

Eesti Rakenduslingvistika Ühing

**Eesti Rakenduslingvistika Ühingu
aastaraamat 6**

Estonian Papers in Applied Linguistics 6

Toimetanud
Helle Metslang, Margit Langemets
ja Maria-Maren Sepper

Eesti Keele Sihtasutus
Tallinn 2010

EESTI RAKENDUSLINGVISTIKA ÜHINGU AASTARAAMAT 6 **Estonian Papers in Applied Linguistics 6**

Toimetajad / Editors: Helle Metslang (Tartu), Margit Langemets (Tallinn),
Maria-Maren Sepper (Tallinn)

Kujundaja / Design: Piia Ruber

Toimetuskolleegium / Advisory board:

Martin Ehala (Tallinn), Auli Hakulinen (Helsinki), Birute Klaas (Tartu), Eino Koponen (Helsinki), Irina Külmoja (Tartu), Maisa Martin (Jyväskylä), Jaan Mikk (Tartu), Hille Pajupuu (Tallinn), János Puszta (Szombathely), Helena Sulkala (Oulu), Urmas Sutrop (Tallinn/Tartu), Eva Toulouze (Tartu), Tiia Tulviste (Tartu), Marilyn Vihman (Bangor, U.K.), Leo Võhandu (Tallinn), Haldur Õim (Tartu)

“Eesti Rakenduslingvistika Ühingu aastaraamat” on registreeritud ja refereeritud järgmistes bibliograafilistes andmebaasides / Estonian Papers in Applied Linguistics is covered by the following abstracting/indexing services:

Arts and Humanities Citation Index (Thomson Reuters)
Central and Eastern European Online Library (C.E.E.O.L)
CSA Linguistics and Language Behavior Abstracts (LLBA)
Directory of Open Access Journals, Lund University (DOAJ)
EBSCO (EBSCO Publishing)
Eesti Rahvusraamatukogu digitaalne arhiiv (DIGAR)
Journal Citation Reports / Social Sciences Edition (Thomson Reuters)
Linguistic Abstracts (Wiley-Blackwell Publishing)
Linguistic Bibliography Online (BL Online)
MLA Directory of Periodicals
MLA International Bibliography
SCOPUS (Elsevier)
Social Sciences Citation Index (Thomson Reuters)
Social Scisearch (Thomson Reuters)

Ühingu aadress / Contact information:

Eesti Rakenduslingvistika Ühing / Estonian Association for Applied Linguistics
Eesti Keele Instituut / Institute of the Estonian Language
Roosikrantsi 6
10119 Tallinn
ESTONIA

ery@eki.ee
www.rakenduslingvistika.ee

“Eesti Rakenduslingvistika Ühingu aastaraamat 6” väljaandmist on toetanud Haridus- ja Teadusministeeriumi riiklik programm “Eesti keel ja kultuurimälu” ja Eesti Keele Instituut.

Autoriõigus / Copyright: autorid ja Eesti Rakenduslingvistika Ühing

ISSN 1736-2563
doi:10.5128/ERYa.1736-2563

SISUKORD

Some remarks on reported evidentiality in French and in Estonian: a contrastive approach <i>Märkusi kaudse evidentsiaalsuse kohta prantsuse ja eesti keeles: kontrastiivne lähenemine</i> Reet Alas, Anu Treikelder	7
Eesti keele ajakategooria omandamisest – ajavormid ja ajasõnad <i>The acquisition of tense and time adverbs in Estonian</i> Reili Argus, Sirli Parm	25
The comprehension of derivational morphemes in early childhood: an experimental study for Lithuanian <i>Tuletusmorfeemide omandamine varajases lapsees: leedu eksperiment</i> Ineta Dabašinskiene	43
Intellektuaalseid võimeid tähistavate sõnade kasutused intellektuaalses tegevuses eesti keele maailmapildis <i>The conceptual roles of intellectual abilities in the worldview of Estonian</i> Larissa Degel	51
Direktiivsed aktipaarid eestikeelsetes infodialoogides ja nende automaatne tuvastamine <i>Directive act pairs in Estonian information dialogues and their automatic recognition</i> Olga Gerassimenko, Riina Kasterpalu, Mare Koit, Andriela Rääbis, Krista Strandson	67
“I puts it away”—early protomorphological ways of inflecting verbs in a child acquiring Saami <i>“Ma paneb selle ära” ehk varased protomorfoloogilised verbimuutestrategiad saami lapsekeeles</i> Johanna Ijäs	87
Mõnede eesti sõnajärjemallide psühholingvistilisest reaalsusest <i>On psycholinguistic reality of some word order patterns of Estonian</i> Annekatriin Kaivapalu	103

Värvingupartiklid eesti- ja saksakeelsetes küsilausetes <i>Abtönungspartikeln in den estnischen und deutschen Fragesätzen</i> Janika Kärk	121
Vähemuskeelte elujõulisus ja arengusuundumused Eesti venekeelsetes keelekeskkondades <i>The viability and development trends of Estonian minority languages in segregative language environments</i> Elvira Küün	135
Narratiivi aeg ja tõlke aeg: diskursiivne käsitus <i>Time of narrative and time of translation: a discursive approach</i> Triin Lõbus, Silvi Tenjes	157
Kesksete lausekomponentide järjestus õppijakeeles: arvutianalüüsi katse <i>Automatic word order analysis of Estonian as a second language: the nuclear sentence</i> Helena Metslang, Erika Matsak	175
Mõnest värvinimetusest ja nende tähendusvahekordadest eesti ja soome keeles <i>On some colour terms and their semantic relationships in Estonian and Finnish</i> Vilja Oja, Mari Uusküla	195
Eestlaste ja venelaste suhtluskäitumine: <i>sina, teie</i> ja keegi veel <i>Politeness behavior among Estonians and Russians: formal and informal second person and the third person</i> Renate Pajusalu, Virve Vihman, Birute Klaas, Karl Pajusalu	207
Õppijakeele kontekstuaalne varieeruvus eesti keele täis- ja osasihitise kasutamisel <i>Contextual variation in learner language: the use of total and partial objects</i> Raili Pool	225
Russification of non-Estonian pupils in Tallinn <i>Mitte-eesti õpilaste venestumine Tallinnas</i> Mart Rannut, Ülle Rannut	243
Language & authoritarianism in the 20th century: the cases of Estonia and Catalonia <i>Keel ja autoritaarsus XX sajandil: Eesti ja Kataloonia juhtum</i> Delaney Michael Skerrett	261
Vahendamata ja vahendatud narratiiv laste kõnearengu hindamisel <i>Self-generated and retold narratives as a tool of language assessment</i> Piret Soodla, Eve Kikas, Renate Pajusalu, Aive Adamka, Sirlil Parm	277

Kas pausid kannavad emotsiooni? <i>Do the pauses in read text carry emotion?</i> Kairi Tamuri	297
Statistilised meetodid murdekorpuse ühendverbide tuvastamisel <i>Statistical methods for phrasal verb detection in Estonian dialects</i> Kristel Uihoaed	307
Kuidas õpetada kõnesüntesaatorile empaatiat? Emotsiooni automaatse tuvastuse võimalustest eestikeelses kirjalikus lauses sisalduva info põhjal <i>How to teach empathy to a speech synthesizer? On the possibility of identifying emotion solely from written Estonian sentences</i> Ene Vainik	327
Conceptual co-presence of motion and emotion in the Estonian terms of personality <i>Liikumise ja emotsioonide kaas-kontseptualiseerimine isikuomaduste sõnavaras</i> Ene Vainik, Toomas Kirt, Heili Orav	349

SOME REMARKS ON REPORTED EVIDENTIALITY IN FRENCH AND IN ESTONIAN: A CONTRASTIVE APPROACH

Reet Alas, Anu Treikelder

Abstract. Among the multiple meanings of the French conditional, the marking of mediated information is generally pointed out in the studies treating this mood in French. In Estonian there is a specialized grammatical(ized) form for marking reported evidentiality—the quotative mood.

Our paper aims to study the contrast between the use of evidentiality strategies in French and Estonian. The analysis will be based on a translated parallel corpus and on different types of texts and studies of monolingual corpora of both languages.

In typological studies the French conditional and the Estonian grammatical evidentials are considered as belonging to the same type of evidentiality, that of reported information (Aikhenvald 2006). However, it appears that they tend to be used in different types of discourse. This observation has served as a starting point for our study: the aim of our contrastive analysis is to find out whether these divergences are confirmed by translated texts, whether they also reveal semantic differences between these evidentiality markers, and which strategies are used in the translations in both directions.*

Keywords: contrastive linguistics, quotative, conditional mood, corpus analysis, French, Estonian

1. Introduction

The aim of our paper is to compare the functioning of two particular forms in genetically unrelated languages: one in French, the so-called “journalistic conditional”, and the other in Estonian, the *quotative* or *oblique* mood, more precisely the quotative mood manifested by the bound morpheme *-vat* suffixed to the first component of

* This research was supported by European Social Fund's Doctoral Studies and Internationalisation Programme DoRa. We thank Antoine Chalvin for his useful comments on the examples.

the verb phrase. Both of these forms have been described as evidentiality markers or strategies in typological and monolingual studies.

We understand here evidentiality as it has been defined by Alexandra Y. Aikhenvald: it is a linguistic category “whose primary meaning is source of information” (2006: 3). The two forms in our focus belong to the same subcategory of evidentiality in the system of Aikhenvald (type A3): the one of **reported evidentiality**, opposing reported or “hearsay” information to “everything else” (2006: 25).

Even if evidentiality is not the primary meaning of the French conditional, one of its particular uses has been regarded as evidential in several studies, including the book of Aikhenvald (2006: 11). This typological similarity as well as the evidence from translated texts has encouraged us to study the correlations between the Estonian quotative mood and the French conditional’s evidential use. The correspondence between these forms can be illustrated by the examples (1) and (2).

- | | | | | |
|------|------------------------------------|--------------|----------------|--------|
| (1a) | Le | Président | est | malade |
| | The | President | be:PRS.IND.3SG | ill |
| | ‘The President is ill.’ | | | |
| | | | | |
| (1b) | Le | Président | ser-ait | malade |
| | The | president | be-COND.3SG | ill |
| | ‘The President is said to be ill.’ | | | |
| | | | | |
| (2a) | President | on | haige | |
| | President | be:PRS.IND.3 | ill | |
| | ‘The president is ill.’ | | | |
| | | | | |
| (2b) | President | ole-vat | haige | |
| | President | be-QUOT | ill | |
| | ‘The president is said to be ill.’ | | | |

The sentences (1a) and (2a) are in indicative in both languages and give as a default reading that the speaker acts as the source of information. The sentences (1b) and (2b), respectively in conditional and quotative, imply that the information has been received from somebody else; the source of information stays unspecified in these examples. The question that arises is whether the sentences (1b) and (2b) convey the same meaning.

Our analysis is based on a parallel corpus of translated texts, including translations in both directions. However, in the present phase, the analysis is confined only to the French equivalents of the Estonian quotative. This is due to the nature of our corpus and difficulties related to the distinction of the meanings of the French conditional. On the other hand, other grammatical tools used in Estonian for expressing evidential values in the predicate (such as the present perfect, the pluperfect or the past participle) have not been taken into account either. The present article is thus only a first attempt to compare phenomena related to evidentiality in Estonian and in French.

Firstly, we shall discuss some general issues concerning the description of the forms in question in the studies of both languages (chapter 2), then the results of the corpus analysis will be presented (chapter 3) and finally we shall observe more

closely some examples allowing, in our opinion, to point to some essential differences between these forms (chapter 4).

2. Some general remarks on the correspondence between the French conditional and the Estonian quotative

2.1. The different status of the French conditional and the Estonian quotative

One of the difficulties of comparing these two forms results from the status they have as evidentiality markers. If the Estonian quotative is unanimously regarded as a marker of evidentiality by the Estonian linguists, in French, the evidential meaning concerns only one of the uses of the conditional mood besides its hypothetical (modal) and temporal use (see e.g. Haillet 2002, who calls it the “conditional of enunciative otherness”, for the distinction of these basic meanings of the French conditional). This form cannot thus be regarded as an evidentiality marker in the strict sense, since evidentiality is not its primary meaning, but rather as an evidentiality strategy according to Aikhenvald (2006). Opinions vary with regard to the relation between the French conditional and the category of evidentiality. Among French-speaking linguists Patrick Dendale (Dendale 1993, Dendale, Coltier 2003) has probably related it most radically to this category (his name for this form is “epistemic conditional”). He considers it explicitly as a marker of evidentiality, since according to him the value of “borrowed” information prevails over the modal value of uncertainty and the alethic value of non-commitment for the truth of the information of this type of conditional¹.

Other linguists who are more concerned with a unified description of the semantics of the French conditional, i.e. the relation of the journalistic conditional to hypothetical and temporal meanings, question the legitimacy of postulating the existence of this category in French. For instance, Agnès Celle in her comparative analysis of French, English and German prefers to discard this category in the study of these languages. According to Celle, this type of conditional is rather a marker of non-commitment to the truth of the information (*non-prise en charge*). (Celle 2006)

On the other hand, the Estonian quotative, formed by adding the suffix *-vat* to the verb stem (*V-vat* for the present and *AUX-vat v-nud* for the past form), is an example par excellence of grammaticalized evidentiality. The latter is composed of the auxiliary verb *olema* ‘to be’ in present quotative followed by the past participle of the core verb. As has generally been suggested, this evidential form developed from the present participle, originally used as a complement of speech-act verbs that later came to act as the independent predicate of any clause (Erelt et al. 2006: 128, see also Sepper 2007 and Kehayov 2008). However, it has to be pointed out that the main marker of the quotative *-vat* is not exclusively reserved for this function: at the same time, Estonian uses it in its original function of present participle and also in the *vat*-infinitive forms.

¹ See Dendale (1993: 175): “Le conditionnel épistémique est avant tout un marqueur *évidentiel* parce que sa valeur de base – c’est-à-dire la valeur qui est toujours présente, qui n’est pas soumise à des variations et qui en plus détermine et explique les autres valeurs – est la valeur *évidentielle* d’emprunt.”

2.2. Semantic correspondences between the French conditional and the Estonian quotative

According to Mati Erelt, Helle Metslang and Karl Pajusalu (2006: 127), the *vat*-marked form, called “reported indicative” in that article, “shows that the speaker is not the source of the message but acts as the mediator of the message received from somewhere”. In addition to this central meaning, also mentioned for the French journalistic conditional by Dendale, Petar Kehayov (2008: 98) refers to its possible correlate meaning of uncertainty (“the cognitive remoteness of the source of information correlates with the degree of certainty”), so there is, as he points out, “an implicational relation between the meanings of evidentiality and epistemicity”. Kehayov (2008: 88) also mentions the correlation between the marking of the source of information and the degree of the speaker’s commitment to the truth of the proposition. All these meanings are pointed out by Laurent Gosselin (2005: 175) who sums up the semantic features assigned to the French journalistic conditional in different studies as follows:

- it appears in the declarative discourse with the aim of passing on information;
- the asserted information is obtained from a source other than the speaker;
- the speaker does not take any responsibility for the information;
- the information is introduced as uncertain;
- the uncertain character of the information is taken as provisional (it is expected to be confirmed).

All these features, except the last one, which is not usually mentioned by French linguists either, are thus common for the two forms in our focus.

2.3. The discursive context of the French journalistic conditional and the Estonian quotative

The name of the “journalistic conditional” refers directly to the type of discourse where it appears most often: it is particularly frequent in the French written media. It also occurs often in the scientific discourse but is relatively rare in fiction (see e.g. Aikhenvald 2006: 106–107 and Kronning 2002).

A recent corpus-based study by Maria-Maren Sepper (2007) of the share and dynamics of different means of expressing reported information in journalistic discourse and fiction shows that in contemporary Estonian the quotative is rare in the former and occurs more often in the latter. It has also been pointed out that the Estonian quotative is rather seldom used in everyday speech where other means occur instead for expressing indirect meaning (see Toomet 2000). It could thus be said that both forms in our focus are mostly used in the written language.

Unfortunately, no French-Estonian translated press corpus is available and a comparative study of these forms in journalistic discourse could thus only be made on the basis of comparable monolingual corpora—a corpus-based study would be a further aim of our research. However, our corpus also contains non-fiction texts, including those representing scientific discourse, which are likely to offer sufficient

evidence for setting out preliminary hypotheses on the equivalences and differences between the two forms in question.

It can be argued that the difference in the distribution of the “journalistic conditional” and the Estonian quotative does not necessarily indicate a semantic difference between these forms: it might also reflect divergences between the conventions of journalistic texts in both languages. For instance, Sepper (2007: 320) suggests in her article that “if the quotative is used, the source of information remains impersonal and the form does not comply with the requirements of journalistic text as a text type”. It is though also justified to argue that the disparity of the discursive context of these forms might result from differences in their meaning, since the rarity of quotative in journalistic discourse does not indicate that no reported information occur in Estonian written media but rather that other linguistic means are used to express it (cf. Sepper 2007).

3. The results of the corpus analysis

Our comparative analysis is based on the Estonian-French parallel text corpus². It consists of 2,140,872 running words and includes 5 different textual categories: Estonian fiction (50% of words), French fiction (24%), Estonian non-fiction (3%), French non-fiction (19%) and EU legislative texts (4%).³

The choice of corpus texts is inspired by 3 main principles: they are supposed to be contemporary, bilingual and diverse/representative, i.e. different authors and genres are used. As was noted before, the journalistic genre is regrettably unrepresented in the corpus due to the lack of translations. Presumably, the journalistic conditional would have been more precisely observable if the written French media texts were included.

The morphological analysis was made by means of ESTMORF search engine⁴ and based on the Estonian quotative retrieval.

3.1. General results of the corpus analysis

After the first selection only 82 quotative forms were collected by ESTMORF, 50 of them (61%) were present, and 32 (39%) past forms.⁵ This result shows that the quotative form does not appear very often even in fiction. As could be expected, it is completely absent in the EU legislative texts.

The principal results are shown in table 1 that first sorts the data by the direction of translation (that is from Estonian to French or vice versa) and secondly by the form of quotative (either present or past).

Despite the semantic and functional similarities assigned to the Estonian quotative and the French conditional in different studies, the equivalents of the Estonian quotative either in French translations or in original texts were very rarely in the conditional form. In total, in our corpus 12 quotative forms (in altogether 10 sentences) were translated into conditional, 7 of them were found in originally Estonian texts and 5 of them in French ones.

² The corpus is available at <http://corpus.estfra.ee/ee> (10.07.2009).

³ The corpus has been created in order to facilitate an extremely labour-intensive undergoing project, the composition of Estonian-French Dictionary.

⁴ The ESTMORF program is a morphological analyser and synthesizer of the Estonian language; for further information: <http://www.eki.ee/keeletehnoloogia/projektid/estmorf/> (30.12.2009).

⁵ One reason for these low frequencies may be the exclusion in the present study of the quotative forms manifested by the lone past participle (the simple past of the quotative mood).

Table 1. The Estonian quotative and its French equivalents in the ESTFRA corpus

Equivalents of Quotative in the Estonian-French translations		Original French forms translated into Quotative in the French-Estonian translations		All findings
Present Quotative in Estonian	41	Present Quotative in Estonian	9	50
Present Conditional	2	Present Conditional	5	7
Present Indicative	17	Present Indicative	1	18
Imperfect	16	Imperfect	1	17
Present Infinitive	2	Present Infinitive	1	3
Present Subjunctive	2	Present Subjunctive	–	2
Present Participle	1	Present Participle	–	1
Pluperfect	1	Pluperfect	–	1
Adjective	–	Adjective	1	1
Past Quotative in Estonian	31	Past Quotative in Estonian	1	32
Past Conditional	5	Past Conditional	–	5
Pluperfect	12	Pluperfect	–	12
Imperfect	6	Imperfect	1	7
Present Indicative	2	Present Indicative	–	2
Past Simple	2	Past Simple	–	2
Past Infinitive	1	Past Infinitive	–	1
Past Perfect	1	Past Perfect	–	1
Past Subjunctive	1	Past Subjunctive	–	1
Preposition	1	Preposition	–	1
	72		10	82

Regarding the French-origin texts, the quotative is rather modestly used by the Estonian translators. It is all the more significant that half of these 10 quotatives were originally the French conditional forms. 4 out of 5 quotative-conditional equivalents were found in the non-fiction texts: 2 in *Martin Luther: un destin* by Lucien Febvre and 2 in *La volonté de savoir. Histoire de la sexualité, I* by Michel Foucault. In a further study it would be interesting to see whether these texts contain journalistic conditional forms that have not been translated into quotative in Estonian.

In the Estonian-source texts, the verbs that occur in the present quotative in Estonian are mainly translated into the present and imperfect indicative forms in French (respectively 41% and 39%). The examples in the present infinitive (2 forms), present subjunctive (2 forms), present conditional (2 forms), present participle (1 form) and the pluperfect (1 form) seemed rather insignificant.

The quotative past form is also primarily translated into indicative forms, such as pluperfect (39%) and imperfect (19%). The past conditional is represented by 5 occurrences (16%), all of them are interpretable as the epistemic conditional, like in (3). This result is partly due to the above-mentioned difference of the usual genre where the Estonian quotative and the French journalistic conditional occur.⁶

It is quite remarkable that one translator seems to use the quotative-conditional relation much more often than the others: all the examples of quotative translated into the evidential past conditional come from Jean-Pascal Ollivry's translation of Karl Ristikivi's historic novel *Põlev Lipp / L'étendard en flammes*, see (3). The

⁶ This conclusion is also observed in the results of another bilingual corpus analysis carried out in 2008 (Alas, Treikelder in press): among the 179 conditional forms found in the French novel *Le Grand Meaulnes* none were classifiable as a 'journalistic conditional' with the principal function to express the reported evidentiality.

present conditional occurrence also comes from Ollivry (this time it concerns the novel by Tammsaare) but it seems to be more likely motivated by the sequence of tenses (6). Only one example out of 7 was found in the Antoine Chalvin's translation.

One can argue that the use of form can also depend on the original author's choices and not only on the translator's preference but we believe that both the original texts and the translator's choices have influenced the result. Firstly, it is clear that in Ollivry's translations the quotative-conditional equivalence is more frequent than in the others; but secondly, evidence suggests that there are also some objective constraints for the choice of the verb form: in the other quotative-rich sources, i.e. *Vastutuulelaev* by Jaan Kross (19 forms) and *Ajaloo ilu* by Viivi Luik (15 forms) there was only case of using conditional as a counterpart for quotative. The scarceness of conditional forms is presumably brought about by the particular genre or context that does not support the use of the epistemic conditional. This difference is observable in the examples (3) and (4) and will be re-discussed in the last chapter.

(3a) *Ühe sellise legendi järgi olevat* [be.AUX:QUOT] Konradin Pavia's uusaasta ööl **õnne valanud** [pour_fortune:PST.PTCP] ja sellest **olevat välja tulnud** [be.AUX:QUOT + come_out:PST.PTCP] üks imelik kuju, mida ta sõbrad igaiüks omamoodi **olevat katsunud** [be.AUX:QUOT + try:PST.PTCP] seletada. Üks **olevat leidnud** [be.AUX:QUOT + find:PST.PTCP], et see **kujutavat** [represent:QUOT] laeva, teine trooni, kolmas katafalki.

(3b) *Selon une de ces légendes*, Konradin **aurait interrogé** [have.AUX:COND.3SG + interrogate:PST.PTCP] le destin à Pavia dans la nuit du nouvel an, et la réponse **aurait revêtu** [have.AUX:COND.3SG + take_on:PST.PTCP] une forme énigmatique que ses amis **auraient tenté** [have.AUX:COND.3PL + try:PST.PTCP] d'interpréter chacun à sa façon. **Pour l'un** [for.PREP + the.SG.M + one.SG.M] **c'était** [be:IMPRF.3SG] un navire, pour l'autre un trône, pour un troisième un catafalque.

'According to one of these legends, Konradin is said to have melted lead [for predicting the future] at New Year's Eve and an enigmatic figure had come out of it that his friends had tried to interpret, everyone in his own way. One of them is said to have guessed that it reminded him of a ship, the second a throne and the third a catafalque.'

(Ristikivi-Ollivry)⁷

In the example (3) which is coming from a historical novel dominated by traditional chronological past tense narration the translator has in most cases chosen to use the epistemic conditional for the quotative forms. On the contrary, in the example (4) none of these forms has rendered the conditional form and the latter is regarded completely inappropriate in this context by the translator. This example comes from a novel representing non-linear narrative with a very complex temporal structure and a constant alternation of tenses.

(4a) See on seesama õun, millest kahekümne kahe aasta pärast talle räägib linalakk Maarja, surnu tütar, seesama, kes on rootsi keeles kirjutanud 280 lehekülge eesti elust ja surmast. Isa tapeti NKVD poolt 1944.?! aastal.

⁷ The source of the example is given according to the author's and the translator's name, for further details see the corpus bibliography. The English translations have been provided by the authors of this article.

Ühes majas, mis on praegugi alles. Seda maja on Maarjale Eestis näidatud. See **olevat** [be:QUOT] igav ja harilik maja, mitte sugugi mõrvapaiga moodi, köögis **tehtavat süüa** [cook:QUOT] ja toas **seisvat** [stand:QUOT] televiisor. Maja taga aias aga **kasvavat** [grow:QUOT] vana kõver õunapuu suurte valgete õuntega, mis **kumavat** [shine:QUOT] nagu kuu.

Selle õunapuu all **olevat** [be:QUOT] üks lapp maad, kus aeg seisab. Seda, et ta seisab, **olevat** [be:QUOT] kohe näha, isegi palja silmaga. Sinna ei saja vihma, ja kui keegi midagi ütleb, siis sinna see ei kosta.

- (4b) C'est de cette même pomme que lui parlera, vingt-deux ans plus tard, la blonde Maarja – fille d'un défunt –, qui a écrit en suédois deux cent quatre-vingts pages sur une vie et une mort estoniennes. Son père a été tué par le NKVD en 1944, dans une maison qui existe encore aujourd'hui. On l'a montrée à Maarja lors d'un séjour en Estonie. C'est [be:PRS.IND.3SG] une maison triste et ordinaire, qui ne ressemble en rien au lieu d'un assassinat: dans la cuisine, on **prépare** [prepare:PRS.IND.3SG] à manger, et il y a [have:PRS.IND.3SG] une télé dans la salle de séjour. Mais dans le jardin, derrière la maison, **pousse** [grow:PRS.IND.3SG] un vieux pommier tordu chargé de grosses pommes blanches qui **luisent** [shine:PRS.IND.3PL] comme la lune.

Sous cet arbre **se trouve** [be_located:PRS.IND.3SG] une parcelle de terre où le temps s'est arrêté. On le **remarque** [notice:PRS.IND.3SG] immédiatement, d'un simple coup d'œil. Il n'y pleut jamais et quand quelqu'un y parle, on ne l'entend pas.

'It is the same apple about which blond Maarja, daughter of the one deceased, will tell him twenty-two years later; the same girl who has written 280 pages of Estonian life and death in Swedish. Her father was killed by NKVD in 1944 ?!. In a house that still exists. Maarja has been shown this house while being in Estonia. **It is** an ordinary and dreary house, not like a murder scene at all; **they cook** in the kitchen and **there is** a TV in the living room. In the backyard **there is** a curvy apple tree with big white apples **that shine** like the moon.

Under the tree **there is** a piece of land where time stops. **It can be seen** right away that time has stopped there. It does not rain in that spot; and if someone speaks, it is not heard.' (Luik–Chalvin)

3.2. The equivalents of the Estonian quotative in French according to the translated corpus

The general results of the corpus analysis have thus revealed that the French conditional form is merely one of the possible equivalents for the Estonian quotative and that in most of the cases the verb is in indicative in French texts. In this section we shall examine more closely what kind of equivalents can be found for quotative in translations. We can distinguish between four **situations**:

- 1) the French equivalent is the conditional form;
- 2) other means (mostly lexical) are used in French for referring to the information being from a remote source;
- 3) French combines conditional with another evidential marker;
- 4) no explicit evidential means are used in French.

The first situation could already be observed in the Ristikivi–Ollivry’s paragraph (3) where the quotative is used 5 times in a row. The past quotative verbs *olevat valanud* ‘is said to have poured’, *olevat välja tulnud* ‘is said to have come out’, *olevat katsunud seletada* ‘is said to have tried to explain’ render the French past conditional form: *aurait interrogé*, *aurait revêtu*, *aurait tenté d’interpréter*. However, the two last ones are translated differently: the verb *olevat leidnud* ‘is said to have found’ has been omitted and replaced by an adverbial expressing the source of information of the next predicate (*pour l’un* ‘according to the first’) and *kujutavat* ‘is said to represent’ renders the indicative imperfect form *était* ‘was’ accompanied by a lexical tool for expressing the source of information *pour l’un*. It is to be noted that the last verb is the only one in **present** quotative and the source of information is rather one of the persons mentioned in the legend—unlike in the case of the first three predicates of the example where the source of information is the legend itself.

The example (3) thus also offers an illustration for situation 2 described above where the verb is in indicative in French and there are other means referring to the source of information. The example also shows that a lexical evidential tool can occur along with the conditional form in French and the quotative in Estonian (situation 3): the whole paragraph is introduced by the construction *ühe sellise legendi järgi / selon une de ces légendes* ‘according to one of these legends’. The co-occurrence of such a lexical tool and the conditional or the quotative form is frequent both in French and in Estonian.⁸

Nevertheless it is important to note that this introduction itself does not evoke the presence of the conditional in French because many similar constructions followed by the indicative tenses are also found in our corpus, see (5) for pluperfect use.

- (5a) *Teise jutu järgi olevat* [be:QUOT] ta aga Pisa laevadega teel Sitsiiliasse.
 ‘According to another legend, he **is said to be** on his way to Sicily with Pisa’s ships.’
- (5b) *Selon d’autres, il s’était embarqué* [board:PLP.3SG] à Pise pour la Sicile.
 ‘According to others, he **had boarded** in Pisa in the direction of Sicily.’
 (Ristikivi–Ollivry)

The use of the conditional as the equivalent of the Estonian quotative can also be due to other reasons than the indication of a different source of information. In the next example, the conditional has rather a temporal meaning, motivated by the rule of the sequence of tenses, considering that the verb in the conditional

⁸ According to Dendale (1993: 170), if the source marker (for example *selon quelqu’un* or *quelque chose*) and the epistemic conditional are used simultaneously, the latter rather expresses the uncertainty of the information than the loan of someone’s words.

form appears in a subordinate clause depending on a main clause in the past tense (*envoya* ‘sent’) and is accompanied by the future time reference (*le lendemain matin* ‘the next morning’).

- (6a) Õhtul *heitis* kirikumees vara magama ja *ajās* ka Milli asemele, sest muidu **ei saavat** [NEG+CAN:QUOT] ta homme maast lahti.
- (6b) Le soir, le bedeau *alla* se coucher tôt et *il envoya* Milli dormir elle aussi, sans quoi elle **ne pourrait pas** [NEG+CAN:FUT.PST.3SG] décoller de terre le lendemain matin.

‘In the evening, the verger *went* to bed early and *sent* Milli to sleep as well because otherwise she **would not come** alive the next morning.’

(Tammsaare–Ollivry)

This conditional form, the so-called “future in the past”, is used twice in our corpus and is the only one that occurs in a complement clause among quotative-conditional equivalents.

In the examples without conditional but containing other evidential means, two kinds of solutions were found: lexical tools already mentioned above (*selon X* or *pour X* ‘according to X’, etc.) or diverse introductory verbs (*dire* ‘to say’, *raconter* ‘to tell’, *accuser* ‘to accuse’, etc.) used with different definite and indefinite sources (*Schmidt* and *Courant* versus *certains* ‘some people’, *un vieil adage* ‘an old saying’, etc.). Mostly, these means were also present in the Estonian texts along with the quotative form.

Among the 10 occurrences of quotative in the French–Estonian translations, in 5 cases the conditional did not occur in the original text (the present, the imperfect, a present infinitive and an adjective were used instead). The Estonian translator also chose to use the quotative form, in addition to the plain introductory verbs and the lexical solutions like *sõnul* ‘after somebody’s words’, *arvamus* ‘(somebody’s) opinion’.

Example (7) shows how in the original text the only evidential tool is the introductory verb of the reported speech, but the Estonian translator uses the quotative mood:

- (7a) Mon oncle Charles Gide me **dit** [tell:PRF.3SG], par la suite, que Tancrède Gide, mon grand-père, dans les derniers temps de sa vie, **consultait** [consult:IMPRF.3SG] toutes sortes de médecastres et de charlatans.
- (7b) Onu Charles Gide **rääkis** [tell:PST.3SG] hiljem, et elu lõpu poole **olevat** [be.AUX:QUOT] minu vanaisa Tancrède Gide **pöördunud** [adress:PST.PTCP] abi saamiseks igat sorti ravitsejate ja posijate poole.

‘My uncle Charles Gide **told** me later that during his last years, my grandfather **called** upon every kind of soothsayers and charlatans.’

(Gide–Tomasberg)

In the example (8) the French original does not have any explicit evidential indication (except the colon), but the translator has still used the quotative.

(8a) Le petit Blaise vint me chercher quelques minutes avant quatre heures: sa tante **désirait** [desire:IMPRF.3SG] me parler.

(8b) Mõni minut enne nelja tuli väike Blaise mind kutsuma: tema tädi **soovivat** [desire:QUOT] minuga rääkida.

‘Little Blaise came to call me some minutes before 4 pm: his aunt **wanted / was said to want** to talk to me.’ (Gide–Tomasberg)

The last example illustrates the fourth situation mentioned in the beginning of this chapter: no explicit evidentiality markers appear in the French text. This situation can also be observed in the example (4) presented above.

In the next chapter we shall discuss some examples in order to find out whether these non-correspondences in translations are due to some specific linguistic constraints or a semantic non-equivalence.

4. Discussion of examples

The juxtaposition of the sentences containing the Estonian quotative and their equivalent sentences in French seems to point to a certain regularity in the syntactic behaviour of these forms. One of the possible French equivalents for the Estonian quotative was the use of an introductory (speech act) verb followed by a complement clause with a predicate in indicative, like in the example (9).

(9a) Nii **olevat** [be.AUX:QUOT] neil **olnud** [be:PST.PTCP] eriline oskus ennast nähtamatuks teha ja sel kombel **olevat** [be.AUX:QUOT] nad prantslaste rünnakutsed alati **nurja ajanud** [ruin:PST.PTCP], kuna nad viimaste piiramistornid hoolsale valvele vaatamata olid saanud põlema pista ja seda tuld **ei olevat saanud** [NEG+be.AUX:QUOT + can:PST.PTCP] **kustutada** muuga kui pühitsetud veega, mida aga raske oli parajal ajal saada.

(9b) *On racontait* [tell:IMPRF.3SG] par exemple qu'ils **étaient** [be:IMPRF.3PL] **capables** de se rendre invisibles et qu'ils **avaient ainsi réussi** [succeed:PLP.3PL] **à déjouer** toutes les tentatives d'assaut des Français en parvenant, malgré une garde serrée, à incendier les tours de siège de ces derniers – et les feux qu'ils allumaient **ne pouvaient** [NEG+can:PLP.3PL] **s'éteindre** qu'avec de l'eau bénite, difficile à trouver au moment voulu.

‘So they **are said to have had / it was said that they had** a special gift to make themselves invisible and that they **had** thus always **ruined** the French attack attempts, because they could set the French siege towers on fire despite the close guard, and (**it is said that**) this fire **could not have been put out** with anything else but the sacred water that was difficult to get at the right time.’ (Ristikivi–Ollivry)

In this sentence, the translator adds *on racontait* ‘it was told’ and the verbs in quotative in the original (*olevat neil olnud* ‘they are said to have had’ and *tuld ei olevat saanud kustutada* ‘it is said that the fire could not have been put out’) are in imperfect indicative (*étaient, pouvaient*). Unlike the example above, in most of the

cases where the introductory verb refers to the source of information, it is present also in Estonian but the predicate is nevertheless in quotative (10).

(10a) Perenaine *arvas* [think:PST.3SG], et kella viie ajal pärast lõunat **ei viibivat** [NEG+be:QUOT] ta seal kunagi.

(10b) Elle nous *répondit* [answer:PRF.3SG] qu'il **ne s'y trouvait** [NEG+be_located.IMPRF.3SG] jamais à cinq heures de l'après-midi.

'The hostess *guessed* that he **would never be** there at 5 p.m.'

(Kross–Moreau)

On the other hand, in the examples where the journalistic conditional is used in French the predicate in that form **never** occurs in a complement clause. The journalistic conditionals of our corpus are used in independent clauses or in relative clauses, see the example (3) that contains both clause types. The only example of conditional in a complement clause seems to offer a temporal reading:

(11a) je me souviens spécialement d'une phrase où il *était dit* [tell:PASS.IMPRF.3SG] que son geste **ne suffirait pas** [NEG+suffice:FUT.PST.3SG], hélas ! à couvrir l'arriéré ...

(11b) iseäranis on mul meeles üks lause, mis *ütles* [tell:PST:3SG], et Abeli žestist paraku **ei piisavat** [NEG+suffice:QUOT] maksevõla kustutamiseks ...

'I can especially remember one sentence that *said* that Abel's gesture **was** unfortunately **not sufficient enough** to erase his debt ...'

(Gide–Tomasberg)

On the contrary, in the cases where the journalistic conditional is not used, the complement clauses are rather frequent—we have found 22 sentences on the whole, plus examples where the predicate is not formally in a complement clause, but accompanied by a parenthetical clause (*disait-il, paraît-il*), like in the example (12).

(12a) ta **kadestavat** [envy:QUOT] mind, sest siin **ei toimuvat** [NEG+happen:QUOT] ammu enam midagi: seisev vesi.

(12b) Il m'*enviait* [envy:IMPRF.3SG], car ici, **disait-il** [tell:IMPRF.3SG], il **ne se passe plus** rien [NEG+happen:PRS.IND.3SG] depuis longtemps: tout est comme une eau dormante.

'He **envies** me because, as he says, **nothing** ever **happens** here: still water.'

(Tode–Chalvin)

The French journalistic conditional appears to behave differently with respect to subordination but it seems that this is not a mere linguistic constraint but rather points to some particular and essential properties of the French conditional in comparison with the evidentiality marking in Estonian.

The Estonian quotatives without conditional equivalents also occur very often in independent clauses. The majority of these clauses represent the so-called "free indirect speech" where the speech or the point of view of a character is presented without explicit speech act verbs and markers of subordination, like in the example (13).

(13a) Nende aktsiate omamine **ei sobivat** [NEG+accommodate:QUOT] küll kokku tema üldiselt pahempoolsete vaadetega, aga lõppude lõpuks **olevat** [be:QUOT] see ükskõik, kelle käes nad on, see **ei muutvat** [NEG+change:QUOT] midagi.

(13b) Ces actions **ne s'accordaient guère** [NEG+accord:IMPRF.3PL] avec ses opinions de gauche, mais en fin de compte l'identité de leur propriétaire **importait** [matter:IMPRF.3SG] peu, cela **ne changeait rien** [NEG+change:IMPRF.3SG] à l'affaire.

'Having these shares is said **not to fit** with his mainly left-leaning political views but in the end it **does not matter** who has them, it **changes** nothing.'
(Tode–Chalvin)

In this example, the predicates (*ei sobivat*, *olevat*, *ei muutvat*) are all in quotative in Estonian and they are all translated by using imperfect which is in the sequence of tenses for marking simultaneity to a past reference point and reckoned to be a form par excellence for expressing subjectivity (see also example (8) for a similar case).

The actions of complement clauses including indirect speech as well as those of “free indirect speech” are interpreted from a point of view internal to the text. Their perception is thus subordinated to this source of information which is represented by a character created in the text and the narrator’s role is merely that of mediator. These utterances often contain other linguistic means only interpretable from the character’s point of view, like the emphatic particle *küll* and the adverbial *lõppude lõpuks* ‘in the end’ in the example (13).

Thus, according to our corpus-study, the non-correspondence between the French conditional and the Estonian quotative seems to result from their relation to the speaker acting as the mediator of the information: the French conditional seems to be incompatible with the internal point of view (particularly frequent in fiction), while the Estonian quotative can be used to express both external and internal points of view.

This observation seems to corroborate Agnès Celle’s viewpoint on the meaning of the French journalistic conditional: she argues that even if this form expresses the non-commitment of the speaker (the enunciator) to the truth of the proposition, the speaker still acts as the main structuring centre of the utterance, and thus, the French conditional, unlike the Estonian quotative, always remains related to the primary source of the utterance.

We recall here the examples (3) and (4) presented above (in the section 3.1.) in which the quotative forms were regularly translated in a different way (epistemic conditional *versus* present indicative) and this difference seemed to refer to a more profound reason than the accidental choice made by the translators. The translated forms found in the Karl Ristikivi’s historic novel (3) report the so-called “hearsay” facts from the past and the journalistic conditional seems a perfect option to express this. On the contrary, the narration in Viivi Luik’s paragraph (4) is presented from the character’s point of view which does not support the intervention of the enunciator as a structuring centre that the use of conditional evokes.

According to Celle (2006: 84), it is not the reported character of the information (the marking of a second point of view) that prevails in the meaning of the French

journalistic conditional but rather the comment of the enunciator.⁹ The Estonian quotative, on the contrary, does not necessarily give this central position to the main source of enunciation because in this form the indirectness of the information is in the foreground and the meanings of uncertainty and non-commitment can merely be considered as corollary values appearing in different degrees according to the context (see also Kehayov 2008: 88). This difference seems to explain the almost total absence of the French journalistic conditional in fiction where the “effacing” of the narrator, i.e. the main origin of enunciation, is a regular procedure. On the other hand, this difference is not sufficient for explaining the rareness of the Estonian quotative in the written media that should be studied on the grounds of a different corpus.

5. Conclusion

In comparison with the various typological and monolingual studies of the Estonian quotative and the French conditional, our corpus-study on the translations of the Estonian quotative form into French has given somewhat surprising results on the alleged semantic and functional affinity between these forms. The corpus-study has revealed that in translations the French conditional appears very rarely as the equivalent of the Estonian quotative. This result might be due to two reasons:

- 1) As both Estonian and French have, besides the forms in question, other tools for expressing the reported character of the information, the translator makes the choice according to his/her personal preferences and/or the discursive or stylistic constraints imposed by the type of text (namely fictional or journalistic/scientific in our case);
- 2) There are linguistic restrictions resulting from the semantic or pragmatic difference between these forms.

Both of these explanations seem to apply to our translated corpus. On one hand, our corpus mainly contains fictional texts where the French conditional is generally not used. Additionally, some translators seem to avoid the conditional form and always use other ways for expressing the reported nature of the information.

A closer look at the examples reveals differences in the behaviour of these forms in relation to the syntactic relations and the phenomena of point of view and focalisation that could also actually explain the discursive and stylistic differences described in the first point. It namely seems that the Estonian quotative is completely neutral with respect to the focalisation of the utterance, allowing the presentation of the actions both from the narrator’s (i.e. the primary source of information) and a character’s point of view (internal to the text). The French conditional, on the contrary, maintains the position of the narrator as the structuring centre of the utterance and is thus incompatible with the internal point of view that supposes the “erasing” of the narrator.

This study could and should be complemented in many aspects. Some of them have already been mentioned above. Studies of different corpora (for instance a journalistic comparable corpus) would certainly offer important supplementary data to the understanding of the meaning and the functioning of the two forms. It

⁹ See Celle (2006: 84): “Ce n’est pas la reprise du point de vue de l’autre qui est au premier plan avec le conditionnel, mais le commentaire de l’énonciateur-origine que ne peut éliminer le décrochage fictive sur une coordonnée”. (It is not the repetition of somebody else’s point of view that is in the foreground in the use of conditional but rather the comment of the enunciator or origin that cannot be eliminated by the fictive disconnection on one of the enunciated coordinates.)

is also obvious that the study of evidentiality in general cannot only be confined to these two forms: both French and Estonian have several other means related to this category that should be studied in a comparative perspective. The findings of our first research in this field will hopefully give some starting points to these further studies.

Abbreviations

3	3rd person	PL	plural
AUX	auxiliary	PREP	preposition
COND	conditional	PRF	perfect
FUT.PST	future in the past	PRS.IND	present indicative
IMPRF	imperfect	PST.PTCP	past participle
M	masculine	PST	past
NEG	negative	QUOT	quotative
PLP	pluperfect	SG	singular
PASS	passive		

Corpus bibliography

- Gide, André 1955. *Si le grain ne meurt*. Paris: Editions Gallimard.
- Gide, André 2006. *Surra, et elada*. Tallinn: Varrak. (Translated by Leena Tomasberg.)
- Kross, Jaan 1987. *Vastutuulelaev*. Tallinn: Eesti Raamat.
- Kross, Jaan 1997. *L'œil du grand tout*. Paris: Editions Robert Laffont. (Translated by Jean-Luc Moreau.)
- Luik, Viivi 1991. *Ajaloo ilu*. Tallinn: Eesti Raamat.
- Luik, Viivi 2001. *La Beauté de l'Histoire*. Paris: Éditions Christian Bourgois. (Translated by Antoine Chalvin.)
- Ristikivi, Karl 1990. *Põlev lipp*. Tallinn: Eesti Raamat. (1st edition: 1961.)
- Ristikivi, Karl 2005. *L'étendard en flammes*. Éditions Alvik. (Translated by Jean Pascal Ollivry.)
- Tammsaare, Anton Hansen 2003. *Tõde ja õigus* (1. osa). Tallinn: Avita. (1st edition: 1926.)
- Tammsaare, Anton Hansen 2009. *La Colline-du-Voleur (Vérité et Justice, 1)*. Larbey: Gaïa éditions. (Translated by Jean Pascal Ollivry.)
- Tode, Emil 1993. *Piiriirik*. Tallinn: Tuum.
- Tode, Emil 1997. *Pays frontière*. Paris: Gallimard. (Translated by Antoine Chalvin.)

References

- Aikhenvald, Alexandra Y. 2006. *Evidentiality*. Oxford: Oxford University Press.
- Alas, Reet; Treikelder, Anu in press. *Remarques sur le conditionnel en français et en estonien: variations modales dans la traduction*.
- Erelt, Mati; Metslang, Helle; Pajusalu, Karl 2006. *Tense and evidentiality in Estonian*. – B. Cornillie, N. Delbecque (Eds.). *Topics in Subjectification and Modalization*. *Belgian Journal of Linguistics*, 20, 125–136.
- Celle, Agnès 2006. *Temps et modalité. L'anglais, le français et l'allemand en contraste*. Bern etc.: Peter Lang.
- Dendale, Patrick 1993. *Le conditionnel de l'information incertaine: marqueur modal ou marqueur évidentiel? – Hilty Gerold (Ed.)*. *XXe Congrès International de Linguistique et Philologie Romanes, Tome I, Section I*. Tübingen: Francke, 165–176.

- Dendale, Patrick; Coltier, Danielle 2003. Point de vue et évidentialité. – Cahiers de praxématique, 41, 105–129.
- Gosselin, Laurent 2005. Temporalité et modalité. Bruxelles: Editions Duculot.
- Haillet, Pierre Patrick 2002. Le conditionnel en français: une approche polyphonique. Paris: Ophrys.
- Kehayov, Petar 2008. An Areal-Typological Perspective to Evidentiality: the Cases of the Balkan and Baltic Linguistic Areas. *Dissertationes linguisticae Universitatis Tartuensis*, 10. Tartu: Tartu University Press.
- Kronning, Hans 2002. Le conditionnel “journalistique”: médiation et modalisation épistémique. – *Romansk forum*, 16, 2, 561–575.
- Sepper, Maria-Maren 2007. Indirectal in literary Estonian. – *Trames*, 11, 3, 299–323.
- Toomet, Piret 2000. Mõnest kaudsuse väljendamise võimalusest tänapäeva eesti keeles. – *Keel ja Kirjandus*, 42 (1), 251–259.

Online references

- Estonian-French parallel text corpus ESTFRA. <http://corpus.estfra.ee/ee> (10.07.2009).
- ESTMORF. <http://www.eki.ee/keeletehnoloogia/projektid/estmorf/> (30.12.2009).

Reet Alas (University of Tartu). Her research interests are comparative linguistics, conditional and modality in Estonian and French.
reet.alas@ut.ee

Anu Treikelder (University of Tartu). Her research interests are comparative linguistics, semantics of tenses, historical linguistics.
anu.treikelder@ut.ee

MÄRKUSI KAUDSE EVIDENTSIAALSUSE KOHTA PRANTSUSE JA EESTI KEELES: KONTRASTIIVNE LÄHENEMINE

Reet Alas, Anu Treikelder

Tartu Ülikool

Artikli eesmärk on võrrelda kahe konkreetse vormi toimimist geneetiliselt kaugetes keeltes: prantsuse keele “ajakirjanduslikku konditsionaali” ja eesti keele kvotatiivi. Mõlemaid vorme on tüpoloogilistes ja kummagi keele uurimustes määratletud kui evidentsiaalsuse markereid või strateegiaid: kui eesti keele grammatiseerunud kvotatiiv on spetsialiseerunud kaudse informatsiooni väljendamisele (vt Erelt jt 2006, Sepper 2007, Kehayov 2008), siis prantsuse konditsionaali puhul on see hüpoteetilise ja ajalise kõrval vaid üks tema põhitähendustest (Haillet 2002).

Tüpoloogilistes uurimustes (Aikhenvald 2006) on mõlemat vormi peetud kaudse evidentsiaalsuse väljendajaks, kuid eesti kvotatiivi ja prantsuse “ajakirjandusliku konditsionaali” puhul on täheldatud (Sepper 2007, Kronning 2002) erinevusi kasutus kontekstis (esimest kasutatakse pigem ilu- kui ajakirjanduses ning teist just pressi- ja teadustekstides). Käesoleva analüüsi aluseks on vormide kirjelduste semantilised sarnasused: lausungi eesmärk on info edastamine, edasiantav info on pärit kelleltki teiselt kui lausujalt, lausuja ei vastuta info eest, info on tõlgendatav ebakindlana (Erelt jt 2006, Kehayov 2008, Gosselin 2005).

Peamiselt erinevaid tekstitüüpe sisaldavast paralleelsest tõlkecorpusest leitud kvotatiivvormid olid ülaltäheldatud semantilistele ja funktsionaalsetele sarnasustele vaatamata väga harva prantsuskeelses versioonis konditsionaalis.

Detailsem lingvistilise konteksti uurimine kinnitas erinevust uuritavate vormide kasutuses vastavalt tekstis avalduvale vaatepunktile: kui prantsuse konditsionaali puhul esitatakse informatsioon alati lausumise põhiallika (ilukirjanduslikus tekstis jutustaja kui informatsiooni vahendaja) vaatepunktist, siis eesti kvotatiivi saab kasutada nii tekstisisese tegelaskuju kui ka jutustaja vaatepunkti väljendamisel. See erinevus andis seega ka võimaliku seletuse prantsuse “ajakirjandusliku konditsionaali” puudumisele ilukirjandustekstides.

Kuna eesti ja prantsuse keeles on lisaks eelnimetatud vormidele evidentsiaalsuse väljendamiseks veel mitmeid võimalusi, võib kvotatiivi konditsionaalis vastete vähesus tuleneda nii tõlkija isiklikest eelistustest kui tekstitüübi diskursiivsetest ja/või stilistilistest piirangutest. Tõlkelause analüüs osutab aga ka semantilistele ja pragmaatilistele erinevustele nende vormide vahel.

Võtmesõnad: kõrvutatav lingvistika, kvotatiiv, konditsionaal, corpuseanalüüs, prantsuse keel, eesti keel

EESTI KEELE AJAKATEGOORIA OMANDAMISEST – AJAVORMID JA AJASÕNAD

Reili Argus, Sirli Parm

Ülevaade. Artikkel käsitleb 135 normaalse keelelise arenguga 4–7-aastase eesti keelt emakeelena kõneleva lapse ajakategooria omandamist. Psühholingvistilise eksperimendi põhjal tehtud analüüs hõlmab nii ajavorme kui ka ajasõnu ning uurimuse eesmärgiks on selgitada, millal areneb eesti lastel välja selge opositsioon mineviku ja mittemineviku vahel, kas ajasõnad aitavad grammatilise aja esitust paremini mõista ning millised ajasõnad on selgemini tajutavad. Analüüsitulemustest selgub, et oleviku ja mineviku eristamine on eesti lastel omandatud kuuendaks eluaastaks, ajasõnad toetavad grammatilise aja mõistmist ja kasutamist ning et eri ajasõnad aitavad grammatilist aega mõista ja kasutada erineval määral. Võrreldes poiste ja tüdrukute tulemusi, võib veel väita, et aja omandamises on soost tulenevad erisused vähemärgatavad.*

Võtmesõnad: keeleomandamine, ajakategooria, ajasõnad, grammatiline aeg, eesti keel

1. Sissejuhatus

Emakeelt omandavate laste ajasüsteem alustab arengut kognitiivsest keskmest *siin ja nüüd*. Mineviku ja oleviku grammatilist eristamist alustavad lapsed paljudes keeltes esialgu mineviku prototüüpselt tuumast ehk teisisõnu: esimesed laste minevikuvormid väljendavad enamasti perfektivsust ning alles mõne aja pärast hakkab aspektikategooria ajakategooriast tasapisi eralduma (vt nt Weist jt 1991: 68 või eesti keele kohta Argus 2007). Inglise keele kohta on ühelt poolt väidetud, et kolmandaks eluaastaks on lastel täiskasvanupärane ajasüsteem omandatud (nt Weist jt 1991: 69), teisalt aga, et mineviku ja oleviku eristamine omandatakse kolmanda ja neljanda eluaasta vahel (Valian 2006: 267), soome keele kohta saadud andmete järgi omandavad soome lapsed ajakategooria mõnevõrra hiljem kui

* Artikli valmimisel on toeks olnud Eesti Teadusfondi grant nr 7492 "3–7-aastaste laste keelelise arengu leksikaalsed ja grammatilised seaduspärad" ning üleeuroopaline projekt COST A33 "Crosslinguistically Robust Stages of Language Development". Täname Kõrveküla Päikeseratta, Roiu Kukupai, Tallinna Mikumanni, Saku Päikesekillu ja Värska lasteaeda.

inglise või poola keelt kõnelevad lapsed (Weist jt 1991: 78). Ajakategoriat hakatakse omandama ajaliste suhete väljendamiseks ning see kujuneb välja üheaegselt verbimorfoloogia (grammatiline aeg ja aspekt) ja ajasõnade tähenduste mõistmise arenguga (Weist 2002: 206). Weist väidab samas, et esmalt hakkavad lapsed kasutama selliseid ajasõnu, mis täpset ajahetke ei väljenda (nt *juba*, *veel*, *varsti*) ja alles seejärel ilmuvad lapse kõneste konkreetsele ajale viitavad sõnad (nt *täna*, *homme*). Võrreldes teise keele omandamise uurimustega (Dietrich jt 1995) saab siinkohal välja tuua erinevuse: teise keele õppijad eelistavad ajasuhete väljendamisel pigem konventsionaalse väärtusega ajasõnu, nagu näiteks *kuu*, *aasta*, *homme*, *eile*, *täna*. Erinevuse põhjus seisneb selles, et keele omandamisel rakendavad lapsed spontaanselt sobivaid operatsiooniprintsiipe (Slobin 1985), et leida ajaliste suhete väljendamiseks verbide morfoloogiline kategooria ja neile vastav leksikaalne väljend. Slobini (1985: 1169–1171) järgi rakendavad lapsed selleks vastenduse (ingl *mapping*) sõnaraamatu printsiipi, mille järgi peab pöörama tähelepanu sellistele sõnadele ja mõistetele, millel on hõlpsasti määratav tähendus, ning memoreerima nende vastendus sobivate grammatiliste vahenditega. Sellest saab järeldada, et komplekssema tähendusega ajasõnade (nt *jälle*) mõistmine ja sobiva verbi ajavormiga vastavusse viimine on laste jaoks keerukam kui teiste selgepiirilistate ajasõnade (nt *praegu*) mõistmine.

Eesti keele spontaanse kõne andmete põhjal võib arvata, et kolmandaks eluaastaks on lastel omandatud ainult ajasüsteemi grammatiline tuumik, olevik ja lihtminevik (vt Vihman, Vija 2006), esimesed täisminevikuvormid ilmuvad lapse kõneste alles kolmandal–neljandal eluaastal (Argus 2007) ning enneminevikuvorme ei ole spontaanse kõne lindistustest üldse leida. Seega võib oletada, et eesti lastel ei pruugi kolmeaastaselt grammatilised aja vormistamise vahendid kaugeltki veel omandatud olla. Samas on andmeid selle kohta, et eesti laps suudab leksikaalsete vahendite ehk ajasõnade abil väljendada nii olevikku (*praegu*), minevikku (*eile*) kui ka tulevikku (*homme*) juba kaheaastaselt (Argus 2008: 51).

Sellised uuringud, kus oleks vaatluse all aja omandamise grammatiliste vahendite kõrval ka leksikaalsed ning kus uuritaks eesti keele ajasüsteemi arengut nelja-aastastel ja vanematel eesti lastel, esialgu puuduvad. Ka ei ole eesti keele ajakategooria omandamise uuringutes seni kasutatud eksperimentaalset metoodikat, mille abil on võimalik saada keeleandmeid just mõne konkreetse kategooria või vormi omandamise kohta eri semantilistes tingimustes ning vältida tervikuna omandatud, näiteks mõne väljendi koosseisus esinevate ajavormide sattumist analüüsitava keelendite hulka. Et teiste keelte omandamise uurimisel kasutatud eksperimendid on andnud keeliti üsna erinevaid tulemusi (vt nt Weist jt 1991), siis on eksperimendi abil saadud eesti keele ajakategooria omandamisandmed eriti väärtuslikud.

Siinse uurimuse **üldisem eesmärk** on selgitada välja, kas ja mis vanuses suudavad eesti keelt kõnelevad lapsed kasutada verbi ajavorme mingit aega tähistavas situatsioonis; millal areneb välja eesti keele ajakategooria põhiliigendust silmas pidades selge opositsioon mineviku ja mittemineviku vahel; mis vanuses asetavad lapsed verbi ajavormid ja ajaväljendid vastavusse; ning kas ajasõnad aitavad grammatilist aega tajuda ja kasutada. Näiteks testisid Laurence B. Leonard ja Laurie R. Krantz inglise keelt emakeelena kõnelevate 4–7-aastaste laste ajasüsteemi omandamist eesmärgiga uurida, kuidas mõjutab minevikuvormide kasutamist see,

kui lauses on muud ajainfot ehk siis ajasõnu. Nad tõestasid, et verbi minevikuvormide omandamine on suuresti sõltuv leksikaalse ajakategooria omandamisest. Leonard ja Krantz leidsid, et ajaadverbide kasutamise puhul kippusid väiksemad lapsed verbi ajavorme vältima, justkui oleks ajasõnast üksi aja väljendamiseks juba küllalt. (Krantz, Leonard 2007: 137)

Uurimuse **alameesmärgid** on teha kindlaks: 1) kuidas lapsed liigendavad mitteminevikku, s.t kas ja mis vanuses eristavad eesti keelt omandavad lapsed tulevikku olevikust; 2) milliseid grammatilisi ja leksikaalseid vahendeid kasutavad lapsed tulevikku väljendamiseks, nt kas lapsed tajuvad ja kasutavad *hakkama*-konstruktsiooni kui tulevikku tähistavat vahendit; 3) kas ajakategooria omandamisel on märgata laste soost tulenevaid erisusi; 4) milliste minevikku, olevikku ja tulevikku paigutuvate leksikaalsete ajaväljendite abil tajutakse ja moodustatakse ajavorme kõige paremini.

2. Eesti keele aja väljendamise grammatilistest ja leksikaalsetest vahenditest

Grammatiline ajasüsteem on eesti keeles üsna väike ja kompaktne: oleviku ja lihtminevikü kõrval on täis- ja enneminevikuvormid küll süsteemi osana olemas, kuid lastele suunatud keeles üsna marginaalselt esindatud (Argus 2007: 20). Olevikuvormide kasutusala on eesti, nagu ka paljudes teistes, nt inglise keeles (Valian 2006: 252) tunduvalt laiem kui minevikuvormide kasutusala. Kuigi verbi ajavormid annavad sündmusele üldise ajaorientatsiooni kõne- ja vaatlushetke suhtes, on sündmuse täpsem lokaliseerimine verbiga etteantud ajapiirkonnas võimalik alles ajamääruse abil ja need täpsustavad sündmuse toimumisaega verbi lihtminevikü ja oleviku vormi kõigi kasutusjuhtude korral (EKG II: 79–80). Tulevikku traditsioonilise grammatilise kategooriana eesti keeles ei ole. Et keeled teevad enamasti vahet minevikü ja mitteminevikü vahel, jääb tulevikule üks osa mitteminevikust. Tulevat aega ei liigendata aga keeltes nii põhjalikult kui möödunud aega. Helle Metslangi väitel peab tulevikust rääkija arvestama küsimusega “aga kust sa tead, et nii juhtub?” (Metslang 1994: 539). Eesti keeles on võimalik tulevat aega väljendada leksikaalselt (nt verbidega *hakkama*, *tuleb*, *läheb*) ning olevikuvormiga, seda täpsustades ajamäärusega (nt ajaadverbidega *pärast*, *hiljem*, *veel*).

Siinse artikli aluseks olevas uurimuses käsitletakse katsetesse valitud ajasõnu grammatilist aega täpsustavate leksikaalsete markeritena, mistõttu ei leia kirjeldamist niivõrd nende sõnade omandamise piiritlemine kui võrd verbi ajavormide ja ajasõnade vastavusse seadmine. Ajasõnade kaks kolmikkomplekti (esimeses ajasõnadega katses *juba-jälle-veel*; teises ajasõnadega katses *enne-praegu-pärast*) on valitud eesmärgiga uurida verbi ajavormide mõistmist ja produtseerimist nii kvantifitseerivate kui ka ajastavate ajamääruste abil. Valitud sõnad on mitmetähenduslikud, välja arvatud *praegu*, mida kasutatakse vaid ajatähenduses. Katsesse valitud sõnade ajatähenduses sisaldub järgnevuse komponent, mis tähendab, et nende tähendusest järeldeb sündmuse, situatsiooni ja/või ruumi järgnevus ja/või selle eelnemine ajateljel.

Lauses ajastava määрусena esinevad ja sündmuse toimumisaega täpsustavad selgepiirilisema tähendusega ajasõnad *enne*, *praegu* ja *pärast* ilmuvad lapse

kõnesse juba enne neljandat eluaastat¹, mis lubab oletada, et nende sõnade mõistmine ja sobiva ajavormiga vastavusse viimine 4–7-aastastele lastele erilisi raskusi ei valmista. Ajasõna *enne* esineb lausete mineviku tingimuses tähenduses ‘varem, varemalt’ ning viitab tegevuse ja ruumi eelnemisele. Ajasõna *praegu* esineb oleviku tingimuses, mille tähendusest järeldeb käimasoleva tegevuse hetke hõlmav ajavahemik. Ajasõna *pärast* esineb tuleviku tingimuses ning selle tähendus viitab järgneval ajalõigul toimuma hakkavale tegevusele.

Lauses kvantifitseeriva määrusena esinevad ja korduvust näitavad hägusama tähendusega ajasõnad on *juba*, *jälle* ja *veel*, neist vähemalt *juba* ja *jälle* esinevad varases lapsekeeles sagedasti (Vider 1995), *veel* ilmub lapse kõnesse esmalt küll hulgatähenduse väljendamiseks (nt *veel piima*), kuid enne neljandat eluaastat esineb see ka aja tähenduses (CHILDES-i andmetel). Ajasõnade katses kasutatud sõnade *juba-jälle-veel* tähendused kutsuvad esile ka eelduse-ootuse komponendi, mille järgi ajasõna tähendusest järeldeb kirjeldatava tegevuse eeldatav toimumis-aeg või -ruum.² Ajasõna *juba* esineb mineviku tingimuses ning selle tähendusest pääseb mõjule implikatuur tegevuse ootuspärasest varasema toimumise kohta ehk kõnealune tegevus on toimunud varem, kui eeldati. Veel implitseerib ajasõna *juba* tähendus ordinaalse korduvuse, mille järgi tegevuse toimumise koht muutub. Korduvust väljendava ajasõna *jälle* tähendusest järeldeb eeldus, et sündmus või situatsioon on varem toimunud ja implitseerib selle (eeldatava või ootuspärase) vahepealse muutuse, mis omakorda on seotud ruumi või tegevuse järgnevusega. Ajasõna *veel* esineb katses tuleviku tingimuses ning selle tähenduses pääseb mõjule implikatuur sündmuse eeldatava muutuse ja korduse kohta. Teatud juhtudel võib *veel*-sõna olla eituslembene (EKG II: 160) ning selle omaduse avaldumist tuleviku tingimuses saab samuti katse andmetel kontrollida.

3. Katse ja informandid

Uuringus kasutatud katse on koostatud rahvusvahelise projekti COSTA A33 “Cross-linguistically Robust Stages of Language Development” aja- ja aspektiuurimise töörühmas ning selle autorid on Bart Hollebrandse ja Margreet van Koert. Tegemist on ühe eelkatsega, mida täiustades on tulevikus kavas luua selline testimisvahend, millega määrata keelelise arengu mahajäämust ajakategooria valdkonnas.

Katse on ülesehituselt võimalikult lihtne: kasutatud ajatingimusi on kolm (minevik, olevik ja tulevik), iga tingimus esineb katses kolm korda ning kokku on katse mõistmis- ja loomisosas mõlemas 9 testüksust.³ Katse on üles ehitatud traditsiooniliselt kahepoolsena ning koosneb seega mõistmis- ja loomistestist. Tuleviku kohta käivates testüksustes on mõistmistestis kasutatud *hakkama*-verbiga konstruktsiooni. *Hakkama*-verb on Metslangi väitel üks tuleviku väljendamise põhivahendeid eesti keeles ning faasiverb *hakkama* on üks neid leksikaalseid vahendeid, millest annab järeldada algava sündmuse paigutamist tuleviku suunas. *Hakkama*-verbi leksikaalne põhitähendus ‘algus’ jääb Metslangi väitel järjest enam tulevikutähenduse varju. (Metslang 1994: 536) Katse andmed annavad võimaluse muu hulgas kontrollida, kuivõrd tajuvad lapsed *hakkama*-verbi kui tuleviku väljendamise vahendit.

¹ Andmed pärinevad CHILDES-i andmebaasist (Child Language Data Exchange System) <http://childes.psy.cmu.edu/> (5.03.2010), mis sisaldab 9 eesti lapse materjale.

² Lähemat käsitlust sõnade tähenduse kohta vt Parm 2007.

³ Tegemist on traditsioonilise 3 x 3 ülesehitusega, vt nt grammatilise aja omandamise eksperimenti van der Feest, van Hout 2002 või Valian 2006.

Katse viidi läbi iga lapsega eraldi. Mõistmis- ja loomistestist koosneva katse sooritasid lapsed samal päeval ja enamasti ilma puhkepausita. Testiüksused on paigutatud juhuslikku järjekorda, näiteks nii, et mõistmistestis ei tekiks kindlat vastuste mustrit. Mittekellelise stiimuli loomisel on katses kasutatud visuaalseid vahendeid ehk lühikesi videoklippe ning ajalise pidepunkti loomiseks ruumi.⁴ Videoklippides on keskel kohal tee, kus on kolm ruumilist pidepunkti ehk maamärki, mille juures üks ja sama tegelane üht ja sama tegevust sooritab.⁵ Grammatilise aja katse (edaspidi A-katse) mõistmistestis küsitakse lapse käest vastavate ajavormidega küsimusi (nt *Kus kuningas kõhis/kõhib/kõhima hakkab?*) ning laps peab ütleva, kus mingi tegevus *toimus/toimub/hakkab toimuma* või osutama ekraanil vastavale kohale ehk maamärgile. Loomistestis palutakse lapsel lause alguseid jätkata, nt *Laua juures kloun...; Taime juures mereröövel...* Leksikaalse aja ehk ajasõnade katsete (A1-katse ja A2-katse) mõistmistestides küsitakse lapse käest küsimusi, milles sisaldub nii verbi ajavorm kui ka sellega semantiliselt sobiv ajasõna ning vastaja ülesanne on sama, mis eespool kirjeldatud ajasõnadeta testis – vastata küsimusele ja näidata ekraanil vastavat kohta. Ajasõnadega loomistestis on lapse ülesandeks lõpetada testija öeldud lause ekraanil toimuva tegevuse kohta, kusjuures lapsele esitatav lause sisaldab ajasõna, nt mineviku tingimuses *Laua juures kloun juba/jälle/veel...; Enne/praegu/pärast laua juures kloun...* Tabel 1 annab ülevaate katse üldisest ülesehitusest kolme kasutatud katsevariandi kaupa.

Tabel 1. Kasutatud katse üldine ülesehitus

	A-katse (ajasõnadeta)	A1-katse (ajasõnad <i>juba, jälle, veel</i>)	A2-katse (ajasõnad <i>enne, praegu, pärast</i>)
Mõistmistesti 9 üksust	Kus kuningas kõhis/ kõhib/kõhima hakkab?	Kus kuningas juba/ jälle/veel kõhis/kõhib/ kõhima (hakkab)?	Kus kuningas enne/ praegu/pärast kõhis/ kõhib/kõhima (hakkab)?
Loomistesti 9 üksust	Laua juures kuningas...	Laua juures kuningas juba/jälle/veel...	Enne/praegu/pärast laua juures kuningas...

Eksperimentis kasutatud ajasõnad on testi lausetesse paigutatud alati kindlates tingimustes: mineviku tingimuses *enne* ja *juba*, oleviku tingimuses *praegu* ja *jälle*, tuleviku tingimuses *pärast* ja *veel*. Kasutatud videoklipid ja nende esitamise järjekord on kolmes katses sama.

Katses on kasutatud kuut agentiivset ja enamasti intransitiivset verbi (*kõhima, aevastama, tantsima, norskama, põlvitama*, (palli käes) *viskama*) ning nende verbidega väljendatud tegevusi sooritavad kolm tegelast: mereröövel, kuningas ja kloun. Et aja- ja aspektikategooria on tihedalt seotud, on keeleomandamiskatsetes nende kahe kategooria lahushoidmine üksjagu keerukas (vt nt Valian 2006: 255). Siinses katses on hoitud ajakategooria omandamise parema jälgimise huvides aspektilisus konstantsena nii, et testiüksustes varieerub küll aeg (ajavorm ja/või ajasõna), kuid aspektilisust silmas pidades on kõik situatsioonid imperfektiivsed.

Ajakategooria omandamise katse viidi läbi kokku 135 normaalse keelise arenguga 4–7-aastase eesti lapsega, iga katsevariandi tegi 45 last. Ülevaate katseisikutena kasutatud laste vanusest, soost ja üldisest hulgast annab tabel 2.

⁴ Aja- ja ruumikategooria seotusele on teiste hulgas tähelepanu juhtinud Helle Metslang, kes on soome *tulla*-verbi tulevikutähtsuse tekke selgitamise juures väitnud, et oluline on ruumitähenduse muutumine ajatähenduseks (Metslang 1994: 545).

⁵ Kaudse eeskujuna on kasutatud Laura Wagneri eksperimenti, kus käpiknukk sooritas ühte ja sama tegevust kolmes eri kohas (Wagner 2001).

Tabel 2. Ajakategooria omandamise katses osalenud lapsed

	4-aastased	5-aastased	6-aastased
Laste arv	15	15	15
A-katse			
Vanusevahemik	4;1–4;11	5;1–5;11	6;2–6;11
Sooline jaotus	N = 9; M = 6	N = 7; M = 8	N = 12; M = 3
A1-katse			
Vanusevahemik	4;1–4;11	5;0–5;11	6;0–6;11
Sooline jaotus	N = 11; M = 4	N = 7; M = 8	N = 6; M = 9
A2-katse			
Vanusevahemik	4;1–4;11	5;2–5;11	6;0–6;11
Sooline jaotus	N = 9; M = 6	N = 9; M = 6	N = 6; M = 9
Laste arv kokku	45	45	45

4. Ajakategooria omandamise katse tulemused

Uurimismaterjal on kogutud eksperimentide läbiviimisega sellisel meetodil, mis võimaldab kvantitatiivset, arvulist mõõtmist ja selle kaudu mõõtmistulemuste saamist. Järeldused on tehtud andmete statistiliste kokkuvõtete põhjal, milleks on kasutatud andmetöötlussüsteemi SPSS 15.0.

Omandatuks on loetud aeg siis, kui kuuest testüksusest on õigesti vastatud vähemalt neli ning kui neid lapsi, kel on rohkem kui neli õiget vastust, on vastava vanusegrupi laste seas enam kui 60%. Tuleviku testüksustes on vastuste hindamisel ajasõnade loomistestides A-1 ja A-2 alati õigeks arvatud need vastused, milles verb on esitatud olevikus (sest koos ajasõnadega *veel* ja *pärast* saab tuleviku aega väljendada) ja ka *hakkama*-konstruktsioon. Eraldi on tuleviku tingimuse vastustes kodeeritud eitav konstruktsioon (nt *ei ole veel tantsinud*), mis on samuti õigeks loetud. A-katses ehk ilma ajasõnadeta katses on aga ainult üksik olevikuvorm *on* loetud valeks vastuseks eelkõige seetõttu, et pelgalt olevikuvormi kasutus tuleviku-situatsioonis ei võimalda öelda, kas laps viitas olevikule või tulevikule. Kui laps on kasutanud testüksuses esitatust erinevat verbi, aga see verb on situatsioonikohases vormis, on vastus loetud õigeks (nt *norskab* asemel *magab*).

4.1. Ajakategooria omandamise üldised tunnused lastel vanuses 4–7 eluaastat

Saadud katsetulemuste põhjal võib väita, et neljandaks eluaastaks ei ole eesti lapsed veel ajakategooriat täielikult omandanud ning just 4.–7. eluaastani toimub ajakategooria aktiivne omandamine. Väite aluseks on ajakatses antud õigete vastuste osakaal, mis suureneb 4.–7. eluaasta vahel märgatavalt. Kasv ulatub mineviku ja oleviku testüksuste puhul viiest kuni kümne protsendini, eriti ilmne on aga õigete vastuste hulga suurenemine tuleviku testüksustes: näiteks on kuueaastaste laste A1-katse tulevikuüksustes umbes 1,5 korda rohkem õigeid vastuseid kui nelja-aastastel ning A2-katses kuueaastastel umbes 1,7 korda rohkem õigeid vastuseid kui nelja-aastastel. Tabelis 3 on antud ülevaade kolme ajakatse tulemustest eri aegade tingimustes ning esitatud vanusegruppide kaupa nende laste osakaal, kellel on aeg omandatud.

Tabel 3. Aegade omandamine 4–7-aastastel lastel ajakatsete põhjal: ühele ajatingimusele 4–6 õiget vastust andnud laste osakaal

Katse nimetus		Minevik			Olevik			Tulevik		
		A	A1	A2	A	A1	A2	A	A1	A2
Vanus	4;1–4;11	80%	87%	93%	80%	87%	80%	13%	47%	40%
	5;0–5;11	67%	87%	100%	75%	67%	80%	20%	67%	80%
	6;0–6;11	93%	87%	100%	80%	80%	93%	40%	67%	93%
Kokku	4;1–6;11	80%	87%	98%	80%	78%	85%	25%	60%	71%

Vaadates kolme ajakatse tulemusi tabelis 3, on näha, et **mineviku** aja tingimustes annavad lapsed kõige rohkem õigeid vastuseid A2-katses, milles oli esitatud ajasõna *enne*. Vanusegrupiti on tulemused selles katses peaaegu võrdsed, seega mõistavad lapsed juba 4-aastaselt ajasõna *enne* tähendust ja oskavad seda vastavusse viia verbi minevikuvormiga. A1-katses, milles oli esitatud ajasõna *juba*, on laste õigete vastuste osakaal pisut väiksem, kuid tulemused on samuti vanusegrupiti võrdsed. Üsna kõrge on selle katsevariandi puhul kõigile kuuele testüksusele õigesti vastanute osakaal, milleks on ligi 60% lastest. Minevikuvormi omandatus kajastub tulemustes, mis saadi A-katses, kus testüksustes ajasõnu üldse ei esinenud. Selle katse tulemused näitavad, et 4–6 õiget vastust on 80% lastest. Kõigile kuuele testüksusele on õigesti vastanud ligi pooled lastest.

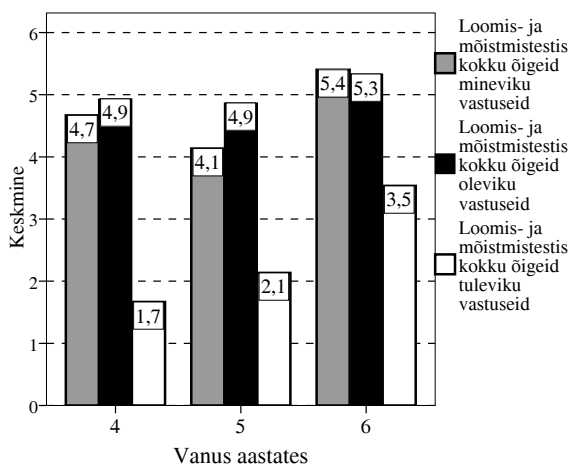
Oleviku aja tingimustes on kõige rohkem õigeid vastuseid samuti A2-katses, selles oli leksikaalse ajainfona esitatud ajasõna *praegu*. Katse tulemused näitavad, et vanusegrupiti on selle katsevariandi tulemustes märkimisväärseid erinevusi vähe. Seega saab väita, et hiljemalt 4-aastaselt suudavad lapsed mõista ajasõna *praegu* tähendust ja viia selle vastavusse verbi olevikuvormiga. Kõik vastused on õiged 73% lastest, kellest suurema osa moodustavad 6-aastased ning 4- ja 5-aastaste hulgas oli kõigi õigete vastuste andjaid võrdselt – mõlemas rühmas pooled lapsed. Üldiste tulemuste poolest sarnanevad A-katse ja A1-katse olevikuaja tulemused: 4–6 õiget vastust oli mõlema katsevariandi puhul umbes 80% lastest. Kuna A-katse ja A1-katse tulemused üldjoontes sarnanevad ega näita väga suurt muutust laste vanuse kasvades, siis võib oletada, et A1-katses esitatud ajasõna *jälle* ei ole oleviku aja tingimuses oluline ajainfo. Selle tõestuseks saab tuua ka kõigile testüksustele õigesti vastanud laste arvu, milleks A1-katse puhul on 17 last 45-st (38%) ja A-katse puhul 26 last 45-st (58%).

Tuleviku aja tingimustes on kõige rohkem õigeid vastuseid A2-katses, milles oli esitatud ajasõna *pärast*. 4–6 õiget vastust oli 71% lastest, kuid vanuserühmade võrdluses saab esile tuua selle, et 4-aastaste hulgas on vaid 6 last 15-st (40% lastest), kes on õigeid vastuseid andnud 4–6 testüksusele. Teistes vanuserühmades on nende laste osakaal tunduvalt suurem. Seega saab väita, et ajasõna *pärast* tähendus on omandatud mitte enne kui 5. eluaastaks ning et alles siis osatakse selle abil tulevikku väljendada. A2-katses olid kõik vastused õiged 60% lastest, lisaks leidis 4 last, kes kasutas loomistestis eitavat konstruktsiooni (nt *Pärast ratta juures kloun üldsegi ei viska / palli ei visanud / ei köhinud / ei köhi*), mis tuleb tuleviku tingimuses samuti põhimõtteliselt õigeks vastuseks lugeda.

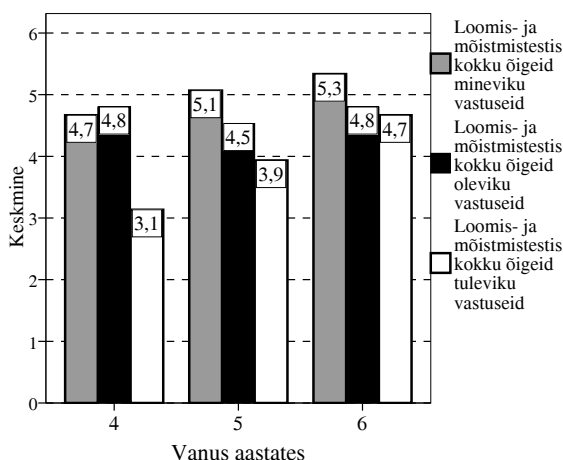
Teises ajasõnadega katses ehk A1-katses oli esitatud ajasõna *veel* ning tulemused näitavad, et 4–6 õiget vastust on 60% lastest ehk vähem kui A2-katses.

Vanuserühmade võrdluses saab välja tuua selle, et 4-aastaste hulgas on neid lapsi, kes vastasid õigesti vähemalt neljale testüksusele, vaid 47%, seevastu 5- ja 6-aastaste hulgas on mõlemas neid 60%. See näitab, et hiljemalt 5. eluaastaks on lapsed õppinud mõistma ajasõna *veel* tähendust ja oskavad selle abil ka tulevikku väljendada. Tuleviku hilisemat omandamist näitab ka asjaolu, et kõigile kuuele testüksusele õigesti vastanute osakaal on alla poole vastanutest (36%). Võrreldes tulemusi A2-katsega, leidis selles katsevariandis enam lapsi, kes olid tuleviku tingimuses vastanud eitava konstruktsiooniga (kokku 10 last).

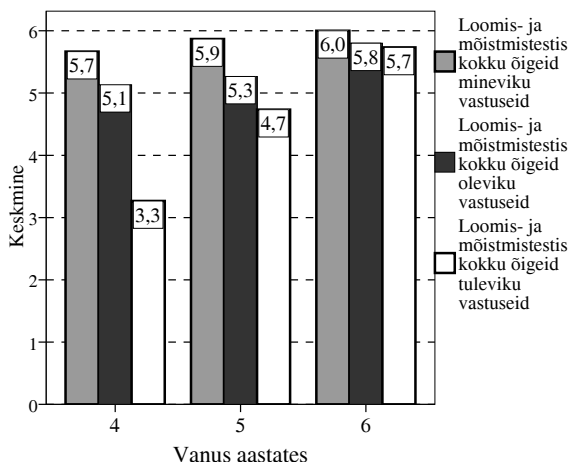
A-katse tulemused näitavad, et 4–6 õiget vastust on tuleviku tingimuses andnud 25% lastest, kellest on kõigile kuuele testüksusele õigesti vastanud vaid neli last, lisaks on seitse last vastanud eitava konstruktsiooniga. Laste vanuse kasvades tulevikutingimuste õigete vastuste hulk küll suureneb, kuid jääb 6-aastaste laste puhul ikkagi ainult 40% lastest. Joonised 1–3 annavad ülevaate õigete vastuste hulga muutumisest lastel vanuses 4–7 eluaastat.



Joonis 1. Keskmine õigete vastuste arv A-katses vanusegrupiti



Joonis 2. Keskmine õigete vastuste arv A1-katses vanusegrupiti



Joonis 3. Keskmine õigete vastuste arv A2-katses vanusegrupiti

Saadud katsetulemuste põhjal võib üldiselt väita, et grammatiline olevik ja minevik on eesti lastel 6. eluaastaks omandatud, tuleviku tähenduse omandatus ei ole aga kindel. Lapsi, kes A-katse tulevikusituatsioonis kasutasid kas *hakkama*-konstruktsiooni või olevikuvormi koos ajasõnaga, oli 40%. Kui neile lisada tinglikult need lapsed, kes kasutasid tulevikusituatsioonile viitamiseks ka lihtsalt olevikuvormi, mis võiks olla põhimõtteliselt aktsepteeritav, tõuseks õigesti vastanute hulk üle 60%. Samas ei anna pelgalt olevikuvormi kasutused siiski alust otsustada, kas laps ikkagi viitab tulevikule või hoopis olevikule.

Aja vormistamise leksikaalsed vahendid on eesti lastel 6. eluaastaks omandatud nii minevikus kui ka olevikus. Ajasõnade *enne* ja *juba* esinemine katsete mineviku tingimuse osas näitab, et 4-aastaste laste jaoks on selgepiirilisema tähendusega ajasõna *enne* paremini tajutatav kui ähmasema tähendusega ajasõna *juba*, kuid 5- ja 6-aastased on omandanud mõlema ajasõna tähenduse ja oskavad neid vastavusse viia verbi minevikuvormiga. Ajasõnade *jälle* ja *praegu* võrdlus oleviku tingimuses näitab, et *praegu* tähendus on lastel omandatud juba 4-aastaselt ja seda osatakse kasutada koos verbi olevikuvormiga. Kuna A-katse ja A1-katse tulemused oleviku tingimuses sarnanesid, siis ei saa midagi väita ajasõna *jälle* tähenduse mõistmise kohta, kuid võib oletada, et selle sõna esinemine ei aita lastel olulisel määral mõista ja kasutada verbi olevikuvormi. Tuleviku tingimuses kasutavad ja mõistavad lapsed ajasõnu *pärast* ja *veel* õigesti alates 5. eluaastast, olles omandanud ka tuleviku väljendamise nende sõnade abil. Kolme loomistesti võrdlus näitab, et lapsed kasutavad tuleviku väljendamisel eitavat konstruktsiooni enim just ajasõna *veel* sisaldavas A1-katses, mida võib selgitada sellega, et tegemist on eesti keeles eitustundliku sõnaga.⁶

Kõige enam oli kolmest katsevariandist õigeid vastuseid A2-katses, mis annab aluse arvata, et ajasõnad *enne*, *praegu* ja *pärast* on laste jaoks kõige selgemad ja nende vastavusse seadmine ajavormiga on omandatud juba 4-aastastel lastel. Kõige vähem oli õigeid vastuseid ilma ajasõnadeta katses, mille põhjal võib järeldada, et grammatiline ajavorm ilma leksikaalse ajainfota on laste jaoks kõige raskem.

⁶ Eitustundlike üksusi eesti keeles on uurinud Leho Paldre (1998), ta on analüüsinud ka sõna *veel* eituslembust.

4.2. Ajakategooria omandamise erisused laste soost lähtuvalt

Katsealuste laste hulk võimaldab teha mõningaid järeldusi ka selle kohta, kuivõrd ja kas üldse sõltub ajakategooria omandamine lapse soost.

Kõigi kolme katsevariandi puhul on ilmne, et poiste ja tüdrukute õigete vastuste hulk ei ole üheski ajatingimuses oluliselt erinev ning jääb üldiselt kuuendiku ehk ühe vastuse piiridesse (kokku oli ühes katses ühes ajatingimuses kuus testiüksust). Nelja-aastaste laste hulgas oli tüdrukutel kõigis ajatingimustes pisut enam õigeid vastuseid (kõige suurem erinevus oli ainult 0,8 vastust mineviku A2-katses).

Seejärel torkab silma, et viieaastaselt on nii mineviku kui ka oleviku tingimustes tüdrukute õigete vastuste arv poiste omast väiksem. Tuleviku tingimuses jääb tüdrukute vastuste hulk poiste omast küll A1-katses suuremaks, kuid A- ja A2-katses väiksemaks ning erinevused ei ületa samuti ühe vastuse piiri. Kuueaastaselt on A-katses kõige suurem vahe tuleviku tingimuses (tüdrukutel keskmiselt 3,8 õiget vastust ja poistel 2,7), minevikus on kuueaastaste poiste ja tüdrukute vahe väga väike (kõige suurem A-katses, keskmiselt 0,4 vastust) ning sama võib väita ka oleviku kohta (kõige suurem 0,8 vastust oleviku A1-katses).

Eelneva põhjal võibki väita, et katsetulemused muutuvad pigem laste vanusest, mitte niivõrd soost sõltuvalt, ning tulemused ei anna ühegi aja puhul alust oletada, et ajakategooria omandataks sooti eri vanuses või eri aegade lõikes erinevalt.

4.3. Ajavaliku vead ehk ajanihked

Laste tehtud vigu võib tinglikult nimetada ajaniheteks. Ajanihke all on siinses kirjutises mõistetud mingi ajavormi ekslikku kasutust, vastust, kus laps kasutas näiteks olevikuvormi minevikku nõudvas kontekstis või osutas mõistmistestis vale ajatingimuse kohta käivale maamärgile. Järgnevalt on ajanihkeid analüüsitud mõistmis- ja loomistestis eraldi, see annab võimaluse teha kindlaks, kas näiteks aspektikatses (vt Argus 2009) märgatud erinevus ühe ja sama struktuuri loomise ja mõistmise omandamises võib olla tunnuslik ka ajakategooria omandamise katsele. Seejärel on ajanihkeid analüüsitud aegade kaupa, et näha, kas aja omandamise puhul võib rääkida esialgsest nn tuumsest oleviku ajasüsteemist, mis tähendab, et lapsed paigutavad ka mineviku- ja tulevikusündmused esialgu olevikku, või pigem ühest, aga ebakindlast ehk liigendamata ajasüsteemist, kus laste valede vastuste hulgas on nihkumisi ka minevikku ja tulevikku.

Üldiselt näitasid katsete tulemused, et laste vanuse kasvades vale aja kasutamine väheneb, enim eksivad 4-aastased ja kõige vähem 6-aastased, 5-aastaste tulemustes on näha kõikuvusi nii sooti kui ka katsevarianditi – tulemused on mõnes katsevariandis isegi halvemad kui 4-aastastel, kuid üldjoontes siiski eksimuste arv väheneb. Tabel 4 näitab, kui palju kõikidest valedest vastustest (kõigi kolme katse loomis- ja mõistmistestid) kaldub minevikku, kui palju olevikku ja kui palju tulevikku. Võrreldes ajanihete hulka eri katsevariantides, ilmneb, et A-katse ehk ajasõnadeta katse on kõige suuremate nihkumiste, s.t eksimuste arvuga.

Tabel 4. Valede vastuste hulk eri aegadesse nihkumiste kaupa

	A1-katse	A2-katse	A-katse
Nihkeid minevikku	81	43	73
Nihkeid olevikku	57	28	139
Nihkeid tulevikku	2	0	1
Nihkeid kokku	140	71	213

Ajanihe **minevikku** tähendab siinses katses seda, et laps on esitanud oleviku või tuleviku kontekstis minevikuvormis vastuse. A-katse ja A1-katse mõistmistestides on kolmandik lastest valinud tuleviku või oleviku asemel mineviku, kusjuures minevikku nihkutakse rohkem oleviku tingimuses ning vanuserühmade võrdluse põhjal saab väita, et eriti just 4-aastased lapsed valivad mõistmistestis olevikus toimuvale tegevusele osutamise asemel hoopis minevikus toimunud tegevuse. Tulenevalt sellest, et A2-katse mõistmistestis oli õigete vastuste hulk võrreldes teiste katsetega märgatavalt suurem, on selles ka ajanihete arv väiksem. Võib öelda, et A2-katses minevikku nihkumist peaaegu ei esine. Loomistestide tulemuste põhjal saab täheldada, et mineviku ülekasutust on A-katses ja A1-katses esinenud pooltel lastest ja mõnevõrra vähem ehk kolmandikul lastest A2-katses. Tavaliselt on lapsed minevikuvormi kasutanud oleviku tingimuses (nt katses eeldatud vastuse *kloun põlvitab* asemel vastas laps hoopiski: *kloun põlvitas*). A1-katse vastustes esineb minevikku kaldumist kõige enam ja selle poolest erinevad need tulemused mõistmistesti omadest. Selline mineviku ülekasutus ja ajasõna *jälle* vastavusse seadmine oleviku ajaga võib tuleneda ajasõna *jälle* tähendusest ja selle sobimisest mitme aja tingimusse (nt *jälle põlvitas*, *jälle põlvitab*, *jälle hakkab põlvitama*). Lisaks võib minevikku kaldumist seostada katse üldise olemusega: tegelik aeg ja katse aeg on omavahel paratamatult pisut nihkes ning arvutiekraanil toimuv tegevus on kõnehetkeks muutunud minevikuliseks.

Ajanihe **olevikku** tähendab, et laps on vastanud mineviku või tuleviku kontekstis olevikuvormiga. A-katse mõistmistestis on mineviku ja tuleviku aja kontekstis olevikuvormi eelistanud pooled lapsed, A1-katses kolmandik lastest ja A2-katses veel vähem ehk viiendik lastest. Olevikuvormi valiku põhjuseks võib olla ka see, et oleviku aega peegeldav maamärk paigutus katse videoklipis ekraani keskosas ja kahtluse korral võis laps just keskse paigutuse pärast seda eelistada. Loomistestis esineb oleviku ülekasutust ehk olevikuvormis verbiga vastamist A-katses pooltel lastest. Selline ajanihe on suures osas ootuspärane, mitteootuspärased olevikuvormis vastused esinevad enamasti tuleviku tingimuses ja võib oletada, et lapsed ei ole selles vanuses leksikaalseid tuleviku väljendamise vahendeid veel päriselt omandanud. Kõige rohkem eksimusi esineb 4-aastaste hulgas, mis annab alust oletada, et leksikaalsed tuleviku väljendamise vahendid omandatakse hiljem, s.t viienda eluaasta paiku. A1-katse ja A2-katse loomistestis oleviku ülekasutust peaaegu ei esine, seega on olevikukesksus omane pigem ajasõnadeta katsele ning katseis kasutatud ajasõnad konkretiseerivad aja, sunnivad valima deiktilisest nullpunktist erineva (markeeritud) aja.

Ajanihe **tulevikku** tähendab, et laps on vastanud oleviku või mineviku kontekstis tulevikuvormiga, kuid sinise kirjutise aluseks olevad katsetulemused ei anna alust väita, et eesti lapsed tulevikku kalduksid. Et tulevik ei ole eesti lastel selles vanuses veel omandatud, on ka ootuspärane, et lapsed ei kipu oleviku ja mineviku kontekstis tulevikuvorme kasutama.

Kokkuvõtvalt võib öelda, et mõistmistestis kaldutakse kõige sagedamini olevikku ja et olevikukeskus on iseloomulik nooremale vanusegrupile ning hakkab 5–6-aastaselt taanduma.

Loomistesti ajanihete tulemuste põhjal saab väita, et erinevalt mõistmistestist kaldutakse siin pigem minevikku. Teisisõnu: sündmuste kohta käivate ajavorimide ja ajasõnade mõistmisel paigutatakse sündmus sagedamini olevikku, kuid ise mingit sündmust kirjeldades kasutatakse sagedamini minevikku. Mineviku eelistamine võib osalt olla tingitud ka katsesituatsioonist – võimalik, et laps tajub, et selleks ajaks, kui ta vastuse valmis saab, ongi ekraanil toimunud tegevus jäänud minevikku. Eri ajasõnade katsete tulemused näitavad, et ajasõnade semantika ise ei tingi nihkumist mingi kindla aja suunas.

Ootuspäraseks võib lugeda tulemuse, et tulevikuvormide ülekasutust nii loomiskui ka mõistmistestis peaaegu ei esine. Võib arvata, et lapsed eelistavad kahtluse korral siiski tuttavamaid aja väljendamise vahendeid, milleks on mineviku- ja olevikuvormid. Juba õigete vastuste analüüsist ilmnes aegade omandamise võrdlusel, et tulevikku nõudvates üksustes on lapsed ebakindlamad ja teevad rohkem vigu. Vaadates lähemalt A-katse loomistestis tulevikuuksuste valesid vastuseid, võib väita, et peale olevikuvormis vastuste, mille puhul pole võimalik kindlaks teha, kas laps pidas silmas tulevikku või mitte, on sageduselt järgmiseks vastusevariandiks eitav konstruktsioon, enamasti täisminevikuvorm, nt *laua juures kuningas ei ole veel kõhinud* või *ei ole veel tantsinud*, kusjuures nelja-aastaste laste hulgas on neid, kes eitavaid konstruktsioone kasutavad, vähem kui kuueaastaste hulgas. See annab alust oletada, et vanuse kasvades mitte ainult ei arene välja opositsioon oleviku ja mineviku vahel, vaid et hakatakse ka mitteminevikku liigendama täpsemalt. Näiteks täisminevikuvormi, nagu *ei ole veel tantsinud*, tulevikuline kasutus on ka üsna loomulik, sest täismineviku puhul saabki vaatepunkt olla ka tulevikus. Eitavate konstruktsioonide suur hulk (kolme katse tulemustes kokku 54 korral 405 võimalikust tulevikuuksuse vastusest) näitab, et kui sündmus on juba kuskil aset leidnud, siis mõtestatakse ka tulevikku pigem mineviku kaudu: *kuskil tegi midagi, kuskil ei ole veel teinud*.

Hakkama-tulevikku leidub laste vastuste hulgas poole vähem kui eitavaid konstruktsioone. *Hakkama*-tuleviku vähene kasutus võib mingil määral olla tingitud ka katsesituatsioonist: videoklipis toimuv tegevus ju tegelikult ei alga järgmise maamärgi juures, sest tegelane on seda tegevust teiste maamärkide juures juba varem teinud. Seega ilmneb, et *hakkama*-verbi tegevuse alguse tähendus on laste jaoks tulevikutähenduse kõrval täiesti aktuaalne.

Ajanihet mineviku suunas mõistmistestides küll leidub, kuid üldjuhul eksivad lapsed selles suunas väga vähe. Rõhutada tuleb aga loomis- ja mõistmistesti erinevust: kõigi katsete loomistestides esineb peaaegu pooltel lastel minevikuvormide ülekasutust, kusjuures enamasti kasutatakse minevikku oleviku asemel. Esmapilgul võiks oletada, et iteratiivsete verbidega (nt *kõhima*, *palli viskama*) testüksustes leidub olevikusituatsioonis minevikku kaldumist rohkem, sest käimasolevat tegevust sooritatakse korduvalt samas kohas ja kuna tegemist ei ole pidevalt ühtlaselt kulgeva tegevusega, vaid üksikute tegevusaktide kordumisega, ongi loomulik, et kord näiteks kõhatanud tegelase tegevuse iseloomustamiseks sobibki kasutada pigem minevikuvormi. Samas ei näita katsete iteratiivsete verbidega testüksuste vastused võrrelduna teiste verbidega (nt *põlvitama*) oleviku situatsiooni testüksuste vastustega sugugi seda, et iteratiivsus tingiks kaldumise minevikku.

Olevikuvormide suunas eksimist esineb mõistmistestides kõigis katsetes peaaegu pooltel lastel. Oleviku kasuks eksimist esineb nii tuleviku- kui ka minevikuaega nõudvates üksustes. Loomistestides leidub aga eri katsevariantide vahel erisusi – A1- ja A2-katses ehk ajasõnadega katsevariantides esineb oleviku ülekasutust vaid mõnel üksikul korral, kuid A-katses on olevikku üle kasutanud umbes pooled vastanutest. Põhjuseks on ilmselgelt eesti keele tuleviku väljendamise vahendite vähesus, olevikuvorm on ka tuleviku ajasõnata üks võimalik tuleviku väljendusvahend.

Seega võib ajanihete analüüsi põhjal väita, et ajasüsteemi omandamisel ei saa alati rääkida esialgselt olevikukesksest tuumikajasüsteemist, vaid pigem eelistatakse eri keelekasutusituatsioonis eri vorme: mõistmisel olevikku ja loomisel minevikku. Selline suundumus võib osutada ühelt poolt sellele, et minevik on sündmustest jutustamisel ehk narratiivses situatsioonis tüüpilisem ning teisalt ka sellele, et lause moodustamine võtab rohkem aega kui mõistmine ning võib oletada, et selleks ajaks, kui laps vastuslause konstrueeritud saab, ongi küsimuse hetkel toimuv tegevus muutunud minevikuliseks ehk juba toimunud tegevuseks.

Samas on nn ühe aja lapsi ehk neid, kes nii mõistmis- kui ka loomistestis vastasid valdavalt ühele kindlale ajale viidates, 135 lapsest 11 ehk 8% – see annab alust oletada, et kuigi üldist nn tuumikajasüsteemi eesti lastel vaadeldud vanuses valdavalt ei esine, on mõnel lapsel siiski just selline tuumikajasüsteem, olgu see siis oleviku- või minevikukeskne, olemas. Põhjalikumad uurimisandmed kolme- ja nelja-aastaste laste ajakategooria omandamise kohta annaksid sellele uurimisküsimusele kindlama vastuse.

4.4. Aja väljendamise vahendite mõistmisest ja loomisest võrdlevalt

Eksperimentaalse meetodika kasutamisel on täheldatud, et ühe ja sama keelestruktuuri mõistmise ja loomise omandamisel võib olla erinevusi ning mitte alati ei saa väita, et mingi keelestruktuur on mõistetav varem, kui seda luua ehk kasutada osatakse. Näiteks võib mõne keelestruktuuri üks osa olla mõistmistesti tulemustel varem ning loomistesti puhul hiljem omandatud (vt nt aspektilisuse omandamise kohta Argus 2009: 39).

Ajakategooria omandamise siinse katse tulemused on mõistmistesti puhul enamasti paremad kui loomistesti omad. Täpsemalt on kõikide aegade nii grammatilised kui ka leksikaalsed vahendid mõistetavad varem või samas vanuses, kui neid ise luua osatakse.

