is dubious for phonological, semantic or other reasons. These stems may be inherited. One stem, *vihka-*(*ma*) 'to hate', may be a Finnic derivative from Indo-Iranian loan stem *viha* 'anger', and one stem (*terve* 'healthy; full') has been considered for an old derivation of the Baltic loan stem *tõrv*. Other dubious Proto-Iranian loan stems (*ahne* 'greedy', *soe* 'warm', *sündi-*(*ma*) 'to be born', *tõug* 'summer grain') may belong to the inherited stems.

4.1.3. Baltic loans

Baltic loans have been acquired starting from the 2nd millennium BC or even the end of the 3rd millennium BC until the 5th century BC from Proto-Baltic, the common ancestor of Lithuanian, Latvian and Prussian. Baltic loans have for the

most part equivalents only in Finnic and Saamic, as Proto-Finnic had already separated from other Uralic branches, but the separate Finnic languages had not yet been formed. About 20 Baltic loans have a possible shared or convergent borrowing in Mordvin, and a few even in Mari. It is possible that Baltic loans arrived into Saami and Mordvin through the mediation of Finnic, but the possibility of direct contacts cannot be excluded (Vaba 2011). According to recent evaluation, of the total 36 Baltic loans in Mordvin some have been borrowed into Pre- or Proto-Mordvin and Proto-Finnic separately (Grünthal 2012). There are 162–235 Baltic loans in Estonian.

Clear Baltic loan stems are e.g. *aas* 'loop, noose', *aeg* 'time', *angerjas* 'eel', *hagu* 'stick, twig', *haljas* 'green; shining', *hammas* 'tooth', *hani* 'goose', *hein* 'hay', *hernes* 'pea', *hirv* 'deer', *hõim* 'kin; tribe', *härg* 'ox', *ihne* 'stingy, mean', *kael* 'neck', *kahv* 'ladle', *kaust* 'upper beam of a sleigh', *kiit-*(*ma*) 'to praise', *kirst* 'chest', *kirves* 'axe', *kold* (→ *kollane* 'yelllow'), *kõrs* 'stalk', *labidas* 'shovel', *lahja* 'lean; watery', *laisk* 'lazy', *liig* 'excess', *luht* 'waterside meadow', *luud* 'broom', *lõhi* 'salmon', *lõug* 'chin', *mõrsja* 'bride', *oinas* 'ram', *puder* 'porridge', *ratas* 'wheel', *rästas* 'thrush', *sein* 'wall', *seeme* 'seed', *sild* 'bridge', *sõsar* 'sister', *tiine* 'pregnant', *tütar* 'daughter', *vagu* 'furrow', *vaha* 'wax', *vehmer* 'shaft of a yoke', *vähk* 'crayfish', *õis* 'flower'.

Among dubious Baltic loans at least 21 stems have alternative loan etymology. They can be earlier Indo-European loans (*leht* 'leaf'), Indo-Iranian loans (*tala* 'scaffolding') and in some cases a Slavic loan etymology has been proposed. The biggest group is formed by stems which have a Germanic origin as other etymology, there are 14 of these (e.g. *kari* 'flock', *roog* 'stalk', *rukis* 'rye', *vai* 'peg'). Of these 8 stems are dubious loans and they may be inherited. 2 stems may be instead Scandinavian loans (*rand* 'beach', *rangi-*(*d*) 'horse collar'). Dubious Baltic loans without cognates are *kartsas* 'ladder' and *õis-*(*vesi*) 'ulcer serum'. 44 stems are dubious Baltic loans because of phonological, semantic or other reasons and these may be inherited stems. The most numerous group is formed by possible Finnic stems (see 3.7). In seven cases there is a possibility of old derivation or variant of stem (*ait* 'granary', *kannel* 'kind of harp', *kõnge-*(*ma*) 'to perish', *laid* 'islet', *ranne* 'wrist', *taim* 'plant') and *tõug* 'breed' may actually be the same stem as *tõug* 'spring grain'.

Some probable Baltic loans have been noted only in the commentaries, among them e.g. *hoone* 'building', *uks* 'door', *voon* 'lamb'.

4.1.4. Germanic loanwords

Germanic loans have been borrowed from Proto-Germanic, the common ancestor of Germanic languages into Proto-Finnic. Some loans are even older and were allegedly borrowed from the so-called Palaeo-Germanic into Pre-Finnic and possibly even from Pre-Germanic into Pre-Finnic (a stage after Proto-Finno-Permic). These loans have entered Proto-Finnic and its earlier stages starting already from around 1900 BC (Kallio 2006, Kallio 2012). Germanic loans came into Finnic until the breakup of Proto-Finnic in the first centuries of the first

millennium AD (after about 200 AD we can speak about Scandinavian loans) (Hofstra 1985:387; LÄGLOS I: XXIII). The amount of Germanic loanwords in Estonian is according to EES 193–315 or 3.57%–5.83% of all stems. As for Finnish the number of 'early Germanic loans' is estimated to be around 500 (Kallio 2012), it can be seen that EES is rather conservative on ascribing Germanic loan etymologies. Again, several possible or probable Germanic loan etymologies are briefly mentioned only in the commentaries, e.g. *halb* 'bad'*, kabi* 'hoof'*, katsuma* 'to touch'*, kord* 'order; turn; layer'*, ots* 'end; head; tip'*, pars* 'bar, pole'*, pöial* 'thumb'*, teadma* 'to know'*, uskuma* 'to believe'*, äke* 'harrow' etc. But even assuming that not all Germanic loans in Finnic have equivalents in literary Estonian, the actual number of Germanic stems could be no less than 300, probably more than 400.

Germanic loans without doubt are in Estonian e.g. *ader* 'plough', *aer* 'oar', *armas* 'beloved', *eit* 'old woman', *hame* 'shirt', *haud* 'grave', *hõlm* 'flap, lappet', *juus* 'hair', *juust* 'cheese', *kade* 'envious', *kalju* 'rock', *kallas* 'shore, bank', *kana* 'hen', *kangas* 'cloth', *katel* 'kettle', *kaunis* 'beautiful', *kaup* 'goods', *kell* 'bell', *king* 'shoe', *kuningas* 'king', *käi-(ma)* 'to walk', *laev* 'boat', *lama-(ma)* 'to lie', *lammas* 'sheep', *leib* 'bread', *luna* 'ransom', *muld* 'earth, soil', *mõrd* 'weel', *mõõk* 'sword', *nael* 'nail', *nahk* 'skin', *narmas* 'thread, thrum', *nõel* 'needle', *paik* 'place', *parras* 'board', *pind* 'surface', *puri* 'sail', *purre* 'footbridge', *põld* 'field', *rada* 'path', *raha* 'money', *rasv* 'fat', *rooste* 'rust', *sadul* 'sattle', *sama* 'same', *suur* 'big', *tarvis* 'needed', *tina* 'tin', *tuba* 'room', *umbes* 'about', *vaev* 'trouble', *varas* 'thief'.

Among dubious Germanic loan stems are two bigger groups: the ones with a possible Baltic and others with a possible Scandinavian etymology beside the Germanic one. Examples of Baltic or Germanic loan stems can be found before (in 4.1.3). There are 17 Germanic or Scandinavian loans, e.g. *ankur* 'anchor', *kibu* 'piggin', *laen* 'loan', *lage* 'bare, bleak', *lai* 'wide', *naast* 'plaque', *nobe* 'fast', *pila-(sta-ma)* 'to sully', *tõbi* 'disease'. It is not always possible to say when the stem has been borrowed and in many cases one has to take into account beside a Germanic or Scandinavian loan source also a possible (Old) Swedish loan, e.g. *att* 'headgear worn by woman of Hiiumaa', *kelk* 'sledge', *ladu* 'store', *madar* 'bedstraw', *napp* 'small wooden bowl'. On 89 cases the loan etymology is considered dubious because of phonological, semantic or other grounds. Hence these stems can be inherited stems, accordingly Estonian-Livonian, Finnic, Finno-Saamic, or Finno-Mordvin stems. The biggest is the group of Finnic stems (see 3.7). Only 2 stems: *natt* 'crawfish net' and *tang(ud)* 'groats', are stems without equivalents in related languages considered to be doubtful Germanic loans.

4.1.5. Scandinavian loanwords

The group of Scandinavian loans is historically later than Germanic loans. These have been acquired from Proto-Scandinavian, common ancestor of Scandinavian languages, starting from the 2nd or 3rd century AD until breaking up into separate Scandinavian languages in the 9th century AD (Hofstra 1985:

387, LÄGLOS I: VIII). There are 29–62 Scandinavian loans in Estonian according to EES.

Clear Scandinavian loan stems are e.g. *haju-*(*ma*) 'to spread', *jõulu-*(*d*) 'Christmas', *kari* 'reef', *koon* 'snout', *kult* 'boar', *küünal* 'candle', *noot* 'seine', *piit* 'jamb', *põrk* 'connecting rod', *raun* 'heap of stones', *riid* 'quarrel', *rooki-*(*ma*) 'to clear, to gut', *saad* 'hay cock', *särk* 'shirt', *taak* 'load, burden', *taud* 'epidemic', *tursk* 'codfish', *tõld* 'coach, carriage', *tänav* 'street', *türn-*(*puu*) 'buckthorn', *vari* 'shadow'.

There is not always enough phonological criteria to distinguish between Germanic and Scandinavian loans. Therefore there are many stems, which besides Scandinavian etymology have a possible Germanic loan etymology (see 4.1.4). Among other dubious stems in this layer are mainly possible Old Swedish and Swedish loans. In 14 cases the borrowing is not sure, e.g. *ent* 'but', *haldjas* 'fairy', *rait* 'huge', *rüüpa-*(*ma*) 'to sip', *siig* 'lavaret', *tila* 'spout', *täbar* 'awkward', and these may be inherited stems.

4.1.6. Old East-Slavic (Old Russian) loanwords

This layer consists of stems borrowed from Church Slavic, Old East Slavic and Old Russian. There is no common view when the East Slavic tribes arrived to the neighbourhood of Finnic tribes in present north-western Russia. It happened somewhere between the 5th and the 8th century, and accordingly these stems have been borrowed from the 2nd half of the 1st millennium until the end of Old Russian period in the 14th century (Mägiste 1962:67; Must 2000:11; Blokland 2009:17). Suggestions of even earlier contacts between Slavic and Proto-Finnic (e.g. Koivulehto 2006) are not very convincing. There are 41–54 Old Slavic and Old Russian loans.

Among clear instances are e.g. *aken* 'window', *hurt* 'greyhound', *kalts* 'rag', *koonal* 'bunch of tow', *lodi* 'small flat-bottomed boat', *lusikas* 'spoon', *määr* 'amount', *nädal* 'week', *raamat* 'book', *rist* 'cross', *saabas* 'boot', *sirp* 'sickle', *tapper* 'battle-ax', *tuhkur* 'ferret', *turg* 'market', *tusk* 'grief', *vaba* 'free', *vaen* 'hostility', *varb-*(*lane*) 'sparrow', *võlu* 'charm', *värav* 'gate', *värten* 'spindle'.

Among dubious Slavic loans are 11 stems with an alternative etymology. 8 of these may be as well Russian loans (*kalts* 'rag', *niit* 'thread', *paja-*(*ta-ma*) 'to tell', *raada-*(*ma*) 'to clear land', *raatsi-*(*ma*) 'to have the heart to', *raja* 'border', *sõir* 'cottage cheese'). In case of these the exact time of borrowing is not certain. There are further 2 possible Germanic loans (*astel* 'thorn', *perv* 'bank') and 1 possible Baltic loan (*lina* 'linen'). 1 stem (*västar* 'fish spear') has only a dubious Old Russian source and no cognates.

4.2. Newer loanwords

Newer loans derive from the time when (North-)Estonian may be considered a separate language. These are borrowed from languages with which Estonian speakers have had contacts. Newer loans have been borrowed into Estonian separately from related languages (parallel loans are possible and frequent).

Among newer loans are loans from Old Swedish, Swedish, Estonian (and Finland) Swedish, Low German, German, Baltic German, Russian, Latvian, and in a few cases from other languages.

4.2.1. Old Swedish loanwords

Old Swedish loan words have been acquired from the earlier stage of development of Swedish, spoken from the 9th century to the first quarter of the16th century (Raag 1988:658). There are 14–33 Old Swedish loans in Estonian.

Clear Old Swedish loans are *agul* 'suburb', *kaal* 'turnip, swede', *kaap* 'man's hat', *kadalipp* 'gauntlet', *kirn* 'churn', *käru* 'barrow', *küüt* 'team of horses', *lauter* 'landing place', *parkal* 'tanner', *piin* 'pain', *pipar* 'pepper', *plett* 'plait', *rootsi* 'Swedish', *siuna-(ma)* 'to curse'.

Among dubious Old Swedish loans are mostly stems where the time of borrowing cannot be determined. These may be older Germanic or Scandinavian loans. It is not always possible to determine if the word has been borrowed from Old Swedish, Swedish or even Low German (like in case of *laadik* 'casket', *moon* 'provisions', *nunn* 'nun', *reede* 'Friday').

4.2.2. Swedish loanwords

Swedish loans entered Estonian mainly in the 16th and 17th centuries when Estonia and Livonia belonged as a whole or partly to the Swedish realm (Raag 1988:661–662). There are 61–140 Swedish loans in Estonian.

Clear Swedish loans are e.g. *halva-(ma)* 'to paralyse', *iil* 'blast', *kann* 'jug', *korp* 'curd cake', *kriim* 'scratch', *kroonu* 'government', *lant* 'spoon bait', *malm* 'cast iron', *moor* 'hag', *näkk* 'water sprite', *pagar* 'baker', *piip* 'pipe', *piisa-(ma)* 'to be enough', *plagu* 'flag', *plika* 'girl', *riiva-(tu)* 'shameless', *rumm* 'nave', *ruut* 'square', *rätsep* 'tailor', *säng* 'bed', *tiidus* 'hasty', *vaar* 'grandpa', *vist* 'probably', *ärtu* 'hearts', *äss* 'ace'.