Kõige suuremad erinevused loomis- ja mõistmistesti vahel on märgatavad ilma ajasõnadeta katses. Eriti suur on erinevus mõistmis- ja loomistesti vahel oleviku testüksustes: mõistmistesti tulemuste järgi on olevik omandatud juba 4-aastastel, kuid loomistestis ajasõnadeta katses alles 6-aastastel, ajasõnadega testis (katsevariandis, milles kasutati oleviku ajasõna *praegu*) 4-aastastel. Kolme katsevariandi võrdlus kinnitab, et ajasõna *jälle* on mõistmistestis selge isegi 4-aastaste jaoks, aga loomistestis ei vii ka veel 6-aastased ajasõna *jälle* tähendust olevikuvormiga vastavusse. Aja väljendamise grammatilised vahendid ehk ajavormid on olevikuaja puhul mõistmistesti tulemuste põhjal omandatud 4-aastaselt, loomistesti tulemuste põhjal alles 6-aastaselt. Seega on lastel ilma leksikaalse ajainfota raske olevikuvormi väljendada, kuigi selle mõistmine probleeme ei valmista. Teisalt võib olevikuvormi

vähem kasutamist loomistestis seletada ka loomistesti enda eripäraga. Nimelt võib olla, et sündmustest jutustades tundub lapsele minevikuvorm kui tüüpiline narratiivis kasutatav vorm loomulikum kui olevik.

Kuigi üldiselt võib siinse katse tulemuste põhjal väita, et erinevalt Krantzi ja Leonardi (2007: 142) tulemustest ei jäta eesti lapsed ajasõna esinemisel lausest verbivormi välja, leidus katsealuste hulgas siiski kolm last, kes käitusid loomistestis samuti nii, et kui olid kuulnud ajasõna, siis ei pidanud nad vajalikuks verbivormi öelda (nt *Kloun taime juures praegu...*, milles olevikuvormis verbi asemel vastati nõustumist väljendav *jah*-iga). Tuleb aga meeles pidada, et Krantz ja Leonard uurisid vaid minevikutingimust koos ajasõnaga ja ilma selleta, siinses katses on aga kasutatud kolme aega ning eri ajasõnad tekitasid selgema ajakontrasti. Võib oletada, et just seetõttu ei jätnud lapsed verbivormi vastusest nii sageli välja.

5. Kokkuvõtte

Ajakategooria omandamise eksperimenti võib ülesehituselt lugeda sihipäraseks: vastuseid, kus laps vastas “ei tea” või jättis üldse vastamata, leidus ainult mõni üksik. Seega oli katseülesanne laste jaoks arusaadav ning laste tehtud vigade puhul ei ole tegemist katsestuaatsioonist tulenevate eksimustega.

Katse tulemustest nähtub, et aja väljendamise grammatilistest vahenditest on verbi olevikuvormid mõistetavad juba nelja-aastaselt, kuid mineviku ja tuleviku ajavormid alles kuueaastaselt. Oleviku- ja minevikuvormide kasutamine õiges ajalises stuaatsioonis omandatakse hiljem, kuueaastaselt. Tulevikuvormide kasutamine ei ole aga ka kuueaastaselt veel päris omandatud.

Ka selge opositsioon mineviku ja mittemineviku vahel areneb eesti lastel siinse katse tulemuste põhjal välja alles kuueaastaselt, siis kui lapsed suudavad nii mõistmis- kui ka loomistestis olevikust minevikku eristada. Laste soost lähtuvad erisused ajakategooria omandamises märgatavad ei ole: mõnes katsevariantis on tüdrukutel paremad tulemused kui poistel ja mõnes teises variantis on see vastupidi, kusjuures õigete vastuste hulga erinevus jääb kõikides ajatingimustes ja katsevariantides ebaoluliseks.

Mittemineviku liigendamist olevikuks ja tulevikuks hakkavad lapsed omandama mitte varem kui kuueaastaselt ning siis oskavad nad vastavaid ajavorme küll eristada, kuid nende vormide kasutamine ehk loomine ei ole ka kuueaastaks veel omandatud. Tuleviku omandamiskäik ilmneb kasutatud katse tulemuste põhjal kõige selgemalt. Arenguhüpe on kõikides katsevariantides märgatav ja õigete vastuste hulk suureneb stabiilselt. Tuleviku väljendamiseks kasutavad lapsed enamasti olevikuvormi koos ajasõnaga, seejärel eitavaid konstruktsioone, mis näitab veenvalt, et tegelikult alustavad lapsed tuleviku väljendamist pigem mineviku kaudu, viidates tulevikusündmusele kui millelegi, mis ei ole veel aset leidnud. *Hakkama*-konstruktsiooni tunnevad lapsed küll tuleviku väljendamise vahendina ära, kuid ise kuigi sageli ei kasuta, selle sagedus jääb tuntuvalt maha eitavate konstruktsioonide sagedusest.

Üldiselt võib katsetulemuste põhjal väita, et ajasõnad aitavad 4–7-aastastel lastel grammatilist aega tajuda ja kasutada. Ajasõnad *enne*, *praegu* ja *pärast* on laste jaoks küll selgemad kui ajasõnad *juba*, *jälle* ja *veel*, kuid mõlemad sõnakomplektid

on oluline ajainfo grammatiliste ajavormide mõistmisel ja loomisel. Kõige suurem ajasõnadevaheline erinevus seisneb oleviku tingimuses kasutatud komplekssema tähendusega sõna *jälle* ja selgepiirilise tähendusega ajasõna *praegu* mõistmisel ning selle vastavusse seadmisel katses esitatud olevikukeskse situatsiooniga: ajasõna *praegu* on omandatud juba nelja-aastastel lastel, kuid *jälle* kasutamine oleviku aja tingimuses tekitab raskusi isegi veel kuueaastastele. See tulemus kinnitab Slobini (1985) vastenduse printsiipi, mille kohaselt lapsed omandavad ja viivad verbimorfoloogiaga vastavusse kõigepealt need sõnad, millel on hõlpsasti määratav tähendus. Kuna mineviku ja oleviku aeg omandatakse tulevikust varem, siis kehtib see omandamiskäik mingil määral ka ajasõnade puhul: minevikule viitav ajasõna *enne* ja olevikule viitav *praegu* ning nende vastavusse seadmine õige ajavormiga on omandatud juba nelja-aastastel lastel, kuid tulevikule viitav ajasõna *pärast* ning selle vastavusse viimine õige ajavormiga omandatakse alles viieaastaselt. Samas vanuses on lapsed omandanud ajasõna *veel* tähenduse ja oskavad selle abil tulevikku väljendada ning loomistestis antud eitava konstruktsiooniga vastuste põhjal võib väita, et lapsed tajuvad ajasõna *veel* ka eitustundlikuna ning tulevikku väljendatakse implitsiitselt.

Viidatud kirjandus

- Argus, Reili 2007. Eesti keele verbi ajamorfoloogia ja aspektilisuse omandamisest. – Emakeele Seltsi aastaraamat, 52 (2006), 7–32.
- Argus, Reili 2008. Eesti keele muutemorfoloogia omandamine. Tallinna Ülikooli humanitaarteaduste dissertatsioonid, 19. Tallinn: TLÜ Kirjastus.
- Argus, Reili 2009. Psühholingvistiline katse eesti keele objekti käändevahelduse omandamise uurimise meetodina. – Emakeele Seltsi aastaraamat, 54 (2008), 22–43.
- Dietrich, Reiner; Klein, Wolfgang; Noyau, Collette 1995. The Acquisition of Temporality in a Second Language. *Studies in Bilingualism*, 7. Amsterdam: John Benjamins.
- EKG II = Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1993. Eesti keele grammatika. II. Süntaks. Lisa: Kiri. Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut. Tallinn.
- Krantz, Laurie R.; Leonard, Laurence B. 2007. The effect of temporal adverbials on past tense production by children with specific language impairment. – *Journal of Speech, Language, and Hearing Research*, 50 (1), 137–148. doi:10.1044/1092-4388(2007)012
- Metslang, Helle 1994. Eesti ja soome – futuurumita keeled? – *Keel ja Kirjandus*, 9, 534–547.
- Paldre, Leho 1998. Eitustundlikud üksused eesti keeles. Magistritöö. Käsikiri Tartu Ülikooli eesti keele õppetoolis.
- Parm, Sirli 2007. Partiklite *veel*, *juba*, *alles*, *jälle* tähendused eesti kirjakeeles. Magistritöö. Käsikiri Tartu Ülikooli üldkeeleteaduse õppetoolis.
- Slobin, Dan Isaac 1985. Crosslinguistic evidence for language making capacity. – D. I. Slobin (Ed.). *The Crosslinguistic Study of Language Acquisition*, Vol. 2. Hillsdale, NJ: Lawrence Erlbaum, 1157–1249.
- Valian, Virginia 2006. Young children's understanding of present and past tense. – *Language Learning and Development*, 2 (4), 251–276. doi:10.1207/s15473341l1d0204_2
- van der Feest, Suzanne; van Hout, Angeliek 2002. Tense comprehension in child Dutch. – *Proceedings of the 26th Boston University Conference on Language Development*. Massachusetts: Cascadilla Press, 734–745.
- Wagner, Laura 2001. Aspectual influences on early tense comprehension. – *Journal of Child Language*, 28, 661–681. doi:10.1017/S0305000901004792

- Weist, Richard 2002. Temporal and spatial concepts in child language: Conventional and configurational. – *Journal of Psycholinguistic Research*, 31 (3), 195–210. doi:10.1017/S0305000901004792
- Weist, Richard; Wysocka, Hanna; Lyytinen, Paula 1991. A cross-linguistic perspective on the development of temporal systems. – *Journal of Child Language*, 18, 67–92. doi:10.1017/S0305000900013301
- Vider, Kadri 1995. 2–3-aastaste laste sõnavara. Diplomitöö. Käsikiri Tartu Ülikooli eesti keele õppetoolis.
- Vihman, Marilyn May; Vija, Maigi 2006. The acquisition of verbal inflection in Estonian. – N. Gagarina, I. Gülzow (Eds.). *The Acquisition of Verbs and their Grammar: The Effect of Particular Languages*. *Studies in Theoretical Psycholinguistics*, 33. Dordrecht: Springer, 269–295.

Reili Arguse (Tallinna Ülikool) uurimisvaldkonnaks on esimese keele omandamine, huvikeskmes on siiani olnud muutemorfoloogia omandamisega seonduv.
Reili.Argus@tlu.ee

Sirli Parm (Tartu Ülikool) uurimisvaldkonnad on lapse keeleline areng, leksikaalne semantika.
sirli.parm@ut.ee

THE ACQUISITION OF TENSE AND TIME ADVERBS IN ESTONIAN

Reili Argus, Sirli Parm

Tallinn University, University of Tartu

The article discusses the acquisition of the category of time by 135 normally developed Estonian children between 4 and 7 years of age. The discussion is based on a psycholinguistic experiment that focused on tense forms and time adverbs. The experiment consisted of three tests designed to show at what age Estonian children develop a clear distinction between the past and non-past, whether the use of time adverbs helps the children to understand the grammatical tense used, and which time adverbs the children find easier to understand.

The results of the experiment show that of the means of grammatical tense marking, the present tense forms are already understood by the four-year-olds while the past and future forms only by the six-year-olds. The situationally correct use of the present and the past tense forms is only acquired at the age of six and that of the future tense forms even later.

The results of the experiment suggest that the Estonian children develop a clear distinction between past and non-past and are only at the age of six able to distinguish between them in comprehension and production tests. There are no significant gender differences in the acquisition of time: in some tests the girls did better than the boys and in the other tests the boys performed better, with insignificant differences in the number of correct answers in all the tenses and test variants.

Estonian children only appear to acquire the division of non-past into present and future at the age of six, when the number of instances of the situationally correct use starts steadily growing. The initial leap and the ensuing increase in correct use were most clearly shown with respect to the acquisition of the future tense. To express the future, the children most frequently used the present tense form with time adverbs. The next most frequent were the negative constructions, which seems to suggest that the children begin to express the future by way of the past—by referring to future events as another form of non-present time. Children also recognise the *hakkama* future but seldom use it themselves.

The results of the experiment suggest that time adverbs help 4–7 year old children to perceive and use tense. Children find the time adverbs *enne* ‘before’, *praegu* ‘now’ and *pärast* ‘after’ easier to understand than those of *juba* ‘already’, *jälle* ‘again’ and *veel* ‘still, yet’.

Keywords: language acquisition, tense, temporal words, Estonian

THE COMPREHENSION OF DERIVATIONAL MORPHEMES IN EARLY CHILDHOOD: AN EXPERIMENTAL STUDY FOR LITHUANIAN

Ineta Dabašinskiene

Abstract. This paper presents the results of an experiment that explores children's tolerance for word form variation, specifically, it compares preschool-age children's tolerance of word form variants, such as suffixes and prefixes in Lithuanian. The stimuli used in this study were pictures of familiar and unfamiliar animals. It appears that Lithuanian children show great tolerance of word form variation, in particular, children are more tolerant of variation in word endings than onsets. This fact suggests that exposure to a richly inflected language, such as Lithuanian, impacts on children's word-learning strategies.

Keywords: language acquisition, morphology, word formation, Lithuanian

1. Introduction

Most children, by the time they have reached primary school age, have become highly efficient word learners, and have acquired a number of word learning strategies that help them in finding the possible referents or meanings of new words (Golinkoff et al. 1992). It was observed that children associate novel word forms with unfamiliar categories of objects. This tendency has been called the principle of contrast (Clark 1993, 1995) or the mutual exclusivity (Markman et al. 2003) and states that a child relates the difference in meaning to the difference in form or the unfamiliar object to the new word.

Our study tries to explain to which extent the use of this word learning strategy is influenced by the typology of the language the child is acquiring.

In Baltic and Slavic languages, children are exposed to word form modifications since early childhood. For example, speakers of Lithuanian, a morphologically rich Baltic language, encounter noun morphology varying along the dimensions of two genders (masculine and feminine), two numbers (singular and plural), and seven

cases (nominative, accusative, genitive, dative, instrumental, locative and vocative). Lithuanian gender categories have quasi-regular correlations with respect to semantic and phonological features of nouns. Thus, most masculine nouns end in *-(i)as*, *-is* or *-ys*, and take the so-called first declension. Most feminine nouns end in *-(i)a* or *-ė*, and take the so-called second declension. The first and the second declensions are the most productive declension types. In addition, some feminine nouns end in *-is* and take the third declension, some masculine nouns end in *-us* and take the fourth declension, and some masculine and feminine nouns end in *-uo* and *-ė* and take the fifth declension. There are masculine nouns referring to males, e.g., *tėtė* ‘father’, which require masculine adjective-noun and pronominal agreement but in form they resemble feminine nouns and take the second declension.

Lithuanian diminutive morphology is also fairly complex. One very prominent feature of the Lithuanian language is the highly productive formation of diminutives from any noun via one or more suffixes. The most frequent and productive suffixes are the masculine *-elis/-ėlis*, *-(i)ukas*, *-utis*, *-ytis*, *-aitis* and their feminine counterparts *-elė/-ėlė*, *-(i)ukė*, *-utė*, *-ytė*, *-aitė*. A variety of suffixes can be found in the formation of diminutives for different lexical semantic groups, and for the majority of nouns at least three or four different suffixes can be attached to the same lemma, e.g., *kepur-ė* ‘cap-FEM’ has diminutive forms *kepur-yt-ė*, *kepur-ait-ė*, *kepur-ėl-ė*, *kepur-iuk-ė*. The diminutive suffixes are interchangeable with no differences in meaning associated with different suffixes. Note that while the word endings of Lithuanian diminutives are fairly homogeneous (*-is* and *-as* for masculine and *-ė* for feminine), there is considerable variety in the phonetic structure of the penultimate syllable and, consequently, more overall variety in the structure of diminutives. Moreover, double suffixation, which reinforces the pragmatic effectiveness or the meaning of smallness associated with the diminutive, is common as well, as evidenced by examples such as *dal-el-yt-ė* ‘particle-DIM-DIM-FEM’, *žmog-el-iuk-as* ‘man-DIM-DIM-MASC’, or *saul-ut-ėl-ė* ‘sun-DIM-DIM-FEM’.

Results from the corpora of spontaneous adult conversations (described in Savickienė 2005, Dabašinskienė 2009) give the following results of the use of diminutives: 3.5%–5% (adult-directed speech or ADS), 53% (child-directed speech or CDS), 48% (child speech or CS). It was observed in longitudinal data that young children as well as adults start to use two or three different diminutive suffixes for the same lemma very early and without any difference in meaning, e.g., *mešk-iukas*, *mešk-utė*, *mešk-ytė* from *meška* ‘teddy-bear’; *med-ukas*, *med-utis* from *medus* ‘honey’ (1;8)¹; *kač-iukas*, *katin-ėlis*, *kat-ytė* from *katė* ‘cat’; *mašin-ėlė*, *mašin-ytė* from *mašina* ‘car’ (1;9); *arbat-ėlė*, *arba-tytė* from *arbata* ‘tea’ (1;10); *kišk-utis*, *kišk-ytis* from *kiškis* ‘rabbit’; *rank-utė*, *rank-ytė* from *ranka* ‘hand’ (2;0). These examples clearly contradict Clark’s principle of contrast (1995). It follows from the Lithuanian data that a child may perceive two different forms of the same word—its basic meaning and the derived diminutive –, which have two different meanings. According to the principle of contrast, it may be assumed that the speaker will have to relate a difference in form with a difference in meaning (Clark 1995: 394). This tendency, however, is not usual in other languages where children learn diminutive suffixes one after another (Gillis 1997, De Marco 1998).

There is virtually no prefixation in Lithuanian noun derivational morphology. However, Lithuanian verb morphology contains a considerable amount of prefixa-

¹ 1;8 means one year and eight months (age of a child).

tion, with 12 verbal prefixes that alter the meanings of basic verbs. For example, the verb *eiti* 'to go' can appear in the prefixed forms *apeiti*, *ateiti*, *įeiti*, *išeiti*, *nueiti*, *paieiti*, *pareiti*, *pereiti*, *praeiti*, *prieiti*, *sueiti*, and *užeiti*, all of which express different modifications of the meaning of the basic verb. To the extent that these differences in meaning are rather subtle and involve dimensions such as aspect, time and direction of movement, one can argue that greater amount of prefixation in the language as a whole leads to greater tolerance for word onset modifications in Lithuanian children. This speculation is supported by the data from Wójcik (2000) who reported a case study of a girl's acquisition of Lithuanian (the same corpus is used in Savickienė's (2003) study of the acquisition of noun morphology). In the input directed to the child, 37% of verbs had prefixed forms (63% were simplex). The child's speech showed similar rates, with 30% of verbs containing prefixed forms. Prefixed verbs are documented in the corpus from the first recording at age 1;7 and have been increasing in frequency, reaching the 50% rate of prefixation of verb forms, which is comparable to estimates for Lithuanian adult-directed speech (57% simplex and 53% prefixed verb forms).

Thus, children are familiar with the prefixed and suffixed word formation strategies since the early period of language acquisition, with only one difference, that is, prefixes are used mostly with verbs and suffixes are used with nouns.

As it was stated above, some first language acquisition studies have shown that children show a strong preference to associate novel word forms with unfamiliar categories of objects. When presented with novel word forms in a forced-choice comprehension test, children readily choose novel objects as opposed to objects with familiar names often described as a "taxonomic response" (Bruening, Brooks 2006). This explanation, though rather robust under controlled conditions of novelty, fails to explain or justify the pervasive use of diminutives in CDS, especially in Baltic, Slavic and Romance languages (see Savickienė 2003, Savickienė, Dressler 2007, Savickienė et al. 2007, Wójcik 1994, Kempe et al. 2007, Olmsted 1994) the speakers of which repeatedly use different morphological derivations to refer to the same objects.

2. Method

Thirty-nine monolingual children (mean age 4;5, range 2;2–6;8, 19 boys and 20 girls) were tested at several day-care centers, schools and at home in Kaunas, Lithuania.

2.1. Material and design

The present research uses a word-learning protocol in which a word-form modification is introduced as a second label of an object. The experiment introduces two contrasting changes in word form involving the same word stem. In one condition the suffix introduces a change to the end of the word and in the other condition the prefix changes the onset of the word. The visual stimuli were 13 x 18 mm pictures of animals, which were collected from nature albums, websites, and other sources. The pictures were arranged into two folders and the material was used to carry out 18 trials, 2 training and 16 test trials. Half of the trials contained pictures of familiar

animals while the other half contained pictures of novel animals. Images of familiar animals were carefully selected to represent objects and their names already known by pre-school children. The novel animals looked unfamiliar to children because of their non-European appearance. Half of the trials involved bi-syllabic animal names and half–tri-syllabic names. Each trial had an introductory page with a single picture of an animal while a second page contained the first image along with the image of a different animal (i.e., two novel animals or two familiar animals). The introductory page served the purpose of naming the animal by its simplex form, by saying, for example, ‘I call this animal a goat’. The test page was used for the forced-choice probe described above. The 18 trials comprised two initial trials to ensure comprehension of the instructions and 16 test trials. The test trials involved the manipulation of two factors: noun type (novel versus familiar) and position of word form modification (beginning versus end of a word).

Half of the nouns were feminine and half were masculine, half of the nouns of each gender had two syllables and the other half had three syllables. The words selected for each stimulus pair were of the same gender and declension class. Two word form modifications to the nouns were selected: (1) diminutive suffixes (*-(i)ukas* for masculine and *-ytė* for feminine nouns) and (2) addition of the invented prefix *pa-* (for the example, see Table 1). The prefix *pa-* is very often used with verbs, but not with nouns. This prefix was chosen to create natural-sounding Lithuanian stimuli and to avoid allophonic variation in word onsets, even though the formation *pa-* plus noun is ungrammatical². There is no such word as *pavarlė* in Lithuanian, but the word *varlė* ‘frog’ exists and the prefix is added to create a novel form.

Position of word form modification was counterbalanced across trials (beginning versus end).

Table 1. An example of stimuli

Stimuli Conditions	Words	Novel words
Simplex	<i>ožys</i> ‘goat’	<i>lodinė</i>
Target end change (diminutive suffix)	<i>ož-iukas</i> ‘goat-DIM’	<i>lodin-ytė</i>
Target beginning change (prefix <i>pa-</i>)	<i>pa-ožys</i>	<i>pa-lodinė</i>
Alternate end change	<i>rykl-iukas</i> ‘shark-DIM’	<i>gokad-ytė</i>
Alternate beginning change	<i>pa-ryklys</i>	<i>pa-gokadė</i>

2.2. Procedure

Each child was tested individually in a single session of approximately 10 to 15 minutes duration. The children were invited to play a game of pointing at the picture with a different animal. They were told that some of the animals look very strange and have unfamiliar names and that they should try to guess which animal might be called by the given name.

Two training trials were used to ensure understanding of the task. Each trial began with the introduction of the first animal picture using the simplex form of the animal name (page one). An experimenter would make a statement (*Čia yra ožys* ‘this is a goat’), and a child was asked to repeat the word. The second page was an experimental one: the animal from the first page was shown and a new picture

² The prefix derivation in nouns is only possible together with inflectional change. The prefix addition to a noun changes not only its inflection, but also the meaning of the new derived noun, for example, *lang-as* ‘window-MASC’ and *pa-lang-ė* ‘windowsill-FEM’.

was added. The second page could contain either two familiar or two unfamiliar animals. A suffix or prefix was attached to the animal names. The child was asked to point at the animal in question (*Parodyk man, kur yra pa-ožys?* ‘Show me where a *pa*-goat is?’). The same order was followed while working with unfamiliar animals and novel words. The children were instructed to point at a required picture, and their responses were recorded and coded. In the case of wrong responses (i.e., pointing) the children were not corrected by the experimenter.

Children’s ability to select the right animal during the experimental trials was measured according to the following parameters: word form modification (beginning vs. end), the structure of the word (two-syllable words vs. three-syllable) and noun familiarity (familiar vs. novel). Age and gender of a child was taken into consideration as well.

3. Results and discussion

An ANOVA was used for the data analysis. The preliminary analyses of the experiment indicated no significant difference in accuracy of performing the task for boys and girls (81.7% correct for boys, 82.7% for girls), and child age was not correlated with the results of accuracy. Preliminary analysis also failed to show a significant difference in accuracy for bi-syllabic versus tri-syllabic nouns (84.8% correct for bi-syllabic nouns, 79.7% correct for tri-syllabic nouns). Preliminary analyses showed significantly higher accuracy for feminine nouns than for masculine nouns (85.5% correct for feminine nouns, 78.9% correct for masculine nouns). Better performance with feminine nouns was not anticipated, as there is a greater number of masculine nouns in Lithuanian (the masculine gender is unmarked), and indicates that for our stimuli (both familiar and novel nouns), the feminine word forms were easier for the children to remember than their masculine counterparts.

The Lithuanian children were significantly more accurate with suffix derivation than variation involving prefixation (85.5% vs. 78.9%), and more accurate with familiar than novel nouns (87.9% vs. 76.6%). The influence of position suggests that Lithuanian children were more sensitive to diminutive derivations as final word form modification than prefix derivation as initial word form modification. The influence of final position, however, could have been expected because of a productive Lithuanian diminutive derivation, as opposed to a non-productive modification of word onset position in the sense that the onset modification was ungrammatical. Lithuanian diminutives tend to be stressed suffixes, which makes them more acoustically salient than other parts of the word.

Another significant factor seems to be familiarity. The children were making more errors in novel nouns than in familiar ones. There was also a tendency noticed to the effect that children make more errors in three-syllable words than in two-syllable words. This holds for both familiar and novel nouns even though the results are not statistically significant. It may perhaps suggest that longer words are more difficult to remember.

One more conclusion follows from the data analysis, and this is that the children did better with the diminutive nouns than with *pa*-prefixed nouns when they hear familiar words. For novel nouns, the children are not better with the diminutives.

The interesting difference between these results and English data is that American children show a strong preference for picking a “different” picture throughout the task (Bruening, Brooks 2006). That is, they do not like morphological variation at all, and tend to select the new picture when they hear the new word form. The Lithuanian children do not show this bias and are equally good at trials where they “should” pick the same picture and when they “should” pick a different picture. This might be explained by the typological differences of the two languages.

In his famous paper on “operating principles” that guide early language learners, Slobin (1973) proposes that the ends of the words are more salient than the beginnings, therefore the former are easier for acquisition. In the case of morphologically rich languages young children encounter modifications of word final position all the time. Starting with a very early period of language acquisition, Lithuanian children hear words that consist of two elements, i.e. “word = stem + ending”, whereas English-speaking children (English is referred to as a “poor morphology language”) get familiarized with a word structure “word = stem”. This word structure is acquired early; therefore, children can identify the modifications of word final position quite easily. Similar results were also provided by research on second language acquisition (Ellis 1997) where the operating principle “pay attention to word end” (Slobin 1973) was indicated as very helpful in the process of language acquisition.

Word learning strategies are influenced by the typology of the language. The cues to systematic features of a language are available in the speech input directed to children and provide children with information about the typology of the language as a whole. Thus, language-learning strategies are shaped by the child’s experience of the language.

Abbreviations

DIM	diminutive
FEM	feminine
MASC	masculine

References

- Bruening, Paul; Brooks, Patricia 2006. Children’s tolerance of word-form variation. – Poster presented at Boston International Conference for the Study of Child Language, July 2006.
- Clark, Eve 1993. *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.
- Clark, Eve 1995. Later lexical development and word formation. – P. Fletcher, B. MacWhinney (Eds.). *The Handbook of the Child Language*. Oxford: Blackwell, 393–412.
- Dabašinskienė, Ineta 2009. Šnekamosios lietuvių kalbos morfologinės ypatybės. [Morphological features of spoken Lithuanian]. – *Acta Linguistica Lituonica*, 60, 1–16.
- De Marco, Anna 1998. The acquisition of diminutives in Italian. *Studies in the acquisition of number and diminutive marking*. – S. Gillis (Ed.). *Antwerp Papers in Linguistics*, 95, 175–192.
- Ellis, Rod 1997. *Second Language Acquisition*. Oxford Introduction to Language Study ELT. Oxford: Oxford University Press.
- Gillis, Steven 1997. The acquisition of diminutives in Dutch. – W. U. Dressler (Ed.). *Studies in Pre- and Protomorphology*. Vienna: Österreichische Akademie der Wissenschaften, 165–179.

- Golinkoff, M. Roberta; Hirsh-Pasek, Kathy; Bailey, Lisa; Wenger, Neil R. 1992. Young children and adults use lexical principles to learn new nouns. – *Developmental Psychology*, 28 (1), 99–108.
- Markman, M. Ellen; Wasow, L. Judith; Hansen, B. Mikkel 2003. Use of the mutual exclusivity assumption by young word learners. – *Cognitive Psychology*, 47, 241–275. doi:10.1016/S0010-0285(03)00034-3
- Kempe, Vera; Brooks, Patricia; Mironova, Natalija; Pershukova, Angelina; Fedorova, Olga 2007. Playing with word endings: Exploring effects of morphological variation in the learning of Russian noun inflections. – *British Journal of Developmental Psychology*, 25, 55–77.
- Olmsted, Hugh 1994. Diminutive morphology of Russian children: A simplified subset of nominal declension in language acquisition. – Charles E. Gribble, et al. (Eds.). *Alexander Lipson: In Memoriam*. Columbus, OH: Slavica Inc., 165–209.
- Savickienė, Ineta 2003. *The Acquisition of Lithuanian Noun Morphology*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Savickienė, Ineta 2005. *Morfopragmatika: deminutyvų vartojimas dabartinėje lietuvių kalboje*. [Morphopragmatics: the use of diminutives in Modern Lithuanian.] – *Kalbotyra*, 54 (1), 91–100.
- Savickienė, Ineta; Dressler, U. Wolfgang (Eds.) 2007. *The Acquisition of Diminutives: A Cross-Linguistic Perspective*. *Language Acquisition and Language Disorders*, 43. Amsterdam, Philadelphia: John Benjamins.
- Savickienė, Ineta; Dressler, U. Wolfgang; Barcza, Virag; Bodor, Peter; Ketrez, Nihan; Korecky-Kröll, Katharina; Palmović, Marian; Stephany, Ursula; Thomadaki, Evangelina 2007. Diminutives as pioneers of derivational and inflectional development – a crosslinguistic perspective. – Sabine Laaha, Steven Gillis (Eds.). *Antwerp Papers in Linguistics*, 112. Antwerp: Universiteit Antwerpen, 81–92.
- Slobin, I. Dan 1973. Cognitive prerequisites for the development of grammar. – Ch. A. Ferguson, D. I. Slobin (Eds.). *Studies of Child Language Development*. New York etc.: Holt, Rinehart, Winston, Inc., 175–208.
- Wójcik, Paweł 1994. Some characteristic features of Lithuanian baby talk. – *Linguistica Baltica*, 3, 71–86.
- Wójcik, Paweł 2000. *The Acquisition of Lithuanian Verb Morphology: A Case Study*. Kraków: Quartis.

Ineta Dabašinskienė (Regional Studies Department, Vytautas Magnus University, Lithuania). Her research interests cover interdisciplinary areas such as socio- and psycholinguistics, especially first and second language acquisition, normal and impaired language development, language use and variation.

i.dabasiniskiene@pmdf.vdu.lt

TULETUSMORFEEMIDE OMANDAMINE VARAJASES LAPSEEAS: LEEDU EKSPERIMENT

Ineta Dabašinskiene

Vytautas Magnuse Ülikool

Artikkel annab ülevaate eksperimentidest, mille abil vaadeldi, kuidas suhtuvad eelkooliealised leedu keelt omandavad lapsed sõnavormi eri osade ehk sufiksrite ja prefiksrite varieerumisse. Katse stiimuliks olid pildid, millel oli kujutatud tuttavaid ja võõraid loomi, nende loomade kohta kasutati sõnavorme, milles varieerus kas sõna algus (sõnale lisati prefiks) või lõpp (deminutiivsufiks *-yt* või *-ukas*). Eesmärk oli teha kindlaks: kas lapsed peavad sufiksiga (deminutiivsufiks) ja prefiksiga vorme uuteks sõnadeks või mitte (prefiks *pa-*, mida tavaliselt kasutatakse koos tegusõnadega, liideti siinses katses nimisõnadele); kas laste oskus neid sõnu korrektselt kasutada on vastuolus väitega, et lapsed eelistavad uusi sõnu kasutada uute objektide tähistamiseks ning kuidas suhtuvad leedu keelt omandavad lapsed sõna lõpu ja sõna alguse varieerumisse. Eksperimentis osales 39 eelkooliealist last vanuses 2;2–6;8 (keskmine vanus 4;5). Tulemustest selgus, et lapsed aktsepteerivad sõnavormis lõpu varieerumist tunduvalt enam kui sõna algusosa varieerumist. Seega mõjutab asjaolu, et leedu keel on rikka muutemorfoloogiaga keel, kus grammatiliselt oluline info on sõnavormi lõpus, laste keeleomandamisstrateegiat.

Võtmesõnad: keeleomandamine, morfoloogia, sõnamoodustus, leedu keel

INTELLEKTUAALSEID VÕIMEID TÄHISTAVATE SÕNADE KASUTUSED INTELLEKTUAALSES TEGEVUSES EESTI KEELE MAAILMAPILDIS

Larissa Degel

Ülevaade. Artiklis uuritakse mõtlemise, mõistmise ja mälu seotud intellektuaalsete võimete rolle eesti keele maailmapildis. Uurimisobjektiks on intellektuaalseid võimeid tähistavad sõnad eesti keeles: *aju(d)*, *ajukäärud*, *aru*, *hallollus*, *intellekt*, *meel*, *mõistus*, *pea*, *taip*, *teadvus*. Uurimuse eesmärk on tabada sõnakasutuses peituvaid “naiivseid” ettekujutusi inimese intellektist. Uurimismaterjal pärineb Tartu Ülikooli kirjakeele korpusest, otsingumootorist Google, kasutatud on ka “Eesti kirjakeele seletussõnaraamatut”. Intellektuaalsed protsessid ja seisundid eesti keele maailmapildis esitatakse kontseptuaalsete metafooride kujul. Ehitatakse üles mõtlemise, mõistmise ja mälu metafoore.

Võtmesõnad: mentaalne maailmapilt, intellektuaalsfäär, naiivteooria, kontseptuaalsed metafoorid, eesti keel

1. Keel, mida me kasutame

Naiivteooriad – tavainimeste maailmavaatelised teadmised, arusaamad, ettekujutused materiaalse maailma nähtustest ja objektidest ning mittemateriaalsetest psüühilistest struktuuridest ja protsessidest – peituvad keele grammatikas, süntaksis, sõnavaras, sõnade tähendustes ja kasutustes. Keelt kasutades ei pane me tähelegi, et eri konstruktsioonidesse on kompaktselt ja kõnelejale nähtamatult “pakitud” maailmavaatelisi teadmisi, konnotatsioone, assotsiatsioone, tähendusvarjundeid. Keelelise maailmapildi teooria järgi näevad, mõistavad, tunnetavad eri keelte kõnelejad maailma pisut erinevalt, omamoodi – läbi oma emakeele prisma. Siinses artiklis käsitletakse intellektuaalset tegevust intellektuaalseid võimeid tähistavate sõnade kasutuste põhjal. Sõnakasutused näitavad, missugused on intellektuaalsed võimed eesti keele maailmapildis, milline on intellektuaalsete võimete roll intellektuaalses tegevuses ning kuidas kulgevad protsessid ja seisundid eesti keele maailmapildi intellektuaalsfääris.

1.1. Uurimuse eesmärk

Seni ilmunud intellektuaalsfääri uurimustes on tähelepanu keskmes olnud eeskätt sõnavara ning intellektuaalsfääri protsessid ja seisundid (D'Andrade 1987, D'Andrade 1995, Apresjan 1995, Fortescue 2001), kuid varju on jäänud intellektuaalsed võimed, nende rollid intellektuaalsfääri protsessides ja seisundites, viisid, kuidas need protsessid ja seisundid kulgevad intellektuaalsfääri sees. Intellektuaalsfääri uurijatest on intellektuaalse tegevuse metafoore seni püüdnud käsitleda ainult Olaf Jäkel (1995), kelle inglise keele andmete põhjal üles ehitatud mudeli (ingl *Mental activity is manipulation*) moodustavad mõtted (*ideas*), *mind* kui teadlikkus (*mind-consciousness*) ja *mind* kui intellekt (*mind-intellect*), mälu (*memory*), teadmised (*knowledge*), probleemid (*problems*) ja lahendused (*solutions*) ning tegevused, nagu mõistmine (*understanding*), probleemi lahendamine (*solving a problem*), mentaalne uurimine (*mental investigation*), mõtete väljakujunemine (*forming ideas*), otsustamine, arutlemine (*judging*), meenutamine, meenumine (*memorizing, remembering*), meelespidamine (*keep in memory*), unustamine, ununemine (*abandoning ideas*). Mudel esitab mentaalset tegevust kui inimese poolt sooritatavat objektide käsitlemist (*physical manipulation of solid objects*), mille puhul intellektuaalsed võimed täidavad instrumendi rolli (*the mind is a tool*) või esinevad töökojana (*the mind is a workshop*).

Intellektuaalsete võimete olemasolust sõltub, kas inimene on tark või rumal. Kui inimesel on intellektuaalne võime olemas (inimene on nt *peaga, ajudega, aruga* või inimesel *on mõistust, nutti*), siis on inimene tark, arukas. Kui inimesel intellektuaalset võimet pole (nt *pole aru peas*, inimene *on mõistusetu, ajudeta*), siis on ta rumal. Intellektuaalsed võimed osalevad intellektuaalsetes protsessides ja seisundites.

Käesoleva uurimuse eesmärk on uurida mõtlemise, mõistmise ja mäluiga seotud intellektuaalsete võimete rolle eesti keele maailmapildis. Kas eesti keele maailmapildis piirduvad rollid instrumendiga ja töökojaga (nagu äsjamainitud Jäkelil)?

1.2. Uurimisobjekt, -materjal, -meetod

Uurimisobjektiks on intellektuaalseid võimeid tähistavate sõnade (*aju(d), ajukäärud, aru, hallollus, intellekt, meel, mõistus, pea, taip, teadvus*)¹ kasutused intellektuaalse võime tõlgenduses. Rolli määratlemisel keskendusin lauses rakendatud metafoorile. Analüüsisin sõnakasutuste põhjal ehitasin üles intellektuaalsete võimete rollide, intellektuaalsete protsesside või seisundite kontseptuaalsed metafoorid, mis süsteemsetes kognitiivsetes projektsioonides peegelduvad konkreetsetelt abstraktselt.

Nagu kõigest muust, mida pole võimalik silmaga näha või käega katsuda, räägime ka intellektuaalsetest protsessidest ja seisunditest metafooride abil: *“It is virtually impossible to think or talk about the mind in any serious way without conceptualizing it metaphorically”* (Lakoff, Johnson 1999: 235). Mõtlemise, mõistmise, mälu protsesside ja seisundite esitamisel ehitatakse metafoor üles järk-järgult (ülevalt alla) ning kõigi komponentide kaupa.

Põhimaterjal pärineb eesti kirjakeele korpusest ja eestikeelsest otsingumootorist Google (kogutud ajavahemikul oktoober 2006 kuni aprill 2007). Kasutatud on

¹ Peale loetletud sõnade tähistavad intellektuaalseid võimeid ka *mutts, oim, oid, nupp, nutt, talk*, aga kuna need sõnad esinesid analüüsitud materjalis vaid üksikutes näitelauasetes, siis neid artiklis ei käsitleta.

ka “Eesti kirjakeele seletussõnaraamatu” (edaspidi EKSS) näitelauseid. Analüüsitud on materjali kogu oma mitmekesisuses: ilukirjanduslikest tekstidest netikommentaarideni, eeskätt tänapäeva keelekasutust. Materjali kogumisel ja analüüsil olen püüdnud hõlmata nii palju kättesaadavat materjali kui võimalik, jättes siiski kõrvale kasutused, mille puhul on Google’is esinenud ainult üks kontekst, s.t mis on vaadeldavas tähenduses ja vormistuses esinenud ainult üks kord. Iga uuritava sõna puhul on vaadeldud selle kõiki muutevorme.

2. Intellektuaalsfääri sisu

Intellektuaalse tegevuse ja intellektuaalsete võimete rollide metafooride ülesehitamisel rakendatakse käesolevas artiklis koondkategoriat *INFORMATSIOON*, mis on metafooride konstrueerimise lähtepunkt. *INFORMATSIOON* hõlmab intellektuaalsfääri täitvat sisu kogu oma mitmekesisuses. *INFORMATSIOON* esineb keelises maailmapildis aina:

INFORMATSIOON ON AINE

Ainena esineval *INFORMATSIOONIL* on palju erinevaid ilmeid ja nimesid: *INFORMATSIOON* võib esineda mõtete, ideede, ootuste, lootuste, soovide, veendumuste, arvamuste, hinnangute, oletuste, unistuste, plaanide, kiindumuste, murede, probleemide, mälestuste, teadmiste vms nime all (1). *INFORMATSIOONI* moodustavad mittemateriaalsed, abstraktsed ideed, probleemid, mõtted jms “esemestuvad” keelises maailmapildis: *INFORMATSIOON* võtab objektide või nähtuste kuju, täidab mahutit, nagu seda teevad esemed, füüsilised objektid (2).

- (1) Loomulikult on nad esimesed **võimalused, mis pähe tulevad.** (SEGAILU/0004)
- (2) --- “Sel ajal olin mina kaheksateistkümnepäevane noor poiss, **pea täis pliksid ja mässuvõitlust.**” (SEGAILU/0024)

George Lakoffi ja Mark Johnsoni (1980: 25) terminoloogias on “esemestumise” puhul tegu nn ontoloogilise metafooriga, mis lubab rääkida abstraktsetest mõistetest ja mittemateriaalsetest nähtustest konkreetsete, füüsiliste objektide kaudu. Lakoffi ja Johnsoni (1999: 235–241) *Embodied Mind*’i metafoorides esinevad mõtted olenevalt käsitletavast fragmendist kui füüsilised objektid (metafooris *The Mind as Body System*, kus mõtlemine on füüsilise keha funktsioneerimine), asukohad (metafooris *Thinking is Moving*, kus mõtlemine on liikumine), vastuvõetavad, tajutavad asjad, objektid (metafooris *Thinking is Perceiving*, kus mõtlemine on vastuvõtmine), käsitletavad objektid (metafooris *Thinking is Object Manipulation*, kus mõtlemine on objektide käsitlemine), toit (metafooris *Acquiring Ideas is Eating*, kus mõtete omandamine on söömine). Mudeli *Mental activity is manipulation* üks esimesi ja põhjanevaid komponente on metafoor *Ideas are solid objects*. Toetudes ingliskeelsele materjalile, näitab Jäkel (1995: 199), et mõtted omavad keeles tahkete füüsiliste objektide kõiki olulisi omadusi: mõte on mitmemõõtmeline, mõtete keerulisus esineb kui kõvadus või sitkus, mõtete tähtsus esineb kui füüsiline kaal, mõtteid ja teadmist esitatakse kui väärisesemeid.

2.1. Intellektuaalsed võimed INFORMATSIOONI mahutina

Intellektuaalsete võimete üks võimalikest rollidest on INFORMATSIOONI mahuti. Mahuti võib olla INFORMATSIOONIGA täidetud või INFORMATSIOONIST tühi. INFORMATSIOON läbib mahutis kolm etappi: ilmub mahutisse, viibib mahutis, lahkub mahutist. INFORMATSIOONI ilmumine, viibimine, lahkumine on üks võimalikke viise, kuidas keel esitab intellektuaalseid protsesse ja seisundeid:

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON MAHUTI, KUS PAIKNEB AINE

INFORMATSIOON ilmub mahutisse kolmel erineval moel: tekib mahuti sees, väljendades oma esmakordset esiletulekut (3), siseneb mahutisse väljastpoolt (4) või seda sisestatakse mahutisse (5). Mahutis viibivat INFORMATSIOONI iseloomustab kaksipidine käitumine – aktiivne ja passiivne. Aktiivne käitumine esineb eeskätt liikumisena (6), passiivne olek on mahutis hoiul viibimine (7). INFORMATSIOON võib inimese tahtest olenemata mahutist lahkuda (8).

- (3) --- minu **peas sündis** kujutlus veiniteatmikust, milline meil võiks leida. (Google; www.blog.ekspress.ee)
- (4) Esimese mõttena **torkas mõistusesse**, et tegu on liiklusummikuga --- (Google; jah.pri.ee)
- (5) Bussisõidu ajal kipuvad vägisi pähe tulema kunagi **ajju topitud** teadmised komparteist --- (Postimees_Extra_2000)
- (6) Eks sellepärast enesetapukatsed poomise läbi nii kergesti korda lähegi, **jookseb mõte läbi aju**. (SEGAILU/0012)
- (7) Ametnik, kellel **maaseadus peas**, dikteeris Mikale ---. (ILU1990/ilu0078)
- (8) Ta oli viivuks sellele mõelnud, kuid siis **lendas** kõik **peast ...** kuni oli Aidanni külatüdruku seltsis puude vahel näinud. (SEGAILU/0107)

2.2. Intellektuaalsed võimed instrumendina

Intellektuaalne tegevus võib esineda keelelises maailmapildis kui intellektuaalsete võimete inimesepoolne kasutamine instrumendi või ruumina. Inimene kasutab intellektuaalseid võimeid instrumendina (9–10), et midagi saavutada või mingit tegevust sooritada:

INTELLEKTUAALNE VÕIME ON INSTRUMENT
INTELLEKTUAALNE TEGEVUS ON INSTRUMENTI INIMESPOOLNE KASUTAMINE

- (9) See küsimus ongi selleks välja mõeldud, et inimene oma **hallollust kasutaks**. (Google; www.baka.ee)
- (10) Hakkab **peaga**² leiba **teenima**. (EKSS)

² Eesti keele komiteativ võib esineda nii kaaslase kui ka vahendina, millele juhib tähelepanu Helle Metslang (2002: 166–167). Näites (10) esineb intellektuaalseid võimeid tähistav sõna komitatiivis, mis toimib instrumentaali funktsioonina, märgib vahendit.

Kasutades intellektuaalseid võimeid instrumendina, sooritab inimene nii mentaalsesse maailmapilti mahtuvaid tegevusi kui ka tegevusi, mis ei kuulu mentaalsesse maailmapilti, nt tegevusi, mille eesmärgiks on materiaalse heaolu saavutamine, materiaalsete hüvede teenimine ja nende suurendamine:

- (11) Neljapäeval, 5. juunil, peaks sul olema võimalik **mõistuse abil** oma **varandust suurendada** ---. (Eesti_Ekspress_1997)

2.3. Intellektuaalsed võimed ruumina

Inimene võib kasutada intellektuaalseid võimeid ka ruumina, kus ta sooritab tegevust:

INTELLEKTUAALNE VÕIME ON RUUM
INTELLEKTUAALNE TEGEVUS ON INIMESE TEGUTSEMINE RUUMI SEES

Teen põhimõttelist vahet intellektuaalsete võimete mahutina esinemise ja ruumina kasutamise vahel. See on tingitud intellektuaalsete võimete eristuvatest rollidest: mahutis paikneb INFORMATSIOON, samas kui intellektuaalset tegevust sooritav inimene viibib intellektuaalsetes võimetes kui ruumis. Jäkel (1995: 217) käsitleb intellektuaalseid võimeid selles rollis töökojana (*The mind is a workshop*), kus ideed on sätitud riulitesse vastavalt nende tähtsusele, jättes seega tähelepanuta töökoja sees toimuva tegevuse. Siin artiklis kasutan Jäkeli (1995) *mental workshop*'i vastena ruumi (mitte mahutit).

Intellektuaalsetes võimetes ruumina võib inimene tajuda, nt näha (12) ja kuulda (13). Intellektuaalsetes võimetes ruumina võib inimene sooritada arvutamist (14).

- (12) --- mida ma oma **peas näen** on kusagil ka tõepoolest olemas ja mingil isemoodi moel olen ma seal ka käinud --- (Google; foorum.keskmaa.ee)
- (13) Vahel **kuulen peas** suurepäraselt muusikat, kuid ei suuda seda paberile panna. (Google; www.poogen.ee)
- (14) Millegipärast jõuab laev Paldiskisse tunduvalt varem, kui oma **peas olin arvutanud**. (Eesti_Ekspress_2001)

2.4. Intellektuaalsed võimed tegijana

Intellektuaalsed võimed võivad eesti keele maailmapildis esineda tegijana ja sooritada tegevust:

INTELLEKTUAALNE VÕIME ON TEGIJA
INTELLEKTUAALNE TEGEVUS ON TEGIJA TEGUTSEMINE

Erinevalt instrumendi rollist (mis hõlmas nii mentaalset kui ka mittementaalset maailmapilti), sooritavad intellektuaalsed võimed tegijana ainult mentaalse maailmapildi tegevusi: nii neid, mis kuuluvad intellektuaalsfääri (15), kui ka neid, mis

hõlmavad emotsioone ja tundeid (16), soove (17), kõnet (18), nt intellektuaalne võime võib tunda, tahta või rääkida.

- (15) Kui meie **mõistus mõistab** elu valesti, siis keha kannatab. (Google; www.luuleviilma.ee)
- (16) Inimese **mõistus tunneb** ennast koduselt ainult siis, kui ta saab pidevalt avastada väikesi ebatäiuslikkusi, ebakorrapära, ebaloogilisust jne. (Google; www.postimees)
- (17) **Mõistus tahab** ise areneda, andke talle see võimalus ja ta vastab teile tänulikkusega. (Google; www.aasiareisid.ee)
- (18) --- aga **minu aru ütleb**, et kui kogu su tegevuse ja olemise aluseks on vabadus ---. (Google; www.newflow.ee)

3. Intellektuaalsete võimete rollid mõtlemisel

Mõtlemine esineb eesti keele maailmapildis nii mahuti sees toimuva, kui ka inimese poolt või intellektuaalsete võimete poolt sooritatava protsessina. Mahuti sees kulgevat mõtlemist mõistetakse kas liikumise, töö(tlemise) või muut(u)misena, kusjuures kõigil juhtudel kehtib metafoor INFORMATSIOON ON AINE. Mõtlemisel-liikumisel toimub mahutis liikumine, nii et kontseptuaalse metafoori võib sõnastada järgmiselt:

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON MAHUTI
MÕTLEMINE ON AINE LIIKUMINE MAHUTI SEES

Mõtlemine esineb keelelises maailmapildis kui inimesest osaliselt sõltumatu ja osaliselt sõltuv. Inimesepoolset kontrolli väljendab INFORMATSIOONI liikumise käivitamine (19–20) või INFORMATSIOONI keerutamine (21).

- (19) Sain kinnitust paljudele asjadele, **mida olin** varem oma **peast läbi lasknud**. (Eesti_Ekspress_1997)
- (20) --- **ma ei lase** seda **teadvusest läbi**, aga ma tahan nendega tuttavaks saada, sinna sisse saada. (Eesti_Ekspress_1999)
- (21) **Keerutan peas** ka romaanifragmente. (Eesti_Ekspress_1998)

INFORMATSIOON liigub mahuti sees erineva kiiruse ja intensiivsusega. Hoogsat, palavikulist, kiiret mõtlemist kirjeldab INFORMATSIOONI kiirem liikumine – lippamine (22), jooksmine (23), tormlemine (24).

- (22) Leena tundis, et mõtted **lippavad** nii kiiresti **läbi pea**, et ta ei jõua mitte ühtegi neist lõpuni mõelda. (SEGAILU\0008)
- (23) Eks sellepärast enesetapukatsed poomise läbi nii kergesti korda lähegi, **jookseb** mõte **läbi aju**. (SEGAILU\0012)
- (24) Väliselt ei öelnud miski mulle, et peale **ajukäärudes tormlevate** ideede võiks --- mehel veel midagi hirmväärtslikku üle ääre pulbitseda. (Google; arhiiv2.postimees.ee8080)

Intensiivne mõtlemine esineb INFORMATSIOONI keerlemise (25), tiirutamise (26), pöörlemisena (27) mahutis. Peale ringlemise võib mõtlemise intensiivsust edastada INFORMATSIOONI monotoonne edasi-tagasi liikumine, tagumine (28), millega kaasneb (tajutav, kuuldav) kloppimine, vasardamine (29), tuksumine, põksumine, trummeldamine (30). Vähem intensiivset mõtlemist iseloomustab INFORMATSIOONI aeglasem, sujuvam ja voolavam liikumine (31–33).

- (25) Kõik see, mida ta selle maailma kohta teada saada jõudnud oli, hakkas ta **teadvuses keerlema**, ja peagi ei jaksanud ta enam isegi muusikat kuulata. (SEGAILU\0004)
- (26) **peas** aga **tiirutab** homme Tallinna-sõit või mõni päevauudistes kuuldud teade. (ILU1990\ilu0246)
- (27) Keegi ei wastanud temale enam, igäühe **ajus pöörles** pime, raske mõte, küümus, et miks on elus nii palju waewa, wiletbußt, nälga? (ILU1910\ilu0077)
- (28) “Välja siit!” **taob** kinnismõte **läbi pea**. (ILU1990\ilu0015)
- (29) Miks just mina, **vasardab** mõne tubli pereisa **ajus**, kes on kolme õlle pärast autoroolist maha võetud ja ränga trahvi saanud. (Google; vabaeuroopa.org)
- (30) --- nüüd **trummeldab hallolluses** pidevalt üks asi – seks, seks, seks (Google; www.ohtuleht.ee)
- (31) Sellised küsimused **sõlmuvad** üksikisiku **meeles** ---. (Eesti_Ekspress_1999)
- (32) --- vaataja **teadvuses heljuvad** kujundid siiski peidavad eneses satori, flashi võimalust ---. (Google; www.postimees.ee)
- (33) Tema **peas uitab** ainult üks mõte - tahaks venda juba näha, nii palju on aega. (Google)

Informatsioon ei pruugi liikuda ainult kindlas suunas (edasi, edasi-tagasi, ringi, läbi või mööda). Kiiret, palavikulist mõtlemist iseloomustab korrapäratus – informatsiooni liikumine võib muutuda kaootiliseks (34), käia korraga mitmes erinevas suunas (35), seguneda (36).

- (34) Sõnad ja mõtted **keerlevad - tiirlevad - joonduvad - liuglevad** tema **peas** nagu purjus baleriinid (Google; www.poogen.ee)
- (35) Need küsimused **ristlesid ajus**. (ILU1930\ram0007)
- (36) Läbielatu, muljed ja mälestused **segunesid** mu **ajus** üheks suureks seikluseks ---. (SEGAILU\0007)

INFORMATSIOON ei pruugi elada aktiivset elu, liikuda ja muutuda, ta võib viibida mahutis staatilises olekus. Staatilisus avaldub kahtepidi: kui seismine/istumine ja kui hoiul olek. INFORMATSIOONI seismine/istumine väljendab mõtlemist – võib oletada, et see on mõtlemise kui INFORMATSIOONI liikumise seiskumine. Seiskumine väljendab

mõtlemise intensiivsust, pidevat, lakkamatut mõtlemist, keskendumist ainult ühele mõttele (37). INFORMATSIOONI mahutis hoiul olek väljendab mäletamist.

- (37) Kui isegi satun korraks lugema, **istub peas mõte**, et jää vait ja kasva suureks ning saa üle juba. (Google: www.parnupostimees.ee)

Peale liikumise võib INFORMATSIOONI aktiivsus avalduda muut(u)misena – mahutis viibides võib INFORMATSIOON (välja) kujuneda (38), valmida (39), küpseda (40) või INFORMATSIOONI võib arendada (41), küpsetada (42). INFORMATSIOON muutub ainult väljakujunemise, arenemise suunas (INFORMATSIOON ei lagune ära ega hävine, vaid areneb, küpseb, valmib). INFORMATSIOONI muutmise on võimalik samuti ainult arendamise, küpsetamise suunas.

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON MAHUTI
MÕTLEMINE ON AINE MUUT(U)MINE MAHUTIS

- (38) Pikapeale **kujunes** tema **teadvuses** kolooniagi mingiks suhtelise vabaduse ja iseseisvuse võrdkujuks, kui koduga kõrvutada. (ILU1970\ilu0028)
- (39) Minu loengud **valmivad** mu **peas**. (Postimees_1997)
- (40) --- tegelikult **küpses** prügikorjamise idee tema **ajus** juba ammu. (Eesti_Ekspress_2000)
- (41) Seda mõtet **arendades** sünnib Theodori **peas** siiski Vabadussõja-ainelise “Noorte kotkaste” idee. (Postimees_1996)
- (42) Enamasti **küpsetab** fotograaf ideed oma **peas** nädalaid või kuidki. (Eesti_Ekspress_2001)

Mõtlemine võib esineda ka inimese poolt sooritatava tegevusena. Kui mõtlemist sooritab inimene, siis kasutab ta intellektuaalseid võimeid kas instrumendi või ruumina. Jäkel (1995: 206) esitab mõtlemist metafoorina *Thinking is working on problem objects with the mind tool*. Lausetes (43–44) esineb mõtlemine inimese poolt sooritatava tegevusena, mille puhul intellektuaalseid võimeid rakendatakse tööle instrumendina.

- (43) --- S. **mõtle aruga** ma saan aru et ta sulle meeldib --- (Google: www.lapsemure.ee)
- (44) Narkoennetuspäev ärgitas noori oma **peaga mõtlema** (Google; www.valgamaalane.ee)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON INSTRUMENT
MÕTLEMINE ON TÖÖ (INIMESE POOLT AINE TÖÖTLEMINE INSTRUMENDI ABIL)

Inimene võib kasutada intellektuaalseid võimeid ka ruumina, s.t mõtlemise sooritamise asukohana:

- (45) **Mõtlen oma peas** igasuguseid asju välja, et Ta saaks jälle öelda, et ma rahuneksin. (Google; blog.delfi.ee)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON RUUM
MÕTLEMINE ON TÖÖ (INIMESE POOLT AINE TÖÖTLEMINE RUUMIS)

Intellektuaalsed võimed võivad sooritada mõtlemist ise, tegijana:

- (46) Tema **aju töötab mitmete erinevate päästmisplaanide kallal**: kas pakkuda töölisele raha, et nad kaju maha ei tõmbaks? (SEGAILU/0024)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON TEGIJA
MÕTLEMINE ON TÖÖ (TEGIJA POOLT AINE TÖÖTLEMINE)

D'Andrade (1987: 135) kirjeldab mõtlemist kui mõtetest mõtete tootmist: *“Thoughts are related to each other. Sometimes, one thought leads to another; sometimes one recognizes inconsistency between thoughts; sometimes one can figure out something from other things one knows or believes”*. Mõtlemine võib ka eesti keeles esineda intellektuaalsete võimete poolt sooritatava INFORMATSIOONI tootmise (47), loomise (48), konstrueerimisena (49).

- (47) No see on muidugi eriti äärmuslik kriisisituatsioon, mille minu pisut väärustunud **mõistus tootis**, aga loodan, et keegi veel saab aru, mida ma silmas pidasin (Google; juulik.wordpress.com)
- (48) Märkan äkilise selgusega, et **mõistus, intellekt, loob** omad probleemid ja teeb seejärel olemise viletsaks, üritades neid lahendada. (Google; teofilosoofia.ee)
- (49) Tema loominguline **aju hakkas** kohe **konstrueerima** oletusi uue teksti kohta ---. (SEGAILU\0115)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON TEGIJA
MÕTLEMINE ON AINE TOOTMINE TEGIJA POOLT

4. Intellektuaalsete võimete rollid mõistmisel

Mõistmine esineb keelelises maailmapildis põhiliselt INFORMATSIOONIGA läheduse saavutamisenä. Jäkeli (1995: 199–205) üles ehitatud mudelis *Mental activity is manipulation* koosneb metafoor *Understanding an idea is establishing physical closeness* neljast komponendist: mõtleja otsib mentaalse vahemaa vähendamise võimalusi, haarab mõttest kinni, võtab selle maast üles, paneb mahutisse.

INFORMATSIOON ON AINE
MÕISTMINE ON AINEGA LÄHEDUSE SAAVUTAMINE

Eesti keele maailmapildis on võimalik nii inimese või intellektuaalsete võimete aktiivne tegutsemine kui ka INFORMATSIOONI enese jõudmine mahutini/mahutisse. Lause (50) esitab mõistmist INFORMATSIOONI jõudmisena mahutisse, lause (51) jõudmisena mahutini. INFORMATSIOONI jõudmist mahutisse võib takistada INFORMATSIOONI

sisu – absurdse (52), imeliku, kummalise (53) sisuga INFORMATSIOON ei mahu mahu-tisse.

- (50) “Loodetavasti **jõuab** nüüd ka vägivaldsete **ajudesse**, millist häbi nad oma käitumisega Inglismaale on teinud,” (Postimees_2000)
- (51) Et ka vanus enamikust mööda ei lähe, **pole jõudnud** veel **ajukääruden**i. (Postimees_1998)
- (52) Mul **ei mahu** see **pähe**, kuidas saab neljakümne häälega pääseda linna volikogusse. (AJAE1990\stako233)
- (53) Punk on midagi kummalist, mis lihtkodaniku **ajudesse ei mahu**. (Google; vooremaa2001.tripod.com)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON MAHUTI
MÕISTMINE ON AINE JÕUDMINE MAHUTISSE/MAHUTINI

Mõistmine esineb eesti keele maailmapildis intellektuaalsete võimete poolt sooritatava INFORMATSIOONI vastuvõtmisena (54). Suutmatuse seda vastu võtta on suutmatuse mõista (55).

- (54) Omandab innukalt kõike, mida ta **taip suudab vastu võtta**. (EKSS)
- (55) K.-i **meel ei võta vastu** teadmist, et tema majapidamises talitas mees, kelle kohus võib tunnistada Eesti ajaloo julmimaks sarimõrvariks (Google; www.ohtuleht.ee)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON TEGIJA
MÕISTMINE ON AINE VASTUVÕTMINE TEGIJA POOLT

Mõistmine võib esineda eesti keeles ka inimese poolt sooritatava INFORMATSIOONI haaramisena, mille puhul intellektuaalseid võimeid rakendatakse instrumendina:

- (56) Venemaa kohta on tavaks öelda, et seda maad ja rahvast pole võimalik **mõistusega haarata**, teda saab uskuda või karta. (Google; www.riigikogu.ee)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON INSTRUMENT
MÕISTMINE ON AINE INIMESE POOLT HAARAMINE INSTRUMENTI ABIL

Mõistmist (kinni)võtmise, (kinni)haaramise kujul esitav metafoor *Understanding is seizing idea objects firmly* on üks komponentidest mudelis *Mental activity is manipulation* (Jäkel 1995: 202). Vakk (1970: 31) juhib tähelepanu sellele, et paljudes indoeuroopa keeltes tuntakse mõistmist väljendavaid verbe, mis algselt tähendasid just (kinni)haaramist, sks *begreifen*, rts *begripen*, pr *comprendre* ‘aru saada, mõistma’ < ld *comprehendere* ‘kinni haarama, kinni võtma, kinni püüdma’.

5. Intellektuaalsete võimete rollid mälu protsessides ja seisundites

Kõigi mälu protsesside ja seisundite asukohaks on intellektuaalsed võimed mahutina. INFORMATSIOONI meeldejäamine esineb jäämise (57) või settimisena (58) mahutisse.

- (57) Need read **jäid** mulle igaveseks **pähe**. (Eesti_Ekspress_1998)
- (58) Teinekord kokkusaamist otsitakse, ja ikka kindla pilguga, valmisolekuga märgata seda, mis kunagi on juba **teadvusse settinud** --- (Google; arhiiv2.postimees.ee8080)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON MAHUTI
MEELDEJÄÄMINE ON AINE MAHUTISSE JÄÄMINE

Meeldejäamisel püüab informatsioon sisenedes pääseda võimalikult sügavamale, võtta mahuti sees kindlamat ja stabiilsemat positsiooni. Selleks, et võtta mahutis kindlamat, püsivamat, stabiilsemat, tugevamat positsiooni, võib INFORMATSIOON käituda sisenemisel üsnagi agressiivselt – tungida nagu nael tungib terava otsaga ihu sisse (59), ajada juured alla, hiilida võimalikult sügavamale põhja, imenduda (60–62). INFORMATSIOONI võib sisestada mahutisse – jätta (63), panna hoiule (64) – hoidmise eesmärgiga.

- (59) Päeva jooksul kuulnud kõnedest olid talle kõige teravamini **meelde lõikunud** Madjaku sõnad ---. (ILU1930\ram0042)
- (60) Tundsin tüdruku kohe ära, tema pilt oli mulle erilisel **meelde sööbinud**. (SEGAILU0040)
- (61) Sõnad on talle ammu **pähe kulunud** ---. (Eesti_Ekspress_1997)
- (62) Üks mittemõistmise juhtum on mul eriti jäävalt **hallolusesse süüvinud**. (SEGAILU0100)
- (63) Ta vaatas tähelepanelikult toas ringi, tahtis täpselt **meelde jätta**, kuidas üks või teine asi seisab, et siis oleks pärast selge, mida ma segi ajanud olen. (SEGAILU/0018)
- (64) Ja mina mäletan kõike nii, nagu oleksin selle momendi fotografeerinud ja pildi **ajudesse hoiule pannud**. (ILU1980\stkto091)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON MAHUTI
MEELDEJÄTMINE ON INIMESE POOLT AINE MAHUTISSE JÄTMINE

Kui INFORMATSIOON viibib *peas* ja *meeles* staatilises olekus, väljendab see INFORMATSIOONI hoiul olekut ehk mäletamist³:

- (65) Puusepal on väga kehv nimemälu, **meeles seisavad** ainult üksikute tuttavate nimed. (Eesti_Ekspress_2001)

³ INFORMATSIOONI staatiline olek väljendab mäletamist ainult *pea* ja *meele* mahutite puhul. Olenevalt lause sisust võib *peas* staatilises olekus viibiv INFORMATSIOON väljendada ka intensiivset mõtlemist, nt lause (37).

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON MAHUTI
MÄLETAMINE ON AINE (HOIUL) OLEK MAHUTI SEES

Mäletamine esineb ka inimesepoolse sihikindla hoidmisena: et väärtuslik vajalik INFORMATSIOON ei lahkuks mahutist, tuleb see mahutis säilitada – pidada (66) ja hoida (67).

(66) --- ta suudab korraga **meeles pidada** kümne partii seise ilma lauale vaatamata. (Eesti_Ekspress_2001)

(67) hommikust saadik juba **meeles hoidnud** ja teistele meelde tuletanud, et täna Aadu muldasängitamise päev. (ILU1990\ilu0510)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON MAHUTI
MÄLETAMINE ON AINE INIMESEPOOLNE HOIDMINE MAHUTIS

Jäkel (1995: 214) kirjeldab meenutamist/meenumist (ingl *remembering*) kui “teadmiskuste tagasitoomist mälu laost mentaalsesse töökohta”. Kui INFORMATSIOON tõuseb mahuti seest, mahuti põhjast (68–69), siis tegu on mahutis olemasoleva INFORMATSIOONI aktiveerumise ehk meenumisega:

(68) Konkurentsilt **tõuseb meelde** Tobias George Smolletti kelmiromaani “Roderick Randomi seiklused” tõlge ---. (Postimees_1996)

(69) Ainult **ajus tõusevad** sisepildid minevikust ---. (ILU1980\stkt0054)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON MAHUTI
MEENUMINE ON AINE MAHUTIS TÕUSMINE

Meenutades liigutab inimene mahutis hoiul olevat INFORMATSIOONI. Liigutamine võib esineda keerutamise (70) kujul.

(70) Juba praegu, veel enne, kui see seiklus siin minevikuks on saanud, hakkab nostalgia vaikselt hinge pugema, **keerutades peas** mälestusi ---. (Google; katutsehhis.blogspot.com)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON MAHUTI
MEENUTAMINE ON AINE INIMESE POOLT LIIGUTAMINE MAHUTIS

INFORMATSIOONI liigutamine omab sarnaseid jooni mõtlemise ja kujutlusega (71). Mälus ja kujutluses võib inimene taastada seda, mida ta on näinud, tunnetanud, tajunud. Mälu erineb mõtlemisest ja kujutlusest INFORMATSIOONI sisu poolest: mälu INFORMATSIOON jääb tervikuna minevikku, mõtlemise ja kujutluse INFORMATSIOON ei seostu ainult mineviku, vaid ka oleviku ja tulevikuga. Mõtlemise tihedat seost teadmise, kujutluse, usu ja mäletamisega mainib Michael Fortescue (2001: 17) mentaalse tegevuse naiivteooria ühe põhjapaneva komponendina: “*When you think, you do so about what you know or imagine, or believe to be the case, or what you remember or intend doing ---*”. Ununemist väljendab INFORMATSIOONI lahkumine mahutist (72).

- (71) olen sageli avastanud end juba trükivalmis lauses mõtlemast, **keerutades peas** sõnu, mida võiks mu järgmisesse blogiposti kirjutada (Google; www.ekspress.ee)
- (72) --- nime peale on mul üldse sant mälu, nende peale ma mihkel ei ole, **kaovad kohe meelest**. (ILU1990/ilu0275)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON MAHUTI
UNUNEMINE ON AINE LAHKUMINE MAHUTIST

Inimene võib proovida (kas ise või millegi vahendusel, kellegi abil) tüütust INFORMATSIOONIST lahti saada – unustada või loobuda mõtlemast, mäletamast (73–74), kuid INFORMATSIOONIST lahtisaamine ei pruugi alluda inimese taatele (75).

- (73) “Enamik noori ja veel kogenematuid suusahüppajaid suudavad kõik ebaseeldiva **peast välja heita**,” ---. (Postimees_1997)
- (74) Vahel osutub raskeks **vabastada aju** muude murede kütkest ja keskenduda õppusele. (TEA1980\tt0053)
- (75) Üks selline sõlm oli seal ka, mis üldse enam **peast välja ei lähe**. (Postimees_1997)

INFORMATSIOON ON AINE
INTELLEKTUAALNE VÕIME ON MAHUTI
UNUSTAMINE ON MAHUTI INIMESE POOLT VABASTAMINE AINEST

6. Intellektuaalsed võimed eesti keele maailmapildis: tegija, mahuti, instrument, ruum

Jäkeli (1995) inglise keele andmete põhjal üles ehitatud mudel *Mental activity is manipulation* esitab mentaalset tegevust kui inimese poolt sooritatavat objektide käsitlemist, mille puhul intellektuaalsed võimed esinevad kas instrumendi või töökojana. Siin analüüsitud sõnade kasutused näitavad, et eesti keele maailmapildis ei piirdu intellektuaalsete võimete esinemine üksnes töökoja või inimese poolt kasutatava (käsitletava) tööriista rolliga. Eesti keele maailmapildis esinevad intellektuaalsed võimed neljas rollis: tegija, mahuti, instrumendi ja ruumina.

Tegijana tegutsevad intellektuaalsed võimed sooritavad intellektuaalset tegevust iseseisvalt, ilma inimese vahendusega. Tegijana figureerivad intellektuaalsed võimed mõtlemisel (INFORMATSIOONI töötlemisel ja tootmisel) ja mõistmisel (INFORMATSIOONI vastuvõtmisel). Tegija rollis sooritavad intellektuaalsed võimed ainult mentaalsesse maailmapilti kuuluvaid tegevusi, s.t nii neid, mis kuuluvad intellektuaalsfääri, kui ka neid, mis hõlmavad emotsioone, soove, kõnet.

Mahutina täidavad intellektuaalsed võimed väga tähtsat rolli mõtlemisel ja mõistmisel, aga eriti mälu protsesside ja seisundite puhul, kus mahuti on meeldejätmise ja -jätmise, mäletamise, meenu(ta)mise, ununemise ja unustamise asukoht.

Instrument ja ruum on kaks viisi, kuidas inimene kasutab oma intellektuaalseid võimeid. Instrumendina rakendab inimene intellektuaalseid võimeid mingi

tegevuse sooritamiseks tööriista, vahendina. Instrumendina kasutatakse intellektuaalseid võimeid mõtlemisel (INFORMATSIOONI töötlemisel) ja mõistmisel (INFORMATSIOONI haaramisel). Instrumendi rolli abil sooritab inimene nii mentaalsesse maailmapilti kuuluvaid tegevusi (sh teadmist, mõtlemist, mõistmist) kui ka tegevusi, mis ei kuulu mentaalsesse maailmapilti (nt tegevusi, mille eesmärgiks on materiaalse heaolu saavutamine, materiaalse hüvede teenimine ja suurendamine).

Ruumina kasutatakse intellektuaalseid võimeid mõtlemisel (INFORMATSIOONI töötlemisel), viibides intellektuaalsete võimete sees. Intellektuaalsete võimete sees võib inimene tajuda, nt näha ja kuulda (selles kontekstis pole tüüpiline intellektuaalsete võimete instrumendina kasutamine), teada, mõelda ja mõista, sooritada arvutamist, rehkendamist, lugemist, kirjutamist.

7. Kokkuvõtteks

Artiklis üles ehitatud metafoorid näitavad eesti keele maailmapildis peituvaid naiivseid teadmisi ja ettekujutusi inimese intellektist. Intellektuaalseid võimeid tähistavate sõnade kasutuste analüüs näitas, et intellektuaalsed võimed esinevad eesti keele maailmapildis tegija, mahuti, instrumendi või ruumi rollides.

Mõtlemisel esinevad intellektuaalsed võimed kõigis neljas rollis: tegija, instrumendi ja ruumi roll toimib INFORMATSIOONI töötlemisel ja tootmisel; mahutina toimub mõtlemine kui INFORMATSIOONI liikumine või muut(u)mine.

Mõistmisel esinevad intellektuaalsed võimed kolmes rollis: tegijana, kes võtab INFORMATSIOONI vastu; instrumendina, mille abil inimene haarab INFORMATSIOONI ja mahutina, kuhu INFORMATSIOON jõuab.

Kõigi mälu protsesside ja seisundite asukohaks on intellektuaalsed võimed mahutina. Kõik mälu protsessid ja seisundid moodustavad kontseptuaalsete metafooride paari, kus ühes metafooris on tegija INFORMATSIOON ja teises inimene: meeldejäamine (INFORMATSIOONI jäämine mahutisse) ja meeldejätmine (inimene sisestab INFORMATSIOONI mahutisse); meenumine (INFORMATSIOONI mahuti põhjast/seest tulek), meenutamine (inimene liigutab INFORMATSIOONI); mäletamine (INFORMATSIOONI mahuti sees hoiul olek ning see, et inimene hoiab INFORMATSIOONI mahuti sees); ununemine (INFORMATSIOONI lahkumine mahutist) ja unustamine (inimene vabastab end INFORMATSIOONIST).

Viidatud kirjandus

- Апресян 1995 = Апресян, Ю. Д. 1995. Избранные труды. Том II. Интегральное описание языка и системная лексикография. Москва: Школа “Языки Русской Культуры”.
- D’Andrade, Roy 1987. A folk model of the mind. – Dorothy Holland, Naomi Quinn (Eds.). Cultural Models of Language and Thought. London: Cambridge University Press, 112–148.
- D’Andrade, Roy 1995. The Development of Cognitive Anthropology. Cambridge: Cambridge University Press.
- Fortescue, Michael 2001. Thoughts about thought. – Cognitive Linguistics, 12 (1), 15–45. doi:10.1515/cogl.12.1.15
- Jäkel, Olaf 1995. The metaphorical concept of mind: “Mental activity is manipulation”. – John R. Taylor, Robert E. MacLaury (Eds.). Language and the Cognitive Construal

- of the World. Trends in Linguistics. Studies and Monographs, 82. Berlin, New York: Mouton de Gruyter, 197–229.
- Lakoff, George 1980. Women, Fire and Dangerous Things: What Categories Reveal about the Mind. Chicago, London: University of Chicago Press.
- Lakoff, George; Johnson, Mark 1999. Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought. New York: Basic Books.
- Metslang, Helle 2002. Grammatikalisatsiooniteooriast eesti keele taustal. – Renate Pajusalu, Ilona Tragel, Tiit Hennoste, Haldur Õim (toim.). Teoreetiline keeleteadus Eestis. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised, 4. Tartu: Tartu Ülikooli Kirjastus, 164–178.
- Vakk, Feliks 1970. Suured ninad murdsid päid ... (Pea ja selle osad rahvalike ütluste peeglis). Tallinn: Valgus.

Allikad

- EKSS = Eesti kirjakeele seletussõnaraamat. Köited I–IV. 1988–2005. Tallinn: Eesti Keele Instituut.
- Google = Google. <http://www.google.ee> (05.03.2009).
- TÜKK = Eesti Kirjakeele Korpus 1980–1990. <http://www.cl.ut.ee/korpused/kasutajaliides> (05.03.2009).

Larissa Degeli (Narva Linnavalitsuse kultuuriosakond) peamised uurimisvaldkonnad on leksikaalne semantika, kõrvutatav keeleteadus.
larissa.degel@narva.ee

THE CONCEPTUAL ROLES OF INTELLECTUAL ABILITIES IN THE WORLDVIEW OF ESTONIAN

Larissa Degel

Narva City Administration, Culture Department

This research focuses on the meaning of the linguistic expressions of intellectual abilities in Estonian that are related to the activities of thinking, understanding and the use of memory. Through this the paper analyzes the worldview reflected in these instances of use. A suitable descriptive tool for analyzing these roles is conceptual metaphors. The aim of the research was to construct conceptual metaphors of the expressions of thinking, understanding and using memory through the notion of intellectual ability. The analysis is based on the Corpus of Written Estonian and the material of Google search.

In the worldview of Estonian an intellectual ability is usually depicted as 1) an agent (a worker), 2) a container of intellectual processes or states, or as 3) an instrument which is used by a person, or 4) a room where a person carries out an intellectual activity. When talking about thinking the intellectual abilities were described by all of the above mentioned roles. When talking about the intellectual abilities in relation to understanding the roles of container, worker and instrument were used. When talking about memory the intellectual abilities were described only by the container of intellectual processes and states metaphor.

Keywords: mental worldview, intellectual shpere, folk theory, conceptual metaphors, Estonian

DIREKTIIVSED AKTIPAARID EESTIKEELSETES INFODIALOGIDES JA NENDE AUTOMAATNE TUVASTAMINE

Olga Gerassimenko, Riina Kasterpalu, Mare Koit, Andriela Rääbis, Krista Strandson

Ülevaade. Analüüsitakse eestikeelseid telefonikõnesid ametiasutustesse, eesmärgiga arendada intelligentset kasutajaliidest, mis võimaldaks suhelda andmebaasidega eesti keeles. Analüüs põhineb Tartu Ülikooli eesti dialoogikorpusel. Käesolevas artiklis keskendutakse dialoogiaktide direktiivsetele naabruspaaridele, mille esiliikme ütleb ametiasutusse helistav klient ja järelliikme ametnik, näiteks klient avaldab infosoovi ja ametnik vastab info andmisega. Kliendi esimene sooviavaldus püstitab eesmärgi, mille saavutamiseks teevad klient ja ametnik koostööd. Mõlemad osalejad võivad algatada alamdialooge, et täpsustada kas soovi või saadud vastust. Artiklis analüüsitakse direktiivsete aktipaaride keelelisi tunnuseid ja tutvustatakse eestikeelsetes dialoogides dialoogiaktide automaatse tuvastamise eksperimente.*

Võtmesõnad: suhtlus arvutiga, dialoogiakt, soov, soovi täitmine, eesti keel

1. Sissejuhatus

Loomulik keel on inimese jaoks kõige hõlpsam suhtlemisvahend, olgu suhtluspartneriks teine inimene või hoopis arvuti. Maailmas on loodud mitmeid kasutajaliideseid elektroonilistele andmebaasidele, mis võimaldavad kasutajatel esitada päringuid loomulikus keeles (McTear 2004, Minker, Bennacef 2004). Kogutakse dialoogikorpust, et uurida inimestevahelist suhtlust ja arvestada seda kasutajaliideste kavandamisel (Switchboard, Verbmobil, BNC jt).

Meie eesmärk on välja töötada kasutajaliides, mille vahendusel saaks andmebaasidega suhelda eesti keeles. Soovime luua sellise liidese, mis suudaks järgida inimestevahelise suhtluse reegleid ja norme. Eesmärgi saavutamiseks uurime eestikeelseid telefonikõnesid ametiasutustesse, kus klient soovib saada infot. Püüame

* Artikli valmimist on toetanud Euroopa Regionaalarengute Fond Eesti Arvutiteaduse Tippkeskuse kaudu, Eesti Teadusfondi grant nr 7503 ning Haridus- ja Teadusministeerium (sihtfinantseeritav teema SF0180078s08 "Loomulike keelte arvutitöötamise formalismide ja efektiivsete algoritmide väljatöötamine ning eesti keelele rakendamine" ja riiklik programm "Eesti keele keeletehnoloogiline tugi"). Täname anonüümseid retsensente.

selgitada, kuidas helistajad formuleerivad oma soove, kuidas neid soove käsitletakse kliendi ja ametniku koostöös, missuguseid keelelisi vahendeid kasutavad osalejad vestluses, ametikõnede üldist ülesehitust jms.

Eelmises artiklis (Gerassimenko jt 2007) keskendusime kliendi soovide keelelise väljendamisele ametikõnedes. Käesolevas artiklis käsitleme üldisemalt dialoogiaktide direktiivseid naabruspaare, kus kliendi soovile või ettepanekule järgneb selle täitmine või vastuvõtmine ametniku poolt. Selline naabruspaar võib olla lihtsaima infodialoogi tuum: näiteks soovib helistaja saada asutuse telefoni-numbrit ja ametnik teatab talle selle. Siin uurime, missugused direktiivsed aktid esinevad dialoogides tegelikult, missugused järelliikmed järgnevad teatud esiliikmetele, direktiivsete naabruspaaride asukohta dialoogis ning keelelisi tunnuseid, mis iseloomustavad neid dialoogiakte. Samuti kirjeldame eksperimente, mis viidi läbi eesmärgiga automaatselt tuvastada dialoogiakte, ning tõlgendame dialoogiaktide automaatselt leitud keelelisi ja dialoogi struktuurist tulenevaid tunnuseid.

Artikkel on üles ehitatud järgmiselt. 2. osas anname ülevaate empiirilisest materjalist – eesti dialoogikorpusest EDiK ja dialoogiaktide tüpoloogias, mida kasutatakse korpuse märgendamisel. 3. osa on pühendatud korpuse analüüsile. Vaatleme kliendi eri tüüpi direktiivseid akte ja ametniku vastuseid nendele. 4. osas käsitleme dialoogiaktide automaatse tuvastamise eksperimente ja 5. osas teeme kokkuvõtteid.

2. Korpus ja dialoogiaktide tüpoloogia

Meie uurimus põhineb eesti dialoogikorpusel EDiK¹ (Koit 2003). Korpuses on seisuga september 2009 kokku 1148 autentset inimestevahelist suulist dialoogi, enamus neist (1017) on telefonikõned, kus kliendid helistavad mitmetesse ametiasutustesse. Korpuses on märgendatud dialoogiaktid. Selleks on Tartu Ülikoolis loodud aktitüpoloogia (Hennoste, Rääbis 2004), mis lähtub vestlusanalüüsi põhimõtetest (Hutchby, Wooffitt 1998). Vestlusanalüüsi kohaselt moodustavad mõned dialoogiaktid naabruspaare, kus esiliikme ütlemine teeb relevantseks järelliikme (nt soov ootab täitmist). Ka arvuti kui dialoogis osaleja peaks suutma eristada naabruspaari esiliiget, mis nõuab teatud reaktsiooni, nn üksikaktidest, mis reaktsiooni ei oota (nt vastuvõtuteade).

Dialoogiaktid on meie tüpoloogias jaotatud üheksasse rühma. Aktide koguarv on 126 (89 naabruspaari- ja 37 üksikakti). Dialoogiaktide nimed koosnevad kahest osast, mis on teineteisest eraldatud kooloniga: 1) esimesed kaks tähte näitavad aktirühma nimetust, nt DI – direktiivid, VR – vabatahtlikud reaktsioonid. Kolmas täht esineb ainult naabruspaariaktide nimetustes ja näitab, kas tegu on esi- või järelliikmega (nt DIE, DIJ); 2) akti pärisnimi, nt DIE: SOOV, DIJ: INFO ANDMINE, VR: NEUTRAALNE VASTUVÕTUTEADE jne.

Selles artiklis pöörame põhitähelepanu direktiivsetele naabruspaaridele. Meie aktitüpoloogia kohaselt saavad sellise naabruspaari esiliikmed olla kas soov, ettepanek, pakkumine, palve oodata või “muu”. Järelliikmeteks on info andmine, info puudumine, keeldumine, nõustumine, mittenõustumine, piiratud nõustumine, edasilükkamine ja “muu” (tabel 1).

Tabel 1. Direktiivsete naabruspaaride esi- ja järelliikmed Eesti dialoogiaktide tüpoloogias

Esiliikmed	Järelliikmed
DIE: soov	DIJ: info andmine
DIE: ettepanek	DIJ: info puudumine
DIE: pakkumine	DIJ: keeldumine
DIE: palve oodata	DIJ: nõustumine
DIE: muu	DIJ: mittenõustumine
	DIJ: piiratud nõustumine
	DIJ: edasilükkamine
	DIJ: muu

Soove on kaht liiki: soov võib väljendada autori kavatsust saada infot (infosoov) või käivitada adressaadi tegevus, nt taksodispetšer peab kliendile saatma takso (tegevussoov). Infosoovide võimalikud järelliikmed on info andmine või info puudumine (1)²; H – ametiasutusse helistav klient, V – ametnik).

- (1) H: ma sooviksin `bussijaama telefoni` numbrit. DIE: SOOV
 V: hh `neli` `neli` `seitse` `üks`, `null` `null` `kaks`.
 DIJ: INFO ANDMINE

Ettepanek ja pakkumine (teha teatavat tegevust) ootavad vastuseks kas nõustumist või keeldumist. Pakkumise puhul on tegevuse tegijaks rääkija ise (*ma helistan teile hiljem*), aga ettepaneku puhul kuulaja (*helistage hiljem*).

Selle artikli jaoks valisime dialoogikorpusest 144 telefonikõnet, mis kuuluvad nelja eri situatsioonitüüpi: 1) kõned infotelefonile, 2) reisibüroosse, 3) polikliiniku registratuuri, 4) takso tellimised (tabel 2). Sedasama alamkorpust kasutasime artiklis (Gerassimenko jt 2007), kus käsitlesime kliendi soove. Siin analüüsisime eri situatsioonitüüpides direktiivsete aktide naabruspaare, kus esiliikme autoriks on ametiasutusse helistav klient ja järelliikme autoriks telefonikõnele vastav ametnik (tabel 2). Eri situatsioonitüüpide alamkorpused on küll erineva mahuga, kuid kuna me uurime direktiivide vormi ja naabruspaaride järelliikmeid igas tüübis eraldi, muu hulgas erinevate vormide esinemist ja osakaalu ühe situatsioonitüübi sees, siis on analüüsi tulemusel leitud esinemiste protsendid ja aktide kasutamise üldised seaduspärad siiski võrreldavad. Nagu näha tabelist 2, on direktiivsete naabruspaaride järelliikmeid ligi kaks korda rohkem kui esiliikmeid. Põhjus on enamasti selles, et me märgendame iga lausungi eraldi aktina. Kui näiteks ametnik teatab vastuseks kliendi soovile pikema telefoninumbri mitmes järjestikuses lausungis (numbrikolmikute või -paaridena), siis saabki üks esiliige mitu järelliiget.

Tabel 2. Kasutatud alamkorpused

Dialoogi tüüp	Dialoogide arv	Tekstisõnade arv	Kliendi DIE-de arv	Ametniku DIJ-de arv
Infotelefon	60	4384	58	170
Reisibüroo	36	12104	39	44
Polikliinik	26	2422	23	28
Takso	22	1028	23	19
Kokku	144	19938	143	261

² Näidetes on kasutatud vestlusanalüüsi transkriptsiooni, märkide loend on artikli lõpus.

Meie uurimuse tulemusi saab rakendada kasutajaliidestest, mis tagavad juurdepääsu mitmetele eestikeelsetele andmebaasidele. Muidugi on mõistlik ja reaalne luua selline kasutajaliides andmebaasile, mis sisaldab telefoninumbreid, aadresse jms, mida küsitakse infotelefonile helistades. Nii saab vähendada inimoperaatorite tihtipeale rutiinset tööd. Ka reisiplaneerimissüsteeme on maailmas juba mitmesuguseid, siingi on arvuti kasutamine asjakohane. Samuti võib arvuti asendada taksodispetšerit, kelle ülesandeks on leida vaba takso ja saata see vajalikul aadressil. See eeldab, et arvutis on interaktiivne linnaplaan ja arvuti saab võtta ühendust taksojuhtidega. Küsitavam on polikliiniku töötaja asendamine, sest nagu näitab meie korpuse analüüs, ei ole ametniku ülesandeks siin ainult arstivisiidiks sobiva aja kokkuleppimine patsiendiga, vaid vahel peab ta suutma hinnata ka patsiendi tervislikku seisundit. Siiski näitavad analüüsi tulemused, et eksisteerivad üldised käitumismallid ja keelelised vahendid, mida kasutatakse kõigis vaadeldud situatsioonitüüpides ühtviisi. Alustame oma analüüsi infotelefonikõnedest ning seejärel kõrvutame neid reisibüroo-, registratuuri- ja taksokõnedega, et leida ühelt poolt eri situatsioonitüüpides asetleidvate dialoogide spetsiifilisi iseärasusi ja teiselt poolt nende universaalseid tunnuseid.

Eesti dialoogikorpuse analüüsimise hõlbustamiseks on välja töötatud veebis kasutatav (parooliga kaitstud) tarkvara, mis muuhulgas võimaldab otsida korpusest ja loendada mitmesuguseid dialoogides esinevaid nähtusi: sõnu või sõnajärjendeid, transkriptsioonelemente, dialoogiakte, andes ette akti nime, osaleja tähise, sõne (Treumuth 2004). Joonisel 1 allpool on näide ühe päringu tulemuse kohta.

3. Direktiivsed aktipaarid: korpuse analüüs

3.1. Direktiivsed naabruspaarid

Infotelefonikõnedes leidub kokku 58 direktiivset naabruspaari, mille esiliikme autoriks on klient ja järelliikme autoriks ametnik (tabel 3).

Tabel 3. Direktiivsed naabruspaarid infotelefonikõnedes

Kliendi direktiiv (DIE)	DIE-de arv	Ametniku vastus (DIJ)	DIJ-de arv	Järelliikmeta DIE-de arv
DIE: SOOV	52	DIJ: INFO ANDMINE DIJ: EDASILÜKKAMINE DIJ: INFO PUUDUMINE	129 31 4	3
DIE: SOOV + TVE: PAKKUMINE	2	DIJ: INFO ANDMINE	2	
KYJ: INFO ANDMINE + DIE: SOOV	1	DIJ: INFO ANDMINE	1	
DIJ: NÕUSTUMINE + DIE: SOOV	2	DIJ: INFO ANDMINE	2	
KYJ: JAH + DIE: SOOV	1	DIJ: INFO ANDMINE	1	
Kokku	58		170	3

Kõik selliste naabruspaaride esiliikmed on soovid. Mõnel üksikul soovil on veel teinegi funktsioon: soov kui teemavahetuse algatamine, nõustumine partneri pakkumise või ettepanekuga, mis on samaaegselt infosoov, jne (tabel 3). Järelliikmeteks on info andmine, info puudumine ja edasilükkamine.

Otsitav tekst: palun

Otsitav aktimärgend: DIE: SOOV

Otsitav osalejatunnus: H

Leiti 15 vastet.

- vaata sõnade sagedusi -

osaleja	lausung	aktimärgend	dialog
H:	palun takso `Ravila `neli `neli.	DIE: SOOV	in_di_te_494_a22_takso_valm05.txt
H:	palun `öelge `Hansa: `Reiside: `telefon.	DIE: SOOV	in_di_te_529_a2_info_valm05.txt
H:	=palun `Tallinna `Tõnismäe `hambapoliiklinik.	DIE: SOOV	in_di_te_455_b6_info_valm05.txt
H:	palun mulle `Viljandis `Grand o`telli tele- tavate- tändab `telefoni numbrit.	DIE: SOOV	in_di_te_455_b8_info_valm05.txt

Joonis 1. Dialogikorpuse tööpink: korpusest otsiti kliendi (H) soove, mis sisaldavad sõna *palun*

Reisibüroodialoogides on selliste naabruspaaride esiliikmeid 39 ja järelliikmeid 44 (tabel 4). Esiliikmeteks on soov ja pakkumine. Ühegi naabruspaari esiliige ei ole ettepanek. Kahel juhul on esiliikmeks DIE: MUU (2). Pooled kliendi soovidest jäävad järelliikmeta, samal ajal kui osa neist saab mitu järelliiget.

- (2) H: .hh ee ma `sooviksin natuke informatsiooni selle `Lapimaa reisi kohta. DIE: SOOV
V: jaa? VR: NEUTRAALNE JÄTKAJA
H: [.hh et=õ] DIE: MUU
V: [millise]=mis `kuupäeva kohta. KYE: AVATUD + VTE: VASTUSE TINGIMUSTE TÄPSUSTAMINE

Tabel 4. Direktiivsed naabruspaarid reisibürookõnedes

Kliendi direktiiv (DIE)	DIE-de arv	Ametniku vastus (DIJ)	DIJ-de arv	Järelliikmeta DIE-de arv
DIE: SOOV	28	DIJ: INFO ANDMINE DIJ: EDASILÜKKAMINE DIJ: INFO PUUDUMINE DIJ: KEELDUMINE + KYE: JUTUSTAV KAS + VTE: VASTUSE TINGIMUSTE TÄPSUSTAMINE DIJ: KEELDUMINE + RY: ÜLEANDMINE DIJ: MUU	17 8 3 2 1 1	15
DIE: PAKKUMINE	4	DIJ: NÕUSTUMINE	4	-
DIJ: NÕUSTUMINE + DIE: SOOV	4	DIJ: INFO ANDMINE DIJ: NÕUSTUMINE	3 1	-
DIJ: PIIRATUD NÕUSTUMINE + DIE: PAKKUMINE	1	DIJ: NÕUSTUMINE	1	-
DIE: MUU	2	DIJ: EDASILÜKKAMINE DIJ: INFO PUUDUMINE	1 2	1
Kokku	39		44	16

Viiel juhul on soov või pakkumine samal ajal ka eelmise direktiivse naabuspaari järelliige: nõustumine kui soov, piiratud nõustumine kui pakkumine (3).

- (3) V: .h aga: aga ma saan teile anda `Hermann
 reise telefoninumbri.= DIE: PAKKUMINE
 H: =jaa DIJ: NÕUSTUMINE + DIE: SOOV
 H: see oleks `hea. DIJ: NÕUSTUMINE + DIE: SOOV

Tabelid 5 ja 6 esitavad direktiivseid naabuspaare kahes ülejäänud situatsiooni-tüübis, **polikliiniku- ja taksokõnedes**. Põhierinevus võrreldes infotelefoni ja reisibürooga on see, et siin ootab klient vastuseks oma soovile ametniku tegevust: arsti vastuvõtuoja reserveerimist või takso saatmist. Tegevusega kaasneb alati info andmine (4) või kommentaar selle kohta, et tegevus tehakse edaspidi. Meie aktiivpoloogias märgendatakse selline kommentaar kui DIJ: MUU (5). Vähestel ebatüüpilistel juhtudel on direktiivse naabuspaari esiliige ettepanek, pakkumine või muu.

Tabel 5. Direktiivsed naabuspaarid polikliinikukõnedes

Kliendi direktiiv (DIE)	DIE-de arv	Ametniku vastus (DIJ)	DIJ-de arv	Järelliikmeta DIE-de arv
DIE: SOOV	19	DIJ: INFO ANDMINE DIJ: EDASILÜKKAMINE DIJ: INFO PUUDUMINE DIJ: MUU	11 4 2 6	2
DIE: PAKKUMINE	1	DIJ: NÕUSTUMINE + DIE: SOOV	1	-
DIJ: NÕUSTUMINE + DIE: SOOV	2	DIJ: INFO ANDMINE DIJ: MUU	2 1	-
DIE: MUU	1	DIJ: MUU	1	-
Kokku	23		28	2

Tabel 6. Direktiivsed naabuspaarid taksokõnedes

Kliendi direktiiv (DIE)	DIE-de arv	Ametniku vastus (DIJ)	DIJ-de arv	Järelliikmeta DIE-de arv
DIE: SOOV	20	DIJ: EDASILÜKKAMINE DIJ: MUU DIJ: NÕUSTUMINE	1 17 1	2
DIE: ETTEPANEK	1			1
DIJ: INFO ANDMINE + PPJ: LÄBIVIIMINE + DIE: SOOV	1			1
DIE: MUU	1			1
Kokku	23		19	5

- (4) H: mina sooviksin neuroloogi juurde e aega
 kinni [panna.] DIE: SOOV
 V: isikukood palun=h. <infovahetuse alamdialog>
 /---/
 V: sis läheb neliteist september kell:
 kümme null null. DIJ: INFO ANDMINE

- (5) H: sooviks takso `Anne seitse- seitsekend `kolm. DIE: SOOV
 V: ja `korterit `number? <infovahetuse alamdialoog>
 H: `korterit number on viiskend `kaheksa.
 V: .hh ja `saadan teile `auto. DIJ: MUU

3.2. Kliendi soovid

Vaadeldavas korpuses on direktiivsete naabruspaaride esiliikmeks ülekaalukalt soov (infotelefonis 100%, polikliinikukõnedes ja taksotellimisel 91%, reisibürookõnedes 82% juhtudest).

Tüüpilises kõnes esitab klient ühe soovi kohe kõne algul ja lõpetab kõne, kui soov on täidetud. Samas on kõnesid, kus klient esitab mitu soovi.

Infotelefonikõnedes esinevatest soovidest on kaks kolmandikku esimesed, kohe kõne algul avaldatud soovid. Kliendid küsivad mitmesugust infot: telefoni- numbreid (75% juhtudest), aadresse, asutuste lahtiolekuaegu, ettevõtete tegevus- alasid jms. Tüüpilise infosoovi keeleline muster on järgmine:

*palun (öelge mulle) X,
 öelge (mulle) X,
 tahaksin teada X,*

kus X tähistab lünka kliendi teadmistes, mille ametnik oma infoga peab täitma.

Üks kolmandik soove esitatakse hilisemas vestluses. Siia kuuluvad need juhud, kus ametnik ei saa täita kliendi soovi ja pakub asendusinfot. Selle asemel et piirduda teatamisega, et soovitav info puudub andmebaasist, püüab ametnik olla võimalikult abivalmis ja pakub asendusinfot (6). Pakkumisele eelneb tüüpiliselt infovahetuse alamdialoog. Mõnikord pole ka klient ise võimeline selgelt oma soovi väljendama, soov on kas liiga üldine või ebaselge.

- (6) /---/
 H: = `ei, ma tahaks teada tegevus `alasad. DIE: SOOV
 /---/
 V: aga: ega siin `täpsemalt ei= `olegi kui siin
 on:=ee .hh ee `tootmine: `kanga `riide `põranda
 `pesu `arjade DIJ: INFO ANDMINE
 /---/
 V: aga:] (.) aga:=mhh te võite `numbri=belt
 `täpsemalt `küside=et DIE: ETTEPANEK
 H: {* ahah *}
 `andke mulle `number [siis {-}] DIJ: NÕUSTUMINE + DIE: SOOV

Pärast soovitud info saamist algatab klient vahel alamdialoogi, milles küsib täiendavat infot (7). Kliendid küsivad üle telefoninumbreid, aadresse, nimesid, täpsustavad ettevõtete tegevusalasid jms.

- (7) V: null= `null. DIJ: INFO ANDMINE
 H: null= `null.
 ja `kus ta `asub. <infovahetuse alamdialoog>
 V: .hh `Tähe `sada kuus `teist,

Reisibürookõnedes on kõige sagedasem esiliige soov (82%), veel esineb pakkumine. On tähelepanuväärne, et tihti (15 juhul) alustab helistaja kõnet reisibüroosse üldise sooviga, mis juhatab sisse teema (8), mitte konkreetse sooviga, nagu see on teistes situatsioonitüüpides. (Ülejäänud kõned algavad küsimusega, mida me selles artiklis aga ei käsitle.) Neljal juhul on helistaja nõustumine ametniku pakkumisega ühtlasi soov (3). Tegelikku eesmärki täpsustatakse järgnevas infovahetuse alamdioloogis. Seega jääb helistaja algne soov kui naabruspaari esiliige ilma järellikmeta, mis aga ei tähenda, et helistajal jääks kokkuvõttes vajalik info saamata. Kohe kõne algul esitatakse 72% soovidest.

- (8) H: ma küsiks nende odavate lennuki `piletite kohta. DIE: SOOV
 V: jaa? VR: NEUTRAALNE JÄTKAJA
 H: kas need tuleb osta kohe `välja. KYE: SULETUD KAS
 (1.0)
 V: reeglina `küll. KYJ: JAH

Ametniku tüüpiline vastus kliendi soovile on info andmine. Telefoninumbreid ja veebiaadresse antakse mitme lühikese lausungi kaupa, mille vahele tuleb partneri akt jätkaja, mis signaliseerib kuuldelolekut (9).

- (9) V: .hh see=on kolm null `üks, DIJ: INFO ANDMINE
 H: jaa? VR: NEUTRAALNE JÄTKAJA
 V: `neli neli neli. DIJ: INFO ANDMINE

Mõlema osalise käitumine on kooperatiivne. Jagades info väikesteks portsjoniteks, on ametnik koostööaldis, sest pikki järjendeid oleks kliendil raske meeles pidada. Jätkajad infoandmiste vahepeal annavad märku, et info on edukalt vastu võetud. Mõnikord eelneb info andmisele edasilükkamine (10).

- (10) V: Austria? DIJ: EDASILÜKKAMINE
 ma kirjutan omale üles `autoga? DIJ: EDASILÜKKAMINE
 mhmh VR: NEUTRAALNE VASTUVÕTUTEADE
 .hh eahh noh üsõnaga me müüme
 `Austrias: < `enamvähem `kõiki >
 suusa `kuurorte. DIJ: INFO ANDMINE

Kui soovitud info puudub andmebaasist, siis teatab ametnik sellest kohe. Kooperatiivne ametnik pakub enamasti asendusinfot (teist bürood või teist reisi, näide 11).

- (11) H: .hh olen uvitatud reisidest Skandi `naaviamaadesse=h. DIE: SOOV
 V: e=`jaah? VR: NEUTRAALNE VASTUVÕTUTEADE
 .hh et asi on `selles=et=e vähemalt `praegu: £ me=ei `korralda neid. £
 DIJ: INFO PUUDUMINE
 (0.5)
 H: ahah VR: NEUTRAALNE INFO OSUTAMINE UUEKS
 (.)
 V: £ et=sis=meil=`on `üks reis
 `Lapimaale kui=te=oleksite `huvitatud. £ DIE: PAKKUMINE

Kolmel juhul keeldub reisiagent helistaja soovi täitmast. Kõiki neid keeldumisi pehmendab agendi lubadus täita soov hiljem (12).

(12) H: .hh ee mul oleks vaja sõita aprillis `Budapesti. DIE: SOOV
/---/
V: saab kuskile `tagasi helistada. DIJ: KEELDUMINE +
KYE: JUTUSTAV KAS +
VTE: VASTUSE
TINGIMUSTE TÄPSUSTAMINE
/---/
V: ma `elistan teile. YA: LUBADUS
ma uusi- uurin kohe `järgi noh, YA: LUBADUS

Polikliiniku- ja taksokõnedes on kummaski 91% direktiivsete naabruspaaride esiliikmetest soovid. Seejuures on polikliinikukõnedes 71% ja taksokõnedes 81% soovidest esimesed, s.t avaldatakse kohe kõne algul.

3.3. Kliendi soovide keelelised tunnused

Varasemas artiklis (Gerassimenko jt 2007) analüüsisime kliendi soovide keelelisi tunnuseid – predikaat-argument-struktuuri ja eelkõige verbide kasutust – eesmärgiga töötada välja reeglid soovide automaatseks tuvastamiseks dialoogsüsteemis. Korpuse analüüs näitas, et infotelefonikõnedes kasutatakse helistaja soovides üsna piiratud hulka verbe (kokku 12). Sagedasemad on *paluma*, *ütleva* ja *soovima*.

Teine tunnus soovide tuvastamiseks on verbi positsioon lausungis. Meie infotelefonikõnedes asub verb lausungi algul 38% juhtudest. 28% soovidest algavad asesõnaga *mina/ma*. Hilisemad soovid algavad partikli, sidesõna või määrsõnaga (*et, aga, äkki*) või lõppevad partikliga *siis*. Verb asub lausungis teisel kohal. Veel üks tunnus on, et verbi *võtma* tingivas kõneviisis kasutatakse ainult hilisemates soovides.

Soovid, millel on mitu funktsiooni, on ebatüüpilised, nad ei saa olla esimesed soovid ja nende keeleline vorm on enamasti määratud lisafunktsiooniga (nt ettepanekuga nõustumine, mis on samaaegselt soov).

Kliendi soovide keelelised tunnused on sarnased ka ülejäänud situatsioonitüüpides. Siiski on siin verbide arv väiksem – 10 reisibüroo-, 5 registratuuri- ja 3 taksokõnedes. Kõige sagedasem verb on *soovima* (vastavalt 14, 6 ja 8 juhtu).

3.4. Järelliikmed

Tabelist 3 nähtub, et **infotelefonikõnedes** saab soov kui direktiivse naabruspaari esiliige järelliikmeks kas info andmise (kõige sagedamini), info puudumise või edasilükkamise.

Pärast kliendi soovi esitamist võib kõne jätkuda kahel viisil: 1) ametnik vastab kohe või 2) esitab täpsustava küsimuse, algatades sellega infovahetuse alamdialoogi. Esimene jätk (1) on infotelefoni puhul tüüpiline: ametnik annab info kohe 60% juhtudest. Ülejäänud juhtudel eelneb vastuse andmisele alamdialoog, kus ametnik täpsustab telefoninumbrit, ettevõtte nime ja asukohta vms.

Reisibürookõnedes, nagu varem juba mainitud, jääb paljudel juhtudel direktiivse naabrusepaari esiliige ilma järelliikmeta. Enamasti on tegu helistaja esimese sooviavaldusega kõne algul, mis üksnes juhatab sisse teema (15 juhtu). Järgnevas infovahetuse alamdialoogis täpsustatakse soovi ja vastatakse küsimustele.

14 juhul järgneb järelliige vahetult esiliikmele (neist esiliikmeist 9 on soovid ja 5 ettepanekud). Kui esiliige on soov, siis võib sellele järgneda 1) info puudumisest teatav vastus, 2) edasilükkamine, et sooritada otsinguid andmebaasides, või 3) kliendi kõne suunamine teisele, pädevamale ametnikule. Seega ei järgne soovile ühelgi juhul vahetult info andmist. Kui esiliige on pakkumine, siis järgneb sellele kohe pakkumisega nõustumine (mittenõustumise juhte meie analüüsitud alamkorpuses ei esine).

Ülejäänud juhtudel alustab ametnik pärast soovi küsimusega infovahetuse alamdialoogi, milles täpsustab reisijate arvu, nimesid, vanuseid jms, s.t kogub infot, mis on vajalik soovi täitmiseks. Seejärel saab helistaja soovitud info või teeb ametnik talle ettepaneku helistada tagasi. Alamdialoogi esinemine või mitteesinemine sõltub ametniku tegevusskeemist ja pädevusest – kas ta taipab kohe, millist reisi klient silmas peab, kas ametnik peab lihtsalt sooritama otsingu andmebaasis või on vaja teha rohkem (näiteks küsida kliendi andmeid).

Ülejäänud situatsioonitüüpides järgneb kliendi soovile enamasti infovahetuse alamdialoog (74% juhtudest registratuuridialoogides ja 75% takso tellimisel). Ametnik täpsustab patsiendi andmeid (nimi, isikukood jms), enne kui saab kirja panna arstiaja (4), või kliendi aadressi ja nime, enne kui saadab takso (5).

Ametnik peab aru saama, millist infot või tegevust temalt nõutakse, ja leidma vajaliku info andmebaasist. Seejärel formuleerib ta vastuse. Ametnik peab kasutama korrektset keelt, häälutama sõnu selgelt, tegema pause info andmisel, et klient suudaks sõnumit mõista ja meelde jätta või üles kirjutada. Ametnikud kasutavad info andmisel lauseid või fraase, sõltuvalt olukorra spetsiifikast ja lähtudes suhtlusökonomiast.

3.5. Järelliikmete keeleline vormistus

Info andmist ei väljendata sugugi mitte alati täislausel (võrdluseks: olemasolevad kasutajaliidesed annavad enamasti vastuseid täislausetena). Vastupidi, infotelefonikõnedes kasutatakse tervelt kolmandikul juhtudest fraase, enamasti telefoninumbrite ütlemlisel. Seega annab ametnik infot ökonoomsel viisil. Kindlasti kasutatakse lauset siis, kui info andmisele eelneb infovahetuse alamdialoog. Sellisel juhul on mõistlik vastata täislausel (13), et siduda vastust varem esitatud päringuga. Teine põhjus kasutada fraasi asemel lauset on siis, kui edastatakse mitu infoühikut, näiteks mitu erinevat telefoninumbrit.

- (13) H: paluks `autohall `Tartus. DIE: SOOV
V: autode `müük või `varuosade numbrit. <infovahetuse alamdialoog>
H: õäääh see mis asub {-} `Ilmatsalu tänaval.
/---/
V: `autode müük on neli=kaks=viis DIJ: INFO ANDMINE

Info puudumist väljendatakse tüüpiliselt lausega, ei piirduta fraasiga. Ametnik selgitab, miks ta ei saa anda nõutud infot (14), või pakub asendusinfot. Ametnik kasutab negatiivset lauset (11) või lauset, mis annab asendusinfot.

- (14) V: =sellist numbrit ei näita meil. DIJ: INFO PUUDUMINE
.hh a=ma vaatan võibla teise numbrit lõpuga

Edasilükkamiseks kasutatakse teatud tüüpväljendeid, mida on piiratud hulk, enamasti üks hetk (*palun*) (15). Edasilükkamist vajab ametnik selleks, et otsida soovitud infot andmebaasist. Vahel sisaldab edasilükkamine kommentaari ametniku jooksvale tegevusele (*vaatame, proovin ühendada, kirjutan üles*) või kordab kliendi eelmist vooru (*oli kahekümn=neljandast.*).

- (15) V: ja: mis linnast helistate. <infovahetuse alamdialoog>
H: Tartust
V: ü:ks=hetk. DIJ: EDASILÜKKAMINE
(.) .hh pea postkontor on kella kaheksast
seitsmeni. DIJ: INFO ANDMINE

Nõustumist väljendatakse positiivse vastuspartikliga *jah, hea, hästi* või verbifraasiga käskivas kõneviisis, mis väljendab pakutavat tegevust (*mõelge järgi, teeme nii*).

Keeldumist vormistatakse jutustava lausega, mis suunab teisele ametnikule (*kohe ma suunan teid Moonikale*), või küsilausega, mis on samal ajal täpsustav küsimus (*kas te mulle mingi telefoni jätate, saab kuskile tagasi helistada.*).

Meie analüüs näitab, et eelistatud reaktsioonid (info andmine ja tegevusettepanekuga nõustumine) on tüüpiliselt lühemad kui mitte-eelistatud (info puudumine, tegevusest keeldumine). Telefoninumbrite või veebiaadresside andmisel kasutatakse nimisõnafraaside järjendit (9). Selliseid juhte on infotelefonikõnedes tervelt 90%. Kui info andmiseks kasutatakse lauset, siis on see tavaliselt lihtlause.

3.6. Järelliikmeta soovid

Infotelefoni-, registratuuri- ja taksokõnedes jääb soov järelliikmeta üksnes vähestel juhtudel (tabelid 3, 5, 6). Need on juhud, kus kliendi soov on kas liiga üldine või siis ei saa ametnik seda soovi otsekohe täita ja algatab infovahetuse alamdialoogi, mille kestel selgitab kliendi täpse infovajaduse ja rahuldab selle võimaluse korral. Esialgne soov jääb otsese järelliikmeta. Ühel juhul on kliendi soov nii hägune, et sellele polegi võimalik vastata, ja klient formuleerib uue, konkreetse soovi. Nii aga jääb esimene soov järelliikmeta.

Reisibüroodialoogides on seevastu üsna tüüpiline, et klient esitab kõne algul just väga üldise soovi, määrates sellega teema. Tema tegelik infovajadus selgub alles järgnevas alamdialoogis ja see enamasti rahuldatakse. Algne soov jääb aga järelliikmeta (vt tabel 4).

3.7. Kokkuvõtte korpuse analüüsist

Ametikõnel on kindel struktuur. Ta algab rituaalse sissejuhatusega, kus osalejad tervitavad teineteist ja ametnik esitleb asutust ja võibolla ka ennast. Telefonile vastates kasutab ametnik tüüpiliselt järgmist mustrit (Rääbis 2006: 146–147):

<asutuse nimi> <ametniku nimi> <tervitus>.

Meie analüüsitud kõnedes ütleb ametnik oma nime ainult infotelefonis ja mõnikord reisibürookõnedes. Ta ei tee seda kunagi registratuuri- ega taksokõnedes. Enamasti klient vastab tervitusele (16). Siis järgneb põhiosa, kus klient esitab oma soovi ja saab vastuse. Vahel kasutatakse alamdialoge nii soovi kui ka saadud vastuse täpsustamiseks (13). Dialog lõpeb rituaalse osaga, kus klient tänab ja ametnik vastab tänuavaldusele. Üks kolmandik kõnedest lõpeb kliendi hüvastijätuga, millele aga alati ei tarvitse järgneda ametniku vastuhüvastijätt (17).

(16) V: Kuku `kliinik <asutuse nimi>
V: tere <tervitus>
H: tere

(17) H: aitüma. <tänamine>
V: palun.
H: nägemist <hüvastijätmine>
V: nägemist

Koostööaldis ametnik annab asendusinfot, kui ta ei saa täita kliendi soovi. Info andmisel eelistatakse lühikesi lauseid. Tehakse pause, et partner saaks vajadusel infot üles kirjutada. Partneri tagasiside on signaaliks, et infoandmine võib pärast pausi jätkuda.

Inimesed kasutavad piiratud hulka tüüpväljendeid tervitamiseks, hüvastijätuks, vastuse edasilükkamiseks ja mõneks muuks dialoogi juhtimisega seotud tegevuseks.

Arendades intelligentset kasutajaliidest, püüame me modelleerida “ideaalset” ametnikku. Tegelikuses on ametnikud vahel klientidega kannatamatud, näiteks registratuuriametnik juhul, kui kliendile ei sobi ükski pakutud aeg. Sellist käitumist me ei jäljenda. Kehtivad üldised käitumismallid ja keelekasutus, mis iseloomustavad head infoagenti.

Dialoogiaktide keeleliste tunnuste väljaselgitamine võimaldab koostada reegleid, mille alusel saab kasutajaliides tuvastada kasutajate dialoogiakte, samuti vormistada oma dialoogiakte keeleliselt sellisel kujul, nagu seda teevad ametnikud tavalises inimestevahelises suhtlemises.

4. Intelligentse kasutajaliidese arendamine

4.1.Eestikeelne infodialoog arvutiga

Loodud on kaks veebis kasutatavat lihtsat kasutajaliidest (Reisiagent ja Teatriagent), mis vahendavad eesti keeles juurdepääsu kahele andmebaasile (Treamuth 2004, Treamuth jt 2006). Üks on Tallinna lennujaamast väljuvate lennukite väljumisaeegade andmebaas ja teine Eesti teatrite mängukavade andmebaas. Arendamisel on kasutajaliides, mis annab infot Tartu kinokavade kohta (Treamuth 2008).

Kasutaja sisestab oma teksti päringuaknasse ja saab vastuse kas tekstina või tehiskõnes. Dialoogihaldurina kasutatakse lõplikku automaati, mille aluseks on dialoogiaktide naabruspaarid (nt tervitus – vastutervitus, soov – info andmine

jms). Kasutaja dialoogiaktide tuvastamine põhineb võtmesõnadel ja/või -fraasidel, piiratud ainevaldkonnas asendab see semantilist analüüsi praktiliseks rakendamiseks piisava tulemuslikkusega. Süsteemi on loimitud eesti keele morfoloogiline analüsaator, mis taandab kasutaja lauses esinevad sõnavormid algvormidele, et otsida nende hulgast võtmesõnu.

Arvuti vastused kasutaja rituaalsetele dialoogiaktidele (tervitus, tänamine, hüvastijätt) valitakse loenditest, mis on moodustatud korpuse analüüsi alusel. Info andmisel kasutatakse etteantud lausemalle, nt

<linna_nimi_ains_sisseütl> **saab lennata** <nädalapäev_ains_alalütl>
kell <kellaeg>.

Etendust <etenduse_nimi> **mängitakse** <kuupäev> **kell** <kellaeg>.

Kasutaja lausungile ainuüksi võtmesõnade põhjal reageerimine võib põhjustada vigu, nt järgnevas katkendis ei tuvastanud arvuti kasutaja hüvastijätku, sest ootas etenduse nime.

(18) Kasutaja: Head aega.

Teatriagent: Hetkel ei toimu etendust "Head ööd, vend". Võimalik, et kava on uuendamata.

Vältimaks mittemõistmisi ja tõstmaks tõrkekindlust, tuleb intelligentes kasutajaliideses kaugemas perspektiivis üle minna lausungite täielikule semantilisele ja pragmaatilisele analüüsile, s.t arvuti peaks tuvastama kasutaja dialoogiakti ja mõistma kasutaja eesmärgi, et adekvaatselt reageerida. Seepärast anname järgnevas lühikese ülevaate sellest, mida on seni tehtud dialoogiaktide automaatse tuvastamise alal.

4.2. Dialoogiaktide automaatne tuvastamine

Dialoogiaktide automaatsel tuvastamisel on kaks võimalikku rakendust: esiteks, intelligentes kasutajaliideses, mis peab tuvastama kasutajate lausungeid (esmajoones on tähtis edukalt tuvastada olulisemaid infoakte – direktiive ja küsimusi), ja teiseks, dialoogiaktide automaatsel märgendamisel uutes, korpusesse lisanduvates dialoogisalvestustes.

Eestikeelsetes dialoogides dialoogiaktide tuvastamiseks ei ole seni veel katsetatud reegleid, mille formuleerimiseni võib viia eelmises peatükis läbiviidud analüüs. Katsetusi on tehtud mõningate statistiliste meetoditega.

Esmalt prooviti tehisnärvivõrke (ühe ja kahe peitkihiga tajureid ning rekurrentseid võrke) ja otsustuspuuid. Tehisnärvivõrgud on nn musta kasti tüüpi meetod, sest treenitud võrgu parameetrid ei ole inimese jaoks otseselt analüüsitavad. Otsustuspuud on aga nn valge kasti tüüpi meetod, nende eeliseks on inimesele hästi arusaadav struktuur. Puu igas tipus kontrollitakse keelelisi tunnuseid ja servadele seatakse vastavusse otsused lausungis tunnuste esinemise või mitteesinemise kohta. Samas on otsustuspuuid suhteliselt keeruline konstrueerida. Tulemused ei olnud kummagi meetodiga paraku kuigi head. Tehisnärvivõrkudes oli dialoogiaktide keskmine tuvastamistäpsus alla 20%. Dialoogiaktide klasse tuvastati küll mõnevõrra paremini, nt edukaimaks osutunud laiendatud rekurrentsed võrgud tuvastasid DIE täpsusega 27,4% ja DIJ täpsusega 48,1%. Otsustuspuude keskmine tuvastamis-

täpsus oli pisut parem, 44,7% (seejuures direktiivsete aktide puhul 43,5%; katsetati veel küsimuste, vastuste ja rituaalsete aktide tuvastamist). (Fišel, Kikas 2006)

Üks otsustuspuudega läbiviidud eksperimentide käigus tehtud tähelepanek oli, et paljud tunnused esinevad lausungites pikemate alamsekventsidenä (kus olulised sõnavormid ei paikne tingimata kõrvuti, vaid nende vahel võib asetseda teisi sõnavorme). Jätkates statistiliste meetodite otsimist, soovisime leida meetodi, mis kasutaks dialoogiaktide tuvastamiseks lausungites esinevaid sõnavormide lüünklikke järjendeid (kus "lünk" võib olla täidetud suvaliste sõnavormidega).

Ühe võimaluse pakuvad sufiksipuud. Sufiksipuu on puukujuline andmestruktuur, mis esitab sõnade (meie juhul lausungite) kõikvõimalikke sufikseid³ kompaktsel viisil. Sufiksipuud võimaldavad leida sarnasusi etteantud sõne ja treeningkorpuse näidete vahel. Tõenäosuslikkus sufiksipuu (Bejerano, Yona 1999), mida meie oma katsetusteks valisime, on servade ja tippudega seotud alamsõnade esinemise ja klassidesse kuulumise tõenäosused. Dialoogiaktide tuvastamisel on klassideks erinevad dialoogiaktid ja iga klassi elementideks tekstiüksuste järjendid (lausungite osad), millega üht või teist dialoogiakti väljendatakse. Näiteks on klassideks DIE: SOOV, DIJ: INFO ANDMINE jne ning nende elementideks vastavalt kõik soovides leiduvad alamsõned, kõik infoandmistes leiduvad alamsõned jne. Puu servade juures on antud üleminekutõenäosused ja lehtede juures väljunditõenäosuste järjendid, kus esimene arv on klassi DIE: SOOV kuulumise tõenäosus, teine arv klassi DIJ: INFO ANDMINE kuulumise tõenäosus jne. Teksti alusel koostatakse sufiksipuu, millest toimub seejärel meid huvitava sõne otsing.

Tõenäosuslike sufiksipuude koostamiseks ja treenimiseks valiti eesti dialoogikorpusest 400 infotelefonikõnet. Treenimine ja testimine toimus 10-kordse ristkontrolli teel. Viidi läbi kaks katsete seeriat. Esimeses arvestati sufiksipuu koostamisel lausungis esinevaid (võimalike lünkadega) alamsekventse, teises lisaks ka eelnevat ja järgnevat dialoogiakti (Kikas 2007). Samuti võeti arvesse kõnelejatunnuseid (klient või ametnik). Dialoogiaktide tuvastamine alamsekventsidi põhjal andis keskmiseks saagiseks ja täpsuseks 46,75%. Eelneva ja järgneva dialoogiakti arvestamine tõstis saagise ja täpsuse 56,49%-ni.

Mõne dialoogiakti tuvastamine oli üsna edukas ka ainult alamsekventsidi kasutamisel, näiteks RIE: TERVITUS (saagis ja täpsus vastavalt 83,24% ja 75,97%), RIE: TÄNAN (95,54 ja 91,46). See on arusaadav, sest nii tervitamisel kui ka tänamisel kasutavad inimesed teatud tüüpväljendeid, mida meetod suutis automaatselt leida.

Naaberaktide arvessevõtmisel tulemused üldiselt paranesid, näiteks akti RIE: TERVITUS tuvastamisel tõusis saagis 93,35%-ni ja täpsus 86,88%-ni. Eelneva ja järgneva akti kasutamine võimaldas paremini eristada akte RIE: TERVITUS ja RIJ: TERVITUS, mille väljendamise keelelised vahendid on ühesugused.

Samas leidis dialoogiakte, mida tuvastati mõlemas katses halvasti, näiteks alternatiivküsimuse tuvastamisel oli esimeses katses saagis vaid 9,41% ja täpsus 19,05%, teises katses ei paranenud tulemused oluliselt (vastavalt 18,82% ja 32,00%).

Elmises peatükis käsitletud aktide soov, info andmine, info puudumine ja edasilükkamine tuvastamise tulemused (saagis ja täpsus) on esitatud tabelis 7 (vt ka Kikas 2007). Tabel 7 näitab, et teises katses on halvenenud akti DIJ: INFO PUUDUMINE tuvastamise saagis ja akti DIJ: EDASILÜKKAMINE tuvastamise täpsus. See on arvatavasti tingitud sellest, et neile aktidele võib eelneva ja järgneva mitmesuguseid erinevaid akte, mille arvessevõtmine lisab sufiksipuusse üleliigseid

³ (Lausungi) sufiksi all mõistame siin tekstiüksuste (sõnavormid, pausid, üneemid) järjendit, millega lausung saab lõppeda.

tunnuseid. Seetõttu ei suuda meetod leida kõiki info puudumise akte (saagis väheneb) või peab mõningaid akte ekslikult edasilükkamise aktideks (täpsus väheneb). Pealegi oli mõlemat akti treeningkorpuses suhteliselt napilt.

Tabel 7. Direktiivsete naabruspaariaktide tuvastamine tõenäosuslike sufiksipuudega: saagis ja täpsus

Dialoogiakt	Arv korpuses	Alamsekventsodega		Alamsekventsude ja kõrvuti asetsevate dialoogiaktidega	
		saagis	täpsus	saagis	täpsus
DIE: SOOV	332	41,57	41,19	62,65	51,61
DIJ: INFO ANDMINE	609	64,37	41,70	67,82	48,42
DIJ: INFO PUUDUMINE	40	22,50	36,00	20,00	44,44
DIJ: EDASILÜKKAMINE	131	67,94	56,33	71,76	43,93

Sufiksipuud on sarnaselt otsustuspuudele valge kasti meetod ja nende üks eelis on see, et nad leiavad lausungites keelelised mustrid⁴, mille alusel toimub dialoogiakti määramine. Analüüsides mustreid, saab lingvist kontrollida oma hüpoteese, pakkuda arvuti leitud mustritele tõlgendusi ja soovitada uusi tunnuseid, mille arvessevõtmine võimaldaks tuvastamist parandada, samuti koostada reegleid.

Tabel 8. Direktiivsete naabruspaariaktide tuvastamine tõenäosuslike sufiksipuudega: sagedasemad mustrid

Dialoogiakt	Mustrid	
	Alamsekventsodega	Alamsekventsude ja kõrvuti asetsevate dialoogiaktidega
DIE: SOOV	<i>ma</i> <i>ma paluksin</i> <i>ma sooviksin</i> <i>ma palun</i>	DIJ: VASTUTERVITUS <i>.hh</i> DIJ: VASTUTERVITUS <i>ma palun</i>
DIJ: INFO ANDMINE	<i>(.) .hh</i> <i>(.)</i>	KYJ: JAH * VR: NEUTRAALNE JÄTKAJA
DIJ: INFO PUUDUMINE	<i>* ei ole antud</i> <i>* ei näita</i>	<i>* ei ole antud</i>
DIJ: EDASILÜKKAMINE	<i>üks hetk</i> <i>üks *</i>	VR: NEUTRAALNE VASTUVÖTUTEADE <i>üks hetk</i> DIJ: INFO ANDMINE * DIJ: INFO ANDMINE

Dialoogiakti DIE: SOOV tuvastamisel leidsid sufiksipuud sellised alamsekventsud, nagu *ma*, *öelge palun*, *ma paluksin*, *ma sooviksin*, *ma palun* lausungi algul. Seejuures algas lausung ligi kolmandikul juhtudest isikulise asesõnaga *ma*. Kõrvuti asuvate aktide arvessevõtmisel leiti lisaks, et kõige sagedamini järgneb DIE: SOOV kõneleja enda (kliendi) vastutervitusele ja lausung algab sissehingamisega (*.hh*) või sõnadega *ma palun*.

Akti DIJ: INFO ANDMINE tuvastamiseks leitud mustrid sisaldavad mitmesuguseid numbrikombinatsioone, mis on tingitud sellest, et infoks, mida ametnik annab, on paljudel juhtudel telefoninumbrid. Kõige sagedamini kasutati üldist mustrit (*.) .hh **, mis tähendab, et info andmine algas pikema pausi ja seejärel sissehingamisega. Kõrvuti asetsevatest aktidest oli sagedasim kombinatsioon KYJ: JAH ja VR: NEUTRAALNE JÄTKAJA, mis tähendab, et info andmine järgnes kliendi

⁴ Mustrites tähistab * suvalist arvu (0 või enam) suvalisi tekstiüksusi. Sufiksipuude leitud mustrid ei koosne ainuüksi lausungite sufiksistest, vaid võivad olla suvalised lausungites esinevate tekstiüksuste järjendid.

üleküsimisele ja selle kinnitamisele ametniku poolt ning info andmisele järgnes omakorda kliendi tagasiside. Sageli järgnes info andmisele ka kliendi vastust pakkuv küsimus (üleküsimine).

Akti DIJ: INFO PUUDUMINE tuvastamiseks väga selgeid mustreid ei avastatud (mis on ka arusaadav, sest selliseid akte oli treeningkorpuses ainult 40). Leitud mustrid sisaldavad sõnu *ei ole antud, ei näita, ei ole meil, meie ei*. Lausungis esineb (eitav) predikaat, s.t tegu on lausega, mitte fraasiga. Kõrvuti asuvate aktide arvestamine ei avastanud kuigi palju uusi mustreid, üksnes kolmes mustris eelneb info puudumisele kliendi KYJ: JAH (mis näitab, et eitava vastuse andmisele on eelnenud infovahetuse alamdialog).

Akti DIJ: EDASILÜKKAMINE tuvastamisel domineeris ülekaalukalt muster *üks hetk*. Aktidest eelnes sagedamini ametniku VR: NEUTRAALNE VASTUVÕTUTEADE ja järgnes ametniku DIJ: INFO ANDMINE (19).

(19) V: jaa?	VR: NEUTRAALNE VASTUVÕTUTEADE
V: üks= `hetk	DIJ: EDASILÜKKAMINE
V: .hh `üld` info, kinnitamata `andmetel	
`seitse `kolm `neli,	DIJ: INFO ANDMINE

Katsetes piirduti suhteliselt lihtsakujuliste sufiksipuudega – ainsaks kasutatud metamärgiks oli *, mis võimaldas lisaks alamsõnedele esitada ka alamsekventse. Üheks edasiarenduse võimaluseks on laiendada sufiksipuid veel teiste metamärkidega. Erinevalt senisest tuleb oluliselt enam rõhku panna optimaalse mustri leidmisele. Siiani kasutatud meetodi korral tuvastatakse dialoogiakt pikima sufiksipuule vastava mustri põhjal. Juhul kui selliseid mustreid on mitu, valitakse neist suurima tõenäosusega muster. Sellisel lähenemisviisil on ilmseid puudusi: eelistuse saavad pikad loetelud, mis üldiselt ei viita akti funktsioonile; vähetõenäolist mustrit eelistatakse ühe sümboli võrra lühemale tõenäolisemale mustrile; arvestamata jäetakse juhud, kus aktile viitab korruga rohkem kui üks muster.

Siiski on saadud tulemused võrreldavad muude, hästi tuntud ja teiste keelte puhul kasutatavate tuvastusmeetoditega. Antud meetodi eeliseks muude meetodite ees on selle hea arusaadavus, kuna dialoogiaktid tuvastatakse treeningkorpusest leitud alamsõnede ja -sekventsides põhjal. Teadaolevalt ei ole sufiksipuid varem dialoogiaktide tuvastamiseks ühegi keele puhul kasutatud.

4.3. Arutelu

Eesti dialoogiaktide tüpoloogia väljatöötamise esmane eesmärk on olnud inimestevahelise suulise suhtluse uurimine. Seetõttu kuulub sellesse väga palju erinevaid dialoogiakte (kokku 126), millest igaüks on asjakohane ühe või teise suhtlusnüansi kirjeldamisel, mis võivad aset leida inimestevahelises suhtluses, kuid mitte inimese ja arvuti vahelises suhtluses. Näiteks on tüpoloogias üksnes rituaalseid naabruspaariakte kokku tervelt 30, sh sellised aktid nagu RIE: SOOV RÄÄKIDA, RIJ: VASTUVÕTMINE, RIJ: TAGASILÜKKAMINE, RIJ: EDASISUUNAMINE, RIJ: MITTEEDASISUUNAMINE jpt, mis vaevalt kunagi esinevad inimese ja arvuti dialoogis. Ka eesti dialoogikorpusesse kogutud inimestevahelistes dialoogides leidub neid akte harva.

Seejuures on mõnede dialoogiaktide määramine isegi kogemustega lingvistidele keeruline, näiteks suletud ja jutustava kas-küsimuse eristamine. Kapa-koeffitsiendi⁵ väärtus, mida tavaliselt kasutatakse märgendajatevahelise üksmeele mõõduna, on eesti dialoogikorpuse märgendamisel olnud parimal juhul 0,8, enamasti aga madalam.

Seetõttu on dialoogiaktide automaatne tuvastamine vastavalt eesti dialoogiaktide tüpoloogiale tõsine väljakutse. Ülesande teeb veel raskemaks asjaolu, et eesti dialoogikorpuse arendamisel on esikohale seatud inimestevahelise suhtluse uurimine, mistõttu on korpusesse kogutud võimalikult eriliigilist suhtlust: erinevaid telefonikõnesid (infotelefon, reisibüroo, bussijaam, polikliiniku registratuur, kauplused, taksodispetšer jm) ja silmast silma vestlusi (kaubandus, teenindus, reisibüroo, teejuhatamine jm). Muidugi peab arvesse võtma ka asjaolu, et inimestevahelise suhtluse salvestamisel tuleb kinni pidada mitmesugustest andmekaitsemeetoditest piirangutest, mis teeb materjali kogumise raskeks.

Siiski saavutati dialoogiaktide tuvastamisel tõenäosuslike sufiksipuudega tulemusi, mis pakuvad praktilise rakendamise seisukohalt huvi. Lisaks annavad tõenäosuslikud sufiksipuud huvitava sissevaate dialoogiaktide moodustamisel kasutatavatesse keelelistesse vahenditesse ja eri dialoogiaktide paiknemisse ametlikes eestikeelsetes telefonikõnedes. Võrreldes tehisnärvivõrkude ja otsustuspuudega saavutati paremad tulemused, mis aga võib osaliselt olla tingitud ka sellest, et homogeenset treeningmaterjali oli rohkem, sest märgendatud korpuse maht oli vahepeal kasvanud.

5. Kokkuvõte ja edasine töö

Artiklis analüüsisime telefonikõnesid ametiasutustesse eesmärgiga arendada intelligentset kasutajaliidest. Valisime eesti dialoogikorpusest 144 dialoogi, kus olid esindatud neli situatsioonitüüpi (infotelefon, reisibüroo, polikliiniku registratuur, taksotellimine). Ametikõne koosneb kolmest osast: 1) rituaalne algus, kus osalejad tervitavad ja tutvustavad ennast, 2) põhiosa, kus klient esitab soovi ja saab vastuse, 3) rituaalne lõpp, kus helistaja tänab ja jätab hüvasti. Helistaja esimene soov püsib eesmärgi, mis saavutatakse koostöös ametnikuga. Me uurisime direktiivseid naabruspaare, mille esiliikme produtseerib helistaja ja järelliikme ametnik. Esiliige seab ootuse kindla järelliikme suhtes. Uuritud naabruspaaride esiliikmed on soov, ettepanek ja pakkumine ning järelliikmed info andmine, info puudumine, nõustumine ja keeldumine. Kõige sagedasemad paarid on soov – info andmine ning pakkumine – nõustumine. Reisibürookõnes on tüüpiline, et klient alustab üldise sooviga, mida hakatakse täpsustama järgnevas alamdialoogis. Kui soovitud info puudub, siis järgneb järelliige esiliikmele kohe.

Kliendid kasutavad kindlaid keelelisi vahendeid, formuleerimaks oma soove (kindlad verbivormid, teatud sõnade järjekord). Ametnik annab infot ökonoomsel viisil, kasutades lühikesi lauseid ja fraase, kui see ei takista arusaamist. Inimesed kasutavad tüüpilisi vahendeid tervitamiseks, tänamiseks jms.

Loodud on kaks veebis kasutatavat lihtsat küsimus-vastussüsteemi, mis suhtlevad inimesega eesti keeles. Üks neist annab infot lennukite väljumisaegade kohta Tallinna lennujaamast ja teine Eesti teatrite mängukavade kohta. Kasutaja sisestab

⁵ Kapa-koeffitsient arvutatakse valemist $k = \frac{P(A) - P(E)}{1 - P(E)}$, kus P(E) on võimalike nõustumiste osakaal ja P(A) on tegelike nõustumiste osakaal.

oma küsimuse päringuaknasse eesti keeles ja saab vastuse kas tekstina või sünteeskõnes. Dialoogihaldurina kasutatakse lõplikku automaati. Kasutaja dialoogiaktide tuvastamine põhineb neis esinevatel võtmesõnadel või -fraasidel, süsteemi on loimitud eesti keele morfoloogiline analüsaator ja tekst-kõnesüntesaator. Arendamisel on intelligentne kasutajaliides, mis vahendab infot Tartu kinokavade kohta.

Edasine töö kulgeb kahes suunas. Esiteks on kavas põhjalikumalt analüüsida sufiksipuude leitud tuvastusmustreid ja võrrelda neid korpuse analüüsi tulemustega, et võtta seda arvesse kasutajaliideses nii kasutaja päringute tuvastamisel kui ka arvuti vastuste ülesehitamisel. Teiseks tuleb põhjalikumalt uurida ametisuhtluse ülesehitust, eriti alamdialoogide esinemise põhjusi ja tingimusi, eesmärgiga muuta arendatav liides võimalikult kasutajasõbralikuks, aga samas ka tõrkekindlaks, eriti kui arvesse võtta, et kõnetuvastuse lisamisel on veatekke võimalusi rohkem.

Dialoogikorpuse lühendid

DIE	direktiivi esiliige
DIJ	direktiivi järelliige
KYE	küsimuse esiliige
KYJ	küsimuse järelliige
PPJ	partneri algatud paranduse järelliige
RIE	rituaali esiliige
RIJ	rituaali järelliige
RY	rituaalne üksikakt
TVE	teemavahetuse esiliige
VR	vabatahtlik reaktsioon
VTE	vastuse tingimuste täpsustamise esiliige
YA	üksikakt

Osalejad

H	helistaja, klient
V	vastaja, ametnik

Transkriptsioonimärgid

.	langev intonatsioon
,	poollangev intonatsioon
?	tõusev intonatsioon
(.)	mikropaus (0.2 sek või lühem)
(0.5)	pausi pikkus sekundites
`sõna	rõhutamine
< >	aeglustatud lõik
* *	vaiksem lõik
e:i	hääliku venitus
.hh	sissehingamine
hh	väljahingamine
si-	pooleli jäänud sõna
[]	pealerääkimine
=	kahe iseseisva üksuse kokkuhääldamine
{sõna}	ebaselgelt kuulnud lõik
{-}	ebaselgeks jäänud sõna
/---/	transkriptsioonist välja jäetud read
£	aktsent

Viidatud kirjandus

- Bejerano, Gill; Yona, Golan 1999. Modeling protein families using probabilistic suffix trees. – Proceedings of RECOMB 1999. ACM Press, 15–24.
- Fišel, Mark; Kikas, Taavet 2006. Dialoogiaktide automaatne tuvastamine. – Mare Koit, Renate Pajusalu, Haldur Õim (toim.). Keel ja arvuti. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised, 6. Tartu: Tartu Ülikooli Kirjastus, 233–245.
- Gerassimenko, Olga; Hennoste, Tiit; Kasterpalu, Riina; Koit, Mare; Rääbis, Andriela; Strandson, Krista; Valdisoo, Maret; Vutt, Evely 2007. Kliendi soovide automaatne tuvastamine eestikeelsetes infodialoogides. – Eesti Rakenduslingvistika Ühingu aastaraamat, 3, 135–154.
- Hennoste, Tiit; Rääbis, Andriela 2004. Dialoogiaktid eesti infodialoogides: tüpologia ja analüüs. Tartu: Tartu Ülikooli Kirjastus.
- Hutchby, Ian; Wooffitt, Robin 1998. Conversation Analysis. Principles, Practices and Applications. Cambridge: Polity Press.

- Kikas, Taavet 2007. Dialoogiaktide tuvastamine eestikeelsetes dialoogides sufiksipuude abil. Magistritöö. Tartu Ülikool, arvutiteaduse instituut. <http://dspace.utlib.ee/dspace/handle/10062/2755> (8.09.2009).
- Koit, Mare 2003. Märgendatud dialoogikorpus kui keeleressurs. – Margit Langemets, Heete Sahkai, Maria-Maren Sepper (toim.). Toimiv keel I. Töid rakenduslingvistika alalt. Eesti Keele Instituudi toimetised, 12. Tallinn: Eesti Keele Sihtasutus, 119–136.
- McTear, Michael F. 2004. Spoken Dialogue Technology: Toward the Conversational User Interface. London: Springer Verlag.
- Minker, Wolfgang; Bennacef, Samir 2004. Speech and Human-Machine Dialog. The Springer International Series in Engineering and Computer Science. Boston, Dordrecht, London: Kluwer Academic Publishers.
- Rääbis, Andriela 2006. Infodialoogi algusrituaalid. – Mare Koit, Renate Pajusalu, Haldur Õim (toim.). Keel ja arvuti. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised, 6. Tartu: Tartu Ülikooli Kirjastus, 143–155.
- Treumuth, Margus 2004. Eesti dialoogikorpus ja selle töötlemise tarkvara. Magistritöö. Tartu Ülikool, arvutiteaduse instituut. <http://dspace.utlib.ee/dspace/handle/10062/1172> (8.09.2009).
- Treumuth, Margus 2008. Kübeke Tartu kultuuri. <http://www.dialoogid.ee/teatriagent/> (8.09.2009).
- Treumuth, Margus; Alumäe, Tanel; Meister, Einar 2006. A natural language interface to a theater information database. – T. Erjavec, J. Žganec Gros (Eds.). Language Technologies, IS-LTC 2006: Proceedings of 5th Slovenian and 1st International Conference, 9–10 October, Ljubljana, Slovenia. Ljubljana, 27–30.

Võrgumaterjalid

Eesti dialoogikorpus EdiK. <http://math.ut.ee/~koit/Dialoog/EDiC.html> (9.02.2010).

Olga Gerassimenko (Tartu Ülikool) on uurinud tagasisidevahendeid eesti ja vene suulises suhtluses. olga.gerassimenko@ut.ee

Riina Kasterpalu (Tartu Ülikool) uurimisvaldkonnad on suuline suhtlus, dialoogi struktuur. riina.kasterpalu@ut.ee

Mare Koit (Tartu Ülikool) on uurinud dialoogi modelleerimist arvutil. mare.koit@ut.ee

Andriela Rääbise (Tartu Ülikool) uurimisvaldkonnad on suuline kõne, telefonisuhtlus, infodialoogide struktuur. andriela.raabis@ut.ee

Krista Strandson (Tartu Ülikool) on uurinud parandusi suulises eesti keeles. krista.strandson@ut.ee

DIRECTIVE ACT PAIRS IN ESTONIAN INFORMATION DIALOGUES AND THEIR AUTOMATIC RECOGNITION

**Olga Gerassimenko, Riina Kasterpalu,
Mare Koit, Andriela Rääbis, Krista Strandson**

University of Tartu

This article studies Estonian human–human phone calls with the further aim of developing a human–computer dialogue system that interacts with the users in the natural language. The analysis is based on the Estonian Dialogue Corpus. Four situation types are considered: directory inquiries, calls to travel agencies, outpatients' offices and taxi services. The most frequent directive act pairs in the dialogues are request – giving information and proposal – agreement. A client's initial request sets up a goal (to get some information or to trigger the agent's action, e.g. sending a taxi to the desired destination) which will be achieved in collaboration with the agent. The study analyzes the linguistic features of the clients' directives and agents' grants. In the dialogues, the officials give the information briefly by using short sentences and phrases. The information-sharing sub-dialogues are initiated by both participants depending on whether a request or a grant needs to be adjusted. The experiments on automatic recognition of dialogue acts in the Estonian Dialogue Corpus are described. Two different probabilistic suffix tree models have been implemented by Taavet Kikas. The first one uses the substrings of the utterances in order to classify the utterances into several dialogue acts. The second model uses the substrings together with the discourse information in the form of preceding and following dialogue act. In the experiments for directive adjacency pairs, the recall of 20%–71% and the precision 36%–56% have been achieved. The most frequent acts in the used training corpus – request and giving information – were recognized with the best recall. In conclusion, the probabilistic suffix trees turn out to be a promising data structure for dialogue act recognition.

Keywords: human–computer interaction, dialogue act, request, grant, Estonian

“I PUTS IT AWAY” – EARLY PROTO-MORPHOLOGICAL WAYS OF INFLECTING VERBS IN A CHILD ACQUIRING SAAMI

Johanna Ijäs

Abstract. In this article, some protomorphological ways of inflect verbs in the acquisition of Saami at age 1;8–3;0 are discussed. The main focus is the forms for first person singular present indicative. The child uses the adult-like forms during all the period since the emergence of first person singular at age 1;9–1;10 but in addition to the adult-like forms, there are three main ways to inflect verbs in the child’s language. Using the stem without the *-n*-suffix is a typical way of inflection at the age of 2;0–2;8. Almost at the same time, at 2;1–2;9, the forms for third person singular present indicative are used instead of the adult-like forms for first person singular. The third difference compared to adult language is the strong gradation of consonant centres in bisyllabic verbs.*

Keywords: bilingual first language acquisition, morphology, verb inflection, North Saami, Finnish

1. Introduction

By the time a child reaches the age of three, (s)he will inflect the most common verb forms in the same way as adult speakers, but how does a child acquire this competence? Acquiring verb inflections, like all morphological acquisition, is regarded as being a progressive and cumulative process, which proceeds from the premorphological stage through the protomorphological stage to modular morphology (Kilani-Schoch, Dressler 2002: 45). Initially, a child acquires new word forms as unanalysed items, almost like new lexemes. This early reproductive stage, when the child cannot yet understand the meaning of different elements of word forms, is called the premorphological stage. In this stage, the child uses normally one inflection of each lexeme (Dressler, Karpf 1995: 102, Laalo 1997: 187, 1999:

* The article is based on my research project on the acquisition of verb inflections in North Saami. The analysis of the data is in progress, so the findings presented in this article should not be considered final. I would like to thank Pekka Sammallahti, Karen Inga Eira and Outi Guttorm for their valuable comments on the earlier version of this article. For the final version of the article I naturally take the full responsibility. Parts of this article are based on my Saami-language paper (see Ijäs 2009). This article is translated from Finnish into English by AAC Global Oy.

354, 362). For example, a Saami-speaking¹ child may use the verb *addit* ‘to give’ in its imperative form only, *atte* ‘give’, and of the verb *boahtit* ‘to come’ only the third person singular, *boahtá* ‘comes’.

As the child’s vocabulary gradually accumulates, (s)he will need several morphological forms in order to express the functions of various word forms. (S)he will begin to use several inflections of each lexeme and start forming individual rules for inflecting words. This stage, when the child has reached an active period in morphological acquisition, is called the protomorphological stage. At this productive stage, rather than imitating word forms as inseparable items, (s)he will start forming her/his own morphological rules. At the protomorphological stage, the child will test the same rules in various contexts, resulting in word forms that deviate from those used by adults. (Dressler, Karpf 1995: 100, Laalo 1998: 362, Bittner et al. 2003: xviii–xix) A child may, for example, use one inflection pattern for words that in fact belong to different inflectional classes. A Finnish-speaking child might form the preterit form of the word *sataa* ‘to rain’ in a similar way as the preterit of the word *heittää* ‘to throw’: *sat-i* (AL²: *sato-i*), cf. *heitt-i* ‘(s)he/it threw’ (see Laalo 2003: 341–345). With Saami-speaking children, the merger of inflectional classes has been observed to take place in, for example, the following cases, in which bisyllabic *i*-stemmed verbs have been inflected according to the inflectional model of *a*-stemmed verbs: *jáma-i* [die-PST.3SG] ‘(s)he/it died’, *vácça-i* [walk-PST.3SG] ‘(s)he/it walked’ (AL: *jámi-i*, *vácci-i*), cf. AL *mana-i* [go-PST.3SG] ‘(s)he/it went’. Gradually, a child’s elementary morphological system will expand to include subclasses for inflections and word formation, and (s)he will begin inflecting words according to adult language systems (Bittner et al. 2003: xix).

The morphological acquisition of several typologically different, albeit mainly Indo-European majority languages, has been discussed in the international Cross-linguistic Project on Pre- and Protomorphology in Language Acquisition³, but so far only a small amount of research has been carried out on, for example, the acquisition of Saami.⁴ I will discuss in this article certain early inflectional models that a Saami-speaking child uses during the process of acquiring verb forms, and I will concentrate specifically on the inflectional variants of the first person singular present indicative, which I will be referring to in short as first person (singular) forms. The reason for focusing on these specific forms is that a child uses them from the very beginning of the verb inflection acquisition process and therefore they reveal the earliest protomorphological inflectional variants employed by a child. On the other hand, the first person singular forms are also interesting, as they are morphologically more complex than a child’s first premorphological, often monomorphemic forms, such as those mentioned above, *atte* and *boahtá*. The acquisition process of the first person forms therefore reveals the means a child employs when solving the problem of the morphological variation of the word stem and the affixation.

¹ With Saami, I am referring to North Saami, which is the variety spoken by 75% of Saami-speaking population, some 17,000 speakers (Sammallahti 1998: [1]).

² AL ‘adult language’.

³ The project actually encompasses 17 different languages. The project was founded by Prof. Wolfgang U. Dressler, and it is financed by the Austrian Academy of Sciences. (<http://www.oeaw.ac.at/ling/preproto.html> 8.2.2010).

⁴ See, however, Bals 2002, 2004.

2. Data

My research is based on data provided by my trilingual daughter, whose first languages are Saami and Finnish. She was born in 2002 in Tromsø in Northern Norway where the majority language is Norwegian. At the age of two she moved with her family to Guovdageaidnu (*Kautokeino* in Norwegian), where 85% of the population of 3,000 are Saami-speaking. Her father speaks Saami to her, while her mother speaks Finnish, and the parents communicate in Saami with each other. For her first 15 months she stayed at home with her mother and heard more Finnish than Saami, but ever since starting her Saami-speaking kindergarten at the age of 1;3, the dominant language in her environment has been Saami. Her father comes from Gáivuotna (*Kåfjord* in Norwegian), but most other adult Saami speakers around her speak the Guovdageaidnu dialect.

The data was recorded on video and includes conversations between my daughter and her father between the age period 1;8 and 3;0. The recordings were made in two of three week intervals, first for 30 minutes at a time, and from the age of 2;5 onwards for an hour at a time, for a total of 21.5 hours of video data.

Owing to the extensive scope of the data, I have restricted my analysis to a maximum of three consecutive occurrences of each new inflection of each lexeme. A new inflection is constituted by a form that is morpho(phono)logically different from the forms registered previously. The data includes 286 verb forms gathered in this manner.

A key concept in this article is *target form*, which I use to refer to the morphologically modularised form a child hears around him/herself and which over time the child will incorporate into his/her speech. There may be more than one target form in some cases, for example, when several dialectal variants exist.

3. First person singular present tense in Saami

There are three stem classes for verbs in Saami: bisyllabic, trisyllabic, and contracted verbs. Bisyllabic and contracted verbs are further divided into sub-classes according to the latus vowel. Particularly the stem of the bisyllabic words undergoes qualitative and quantitative changes when inflected. The quality and/or the length of the vowel centre and the consonant centre of the stem alternates in the different inflected forms, as does also the latus vowel, as shown in the following examples: *boahht-i* [come-IMP.2DU] ‘come!’ / [come-PTCP.PRS] ‘coming’; *boahťá-n* [come-PTCP.PST] ‘(has/have/had etc.) come’ / [come-CONN.PST] ‘(did not) come’; *boadá-n* [come-1SG] ‘I come’; *boahťi-t* [come-INF] ‘to come’ / [come-1PL] ‘we come’; *bohte-t* [come-3PL] ‘they come’ / [come-PST.2SG] ‘you came’ / [come-IMP.2PL] ‘come!’ etc. In trisyllabic and contracted verbs there is no gradation in the consonant centre nor any alternation in the vowels in first and second syllables.

The personal suffixes in first person singular are *-n* and *-an*. The suffix *-n* is appended to the bisyllabic and contracted verbs and the suffix *-an* to trisyllabic verbs. In present-tense bisyllabic verbs the suffix is appended to the weak stem, for example, *bora-n* (< *borra-t* ‘to eat’), *boadá-n* (< *boahťi-t* ‘to come’). The forms can appear either in the longer largo form or the shorter allegro form. In the largo form,

the unstressed vowel in the second syllable is unshortened, for example, *boadán* ‘I come’, *oaččun* ‘I get’, and in the allegro form, the same vowel has been shortened: *boadán* ‘I come’, *oaččon* ‘I get’. In the present-tense contracted verbs the suffix -*n* is appended to the bisyllabic vowel stem: *čohkká-n* ‘I sit’ (< *čohkká-t* ‘to sit’). In trisyllabic verbs, the personal suffix -*an* is appended to the consonant stem: *veahkeh-an* ‘I help’ (< *veahkeh-it* ‘to help’). (See Sammallahti 1998: 41–42, 77)

A child learning the language has, in other words, to differentiate between the number of syllables in the target form and the vowel and consonant variation in the stem when learning the stem of the first person singular inflection as well as to identify the inflectional morpheme appended to the stem.

4. The first occurrences of the first person singular form

Based on my data, the first person singular form first appears in a child’s language at the age 1;9–1;10, which can be regarded as a typical acquisition age in other languages also. According to Jorma Toivainen (1980: 49–50) a Finnish-speaking “median child” acquires this form at the age of 2;2. There is, however, considerable variation between individuals: while one child was recorded producing the form at 1;4, two other children did so at 2;5. The first person form appeared in Estonian-speaking Andreas’s speech at 1;10 and from the age of 2;1 he used it productively (Argus 2007).

In my data, the first occurrence of the first person singular form was one belonging to the stem class of trisyllabic verbs, *báhkin* (the figure after each example refers to the age at which the form was first recorded):

- (1) [*mun*] ***báhki-n*** (AL: *bálkest-an*) [*dohko*]⁵ (1;9)
 1SG throw-1SG there
 ‘I throw it there’

The target form is the trisyllabic *bálkestan* ‘I throw’, and the form presented above contains the target morpheme structure in the sense that it is possible to discern both the stem (*báhki*) and the inflection (*n*), but in a child’s language it has become shortened so as to seem a bisyllabic verb (see 5.4.).

The first form in the bisyllabic stem class to appear in the data was *áiggun* [will/ be going to-1SG] ‘I will/am going to’, which was in the target form. A couple of months later, forms of other lexemes were also registered in the data (cf. the acquisition ages of the first person forms with the Estonian-speaking child mentioned above):

- (2) ***áiggu-n*** [*mun dan gáhku*] (1;10)
 will-1SG 1SG it.GA cake
 ‘I want that cake’
- (3) [*mun*] ***bija-n*** [*dohko doavji-i*] (2;1)
 1SG put-1SG there tummy-ILL
 ‘I put (it) there in the tummy’

⁵ If the other words in expressions have no significance in interpreting the finite verb, I have only paid close attention to the finite forms when transcribing the data, and therefore the other elements in the expression (and for practical reasons also other than finite forms in questions) have been marked in square brackets. The main focus in the transcription of the data was to transcribe the morphology in the child’s verb forms, so that their development could be chronologically monitored. Other words appearing in registered expressions cannot therefore be used as evidence for, for example, the acquisition of other word forms.

- (4) [mun] **geahčá-n** (.) [gosa mana-i] (2;1)
 1SG watch-1SG where go-PST.3SG
 'I watch where (it) went'
- (5) [mu(n?)] **boahtá-n** (AL: boadá-n) [bargga] (2;1)
 1SG come-PTCP.PST/CONNeg.PST⁶ soon
 'I come soon'

The potential use of first person inflections is reduced by the fact that the child occasionally uses the noun *biigá* to refer to herself⁷ or a shortened version of her own name (Eliinná in Saami, Elina in Finnish) *Enná* then adding a third-person verb form: *biigá borrá* 'biigá eats', *biigá ohcá* 'biigá looks', *biigá ballá* 'biigá is afraid', *Enná geargá* 'Enná has time to'.⁸ At the age of approximately two years, the child begins to increasingly use the first person pronoun and the congruent verb form: *mun bijan* 'I put', *mun geahčan* 'I watch'. This method of reference will become established nearer to the age of 2;6, but the verb inflection used may still occasionally be the third person form (see 5.2.).

5. The inflectional variants of the first person singular form

According to my data, what is typical to the acquisition process of the first person singular inflection category is that although target forms have been registered throughout almost the entire data collection period from the time they first appear in a child's speech, she will use other forms alongside the target forms. Furthermore, while she uses a variety of morphologically different verb forms, which variant is dominant at a given time varies throughout the course of the collection period.

5.1. Non-suffixed form

A typical inflectional variant in the first personal singular is the non-suffixed word stem, appearing in a child's speech at the age of two and which remains in use for approximately eight months (2;0–2;8).

- (6) [m m mun] **bija** [áhči jusa da] (2;4)
 1SG put daddy.GA with here
 'I put here with dad'
- (7) [mu] **bijá** [dákkár-a -] (2;7)
 1SG.GA put this-GA
 'I put this'

⁶ See 5.3.

⁷ *biigá* originally means 'maid, servant', but is also used as a term of endearment to refer to little girls who are close to the speaker, often instead of a proper name.

⁸ The reason for using the third person form may lie in the fact that adults often use third-person forms when talking with children when referring to either themselves or their companion (see Toivainen 1980: 46). This type of speech is also registered in my data, where the father refers to the child as well as to himself with a noun (f = father, E = child)

f *biigá sárgu lámppa* ((question)) 'biigá draws a lamp'

E *ii'no'*

f *ii okei (,) áhčči sárgu lámppa* 'no okey (,) daddy draws a lamp' (1;8)

- (8) [mu] **oainná** [duoika] (2;1)
 1SG.GA see mosquito
 'I see a mosquito'
- (9) [á á] **áiggu** [mu mu mun dearpa-t] (2;5)
 will 1SG beat-INF
 'I will beat'

As the previous examples show, the *latus* vowel in the stem is either in *largo* or *allegro* form according to the structure of the target form, and the consonant centre is in weak grade. The use of the personal pronoun *mun* [1SG.NOM] 'I' without the /n/ sound in some of the examples above may be connected to the non-suffixed verb form, but it is hard to tell for certain what triggers such an inconsequent use of a *mu*-form that resembles the genitive-accusative case, cf. *mu* [1SG.GA] 'my; me'. The child uses the stem as the first person singular form particularly at the age of two and a half.

Non-suffixed forms also appear in trisyllabic and contracted verb categories:

- (10) [de mu] **geahččada** (2;5)
 then 1SG.GA try/try.3SG
 'then I try'
- (11) [mun] **háliida** (2;8)
 1SG want/want.3SG
 'I want'
- (12) [mu mu mu] **čohkká** [unna beavddi-s ná] (2;5)
 1SG.GA sit/sit.3SG small table-LOC like this
 'I sit by a little table like this'
- (13) [mun] **lea** [gáhhttu] (2;5)
 1SG be/be.3SG cat
 'I am a cat'

Interpreting these forms is not as unambiguous as with bisyllabic forms. This is because the lack of a personal suffix makes them identical with the third person singular forms (see 5.2.).

Another non-suffixed form is the strong-grade *máhtte*, in which the *latus* vowel is, in this context, unusual, *e*. The form was registered in the collected data twice, the other one presented in the following example:

- (14) E [do don máhte-t (AL: máhtá-t) bidja-t dá]
 2SG can-2SG put-INF here
 'you can put here'
- f [mon (.) jáhkátgo]
 'me ((question)), do you think'
- E [i-n mu] máhtte (AL: máhte)
 neg-1SG 1SG.GA can.CONNEG
 'I can't'
- [nie bidja-t diekko (.) á áhčči (.) mu mun]
 so put-INF there dad 1SG
 'so put there, dad I'

<i>máhtte</i>	[<i>diekkár-a</i>	<i>bidja-t</i>	<i>lávka-i</i> ((?)) (2;7)
can	that-GA	put-INF	bag-ILL
'I can put that in a bag'			

The example shows that the child uses a strong-grade *máhtte* deviant from adult language first, instead of a weak-grade present negative *máhte*, and immediately after this, the same form is used to refer to the child herself in a positive sentence. As mentioned above, the lack of inflectional morpheme is commonplace at this stage of child language acquisition, but replacing the latus vowel with *e* was registered on three occasions only: once in the form *máhtte-n* and twice in *máhtte*. In all three cases, the child and/or her father had shortly before used the negative form *máht(t)* *e*, the latus vowel in which seems to have caused the use of the non-suffixed *máhtte* form as a finite form.

It is also noteworthy that the child uses the vowel *e* on the first line of the above examples as the latus vowel in *máhte-t* (AL: *máhtá-t*), see also the child's other usage that deviates from adult language, using *e* as the latus vowel (which are not, however, very productive based on the data), such as *bease-t*⁹ (AL: *beasa-t* [get-2SG] 'you get'), *bidje* (AL: *bidjá* [put.3SG] 'puts') and *galgge* (AL: *galgga* [CONNeg] 'must (not)'). The stem with the latus vowel *e* may therefore be a kind of "universal stem" that the child uses in a number of different contexts. Forms in which, contrary to the general rule, the vowel in the unstressed syllable is *e*, also exist in the language of those adult speakers whose Saami skills are not very strong, for example, *sáhte* [can.3SG] '(s)he can', cf. AL: *sáhtá* [can.3SG] '(s)he can' (cf. AL: *sáhte* [CONNeg] 'can (not)'); *sáhtte* [can.1PL] '(we) can', cf. AL: *sáhtti-t* [can-1PL] (cf. AL: *sáhtte* [can.2DU/ PST.3PL] '(we) can/ they could').

The data features examples of the usage of weak-grade stems also in other inflections: the child uses the forms *bijá* [put], *boadá* [come] and *dáhta* [can] in second person singular, but each of these forms were registered only once, which would indicate that in practice, the use of the non-suffixed form is "reserved" for one purpose, the first person singular. The non-suffixed form is used only in the present tense, because in the preterit forms, the first of which appears at the age of 2;3, the child uses the *n*-suffix throughout the data collection period.

The use of the stem as a non-suffixed inflectional form has also been documented in Finnish child language. A Finnish-speaking child may use the noun stem without the genitive suffix *-n* and express the morphological characteristics of the word through gradation only: *mekko* [dress.NOM] : *meko* [dress.GEN], cf. AL: *mekko* [dress.NOM] : *meko-n* [dress-GEN]. The reason behind this is assumed to be that the suffix *-n* is not as salient as gradation. (Laalo 1997: 199–200, 2002: 92, Karjalainen 1979: 25–26.) The word-final *-n* is not always easily discernible in Saami either, and therefore it is possible that gradation is a more salient feature for a Saami child than an inflectional suffix, cf. second person singular suffix /*h(t)*/; cf. also first person singular preterit forms in the data, which all have the suffix *-n*.

⁹ In the form *bease-t* the vowel in the unstressed syllable could, in principle, be interpreted as being influenced by the second syllable vowel in the form *päáse-t* [get-2SG] '(you) get' in Finnish, the child's other first language. See also the latus vowel in the present negative *bease* in the following example written by an adult: *Dán jagi juovlabasáras Guovdageainnus ii bease* [[CONNeg] < *beassa-t* [get-INF]] *dušše vuotlihku geahččalit, doppe beasat maid vásihit báikkálas juigiid ja lávluid.* 'In this year's Christmas bazaar in Guovdageaidnu you get not only to try your luck in the draw but also to listen to local yoikers and singers.' (Áššuu 18.12.2007 p. 4.)

5.2. Third person forms

Nearly concurrently with non-suffixed forms, at the age of 2;1–2;9, the child also uses the third person singular forms for the first person inflectional forms, as evident in the following examples:

- (15) [mun] **manná** [gávpog-ii (.) (>)]
 1SG go.3SG Town-ILL
 ‘I go to town’
 [(>) *biigá* *boahtá* *bargga*] (2;1)
 “*biigá*” come.3SG soon
 “*biigá*” comes soon’
- (16) [*gea* (.) *dieppe*] **galgá** [mun *dahka-t* (.) (>)]
 look there must.3SG 1SG do-INF
 ‘look, there I must do’
 [(>) *mun*] **bidjá** [ná ná ná *nie* (.) *ale* (.) (>)]
 1SG put.3SG like this like this so NEG.IMP.2SG
 ‘I put this way so, don’t’
 [(>) -- *nie* (.) *mun*] **bidjá** [*eret* (.) *don* *i-t* (>)]
 -- so 1SG put.3SG away 2SG NEG-2SG
 ‘-- so I put away, you must’
 [(>) *oaččo* *nie* (.) *mun*] **dahká** [*nie* *ná* *nie*] (2;4)
 can.CONNEG so 1SG do.3SG so like this so
 ‘not do so, I do so, like this, so’
- (17) [*gea* *ale* (.) *mun*] **duhtta** (2;4)
 look NEG.IMP.2SG 1SG get angry.3SG
 ‘look, don’t, I get angry’

The use of the third person singular instead of the first person fades after the age of two and half years. As a negative verb, the child usually uses the *in*-form as in adult language from its very first appearances onwards at the age of 2;0, but at a little under two years and six months, the child also starts using the third person singular form *ii* alongside the *in*-form. Its use alongside the target form continues sporadically up until the age 2;9:

- (18) [*m-m mun*] **ii** [*lea(-t?)* *váltán*] (2;4)
 1SG NEG.3SG be(-CONNEG) take-CONNEG.PST
 ‘I have not taken’
- (19) [*mu mu mu (?)*] **ii** [*gille*] (2;5)
 1SG.GA NEG.3SG want.CONNEG
 ‘I don’t want’

Finnish-speaking children may also use the third person singular verb form with a first person singular pronoun, according to Toivainen, until the age of two years and six months, for example, *minä* [1SG] *hakee* [get.3SG] ‘I get’ (cf. AL: *hae-n* [get-1SG] ‘I get’) (Toivainen 1980: 46–47).

With trisyllabic and contracted verbs, the third person singular form and the non-suffixed form converge so that in the examples (11) and (12) above, the verb forms *háliida* and *čohkká* are ambiguous.

5.3. The consonant centre in bisyllabic strong-grade stems

The choice of the grade in the consonant centre does not cause too many problems for a two-year-old, although at that age the following forms with the weak instead of the strong grade were registered: *galgá-n* [must-1SG] ‘I must’ (cf. AL: *galgga-n* [must-1SG] ‘I must’); *boahtá-n* [come-1SG] ‘I come’ (cf. AL: *boadá-n* [come-1SG] ‘I come’; see Chapter 4.) and *man(n)a-n* [go-1SG] ‘I go’ (cf. AL: *mana-n* [go-1SG] ‘I go’). The central consonants in the forms used by the child and the target forms are usually identical, but especially from the age 2;7 until the end of the data recording period, the child sometimes uses forms with a strong-grade consonant centre, in addition to the target forms. It follows that the forms converge with other forms in the adult language paradigm, as the following examples show:

- (20) [*mu mu mu mun*] ***dahka-n*** (AL: *daga-n*) (2;7)
 1SG do-PTCP.PST/CONNeg.PST/ACT.NOM/VN
 ‘I do’
- (21) [*juo mun*] ***galga-n*** (AL: *galgga-n*) [*manna-t*] (2;8)
 yes 1SG must-PTCP.PST/CONNeg.PST/ACT.NOM/VN go-INF
 ‘yes I must go’
- (22) [*go dál mun*] ***váldá-n*** (AL: *válddá-n*) [*eret (>)*]
 because now 1SG take-PTCP.PST/CONNeg.PST away
 ‘because now I take (it) away’
 [(>) *dá-i-d go dat lea juhka-n*] (2;10)
 this-PL-ACC when it be.3SG drink-PTCP.PST
 ‘these when it has drunk’
- (23) [*ja*] ***bidja-n*** (AL: *bija-n*) [*duo-sa*] (2;11)
 ja put-PTCP.PST/CONNeg.PST/ACT.NOM/VN that-ILL
 ‘and I put there’

All the above forms (apart from *dahkan*, which was registered three times) are exceptions to the forms usually used by the child. For example, the gradation of the consonant centre *dj : j* was registered as early as at 1;10: *bidjá : bija* (see also Bals 2002: 49; according to her, *dj : j* gradation is one of the earliest gradation types children acquire).

When a child uses the stem (e.g. *bijá*, cf. AL: *bija-n*) in parallel with the form with a strong-grade consonant centre (e.g. *bidja-n*, cf. AL: *bija-n*), it would also be possible to categorise strong-grade, non-suffixed forms such as *manná* and *máhtá*, which I have now interpreted as being third person forms, as stems, in which the grade of the consonant centre deviates from that of the target form, cf. *bijá* and *bidjá*. This interpretation is supported in the case of *a*-stemmed forms by the fact that in the target forms of the Guovdageaidnu and Gáivuotna dialects, for example

the one which according to standard orthography is spelled *manan*, the latus vowel *a* is lengthened when the vowel centre of the word is short and the consonant centre is in quantity I or II: *manán*, *dagán*, *logán* etc. (see Sammallahti 2006: 175–176).

5.4. The shortening of trisyllabic stems

The number and frequency of inflectional forms of trisyllabic verbs registered in the data was substantially lower than those of bisyllabic verbs. However, certain main characteristics are discernible, as shown by examples in Sections 5.1. and 5.2. Trisyllabic stems may also be shortened into bisyllabic ones to begin with, either so that the word form used by the child does not contain the second syllable present in the adult language form or that the third syllable in the adult language form, and thereby also the inflectional suffix, is lacking in the form used by the child:

(24) [*mh mun*] ***báhki-n*** (AL: *bálkest-an*) [*dohko*] (1;9)
 1SG throw-1SG there
 ‘I throw it there’

(25) ***háji*** (AL: *háliid-an*) [*bojja-t vel eambbo*] (2;6)
 want eat-INF even more
 ‘I want to eat even more’

(26) [*mun*] ***veahk-an*** (AL: *veahkeh-an*) [*du*] (2;7)
 1SG help-1SG 2SG.GA
 ‘I help you’

The shortening of trisyllabic words into bisyllabic ones has also been documented by Berit Anne Bals Baal, in whose data the youngest informants, two and two-and-a-half-year-old children, only used bisyllabic words, for example *rahhka*, cf. AL: *ráhkad-it* [make-INF] ‘to make’, *baači*, cf. AL: *bávččag-ii* [hurt-PST.3SG] ‘(it) hurt’, *věahhke*, cf. *veahket* [help.IMP.2SG] ‘help!’ (Bals 2002: 69–70, 53)¹⁰. In one of the early stages of word acquisition, common to many languages, words longer than two syllables are shortened. According to studies on the phonological development of Finnish-speaking children, words may be shortened also during a stage known as the systematic language acquisition, so that words with three or more syllables are shortened into bisyllabic ones. (Kunnari 2000, Savinainen-Makkonen 2001, Savinainen-Makkonen, Kunnari 2004: 86)

5.5. Non-frequent inflections

In addition to the principal inflection types discussed above, the child also uses other inflectional forms instead of first person forms. Most of the forms were single occurrences and exceptions to the forms the child usually uses, and I will therefore not discuss those in detail in this conjunction. One of these occurrences is interesting because it may be a case where the child’s other first language Finnish affects the Saami inflection. Namely in one verb form, the child uses the latus vowel, which deviated from the vowel in the unstressed syllable in the target form:

¹⁰ See also the trisyllabic and quadrosyllabic forms in my data, registered at the age 1;11–2;1: *čohkkodit* (AL: *čohkked-it*) [sit down-INF] ‘to sit down’, *billahuvvá* [break, break down.3SG] ‘breaks, breaks down’, *bálkest-it* [throw-INF] ‘to throw’, *bávččag-ii* [hurt-PST.3SG] ‘(it) hurt’, *oahtti-me* (AL: *oasti-me*) [buy-ACT.ESS] ‘buying’.

- (27) [go mun] **lávlla-n** (AL: *lávllu-n*) [dan nuppe --] (2;11)
 when 1SG sing-1SG it.GA other.GA
 ‘when I sing the other --’

Here, the form *lávllan* ‘I sing’ seems to have been affected by the Finnish form *laulan* ‘I sing’. In adult language the verb *lávlut* ‘to sing’ belongs to bisyllabic *u*-stemmed verbs, but the vowel *a* in the second syllable of the Finnish first person singular form may be the reason for the child inflecting the verb according to model for bisyllabic *a*-stemmed verbs, cf. for example Saami *guovla-t* [peek-INF]: *guovlla-n* [peek-1SG]. The use of the vowel *a* may also have been motivated by the fact that there are far more *a*-stemmed verbs in Saami than *u*-stemmed ones, and that they are more frequent than *u*-stemmed verbs. The form *lávllan* appeared in the data twice towards the end of the recording period, but in the same recording, one minute after the *lávllan* forms, also the target form appears:

- (28) [kánske mun] **lávllu-n** [dien] (2;11)
 maybe 1SG sing-1SG it.GA
 ‘maybe I sing it’

The *lávllan* form can in fact be regarded as an exception among the verb forms used by the child, since there are no other examples in which the present-tense form of the bisyllabic verb subclass would incorporate the latus vowel of another subclass.

6. Conclusion

The target forms of first person inflectional forms began to establish in the child informant’s language at the age of around two and a half years, but she would use other inflectional variants alongside them almost up until the end of the data recording period.

The main stages of the acquisition process of the target forms are illustrated in Diagram 1, which gives the various variants of the form *bija-n* [put-1SG] ‘I put’ and the age at which they were used. On the vertical axis is the child’s age (year, month and day) at which each form was recorded on video, while the horizontal axis gives the form variants in chronological order. The vertical columns show when a given form was used and how many times. The figure in brackets refers to the number of times each variant was registered on the tape recorded at that particular age. The plus sign means if the form was used at least as many times as indicated by the number in brackets. For example the child may have used the *bijan* form countless times from the age 2;8.27–28 onwards until the form *bidjan* appeared at the age 2;11.1–2, but the diagram only shows the first three consecutive occurrences of each variant (in this case: *bijan*), which are different from the preceding variant (in this case *bijá* at the age 2;7.10), namely one at the age 2;7.30 and two at the age 2;8.27–28.

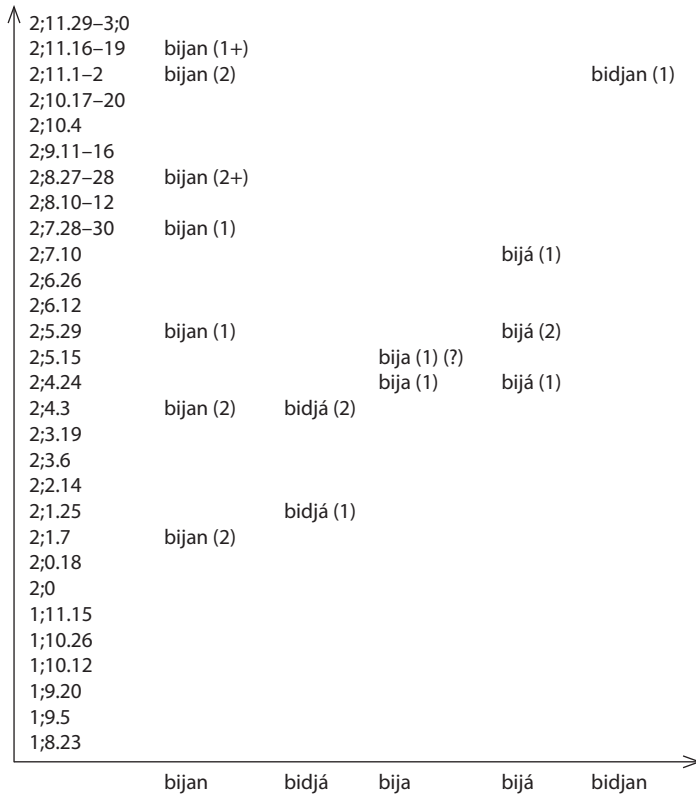


Diagram 1. Acquisition of the form *bijan*

The Diagram 1 shows that the target form *bijan* was first registered at the age of roughly two years and that the child uses it almost throughout the entire data recording period. However, a few weeks after the first occurrence, the target form is replaced by a third person form *bidjá*, but it disappears a few months later, when a non-suffixed form appears in the child's language, the latus vowel being either in largo or allegro form. The child uses the non-suffixed form and the target form in parallel for a few months until the target form becomes established as the sole form after the age of two and a half years. However, when the child is nearly three years old, one occurrence is still registered of a form with the consonant centre in strong grade.

The diagram shows that in addition to the target form *bijan*, the child uses three other inflectional variants deviating from the morphophonology of the target form: (1) the non-suffixed forms *bija* and *bijá*, (2) the third person singular form *bidjá* and (3) the strong-grade form *bidjan*.

Based on all the first person singular forms featuring in the data, it appears that non-suffixed forms and the third person singular forms are used almost in parallel: the former at the age 2;0-2;8, the latter at the age 2;1-2;9. The data has five lexemes that have been used both in a non-suffixed form and the third person singular form, and based on these five lexemes it can be assumed that when acquiring the language the child may use third person forms and non-suffixed forms consecutively (as in the forms of the lexemes *BIDJAT* and *DAHKAT*) or in parallel (as

in the forms of the lexemes GALGAT, MANNAT and ÁIGUT). It would, however, require a more extensive sample of data to ascertain whether it is more common for the child to use these forms consecutively or in parallel.

The child uses the word stem particularly for the first person singular, but the third person singular form is used, according to the data, instead of the target form also in other persons and numbers, such as the second person singular and third person plural.

The child uses strong-grade forms particularly in the latter part of the data recording period, at the age 2;7–2;11, and they are mostly isolated incidences, appearing alongside target forms. Usually the consonant centre is in the weak grade as per the target form model, but like in other inflections, the grade may also deviate from the grade of the consonant centre in the target form.

In addition to the three main inflectional forms that deviate from the target form, as mentioned above, the child may also use other inflections. Although most of these occurrences were registered only once or twice, they reinforce the view that the acquisition process of the verb inflection system is a complex one and show that during the protomorphological phase, the child experiments with combining various morphophonological rules before eventually acquiring the target forms.

Abbreviations

1	first person	INF	infinitive
2	second person	LOC	locative
3	third person	NOM	nominative
ACT	the so-called action	PL	plural
CONNÉG	connegative	PRS	present
DU	dual	PST	past
ESS	essive	PTCP	participle
GA	genitive-accusative	SG	singular
ILL	illative	VN	verbal noun
IMP	imperative		

References

- Argus, Reili 2007. Morphological development of three-year-old Estonian speaking children. Paper presented at the Kielitieteen päivät conference at the University of Oulu 24–25 May 2007.
- Áššu [Newspaper. Guovdageaidnu, Norway].
- Bals, Berit Anne 2002. The Acquisition of Grade Alternation in the Kautokeino Dialect. Unpublished MA Thesis. [Tromsø]: Linguistic Department, University of Tromsø.
- Bals, Berit Anne 2004. The acquisition of grade alternation in North Saami. – Tromsø Working Papers in Language Acquisition. Nordlyd. Tromsø University Working Papers in Language & Linguistics, 32, 1 (2004), 1–27. <http://www.ub.uit.no/baser/nordlyd/viewissue.php?id=7> (22.09.2009).
- Bittner, Dagmar; Dressler, Wolfgang U.; Kilani-Schoch, Marianne 2003. Introduction. – Dagmar Bittner, Wolfgang U. Dressler, Marianne Kilani-Schoch (Eds.). Development of Verb Inflection in First Language Acquisition. A Cross-Linguistic Perspective. Studies in Language Acquisition, 21. Berlin & New York: Mouton de Gruyter, vii–xxxvii.
- Dressler, Wolfgang U.; Karpf, Annemarie 1995. The theoretical relevance of pre- and protomorphology in language acquisition. – Geert Booij, Jaap van Marle (Eds.). Yearbook of Morphology 1994. Dordrecht: Kluwer Academic Publishers, 99–122.

- Ijäs, Johanna 2009. *Mun válddá dán* – giellaoččodeaddji máná ovddemus árramorfologalaš vearbasojahanvuogit. – Johanna Ijäs, Nils Øivind Helander (Eds.). Sáhkavuoruiin sáhkkan. Sámegeiela ja sámii girjjálašvuoda muhtin áigequovdilis dutkanfáttát. Diedut 1/2009. Guovdageaidnu: Sámi allaskuvla, 54–69.
- Karjalainen, Merja 1979. Lapsenkielen genetiivi. Yksi-, kaksi- ja kolmivuotiaiden lasten genetiivin käytön tarkastelua. Unpublished MA Thesis. [Oulu]: Oulun yliopiston suomen ja saamen kielen laitos.
- Kilani-Schoch, Marianne; Dressler, Wolfgang U. 2002. The emergence of inflectional paradigms in two French corpora: an illustration of general problems of pre- and protomorphology. – Maria D. Voeikova, Wolfgang U. Dressler (Eds.). Pre- and Protomorphology. Early Phases of Morphological Development in Nouns and Verbs. LINCOM Studies in Theoretical Linguistics, 29. Muenchen: Lincom Europa, 45–59.
- Kunnari, Sari 2000. Characteristics of Early Lexical and Phonological Development in Children Acquiring Finnish. Acta Universitatis Ouluensis B 34. Oulu: Oulu University Press.
- Laalo, Klaus 1997. Alkukatsaus lapsenkielen esi- ja varhaismorfologiaan. (English summary: Introduction to pre- and protomorphology. <http://www.kotikielenseura.fi/virittaja/hakemistot/jutut/vir97laalo.html> 22.09.2009.) – Virittäjä, 101, 186–207.
- Laalo, Klaus 1998. Välikatsaus lapsenkielen varhaismorfologiaan. (English summary: Interim review of protomorphology in child language. <http://www.kotikielenseura.fi/virittaja/hakemistot/jutut/vir98laalo.html> 22.09.2009.) – Virittäjä, 102, 361–385.
- Laalo, Klaus 1999. Ensianoista ja esimorfologiasta varhaismorfologiaan. Lapsen sija-järjestelmän ja verbintaivutuksen alkuvaiheita. (English summary: From first words and premorphology to protomorphology. <http://www.kotikielenseura.fi/virittaja/hakemistot/jutut/vir99laalo.html> 22.09.2009.) – Virittäjä, 103, 354–377.
- Laalo, Klaus 2002. Acquisition of case in Finnish: A preliminary overview. – Maria D. Voeikova, Wolfgang U. Dressler (Eds.). Pre- and Protomorphology. Early Phases of Morphological Development in Nouns and Verbs. LINCOM Studies in Theoretical Linguistics, 29. Muenchen: Lincom Europa, 83–103.
- Laalo, Klaus 2003. Early verb development in Finnish: A preliminary approach to mini-paradigms. – Dagmar Bittner, Wolfgang U. Dressler, Marianne Kilani-Schoch (Eds.). Development of Verb Inflection in First Language Acquisition. A Cross-Linguistic Perspective. Studies in Language Acquisition, 21. Berlin & New York: Mouton de Gruyter, 323–350.
- Sammallahti, Pekka 1998. The Saami Languages. An Introduction. Kárásjohka: Davvi Girji.
- Sammallahti, Pekka 2006. Jietnadataohpa vuoddogursa. Unpublished manuscript.
- Savinainen-Makkonen, Tuula 2001. Suomalainen lapsi fonologiaa omaksumassa. (English summary: Finnish children acquiring phonology.) Publications of the Department of Phonetics, 42. [Helsinki]: University of Helsinki.
- Savinainen-Makkonen, Tuula; Kunnari, Sari 2004. Ensisanojen rakenteet. – Sari Kunnari, Tuula Savinainen-Makkonen (Eds.). Mistä on pienten sanat tehty. Lasten äänteellinen kehitys. Helsinki: Werner Söderström Osakeyhtiö, 84–87.
- Toivainen, Jorma 1980. Inflectional Affixes Used by Finnish-speaking Children Aged 1–3 years. Suomalaisen Kirjallisuuden Seuran toimituksia, 359. Helsinki: Suomalaisen Kirjallisuuden Seura.

Johanna Ijäs (Sámi allaskuvla/ Sámi University College, and Giellagas Institute, University of Oulu). Her research interests include language acquisition, morphology, and child bilingualism. She has published scientific papers in the field of acquisition of morphology in Saami, and bilingual language acquisition. She has edited several scientific publications.
johanna.ijas@samiskhs.no

“MA PANEB SELLE ÄRA” EHK VARASED PROTOMORFOLOOGILISED VERBIMUUTE- STRATEEGIAD SAAMI LAPSEKEELES

Johanna Ijäs

Sámi allaskuvla / Sámi University Collage ja Giellagas Instituut, Oulu Ülikool

Artikkel kirjeldab mõnesid protomorfoloogilisi verbimuuteviise põhjasaami lapsekeeles vanuses 1;8–3;0. Uuringu keskmeks on valitud kindla kõneviisi oleviku ainsuse esimese isiku vormid. Laps kasutab neid juba alates verbimuutesüsteemi õppimise algusest, seetõttu peegeldavad just need vormid kõige varasemaid protomorfoloogilisi infleksioonivariante, mida laps kasutab. Need vormid on sageli ka morfoloogiliselt keerukamad kui lapse esimesed ühemorfeemilised vormid.

Laps kasutab täiskasvanule sarnaseid muutevorme kogu keelearengu perioodil alates ainsuse esimese isiku vormi tekkest vanuses 1;9–1;10, kuid lisaks täiskasvanutega sarnastele vormidele leidub veel kolm põhilist viisi, kuidas laps verbe muudab. Vanuses 2;0–2;8 kasutab laps verbe sageli ilma *n*-lõputa, nt *bija* (vrd täiskasvanukeeles *bija-n* ‘panema-SG1PRES’). Umbes samal ajal, vanuses 2;1–2;9, kasutatakse esimese isiku asemel kindla kõneviisi oleviku ainsuse kolmanda isiku vormi, nt *bohtá* (vrd täiskasvanukeeles *boadá-n* ‘tulema-PRES.SG1’). Kolmas erinevus täiskasvanust on tugeva astme kasutamine kahesilbiliste sõnade konsonantses keskmes, nt *váldán* (vrd täiskasvanukeeles *válddá-n* ‘võtma-PRES.SG1’).

Lisaks kolmele ülalnimetatud muutelõpule kasutavad lapsed ka muid lõppe. Kuigi enamik muudest juhtudest esines vaid 1–2 korral, kinnitavad need seisukohta, et verbimuutesüsteemi omandamise protsess on keerukas, ning et protomorfoloogilises faasis eksperimenteerib laps erinevate morfoloogiareeglite kombineerimisega, enne kui lõpuks sihtvormid omandab.

Võtmesõnad: kakskeelne esimese keele omandamine, morfoloogia, verbivormistik, põhjasaami keel, soome keel

MÕNEDE EESTI SÕNAJÄRJEMALLIDE PSÜHHOLINGVISTILISEST REAALSUSEST

Annekatriin Kaivapalu

Ülevaade. Projekti “VAKO: Eesti vahekeele korpuse keeletarkvara ja keeletehnoloogilise ressursi arendamine (2008–2010)” üheks eesmärgiks on automaatse vealeidja prototüübi loomine. Alustati eesti õppijakeele kõige sagedasemast veatüübist, milleks korpuse veastatistika põhjal on sõnajärg. Kuna eesti keelele omane suhteliselt vaba sõnajärg võimaldab aktsepteerida paralleelselt mitmeid sõnajärjemalle, on ka sõnajärjevigade interpreteerimine väga varieeruv.

Käesoleva artikli eesmärgiks on selgitada, milliseid sõnajärjemalle peavad emakeelekasutajad ja teise keele õppijad loomulikuks ning milliseid mitte, et saadud tulemuste põhjal kirjeldada mõnede eesti keele sõnajärjemallide psühholingvistilist reaalsust ning määratleda sõnajärjevigade prioriteetsus vealeidja loomise seisukohalt. Selleks koostati sõnajärjemallide loomulikkuse hindamise test. Sõnajärjemallide valiku aluseks olid need korpuse sõnajärjevead, mille osas märgendajate vahel eriarvamusi ei olnud. Sõnajärjemallide loomulikkust paluti hinnata Tallinna Ülikooli eesti- ja venekeelsetel üliõpilastel.

Uurimuse tulemusena selgus, et V2-reegel on eestikeelse keelekasutaja jaoks psühholingvistiliselt reaalne ning kõige häirivam on laiendi paiknemine lausealgulise subjekti ja verbi vahel. Venekeelse õppija keeletaju seisukohalt ei ole V2-reegli järgimine või rikkumine aga distinktiivne: ühtviisi loomulikuna tajuti sõnajärjemalle subjekt-verb-laiend-laiend-(laiend), laiend-verb-subjekt-laiend-(laiend), laiend-subjekt-verb-laiend-(laiend) ja subjekt-laiend-verb-laiend-(laiend).*

Võtmesõnad: õppijakeel, korpuspõhine veaanalüüs, sõnajärg, emakeelne keelekasutaja, eesti keel, vene keel

* Uurimus on teostatud riikliku programmi “Eesti keele keeletehnoloogiline tugi (2006–2010)” projekti “VAKO: Eesti vahekeele korpuse keeletarkvara ja keeletehnoloogilise ressursi arendamine (2008–2010)” raames.

1. Taustast

Riikliku programmi “Eesti keele keeletehnoloogiline tugi (2006–2010)” projekti “VAKO: Eesti vahekeele korpuse keeletarkvara ja keeletehnoloogilise ressursi arendamine (2008–2010)” tegevuse üheks eesmärgiks on automaatse vealeidja prototüübi loomine. Töö käigus on tulnud leida vastuseid mitmetele küsimustele, sh välja selgitada, mis tüüpi vead õppijakeeles kõige sagedamini esinevad, ning määratleda veatüübi olulisus kommunikatiivse keeleoskuse omandamise seisukohalt, kuna sage veatüüp ei tarvitse ilmtingimata suhtluseesmärgi saavutamist takistada. Samuti tuli otsustada, milliseid meetodeid ja tarkvara kasutades vigu leida: kas lähtuda reeglipõhisusest või veastatistikast või eelistada erinevate lähenemiste sünteesi. Programmeerimise seisukohalt oli oluline määratleda, missugust lingvistilist analüüsi vealeidja loomine vajab, et statistikapõhine programm hakkaks vigu leidma. Omaette küsimus on, millist osa lausest peaks vealeidja õppima analüüsima. Ühelt poolt võivad õppijakeele vead ilmned normatiivsest lausestruktuuri märgendusest väljaspool, teisalt aga pole normivastatus võõras ka loomulikule keelekasutusele. Analoogilistele küsimustele on vastuseid otsitud ka näiteks rootsi õppijakeele vealeidja väljatöötamisel (Kann 2002, Borin 2002, Bigert jt 2005).

Eesti õppijakeele vealeidja prototüübi loomist alustati õppijakeele korpusanalüüsist. Selleks, et välja selgitada, milline vealiik on õppijakeeles sage, koostati Eesti vahekeele korpuse¹ (edaspidi EVKK) kesktaseme tekstidest pilootkorpus (Pille Eslon), mis sisestati ESTCG parserisse (Müürisep 2000) ning ühestati käsitsi (Helena Metslang, Pille Eslon, Annekatrin Kaivapalu). Seejärel seoti ESTCG märgendus EVKK veataksonomiaga (Erika Matsak, Vahur Rebas), mis võimaldas ühestajatel jätkata pilootkorpuses esinevate keelevigade käsitsi märgendamisega. Märgendatud vigade statistikast selgus pilootkorpuse vealiikide sagedus: kõige sagedasemateks osutusid sõnajärvead.

Seega oli tegemist korpuspõhise veaanalüüsiga, mis kahtlemata aitab paremini mõista õppijakeele olemust ja arengustaadiume ning omab olulisi pedagoogilisi (Granger 2002: 14) ja keeletehnoloogilisi rakendusi. Siiski kätkeb veaanalüüs meetodina endas paljusid metodoloogilisi küsitavusi, mida on põhjendatult kritiseeritud (Borin, Prütz 2004: 68, Ellis, Barkhuizen 2005: 70). Veaanalüüsile on ette heidetud lähteandmete heterogeensust, veakategooriate laialivalgust, keeleomandamise käsitlemist staatilisena, piirdumist ainult sellega, mida õppijad teha ei suuda. Olulisemaid vajakajäämisi on väga juhuslik teave produtseerimisprotsessi ning õppijate strateegiate kohta (Kaivapalu 2008). Samuti jääb pelgalt veaanalüüsi põhjal selgusetuks, milliseid keelendeid õppija veana tunnetab ja kas need vead on häirivad ka emakeelse keelekasutaja jaoks. Keelend, mis ei vasta kirjakeele normile, ei tarvitse tingimata olla vastuolus emakeelse keelkasutaja keeletajuga: normivastatus ei ole emakeelsele keelekasutusele sugugi võõras. Kuna eesti keelele omane vaba sõnajärg võimaldab aksepteerida paralleelselt mitmeid sõnajärjemalle, on sõnajärvevigade interpreteerimise diapason väga varieeruv. Seetõttu on poolautomaatse vealeidja prototüübi väljatöötamisel lisaks vealiikide sagedusele oluline arvestada ka nende prioriteetsusega. Keeleõppe seisukohalt on esmatähtis, et õppija keelekasutus oleks emakeelse kõneleja keelekasutusele võimalikult lähedane. Sagedasti esinev viga ei tarvitse emakeelse keelekasutaja jaoks olla ilmtingimata häiriv, küll aga on

vealeidja prototüübi väljatöötamisel esmajoones põhjust keskenduda vigadele, mida emakeelne keekekasutaja veana tunnetab.

Käesoleva uurimuse eesmärgiks on määratleda vealeidja loomise tarvis sõnajärjevigade prioriteetsus. Selleks selgitatakse kirjeldava statistika abil välja eesti õppijakeeles sagedamini esinevate normikohaste ja normivastaste sõnajärjemallide loomulikkus ja ebaloomulikkus nii emakeelse keekekasutaja kui keeleõppija arvates. Eesmärgiks ei ole statistiline tõeväärsus, vaid suundumuste kirjeldamine antud valimi piires. Artikli eesmärgiks ei ole ka strukturalistlik-biheivioristliku ja veaanalüüsi teooria vaimus teha keelestruktuuride formaallingvistilise võrdluse põhjal oletusi EVKK-s esinevate sõnajärjevigade põhjuste kohta. Samuti ei käsitleta emakeelsete keekekasutajate ja keeleõppijate hinnangute põhjusi, sest põhjuslik analüüs saab tugineda ainult introspektiivsele ainekule. Niisiis on tegemist emakeelsete keekekasutajate ja keeleõppijate hinnangute esialgse kaardistamisega edasise uurimistöö tarvis.

2. Õppijakeel, sõnajärg ja V2-reegel

Sõnajärg kuulub nende sihtkeele nähtuste hulka, mille vastu eksivad isegi väga hea keeleoskusega õppijad. Ka need eriti edasijõudnud keeleõppijad (ingl *near-native speakers*), kelle keeleoskus muus osas emakeelse keekekasutaja omast ei erine, eristuvad sageli vaid sihtkeelest erineva sõnajärje poolest. Sihtkeeltes, mille sõnajärjele on omane verbi paiknemine lauses teisel kohal (V2-reegel), nagu näiteks rootsi, saksa ja arvatavasti viimase mõjul ka eesti keeles (Aavik 1912), on õppijakeele lausetasandi sõnajärje valdavaks probleemiks V2-reegli järgimine: õppijad ei tunnetata sageli verbi asukohta lauses normikohase ja -vastase keekekasutuse eristajana. Kui eriti edasijõudnud rootsi keele õppijatel, kes igapäevases suulises suhtluses ei erinenud emakeelsetest keekekasutajatest, paluti hinnata mõnede morfosüntaktilise keelenähtuste osas (V2-reegel, refleksiivsed possessiivpronoomenid, ühildumine soos ja arvus, adverbide paiknemine kõrvallauses) vigu sisaldavate lausete grammatilist õigsust, põhjustasid probleeme just V2-reegli vastu eksimised, mida vigadena ei tunnetatud (Hyltenstam, Abrahamsson 2003, Abrahamsson, Hyltenstam 2009). Ka EVKK sõnajärjevigade hulgas domineerisid V2-reegli vead (muudest EVKK sõnajärjevigadest vt Metslang, Matsak 2010).

Kuigi esmapilgul võib tunduda, et V2-reegli omandamine sihtkeeles on eriti keerule nende õppijate jaoks, kelle emakeeles see puudub, ei ole keeletöötlusprotsess² kunagi tõlgendatav ainult lingvistilises võtmes. Järelduste tegemine pelgalt lähte- ja sihtkeele struktuuride, sh sõnajärjereeglite, või lähtekeele struktuuri ja õppija sihtkeele kasutuse võrdluse põhjal on ilmne lihtsustamine (lähtekeele mõju uurimise ühtsest metodoloogilisest raamistikust vt Jarvis 2000). Uurimistulemused osutavad lisaks emakeelele teiste varem omandatud keelte mõju V2-reegli järgimisele ja eiramisele (Cedden, Audin 2009), samuti on tõendeid nii otsese keeleülekande vastu (Pienemann 1988, Håkansson 2001) kui poolt (Cedden, Audin 2009). Pärilise probleemi ei ole V2-reegli omandamine ka nende õppijate jaoks, kelle emakeeles antud reegel olemas on (Bochnacker 2006), samas aga võib otsene keeleülekande puududa juhtudel, kus seda võiks keelestruktuuride võrdlusel eeldada (Håkansson jt 2002).

² Artiklis kasutatakse esmakordselt kolleeg Peep Nemvaltsi ettepanekul terminit *keeletöötlus* ingliskeelse termini *processing* ja soomekeelse termini *prossointi* eestikeelse vastena.

Käesolevas artiklis on vaatluse all eestikeelsete keelekasutajate ja venekeelsete eesti keele õppijate hinnangud mõnede eesti kirjakeele ja õppijakeele sagedasemate sõnajärjemallide loomulikkusele. Eesti keel on antud uurimuses seega nii lähte- kui sihtkeele, vene keel ainult lähtekeele rollis. Nii eesti kui vene kuuluvad põhisõnajärje poolest SVO-, üldisemalt SVX-keelte hulka. Siiski erinevad eesti ja vene keel põhisõnajärje vabaduse astme poolest: eesti keel on määratletud kui paindlik, vene keel kui väga paindlik SVO-keel (Koptjevskaja-Tamm, Wälchli 2001). Mõlema keele sõnajärje aluseks on pigem infostruktuuri kui süntaktilised printsüübid, mille kohaselt lause algusesse paigutub teema ja lõppu fookus (Tael 1988: 40, Štšadneva 2002). Iga konkreetse lause puhul mõjutavad sõnajärje valikut lisaks infostruktuurile ka süntaktiline raskus, tekstitüüp ja afekteeritus, kusjuures ei ole kindlasti tegemist lõpliku mõjurite loendiga (Remmel 1963: 217, Lindström 2005: 22–35).

Erinevalt vene keelest määrab eesti keeles subjekti paiknemise lauses süntaktilise piiranguna paljuski V2-reegel, mis on pragmaatiliste tendentside suhtes primaarne ja piiritleb nende toimimise raamid (Huumo 1995: 51). Eriti märgatav on see tendents eesti ilukirjanduskeeles, suulises kõnes seevastu on subjekti ja verbi otsejärg tunduvalt sagedasem (Lindström 2005: 18–19). Aluse ja öeldise vastastikust asendit võivad potentsiaalselt mõjutada esimese lauseliikme süntaktiline funktsioon, lausetüüp, subjektifraasi semantilised ja grammatilised ning infostruktuurilised omadused (Lindström 2005: 90). Otsejärge eelistatakse suulises eesti keeles eriti juhul, kui lause algab sihitise, öeldistäite, põhjus- või viisimäärusega või kui subjektifraasi moodustab (peamiselt 1. ja 2. isikule viitav) pronoomen. Kui lause algab valdajamäärusega või subjektiks on täisnimisõnafraas, kasutatakse lauses sagedamini subjekti ja verbi pöördjärge. (Lindström 2005: 173–174) Kuigi kohamäärusega algava lause puhul on sagedamini tegemist pöördjärjega, ei ole see tegur määrav pöörd- või otsejärje valikul; samuti ei mõjuta pöörd- või otsejärje valikut ajamäärus lause alguses (Lindström 2005: 91). Küll aga mõjutab pöörd- või otsejärje valikut subjektifraasi elusus või elutus: ilukirjanduskeeles on 67,4% otsejärjega lausetest elusale viitava subjektiga ja 42,8% pöördjärjega lausetest elutule viitava subjektiga (Huumo 1995: 85); suulises kõnes vastavalt 66,5 ja 83,4% (Lindström 2005: 95).

Õppijakeele analüüsil tuleb arvesse võtta nii kirjaliku kui suulise keele sõnajärje spetsiifikat, sest õppijate kirjalike tekstide sõnajärg on sageli suulisest kõnest mõjutatud. Eri registritest pärit sisendi eristamine on õppijatele tavaliselt kuni kõrgeimate keeletasemeteni tõsine väljakutse ja nii kalduvad kirjaliku ja suulise keele erijooned EVKK tekstides omavahel põimuma. Õppijakeele tekstid erinevad ilukirjandustekstidest ka lausete pikkuse ja keerukuse poolest. Seetõttu on vealeidja loomisel esmatähtis lause predikaatne kese ning verbilaiendite ja täiendite rohkus ei mängi nii olulist rolli.

3. Õppijakeele korpuspõhine veaanalüüs ja vigade psühholingvistiline reaalsus

Andrus Saareste (1949: 46, viidatud Remmel 1963: 216 järgi) sõnul on lauseliikmete järjestus üks neid keeleteaduse valdkondi, mis allub keerulistele psühholoogilistele teguritele, ja võimatu oleks kujutleda, et suudetakse selgitada kõiki äärmisi peensusi, mis esinevad väljendusviisi keerukais normides. Sõnajärjevigade prioriteetsus sõltub otseselt nende *psühholingvistilisest reaalsusest*, mis on vigade klassifitseerimine ja interpreteerimise kõrval nii traditsioonilise kui ka korpuspõhise veaanalüüsi peamisi probleeme (Kaivapalu 2009). Keelendite (sh keelevigade) psühholingvistilise reaalsuse all mõistetakse keelekasutaja või -õppija arusaama keelendi loomulikkusest või ebaloomulikkusest, normikohasusest või normivastasusest; keelte võrdluse korral keeltevahelistest sarnasustest ja erinevustest (Kaivapalu 2004, Sajavaara 2006). Ligikaudu samas tähenduses on kasutatud ka termineid *psühhotüpoloogia* (Eckman 2004: 517, Ribbert, ten Thije 2007: 78) ning *keeltevaheline tajutus ja oletatav sarnasus* (Ringbom 2007).

Vigade psühholingvistilise reaalsuse varieerumine on tingitud lingvistiliste, psühholoogiliste ja sotsiaalsete tegurite omavahelisest põimumisest keele omandamisest ja kasutamisel. Põhiküsimused on seejuures, kas ja kuidas viga tunnetatakse, kelle arvates ja mille suhtes on tegemist veaga. Nii nagu võivad erineda emakeelse keelekasutaja ja keeleõppija veatunnetus, võib vigade psühholingvistiline reaalsus sõltuda vanusest, soost ja haridustasemest. Samuti ei tarvitse kirjakeele normi vastane keelend olla seda suulise keele normi või tegeliku keelekasutuse seisukohalt.