Most of the dubious Swedish loans have an alternative loan etymology. Among these are the aforementioned older Germanic, Scandinavian and Old Swedish loans (see. 4.1.4, 4.1.5, 4.2.1). Most numerous are stems which may be loans from Middle Low German (42), e.g. *just* 'just, exactly', *lahing* 'battle', *munk* 'monk', *pung* 'purse', *puuk* 'mythological thief', *pärssi-(ma)* 'to drag', *pöök* 'beech', *rünk* 'block', *taldrik* 'plate', *telling* 'scaffolding', *tikk* 'splinter, peg', *tipp* 'tip', *tiss* 'tit', *tokk* 'skein', *trükkal* 'printer', *tull* 'oarlock', *vürst* 'duke', *üür* 'rent'. There are no clear linguistic criteria to distinguish the loan sources. Multiple borrowing from different sources is also one possibility. In the case of 8 words there is a third possibility of borrowing from (High) German: *ankur* 'cask', *kahvel* 'gaff', *pootshaak* 'boat hook', *pootsman* 'boatswain', *tali* 'pulley', *toppa-(ma)* 'to stuff', *tragi* 'grapnel'. Usually they are maritime terms which are known in many languages and there are no linguistic criteria to distinguish the source of the borrowing. In the case of 23 stems there is an alternative German loan etymology, e.g. *hiiva-(ma)* 'to heave', *kindral* 'general', *mall* 'protractor', *mamma* 'mum', *tanta* 'auntie', *tubakas*

'tobacco', *tükk* 'cannon', *vant* 'shroud'. 6 stems may have been borrowed through Finnish, in these cases the Finnish word is borrowed from Swedish: *jaala* 'yawl', *korp* 'raven', *murjan* 'Moor, negro', *norss* 'smelt', *rulla-(ma)* 'to roll', *topsel* 'topsail'.

In the case of 5 stems etymology is dubious on phonological or other grounds, but the present stage of research has no other etymology: *kentsa-(kas)* 'strange, odd', *klutt* 'urchin, brat', *marakratt* 'romper', *pliiats* 'pencil', *riisk* 'clamp'.

4.2.3. Estonian Swedish and Finland Swedish loanwords

These loans have been acquired from Estonian Swedes, who settled on the islands in western Estonia and coastal areas of north-western Estonia in the 13th to the 15th century. A few loans are probably from Finland Swedish spoken on the northern coast and islands of Finland. There are altogether 21–29 stems of Estonian or Finland Swedish origin.

Clear Estonian or Finland Swedish stems are e.g. *haala-(ma)* 'to haul', *hauskar* 'bailer', *julla* 'yawl', *kepp* 'stick', *klibu* 'shingle', *klomp* 'lump', *nugi-(ma)* 'to parasitize', *räim* 'Baltic herring', *tont* 'ghost', *viiger* 'marbled seal'.

In case of dubious Estonian Swedish loans either a Swedish source is possible (*tups* 'tassel'), or Swedish or Finland Swedish (*holm* 'reef'), Finland Swedish (*rool* 'steering wheel'), Low German (*viik* 'bay'), Baltic German (*karbus* 'hood'), Low German or Baltic German (*porss* 'gale'), or Finnish (*päss* 'ram').

There are up to 6 Finland Swedish loan stems. These have been borrowed from Swedish- speaking people living in southern Finland. The word *simman* 'village hop' has Finland Swedish source as the only etymology.

4.2.4. Low German loanwords

Low German loans were borrowed mainly from the Middle Low German, which was spoken by most of the crusaders who invaded at the beginning of the 13th century and clergy, officials, traders and craftsmen, who followed the crusaders. Low German was the official language of Estonia until the 2nd half of the 16th century, when High German speakers started to arrive and little by little Low German was pushed into the backround. Low German was preserved as a spoken language until the 2nd half of the 18th century and locally even later (Ariste 1937:135). There are 476–659 Low German loan stems in Estonian and it is the biggest layer of loan stems with 8.81%–12.20% of all Estonian stems.

Clear Low German loans are e.g. *aam* 'cask', *amet* 'profession', *arst* 'physician', *haamer* 'hammer', *heegel-(da-ma)* 'to crochet', *hunt* 'wolf', *höövel* 'plane', *ingel* 'angel', *jaht* 'hunting', *jakk* 'jacket', *jünger*, *kabel* 'chantry', *kamber* 'chamber', *kapp* 'cabinet', *kast* 'box', *kelm* 'rogue', *kiht* 'layer', *kink* 'present', *klaas* 'glass', *kokk* 'cook', *kook* 'pastry', *kool* 'school', *korsten* 'chimney', *korv* 'basket', *kriit* 'chalk', *kruus* 'cup', *köök* 'kitchen', *kütt* 'hunter', *laat* 'market', *liim* 'glue', *meister* 'master', *müür* 'wall', *naaber* 'neighbour', *neer* 'kidney', *orel* 'organ, *paar* 'pair', *prii* 'free', *püss* 'firearm', *ruum* 'room', *saag* 'saw', *tool* 'chair', *undruk* 'skirt', *vorst* 'sausage', *õli* 'oil'.

Dubious Low German loans mostly have an alternative loan etymology. As in the case of Swedish loans, there is not always enough linguistic criteria to distinguish the source language. Because of that the biggest groups of dubious Low German loans are those with a possible German (116 stems) and a Swedish etymology (42 stems). Low German or German loans are e.g. *ahv* 'monkey', *almus* 'alms', *haak* 'hook', *kang* 'archway', *kiiker* 'spyglass', *kiil* 'keel', *klaar* 'clear', *lekki-*(*ma*) 'to leak', *luuk* 'hatch', *mast* 'mast', *müts* 'cap', *niker-*(*da-ma*) 'to carve', *nüppel* 'cudgel', *pall* 'ball', *pekk* 'lard', *pilt* 'picture', *punn* 'plug', *raa* 'yard', *rant*, *rükki-*(*ma*), *saal*, *taaler*, *teemant*, *tekk* 'deck', *telli-*(*ma*) 'to order', *tokk* 'stick', *tonn* 'ton', *toon* 'tone', *trahv* 'fine, penalty', *vamm* 'dry rot', *vanku-*(*ma*) 'to shake', *vokk* 'spinning wheel'. On Low German or Swedish loans see 4.2.2.

In dubious cases an alternative Old Swedish loan etymology (see 4.2.1), Baltic German (*kalkun* 'turkey', *timp* 'cross bun', *trits* 'skate', *nüke* 'trick', *porss* 'gale'), or in a few cases some other loans are possible.

9 dubious Low German loan stems (four of these may be High German loans) may be descriptive stems Finnic (with possible cognates in other Finnic languages, e.g. *tömp*) or Estonian stems (with possible cognates in other Finnic languages, e.g. *krimps* 'wrinkle', *tipp* 'point'). 7 dubious loans may be variants or derivations of other stems. In case of 13 stems Low German etymology is dubious on phonological or other grounds, but it is the only etymology offered so far: *karp* 'case, box', *kemp-*(*le-ma*) 'to scuffle', *kiivas* 'not straight', *kuub* 'coat', *lamu* 'seaweed', *ork* 'broach', *peldik* 'toilet', *redu* 'hiding place', *tempi-*(*ma*) 'to mix', *trulling* 'loach', *tutt* 'tassel', *tuur* 'sturgeon', *väär* 'gallery'.