EVKK-s on keeleviga määratletud grammatikareeglile või kommunikatiivsele eesmärgile mittevastava keelekasutusena, mille hulka ei kuulu väsimusest ja hooletusest põhjustatud eksimused ning keelevääratused (Eslon 2007: 106). Grammatikareeglile või kommunikatiivsele eesmärgile vastavuse tajumine võib aga varieeruda isegi sarnase struktuuriga lausete puhul, nagu ilmneb näidetes (1)–(3) esitatud sõnajärjevigade puhul.

- (1) 3 aastat tagasi **ma lõpetasin** 10 klassi Narva Hummanitaar gümnaasiumis.
- (2) Juba koolis **ma unistasin**, et tulevikus lähen TTÜ-sse õppima.
- (3) Üks kord nädalas **tüdruk sõidab** linna, et käia kinos.

Tegemist on lausetega, milles kõigis on rikutud V2-reeglit: iseseisvas jaatavas väitelause on öeldisverb (täpsemalt öeldisverbi pöördeline osa) tavaliselt teisel kohal, teemana toimiva moodustaja järel (EKK 2007: 525). Hoolimata lausestruktuuri sarnasusest suhtusid märgendajad sõnajärjevigadesse erinevalt: kui esimeses lauses märgendasid normivastase sõnajärje määrus-alus-öeldis sõnajärjeveaks kõik kolm märgendajat, siis teises lauses kaks ja kolmandas lauses vaid üks märgendaja. Selline tulemus on seda üllatavam, et lausetes (1) ja (2) on aluseks rõhutu asesõna, mille puhul peetakse otsejärke suhteliselt tavaliseks (EKK 2007: 526). Nii tekibki küsimus, millest sõltuvad märgenduseelistused, ehk teisisõnu, kas kirjakeele normile vastav V2-reegel on psühholingvistiliselt reaalne ka tegeliku keelekasutuse seisukohalt. Lihtsaim võimalus uurida, milline sõnajärjemall on emakeelse keelekasutaja

või keeleõppija jaoks keeleomane ja loomulik, on küsida seda keelekasutajalt või -õppijalt endalt.

4. Ainestik ja meetod

Sõnajärjemallide psühholingvistilise reaalsuse väljaselgitamiseks kasutati tajutesti, milles tuli hinnata keelendi loomulikkust neljaastmelisel Likert-tüüpi skaalal *täiesti loomulik – pigem loomulik – pigem ebaloomulik – täiesti ebaloomulik*. Tegemist oli niisiis kategoriaalse, mittevõrdleva skaala tüübiga, mis määrab ära vastajate nõustumise astme teatud väidetega, mis on seotud mingi hoiaku hindamise või mõõtmisega.³ Uurimuses osalejatel paluti vastata nii, nagu nende enda keeletunne neile ütleb, ega mõelda sellele, kas vastused on õigekeelsuse seisukohast õiged või valed, s.t lähtuda oma keeletundest, mitte aga keelendite normikohasusest. Testi juhendis rõhutati, et testitavad hindaksid just sõnajärje, mitte aga kogu lause loomulikkust. Siiski ei ole ilma introspektiivse ainestikuta võimalik kindlalt väita, kuivõrd suutsid testitavad oma hinnangutes keskenduda puhtalt sõnajärjele, jättes tähelepanu alt välja lause kui terviku. Uurimuses osalesid Tallinna Ülikooli 19–37 aasta vanused üliõpilased. Osalejate hulgas oli 50 eestikeelset keelekasutajat ja 50 venekeelset eesti keele õppijat, kes olid koolis õppinud eesti keelt 6–12 aastat. Uurimuses ei kasutatud nende venekeelsete informantide vastuseid, kes olid õppinud eestikeelses koolis või kelle kodune keel oli eesti keel.

Testilauseste koostamisel oli aluseks korpusest tulenev lähenemisviis (ingl *corpus-driven approach*). Erinevalt korpuspõhisest analüüsist (*corpus-based language analysis*), mille puhul kasutatakse korpusainestikku allikmaterjalina keelekirjelduste ja teooriate paikapidavuse tõendamisel, tõlgendatakse korpusest tuleneva lähenemisviisi korral korpust keeleainesena, millel uurija ennast juhtida laseb (Tognini Bonelli 2002: 73–75). Seetõttu moodustati testilauseid pilootkorpuse neljateistkümne sõnajärjeviga sisaldava lause põhjal, mille osas kolm märgendajat olid üksmeelsed (näited 4–17).

- (4) Aasta jooksul nad kasvavad põrsast, et Jõulupüha eel tappa teda ja ära süüa.
- (5) Ta alati riietub pidulikult, nagu väike nukk.
- (6) Põhiteema mis oli käsitud peolauas oli elu planeedis Marsil tõenäosus.
- (7) Zaura kinkis mulle siili, tehtud paberist.
- (8) Energia on vaja inimestele.
- (9) Oma huvidest võin määrata sport, kuid seoses seega, et minu elu pragu näeb välja järgmiselt.
- (10) TTÜ kohta võin öelda, et minu arvates see on ainuke ülikool, mis üldse välja näeb nagu ülikool: suured auditooriumid, laboratooriumid j.n.e.
- (11) Muidugi, elamises õpilaskodus on ka oma miinused.
- (12) 3 aastat tagasi ma lõpetasin 10 klassi Narva Humanitaar gümnaasiumis.

- (13) Pärast seda ma koolisin Peterburis, kus ma õppisin kolidžis.
- (14) Aga oli ka palju häid õpetajad, kes armastasid lapsi, oma tööd ja põnevalt jutustasid oma õppeainet, just need õpetajad jäävad mäalu.
- (15) Pärast kooli ma sain hea keeleoskuse, ja humanitaarained, mis väga aitas mind järgmises koolituses.
- (16) Ja et aru saada, millest räägitakse ma pean sügavalt õppima eesti keel juba praegu.
- (17) Esimest korda toimunud aastal 1869 laulupeod muutusid ajaloolise tähtsusega sündmusteks ja olid väga populaarsed eestlaste seas.

Korpusest tuleneva lähenemisviisi korral on eelduseks, et ainestikku kasutatakse algupärasel kujul. Seetõttu säilitati testilauseid koostades maksimaalselt korpuse lausete struktuur ega muudetud testilauseite ja lause moodustajate ehitust. Samuti lähtuti testilauseite koostamisel nendest lausetüüpideist, mida vealeidja vaatleb (vt lähemalt Metslang, Matsak 2010): neutraalse sõnajärjega isikulises tegumoes jaatavas kõneliigis väit- ja käsklauseid; lihtlauseid või liitlause pealauseid, kui need on vasakult esimesed osalauseid; ühend- ja väljendverbidega laused. Vealeidja ei vaatle kõrvallauseid, sh vasakpoolsed osalauseid, mis algavad: *et, kuna, nagu, kui, siis kui* jne; umbisikulises tegumoes lauseid; lauseid, mis ei sisalda finiiitverbi, ega küsilauseid. Kui pilootkorpuse sõnajärjeviga esines kõrvallauses, moodustati sellest tajutesti tarvis elementaarlause.

Õppija vigasest lausest moodustati “Eesti keele käsiraamatu” (EKK) alusel normikohane lause nagu lause (12) alusel lause (18).

- (18) Kolm aastat tagasi lõpetasin ma kümnenda klassi Narva Humanitaar-gümnaasiumis.

Normikohase lause sõnajärge muudeti nii lause moodustajate ümberpaigutamise teel kui ka sõnavara varieerimise teel. Igast sõnajärjeviga lause tüübist moodustati sel viisil veel 2–4 varianti, näiteks lause (12) põhjal variandid (19)–(20).

- (19) Mõne aasta pärast alustavad nad magistriõpinguid välismaal.
- (20) Ta kirjutas lõpukirjandi paari aasta eest oma kodukoolis.

Koos pilootkorpuse õppijakeele lausega võis variantide arv niisiis olla 3–5. Lausete vastavus kirjakeele normile oli varieeruv. Kokku esitati testis juhuslikus järjekorras 50 kontekstivaba elementaarlauseit (lähemalt testist vt Kaivapalu 2009). Käesolevas artiklis tuleb 50 testilausest vaatluse alla 16, mis on käsitletavad nelja sõnajärjemallina. Iga malli puhul tuginetakse neljale sama põhistruktuuriga testilausele, mis erinevad kohati verbilaiendite liigi ja moodustajate struktuuri poolest. Valiku aluseks oli verbi ja selle laiendite paiknemine lauses: põhitähelepanu keskendus V2-sõnajärjele. Ülejäänud tegurite varieeruvust püüti minimaliseerida niivõrd, kui korpusest tulenev lähenemisviis seda võimaldas. Kõik artiklis vaatluse alla tulevad laused on normaallauseid, lausealguliseks verbieliseks laiendiks on ajaadverbiaal, subjektiks enamasti personaalpronoomen, verbijärgseteks laienditeks objekt või aja- või kohaadverbiaal või mõlemad. Artiklis käsitletavaid lauseid on lähemalt

kirjeldatud peatükis 5. Sõnajärjemallide ülesmärkimisel on Kaja Taela (1988: 5) eeskujul kasutatud tähiseid S – subjekt, V – verb ja X – verbi laiend, kuid erinevalt mainitust on kõrvuti asetsevad lause kõrvalliikmed eraldi tähistatud. Üksiklausetehistamisel on täpsustatud ka, kas lause kõrvalliikme puhul on tegemist objekti (O) või adverbiaaliga (A). Sarnase struktuuriga lauseteh puhul on kasutatud täiendavaid tähiseid 1 ja 2.

5. Sõnajärjemallide psühholingvistiline reaalsus

Eesti sõnajärjemallide sageduses ilmneb vaba sõnajärje printsiip väga selgesti. Kaks statistiliselt kõige sagedasemat sõnajärjemalli esinevad neutraalseks peetud populaarteaduslikes ja ajakirjanduslikes tekstides peaaegu võrdse sagedusega: otsejärje SVX esinemissagedus on 25% ja pöördmalli XVS⁴ esinemissagedus 24% (Tael 1988: 6). Kui SVX-malli esinemissagedus on enam-vähem sama (23...25% kõigist lausetest) peaaegu kõigis allkeeltes, siis kirjakeele kaht sagedast malli XVS ja XVSX esineb teistes allkeeltes tunduvalt vähem, 1...11% kõigist lausetest (Lindström 2005: 194). Kuna markeerituse seisukohalt on kõige neutraalsemaks teemaks subjekt, on eesti keele põhisõnajärjeks⁵ peetud SVX-järge (Ehala 2001: 24). Samuti vastab SVX-mall kõige paremini infostruktuuri printsiipidele (Tael 1988: 40).

Õppijakeeles sagedasti esinevad sõnajärjemallid XSVX ja SXVX ei ole eesti keeles kuigi sagedased: XSV-lauseid esineb ilukirjanduslikes proosatekstides 2%, XSVX-lauseid alla 1%, SXV-lauseid 3% (Tael 1988: 6). Suulises kõnes on otsejärje sagedus pisut suurem: XSV-malli esineb 6,1% ja XVSX-malli 3,2%. SXV-malli esinemissagedus on suulises kõnes 3,8% ja SXVX-malli 1,1%. (Lindström 2005: 18–19)

Sihtkeele keelendite esinemissagedus on keele omandamise ja keeletaju kujunemise seisukohalt oluline: esmajoones omandatakse sagedamini esinevad keelendid, neid kasutatakse enam ja ka tajutakse loomulikumana. Samuti mõjutab õppija keeletaju muude tegurite hulgas emakeel. Ka vene keeles on kõige neutraalsem SVX-mall, mida kasutatakse eelkõige ametlikus ja teadusstiilis, kui eesmärgiks on informatsiooni võimalikult täpne ja tõhus vahendamine. Pöördjärg on omane esmajoones ekspressiivsele stiilile, sageli on sellel kõnekeelne või rahvaluulepärase varjund. Sellisel juhul on lauserõhul oluline roll. (Štšadneva 2002)

Alljärgnevalt kirjeldatakse eestikeelsete keelekasutajate ja venekeelsete eesti keele õppijate hinnanguid õppijakeeles esinevate sõnajärjemallide SVXX(X), XVSX(X), XSVX(X) ja SXVX(X) loomulikkusele. Et esinevaid suundumusi selgemini välja tuua, on hinnangutepaari *täiesti loomulik* ja *pigem loomulik*, samuti hinnanguid *täiesti ebaloomulik* ja *pigem ebaloomulik* käsitletud koos.

5.1. Sõnajärjemall SVXX(X)

Sõnajärjemalli SVXX(X) esindasid testis neli normaallauset (näited 21–24).

(21) Nad õpivad terve nädala matemaatikat (SVAO)

(22) Nad lendasid mõne aja pärast Indiasse (SVAA)

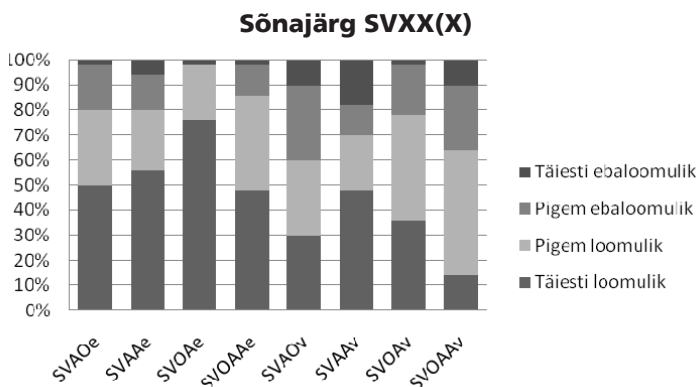
⁴ Taela (1988: 6) käsitluses on kõik pöördjärje allmallid koondatud malliks XVS.

⁵ Põhisõnajärje ja baassõnajärje erinevustest vt Ehala 2001: 24–25.

(23) Ma saan uue võimaluse alles pärast vaheaega (SVOA)

(24) Ta kirjutas lõpukirjandi paari aasta eest Tallinnas (SVOAA)

Testilauseid on ühesugused nii infostruktuuri kui ka subjekti süntaktilise raskuse (asesõna) poolest, erinevus seisneb verbijärgsete laiendite liigis ja järjekorras. Lauses (21) järgnevad verbile ajaadverbiaal ja objekt, lauses (22) aja- ja kohaadverbiaal, lauses (23) objekt ja ajaadverbiaal, lauses (24) objekt, aja- ja kohaadverbiaal. Eestikeelsete keeleteadajate ja venekeelsete eesti keele õppijate hinnangud lausetel (21–24) loomulikkusele on esitatud joonisel 1.



Joonis 1. Sõnajärjemalli SVXX(X) psühholingvistiline reaalsus

Jooniselt ilmneb, et statistiliselt sagedaseim sõnajärjemall on eestikeelsete keeleteadajate jaoks ka psühholingvistiliselt reaalne: kõiki nelja testilauseid on vähemalt 80% vastajatest hinnanud loomulikuks või pigem loomulikuks. Ka venekeelsete eesti keele õppijate jaoks on SVXX(X)-mall valdavalt loomulik või pigem loomulik. Siiski on eestikeelsete keeleteadajatega võrreldes veidi rohkem neid venekeelseid vastajaid, kes peavad antud sõnajärge pigem ebaloomulikuks või täiesti ebaloomulikuks.

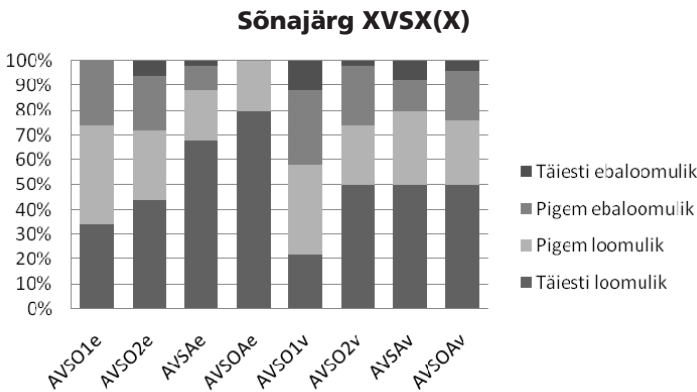
Kui võrrelda sõnajärjemalli SVXX(X) testilauseid omavahel, selgub, et nii eestikeelsete keeleteadajate kui venekeelsete õppijate arvates on teistest loomulikum lause (23), milles verbile järgneb objekt ja siis ajaadverbiaal. Tegemist on ka lausega, milles ainsana on objekt mitmesõnaline. Väike erinevus ilmneb eestikeelsete keeleteadajate ja venekeelsete õppijate testilauseid loomulikkuse järjekorras. Kui eestikeelsete keeleteadajate jaoks järjestuvad testilauseid loomulikkuse alusel SVOA, SVOAA ning seejärel võrdselt SVAO ja SVAA, siis õppijate järjestus on SVOA, SVAA, SVOAA ja SVAO. Seega kaldub eestikeelsete keeleteadajate jaoks loomulikum olema objekti paiknemine verbi järel, venekeelsete keeleteadajate jaoks aga lause lõpetamine adverbialiga.

5.2. Sõnajärjemall XVSX(X)

SVXX(X)-sõnajärjega esinemissageduselt peaaegu võrdset malli XVSX(X) järgivaid testilauseid oli samuti neli (näited 25–28).

- (25) Pool päeva lugesime me kriminulle (AVSO1)
- (26) Tehnikumi järel sai ta hea töökoha (AVSO2)
- (27) Ülikooli lõpetamise järel sõitis Henrik Londonisse (AVSA)
- (28) Mõne aasta pärast alustavad nad magistriõpinguid välismaal (AVSOA)

Kaks esimest lauset, (25) ja (26), on samasuguse ehitusega, kuid erinevad objekti fraasistruktuuri poolest: lauses (25) on objektiks laiendamata, lauses (26) laiendatud noomenifraas. Lause (27) erineb ülejäänutest subjektifraasi ehituse poolest: asesõnalise subjekti asemel on subjektiks nimi. Nii nagu ilukirjanduskeeles (Huumo 1994: 280–285) on ka õppijakeele XVSX(X)-lausetes lause alguses enamasti adverbiaal, antud juhul ajaadverbiaal. Lauseid (27) ja (28) erinevad lausetest (25) ja (26) lausealgulise ajaadverbiaali pikkuse poolest. Lauseid (27) ja (28) lõppevad kohaadverbiaaliga, mis lauses (27) paikneb subjekti, lauses (28) aga subjekti ja objekti järel. Eestikeelsete keekekasutajate ja venekeelsete õppijate hinnangud XVSX(X)-malli loomulikkusele on esitatud joonisel 2.



Joonis 2. Sõnajärjemalli XVSX(X) psühholingvistiline reaalsus

Pöördjärg kaldub niisiis olema eestikeelsete keekekasutajate jaoks samuti psühholingvistiliselt reaalne: hinnangud malli XVSX(X) loomulikkusele ei erine märkimisväärselt hinnangutest SVXX(X)-malli kohta. Ka suurem osa venekeelsetest eesti keele õppijatest tajus pöördõnajärge loomulikuna, kuid venekeelsete õppijate jaoks oli vahe pigem loomuliku ja pigem ebaloomuliku vahel pisut väiksem kui emakeelsetel keekekasutajatel.

Võrreldes hinnanguid kahe sama struktuuriga (AVSO) lause (24) ja (25) kohta, on näha, et eestikeelsed keekekasutajad tajusid neid üsna ühtmoodi: hinnangute *täiesti loomulik* ja *pigem loomulik* summa erines vaid 2% võrra. Lauseid (24) hinnati siiski enam pigem loomulikuks, lauset (25) enam täiesti loomulikuks. Venekeelsete õppijate puhul oli vahe kahe lause vahel suurem: lauset (25) hinnati loomulikumaks. Viimane asjaolu on tõenäoliselt põhjustatud mitmuse esimese pöörde lõpu ja perso-

naalpronoomeni järjestikusest asetusest, mis eestikeelseid keelekasutajaid vähem häiris. Kõige loomulikumaks hindasid eestikeelsed keelekasutajad kahe verbijärgse laiendiga, objekti ja adverbiaaliga lauset: kõikide eestikeelsete keelekasutajate jaoks oli see kas täiesti loomulik või pigem loomulik. Venekeelsete õppijate jaoks olid teistest loomulikumad adverbiaaliga lõppevad laused, kuigi vahe objektiga lõppeva lausega (25) ei olnud suur.

Niisiis ei erinenud oluliselt ei eestikeelsete keelekasutajate ega venekeelsete õppijate hinnangud kahele eesti keeles sagedamini esinevale sõnajärjemallile.

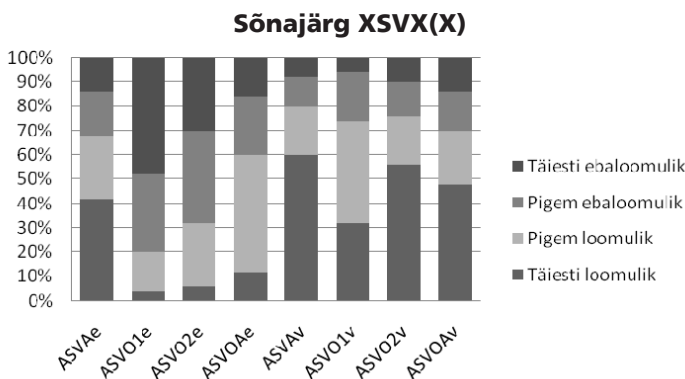
5.3. Sõnajärjemall XSVX(X)

Õppijakeeles sagedasti esinev sõnajärjemall XSVX(X) esindab otsejärge, mis esineb tavapärase pöördjärje XVS asemel. Enamuses pöördjärjega mallides on tegemist inversiooniga, mille puudumisel, eeldades, et topik jääb samaks, oleks lause vastuvõetamatu. Samas ei muutu kõik invertteeritud sõnajärjemallid inversiooni puudumisel eagrammatiliseks (Tael 1988: 10–11). Suhteliselt tavaliseks peetakse otsejärge juhul, kui subjektiks on rõhutu asesõna, samuti kasutatakse järjekindlalt otsejärge juhul, kui lause algab viisiadverbiaaliga (EKK 2007: 526). Õppijakeele otsejärjega lausetes esineski subjektina valdavalt rõhutu asesõna, topikuks oli ajaadverbiaal (laused 29–32).

- (29) Pärast seda ma kolisin Peterburi (ASVA)
- (30) Aasta jooksul nad kasvatavad pörsast (ASVO1)
- (31) Pärast kooli ma sain hea keeleoskuse (ASVO2)
- (32) Kolm aastat tagasi ma lõpetasin kümnenda klassi Narva Humanitaargümnaasiumis (ASVOA)

Lause (29) lõpeb adverbiaaliga, laused (30) ja (31) objektiga, kusjuures lauses (30) on objektiks laiendamata, lauses (31) laiendatud noomenifraas. Lauses (32) järgnevad verbile nii objekt kui adverbiaal, samuti on verbielne ajaadverbiaal pikem kui teistes lausetes.

Eestikeelsete keelekasutajate hinnang invertteerimata lausete loomulikkusele oli lauseti üsna erinev (joonis 3).



Joonis 3. Sõnajärjemalli XSVX(X) pühholingvistiline reaalsus

Selgelt eristuvad ülejäänud lausetest mõlemad objektiga lõppevad laused (30) ja (31), mille invertteerimata sõnajärge hinnati teistest oluliselt ebaloomulikumaks. Eriti ebaloomulikuks osutus lause (30). Kõikides lausetes on subjektiks rõhutu määrsõna, kuid lauses (30) ei ole seda võimalik erinevalt ülejäänud lausetest ära jätta, ilma et tähendus muutuks. Pöördjärje asemel esinev otsejärg on seega eestikeelse keelekasutaja jaoks vähem häiriv juhul, kui verbi pöördeline vorm viitab 1. või 2. isikule, nii et asesõnalise subjekti puudumine lausest ei muudaks selle tähendust. Seevastu kohaadverbiaaliga lõppevaid lauseid (29) ja (32) tajuti vähemalt 60 protsendil juhtudest täiesti või pigem loomulikuna. Sõnajärjemalli XSVX(X) otsejärje loomulikkus näib siis sõltuvat muuhulgas verbilaiendite arvust ja liigist: adverbialiga lõppevad laused osutusid eestikeelsete keelekasutajate jaoks loomulikumaks kui objektiga lõppevad laused. Lisaks lause viimase moodustaja lauseliikmelisele funktsioonile võib rolli mängida ka asjaolu, et adverbialiga lõppevad laused (29) ja (32) on sisu poolest loomulikumad ja terviklikumad kui objektiga lõppevad laused (30) ja (31). Täiesti võimalik, et eestikeelsete keelekasutajate hinnanguid on mõjutanud nende lausete kui tervikute ebaloomulikkus: on ju raske pidada sõnajärge loomulikuks, kui lauset tervikuna tajutakse ebaloomulikuna.

Seevastu venekeelsete õppijate jaoks on XSVX(X)-mall loomulik (joonis 3): lauseid (29)–(32) pidas täiesti või pigem loomulikuks 70–80% vastajatest, kusjuures lausetevahelised erinevused olid väikesed. Niisiis võib järeldada, et venekeelsete õppijate jaoks ei ole laiendiga, antud juhul adverbialiga algava lause subjekti ja verbi järjekord distinktiivne: ühtviisi loomulikuks peeti nii XVSX(X)- kui XSVX(X)-malli.

5.4. Sõnajärjemall SXVX(X)

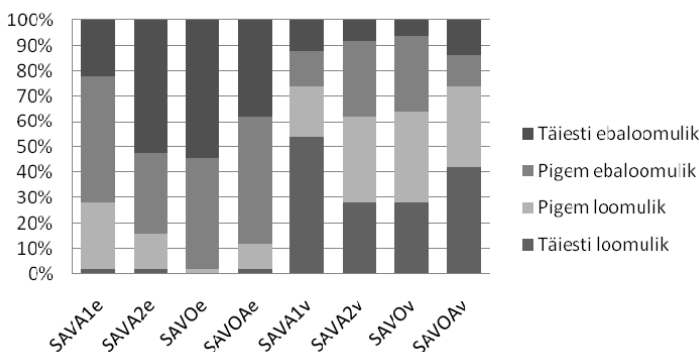
Kui eesti kirjakeeles on SXV-malli subjekti ja verbi vaheline laiend enamasti objekt, siis õppijakeele lausemallis tähistab X adverbial (näited 33–36).

- (33) Ta alati riietub pidulikult (SAVA1)
- (34) Mu elu praegu näeb välja järgmiselt (SAVA2)
- (35) Nad põnevalt jutustasid oma õppeainet (SAVO)
- (36) See väga aitas mind järgmistes koolitustes (SAVOA)

Lausetes (33) ja (35) on subjektiks personaalpronoomen, lauses (36) demonstratiivpronoomen, lauses (34) laiendatud noomenifraas. Lause (34) erineb ülejäänutest ka selle poolest, et predikaadiks on ühendverb, mille adverbialne komponent ei asetse päris lause lõpus, vaid eespool. Lausetes (33) ja (34) järgneb verbile viisiadverbiaal, lausetes (35) ja (36) objekt, kusjuures lauses (36) järgneb objektile veel adverbial. Eestikeelsete keelekasutajate ja venekeelsete õppijate hinnangud sõnajärjemalli SXVX(X) loomulikkusele on esitatud joonisel 4.

Tajutesti tulemustest ilmneb selgesti, et vaadeldud sõnajärjemallidest on SXVX(X)-mall eestikeelsete keelekasutajate jaoks kõige ebaloomulikum: vaid väike osa vastajatest oli hinnanud adverbialiga lõppevaid lauseid (33), (34) ja (36) pigem loomulikuks.

Sõnajärg SXVX(X)



Joonis 4. Sõnajärjemalli SXVX(X) psühholingvistiline reaalsus

Venekeelsete eesti keele õppijate vastused erinesid aga eestikeelsete keelekasutajate vastustest oluliselt. Venekeelsete õppijate keeletajus on SXVX(X)-mall üsna loomulik: vähemalt 60% vastajatest pidas lauseid (33–36) täiesti või pigem loomulikuks. Venekeelsete õppijate hinnang SXVX(X)-malli loomulikkusele ei erine seega kuigi-võrd hinnangutest SVXX(X)-, XVSX(X)- ja XSVX(X)-malli loomulikkusele. Antud asjaolu näitab veel kord subjekti ja verbi lauses paiknemise ebaolulisust venekeelse eesti keele õppija keeletaju seisukohalt.

Lausete omavahelisest võrdlusest selgub, et eestikeelsete keelekasutajate jaoks olid teistest loomulikumad adverbiaaliga lõppevad laused, kusjuures loomulikkust vähendas veidi nii laiendatud noomenifraas kui elutule viitav demonstratiivpronoomen subjektina. Objektiga lõpevat lauset aga tajuti ebaloomulikumana. Venekeelsete õppijate arvates olid ülejäänutest veidi loomulikumad need adverbiaaliga lõppevad laused, milles subjektiks pronoomen; laiendatud noomenifraas subjektina vähendas SXVX-lause loomulikkust.

6. Metodoloogilisi edasiarendusi

Käesolevas uurimuses kasutatud korpusest tulenev lähenemisviis võimaldas kaardistada õppijakeele sagedasemad ja olulisemad sõnajärjevead ning leida vastus küsimusele, millised nendest vigadest on emakeelse keelekasutaja keeletaju seisukohalt häirivamad ning millisena keeleõppija neid vigu tajub. Samas tekkis töö käigus mitmeid, peamiselt metodoloogilisi küsimusi, mis vajaksid edasist analüüsi ja lahendamist. Osa neist küsimustest tuleneb psühholingvistilise tajutesti iseärasustest, osa testilause struktuurist ja kontekstivabadusest.

Psühholingvistilise tajutesti puhul tekib esmalt küsimus, kuidas tagada, et uurimuses osalejad annaksid hinnangu just nimelt sõnajärje loomulikkusele, mitte aga lause loomulikkusele mõnest muust vaatenurgast. Lisaks selgitustele vahetult enne testi läbiviimist on seejuures abiks introspektiivsed meetodid: valjusti mõtlemine ja retrospektiivne intervjuu, mis annavad lisainformatsiooni uurimuses osalejate mõttekäigu ja hinnangute põhjenduste kohta (Kaivapalu 2009). Alles introspektiivsed meetodid võimaldavad välja selgitada emakeelsete keelekasutajate ja õppijate

hinnangute tegelikke põhjusi, vastasel juhul on tegemist pelgalt oletustega. Samuti on introspektiivsete meetodite abil võimalik selgitada, kuidas informandid tajuvad Likert-skaala kontiinumil *täiesti loomulik – pigem loomulik – pigem ebaloomulik – täiesti ebaloomulik*, ehk teisisõnu, kust läheb vastaja jaoks ligikaudne piir näiteks *pigem loomuliku ja pigem ebaloomuliku* vahel.

Teine rühm küsimusi on tingitud asjaolust, et õppijakeele korpuse tekstides esinevaid kontekstisiduseid lauseid esitati vealeidja loomisest tulenevalt tajutestis kontekstivabade üksiklausetena. Kuna lause sõnajärg sõltub eriti vasakpoolsest kontekstist, on edaspidi põhjust uurida, kas ja kuidas mõjutab kontekst emakeelsete keelekasutajate ja õppijate hinnanguid lause sõnajärje loomulikkuse kohta. Samuti oleks oluline selgitada, kas eesti ilukirjanduskeeles ja suulises kõnes lause sõnajärge mõjutavatel teguritel on oma osa ka emakeelsete keelekasutajate ja õppijate hinnangutes. Antud tajutulemustest selgus, et hinnangut sõnajärje loomulikkusele mõjutas muuhulgas, kas verbijärgne laiend on objekt või adverbiaal. Seegi asjaolu vajaks edaspidi põhjalikumat uurimist nagu ka fraasisisese sõnajärje psühholingvistiline reaalsus.

7. Järeldused ja väljund

Psühholingvistilise tajutesti tulemused näitasid niisiis, et eesti keele kaks statistiliselt kõige sagedasemat sõnajärjemalli SVXX(X) ja XVSX(X) kalduvad eestikeelsete keelekasutajate jaoks olema psühholingvistiliselt reaalsed: neid sõnajärjemalle tajuti enamasti täiesti või pigem loomulikuna, kusjuures mallisisene homogeensus oli küllalt suur. Samas tajusid eestikeelsed keelekasutajad inversiooni puudumist pöördjärjega lausetes lauseti üsna erinevalt. Otsejärje loomulikkus sõltus muu hulgas verbijärgse laiendi süntaktilisest funktsioonist: verbijärgse objektiga lõpevaid lauseid tajuti ebaloomulikumana kui adverbiaaliga lõpevaid lauseid. Enamasti väljendab V2-reegli järgimine siiski eestikeelse keelekasutaja jaoks loomulikku ning selle rikkumine pigem ebaloomulikku keelekasutust. Selgelt kõige ebaloomulikum oli eestikeelsete keelekasutajate jaoks SXVX(X)-mall.

Venekeelsete eesti keele õppijate keeletaju seisukohalt ei ole subjekti ja verbi järjekord lauses aktsepteeritavaid ja mitteaktsepteeritavaid lauseid eristav tegur. V2-reegli rikkumine on venekeelsete keelekasutajate jaoks pigem loomulik kui ebaloomulik ning sõnajärjemalle SVXX(X), XVSX(X), XSVX(X) ja SXVX(X) tajuti üldjoontes ühtviisi loomulikuna. Seega ei taju keeleõppijad piisavalt nende sõnajärjemallide erinevust ning seetõttu on poolautomaatse vealeidja prototüübi väljatöötamisel tõepoolest põhjust keskenduda subjekti ja verbi paiknemisele lauses. Siiski ei ole kõik sõnajärjevead eestikeelse keelekasutaja jaoks võrdselt häirivad. Psühholingvistiline tajutest andis ülevaate sellest, kuidas sõnajärjemalle tegelikult tunnetatakse. See aga võimaldab määratleda prioriteetsed reeglid, millest vealeidja väljatöötamisel esmajoones lähtuda.

Lühendid

- A adverbiaal
O objekt
S subjekt
V verb
V2 reegel, mille põhjal paikneb verb lauses teisel kohal
X verbi laiend

Viidatud kirjandus

- Aavik, Johannes 1912. Kõige suurem germanismus eesti keeles. – Eesti Kirjandus, 9, 353–369.
- Abrahamsson, Nicklas; Hyltenstam, Kenneth 2009. Age of L2 acquisition and degree of nativelikeness listener perception vs. linguistic scrutiny. – *Language Learning*, 58, 249–306. doi:10.1111/j.1467-9922.2009.00507.x
- Biggert, Johnny; Kann, Viggo; Knutsson, Ola; Sjöbergh, Jonas 2005. Grammar checking for Swedish second language learners. <http://www.csc.kth.se/tcs/projects/xcheck/rapporter/bigkankusjoo5.pdf> (15.9.2009).
- Bohnacker, Ute 2006. When Swedes begin to learn German: from V2 to V2. – *Second Language Research*, 22 (4), 443–486. doi:10.1191/0267658306sr2750a
- Borin, Lars 2002. CrossCheck project status report. <http://www.csc.kth.se/tcs/projects/xcheck/lagesrapport-021015.pdf> (17.9.2009).
- Borin, Lars; Prütz, Klas 2004. New wine in old skins? A corpus investigation of L1 syntactic transfer in learner language. – Guy Aston, Sylvia Bernardini, Dominic Stewart (Eds.). *Corpora and Language Learners. Studies in Corpus Linguistics*, 17. Amsterdam, Philadelphia: John Benjamins, 67–87.
- Cedden, Gülay; Audin, Ozgür. “Meistens ich weiß nicht, wo ich muss das Verb gebrauchen.” Die Problematik des V2 Phenomens bei türkischen Lernenden der deutschen Sprache. – *Zeitschrift für Interkulturellen Fremdsprachenunterricht* / [Online] 12:3. http://zif.spz.tu-darmstadt.de/jg-12-3/beitrag/Ceden_Aydin.htm (4.12.2009).
- Eckman, Fred 2004. From phonemic differences to constrant rankings: Research on second language phonology. – *Studies in Second Language Acquisition*, 26, 514–539. doi:10.1017/S027226310404001X
- Ehala, Martin 2001. Eesti keele baassõnajärjest. – Reet Kasik (toim.). *Keele kannul: pühenusteos Mati Erelti 60. sünnipäevaks 12. märtsil 2001*. Tartu Ülikooli eesti keele õppetooli toimetised, 17. Tartu: Tartu Ülikooli Kirjastus, 24–41.
- EKK = Erelt, Mati; Erelt, Tiiu; Ross, Kristiina 2007. *Eesti keele käsiraamat*. Tallinn: Eesti Keele Sihtasutus.
- Ellis, Rod; Barkhuizen, Gary 2005. *Analyzing Learner Language*. Oxford University Press.
- Eslon, Pille 2007. Õppijakeelekorpused ja keeleõpe. – Pille Eslon (toim.). *Tallinna Ülikooli keelekorpusete optimaalsus, töötlemine ja kasutamine*. Tallinna Ülikooli eesti filoloogia osakonna toimetised, 9. Tallinn: TLÜ Kirjastus, 87–120.
- Granger, Sylviane 2002. A bird’s eye view of learner corpus research. – Sylviane Granger, Joseph Hung, Stephanie Petch-Tyson (Eds.). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. *Language Learning & Language Teaching*, 6. Amsterdam, Philadelphia: John Benjamins, 3–33.
- Huumo, Tuomas 1994. Kontrastiivinen tutkimus suomen ja viron sanajärjestyksestä. Lisen-siaatintyö. Turun yliopiston suomen ja yleisen kielitieteen laitos.
- Huumo, Tuomas 1995. Remarks on the Functions of word order in Finnish and Estonian. – Heikki Leskinen, Sándor Maticsák, Tõnu Seilenthal (toim.). *Congressus Octavus Internationalis Fenno-Ugristarum, Jyväskylä 10.–15.8.1995. Pars IV*. Jyväskylä: Moderatores, 47–51.

- Håkansson, Gisela 2001. Against full transfer: evidence from Swedish learners of German. – Lund University Department of Linguistics Working Papers, 48, 67–86.
- Håkansson, Gisela; Pienemann, Manfred; Sayehli, Susan 2002. Transfer and typological proximity in the context of second language processing. – *Second Language Research*, 18 (3), 250–273. doi:10.1191/0267658302sr2060a
- Hyltenstam, Kenneth; Abrahamsson Niclas 2003. Age of onset and ultimate attainment in near-native speakers of Swedish. – Kari Fraurud, Kenneth Hyltenstam (Eds.). *Multilingualism in Global and Local Perspectives. Selected Papers from the 8th Nordic Conference on Bilingualism*, November 1–3, 2001, Stockholm Rinkeby. Stockholm: Centre for Research on Bilingualism, Stockholm University, and Rinkeby Institute of Multilingual Research, 319–340.
- Jarvis, Scott 2000. Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. – *Language Learning*, 50 (2), 245–309. doi:10.1111/0023-8333.00118
- Kaivapalu, Annekatrin 2004. Kui sarnane on sarnane? Eesti ja soome mitmusevormide psühholingvistilisest reaalsusest. – Helena Sulkala, Heli Laanekask (toim.). *VIRSU II. Suomi ja viro kohdekielinä. Oulun yliopiston suomen ja saamen kielen ja logopedian laitoksen julkaisuja*, 24. Oulu: Oulun yliopisto, 62–71.
- Kaivapalu, Annekatrin 2008. Lähtekeele mõju korpuspõhine uurimine. – Pille Eslon (toim.). *Õppijakeele analüüs: võimalused, probleemid, vajadused*. Tallinna Ülikooli eesti filoloogia osakonna toimetised, 10. Tallinn: TLÜ Kirjastus, 93–119.
- Kaivapalu, Annekatrin 2009. Õppijakeele korpusanalüüsi täiendavatest meetoditest. – Pille Eslon, Katre Õim (toim.). *Korpusuuringute metodoloogia ja märgendamise probleemid*. Tallinna Ülikooli eesti filoloogia osakonna toimetised, 11. Tallinn: TLÜ Kirjastus, 72–98.
- Kann, Viggo 2002. CrossCheck – a grammar checker for second language writers of Swedish. <http://www.csc.kth.se/tcs/projects/xcheck/ansokano2-1.pdf> (15.9.2009).
- Koptjevskaja-Tamm, Maria; Wälchli, Bernhard 2001. The Circum-Baltic languages: An areal-typological approach. – Östen Dahl, Maria Koptjevskaja-Tamm (Eds.). *Circum-Baltic Languages. Vol. 2. Grammar and Typology. Studies in Language Companion Series*, 55. Amsterdam, Philadelphia: John Benjamins, 615–750.
- Lindström, Liina 2005. Finiitverbi asend lauses. Sõnajärg ja seda mõjutavad tegurid suulises eesti keeles. *Dissertationes philologiae estonicae Universitatis Tartuensis*, 16. Tartu: Tartu Ülikooli Kirjastus.
- Metslang, Helena; Matsak, Erika 2010. Kesksete lausekomponentide järjestus õppijakeeles: arvutianalüüsi katse. (Käesolevas kogumikus.)
- Müürisep, Kaili 2000. Eesti keele arvutigrammatika: süntaks. Tartu: Tartu Ülikooli Kirjastus.
- Pienemann, Manfred 1998. *Language Processing and Second Language Development: Processability Theory*. *Studies in Bilingualism*, 15. Amsterdam: John Benjamins.
- Rommel, Nikolai 1963. Sõnajärjestus eesti lauses. – Eesti keele süntaksi küsimusi. Keele ja Kirjanduse Instituudi uurimused, VIII. Tallinn: Eesti Riiklik Kirjastus, 216–389.
- Ribbert, Anne; ten Thije, Jan D. 2007. Receptive multilingualism in Dutch-German intercultural team cooperation. – Jan D. ten Thije, Ludger Zeevaert (Eds.). *Receptive Multilingualism. Linguistic Analyses, Language Policies and Didactic Concepts*. *Hamburg Studies on Multilingualism*, 6. Amsterdam, Philadelphia: John Benjamins, 73–101.
- Ringbom, Håkan 2007. *Cross-linguistic Similarity in Foreign Language Learning*. Clevedon: Multilingual Matters LTD.
- Saareste, Andrus 1949. *Sur la Place du Verbe Attributif dans la Proposition Subordonnée Estonienne*. *Aphoporeta Tartuensia*. Stockholm.

- Sajavaara, Kari 2006. Kontrastiivinen analyysi, transfer ja toisen kielen oppiminen. – Annekatriin Kaivapalu, Külvil Pruuli (toim.). Lähivertailuja, 17. Jyväskylä Studies in Humanities, 53. Jyväskylä: Jyväskylän yliopisto, 9–26.
- Štšadneva 2002 = Щаднева Валентина 2002. О «свободном» порядке слов в русском языке. – Труды по русской и славянской филологии. Лингвистика. Новая серия, 8. Tartu: Tartu Ülikooli Kirjastus, 306–321
- Tael, Kaja 1988. Sõnajärjemallid eesti keeles (võrrelduna soome keelega). Preprint KKI-56. Tallinn: Eesti NSV Teaduste Akadeemia Keele ja Kirjanduse Instituut.
- Tognini Bonelli, Elena 2002. Functionally complete units of meaning across English and Italian: Towards a corpus-driven approach. – Bengt Altenberg, Sylviane Granger (Eds.). Lexis in Contrast. Corpus-based Approaches. Studies in Corpus Linguistics, 7. Philadelphia: John Benjamins, 73–95.

Võrgumaterjalid

EVKK = Eesti vahekeele korpus. <http://evkk.tlu.ee> (4.03.2010).

Mõistete sõnastik. TNS Emor. <http://www.emor.ee/index.html?id=432> (4.03.2010).

Annekatriin Kaivapalu (Tallinna Ülikool) on uurinud kõrvutavalt eesti ja soome muutemorfoloogiat, eesti ja soome üliõpilaste arvamusi emakeelest, soome ja eesti õppijakeelt, lähtekeele mõju sihtkeele morfoloogia omandamisele, absoluutset ja suhtelist morfoloogilist kompleksust, keelendite psühholingvistilist reaalsust; käsitlenud korpuspõhise õppijakeele uurimise võimalusi ja piiranguid. kaivapa@tlu.ee

ON PSYCHOLINGUISTIC REALITY OF SOME WORD ORDER PATTERNS OF ESTONIAN

Annekatrin Kaivapalu

Tallinn University

One of the aims of the project VAKO (*Development of language software and language technology resources for the Estonian Interlanguage Corpus*) is to create the grammar checker prototype for second language writers of Estonian. According to the error statistics of the Estonian Interlanguage Corpus the most frequent errors are word order errors. Estonian is regarded as a free word order language, as usually more than one word order pattern is acceptable. Although the tagging of the word order errors in EIC is based on the norms of Standard Estonian, different annotators can sometimes still interpret the errors differently. At the same time the real use of word order by native speakers of Estonian is not always in accordance with the norms of Standard Estonian. From the perspective of teaching Estonian as a second language it is of crucial importance for the learner language to be as similar to the language use of native speakers of Estonian as possible.

The aim of this paper is to explore the psycholinguistic reality of some word order patterns of Estonian from the point of view of Russian learners of Estonian as well as from that of native Estonian speakers. The perception of the following word order patterns of Estonian is discussed: 1) SVXX(X), 2) XVSX(X), 3) XSVX(X), 4) SXVX(X) where S stands for subject, V for verb, and X for any complement or modifier other than subject.

The study concludes that the main difference in the word order perception of Estonian native speakers and Russian learners of Estonian concerns the V2-rule. The pattern XVSX(X) was perceived as natural by most Estonian native speakers and the pattern XSVX(X) as rather unnatural. The most unnatural pattern for Estonian native speakers was SXVX(X). The Russian speakers of Estonian did not perceive either order of subject and verb as distinctive but considered the patterns SVXX(X), XVSX(X), XSVX(X) and SXVXX as equally natural.

Keywords: learner language, corpus-based error analysis, word order, native speaker, Estonian, Russian

VÄRVINGUPARTIKLID EESTI- JA SAKSAKEELSETES KÜSILAUSES

Janika Kärk

Ülevaade. Artiklis käsitletakse sagedasemaid eesti- ja saksakeelsetes küsilauses esinevaid värvingupartikleid. Näidenditekidest pärit küsilauses korpuse põhjal võrdleb autor saksa ja eesti keeles esinevate värvingupartiklite (kommunikatiivset) funktsiooni eri liiki küsilauses. Sagedamini esinevate partiklipaaride *doch/ju, gar/-gi, -ki* ja *denn/siis* võrdlemisel tuuakse esile sarnasused ja erinevused värvingupartiklite kasutamisel mõlemas keeles. Ühe aspektina vaadeldakse värvingupartiklite tõlgitavust eesti keelest saksa keelde ja vastupidi. Partiklite kirjeldamise käigus joonistuvad välja kriteeriumid, mida perspektiivis on võimalik partiklite uurimisel aluseks võtta.

Võttesõnad: komparatiivne lingvistika, näidendid, tõlkimine, eesti keel, saksa keel

1. Sissejuhatus

Viimastel aastatel on värvingupartikleid (VP) eesti keeles nii mõneski käsitluses pikemalt ja põhjalikumalt uuritud. Peamiselt on piirdutud üksikute partiklite kirjeldamisega, puudub süstemaatilisem analüüs värvingupartiklite esinemisest. Siinses artiklis vaadeldakse värvingupartikleid ühes lausetüübis, nimelt küsilauses, ja võrdluses saksakeelsete paralleellausetega kirjeldatakse mõningaid eesti ja saksa keele sagedasemaid partikleid. Peamiselt keskendutakse partiklite funktsioonide ja kasutustingimuste kirjeldamisele. Kuna tegemist on kontrastiivse uurimusega, püütakse kahte keelt võrreldes ammutada uusi teadmisi partiklite kasutamise kohta eesti keeles.

Me vaatasime ju silma? / Wir haben uns doch in die Augen gesehen? küsib Tüdruk Jaan Tätte näidendis “Sild” (JTe, JTs) lootusrikkalt. Küsimuse tegelikku eesmärki – kinnituse ootust, lootust positiivsele vastusele, isegi tüdruku suhtumist vestluspartnerisse – väljendavad eeltoodud näites nimelt värvingupartiklid, mida järgnevalt vaatlemegi.

Käesolev analüüs püüab vaadelda värvingupartiklite esinemist ühes lausetüübis ning leida kriteeriumid värvingupartiklite kirjeldamiseks. Autori eesmärgiks on luua raamistik edasiseks partiklite süstemaatiliseks analüüsiks. Lähtudes saksa keele partiklitest vaadeldakse, mis neile eestikeelsetes tekstides vastab, ja vastupidi, partikleid sisaldavat eestikeelset lauset võrreldakse saksakeelse tõlkega. Uurimuses kõrvutatakse eesti- ja saksakeelseid ilukirjanduslikke tekste. Ilukirjandus, eriti käesolevas uurimuses aluseks võetud näidendid, pakub piisavalt materjali partiklite kirjeldamiseks, kuivõrd sisaldab elulähedasi, igapäevakeelele sarnaseid dialooge. Näidenditeks kitsamalt on kirjutatud eesmärgiga neid suuliselt ette kanda. Autori püüd on seejuures väljendada võimalikult täpselt kõneleja emotsioone ja suhtumisi, võttes muude vahendite kõrval abiks just (värvingu)partiklid.

Kõrvutavaid uurimusi, eriti saksa keelt puudutavaid, on viimaste aastakümnete aktiivse partikliuurimise käigus ilmunud ohtralt. Juba partikliuurimise üheks teerajajaks peetud Harald Weydt (1969) näitas traditsioonilise kontrastiivse uurimismetodi tulemuslikkust keelesüsteemi kirjeldamise jaoks, kõrvutades prantsuse ja saksa keele partikleid. Eesti keele partiklite *-ki*, *-gi* ja *ka* funktsioone on soome keelega põhjalikult võrrelnud Helle Metslang (2002). Nii eesti kui saksa keelt võib pidada partiklirikaste keelte hulka kuuluvaks, seega võiks oletada, et paralleeltekstid pakuvad piisavalt võrdlusmaterjali. Siinse uurimuse põhieesmärk on seega kontrastiivsele uurimusele toetudes luua uusi teadmisi mõningate partiklite kasutamise kohta, eelkõige aga aspektide kohta, mida peaks partiklite kirjeldamisel silmas pidama.

2. Värvingupartiklitest

Käesolevas uurimuses keskendutakse värvingupartiklitele. Värvingupartiklite all mõeldakse morfoloogiliselt muutumatuid sõnu, millel puudub või on vähene iseseisev leksikaalne tähendus, kuid mis modifitseerivad lauseliikmeid, väljendavad kõneleja suhtumist ning esinevad pragmaatiliste indikaatoritena (Hennoste 2000).

Tiit Hennoste (2000: 71–72) liigitab värvingupartiklid vabalt liikuvate partiklite rühma, võttes aluseks nende süntaktilise käitumise. Kogu partiklite süsteem on seejuures järgmine.

1. Üksiesinevad partiklid ehk suhtluspartiklid (*ahah, jah* jms).
2. Tekstipartiklid (piiripartiklid *vä/või, siis/sis, no, nii*; vabalt liikuvad partiklid (värvingu-, modaal-, fookus-, intensiivsuspantiklid)).
3. Toimetamispartiklid ehk n-ö tekstitegemispartiklid (*noh, ah, tähendab*).

Käesolevas uurimuses toetutaksegi Hennoste liigitusele, mida täiendatakse “Eesti keele grammatika” (EKG II) käsitlusega. Nii saadakse võrdlusalus saksakeelses teoreetilises kirjanduses levinud arusaamadele.

Saksakeelses lingvistikakirjanduses on üldlevinud partiklite jaotamine modaal- ehk värvingupartikliteks (*Modalpartikeln, Abtönungspartikeln*), rõhupartikliteks (*Gradpartikeln, Fokuspartikeln*) ja intensiivistavateks partikliteks (*Intensivpartikeln*) (Helbig, Buscha 1994, Hentschel, Weydt 1983, Engel 1991, Erben 1972). Kui varasemates, enne 1995. a ilmunud käsitlustes on partiklid reeglina adverbide alaliik, erifunktsioon kommunikatiivsel tasandil (Eisenberg 1985, Thurmair 1989, Helbig, Buscha 1994), siis hilisemates teoreetilistes käsitlustes moodustavad par-

tiklid iseseisva sõnaliigi (vrd Kärnä 2005). Just pragmaatiline funktsioon eristabki partiklid nii saksa kui eesti keeles adverbidest. Nimelt võib esineda, et sama sõna toimib lauses kas adverbi või partiklina. Nii esineb partikkel *siis* lauses *Kõigepealt sõime, siis läksime kinno* temporaaladverbina, lauses *Mis sa siis läksid kaasa!* etteheidet rõhutava partiklina. Seega on erinevatel partiklitel, sõltuvalt partikliliigist, oma roll peamiselt metakommunikatiivsel tasemel. Partiklite suur semantiline, süntaktiline ja pragmaatiline variatiivsus on põhjustanud arvukalt erinevusi partiklite võimalikes kirjeldustes ja alajaotustes (pikemalt vt Brjantsev 2001: 12 jj).

Miks on raske partikleid iseseisva sõnaliigina defineerida? Sõnaliigi defineerimisel tuleb arvestada süntaktilisi, morfoloogilisi ja semantilis-pragmaatilisi aspekte. Morfoloogiliselt on partiklid muutumatud sõnad. Semantilis-pragmaatilisest aspektist võivad partiklid vastavalt situatsioonile omada erinevat funktsiooni. Vaatamata püüdlustele on osutunud keeruliseks määrata mingi partikli põhitähendust, sest see võib sõltuda oluliselt kontekstist. Lisaks sellele käituvad erinevad partikliliigid erinevalt. Intensiivsuspartiklid (*nii, väga* jms) on seotud näiteks adjektiiviga või adverbiga lauses ja intensiivistavad selle tähendust (1). Rõhupartiklid (*ka, just, eriti* jms) tõstavad esile või rõhutavad mingit lauseosa (2). Värvingupartiklid (3) väljendavad kõneleja ootusi, suhtumist vm (*ju, siis, nagu* jms). Seega on semantilis-pragmaatilise kriteeriumi põhjal raske teha üldistusi, mis kehtiksid kõikide partiklite kohta. Võrreldes teiste partikliliikidega tuleb esile tõsta värvingupartiklite kandva tähenduse suuremat sõltuvust kontekstist.

- (1) Ma olen ju **nii** vana. (MKe)
- (2) **Eriti** kahju on poisist, kel tõesti raha ei ole, korra või paar käib peol, tahab mingit tüdrukut, teeb talle välja, tüdruk ütleb aitäh, ja siis juba enne peo lõppu läheb teise juurde. (MKe)
- (3) Esimest korda oli nii, et läksin sõbra juurde, tuttav tuli vastu ja ütles, et ma annan sulle, pärast maksad ära. Pistis mulle taskusse, mõtlesin, et võiks **ju** ära proovida. (MKe)

Kolmandaks tuleb arvestada süntaktilise kriteeriumiga. Süntaktilisest aspektist sõltub partiklite esinemine partikliliigist, aga ka partiklist endast. Nagu eespool mainitud, liigitab Hennoste värvingupartiklid vabalt liikuvate partiklite rühma, s.t nad on süntaktiliselt sõltumatud ning tegutsevad lauselaiendina. Lauselaiend on “Eesti keele grammatika” järgi üks üldlaiendite liik.

Üldlaiend on lauseelement, mis modifitseerib kogu situatsiooni tähendust. Tal ei ole situatsiooni komponendi tähendust ja seetõttu ei saa tema kohta lauses ka küsimust esitada. Leksikaaltähendus üldlaiendil a) puudub üldse või b) kuulub situatsiooni suhtes metatekstilisele tasandile .. Üldlaiendina toimivad partiklid, harvemini ka muudesse sõnaliikidesse kuuluvad sõnad ja nende ühendid .. Üldlaiendid jagunevad LAUSE- ja FRAASILAIENDITEKS. Lauselaiendi mõjualaks on lause tervikuna ning ta paikneb enamasti lause alguses või kuulub vormiliselt öeldise juurde. Nt **Küllap** tuleb tänavu varajane kevad .. Fraasilaiendi mõjualaks on mingi fraas lauses ning ta asub tavaliselt vahetult selle fraasi ees, harvem järel .. Nt *Võta mind ka kaasa.* (EKG II: 99, kursiivi kasutus siinselt autorilt.)

Kui intensiivsus- ja rõhupartiklid on otseselt seotud enamasti omadus- või määr sõnaga, paiknedes alati vahetult enne sõna, mille tähendust nad modifitseerivad (1), (2), siis värvingupartiklid (3) on seotud otseselt verbiga lauses ja nende tähendus laieneb seega kogu lausele, väljendades midagi, mis ei lähtu üksikute sõnade semantilisest sisust. Erinevalt värvingupartiklitest ei saa intensiivsus- ega rõhupartiklid üldjuhul lauses kohta muuta, ilma, et lause muutuks grammatiliselt valeks.

“Die Partikeln sind dadurch syntaktisch gekennzeichnet, dass sie keine Satzglieder bilden, sondern als Satzgliedteile auftreten, ganz ausserhalb eines Satzverbandes stehen, oder aber nahtlos in den Satz integriert sind,” kirjeldab Aino Kärnä (2005) nimetatud partikliliike nende süntaktilisest käitumisest lähtuvalt.

Käesoleva uurimuse jaoks on relevantne partiklite omadus esineda kommunikatiivsetes lausetüüpides erinevalt. Saksa keelt on selles osas oluliselt põhjalikumalt uuritud.

Das Vorkommen der Partikeln mit vorwiegend kommunikativem Wert ist in starkem Maße an zwei Faktoren gebunden: (1) an die formalen Satzarten Aussagesatz, Fragesatz, Aufforderungssatz, beim Fragesatz wieder differenziert in Entscheidungs- und Ergänzungsfrage, (2) an die kommunikativen Sprecherintentionen (Aussage, Frage, Aufforderung). (Helbig, Buscha 1994: 486. Vrd ka Hentschel, Weydt 1994: 287, Thurmair 1989: 49)

Mõned värvingupartiklid (nt *doch*) esinevad mitmes lausetüübis, teised vaid seotuna ühe lausetüübiga (nt *etwa*). Seni teadaolevalt ei ole läbivat semantilist seletust seotusele erinevate lausetüüpidega. Seejuures sõltub partiklite esinemine saksa keeles nii formaalsest kriteeriumist (väitlause, küsilause, hüüdlause, käsklause) kui ka lause kommunikatiivsest eesmärgist (küsimus, üleskutse, jutustav lause vm) lähtuvalt.

Gerhard Helbigi ja Joachim Buscha (1994) andmetel esinevad eelkõige kommunikatiivsel tasandil tähendust omavad partiklid, mille hulka kuuluvad kahtlemata ka vaatluse all olevad värvingupartiklid, saksa keeles järgnevalt.

Tabel 1. Mõningate partiklite esinemine lausetüübiti saksa keeles (Helbig, Buscha 1994: 487)

Partikkel	Väitlause		Üldküsilause		Eriküsilause			Hüüdlause
	Väitlause funktsioonis	Hüüdlause funktsioonis	Küsilause funktsioonis	Hüüdlause funktsioonis	Küsilause funktsioonis	Hüüdlause funktsioonis	Retooriline küsimus	
<i>aber</i>		x		x		x		x
<i>auch</i>	x		x			x	x	x
<i>bloß</i>	x		x	x	x	x	x	x
<i>denn</i>			x		x		x	
<i>doch</i>	x	x		x		x		x
<i>eben</i>	x	x						x
<i>etwa</i>			x					
<i>halt</i>	x	x						x
<i>ja</i>	x	x		x				x
<i>mal</i>	x							x
<i>schon</i>	x						x	x
<i>wohl</i>	x		x		x		x	

Tabel 1 näitab värvingupartiklite funktsioonide paljusust; näiteks selgub, et partikkel *aber* esineb ainult lausetes, millel on hüüdlause funktsioon, sõltumata formaalsest lausetüübist, partikkel *denn* esineb ainult lauses, millel on küsilause funktsioon, ning partikkel *wohl* võib esineda nii jutustava lause kui küsilause funktsiooni omavas lauses. Tabelis esitatud partiklitest peatutakse käesolevas uurimuses pikemalt partiklitel *gar*, *denn*, *nun* ja *doch* ning nende võimalikel eestikeelsetel vastetel. Kuivõrd ja millised partiklid eesti keeles on seotud kindlate lausetüüpidega, ei võimalda sinne korpus piisavalt uurida, esialgu võib vaid oletada, et teatud piirangud värvingupartiklite kasutamises lausetüübiti esinevad.

Küsilause te jaotuses on siin toetutud “Eesti keele grammatika” (EKG II: 168) küsilause te liigitusele:

- 1) alternatiiv- ehk valikküsilause (küsiiva partikliga *kas*, *ega*; inversiooni abil moodustatud küsilause te on finiiitverb lause alguses), mis omakorda liigitatakse üldküsilauseks (jah/ei-vastus, küsimuse mõjualaks on kogu lause), ülderiküsilauseks (jah/ei-vastus, aga küsituumas olemasolu) ja loetlevaks valikküsilauseks (ei eelda jah/ei-vastust ega mõju ka kogu lausele);
- 2) eriküsilause (küsisõna või küsifraasiga).

Nagu eesti, nii ka saksa keeles algab eriküsilause küsisõnaga (siit saksakeelne nimetus *w-Fragesatz*). Valikküsilause moodustub saksa keeles finiiitse verbi toomise abil lause algusesse (4); ka eesti keeles on võimalik sama moodustusviis (5).

(4) Gefalle ich ihnen denn nicht? (FDs)

(5) Olin vist veidi äge? (FDe)

Peamisteks eesti keele valikküsilause markeriteks on *kas* ja *ega* ning lauselõpuline *või*.

(6) Kas käsi saaks pesta? (JTe)

(7) Ega te pahanda, et sellest räägin? (JTe)

(8) Nalja teed või? (SSe)

Iga liiki küsilause võib olla eelistusküsilause, jagunedes oletus- ja kahtlusküsilauseks (sisaldab eelnevat hinnangut, nt *Kas ta minust enam hoolibki?*) (Metslang 1981).

Kahtlusküsilause te väljendavad kahtlust

- a) partiklid *-gi/-ki*,
- b) jaatavas lauses eituslembesed sõnad ja sõnaühendid (*ealeski, iialgi, üldsegi enam, hoolimata...*),
- c) eitavas lauses jaatuselembesed sõnad ja sõnaühendid (*juba, hoopis, kenasti, päris, üsnagi tublisti, üpriski...*) (Vrd EKG II: 172)

Oletus- ja kahtlusküsilause te sisaldavad teataval määral retoorilisust, s.t alati ei ole võimalik üheselt eristada oletus- ja kahtlusküsilause te retoorilisest küsimusest. Siin ja järgnevalt on neid eristatud ainult selgetel juhtudel.

Saksa keeles võib värvingupartikkel olla selleks indikaatoriks, mis muudab nt eriküsilause te retooriliseks küsimuseks. Partiklita küsilause te võivad kontekstist sõltuvalt väljendada informatsiooniküsimust või retoorilist küsimust. Modaalpartikli lisandudes võib olukord muutuda, vrd *Wer ist schon für die Wiedervereinigung?* ja *Wer ist für die Wiedervereinigung?* (Meibauer 1991: 227 jj). Selline kasutus on

võimalik ka eesti keeles: *Kes siis taasühinemise poolt on?* (retooriline küsimus), vrd *Kes on taasühinemise poolt?* (informatsiooniküsimus).

Partikleid on eestikeelsetes küsilausestes suhteliselt vähe uuritud. Enkliitikut *-ki/-gi* küsituumaa fokuseerijana eriüldküsilausest kirjeldab Metslang (2002), põgusalt puudutatakse partikleid oletus- ja kahtlusküsilausestes sama autori ülevaates “Küsilause eesti keeles” (Metslang 1981). Analüüsidest partikli *siis* kasutusviise, on Airi Jansons kirjeldanud 91 juhtu (kokku 1772-st), kus partikkel *siis* esines värvingupartiklina:

Kõik partikli *siis* abil imestust avaldavad lausungid olid vaadeldavates tekstides küsimused, millega enamasti reageeriti partneri voorule või tegevusele ..., andes mõista, et kõneleja ise on erineval arvamusel (viidates tekstivälisele asjaolule, nt kõneleja mõttekäigule), ja oodates selgitust. (Jansons 2002: 627)

Küsilausest võib partikkel muuta lause pragmaatilist funktsiooni (nt informatsiooniküsimusest retooriliseks), intensiivistada küsimust või muuta seda sõbralikumaks, väljendada imestust või etteheidet, apelleerida ühisele teadmisele. Järgnevalt võetaksegi vaatluse alla mõningad sagedasemad värvingupartiklid, mis võivad esineda küsilausest, ja vaadeldakse, milliseid funktsioone üksikud partiklid küsilausest konkreetselt omavad ja millistest teguritest värvingupartikli kasutamine sõltub. Vaadeldakse korpuses esinenud küsilausestüüpe: eriküsilausest (*Ergänzungsfrage*, *w-Fragesatz*) ja üldküsilausest (*Entscheidungsfrage*). Viimase alla on loetud ka ülderiküsilausest.

3. Analüüsimaterjal

Uuritava korpuse moodustavad ca 100 küsilausest eesti- ja saksakeelsetest näidenditeksidest. Valikul on lähtutud põhimõttest, et saksa keelest eesti keelde ja eesti keelest saksa keelde tõlgitud näidenditeksidest oleks vaatluse alla võetud lähedane arv lauseid (59 ja 40), vähendamaks uurimistulemuste sõltuvust tõlkija eelistustest. Igal näidendil on eri tõlkija. Viiest vaatluse alla võetud näidendist on kaks tõlgitud eesti keelest saksa keelde ja kolm saksa keelest eesti keelde. Vaatluse alla on võetud küsilausest mõlemas keeles, kus esineb vähemalt üks värvingupartikkel. Tabelis 2 on näendid nimetatud lähtekeelse pealkirjaga. Nii näiteks on Jaan Tätte näidendist “Sild” võetud vaatluse alla 12 küsilausest, milles eestikeelses lauses esinev värvingupartikkel on saksa keelde tõlgitud samuti värvingupartiklina; 15 näidet on küsilausestest, milles eestikeelne originaal ei sisaldanud värvingupartikleid, küll aga tõlge, ja 15 on sellised küsilausest, kus värvingupartikkel on jäänud tõlkimata. Korpuse maht ei võimalda küll teha kaugeleulatuvaid järeldusi värvingupartiklite tõlkimisvõimalustest, siiski nähtub allpool toodud arvudest, et erinevused tõlgitud partiklite osas on tõlkijate vahel suured. Nii näiteks ei esine Merle Karusoo näidendis “Laste ristiretk” ühtegi näidet, milles eestikeelse küsimuse värvingupartikkel oleks ka tõlkes värvingupartikliga edasi antud. (Siinkirjutaja hinnangul on nimetatud näidendi tõlkes ka muus osas vajakajäämisi.) Võrdluseks võiks tuua Helmut Krausseri näidendi “Lederfresse”, milles suur osa originaalis esinevatest värvingupartiklitest on lähtekeelde (antud juhul eesti keelde) tõlgitud (11 partiklit 16-st), vaid üks värvingupartikkel on jäänud tõlkimata ja neljal juhul on tõlkija

otsustanud värvingupartikli kasuks, kuigi seda originaalkeeles ei esine. Mõningate sagedasemate partiklite tõlkimisvõimaluste kohta pikemalt edaspidi.

Tabel 2. Korpuse mahuline koosseis

Näidendi autor ja pealkiri (lähtekeeles)	Küsilauseid, milles VP on lähtekeelde tõlgitud	Küsilauseid, milles esines VP ainult saksa keeles	Küsilauseid, milles esines VP ainult eesti keeles
Jaan Tätte "Sild"	12	15	15
Friedrich Dürrenmatt "Die Physiker"	8	8	11
Merle Karusoo "Laste ristiretk"	–	1	7
Helmut Krausser "Lederfresse"	11	1	4
Andreas Sauter, Bernhard Studlar "A. ist eine Andere"	4	–	2
Kokku	35	25	39

4. Sagedasemad värvingupartiklid küsilausestes

Järgnevalt keskendutakse mõningate värvingupartiklite funktsioonile eesti- ja saksakeelsetes küsilausestes. Partiklipaaride pragmaatilise funktsiooni kõrval püütakse jälgida ka süntaktilisi erisusi, seost kontekstiga jt piiranguid partiklite kasutamises.

Sagedasim värvingupartikkel saksakeelsetes küsilausestes on partikkel *denn*, korpuses esinesid veel *gar*, *wohl*, *nun*, *doch*.

Hennoste (2000) loeb värvingupartiklite hulka eesti keeles *ikka*, *nüüd*, *siis*, (*sis*) *alles*, *küll*, *eks*, *tõesti*, *just*, *õige*, *alati*, *vaid*, *ju muidugi*, *tegelt*, *nagu*, *ütleme*, *-gi/-ki*. Nimetatud partiklitest esinevad korpuse küsilausestes *siis*, *-gi/-ki*, *ju*, *õige*.

- (9) Mis ta **siis** teeb? (JTe)
- (10) Sa ei teagi, kellega sa abiellusid? (JTe)
- (11) Me vaatasime **ju** silma? (JTe)
- (12) Sa arvasid kindlasti, et ma olen selle raha mõttetult tulude loopinud, **eks**? (HKe)
- (13) Ja keegi ei tea veel, mis on inimese võimalused. Kuidas **õige** saaks E–LA–DA? (MKe)

Lisaks nimetatutele võivad küsilausestes esineda veel *ikka* (*On ta ikka kohal?*), *küll* (*Miks ta küll ei tulnud?*), retoorilises küsimuses *õige* (*On ta õige hulluks läinud?*), kõnekeeles *nagu* (*On ta nagu normaalne vä?*). Lisaksin siia loetelusse veel partikli *ka*, mis võib küsilausestes intensiivistada olemasolevat retoorilisust:

- (14) Õigupoolest – miks **ka** mitte? (SSe)
Warum eigentlich nicht? (SSs)

Partikkel võib muuta küsimuse ka sõbralikumaks, mõjudes pigem vestluse arendusena kui olulise informatsiooni hankimise eesmärgil mõeldud küsimusena:

- (15) Mida [mis toitu] **siis** meie auväärt haiged **ka** saavad? (FDe)
Was kriegten **denn** unsere lieben Kranken? (FDs)

Eesti kirjakeele seletussõnaraamat (EKSS) rõhutab küll partikli *ka* intensiivistavat funktsiooni, seda nii rõhulises kui rõhuta asendis, kuid tundub siiski, et partikkel võib esineda ka küsimust pehmendavas funktsioonis.

Järgneva korpus esinenud partikleid analüüsiva osa näidetes on tõlgete selgema eristamiseks esitatud esimesena lause originaalkeeles, sellele järgneb tõlge.

Lähemalt analüüsitakse saksa partiklite *gar*, *doch*, *denn ja nun* ning nende eestikeelsete vastete *-gi/-ki*, *siis, ju* ning *õieti* kasutamist küsilausestes. Vaatluse alla võetud partiklipaarid esinesid korpuses korduvalt ning nende puhul võib täheldada võrreldavaid pragmaatilisi funktsioone.

4.1. Saksa *gar* – eesti *-gi/-ki*

Tundub, et väga sarnased oma funktsioonilt küsilauses on saksakeelne partikkel *gar* ja eestikeelne enkliitik *gi/-ki*. Nimelt on kõikidel kasutusjuhtudel võimalik nimetatud partikleid mõlemas keeles samas funktsioonis kasutada: partiklid väljendavad kõneleja imestust. Korpuses esinevates näidetes esineb partikkel *gar* eituslembelisena ja intensiivistab imestust.

- (16) Või ei ole**gi** see ruum? (JTe)
Ein seltsamer Raum. Oder ist es **gar** kein Raum? (JTs)

- (17) Sa ei ole**gi** mind näinud (JTe)
Du hast mich also noch **gar** nicht gesehen? (JTs)

Näidetes (16) ja (17) on partikkel *-gi* peaaegu obligatoorne. Ehk siis – kinnitust ootavas küsimuses väljendab partikkel *-gi* imestust kuulaja reaktsiooni (või selle puudumise üle) ja ootust, et kuulaja kinnitaks kõneleja vastuse-eelistust. Metslangi (2002: 61) liigituse järgi on tegemist *ootuspärasust eitava teatestruktuuripartikliga*. Saksakeelne partikkel *gar* oleks ka näites (17) võimalik, aga mitte kohustuslik. Tundub, et partikli *-gi/-ki* kasutusvõimalused on palju laiemad (nt ka partikli *doch* vastena väitluses, apelleerides ühisele teadmisele; kõneleja kannab oma arvamust üle kuulajale, eeldab konsensust; väljendab etteheidet, õigustust, selgitust).

Erinevalt saksakeelsest partiklist *gar* ei ole eestikeelne kliitik *-gi/-ki* eituslembene, vrd *Sa ei ole**gi** rõõmus? Sa oled**ki** siin?* vs. *Du bist **gar** nicht froh? *Du bist **gar** hier?*

4.2. Saksa *doch* – eesti *ju*

Saksa keele sagedasim partikkel *doch* esines korpuses eranditult vaid oletusküsilausestes (või oletusküsimust väljendavas väitlausestes). Partikli *doch* põhilisele pragmaatilisele funktsioonile – kinnituse otsimisele – vastab eestikeelne partikkel *ju*:

- (18) See käib **ju** vindiga. Või? (JTe)
Das geht **doch** mit Gewinde? (JTs)
- (19) Me vaatasime **ju** silma? (JTe)
Wir haben uns **doch** in die Augen gesehen? (JTs)

Samas funktsioonis oleks võimalik veel partikkel *ometi* ja *ikka*, korpuses näidet kahjuks ei esine.

(20) Sa elad **ju [ometi/ikka]** siin vastasmajas? (JTe)

Neist kolmest (*ju, ometi, ikka*) kinnitavast värvingupartiklist saab *ikka* ainsana esineda ka *kas*-küsilause, mis toimib sel juhul aga kahtlusküsilause (21). Seevastu ei sobi *kas*-küsilause modifitseerima *ometi* ja *ju*: **Kas te olete mind ju näinud?* **Kas te olete mind ometi näinud?*

(21) Kus... kas te olete mind **ikka** enne näinud? (JTe) (kahtlusküsilause)

Wo... Sie haben mich **doch** bestimmt schon mal gesehen? (JTs) (oletusküsilause)

Korpuses leidub kolm juhtu, kus eesti keeles küsisõnaga *ega* moodustatud oletusküsilause on tõlgitud saksa keelde partikli *doch* lisamisega väitlausele (22).

(22) Ebe, **ega** sa mu peale pahane ole? (JTe)

Ebe, du bist **doch** nicht böse mit mir? (JTs)

4.3. Saksa *denn* – eesti *siis*, *õieti*

Värvingupartikkel *denn* on saksa keele kõige sagedasem partikkel argivestluste küsilausestes. Elke Hentscheli ja Harald Weydti andmetel esines seda 21723 sõnast (23 argivestlusest) koosnevas korpuses 30% kõikidest eriküsilausestest ja 8% üldküsilausestest. Varasemates töodes on püütud eristada partikli *denn* pragmaatilist mehhanismi eri- ja üldküsilausestes. Nimelt väljendavat *denn* eriküsimustes sõbralikkust ja üldküsilausestes imestust. Kui üldküsilauses mängib rolli väline põhjus, et lausungis kahelda (*Sprechen Sie denn japanisch?* ‘Kas te räägite siis jaapani keelt?’), siis eriküsilause puhul haaratakse kõnetatav kaasa ja see mõjub sõbralikult (*Wo waren Sie denn im Urlaub?* ‘Kus te siis ka puhkusel käisite?’). (Hentschel, Weydt 1983: 263, 268)

Siiski erinevad keeleteadlaste arvamused, mis puudutab vestlusega seotud ootusi või hoiakuid. Nii käsitleb Jutta Lütten (1977) partiklit *denn* fakultatiivse keelevahendina, küsimuse variandina, grammatiseerunud partiklina. Ka Angelika Redder (1990) kritiseerib Hentscheli ja Weydti liialt lihtsustatud teooriat ning toob näiteid, kus *denn* eriküsilauses mingil juhul sõbralikult ei mõju (*Pubertäre Tochter zur Mutter: Halt dich aus meinen Freundschaften raus! Mutter: Sag mal, wie redest du denn mit mir?* ‘Puberteetikust tütar emale: hoia minu sõpradest eemale! Ma: Kuule, kuidas sa **ometi** minuga räägid?’) (Redder 1990: 27). Helbig (1994: 105 jj) eristab seitse värvingupartikli *denn* pragmaatilist funktsiooni:

- 1) üldküsilauses väljendab imestust või üllatust;
- 2) üldküsilauses väljendab etteheidet;
- 3) eriküsilauses väljendab sõbralikkust;
- 4) eriküsilauses väljendab etteheidet (seotud negatiivse hinnanguga);
- 5) eriküsilauses seotud eelmise lausungiga/mõttega;
- 6) retoorilistes küsimustes intensiivistab retoorilisust;
- 7) hüüdlausestes väljendades imestust, reaktsiooni, intensiivistades etteheidet.

Võimalikke eestikeelseid vasteid partiklile *denn* on mitmeid (*siis, küll, ka, ometi*). Sõna *siis* funktsioone eesti keele suulises kõnes võib Jansonsi (2002: 614) järgi jagada kuude suurde rühma: 1) ajale ja järgnevusele viitamine, 2) asjaolule viitamine, 3) teksti liigendamine, 4) lisamine, 5) tekstile värvingu andmine, 6) parandusmehhanismis osalemine. Värvingupartikkel *siis* väljendab (küsilauses) üllatust (reaktsioon, kui kõneleja on eriarvamusel ja ootab selgitust), imestust, tõrjuvat suhtumist või kõneleja pahameelt. (Jansons 2002: 627) Võrreldes mõlema partikli pragmaatilist funktsiooni, leiame seega ühtelangevusi. Vaatleme järgnevalt küsilause tüüpide kaupa, milliseid pragmaatilisi funktsioone täidavad lauses partiklid *denn/siis* sarnaselt.

1. Eriküsilause. Küsimusele oodatakse vastust, soovitakse täpsustada. Küsimus on sõbralik, n-õ “jutu jätkuks”, seostub sisuliselt eelneva lausungiga, näiteks lauses (23) jutuks oleva maailmameistrivõistluste teemaga.

(23) Wer hat **denn** gespielt? (SSs)

Kes **siis** mängis? (SSE)

(24) Mis ta **siis** teeb? (JTe)

Was macht sie **denn**? (JTs)

Sõbralikumaks võiks eelolevad näited muuta partiklipaar *siis ka*. Kes **siis ka** mängis? või *Mis ta siis ka teeb?* mõjuks sarnaselt saksakeelsetele näidetele (23) ja (24) vastutulelikumalt kui samad lausungid ilma partiklita.

Reaktsioonina eelnevale lausele intensiivistab partikkel etteheidet, väljendab arusaamatust teise käitumisest või ütlustest.

(25) Wie soll ich es **denn** sonst sagen? (FDs)

Aga kuidas ma [siis] peaksin rääkima? (FDe)

(26) Wie viel verlangst du **denn** von mir? (HKs)

Kui palju sa **siis** mult nõuad? (HKe)

2. Üldküsilause. Küsimuses peegeldub etteheide, imestus, lisaks üleküsimine, kinnituse ootus.

(27) Käite [siis] siin silla peal? (JTe)

Sind Sie **denn** sonst manchmal hier auf der Brücke? (JTs)

(28) Liebst du mich **denn** gar nicht? (FDs)

Te ei armasta mind **siis** sugugi? (FDe)

3. Retooriline küsimus. Korpusest leidunud näide (29) on retooriline eriküsilause.

(29) Wenn wir die Bullen nicht hätten, wer würde uns **denn** beschützen vor Leatherface und Henry und Jason und Michael Myers, wer **denn**? (HKs)

Kui meil mente poleks, kes kaitseks meid **siis** Leatherface'i ja Henry ja Jasoni ja Michael Myersi eest, kes **siis**? (HKe)

Eelnevast järeldub siiski põhimõtteline erinevus eestikeelse partikli *siis* ja saksakeelse *denn* vahel. Eesti keeles ei väljenda partikkel *siis* antud materjalis sõbralikkust, viisakust. Tundub, et saksakeelne *denn* on kinnistunud viisaka küsimuse

formuleerijana (vähendab distantsi). See funktsioon puudub eestikeelsel partiklil *siis*. Kontrollimist vajaks oletus, et eesti keeles puudub partikkel, mis muudaks küsilauset sõbralikumaks, pehmemaks. Eriti silmatorkav on see vestluse algfaasis. Sama kinnitab ka algajatele mõeldud saksa keele õpikute analüüs (Kärk 2008). Nimelt näitas nimetatud uurimus, et saksa keelt emakeelena rääkivate autorite õpikutes esines oluliselt rohkem värvingupartikleid (peamiselt partiklit *denn*) küsilauses kui eesti keelt kõnelevate autorite koostatud saksa keele õpikuis algajatele. Ilmselt on siinkohal tegu emakeele mõjuga partiklikasutusele võõrkeeles. On ju partiklid teadaolevalt keelendid, mille kasutamise täielik omandamine võõrkeeles on teiste leksikaalsete vahenditega võrreldes oluliselt keerulisem. Korpuse põhjal ilmnes, et kui saksa partiklil *denn* ei ole seos eelnevaga obligatoorne, siis eestikeelne partikkel *siis* viitab enamasti millelegi, millest on eelnevalt juttu olnud, ja see välistab partikli esinemise vestluse algfaasis.

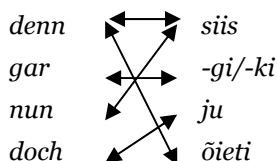
- (30) Was gibt es **denn**? (FDs)
 Mis siin toimub? (FDe)
 (*Mis siin **siis** toimub?)

Vaadeldes, millistel juhtudel on partikkel *denn/siis* jäänud tõlkimata, ilmneb, et tõlkides saksa keelest eesti keelde, on tõlkija kokku 10 korral lisanud partikli *siis*, vastupidise tõlke korral seevastu on 9 korral jäänud partikkel *siis* tõlkimata. Enamikul nimetatud juhtudest väljendab partikkel *siis* kannatamatust, otsustavust ja soovi eelnev jutt kokku võtta. Näites (31) täidab saksa keeles seda funktsiooni lausealguseline *also* 'niisiis'. Kuna tegemist on ühe näitega, võib siin näha ka tõlkija eelistust, kuna partikkel *denn* oleks küsimust intensiivistavas funktsioonis samuti võimalik. Lausealguseline *also* tundub antud näidendis olevat sage vahend lause sissejuhatamisel.

- (31) Olen **siis** ilus või? (JTe)
 Also, bin ich schön? (JTs)
 (Bin ich **denn** schön?)

5. Kokkuvõtteks

Korpuses ilmnenud saksa ja eesti partiklite küsilauses kasutamise vastavused on toodud joonisel 1.



Joonis 1. Saksa ja eesti partiklite peamised vastavused küsilauses

Sagedaseks värvingupartiklipaariks küsilauses osutus *denn/siis*. Mitmete sarnasuste juures partiklipaari *denn/siis* kasutamisel ilmnes üks põhimõtteline erinevus: eestikeelse partikli *siis* kasutamine eeldab seost eelneva kontekstiga, seetõttu ei saa seda kasutada vestluse algfaasis. Küsimust pehmemdavas funktsioonis võib eesti

keeles esineda muuhulgas partikkel *ka*. Antud korpus jääb siinkohal järelduste tegemiseks siiski napiks. Ilmnes, et küsides kinnitust imestust äratavale faktile, on partikli *-gi/-ki* kasutamine peaaegu obligatoorne (saksa *gar* seevastu vaid võimalik ja sedagi eitava lause puhul), ka see väide vajaks põhjalikumat analüüsi.

Joonistus välja hulk kriteeriume, mille alusel värvingupartiklite kasutamist võiks analüüsida:

- 1) esinemine lausetüübiti (küsi- jm laused; küsilause tüübid);
- 2) laused oma otseses kommunikatiivses või nihkunud funktsioonis (nt küsilausega väljendatav väide või käsk; eelistusküsilause);
- 3) esinemine eitavas ja/või jaatavas lauses;
- 4) seotus kindla lauseliikmega (lausepartikkel või fraasipartikkel);
- 5) seos eelneva kontekstiga või selle puudumine;
- 6) tekstifunktsioonid.

Lisaks leidis kinnitust, et kuigi partiklite kasutamine, eriti mis puudutab nende pragmaatilist funktsiooni, võib keeliti olla erinev, võib saksa ja eesti keele partiklite kirjeldamisel toetuda samalaadsetele kriteeriumitele. Sarnaselt saksa keelega võivad värvingupartiklid ka eesti keeles nihutada lause kommunikatiivset funktsiooni (näiteks informatsiooniküsimusest retooriliseks) ning kanda edasi erinevaid tekstifunktsioone (etteheide, imestus jne). Loetletud kriteeriumid annavad lähtealuse üksikute partiklite edaspidiseks põhjalikumaks käsitlemiseks.

Materjali allikad

JTe = Tätte, Jaan 2000. *Sild. Eesti Teatri Agentuuri raamatukogu digitaalkogu.*

JTs = Tätte, Jaan 2000. *Die Brücke. Eesti Teatri Agentuuri raamatukogu digitaalkogu.* (Saksa keelde tõlkinud Irja Grönholm.)

FDs = Dürrenmatt, Friedrich 1986. *Die Physiker.* Zürich: Diogenes.

FDe = Dürrenmatt, Friedrich 1964. *Füüsikud.* Loomingu Raamatukogu, 14. (Eesti keelde tõlkinud Aleksander Kurtna.)

MKe = Karusoo, Merle 1997. *Laste ristiretk.* Võrumaa. Eesti Teatri Agentuuri raamatukogu digitaalkogu.

MKs = Karusoo, Merle 1997. *Kindergefährzug.* Eesti Teatri Agentuuri raamatukogu digitaalkogu. (Saksa keelde tõlkinud Eili Heinmets.)

HKs = Krausser, Helmut 1994. *Lederfresse.* Theater heute, 6, 44–49.

HKe = Krausser, Helmut. *Nahkmask.* Käsikiri. Eesti Teatri Agentuuri raamatukogu. (Eesti keelde tõlkinud Kleer Maibaum.)

SSs = Sauter, Andreas; Studlar, Bernhard 1999/2000. *A. ist eine Andere.* Berlin. Käsikiri. Eesti Teatri Agentuuri raamatukogu.

SSe = Sauter, Andreas; Studlar, Bernhard 2000. *A. on keegi teine.* Eesti Teatri Agentuuri raamatukogu digitaalkogu. (Eesti keelde tõlkinud Maria Taimre.)

Viidatud kirjandus

Brjantsev (Kärk), Janika 2001. *Zu den Veränderungen in der Partikelfrequenz in der jüngeren deutschen Sprachgeschichte. Eine Untersuchung anhand persönlicher Briefe des 18., 19. und 20. Jahrhunderts.* Magistritöö. Talinna Pedagoogiline Ülikool.

EKG II = Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1993. *Eesti keele grammatika. II. Süntaks.* Lisa: Kiri. Eesti Teaduste Akadeemia Eesti Keele Instituut. Tallinn.

- EKSS = Langemets, Margit; Tiits, Mai; Valdre, Tiia; Veskis, Leidi; Viks, Ülle; Voll, Piret (toim.) 2009. Eesti keele seletav sõnaraamat I–VI. Eesti Keele Instituut. Tallinn: Eesti Keele Sihtasutus. <http://www.eki.ee/dict/ekss> (28.02.2010).
- Eisenberg, Peter 1985. Grundriß der deutschen Grammatik. 3., überarbeitete Auflage. Stuttgart Weimar: Metzler.
- Engel, Ulrich 1991. Deutsche Grammatik. Heidelberg: Groos.
- Erben, Johannes 1972. Deutsche Grammatik. München: Hueber.
- Helbig, Gerhard 1994. Lexikon Deutscher Partikeln. Berlin und München: Langenscheidt.
- Helbig, Gerhard; Buscha, Joachim 1994. Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht. Leipzig: Langenscheidt.
- Helbig, Gerhard; Buscha, Joachim 2005. Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht. Berlin und München: Langenscheidt.
- Hennoste, Tiit 2000. Sissejuhatus suulisse eesti keelde. IV Suulise kõne erisõnavara. III Partiklid. – Akadeemia, 8, 1772–1806.
- Henschel, Elke; Weydt, Harald 1983. Der pragmatische Mechanismus: *denn* und *eigentlich*. – Harald Weydt (toim.). Partikeln und Interaktion. Tübingen: Niemeyer.
- Henschel, Elke; Weydt, Harald 1994. Handbuch der deutschen Grammatik. 2. Auflage. Berlin: Walter de Gruyter.
- Jansons, Airi 2002. Partikli *siis* funktsioonid eesti suulises kõnes. – Keel ja Kirjandus, 9, 612–629.
- Kärk, Janika 2008. Zu den frequentesten Abtönungspartikeln in den Lehrwerken für Deutsch als Fremdsprache. Mit besonderer Berücksichtigung der Modalpartikel “denn” in den Fragesätzen. – Mari Tarvas (toim.). Tradition und Geschichte im literarischen und sprachwissenschaftlichen Kontext. Frankfurt am Mein: Peter Lang, 131–141.
- Kärnä, Aino 2005. Ein altes Problem: Partikeln in der Grammatik – ja, aber wie? Linguistik-online.de/22_05 (19.10.2006).
- Lütten, Jutta 1977. Untersuchungen zur Leistung der Partikeln in der gesprochenen deutschen Sprache. Göppinge: Verlag Alfred Kümmerle.
- Meibauer, Jörg 1991. Existenzimplikaturen bei rhetorischen W-Fragen. – Marga Reis, Inger Rosengren (toim.). Fragen und Fragesätze. Tübingen: Niemeyer.
- Metslang, Helle 1981. Kõsilause eesti keeles. ENSV Teaduste Akadeemia Keele ja Kirjanduse Instituut. Tallinn: Valgus.
- Metslang, Helle 2002. *-ki/-gi, ka* ja nende soome kaimud. – Reet Kasik (toim.). Lähivertailuja 12. Soome-Eesti kontrastiivseminar 30.05.–1.06.2001, Kääriku. Tartu: Tartu Ülikool, 57–81.
- Redder, Angelika 1990. Grammatiktheorie und sprachliches Handeln: denn und da. Tübingen: Niemeyer.
- Thurmair, Maria 1989. Modalpartikeln und ihre Kombinationen. Tübingen: Max Niemeyer.
- Weydt, Harald 1969. Abtönungspartikel: die deutschen Modalwörter und ihre französischen Entsprechungen. Bad Homburg v.d.H. etc.: Gehlen.

Janika Kärk (Tallinna Ülikool). Põhiliseks uurimisvaldkonnaks on partiklid saksa ja eesti keeles. janika.kark@mail.ee

ABTÖNUNGSPARTIKELN IN DEN ESTNISCHEN UND DEUTSCHEN FRAGESÄTZEN

Janika Kärk
Tallinn University

Im vorliegenden Beitrag werden die Abtönungspartikel in den estnischen und deutschen Paralleltexten kontrastiv unter Betracht gezogen. Anhand eines aus Schauspieltexten bestehenden Korpus wird erforscht, ob und wie die Abtönungspartikel übersetzt werden. Dabei konzentriert man sich vor allem auf die kommunikative Funktion der Abtönungspartikel in den Fragesätzen. Man vergleicht, ob es im Estnischen vergleichbare Äquivalente gibt, welche Restriktionen hinsichtlich der Verwendung der Abtönungspartikel in den beiden Sprachen gibt, wie die Abtönungspartikel sich syntaktisch verhalten und wie sie die Sprechintention beeinflussen. Während die Abtönungspartikel kontrastiv beschrieben werden, lassen sich Aspekte festlegen, die man bei der Beschreibung der Abtönungspartikel berücksichtigen sollte: 1. Restriktionen auf Satzart; 2. kommunikative Funktion des Satzes (darunter die Fähigkeit der Partikel, die kommunikative Funktion des Satzes zu ändern); 3. Vorkommen im bejahenden oder verneinenden Satz; 4. Bezogenheit auf bestimmte Satzglieder; 5. Bezug auf Kontext; 6. Textfunktionen.

Aus der Untersuchung geht hervor, dass es in beiden Sprachen Partikel gibt (*denn*, *siis*), die die kommunikative Funktion des Satzes ändern können (aus der Informationsfrage wird eine rhetorische Frage). Vergleichbar können die Abtönungspartikel in beiden Sprachen auf der metakommunikativen Ebene Empörung, Vorwurf, Überraschung oder andere Emotionen und Einstellungen des Sprechers ausdrücken. Aus dem Korpus geht jedoch nicht hervor, ob es im Estnischen eine Partikel gibt, die eine Frage, vor allem in der Anfangsphase eines Gesprächs, freundlicher gelten ließe, wie es die deutsche Partikel *denn* tut. Wenn in den meisten untersuchten Fällen die estnische Partikel *siis* vergleichbare illokutive Funktion trägt, schließt der obligatorische Bezug auf die vorher gesagte Proposition die Verwendung der Partikel *siis* aus. Im Allgemeinen kann man sagen, dass es im Estnischen zwar oft vergleichbare Äquivalente zu den deutschen Partikeln gibt, aber beim Übersetzen jedoch gehören Partikeln wegen ihrer komplizierten semantisch-pragmatischen Struktur zu den schwierigen Fällen und ihre Verwendung kann vom Übersetzer abhängig sein.

Keywords: kontrastive Analyse, Schauspielstücke, Übersetzen, Estnisch, Deutsch

VÄHEMUSKEELTE ELUJÕULISUS JA ARENGUSUUNDUMUSED EESTI VENEKEELSETES KEELEKESKKONDADES

Elvira Küün

Ülevaade. Artiklis on kirjeldatud vähemuskeelte situatsiooni Eesti venekeelsetes keelekeskkondades. Vaatluse all on informantide eelistatavad keeled nii kodus kui ka väljaspool kodu ning lugemisel. Eesmärk on anda ülevaade vähemuskeelte elujõulisusest, järjepidevusest ja arengust mitme põlvkonna vältel. Ühtlasi on eesmärgiks Eesti eri keskkondades tarvitataivate keelte kasutamise kaardistuse põhjal luua andmebaas nende piirkondade vastavate keelekeskkondade kohta ning andmestu tutvustamine. Segregatiivsete keelekeskkondade andmebaasi kaudu saab ülevaate, mis alustel toimub üldiselt keele-eelistus, millest on tingitud mingi keele dominatsioon teiste keelte üle, mis keeli kasutatakse segregatiivsetes keelepiirkondades eri valdkondades kodus, koolis ja mujal. Samuti on andmebaasis kirjas, mis tasemel vallatakse kasutatavaid keeli. Antud uurimus võib saada oluliseks lähtepunktiks segregatiivsete keelekeskkondade keelekõnelejate haridusvajaduste määramisel ja sellega seoses ka vastavate haridusotsustuste mõjutamisel Eesti haridusmaastikul.

Uurimuse läbiviimise meetodina on kasutatud kirjalikku küsitlust ja intervjuud.*

Võtmesõnad: hariduspoliitika, keelepoliitika, keelevalik, keelepädevus, kodukeel, eesti keel, vene keel, teised keeled

1. Sissejuhatuseks

20. sajandile üks iseloomulikumaid jooni on rahvaste migratsioon üle rahvuslike ja geograafiliste piiride. Paljusid maailma kultuure ja keeli – eriti neid, mis on arvu- liselt väiksemad – ähvardab sulandumine teistesse, domineerivatesse keeltesse ja kultuuridesse. Tuhanded keeled on paari viimase sajandi jooksul juba kadunud.

* Käesolevat uurimust on rahastanud riikliku programmi “Eesti keel ja kultuurimälu 2009–2013” (EKKM09-186) projekt “Segregatiivsete keelekeskkondade andmebaasi korrastamine ja avalikustamine”.

Maailmas valitseb tendents, kus surve muutuda homogeenseks puudutab nii rahvuslikku assimilatsiooni kui ka majanduslikku üleilmastumist (Edwards 1996).

On arvatud, et 21. sajandil tabab 90% keeltest keelesurm (Krauss 1992). Keelesurm on keelevahetuse viimane staadium, mis saab alguse sihtkeelekõnelejate arvu vähenemisest, mis toob endaga kaasa keeleoskuse kadumise või selle tarvitamise kahanemise eri valdkondades (Baker 2006: 75). Keele säilitamine aga tähendab keele jätkuvat võitlust regionaalselt ja sotsiaalselt võimsama keelega (*ibid.*). Keele saatust mõjutavad erisugused makro- ja mikrotasandi mõjurid. Conklin ja Lourie (1983) on välja pakkunud järgmised poliitilised, kultuurilised ja keelelised tegurid, mis aitavad keelt kas säilitada või välja tõrjuda.

- 1) Poliitilised, sotsiaalsed ja demograafilised tegurid:
 - sidepidamine kodumaaga ja kodumaa külastamine;
 - etnilise grupi identiteedi säilitamine enamusgrupi identiteedi omaksvõtu asemel;
 - suhtlusvõrgustike roll keele säilitamisel või keelevahetusel.
- 2) Kultuurilised tegurid:
 - eksisteerivad emakeelsed institutsioonid (s.o koolid, kogukonna organisatsioonid, massimeedia, vabaaja tegevused);
 - kultuuritegevused ja religioonilised kombetalitused viiakse läbi oma kodukeeles;
 - etniline identiteet on tugevalt seotud kodukeelega;
 - üksikindiviidist lähtuvad keelehoiakud ning suhtumine päritolukeelde ja -kultuuri.
- 3) Keelelised tegurid:
 - emakeel on standardiseeritud ja eksisteerib kirjalikus vormis;
 - kodukeelel on rahvusvaheline staatus;
 - on olemas emakeelne kirjaoskus, mida kasutatakse kogukonnas ja kodumaal.

Keel võib kaduda väga kiiresti, kui seadus keelab seda koolis kasutada ning olukorras, kus kodune keeleülekanne mehhanism enam ei toimi. Keele taastähtsustamine kooli kaudu on väga aeglane protsess. Keelekasutuse vähenemine võib aega võtta vaid paarkümmend aastat, keele taassündimiseks läheb palju rohkem aega (Baker 2006). Keele taassünd hariduses ei alga väikestest koolilastest, vaid vajab õpetajate valmidust ja vastavat väljaõpet – õpetajad peavad tegema lapsevanemate ja keelekorraldajatega koostööd, et keelt päästa. Hyltenstam ja Stroud (1996) lisavad, et analüüsides keelevahetust, on sotsiaalse ja keelekõnelejate kogukonna dimensiooni kõrval väga tähtis ka individuaalne, isiklik tasand vähemuskeele säilitamisel (selle esindajate suhtumine päritolukeelde).

On leitud, et kui on loodud võimalus saada haridust oma päritolukeeles, siis suurenevad inimese eneseaustus ja keeleteadmised, samuti paranevad intellektuaalsed võimed (Baker, Jones 1998: 517). Alternatiivselt on arvatud, et kus võimalik, peaks koolisüsteem piisavalt paindlikult liitma endaga vähemusgruppide kodukeele ja kultuuri (*ibid.*).

1.1. Keelevähemuste keeleline identifitseerimine

Emakeelel on tugev mõju lapse keelevelikule, isa esimesel keelel selles tavaliselt nii suurt osa pole leitud (Bayley, Schecter 2003: 18), kuigi samas ei pea ema- ja kodukeel tingimata ühtima (Baker, Jones 1998). Terminit *kodukeel* on mõned lingvistid eelistanud märkimaks iseloomulikke dialekte ja keeli, mida kasutatakse sageli ainult kodukontekstis ja neid keeli võidakse paljudel juhtudel järgmistele põlvkondadele edasi anda ainult suulise pärimusena (Moon jt 2000: 775).

Keelevähemusi aga võib määratleda kui üksikisikuid, kelle kodudes kasutatakse aktiivselt enamkasutatavast erinevat keelt ja kellel seepärast on olnud võimalus tõsta oma keelepädevuse taset selles keeles (Goldenberg jt 2006: 21). Domineerivaks keeleks on nimetatud keelt, milles on keelekõnelejal kõige paremad teadmised või mida ta kõige rohkem kasutab (Baker, Jones 1998).

Et keel esindab ja vahendab otsustavat elementi inimese sotsiaalses identiteedis, sellele on osutanud mitu uurijat (nt Hoffmann, Ytsma 2003, Skutnabb-Kangas 2000, Iskanius 2005, Ü. Rannut 2005). Vähemusgruppide keeleidentiteeti on vaadeldud kui mitteametliku keele omaksvõttu emakeelena või kodukeelena (Li 2001: 137). See ei tähenda seda, et need, kes samastavad end mitteametliku keelega kui ema- või kodukeelega, tunnevad tingimata tugevat kuuluvustunnet sellesse gruppi (*ibid.*).

Tööturg on keelevahetuse ja -hoiakute üks tugevamaid mõjutegureid, mis mõjutab ka keelelisi valikuid ning sellega koos ka keelelist identiteeti (Iskanius 2005). Shapiro ja Stelcner (1997) on väitnud, et harilikult kalduvad uusimmigrandid, tulles sihtmaale, alguses kas ainult või sagedamini kasutama oma emakeelt, misjärel aja möödudes hakkab see tendents muutuma. Järelikult tuleb tõdeda, et valitsev tööturusihtuolus ei aita vähemuskeelte säilitamisele kaugeltki kaasa. Teine oluline tegur on perekond: nendel etnilistel gruppidel, kes rõhutavad peresidemeid, paistab olevat tugev mõju oma laste hoiakutele õppida ja kasutada nende enda etnilist keelt kodus (Gans 1997).

Keel on kultuuri kandjaks, seepärast on keelelisel identiteedil oluline osa etniiteedi moodustumisel (Skutnabb-Kangas 2000). Ühise keeleruumi kaudu ühtlustub identiteet, kuid osa kollektiive on kakskeelsed, kus kõneleja valib keele olenevalt olukorrast või valdkonnast (diglossia, vt Vahtin 2004). Näiteks keelekogukond võib kasutada vähemuskeelt kodus, naabruses, religioossetel eesmärkidel ja sotsiaalsetes tegevustes, kuid samas enamuskeelt töökohal, hariduses ja massimeediat tarbides (Hudson 2001). Harilikult on enamuskeel kasutusel ametlikus suhtluses ja vähemuskeel mitteametlikus (Baker 2006). See kontrast kõrge ja madala ühiskondliku staatusega keele vahel võib olla tingitud pigem keeleprestiižist ja keelevõimust, mitte keelevariatiivsusest (*ibid.*).

Väga oluline on suhe keeleveliku ja identiteedi vahel. Mingi keele valimine markeerib teatud identiteeti ja võib otsustada isegi seda, kas indiviid võetakse mingisse gruppi vastu või mitte. Inimesed on keelevelikuga samal ajal silmitsi ka oma identiteedi või kogukonna valikuga (Mills 2001: 400).

1.2. Vähemusrühmade etnolingvistiline vitaalsus

Keelekontaktide situatsioonis saame mis tahes keele puhul rääkida keelelisest lojaalsusest (Fishman 1991). Keelekontakt on alati olemuselt unikaalne ja seda hinnatakse selle karakteristikute ja dünaamilisuse põhjal. Teiseks, tähtis on arvestada sotsiaalpsühholoogilisi faktoreid, mis võivad mõjutada või kujundada keelelojaalsust, nagu ka väljastpoolt tulenevad, objektiivsed tegurid (Komondouros, Mcentee-Atalianis 2007). Keelt on sageli vaadeldud kui lojaalsuse tundemärki, sümbolit, millesse suhtutakse kui mingi rahvuse hulka kuulumise tähistajasse (Baker 2006). Etnolingvistilise vitaalsuse probleeme on Eesti tingimustes uurinud nt ka Martin Ehala, Anna Verschik, Kadri Koreinik, Karl Pajusalu jt.

Etnolingvistilise vitaalsuse kontseptsioon rõhutab, et individuaalse keelevaliku sotsiaal-psühholoogilised aspektid ja assotsiatsioonid on seotud väliste mõjuritega, mis on sotsioökonomilised, geograafilised ja poliitilised keelekontakti situatsioonis (Giles jt 1977). Etnolingvistilist vitaalsust on defineeritud kui midagi, mis paneb inimrühma käituma end teistest eristavalt ja rõhutama kollektiivset identiteeti rühmadevahelistes dialoogides (Giles jt 1977: 308). Seda saab hinnata kolme aspekti järgi: staatus, demograafia ja institutsionaalne toetus (*ibid.*). Gilesi ja Byrne'i (1982) järgi sisaldab etnolingvistiline vitaalsus järgmisi aspekte:

- 1) etnilise rühma majanduslik, ajalooline, sotsiaalne, poliitiline ja keeleline staatus;
- 2) etnilise grupi suurus ja levik, segaabelud;
- 3) institutsionaalne tugi etnilistele rühmadele (s.o meedia, haridus, tööstus, religioon, teenused, kultuur ja riigivalitsemine).

Bakeri ja Jonesi (1998) järgi tähendab keelevitaalsus seda määra, mille piires keelevähemused jõuliselt säilitavad ja laiendavad oma igapäevase keele kasutamist ja korrastavad selle funktsioone. Keele elujõulisuse määramisel on olulised selle keele staatus, institutsionaalne toetus, majanduslik väärtus ja seda keelt kõnelejate arv ja levik. Keelelojaalsus peegeldab individuaalseid ja sotsiaalseid jõupingutusi säilitada etniline identiteet sünni- või pärandkeele jätkuva kasutamise abil (Wiley 1996), lojaalsus mingile keelele on selle säilitamise vahend.

Tegelikkuses võivad indiviide survestada kodukeelt vahetama domineeriva enamusrühma keele vastu mitu tegurit, mis võivad olla nii positiivsed (s.o tugev soov assimileeruda riigikultuuri või suurendada enda kultuurilist või majanduslikku kapitali) kui ka negatiivsed (s.o reaalne ja tajutud diskrimineerimine, keelepoliitika, mis teeb kohustuslikuks ametliku keele ja kitsendab domineerivast keelest erinevate keelte kasutust). Uurijad (nt Urzúa, Gómez 2008) on toonud välja, et keelevahetus kaldub kiirenema, kui hoiakud pärandkeele suhtes on negatiivsed (s.o pärandkeelele on omistatud väike väärtus) ja kui pole piisavalt võimalusi sotsiokultuuriliseks keelekaitseks. Mis tahes jõud selle taga ka pole, viib keelevahetus tavaliselt keelekaoni nii üksikindiviidide puhul kui ka rühma sees.

Kogukondades, kus on indiviididel tugevad sotsiaalsed võrgustikud, positiivsed vaated kogukonna keelele ja täpselt ära määratud kodukeele kasutamise reeglid, on keelemuutuste protsess aeglane, võttes aega umbes neli põlvkonda (Holmes, Aipolo 1993). Kogukondade ja indiviidide seas aga, kus võrgustikud on nõrgad, on hoiakud keele alalhoidmise suhtes ambivalentsetes ning seega võib ühiskonnas domineeriv keel muutuda kodus kasutatavaks keeleks, lisaks on keelemuutus

kiire ning võib toimuda isegi vähem kui kahe põlvkonna jooksul (Starks 2005: 541). Põlvkondadevaheliseks keelesäilitamiseks on vajalik kasutada vähemuskeelt perekonnaringis, sest keelel, mis on õpitud kodus, on parimad võimalused säilida ka väljaspool koduseinu (Broeder, Extra 1999).

1.3. Vähemuskeelsete laste õpe: koolikeskkond ja vastutus

Lapsed, kellel on enamikust erinev etniline tagapõhi, võivad end leida kahe kultuuri vahelt, kus nende identiteet on väga tugevalt seotud nende keelevelikuga. See võib teha neid koolis oma kodukeele kasutamise suhtes tõrksaks, juhul kui neid pole kindlustatud tõeliselt mitmekultuurilise koolikeskkonnaga (Siraj-Blatchford, Clarke 2000: 29). Pedraza ja Pousada (1992: 253) järgi sõltuvad lapsed omavahel võistlevatest mõjudest, mida avaldatakse neile kodust, koolist ja kogukonnast. Kaasamise mõistega toonitatakse seda, et kohandumisele kaasaaitamine kuulub kooli vastutusalasse: kool peab muutma oma keskkonda ja poliitikat, et mahtuda õpilaste vajadustesse (DfEE 1999).

Maailmas on loodud vabatahtlikkuse alusel moodustatud keeleklasse kooliealistele lastele. Kui kool ei toeta immigrantide vähemuskeele õpet, siis võib keele n-ö taastootmine olla mitteküllaldane selleks, et vähemuskeel säiliks. Seepärast on kohalikud kogukonnarühmad loonud võimalusi andmaks oma lastele vastavat koolitust. Näiteks Inglismaal ja Kanadas on moodustatud vähemuskeele klasse, õhtukoole, koolivaheajal töötavaid õpperühmi, laupäeva- ja pühapäevakoole erinevatele kogukondadele selleks, et õpetada lastele nende vanemate ning vanavanemate keelt ja kultuuri (Baker 2006: 125–126). Sarnane vabatahtlik tegevus võib olla samas ka religioosse, kultuurilise, sotsiaalse, integratiivse ja etnilise vähemuse elujõulisuse tugitalaks (*ibid.*). Eksisteerivad aga ka ametlikult kakskeelse õppeprogrammiga koolid, kus ühel päeval kasutatakse ühte keelt ja teisel teist, ning leidub ka selliseid õppeasutusi, kus kasutatakse õppekeelena pool päeva ühte keelt, päeva teisel poolel aga teist keelt (Baker, Jones 1998: 588).

On leitud, et süstemaatiline kodukeele õppekeelena kasutamine eelkoolieas toetab hilisemat akadeemilist võimekust ja avaldab positiivset mõju ka kooliealistele lastele, kui nad õpivad kakskeelsete õppeprogrammide järgi (vt nt Kohnert jt 2005: 257). Lõppkokkuvõttes on lapsevanemate võimuses julgustada ka esimest keelt omandama, seda edasi arendama ja säilitama ning ka kirjalikus vormis edasi andma (Seville-Troike 2000: 165). On oluline uurida, millised perekonna keelepoliitika komponendid, nagu ka sotsiolingvistilised tegurid, võivad kiirendada või aeglustada kodukeele alalhoidu.

1.4. Informaalne keeleplaneerimine

Palju on diskuteeritud selle üle, kas üldse on vaja toetada kodukeele arengut, kui see pole enamuskeel ja seda ei kasutata haridussüsteemis. Süstemaatilise keelekõnelemise võimaluse andmise positsioon on motiveeritud ja seotud lapse sotsiaalse, emotsionaalse ja kognitiivse arenguga perekonna kultuurikontekstis (Tuijl jt 2001). Sotsiaalsete, emotsionaalsete, kognitiivsete ja kommunikatsiooni-

oskuste areng on lastel vastastikku sõltuv. Need üksteisega suhestuvad oskused arenevad kultuurikonteksti sees ja primaarne kultuurikeskkond ongi lapse jaoks vahetu perekond (Kohnert jt 2005).

Alternatiivne perspektiiv kodukeele säilitamiseks on keelestrateegia, kus lapsevanemad kasutavad oma lastega suheldes aktiivselt oma vähemuskeelt (Goldenberg jt 2006). Üks põhjus võib selleks olla, et kindlustada parim võimalik lingvistiline mudel laste jaoks. On täheldatud, et ühes keeles omandatud oskused ja teadmised kanduvad üle teise keelde (Wong Fillmore 1991). Teine põhjus esimest keelt kodus säilitada on selle olulisus kultuurilistel, kognitiivsetel ja pragmaatilistel põhjustel, mis moodustab omakorda olulise osa kultuuripädevusest (Simpson 1991). Pole aga selget vastust küsimusele, mis keeles peaks kodus püüdma laste kirjaoskuse arengut ergutada.

Piller (2002: 62) viitab samuti mitteametlikule keeleplaneerimisele ning toob välja, et paljud abielupaarid ei tee teadlikku otsust, mis keelt peaks kodus kasutama, seega on keelevalik juhuslik. Samas aga leidub ka neid, kes hoiavad teadlikult kahte keelt eri situatsioonide puhul lahus ja määravad ära täpsed strateegiad, mis keeles rääkida nii omavahel kui ka lastega. Pavlenko (2004: 184) toob samuti välja, et lapsevanemad teevad keelevalikuid perekonnas kas teadlikult, alateadlikult või spontaansete otsuste põhjal. Vanemate keelevalikule avaldavad mõju nii sotsiaalne kontekst, milles perekond asub, kui ka lokaalne kontekst (nt kogukonna keele-eelistus). Oma lastega suhtlemiseks võivad lapsevanemad valida mõlemad ühe kindla keele. Teisel juhul võib aga ema valida lapsega suhtlemiseks ühe keele, isa teise keele. Paljud pered säilitavad tasakaalu kahe keele vahel, näiteks vanemad räägivad omavahel ühes kindlas kokkulepitud keeles, lastega rääkides aga kasutatakse teist või isegi kolmandat keelt. On registreeritud juhtumeid, kus keel ja keelestrateegia perekonnas valitakse lähtuvalt emotsioonidest, näiteks võidakse ühes keeles väljendada positiivseid, teises negatiivseid tundeseisundeid (Pavlenko 2004).

On avastatud, et lapsed ise võivad üllatavalt tugevalt oma keelevalikut juhtida (Tuominen 1999). Samuti näitab laps üles valmidust kasutada kahte keelt paralleelselt. Koodivahetus, milleks on keelevariandi vahetamine ühe lause sees või lausete vahel (Verschik 2004), leiab tavaliselt aset informaaalses suhtlusolukorras (Zentella 1999). Ka lapsevanemad kasutavad aeg-ajalt koodivahetust, näiteks kui ei suudeta meenutada sõna või fraasi ühes keeles. Täiskasvanute puhul aga võib koodivahetusel ja koodivaheldamisel olla ka pragmaatilisi eesmärke, näiteks käsu tugevdamiseks (Kaur, Mills 1993). Muuhulgas on tehtud tähelepanekuid, et vähemusrühmade seas kasutatav koodivahetus on sisuliselt samm enamuse keelse keelevariandi suunas (Baker, Jones 1998: 587).

Vähemuskeelsete laste kodukeele tase koos selle sotsiaalsete, emotsionaalsete ja kultuuriliste lülidega on paraku ohus ka vastavate keelte kodus suhteliselt aktiivsel kasutamisel. Võime säilitada ja arendada oskusi vähemuskeeles peab olema vastavuses kooli süstemaatilise toetuse ja rikastamisprogrammidega kodukeeles. Kui kavandatud tugi kodukeele jaoks pole kättesaadav, siis soovivad vähemuskeelsed lapsed tõenäoliselt palju vähem kasutada oma vanemate ja teiste lähedaste pere liikmetega suheldes oma kodukeelt (Anderson 2004). Järelikult suureneb sellega seoses nende laste jaoks koormus nende sotsiaalsele, emotsionaalsele ja akadeemilisele arengule.

2. Vähemuskeelte uuringu tähtsus Eesti kontekstis

Eesti kui Euroopa Liidu liikmesmaa jaoks on oluline välja selgitada, milline on praegune reaalne keelesituatsioon ja mis keeli Eesti elanikud kasutavad – see lähtub Euroopa Liidu keelepoliitikast, mis propageerib mitmekeelsust (Commission ... 2005). Samavõrra tähts on järele uurida eri keelte tõeline keelekasutuse ulatus ning keelte varieeruvus avalikus ja erasfääris. Toetudes Euroopa konventsioonile (2003), peab Euroopa Liit aktsepteerima mis tahes kultuurilisi, religioosseid ja lingvistilisi erinevusi (Baldauf, Kaplan 2006).

Vastavalt 2000. a rahva- ja eluruumide loenduse andmetele elab Eestis 142 etnilise rühma esindajaid, kes kõnelevad kokku 109 emakeelt (Eesti ... 2000). Need registriandmed näitavad, mis keeli väidetakse olevat emakeel. Selgusetu on aga see, mis keeli ja keelevariante kasutatakse kodus kui erasfääris.

Uurides kodu- ja vähemuskeeli, võib saada olulist lähtematerjali hariduspoliitika, et aidata planeerida kodukeele õpet – igal inimesel on ju õigus oma emakeelele. Oma keele ja kultuuri väärtustamine kujundab ka positiivset suhtumist sihtkultuuri. Eestis on praegu 21 pühapäevakooli (Muldma 2009), kus õpetatakse Eesti vähemuskeelsetele lastele nende päritolumaa keelt ja kultuuri. Kõige selle juures aga säilib ikkagi oht, et keel kas hääbub või sureb. Selle põhjuseks on sageli just kultuurisurve ja keele prestiiži langus kõnelejate endi silmis. Alati ei olegi selle puhul määrav keelekõnelejate arv, põhiline on siiski suhtumine (M. Rannut jt 2003).

2.1. Uuringu eesmärk ja meetodid

Uuringu eesmärk oli välja selgitada, mis keeled on uuritavates venekeelsetes keelekeskkondades kasutusel kodus ja koolis, mis tasemel vallatakse vastavaid keeli, mis on keeleveliku ja -elistuse aluseks ning missugune on keelerepertuaar nii koolis kui ka kodus. Ühtlasi uuriti informantidest lastevanemate arvamust selle kohta, kas nemad sooviks, et koolis õpetataks nende vähemuskeelset päritolukeelt. Uuringud korraldati Eesti linnades, mida iseloomustab nende mitmekeelne ja -kultuuriline rahvastik.

Kohtla-Järve linnas viidi läbi kvantitatiivne uurimus, Maardu linnas kasutati lisaks kvalitatiivset meetodit. Maardu puhul olid kooliõpilaste kõrval informantidena kaasatud ka nende vanemad, kelle kaudu omakorda saadi informatsiooni ka nende vanemate kodukeelte kohta. See võimaldas uurida keeleülekannet mitme põlvkonna vältel ja seega vähemuskeelte elujõulisust pikemal ajavahemikul. Uuringu ühe eesmärgina püüti ühtlasi välja selgitada koolides õpetatavad keeled ning välja uurida, kas selgub vajadus õpetada peale nende veel muid keeli. Üks eesmärk oli vaadelda ka keelevähemuste etnilise identiteedi säilimist või muutumist pikema aja vältel ning võrrelda nende kahe uuringu tulemusi Tallinna linnas korraldatud kodukeelte uurimusega (vt Ü. Rannut, M. Rannut 2007).

Uurimismeetodina kasutati anketeerimist ja intervjuud. Anketeerimise abil saadi informatsiooni muuhulgas õpilaste ja nende vanemate päritolumaa kohta, samuti kodukeele kasutuse ulatuse, kodukeele pädevuse ja suhtluskeele valiku kohta pereliikmetega. Uurimise all oli ka õpilaste keele-elistus väljaspool kodu. Uuritava keelerepertuaari moodustasid ühtlasi koolis õpitavad keeled. Käesoleva

andmebaasi abil saab iga keelerühma kaupa eraldi oletada nende elujõulisust ja kestust põlvkondade lõikes.

Statistiliseks töötluks kasutati statistilist andmetötluspaketti SPSS 13.0. Kasutati korrelatsiooni, T-testi ning χ^2 -testi. Korrelatsiooni eesmärgiks oli välja selgitada, kas kaks tunnust on omavahel seotud või mitte ning kui tugev see seos on. T-test võimaldas kindlaks teha, kas erinevus tunnuste keskväärtuste vahel on oluline või mitte. χ^2 -testi kasutati üldkogumi puhul, kontrollimaks, kas jaotus on ühtlane.

2.2. Kahe venekeelse piirkonna areng: Kohtla-Järve ja Maardu linn

Keelelisest küljest on Eestis küllalt omapärasteks piirkondadeks kujunenud Kohtla-Järve ja Maardu linn. Praeguse Kohtla-Järve kohale ei olekski arvatavasti kerkinud linna, kui poleks olnud põlevkivi, mida hakati seal ulatuslikult kaevandama. Põlevkivi baasil arenes välja keemiatööstus. 1947. aastal muudeti Kohtla-Järve vabariikliku alluvusega linnaks, mis tõi kaasa täiendavate investeringute saabumise sinna (Kohtla-Järve ... 2008). Seoses tööstuse arenguga toimus Kohtla-Järvel pärast II maailmasõda immigrantide sissevool Nõukogude Liidu teistest regioonidest (Helemäe jt 2000: 17). Nii saabus siia kümneid tuhandeid mitte-eestlasi, kellest suur osa siirdus Kohtla-Järvele kui tööstuspiirkonda. Selle tulemusel ongi kujunenud olukord, kus Kohtla-Järve linna rahvastikust enamiku moodustab venekeelne elanikkond (*ibid.*). Linna elanike arv on natuke üle 50 000 (Kohtla-Järve ... 2009). Eestlasi oli 2000. a rahvaloenduse andmeil Kohtla-Järve linnas 17,8%.

Oma arengu poolest on sarnane Maardu linn, mis asub vahetult Tallinnast idas ja kus elab üle 16 000 inimese. 1939. a loodi riiklik aktsiaselts Eesti Fosforiit uue kaevanduse ja rikastusvabriku rajamiseks ning uus tööstuskompleks. Pärast II maailmasõda jätkati Maardu fosforiiditootmise arendamist – rajati väävelhappetsehhi ja superfosfaaditsehhi ning samal ajal toimus märkimisväärne mitte-eestlaste sisse-ränne (Maardu ... 2009). Vaevalt leidub Eestimaal teist samasugust linna nagu Maardu, kus elab nii palju erinevate rahvuste esindajaid. Enamiku elanikkonnast moodustavad venelased, eestlaste protsent on umbes 20%. Linnas on esindatud 41 rahvust, kes kuuluvad erinevatesse konfessioonidesse (*ibid.*).

3. Kohtla-Järve linna keelteuuring

Küsitlus korraldati kümnes Kohtla-Järve koolis, sihtrühmaks olid 8–12-aastased õpilased (erinevalt Maardust polnud kaasatud lapsevanemaid). Uuringus osales 80,5% kõigist Kohtla-Järve 2.–5. klasside õpilastest (kokku vastas ankeedile 1002 õpilast, neist 774 õpilast vene ja 228 eesti õppekeeleaga koolidest).

Uuringus osalenud Kohtla-Järve koolide õpilased on pärit kahest ning nende vanemad 16 riigist (vt tabel 1). Suurem osa informante oli sündinud Eestis: õpilastest peaaegu kõik (99,5%), emadest oli Eestis sündinuid 82,63% ning isadest 81,53%. Venemaal oli sündinud 5 õpilast, emasid põlvneb sealt 12,97% ja isasid 13,17%. Seega kuulub suurem osa antud õpilasarühmast teise põlvkonna immigrantide hulka.

Tabel 1. Õpilaste ja lapsevanemate sünnimaad¹

Sünnimaa	Õpilased		Emad		Isad	
	Arv	%	Arv	%	Arv	%
Eesti	997	99,5	828	82,63	817	81,53
Venemaa	5	0,5	130	12,37	132	13,17
Ukraina			15	1,5	16	1,6
Valgevene			13	1,3	17	1,7
Usbekistan			4	0,4	0	0
Kasahstan			4	0,4	1	0,1
Aserbaidžaan			1	0,1	4	0,4
Läti			2	0,2	2	0,2
Armeenia			2	0,2	3	0,3
Türkmenistan			1	0,1	2	0,2
Leedu			1	0,1	2	0,1
Poola			1	1,1	0	0
Tadžikistan					1	0,1
Taani					1	0,1
Soome					3	0,3
Itaalia					1	
Kokku	1002	100	1002	100	1002	100

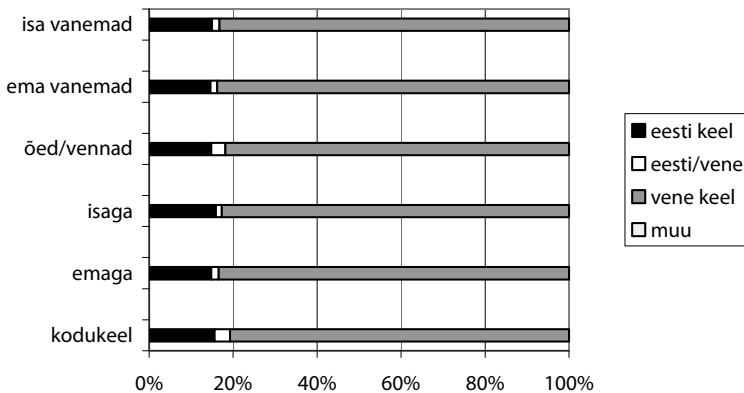
3.1. Sünnimaa ja kodukeele seos

Õpilaste vanemate sünnimaa ei pruugi tingimata mõjutada kodukeele valikut. Kohtla-Järve õpilased kasutasid pereliikmetega suheldes ainult kas eesti või vene keelt, olenemata vanemate sünnimaast (vt joonis 1). Vene õppekeelega koolide õpilastel oli kodukeeleks vene keel isegi sel juhul, kui vanemad olid sündinud Eestis või mõnes teises riigis.

Sünnimaade järgi suurema rühma moodustanud informantidest, kes olid pärit Ukrainast ja Valgevenest (vt tabel 1), suhtlesid vene keelt kodus kasutavad õpilased nii ema- kui ka isapoolsete vanavanematega vene keeles. Sisuliselt on see märk keelekaost. Teisalt võib see näidata ka vanemate väljarännet Ukraina ja Valgevene venekeelsetel aladel. Keelekasutust põlvkonniti võrreldes on näha, et keelevahetus on ilmnunud juba eelmises põlvkonnas. Sellest lähtudes võib väita, et nende rahvuste keeled on Eestis vähemalt selles Kirde-Eesti piirkonnas vene keelde assimileerunud.

Eesti õppekeelega koolide õpilasi oli uuritud õpilaste koguarvust 22,75%. Nendest rääkis eesti keelt kodukeelena 68%, vene keelt aga 15% ja nii eesti kui ka vene keelt 16%. Seda võib seletada asjaoluga, et osa venekeelse taustaga lapsi on pandud õppima koolidesse, kus on õppekeeleks eesti keel. Mõnel neist on üks vanem eestlane ning kodukeelena kõneldakse nii eesti kui ka vene keelt.

¹ Suurem osa väljatoodud riikidest on endise Nõukogude Liidu vabariigid. Siin ja järgnevas on neid nimetatud nende tänapäevaste nimedega.



Joonis 1. Kodukeele valik pereringis

Vene õppekeelega koolides õppijad tarvitavad pereliikmetega lävimisel vene keelt. Eesti koolide õpilastest kasutavad eesti keelt emaga suheldes 148, isaga kõneldes 159 ning õdede ja vendadega 148 õpilast. Vene keel on suhtluskeeleks emaga 62 juhul, isaga 55 ning õdede ja vendadega rääkides 46 õpilasel. Nendes peredes on mõlemad vanemad venekeelsed, kuigi lapsed õpivad eesti õppekeelega koolis. Vanemad soovivad nii täiustada oma laste eesti keele oskust ja kiirendada eesti keele selgekssaamist, mis loob paremaid eeldused edukuseks hilisemas elus.

Kuna venekeelsete koolide õpilastel on kodukeeleks ainult vene keel, siis statistiliselt analüüsitakse eestikeelsete koolide õpilaste kodukeele ja vanemate sünnimaa seoseid.

Kodukeele seos (korrelatsioon) lapsevanemate sünnimaauga on statistiliselt nõrk. Seda näitavad ka T-testi ja χ^2 -testi tulemused:

- seos ema puhul: $r = 0,21$ ($p = 0,004$), $t = 1,846$ ($p = 0,066$), $\chi^2 = 56,12$ ($p = 0,000$);
- seos isa puhul: $r = 0,34$ ($p = 0,001$), $t = 1,619$ ($p = 0,12$), $\chi^2 = 51,27$ ($p = 0,000$).

Vanavanematega suheldes kasutasid vene koolides õppivad õpilased eranditult ainult vene keelt. Eesti õppekeelega koolide õpilased kasutasid ema- ja isapoolsete vanavanematega suheldes nii eesti kui ka vene keelt. Eesti õppekeelega koolide 146 õpilasel oli emapoolsete vanavanematega kõneldes suhtluskeeleks eesti keel, vene keel aga 66 õpilasel ning nii eesti kui ka vene keelt rääkis nendega 16 õpilast. Isapoolsete vanavanematega kõneldi eesti keelt (150 õpilast), vene keelt (60 õpilast) ja paralleelselt nii eesti kui ka vene keelt (18 õpilast). Juhul, kui vanavanemad olid mingist muust rahvusest kui eestlane, siis oli suhtluskeeleks ainult vene keel. Sellest nähtub, et keelevahetuse protsess on alanud juba varasemas põlvkonnas ning muud keeled on sulandunud vene keelde. Päritolumaa ei mõjuta antud juhul suhtluskeele valikut, kuid seda mõjutab eesti õppekeelega koolides õppijate kodukeel.

Vanavanematega kõneldava keelega on kodukeel statistiliselt tugevas korrelatsioonis:

- seos ema vanematega: $r = 0,75$ ($p = 0,000$), $t = 27,75$ ($p = 0,000$), $\chi^2 = 146,00$ ($p = 0,000$);
- seos isa vanematega: $r = 0,77$ ($p = 0,000$), $t = 27,65$ ($p = 0,000$), $\chi^2 = 146,00$ ($p = 0,000$).

Suuruselt kolmas päritoluriik oli Ukraina, ometi ukraina keelt kodukeelena ei kasutatud. See näitab, et ukraina keel on Kohtla-Järve piirkonnas hääbunud. Muuhulgas võib siin leida segaabelude tugevat mõju. Lisaks on immigrandid siia elama asunud ammu, seega pole enam tegemist uusimmigrantidega ning ka see on üks tegur, mis on soodustanud nende venestumist.

Põlvkondadevahelist keelekasutust vaadeldes on näha, et keel on hakanud hääbuma juba varasemas põlvkonnas ja see tendents on põlvest põlve üha süvenenud. Kui leidub piisav hulk õpilasi, kes sooviks vastavat keelt süvendatult õppida, võiks luua kas või valikaine korras koolide juurde õpperühmi, et tõkestada ukraina keele kadumist selles piirkonnas.

Väljaspool kodu kasutasid riigikeelt eesti koolides õppivatest lastest kokku 177, kõik vene õppekeelega koolides õppijad suhtlesid koolis ja sõpradega vene keeles ning lisaks neile kõnelesid 51 eesti koolides õppivat õpilast väljaspool kodu vene keeles, kokku 825 õpilast. Kaasõpilastega lävides oli suhtluskeeleks üldjuhul vene keel, eesti õppekeelega koolides aga 185 juhul eesti keel (eestikeelsetest õpilastest 81,14%). Sõpradega kõneldes kasutati eesti keelt veel harvem (15,67% kogu informantide hulgast). Seega domineerib vene keel märkimisväärselt teiste keelte üle, kusjuures muid keeli informandid väljaspool kodu ei kasutanud.

3.2. Keelteoskuse tase

Õppijate keele-eelistust (keele valdamise nelja osaoskuse põhjal) iseloomustavad järgmised asjaolud. Oma emakeelt, mida ei kasutata kooli õppekeelena, oskavad sisserännanud suuliselt paremini kui kirjalikult. Vene õppekeelega koolide lapsed ja eesti koolide eesti taustaga peres elavad õpilased suudavad oma emakeeles hästi ka lugeda ja kirjutada. Seda ei saa öelda eesti koolides õppivate vene õpilaste kohta, kelle kodukeeleks on vene keel: osa neist oskab oma emakeeles lugeda ja kirjutada vaid üsna madalal tasemel. Seda võib seletada asjaoluga, et pereliikmetega suheldes kasutatakse kodukeelt suuliselt, samal ajal kui (eesti) koolis seda ei õpetata või mängivad selle omandamisel rolli muud tegurid. Lugemis- ja kirjutamisoskust emakeeles saaks arendada ka pereringis, kus tulemust mõjutab eelkõige see, mil määral vanemad väärtustavad lapse emakeelt.

Silma torkab ja suunab edasisele arutelule tõik, et Kohtla-Järvel ei toonud uuritud õpilaste hulgast ükski kodus kasutatava keelena välja selle riigi keelt, kust on sisse rännanud nende vanemad (v.a Eesti- ja Venemaa). Kuna kirjaoskuse arendamisel on eriti oluline funktsioon keele säilitamise piiudel, võiks lapsi õpetada nende kodukeeles ka lugema ja kirjutama ning võtta arutluseks nende keelte õpetamise. See toetaks muuhulgas ka teiste keelte õpet ning õpilase positiivse mina-pildi kujunemist.

Kõige paremini vallatavaks keeleks pidasid õpilased vene keelt (1002 informandist 809), vrd eesti keelt oskas kõige paremal tasemel 193 õpilast. Eesti keelt eelistas kasutada 18%, ülejäänud kasutasid vene keelt (81%). Seega on nende

õpilaste keele-eelistuseks vene keel. Peale selle õpib 949 last inglise keelt, saksa keele õppijaid on vähem – 53. Soovitakse kindlasti täiustada oma teadmisi nii eesti, inglise kui ka saksa keeles. Lisaks soovitakse õppida veel ka prantsuse, itaalia, hispaania, soome ja rootsi keelt ning paar õpilast oleks soovinud õppida jaapani ja araabia keelt. Juba I kooliastme õpilaste seas on populaarseks saanud Euroopa Liidu keeled, eriti prestiižseks loetakse inglise ja prantsuse keele tundmist.

4. Maardu linna vähemuskeelte uuring

Maardu linnas uuriti vene õppekeelega Maardu Gümnaasiumi algklasside õpilasi ja nende vanemaid. Esiteks anketeeriti neid selleks, et välja selgitada vähemuskeelt kodukeelena kõnelevad õpilased. Teise etapina intervjueriti neid suuliselt (9 last). Lapsevanemad vastasid oma abikaasade/elukaaslaste ja omakorda mõlema vanemate kohta käivatele küsimustele.

Uuringus osales 177 õpilast (sh 84 tüdrukut ja 93 poissi). Neist 9 lapse kodukeel oli (lisaks vene keelele) mingi muu keel peale eesti keele. Lapsevanemaid osales 152 (sh 133 ema ja 19 isa). Uuringus osalenud informandid on pärit 11 riigist (vt tabel 2). Suurem osa on sündinud Eestis: 96,6% õpilastest ning isadest 64,9% ja emadest 60,5%.

Tabel 2. Õpilaste ja nende vanemate sünnimaad

Sünnimaa	Õpilased		Isad		Emad	
	Arv	%	Arv	%	Arv	%
Eesti	171	96,6	107	60,5	115	64,9
Venemaa	2	1,1	46	25,9	41	23,1
Ukraina	4	2,3	12	6,8	11	6,2
Valgevene			3	1,7	3	1,7
Leedu			1	0,6	1	0,6
Läti			5	2,8		
Kasahstan					4	2,3
Armeenia					1	0,6
Ungari					1	0,6
Kõrgõzstan			2	1,1		
Aserbaidžaan			1	0,6		
Kokku	177	100	177	100	177	100

Suurema osa väljaspool Eestit sündinud vanemate päritolumaaks on Venemaa ja Ukraina. Venemaal sündinud isad ja emad on vastavalt 46 ja 41 ning Ukrainas vastavalt 12 ja 11. Õpilastest oli Venemaal sündinud 2 last ja Ukrainas 4.

Lapsevanemate rahvusteks olid kolme suurema rühmana venelased (isad 138 ja emad 150), ukrainlased (15 isa ja 11 ema) ning valgevenelased (13 isa ja 9 ema). Lisaks oli küsitlute seas 6 eestlasest isa ja 4 ema. Leedulasi oli küsitletud lapsevanemate hulgas 1 (ema), tatarlasi 2 (1 isa, 1 ema), armeenlasi 1 (isa), poolakaid samuti 1 (isa). Lisaks oli informantide seas esindatud 2 soomlast (1 isa ja 1 ema).

4.1. Sünnimaa ja kodukeele seos

Maardu õpilased kasutavad 91,5% ulatuses kodukeelena vene keelt, 8,5% õpilasi tarvitab kodukeelena vene keele kõrval mingit muud keelt (vt tabel 3). Segaperedes on tavaline, et koduse keelena kasutatakse vene keelt (Ü. Rannut 2002). Kodus kasutatavateks teisteks keelteks on ukraina, leedu, tatari ja eesti keel. Samuti kõneldakse eesti keeles eestikeelse taustaga vanema ja vanavanematega (eesti ja vene keeles segamini), kuid õdede ja vendadega kõnelemisel eelistatakse vene keelt, kuna seda keelt vallatakse kõige paremini. Need vastused on tüüpilised vene keelekeskkonnas viibivatele lastele. Vanemate vastustest selgub, et ka nemad on õppinud vene õppekeelega koolis ning osa neist on sündinud Venemaal.

Tabel 3. Maardu Gümnaasiumi õpilaste kodukeel

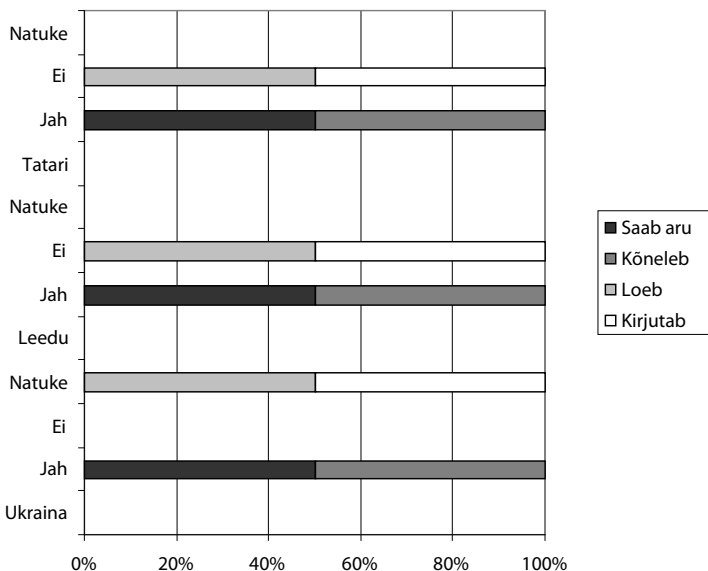
Keel	Õpilased		Märkused
	Arv	%	
eesti	6	3,4	3 ema eestlane, 3 isa eestlane
vene	162	91,5	venelased, valgevenelased, ukrainlased jt
ukraina	7	3,9	6 ema ukrainlane, 1 isa ukrainlane
leedu	1	0,6	ema leedulane
tatari	1	0,6	mõlemad vanemad tatarlased

Vene keele järel moodustab suuruselt teise rühma ukraina keel. Oma teiseks kodukeeleks pidas ukraina keelt 3,9% õpilastest (7), neist kuuel oli ema ukrainlane, ühel isa. Emadest kõneles ukraina keeles 9, isadest 1. Ukrainas sündinud lapsevanematest kõneles 43,5% kodus ukraina keeles, kuid teise keelena. Põhikeeleks jäi vene keel. Väljaspool kodu ei kasutanud ukraina keelt ei lapsed ega ka nende vanemad.

Kui võrrelda põlvkondadevahelist keelekasutust, siis võib märgata, et ukraina keele kasutamine on aja jooksul vähenenud. Osa lapsevanemaid (3 ema) oli kõnelema lapsepõlves oma vanematega ukraina keeles, kuid nad olid õppinud vene õppekeelega koolis, ja oma lastega kõneldakse praegu ainult vene keeles. Suur osa küsitletud Ukrainas sündinud lapsevanematest suhtleb ka oma vanematega vene keeles (3 ema ja 10 isa). Võimalik, et ukraina keele toetamiseks oleks vaja õpperühmi või klasse, et takistada ukraina keele lõplikku sulandumist vene keelde.

4.2. Vähemuskeelte oskuse tase

Kodukeele pädevust on vaadeldud ukraina, leedu ja tatari keele puhul skaalal arusaamine–rääkimine–lugemine–kirjutamine (vt joonis 2). Ukraina keelest said aru ja oskasid selles keeles end suuliselt väljendada kõik 7 õpilast, lugeda ja kirjutada oskasid nad ainult natuke. Leedu ja tatari keele puhul saadi nendest keeltest aru ja osati neid kõnelda, kuid lugeda ega kirjutada mitte.



Joonis 2. Vähemuskeele valdamine

Seega on nende kodukeelte suuline oskus (arusaamine ja rääkimine) parem kui kirjalik (lugemine, kirjutamine). See on tingitud sellest, et kodus kasutatakse keelt suuliselt ja neid keeli koolis ei õpetata. Lugemis- ja kirjutamisoskuse tase sõltub põhiliselt sellest, kas keelt koolis õpetatakse, samuti on oluline, kas nende keelte valdamist kõrgemal tasemel peavad lapsevanemad tähtsaks. Keele säilitamise eeldus on kirjaoskus ning seepärast võiks sisse seada teatud keeleõppe tugirühmad, mis aitaksid omandada lugemis- ja kirjutamisoskuse vastavates keeltes.

4.3. Vajadus vähemuskeelte õppe järele

Lapsevanematelt küsiti arvamust vähemuskeelsete õpperühmade või klasside vajalikkuse kohta. Vastused olid valdavalt eitavad, kusjuures domineeris praktiline vaatepunkt, näiteks:

- “Eestis elades pole sellist vajadust”
- “Minu laps ei lähe arvatavasti kunagi elama Ukrainasse”
- “Ukraina keel pole Eestis tähtis”

Küsimuse puhul, kas lapsevanemad saadaksid oma lapse õppima, kui lähedal oleks või koolis avataks vähemuskeele õpperühm, vastas üks lapsevanem eitavalt (“Sellel pole mõtet, kuna ukraina kultuur ei erine oluliselt vene kultuurist, seega saab laps vajalikud teadmised kultuurist ja keelest ka vene õppekeele koolis, ukraina kultuuri võib aga õpetada kodus”), aga ülejäänud pidasid säärase õpperühma avamist heaks mõtteks. Osa lapsevanemaid arvas, et päritolukeelt oleks mõistlik õppida valikaine raames paar korda nädalas. Mõned rõhutasid juurte tundmist, mis toetaks lapse kultuuriidentiteedi säilitamist, ning mõned apelleerisid keeleõigusele – õigusele tunda oma etnilist kuuluvust, kultuuri ja keelt.

Seega leiab osa Eestis elavaid ukrainlasi, et siin elades ei ole ukraina keel vajalik ega tähtis. Osa arvab isegi, et ukraina kultuur ei erine märkimisväärselt vene omast. Rohkem leidub siiski neid, kes tahavad, et nende lapsed õpiksid tundma oma esivanemate keelt ja kultuuri. Seega oleksid ilmselt vajalikud ukraina keele õpperühmad või -klassid õpilastele, kui leidub vajalik hulk seda õppida soovijaid.

Kodukeelena oli nimetatud ka leedu ja tatari keelt (vt tabel 3). Leedu keele puhul oli ema leedulane (sündinud Leedus), isa aga venelane. Koduse keelena kasutati vene keelt ja teise keelena leedu keelt. Õpilane ja tema isa olid sündinud Eestis. Selle perekonna ema oli oma lapsepõlves koduse keelena kasutanud nii vene kui ka leedu keelt ja õpetas ka oma lastele leedu keelt. Tema laps kasutab ema ja õega suheldes nii vene kui ka leedu keelt. Lapsed kõnelevad ema vanematega leedu ja veidi ka vene keeles, kuid isapoolsete vanavanematega rääkides kasutatakse vene keelt. Küsimusele, mis puudutas leedu keele osaoskuste valdamise taset, vastas õpilane, et saab leedu keelest aru ja oskab seda rääkida, kuid lugeda ega kirjutada ei oska. Siit nähtub, et selles segaperekonnas, kus isa on venelane, kasutatakse kodukeelena vene keelt ja teise keelena leedu keelt. Ühel juhul, kui isa oli leedulane ja ema venelane, kasutati kodukeelena ainult vene keelt. Seega on ema mõju kodukeele valikule antud uurimisaluste puhul suurem kui isal. Leedulasest ema oma lapsi leedukeelse õppega kooli Eestis õppima ei paneks, kuna lapsed valdavad vene keelt paremini ja kodus kõneldakse samuti rohkem vene keeles – seega on lastel kergem õppida vene õppekeelega koolis. Samas oleks ta nõus saatma oma lapse koolis õppima leedu keelt valikainena (“Lapsel oleks vaja teada ka vanavanemate keelt ja kultuuri, kuna see aitab lapsel kultuursemaks inimeseks kasvada ja oma päritolu paremini teadvustada ning selle üle ka uhkust tunda”).

Tatari keel oli küsitlute hulgas esindatud kogu perekonna ulatuses. Mõlemad uuritud vanemad olid tatarlased, kuid sündinud Eestis. Õpilase vanavanemad (nii ema- kui ka isapoolsed) olid tulnud Eestisse tööd otsima 1949. aastal. Vanemad olid õppinud vene õppekeelega koolis, kuid koduseks keeleks kujunes välja vene keele kõrval ka tatari keel. Perekonna ema ja isa kõnelesid oma vanematega nii vene kui ka tatari keeles, samuti suhtles õpilane nii ema, isa, venna ja emapoolsete vanavanematega kui ka isapoolsete vanavanematega vene ja tatari keeles. Selle perekonna puhul on näha, kui hästi on säilinud järjepidevus tatari keele kõnelemisel põlvkondade kaupa. Kaugeleulatuvaid järeldusi ei saa muidugi teha, kuna informantide hulk on väike, ainult 0,6% selles koolis kõneldavatest keeltest, kuid suundadest annab see siiski aimu.

Tatari keele õppimise kohta arvas lapsevanem, et tatari keelega pole Eestis midagi peale hakata. Kui aga oleks olemas huvialagrupp või avataks vastav valikaine koolis, siis sinna paneks lapsevanem oma lapse õppima, et ta tutvuks tatari kultuuriga ja õpiks tatari keelt, “sest enda päritolu ja juuri tundev inimene on haritud inimene”. Tatari keelest saab uuritud õpilane aru ja ka räägib seda, kuid lugeda ega kirjutada ei oska.

Kahel juhul nimetasid lapsevanemad valgevene keelt ja ühel juhul tšuvaši keelt. Ise olid nad neid kasutanud ainult oma lapsepõlvkodus kodukeelena, nad olid õppinud venekeelses koolis ja venestunud. Oma lastele nad oma keelt enam ei õpetanud, kodukeeleks oli vene keel.

Euroopa Liidu keeltest nimetasid lapsevanemad kahel juhul soome ja ühel juhul poola keelt, milles nad olid lapsepõlves kõnelenud oma vanematega, kuigi olid õppi-

nud vene õppekeelega koolis. Koduse suhtluskeelena on neil praegu kasutusel vene keel ja oma rahvuseks märkisid nad: venelane. Siin on näha, et keele järjepidevus on kadunud. Lisaks kõneldi väljaspool kodu kõikidel juhtudel vene keeles.

5. Kohtla-Järve, Maardu ja Tallinna linna õpilaste vähemuskeelte võrdlus

Uuringus osalenud Kohtla-Järve õpilased on pärit kahest ja lapsevanemad 16 riigist. Suurem osa informantidest on sündinud Eestis. Seega kuulusid õpilased teise põlvkonna immigrantide hulka. Venemaal sündinuid oli 5 õpilast, 130 ema ja 132 isa. Maardu õpilased ja nende vanemad on pärit 11 riigist. Suurem osa informantidest on sündinud Eestis ja enamik õpilastest on seega samuti teise põlvkonna sisserännanud. Võrreldes Tallinnas korraldatud uuringu tulemustega (vt Ü. Rannut, M. Rannut 2007) selgub, et Tallinna õpilased on sündinud koguni 28 ja nende vanemad 52 riigis. Tallinna tööturg on laiem kui kusagil mujal Eestis, sealjuures on see atraktiivsem ka välismaalastele, kelle kodukeel üldjuhul erineb kohalike keeltest.

Sarnaselt Kohtla-Järve venekeelsete koolide õpilastega kasutavad Maardu Gümnaasiumi 2.–5. klasside õpilased domineeriva kodukeelena vene keelt. Maardus oli kodukeeleks vene keele kõrval ka mingi teine keel (ukraina, leedu, tatari ja eesti keel), kusjuures põhikeeleks jäi siiski vene keel. Võrdluseks: Tallinna õpilased kasutasid kodudes 22 keelt (õpilastest pidas teisi keeli oma kodukeeleks 2%), kusjuures küllalt levinud oli muuhulgas inglise keel, juhul kui mõlemad vanemad või üks neist oli pärit väljastpoolt Eestit. Kolm Tallinna õpilast kasutasid ainsa kodukeelena inglise keelt. Ei Kohtla-Järve ega Maardu õpilaste puhul sellist inglise keele kasutust ei täheldatud, kuigi Kohtla-Järvel oli üks isa pärit Taanist, kolm Soomest ja üks Itaaliast.

Ukraina päritolu lapsevanematest kasutas Tallinnas vaid 7% kodukeelena ukraina keelt, mis näitab samuti nagu Kohtla-Järve ja Maardu uurimusegi puhul märkimisväärset keelevahetuse määra, mille üheks tõukejõuks võivad olla ka segaabiellud. Vastupidise näite võib tuua Tallinnasse elama asunud Aserbaidžaanist pärit perede kohta: 43 lapsevanemast 30 kõneleb kodus aserbaidžaanis keeles, mis on suur arv, võrreldes Eestis elava ukraina kogukonna suurusega (Eesti ... 2000). Aseritest uusimmigrandid näivad uude riiki elama asudes väärtustavat oma keelt rohkem kui ukrainlased. Valdavalt on siiski ka Tallinnas, nagu Kohtla-Järvel ja Maarduski, kodukeeleks vene keel, mida kasutab 61% segaperede õpilasi ja teise keelena 27% (Ü. Rannut, M. Rannut 2007).

Vanavanematega kõnelesid Kohtla-Järve vene koolide õpilased vene keeles. Eesti keeles suheldi eestlastest vanavanematega. Maardu õpilased kõnelesid vanavanematega samuti vene keeles. Õpilased, kelle vanavanemad olid eestikeelse taustaga, kõnelesid nendega paralleelselt eesti ja vene keeles. Üks Maardu õpilane kasutas vanavanematega suhtlemisel ka leedu ja teine tatari keelt. Kui võrrelda põlvkondadevahelist keelekasutust, võib märgata, et ukraina keele kasutamine on põlvest põlve vähenenud. Keelevahetus ei ole toimunud hiljuti, vaid juba varasemas põlvkonnas. Kui vaadelda mitme põlvkonna vahelist keelevahetust, siis kasutas Tallinnas 80% vanavanematest lapselastega kõneldes vene keelt.

Kõigi kolme linna puhul ilmneb, et keelevahetus on toimunud juba eelmises põlves või on väljaränne toimunud Ukraina ja Valgevene venekeelsetest piirkondadest.

6. Kokkuvõte

Uurimusest nähtub väga selgelt, et Eesti venekeelsete piirkondade õpilased kasutavad nii kodus kui ka väljaspool peamiselt vene keelt. Vähemuskeelte osaks Eesti segregatiivsetes keelekeskkondades on käesoleva uurimuse põhjal kadumine. Keelevähemused assimileeruvad ja venestuvad.

Keele elujõulisuse ainus mõõt ei ole keelerühma suurus, vaid olulised mõjurid on ka keele staatus, segaabielse mõju keelevalikule jm. Kusjuures määravaks teguriks on inimeste suhtumine oma päritolukeelde. Keelerühmad varieeruvad oma suhtumise poolest – osa vähemusrahvustest väärtustab enda etnilist kuuluvust ja oma päritolukeelt üha rohkem ning püüab teadmisi oma lastele ka edasi anda, hoolitsedes nõnda keele järjepidevuse ja elujõulisuse eest.

Eesti väikekeeled vajavad Eesti keelepoliitika tuge.

Viidatud kirjandus

- Anderson, Raquel 2004. First language loss in Spanish-speaking children: Patterns of loss and implications for clinical practice. – Brian A. Goldstein (Ed.). *Bilingual Language Development and Disorders in Spanish-English Speakers*. Baltimore: Brookes, 187–212.
- Baker, Colin 2000. *A Parent's and Teacher's Guide to Bilingualism*. 2nd edition. Clevedon, Boston, Toronto, Sydney: Multilingual Matters Ltd.
- Baker, Colin 2006. *Foundations of Bilingual Education and Bilingualism*. 4th edition. Clevedon: Multilingual Matters.
- Baker, Colin; Jones, Sylvia Prys 1998. *Encyclopedia of Bilingualism and Bilingual Education*. Clevedon: Multilingual Matters.
- Baldauf, B. Richard; Kaplan, B. Robert 2006. *Language Planning and Policy in Europe: the Czech Republic, the European Union and Northern Ireland*. Clevedon: Multilingual Matters.
- Bayley, Robert; Schecter, Sandra R. 2003. *Language Socialization in Bilingual and Multilingual Societies*. Clevedon: Multilingual Matters.
- Broeder, Peter; Extra, Guus 1999. *Language, Ethnicity and Education: Case Studies of Immigrant Minority Groups and Immigrant Minority Languages*. Clevedon: Multilingual Matters.
- Commission of the European Communities 2005. *Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions. A New Framework Strategy for Multilingualism*. COM(2005)596 final. http://ec.europa.eu/education/policies/lang/doc/com596_en.pdf (17.12.2009).
- Conklin, Nancy; Lourie, Margaret 1983. *A Host of Tongues*. New York: The Free Press.
- DfEE 1999. *The National Curriculum Key Stages 1 and 2*. DfEE, QCA: London.
- Edwards, Viv 1996. *The Other Languages: A Guide to Multilingual Classrooms*. Reading, England: National Centre for Language and Literacy.
- Eesti rahvaloendus 2000 = Eesti Statistikaameti rahvaloendus 2000. http://www.stat.ee/gatekeeper.stat.ee:8000/px-web.2001/Database/Rahvaloendus_regionaalne/Rahvaloendus_regionaalne.asp (17.08.2009).

- Fishman, Joshua A. 1991. *Reversing Language Shift. Theoretical and Empirical Foundations of Assistance to Threatened Language*. Clevedon, Philadelphia, Adelaide: Multilingual Matters Ltd.
- Gans, Herbert J. 1997. Toward a reconciliation of 'assimilation' and 'pluralism': the interplay of acculturation and ethnic retention. – *International Migration Review*, 31 (4), 875–892. doi:10.2307/2547417
- Giles, Howard; Bourhis, Richard Y.; Taylor, Donald M. 1977. Towards a theory of language in ethnic group relations. – Howard Giles (Ed.). *Language, Ethnicity and Intergroup Relations*. London: Academic, 307–348.
- Giles, Howard; Byrne, Jane L. 1982. An intergroup approach to second language acquisition. – *Journal of Multilingual and Multicultural Development*, 3 (1), 17–40.
- Goldenberg, Claude; Rueda, Robert S.; August, Diane 2006. Sociocultural contexts and literacy development. – Diane August, Timothy Shanahan (Eds.). *Developing Literacy in Second-language Learners. Report of the National Literacy Panel on Language-Minority Children and Youth*. National Literacy Panel on Language-Minority Children and Youth (U.S.), National. Mahwah, New Jersey: Lawrence Erlbaum Associates, 247–346.
- Helemäe, Leena; Plotkin, Aleksander; Semjonov, Aleksei; Vöörmann, Rein 2000. Identiteedi kujunemise probleemid post-sotsialistlikus keskkonnas (Tallinna rahvusvähemuste kogemuse alusel). *Uuringu aruanne, Tallinn 1999–2000*. Tallinn: Inimõiguste Teabekeskus.
- Hoffmann, Charlotte; Ytsma, Jehannes (Eds.) 2003. *Trilingualism in family, school, and community*. Language Arts & Disciplines. Clevedon: Multilingual Matters Ltd.
- Holmes, Janet; Aipolo, Anahina 1993. The Wellington Tongan community: Prospects for language maintenance. – *Wellington Working Papers in Linguistics*, 1, 1–16.
- Hudson, Alan 2001. Diglossia. – Rajend Mesthrie (Ed.). *Concise Encyclopedia of Sociolinguistics*. Oxford: Elsevier Science.
- Hyltenstam, Kenneth; Stroud, Christopher 1996. Language maintenance. – Hans Goebel, Peter H. Nelde, Zdenek Sary, Wolfgang Wölk (Eds.). *Contact Linguistics: An International Handbook of Contemporary Research*. Berlin: Walter de Gruyter, 271–286.
- Iskanius, Sanna 2005. Venäjänkielisten maahanmuuttajaopiskelijoiden kieli-identiteetti. *Jyväskylä Studies in Humanities*, 51. Jyväskylä: Jyväskylän Yliopisto.
- Kaur, Sukhwant; Mills, Richards 1993. Children as interpreters. – Richard W. Mills, Jean Mills (Eds.). *Bilingualism in the Primary School*. London: Routledge, 112–129.
- Kohnert, Kathryn; Yim, Dongsun; Nett, Kelly; Kan, Pui Fong; Duran, Lilian 2005. Intervention with linguistically diverse preschool children: a focus on developing home language(s). – *Language, Speech, and Hearing Services in Schools*, 36, 3, 251–263. doi:10.1044/0161-1461(2005/025)
- Kohtla-Järve ametlik kodulehekülj. <http://www.kjlv.ee/?lang=et&page=commoninformation/history> (11.08.2009).
- Komondouros, Markos; McEntee-Atalianis, Lisa 2007. Language attitudes, shift and the ethnolinguistic vitality of the Greek orthodox community in Istanbul. – *Journal of Multilingual and Multicultural Development*, 28, 5, 365–384. doi:10.2167/jmmd483.1
- Krauss, Michael 1992. The World's languages in crisis. – *Language*, 68, 4–10.
- Li, Peter S. 2001. The economics of minority language identity. – *Canadian Ethnic Studies*, 33 (3), 134–154.
- Maardu linna ametlik lehekülj. <http://www.maardu.ee/index.php?page=65&> (28.08.2009).
- Mills, Jean 2001. Being bilingual: perspectives of third generation Asian children on language culture and identity. – *International Journal of Bilingual Education and Bilingualism*, 4 (6), 383–402.
- Moon, Bob; Ben-Peretz, Miriam; Brown, Sally 2000. *The Routledge International Companion to Education*. London, New York: Routledge.

- Muldma, Maia (Ed.). 2009. Dialogue of cultures – possibility or inevitability? II. / Kultuuride dialoog – võimalus või paratamatus? II. Tallinn: Tallinn University Press.
- Pavlenko, Aneta 2004. “Stop doing that, *ja komu skazala*”: language choice and emotions in parent-child communication. – *Journal of Multilingual and Multicultural Development*, 25 (2, 3), 179–203. doi:10.1080/01434630408666528
- Pedraza, Pedro; Pousada, Alicia 1992. Bilingualism in and out of school: ethnographic perspectives on the determination of language dominance. – Marietta Saravia-Shore, Steven F. Arvizu (Eds.). *Cross-cultural Literacy: Ethnographies of Communication in Multiethnic Classrooms*. New York: Garland Publishing, 253–272.
- Piller, Ingrid 2002. *Bilingual Couples Talk: the Discursive Construction of Hybridity*. Amsterdam: John Benjamins.
- Rannut, Mart; Rannut, Ülle; Verschik, Anna 2003. Keel, võim, ühiskond. Tallinn: Tallinna Pedagoogikauilikooli kirjastus.
- Rannut, Ülle 2002. Muukeelsete õpilaste integratsioon eesti koolis. Kohtla-Järvel, Tallinnas, Valgas ja Sindis muukeelsete õpilastega töötavate õpetajate, muukeelsete õpilaste ja nende vanematega läbi viidud uuringu tulemused 20.03–08.05.2002. Tallinn: Mitteestlaste Integratsiooni Sihtasutuse Haridusprogrammid.
- Rannut, Ülle 2005. Keelekeskkonna mõju vene õpilaste eesti keele omandamisele ja integratsioonile Eestis. Tallinna Ülikooli humanitaarteaduste dissertatsioonid, 14. Tallinn: Tallinna Ülikooli kirjastus.
- Rannut, Ülle; Rannut, Mart 2007. Tallinna õpilaste kodukeel. – *Haridus*, 3–4, 7–10.
- Seville-Troike, Muriel 2000. Causes and consequences of language maintenance/shift. – Elite Olshtain, Gabriel Hornczyk (Eds.). *Language, Identity, and Immigration*. Jerusalem: The Magnes Press, Hebrew University, 159–171.
- Shapiro, Daniel, M.; Stelcner, Morton 1997. Language and earnings in Quebec: trends over twenty years, 1970–1990. – *Canadian Public Policy/Analyse de politiques*, 23 (2), 115–140.
- Simpson, Alan 1991. The uses of “cultural literacy”: a British view. – *Journal of Aesthetic Education*, 25, 4, 65–73. doi:10.2307/3332904
- Siraj-Blatchford, Iram; Clarke, Priscilla 2000. *Supporting Identity Diversity and Language in the Early Years*. Buckingham: Oxford University Press.
- Skutnabb-Kangas, Tove 2000. *Linguistic Genocide in Education – or Worldwide Diversity and Human Rights?* Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Starks, Donna 2005. The effects of self-confidence in bilingual abilities on language use: perspectives on Pasifika language use on South Auckland. – *Journal of Multilingual and Multicultural Development*, 26, 6, 533–550. doi:10.1080/01434630508668424
- Zentella, Ana Celia 1999. *Growing up Bilingual*. Malden, MA: Blackwell.
- Tuominen, Anne 1999. Who decides the home language? A look at multilingual families. – *International Journal of the Sociology of Language*, 140, 59–76. doi:10.1515/ijsl.1999.140.59
- Urzúa, Alfredo; Gómez, Edwin 2008. Home style Puerto Rican: a study of language maintenance and use in New England. – *Journal of Multilingual and Multicultural Development*, 29 (6), 449–466. doi:10.1080/01434630802147999
- Vahtin 2004 = Вахтин, Николай 2004. Социоллингвистика и социология языка. – *Язык и общество*. Санкт-Петербург: Гуманитарная Академия, 33–39.
- Van Tuijl, Cathy; Leseman, Paul P. M.; Rispens, Jan 2001. Efficacy of an intensive home-based educational intervention programme for 4- to 6-year-old ethnic minority children in the Netherlands. – *International Journal of Behavioral Development*, 25, 148–159. doi:10.1080/01650250042000159
- Verschik, Anna 2004. Koodivahetus meil ja mujal. – *Keel ja Kirjandus*, 1, 25–45.

- Wiley, Terrence, G. 1996. Language planning and language policy. – Sandra L. McKay, Nancy H. Hornberger (Eds.). *Sociolinguistics and Language Teaching*. Cambridge: Cambridge University Press, 103–147.
- Wong Fillmore, Lily 1991. When learning a second language means losing the first. – *Early Childhood Research Quarterly*, 6, 323–346. doi:10.1016/S0885-2006(05)80059-6

Elvira Küüni (Tallinna Ülikool) peamised uurimised on Eesti venekeelse elanikkonna etniline ja keeleline identiteet, vene emakeelega õpilaste kohanemine eestikeelses üldhariduskoolis ja töökeskkonnas, Eesti vähemuskeelte arengutendentsid.
elvira.kuun@tlu.ee

THE VIABILITY AND DEVELOPMENT TRENDS OF ESTONIAN MINORITY LANGUAGES IN SEGREGATIVE LANGUAGE ENVIRONMENTS

Elvira Küün

Tallinn University

The language policy of the European Union promotes diversity. Consequently, it is an important matter for Estonia as one of its member states to find out what the current language situation in Estonia is. It is necessary to get an overview of what languages the language minority groups in Estonia use in the public sphere and in informal communication, and whether there is a need to teach minority languages more extensively at Estonian schools. The government's education policy should take this information into account.

According to the Estonian population registry—the official source of information on what the people in Estonia claim to be their mother tongues—the 142 ethnic groups living in Estonia speak 109 different mother tongues (the biggest minority group is Russian). Unfortunately, the register does not contain more specific data: for instance on what languages are used at home, i.e. in the informal environment. The speakers of minority languages in Estonia are immigrants who have mainly come to live and work here in the industrial areas and are currently forming segregated groups. Although at present there are quite many Sunday schools operating in Estonia where the children of minority language group representatives are taught their mother tongue and their culture of origin, there still remains a threat to their languages' existence in Estonia. Common reasons for this are cultural pressure and the lower prestige the minority language has in the eyes of the speakers themselves. The speakers' number is not always decisive here—their attitude towards their mother tongue plays the main role in this case.

Language transfer between generations can be predicted from the primary school pupils' language use. Therefore, the language choices of the school pupils of this age and their parents were studied in the article. The results of the survey showed that the oral competence of the minority language group representatives in the minority language is better than their literacy. However, literacy has the key role in language maintenance. The non-Russian minority language speakers of the Estonian segregative language environments tend to assimilate in the Russian community. In order to slow down the assimilation process of different population groups the article suggests to consider offering facultative subjects on the minority language and culture at the schools where a sufficient number of children from a minority language are taught.

Keywords: education policy, language policy, language choice, language competence, home language, Estonian, Russian, other languages

NARRATIIVI AEG JA TÕLKE AEG: DISKURSIIVNE KÄSITLUS

Triin Lõbus, Silvi Tenjes

Ülevaade. Kognitiivses narratoloogias mõistetakse ilukirjandusliku narratiivi interpreteerimist lugeja poolt kui loomaailmast mentaalsete representatsioonide konstrueerimist tekstis leiduvate keeleliste markerite põhjal. Käesolevas artiklis uuritakse aspektikategooria rolli selles protsessis, võrreldes hispaaniakeelseid narratiive nende eestikeelsete tõlgetega. Lähtutakse sotsiokognitiivsest käsitlusest, mille kohaselt kirjanduslikku narratiivi vaadeldakse sotsiokultuurilise praktikana, mis eeldab teatud konventsioonide järgimist tähenduse loomisel. Analüüsitakse nn narratiivse imperfekti rolli narratiivi ajalise mõõtme kujundamisel ning samuti ruumi ja eelkõige vaatepunkti konstrueerimisel hispaania keeles ning käsitletakse võimalusi samasuguste modifikatsioonide väljendamiseks eestikeelses tõlkes.*

Võtmesõnad: diskursusanalüüs, tekstianalüüs, narratoloogia, grammatiline aeg, aspekt, eesti keel, hispaania keel

1. Sissejuhatus

Käesolevas artiklis uurime ajakäsitlusi ilukirjanduslikus narratiivis hispaaniakeelse originaali ja eestikeelse tõlke võrdluses. Analüüsime nn **narratiivse imperfekti**¹ rolli hispaania keele jutustamistehnilise ja stilistilise vahendina ning probleeme ja võimalusi kõnealuse vormi edasiandmisel eesti keeles nii, et tõlke lugeja võiks tekstis kodeeritud loomaailma kogeda võimalikult samal viisil kui autor on selle konstrueerinud hispaania keeles. Narratiivseks imperfektiks nimetatakse romaani keeltes narratiivses diskursuses esinevat eripärast aspektikasutust, mille puhul imperfektiivset verbivormi kasutatakse selliste sündmuste kujutamiseks, mille ajalis-aspektiline iseloom eeldaks perfektiivset aspekti.

Laiemas mõttes käsitleb meie analüüs küsimust, missuguseid aja kategoriseerimise põhimõtteid rakendavad inimesed oma intuitiivses maailmanägemises ja

* Artikkel on valminud Eesti Teadusfondi grandid nr 8008 toel. Autorid tänavad kaht anonüümset retsensenti väärtusliku tagasiside eest.

¹ Erinevalt eesti grammatikatraditsioonist, kus terminit *imperfekt* kasutatakse üldise lihtmineviku tähenduses, mõistame käesolevas artiklis selle all hispaania keele imperfektiivse aspektitähendusega minevikuvormi.

kui palju sellest annab ette konkreetne keel. Aeg ja ruum on maailma eksistentsi määraavad parameetrid ja inimliku kogemuse põhivormid. Inimene juhindub oma teadvuses ja tegevuses ühtedest või teistest maailmapildi põhikategooriatest, mis on kätketud keelde, ja mõelda maailmast nende kategooriateta on niisama võimatu nagu mõelda väljaspool keele kategooriaid (Gurevitš 1992). Kui kirjandusteksti autor on kategoriseerinud teatud seotud sündmusi keeleliste kategooriate vahendusel, võib valikut käsitleda kategoriseerimisaktina, mis näitab keelekasutaja mõtlemisteid.

Narratiivi mõistmine tekstis leiduvate keeleliste markerite ja nende põhjal tehtavate järelduste abil on üheks peamiseks uurimisobjektiks nn kognitiivses narratoloogias, kus seda protsessi käsitletakse kui **loomaailma** (re)konstrueerimist. See tähendab, et narratiivi interpreteerimise käigus lugeja loob ja aktualiseerib mentaalseid mudeleid maailmast, millest loos räägitakse. Loomaailma võibki vaadelda kui globaalset mentaalset representatsiooni sellest, kes tegi mida kellele, millal, kus, miks ja mil viisil. (Herman 2000, 2002)

Hispaania ja eesti keele võrdlus on huvipakkuv sellepärast, et hispaaniakeelse narratiivi autor kategoriseerib sündmusi reeglipäraselt grammatilise aspektikategooria vahendusel, mis eesti keele ajakogemuses samal viisil ei kajastu, kuid mis mängib olulist rolli narratiivi ajalise struktuuri väljendamisel ja interpreteerimisel. Oma analüüsis vaatlemegi narratiivse imperfekti rolli loomaailma ajastruktuuri konstrueerimisel. Samuti pöörame tähelepanu ajastruktuuri seostele ruumisuhete ja narratiivi **vaatepunktiga**, kust loomaailma sündmusi implitsiitselt nähakse, kogetakse, mõistetakse ja hinnatakse. Põhieristusena võib vaatepunkt asuda kas loomaailma sees või sellest väljaspool (Uspenskij 1970). Esimesel juhul on vaatepunkt omistatav mõnele tegelastest, kujutades loomaailma sündmusi tema subjektiivsest füüsilisest ja psühholoogilisest perspektiivist lähtuvalt. Teisel juhul luuakse objektiivne vaatepunkt, mis võib olla rohkem või vähem piiratud, nn kõikeadvast jutustajast kuni kõrvalseisva vaatleja perspektiivini, kust esitatakse vaid kujuteldavas loomaailmas väliselt nähtavat.

Traditsiooniliselt on grammatilise aja ja ka aspekti käsitlemisel lähtunud Hans Reichenbachi (1947) analüüsist, kirjeldades kõnealuste vormide semantikat kõnehetke, sündmuse hetke ja viitehetke vaheliste suhete kaudu (Fleischman 1991, Comrie 1993). Eesti keele verbiaegu on uurinud eelkõige Helle Metslang (1994), kes lähtub samuti samast formalistlikust suunast ning piirdub oma käsitluses lihtlausega; eesti keele ajakasutust tekstis võrdluses soome keelega on käsitlenud Valma Yli-Vakkuri (1994). Erinevalt Reichenbachist lähtuvast traditsioonist on käesoleva analüüsi aluseks kognitiivne lähenemisviis, milles võetakse arvesse ka ilukirjandusliku narratiivi kui diskursuseliigi sotsiokultuurilisi aspekte. Meie käsitlus kognitsioonist ühtib Lenny Shedletsky definitsiooniga, mille järgi kognitsioon hõlmab 1) mitmesuguseid mentaalseid entiteete (ingl *mental entities*, eesti keeles ka ~ meeles paiknevaid), nagu mõtted, tähendused, ideed, hoiakud, teadmised, uskumused, tajumused, kavatsused, mälestused, kujutluspildid, unistused, meelesisendid, teadvus ja alateadvus, ning 2) mentaalseid operatsioone (*the operations of the mind*, eesti keeles ka ~ meeles teostatavaid), nagu seostamine, infootsing, võrdlemine, tähelepanu ja järeldamine (Shedletsky 2008: 541). Käesolevas töös peetakse kognitiivse lähenemisviisi all silmas looja- ja lugejapoolset loomaailma konstrueerimist, narratiivi esitust ja interpreteerimist.

2. Aeg, aspekt ja narratiiv

Mitmed uurijad (Bruner 1986, Talmy 1995, Herman 2000) on oletanud, et narratiiv üldises mõttes kui teatud mõtlemise strateegia on omaette kognitiivne süsteem, mis võimaldab meil kogemust struktureerida ja mõista. Narratiivse kognitiivse süsteemi funktsioon on seostada ja integreerida teadvussisu teatud komponendid läbi aja koherentseks struktuuriks (Talmy 1995: 422), muutes meie reaalsuskogemuse mõistetavaks ja manipuleeritavaks.

Olulisim narratiivi defineeriv omadus on just selle ajalikus. Narratiivi võib defineerida kui sündmuste järgnevuse järjestikuselt organiseeritud representatsiooni (Genette 1972, Chatman 1990). Käesoleva artikli uurimisteema seisukohalt on oluline, et ajaline järgnevus on lahutamatult seotud sündmuste aspektuaalse analüüsiga. Germaani ja romaani keelte ajasüsteemis ei ole aspektile pööratud nii palju tähelepanu kui grammatilisele ajale, sest eraldiseisev grammatiline aspektikategooria puudub. Hispaania keeles on verbi morfoloogias avalduvatesse ajatähendustesse kodeeritud nii olevik, minevik kui ka tulevik ning aspekt on grammatilise ajaga põimunud sel viisil, et minevikuvormides on samaaegselt väljendatud ka **perfektiivsus** või **imperfektiivsus**. Traditsioonilises grammatikas käsitletakse neid vorme ühtviisi verbiaegadena, kuid Joan L. Bybee ja Östen Dahl peavad aspektuaalset eristust peamiseks. Hispaania keele, nagu ka teiste romaani keelte ajalise-aspektuaalsed põhikategooriad moodustavad seega kolmiksisüsteemi, kus 'minevikuline perfektiiv' on opositsioonis 'minevikulise imperfektiivi' ja 'mitte-minevikulise imperfektiiviga' (mitte 'olevik' opositsioonis 'perfektiivse mineviku' ja 'imperfektiivse minevikuga') (Bybee, Dahl 1989: 85). Käesolevas analüüsis huvitabki meid perfektiivi-imperfektiivi eristus minevikuvormide vahel, sest need on peamised ilukirjanduslikus narratiivis kasutatavad vormid.

Imperfektiivsuse-perfektiivsuse opositsiooni semantikat kirjeldatakse enamasti kas piiritletuse-mittepiiritletuse või totaalsuse-partsiaalsuse kaudu. Viimast käsitusviisi esindab ka Bernard Comrie (1995) paljutsiteeritud definitsioon, mille kohaselt perfektiivne aspekt kujutab sündmust ühtse tervikuna, eristamata seda moodustavaid erinevaid faase (totaalne kujutusviis), ja imperfektiivne aspekt viitab sündmuse sise-misele ajalisele struktuurile, seda vaadeldakse seestpoolt (partsiaalne kujutusviis). Nils B. Thelin (2002) omistab aspektuaalsele analüüsile keskse koha tajus ja kognitsioonis, mille aluseks on meid ümbritseva kaose osadeks jaotamine. See primaarne aspektuaalne analüüs seisnebki osa-terviku eristuses, mille abil eristame ja suhestame omavahel objekte ja sündmusi meid ümbritsevas kaoses. Suhteid objektide vahel mõistetakse kui suhteid ruumis ja muutused nendes suhetes on põhjustatud objektide liikumisest ruumis. Liikumist omakorda mõistetakse toimuvana ajas, kuna ruumis toimuvat muutust reinterpreteeritakse muutusena ühest olukorrast teise (vanast uude). Nii tekib algse sündmuste ajalise-ruumilise ehk aspektuaalse eristamise baasilt primaarne ajaline eristus enne-pärast. Seega sündmuste järjestamine eeldab nende aspektuaalset analüüsi, sest järjestamiseks tuleb need piiritleda alguse ja lõpuga, s.t vaadelda neid totaalselt. Sellega kooskõlas funktsioneeribki aspekt narratiivi ajasuhete väljendamisel (vt jaotis 4). Oluline on märkida, et ajalise-aspektuaalsed suhted ei ole absoluutsed, vaid sõltuvad perspektiivist (Thelin 2002). Kirjanduslikus narratiivis kajastub see autori võimaluses kasutada ka loomaailma kujutamisel erinevat perspektiivi, mõjutades soovitud viisil lugeja kogemust.

3. Kirjanduslik narratiiv kui sotsiaalne praktika. Kuidas mõista fiktiivset aega?

Oma uuringutes võtame aluseks sotsiokognitiivsusega seotud käsitlused (van Dijk 1979, 2006, Geisler 1991), vaadeldes narratiivi konstrueerimise ja interpreteerimisega seotud aspekte sotsiaalse interaktsiooni kontekstis. Vastavalt sellele läheneemisviisile paigutuvad kõik diskursuse vormid, nii suulised kui ka meid siin huvitav kirjalik diskursus, teatud sotsiokultuurilisse ja -ajaloolisse konteksti. Kirjalikud tekstid ja kirjutamispraktikad konstrueeritakse sotsiaalselt. Korduva interaktsiooni läbi kujundab kogukond tekstide suhtes teatud positsiooni või annab neile teatud väärtuse ning nii tekib kogukonna kirjutamiskultuur, s.t kujunevad tekstilised praktikad ehk viisid, mille abil luua ja edastada tähendust. See on kahesuunaline protsess – tekstid ja kirjutamispraktikad tekivad sotsiokultuurilise konteksti baasil, kuid samal ajal osutavad sellele kontekstile ning aitavad seda luua. (Kostouli 2005)

Spetsiifilise sotsiokultuurilise praktikana eeldab ka kirjanduslik narratiiv seega lähtumist teatud konventsioonidest tähenduse loomisel ja selle interpreteerimisel. Juba Mihhail Bahtin (1987) rõhutab vajadust uurida kirjandusžanre mitte ainult kirjanduse piirides kunsti spetsiifika seisukohalt, vaid ka kui teatud diskursuse tüüpe, mida ta oma käsitluses nimetab kõnežanrideks. Eriti oluliseks peab ta primaarsete ja sekundaarsete kõnežanride erinevusi. Primaarsed kõnežanrid (nt argidialoogi repliik, kiri) on kujunenud otse kõnesuhtluses; sekundaarsed žanrid (nt teaduslik uurimus, publitsistika, kirjandusžanrid) tekivad keerukamas kultuurisuhtluses, transformeerides kujunemisprotsessis primaarseid žanre. Teisisõnu on üks peamisi küsimusi, mille Bahtin tõstatab ja mis on jätkuvalt aktuaalne, kirjandusliku diskursuse suhestumine standardse suulise kõnesituatsiooniga. Üldiselt nähakse igasuguse kirjaliku diskursuse põhiomadusena kirjutamise ja lugemise kontekstide lahutatust (Olson 2001). Keelelisi vorme kasutatakse nii, et teksti võiksid lugeda tundmatud lugejad kontekstides, mida kirjutaja täielikult ei kontrolli. Tekst omandab nii spetsiifilise vormi, mis nõuab spetsiifilist interpreteerimisviisi. Ilukirjanduse puhul on see eemaldumine kõneleja-kuulaja vahetu suhtlussituatsiooni mudelist ilmselt kõige teravam ja olemuslikum. See kajastub selgelt näiteks epistemoloogilises plaanis ja ka keelelistes erijoontes, mis eristavad sellist diskursust suulisest suhtlusest. Erinevalt reaalsusest, kus meil ei ole otsest ligipääsu teise inimese mõtetele ja tunnete, on ilukirjanduses võimalik vahetult kujutada kolmanda isiku mentaalseid seisundeid, mõtteid, kavatsusi, tegevusmotive jne.

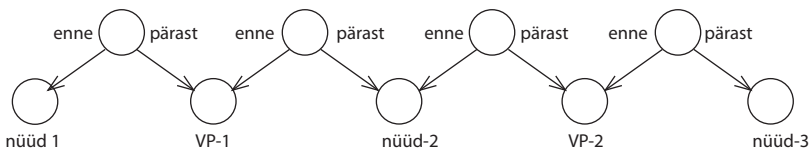
Žanri kujunemise seisukohalt on huvitav jälgida, kuidas keskaegsed autorid, kes lähtuvad veel analoogiast tavapärase suulise suhtluse kogemusega, vältisid tegelase subjektiivsuse kujutamist või lasid tegelasel oma mõtteid avaldada monoloogis (valjusti mõtlemine), mis võimaldab jutustajal justkui õigustada oma teadmist teise inimese sisemaailmast (Galbraith 1995). Kirjaliku narratiivi arenedes sellised epistemoloogilised kõhklused kaovad ning lugejad ja kirjutajad hakkavad võõra subjektiivsuse kujutamist aktsepteerima sama tõepäraselt kui maastiku kirjeldust. Keelelisel tasandil kajastub see muutus siirdkõne tekkimises 17.–19. sajandil ilukirjanduslikule narratiivile ainuomase süntaktilis-stilistilise vormina just tegelase teadvuse vahetuks kujutamiseks (Banfield 1993). Siit jõuame taas loomaailma mõiste juurde (vt jaotis 1), mis just kirjandusliku narratiivi maailmaloovat või-

met rõhutades (Herman 2002) väljendab ja samas seletab kõige paremini sellise diskursuse autonoomiat reaalse suhtlussituatsiooni suhtes. Lugu loob uue, meie reaalsest aegruumist eraldi seisva maailma, kuhu me siseneme, võttes seal toimuvat reaalsusena, mis lihtsalt eksisteerib.

Inimese kognitsioon on välja kasvanud tema bioloogilisest olemusest, kogemuse projitseerimise, osutamise ja viitamise ning interpreteerimise ehk teiseks tõlkimise võimest. Just kognitiivne narratoloogia saab arvestada neid võimeid nii tekstide looja kui lugejaga seoses. Tänu oma maailmalooval võimele, mis võimaldab luua kõneleja minast erineva teadvussubjekti, konstrueeritakse loomaailma sisenedes ka vastav uus ajaline ja ruumiline väli. David A. Zubin ja Lynne E. Hewitt kirjeldavad seda kui deiktulist siiret, mis tähendab, et mina-siin-praegu-koordinaadid ehk deiktiline kese kantakse üle teatud punkti diskursuse maailma sees ning interpreteeritakse teksti sellest perspektiivist (Zubin, Hewitt 1995). Ajalisel teljel tähendab see, et kuigi narratiiv on kirjutatud mineviku ajavorme kasutades, ei leia need sündmused aset reaalses minevikulises ajas. Lugeja mõistab minevikuvorme kui üht signaali, et tuleb teostada deiktiline siire ning luua kirjeldatud sündmuste jaoks fiktiivne olevikuhetk loo sees ehk narratiivi **nüüd-punkt**, mis ei ole deiktiliselt seotud lugeja enda reaalse ajaga. See fiktsiooni tasandil konstrueeritud nüüd-punkt on deiktiline kese, mida lugeja võtab kui enda oma ning seeläbi tekib illusioon loo maailma vahetust kogemisest. Et narratiivile on omane ajaline edasiliikumine, liigub ka nüüd-punkt loo arenedes edasi. Vastavate keeleliste markeritega saab selle liikumise ka peatada või katkestada stabiilse ajalise edasiliikumise, nihutades nüüd-punkti hoopis uude ajalisel raamistikku. Lugeja loob loomaailma ajastruktuurist mentaalseid representatsioone, jälgides nüüd-punkti dünaamikat, ning – asetades end sellega määratletud deiktilisse keskmesse – liigub looga kaasa ning kogeb esitatavaid sündmusi sellest punktist lähtuvalt.

4. Narratiivne imperfekt ja imperfektiivne narratiiv

Aspektikategooriate funktsioneerimine narratiivis põhineb niisiis sellel, kuidas need suhestuvad nüüd-punktiga ja kuidas seda mõjutavad. Michael J. Almeida (1995) mudeli järgi funktsioneerib aspekt narratiivis järgmiselt: perfektiivselt kujutatud sündmuse ajaline viitepunkt asetub pärast kehtivat nüüd-punkti ja uus nüüd-punkt asetatakse vahetult pärast viitepunkti; seega perfektiivne aspekt liigutab nüüd-punkti edasi, kandes loo ajalist dünaamikat (vt joonis 1). Imperfektiivsed sündmused hõlmavad ajaliselt antud hetkel kehtivat nüüd-punkti; kujutatav sündmus on nüüd-punktiga samaaegne ning ajaline edasiliikumine tühistatakse.



Joonis 1. Loo ajaline dünaamika (VP = viitepunkt) (Almeida 1995: 172)

Kirjeldatud põhimõttel toimib aspekt nii keeltes, kus see on grammatiliselt väljendatud, nagu hispaania keel, kui ka eesti keeles, kus perfektiivsus või imperfektiivsus ei avaldu reeglipäraselt, vaid põhineb eelkõige verbi semantilistel aspektuaalsetel omadustel.

19. saj keskel tekib aga esmalt prantsuse- ja hiljem ka hispaaniakeelsetes kirjandustekstides uus imperfekti kasutusviis, mis ei vasta kuidagi kõnealuse vormi kanoonilisele ajalise-aspektilisele tähendusele, vaid imperfekti kasutatakse justkui perfektiivse vormi koha peal. Sellist keelekasutust on hakatud nimetama narratiivseks imperfektiks, kuigi puudub üksmeel selles osas, millised juhtumeid täpselt see mõiste hõlmab. Käesolevas uurimuses lähtume Jacques Bres' (2005) definitsioonist, mille kohaselt narratiivse imperfekti all mõeldakse imperfektiivse vormi kasutamist ajalisel edasiliikuvate sündmuste kujutamisel. Bres on seisukohal, et narratiivses kasutuses ei muutu imperfekti aspektuaalne tähendus, vaid tegemist on perfektiivset aspekti nõudva konteksti ja imperfektiivse aspekti interaktsiooniga. Sellest vastuolust tekibki eriline stilistiline efekt.

Kuna sellist võtet kasutati algselt tüüpiliselt loo, peatüki või episoodi lõpus, on Harald Weinrich (1964) seletanud seda nähtust kui narratiivitehnilist võtet, mis aitab lugejat üleminekul loomaailma tasandilt tagasi igapäevasesse reaalsusse. Sarnaselt imperfekti tüüpilisele kasutamisele loo alguses, mis aitab teostada deiktulist siiret loomaailma sisse, juhatab loo lõppedes imperfekt ka sujuvalt loomaailmast välja. Weinrich toob välja huvitava paralleeli Euroopa vanemas kirjanduses esineva raamjutustusega, mida võib vaadelda kui teistsugust jutustamistehnikat sama funktsiooni täitmiseks. Näiteks *exemplum*-i žanris kasvavad lood välja eelnevast vestlusest või arutelust, illustreerides seda, ning neile järgneb jällegi kuulajate kommentaar. Selline tehnika kajastab teadlikkust sellest, et loomaailm seisab reaalsest aegruumist mingil viisil eraldi, ja näitab samas, et üleminekut ühest teise peetakse probleemseks (vrd jaotises 3 käsitletud epistemoloogilisi küsimusi). Narratiivse imperfekti tekkes võibki Weinrichi arvates oletada kajastust sellest varasemast arusaamast, et fiktsiooni ja reaalsuse vaheliseks siirdeks on vaja mingil kujul vaheetappi.

Hiljem ei piirdu narratiivne imperfekt enam tekstiosa lõpupositsiooniga ning omandab väga mitmesuguseid kasutusi, mille narratiivitehnilist funktsiooni on raske üheselt kirjeldada. Paljude narratoloogide tähelepanu on pälvinud näiteks Anton Tšehhovi jutustamislaad, mida on nimetatud imperfektiivseks narratiiviks, kuna tavapärase perfektiivse aspekti asemel on loo kandvaid sündmusi kujutatud imperfektis, n-ö mitte-sündmustena, mis ei tõuse lugeja interpretatsioonis esiplaanile (Thelin 1990, Björklund 1993).

5. Narratiivne imperfekt hispaania keeles ja tõlkevõimalused eesti keeles

Uurisime narratiivse imperfekti jutustamistehnilist rolli loomaailma konstrueerimisel Benito Pérez Galdósi romaanis "Osatu", Isabel Allende romaanis "Vaimude maja", Gabriel García Márqueze romaanis "Armastus koolera ajal", Juan Rulfo romaanis "Pedro Páramo" ja Pío Baroja novellides, ning probleeme ja võimalusi selle funktsioonide edasiandmisel eestikeelses tõlkes. Teoste valikul on piiratud

kolmandas isikus esitatud narratiividega, sest minajutustuse teistsugune deiktiline ülesehitus eeldaks eraldi käsitlust. Nimetatud tekstides esines 143 narratiivse imperfekti kasutamise juhtu ning nende analüüsimisel lähtusime ajalise edasiliikumise interpreteerimise viisist (ajamäärus, põhjuslikud suhted) ja verbi semantilisest aspektist. Semantiline aspekt seab teatud piirangud aspektuaalse perspektiivi valikuks sündmuse kujutamisel. Antud juhul puudutab see eelkõige momentaanseid sündmusi, mis sisemise ajalise struktuuri puudumise tõttu eeldavad kujutamist totaalselt, s.t perfektiivselt.

Järgnevalt esitame näidete varal analüüsi tulemused, tuues välja peamised viisid, kuidas narratiivne imperfekt mõjutab loomaailmast loodavaid representatioone, ning kuidas on see edasiantav eesti keele vahenditega. Esinemisjuhud oleme jaotanud kolme rühma vastavalt sellele, kuidas paigutub narratiivse imperfektiga väljendatud sündmus narratiivi ajateljele.

Esitatud näidetes on narratiivsed imperfektid ja nendele vastavad verbid eestikeelses tõlkes esile tõstetud. Hispaaniakeelsetes näidetes on märges verbi grammatilise aspekti kohta (IMPERF = imperfektiivne aspekt).

5.1. Määrusega tähistatud ajaline nihe

Vaadakem kõigepealt näiteid, kus narratiivne imperfekt tekib imperfektiivse aspekti kombinatsioonis lausealgulise ajamäärusega, mis tähistab distantsi eelmisest ajalisest punktist (nt *poco después / veidi aega hiljem*) või lokalisatsiooni ajateljel (nt *a las ocho / kell kaheksa*). Lausealguline ajamäärus on vahend nüüd-punkti nihutamiseks uude kohta (Zubin, Hewitt 1995). Nihe katkestab loo aja stabiilse edasiliikumise, tehes suurema või väiksema hüppe uude ajalise raamistikku. Tegemist on ellipsiga, mis tekitab lahknevuse lugeja aja (teksti vastuvõtmise ajaline järgnevus) ja loo aja (loomaailma kujuteldavate sündmuste järgnevus) vahel, sest teatud loomaailma sündmusi ei ole tekstis eksplitsiitselt kujutatud, s.t teatud ajaline lõik loo tasandil jäetakse teksti tasandil vahele. Lugeja interpretatsioonis tekitab see mulje loo aja kiirest edasiliikumisest, mida veelgi süvendab sündmuste kujutamine imperfektiivsest perspektiivist.

- (1) Le embarazaron, oprimiéndole en fuerte anilla horizontal de hierro sujeta a la pared, y allí, sin defensa posible, desnudo, recibió la acometida. *Poco después **yacia*** (IMPERF) aletargado en una cama con visibles apariencias de bienestar. Al fin, durmió profundamente. (BPGH: 22)

Teda hoiti kinni, alasti suruti ta seina külge kinnitatud tugevasse horisontaalsesse raudrõngasse ja seal, suutmata end kaitsta, alistus ta rünnakule. *Veidi aega hiljem **lebas*** ta uimasena, ent ilmse mõnutundega voodis. Lõpuks uinus ta sügavasti. (BPGE: 18)

Näites (1) võib näha, kuidas tänu imperfektiivsele kujutamiseviisile (*yacia*) uues narratiivi nüüd-punktis ei koge lugeja hispaania keeles sündmusi sujuvalt kulgevatena, vaid tekib hüppeline olukorra muutus. Imperfekt kujutab sündmust partsiaalselt, s.t kirjeldatavasse sündmuste järgnevasse asetub selle sisemine faas algus- ja lõpp-punktile viitamata. Et uue olukorra algust ei ole tekstis kodeeritud, kogeb lugeja seda narratiivi uues nüüd-punktis juba toimuvana, mis jätab mulje sündmuste kiirest

vaheldumisest loo ajas ja tõstab esile kontrasti eelmise situatsiooniga. Antud näide on tüüpiline, kuna imperfekti on sellise efekti loomiseks kasutatud dramaatilise stseeni ja samuti kogu peatüki lõpus. Ka eesti keeles on siin kasutatud seisundiverbi (*lebas*), mida tüüpiliselt mõistetakse imperfektiivsena, mistõttu vastav sündmuste ajalis-aspektiline kujutusviis on hästi edasiantav.

- (2) Se fueron de viaje a Italia y a los dos días de embarcarse, Esteban **se sentía** (IMPERF) enamorado como un adolescente, [---] (IAh: 89)

Nad läksid pulmारेisile Itaaliasse ja *kaks päeva pärast laevale asumist tundis* Esteban, et oli armunud nagu nooruk, [---] (IAe: 111)

Erinevalt eelmisest näitest jääb näites (2) eesti keeles sündmuse aspektuaalsus ebamääraseks, sest *tundma* võib tähistada nii momentaanset sündmust kui ka staatilist seisundit, millest tulenevalt tajutakse sündmust vastavalt kas pigem perfektiivselt või imperfektiivselt. Seetõttu kogeb lugeja eestikeelse teksti põhjal stseeni neutraalsemalt. Et lugeja looks sellest täpsemini originaalile vastava mentaalse representatsiooni, võiks eesti keeles lisada *juba*, mis rõhutaks muutuse kiirust, vrd (4). Loo ülesehituse seisukohast ei ole siiski ka hispaania keeles siin kontrast nii terav kui näites (1), sest narratiivne imperfekt ei rõhuta siin piiri sündmuste arenguliinide, stseenide või tekstiosade vahel.

- (3) Se había ido a errar por las escolleras, y estuvo recitando versos de amor contra el viento, llorando de júbilo, hasta que acabó de amanecer. A las ocho **estaba sentado** (IMPERF) bajo los arcos del Café de la Parroquia, alucinado por la vigilia, tratando de concebir un modo de hacerle llegar su bienvenida a Fermina Daza, cuando se sintió sacudido por un estremecimiento sísmico que le desgarró las entrañas. (GGMh: 148)

Ta tuias muulil edasi-tagasi, luges lõõtsuvasse tuulde lembevärssse ja nuttis rõõmust, kuni läks valgeks ja hommik kätte jõudis. *Kell kaheksa istus* ta kohviku võlvlaega terrassil, magamatusest uimane, ja püüdis välja mõelda, kuidas peaks Fermina Dazale kätte toimetama oma tervitused – kui äkki oleks nagu maavärin terve tema sisikonna segamini raputanud. (GGMe: 99)

Ka näites (3) võib narratiivse imperfekti jutustamistehniliseks funktsiooniks pidada kontrasti rõhutamist, kuigi siin mitte muutuse kujul võrreldes eelneva olukorraga, vaid loo arengus äärmiselt olulise *järgneva* sündmuse suhtes. G. García Márqueze jutustamislaad romaanis kujutab loomaailma üldiselt neutraalsest, märkamatuist vaatepunktist, tungimata tegelaste sisemisse subjektiivsesse ruumi ning kasutamata ka loomaailmast selgelt väljaspool asuva vaatepunkti võtet. Lugeja aeg kulgeb üldiselt paralleelselt loo ajaga. Narratiivne imperfekt on seetõttu kõnealuses tekstis üsna erandlik. Käsitletaval juhul see justkui aeglustab tegelase ruumiliste liikumiste ja tegevuste järgnevuse, sest ajalisel liigub lugu edasi (*a las ocho / kell kaheksa*), kuid tegevus on kujutatud staatiliselt (*estaba sentado*, samuti eesti keeles kasutatud seisundiverb *istus*). See eelmiste perfektiivsete sündmuste jadas esilekerkiv imperfektiiviga vormistatud stseen ja selle staatika valmistavad lugeja ette järgmiseks ootamatuks sündmuseks, mis moodustab kogu eelmise tegevusliini kulminatsiooni. See tõuseb lugeja loomaailma tajumisel jõuliselt esile, pannes teda

sel viisil vahetumalt kogema tegelase vapustust. Huvitav on see näide seetõttu, et eesti tekstis on tõlkija kirjeldatud efekti tugevdamiseks kasutanud mõttekriipsu, mis tõstab järgnevat esile.

- (4) El Majito se dejó ir con grave paso por la calle de Moratines abajo. Era el día ventoso, frío y seco, hijo maldito de la malditísima primavera de Madrid. La pluma del ros del Majito – porque una pluma de pavo tenía – se torcía con la fuerza del viento. [---]

*En la calle de Ercilla **tenía** (IMPERF) ya un séquito de seis muchachos; en la del Labrador ya se le había incorporado una partida de diecisiete, entre hembras y varones, siendo las primeras, icosa extraña!, las que más bulla metían.* (BPGH: 92–93)

Keku läks tähtsal sammul mööda Moratinese tänavat. Päev oli tuuline, külm ja kuiv, jõleda Madriidi kevade jõle sünnitis. Sulg Keku mütsil – seal oli nimelt paabulinnusulg – kõverduis tuule käes. [---]

Ercilla tänavas oli tal juba kuuest poisist koosnev saatjaskond; Labradori tänavas olid temaga ühinenud juba seitseteist tüdrukut ja poissi, kusjuures esimesed – imelik asi! – kõige rohkem lärmi tegid. (BPGe: 75–76)

Et deiktilise välja erinevad parameetrid (aeg, ruum, psühholoogiline vaatepunkt) on omavahel tihedalt seotud, mõjutavad ühe komponendiga eksplitsiitselt teostatavad operatsioonid sageli implitsiitselt ka teisi. Eriti otsene seos on ruumilise ja ajalise asukoha vahel liikumise puhul, kuna liikumine toimub ajas. Näites (4) on fookuses loomaailma ruumis liikuv tegelane ja lugeja representatsioon antud stseenist lähtub temaga kaasaliikuvast deiktilisest punktist. Kui lausealguline määrus nihutab deiktilise keskpunkti ruumis uude kohta (*Ercilla tänavas*), toimub nihe ka ajalisel teljel. Seega võib nüüd-punkti edasi nihutada ka kohamäärus ning koos imperfektiivse aspektiga (nii hispaania keeles kui ka eesti keeles tänu seisundiverbi semantilisele aspektile) saavutatakse samasugune kujutusviis, nagu nägime eelmistes näidetes. Näites (4) ei ole narratiivne imperfekt stseeni või tekstiosa üleminekut või pöördepunkti tähistavas positsioonis, mis seda veelgi esile tõstaks, kuid olukorra kiire muutuse rõhutamiseks on kasutatud ka määrust *ya / juba*.

- (5) *Y poco después se **volvía a cerrar** (IMPERF) la triste alcoba, y retirándose personas y luces, todo **quedaba** (IMPERF) en silencio y soledad tristísima.* (BPGH: 154)

Ja varsti pärast seda **suleti** taas kurb magamistuba, ning et inimesed ja valgus lahkusid, **vajus** kõik vaikusse ning ülikurba üksindusse. (BPGe: 125)

Erinevalt eelmistest näidetest eeldab näites (5) sündmuste aspektiline iseloom (momentaansus) nende kujutamist perfektiivselt. Nii mõistetakse nende kujutusviisi ka eestikeelses tõlkes (*suleti* puhul märgib täissihitis selgelt perfektiivsust), kuid hispaania keeles on neid kirjeldatud imperfektiivse vormiga. Hispaania keeles tekib nii kontrast imperfekti ajalis-aspektuaalse tähenduse ja konteksti vahel, kus ajamäärus viib lugeja ajalisel edasi uude nüüd-punkti ning sündmused on mõistetavad totaalselt, lõpule jõudnutena, ning põhjuslikult üksteisele järgnevatena. Kuid

imperfekti vorm peatab ajalise edasiliikumise määrusega tähistatud nüüd-punktis; järgmiste verbidega kirjeldatu ei asetu ajateljel edasi selle punkti järele (nagu seda teeks eeldatav perfektivne aspekt), vaid need jäävad seotuks sama ajalise punktiga. Loo dünaamiline kulg pidurdub, fookus ei ole enam sündmuste järgnevusel, vaid justkui distantseerutakse narratiivi ajateljest ning stseen omandab staatilise iseloomu. Nii võtab narratiivne imperfekt kokku ja lõpetab eelneva dramaatilise süžeeleini ning samuti kogu peatüki. Selline funktsioon on teksti lõpupositsioonis esinevale narratiivsele imperfektile tüüpiline (vt jaotis 4).

Kuna narratiivne imperfekt ilmub kontrastina perfektivset aspekti eeldava kontekstiga, võib selle eripärane tähendus avalduda suuremal või vähemal määral, sõltuvalt kontekstiliste eelduste tugevusest (Bres 2005). Verbi momentaanne semantiline aspekt loob tugeva eelduse sündmuse kujutamiseks perfektivselt, seetõttu on näite (5) puhul tekkinud kontrast teravam kui eelmistes näidetes ning ka selle mõju ajalisele ja vaatepunktilisele interpretatsioonile tugevam.

5.2. Põhjuslike suhete alusel järeldatav ajaline edasiliikumine

Sageli ei ole edasiliikumine narratiivi ajateljel küll kodeeritud keeleliste markerite abil, kuid selline interpretatsioon on siiski järeldatav teadmiste põhjal, mis lugejal on maailmast, s.t sündmustest ja nende omavahelistest seostest, eriti põhjussuhetest. Nendel juhtumitel ei ole tegemist hüppelise ajalise nihkega, vaid sujuva ajalise edasiliikumisega. Kuigi imperfekti vorm tüüpiliselt paigutab sündmuse kehtivasse nüüd-punkti, tühistades edasiliikumise, järeldub käsitletavates näidetes põhjusseostest sündmuste ajaline järgnevus.

- (6) Isidora puso atención, y, en efecto, del fondo invisible **venía** (IMPERF) un rumor hondo y persistente, [---] (BPGH: 48)

Isidora pingutas tähelepanu, ja tõepoolest, ruumi nähtamatust kaugusest **tuli** pidev, sügav heli [---] (BPGe: 39)

Näites (6) paigutub kirjeldatud heli lugeja mentaalses representatsioonis pärast sündmust *Isidora pingutas tähelepanu*. Loo ülesehituses ei tõuse kõnealune imperfektikasutus kuigivõrd esile, kuid näitab ilmekalt narratiivi ajastruktuuri seost vaatepunktiga. Kuna sündmuse kujutatakse tegelase vaatepunktist, hakkab ka lugeja kirjeldatud heli justkui kuulma alles siis, kui tegelane on selle olemasolu märganud. Samasuguse interpretatsiooni annab ka eestikeelne tõlge.