4.2.5. German loanwords

German loans have been acquired from High German, which was increasingly used beside Low German starting from the second half of the 16th century and finally replaced it (Rätsep 1983:547). High German was used first as a literary and official language, as a spoken language it was initially used by the higher and more educated classes. The spread of High German was facilitated by officials who arrived from Germany after a black plague epidemic at the beginning of 17th century. (Ariste 1940:36–37). There are 356–506 German loans in Estonian or 6.59%–9.37% of all Estonian stems.

Clear German loans are e.g. *aabits* 'primer', *hakki-*(*ma*) 'to chop', *hangel-*(*da-ma*) 'to profiteer', *hekk* 'hedge', *jope* 'jacket', *jääger* 'hunter', *kamm* 'comb', *kartul* 'potato', *kett* 'chain', *kips* 'gypsum', *kirss* 'cherry', *klamber* 'clamp', *kleepi-*(*ma*) 'to paste', *kleit* 'dress', *kohver* 'suitcase', *kork* 'cork', *kurk* 'cucumber', *kutsar* 'coachman', *number* 'number', *parun* 'baron', *pirn* 'pear', *plats* 'square', *pross* 'breech', *rehken-*(*da-ma*) 'to reckon', *sahtel* 'drawer', *sink* 'ham', *tass* 'mug', *vürts* 'spice'.

Dubious Baltic German loans usually have an alternative loan etymology. Two biggest groups are formed by stems, which alternatively may be borrowed from Low German or Swedish. Examples of these can be found before (see 4.2.2 and

4.2.4). 12 stems may be borrowed from Russian, e.g. *nuut, palitu, pitsat, poolakas, sall, sits*. Stem of 5 words may be identical with other stems (*nips, plätser-(da-ma), tarre-(ta-ma), iivel-(da-ma), klaps*). 2 stems may be considered as descriptive Estonian stems (*nutsi-(ma), niku-(ta-ma)*. In case of 3 stems the dubious German etymology is the only one proposed so far (*leesk-(putk), maat-(uss), mumm* 'dead person').

4.2.6. Baltic German loanwords

Baltic German loans derive from the German dialect which developed from a High German dialect with Low German influences in Livonia and Estonia during the 17th and 18th centuries and was in general use until the first decades of the 20th century (Kiparsky 1936:10–12; Rätsep 1983:547). There are 43–56 Baltic German loans, e.g. *aasi-(ma)* 'to rally', *juker-(da-ma)* 'to fuss', *kuul* 'ball, bullet', *latter* 'stall', *nagi* 'coat-hook', *redel* 'ladder', *redis* 'radish', *sahver* 'pantry', *sirel* 'lilac', *soni* 'kepi', *tengelpung* 'purse', *traagel-(da-ma)* 'to baste', *traks* 'braces', *tärklis* 'starch', *villi-(ma)* 'to bottle'.

Most of dubious Baltic German loans have alternative loan etymology, either from Low German (see 4.2.4), German (*pits* 'dram', *pressi-(ma)* 'to press'), Swedish (*klimp* 'dumpling', *purk* 'jar'), Estonian Swedish (*porss* 'bayberry') or Russian (*troska* 'droshky'). One dubious Baltic German loan (*riiv* 'door bolt') is dubious on phonological grounds, but there is also no other etymology.

4.2.7. Russian loanwords

Russian loans have entered Estonian starting from the 15th century (Must 2000:11). There are 228–274 Russian loans covered in EES (until the end of 19th century).

Clear Russian loans are e.g. *arssin* 'arshin', *jaam* 'station', *kamp* 'troop', *kapsas* 'cabbage', *kasukas* 'fur coat', *kiisu* 'cat, kitten', *kobras* 'beaver', *kopsik*, *kõrts, lootsik, majakas* 'lighthouse', *munder* 'uniform', *nari* 'plank bed', *nihu* 'mishap', *pintsak* 'jacket', *pirukas* 'Russian pasty', *prussakas* 'cockroach', *prää-nik* 'biscuit', *puravik* 'boletus', *putka* 'small booth', *ranits* 'schoolbag', *riisikas* 'milk-cap', *räni* 'silicon; flint', *sinel* 'greatcoat', *suli* 'crook', *tassi-(ma)* 'to carry', *tatar* 'buckwheat', *tita* 'small child', *tubli* 'good and suitable', *tuus* 'ace', *tõlk* 'translator', *vurle* 'fop, dandy'.

Among dubious Russian loans there are 8 possibly earlier, Old-Russian loans (see 4.1.5), 12 possibly German loans (see 4.2.5) and a few other loans. 4 possible loans may be derivations of other stems (*torka-(ma)* 'to poke', *nehku-(gi)* 'nothing', *roit-(ma)* 'to snoop', *räbik* 'white long-coat of a Setu woman'). In the case of 14 stems is Russian etymology dubious on phonological or other grounds, but it is only etymology offered so far: *junts, liisk, mehka, räsi-(ma), sohk, surnuk, timukas* 'hangman', *trett, tuhknai, vatsk, vuhva*.

4.2.8. Latvian loanwords

Latvian loans have entered Estonian starting from the 8th century, but the contacts grew more tight probably from the 13th century (Rätsep 1983:546, Vaba 1997:504). There are 31–43 Latvian loans in literary Estonian.

Clear loans from Latvian are e.g. *kanep* 'cannabis', *kauss* 'bowl', *kiin* 'chopper', *kuut* 'cote', *kõuts* 'tomcat', *lääts* 'lens', *magun* 'papaver', *nuum* 'fattening', *palakas* 'sheet', *pastel* 'soft shoe', *rauts* 'type of scythe', *rääts* 'buckle', *sõkal* 'beard, glume', *viisk* 'soft shoe'.

Among dubious Latvian stems is 1 possible Baltic German loan (*kukkel* 'bun') and 1 possible Russian loan (*ting*). 1 stems may be Finno-Mordvinic stem (*tsura* 'boy'). In the case of 9 stems (*karask* 'barley bread', *kaugas* 'pocket, pouch', *kippel* 'spade', *konu* 'larva', *nägal* 'greedy', *pälvi-*(*ma*) 'to merit', *rask* 'puttee', *sutt* 'lamprey', *tumm* 'gruel') is Latvian etymology dubious on phonological or other grounds, but it is the only etymology offered so far.

4.2.9. Finnish loanwords

Most of the Finnish loans are consciously brought into literary language. This process began already in the second half of the 19th century but the biggest influx of Finnish loans was in the course of language renovation in the first quarter of the 20th century; a few Finnish loans have been borrowed also later (Rätsep 1976: 211–212). There are altogether 197–212 Finnish loanstems or 3.65%–3.92% of Estonian stems.

Clear Finnish loans are e.g. *aare* 'treasure', *aine* 'matter', *alista-*(*ma*) 'to subdue', *ehe* 'ornament', *hetk* 'moment', *hooma-*(*ma*) 'to notice', *huvi* 'interest', *julm* 'cruel', *häiri-*(*ma*) 'to disturb', *hääbu-*(*ma*) 'to fade away', *julm* 'cruel', *koge-*(*ma*) 'to experience', *kummaline* 'strange', *kuulu-*(*ma*) 'to belong', *lakka-*(*ma*) 'to end, to stop', *lelu* 'toy', *levi-*(*ma*) 'to spread', *lohe* 'dragon', *loobu-*(*ma*) 'to give up', *maini-*(*ma*) 'to mention', *matk* 'trip', *masenda-*(*ma*) 'to deject', *menu*, *mugav* 'comfortable', *nauti-*(*ma*) 'to enjoy', *orb* 'orphan', *raev* 'rage', *reibas*, *retk* 'trip', *rivi* 'row', *ründa-*(*ma*) 'to attack', *saabu-*(*ma*) 'to arrive', *seikle-*(*ma*) 'to gad about', *solva-*(*ma*) 'to insult', *sääst-*(*ma*) 'to spare', *uje* 'timid', *vaist* 'instinct', *veet-*(*ma*) 'to spend', *üllata-*(*ma*) 'to surprise'.