- (7) El sol se fue volteando sobre las cosas y les devolvió su forma. La tierra en ruinas estaba frente a él, vacía. El calor caldeaba su cuerpo. Sus ojos apenas se movían; saltaban de un recuerdo a otro, desdibujando el presente. De pronto su corazón **se detenía** (IMPERF) y parecía como si también se detuviera el tiempo. Y el aire de la vida. (JRh: 280)

Päike tiirles asjade kohal ja moondas nende kuju. Varemets, tühi maa oli ta ees. Kuumus soojendas ta keha. Ta silmad liikusid vaevu; nad hüppasid ühelt mälestuselt teisele, pühkides oleviku minema. Äkki **jäi** ta süda **seisma** ja paistis, nagu oleks ka aeg ja eluhingus seisma jäänud. (JRe: 90)

Erinevalt näitest (6) on näites (7) nii hispaania keeles kui ka eestikeelses tõlkes tegemist momentaanse verbiga (*se detenia / jäi seisma*), mis eeldaks tüüpiliselt kujutamist perfektiivsena. Siiski on hispaania keeles kasutatud imperfekti vormi. Selle vastuolu tõttu tõuseb sündmus, milleks on peategelase Pedro Páramo surm, loo arengus esile. Kuna imperfekt ei loo narratiivi kujutamisel uut nüüd-punkti, peatab see ajalises plaanis romaani lõpus loo kulu. Ka siin võib samas näha imperfekti ajalise tähenduse seost vaatepunktiga, paigutades selle tegelase subjektiivsesse ruumi. Üldiselt on romaani vältel Pedro Páramot kujutatud pigem objektiivselt vaatepunktist, kuid lõpustseenis samastutakse tema subjektiivse teadvusega, mis omakorda tõstab kõnealuse stseeni loomaailma kujutamisel esile. Eestikeelses tõlkes selline aspektiline, ajaline ja vaatepunktiline kontrast ei avaldu.

Loomaailma representatsiooni eriti tugevalt mõjutavad ajalised ja sellega seotud ruumilised ning eelkõige vaatepunktilised modifikatsioonid tekivad narratiivsete imperfektide pikemate jadade korral. Vaatepunkti konstrueerimisel mängib lisaks teatud markerite olemasolule lauses rolli ka selle paiknemine kontekstis, s.t. eelmiste lausete subjektiivsus või objektiivsus (Wiebe 1994). Imperfektiivne aspekt on kontekstitundlik vaatepunkti marker, mistõttu pikemas imperfektiga kujutatud tekstilõigus avaldub selle vaatepunkti loov potentsiaal tugevamalt kui üksiku verbi puhul.

(8) El viejo se asomó a la boca de la caverna.

– ¡Zagal, zagal! – gritó, con desesperación.

Nada, **no se oía** (IMPERF) nada.

– ¡Zagal! ¡Zagal!

Parecía oírse (IMPERF) mezclado con el murmullo del viento un balido doloroso que subía desde el fondo de la caverna.

Loco, trastornado, durante algunos instantes, el pastor **vacilaba** (IMPERF) en tomar una resolución; luego se le ocurrió pedir socorro a los demás cabreros, y echó a correr hacia el castillo. (PBh: 74)

Vanamees kummardus üle kuristikuserva.

“Poiss, poisu!” karjus ta meeleheitlikult.

Mitte midagi **polnud kuulda**, mitte ühtegi häält.

“Poiss! Poiss!”

Alt kuristiku sügavusest, segamini tuule sosinaga, **näis kostvat** üksainus valus oie.

Juhmistunud, peaaegu mõistuse kaotanud, **oli** karjus mõne silmapilgu vältel **võimetu** midagi otsustama. Siis sähvas ta peast läbi, et tuleb teisi kitsekarjuseid appi kutsuda, ja ta pistis lossi poole jooksu. (PBe: 40)

Näites (8) kujutatakse novelli pöördepunktiks olevat dramaatilist sündmust, kus vana karjuse poeg kukub kuristikku. Kuna üks sündmus toimub reaktsioonina teisele, on järeldatav, et need paigutuvad ajas üksteise järele, kuid hispaania keeles on neid kujutatud imperfekti vormis. Selle tulemusel loo ajaline kulg justkui peatub, stseen omandab staatilise iseloomu, mis teeb lugejale tajutavaks valitseva pinge. Lugeja interpreteerib stseeni tegelase subjektiivse deiktilise keskmega samastudes, kogedes tema meeleheidet ja peataolekut. Eesti keeles ei tule subjektiivne vaatepunkt nii selgesti esile.

Järgnevalt esitame kaks näidet B. Pérez Galdósi romaanist “Osatu”. Selles romaanis loob autor fiktiivse jutustaja kuju, pöördudes aeg-ajalt otseselt lugeja

poole, esitades oma teadmisi loost piiratuna (*ma ei tea kindlasti, kas see oli samal päeval*) jne. Sel viisil konstrueeritakse loost väljaspool asuva jutustaja deiktiline kese. Selle tulemusel jutustatakse sündmustest suurel määral justkui nähtuna loomaailmast väljaspool asuvast vaatepunktist. Lugeja distantseeritakse seega loomaailmast. Sama funktsiooni täidab romaanis sageli ka narratiivne imperfekt.

- (9) ¡Con qué inocente confianza y abandono iban los dos, en familiar pareja, por los senderos torcidos que conducen desde el camino de Aragón a Pajarillos! **Bajaban** (IMPERF) a las hondonadas de tierra sembrada de mies raquílica; **subían** (IMPERF) a los vertederos, donde lentamente, con la tierra que vacían los carros del Municipio, se van bosquejando las calles futuras; **pasaban** (IMPERF) junto a las cabañas de traperos hechas de tablas, puertas rotas o esteras, y blindadas con planchas que fueron de latas de petróleo; **luego se paraban** (IMPERF) a ver muchachos y gallinas escarbando en la paja; **daban vueltas** (IMPERF) a los tejares; **se detenían** (IMPERF), **se sentaban** (IMPERF), **volvían a andar** (IMPERF) un poco, sin prisa, sin fatiga. (BPGH: 71)

Millise süütu usalduslikkuse ja sundimatusena läksid nad nagu kaks lähedast sõpra mööda käänulisi radu, mis viivad Aragóni teelt Pajarillosele! Nad **laskusid** nõgudesse, kus kasvas kidurat teravilja, **tõusid** prügimägedele, kus muld, mida siia vedasid linnavalitsuse vankrid, hakkas pikkamisi kujundama tulevase tänavaid; **möödusid** kaltsukorjajate onnidest, kaetud vanade petrooleumivaatide plekiga; nad **peatusid**, et vaadata poisikesi ja õlgedes siblivaid kanu; **möödusid** tellisetehastest; **jäid** siin-seal **peatuma**, **võtsid istet**, **läksid jälle edasi**, ruttamata, väsimata. (BPGe: 58)

Näites (9) kujutatud tegevused paiknevad ajas üksteise järel. Kuna hispaania keeles on neid kujutatud imperfektiivses aspektis, on nad tajutavad kui samaaegsed ühe ja sama nüüd-punktiga, s.t nad kõik iseloomustavad teatud üht ajalist punkti loos. Erinevad tegevused on koondatud ühtsesse ajalisse raami, mis on fookuses, ning sündmuste omavahelised suhted selle raamistuse sees jäävad tagaplaanile. Selliselt modifitseeritud ajalised suhted avaldavad mõju ka ruumilisele mõõtmele ja vaatepunktile. Lugeja ei liigu tegelaste ja nende liikumisega kaasa, vaid näeb kogu episoodi justkui tervikuna kuskilt eemalt. Seda võiks piltlikult võrrelda kõrgusest linnale avaneva vaatega, kus tänavad on kõik samaaegselt näha nagu staatilisel pildil. Kui aga asume linnas sees, saame erinevaid ruumilisi punkte näha ainult ajaliselt üksteise järel, ühest teise liikudes. Teisisõnu deiktiline kese, kust lugeja sündmusi näeb, on konstrueeritud sündmusteliinist väljaspool, distantseerudes sellest. Loo struktuuris täidab kõnealune ajaliselt ja vaatepunkti poolest modifitseeritud lõik lisaks tüüpilist rolli märkida üleminekut uuele teemaarendusele. Ka eestikeelses tõlkes jääb kirjeldatud tegevuste täpne omavaheline järjestus tegelikult määramatuks, neid on kujutatud läbisegi (tõlkija on ilmselt seepärast välja jätnud originaalis esineva *luego* 'siis, seejärel', mis märgiks täpset ajalist suhet). Hispaaniakeelses tekstis tekkiv ajaline ja vaatepunktiline efekt ei ole siiski edasiantav.

- (10) Don José, a quien las horas se le hacían siglos, no pensaba en apuntar en el Diario ni en el Mayor los gastos extraordinarios de aquel día. Por

la tarde **ocupábase** (IMPERF) de instalar la mesa en la sala, por ser el comedor muy pequeño para tan gran festín. Después **se miraba** (IMPERF) diecinueve veces al espejo, **se acicalaba** (IMPERF), y en el colmo ya del regocijo, les **quitaba** (IMPERF) a los chicos del tercero el tambor con que atronaban la casa toda, y **tocaba** (IMPERF) por los pasillos con furor y denuedo, seguido de la turba infantil y por ésta con alegres chillidos aclamado. (BPGh: 192)

Don José, kellele tunnid tundusid sajanditena, ei mõtelnudki kanda majapidamistarvete raamatusse ega pearaamatusse selle päeva erakordseid kulusid. Õhtupoolikul **oli ta ametis** söögilaua ülesseadmisega saali, sest söögituba jäi nii suure pidusöömingu jaoks väikeseks. Siis **vaatas** ta oma parkümmend korda peeglist, **klanis** end, ja kui rõõmutuju haripunkti tõusis, **võttis** neljanda korruse poistelt trummi, millega need kõrvulukstavaltpõristasid, nii et kogu maja kajas, ja **mängis** kõigis koridorides tohutu vaimustuse ning hooga, saadetuna lastekarjast, kes talle rõõmsa kilkamisega oma heakskiitu avaldas. (BPGe: 157)

Näites (10) on samuti tegemist sündmuste järgnevusega, mida on hispaania keeles kujutatud imperfekti vormis. Eelmise näitega samal viisil tekib ka siin distantseerunud vaatepunkt. Kirjeldatud episoodi ei näe lugeja selle sujuvas kulgemises, narratiivi ajateljel tegelasega kaasa liikudes ja tema teadvuse ja tunnetega samastudes, vaid tegelast jälgitakse justkui eemalt, ilma empaatiata, mis tõstab esile tema õhina naeruväärsuse. Samal ajal on imperfekti kasutamisel antud juhul ka narratiivi ülesehitav roll, tõstes tekstis stilistiliselt esile stseeni, mida võib pidada eelnevalt kujutatud sündmuste kulminatsiooniks. Eestikeelse teksti lugeja interpreteerib stseeni neutraalse sündmuste järgnevusena.

5.3. Narratiivne imperfekt ja habituaalsus

Analüüsitud tekstides leidis juhtumeid, kus hispaania keeles imperfekti vormis kujutatud stseen eemaldub narratiivi ajateljelt. Selle ajaline lokalisatsioon teiste loo sündmuste suhtes hägustub ning lükkub ajaliselt määramatule tagaplaanile, omandades habituaalse tähenduse. Habituaalsus kuulub imperfektiivse aspekti tüüpiliste tähenduste hulka, kuid käsitletavatel juhtudel ei ole siiski tegemist imperfekti standardse kasutusega, sest üleminek toimub sujuvalt ning õieti on kogu stseen tajutav justkui seguna konkreetsest ühekordsest sündmustejadast ja habituaalsest kirjeldusest.

- (11) Sorprendamos a don Manuel José Ramón Pez (o del Pez) cuando, recién abandonadas las ociosas plumas, **entraba** (IMPERF) en su despacho a enterarse de varios asuntos, ajenos a su empleo, [---]. Envuelto en su abrigadora bata, calados los lentes o quevedos, [---], **revolvía** (IMPERF) cartas, **consultaba** (IMPERF) notas, **hojeaba** (IMPERF) memorándums, **ordenaba** (IMPERF) *in mente* lo que no tenía orden, **hacía** (IMPERF) cálculos, **esbozaba** (IMPERF) proyectos, **trazaba** (IMPERF) planes. La frase y el guarismo se entrecruzaban en su cerebro, [---]; **abismábase** (IMPERF) en meditaciones; *después*, tarareando una cancioncilla, **pasaba** (IMPERF)

la vista por los periódicos de la mañana, **daba** (IMPERF) algunas órdenes a sus escribientes y **se ocupaba** (IMPERF) un poco de teatros y diversiones. (BPGh: 176)

Vaadakem don Manuel José Ramón Pezi (või del Pezi), kui ta *hommikuti* lahkus laiskusele meelitavalt sulgpatjadelt ja **astus** oma kabinetti tutvuma mitmesuguste tema ametile võõraste asjadega, [---]. Soe hommikumantel ümber, näpitsprillid ninal, [---], **sois** ta kirju, **vaatas** ülestähendusi, **lehitses** märkmikku, **korraldas** *in mente* seda, milles polnud korda, **arvutas**, **visandas** projekte, **nuputas** plaane. Sõnad ja arvud ristusid tema ajus, [---]; ta **süvenes** mõtisklustesse; *seejärel heitis* laulu ümisedes pilgu hommikulehtedesse, **tegi** mõned korraldused oma kirjutajaile ja **tegeles** *siis* pisut teatrite ning meelelahutuskohtadega. (BPGe: 144)

Näites (11) kirjeldatakse tegevusi, mis loogiliselt toimuvad üksteise järel. Nende järjekord jääb kas määramatuks või on eksplitsiitselt väljendatud määrusega *después/seejärel*; tõlkija on eesti keeles lisanud ka *siis*. Erinevalt näidetest (9) ja (10), kus on samuti tegemist üksteisele järgnevate sündmustega, võib siin kirjeldatud tegevustele anda habituaalse tõlgenduse, kuna tegemist on uue peatüki algusega ja stseeni ajaline paigutus loomaailmas jääb ebamääraseks. Tõlkija on eesti keeles habituaalsust eksplitsiitselt väljendanud, lisades määruse *hommikuti*. Samas luuakse pöördumisega lugeja poole (*sorprendamos .../vaadakem ...*) nii hispaania kui ka eesti keeles konkreetne vaatepunkt, justkui asuks lugeja koos jutustajaga kusagil lähedal ja jälgiks stseeni sealt. See loob mulje, et vaatame tegelast ühel konkreetsel, loomaailma ajateljel teatud kohta omaval hommikul, mis on vastuolus habituaalsusega. Hispaania keeles eksisteerivad paralleelselt mõlemad tõlgendusvõimalused. Eesti keeles on tõlkija otsustanud habituaalsuse kasuks. Vastavat määrust lisamata ei kerkiks habituaalne interpretatsioonivõimalus esile.

6. Kokkuvõte

Diskursiivne lähenemine kognitsioonile kui inimese meelte, taju ja võimete uurimisele tähendab tekstide loomisel ja tõlkimisel ning lugeja mentaalsete representatsioonide kujunemisel arvestamist teadmistega maailma kohta ning sotsiokognitiivsete aspektidega. Kognitiivse narratiivi vaimus peaksid nii originaaltekst kui tõlge tooma lugejani **kogemuse** tõlgendamist võimaldavast tähenduste paljususest ja loomaailma dünaamilisest vastasmõjust.

Analüüsitud näidete põhjal võime narratiivse imperfekti üldiseks iseloomustamiseks öelda, et selle jutustamistehnilised omadused hispaania keeles põhinevad kontrastil, mis väljendub nii paradigmaatilises lõikes, olles vastuolus antud kohal eeldatava perfektiivse aspektiga, kui ka süntagmaatilisel, tõstes sündmuse või stseeni esile ümbritseva kontekstiga võrreldes. Kontrasti tulemuseks ei ole siiski vaid lugeja tähelepanu teravdamine loo kulminatsioonis või pöördepunktis, vaid narratiivne imperfekt annab sündmusele imperfektiivse perspektiivi ja seega toob kaasa muutused selle suhestumises narratiivi nüüd-punktiga. Tulemuseks võib olla sündmuste kiirendatud esitus või vastupidi – loo kulgemise aeglustumine ja

distantseerumine selle ajateljest. Muutunud ajaline struktuur toob kaasa muutusi ka loo interpretatsiooni teistes parameetrites, eelkõige vaatepunktis, mis mõjutab loo lugejapoolset kogemist enamgi kui sündmuste ajaline paigutus. Narratiivse imperfektiga luuakse sageli neutraalsest märkamast perspektiivist erinev vaatepunkt, mis võib samastuda tegelasega või kuuluda justkui kõrvalseisjast vaatlejale. Eesti keeles on narratiivse imperfektiga loodud sündmuste kujutusviis üldiselt paremini edasiantav juhul, kui kontrast kontekstiliste eeldustega on nõrgem (nt seisundiverbide puhul). Mõnikord saab kontrasti väljendada teiste vahenditega, isegi kirjavahemärkidega. Eesti keeles väljendamatud modifikatsioonid tekivad eriti siis, kui vastuolu kontekstiliste eeldustega on suur (momentaansed verbid), ja pikemate narratiivses imperfektis olevate järgnevuste korral.

Kuigi käesolevas artiklis käsitletud näidete põhjal võiks väita, et grammatilise aspektikategooria olemasolu annab hispaania keelele suuremad väljendusvõimalused kui on eesti keelel, tuleb märkida, et teisest küljest tähendab aspekti väljendamise vältimatus ka teatud piirangut. Selle mõju narratiivi loomisel oleks huvitav uurida eesti-hispaania suunalise tõlkevõrdluse kaudu, kuid otsetõlgete puudumise tõttu ei ole see praegu kahjuks võimalik.

Materjali allikad

- BPGe = Pérez Galdós, Benito 1979. *Osatu*. Tallinn: Eesti Raamat. (Eesti keelde tõlkinud Aita Kurfeldt.)
- BPGh = Pérez Galdós, Benito 2005. *La desheredada*. Madrid: Alianza Editorial.
- GGMe = García Márquez, Gabriel 1995. *Armastus koolera ajal*. Tallinn: Varrak. (Eesti keelde tõlkinud Marin Mõttus.)
- GGMh = García Márquez, Gabriel 2006. *El amor en los tiempos del cólera*. Barcelona: Debolsillo.
- IAe = Allende, Isabel 1992. *Vaimude maja*. Tallinn: Eesti Raamat. (Eesti keelde tõlkinud Ruth Lias.)
- IAh = Allende, Isabel 1992. *La casa de los espíritus*. Barcelona: Plaza & Janés.
- JRe = Rulfo, Juan 1979. *Pedro Páramo*. Tallinn: Perioodika. (Eesti keelde tõlkinud Tatjana Hallap.)
- JRh = Rulfo, Juan 1968. *El llano en llamas*; Pedro Páramo. La Habana: Casa de las Américas.
- PBe = Baroja, Pío 1979. *Kuristik*. – Vagabund Elizabide. Tallinn: Perioodika. (Eesti keelde tõlkinud Asta Põldmäe.)
- PBh = Baroja, Pío 1988. *La sima*. – Mari Belcha y otros cuentos. Madrid: Ediciones de la Torre.

Viidatud kirjandus

- Almeida, Michael J. 1995. Time in narratives. – Judith F. Duchan, Gail A. Bruder, Lynne E. Hewitt (Eds.). *Deixis in Narrative: A Cognitive Science Perspective*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 159–189.
- Bahtin, Mihhail 1987. *Valitud töid*. Tallinn: Eesti Raamat.
- Banfield, Ann 1993. Where epistemology, style, and grammar meet literary history: The development of represented speech and thought. – John A. Lucy (Ed.). *Reflexive Language, Reported Speech and Metapragmatics*. Cambridge: Cambridge University Press, 339–364.

- Björklund, Martina 1993. Narrative Strategies in Čechov's *The Steppe*: Cohesion, Grounding and Point of View. Åbo: Åbo Akademi University Press.
- Bres, Jacques 2005. *L'imparfait dit narratif*. Paris: CNRS-Éditions.
- Bruner, Jerome 1986. *Actual Minds, Possible Worlds*. Cambridge, Mass.: Harvard University Press.
- Bybee, Joan L.; Dahl, Östen 1989. The creation of tense and aspect systems in the languages of the world. – *Studies in Language*, 13 (1), 51–103.
- Chatman, Seymour 1990. *Coming to Terms: The Rhetoric of Narrative in Fiction and Film*. Ithaca, London: Cornell University Press.
- Comrie, Bernard 1993. *Tense*. Cambridge: Cambridge University Press.
- Comrie, Bernard 1995. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.
- Dijk, Teun A. van 1979. Cognitive processing of literary discourse. – *Poetics Today*, 1, 143–160.
- Dijk, Teun A. van 2006. Discourse, context and cognition. – *Discourse Studies*, 8 (1), 159–177.
- Fleischman, Suzanne 1991. Toward a theory of tense-aspect in narrative discourse. – Jadranka Gvozdanovic, Theo A. J. M. Janssen (Eds.). *The Function of Tense in Texts*. Amsterdam etc.: North-Holland, 75–97.
- Galbraith, Mary 1995. Deictic shift theory and the poetics of involvement in narrative. – Judith F. Duchan, Gail A. Bruder, Lynne E. Hewitt (Eds.). *Deixis in Narrative: A Cognitive Science Perspective*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 19–59.
- Geisler, Cheryl 1991. Toward a sociocognitive model of literacy: Constructing mental models in a philosophical conversation. – Charles Bazerman, James Paradis (Eds.). *Textual Dynamics of the Professions: Historical and Contemporary Studies of Writing in Professional Communities*. Madison: University of Wisconsin Press, 171–190.
- Genette, Gérard 1972. *Figures III*. Paris: Éditions du Seuil.
- Gurevič, Aron 1992. *Keskaja inimese maailmapilt*. Tallinn: Kunst.
- Herman, David 2000. Narratology as a cognitive science. – *Image & Narrative*, 1 (1). <http://www.imageandnarrative.be/index.htm> (07.09.2009).
- Herman, David 2002. *Story Logic: Problems and Possibilities of Narrative*. Lincoln: University of Nebraska Press.
- Kostouli, Triantafillia 2005. Making social meanings in contexts. – Triantafillia Kostouli (Ed.). *Writing in Context(s). Textual Practices and Learning Processes in Sociocultural Settings*. New York: Springer, 1–26.
- Metslang, Helle 1994. Temporal relations in the predicate and the grammatical system of Estonian and Finnish. *Oulun yliopiston suomen ja saamen kielen laitoksen tutkimusraportteja*, 39. Oulu: University of Oulu.
- Olson, David R. 2001. What writing is. – *Pragmatics & Cognition*, 9 (2), 239–258. doi:10.1075/pc.9.2.040ls
- Reichenbach, Hans 1947. *Elements of Symbolic Logic*. New York: Macmillan.
- Shedletsky, Lenny 2008. Cognition. – Wolfgang Donsbach (Ed.). *The International Encyclopedia of Communication*. Vol. II. Malden, MA, USA, Oxford, UK, Carlton, Victoria, Australia: Blackwell Publishing Ltd., 541–543.
- Zubin, David A.; Hewitt, Lynne E. 1995. The deictic center: A theory of deixis in narrative. – Judith F. Duchan, Gail A. Bruder, Lynne E. Hewitt (Eds.). *Deixis in Narrative: A Cognitive Science Perspective*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 107–128.
- Talmy, Leonard 1995. Narrative structure in a cognitive framework. – Judith F. Duchan, Gail A. Bruder, Lynne E. Hewitt (Eds.). *Deixis in Narrative: A Cognitive Science Perspective*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 421–460.

- Thelin, Nils B. 1990. Verbal aspect in discourse: On the state of the art. – Nils B. Thelin (Ed.). *Verbal Aspect in Discourse*. Amsterdam: Benjamins, 3–88.
- Thelin, Nils B. 2002. Biopragmatism, space/time cognition, and the sense of language. – Michael Shapiro (Ed.). *The Peirce Seminar Papers*, Vol. 5. New York: Berghahn Books, 1–68.
- Uspenskij 1970 = Успенский, Борис Андреевич 1970. Поэтика композиции: структура художественного текста и типология композиционной формы. Москва: Искусство.
- Weinrich, Harald 1964. *Tempus: besprochene und erzählte Welt*. Stuttgart: Kohlhammer.
- Wiebe, Janyce 1994. Tracking point of view in narrative. – *Computational Linguistics*, 20 (2), 233–287.
- Yli-Vakkuri, Valma 1994. *Tempuksenkäyttö suomalaisessa ja virolaisessa fiktiivisessä tekstissä*. – Karl Pajusalu, Valma Yli-Vakkuri (toim.). *Lähivertailuja 7: Suomalais-virolainen kontrastiivinen seminaari Tammivalkamassa*, 5.–7.1993. Turku, 136–149.

Triin Lõbus (Tartu Ülikool) on õppinud hispaania filoloogiat, tõlkinud hispaania keelest eesti keelde, uurinud aega ja aspekti narratoloogias ja kontrastiivse lingvistika raamistuses.
triin.lobus@mail.ee

Silvi Tenjes (Tartu Ülikool) on uurinud käeliigutusi, kõnekeelt ja kommunikatsiooni, metafoorsust, ruumisemantikat. Uurib suhtlust multimodaalses kommunikatsioonis ning sotsiaalses interaktsioonis. Huvitub tegevusrepresentatsiooni seosest tajuruumiga liigutuste sooritamiseks ning õppimiseks.
silvi.tenjes@ut.ee

TIME OF NARRATIVE AND TIME OF TRANSLATION: A DISCURSIVE APPROACH

Triin Lõbus, Silvi Tenjes

University of Tartu

The current article explores temporal relations in fictional narratives by comparing original Spanish works with their translations in Estonian. The focus lies on the role of the so-called narrative imperfect as a device of style and narrative technique in Spanish. Furthermore, the article endeavours to address the problems and possibilities associated with transmitting the narrative imperfect in Estonian in such a way that the reader is able to experience the storyworld in the most similar way possible with respect to its original construction by the author in Spanish. The study is based on a sociocognitive approach. Accordingly, fictional narrative is seen as a sociocultural practice expected to follow certain conventions in creating and interpreting meaning.

Interpreting temporal relations in fictional narrative requires relocation to the narrative now-point located within the storyworld. Therefore, the functioning of aspectual categories in a narrative concerns its interaction with this now-point. The narrative imperfect involves the rather exceptional use of the imperfective aspect in the case of a temporal progression of narrative now-point. The present study is intended to point out the main ways of how narrative imperfect can affect mental representations of the storyworld. In addition, the possibilities of expressing the same modifications by means of the Estonian language are discussed. The results of the analysis show that the use of narrative imperfect in Spanish is generally characterized by its contrast with the presupposed use of perfective aspect and with the surrounding context, resulting in modifications in the temporal structure as well as in the spatial dimension and particularly the viewpoint. The special representation of events created by the narrative imperfect is more easily transmittable to Estonian in case of a weak contrast. On the other hand, translation problems increase in the presence of a sharp contrast.

Keywords: discourse analysis, text analysis, narratology, tense, aspect, Estonian, Spanish

KESKSETE LAUSEKOMPONENTIDE JÄRJESTUS ÕPPIJAKEELES: ARVUTIANALÜÜSI KATSE

Helena Metslang, Erika Matsak

Ülevaade. Artikkel käsitleb eesti keele lihtlause sõnajärje arvutianalüüsi katset, mille eesmärgiks on õppijakeele sõnajärje vealeidja loomine. Katse käigus koostati eesti keele sagedaste sõnajärjetüüpide mallid, mis kirjeldasid lihtlause ja mõne lihtsama liitlause tüübi verbi, tuumargumentide ning nende järge mõjutavate moodustajate või sõnade järge (põhiliselt subjekt, objekt, predikaat, adverbiaal lause algul või seotud laiendina, üldlaiend). Mallid leiti Tartu Ülikooli kirjakeele korpuse põhjal. Saadud mallide katvust hinnati kirjakeele ja õppijakeele korpuste peal spetsiaalselt loodud programmi abil. Artiklis kirjeldatav programm, mis on kasutatav koos mallide koguga, analüüsib õppijakeelt, märkides küsitavaks laused, mis ühelegi mallile ei vasta. Artikkel tutvustab mallide kogu loomise protsessi ja tekstilause sõnajärge hindavat programmi. Antakse ka ülevaade programmi efektiivsusest õppijakeele tekstide analüüsil ning vealeidja edasise arendamise vajadustest. Õppijakeele analüüsil kasutati Tallinna Ülikooli eesti vahekeele korpust, mis koondab ligi 740 000 sõne mahus eesti keele õppijate loovkirjutisi ja harjutusi.*

Võtmesõnad: sõnajärg, korpuslingvistika, teise keele omandamine, eesti keel

1. Sissejuhatus

Teise keele õppimine on keerukas protsess ja õppija peab olema vapper ega tohi karta emakeelekõnelejate muigvel suid, kui ta on taas mõne iseäranis kohmaka lause koostamisega hakkama saanud. Artikli allikmaterjalid leidsid neid rohkesti, näiteks *Ühel hetkel mul tundakse, et olin Vana Egiptus, Pärast puhkemist ma uuesti sõitsin TTÜsse.*

* Tööd on toetanud riiklik programm "Eesti keele keeletehnoloogiline tugi (2006–2010)", projekt R0807 "VAKO – Eesti vahekeele korpuse keeletarkvara ja keeletehnoloogilise ressursi arendamine" ning sihtteema SF0180084s08 "Eesti keele morfosüntaktiline ehitus ja areng". Täname Vahur Rebast, kes aitas vealeidja prototüüpi programmeerida, Kaili Müüriseppa hea nõu eest morfosüntaktilise arvutianalüüsi küsimustes, Raili Pooli, kes võimaldas meile ligipääsu oma rektsioonisõnastiku (Pool 1999) elektroonilisele versioonile, ning anonüümseid retsensente kasulike tähelepanekute eest.

Üks keerukaid valdkondi keeleõppijale on sõnajärg. Käsitleme siin õppijakeele tuumlause sõnajärje automaatse tuvastamise küsimust, mis omakorda on tihedalt seotud kohustuslike lauseelementide õige määramisega. Keelespetsialisti ja informaatiku koostöös sündinud artikkel kirjeldab õppijakeele sõnajärge, tuumlause ehituse probleeme ning meetodit, kuidas oleme neid nähtusi automaatselt, süntaksist lähtuvalt analüüsinud. Kuna lause sõnajärg on kompleksne ning mitte täiesti reeglustatav keelevaldkond, siis kirjeldame siinkohal vaid esimesi samme nende ülesannete lahendamisel. Artikli sissejuhatav osa annab ülevaate olulisematest eesti sõnajärjeuringutest ning meie lingvistilistest lähtekohtadest ja eesmärkidest. Seejärel iseloomustame Tallinna Ülikooli eesti vahekeele korpuse (edaspidi EVKK)¹ materjali põhjal õppijakeele lausestruktuuri raskuskohti. Järgnevas osas tutvustame sõnajärje vealeidjat ja mallide kogu ning nende loomise käiku. Artikli lõpuosas hindame vealeidja töö edukust, seda mõjutavaid tegureid ja kirjeldame programmi töö tulemuste parandamise võimalusi.

2. Eesti keele sõnajärjest

Sõnajärge võib vaadelda tasanditi. Fraasitasandi sõnajärjes mängib rolli fraasituuma ja selle laiendite järjestus (nt nimisõnafaasis adjektiivfraaside ning neis omakorda adjektiivi laiendavate adverbide järjestus). Lihtlauses on lausetasandi sõnajärje puhul oluline subjekti, predikaadi (või mitmesõnalise predikaadi osade), seotud ja vabade laiendite ning ka muude tekstilause sõnajärge mõjutavate lauseelementide paiknemine (näiteks üldlaiendid). Kuigi paljudes kontekstides on eesti lausetasandi sõnajärg üsna vaba, kehtib V2-reegel enamasti nii normaal-, olemasolu- kui kogeja-omajalaluses (Helle Metslang jt 2003: 92–93). Eesti keele sõnajärjemallide sagedusi on vaadelnud Kaja Tael (1988) ja leidnud, et eesti lausete sõnajärjes on domineerivateks kombinatsioonideks subjekt-verb-X (25%), X-verb-subjekt (10%, verb asub sundpositsioonil, verbi muu asukoht muudaks tekstilause vastuvõetamatuks), X-verb-X (10%), verb-X (10%), X-verb-subjekt-X (9%). X tähistab kõigis mallides verbi laiendit, lause kõrvalliiget, mida võib olla üks või enam. (Tael 1988: 3–5)

Tael on tekstilause sõnajärje analüüsil lähtunud järgmistest näitajatest:

- fraaside süntaktilised omadused (laiendite liik ja asukoht, sõnade arv tarindis);
- fraaside semantilised omadused (fraasi semantiline roll ja semantiline tüüp);
- fraaside infostruktuurilised omadused (nt fraasi kuulumine tugevasse või nõrka teemasse või reemasse ning fraasi poolt kantud info olulisus).

Oma töös on ta kasutanud tervelt 67 sõnajärje varieerumisega seotud muutujat (nt *subjekt*, *topikalisatsiooni põhjus*) ning neist igäihe all on 0–15 alakatagooriat (nt *elusolend*, *emfaas*, *kontrast*). (Tael 1988: 3, 47)

Eesti sõnajärje aluseks on pigem infostruktuurilised kui süntaktilised printsiibid, eriti suulise kõne puhul (Lindström 2005: 173). Nii võib see olla ka õppijakeele puhul, kes enamasti kuni üsna kõrge keeleoskustasemeni ei erista hästi eri registritest pärit sisendit ning kelle kirjalike tekstide sõnajärg võib olla suulisest

kõnest mõjutatud. Siiski leiame, et mitmekesiste süntaktilise analüüsi võimaluste olemasolul² tasub uurida, kui palju toetab süntaktiline info sõnajärje ja tuumlause argumentide uurimist.

Põhilised eesti sõnajärje tendentsid on süstematiseerinud Nikolai Rimmel (1963), tuues välja neutraalse lause erinevate liikmete paiknemise seaduspärad (vt lähemalt jaotis 4). Ta sõnastas lause terviklikkuse seaduse, mis aitab edukalt selgitada ka õppijakeele sõnajärjeküsitavusi:

.. sõnal, mis lausele terviklikkuse (või lõpliku terviklikkuse) annab, on tendents pealause lõppu asetuda, kuna aga öeldise laiend, mida kõige hõlpsamini saab lausest välja jätta, püüab, vastupidi, vahetult öeldise järele paigutada .. (Rimmel 1963: 260–261)

Rimmeli (1963: 259) järgi annab rahulikule neutraalsele lausele terviklikkuse näiteks sihitis, kui see laiendab sihilist verbi, mis ei vaja peale sihitise muud laiendit, nt *Ta lõpetas eile töö*. Valdavalt deskriptiivne käsitlus annab ülevaate kesksete lauseliikmete paiknemisest pealauses, abimäärsõnade ja verbi infiniitsete vormide paigutusest ning ka sellistest spetsiifilisematest kontekstidest nagu kõrvallused, küsilused, emotsionaalsed ja erirõhuga laused.

Meie vaadeldava lausetüübi ja süntaksianalüüsiga (vt lähemalt jaotis 3) seostuvad ka järeldused, mille Liina Lindström on teinud suulise kõne sõnajärjest. Kui lause alguses on sihtis, öeldistäide, põhjus- või viisimäärus või kui subjektiks on pronoomen (peamiselt esimesele või teisele isikule viitav), siis on lauses sagedamini subjekti ja verbi otsejärg. Pöördjärjestust eelistatakse juhtudel, kui lause algab valdajamäärusega või kui subjektiks on täisnimisõnafraas. (Lindström 2005: 173–174) Eestikeelsete ilukirjandustekstide sõnajärje kohta on teada, et verbi ees kaldub verbifraasis paiknema adverbiaal – objekt ja predikaatiiv on seal harvem. Predikaativilausetes on sõnajärg tavaliselt SVX, muudes *olema*-verbiga lausetes XVS (eksistentsiaallused ja kogeja-omajalused). (Huumo 1994: 280–285) Osalauseste paiknemisest eesti liitlausest on ilmunud doktoriväitekirj Kirsi Höglundilt (2006).

3. Sõnajärje arvutianalüüsi eesmärgid

Kavandasime esmaseks uurimisobjektiks lihtlause (või ka võimalikult neutraalse osalause), mis oleks võimalikult terviklik ning mille tuumargumentide struktuur ja sõnajärg oleksid vähe mõjutatud teisestest grammatilistest teguritest. Näiteks ei soovinud me uurimusse kaasata põimlause kõrvallauset, sest sellest võib kontekstiellipsi tõttu puududa mis tahes komponent, ka predikaat, selle verbivorm võib olla infiniitne ning sõnajärg võib olla eripärane (EKG II: 276). Samas soovisime, et laused oleksid tekstilused, sest üldjuhul tegelevad õppijad just kontekstisidusate lausete harjutamisega. Lisaks täislausetele olid siiski vaatluse all ka narratiivses tekstis laialt levinud aluseta lünklaused. Tähelepanu alt jäid välja umbisikulises tegumoes laused ning hüüd- ja küsilused ning võimalusel ka mitmesõnalisi predikaate sisaldavad laused.

Ennekõike pakkus huvi lause struktuurne kese: predikaat ja ta seotud laiendid, mis sõltuvad otseselt verbi semantilisest tähendusest ning mille grammatilised ja

² Vt nt www.keeleeveeb.ee (10.02.2010) ja www.cl.ut.ee (10.02.2010).

semantilised omadused tulenevad verbi argumentstruktuurist. Võtsime arvesse nii obligatoorsed kui fakultatiivsed seotud laiendid, verbi vabad laiendid püüdsime vaatluse alt välja jätta, välja arvatud juhtudel, kui nad olid lause algul ja mõjutasid otseselt lause kohustuslike elementide sõnajärge, nt ***Kord hommikupoolikul helistas mulle sõber Joonas***. Nõnda jäid küll kõrvale mitmed õppijaile raskusi valmistavad lausetasandi sõnajärje probleemid (sh määruste omavaheline paiknemine), kuid see valik kitsendas pisut meie ees seisvat üsna ulatuslikku ülesannet. Kuna tekstides tuli sageli ette komplementlauseid, milleta pealause oleks poolik, siis võtsime arvesse ka mõned osalausekujulised verbilaiendid.

Käesolevas artiklis nimetame oma uuritavat struktuuri tuumlauseks. Lähene mine eesti keele sõnajärje korrektsuse hindamisele põhines kirjakeele korpusel kui emakeelekõnelejate poolt kirjapandud ja toimetatud tekstil, sõnajärje uurimustes esiletoodud seaduspärasustel ning autorite introspektsioonil. Meie koostatud õigete (paremate) sõnajärjemallide kogu esindab pigem tendentse kui reegleid, neist kõrvalekaldumist ei saa sageli veaks lugeda, sest neid võib ette tulla ka emakeelsete kõnelejate tekstides. Hinnates õppijakeele lauseid, otsustasime kahtluse korral lausejärje siiski tinglikult küsitavaks (ebasoovitavaks) hinnata, sest leidsime, et on parem, kui jääb alles võimalus nendele kohtadele tähelepanu juhtida.

Meie töö oli kolm **eesmärki**:

- 1) püüda automaatselt tuvastada õppija tuumlausete struktuuri;
- 2) püüda automaatselt hinnata tuumlause argumentide jm elementide järje vastavust mallile (õigsust);
- 3) saada ülevaade teguritest, mis õppijakeele sõnajärje arvutianalüüsi positiivselt ja negatiivselt mõjutasid.

Võtsime korrektse eesti keele sõnajärje leidmisel näitematerjaliks Tartu Ülikooli kirjakeelekorpuse (TÜKK)³ ilukirjanduse osa. Tegime õppijakeele sõnajärje vigade analüüsimisel selle valiku, sest ilukirjanduse sõnajärge on eesti keeles põhjalikumalt uuritud (nt Remmel 1963, Huumo 1994, Lindström 2005). Lisaks lausete tekstisidususele sobisid ilukirjandustekstid tänu narratiivsele ja arutlevale tekstiliigile (kasutasime enam jutustavaid ja võimalikult vähe dialoogi sisaldavaid tekste), inimese elu ja vaba aega puudutavale temaatikale ja hinnangute rohkusele. Ka vabakirjutuslikes õppijatekstides jutustatakse sageli autori elust ja kogemustest ning antakse mitmesugustele nähtustele hinnanguid. Ilukirjandustekstid erinevad õppijate tekstidest lausete pikkuse ja kohatise keerukusega: õppijakeele lausete hindamise juures ei ole oluliselt abi rohkete osalause, täiendite ja verbilaiendite komplekssest analüüsist.

4. Õppija sõnajärje- ja tuumargumentidevad

Kõigi vajalike ja õiges vormis tuumargumentide valimine lausesse ning nende õige järjestamine on edasijõudnud keeleõppija jaoks sageli selleks “koera sabaks”, millest ei jõuagi üle astuda ning mis võibki jääda eristama õppija keelt emakeelekõneleja omast. Samas leidub EVKK-s loomulikult ka rohkesti näiteid kõigi soovituslike sõnajärjemallide kasutamisest sarnaselt emakeelekõnelejatele.

³ [http://www.cl.ut.ee/korpused/baaskorpus/\(10.02.2010\)](http://www.cl.ut.ee/korpused/baaskorpus/(10.02.2010)).

Toome siinkohal valiku probleemidest, mis õppijail tekkisid, ning seletame neid R Emmeli (1963) esitatud eesti keele sõnajärje seaduspärasuste alusel (vt ka Helle Metslang jt 2003: 118–121, 139–141, 149–151). Vajadusel lisame selgituseks praegu käibivad terminid “Eesti keele grammatika” (EKG II) mõistestikust. Siinse jaotise lausenäited pärinevad EVKK-st. Vealeidja edukust vaadeldud eksimustega toimetulekul vaatleme jaotises 6.

Õppijakeele lausetasandi sõnajärje kõrvalekalletest moodustas V2-malli mittejärgimine üle 2/3, näiteks: *Ta alati kuhugi hilineb* (parem: *Ta hilineb alati kuhugi*). Inversiooniga lauseid leidis õppijatekstides üldse alla 1% (mis on tunduvalt vähem kui emakeelsete kirjutajate tekstides, vrd Taela andmed jaotises 2). Õppija-lauses esines lisaks tavalisele V2-reegli vastu eksimisele järgmisi probleeme.

Üldjuhul põhjustab kõrvalekaldeid V2-st liigse(te) moodustaja(te) toomine predikaadi ette, kuid rohkelt leidub ka näiteid, kus üheks predikaadi ees paiknevaks lauseelemendiks on üldlaiend: *Vabal ajal Narvas ma käin discol*.. (parem: *Narvas käin ma vabal ajal diskol*); *Minu arvates see on inemeste tavad ja traditsioonid* (parem: *Minu arvates on need inimeste tavad ja traditsioonid*); *Aga samuti ma olin loomaaias* (parem: *(Aga) samuti olin ma loomaaias*).

Öeldistäite neutraalseim asukoht on lause lõpus (Remmel 1963: 261, 323). *See punkt on väga tähtis minu jaoks* (parem: *See punkt on minu jaoks väga tähtis*). *Otsekohe tundus teissugune õhkkond* (parem: *Otsekohe tundus õhkkond teistsugune*). Viimase näite ajamääruse õiget asukohavalikut kinnitavad järgmised seaduspärad. Pealause öeldise laiend võib R Emmeli (1963: 258) järgi paikneda nii lause alguses, öeldise järel kui aluse ja öeldise vahel. Kui lauses on kaks öeldise laiendit, üks neist küsimusele *millal?* vastav ajamäärus, siis nende lahutamise korral saame kõige neutraalsema lause, kui algusse paigutame just ajamääruse (Remmel 1963: 227).

On verbe, mis väljendavad terviklikku mõtet ainult koos teise, ma- või da-infinitiivi vormis oleva verbiga. Seetõttu on neid verbe laiendaval infinitiivil tendents lause lõppu asetuda. (Remmel 1963: 264) *Ning ei ole vaja oodata neid koolist* (parem: *(Ning) ei ole vaja neid koolist oodata*).

Remmel on tähendanud, et **kui lauses on öeldistäitemäärus** (EKG II: latiivne seisundimäärus), **siis asetub ta lõppu, sest ta annab lausele terviklikkuse**. Lisaks osutab ta, et **väliskohakäändes olev sihitismäärus** (EKG II: valdajamäärus või sõltuvusmäärus) **ei anna lausele terviklikkuse iseloomu ja asetub seepärast teistest öeldise laienditest ettepoole**. (Remmel 1963: 260) *Kõik muutused said tõeliseks sokiks inimestele*.. (parem: *Kõik muutused said inimestele tõeliseks šokiks*).

Järgmine lause esindab kõrvalekallet eesti keeles väga sagedase *vaja olema*-lause tavajärjest: *Energia on vaja inimestele* (parem: *Energiat on inimestele vaja*). Remmel (1963: 238–239) **loeb kogejalause semantilise subjekti sihitismääruseks** (EKG II: valdajamäärus) **ning on leidnud, et see paikneb öeldise suhtes üldiselt samas asendis nagu tavalause alus, s.t pealause öeldisega kõrvuasendis**.

Kui personaalses konstruktsioonis on öeldiseks sihiline verb, mis selleks, et lause terviklik oleks, ei vaja muud laiendit peale sihitise, asub sihitis lause lõppu ja teised öeldise laiendid sellest ettepoole (Remmel 1963: 259). Järgmises õppijalauses, kus predikaadiks vaid üht seotud laiendit

nõudev verb, on sihitise asukoht lauses küsitav: *Ta pakkub palju võimalusi edaspidi* (parem: *Ta pakub edaspidi palju võimalusi*)⁴.

Keerulisem on põhjendada, miks mõjub järgmise lause järg küsitavalt: *Aga samuti reisimine annab võimalusi inimestele puhkama ja nautima loodusega* (parem: *(Aga) samuti annab reisimine inimestele võimalusi puhata ja loodust nautida*). Siin on probleemiks nimisõnalise peasõna ning järeltäiendi lahutamine. Kuid see selgitus ei anna vastust küsimusele, miks on parem tuua määrus sihitisest ettepoole – siin on vaja vaadata mitut R Emmeli sõnastatud seadust koos. R Emmel (1963: 259) toob välja tendentsi, et **lauses, kus peale sihitise on** (seotud laiendina – autorite kommentaar) **ka liikumise sihtkohta märkiv kohamäärus, asub viimane lõpus ja sihitis vahetult selle ees**. Samas toob ta näiteid, kuidas **sihitis liigub lõpu poole keerukama konstruktsiooniga lauses, näiteks kui nimisõna laiend seisab oma põhisõna järel** (1963: 266, 327). Need reeglid näivad laienevat ka muudele sihti märkivatele määrustele, nt valdamäärustele, ning võiksid siinset juhtumit selgitada⁵.

Vaegisikulise konstruktsiooni kasutamise kaasnep kõrvalekalle on: *Võib võrrelda nende Big Ben'i meie Oleviste kirikuga* (parem: *Nende Big Beni võib võrrelda meie Oleviste kirikuga*).

Tuumlause sõnajärje probleemist on lahutamatu tekstilause **kõigi vajalike tuumelementide olemasolu** küsimus lauses. Õppijatekstides tuleb ette nende ekslikku väljajätmist, mittevajalike elementide lisamist ning vale süntaktilise rolli kasutust (näiteks objekti asemel adverbiaal). Õppijate lausemoodustusraskusteks olid meie materjalis predikaadi puudumine (*Kõrval seina juures diivan*, parem: *Kõrval seina juures on diivan*); koopula puudumine (*See on ei ole suur tuba, aga päikesepaisteline*, parem: *See ei ole suur tuba, kuid on päikesepaisteline*); subjekti ja verbi väljajätt predikatiivilauses (*Väga raske küsimus*, antud kontekstis olnuks parem: *See on väga raske küsimus*); kohustusliku verbilaiendi puudumine (*Ma tahan alustada, et ...*, parem: *Ma tahan alustada sellega, et ...*); objekti kasutamine reksiooniadverbiaali asemel (*Puhkuse lõpuks mina saanud aru üks asi ...*, parem: *Puhkuse lõpuks sain ma ühest asjast aru ...*); lisaks ka puuduv sidend komplementlause algul (*Aga ma arvan küll, kõige parem ja armas koht on ...*, parem: *(Aga) ma arvan küll, et kõige parem ja armsam koht on ...*).

5. Sõnajärje mallikogu ja vealeidja prototüübi loomine

Meetodeid, mida sõnajärje vealeidja programmi loomise aluseks valida, leidub palju, näiteks lingvistilistel reeglitel (vt nt Arppe 2000) ning statistikal põhinevad (Athanaselis jt 2006). Meie valikuks oli reeglipõhine lähenemine, et kasutada juba olemasolevaid eesti keele ressursse ja tarkvara. Võtsime vealeidja loomisel aluseks Kaili Müürisepa loodud kitsenduste grammatika süntaksianalüsaatori (edaspidi: parser)⁶ ja morfosüntaktiliselt märgendatud ilukirjanduskorpuse⁷. Töö koosnes järgmistest etappidest: 1) esialgse mallide kogumi loomine ilukirjanduskorpuse ja

⁴ Selgitus kehtib, kui lugeda võimalikuks, et *pakkuma* võib eesti keeles esineda nii mono- kui ditransitiivse verbina (vrd: *Ta pakub edaspidi palju võimalusi*. *Ta pakub mulle edaspidi palju võimalusi*).

⁵ Pooli (1999) sõnastikus on ligi 106 sellist verbi, millest esialgsel vaatlusel 79 puhul mõjub sihti märkiva määrusega lõppev lause tavaliselt tõepoolest neutraalsemalt (*Kohandasin artikli nõuetele*) ja vaid 27 juures on neutraalsem sihitisega lõppev lause (*Selgitasin sõbrale oma ideed*). Vaatluse alt jäid välja kahe laiendiga verbid, millest üks oli sulgude abil vabaks laiendiks märgitud.

⁶ EstCG Parser 1.0a, <http://www.cs.ut.ee/~kaili/parser/>, <http://lepo.it.da.ut.ee/~kaili/grammatika/> (19.02.2010).

⁷ <http://www.cl.ut.ee/korpused/syntaksikorpus/> (11.02.2010).

EVKK väljavõtete põhjal, 2) õigete ja valede mallide automaatne lisamine kogusse ning käsitsi kontrollimine, 3) kogus olevate õigete mallide kasutatavuse hindamine kirjakeele ja õppijakeele korpusel, 4) vealeidja prototüübi loomine, mis annab mallikoguga võrdluse põhjal õppijalauseste sõnajärjele hinnangu ja 5) vealeidja otsuste pisteline käsitsi kontroll ning tulemuste hindamine.

5.1. Sõnajärje mallid ja nende tuvastamise võimalused

Alustasime tööd emakeelekõnelejate lausetest pärinevate mallide kogumisega. Mallikogu koostamiseks oli vaja uurida lauseid nende formaalse esituse kujul, milleks sobis lausete esitamine morfosüntaktiliste märgendite abil. Parseri väljund kujutabki endast lause põhjalikku tõlgendust, kus on esitatud kõigi sõnade sõnaliik, vorm ning süntaktiline roll (parseri analüüs on pindsüntaktiline ning ei kajasta lause hierarhilist struktuuri). Formaalseks ning otsitavaks malliks peame siinkohal järjestatud komplekssete morfoloogiliste ja süntaktiliste tähiste (märgendite) jada, mis esitab lause struktuuri ning milles võivad esineda ainult vajalikud ehk malli eristamise seisukohalt olulised osad. Mallid koosnevad parseri väljundist valitud grammatilistest tähistest, nende selgitused on lisatud artikli lõppu.

Mallide otsimise etapil olid probleemiks traditsiooniliste arvutuslike meetodite kasutamise raskused. Olemasolevad tuntud keeleanalüüsi programmid on mõeldud korpusuuringuteks ja on kohandatud töötleva tekstimaterjali, mis on esitatud tähtede ja kirjavahemärkidega ega sisalda muid sümboleid. Kui sõnadest koosnevas tekstis saab niisugune tarkvara (näiteks WordSmith Tools 5.0) eraldada näiteks sõnadest koosnevaid kollokatsioone, siis formaalsete tähistega see kahjuks ei tööta. Loomuliku keele sõnalõpu tunnuseks on tühik või kirjavahemärk. Morfosüntaktilise märgendi lõpu tunnuseks ei pruugi tühik sobida, sest tühikud paiknevad ka märgendi sees. Näiteks on lauses *Mul on mälestused koolipõlvest* vormi *koolipõlvest* morfosüntaktiline esitus järgmine:

kooli_põli+st // _S_ com sg el // @<NN.

Kaalusime võimalust kasutada tuumargumentide järjestuse hindamisel n -gramm-meetodit (vt nt Alam jt 2006) ning vastavat tarkvara.⁸ Nägime siin raskuskohtadena süsteemi madalat töökiirust ning suurt varieeruvust õppijakeele ortograafias ja vormimoodustuses. Lisaks esitab vaid osa sõnale vastava morfosüntaktiliste tähiste kompleksi märgenditest tuumlause sõnajärje jaoks olulist informatsiooni. Need esinevad teiste, ebaoluliste märgendite vahel ning see struktuur ei sobi n -grammi moodustamiseks. On vaja algoritmi, mis lubaks selekteerida välja olulised märgendid ning uurida nende järjestust ning korduvust.

Esmärgiks on leida suurim võimalik kogus nii õigeid kui ka küsitavaid (edaspidi tinglikult: valesid või ebakorrektsed) malle. Selline lähenemine annab võimaluse programmi sisendi täpsemaks liigitamiseks sõnajärje korrektsuse osas. Formaalse mallide otsimiseks vajasime abivahendeid, mis tooksid esile korduvad grammatiliste märgendite kooslused. Neid kirjeldab jaotis 5.2.

⁸ n -gramm on n järjest paikneva tähe või vajadusel sõna jada. Näiteks lausest *Poiss leidis võtmed üles* võib eraldada järgmised bigrammid (kahest osast koosnev jada): *poiss leidis*, *leidis võtmed*, *võtmed üles*. Korpusel leitud n -grammide sageduse põhjal on võimalik prognoosida, millise tõenäosusega ilmub vaadeldav osa, samuti millise tõenäosusega ilmub järgnev sõna, kui esimene sõna on fikseeritud. Meetodi rakendamiseks on võimalik kasutada spetsiaalset tarkvara, näiteks kfngram: <http://kwicfinder.com/kfngram/kfngramHelp.html> (11.02.2010).

5.2. Mallileidja töö ettevalmistus

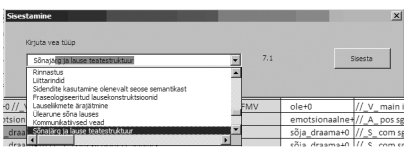
Kasutasime malliotsija prototüübi sisendina parseri väljundit. Õigete mallide otsimise juures kasutasime abivahendina tabelit, kus kirjakeelsete lausete järele olid ritta paigutatud kõigi lause sõnade morfosüntaktilised märgendid (ebakorreksete märgendijadade otsimiseks kasutasime sarnast tabelit normist kõrvalekalduvaks märgitud õppijalausetest ning nende märgenditest). Kuid tabelite loomine eeldas parseri väljundi eelnevat kohandamist sobivale kujule.

Märgendijadade kogu loomiseks kopeeritakse eelnevalt parseri väljund etteantud tabelisse, mis on loodud Microsoft Exceli keskkonnas Visual Basic for Applications programmeerimisvahendite abil. Teksti sisestamiseks õigesse tabeli osasse on loodud spetsiaalne vorm (nupp “Sisesta tekst”). Seejärel on võimalik programmeeritud makrode abil teha parseri väljund korda nii, et sõnad, sõnade vormid ning morfosüntaktilised tähised oleksid eraldi tulpades (joonis 1).

	A	B	C	D	E	F	G	H	I	J	K	L
1	Sisestetekst											
2	Vea tüüp											
3	Tekst korda	Paiguta	Süntaksianalüsaatori väljund	Vea tüüp	Sõna vorm	Morfosüntaktiline analüüs	1	1.1	1.1.1	1.2	1.2.1	1.2.2
4	1	SLAS	kool+st // S_com sg el #cap // **CLB @ADVL		kool+st	// S_com sg el #cap // **CLB @ADVL	0	0	0	0	0	0
5	2	Koolist	kool+st // S_com sg el #cap // **CLB @ADVL		kool+st	// S_com sg el #cap // **CLB @ADVL	0	0	0	0	0	0
6	3	ma	mina+0 // P_pers ps1 sg nom // @SUBJ	7.1	mina+0	// P_pers ps1 sg nom // @SUBJ	0	0	0	0	0	0
7	4	ei	ei+0 // V_aux neg // @NEG		ei+0	// V_aux neg // @NEG	0	0	0	0	0	0
8	5	või	võ+0 // V_mod indic pres ps neg #FinV #Intr // @+FCV		võ+0	// V_mod indic pres ps neg #FinV #Intr // @+FCV	0	0	0	0	0	0
9	6	midagi	mis+vdagi // P_indef sg part // @ADVL		mis+vdagi	// P_indef sg part // @ADVL	0	0	0	0	0	0
10	7	eriti	erit+0 // D // @ADV		erit+0	// D // @ADV	0	0	0	0	0	0
11	8	halba	halb+0 // A_pos sg part // @CAN		halb+0	// A_pos sg part // @CAN	0	0	0	0	0	0
12	9	uttelda	uttelda+0 // S_com sg nom #? // @PRD		uttelda+0	// S_com sg nom #? // @PRD	0	0	0	0	0	0
13	10	S	.		.		0	0	0	0	0	0
14	11	SLS	.		.		0	0	0	0	0	0
15	sisesta											

Joonis 1. Mallide otsimise eeltöö. Parseri väljundi kohandamine Microsoft Exceli keskkonda

Kui tegu on õppijakeele lausetega, siis järgmise etapina on vaja tekst märgendada, et tuua esile ühe või teise veaga seotud kohti. Töö lihtsustamiseks on lisatud vorm, mis esimeste sisestatud tähtede järgi pakub veaklassifikaatori numברי (joonis 2; EVKK loomisel on vorm kasutusel ka teiste vealiikide märgendamisel). Õppijalausetete korrektsust hindas keeletespialist samas tabelis käsitsi. Valede mallide otsingul olid kasutusel vaid sõnajärje vea märgendi saanud laused.



Joonis 2. Vorm veaklassifikaatorite lisamiseks

Lause ning sellele vastavad töödeldud morfosüntaktilised tähised paigutatakse automaatselt eraldi Exceli töölehele (joonis 3).

	A	B	C	D
1	Ormeti ma ei saanud oma vana s // D_ #cap // **CLB @ADVL		// P_pers ps1 sg nom // @SUBJ	// V_aux neg // @NEG
2	Smerdjakov oli Kesk-Venemaalt // S_prop sg nom #cap #? // **CLB @SUBJ		// V_main indic impf ps3 sg ps af #FinV #Intr // @+FMV // S_prop sg abl #cap // @ADVL	
3	See on professionaalne vilumus // P_dem sg nom #cap // **CLB @SUBJ		// V_main indic pres ps3 sg ps af #FinV #Intr // @+FMV // A_pos sg nom // @AN>	
4	Smerdjakov oli mulle mõjunud s // S_prop sg nom #cap #? // **CLB @SUBJ		// V_aux indic impf ps3 sg ps af #FinV #Intr // @+FCV // P_pers ps1 sg all // @ADVL	
5	Ta sarnanes saagihanele õngem // P_pers ps3 sg nom #cap // **CLB @SUBJ		// V_main indic impf ps3 sg ps af #FinV #Intr // @+FMV // A_pos sg all // @AN>	

Joonis 3. Mallide loomisele aluseks olevate morfosüntaktiliste märgendijadade tabel

Korrektsete lausete korral saab otsida, mis kombinatsioonid korduvad, ja selliste korduvate osade abil kirjutada malle õige lause tuvastamiseks. Sõnajärvevigadega lausetest leitud mallid lisasime valede mallide kogusse.

Mallide otsimiseks on vaja käsitsi moodustada a) tekstisõnade ning b) morfo-süntaktiliste märgendite hulgad (kasutame terminit *hulk* siin matemaatilise hulga mõistes), millega sõnajärje uurimisel ning mallide tuvastamisel ei pea arvestama.

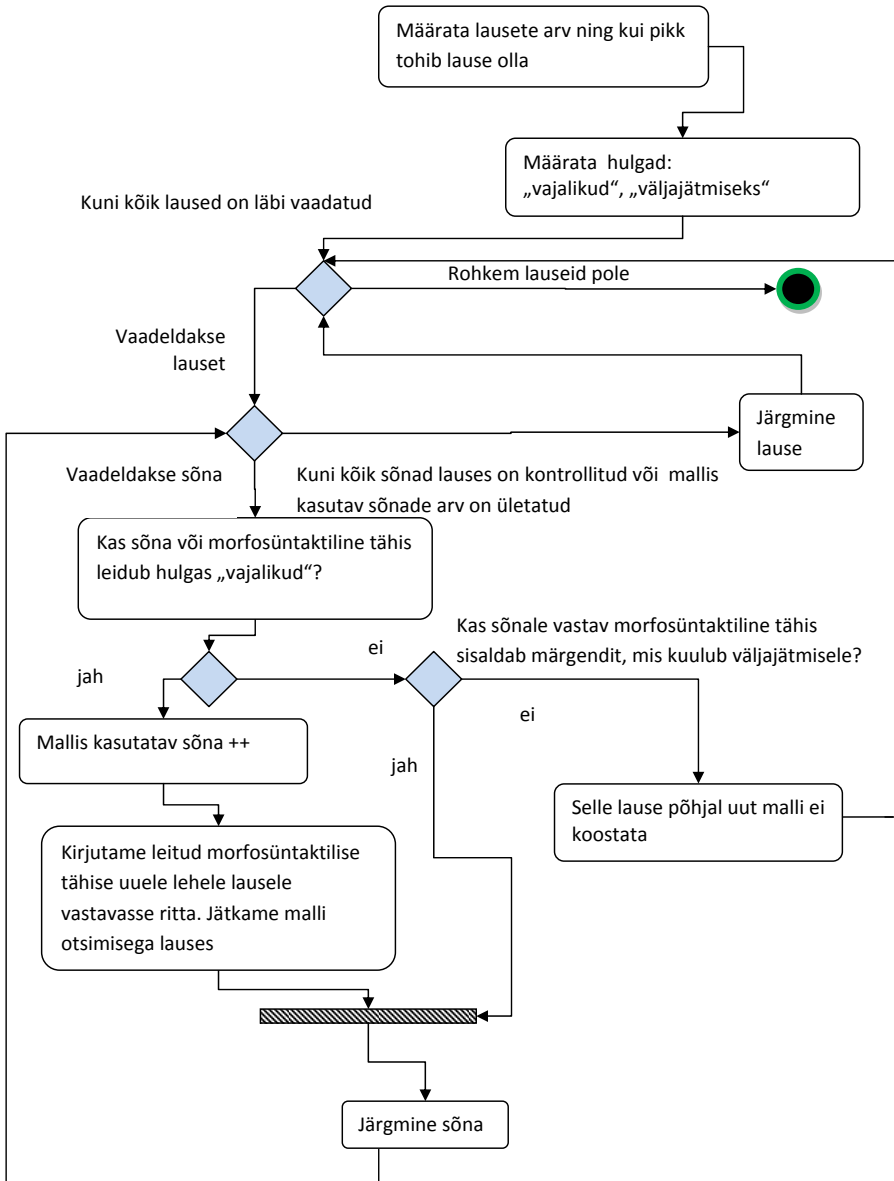
Fraasisisesed laiendid (moodustajatüüpidega nt omadussõnalised eestäiendid märgendiga @AN>, kvantori järellaiendid @<Q; loomuliku keele sõnaloenditena nt adjektiivi ja adverbilise laiendid nagu *päris*, *väga*) jäävad vaatluse alt välja. Teises väljajätavate lauseelementide hulgas olid üldlaiendina esinevad ühendid ja rõhumäärsõnad (nt *minu meelest*, *ju*, *siis*, *ka*, *hoopis*) ning sidesõnad, mis ei pruugi põhjustada subjekti ja predikaadi inversiooni ning millega on vaja seetõttu tavapärase V2-reegli puhul arvestada. Selle ning ka adjektiivi- ja adverbialifraasi laiendliikmete hulga reeglistik vajab edaspidi veel täpsustamist, sest nende sõnade käitumine on sõnajärje suhtes varieeruv ning nad võivad esineda ka erinevates funktsioonides (nt *siis* võib olla ka ajamääruse rollis). Lisaks on vaja moodustada hulk nendest märgenditest, mis on vajalikud ning mille abil on võimalik hinnata, kas lause sõnajärg on korrektne või mitte. Tuumlause sõnajärje analüüsi jäetakse sisse predikaat, ta argumendid (nt @SUBJ) ja vabad määrused.

5.3. Mallide otsimise algoritm

Mallide otsimiseks kontrollitakse märgendijadade tabelis kõigepealt lausete arvu ning arvutatakse maksimaalne lause pikkus vaadeldavas lausetekogumis. See informatsioon on vajalik, et tagada, et iga lause ning iga sõna saaks töödeldud. Järgnevalt loetakse sisse nii vajalike kui väljajätavate sõnade hulgad. Eri tüüpi hulkade olemasolust tulenevalt peab lauset olema samaaegselt võimalik analüüsida nii loomuliku keele sõnajärjendi kui ka märgenditejada kujul (vt jaotis 5.2). Programmil tuleb analüüsida nii terviklikke lihtlauseid kui liitlause osalauseid (vt jaotis 3) – seetõttu lisasime algoritmile ka lihtsamad osalause piiri reeglid.

Mallide tuvastamise protsessi käigus, kuni kõiki lauseid ei ole veel läbi vaadatud, võtab programm uurimiseks järjekorras järgmise lause (vt joonis 4). Lausesiselt vaadeldakse sõnu üksteise järel. Kõigepealt otsitakse sõna väljajätmiseks lubatud loomuliku keele sõnade hulgast. Kui sõna sealt ei leita, siis kontrollitakse vajalike märgendite hulka. Kui sellesse kuuluv sobiv märgend leitakse, siis kirjutatakse märgend välja vastavasse lahtrisse ning minnakse järgmise sõna analüüsi juurde. Kui sõna või märgend leidub aga väljajätavates hulkades, siis ei kirjutata seda vastavasse lahtrisse välja, vaid lihtsalt jätkatakse analüüsi, minnes järgmise sõna juurde. Juhul kui nimetatud hulkadest vastet ei leita, lause analüüs katkestatakse. Näiteks osalause piiri märgend CLB ei kuulu ülaltoodud hulkadesse ning kui tegu ei ole just lause vasakpoolseima osalause alguse märgendiga, siis selle lause uurimine lõpetatakse ning minnakse järgmise lause analüüsi juurde.

Tabel 1 esitab näitena lause *Juba paar nädalat valitses põud* (TÜKK) analüüsi.



Joonis 4. Mallide otsimise algoritm. Joonis näitab malliotsija otsuseid ühe lause ning ka kogu vaadeldava teksti töötlemisel. Plokkskeemil kujutatud protsessis kirjutatakse sobivate lausete põhjal mallikogusse uued mallid, mittesobivad laused jäetakse vahele. Rombidega esitatakse hargnemisi ja tsükleid, tehtega ++ näidatakse sammu (muutuja) suurenemist, programmi lõppu tähistatakse ringiga

Tabel 1. Mallileidja sisend

Juba	paar	nädalat	valitses	pöud	.
//_D_ #cap // **CLB @ADVL	//_N_ card sg nom I // @ADVL	//_S_ com sg part // @<Q	//_V_ main indic impf ps3 sg ps af #FinV #Part-P // @FMV	//_S_ com sg nom // @SUBJ	//_Z_ Fst //

Sõna *juba* leitakse loomuliku keele sõnade hulgast, mis on lubatud välja jätta, seda välja ei kirjutata ning minnakse järgmise sõna juurde. Järgmine sõna *paar* ei asu loomuliku keele sõnade hulgas ning tema morfosüntaktiline info ei sisalda ka märgendit, mis oleks lubatud väljajätmiseks. Seejuures aga leitakse märgend @ADVL, mis on sõnajärje seisukohalt oluline ning asub hulgas “vajalikud”. See märgend kirjutatakse vastavasse lahtrisse välja, seejärel minnakse järgmise sõna juurde. Sõna *nädalat* on lubatud väljajätmiseks, sest selle märgend @<Q kuulub vastavasse hulka. Märgendit ei kirjutata välja ning vaatluse alla võetakse järgmine sõna. Selleks on *valitses*, mille morfosüntaktilises analüüsis leidub märgend @FMV, mis on pärit hulgast “vajalikud”. Märgend kirjutatakse välja ning asutakse uurima sõna *põud*. Märgend @SUBJ on samuti hulgas “vajalikud” ning kirjutatakse välja. Järgmisel kohal asub punkt ning sellele vastav märgend, millega analüüs lõpetatakse, kuna see ei kuulu ühtegi hulka. Tulemuseks saame malli: ['@ADVL' '@FMV' '@SUBJ'].

Elkirjeldatud algoritmi abil saadud mallid läbisid keelespetsialisti kontrolli, mille käigus täiendati hulkasid ja täpsustati märgendeid mallides.

5.4. Õigete mallide kogu

Ilukirjanduskorpuse tekstide põhjal eraldati 148 malli, neile lisati lingvisti poolt käsitsi moodustatud 94 malli. Viimaste eesmärk on parandada süsteemi suutlikkust tulla toime eri tüüpi ühend- ja väljendverbe sisaldavate lausetega ning subjekti-ellipsiga lausetega.

Mallikogus domineerivad lihtsad süntaktilised rollid (nt @SUBJ), kuid sõnu on ka piiritletud näiteks ainult sõnaliigi (ühendverbi koosseisu kuuluv adverb) või käände abil (kogejalause esimese sõna vormiks on adessiiv või allatiiv). Tabel 2 iseloomustab meie mallikogu lähemalt.

Tabel 2. Lauseliikmete positsiooneelistused mallikogus (242 malli)⁹

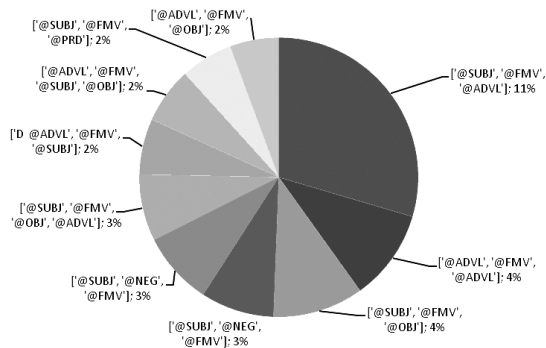
Lauseliige	Esinemus kokku	Positsioon lauses					
		1.	2.	3.	4.	5.	6.
@SUBJ	130	88	4	19	15	4	–
@OBJ	133	1	18	37	53	20	4
@PRD	10	1	2	4	1	2	–
@ADVL	147	29	6	57	35	17	3
@FCV	63	12	48	3	–	–	–
@FMV	95	16	59	15	5	–	–
@IMV	82	–	12	21	29	20	–

Mallide sees esinevad esimesel positsioonil kõige sagedamini subjekt (88 korda) ning adverbiaal (29). Malli teisel positsioonil on ootuspäraselt populaarseimad lauseliikmed finiiised predikaadid ning mitmeosaliste predikaatide finiiised osad (vastavalt @FMV 59 korda ning @FCV 48 korda). Malli kolmandal positsioonil domineerivad adverbiaal (57, põhiliselt abimäärsõnad) ja objekt (37). Predikatiiv esineb kõige rohkem kolmandal kohal (4 korda). Tabeli alusel on edaspidi võimalik

⁹ Mallikogus jagunevad adverbiaalid kaheks: adverbist peasõnaga määrused (eelkõige ühendverbide komponendina) ning määrused, mille fraasi peasõna sõnaliiki pole täpsustatud. Adverbist peasõnaga määrused eelistavad 3. ja 4. positsiooni, täpsustamata peasõnaga adverbiaalid esinevad põhiliselt 1. positsioonil.

moodustada ja süsteemiga liita reegliteblokk lauseliikmete positsioonidest, mis on suure tõenäosusega ebasoovitavad, nt subjekt teisel kohal ning mitmeosalise predikaadi infiniitvorm lause esimesel positsioonil.

242 malli leidmise järel kontrolliti suurema ilukirjanduse valimiga (681 lauset) nende esinemissagedust. Joonis 5 näitab protsentuaalset jaotust sagedasemate mallide vahel (mida esines 681 lause seas üle 1%). Uuritud valimis osutusid kõige populaarsemateks järjestusteks kolmikud ['@SUBJ', '@FMV', '@ADVL'] (73 lauset ehk 681-st mallist 11%, nt *Smerdjakov tuli ujumast*) ning ['@ADVL', '@FMV', '@ADVL'] (26 lauset ehk 4%, nt *Pärast istusime sauna ees terrassil* – mallide esinemissageduse hindamisel ja vealeidja töös lubasime malle sobitada lausega ka siis, kui adverbiaalide arv täpselt ei klappinud – vt lähemalt jaotis 7). Enamik sageli kasutatavatest mallidest ongi kolmekohalised. Seejuures on aga populaarsed ka mõned neljakohalised mallid: ['@SUBJ', '@FMV', '@OBJ', '@ADVL'], mille esinemissagedus on 3% (19 lauset, nt *Ta juhtis meie tähelepanu pääsukestele*) ning ['@ADVL', '@FMV', '@SUBJ', '@OBJ'] (*Seepeale võttis ohvitser relva*). 2%-line esinemissagedus on ka määrsõnalise adverbiaaliga (sageli üldlaiendiga) algaval mallil ['_D_..@ADVL', '@FMV', '@SUBJ'] (16 lauset, nt *Samas aga kõlasid lasud*). Suurem osa malle olid aga haruldased, esinedes vaid 1–2 korda. Seetõttu ei anna vealeidja arendamisel mallidehulga kasvatamine enam olulist efekti. Tulemuste edasiseks parandamiseks tasub täiustada mallide lausetega ühendamise viise.

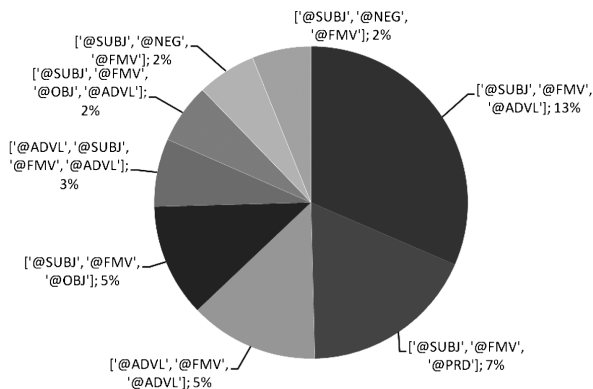


Joonis 5. Sõnajäremallide esinemus kirjakeeles (681 TÜ ilukirjanduskorpuse lause seas)

5.5. Kirjakeele mallide esinemus õppijakeelekorpuses

Võrdlesime kirjakeele korpusest leitud õigete mallide kogu õppijakeelega. Selleks võtsime õppijakeele korpusest juhumeetodil välja 4743 lausest koosneva valimi. Üldjoontes on tulemused võrreldavad ilukirjanduse omadega, kuid nad on ebatäpsemad lausetes esinevate õppijavigade ja mõningate keerukamate süntaksiküsimuste tõttu, mille lahendamine malle otsivale programmile hetkel üle jõu käib (vt jaotis 7). Toome siinkohal kokkuvõtte õppijate eelistustest tuumlause sõnajärjele.

Materjalis leiduvad õppijakeele lausestruktuurid vastasid vaid osale 242-st ilukirjanduse põhjal leitud mallist: õppija kasutas neist vaid 83, mis katsid ligi 3/4 õppijakeele lausetest.



Joonis 6. Õppijate eelistused sõnajärjemallide suhtes

Populaarsemate mallide osas oli nii kattuvusi kui erinevusi. Sarnaselt ilukirjanduskeelele oli õppijakeele valimis kõige sagedamini kasutatavaks malliks ['@SUBJ', '@FMV', '@ADVL'], mille osakaal on 13% (564 lauset, nt *Zaura õpib kohalikus koolis*). Teisel kohal on aga mall ['@SUBJ', '@FMV', '@PRD'] osakaaluga 7% (324 lauset, näiteks *Aga arvutite kasutamine on ka ohtlik*), mis ilukirjanduses oli ainult 2% osakaaluga. Kolmandal kohal on mall ['@ADVL', '@FMV', '@ADVL'] (5%, 238 lauset, nt *Esialgu olime paanikas*).

5.6. Õppijakeele sõnajärje vealeidja prototüüp

Jaotistes 5.4 ja 5.5 iseloomustasime sõnajärje mallide kogu, mis on sõnajärje vealeidjale sisendmooduliks. Järgnevalt anname ülevaate praeguseks väljatöötatud esimesest vealeidja prototüübist, mis on mõeldud rakendamiseks EVKK õppijakeelekorpuses.

Vealeidja eesmärk on anda automaatselt hinnang õppijateksti lausetele (vaid antud projektis vaadeldavatele, mitte kõigile – vt jaotis 3), mis on programmi sisestatud: kas nende tuumelementide sõnajärg on korrektne või suure tõenäosusega ebasooviv.

Prototüüp kasutab Unixi-põhist kitsenduste grammatika süntaksianalüsaatorit (parserit), mis on installeeritud korpusega seotud serverisse. Vealeidja hinnang põhineb õppijalausete võrdlusel õigete sõnajärjemallide koguga (tulevikus ka valede sõnajärjemallide koguga ning lauseliikmete tüüpiliste positsiooneelistuste infoga). Programmi töö koosneb järgmistest sammudest:

- 1) morfosüntaktiliste tähiste jadade lisamine õppijalausetele parseri abil;
- 2) õppijalause sõnavormide ning morfosüntaktiliste tähiste võrdlemine mittevajalike sõnavormide ja märgendite hulkadega; neisse kuuluvate sõnade kustutamine vaatlusalustest järjenditest;
- 3) allesjäänud morfosüntaktiliste tähiste seast hulka “vajalikud” kuuluvate märgendite esiletõstmine (üle jäävad (osa)lausepiiri ja kirjavahemärkide märgendid);
- 4) esiletõstetud märgendite jadade võrdlemine õigete mallide koguga, võttes arvesse mallide rakendamise prioriteete (näiteks kui sama lause jaoks on

tuvastatud kaks malli ning üks neist koosneb kolmest ning teine neljast osast, siis kasutatakse pikemat malli). Valede mallide kogu käesoleva katse käigus ei rakendatud;

- 5) malliga ühendatud lause puhul malli katvuse hindamine lause piires (nt 100% või 75%); malliga vaid vähesel määral kaetud lausete küsitavaks märkimine;
- 6) malliga mitteseostatud lausete küsitavaks märkimine.

6. Vealeidja tulemused

Vealeidja töö hindamiseks kontrollisime 100 õppijalauset, mida programm oli analüüsinud süntaksianalüsaatoriga ja ühendanud olemasoleva malliga või tuvastanud vastava malli puudumise mallikogust. 100 lause seas oli 26 tuumlause sõnajärje veaga lauset. Neist 16 puhul tuvastas programm, et lause ei vasta mitte ühelegi sõnajärje mallile. 8 puhul ühendas programm lause olemasoleva malliga ekslikult. 100 lausest 26 puhul viis programm korrektse lause kokku õige sõnajärje malliga ja 36 puhul tuvastas õigesti vastava malli puudumise mallikogus. Lausete võrdlemisel mallikoguga tegi programm 21 juhul vea õppijavea tõttu, 4 korral eelkõige süntaksianalüsaatori eksliku analüüsi tõttu ning 10 korral keerukamate grammatikanähtuste esinemisel, millega ta toime ei tulnud. 3 lauset jäid analüüsist välja, kuna need ei kuulunud vealeidja vaatlusaluste lausetüüpide alla, kuigi vealeidjale on ette antud ka reeglid eemaldamiseks analüüsist mitte-vaatlusalused laused: näiteks need, mis sisaldavad impersonaali ja algavad võrdlusega.

Parseri ekslikke analüüse ja seeläbi ka vealeidja valetõlgendusi põhjustasid järgmised õppija vead:

- ortograafiaviga (***Koguaeg** ma olin maganud ainult kolm või neli korda olin ärkanud; Muidugi ka Eesti rahvas mängis isesesvumisprotsessis **surt** osa*);
- ortograafiaviga koos objektikäände veaga (***Stokholmis** olen näinud väga huvilise **asi***);
- ajavormi viga (*Mul **olnud** vähe aega, et hästi uurida seda kohta; Puhkuse lõpuks mina **saanud** aru üks asi, et ma olen linna elanik ja eluga külas kohanematu*);
- ekslik infiniitivivormi valik (*Meie ajal tehnikaareng annab palju võimalusi **reisima***);
- objektikäände viga (*Edasi ma olen näinud surepärane **linn**; Oma **matk** ma alustasin veekogust*);
- ühildumisviga (*Seal mulle **meeldis lövid**, sest ma olen näinud nende esimene kord elus*);
- reksiooniviga (ekslik adverbiaali kasutamine objekti asemel: *Veel ma meenutan kultuursetest **inimestest***);
- osalausepiiri märkimise viga (*Õues jalutasid koduloomad mina esimene kord silmasin lehm ei ole televisorist*);
- osalausepiiri märkimise viga ja predikaadi ekslik puudumine (*Estlased väga toredad inimesed mina tutvustan ja suhtlen nendega hea meelega*).

Parser suutis hoolimata vahelejäetud komast osalause piiri mitmel korral tuvastada, kui õppija oli seal kasutanud sidendit (*Sellepärast ma ei saa midagi öelda mida ma olen näinud esmaspäeval*).

Lisaks põhjustasid süntaksianalüsaatori ekslikke analüüse subjektita kogejalause (*Kõigepealt mulle meeldis selles linnas*) ja tundmatu lühend.

7. Vabad ja seotud adverbiaalid

Vealeidja saab parseri abil õppijakeelega sageli hästi hakkama, leides lausest viiteid ka tundmatute või ebakorreksete kohtade analüüsiks. Keerukamad probleemid õigete tuumargumentide olemasolu ja nende järjestuse tuvastamisel tekkisid nii õppijakeele kui kirjakeele tekste analüüsides ennekõike kahel juhul: komplekssemate rinnastusjuhtudega (nt *varased rongile tõttajad ja hilja peale jäänud pidulised*) ning arvukate vabade adverbiaalidega lausete korral. Nimetatud adverbiaale on raske automaatselt ära tunda ning nad segavad lausete kokkuvüümist õigete mallidega.

Kõige teravamaks probleemiks oligi seotud adverbiaalide eristamine vabadest. Kui seotud laiend on nii sisult kui ka vormilt põhjast tingitud, siis vaba laiendi tingitus põhjast on minimaalne: laiend võib esineda täistähenduslikult väga erinevate põhjadega. Püüdsime mallide koostamisel vabu adverbiaale vältida, sest nende arv on potentsiaalselt lõpmatu. Verbilaiendite seotuse üle otsustamine võib lisaks sõltuda ka subjektiivsest keeletajust. Seotud laiendite äratundmiseks on vaja teada eri verbide reksioone, mis võiks lisaks aidata hinnata ka sõnajärje õigsust: Huno Rätsepa (1978: 218–221) järgi on vähemalt paljude kontekstiväliste lausete puhul kehtiv, et sõnajärje tüübi valiku määrab verb (vrd *lendama* ja *leiduma*: *Lind lendab*; *Leidus vabatahtlikke*).

Kuna oluline osa kohustuslikest verbilaienditest kajastub Raili Pooli sõnastikus “Eesti keele verbireksioone” (objektid ja seotud adverbiaalid) ning eesti keele verbikesksete püsiühendite andmebaasis (ühend- ja väljendverbide osad)¹⁰, võiks adverbiaalide seotuse üle otsustamisel olla nende allikate kasutamisest abi. Katse käigus läbivaadatud 50 õppijalausest esines 108 adverbiaali, millest 70 olid vabad adverbiaalid või üldlaiendid (neist 27 lausealgulised ning seetõttu üldjuhul mallidesse kuuluvad)¹¹, 8 abimäärsõnad (7 neist on olemas ka eesti keele verbikesksete püsiühendite andmebaasis), 25 seotud adverbiaalid (17 neist on olemas Pooli sõnastikus), 2 adverbiaalatribuudid ning 3 relatiivlause alguse sidendid.

Kirjeldame lähemalt Pooli “Eesti keele verbireksioone” (1999) kasutamise võimalusi adverbiaalide automaatanalüüsil. Sõnastik esitab reksioonid, sh objektiikäänded 563 verbile ning võimaldab lihtsat tekstilause kohustuslike elementide tuvastamist. Verbireksioonide sõnastiku kasutamise väärtuseks on suure hulga vabade adverbiaalide analüüsist kõrvaldamine, mida muude formaalsete tunnuste abil ei saa teha. Kuigi vabade laiendite eraldamine seotutest polnud ilmselt raamatu autori eesmärk, toetab see spetsiaalsete vahendite puudumisel robustset tuvastust. Järgnevalt kirjeldame näidet, kuidas sõnastikku “Eesti keele verbireksioone” on võimalik selleks kasutada.

Korrektset moodustatud lause sisaldab malli järgi lause alguse adverbiaali või adverbiaale (kui on), üldjuhul subjekti, verbi, sõnastiku verbikirjes toodud verbilaiendit ning malli kõiki muid vajalikke elemente, mis on leitud viisil, nagu käesolevas artik-

¹⁰ <http://www.cl.ut.ee/ressursid/pysiyhendid/index.php?lang=et> (2.01.2010).

¹¹ Parseri märgenduspõhimõtetest lähtuvalt on loetud adverbiaalide hulka ka üldlaiendid.

lis kirjeldatud. Muude elementide hulka ei tohiks kuuluda adverbiaalid, sest liigsete vabade adverbiaalide eemaldamine on üks sõnastiku kasutamise eesmärkidest.

Näiteks on sõnastikus verbi *ahvatlema* juures kirjed:

objekt + millele? (Poiss ahvatles sõbra vargusele)
objekt + *ma*-infinitiiv (Reklaam ahvatleb inimesi ostma)

Lähtudes sõnastiku kirjetest, peab vealeidja verbi kohustuslikeks laienditeks ükskõik mis käandes objekti ning allatiivis määrust (nt *vargusele*) või *ma*-infinitiivi (nt *ostma*). Sõnastiku kasutamine võimaldab aga lausest *S ahvatles L-i parki koristama šampanjaga*¹² eemaldada adverbiaali *šampanjaga* kui vaba laiendi, sest komitatiivset verbilaiendit verbikirjes pole. Samuti on näiteks õppijakeele lause *Ma tahan alustada, et igal suvel ma armastan kuhugi reisida* (EVKK) puhul võimalik tänu sõnastikule tuvastada lause tuumelemendi (komitatiivne määrus verbi *alustama* juures) puudumise viga.

Sõnajärje vealeidja peab arvestama ka võimalusega, et lause alguses võib esineda vabu adverbiaale, sest ka need mõjutavad lause tuumelementide sõnajärge. Näiteks: ***Sinisinise veikleva veepiiri taga ahvatles vanderselle valge liivaranna ja männimetsaga väike püsielaniketa Pedassaar.***

Kui ühendada programmiga Pooli sõnastik, valmistaksid endiselt raskusi sõnastikust puuduvad seotud laiendid. Muud probleemid, mida sõnastiku kasutamine lahendada ei pruugi, on järgmised.

- Lause algul on rohkem kui üks vaba adverbiaal. Vahel on selline sõnajärg õige (***Ühe meremiili kaugusel sinisinise veikleva veepiiri taga ahvatles vanderselle valge liivaranna ja männimetsaga väike püsielaniketa Pedassaar***), vahel aga mõjub ebaloomulikuna (***Vabal ajal Narvas ma käin discol ..*** (EVKK)).
- Valik ühe verbi kirjete vahel nõuab aeg-ajalt semantilist infot, mis on potentsiaalsete vealeidja vigade allikas (*arvestama + kellega? mida? – elusolendit tähistav laiend ei saa üldjuhul olla partitiivis*).
- Mitmesugused spetsiifilisemad konstruktsioonid, näiteks verbiga *olema* (*Oli **kuldilus** suvise pööripäeva järgne päev* – atribuut on kohustuslik, kuid sõnastik selliseid konstruktsioone ei kirjelda).

8. Kokkuvõte

Keeleõppija suutlikkus moodustada õigesti sihtkeele tuumlauset on vajalik kogu lause terviktähenduse edukaks edasiandmiseks ja sellistegi vigade vältimiseks, mida teevad kõige kõrgemal tasemel teise keele kõnelejad. Artiklis on kirjeldatud eesti keele kui teise keele õppija raskusi lause tuumelementide valikul, vormistamisel ja järjestamisel, antud ülevaade esimestest sammudest õppijakeele sõnajärje arvutianalüüsil, tutvustatud esimest sõnajärje vealeidja prototüüpi ja analüüsitud tegureid, mis selle tööd mõjutavad.

Vealeidja prototüübi väljatöötamisel loodi Tartu Ülikooli kirjakeelekorpuse põhjal sõnajärje mallide kogu, kasutades selleks süntaksianalüsaatorit ja Exceli keskkonnas loodud vahendeid. Tutvustatud vealeidja eesmärk on tuvastada õppijalause tuumargumendid, nende süntaktiline roll ja järjestuse õigsus. Vealeidja

annab õppijalauseste sõnajärje korrektsusele hinnangu, lähtudes peamiselt nende morfosüntaktiliste analüüside vastavusest mallidele.

Kirjeldataud katses sisestati vealeidjasse 100 EVKK lauset ning kontrolliti programmi väljundi õigsust. 62 korral tegi programm õige otsuse: 26 korral viis vealeidja lause kokku õige malliga ja tuvastas lause korrektsuse, 36 korral tuvastas vealeidja, et lause ei vasta mallile. Neist 36 lausest sisaldasid 16 õppijapoolset kõrvalekallet loomulikust sõnajärjest.

Lahendamist vajavad märkimata osalause piiride, keerukamate rinnastusjuhtumite ning seotud adverbiaalide tuvastamise küsimused. Seotud adverbiaalide eristamiseks vabadest adverbiaalidest pakkusime võimaliku lahendusena Pooli sõnastiku “Eesti keele verbireksioone” ja eesti verbikesksete püsiühendite andmebaasi kasutamist. Edaspidi oleks õppija sõnajärje arvutianalüüsil kasu näiteks lähtumisest kõrvalekaldumistest R Emmeli sõnajärjereeglitest ning kirjakeele lause liikmete positsioonielistustest. Et lause sõnajärge efektiivsemalt kontrollida, saab programmi lisada tüüpiliste vigade mooduli. Sagedasemate vigade tuvastamine enne lausete kontrollimist mallikogust tõstab programmi töö täpsust ja kiirust. See on tõenäoliselt vähem töömahukas lahendus, kui oleks võtta kasutusele pragmaatiline info, mis on tegelikult eesti keele sõnajärje peamiseks mõjutajaks.

Lühendid ja sümbolid¹³

@ADVL	adverbiaal	<u> </u> A <u> </u>	adjektiiv
@AN>	adjektiiv eestäiendina	<u> </u> D <u> </u>	adverb
@FCV	mitmeosalise predikaadi	<u> </u> P <u> </u>	pronoomen
	finiitne osa (<i>olema</i> liitaegades jms)	<u> </u> S <u> </u>	substantiiv
		<u> </u> V <u> </u>	verb
@FMV	finiitne predikaat	CLB	osalause piir
@IMV	infiniitne predikaat	V	finiitverb
@OBJ	objekt	V2,	
@PRD	predikatiiv	V2-reegel	<i>verb teisel kohal</i> -reegel
@SUBJ, S	subjekt	X	verbi laiend, lause kõrvalliige,
@<NN	nimisõna järeltäiendina		mida võib olla üks või enam
@<Q	kvantori järellaiend		

Viidatud kirjandus

- Alam, Md. Jahangir; UzZaman, Naushad; Khan, Mumit 2006. N-gram based statistical grammar checker for Bangla and English. – Proceedings of 9th International Conference on Computer and Information Technology (ICCIT 2006), Dhaka, Bangladesh. http://www.pan11on.net/english/outputs/Working%20Papers/Bangladesh/Microsoft%20Word%20-%2019_N_112.pdf (28.02.2010).
- Arppe, Antti 2000. Developing a grammar checker for Swedish. – Torbjorn Nordgard (Ed.). Proceedings from the 12th Nordiske datalingvistikkdager, Trondheim, December 9-10, 1999. Department of Linguistics, Norwegian University of Science and Technology (NTNU). Trondheim: University of Trondheim, 13–27. <http://www.ling.helsinki.fi/~aarppe/Publications/Nodalida-99.pdf> (02.01.2010).
- Athanaselis, Theologos; Bakamidis, Stelios; Dologlou, Ioannis 2006. A fast algorithm for words reordering based on language mode. – 16th International Conference, Athens,

¹³ Loendis on toodud vaid artiklis sagedamini kasutatud märgendid. Täielik kitsenduste grammatika märgendite ülevaade asub veebiaadressidel <http://math.ut.ee/~kailli/parser/demo/morftags.html> ja <http://math.ut.ee/~kailli/papers/syntax.html> (12.02.2010).

- Greece, September 10-14, 2006. Proceedings, Part II, 943–951. <http://www.springerlink.com/content/q646768285871122/fulltext.pdf> (02.01.2010).
- EKG II = Ereht, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1993. Eesti keele grammatika II. Süntaks. Lisa: Kiri. Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut. Tallinn.
- Huumo, Tuomas 1994. Kontrastiivinen tutkimus suomen ja viron sanajärjestyksestä. Litsensiaaditöö. Käsikiri Turu Ülikooli soome ja üldkeeleteaduse õppetoolis.
- Höglund, Kirsi 2006. Estnisk satsföljd och meningsstruktur. Doktoriväitekiiri. Uppsala.
- Lindström, Liina 2005. Finitiverbi asend lauses. Dissertationes philologiae estonicae Universitatis Tartuensis, 16. Tartu: Tartu Ülikooli kirjastus.
- Metslang, Helle; Krall, Ingrid; Pajusalu, Renate; Saarso, Kristi; Sõrmus, Elle; Vare, Silvi 2003. Keelehärm: eesti keele probleemseid piirkondi. Tallinn: Tallinna Pedagoogikaülikool.
- Pool, Raili 1999. Eesti keele verbirektsioone. Tartu: Tartu Ülikooli kirjastus.
- Rommel, Nikolai 1963. Sõnajärjestus eesti lauses. – Eesti keele süntaksi küsimusi. KKI uuri-
mused, VIII. Tallinn: Eesti Riiklik Kirjastus, 216–381.
- Rätsep, Huno 1978. Eesti keele lihtlausete tüübid. Emakeele Seltsi toimetised, 12. Tallinn: Valgus.
- Tael, Kaja 1988. Sõnajärjemallid eesti keeles (võrrelduna soome keelega). Preprint KKI-56. Tallinn.

Võrgumaterjalid

- Eesti keele verbikesksete püsiühendite andmebaas. <http://www.cl.ut.ee/ressursid/pysiyhendid/index.php?lang=et> (2.01.2010).
- EstCG Parser 1.0a. <http://www.cs.ut.ee/~kaili/parser/>, <http://lepo.it.da.ut.ee/~kaili/grammatika/> (19.02.2010).
- EstCG Parseri morfoloogilised märgendid. <http://math.ut.ee/~kaili/parser/demo/morftags.html> (12.02.2010).
- EstCG Parseri süntaktilised märgendid. <http://math.ut.ee/~kaili/papers/syntax.html> (12.02.2010).
- EVKK = Eesti vahekeele korpus. evkk.tlu.ee (20.10.2009).
- Keeleveeb. www.keeleveeb.ee (10.02.2010).
- kfNgram. <http://kwicfinder.com/kfNgram/kfNgramHelp.html> (11.02.2010).
- Süntaktiliselt analüüsitud ja ühestatud tekstikorpus. <http://www.cl.ut.ee/korpused/syntaksikorpus/> (10.02.2010).
- Tartu Ülikooli Arvutilingvistika Uurimisrühma koduleht. www.cl.ut.ee (10.02.2010).
- TÜKK = Tartu Ülikooli kirjakeele korpus. <http://www.cl.ut.ee/korpused/baaskorpus/> www.cl.ut.ee (20.10.2009).

Helena Metslangi (Tallinna Ülikool, Tartu Ülikool) teaduslikeks huvialadeks on süntaks, eriti lauseliikmete vahelised üleminekuvalad, ja õppijakeele korpused.
helena.metslang@gmail.com

Erika Matsaki (Tallinna Ülikool) uurimisvaldkonnad on loogilised konstruktsioonid eestikeelsetes tekstides, keeletehnoloogia.
erika.matsak@tlu.ee

AUTOMATIC WORD ORDER ANALYSIS OF ESTONIAN AS A SECOND LANGUAGE: THE NUCLEAR SENTENCE

Helena Metslang, Erika Matsak

Tallinn University, University of Tartu

This article gives an overview of our work on the automatic analysis of second language word order. For this purpose, an error analyzer and a set of correct word order patterns found from the fiction sub-corpus of Tartu University's Corpus of Written Estonian were created. It is important to be able to form the nuclear sentence of the target language well (incl. subject, finite verb, obligatory modifiers of the verb and other elements influencing the sentence word order) because a well-formed core clause conveys the integral meaning of the whole sentence and helps to avoid the errors that even the very high level learners make.

The article describes the learner's difficulties in choosing, inflecting and ordering the core elements of the sentence (in the data of EVKK – Estonian Interlanguage Corpus). It gives an overview of the first steps of the automatic analysis of learner language word order, introduces the set of correct word order patterns and the prototype of the word order error analyzer and analyzes the factors influencing the success of its performance.

The task of the error analyzer is to detect the core arguments, estimate their syntactic role and assess the correctness of their order. In the test described in the article 100 sentences were analyzed with the error analyzer and the output was assessed. The program made the correct choice 62 times of which 26 times the error analyzer connected the correct sentence with the right rule and assessed that the word order in the clause was correct. In the other 36 times the analyzer found that the input clause didn't correspond to any rule in the pre-defined set. 16 clauses from these 36 contained a word order error.

There are a number of problems that still need to be solved including the erringly unmarked clausal border, more complex cases of coordination and imprecise adverbial analysis. The article suggests the use of the valency dictionary of Estonian verbs (Pool 1999) and the Database of Estonian verbal multi-word expressions as a possible solution of how to improve the program's ability to distinguish free and bound adverbials. In the future it would be useful to integrate Remmel's (1963) word order laws and the typical position preferences of parts of sentences into the system. To check the learners' word order more effectively it is possible to add the module of typical errors to the programme. Identifying the more frequent errors of the clause, before searching within the set of correct patterns, would raise the precision and speed of the programme. This would likely be less time consuming than supplementing the error analyzer with pragmatic information (which in fact is one of the most important factors influencing Estonian word order).

Keywords: word order, corpus linguistics, second language acquisition, Estonian

MÕNEST VÄRVINIMETUSEST JA NENDE TÄHENDUSVAHEKORDADEST EESTI JA SOOME KEELES

Vilja Oja, Mari Uusküla

Ülevaade. Artiklis analüüsitakse lilla-roosa värvigamma nimetusi eesti ja soome suulises keeles. Käsitleva murdeaines kajastab peamiselt 20. saj 1. poole keelt ning tänapäevaste nimetustena arvestatakse 1990-ndatel ja 2005–2006. a läbi viidud küsitluste tulemusi. Lillat ja roosat värvi on eesti ja soome keeles tähistatud põhivärvi (sinist või punast) märkiva sõnaga ja nendest moodustatud ühenditega, samuti võrdlustest lähtuvate liitnimetustega. Nende kõrval esineb rida indoeuroopa tüvedega laentermineid, nagu *lilla*, *violett*, *maliina*, *aniliin*, *purpur*, *kretliini*, *roosa* ja *pinkki*. Keelejuhtide ütluste põhjal on määratletud värvigamma osa või konkreetne toon, mida mingi nimetusega väljendatakse. Selgus, et indoeuroopa keelest laenatud terminite kasutus erineb ühelt poolt laenuandja keele omast ning teiselt poolt keeliti/murdeti läänemeresoome alal. Sõnavara võrdlus näitas ühtlasi viimase saja aasta jooksul nende värvimõistete nimetamises toimunud muutusi. Muuhulgas ilmnes selgelt tendents laennimetuste järjest suuremale eelistusele.*

Võtmesõnad: värvinimetused, roosa, lilla, kõnekeel, murdekeel, eesti keel, soome keel

1. Sissejuhatuseks

Kas kanarbikuõied on roosad või lillad? Mis värvi on nõmmeliivatee? Kuidas tõlkida *sirelililla* inglise keelde? Missugust värvust väljendab *roosa* eesti ja missugust soome keeles? Need ja paljud sarnased küsimused tekivad sageli tõlketöös ja keeleõppes, samuti neil, kel vaja värve või värvilisi objekte (nt lilled, tekstiilesemad) kirjeldada. See ajendas käesoleva kirjutise autoreid lähemalt uurima lilla-roosa värvigamma nimetuste vahekordi kahes suurimas läänemeresoome keeles.

* Uurimistöö on valminud Eesti Keele Instituudi baasfinantseerimise ning Eesti Teadusfondi grantide nr 6744 ja 7717 osalisel toel.

Lähtematerjaliks on ühelt poolt tänapäeva suuline eesti ja soome keel ning teiselt poolt varasem, 20. sajandil kogutud murdeaines mõlemast keelest. Tänapäeva eesti keele värvinimetuste osas on aluseks võetud Urmas Sutropi 1990-ndatel aastatel Tallinnas ja Tartus läbi viidud küsitluste tulemused ning soome värvinimetuste osas 2005 ja 2006. a Helsingis, Espoos, Tuusulas, Turus, ja Lempääläs Mari Uusküla küsitlustele saadud vastused. Mõlemad küsitlajad töötasid Ian Daviese ja Greville Corbett' välimeetodil (1995: 27–28), eesmärgiks oli välja selgitada vastava keele põhivärvinimed (vt Sutrop 1995: 798–799, 2002: 58–59, Uusküla 2008: 30–31). Esmalt paluti keelejuhtidel vabalt loetleda nii palju värvinimetusi, kui neile meenus, seejärel näidati ükshaaval 65 värvitahvli ning lasti nimetada iga konkreetset värvi. See meetod võimaldab küllalt täpselt määrata värvinimetuste sisu ning eri keelte nimetusi omavahel võrrelda.

Suur osa käsitlevast eesti murrete materjalist pärineb Tallinnas Eesti Keele Instituudis koostatud eesti murdesõnavara koondkartoteegist ning soome murdeaines Helsingis Kotimaisten kielten tutkimuskeskuses säilitatavast soome murdesõnavara arhiivist. Nendes murdekogudes talletatud sõnavara kajastab valdavalt 20. sajandi 1. poole suulist keelesituatsiooni (hiljem kogutud materjalide puhul on keelejuhtideks valitud eranditult vana murde kõnelejad). Murdekogudest saadud värvinimetuste semantika määratlemisel on aluseks võetud ühelt poolt sõna kasutus näitelauses ja -fraasides ning keelejuhi selgitused, teiselt poolt koguja kommentaarid. Murdenäited on artiklis edasi antud kirjakeele vahenditega, lisatud on üksnes palatalisatsioonimärk konsonandi kohal või kõrval (nt *ń*, *l'*).

Lilla ja roosa värvi eesti- ja soomekeelsed nimetused võib laias laastus liigitada kahte rühma: omasõnavara ja laenatud terminid. Järgnevalt käsitletakse kummagi rühma sõnavara eraldi, kuid pööratakse erilist tähelepanu nimetuste omavahelistele suhetele ning nende tähendusnüansside sarnasustele ja erinevustele.

2. Omasõnavara

Omasõnavaraks on siin arvatud lilla ja roosa nimetused, millel vastavalt eesti või soome keeles on olemas muu varasem otstarve, nagu põhivärve (sinist ja punast) märkivad sõnad ja nende baasil moodustatud ühendid (vt 2.1). Siia kuuluvad ka nn kirjeldavad ühendid, mis vahendavad teavet värvi kohta võrdluse teel ehk võrdlusel baseeruvad värvinimetused (vt 2.2).

2.1. Põhivärvidest lähtuvad nimetused

Eesti ja soome keeles on tavaline väljendada kahe värvi segatoone põhivärvide nimedest moodustatud liitsõnaga. Lillat märgivad kahe värvinimetuse kombinatsioonid, nagu eesti *sinakaspunane*, (Vastseliina) *sinineverrev* jt, soome *sini-punainen*, murdeis *sinisempunane*, *sinipunerva*, *punasesinertävä*, *punasininen*, *liusunpunainen* (*liusu* 'sinine') jms. Soome kirjakeeles on tänini lilla värvi nimena enim soositud liitsõna *sinipunainen* (< *sininen* + *punainen*) ja roosa värvi nimena *vaaleanpunainen* (< *vaalea* 'hele' + *punainen*) (vt nt Kerttula 2007: 160, Uusküla 2007: 383 jj). Viimasega sarnaneb lõunaeesti murdeline *val'ssverrev* < *val'ss*

'hele' + *verrev* 'punane' (vt Oja 2001: 66–67). Soome liitsõna *vaaleanpunainen* on sedavõrd üldine roosa värvi nimetus, et sellest on omakorda moodustatud ühendid *tummanvaaleanpunainen* 'tumeroosa' ja *vaaleanvaaleanpunainen* 'helerosa' (vt Uusküla 2009: 270). Värvinime *sinipunainen* kasutus kõnekeeles on sel sajandil tunduvalt vähenenud ning lillat värvi on hakatud tähistama laenterminitega.

Sageli on lillat või roosat värvi nimetatud vaid ühe – emma-kumma põhivärvi järgi, näiteks lillakasroosa kartulisordi kohta on eestlased öelnud nii *sinine* kui *punane* (lõunaeesti keeles ka *verev*). Murdeaines näitab, et lillat seostatakse pigem sinise kui punase värviga, nt eesti (Muhu) *siniõllud* 'jaanililled' o *sinised*; (Käina) *viiolett on sinine*; (Karuse) *sinised kaapsad* 'punased kapsad' *olid mõisates*; (Leivu) *šinnin eirelääts* 'lilla hiirehernes', *šinnin upinain* 'punane ristik'; (Torma) *sinine ja lilla on ühe näu pial*; soome (Ylikiiminki, Paavola) *liusi* = *sinipunainen*. Roosat värvi väljendatakse ka punase värvi moderatiivse nimetusega, nagu soome *punertava* 'punakas', lõunaeesti *verkjass* id.

2.2. Võrdlusel baseeruvad värvinimetused

Eelmistest täpsema värvitooni määratluse annavad liitnimetused, mille nimisõnaline täiendosa tähistab üldtuntud värvusega võrdlusalust. Sellise ühendi põhisõna on kas värvinimetus või pelgalt parameetersõna, mis osutab, et tegemist on värviga.

Värvinimetus on põhisõnaks sellistes igapäevastes lillat väljendavais liitsõnades nagu eesti *peedipunane*, *kannikesesinine*, soome *orvokinsininen* (< *orvokki* 'kannike' + *sininen*) jt. Roosa värvi nimetustest kuuluvad siia näiteks eesti *vaarika-punane* ja soome *ruusunpunainen*, sõnasõnalises tõlkes 'roosi (= kibuvitsa)' + 'punane'. Põhisõnana esinevad loomulikult ka lillat ja roosat tähistavad laenulised värvinimetused (nt *sirelililla*), kuid neid käsitletakse allpool vastava põhisõna juures.

Parameetersõna ühendeist on ehk tuntuim eesti *sirelikarva*, soome *sireeninvärinen* (*sireeni* 'sirel' + *värinen* 'värvi'). Sama konstruktsiooniga on Setu murraku roosa värvi nimetus *marañaolinõ* (sõnasõnalt 'marja välimusega' ehk 'marja värvi'), kusjuures ühendit *muštjanõ marañaolinõ* (*muštjanõ* 'tume') on setud kasutanud tähenduses 'lilla'. Mõnikord on murdekeeles värvust iseloomustatud ka võrdluse alusel nimisõnast moodustatud omadussõnatuletisega, nt (Kodavere) *kirsikäs* 'roosa'.

3. Laenatud terminid

Lisaks eelkirjeldatud sõnadele väljendatakse mõlemas kõnealus keeles lillat ning roosat indoeuroopa keelest laenatud terminitega. Eesti keeles on niisugused laenimetused *lilla*, *violet*, *maliina*, *aniliin*, *purpur* ja *roosa*, soome keeles *liila*, *violetti*, *kretliini*, *aniliini*, *purppura*, *roosa* ja *pinkki*. Ehkki mitu samatüvelist sõna esineb nii eesti kui ka soome keeles, ei kasutata neid alati keeliti/murdeti ühtmoodi.

3.1. Lilla

Eesti keelele kõige omasem lilla värvi nimetus on saksa laensõna *lilla* (< sks *lila* 'lilla') (Mägiste 1982–1983: 1312). Nii tänapäeva keeles kui ka eelmise sajandi murdekõnes on selle sõnaga väljendatud igasuguste heledus-tumedus- ja kirkusastmetega ning eri toonides sinise ja punase segavärvust. Sõnaga *lilla* (Muhu murrakus *lille*) kirjeldatakse mitmesuguste taimede õite värvust (nt kanarbik, jaanilill, priimula, ohakas, naistenõges, hiirehernes, kurereha, sirel, palderjan, kellukad), aga ka palju muud (nt riie, lõng, kaaliumpermanganaat, taevas, lepaseened, kartulid, marjad). Eesti keeles märgib sõna *lilla* niivõrd laia värviskaalat, et erinevate toonide ning heledus-tumedusastmete täpsustamiseks kasutatakse tihti liitsõnu, nt *tumelilla*, *helelilla*, *kahvatulilla*, *erelilla*, *sügavlilla*, *erklilla*, *punakaslilla*, *sinakaslilla*, *mustjashilla*, *valkjashilla*; (Iisaku rannikumurdes) *vaalakaslilla*; (Juuru) *kaks lillat värvi: teene punalilla, teene sinililla värv*; (Karksi) *vanast õigati värmi lilla sinine, lilla punane*.

Soome keeles päris sama kujuga omadussõna pole. Nimisõna *lilla* märgib soome murdeis küpsset soomurakat ning sellest võrdluse alusel moodustatud liitsõna *lillapunane* on kasutatud lillakaspunase värvi nimetusena, nt (Kivennapa) *lillapunane = punavioletti, sinipunanen; siniseen vivahtava punanen kutsuttiin lillapunaseksi*.

Tänapäeva soome keelde on rootsi keelest laenatud adjektiiv *lila* või *liila* (< rts *lila* 'lilla'), mis väljendab tavaliselt sirelikarva helelillat värvust. Samas tähenduses esineb liitsõna *liilanvärinen*. Ehkki eesti *lilla* ja soome *li(i)la* on laenatud eri keelte kaudu, pärinevad need samast algtüvest: << pr *lilas* (< *lilac*) 'sirel' < hisp *lilac* < ar *lilāc* id. < pärsia *lilac* ~ *nīlak* 'sinakas' < *nīl* 'sinine, indigo' (SSA 2: 73). Tänapäeva soome kõnekeeles on eriti nooremal põlvkonnal levinud *liila* enam kui *violetti*, kuigi esimese värv on heledam, teisel tumedam. Samas kasutatakse kahte laentermini ka sünonüümselt. Eesti *lilla* sisuliseks vasteks on soome keeles tavaliselt kahest värvisõnast koosnev liitnimetus *sinipunainen*, paljudes murretes *kretliini* vms. Seega ei ole eesti *lilla* ja soome *li(i)la* sünonüümid. Kui soome sõna ingliskeelne vaste on *lilac*, siis eesti keelde võiks selle tõlkida liitsõnadega *helelilla* või *sirelililla* ning üldnimetuse *lilla* ingliskeelseks vasteks sobib kõige paremini *purple* (vt Oja 2002: 256–258).

3.2. Violet

Romaani algupäraga lillenimest tekkinud värvisõna (vrd pr *violette* 'kannike; lilla' < ld *viola* 'kannike; lilla') märgib punase ja sinise segavärvi nii eesti kui ka soome keeles. Eesti keelde on *violett* laenatud saksa keele kaudu (< sks subst *Violett*; adj *violett*), soomekeelne *violetti* on saadud rootsi keele vahendusel (< rts *violett* 'violetne') (Mägiste 1982–1983: 3870, SSA 3: 454). Murdekeeles on värvinimetuse levila üsna ulatuslik. Laensõnale omaselt esineb mõlemas keeles rohkesti erinevaid häälikulisi variante. Eestlased on võõrapärase nimetuse rahvakeeles sageli asendanud tuttava sõnaga *viilul* ~ *viiluldi* või mõne muu mugandatud variandiga (vt nt Oja 2001: 29, 2007: 205). Kirjakeelset sõnakuju *violet(t)* on kuulnud vaid paaril korral uuemas murdekeeles. Soome murdesõnades varieerub alguskonsonant: *v-/f-* ja 1. silbi diftongi järelois, labiaalvokaal: *-io-/-iu-/-iü-*.

Eesti murretes esineb sõna peamiselt substantiivsena – lilla värvaine ja roosa kartulisordi nimetusena (nt *viuli värv*; *viuliti kardula*, *viuletid*) ning adjektiivsena iseloomustab kõige sagedamini nendesamade objektide värvust. Murdekõnelejate jutus on selgesti tajutav asjaolu, et algselt on eesti keelde laenatud substantiiv – Lääne-Euroopast toodud värvaine nimetus. Puutaolise vormi järgi nimetati värvaine tükki: *viulipuu*, *viuldi puu*, *violiti värvi puu* vms (Oja 2007: 204). Hiljem on lisandunud adjektiivne funktsioon – sama sõnaga on hakatud nimetama ka selle aine värvust ja sellega värvimisel saadud punase-sinise vahelist värvitooni. Näiteks Rannu murraku keelejuht väidab, et lõnga sai violetiks värvida üksnes vabrikuvärviga: *viuli sinine ol'i poodi värm, ega seräst mõtsa värmega es saa tetä*.

Soome murretes on adjektiivset kasutust rohkem. Mõlemas keeles tarvitatakse ka *ne/nen*-sufiksiga adjektiive ning liitnimetusi, mille põhisõnaks sinist või punast värvi märkiv adjektiiv: nt eesti *violetne*, *violetisinine*, *viulipunane* jt, soome *violetinen*, *violettisininen*, *violetin sininen*, *violettipunainen* jt. Tänapäeva soome keeles on *violetti* kujunenud peamiseks lilla värvi nimetuseks ning suhteliselt uus laen *lila* või *lila* on selle hüponüüm (Kerttula 2007: 159–160). Eesti keeles on *violett* ning *violetne* praeguseks samuti saanud üldkasutatavaiks värvisõnadeks. Ometi näitab Sutropi kogutud materjal, et sõna *violett* on kõnekeeles oluliselt harvem kasutatav kui *lilla*, kuigi esineb *lilla* sünonüümina (2002: 74 jj).

3.3. Kretliini

Soome kõnekeeles ja paljudes murretes märgib lillat värvust rootsi laen *kretliini* < rts *gredelin*, *gredlin* < pr *gris de line* 'linakarva hall' (Hellquist 1922: 199, SSA 1: 418; vt ka Koski 1983: 230–231). Värvinimetus *retelijini* esines juba 18. saj soome keeles (vt Ganander 1997: 785). 20. saj murdekeeles oli sõna laialt levinud, vähem näiteid on siiski registreeritud Kesk-Soome murdest ja Põhja-Soomest ning ühtki teadet pole ingeri murdeist. Häälikulisi murdevariante on laensõnale omaselt rohkesti: *krätliini*, *kreteliini*, *reliini*, *retlinki*, *krekliini*, *prekliini* jt, lisaks adjektiivisufiksiga tuletised (*k*)*retliininen* ja *retuliininen*. Laensõna esineb nii ühendisparameetersõnaga (nt *retliininkarvaine*, *retliinivärinen*) kui ka punase ning sinise värvi nimetustega (nt *kretliinisinen*, *retliinipunainen*). Sagedamini kirjeldatakse nõnda värvitud esemeid, peamiselt lõnga või riidet, aga ka muude objektide, näiteks kanarbiku, kartuli, hortensia ja varesemarja õite lillat või lillakasroosat värvust. Sünonüümina on enim mainitud liitnimetust *sinipunainen*, kuid ka sõnu *punasinen*, *violetti*, *violetin punainen* jm. Samas on laensõna eelistatud isegi omapärasele liitnimetusele, käsitades viimast vaid kirjeldava seletusena, kusjuures tõeline värvinimi on *kretliini*, nt (Laitila) *sinipunase färine on kretliini*. Mõned keelejuhid on 1930-ndail aastail väitnud, et sõna on tema kodumurdes vananenud. Käesoleval sajandil läbi viidud küsitlus näitas, et mõne vanema keelejuhi idiolekti kuulub endiselt sõna *retliini*, *retuliini* vms, mis märgib tumedamat või sinakaslillat värvust. Olulisi piirkondlikke erinevusi sõna kasutajaskonnas ei täheldatud (Uusküla 2009: 270). Tänapäeva soome keele sõnaraamatuid leiame märksõnad *gredliini* (kõnekeeles *kretliini*) (NS 1: 308) ja *kretliini* märgendiga "vananenud" (PS 1: 555). Soome murrete sõnaraamatu tegijad on otsustanud sõnavariandid koondada märksõna *retliini* alla.

3.4. Liina

Eesti läänemurde alal on kellukate, kanarbiku jt lillade ning roosade õite värvi kirjeldatud adjektiiviga *liina* või *liine*. Kahtlemata on tegemist lühenenud sõnakujudega, ent pole päris selge, millisest laensõnast need on tulenenud. Mauno Koski oletus, et *liina* ja *liine* võiksid olla lühendused rootsi lilla värvi nimetusest *gredlin* (1983: 235) tundub ebausutavana, kuna rootsi sõna äratuntavaid variante pole eesti murretes registreeritud.

Mari Must on need sõnad ühitanud vene laenuga *maliina* (murdeis ka *maliine*, *maleena*) < *vn малина* 'vaarikas; vaarikapunane' (2000: 191). Ei ole aga võimalik, et *liina* ja *liine* on mugandused hoopis sõnast *aniliin*. Eesti keelde värvaine nimena laenatud võõrsõnal on rahvakeeles palju variante, nagu *aneliin*, *aaniliin*, *aanilin(e)*, *anililla* jt. Soome 20. saj murdekeeles märkis *aniliini* (*aneli*, *aliini* jms variandid) samuti eelkõige värvainet, harvem värvust. Adjektiivi funktsiooni täitis sageli liitnimetus, nagu eesti (Kanepi) *aanilin verev*, (Urvaste) *anilillisinine*, soome *aneliinipunainen*, *annelipunanen*, *anoliinivärine* jms.

Aniliinivärviga värviti lõnga lillaka varjundiga erkroosaks või lillakaspunaseks. Seetõttu on mõistetav, miks eesti murdekõnelejail kipuvad võõrad sõnad *aniliin* ja *maliina* omavahel segunema, nt (Kadrina) *maliina on ikke tema õige nimi, aneliiniks aga üiti*. Mõlemad laenud esinevad ka 19. saj eesti murdekeelt kajastavas Wiedemanni sõnaraamatus (1973: 34 ja 580). Samasugust värvi väljendavad soome adjektiivid *aneliinipunainen*, eesti *anililla*, *liina*, *maliina* jt variandid, nt soome (Hailuoto) *Anellia oli paha muuttumaan. Punasen sinistä väriltä, kulkukauppiaat möijät väriä* 'Aniliin ei vahetanud kergelt värvi. Värvilt lilla, rändkaupmehed müüsid värvi'. Uusküla 2005–2006. a välitöödel selgus, et värvisõna *aniliini* ja ühend *aniliinipunainen* või *aniliininpunainen* väljendavad erkroosat ning lillakaspunast värvi nii noortel kui ka vanematel Soome eri paigust pärit inimestel (nimetust mainisid 17–52-aastased inimesed).

3.5. Purpur

Eesti *purpur*, *purpurne*, *purpurpunane* ja soome *purppura*, *purppuranpunainen* pärinevad tõenäoliselt Piiblist. Sama nimetust kandis tööstuslikult toodetud värvaine. Võimalik, et Piibli tõlgetesse tuli värvinimetus otse ladina keelest (< *ld purpura*), kuid eesti keelde võib sõna olla laenatud ka saksa või alamsaksa keele kaudu, vrd sks *Purpur*, asks *purpur*, ja soome keelde rootsi keele vahendusel, vrd rts *purpur* (Mägiste 1982–1983: 2241, SSA 2: 437).

Eesti ja soome rahvakeeles teatakse võõrsõna enamasti tähendustes 'punane' ja 'erepunane'. Mõned keelejuhid on kinnitanud, et *purpur* / *purppuranpunainen* on 'tumepunane' ja üksikud isikud on väitnud, et '(tume) lillakaspunane' (Kerttula 2007: 160, Oja 2001: 162–163). Käesoleval sajandil Soomes läbi viidud küsitlusel kasutasid värvinime *purppura* samuti vaid üksikud keelejuhid tähenduses 'punakaslilla', liitsõna *purppuranpunainen* viitas aga tavaliselt tumepunasele, punasele või lillakaspunasele värvitoonile. Üks katseisik kasutas ka liitsõnalist kombinatsiooni *purppuravioletti* tähenduses 'punakaslilla'. Nende adjektiividega on kirjeldatud näiteks punetavat taevast ja pilvi, loojuva päikese ja tulekahju kuma, vere värvi, lilleõite, lõnga ja riide värvi.

“Eesti kirjakeele seletussõnaraamatus” (EKSS 4: 484) on värvinimetuse tähenduseks märgitud ‘lillakaspunane’, soome keele sõnaraamatuis ‘lilla või sinise helgiga punane värv’ (NS 4: 496, PS 2: 553). Mõnikord on ingliskeelse värvisõna *purple* eestikeelseks vasteks pakutud ‘purpurne’. Sellisest tõlkest on soovitatav hoiduda, kuna see ei anna edasi õiget värvi tooni. Parim eestikeelne tähendusvaste inglise omadussõnale *purple* on *lilla* (vt 2.1.).

3.6. Roosa

Eesti kirjakeeles ja kõigis murretes on roosa värvi üldnimetus sakslastelt laenatud omadussõna *roosa* (< sks *rosa*) ning rohkesti leidub ka sellise põhisõnaga liitvärvinimetusi, nagu *helerosa*, *halerosa*, (Iisaku rannikumurdes) *vaalakas roosa*, *tumerosa*, *vanarosa*, *erkroosa*, *lillakasroosa* jt. Murdekeeles on kasutatud ka adjektiivisufiksiga tuletisi *roosane* või *roosaline*. Wiedemanni sõnaraamat (1973: 979) näitab, et 19. saj keskaigu oli see värvisõna eesti keeles kodunenud.

Soome keeles on *roosa* (< rts *rosa* ‘roosa’) sedavõrd uus, et 20. saj keskel ilmunud “Nykysoomen sanakirja” (NS) seda veel ei tunnista, kuid sajandi lõpul ilmunud “Suomen kielen perussanakirja” esitab sõna koos seletusega (*vaalean ruusunpunainen* (PS 2: 665)). Mõlemas keeles esineb sõna üksnes värvinimetuseks, kuna taime (roosi ehk kibuvitsa) tähistab varasem laen (ee *roos*, sm *ruusu*). Erinevalt eesti keelest on tänapäeva soome keeles suhteliselt uus värvinimetuseks *roosa* nn moesõna, mis on paljudele tuttavaks saanud Soomes suurt populaarsust pälvinud kataloogikaubandusest (Koski 1983: 241–244, Uusküla 2009: 268–269). Aastail 2005–2006 Soomes küsitletud 68-st keelejuhist kasutas sõna *roosa* ligi 11%, tähistades sellega erinevaid värvi toone (beež, kahvatu-roosa, lõheroosa, hallikas-roosa).

Eesti keeles väljendatakse sõnaga *roosa* ning selle tuletistega valge ja punase segavärvust, aga ka lillakasroosat ja kahvatulillat värvust. Nii *roosa* kui *lillana* on kirjeldatud näiteks kanarbiku, jaanilille ja sireli õisi ning punaka koorega kartuleid. Võrreldes omavahel kahes keeles sama sõnaga väljendatud värvusi, näeme, et soome *roosa* märgib kas hallikamat ja kahvatumat, või hoopis oranžikamat tooni kui eesti *roosa* ning seega ei ole eesti ja soome *roosa* täissünonüümid.

3.7. Pinkki

Tänapäeva soome keeles kasutatakse roosa värvi tähistamiseks teistki uut laensõna *pinkki* << ingl *pink* (Kerttula 2007: 160, SSA 2: 368). Sõna on soome keeles levinud kõigepealt kaubandusliku terminina, märkides värvide ja mitmesuguste esemete värvust (Koski 1983: 241–244). Nagu *roosa*, on ka *pinkki* soome keeles omaks võetud alles 20. saj lõpul ning esineb kirjakeele sõnaraamatus tähendusseletusega *vaalean neilikanpunainen*, sõna-sõnalt ‘nelgi värvi roosa’ (PS 2: 477). Väga olulisi tähenduslikke erinevusi soome värvinimetuste *pinkki* ja *roosa* vahel möödunud sajandi keeles ei täheldatud (Kornerup, Wanscher 1961: 214). Sel sajandil tehtud küsitluste põhjal võib järeldada, et olukord on muutunud. Esiteks on omadussõna *pinkki* laiemalt levinud kui *roosa*, kuuludes suurema hulga inimeste idiolekti. Teiseks väljendatakse värvisõnaga *pinkki* väga konkreetset lillakasroosat tooni

(vrd nurmelgi, *Dianthus deltoides* õite värviga). Nooremate soomlaste kõnes on see tavaline omadussõna, esinedes ka liitsõna osana, nt *tummanpinkki* ja *vaaleanpinkki* (Uusküla 2009: 269–270). Kuidas seda värvi võorkeeltes nimetada nii, et säiliks õige tähendus, peab iga tõlkija konkreetse teksti puhul otsustama.

4. Kokkuvõtteks

Asjaolu, et lilla värv sisaldab spektri kaht põhivärvi, kajastub otsesõnu liitnimetustes nagu eesti *punakassinine*, soome *sinipunainen* jms. Teiselt poolt soovivad inimesed nimetusega ka segavärvi domineerivat värvust edasi anda. Lillat ja roosat tähistavaid sõnu on püütud tähenduslikult diferentseerida ühelt poolt täiendava omadussõna abil (nt eesti *punakaslilla*, soome *vaaleanpunainen*), teiselt poolt võrdlusega (nt ee *löheroosa*, sm *sireeninvärinen*).

Eesti keeles on kõige üldisema sisu ja kasutajaskonnaga laennimetused *lilla* ja *roosa*. Praeguseks on ka *violett* ning *violetne* saanud üldkasutatavaiks värvisõnadeks ning neid võib lugeda *lilla* sünonüümideks. Kõnekeeles eelistatakse siiski nimetust *lilla*. Soome keeles on *violetti* kujunenud peamiseks lilla värvi nimetuseks ning suhteliselt uus laen *liila* või *lila* on selle hüponüüm. Nooremad kõnelejad kasutavad aga omadussõna *liila* rohkem, kusjuures see tähistab varasemast laiemat spektriala. Eesti murretes märgivad *violett* ja selle variandid pigem värvainet, sellega võrreldes on soome murretes adjektiivset kasutust rohkem. 20. sajandi alguses ütlesid paljud soomlased kõigi sinise ja punase vahele jäävate toonide kohta *kretliini* või *sinipunainen*. Tänapäeval on nende omadussõnade kasutus vähenenud, noorte kõnepruugis võivad sõnad üldse puududa. Roosa värvi nimetusena on eesti keeles üldlevinud laensõna *roosa*, soome keeles käibib seni omakeelne liitnimetus *vaaleanpunainen*, mille kõrvale on siiski tekkinud uued laensõnad *roosa* ja *pinkki*.

Võrreldes kahes keeles sama sõnaga väljendatud värvusi omavahel, näeme, et soome *roosa* märgib kas hallikamat ja kahvatumat, või hoopis oranžikamat tooni kui eesti *roosa* ning seega ei ole eesti ja soome *roosa* täissünonüümid. Ka eesti *lilla* ja soome *li(i)la* ei ole enamasti sünonüümid. Soome murretes on eesti *lilla* vasteks tavaliselt kahest värvisõnast koosnev liitnimetus *sinipunainen* või laensõna *kretliini* variandid. Uus laen *liila* märgib tänapäeva soome keeles helelillat värvust, mille kohta eestlased ütlevad *helelilla* või *sirelililla*. Selle ingliskeelne vaste on *lilac*, kui eesti *lilla* ingliskeelseks vasteks sobib kõige paremini *purple*. Üks oluline põhjus, miks samast algtüvest lähtunud laennimetusi kasutatakse eesti ja soome keeles erinevalt, on kindlasti erinev laenuallikas – eesti keelde tuli enamik neist sõnadest saksa keele kaudu, soome keelde aga rootsi keele vahendusel. Oma mõju avaldab ka laenamise aeg ehk laenu vanus.

Teine osa käsitletud laensõnu esineb mõlemas keeles enam-vähem samas tähenduses. Värvaine *aniliin* nime on murdekõnelejad mitut viisi mugandanud, sealhulgas esinevad rahvaetümoloožiast ajendatud variandid eesti *anililla* (*hani* + *lilla*), soome *anneli* jt. Mõlemas keeles tähistatakse nendega erkroosat ning lillakaspunast värvi. Laensõnu *purpur* jt samast tüvest tulenenud sõnu on enamik eestlasi ja soomlasi pidanud erepunase värvi nimetuseks. Ehkki mõlema keele sõnaraamatud annavad tähenduseks ka 'lillakaspunane', ei ole see läänemeresoomlaste keelepruugis inglise *purple* vaste.

Roosa ja lilla värvi kokkupuutepunkte värviskaalal tähistavad ühelt poolt eesti *liina*, *maliina*, *aniliin* ja soome *aniliini*, *pinkki*. Teiselt poolt märgivad mõned värvisõnad niivõrd laia värviskaalat, et sama värvi kohta on kasutatud nii roosa kui ka lilla värvi nimetusi, nt eesti *lilla* ja *roosa*. Uurimuse tulemustele toetudes on tabelis 1 välja toodud soovituslikud eestikeelsed tähendusvasted soomekeelsetele lilla ja roosa värvi nimetustele.

Tabel 1. Soome sõnade soovituslikke tähendusvasteid tänapäeva eesti keeles

Soomekeelne värvinimetus	Eestikeelne värvinimetus
violetti	lilla, violett
liila	helelilla, sirelililla
kretliini	lilla, tumelilla, sinakaslilla
sinipunainen	lilla
vaaleanpunainen	roosa
pinkki	lillakasroosa, tumeroosa, erkroosa
roosa	roosa, lõheroosa, kahvaturoosa
aniliini	punakaslilla, erkroosa
purppura	lillakaspunane, punakaslilla

Viidatud kirjandus

- Davies, Ian; Corbett, Greville 1995. A practical field method for identifying probable basic colour terms. – *Languages of the World*, 9, 25–36.
- EKSS = Eesti kirjakeele seletussõnaraamat I–VI. Eesti Keele Instituut. Tallinn: Eesti Keele Sihtasutus, 2009.
- Ganander, Chrisfrid 1997. Nyt finskt lexicon. Alkuperäiskäsikirjoituksesta [1786–1787] toimittanut Liisa Nuutinen. *Suomalaisen Kirjallisuuden Seuran toimituksia*, 676. *Kotimaisten kielten tutkimuskeskuksen julkaisuja*, 95. Helsinki: Suomalaisen Kirjallisuuden Seura, Kotimaisten kielten tutkimuskeskus.
- Hellquist, Elof 1922. *Svensk etymologisk ordbok*. Lund: Gleerups.
- Kerttula, Seija 2007. Relative basicness of color terms: Modeling and measurement. – Robert E. MacLaury, Galina V. Paramei, Don Dedrick (Eds.). *Anthropology of Color. Interdisciplinary multilevel modeling*. Amsterdam, Philadelphia: John Benjamins, 151–169.
- Kornerup, Andreas; Wanscher, Johan Henrik 1961. *Värien kirja*. Kaarle Hirvonen (toim.). Porvoo: WSOY.
- Koski, Mauno 1983. Värien nimitykset suomessa ja lähisukukielissä. *Suomalaisen Kirjallisuuden Seuran toimituksia*, 391. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Must, Mari 2000. *Vene laensõnad eesti murretes*. Tallinn: Eesti Keele Sihtasutus.
- Mägiste, Julius 1982–1983. *Estnisches etymologisches Wörterbuch*. Helsinki: Finnisch-Ugrische Gesellschaft.
- NS = *Nyckysuomen sanakirja*. Porvoo, Helsinki: Suomalaisen Kirjallisuuden Seura, Werner Söderström osakeyhtiö, 1951–1961.
- Oja, Vilja 2001. *Linguistic studies of Estonian colour terminology*. *Dissertationes philologiae estonicae Universitatis Tartuensis*, 9. Tartu: Tartu University Press.
- Oja, Vilja 2002. Some semantic problems in the translation of colour terms. – Henrik Gottlieb, Jens Erik Mogensen, Arne Zettersten (Eds.). *Symposium on Lexicography X. Proceedings of the Tenth International Symposium on Lexicography May 4–6, 2000, at the University of Copenhagen*. *Lexicographica. Series Maior*, 109. Tübingen: Niemeyer, 253–260.

- Oja, Vilja 2007. Colour naming in Estonian and cognate languages. – Robert E. MacLaury, Galina V. Paramei, Don Dedrick (Eds.). *Anthropology of Color. Interdisciplinary multi-level modeling*. Amsterdam, Philadelphia: John Benjamins, 189–209.
- PS = Suomen kielen perussanakirja. Kotimaisten kielten tutkimuskeskuksen julkaisuja, 55. Helsinki: Edita, Kotimaisten kielten tutkimuskeskus, 1990–1994.
- SSA = Suomen sanojen alkuperä. Etymologinen sanakirja. Suomalaisen Kirjallisuuden Seuran toimituksia, 556. Kotimaisten kielten tutkimuskeskuksen julkaisuja, 62. Helsinki: Suomalaisen Kirjallisuuden Seura, Kotimaisten kielten tutkimuskeskus, 1992–2000.
- Sutrop, Urmas 1995. Eesti keele põhivärvinimed. – *Keel ja Kirjandus*, 12, 797–808.
- Sutrop, Urmas 2002. The vocabulary of sense perception in Estonian: Structure and history. *Opuscula Fenno-Ugrica Gottingensia*, 8. Frankfurt am Main: Peter Lang.
- Uusküla, Mari 2007. The basic colour terms of Finnish. – *SKY Journal of Linguistics*, 20, 367–397.
- Uusküla, Mari 2008. Basic colour terms in Finno-Ugric and Slavonic languages: Myths and facts. *Dissertationes linguisticae Universitatis Tartuensis*, 9. Tartu: Tartu University Press.
- Uusküla, Mari 2009. Some remarks on colour naming differences in Finnish. – Marja Kallasmaa, Vilja Oja (toim.). *Kodukeel ja keele kodu. Home language and the home of a language*. Pühendusteos Helmi Neetarile 75. sünnipäevaks 29. jaanuaril 2009. Eesti Keele Instituudi toimetised, 13. Tallinn: Eesti Keele Sihtasutus, 265–272.
- Wiedemann, Ferdinand Johann 1973. Eesti-saksa sõnaraamat. *Estnisch-deutsches Wörterbuch*. Neljas, muutmata trükk teisest, Jakob Hurda redigeeritud väljaandest. *Vierter unveränderter Druck nach der zweiten, von J. Hurt redigierten Auflage [1893]*. Tallinn: Valgus.

Vilja Oja (Eesti Keele Instituut) peamised uurimisvaldkonnad on dialektoloogia, leksikoloogia, geolingvistika, soome-ugri keeled, läänemeresoome keelte vastastikused suhted ja keelekontaktid. vilja.oja@eki.ee

Mari Uusküla (Eesti Keele Instituut) peamised uurimisvaldkonnad on värvinimed, võrdlev leksikoloogia, soome-ugri keeled ja antropoloogiline lingvistika. mari.uuskyla@eki.ee

ON SOME COLOUR TERMS AND THEIR SEMANTIC RELATIONSHIPS IN ESTONIAN AND FINNISH

Vilja Oja, Mari Uusküla

Institute of the Estonian Language

The article discusses the terms for hues from purple to pink used in spoken Estonian and spoken Finnish. Most of the dialectal material illustrates the usage of the first half of the 20th century, while the results of the field work conducted in the 1990s and 2005–2006 are treated as indicative of the modern state of the art. In Estonian as well as in Finnish, both purple and pink have been referred to by words for the main hue (blue or red) or by their combinations, or by comparison-based compound terms.

Beside those there are a number of borrowed terms with Indo-European stems, such as *lilla*, *violett*, *maliina*, *aniliin*, *purpur*, *kretliini*, *roosa* and *pinkki*. With the help of the native speakers, the part of the spectrum area or concrete hue referred to by this or that term has now been specified. It turns out that the use of the terms borrowed from Indo-European languages differs, on the one hand, from that of the donor language and, on the other hand, across the languages/dialects spoken in the Finnic area. In addition, our lexical comparison revealed the changes in reference to the colour concepts in question over the past century. Among else, there is a clear tendency towards an increasing preference of borrowed terms.

Keywords: colour terms, pink, purple, spoken language, dialect speech, Estonian, Finnish

EESTLASTE JA VENELASTE SUHTLUS- KÄITUMINE: *SINA, TEIE* JA KEEGI VEEL

Renate Pajusalu, Virve Vihman,
Birute Klaas, Karl Pajusalu

Ülevaade. Artikli eesmärgiks on võrrelda Eestis elavate eestlaste ja venelaste suhtluskäitumist. Materjal põhineb sotsiolingvistilistel üliõpilaste küsitlustel. Eelkõige keskendutakse *sina/teie* valikule, kuid vaadeldakse ka muid suhtluse aspekte, näiteks suhtluskäitumise erinevuste teadvustamist ja kolmanda isiku nimetamisvormi valikut.

Võtmesõnad: viisakus, pöördumine, pragmaatika, eesti keel, vene keel

1. Sissejuhatus

Suhtluse kultuurilisi norme määravad mitmed parameetrid. Penelope Browni ja Stephen Levinsoni viisakusteooria (1987) raames on esitatud tuntud kõneakti ohustamise määra (*W*) valem $W(FTA^1) = R + D + P$. Selle alusel võib öelda, et kõnelejad lähtuvad oma kõneaktide sobitamisel viisakusnormidega käimasoleva kõneakti kaalust selles kultuuris (*rate*, *R*), kõnelejatevahelisest sotsiaalsest distantsist (*D*) ja võimusuhetest (*power*, *P*). Kaalu all mõeldakse siin käimasoleva kõneakti staatust vaadeldavas kultuuris, distantsi all eelkõige vanuselist erinevust ja/või tuttavust-võõrust, mõnes kultuuris võib distants sõltuda ka kõnelejate soost ja muudest sotsiaalsetest kriteeriumidest. Võimusuhe kajastab eelkõige suhtlejate vahelist erinevust võimускаalal, s.t ülemuse-alluva suhteid jms. Selle valemi üle on korduvalt diskuteeritud, nt on esile toodud, et kolme parameetrikimpu (*R*, *D* ja *P*) ei saa vaadelda lihtsalt kujutletavate väärtuste summana, vaid pigem käivitab igaüks neist oma tervikliku käitumismallistiku, mis ei sõltu teistest parameetritest (Brown, Levinson 1987, Vogelberg 2002). Kõigi kolme parameetri väärtused kujunevad vestluse käigus, s.t suhtlusolukorra ja partneri reaktsioonide põhjal otsustatakse vastastikku, kuidas tegelikult käitutakse.

¹ Mõiste *face threatening act* 'nägu ohustav akt' (lühendina *FTA*) on Browni ja Levinsoni (1987) viisakusteooria üks nurgakivisid, mille all mõeldakse igasugust kõneakti, mis võib mõjutada kõneleja või kuulaja näo (ehk minapildi) turvalisust. Viisakusvahendid ongi keeles selleks, et seda ohtu vähendada.

Käesolevas artiklis² käsitleme lähemalt teietamise fenomeni, mis väga vähesel määral (kui üldse) sõltub kõneakti sisust, kuna samad inimesed meie kultuuri-ruumis reeglina teietavad üksteist alati, sõltumata kõne all olevast teemast. Oma küsitluskava koostamisel lähtusime aga oletusest, et nii eesti kui ka vene kultuuris on kõige olulisemaks *sina/teie* valiku määräjaks sotsiaalne distantss inimeste vahel. Sotsiaalse distantssi põhilised moodustajad on vanus ja tuttavus. Vanust ja tuttavust oskavad suhtlejad näost näkku suhtluses alati hinnata, vähemalt sel määral, mis suhtlusstiili valikuks vajalik on. Sealjuures võib vanuseline parameeter olla nii suhteline (vanuseline erinevus kahe inimese vahel) kui absoluutne (lastesse suhtutakse alati teisiti kui täiskasvanutesse).

Käesolev uurimus põhineb küsitlustel, mis seab uuritavale materjalile küllaltki olulised piirangud. Kõigepealt saame me küsitlustest teada, mida inimesed arvavad end teatud olukordades tegevat, mitte seda, mida nad tegelikult teevad. Kogu viisakuse valdkond on ju argikeeles defineeritud hoopis teisiti kui keeleteaduslike viisakusteooriate raames. Kui argikõnes peetakse viisakaks formaalset käitumist (nn *first-order (im)politeness*, Watts 2003), siis keeleteaduses viisakust kui nähtust tervikuna käsitledes räägitakse igasugustest vestluspartnerite sotsiaalsest suhtest tingitud verbaalse käitumise markeeringutest (*second-order (im)politeness*, Watts 2003). Seega kui informandid midagi viisakaks või ebaviisakaks nimetavad, ei kehti samasugune hinnang automaatselt ka keeleteadusliku viisakusteooria raames.

Teine küsitavus seostub sellega, mida inimesed argielus tähele panevad. Ilmselt jääb palju rohkem meelde ebaviisakus, s.t olukord, kus kõik suhtluse parameetrid ei vasta ootustele. Richard Watts (2003: 5 jj) märgib, et ebaviisakust on palju vähem uuritud kui viisakust.

Samuti ei ole küsimustikuga võimalik modelleerida väga palju erinevaid ega ka mitte väga ootamatuid olukordi, sest küsitletavad ei jaksa ega oska väga pikalt vastamise süveneda. Siiski on selline küsitlus üks viis võrrelda sotsiaalseid grupe omavahel ja teha järeldusi, mis ulatuksid üle konkreetsete olukordade mikroanalüüsi. Oleme ka samal arvamusel kui Watts, kui ta ütleb, et viisakust ei saa vaadelda väljaspool (eba)viisakuse argikäsitlusi. Seega, see mida tavaline inimene ütleb ja arvab viisakusest, on ülimalt oluline.

Kuna viisakuskäitumist seostatakse tihti suhtlejate sooga, siis koostasime küsimustiku nii, et soolised erinevused esile tuleksid. Naiste ja meeste keelelise käitumise erinevustest on tehtud palju uurimusi, kuid kokkuvõttes näib olevat nii, nagu Sara Mills (2003: 235) on öelnud: sugu ei määra suhtluse viisi mingil üksühesel viisil vaid on ainult lisafaktor kõikide muude konteksti omaduste hulgas.

2. *Sina ja teie* eesti ja vene keeles: varasemad uurimused

Eesti keele *sina/teie*-kasutust on ka varem uuritud sotsiolingvistikulise küsitluse abil. Leelo Keevallik (1999) on küsitlenud eri vanuses ja eri kodukohaga kooliõpilasi. Selgus, et horisontaalsed suhted (s.t eelkõige tuttavusaste) on *sina/teie* valikul olulisemad kui vertikaalsed. Siinses artiklis kirjeldatud uurimus on osaliselt inspireeritud Keevalliku küsimustikust.

Eesti kultuuris on *sina/teie* valiku probleem küllaltki hästi teadvustatud, selle üle arutletakse ajakirjanduses (näiteks toimus ajalehes Postimees aastal 2008 diskussioon *sina* ja *teie* sobilikkusest praeguses ühiskonnas) ja omavahel, tõenäoliselt õpetatakse seda mingis vanuses lastele otseste instruksioonide abil (mitte kaudselt, nagu paljusid muid keelelisi parameetreid).

Vene keele viisakuse uurimused võib kokku võtta umbes nii, nagu seda teeb Watts (2003: 15): ühelt poolt on vene kultuuris viisakus selgelt seotud vulgaarse keelekasutuse vältimisega, teiselt poolt on vene keeleruumis kombeks paljusid asju ütelda palju otsesemalt kui läänemaailmas ja ühiskond väärtustab tihti just lähidistantsi-suhtlust. Mitmed uurijad rõhutavad vene kultuurile omast südamlikkust ja otsekohesust (Rathmayr 1999). Üks põhjusi, miks vene keel võib endale lubada pehmemdamata kõneakte, on see, et vajalik formaalsusaste saavutatakse teietamise abil ja muid viisakusvahendeid formaalses olukorras ei vajatagi (vt nt Vogelberg 2002).

Vene viisakuskäitumist on teoreetilisel tasandil uuritud üllatavalt vähe. Krongauz (2004: 165) tõdeb, et kui vene etiketist midagi ongi kirjutatud, siis on tegemist preskriptiivse kirjeldusega. Ilmselt tuntuim autor vene etiketi osas on Formanovskaja (nt 1989), kes on väitnud, et *sina/teie* valik toimub järgmiste parameetrite järgi: sotsiaalne staatus, lähedus (kui hästi kõnelejad tunnevad üksteist), suhte tüüp, olukorra tüüp (ametlik, mitteametlik) ning sotsiaalsete rollide sarnasus või erinevus.

Krongauz (2004: 174–175) üldistab vene traditsioonilise viisakuskoodi omapära järgmiselt. Kui tegemist on formaalse vestlusega (näiteks ostja ja müüja poes), on vene kultuuris suurem anonüümsus ja distants kui “Euroopa” omas, kui aga on tegemist “tegeliku” suhtlusega, siis väiksem; vene traditsioonilist viisakust iseloomustab eelkõige suur variatiivsus, erinevates olukordades käitutakse väga erinevalt. Viimasel aastakümnel on olukord siiski oluliselt muutunud, räägitakse nn uuest vene viisakusest (Rathmayr 2008), mis on seotud teenindussfääri “läänelikustumisega”.

Eesti eestlaste ja venelaste erinevat viisakuskäitumist on võrdlevalt uuritud mitmel eri viisil. Olga Puusepp (1999) on lähtunud Leelo Keevalliku (1999) küsitlusest eesti kooliõpilaste seas ja viinud läbi sarnase küsitluse Eesti eri piirkondade vene koolides. Suurim erinevus eestlaste ja venelaste vahel on selle uurimuse põhjal see, et *teie* on venelase jaoks alati austav ja positiivne, samas kui eestlase jaoks võib see olla ka tõrjumise vahend. Maria Gonorova on eesti ja vene koolide õpetajaid intervjuuerides jõudnud tulemuseni, et kuigi mõlema õppekeele koolides on selge tendents suurema sinatamise poole, on vene koolide õpetajad tunduvalt rohkem selle vastu ja leiavad, et “ei tohi levivat sinatamist koolis lubada, sest siis võib kord käest minna” (Gonorova 2007: 52). Kõik, kes eesti-vene küsitlusi läbi on viinud, tõdevad, et venelased teietavad rohkem kui eestlased (Puusepp 1999, Borovkova 2002, Gonorova 2004).

Üks olulisi erinevusi *sina/teie* kasutamises eesti ja vene kultuuri vahel on see, et eesti kultuuris üldjuhul ei kasutata ebasümmeetrilist süsteemi, v.a siis, kui täiskasvanu räägib lapsega (Keevallik 2004: 205). Roger Brown ja Albert Gilman (1960: 257–261) on juba 1960. aastatel märkinud, et võimusuhted muutuvad Euroopa keelte pöördumistes üha vähem oluliseks. Sellega kaasneb ka ebasümmeetriliste kasutuste kadumine. Vene kultuuris on ebasümmeetria aga võimalik ka täiskasvanute puhul, näiteks kui kõnelejate vahel on suur vanusevahe (Rathmayr 1999).

Teine oluline erinevus, mida eesti ja vene kultuuri vahel on täheldatud, on see, et eesti keeles on *sina/teie* kasutus püsiv, s.t kui ükskord on sinatama hakatud, siis sinatatakse kogu elu (Keevallik 2004: 206). Vene keeles aga sõltub *sina/teie* kasutus rohkem olukorrast.

Eestis on tegemist mitmekeelse suhtluskeskkonnaga, kuigi üldine seisukoht on see, et eestlased ja venelased elavad suhteliselt eraldatuna. Vanemate inimeste teietamine on vene kultuuris väga selgelt seotud seisukohaga, et haritud inimene teietab endast vanemat. Seetõttu keelduvad vene emakeelega inimesed mõnikord võtmast omaks eesti ühiskonna *sina/teie* malle ka siis, kui nad tegelikult teavad hästi, kuidas eestlased sinatavad ja teietavad. Nähtust, kus inimene teadlikult keeldub valimast teise keele kultuuriruumi kuuluvat suhtlustiili, on tähele pandud ka teiste kultuuride puhul (Kasper, Blum-Kulka 1993: 18).

Käesoleva artikli üheks algtoukeks oli Olga Borovkova (2002) küsitluse tulemusel püstitatud hüpotees, et vähemuses elavate rühmade *sina/teie* valiku seaduspärasused erinevad enamuses elavate samast rahvusest rühmade valikust: Tartu venelased ja Kohtla-Järve eestlased käitusid küsitluses sarnaselt, mis Borovkova hinnangul on tingitud sellest, et mõlemad elavad vähemuses.

3. Materjal

Eesti ja vene viisakuskäitumise uurimiseks koostasime sama küsimustiku nii eesti kui vene keeles (vt lisa 1). Üldeesmärk oli esitada nii selliseid küsimusi, mille vastused oleksid analüüsitavad statistiliselt, kui ka selliseid küsimusi, millele vastataks vabalt. Viimaste eesmärk oli selgitada välja üldisi tendentse, ja kuigi neid on tunduvalt raskem arvuliselt võrrelda, annavad nad siiski olulist informatsiooni suhtluse stereotüüpide kohta.

Kokku analüüsisime 159 küsimustikku, nende hulgas oli 62 eestlaste ja 97 venelaste vastatut. Tartu ülikoolist saime 97 täidetud küsimustikku (45 eestlast ja 32 venelast). Kohtla-Järvel asuvast Virumaa kolledžist saime kokku 55 täidetud küsimustikku, neist 17 eestlaste ja 38 venelaste oma. Saamaks võrdlusmaterjali ka Venemaal elavate venelaste käest palusime küsimustikele vastata ka 27 Peterburi ülikooli üliõpilasel. Vastanud eestlastest oli mehi 31 ja naisi samuti 31, Eesti venelaste hulgas oli naisi 44 ja mehi 26. Peterburi võrdlusrühm koosnes ainult neidudest. Küsimustikele vastati auditooriumides, seega oli aktualiseeritud eelkõige õppeolukord.

4. *Sina ja teie* pöördumises erinevate inimeste poole

Küsimustiku alguses paluti vastata küsimusele: *Kas Te teietate allpool loetletud inimesi? Tõmmake ring ümber kas numbrile 1, 2 või 3, kusjuures 1 – alati, 2 – mõnikord, 3 – mitte kunagi.*

Inimesed, kelle suhtes tuli otsustada, jagunesid seitsmesse rühma. Küsitluses endas neid rühmi küll ei eristatud, kuid nende märgatavust polnud ka välditud, sest ühe grupi esindajad järgnesid üksteisele. Allpool käsitleme vastuseid esimesele küsimusele rühmade kaupa. Oleme arvutanud vastuste keskmise väärtuse (vt tabel 1).

1. Sõprade (*sõber* ja *sõbranna*) osas oletasime, et vastajad käsitlevad seda kategooriat kui omaealist ja tuttavat (sotsiaalselt distantsilt eriti lähedast) inimest. Kõikides rühmades oli valdav sinatamine ja keskmine jäi 2.7 ja 3.0 vahele. Eristus ainult üks grupp, milles keskmine väärtus oli 2.0, need olid Tartus elavad vene mehed. Nähtust on raske seletada, ainuke oletus on see, et noored vene mehed kasutavad mõnikord sõpradega suheldes humoristlikku teietamist, mille korral teadlikult rikutakse tavalisi vestlusnorme ja mis tegelikult osutab suurt lähedust suhtlejate vahel.

2. Generatsiooni võrra vanemaid, kuid suhteliselt lähedalt eakaaslase kaudu seotud inimesi (*sõprade vanemad, oma poisi/tüdrukute vanemad, vanemate nais- ja meestuttav*) teietavad venelased kõikides rühmades. Selle põhjuseks on ilmselt vene kultuuris oluline vanuseline distants. Eestlastel on kõikides informandirühmades suuremad keskmised (eestlastel vahemikus 1.4–2.0, venelastel 1.0–1.2), seega eestlased sinatavad selle rühma inimesi venelastest rohkem, kuid siiski vähem kui sõpru. Eestlaste hulgas on nõrk tendents selle poole, et naissoost informantidel on pisut kõrgemad keskmised, mis võib olla tingitud tüdrukute tihedamast suhtlusest vanemate inimestega (näiteks sõprade vanematega).

3. Sugulasi (küsitluses *onu* ja *tädi*) eestlased sinatavad (keskmine selles rühmas on 2.3–2.9), vene informandid aga mõnikord sinatavad, mõnikord teietavad (keskmine vahemikus 1.5–2.1). Onu ja tädi, kes vanuseliselt kuuluvad eelmise grupi inimestega samasse generatsiooni, on tüüpiliselt lähedasemad, kuna kuuluvad perekonda ja kuna neid tuntakse tõenäoliselt kauem, kui eelmise rühma inimesi. Erinevus muudest vanema generatsiooni inimestest (s.t 2. grupist) on aga näha vaid eestlaste puhul, venelastel on keskmised suures plaanis samad ja erinevust ei paista olevat. Siin on näha selge kultuuriline erinevus: kui vene kultuuris on oluline vanusevahe ja seetõttu teietatakse ka vanemaid sugulasi, siis eesti kultuuris sinatatakse sugulasi ka siis, kui nad on vanemad. Nõrga tendentsina on näha Peterburis elavate venelaste suuremat *sina*-kasutust kui Eesti venelastel.

4. Neljandasse rühma koonduvad võõrad inimesed ühistranspordis (*võõras vana/keskealine/noor naine bussis, võõras vana/keskealine/noor mees bussis*). Selle rühma puhul oletasime, et vastajatel meenub suhteliselt tüüpiline olukord, mida kõik on kogenud – lühike vestlus juhusliku tuttavaga ühistranspordis või üksik kõneakt (näiteks palve komposteerida pilet või küsimus selle kohta, kas inimene väljub või mitte). Võõrast, endast vanemat inimest teietavad bussis meie küsitluse andmetel nii eestlased kui ka venelased. Kui aga tegemist on samaealise noorega, on tulemused eri rühmades küllaltki erinevad, kuigi keskmised jäävad siiski pigem teietamise poolele – 1.4 ja 2.1 vahele. Tartu eesti naiste vastused olid kõige lähemal sinatamisele (keskmine 2.1). Nii Tartu kui Kohtla-Järve venelaste hulgas on näha tendents selles suunas, et naised teietavad sellises olukorras rohkem (keskmine 1.5), mehed aga vähem (keskmine 1.7–2.2). Peterburi venelaste rühm koosnes ainult tüdrukutest ja sarnaneb oodatult vene neiudega Eestis.

5. Viienda rühmana käsitlesime võõrast (nais- või mees-)üliõpilast auditooriumis, kes on eeldatavasti tuttavusskaalal sama võõras kui tundmatu bussis, kuid kes kuulub vastajatega samasse sotsiaalsesse rühma, mistõttu vastustes oleks võinud olla kollegiaalset solidaarsust. Kõikides rühmades olid keskmised võrreldes busiolukorraga mõnevõrra sinatamise suunas. Venelased vastasid siiski nii, et võime oletada vene üliõpilaste vahel suuremat solidaarsust kui eesti üliõpilaste vahel: vene gruppide keskmised olid vahemikus 2.1–2.6, eesti gruppidel aga 1.7–2.3. Nii Tartu kui Kohtla-Järve meestel olid keskmised pisut kõrgemad (2.4–2.6) kui naistel

(2.1–2.3). Peterburi rühm sarnanes keskmiste poolest (1.9–2.0) aga pigem eestlaste kui Eesti venelastega. Võime seega oletada, et Eestis elavad vene noored sinatavad samaealist kolleegi sagedamini kui eestlased ja kui Peterburi venelased.

6. Eraldi grupina käsitleme teenindajaid, esitatud suhtluspartnerite nimekirjas on müüja ja ettekandja tüüpiliselt naised, politseinik mees, meesametniku puhul oli sugu üheselt määratud. Selles rühmas oli teietamine valdav nii vene kui eesti kultuuris (keskmised 1.0–1.3).

7. Viimasena küsisime seda, kuidas pöördutakse naabrite poole. Naabrid on küllaltki hägus kategooria, sõltuvalt elukeskkonnast varieerub suhe naabritega palju. Venelased teietasid naabreid kõikides rühmades rohkem kui eestlased, eriti suur oli erinevus venelaste ja eestlaste vahel Kohtla-Järvel, kus venelaste keskmised olid 1.1–1.4, eestlaste omad aga 2.2–2.3. Tartu venelased sinatasid pisut rohkem kui Kohtla-Järve eakaaslased ja neil ei olnud Tartu eestlastega olulist erinevust. Peterburi venelased vastasid siin samamoodi nagu Kohtla-Järve venelased.

Tabelis 1 on näha keskmised tulemused, mille puhul suurem number näitab suhteliselt suuremat sinatamist.

Tabel 1. Vastuste keskmised väärtused küsimustiku küsimusele *Kas te teietate nimetatud inimesi?*
1 – teietan alati; 2 – teietan mõnikord; 3 – ei teieta kunagi

	TEM	TEN	TVM	TVN	KEM	KEN	KVM	KVN	PVN
kokku vastanuid	21	24	17	15	10	7	9	29	27
sõber	2.8	3.0	2.0	2.7	2.9	3.0	2.8	2.8	2.7
sõbranna	2.8	3.0	2.0	2.9	3.0	3.0	2.8	2.8	2.7
sõprade vanemad	1.6	1.8	1.0	1.0	1.4	1.6	1.2	1.0	1.0
oma poisi/tüdruku vanemad	1.8	2.0	1.1	1.0	1.7	1.7	1.0	1.0	1.1
vanemate naistuttav	1.6	2.0	1.2	1.0	1.5	2.0	1.1	1.0	1.2
vanemate meestuttav	1.6	1.8	1.2	1.0	1.6	1.9	1.1	1.0	1.2
onu	2.7	2.9	1.9	1.9	2.3	2.7	2.1	1.6	2.1
tädi	2.7	2.9	1.9	1.8	2.3	2.7	1.9	1.5	2.1
võõras vana mees bussis	1.1	1.0	1.0	1.0	1.2	1.1	1.0	1.0	1
võõras vana naine bussis	1.1	1.0	1.0	1.0	1.1	1.1	1.0	1.0	1
võõras keskealine mees bussis	1.1	1.1	1.0	1.0	1.3	1.1	1.0	1.0	1
võõras keskealine naine bussis	1.1	1.1	1.0	1.0	1.2	1.1	1.0	1.0	1
võõras neiu bussis	1.4	2.1	1.8	1.5	1.5	1.6	1.7	1.5	1.4
võõras noormees bussis	1.5	2.1	2.2	1.5	1.8	1.6	2.1	1.5	1.4
võõras meesüliõp. loenguruumis	2.0	1.9	2.6	2.1	1.7	2.3	2.6	2.3	2.0
võõras naisüliõp. loenguruumis	2.1	1.9	2.4	2.1	1.7	2.3	2.4	2.3	1.9
müüja	1.1	1.2	1.2	1.0	1.2	1.1	1.2	1.0	1.0
ettekandja	1.1	1.2	1.2	1.0	1.3	1.3	1.1	1.0	1.0
politseinik	1.1	1.1	1.0	1.0	1.1	1.0	1.0	1.0	1
panga meesametnik	1.1	1.2	1.0	1.0	1.1	1.1	1.0	1.0	1
naabrimees	2.1	1.9	1.9	1.6	2.2	2.3	1.3	1.4	1.4
naabrinaine	2.1	1.9	1.9	1.7	2.2	2.3	1.1	1.4	1.4

T – Tartu ülikool, K – Kohtla-Järve kolledž, P – Peterburi ülikool
V – vene rühm, E – eesti rühm, N – naine, M – mees

5. Vabad vastused

Küsitluse küsimused 2–8 olid esitatud samuti kirjalikult, küsitluslehel oli iga küsimuse järel ruumi umbes 4–5 rida. Vastused olid suhteliselt lühikesed, mis näitab, et eesti ja vene suhtluserinevused ei ole teadvustatud probleem, millele inimesed oskaks kohe detailselt vastata. Vastuste lühidus võimaldas aga neid hõlpsamini grupeerida ja nii oli võimalik teha üldistusi, mis kirjeldavad küsitletavate rühma tüüpilisi suhtumisi.

5.1. Õppejõu ja üliõpilase vastastikune *sina/teie*

Kaks vaba vastusega küsimust (2 ja 3) puudutasid õppejõu ja üliõpilase vahelist suhtlust, mis oli ühtviisi tuttav kõigile vastanuile. Muidugi ei olnud ka siin tege mist absoluutselt sarnaste suhetega: kolme kõrgkooli (Tartu ülikool, Kohtla-Järve kolledž, Peterburi ülikool) sisekliima ja kombed võivad olla erinevad mitte niivõrd kõneldava keele viisakusnormide tõttu kui muudel põhjustel.

Teine küsimus koosnes kahest poolest: *Kas teietate õppejõude?* ja *Kas oleks tore sinatada?* Selgus, et kõik vastanud, nii eestlased kui ka venelased, üldreeglina teietavad õppejõudu. Siiski leidus igas Eesti rühmas mõni üliõpilane, kes vastas *enamasti*, seega vähemalt üksikute jaoks on teatud olukorras võimalik õppejõudu ka sinatada. Selles mõttes vastasid pisut erinevalt Peterburi vene üliõpilased: mitte keegi neist ei maininudki võimalust, et õppejõudu võiks sinatada.

Küsimuse teisele poolele (*Kas oleks tore sinatada (õppejõudu)?*) oli rohkem erinevaid vastuseid, kuid eestlased siiski valdavalt õppejõudu sinatada ei soovi. Need eestlased, kes pooldasid teietamist, mainisid põhjusena eelkõige suhtlusstiili: teietamine on korrektsem, akadeemilisem, mitte liiga familiaarne jms. Vanus tuleb kriteeriumina esile ainult nendel vastajatel, kes on öelnud, et nooremaid (meeldivaid, sõbralikke) õppejõude võiks ka sinatada. Eestlased tervikuna näivad seega olevat üliõpilase-õppejõu suhtlusmudelisse seotud suhtlusolukorraga, akadeemilise sfääriga, mitte niivõrd kahe isiku vahelise sotsiaalse distantsiga või võimusuhtega. Eestlaste arvamused on kõikuvad ja neil ei ole selliseid kindlaid kategooriaid nagu *lugupidamine* Eesti venelastel (vt allpool). Üllatuslikult on eestlaste vastused sarnased just Peterburi venelaste omadega.

Venelased nii Tartus kui Kohtla-Järvel vastasid küsimuse teisele poolele absoluutselt eitavalt, seega ei ole neil mingit soovi õppejõudu sinatada. Sealjuures nimetatakse põhjusena eelkõige vanusevahet (18 korda) ja lugupidamise (*uvaženie*) osutamise vajalikkust (25 korda). Mõnevõrra on ka üldisemaid vastuseid (*teie on korrektsem, kasvatus ei luba sinatada, sinatamine oleks ebakultuurne* jms). Paaris vastuses on mainitud, et sinatades oleks küll kergem suhelda, aga siiski oleks see “vale”. Eesti venelastega võrreldes oli Peterburi informantidel üldine suhtumine küll sama, kuid põhjendused olid sõnastatud teisiti. Kui Eesti venelastel oli lugupidamine (*uvaženie*) just selle sõnaga mainitud väga tihti, siis Peterburi vastustes esines see sõna ainult üks kord. Peterburi üliõpilased soovivad õppejõude teietada pigem see tõttu, et *teie* vastab õppejõu staatusele või et see kuulub õppeprotsessi juurde.

Seega on Eesti venelaste käitumisteadvuses üliõpilase ja õppejõu suhtes kaks olulist kategooriat: vanuseline distants ja lugupidamine, samas kui eestlastel selget

teietamise põhjust kas ei ole üldse või on (akadeemilises) suhtlusvaldkonnas. Vene *uvaženie* 'lugupidamine' on oma esinemissageduse tõttu vastustes ilmselt spetsiifiline kultuurikategooria. Osaliselt on see kindlasti seotud vanuselise distantsiga, sest vanemast inimesest tuleb lugu pidada. Kuid vanus ei ole ainuke lugupidamise põhjus; mida aga lugupidamiseks veel vaja on, see vastustest otseselt ei selgu. Paaril korral on küll nimetatud staatust. Võime oletada, et lugupidamine on tervikkategooria, mida vastajad ei pea tarvilikuks lahti seletada.

Kolmas küsimus oli *Kas õppejõud teietavad teid?* Eestlased vastasid, et enamasti teietavad, kuid mõnikord ka sinatavad. Nii Eesti kui Peterburi venelased vastasid enamasti, et teietavad, kuid siin oli pisut rohkem kui eelmise küsimuse puhul vastuseid *nii ja naa* või *enamasti* või *sõltuvalt õppejõust*. Erinevad olid eestlaste ja venelaste vastused küsimuse teisele poolele: *Kas teile meeldiks, kui nad sinataksid?* Eestlased aktsepteerisid õppejõu-poolset sinatamist suhteliselt sageli, paljud leidsid, et sinatamine loob vabama õhkkonna. Eestlaste vastustest ei olnud näha, et sinatamine võiks olla solvav.

Venelaste rühmades oli erinevaid vastuseid, mõnele sinatamine meeldiks, kuna see looks sõbralikumad ja sundimatamad suhted, teistele ei meeldiks. Need, kellele õppejõu-poolne sinatamine ei meeldiks, on põhjustena eelkõige välja toonud selle, et *me oleme ikkagi täiskasvanud*, samuti on mainitud ka lugupidamist, seega ei ole lugupidamine üheselt seotud vanema inimesega, vaid seda peaks tundma ka õppejõud üliõpilase vastu. Väike erinevus tuli välja ühelt poolt Tartu ja Peterburi ja teiselt poolt Kohtla-Järve vene naiste gruppide vahel: esimesed lubaksid sinatamist rohkem kui teised. Seega näeme Kohtla-Järve vene naiste hulgas rohkemat orienteerumist traditsioonilisele suhtlusmallile, Tartus ja Peterburis see mall aga mõnevõrra kõigub. Võimalik, et Tartu venelaste puhul on see põhjustatud sellest, et nad on rohkem mõjutatud eestlaste suhtlusmallidest, ka nende õppejõud on ju üldiselt eestlased. Peterburis aga võib olla suure linna fenomen, kus käitumistraditsioonid ei ole nii üheselt aktsepteeritud. Mehed olid kõikides rühmades selles küsimuses suhteliselt ükskõiksemad, nende vastused on lühemad ja erinevalt vene neidude rühmast ei väljendanud nad otsesest solvumist sinatamise puhul (kuigi ei tahtnud seda aktsepteerida).

5.2. Sina/teie üldisemaid aspekte: muutus pöördumises ja kõhkklus valikul

Nii vene kui ka eesti kultuuris on põhimõtteliselt võimalik siirdumine teietamiselt sinatamisele. Seetõttu küsisime: *Kuidas minnakse suhtluses üle teietamiselt sinatamisele? Kas selleks on olemas reeglid ja kas te ise järgite neid?*

Mõistest "reegel" on küsitletavad aru saanud erinevalt. Ühed ütlevad, et reegleid pole ja esitavad seejärel kirjelduse reeglitest, teised ütlevad, et reeglid on, aga ei selgita neid. Vaatleme lähemalt neid vastuseid, kus tõepoolest seletati muutust pöördumistes. Eelkõige nimetati järgmisi suhtluse pöördumist muutvaid omadusi:

- suhtluse pikaajalisus: mida kauem suheldakse, seda suurem tõenäosus on minna üle *sinale* – vastus oli tüüpiline eestlastele (13 vastajat), kuid esines ka venelastel (3 vastajat), üks vene naine kirjutas, et *kui on sama vana ja planeerin pikaajalist suhtlust, siis hakkab kohe sinatama*, seega ilmnis,

et inimesed mõnikord (omavanusega suheldes) lähtuvad juba ette sellest, et suhtlus võib kujuneda pikaajaliseks;

- suhtluse sagedus ja harjumuslikkus: mida sagedamini suheldakse, seda tõenäolisemalt minnakse *sinale* üle – nii vastas üks eestlane ja kolm venelast;
- suhtlejatevaheline lähedus: mida lähedasemaks suhtlus muutub, seda suurem on tõenäosus minna üle *sinale* – nii arvas 11 eestlast ja 22 venelast, seega paistab see vastus olevat tavalisem just venelaste hulgas (tõenäoliselt siiski ainult omavanustega suheldes, kuidas muidu seletada sugulaste teietamist);
- sundimatus (kergus, vabadus, ladusus): mida ladusamalt suhtlus õnnestub, seda tõenäolisem on üleminek *sinale* – nii arvasid kolm eestlast; üks venelane nimetas veel eraldi *pärast barjääri ületamist*;
- koostöö määr: mida rohkem on ühiseid teemasid või eesmärke, seda suurem on tõenäosus minna üle *sinale* – nii arvasid kaks eestlast ja kaks venelast, lisaks ütles üks eesti mees, et *sina* hakatakse kasutama *positiivse tooniga vaidluse* jooksul, mida võib ka pidada koostöö ilminguks.

Üks rühm seletusi pöördumisviisi muutmise kohta seisnes selles, et nimetati mingi kõneaktitüüp, mis tuleks sooritada, et vahetada pöördumisvormi. Enamasti kirjutati, et üleminekuks tehakse ettepanek (7 eestlast ja 7 venelast), selgitamata sealjuures, kes ettepaneku teeb. Mõnel juhul oli siiski mainitud ka seda, kumb suhtlejatest ettepaneku teeb: üks eestlane ütles, et ettepaneku teeb vanem, üks venelane, et vanem ja austatum; kaks eestlast kirjutasid, et ettepaneku teeb see, kellel on ebamugav. Üks eestlane ja kaks venelast nimetasid samuti ettepanekut, kuid täpsustasid, et ettepaneku teeb partner. Kuna tegemist oli noorte vastajatega, võib oletada, et prototüüpne situatsioon on tõesti selline, et ettepaneku teeb vanem suhtluspartner. 15 eestlast ja 5 venelast vastasid, et *teielt sinale* üleminekus lepiketakse kokku, täpsustamata, kuidas just. Kohtla-Järve vene rühmas oli kolm naist, kes ütlesid, et nad *kunagi ei lähe üle sinale täiskasvanutega*.

Ainult kaks inimest nimetasid spetsiifilist rituaali, millega *teielt sinale* üle minnakse, üks venelane nimetas sinasõpruse joomist (*pit' na bruderšaft*), üks eestlane nimepidi tutvustamist, mis oletatavasti tähendab tutvustamisel eesnime nimetamist. Üks eestlane esitas ka väite, et *omavanuseid esimesel korral teietan, teisel juba sinatan*. Eestlaste vastustes oli ka selliseid, kus muutust küll aktsepteeriti, s.t oldi nõus sellega, et mingil hetkel pöördumisvorm muutub, kuid ei nähtud selles mingeid seaduspärasusi ja vastati *kuidas juhtub* või *omavanustega läheb sujuvalt*.

Vaid üksikutes küsitlustes tuli ilmsiks vene viisakuskoodile omaseks peetud (Rathmayr 1999) *sina* ja *teie* situatsiooniline vahetumine, nimelt et vabal ajal võib öelda *sina*, kuid tööl peab ütleva *teie*.

Mõnikord oli nimetatud inimese omadusi, kellega suhtlemisel *sinale* üle minnakse. Nimetati, et meeldivat inimest sinatatakse, ebameeldivat või distantseeritud teietatakse.

Nagu eespool öeldud, ei olnud kõikide venekeelsete informantide vastustest võimalik välja lugeda seda, et nad oleksid silmas pidanud mingit muutust suhetes. Kõige rohkem (11) oli selles rühmas viiteid vanusele, umbes sellises formuleeringus:

kui on sama vana, siis sinatan. Sealhulgas oli üks täpsem kommentaar: *kui on vähemalt 10 aastat vanem, siis kindlasti teietan*. Nimetati ka võrdsust (*kui on võrdne, sinatan*) ja mitteformaalseid suhteid. Vanusevahe on aga püsiv suurus, nii et seda tüüpi vastusest ei ole võimalik välja lugeda seda, kuidas olukorda muudetakse. Samuti on raskesti interpreteeritav vastusetüüp *vastastikuse nõusoleku korral või kui on kergem öelda sina*, kuna siit ei selgu, kas nõusolek peaks olema kohe suhtluse alguses ja kui ei, siis kuidas see tekib.

Kokkuvõtteks võib öelda, et selgeid käitumisnorme pöördumisvormi vahetamiseks ei ole ei eestlastel ega venelastel. Väljapoolt siinset küsitlust on teada, et vanemate inimeste hulgas on levinud arusaam, et *sinale* üleminekut pakub vanem nooremale ja naine mehele. Selle reegli esimene pool kajastub ka meie küsitluse vastustes, sest mõnedes vastustes on see otseselt näha, mõnedes kaudselt ja vähemalt pole kordagi öeldud vastupidist. Sooliste reeglite kohta pole aga kogu küsitluste hulgas ühtegi märki. Näib nii, et kuigi nii eestlased kui venelased nimetavad ülemineku põhjuseks suhtluse kestust ja lähedust, esines esimene kriteerium eestlastel ja venelastel enam-vähem võrdsel määral (13 ja 11), venelastel oli aga tunduvalt rohkem nimetatud suhtluse lähedust (3 ja 22).

Teine üldisem küsimus *sina/teie* kasutuse kohta puudutas seda, kas alati on teada, kumba vormi sobib kasutada. Sõnastatud oli küsimus nii: *Kas mõnikord on juhtunud, et Te ei tea, kas sinatada või teietada? Mida Te siis teete?* Eesti kultuuris on selline olukord tavaline ja ka vastanute hulgas oli praktiliselt kõigil eestlastel kogemusi sellega, et nad ei tea, kumba vormi kasutada. Ainult kolm eestlast vastas eitavalt, s.t nad ei teadnud, et neil oleks kunagi olnud probleemi. Venelastel oli sellist probleemi elus esinenud samuti paljudel, kuid siiski vastas eitavalt 20 venelast, kellest vaid 2 oli Peterburi informantide hulgast. Nii eesti kui ka vene gruppides oli sagedasim arvamus, et kui on kahtlus, teietatakse.

Suhteliselt vähe oli kirjeldatud olukordi, milles kõhklused tekivad. Mõned vastajad siiski märkisid, et probleem tekib siis, kui partner on sama vana või natuke vanem. Kordagi ei nimetatud seda, et probleeme oleks rohkem vastassoo esindajatega. Vastajad ei kirjeldanud eriti täpselt seda, mida nad kahtluse korral teevad. Nii eesti kui ka vene keeles on näiteks umbisikulisi ja geneerilisi vorme, millega lühikeses vestluses on võimalik toime tulla ilma otseselt pöördumata (ülevaadet eesti keele pöördumist “varjavatest” vormidest vt Erelt 1990). 8 eestlast ja 4 venelast väitsidki, et nad väldivad otsest pöördumist. Ühtegi vastavat näidet siiski ei esitatud. Mõned märkisid, et nad kohmetuvad, mõned, et kasutavad eri pöördumisviise vaheldumisi. Neli eestlast ja üks venelane väitsid, et nad küsivad otse, kuidas pöörduda. See, et sellist varianti nägid võimalikuna ainult 5 inimest, näitab, et kummaski kultuuris ei ole tegelikult normaalne küsida, kuidas pöörduda. Ilmselt on põhjus selles, et niisugune küsimus toob endaga kaasa surve partnerile, kuna ta peaks pakkuma lähedasemat ja sõbralikumat *sina*-varianti ka siis, kui see talle tegelikult ei meeldikski.

Mõned on vastanud nii, et sellest võib järeldada nende kõrgendatud tähelepanu partneri käitumise suhtes “probleemaatilises” suhtluses. Näiteks on üks eesti naine kirjutanud, et ebakindluse korral ta teietab ja jälgib kogu aeg, kuidas partner reageerib, üks eesti mees aga, et ta sinatab ja jälgib siis partnerit.

Võib öelda, et kuigi nii eestlaste kui venelaste grupis on umbes võrdselt neid, kes tunnistavad probleemsete situatsioonide olemasolu, leidub siiski ka erinevusi kahe rahvusgrupi vahel. Eelkõige seisneb see selles, et venelaste hulgas on

tunduvalt rohkem neid, kes vastasid, et nad oskavad alati õige pöördumisvormi valida. Kõhkluste olemasolu korral käituvad nii eestlased kui venelased üldjuhul ühtemoodi: pöörduvad *teie*-vormis või püüavad pöördumist vältida, kuid üldjuhul ei küsi partnerilt, kuidas tuleks pöörduda.

5.3. Muud suhtluskombed: eestlased venelaste arvates ja vastupidi

Kuna kõik uuritavad rühmad (v.a Peterburi kontrollgrupp) elavad kakskeelses ühiskonnas, kus on pidevaid vastastikuseid kokkupuuteid, küsisime ka arvamusi teise rühma kohta: eestlastel paluti vastata küsimustele venelaste kohta ja vastupidi. Kohe tuleb öelda, et erinevusi ei osatud eriti palju esitada, olgu seetõttu, et suhtlus eestlaste ja venelaste vahel on vähene, või siis seetõttu, et suhtluskäitumine tundubki väga sarnane.

Rohkem kui muudes rühmades teadvustasid erinevust Kohtla-Järve vene naised, kelle hulgas umbes pooled arvasid, et eestlased sinatavad rohkem. Ka oli selles rühmas selliseid vastuseid, millest kaudselt võib näha suhtumist *sina/teie* valikusse. Näiteks arvas üks informant, et *eestlased tahavad, et isegi venelased ütleksid sina*, millest on näha, et eestlased vastaja arvates sinatavad, venelased aga mitte. Vastati ka, et *venelased peavad reeglitest rohkem kinni ja haritud inimesed teavad alati, et peab teietama*. Tartu vene naistel oli samuti palju arvamusi, et eestlastel on rohkem sinatamist, kuid seda ei olnud eriti kommenteeritud. Üks informant oli kirjutanud, et *eestlastel on lihtsam sina peale üle minna*. Nii Kohtla-Järve eesti kui ka vene mehed ei osanud üldiselt erinevusi välja tuua. Seda võib interpreteerida kas nii, et naised on tähelepanelikumad suhtluse nüansside suhtes, või et eesti ja vene mehed käituvadki sarnasemalt.

Eestlaste hulgas oli samuti nii, et mehed ei osanud õieti erinevustest vene suhtluskommetega midagi arvata, naistel oli välja toodud pisut rohkem erinevusi. Näiteks mõningaid seletusi: *venelased teietavad rohkem vanemaid, noored omavahel aga semutsevad, venelased teietavad vanavanemaid; vene noored on vanemate vastu viisakamad; venelased räägivad kõvema häälega, kiiremini*. Oli siiski ka üksikuid arvamusi, et venelased sinatavad rohkem, kaks Tartu vene naist arvasid näiteks, et venelased teietavad vähem ja on sõbralikumad.

Teine küsimus üldisemate arvamuste grupis oli: *Kas eesti ja vene suhtluskommetes on muid erinevusi?* Eestlased arvasid tihti, et venelased on emotsionaalsemad ja avatumad, üldiselt olid vastused positiivses toonis. Mõned Tartu eesti mehed on siiski märkinud ka venelaste agressiivsust ja familiaarsust. Kokkuvõttes esines kõige rohkem arvamust, et venelased kätlevad rohkem, mõned olid ka kirjutanud, et puudutavad rohkem (*kallistavad, suudlevad, patsutavad*). Ilmselt on venelaste füüsiline suhtlusdistant (vähemalt eestlaste arvates) lähedasem kui eestlastel. Mitmed eestlased tõid välja selle, et venelased kasutavad isanime.

Venelastel on eestlaste kohta rohkem konkreetsemaid arvamusi, vähem on üldistusi, mida võiks interpreteerida rahvusliku iseloomu kirjeldamisena. Mõned Tartu vene naised on siiski arvanud, et eestlased on suletumad ning vaoshoitumad ja jämedamad (*grubyy*). Kõige vähem erinevusi on välja toonud Kohtla-Järve vene mehed. Kõige sagedamini korduv tähelepanek eestlaste vestluskäitumise kohta oli see, et eestlased ei kasuta isanime ja kooliõpilased pöörduvad õpetaja poole

õpetaja. Keegi polnud küll kommenteerinud seda sellest seisukohast, kas isanime mittekasutamine ka häirib. Tõenäoliselt on venelased tänapäeval harjunud sellega, et väljaspool oma kultuuri isanime ei kasutata.

Küsisime ka: *Kas Teil on olnud probleeme suheldes vene keelt kõnelevate inimestega selle tõttu, et Te ei oska valida õiget suhtlusstiili? Tooge mõni näide.* Sellele küsimusele oli üksmeelne vastus peaaegu kõigi poolt, kes üldse vastasid: probleemid on keeleoskuses, mitte suhtluses (keeleoskuse vähesust on tõdenud ka nt Triin Vihalemm 1999).

6. Kolmanda isiku nimetamine

Kolmanda (s.t kohal mitte viibiva) isiku nimetamine kuulub sama palju suhtlus-situatsiooni omaduste hulka kui otsene pöördumine. Nii eesti kui vene kultuuris sõltub see tõenäoliselt mõnevõrra nimetatava staatusest, kuid rohkem siiski suhtle-
lejate omavahelisest suhtest: ametlikus olukorras kasutatakse ka lähedase sõbra kohta ametlikku nimetust, sõpradevahelises argisuhtluses ei nimetata ka presidenti ametlikult.

Võimalikke nimetamisviise on igas kultuuris palju. Vene kultuur on nende poo-
lest aga eriti rikas (vt nt Krongauz 2004: 171–173): lisaks ees- ja perekonnanimele võidakse kasutada ka isanime; eesnimest on peale selle olemas ka lühivariandid ja nende deminutiivid (*Mihhail – Miša, Miška, Mišulka*), teatud suhtlemisstiili juures on lühivariandid ka isanimel (*Mihhailovitš – Mihalõtš*).

Küsisime, *kuidas nimetate sekretärile dotsenti, keda on vaja leida?* Vastust ootasime nelja isiku kohta, kelle nimede järgi oli võimalik otsustada, kas tegu oli eesti mehe, eesti naise, vene mehe või vene naisega. Andmetena oli antud eesnimi, isa nimi (algvormis) ja perekonnanimi. Vastuseks tuli nimi välja kirjutada, seega isanimi tuli kasutamise korral ka moodustada. Küsimustikku koostades lootsime, et situatsioon (õppejõu otsimine) on üliõpilastele väga tuttav ja seetõttu on loota ka tõepäraseid tulemusi.

Esitame näitena naiste (*Mari Mänd ja Tatjana Ivanova*) nimetamisvormid tabelites 2 ja 3 (tabelis 2 ei ole esitatud Peterburi venelaste rühma, kuna Eestist väljaspool ei saa neil olla väljakujunenud viisi eestlast nimetada). Nii eestlastel kui ka venelastel oli kõige tavalisem nimetamisvorm *eesnimi + perekonnanimi* (nt *Mari Mänd*). Mõned eestlased olid proovinud vene nimedele ka isanime lisada, kuid seda esines siiski väga vähe ja mitte alati ei olnud see grammatiliselt õnnestunud. Kui venelased nimetasid venelast, siis olid tavalisemateks variantideks *eesnimi + perekonnanimi* (*Tatjana Ivanova*) ja *eesnimi + isanimi + perekonnanimi* (*Tatjana Sergejevna Ivanova*). Tartu venelaste vastustes olid isanimega ja isanimeta varian-
did esindatud enam-vähem võrdselt, Kohtla-Järve venelaste hulgas oli isanimega nimekuju palju rohkem kui isanimeta.

Tiitlit *dotsent* kasutasid eestlased sagedamini kui venelased, *härä/proua* ja *gospodin/gospoža* olid haruldased, kuid siiski pisut levinumad venelaste vas-
tustes. Eesti venelaste hulgas oli märgata nõrka tendentsi, et *gospoža* ja *gospodin* lisati eelkõige eesti nimede ette. Võimalik, et eesti nimede patronüümitus võib venelaste jaoks tekitada probleeme õige nimetamisvormi valikul ja selle vältimi-
seks kasutavad nad pöördumissõnu. Üks erinevusi vene gruppide vahel oli see, et

Kohtla-Järve vene naiste hulgas oli kõige sagedasem nime perekonnanimega algav variant (*perekonnanimi + eesnimi + isanimi*, *Ivanova Tatjana Sergejevna*), mida peetakse bürokraatliku stiili elemendiks³. Eesti ja Peterburi venelaste käitumine erines selle küsimuse osas eelkõige selle poolest, et Eesti venelastel oli kõikides rühmades üks võimalikke nimetamisvormi *eesnimi + perekonnanimi*, mida Peterburi venelastel vähemalt vene nimede osas kordagi ei esinenud. Võimalik, et tugeva eesti nimekasutuse mõju tagajärjel on Eestis ka venelaste hulgas patronüümide kasutamine vähenemas.

Tabel 2. Kolmanda isiku nimetamine (kui informant oli pakkunud mitu varianti, siis on iga variant arvesse võetud)

Mari Mänd	KEM	KEN	KVM	KVN	TEM	TEN	TVM	TVN
vastanuid	9	7	9	17	16	23	10	18
<i>dotsent + nimi</i>	1	5	0	1	3	9	0	3
<i>proua/gospoža + nimi</i>	1	2	0	3	0	0	1	3
E+l	0	0	0	0	0	0	1	0
E+l+P	0	0	0	0	0	0	1	0
E+P	3	1	7	12	11	14	6	12
P+E+l	0	0	0	0	0	0	0	0
P	4	0	0	0	1	0	1	0
E	0	0	2	0	0	0	0	0
P+E	0	0	0	1	0	0	0	0

Tabel 3. Kolmanda isiku nimetamine (kui informant oli pakkunud mitu varianti, siis on iga variant arvesse võetud)

Tatjana Sergejevna Ivanova	KEM	KEN	KVM	KVN	TEM	TEN	TVM	TVN	P
vastanuid	8	7	8	26	15	22	10	17	27
<i>dotsent + nimi</i>	1	5	0	1	4	8	0	3	2
<i>proua/gospoža + nimi</i>	0	2	0	0	0	0	0	2	1
E+l	0	1	3	3	0	2	3	1	10
E+l+P	0	0	2	4	0	1	1	6	13
E+P	3	0	2	4	9	11	4	6	0
P+E+l	0	0	1	16	0	0	1	3	5
P	4	0	0	0	1	0	1	0	1

P – perekonnanimi, E – eesnimi, l – isanimi

7. Kokkuvõte

Analüüsisime käesolevas artiklis Eesti eestlaste ja venelaste (ning võrdlusgrupina Peterburi venelaste) suhtluskäitumise mõningaid jooni. Eelkõige oli vaatluse all pöördumisvormi (*sina* või *teie*) valik ja kolmanda isiku nimetamine.

Kokkuvõtlikult võib öelda, et eestlaste ja venelaste pöördumisvormi valikut juhivad suures osas samad printsiibid: teietatakse vanemat ja võõrast. Samas on näha, et eestlastele on olulisim parameeter võõrus (mida on näidanud juba Keevallik 1999), venelastele vanus. Kõige ilmekamalt avaldub see pöördumises vanemate

³ Irina Külmoja, isiklik vestlus.

sugulaste poole, kus venelased teietavad oluliselt rohkem kui eestlased, samas kui auditooriumis kohatud võõra üliõpilase poole pöörduvad teietades sagedamini eestlased. Eesti venelaste suhtluskäitumises on teietamise põhjused rohkem teadvustatud sageli kasutatava mõiste *lugupidamine (uvaženie)* kaudu.

Eestlastel näib suhtluskäitumist määravat ka laiem suhtlustaust, näiteks akadeemiline suhtlussfäär tingib nii *teie*-vormi kasutuse kui ka akadeemilise pöördumisvormi *dotsent*, samas kui vene vastajatel ei tundunud see aspekt olevat eriti oluline ei *teie*-vormi ega nimetamisvormi valikul (kolmandat isikut nimetati pigem valdkonnast sõltumatu pöördumisnimetusega *härä* või *proua* kui akadeemilise tiitliga).

Pöördumisvormi valik tekitab mõnikord probleeme nii eestlastele kui venelastele, kõhkluse korral eelistatakse teietamist. Üleminek sinatamisele ei ole eriti reeglilistatud, s.t vastajad ei osanud selgelt välja tuua, millal ja mis tingimustel see juhtub. Eestlaste olulisim faktor tundub pöördumisvormi muutuses olevat suhtluse kestus, venelastel aga lisaks kestusele ka lähedus.

Kultuuridevahelised erinevused suhtluskäitumises ei olnud vastajatel eriti teadvustatud. Vastastikku on teada, et eestlased sinatavad rohkem, venelased kasutavad pöördumisel isanime. Venelaste jaoks näis olevat eriti hästi teadvustatud fakt, et eesti lapsed pöörduvad koolis õpetaja poole *õpetaja*.

Käsitletud küsitlus näitas, et erinevused eestlaste ja venelaste viisakuskäitumises ei ole väga suured, kuid mõnes aspektis (vanuselise distantsi olulisus, lugu pidamine teietamise põhjusena, nimetamisvormide valik jm) olulised.

Viidatud kirjandus

- Borovkova, Olga 2002. Viisakuskood eesti ja vene keeles: pöördumised. Bakalaureusetöö. Käsikiri. Tartu Ülikool.
- Brown, Roger; Gilman, Albert 1960. The pronouns of power and solidarity. – Thomas A. Sebeok (Ed.). *Style in Language*. Cambridge, Mass.: MIT Press, 253–276.
- Brown, Penelope; Levinson, Stephen 1987. *Politeness: Some Universals of Language Use*. Cambridge: Cambridge University Press.
- Erelt, Mati 1990. Kõneleja ja kuulaja kaudse väljendamise võimalusi eesti keeles. – Keel ja Kirjandus, 1, 35–39.
- Formanovskaja 1989 = Формановская, Н. И. 1989. *Речевой этикет и культура общения*. Москва: Высшая школа.
- Goborova, Maria 2007. Pöördumiste eripärad Tallinna eesti ja vene koolide õpilaste seas. Magistritöö. Käsikiri. Tartu Ülikool.
- Kasper, Gabriele; Blum-Kulka, Shoshana (Eds.) 1993. *Interlanguage Pragmatics*. Oxford: Oxford University Press.
- Keevallik, Leelo 1999. The use and abuse of singular and plural address forms in Estonian. – *International Journal of the Sociology of Language*, 139, 125–144. doi:10.1515/ijsl.1999.139.125
- Keevallik, Leelo 2004. Politeness in Estonia: a Matter of Fact style. – Leo Hickey (Ed.). *Politeness in Europe*. Clevedon, England: Multilingual Matters, 203–217.
- Krongauz 2004 = Кронгауз, М. А. 2004. Русский речевой этикет на рубеже веков. – *Russian Linguistics*, 28, 163–187.
- Mills, Sara 2003. *Gender and Politeness*. Studies in Interactional Sociolinguistics. Cambridge: Cambridge University Press.

- Pajusalu, Renate; Vihman, Virve; Klaas, Birute; Pajusalu, Karl 2010. Forms of address across languages: Formal and informal second person pronoun usage among Estonia's linguistic communities. – *Intercultural Pragmatics* 1, 7, 75–101.
- Puusepp, Olga 1999. *Sina ja teie eesti ja vene keeles*. Bakalaureusetöö. Tallinna Pedagoogiline Ülikool.
- Rathmayr, Renate 1999. Métadiscours et réalité linguistique: L'exemple de la politesse russe. – *Pragmatics* 1, 9, 75–96.
- Rathmayr, Renate 2008. Intercultural aspects of new Russian politeness. – *WU Online Papers in International Business Communication / Series One: Intercultural Communication and Language Learning*, 4, February 2008.
- Vihalemm, Triin 1999. Group identity formation processes among Russian-speaking settlers of Estonia: a view from linguistic perspective. – *Journal of Baltic Studies*, 1, 30, 18–39.
- Vogelberg, Krista 2002. Keelelise viisakuse mudelite mõnedest vaieldavatest aspektidest eesti, vene ja inglise keele võrdlevate uuringute valguses. – Renate Pajusalu, Ilona Tragel, Tiit Hennoste, Haldur Õim (toim.). *Teoreetiline keeleteadus Eestis*. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised, 4. Tartu, 297–312.
- Watts, Richard J. 2003. *Politeness. Key Topics in Sociolinguistics*. Cambridge: Cambridge University Press.

Renate Pajusalu (Tartu Ülikool) uurimisvaldkonnad on semantika, pragmaatika, suhtluslingvistika, keele omandamine.
renate.pajusalu@ut.ee

Virve-Anneli Vihmani (Tartu Ülikool) uurimisvaldkonnad on süntaks, semantika, keele omandamine.
virve.vihman@ut.ee

Birute Klaasi (Tartu Ülikool) uurimisvaldkonnad on keeletüpoloogia, eesti, soome ja balti kontrastiivsed uuringud, eesti keel teise keelena.
birute.klaas@ut.ee

Karl Pajusalu (Tartu Ülikool) uurimisvaldkonnad on keele varieerumine, sotsiolingvistika, eesti murded ja läänemeresoome keeled.
karl.pajusalu@ut.ee

Lisa 1.

Tartu Ülikooli keeleteadlased paluvad Teil vastata järgmistele küsimustele. Vastused on anonüümsed ja neid kasutatakse ainult teaduslikuks uurimistööks.

Alguses palume kirjutada enda kohta mõned meile olulised andmed:

sugu MEES NAINA
vanus
ema keel
isa keel
kodune suhtluskeel
õppekeel koolis
elukoht

Kas kasutate suhtluses vene keelt? EI MÕNIKORD TIHTI

Nüüd palume vastata järgmistele küsimustele:

1. Kas Te teietate allpool loetletud inimesi? Tõmmake ring ümber kas numbrile 1, 2 või 3.

1 – alati, 2 – mõnikord, 3 – mitte kunagi.

sõber	1	2	3
sõbranna	1	2	3
sõprade vanemad	1	2	3
oma poisi/tüdruku vanemad	1	2	3
vanemate naistuttav	1	2	3
vanemate meestuttav	1	2	3
onu	1	2	3
tädi	1	2	3
võõras vana mees bussis	1	2	3
võõras vana naine bussis	1	2	3
võõras keskealine mees bussis	1	2	3
võõras keskealine naine bussis	1	2	3
võõras neiu	1	2	3
võõras noormees bussis	1	2	3
võõras meesüliõpilane loenguruumis	1	2	3
võõras naisüliõpilane loenguruumis	1	2	3
müüja	1	2	3
ettekandja	1	2	3
politseinik	1	2	3
panga meesametnik	1	2	3
naabrimees	1	2	3
naabrinaine	1	2	3

2. Kas Te teietate õppejõude?

Kui jah, kas oleks tore, kui võiks sinatada? Miks?

3. Kas õppejõud teietavad Teid?

Kui jah, siis kas oleks tore, kui nad sinataksid? Miks?

4. Kuidas minnakse suhtluses tavaliselt üle teietamiselt sinatamisele? Kas selleks on mingeid reegleid ja kas Te ise järgite neid?

5. Kas mõnikord on juhtunud, et Te ei tea, kas sinatada või teietada? Mida Te siis teete?

6. Kas eesti ja vene kultuuris on sinatamise ja teietamise suhtes erinevusi? Milliseid?

7. Kas eesti ja vene kultuuris on suhtluskommetes muid erinevusi? Tooge mõni näide!

8. Kas Teil on olnud probleeme suheldes vene keelt kõnelevate inimestega selle tõttu, et Te ei oska valida õiget suhtlusstiili? Tooge mõni näide.

9. Kuidas nimetate sekretärile dotsenti, keda on vaja leida?

Õppejõu andmed

Mina nimetan nii

(kasutage ainult neid, mida peate vajalikuks)

Eesnimi: Mari

Isa nimi: Jüri

Perekonnanimi: Mänd

Eesnimi: Jaan

Isa nimi: Tarmo

Perekonnanimi: Tamm

Eesnimi: Tatjana

Isa nimi: Sergei

Perekonnanimi: Ivanova

Eesnimi: Georgi

Isa nimi: Ivan

Perekonnanimi: Trofimov

POLITENESS BEHAVIOR AMONG ESTONIANS AND RUSSIANS: FORMAL AND INFORMAL SECOND PERSON AND THE THIRD PERSON

Renate Pajusalu, Virve Vihman, Birute Klaas, Karl Pajusalu

University of Tartu

This paper examines linguistic politeness among Estonians and Russians in Estonia. We raise the question of how groups living in different conditions of societal bilingualism acknowledge different politeness norms and hence exhibit different politeness behavior. Our study focuses especially on the choice of the formal or informal second person pronoun (the T/V phenomenon), but we also investigate other aspects of interaction (such as respondents' metalinguistic awareness of forms of address and reference to an absent third person). The analysis is based on university students' responses to a sociolinguistic questionnaire. Data from native Russian speakers in St. Petersburg are used for comparison with data from Russian speakers in Estonia.

On the basis of the findings from this study, we conclude that the communicative culture of Russians and Estonians is not greatly different, although certain particularities exist which differentiate them. Both Russians and Estonians use the formal V pronoun when addressing strangers older than themselves. The most notable difference in making the choice between forms of address is that Russians choose the formal V form primarily on the basis of age difference, addressing even close acquaintances with V if they are older. This behavior is explained by the speakers with the concept of *uvaženie*, or "respect".

For Estonians, age differences are less significant than familiarity, and close acquaintances or relatives are usually addressed with T, even if they are much older. There is no clear, unified stereotypical explanation among the Estonian respondents of why one should use V with anyone (for instance, a lecturer). Explanations tend to appeal to the context of discourse (e.g. it is habitual to use V in the context of academic discussion). Because the norms for using V are clearer amongst Russian speakers, Russians seem to have fewer dilemmas than Estonians regarding which form to use.

Keywords: politeness, forms of address, pragmatics, Estonian, Russian

ÕPPIJAKEELE KONTEKSTUAALNE VARIEERUVUS EESTI KEELE TÄIS- JA OSASIHITISE KASUTAMISEL

Raili Pool

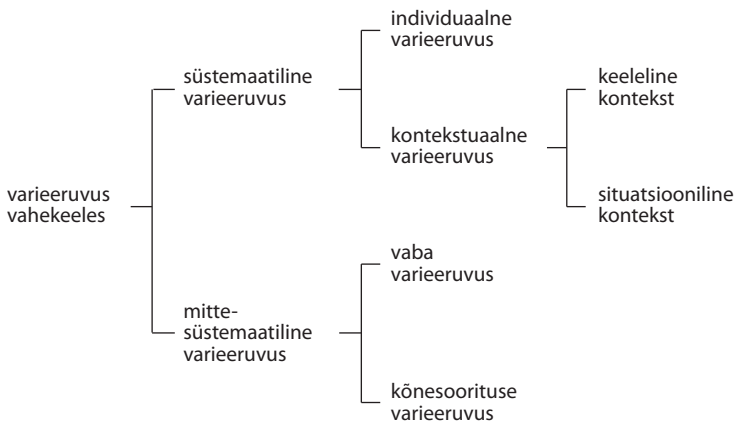
Ülevaade. Artiklis antakse ülevaade õppijakeeles esineva varieeruvuse liikidest ning keskendutakse kontekstuaalsele varieeruvusele, mis avaldub nii keelelises kui ka situatsioonilises kontekstis. Üldteada on asjaolu, et eri keelekasutussituatsioonidest kogutud õppijakeele materjal võib anda erinevaid tulemusi, kuid eesti keele osas on vastavaid empiirilisi uurimusi seni tehtud väga vähe. Artiklis käsitletakse vene emakeelega eesti keele õppijate täis- ja osasihitise kasutuse erinevusi vabas suulises ja kirjalikus keelekasutuses. Vaadeldav informandirühm saavutas suulises materjalis eesti keele grammatikast lähtuvalt märgatavalt paremaid tulemusi kui kirjalikes tekstides. Põhjuseks tuleb pidada suulise ainestiku lihtsamat keelelist konteksti, mis avaldub süntaktiliste tarindite väiksema keerukuse ning piiratuma verbide valiku koosmõjus.*

Võtmesõnad: situatsiooniline kontekst, grammatiline korrektsus, keeleline kontekst, süntaks, sõnavara, vene keel, eesti keel

1. Sissejuhatus

Teise keele õppijate poolt kõneldav-kirjutatav õpitava keele variant ehk õppijakeel (vrd vahekeel) on oma olemuselt ühtaegu nii süstemaatiline kui ka varieeruv, kusjuures ka varieeruvus võib õppijakeeles olla nii süstemaatiline kui ka mittesüstemaatiline. Joonis 1 annab ülevaate õppijate keelekasutuses esineva varieeruvuse jagunemisest (Ellis 1990: 76).

* Artikli valmimist on toetanud riiklik programm "Eesti keel ja kultuurimälu" ning sihtfinantseeritav teadusteema SF0180056s08.