Of dubious Finnish loans 6 may be loans from Swedish (see 4.2.2.), 1 from Finland Swedish and one from Estonian or Finland Swedish (*päss* 'ram'). 1 word (*murjan*) may be Low German or Swedish loan and 1 word (*topsel*) may be also a German or Russian loan. In the case of 5 words the fact of borrowing is not sure. These may be old Finnic stems (*ahker, tohu-*(*tu*)), or they have been formed from other stems ore are identical with them (*kangasta-(ma)*, *lapi*, *mais-(maa)*).

4.2.10. Other loanwords

There are only a few stems borrowed from other languages which have been included in EES, altogether 6–18.

There are 6 clear loans from other languages: *jaana-*(*lind*) 'ostrich' (from Hebrew), *jospel* 'young shaver' (from Yidish), *mangu-*(*ma*) 'to beg' (from Roma), *morn* 'sullen' (from French), *nulg* 'fir' (from Mari) and *vutt* 'football' (from English). Some of these have been consciously borrowed into literary language in the period of language reform. Several of these are not found in Wiedemann's dictionary (*jospel, morn, nulg, vutt*).

Dubious loans from other languages are mainly stems which may belong to some other loan groups, which were dealt with before. Among them are several words of sailing, which may be borrowed directly from Dutch: *kaljas* 'schooner', *loovi-*(*ma*) 'to tack', *madrus* 'sailor' or English: *luhva-*(*ma*) 'to luff' and words which phonetical shape points to possible direct loan from Latin: *ordu* 'order', *risti-* 'Christian', *tiisikus* 'tuberculosis'. In a few cases other languages have been proposed as sources of loans, such as Old High German (*kirik* 'church') and Middle High German (*poola* 'Polish).

4.3. Artificial stems

Artificial stems are created to enrich a literary language. There are 97 such stems included in EES, which makes up 1.8% of all stems. To this group belong e.g. *aabe* 'letter', *eira-*(*ma*) 'to neglect', *emba-*(*ma*) 'to caress', *hõiva-*(*ma*) 'to occupy', *hõlva-*(*ma*) 'to make useful', *kolp* 'skull', *küül-*(*ik*) 'rabbit', *laip* 'dead body', *laup* 'forehead', *liibu-*(*ma*) 'to cling', *lünk* gap', *malbe* 'modest', *meede* 'measure', *mõrv* 'murder', *mürsk* 'shell; guided missile', *naas-*(*ma*) 'to return', *nenti-*(*ma*) 'to state', *nõme* 'ignorant', *raal* 'computer', *reet-*(*ma*) 'to betray', *relv* 'weapon', *roim* 'crime', *selve* 'self-service', *siiras* 'sincere', *solge* 'lithe', *sulnis* 'delicious', *tarni-*(*ma*) 'to provide', *tõik* 'fact', *ulm* 'dream', *vandel* 'ivory', *veen-*(*ma*) 'to convince', *väisa-*(*ma*) 'to visit'. In several cases the words of some other language have been used as a model, but if a clear correspondence between stems is missing, these are not considered as borrowings. Among artificial stems there are ones that have been created by artificial contraction of stems (e.g. *sudu* 'smog' < *suits + udu, tõik* 'fact' < *tõsi + seik*).

5. Stems of unknown origin, onomatopoetic and descriptive stems

5.1. Onomatopoetic and descriptive stems

The notion of descriptive stems has been problematic for the Finno-Ugric etymological science. It has been kind of a refuge, where most of the stems, which do not have a credible etymology, have been located. In EES the term *häälikuliselt ajendatud* 'phonetically motivated' has been used to cover onomatopoetic and descriptive stems without making distinction between them. They do not have credible cognates. There are altogether 595 of such stems or 11% of all stems. The difference between these and the stems of unknown origin (5.3.) is very subjective.

5.2. Stems formed by the contraction of other stems

There are 36–44 stems, which have been formed through the contraction of word compounds or through mixture of several words.

Among clear instances of such stems are e.g. *aasta* 'year', *aituma ~ aitäh* 'thank you', *külimit* 'a corn measure', *künnap* 'sinew', *millal* 'when', *nagu* 'like', *ning* 'and', *nõnda* 'so', *paharet* 'demon', *praegu* 'now', *teistre* 'neighbouring', *tänavu* 'this year', *uibu* 'apple tree', *veski* 'mill', *ööbik* 'nightingale'.

5.3. Stems of unknown origin

252 stems are considered to be of unknown origin. Among these are stems, which have no possible etymology offered (65), e.g. *helpi-*(*ma*) 'to gulp', *kaan* 'leech', *kirre* 'northeast', *krohv* 'plaster', *kunel* 'dormouse', *luide* 'dune', *noomi-*(*ma*) 'to reprove', *nurg* 'white bream', *pistr-*(*ik*) 'falcon', *ruuge* 'light brown', *suir* 'beebread', *tirel* 'somerssult', *võllas* 'gallows', *üla-*(*ne*) 'anemone'. The other group in this layer is formed by stems that have been considered descriptive, but this is not convincing, and they are best described as stems of unknown origin. There are 74 of such stems, e.g. *aps* 'mistake', *jäle* 'yucky', *kaba* 'butt, bowl', *luni-*(*ma*) 'to solicit', *mutt* 'mole', *nuga* 'knife', *onn* 'hut', *plonn* 'adobe', *praht* 'rubbish', *põis* 'bladder', *rong* 'train', *sirvi-*(*ma*) 'to browse', *tiir* 'round', *äsa-*(*ma*) 'to hit' (some of these have dubious cognates). The third group contains stems, which have only very improbable guesses about their origin, such as loan sources which are not registered from written sources or improbable reconstructions. Among 111 of such stems are e.g. *alp*, *igri-*(*tse-ma*), *jube*, *kalk-*(*vel*), *kena*, *kesv*, *kiiba-*(*kas*), *naaskel*, *nori-*(*ma*), *puskar*, *põll*, *pürg*, *rõhu-*(*ma*), *tardu-*(*ma*), *trimpa-*(*ma*), *uit*, *vassi-*(*ma*), *vemp*, *äsja*.

6. Statistical analysis and conclusions

6.1. The division of Estonian stems between layers

Based on EES with a few additions the division of Estonian stems in numbers by historical-etymological layers is in Table 1. The comparison of inherited stem layers is presented in Figures 1 and 2. The distribution of loan stems by layer in Estonian according to EES is presented in Figure 3 and with comparative data from Rätsep 1983 and Rätsep 1986 in Figure 4.[6]

The data of Rätsep 1983 and 1986 is not strictly comparable to the data of EES, because the source of the Rätsep analyses is not clear. He has probably included more dialectal words and late 20th century borrowings from German and Russian. The significant difference in Low German loans (according to Rätsep 1983 data there were 771 or 13.92% (Rätsep 1983:546), but now only 476 or 8.8% of all stems) can be explained partly by the same reason, but part of these are redefined

---

[6] Finno-Volgaic is a cover term for comparisons purpose, comprising Finno-Maric and Finno-Mordvinic.

as (High) German loans already in Rätsep 1986. The difference in the number of Baltic loans reflects the development of research in Baltic loans. The difference in Finnish loans is explained by the selection of data: Rätsep has included only stems, which were missing in Estonian, but in EES such Finnish loans are often included, e.g. *kuuluma* 'to belong' < Finnish *kuulua* 'id.' which is derivation of Finnish *kuulla* (Estonian *kuulma* 'to hear').

Table 1. Division of Estonia stems by historical-etymological layers

|	
Uralic stems	107–169
Finno-Ugric stems	134–206
Finno-Permic stems	43–85
Finno-Maric stems	38–54
Finno-Mordvinic stems	65–99
Finno-Saamic stems	110–174
Finnic stems	960–1368
South-Finnic stems	90–130
Estonian-Livonian-Votic stems	9–11
Estonian-Livonian stems	58–86
Estonian-Votic stems	23–33
Stems found only in Estonian	746–816
stems formed by contraction	36–44
"descriptive stems"	458–520
of unknown origin	252
Indo-European loans	16–40
Indo-Iranian and Iranian loans	22–39
Baltic loans	162–235
Germanic loans	193–315
Scandinavian loans	29–62
East Slavic and Old Russian loans	41–54
Old Swedish loans	14–33
Swedish loans	61–140
Estonian and Finland Swedish loans	22–35
Low German loans	476–659
German loans	356–506
Baltic German loans	43–56
Russian loans	228–274
Latvian loans	31–43
Finnish loans	197–212
Other loans	6–18
Artificial stems	97
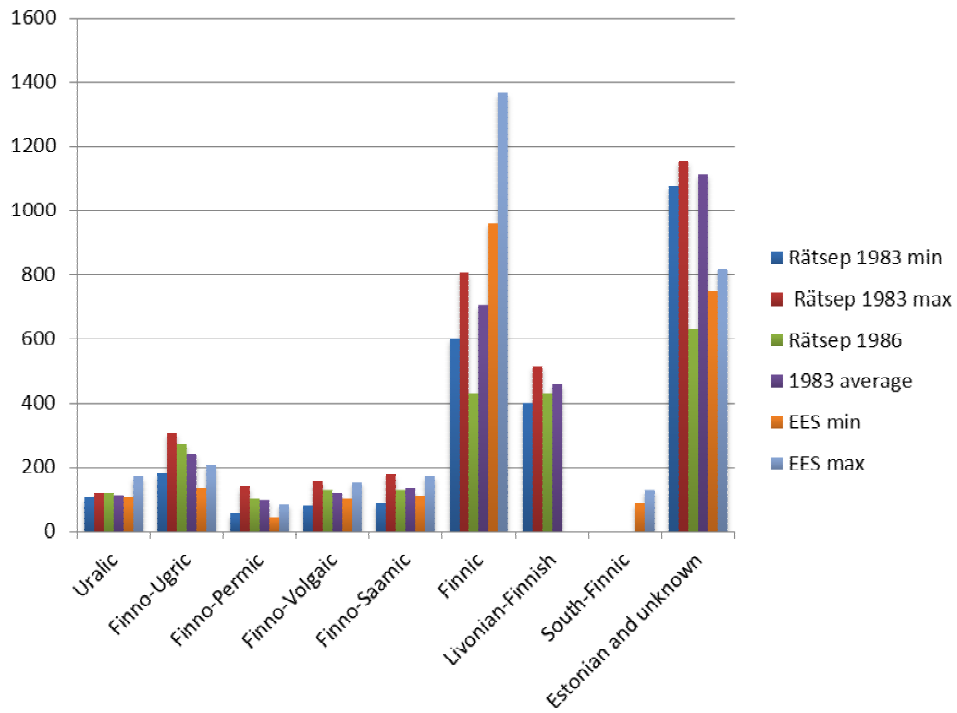
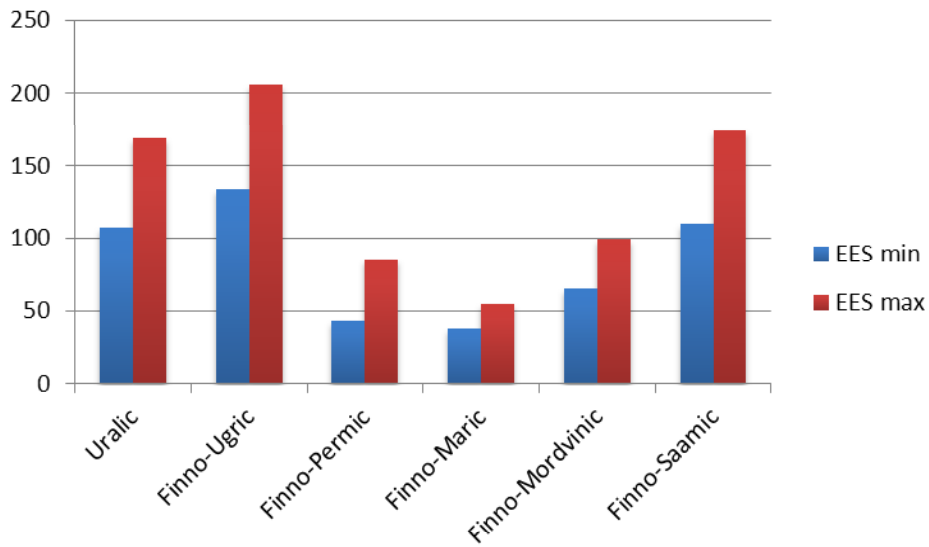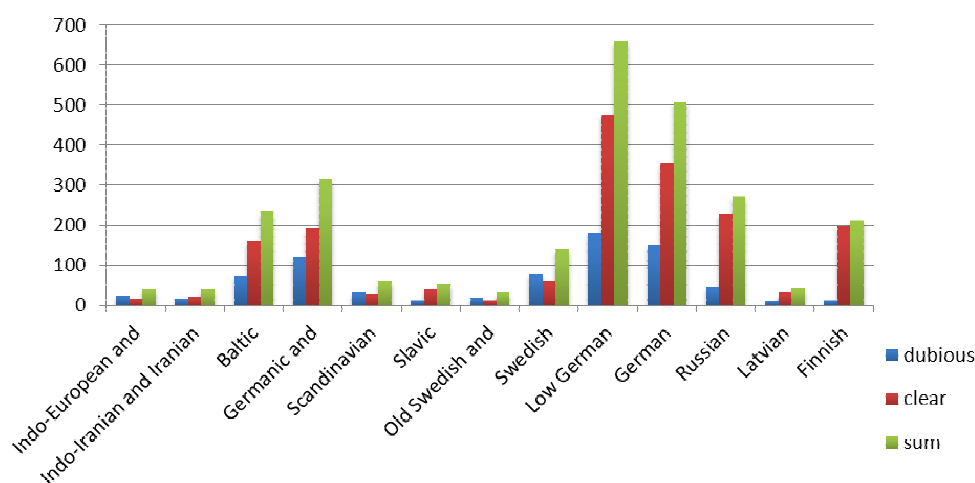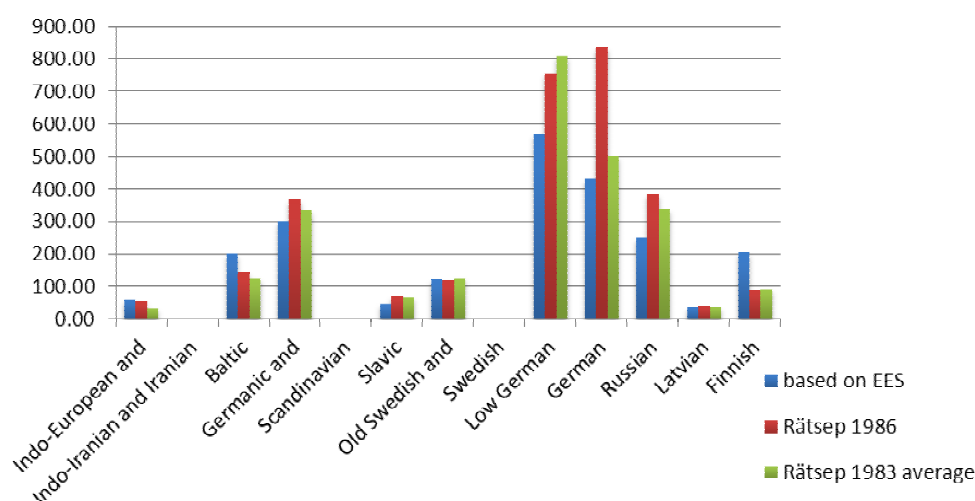**Figure 1.** Inherited stems by layer in Estonian.

Figure 2. Inherited stems older than Finnic by layer in Estonian.

Figure 3. Loan stems by layer in Estonian (according to EES).

Figure 4. Comparison of loan stems in EES with Rätsep 1983 and Rätsep 1986.

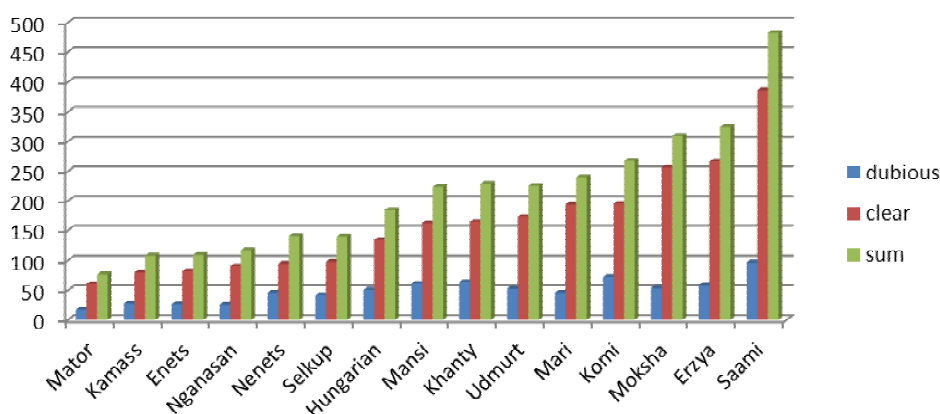## 6.2. The number of cognates in related languages

It is clear that in every language a reverse process to the addition of new stems is going on – old stems disappear. Because of this there are gaps in cognate languages and rarely a stem has been preserved in every cognate language. Among Uralic stems there are only a few with cognates in all languages. Still, it is probable that the closer the two languages are, the more common stems they should have as less time has passed since their break-up from the common proto-language. In order to check

this up I have counted all cognates in related languages of Estonian (except Finnic languages). The results are presented in Table 2 and Figure 5.

Table 2. Number of cognates of Estonian stems in related languages (excluding Finnic languages)

Samoyed:	Mator	Kamass	Enets	Nganasan	Nenets	Selkup			
clear	60	80	82	91	95	98			
dubious	17	28	27	26	46	42			
sum	77	108	109	117	141	140			
Finno-Ugric	Hungarian	Mansi	Khanty	Udmurt	Mari	Komi	Moksha	Erzya	Saami
clear	134	163	165	173	194	195	256	267	386
dubious	51	61	64	53	46	73	54	58	97
sum	185	224	229	226	240	268	310	325	483

Figure 5. The number of cognates of Estonian stems in related languages (excluding Finnic languages).

The data reveal that there is in most cases more cognates in closer languages (that have separated later from the common proto-language). Still, some divergences deserve attention. In Komi there are more cognates than in Mari and at the same time, there are considerably fewer cognates in Mari than in Mordvin languages. This offers another proof against the theory of Finno-Volgaic node. The other Permic language, Udmurt, has considerably fewer cognates than Komi (accordingly 173 clear and 53 dubious cognates and 195 clear and 73 dubious), the difference is 22 clear and 20 dubious stems or 11.3% and 27.4% less than in Komi. Such a difference indicates that Komi is at least lexically closer to Estonian than Udmurt. Another possible explanation is, that Komi is lexically and etymo-logically better researched than Udmurt. There is a short Komi etymological

dictionary (Lytkin, Gulyaev 1970). Samoyedic languages have a clear distance from Finno-Ugric languages but somewhat surprising is that the South Samoyed language Selkup has most cognates with Estonian. Logically, fewer cognates has Mator, as the first to die out and the poorest documented of these languages. On the other hand Kamass, which also has vanished, has almost the same amount of cognates than northern languages Enets and Nganasan.

Acknowledgements

Address:
    Sven-Erik Soosaar
    Roosikrantsi 6
    Institute of the Estonian Language
    10119 Talllinn, Estonia
E-mail: svenerik@eki.ee

References

Aikio, Ante (2002) "New and old Samoyed etymologies". *Finnisch-Ugrische Forschungen* 57, 9–57.

Aikio, Ante (2006) "New and old Samoyed etymologies (Part 2)". *Finnisch-Ugrische Forschungen* 59, 9–34.

Ariste, Paul (1937) "Mõningaid alamsaksa laensõnu". [Some Low German loanwords.] *Eesti Keel* 5, 135–140.

Ariste, Paul (1940) "Georg Mülleri saksa laensõnad". [German loanwords of Georg Müller.] (*Acta et Commentationes Universitatis Tartuensis,* B 46:1.) Tartu.

Bereczki, Gabor (1974) "Suščestvovala pravolžskaja obščnost' finno-ugrov?". *Acta Linguistica Hungarica* 24, 81–85.

Bereczki, Gabor (1988) "Geschichte der wolgafinnischen Sprachen". In *The Uralic languages: description, history and foreign influences*, 314–350. Denis Sinor, ed. (Handbuch der Orientalistik. 8. Abt. Handbook of Uralic studies, 1.). Leiden, New York, København, and Köln: E. J. Brill.

Blažek, Václav (2012) "Was there a Volgaic unity within Finno-Ugric?". *Finnisch-Ugrische Forschungen* 61, 1–3, 29–91.

Blokland, Rogier (2009) *The Russian loanwords in literary Estonian*. (Veröffentlichungen der Societas Uralo-Altaica, 78.) Wiesbaden: Harrassowitz Verlag.

EES = Metsmägi, Iris, Meeli Sedrik, and Sven-Erik Soosaar (2012) *Eesti etümoloogiasõnaraamat.* [Etymological dictionary of Estonian.] Tallinn: Eesti Keele Sihtasutus.

EEW = Mägiste, Julius (1983) *Estnisches etymologisches Wörterbuch.* 12 Bde. Helsinki: Finnisch-Ugrische Gesellschaft.

Grünthal, Riho (2012) "The evidence of Baltic loanwords in Mordvinic". In *A linguistic map of prehistoric northern Europe*, 297–343. Riho Grünthal and Petri Kallio, eds. (Suomalais-Ugrilaisen Seuran Toimituksia, 266.) Helsinki: Suomalais-Ugrilainen Seura.

Helimski, Eugene (1999) "Ural. *läs* 'bei, in der Nähe' im Samojedischen (und im Ungarischen?)". *Finnisch-Ugrische Mitteilungen* 21–22, 77–81.

Häkkinen, Jaakko (2007) *Kantauralin murteutuminen vokaalivastaavuksien valossa*. [Dialectal diversification of Proto-Uralic in the light of vowel correspondencies.] Pro-gradu työ. Helsinki: Helsingin Yliopiston suomalais-ugrilainen laitos.

Häkkinen, Jaakko (2012) "After the protolanguage: invisible convergence, false divergence and boundary shift". F*innisch-Ugrische Forschungen* 61, 1–3, 7–28.

Hofstra, Tette (1985) *Ostseefinnisch und Germanisch. Fru he Lehnbeziehungen im nördlichen Ostseeraum im Lichte der Forschung seit 1961*. Groningen: Rijksuniversiteit te Groningen.

Kallio, Petri (2006) "Suomen kantakielten absoluuttista kronologiaa". Virittäjä 110, 2–25.

Kallio, Petri (2012) "The Prehistoric Germanic Loanword Strata in Finnic" in *A Linguistic Map of Prehistoric Northern Europe*. Ed. by Riho Grünthal and Petri Kallio. SUST 266. Helsinki, 225–238.

Kiparsky, V. (1936) *Fremdes im Baltendeutsch.* (Mémoires de la Société Néo-philologique de Helsingfors, 11.) Helsinki: Société Néo-philologique.

Koivulehto, Jorma (1999) "Varhaiset indoeurooppalaiskontaktit: aika ja paikka lainasanojen valossa". In *Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan*, 207–236. Paul Fogelberg, ed. (Bidrag till kännedom av Finlands natur och folk, 153.) Helsinki: Finnish Society of Science and Letters.

Koivulehto, Jorma (2006) "Wie alt sind die Kontakte zwischen *Finnisch-Ugrisch und Balto-Slavisch?*". *Slavica Helsingiensia* 27, 179–196.

LÄGLOS I = Kylstra, A. D., Sirkka-Liisa Hahmo, Tette Hofstra, and Osmo Nikkilä (1991) *Lexicon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen*. Bd. I: *A–J.* Amsterdam and Atlanta, GA: Rodopi.

Lytkin, V.I., Guljaev, E.S. (1970) *Kratkij ètimologičeskij slovar' komi jazyka.* Moscow: Izdatel'stvo Nauka.

Metsmägi, Iris, Meeli Sedrik, and Sven-Erik Soosaar (2013) "Eesti kirjakeele tüvevara päritolu arvudes". [Etymology of the stems of the literary Estonian in number.] *Keel ja Kirjandus* 5, 313–332.

Michalove, Peter A. (2002) "The classification of the Uralic languages: lexical evidence from Finno-Ugric". *Finnisch-Ugrische Forschungen* 57, 58–67.

Must, Mari (2000) *Vene laensõnad eesti murretes.* [Russian loanwords in Estonian dialects.] Tallinn: Eesti Keele Sihtasutus.

Mägiste, Julius (1962) "Äldre ryska lånord i estniskan särskilt i det gamla estniska skriftspråket". (Lunds Universitets Årsskrift. N. F. Avd. 1, 55, 1.) Lund: CWK Gleerup.

Raag, Raimo (1988) "*Nunn, prilla, koka* ja teised. Eesti keele rootsi laensõnadest." [*Nunn, prilla, koka* and others. On Swedish loanwords in Estonian.] *Keel ja Kirjandus* 11 and 12, 655–664 and 725–732.

Rätsep, Huno (1976) "Soome laenudest eesti kirjakeeles". [On Finnish loans in literary Estonian.] *Keel ja Kirjandus* 4, 210–216. [also in his *Sõnaloo raamat*, 229–245. Tartu: Ilmamaa, 2002.]

Rätsep, Huno (1983) "Eesti kirjakeele tüvevara päritolu". [The origin of Estonian word stems.] *Keel ja Kirjandus* 10, 539–548. [also in his *Sõnaloo raamat*, 59–77. Tartu: Ilmamaa, 2002.]

Rätsep, Huno (1986) "Eesti kirjakeele sõnatüvede tuletuskoormus." [The derivational potential of the stems in the literary Estonia language.] *Keel ja Kirjandus* 11, 662–667. [also in his *Sõnaloo raamat*, 84–96. Tartu: Ilmamaa, 2002.]

Raun, Alo (1982) *Eesti keele etümoloogiline teatmik.* [Etymological reference book of Estonian.] Rooma and Toronto: Maarjamaa.

SSA = *Suomen sanojen alkuperä. Etymologinen sanakirja.* [Origin of the Finnish words. An etymological dictionary.] Erkki Itkonen and Ulla-Maija Kulonen, eds. (Suomalaisen Kirjallisuuden Toimituksia, 556.) Helsinki: Suomalaisen Kirjallisuuden Seura and Kotimaisten kielten tutkimuskeskus, 1992-2000.

Vaba, Lembit (1997) *Uurimusi läti-eesti keelesuhetest.* [Studies on the Latvian-Estonian loan relationships.] Tallinn and Tampere: Eesti Keele Instituut and Tampereen yliopiston suomen kielen ja yleisen kielitieteen laitos.

Vaba, Lembit (2011) "Balti laenude uurimine avab meie kauget minevikku". [Research on Baltic loans opens up our distant past.] *Keel ja Kirjandus* 10, 734–763.

ÕS 2006 = Erelt, Tiiu, ed. (2006) *Eesti õigekeelsussõnaraamat ÕS 2006.* [Estonian orthographical dictionary.] Koostanud Tiiu Erelt, Tiina Leemets, Sirje Mäearu, and Maire Raadik. Tallinn: Eesti Keele Sihtasutus. Available online at <http://www.eki.ee/dict/qs2006/>. Accessed on 03.07.2013.

Editorial note

This article is largely based on the structure of an earlier publication by three authors, Iris Metsmägi, Meeli Sedrik, and Sven-Erik Soosaar, "Eesti kirjakeele tüvevara päritolu arvudes" [The origin of stems of standard Estonian in numbers], and repeats some parts of its text. The article has been published in *Keel ja Kirjandus* (Tallinn) 2013, 5, 313–332 (available online at http://kjk.eki.ee/ee/issues/2013/5/352). In the printed version of this paper, the article is mistakenly referred to as Metsmägi, Sedrik, and Soosaar 2012.