Joonis 1. Õppijakeele varieeruvuse tüübid

Artiklis tuleb vaatluse alla õppijakeele kontekstuaalne varieeruvus ehk täpsemalt keelelise ja situatsioonilise konteksti roll õppijakeele varieeruvuse seletamisel. Lähte-eelduseks on teadmised, et õppijakeele andmete kogumise viis võib mõjutada saadavaid uurimistulemusi, kuna erinevates keelekasutussituatsioonides võidakse rakendada eri tüüpi keeleteadmisi ning kasutada erinevaid keelelisi stiile. Situatsioonilise konteksti mõju õppijakeele stiilile kujutab Elaine Tarone (1983: 152) kontiinumina, mille ühes ääres on vaba stiil, mis tuleb esile spontaanses suulises keelekasutuses, teises ääres aga ettevaatlik stiil, mille puhul on õppijate kontroll oma keelekasutuse üle kõige tugevam ja mis avaldub kõige kujukamalt grammatilise korrektsuse hindamise testides (vt joonis 2). Vaba ja ettevaatliku stiili vahele jääb veel määramata hulk erineva enesekontrollitasemega stiile, mis tulevad esile erinevates keelekasutussituatsioonides.

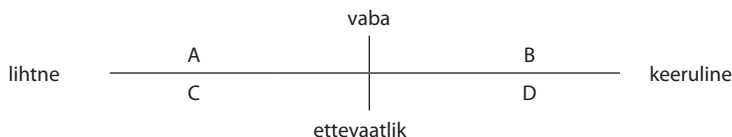
vaba stiil	stiil 2	stiil 3	stiil 4	stiil n	ettevaatlik stiil
spontaanse kõne materjal	planeeritud kõne materjal		erinevad ülesanded (nt imiteerimine, lausete kombineerimine)		grammatilise korrektsuse hindamine

Joonis 2. Vahekeele stiilide kontiinum

Joonisel 2 kujutatud vahekeele stiilide kontiinumit võib vaadelda ka õppijate implitsiitsete ja eksplitsiitsete keeleteadmiste avaldumise kontiinumina. Kuna implitsiitsed keeleteadmised on sellised, mida teise keele kõneleja kasutab automaatselt, teadvustamata, avalduvad need tõenäoliselt kõige kujukamalt spontaanses suulises keelekasutuses ehk vabas stiilis. Eksplitsiitseid teadmisi suudavad õppijad kontrollida ja verbaliseerida, mistõttu need on lähedased joonisel 2 kujutatud kontiinumini ettevaatlikule stiilile, mis avaldub kõige ilmsemal kujul mitmesugustes metalingvistilisi teadmisi kontrollivates ülesannetes.

Keelelise konteksti mõju õppijate keelekasutuse varieerumisele avaldub juhtudel, mil kaks erinevat keelelist konteksti põhjustavad õppijakeeles kahe erineva vormi kasutuse, isegi kui sihtkeeles eeldaksid need kaks konteksti tegelikult sama vormi kasutust (Ellis 1990: 83). Näiteks võib siinkohal tuua eesti sihitise tarvituse,

mille korral sihtkeelepärane partitiivobjekt nt ühendites *loen raamatut* ja *loetakse raamatut* võib õppijakeeles avalduda esimesel juhul samuti partitiivsena, kuid impersonaalse öeldisverbi puhul sageli just verbi morfoloogilisest vormist lähtuvalt nominatiivsena ehk kujul *loetakse raamat*. Keelelise ja situatsioonilise konteksti mõju õppijakeele varieeruvuse seletamisel võtab Rod Ellis (1990: 84) kokku nii, nagu see on kujutatud joonisel 3: keelelised kontekstid avalduvad kontiinumina, mille äärmised otsad on lihtne ja keeruline keeleline kontekst, ning seda kontiinumit vaadeldakse seoses õppijakeele stiili kontiinumiga, mille äärmised avaldumisvormid on vaba ja ettevaatlik stiil. Õppijakeele uurimisel paigutub vaadeldav keeleline vorm ühte neljast lahtrist A–D, näiteks lahtrisse A paigutuvad vormid, mida kasutati lihtsates keelelistes kontekstides ning vabas stiilis.



Joonis 3. Keelelise ja situatsioonilise konteksti mõju vahekeele varieeruvusele

Erinevatest keelekasutussituatsioonidest kogutud materjalil põhinevaid uurimusi eesti keele kui teise keele omandamise kohta on praeguseks tehtud väga vähe. Suulist ja kirjalikku õppijakeelt on võrreldud kahes Tartu Ülikoolis kaitstud bakalaureuse-töös: Alina Fjodorova (2004) jõudis sõnavaravigu käsitlevas töös tulemuseni, et vabas suulises keelekasutuses tegid õppijad ajalise surve tõttu rohkem sõnavaravigu kui kirjandites, Olga Pastuhhova (2004) vaatles emakeele mõjust tingitud süntaksivigu vabas suulises õppijakeeles ning kirjalikes tõlkelausetes, kuid kirjaliku materjali vähesuse tõttu ei saanud eri keelekasutussituatsioonide kohta suuremaid järeldusi teha. Eri tekstitüüpides kasutatud sõnavara rikkust L1- ja L2-kõnelejalatel on võrrelnud Hille Pajupuu, Krista Kerge ja Pilvi Alp (2009), kelle poolt uuritud rühmadel suurenes sõnavara rikkus skaalal suuline dialoog, suuline monoloog, kirjalik essee ning L2 sõnavara oli iga tekstitüübi puhul L1 omast märkimisväärselt vaesem, kuigi mitte-estlaste puhul oli tegemist eesti keele kõrgtaseme eksami edukalt sooritanud kõnelejatega. Soome keele osas on suulise ja kirjaliku õppijakeele erinevusi uurinud Annekatrin Kaivapalu, kes oma soome keele muutemorfoloogia omandamist käsitlevas doktoriväitekirjas võrdles muuhulgas ka kirjaliku ja suulise üksiksõnade käänamistesti tulemusi (Kaivapalu 2005). Tulemused olid seotud õppijate emakeelega: eestikeelsed õppijad käänasid ootuspäraselt peaaegu kõiki testisõnu kirjalikult paremini kui suuliselt, kuid venekeelsed õppijad saavutasid kõigi sõnade puhul suuliselt paremad tulemused (Kaivapalu 2005: 206, 232). Autor põhjendab seda venekeelsete õppijate eestikeelsetest erinevate käänamisstrateegiate kasutamisega (Kaivapalu 2008).

Artiklis käsitlen eesti keele täis- ja osasihitise kasutamist eesti keelt kõrgtasemel (võrreldav keeleoskustasemega C1) valdavate vene emakeelega üliõpilaste vabas suulises ja kirjalikus eesti keeles. Lähtun eeldusest, et vaadeldav materjal paigutub situatsioonilise konteksti poolest vahekeele stiilide kontiinumis eri kohtadesse: a) suulise keelekasutuse materjal on kogutud informaalsetest vestlustest, mille puhul informandid hinnangulist tagasisidet ei saanud ning võib eeldada,

et nende kontroll oma keelekasutuse üle oli nõrgem – see materjalikogu esindab õppijakeele vabamat stiili; b) kirjalik materjal on pärit informantide jaoks olulise tähtsusega eksamikirjanditest, mille puhul oli vigade vältimine väga oluline ning võib eeldada, et informantide kontroll oma keelekasutuse üle oli suhteliselt suur – see ainek on joonisel 2 kujutatud skaalal lähemal ettevaatlikule stiilile. Artiklis toon välja kahe materjalikogu analüüsimisel saadud tulemused ning vaatlen, kas suulise ja kirjaliku õppijakeele erinevuste seletamisel on võimalik aluseks võtta keelelise konteksti mõju.

2. Materjal ja informandid

Artiklis analüüsitan keeleline materjal on kogutud suurema täis- ja osasihitise kasutust käsitleva uurimuse tarbeks, mille seniseid tulemusi on kirjeldatud mitmes artiklis (Pool 2005, 2006b, 2007b, 2007a). Uurimuse informantideks on Tartu Ülikooli eesti keele (võõrkeelena) eriala vene emakeelega üliõpilased, kes olid materjali kogumise hetkeks jõudnud oma studiumi lõpuossa ning olid sooritanud ka eesti keele kõrgtaseme eksami. Muudest eesti keelt kõrgtasemel valdavatest mitte-eestlastest moodustavad nad mõneti erineva rühma seetõttu, et tegemist on filoloogiaüliõpilastega, kes olid lisaks praktilise eesti keele õpetusele saanud ka põhjaliku ning süsteemse teoreetilise eesti keele alase ettevalmistuse. Informandid moodustavad nii emakeele, vanuse, eesti keele oskuse kui haridustaseme poolest homogeense rühma. Artiklis analüüsitan suuline ja kirjalik keeleaines pärineb samadelt informantidelt ning jaguneb järgmiselt.

- 1) Kõrgtaseme kirjalik keelekasutus: 28 eesti keelt kõrgtasemel valdava vene emakeelega üliõpilase eesti keele lõpueksami kirjandiosast kogutud objektid. Kirjandite keskmine pikkus on 369 sõna, tekstikogu maht kokku 10 332 sõna, sh 581 objekti.
- 2) Kõrgtaseme suuline keelekasutus: 15 eesti keelt kõrgtasemel valdava vene emakeelega üliõpilasega läbiviidud suuliste vestluste litereeringutest kogutud objektid. Lindistusi on kokku 10 tundi, millest on registreeritud 882 objekti.

Uurimus on oma olemuselt kvalitatiivne ning meetodina kasutan sihtkeelepärase kasutuse analüüsi, mille korral vaadeldakse õppijakeele grammatilist korrektsust sihtkeele reeglite seisukohast, kusjuures lisaks uuritava keelevormi esinemisele obligatoorsetes kontekstides võetakse arvesse ka ülekasutuse juhud (vt Pica 1984, Ellis, Barkhuizen 2005: 79–80). Mingi keelelise vormi obligatoorseks kontekstiks peetakse keelelist konteksti, milles emakeelekõnelejad kasutaksid grammatiliselt korrektse lausungi produtseerimiseks just seda vormi (Tarone, Yule 1991: 126, Tarone 1988: 123). Õppijakeele analüüsimisel rakendan lingvistilist lähenemisviisi. Artikli eesmärkideks on

- 1) tuua välja suulise ja kirjaliku õppijakeele objektikasutuse erinevused ehk teisisõnu jälgida situatsioonilise konteksti mõju õppijakeele sihtkeelepärasele korrektsusele,
- 2) selgitada välja keelelise konteksti mõju suulise ja kirjaliku objektikasutuse erinevustele, võttes arvesse nii grammatilisi kui ka leksikaalseid õppijakeele erijooni.

3. Sihitisekasutuse sihtkeelepärane korrektsus suulises ja kirjalikus õppijakeeles

Suulise ja kirjaliku õppijakeele materjalikogust on eraldatud eitavate konstruktsioonide objektid, kuna need esinesid reeglipäraselt partitiivis ning ei paku objekti käändevahelduse seisukohast huvi. Välja on jäetud ka vormihomonüümiaga objektid, mille puhul pole võimalik üheselt kindlaks määrata, millist objektikäänet kolmest võimalikust (nominatiiv, genitiiv või partitiiv) on informandid soovinud kasutada.¹ Artiklis on vaatluse all vaid need käändsõnalised objektid, milles on võimalik kindlaks määrata eesti keele kui sihtkeele grammatika seisukohast korrektsete ja ebakorreksete kasutusjuhtude vahet, andmed on esitatud tabelites 1 ja 2.²

Tabel 1. L2-kõnelejate objektitarvitus vabas kirjalikus eesti keeles (28 üliõpilast)

	Kokku	Korrektsed	Ebakorrektsed või küsitavad
Partitiiv	378	338 (89,4%)	40 (10,6%)
Genitiiv	42	33 (78,6%)	9 (21,4%)
Nominatiiv	32	20 (62,5%)	12 (37,5%)
Kokku	452 (100%)	391 (86,5%)	61 (13,5%)

Tabel 2. L2-kõnelejate objektitarvitus vabas suulises eesti keeles (15 üliõpilast)

	Kokku	Korrektsed	Ebakorrektsed või küsitavad
Partitiiv	532	513 (96,4%)	19 (3,4%)
Genitiiv	91	87 (95,6%)	4 (4,4%)
Nominatiiv	92	70 (76%)	22 (24%)
Kokku	715 (100%)	670 (93,7%)	45 (6,3%)

Arvandmetest on näha, et informandid peavad partitiivi eesti keele sihitise põhikäändeks, seejuures on partitiivobjekti osakaal täissihitise käänetega võrreldes suurem kirjalikus õppijakeeles. Tabelites 1 ja 2 toodud õppijakeele andmed annavad ülevaate sihitisekäänete ülekasutusest (nt kirjalikus materjalis esines 40 partitiivobjekti ülekasutuse juhtu, mis avalduvad sihtkeele seisukohast vigadena), samas ei paku need andmed infot selle kohta, millist objektikäänet oleks neil ülekasutuse juhtudel tegelikult tulnud kasutada ehk missugune objektikäänne on jäänud kasutamata. Nii oleks näites (1) pidanud eesti keele grammatikareeglite kohaselt olema partitiivi asemel genitiiv, näites (2) genitiivi asemel nominatiiv, näites (3) nominatiivi asemel partitiiv, näites (4) nominatiivi asemel genitiiv jne.

- (1) See mees unustas oma **perekonda** (pro *perekonna*), mis jäi kodumaale. (kirjalik)
- (2) Milleks võtta tööle **kellegi suvalise kõrgharidusega inimese** (pro *keegi suvaline kõrgharidusega inimene*), kui kõrvalmajas elav onupoeg või vennatütar otsib parajasti tööd? (kirjalik)
- (3) vot Tartuga juhtus niimoodi et me tulime **dokumendid** (pro *dokumente*) ära anda (.) ära andma (suuline)

¹ Materjalis esinenud eitavate konstruktsioonide objekte ning vormihomonüümseid objekte ja nende analüüsimise problemaatikat on pikemalt käsitletud artiklites Pool 2005, Pool 2006a ja Pool 2007a, mistõttu pole otstarbekas sama informatsiooni siinses kirjutises korrata.

² Õppijakeele analüüsimisel olen lähtunud grammatikareeglitest ning iseenda kui emakeelkõneleja keeletajust. Ühte rühma olen liigitanud nii eesti keele seisukohast ilmsed vead kui ka küsitavad piiripealsed juhtumid.

(4) ma küll ei teadnud et peab palju varem muretsema **ühiselamukoht** (pro *ühiselamukoha*) (suuline)

Õppijakeele sihtkeelepärasesest korrektsusest täieliku ülevaate saamiseks tuleks üht ja sama sihtkeele seisukohast ebakorrektsust või küsitavat vormi (tabelitest 1 ja 2) käsitleda kahest vaatenurgast – ühe objektikäände üle- ning teise objektikäände alatarvituse juhtumina.

Sihtkeele seisukohast nii õigete kui ka valede (sealhulgas nii ületarvitatud kui ka vajalikus kontekstis tarvitamata jäetud) vormide kaasahaaramiseks grammatilise korrektsuse protsendi arvutamisel saab kasutada Teresa Pica (1984) valemit:

$$\frac{\text{korrektsete tarvitajate arv obligatoorses kontekstis}}{\text{obligatoorsete kontekstide arv} + \text{mitteobligatoorsetes kontekstides tarvitatud arv}} \times 100 = \text{korrektsuse \%}$$

Seda valemit kasutades saame suulise ja kirjaliku õppijakeele objektikasutuse sihtkeelepärase korrektsuse kohta täpsemad protsentuaalsed andmed, tulemused on esitatud tabelis 3.

Tabel 3. Õppijakeele objektikasutuse sihtkeelepärane korrektsus

	L2 kirjalik	L2 suuline
Partitiiv	87,1%	92,8%
Genitiiv	45,2%	81,3%
Nominatiiv	38,5%	70%

Tabeli 3 andmete põhjal on vaadeldava informandirühma objektikäänete omandatuse hierarhia partitiiv > genitiiv > nominatiiv ning see langeb kokku eesti keele sihitisekäänete markeerituse hierarhiaga³. Teise keele omandamise uurimisel on tavaks pidada mingit struktuuri keeleõppijate poolt omandatuks, kui selle sihtkeelepärase korrektsuse protsent on 90 või suurem (taustaks on tõsiasi, et ka emakeelekõnelejad ei pruugi alati saavutada saajaprotsendilist korrektsust) (Ellis, Barkhuizen 2005: 84–85). Seda arvestades võib informandirühma kui tervikut vaadeldes omandatusest rääkida vaid partitiivobjekti puhul, mille kasutamine on suulises keelekasutuses õnnestunud veidi paremini kui kirjalikus. Tähelepanu väärib täissihitise käänete – nominatiivi ja genitiivi – sihtkeelepärase korrektsuse sõltuvus situatsioonilisest kontekstist: suulise materjali ainekus on täissihitise korrektsuse protsent märgatavalt suurem kui kirjalikus. Seesugune tulemus on huvitav ka seetõttu, et kirjaliku keelekasutuse aines on kogutud informantide jaoks nn kõrgete panustega eksamitööst, mille hinne oli neile otsustava tähtsusega ja mille kirjutamisel püüti igati vältida.⁴ Suuline materjal on seevastu pärit vaba ülesehitusega intervjuudest, mille puhul hinnangulise tagasiside võimalus puudus ning informantidel ei olnud vajadust oma keelekasutuse korrektsust pingsalt jälgida. Siinsed andmed ei kinnita mitmete varasemate uurimuste tulemusi, mille kohaselt saavutasid teise keele õppijad ettevaatliku stiili korral vabast stiilist suurema grammatilise korrektsuse, kuna see võimaldab vormile suuremat tähelepanu pöörata (Larsen-Freeman, Long 1991: 84–85, Tarone 1988: 82–87).

Käsitatud keeleainese põhjal võib öelda, et situatsioonilisel kontekstil on täis- ja osasihitise kasutamisele mõju: vaba suulise keelekasutuse andmed näitavad siht-

³ Markeerituse seosest teise keele kõnelejade objektikäänete omandamisega vt Pool 2007b.

⁴ Täis- ja osasihitise vältimise võimalustest lähemalt vt Pool 2007a.

keele grammatika seisukohast oluliselt paremat tulemust kui kirjalikus materjalis, mille puhul oli informantide kontroll tõenäoliselt oma keelekasutuse üle tugevam. Kirjaliku materjali allikaks olevates eksamitöödes rakendasid informandid eeldatavasti informaaalsest suulisest kõnest rohkem eksplitsiitseid keeleteadmisi, mis on teadvustatud ning hõlmavad teadlikku reeglikasutust (vt Ellis, Barkhuizen 2005: 5–6). Ometi ei ole see viinud sama heade tulemusteni kui suulises keelekasutuses, seda eriti nominatiiv- ja genitiivobjekti puhul, mille tarvitamine on eesti keeles reguleeritud teise keele kõneleja jaoks keeruliste reeglitega. Artikli järgmises osas vaatlen, kas suur erinevus suulise ja kirjaliku materjali täissihitiste kasutuses võib tuleneda eri keelekasutussituatsioonide keelelise konteksti mõjust.

4. Keelelise konteksti seos objektikänete sihtkeelepärase korrektsusega

4.1. Grammatiline kontekst: sihitis finiiitse ja infiniitse verbivormi laiendina

Järgnevas analüüsis püüan välja selgitada, kas suulise materjali sihitisekasutuse suurem grammatilise korrektsuse protsent võiks tuleneda sellest, et suulise kõne objekte oli kasutatud lihtsamal keelelises kontekstis ning kas kirjalikus materjalis esinenud vigade suur suhtarv võiks olla seotud keerukama keelelise kontekstiga. Niisugusel kujul püstitatud hüpoteesi kohaselt peaksid suulise ja kirjaliku õppijakeele objektitarvitusjuhud paiknema situatsioonilise ja keelelise konteksti ristumise skeemil (joonis 3) keelelise konteksti skaalal vähemalt osaliselt erinevates punktides: suulise ainestiku objektid eeldatavasti rohkem lõigus A kui B ning kirjaliku materjali objektid rohkem lõigus D kui C.

Õppijakeeles tarvitatud sihitiste keelelise konteksti keerukust määratledes on mõttekas lähtuda objekti sisaldavate lausete süntaktilisest struktuurist. Järgnevas jaotuses olen lähtunud esmalt sellest, kas sihitis laiendab finiiitset või infiniitset verbivormi. Seejuures võib kuulumist finiiitse verbivormi juurde pidada sihitise puhul lihtsamaks keeleliseks kontekstiks kui infiniitse verbivormi laiendamist. Sihitise kuulumist infiniitse verbivormi juurde võib teise keele kõnelejal olla üldse keeruline tajuda ning eriti *da*-infinitiivi laiendavate täissihitiste kasutamise aluseks on eesti keele grammatikas väga komplitseeritud reeglid, mis põhinevad *da*-infinitiivi lauseliikmelisel funktsioonil, nõudes teise keele kõnelejatelt detailseid süntaksialaseid teadmisi. Tabelid 4–6 sisaldavad arvilisi kokkuvõtteid täis- ja osasihitise esinemusest vastavalt finiiitse ja infiniitse verbivormi laienditena. Kuna suulise ja kirjaliku materjali hulk on erinev, sobivad võrdlemiseks eelkõige protsentuaalsed andmed.⁵

Tabel 4. Partitiivobjektid finiiitse ja infiniitse verbivormi laiendina

	L2 suuline	L2 kirjalik
Finiitne	412 (77,4%)	273 (72,2%)
Infiniitne	120 (22,6%)	105 (27,8%)
Kokku	532 (100%)	378 (100%)

⁵ Kuna tegemist on kvalitatiivse analüüsiga, on arvandmetel siin peamiselt illustratiivne tähendus. Oluline on ka märkida, et kõik esitatavad järeldused kehtivad ainult selle informandirühma keelekasutuse kohta ega laiene muudele eesti keele õppijatele.

Tabel 5. Genitiivobjektid finiiitse ja infiniitise verbivormi laiendina

	L2 suuline	L2 kirjalik
Finiitne	81 (89%)	28 (66,7%)
Infiniitne	10 (11%)	14 (33,3%)
Kokku	91 (100%)	42 (100%)

Tabel 6. Nominatiivobjektid finiiitse ja infiniitise verbivormi laiendina

	L2 suuline	L2 kirjalik
Finiitne	75 (81,5%)	20 (62,5%)
Infiniitne	17 (18,5%)	12 (37,5%)
Kokku	92 (100%)	32 (100%)

Tabelitest 4–6 on näha, et suulises ja kirjalikus õppijakeeles kasutatud sihitiste keeleline kontekst on osaliselt erinev: kirjalikus ainekust esines protsentuaalselt rohkem infiniitseid verbivorme laiendavaid sihitisi kui suulises, seejuures on erinevused eriti ilmsed genitiiv- ja nominatiivobjektide puhul. Nende andmete alusel on küll võimalik väita, et suulise ja kirjaliku õppijakeele ainekust paigutub keelise konteksti skaalal keerukuse alusel eri kohtadesse – suuline materjal lihtsama ning kirjalik vastavalt keerulisema konteksti poole –, kuid tabelid 4–6 ei anna infot selle kohta, kuidas erinev keeleline kontekst mõjutab objektikasutuse sihtkeelepärast korrektsust.

Järgevalt on vaja leida vastus küsimusele, kas infiniitset verbivormi laiendavate sihitiste suurem hulk kirjalikus õppijakeeles mõjutab vigade hulka. Selleks on võimalik eristada tabelites 1 ja 2 väljatoodud ebakorrektsete ja küsitavate objektikasutusjuhtude (teisisõnu ületarvitusjuhtude) hulgas vastavalt finiiit- ja infiniitverbi juurde kuuluvad sihtkeele seisukohast valesti kasutatud objektid. Tabelites 7 ja 8 esitatakse vales käändes tarvitatud objektide andmed (käände nimi tähistab juhtumeid, kui oli kasutatud seda, kuid pidanuks olema muu objektikääne). Samas tuleb arvestada, et nii väikese materjalihulga puhul võib ka lihtsalt juhuslikkusele olla üsna suur roll ning arvulistele andmetele ei saa omistada liiga suurt tähtsust.

Tabel 7. Kirjaliku õppijakeele objektivigade jagunemine keelise konteksti järgi

	Finiitverbi laiend	Infiniitverbi laiend	Kokku
Partitiiv	30 (75%)	10 (25%)	40 (100%)
Genitiiv	4 (44,4%)	5 (55,6%)	9 (100%)
Nominatiiv	7 (58,3%)	5 (41,7%)	12 (100%)

Tabel 8. Suulise õppijakeele objektivigade jagunemine keelise konteksti järgi

	Finiitverbi laiend	Infiniitverbi laiend	Kokku
Partitiiv	18 (94,7%)	1 (5,3%)	19 (100%)
Genitiiv	2 (50%)	2 (50%)	4 (100%)
Nominatiiv	15 (68,2%)	7 (31,8%)	22 (100%)

Tabelis 7 esitatud andmed näitavad, et $\frac{3}{4}$ **kirjaliku õppijakeele** partitiivobjektide vigu (s.o juhte, kus partitiivi asemel oleks pidanud olema muu objektikääne) on tehtud finiiitverbi, $\frac{1}{4}$ infiniitverbi laiendavates sihitistes. Tabelist 4 on näha, et

selles materjaliosas oligi umbes kolmveerand kõigist partitiivobjektidest kasutatud finiiitse ning veerand infiniitse verbivormi laiendina, seega võib öelda, et keerulisem keeleline kontekst ei ole mõjutanud partitiivobjektide vigade hulka kirjalikus õppijakeeles. Näited (5) ja (6) illustreerivad kirjalikus õppijakeeles finiiitverbi laiendina kasutatud partitiivobjekte, näited (7) ja (8) vastavalt infiniitverbide objekte.

- (5) Ootasin väga **esmatutvumist** Tartu linnaga. (kirjalik)
- (6) Jah, **sedä** (pro *selle*) mõtles inimene välja. (kirjalik)
- (7) Ning on raske kohe öelda ja väljendada **kõiki mõtteid ja tundeid**, mis on seotud ülikooliga. (kirjalik)
- (8) Lõpuks ma tahaks öelda, et ülikooli astumisega täitusid minu unistused: astuda Tartu ülikooli ning õppida **eesti keelt** (pro *eesti keel*) ära. (kirjalik)

Genitiiv- ja nominatiivobjekte esines kogu materjalis partitiivobjektidest tunduvalt vähem, samuti on nende käänete kasutamisel tehtud vigade hulk vastavalt väiksem.

Kirjalikus õppijakeeles täissihitise kasutamisel tehtud vead (juhud, kus kasutatud täissihitise käände asemel oleks pidanud olema kas teine täissihitise kääne või osasihitis) jagunevad peaaegu võrdselt finiiit- ja infiniitverbide laiendite vahel, genitiivobjektide vigu oli pisut enam infiniitse, nominatiivobjektide vigu aga finiiitse verbi laiendite hulgas. Kui võrrelda tabeli 7 andmeid tabelites 5 ja 6 esitatud andmetega kõigi täissihitiste jaotumisest finiiit- ja infiniitverbide laienditeks, on võrreldes partitiivobjektide kasutusega näha erinevusi. Seos keelelise konteksti keerukuse ja vigade arvu vahel on kõige ilmsem genitiivobjektide puhul: üle poole selle sihitisekäände vigadest oli tehtud infiniitse verbi laiendites, samas laiendas infiniitverbi vaid kolmandik kõigist kirjaliku õppijakeele genitiivobjektidest. Seega on põhjust oletada, et kui kirjalikus materjalis oleks genitiivobjekti kasutatud vähem infiniitse ja rohkem finiiitse verbivormi laiendina, oleks selle sihitisekäände tarvitamisel tõenäoliselt tehtud ka vähem vigu. Näited (9) ja (10) illustreerivad genitiivobjektide ületarvitusest tingitud vigu infiniittarindites.

- (9) Mul on unistus **hispaania keele** (pro *hispaania keel*) ära õppida. (kirjalik)
- (10) Kolmas grupp koosneb nendest, kes astub ülikooli, et üliõpilase **staatuse** (pro *status*) saada ja välismaale tööle sõita. (kirjalik)

Näidetes (9), (10) ja eespool toodud näites (2) on informandid valinud genitiivi kui primaarse täissihitise käände, tehes seega sihtkeele seisukohast õige valiku täis- ja osasihitise vahel, kuid eksides täissihitise vormi valikul. Näidetes kasutatud infiniittarindites on eesti keele grammatika kohaselt täissihitise käändena võimalik vaid nominatiiv, mille kasutuse tingib lausete süntaktiline struktuur: näites (9) laiendab sihtis atribuudi funktsioonis *da*-infinitiivi, näidetes (10) ja (2) iseseisva predikaadina talitlevat *da*-infinitiivi. Genitiivse täissihitise tarvitamine *da*-infinitiivi laiendina on võimalik vaid vähestel juhtudel: kui *da*-infinitiiv on lauses sihitiseks (11) või esineb koos modaalverbiga liitöeldise koosseisus (12).

(11) .. selle abil oskaksid need tulevikus **parema töö** leida. (kirjalik)

(12) Arvan, et sellele küsimusele võib anda **jaatava vastuse**. (kirjalik)

Teise keele kõnelejatelt eeldab täissihitise käänete sihtkeelepärane kasutus infiniit-tarindites põhjalikke süntaksialaseid teadmisi, mis teeb õige käände valiku sageli äärmiselt keeruliseks.

Kirjaliku õppijakeele materjali ekslikest nominatiivobjektidest esines infiniitse verbi laiendina pisut alla poole (41,7% e 12-st juhust 5, vt tabel 7). Kuna kõigist selle materjalikogu nominatiivobjektidest laiendas infiniitverbi 37,5% (vt tabel 6), ei ilmne siin märgatavat seost nominatiivobjektide keelelise konteksti keerukuse ja ületarvitusvigade arvu vahel. Uurimuse materjali alusel ei ole põhjust oletusteks, et juhul, kui kirjalikus materjalis olnuks kasutatud vähem infiniitseid verbivorme, oleks ka nominatiivobjektide kasutamisel tehtud vähem vigu, kuna materjalis sisaldunud vead jagunevad enam-vähem proportsionaalselt finiiit- ja infiniitver-bide laiendite vahel. Näide (13) illustreerib nominatiivobjekti ületarvitust *da*-infinitiivi laiendina, näide (14) aga pöördelise verbivormi laiendina.

(13) Aga ikkagi meie kõik püüame rohkem teenida, leida **töökoht** (pro *töökohta* v *töökohta*), kus on hea palk, sest tahame paremini elada. (kirjalik)

(14) Näen oma **vanemad** (pro *vanemaid*) väga harva. (kirjalik)

Näites (13) on kasutatud nominatiivobjekti lauses sihitisena talitleva *da*-infinitiivi laiendina, mille korral on grammatikareeglite kohaselt võimalik vaid genitiivne täissihitis (konkreetsel juhul on paralleelvariandina võimalik ka osasihitise kasuta-mine). Käesoleva uurimuse informantidelt on kogutud ka introspektiivsed andmed sihitisekasutuse kohta (vt Pool 2006b), mis annavad tunnistust sellest, et vaadelda-vad üliõpilased seostavad *da*-infinitiiviga sageli ainult nominatiivobjekti, tegemata vahet *da*-infinitiivi süntaktilisel funktsioonil lauses. Näites (14) sisalduv partitiivverb *nägema* eeldab laiendina partitiivobjekti kõigis süntaktilistes struktuurides.

Kokkuvõtlikult võib öelda, et keerukamat keelelist konteksti kirjaliku õppija-keele materjalis võib pidada vigade arvukuse mõjuriks täissihitise (eriti genitiiv-objekti) puhul. Partitiivobjekti kasutamisel ei mängi keeleline kontekst vigade seisukohast erilist rolli, kuna finiiit- ja infiniitverbe laiendavates sihitistes tehtud vigade arv oli proportsionaalselt võrreldav kõigi partitiivobjektide jagunemisega finiiit- ja infiniitverbide laienditeks.

Suulise õppijakeele ainestikus esines kirjalikuga võrreldes vähem vigu ning ka oluliselt vähem infiniitseid verbe laiendavaid sihitisi. Vastust vajab küsimus, kas vigade vähene arv võib tuleneda lihtsamast keelelisest kontekstist ehk kas on tõenäoline, et kui suulises kõnes oleks tarvitatud rohkem infiniitseid verbivorme, olnuks ka objektivigade arv vastavalt suurem.

Partitiivobjektide puhul ei mängi keelelise konteksti keerukus rolli, kuna ületarvitusvigadest vaid üks esines infiniitse verbivormi, liitpredikaadi koosseisu kuuluva *ma*-infinitiivi laiendina (15).

(15) ma pidin esitama enne eksamile minekut selle välismaalase passi aga mul seda ei olnud (.) ma sellest ei teadnud midagi et ma pean enne **sedä** (pro *selle*) muretsema (suuline)

Võrreldes tabelis 8 esitatud andmeid genitiiv- ja nominatiivobjektide ületarvitusest põhjustatud vigade jagunemisest finiiit- ja infiniitverbi laiendite vahel tabelites 5 ja 6 toodud kokkuvõtetega kõigi suulise materjali genitiiv- ja nominatiivobjektide vastavast jaotusest, on näha, et keerulisem keeleline kontekst on suulise kõne täissihitiste vigade arvu mõjutanud. Kuigi vigade üldarv on materjali suurt hulka arvestades väike, on proportsionaalselt suur osa neist tehtud infiniitse verbivormi laiendites. Nii esines tabeli 5 kohaselt kõigest suulise kõne genitiivobjektidest infiniitverbi laiendina vaid 11% (91-st 10), samas oli infiniitse verbi laiendites tehtud 50% (4-st juhust 2) genitiivobjekti ületarvitusvigadest. Kõigest nominatiivobjektidest esines tabeli 6 järgi infiniitverbi laiendina 18,5% (92-st 17), valedest kasutusjuhtudest (vt tabel 8) oli infiniitverbi laiendeid 31,8% (22 juhust 7). On tõenäoline, et kui suulises materjalis oleks kasutatud rohkem infiniitseid verbivorme, oleks see viinud ka suurema vigade arvuni. Analüüsitud materjalist ilmneb seega seos suulise õppijakeele lihtsama keelise konteksti ja väiksema vigade arvu vahel.

Näited (16) ja (17) illustreerivad genitiivobjektide ületarvitusest tingitud viga suulises materjalis: näites (16) laiendab omastavaline sihitis *et-kõrvallause da-*infinitiivset predikaati, näites (17) finiiitset verbivormi *vahetas*.

(16) keegi nagu ei usu (.) arvavad et ma andsin raha et **selle viisa** (pro *seda viisat*) saada (suuline)

(17) no talle meeldis seal rohkem ja no ta vahetas **töökoh-** (.) **töökoha** (pro *töökohta*) ja --- (suuline)

Näite (16) süntaktiline struktuur välistab genitiivse täissihitise (täissihitise käänetest oleks siin võimalik vaid nominatiiv, konkreetnes lausungis on emakeelekõneleja seisukohast loomulikum siiski partitiivobjekt). Näites (17) eeldab partitiivobjekti kasutust öeldisverbi *vahetama* ja objektnoomeni *töökoht* semantika koosmõju. Informant on kasutatud genitiivvormini jõudnud arutluse tulemusena ning läh-
 tunud tõenäoliselt tegevuse tulemuslikkusest.

Näites (18) on nominatiivobjekti ületarvitusest tingitud viga infiniitse verbivormi juures (ka *ma-*infinitiivi enda vorm ei ole korrektne), näites (19) on nominatiivne sihitis partitiivverbide hulka kuuluva finiiitverbi *näitama* laiendiks.

(18) ma käisin Soomes --- **maasikad** (pro *maasikaid*) korjama (suuline)

(19) siis tollis olin nii rõõmus kui Eesti piirivalvur ütles mulle et (.) tere tulemast Eestisse ja näidake oma **dokumendid** (pro *dokumente*) (suuline)

Siiani olid vaatluse all ainult ületarvitusest tingitud objektivead ja nende seos õppijakeele grammatilise kontekstiga. Tabelis 9 on kokkuvõtlikult esitatud finiiit- ja infiniitsete verbe laiendavate objektide sihtkeelepärane korrektsuse protsent, arvestades nii objektikäänete üle- kui ka alatarvitust (vrd tabel 3).

Tabel 9. Finiiit- ja infiniitverbe laiendavate objektide sihtkeelepärane korrektsus kirjalikus ja suulises õppijakeeles

	Finiitverbi laiendid		Infiniitverbi laiendid		
	Kirjalik	Suuline		Kirjalik	Suuline
Partitiiv	86,8%	92,3%	Partitiiv	84,8%	94%
Genitiiv	48%	86%	Genitiiv	50%	61,5%
Nominatiiv	46,4%	73%	Nominatiiv	33,3%	59%

Materjali analüüsi põhjal on võimalik väita, et suurem vigade hulk kirjaliku õppijakeele ainestikis on täissihitiste puhul mõjutatud keerulisemast keelelisest kontekstist (e infiniitsete verbivormide rohkusest), suulise õppijakeele väiksem vigade hulk aga vastavalt lihtsamast keelelisest kontekstist (e infiniitverbide väiksemast osakaalust). Samas ei anna keelelise konteksti süntaktiline keerukus või lihtsus vastust küsimusele, miks esines kirjalikus materjalis suulisest protsentuaalselt enam partitiivobjektide ületarvitusjuhtumeid ning miks on tabelis 9 välja toodud suulise objektikasutuse sihtkeelepärase korrektsuse protsent nii finiiit- kui ka infiniitverbe laiendavate sihitiste puhul suurem kui kirjalikus. Keelelise konteksti mõju täpsemaks hindamiseks on seega lisaks grammatilisele poolele tarvilik vaadelda ka õppijakeele leksikaalset külge.

4.2. Sihitise põhjaks olevate verbide mitmekesisus suulises ja kirjalikus õppijakeeles

Suulise ja kirjaliku õppijakeele sihitisekasutuse erinevuste analüüsi täpsustamiseks on mõttekas laiendada artiklis rakendatavat keelelise konteksti mõistet ning võtta lisaks lausete süntaktilisele struktuurile vaatluse alla ka sihitise põhjaks olevate verbide olemus, eriti verbivaliku mitmekesisus. Tabelis 9 esitatud andmetest on näha, et suulises materjalis oli sihtkeelepärase korrektsuse protsent kirjalikust kõrgem nii finiiit- kui ka infiniitverbide laienditena kasutatud objektide puhul. Järgnevalt püüan leida vastust küsimusele, kas suulise keelekasutuse parem ning kirjaliku vastavalt nõrgem tulemus võib olla seotud vestlus- ning kirjanditeemadest tuleneva verbivalikuga, võimaluse korral toon esile tähelepanekuid ka sagedasematest objektnoomenitest.

Suulised vestlused hõlmasid kõigi 15 informandi puhul sama teemaatikat: informantide päritolu ja kooliaeg, ülikoolis õppimisega seonduv, eredamad mälestused ülikooliajast, senised töökogemused, tulevikuplaanid jms. Informantide vestluspartnerina suunasin küll vestluse kulgu ja esitasin täiendavaid küsimusi, kuid üliõpilased olid teemaderingi piires vabad rääkima just sellest, mis neile meeldis ja millest rääkimisel nad ennast kindlalt tundsid. Seega pole välistatud mõningane vältimisstrateegiate rakendamine: tuttavatest asjadest rääkides välditakse tegelikult keerulisemaid ja elus harvemini ette tulevaid teemasid, mis omakorda võib viia lihtsama keelelise konteksti ehk lihtsamate süntaktiliste struktuuride ja piiratumate sõnavara kasutamiseni.

Kirjalik materjal pärineb eksamikirjanditest, mille teemad olid ette antud. 28 informandi tööd jagunesid paljude teemade vahel, mis käsitlesid hariduse rolli elus, raha mõju inimesele, ülikooli tähtsust kirjutaja elus, keelte rolli inimeste ühendajana, ajakirjanduse rolli ühiskonnas, religiooni küsimusi, edukuse kriteeriume jm. Kuigi informandid said kirjanditeema valida mitme variandi seast, pidid nad kogu kirjutise vältel valitud teemast kinni pidama ning kirjanditeema eeldab ka sõnavara varieerimist. Seetõttu on põhjust eeldada, et kirjalikus materjalis oli lisaks keerukamatele süntaktilistele struktuuridele kasutatud ka vaheldusrikkamat sõnavara. Eelkirjeldatud taustateadmistest lähtudes saab püstitada hüpoteesi, mille kohaselt võib suulise õppijakeele suurem sihtkeelepärane korrektsus sihitise kasutusel tuleneda sellest, et suulises kõnes kasutati sihitise põhjana vähesemaid verbe ning esines ka mitmeid väga sageli korduvaid objektnoomeneid, mille kasutamine informanti-

dele raskusi ei valmistanud, kirjalikus materjalis aga võib eeldada leksika suuremat varieerumist, mis võib olla üheks objektivigade rohkust mõjutavaks teguriks.

Materjali võrreldavuse huvides on järgnevas kokkuvõttes arvesse võetud võrdne hulk objektikasutusjuhte nii suulisest kui kirjalikust materjalikogust – kummastki 450 objekti. Tabelis 10 on ära toodud selles ainestikus sihitise põhjana kasutatud verbid esinemissageduse järjekorras, sulgudes on lisatud esinemiskordade arv 450 objektikasutusjuhust.

Tabel 10. Sihitise põhjana esinenud verbid esinemissageduse järjestuses

Suuline õppijakeel (450 näidet)	Kirjalik õppijakeel (450 näidet)
1. <i>õppima</i> (42), sh <i>ära õppima</i> (2)	1. <i>andma</i> (39), sh <i>ära andma</i> (1)
2. <i>tegema</i> (30)	2. <i>tegema</i> (26)
3. <i>rääkima</i> (29), sh <i>ära rääkima</i> (1)	3. <i>leidma</i> (24)
4. <i>võtma</i> (28), sh <i>ära võtma</i> (1)	4. <i>õppima</i> (18), sh <i>ära õppima</i> (1)
5. <i>saama</i> (23)	5.–6. <i>saama</i> (14), <i>tundma</i> (14)
6. <i>õpetama</i> (19)	7. <i>võtma</i> (10)
7. <i>andma</i> (18)	8. <i>saavutama</i> (9)
8.–9. <i>valima</i> (17), <i>kirjutama</i> (17)	9. <i>tahtma</i> (8)
10. <i>lõpetama</i> (14), sh <i>ära lõpetama</i> (2)	10.–15. <i>kasutama</i> (7), sh <i>ära kasutama</i> (1), <i>õpetama</i> (7), <i>näitama</i> (7), <i>tähendama</i> (7), <i>sooritama</i> (7), <i>teadma</i> (7)
11. <i>leidma</i> (12)	16. <i>nägema</i> (6)
12. <i>teadma</i> (11)	17.–25. <i>ümbritsema</i> (5), <i>valima</i> (5), <i>aitama</i> (5), <i>looma</i> (5), <i>hindama</i> (5), <i>vaatama</i> (5), <i>lahutama</i> (5), <i>vajama</i> (5), <i>otsima</i> (5)
13. <i>nägema</i> (8)	26.–35. <i>ootama</i> (4), <i>tooma</i> (4), <i>arendama</i> (4), <i>nimetama</i> (4), <i>lahendama</i> (4), <i>pidama</i> (4), <i>rääkima</i> (4), <i>omandama</i> (4), <i>valdama</i> (4), <i>oskama</i> (4)
14.–15. <i>panema</i> (7), <i>vahetama</i> (7), sh <i>ära vahetama</i> (1)	36.–44. <i>kuulama</i> (3), <i>realiseerima</i> (3), <i>teada saama</i> (3), <i>toetama</i> (3), <i>lugema</i> (3), <i>liitma</i> (3), <i>mõjutama</i> (3), <i>esitama</i> (3), <i>panema</i> (3)
16.–17. <i>ostma</i> (6), <i>oskama</i> (6)	45.–73. <i>mängima</i> (2), <i>eelistama</i> (2), <i>arvestama</i> (2), <i>unustama</i> (2), <i>ostma</i> (2), <i>tooma</i> (2), <i>meenutama</i> (2), <i>parandama</i> (2), <i>rakendama</i> (2), <i>veetma</i> (2), <i>hankima</i> (2), <i>ütleva</i> (2), <i>kaotama</i> (2), <i>täiendama</i> (2), <i>muutma</i> (2), <i>organiseerima</i> (2), <i>mainima</i> (2), <i>jätkama</i> (2), <i>kirjutama</i> (2), <i>moonutama</i> (2), <i>täitma</i> (2), <i>segama</i> (2), <i>hoidma</i> (2), <i>õigustama</i> (2), <i>välja mõtlema</i> (2), <i>soovima</i> (2), <i>silmas pidama</i> (2), <i>armastama</i> (2), <i>jagama</i> (2)
18.–20. <i>vaatama</i> (5), <i>sooritama</i> (5), sh <i>ära sooritama</i> (1), <i>ütleva</i> (5)	74.–142. <i>kaitsma</i> (1), <i>tõestama</i> (1), <i>rikkuma</i> (1), <i>rahuldama</i> (1), <i>kaebama</i> (1), <i>suunama</i> (1), <i>väärtustama</i> (1), <i>kulutama</i> (1), <i>kinkima</i> (1), <i>kandma</i> (1), <i>ravima</i> (1), <i>korjama</i> (1), <i>avaldama</i> (1), <i>arvama</i> (1), <i>maha jätma</i> (1), <i>jooma</i> (1), <i>arutleva</i> (1), <i>saatma</i> (1), <i>edasi andma</i> (1), <i>pakkuma</i> (1), <i>vastu võtma</i> (1), <i>ühendama</i> (1), <i>lähima</i> (1), <i>abistama</i> (1), <i>ületama</i> (1), <i>pöörama</i> (1), <i>praktiseerima</i> (1), <i>kohtama</i> (1), <i>viskama</i> (1), <i>tassima</i> (1), <i>sundima</i> (1), <i>võrdleva</i> (1), <i>maitsma</i> (1), <i>koolitama</i> (1), <i>nautima</i> (1), <i>külastama</i> (1), <i>ette kujutama</i> (1), <i>kajastama</i> (1), <i>levitama</i> (1), <i>müüma</i> (1), <i>eristama</i> (1), <i>kartma</i> (1), <i>tunnistama</i> (1), <i>tagama</i> (1), <i>pooleli jätma</i> (1), <i>päästma</i> (1), <i>petma</i> (1), <i>uskuma</i> (1), <i>tekitama</i> (1), <i>maksma</i> (1), <i>austama</i> (1), <i>omaks võtma</i> (1), <i>eraldama</i> (1),

Suuline õppijakeel (450 näidet)	Kirjalik õppijakeel (450 näidet)
	<i>väärima (1), moodustama (1), alustama (1), ümber tege- ma (1), märkama (1), proovima (1), omama (1), mõist- ma (1), puudutama (1), alla suruma (1), väljendama (1), tõstma (1), korraldama (1), säilitama (1), sünnitama (1), avastama (1)</i>
21.–25. <i>viima (4), sh ära viima (1), unusta- ma (4), sh ära unustama (2), tundma (4), küsima (4), vastu võtma (4)</i>	
26.–37. <i>näitama (3), korraldama (3), ootama (3), lugema (3), tähendama (3), kuulma (3), meenutama (3), jutustama (3), üürima (3), esitama (3), tooma (3), korja- ma (3)</i>	
38.–58. <i>koguma (2), nimetama (2), kõne- lema (2), valdama (2), mõtlema (2), tõlki- ma (2), lindistama (2), kuulama (2), joo- nistama (2), pakkuma (2), omandama (2), sööma (2), kasutama (2), välja mõtlema (2), muretsema (2), nõudma (2), armastama (2), jätma (2), saatma (2), jälgima (2), vabas- tama (2)</i>	
59.–107. <i>pidama (1), välja saatma (1), väljendama (1), kontrollima (1), jooma (1), häirima (1), ajama (1), ette kujutama (1), kasvatama (1), aitama (1), edasi and- ma (1), tahtma (1), keelama (1), valmistama (1), kokku panema (1), moonutama (1), sünnitama (1), asendama (1), alustama (1), kordama (1), huvitama (1), jagama (1), avaldama (1), kohtama (1), rakenda- ma (1), äratama (1), lisama (1), paluma (1), sundima (1), välja töötama (1), viskama (1), tellima (1), tähistama (1), proovile panema (1), maksma (1), alla kirjutama (1), ületama (1), pöörama (1), välja andma (1), küüditama (1), segamini ajama (1), välja tõmbama (1), lehitsema (1), avama (1), lubama (1), arvama (1), kandma (1), esile tooma (1)</i>	

Tabelist 10 on näha, et verbivalikut puudutav hüpotees pidas paika: suulises materjalikogus sihitise põhjana tarvitatud erinevate verbide hulk on palju väiksem (suulises 107, kirjalikus 142 eri verbi). Tähelepanu väärib enamasti partitiivobjekti võtvate verbide *õppima*, *rääkima* ja *õpetama* suur hulk suulises materjalis. Nende verbide näol pole küll tegemist puhtalt partitiivverbidega, kuid ilma resultatiivsete laienditeta tarvitatakse neid enamasti koos partitiivobjektiga. Vestlustes oli neid kasutatud valdavalt partitiivi eeldavates kontekstides ja piiratud hulga objektnoomenitega, nt *õppima keelt*, *rääkima keelt*, *õpetama keelt*, nagu näidetes (20)–(22).

(20) sest ma hakkasin õppima **eesti keelt** esimesest klassist (suuline)

(21) ainuke kes meil **eesti keelt** räägib on vanaema (suuline)

- (22) aga siis avati meil eesti koolis selline eksperimentaalklass kus võeti vene-
lasi vastu ja õpetati **eesti keelt** nendele (suuline)

Esitatud näited illustreerivad kontekste, mille puhul võib kõrgtasemel eesti keele kõnelejatel objekti vormivalikut pidada juba automaatseks ning neid öeldisverbe võib pidada sihitisekasutusel lihtsamaks keeleliseks kontekstiks. Viimast väidet kinnitab ka asjaolu, et nimetatud verbe laiendavate objektide hulgas esines vaid üks viga, vt näidet (23), kus partitiivi asemel oli kasutatud nominatiivi.

- (23) ja põhimõtteliselt saab öelda et nagu **vormiõpetus** (pro *vormiõpetust*)
nagu ise kodus õppisin ja (suuline)

Kirjaliku materjali hulgas olid kolm sagedasimat verbi aspektverbide hulka kuuluvad *andma*, *tegema* ja *leidma*, mida võib pidada sihitisekasutusel keerulisemaks keeleliseks kontekstiks, kuna igal üksikjuhtumil on vaja mõelda sihitise vormivaliku tingimustele. *Tegema* oli nii kirjalikus kui suulises ainekis esinemissageduselt teisel kohal, põhjuseks on ilmselt selle verbi kuulumine eesti keele tuumverbide hulka (Tragel 2003), millel on keeles väga laialdased kasutamismõimalused. Kirjalikus õppijakeeles kõige sagedamini sihitise põhjana esinenud verb *andma* osutus objektikasutuse seisukohast veaohhtlikuks: partitiivi ületarvitusest tingitud viga kirjalikus materjalis oli üldse kokku 40, millest 8 oli just ühendis *andma*-verbiga, nagu näidetes (24) ja (25).

- (24) Tartu ülikool annab üliõpilasele **head võimalust** (pro *hea võimaluse*)
ennast igakülgsest arendada. (kirjalik)

- (25) Ning vahel juhtub nii, et fakte moonutatakse, antakse nendele **teistsu-
gust värvingut** (pro *teistsugune värving*) ja spekulereeritakse nendega.
(kirjalik)

Andma-verbi nii suur esinemissagedus võib tuleneda ühelt poolt sellest, et tege-
mist on põhisõnavarasse kuuluva tuumverbiga (Tragel 2003), kuid määravam on
kindlasti kirjandite teemavalik: mitme kirjandi teemaks oli “Mida on mulle andnud
ülikool?”.

Kirjalikus ainekis sageduselt kolmandal kohal olev verb *leidma* pole samuti
keeleõppijatele sihitise põhjana lihtne (seda laiendavate sihitiste hulgas oli kokku
neli viga). Vaadeldavas materjalis paistab silma informantide soov kasutada seda
verbi sageli just infiniitvormis, mis muudab keelelise konteksti veelgi keerulisemaks.
Probleemid tekivad eelkõige täissihitise käände valikuga, nagu näites (26), milles
lause süntaktiline struktuur võimaldab täissihitise käändena vaid nominatiivi tarvi-
tamist, paralleelvariandina aga kindlasti (isegi eelistatumalt) ka partitiivobjekti.

- (26) Et nendele **vastuse** (pro *vastus* v *vastust*) leida, tuleb mõelda selle peale,
kust religioon üldse tuleb. (kirjalik)

Nii kirjalikus kui ka suulises ainekis sageduselt teisel kohal olev tuumverbide
hulka kuuluv *tegema* on aspektverbina põhjustanud nominatiivobjekti ületarvitust,
mida illustreerivad näited (27) ja (28).

- (27) Ülikool tegi minust **täiskasvanud inimene** (pro *täiskasvanud inimese*),
kes teab, mida tahab. (kirjalik)

- (28) mina üritan **see** (pro *sedā*) teha nii eesti keelest vene keelde nagu õpetajatele --- kui ka vene keelest eesti keelde kui saadetakse mingi kirju või käskkirju või (.) (suuline)

Sihitise põhjana kasutatud verbide väiksem valik suulises õppijakeeles lihtsustab keelelist konteksti märgatavalt ning verbikasutus avaldab mõju ka objektnoomenite valikule. Õppijakeele leksika vaheldusrikkus on seega grammatilise korrektsuse saavutamist mõjutav tegur.

5. Kokkuvõtteks

Kahest eri keelekasutussituatsioonist kogutud õppijakeele ainekogu analüüs näitas, et täis- ja osasihitise kasutamise sihtkeelepärane korrektsus varieerub vastavalt õppijakeele situatsioonilisele kontekstile: vabamat stiili esindavas suulises materjalis oli korrektsuse protsent tunduvalt suurem kui ettevaatlikumat ning suurema enesekontrolliga stiili esindavas kirjalikus materjalis. Suulise õppijakeele sihtkeele seisukohast paremini õnnestunud sihitisekasutus on seotud lihtsama keelelise kontekstiga: suulises aineses esines kirjalikuga võrreldes vähem infiniitse verbivormi laiendina kasutatud objekte ning sihitise põhjana kasutatud verbide valik on väiksem, hõlmates mitmeid väga suure esinemissagedusega verbe, millega seoses kasutati piiratud hulka objektnoomeneid. Kirjaliku ainese suurem objektiivigade hulk suuliseiga võrreldes on samuti seotud nii lausete süntaktilise struktuuriga (keerulisemad tarindid) kui ka kirjandžanri eripärast tuleneva sõnavaralise mitmekesisusega: sihitiste põhjana esinevate verbide valik on suulise materjalikoguga võrreldes märgatavalt suurem. Just verbide mitmekesisus ning asjaolu, et neist sagedasimate hulka kuulub mitmeid aspektverbe, on vigade rohkuse seisukohast olulise tähtsusega: niisugustes tingimustes on objektikäände automaatse kasutuse võimalus väike (eriti võrreldes suulise materjali sagedasimate verbidega, nt *rääkima* + objekt partitiivis) ning keeleõppijad on sunnitud pidevalt mõtlema objekti vormi valikut määravatele semantilistele ning süntaktilistele tingimustele.

Viidatud kirjandus

- Ellis, Rod 1990. *Understanding Second Language Acquisition*. Sixth impression. Oxford: Oxford University Press.
- Ellis, Rod; Barkhuizen, Gary 2005. *Analyzing Learner Language*. Oxford: Oxford University Press.
- Fjodorova, Alina 2004. Vene üliõpilaste sõnavarakasutus suulises ja kirjalikus eesti keeles. Bakalaureusetöö. Tartu Ülikooli eesti keele (võörkeelena) osakond.
- Kaivapalu, Annekatrin 2005. Lähdekieli kielenoppimisen apuna. *Jyväskylä Studies in Humanities*, 44. Jyväskylä: Jyväskylän yliopisto.
- Kaivapalu, Annekatrin 2008. Mida räägivad sõnasisesed pausid ja kordused sihtkeele käänamisprotsessist? – Eesti Rakenduslingvistika Ühingu aastaraamat, 4, 35–52.
- Larsen-Freeman, Diane; Long, Michael H. 1991. *An Introduction to Second Language Acquisition Research*. London, New York: Longman.
- Pajupuu, Hille; Kerge, Krista; Alp, Pilvi 2009. Sõnavara loomulik rikkus haritud keeleoskaja tekstides. – Eesti Rakenduslingvistika Ühingu aastaraamat, 5, 187–196.

- Pastuhhova, Olga 2004. Interferentsivead vene üliõpilaste suulise ja kirjaliku eesti keele süntaksis. Bakalaureusetöö. Tartu Ülikooli eesti keele (võõrkeelena) osakond.
- Pica, Teresa 1984. Methods of morpheme quantification: their effect on the interpretation of second language data. – *Studies in Second Language Acquisition*, 6, 69–78.
- Pool, Raili 2005. Täis- ja osasihitise kõrgtasemel eesti keelt teise keelena kõnelejate kirjalikus keelekasutuses. – Raili Pool (toim.). *Teine keel. Uurimusi eesti keele kui teise keele omandamisest*. Tartu Ülikooli eesti keele (võõrkeelena) õppetooli toimetised, 5. Tartu: Tartu Ülikooli Kirjastus, 8–74.
- Pool, Raili 2006a. Sihitise vormihomonüümia – keeleuurija probleem, keeleõppija tugi. – Külvi Pruuli, Annekatrin Kaivapalu (toim.). *Lähivertailuja 17*. Jyväskylä Studies in Humanities, 53. Jyväskylä: Jyväskylän yliopisto, 85–101.
- Pool, Raili 2006b. Täis- ja osasihitise kasutamisest eesti keelt teise keelena õppijate pilgu läbi. – *Eesti Rakenduslingvistika Ühingu aastaraamat*, 2, 185–202.
- Pool, Raili 2007a. Keeleõppijate vältimisstrateegiatest eesti keele täis- ja osasihitise näitel. – *Eesti Rakenduslingvistika Ühingu aastaraamat*, 3, 235–252.
- Pool, Raili 2007b. Täis- ja osasihitise omandamise ning markeerituse seosest. – *Keel ja Kirjandus*, 1, 35–50.
- Tarone, Elaine 1983. On the variability of interlanguage systems. – *Applied Linguistics*, 4/2, 143–163.
- Tarone, Elaine 1988. *Variation in Interlanguage*. London: Edward Arnold.
- Tarone, Elaine; Yule, Georg 1991. *Focus on the Language Learner. Second impression*. Oxford: Oxford University Press.
- Tragel, Ilona 2003. *Eesti keele tuumverbid*. *Dissertationes linguisticae Universitatis Tartuensis*, 3. Tartu: Tartu Ülikooli Kirjastus.

Raili Pooli (Tartu Ülikool) peamiseks uurimisvaldkonnaks on eesti keele kui teise keele omandamine ja kasutus.
 raili.pool@ut.ee

CONTEXTUAL VARIATION IN LEARNER LANGUAGE: THE USE OF TOTAL AND PARTIAL OBJECTS

Raili Pool

University of Tartu

The article discusses the use of total and partial objects by speakers of Estonian as a second language in spontaneous speech and writing. The aim of the article is to identify differences in spoken and written use of objects and explain these differences by analysing the linguistic context. The informants of the study were students of University of Tartu whose mother tongue is Russian and who possess an advanced knowledge of Estonian (their proficiency in the language corresponds to the proficiency level C1). Written language data consisted of the informants' examination essays (the essays contained 581 instances of object use by 28 students) which, because of the significance of the examination, reflect a careful use of language by the informants. The spoken language data, which demonstrate a freer use of language by the learners, contained informal conversations between informants and the author of the article (conversations with 15 students totalled 882 instances of object use).

Analysis of the data showed that situational context influences the grammaticality of learner language: the percentage of correct use of total and partial objects was significantly higher in spontaneous speech data than in written data (written data: partitive 87.1%, genitive 45.2%, nominative 38.5%; spoken data: partitive 92.8%, genitive 81.3%, nominative 70%).

The linguistic context in learner language was taken into consideration when differences in the use of total and partial objects in speech and in writing were analysed. The article examines the occurrences of total and partial objects functioning as complements of finite and non-finite verb forms and the variety of verbs that subcategorise for the object. The analysis of the data shows that the higher percentage of grammaticality in the use of objects in spoken learner language is partly due to its simpler linguistic contexts: speech data contained considerably fewer objects that functioned as complements of non-finite verb forms and had a far smaller variety of transitive verbs subcategorising for the object than in the written data. The written data demonstrated more complex linguistic contexts (there were more non-finite constructions and a greater variety of verbs, including many aspectual verbs), which reflects the requirements of the essay genre and the variation in essay titles.

Keywords: situational context, grammaticality, linguistic context, syntax, vocabulary, Russian, Estonian

RUSSIFICATION OF NON-ESTONIAN PUPILS IN TALLINN

Mart Rannut, Ülle Rannut

Abstract. In the article home languages, their qualities and use by pupils in grades 2 to 5 in Tallinn are analyzed. Two larger and more characteristic groups (Ukrainians and Belarusians) and two groups with Islamic backgrounds (Azerbaijanis and Tatars) were compared with the aim of analysing different language loss and preservation factors. Although Ukrainians and Belarusians form much larger ethnic groups within Estonia, they have not managed to preserve their language anywhere near as successfully as the Azerbaijanis and Tatars. Comparison of the data suggests that language preservation in Estonia and the ability to withstand Russification are most influenced by the degree of difference of the language from Russian and the religious difference from the Russian Orthodox faith (and Soviet era atheism). Home languages are characterised by their use mainly in the verbal form, with skills in the literary language generally low or even non-existent. Languages are for the most part learned informally, primarily through communication with parents and/or grandparents. The result is a supersaturated vernacular with a large degree of variation and loans from superstratum languages which falls far from the standard of the language in question.

Keywords: language education policy, minorities, home language, Estonian, Russian

1. Introduction

Language planning in education in Estonia to date has seen broader treatment of its formal structure (M. Rannut, Ü. Rannut 1995, Ü. Rannut 2003, M. Rannut 2006, Vare 2002a, 2002b, EMER 2008) with less attention paid to the non-formal environment and the attitudes stemming from it (Ü. Rannut 2005, M. Rannut 2008, M. Rannut, Ü. Rannut 2007, Vihalemm, Masso 2003, Kempainen, Ferrin 2002).

This article primarily focuses on this field, investigating the home languages of pupils from Tallinn and examining a variety of factors which influence the language use of students with different ethnic backgrounds and the formation of their linguistic attitudes. Studied in closer detail are pupils from countries which formed part of the former Soviet Union and the home languages they speak, in which the influence of Russification and the displacement of minority languages continue to be observed. The data for the study was collected via questionnaires from 8340 students aged 8 to 12. 34 schools with Estonian as the language of instruction and 19 schools with Russian as the language of instruction took part in the study, which covered all of the districts of the city (M. Rannut, Ü. Rannut 2007).

Home language (cf. Extra, Yagmur 2004) means the language used most frequently or regularly at home by the person being interviewed at the time of the study. As such, a person's home language does not necessarily have to be the same as the language of their community (their everyday language; the one which covers the wider communication network they form part of) or their native language (which could be the language of the country of origin or another language with which they identify themselves). A wide range of other factors was also taken into consideration in the survey. Particular emphasis was given to different aspects of the home languages of pupils from former Soviet countries, where Russification and the loss of ethnic languages of origin are a continuing process. This has not taken place in a linguistically neutral environment, but in conditions of a powerful external influence – in this case a language policy consciously designed to achieve this result – which constitutes the notion of 'Russification'. One of the aims of the study is therefore to describe Russification processes, which continue unabated; while another, as a counterbalance to them, is to highlight the need for and measures of a linguistic support environment and the need for a minority policy in order to ensure the preservation of language diversity and to make languages less endangered.

Russification as a form of linguistic imperialism refers to both an official and unofficial ethnic and language policy which was implemented (and continues to be implemented) by Tsarist Russia, the Soviet Union and today's Russian Federation against national minorities with the aim of russifying other peoples and making them speak Russian.

The direct influence of Russification on Estonians in Estonia during the Soviet occupation was modest; in the census of 1989 just 1% of Estonians nominated Russian as their native language. The reason for this was first and foremost the well-functioning education and language maintenance system, which operated despite strong pressure from Russification policy, and which maintained the high status of Estonian (M. Rannut 1994, 2001, 2004). Nevertheless, other peoples and languages (besides Estonians and Russians) were smothered by Russification, as the Soviet authorities did not allow them to set up any schools with instruction in their own languages or to officially teach the next generation their ethnic languages. (In the Republic of Estonia in the pre-war years this was not only permitted, but promoted by the state.) The result of it was that Russian became the only written language and the primary spoken language among immigrants, especially among Belarusians and Ukrainians as speakers of closely related languages and among mixed marriage families.

The demographic aspect of Russification in Estonia is very clear: although the census (Statistical Office of Estonia 2000) lists representatives of 142 different ethnic groups claiming 109 different native languages between them, there are only two linguistically and ethnically homogenous groups in Estonia: Estonians, who comprise the majority speaking the titular language; and Russians. This article will not examine Estonia's two largest ethnic groups: indigenous Estonians, who largely managed to preserve their language and culture even during the Soviet occupation; and Russians, who benefited most both linguistically and culturally from the conditions of the period (their numbers increasing twelve-fold between 1944 and 1989, and the number of Russian speakers increasing almost 20 times; see Katus, Sakkeus 1992). The third largest group, Ukrainians, can be more or less divided into two groups based on language: those who have retained Ukrainian as their native language (41%) and those who have adopted Russian as their first language (57%). The situation is even worse among Belarusians, where only 29% consider Belarusian as their native language and 70% speak Russian as their first language. The dominance of Russian can also be seen among other ethnic groups. It can therefore be said that ethnic identity does not necessarily depend on home language, as this is influenced by a number of factors, including the prestige of the languages spoken around people and the ideologies related to and attitudes towards these languages.

2. Home language of pupils in Tallinn

Estonia's pupils have 30 different home languages, comprising 28% of all of the native languages spoken in Estonia (EHIS, data for 2003). The data collected as part of this study indicates that pupils from the 2nd to the 5th grades in Estonian- and Russian-language schools in Tallinn speak 22 different languages as home languages, which represents 20% of the total number of languages spoken in Estonia according to the Statistical Office and indicates severe reduction in the number of languages spoken by younger generations. The two main languages spoken at home by pupils are Estonian and Russian. Estonian was spoken as the only home language by 55% of the pupils surveyed, with 4.9% speaking Estonian at home in addition to another language. These figures are in line with the proportion of Estonians in Tallinn (54.9%; Statistical Office 2007). Russian is spoken as the only home language by 39.6%, with a further 4.6% speaking Russian at home in addition to another language. This figure is significantly larger than the proportion of ethnic Russians in Tallinn (36.5%), indicating Russification of other ethnic groups. 96% of the pupils who speak only Russian at home were born in Estonia; 3% were born in Russia and the remainder in Ukraine (16), Belarus (7), Latvia (7) and other former Soviet countries. 66% of the mothers and 63% of the fathers of pupils who only speak Russian at home were born in Estonia; 21% of parents are from Russia, 6% from Ukraine and the remainder from former Soviet countries.

The widespread use of Russian as the only home language is a sign of extensive linguistic segregation. Only around one in every ten users of Russian as a home language is to some extent bilingual, though only a marginal proportion of them were born outside Estonia. Two thirds of the parents of Russian-speaking pupils

were born in Estonia. This means that legally they can no longer be viewed as immigrants or the family members of immigrants, even though their linguistic behaviour continues to indicate otherwise. It is therefore possible in Estonia, where Estonian is the only official language, to preserve Russian as the only home language for generations, which highlights the weakness of Estonia's language and integration policy to date. Just 2% of pupils speak a language other than Estonian or Russian at home. The most widely used languages in these cases are English, Azerbaijani, Ukrainian, Finnish, Armenian, Italian, Lithuanian, Latvian, Tatar, Swedish and German. Other home languages include Belarusian, Spanish, French, Romanian, Turkish, Bashkir, Georgian, Hebrew, Korean, Hungarian, Setu, Arabic, Chinese, Croatian, Portuguese and a sign language.

The study revealed that 39.6% of pupils from the respective cohort in Tallinn speak Russian as their only home language and 4.6% as a second language. Just 3% of them and 21% of their parents were born in a republic of the Russian Federation (in this case Chechnya, Tatarstan, Mordovia, Komi, Bashkortostan, Northern Ossetia-Alania and Adygeya) where national languages remain spoken to different extents. Interviews with 224 Russian-speaking pupils revealed that of those who considered themselves to be Russian, one-third had grandparents with a non-Russian background, incorporating very different native languages and national identity. All other ethnic groups showed a declining tendency between generations. In the case of parents with Belarusian origins, 7.2% of grandparents considered themselves to be Belarusian, but just 2.9% of the pupils' parents identified themselves in the same way. There were also several times more grandparents with Ukrainian origins (6.5%) than parents of pupils who identified themselves as Ukrainian (just 1.4%). This represents a 60% drop in Belarusians between generations and a 78% drop in Ukrainians from grandparents to parents. Tatars declined by 33%.

Below we examine some of the larger and more characteristic groups from former Soviet countries in our home language study: Ukrainians, Belarusians, Azerbaijanis and Tatars. What leaves Ukrainians and Belarusians open to particular influence is the fact that both of their languages are part of the Eastern Slavic group of languages and therefore closely related to Russian. They are also close to Russians in terms of predominantly belonging to the Russian Orthodox Church. The second group looks at the predominantly Islamic Azerbaijanis and Tatars, whose languages both form part of the Turkic group of languages, and who are therefore much more removed from Russians in terms of religion and language.

2.1. Ukrainians

The Ukrainians who immigrated to Estonia during the Soviet era were already markedly Russified. Of those living in the country in 1989, 65% (30,253) spoke Ukrainian while 94.2% (45,472) could speak Russian (Issakov, Tõnurist 1999a: 482–483). The restoration of independence in Estonia saw a fall in the number of Ukrainians by more than a third, although to date they still form the third largest ethnic group in the country: 28,158 Ukrainians, of whom 51% live in Tallinn (Statistical Office of Estonia 2000). 38% of those living in Tallinn claim Ukrainian as their native language, compared to 41% throughout the country; while 61% of those living in Tallinn name Russian as their native language, compared to 57% throughout the

country. According to the Statistical Office, just 108 Ukrainians in Tallinn and 481 in Estonia overall consider Estonian to be their native language.

The number of students with Ukrainian origins in the study is 433. There are only five families in which both the parents and children were born in Ukraine, while there are 35 cases of both parents being born in Ukraine but the children being born in Estonia or elsewhere outside of Ukraine. The majority of families (382 or 88%) constitute mixed marriages where only one of the parents was born in Ukraine. There are also five families in the study where only the child was born in Ukraine, and families where only the grandparents are from the country. Mixed families in which only one of the parents was born in Ukraine are the most common and that they mostly speak Russian (83% Russian only; 9% Estonian and Russian; and 3% Russian and Ukrainian). However, Russian is also dominant in families where both parents were born in Ukraine: the study revealed that 91% of parents born in Ukraine and 50% of new immigrants with Ukrainian origins only speak to their children in Russian.

The study shows that fathers use Ukrainian more than mothers when communicating with their children. 1.2% of mothers and 2.1% of fathers with Ukrainian origins used Ukrainian at home as the only language of communication, while 2% of pupils used it as a second language with their mothers and 3% with their fathers. At the same time, the number of mothers with Ukrainian origins is higher than that of fathers (244 and 216 respectively). Women can therefore be seen to be more easily assimilated than men, as confirmed in previous interviews (Ü. Rannut 2005: 112–113) and studies (Dorian 1981, Gal 1979), which show that the main figure of influence in the determination of home language is the mother, who also leads the family in its use.

81% of families use Russian alone as their home language. A total of only four families (0.9%) use Ukrainian as their first, i.e. home language, while 22 (5.1%) use it as a second language. There were a total of 26 pupils in the study who speak Ukrainian as their home language, but 36 parents (14 mothers and 22 fathers). Just 7% of parents born in Ukraine speak Ukrainian at home, which shows the remarkable loss of the language among first generation immigrants and their assimilation into Russian-speaking society. 75% of male and female grandparents use Russian alone when communicating with their grandchildren, which shows that language loss has generally already occurred in previous generations. 36 grandmothers (9%) and 39 grandfathers (9%) used Ukrainian exclusively when speaking to their grandchildren. A further 22 grandparents (2.5%) use Ukrainian as a second language. It is notable that a lot more pupils use Ukrainian when speaking to their grandparents than they do when speaking to their parents (81 pupils or 18.7% and 26 pupils or 6% respectively). A total of 66 pupils use Ukrainian alone when communicating with their grandmothers or grandfathers, 7 of whom use Ukrainian when speaking to both their grandmother and grandfather. A further 15 pupils use Ukrainian as a second language when communicating with one or both of their grandparents.

The use of Ukrainian decreases from generation to generation. Just 4 pupils (0.9%) speak Ukrainian exclusively with their brothers and sisters, while 7 pupils (1.7%) use it as a second language. That these pupils use Russian as a home language can also be seen in the fact that of those questioned, 360 of them study in schools with Russian as the language of instruction and 68 in schools with Estonian as the

teaching language. The influence of the language of instruction of the school can also be seen on the children's choice of language in communicating with one another. The study shows that the use of Estonian has risen among children studying at schools with Estonian as the language of instruction. Half of the sixty-eight pupils with Ukrainian origins studying in schools with Estonian as the teaching language use Estonian as a first or second language with their brothers and sisters, with 28% of pupils using Estonian alone, 32% using Russian alone and none using Ukrainian. Just ten of the 360 pupils with Ukrainian origins studying in schools with Russian as the language of instruction (3%) speak Estonian as a first or second language. The majority of pupils (84%) use Russian exclusively at home with their brothers and sisters, while 7 pupils (2%) use Ukrainian. The teaching language of the school clearly influences linguistic attitudes: the study revealed that just 25 of the 433 pupils with Ukrainian origins prefer to speak Ukrainian, whereas 298 nominated Russian as their preferred language, 69 chose Estonian and 46 said English was their favourite.

The influence of a school's language of instruction on the language of communication between children has been substantiated in previous interviews with and studies into parents and children with Ukrainian origins, where Russian had become the first language of the children at home but the influence of Estonian-language schools had also seen the children speaking to each other in Estonian at home (Ü. Rannut 2003: 109–110). Those speaking to their brothers and sisters in Ukrainian were mostly found among new immigrants, where the pupils had older brothers and sisters, both parents were of Ukrainian origin and the whole family used Ukrainian exclusively as their language of communication (Ü. Rannut 2005: 111).

The main factors influencing use of language among Ukrainians are the time of migration (i.e. how many generations of immigrants are involved), mixed families and linguistic attitudes. Interviews with mixed Russian/Ukrainian families in Estonia have shown that the use of Ukrainian as a home language has decreased over the years, as little value is placed on it. As a result, Ukrainian language skills have also suffered: while older children can still speak the language to some extent, younger children no longer understand it.

The study revealed that 19.2% (83) of pupils with Ukrainian origins communicate in Ukrainian; 18.7% (81) pupils use Ukrainian when speaking to their grandparents; 6% (26) use it when speaking to their parents; and 2.5% (11) use it when speaking to their brothers and sisters. The decline in the use of Ukrainian from one generation to the next is clear to see. Moreover, use of Ukrainian does not necessarily mean competence in Ukrainian. 83 pupils who use Ukrainian at home or when communicating with their grandparents were asked to give an assessment of their Ukrainian language skills. Their responses reflected a downward trend in spoken and written skills: spoken Ukrainian is understood most (and best), with 69% of pupils responding to this positively, 29% understanding some and a couple of pupils (2%) understanding none. Two pupils who claimed not to understand Ukrainian fell into the category of those who are spoken to in Ukrainian by their grandparents but who themselves use Russian when speaking to their grandparents. 57% of pupils said they could speak Ukrainian; 37% that they could speak some Ukrainian; and 6% claimed not to be able to speak any Ukrainian at all, even

though they are spoken to at home or by their grandparents in Ukrainian. 52% of pupils said they could read Ukrainian, while 33% said they could also write in the language. 20% claimed they could not read the language at all and 31% claimed they could not write in it. Such results in reading and writing are in fact noteworthy, as no pupils claimed to be learning Ukrainian either at school or elsewhere.

2.2. Belarusian

As with Ukrainians, Belarusians mostly began to emigrate to Estonia after World War II. By 1959 their numbers in the country had reached 10,030. After thirty years of Soviet occupation this figure had almost tripled reaching 27,711 by 1989 (Issakov, Tõnurist 1999b: 510). Belarusian was the native language of 8841 (32%) of them and the second language for a further 4703, while Russian was the native language of 18,591 (67%) and the second language of 8283, thus revealing a remarkable amount of Russification. Belarusian as a language held very little status, being seen as little more than a dialect and poor cousin of Russian, and it therefore never became a literary language. The Soviet authorities were intolerant of Belarusian culture: those who championed the language and culture were arrested, with some of them sent to hard labour camps, where they died, and some of them exiled. The wave of migration after World War II was the result of Russification policy, which redirected workers of different nationalities to occupied territories where their common language would be Russian. Already weak, the ethnic identity of the Belarusians diminished further under the influence of this forced intermingling and with Russian as the *lingua franca*, with the result that every consecutive census has seen fewer and fewer people nominating Belarusian as their native language. Just 27% of the Belarusians in Estonia considered Belarusian to be their native language in the 2000 census. Since the only schools that existed in Soviet times alongside Estonian-language schools were those with Russian as the language of instruction, this deepened the extent of the assimilation of Belarusians into Russian-speaking society and the acceptance of Soviet identity, as a consequence of which they began to class themselves as Russians.

After Estonia regained its independence there was a new national awakening among Belarusians and recognition of the need for them to distinguish themselves from the rest of the Russian-speaking population. As a result, many Belarusians left the country; their numbers in Estonia dropped by more than a third, and of the 16,133 that remain, 7593 (47%) live in Tallinn. Although a mere 27% of Estonia's Belarusians consider Belarusian to be their native language and 70% speak Russian as a first language, more Belarusians are starting to place greater value on their language and culture (Statistical Office of Estonia 2000).

There were 154 pupils (or one of their parents) in this study who were born in Belarus, plus 2 pupils who were born (along with their parents) outside of Belarus, but who speak to their grandparents in Belarusian. Only 3 pupils in this group (1.9%) use Belarusian in addition to Russian and Estonian at home when communicating with their parents. Russian is predominantly used, with Estonian also used in individual cases, resulting from the choice of school language or the native language of the parents. 86.5% of families only speak Russian at home,

while 15 pupils (9.6%) use Russian as a second language or as the main language of communication, which brings the total number of speakers of Russian as a home language in this group to 150 pupils (96.1%). It is remarkable to note that only fathers communicate in Belarusian with their children; that in only two cases is Belarusian used exclusively; and that in one case it is only used as a second language after Russian. Russian dominates as the language of communication with both parents and grandparents in 87% of cases. Just 22 pupils communicate with their parents and grandparents in Belarusian. No pupils use Belarusian when speaking to their brothers and sisters, which is only to be expected given that there are no families in the group in which Belarusian is exclusively spoken. 80% of pupils (125) only use Russian with their brothers and sisters, while 7 pupils (4%) use Estonian and the remaining 6% use Estonian or other languages (English, German) in addition to Russian. It can therefore be claimed that 86% of pupils speak to their brothers and sisters at home in Russian.

As with the Ukrainian group there was also a marked decline in language skills, with spoken skills faring better than written skills. 65% of the twenty pupils who use Belarusian understand it and are able to speak it, while 60% claimed also to be able to read it and 50% to write in it. None of the 156 pupils studies Belarusian at school or outside of school, and only 8 of them (5%) expressed an interest in learning it. It can therefore be seen that the use of Belarusian among Estonia's Belarusian community is virtually non-existent, as reflected in the pupils' assessments of their language skills.

In summary, while Belarusians comprise the fourth largest ethnic group in Estonia and 46% (7938) of them live in Tallinn, Belarusian as a language has basically died out among pupils in the city. Just 22 of the 156 pupils with Belarusian origins involved in the study (14%) use Belarusian, primarily when communicating with their grandparents. No pupils speak to their brothers or sisters in Belarusian, since only 3 parents use the language as one of their home languages.

2.3. Azerbaijanis

The third group is Azerbaijanis, who do not form a large ethnic group within Estonia according to Statistical Office figures (2001) but who have nevertheless managed to preserve their language with much greater success than the Ukrainians (see also Mamedov 1999). 1238 Azerbaijanis had come to live in the country by 1989, of whom 869 (70.2%) spoke Azerbaijani as their native language. 342 spoke Russian as a first language and 686 as a second language, meaning a total of 83% of Azerbaijanis used Russian in a lingua franca function. The independence gained by Azerbaijan in 1991 upon the collapse of the Soviet Union and the establishment of Azerbaijani as the only official language in the country of 8 million people has undoubtedly contributed to the preservation of their language today. Another factor influencing preservation both in Azerbaijan and among diaspora communities is religion: whereas the Russian Orthodox faith as one of the most international denominations of Christianity is widespread among Russians, Islam is the religion of 96% of all Azerbaijanis (80% of them Shi'a and 20% Sunni) and forms a major part of their national identity. Census data from 2001 had 880 Azerbaijanis living in Estonia, 57% of them (505) in Tallinn. 69% (347) spoke Azerbaijani, 29% (145)

Russian and 8% (4) Estonian as their first language (Statistical Office of Estonia 2000). Compared to the Ukrainians and Belarusians, this is a group which has been much more successful in preserving its native language.

In our study there were 48 pupils in Tallinn with Azerbaijani origins. In five cases both the pupil and his or her parents were born in Azerbaijan (10%); 13 pupils' parents are both from Azerbaijan (27%); and 28 pupils (58%) are part of mixed marriage families where only one of the parents is from Azerbaijan. 13% of the families speak Azerbaijani as their only home language, with a further 21% speaking it as a second language (16 pupils in total). 16 pupils (33.3%) use Azerbaijani when communicating with their parents, of whom 6 (12.5%) use it as their only home language. The language is used to speak to 18 fathers (38% of fathers) and 14 mothers (29.2% of mothers).

Azerbaijani is the sole home language only of those pupils whose parents were both born in Azerbaijan and who, for the most part, were also born there themselves. These people are therefore new immigrants who value their native language. Estonian is spoken as the only home language by 2 pupils and by a further 2 pupils as a second language (8% in total), as reflected in the choice of school teaching languages. Russian is most dominant in mixed marriage families, where 75% of pupils (21) use Russian as the only home language when communicating with their parents and 14% (4) use it as a second language in communicating at home.

Regardless of the fact that their native language has been comparatively well preserved, Azerbaijanis exhibit the same tendency as Ukrainians in that more pupils speak Azerbaijani with their grandparents than with their parents. A total of 20 pupils (42%) use Azerbaijani to communicate with their grandparents; 13 use the language when speaking to both their grandmothers and grandfathers. It should nevertheless be noted here that the difference between generations is not as great as was observed with the previous groups.

Unlike the groups examined earlier, there is also more communication with brothers and sisters in Azerbaijani. 14.6% of pupils (7) use Azerbaijani as a second language and 1 pupil uses it as the only language of communication with their siblings. However, more than half of the pupils in this group too (27, 56%) use Russian, and 10.4% (5) speak with their brothers and sisters in Estonian, because of their school language.

As with Ukrainians and similar to previous study results (Ü. Rannut 2005), the choice of language used with brothers and sisters is primarily influenced by the language of instruction of the pupils' schools and the resulting language that they use with their peers both at school and outside of school. The study showed that ten of the eleven Azerbaijani pupils studying in Estonian-language schools speak Estonian exclusively with their classmates; outside of school, however, Russian dominates. A similar tendency could be seen in Russian-language schools, where 34 of the 37 pupils use only Russian in school, and 29 pupils also use it as their sole language of communication outside of school.

20 pupils who use Azerbaijani when speaking to their parents or grandparents were asked to give an assessment of their Azerbaijani language skills. As with the Ukrainian pupils, their responses reflected a downward trend in spoken and written skills: 85% of the pupils understood Azerbaijani and 80% of them could speak the language, outstripping the same figures for the Ukrainian pupils (69% of whom claimed to understand the language and 57% of whom said they could speak it).

None of the pupils who use Azerbaijani as a home language claimed that they could not speak the language, although one did claim that they could not understand it. 45% of pupils said they could read and write in Azerbaijani, while 30% said they could not read the language at all and 35% said they could not write in it. The fact that some pupils who speak Azerbaijani as their sole home language cannot read in their native language when they start school has also emerged in interviews with both pupils and teachers (M. Rannut, Ü. Rannut 2003).

2.4. Tatars

Tatars form the sixth largest ethnic group in Estonia, and like the Azerbaijanis are largely Islamic, different therefore from both Estonians and Russians in terms of religion. Nevertheless, Statistical Office data has only a third of Estonia's Tatars as Muslims, with 8% subscribing to the Russian Orthodox faith and the majority (58%) claiming no religion, likely as a result of the atheist propaganda of the Soviet Union. 404 of the 754 Tatar Muslims in Estonia live in Tallinn. The Estonian Islamic Congregation was re-registered in 1989.

The Estonian Islamic Congregation and the Tatar Cultural Society initially overlapped in terms of membership. The Islamic Congregation was restructured in 1994 and set itself the goal of bringing together and organising the activities of all of the Muslims living in Estonia. Among the members of the congregation are people from Azerbaijan, Uzbekistan, Kyrgyzstan, Kazakhstan and other Islamic countries (Estonian Bureau for Lesser-Used Languages (EstBLUL)¹). After World War II the number of Tatars in Estonia grew rapidly, having risen to 4058 by 1989 (Ahmetov, Nisamedtinov 1999: 450). After Estonia regained its independence many of them left the country, and by 2001 their number had dropped by almost half to 2582, 1265 of whom (49%) live in Tallinn. A further 245 live nearby in Maardu. According to the data of the Statistical Office, 50% of the Tatars in Estonia speak Russian and 47.6% speak Tatar as their first language. They are therefore ranked higher in terms of their native language than Ukrainians and Belarusians, but lower than Azerbaijanis.

A total of just 13 pupils with Tatar origins were found as part of this study, which is surprisingly few, given the size of the ethnic group in Estonia. None of the pupils were themselves born in Tatarstan; as an exception, they were all born in Estonia. There were also no families where both parents were from Tatarstan, which shows that there are no new immigrants to be found among the group and that mixed marriage families predominate.

From interviews it could be concluded that there may in fact have been more pupils with Tatar origins, as some parents who were born elsewhere in Russia (there are more ethnic Tatars outside of Tatarstan than there are in Tatarstan) consider themselves to be ethnic Tatars but have adopted Russian as the language they use amongst themselves.

5 pupils (38.5%) use Tatar as one home language when communicating with their parents, while 8 only use Russian at home. The same number of pupils (5) use Tatar when communicating with their grandparents, but three of the five pupils use Tatar in this case as their only language of communication. It can therefore also

¹ Tatar language, <http://www.estblul.ee/ENG/Languages/tatar.html> (20.02.2010).

be said of this group that more Tatar is used when communicating with grandparents.

A total of six pupils (46.2%) speak Tatar as their only language or as a second language when communicating with their parents or grandparents. Numbers-wise this would seem to indicate that the Tatar-language community is diminishing, but in terms of percentages they have in fact managed to maintain their native language better than the other groups examined in this study. The only cause for concern is that Tatar is not used as the sole home language by any family, reflecting its displacement by Russian and the rapid assimilation which took place among the previous generation as a result of Russification and enforced atheism. Here it is important to note the status factor which operated in Russian-speaking environments: the Tatars were particularly badly affected by Russian chauvinism, associated (unfoundedly) with the Mongols of history, as a result of which many attempted to conceal their ethnic identity and, if possible, amended their 'passport nationality' to Russian. In the hierarchy of religious status Islam was also poorly placed and connected with an even greater lack of education and financial retardation than the Russian Orthodox faith. At the same time, these attitudes never existed in Estonian-speaking society.

As with the other ethnic groups, the national language is used as a home language least often in communication between brothers and sisters, which is dominated by Russian. Only one pupil also uses Tatar. This is undoubtedly influenced by the language of instruction of the pupils' schools, since 69% of the pupils go to Russian-language schools. It should also be noted though that those students who go to schools with Estonian as the language of instruction (31%) also use Russian for home communication, with again one pupil also using Tatar. The impression gained from comparing home languages and school languages was that children are only placed in Estonian-language schools if Estonian is one of the languages used at home. Since Russian is usually also added to this, it is difficult to determine whether it is a case of the native language of one of the parents or simply use of Estonian as a foreign language related to the children's studies, since Estonian clearly enjoys high status in Estonian-language schools, influencing the attitudes of both children and their parents.

All 6 pupils who use Tatar when communicating with their parents or grandparents claimed to be able to speak and understand the language. However, in terms of reading and writing skills the study showed a similar decreasing tendency to that of other groups. One pupil nevertheless claimed to be able to both read and write in Tatar, while another claimed to be able to do so a little.

3. Discussion and summary

The Soviet occupation of Estonia brought with it an all-encompassing policy of Russification whose consequences are continuing to have a negative impact on the sustainability of different national groups. Two larger and more characteristic groups (Ukrainians and Belarusians) and two groups with Islamic backgrounds (Azerbaijanis and Tatars) were compared with the aim of analysing different language loss and preservation factors.

The social status of the language in the country or state of origin and that country or state's political independence from Russia only affects the group internally, thus producing differences between the Ukrainians and Belarusians and between the Azerbaijanis and the Tatars. It is for this reason that (for example) the Ukrainians, who enjoy greater political independence from Russia and have one national language, have managed to preserve their native language more successfully than the Belarusians, whose country is bilingual and markedly more politically dependent on Russia. Pupils from monolingual Azerbaijan have also maintained their language more effectively than those from Tatarstan, which is bilingual and forms a part of the Russian Federation.

What leaves Ukrainians and Belarusians open to particular influence is the fact that both of their languages are part of the Eastern Slavic group of languages and therefore closely related to Russian. The similarity is increased further by the verbal use of the option used as the home language, which is generally far from the literary standard, with Belarusian/Russian and Ukrainian/Russian patois (known as *Trasyanka* and *Surzhyk*, respectively) used in its place. They are also close to Russians in terms of predominantly belonging to the Russian Orthodox church. The lack and uncertainty of specific barriers preserving nationality, but also the lack of the national languages in a number of functions (such as teaching and literary languages), has enabled Russian to become dominant among these ethnic groups and for their native languages to be replaced. In assessing Estonia's Ukrainians it could be said that their language has all but died out under pressure from Russification and that for new life to be breathed into it and for further Russification to be prevented home language classes and groups must be established, especially in the case of such a large group of people. The situation with Belarusian is even direr, having practically disappeared within its own community: just 22 of the 156 pupils with Belarusian origins (14,1%) use the language, mostly when speaking with grandparents and at a level of quality far removed from the literary norm. The study also revealed that no pupils with Belarusian origins speak to their brothers and sisters nor to their friends in Belarusian. Since 88% of the pupils go to Russian-language schools, it is natural that 83% communicate exclusively in Russian at school and outside of school, with the remainder speaking a mixture of Russian and Estonian.

As the preservation of Belarusian is even problematic in Belarus due to strong pressure from Russification (and despite the fact that with an effective language policy it would have more than 10 million speakers!), none of the much-needed support for the language is coming from that direction either. In contrast, the flexing of political muscle that is going on between Ukraine and Russia at the moment and the fact that the Soviet era remained an insufficient period of time to truly russify the Ukrainians means that things are quite different for this ethnic group: national and linguistic identity is being reinforced. The extent to which these nationalist sentiments are reaching Estonia's Ukrainian population is debatable: the results of the 2007 parliamentary elections in Ukraine show that the majority of the Ukrainian community here supports its nation's Russian-minded party. This is a signal of the crossing of a critical line in views, where dominant attitudes are adopted among speakers of an encroaching language even if skills in their own language have been preserved to some extent (Fishman 2001). Given the number

of pupils with Ukrainian or Belarusian origins and their tendency to live in the same northern Estonian towns and cities and districts of Tallinn, it would make sense to establish classes in support of these groups' languages and cultures. To date this has been very limited, despite a number of events, mainly because parents continue to prefer Russian-language education for their children. Even among the pupils themselves the level of interest in learning the language of their country of origin is modest at best.

The second group examined as part of the study are the Turkic language-speaking and predominantly Islamic Azerbaijanis and Tatars. The objective in studying these groups was to enable an analysis of the linguistic and cultural assimilation due to Russification of ethnic groups religiously and linguistically different from Estonians and Russians and an analysis of the current situation of their languages and the opportunities that exist for preserving them. Slightly fewer than half of the Tatars in Estonia (47.6%) use Tatar as their first language. They are therefore ranked higher in terms of their native language than Ukrainians and Belarusians, but lower than Azerbaijanis, who are mostly new immigrants. Nevertheless, 6 pupils (38.5%) of the children with Tatar origins speak Tatar as a first or second language and mostly when communicating with their grandparents. There are no opportunities for learning Tatar at present in Tallinn. The Sunday school activities which have been relaunched are insufficient. Organising proper education in the language is being hindered by the lack of qualified teachers, but no less by the fact that the Tatar alphabet was changed a number of times during the Soviet era, the Arabic alphabet being replaced in the 1920s with the Latin alphabet and it being replaced in the 1930s with the Cyrillic alphabet due to Russification. The impression gained from the interviews with Estonia's Tatars was that although the parents use Russian as the only language of communication with their children, they place high value on Tatar and identify themselves with it.

Unlike the Tatars, the Azerbaijanis have their own country of 8 million people and their own national language, which have undoubtedly contributed greatly to the preservation of the language today. However, in spite of their native language remaining well preserved, a loss of Azerbaijani can still be seen to be occurring between generations. The language is used as a first or second home language by 41.7% of children when speaking with their grandparents, by 33.3% when speaking with mothers and fathers, and by 8 pupils (17.7%) when speaking to siblings. Mixed marriages are a significant factor in the loss of Azerbaijani, with 75% of pupils (21) using Russian as their only home language when communicating with their parents and 14% (4) using it as a second language in communicating at home. Compared to the Tatars, the main difference would appear to lie in the time of emigration more than the extent of Russification. Azerbaijani is the sole home language only of those pupils whose parents were both born in Azerbaijan and who, for the most part, were also born there themselves.

The results of the study highlight the glaring differences between the two groups: first, Ukrainians and Belarusians; secondly, Azerbaijanis and Tatars. Although Ukrainians and Belarusians form much larger ethnic groups within Estonia, they have not managed to preserve their language anywhere near as successfully as the Azerbaijanis and Tatars. Comparison of the data suggests that language preservation in Estonia and the ability to withstand Russification are most influenced by

the degree of difference of the language from Russian and the religious difference from the Russian Orthodox faith (and Soviet era atheism).

Also having a negative influence on the preservation of the languages among the younger generation (pupils) in all groups is the language of instruction in schools and the lack of opportunities for the pupils to learn their national languages either at school or outside of it.

Estonia's regaining of its independence has led to a rise in national self-awareness and through it a newfound respect for national language. Despite the different statuses of Russian, the de-Russification trend is the same for all compared to Soviet times. Both the drop in status of Russian in the countries of origin and the favourable attitude of Estonia's language policy to the preservation of languages with small numbers of speakers have changed the linguistic attitudes of new immigrants with different native languages. Interviews with Russian-speaking parents (Ü. Rannut 2005) show that representatives of those national groups who migrated to Estonia during the Soviet era have begun to place higher value on their language and culture, even if their nation remains part of the Russian Federation, and have started to use their native languages as their home languages wherever possible.

Home languages are characterised by their use mainly in the verbal form, with skills in the literary language generally low or even non-existent. Interviews revealed that in preschool preparations parents focus wholly on the language of instruction of the future school (for example Estonian), as a result of which teaching their children to read and write in their native language is not considered imperative. Languages are for the most part learned informally, primarily through communication with parents and/or grandparents. The result is a supersaturated vernacular with a large degree of variation and loans from superstratum languages which falls far from the standard of the language in question.

The member of the family wielding the greatest influence in the determination of home language is the mother, who also uses the language predominantly in the family; the main factor in the determination of the language is the relationship network (function) and status (for a general theoretical treatment see Dorian 1981, Gal 1979). Women can therefore be seen to be more easily assimilated than men, as confirmed in previous interviews (Ü. Rannut 2005: 112–113).

Compared to Western Europe, the extent of mixed marriages between nationalities stands out, often taking in more than three-quarters of the national groups. The reasons behind this are various and mostly connected to the nature of the Soviet system. One was the lack of private ownership and choice of where to live, which did not necessarily allow people to live with people from their own area and which was designed to create city districts with national flavours. The lack of private and state schools (apart from Russian and Estonian ones) and national societies also shepherded young people with different national backgrounds into the same, Russian-language environment of communication and way of thinking, which produced favourable conditions for mixed marriages, which in turn promoted the use of Russian. At the same time, the number of mixed marriages involving Estonians was marginal, accounting for just 6 to 8% of all marriages.

The influence of Russification in the case of the home languages examined in this study can be seen most prominently in the lack of availability of formal education in people's native languages, which may still allow them to use them as

spoken languages, but not in their status as literary languages. (There is little point in discussing native language language technology resources in the case of former Soviet non-Russian autochthonous peoples.) Other key factors are the time at which the people emigrated, the extent to which they did so, the size of the group, how dispersed they are, the religious barrier in the case of Islam and respect for it (including for the tradition of bringing a wife from the person's home village in their country of origin or from within their tribe), the negative attitude of Russians towards specific national groups (forcing people to either conceal or flaunt their origins or ethnic identity) and the level of prestige of languages. Arriving in Estonia from somewhere outside of their ethnic territory is itself a sign of people have adopted another language, while mixed marriages are also a catalyst for adopting another language.

The data collected during the study shows that Estonian schools have an important role to play in changing and shaping people's attitudes and values so that solutions other than just a transition to Russian, to date seen as unavoidable, can be found in the homes of pupils with languages other than Estonian in the choice of which language to use at home. Estonian society is also different in terms of its greater ethnic and linguistic tolerance, as a result of which a number of groups here who were persecuted under the Soviet Union (such as the Tatars and people from the Caucasus) no longer feel ashamed of their native languages or ethnic identities. This supports the use and sustainability of a range of home languages.

Studying the use of language and linguistic attitudes of pupils as the younger generation and the connections that exist between them and the use of language of older generations also enables us to assess Estonia's future multicultural development. Such forecasts are generally pessimistic, predicting that the majority of languages will die out completely over the next couple of generations. Customising the current education system to take the needs of home languages into account would create the required support and give new hope for the preservation of linguistic and cultural diversity in Estonia.

References

- Ahmetov, Dajan; Nisamedtinov, Ramil 1999. Tatarlased. – Jüri Viikberg (toim.). Eesti rahvaste raamat. Rahvusvahemused, -rühmad ja -killud. Tallinn: Eesti Entsüklopeediakirjastus, 449–452.
- Dorian, Nancy 1981. *Language Death: The Language Cycle of a Scottish Gaelic Dialect*. Philadelphia: Univ. Penn. Press.
- EHIS = Eesti Hariduse Infosüsteem [= National Registry of Estonian Education (with restricted access); former Andmevara]. <http://www.ehis.ee> (20.02.2010).
- Estonian Bureau for Lesser-Used Languages (EstBLUL). Tatar language, <http://www.estblul.ee/ENG/Languages/tatar.html> (20.02.2010).
- Extra, Guus; Yagmur, Kutlay (Eds.) 2004. *Urban Multilingualism in Europe. Immigrant Minority Languages at Home and School*. Clevedon: Multilingual Matters.
- Gal, Susan 1978. Peasant men can't get wives: Language change and sex roles in a bilingual community. – *Language in Society*, 7, 1–16. doi:10.1017/S0047404500005303
- Issakov, Sergei; Tõnurist, Igor 1999a. Ukrainlased. – Jüri Viikberg (toim.). Eesti rahvaste raamat. Rahvusvahemused, -rühmad ja -killud. Tallinn: Eesti Entsüklopeediakirjastus, 479–495.

- Issakov, Sergei; Tõnurist, Igor 1999b. Valgevenelased. – Jüri Viikberg (toim.). Eesti rahvaste raamat. Rahvusvahemused, -rühmad ja -killud. Tallinn: Eesti Entsüklopeediakirjastus, 508–515.
- Katus, Kalev; Sakkeus, Luule 1992. Foreign-born Population in Estonia. Rahvastiku-uuringud. Seeria B, 19. Tallinn: Eesti Kõrgkoolidevaheline Demouuringute Keskus.
- Kemppainen, Raija; Ferrin, Scott Ellis 2002. Parental choice and language-of-instruction policies and practices in Estonia. – *Education and Urban Society*, 35 (1), 76–99. doi:10.1177/001312402237215
- Mamedov, Vidadi 1999. Aserbaidžaanid. – Jüri Viikberg (toim.). Eesti rahvaste raamat. Rahvusvahemused, -rühmad ja -killud. Tallinn: Eesti Entsüklopeediakirjastus, 33–36.
- Rannut, Mart 1994. Beyond linguistic policy: the Soviet Union versus Estonia. – Tove Skutnabb-Kangas, Robert Phillipson (Eds.) in collaboration with Mart Rannut. *Linguistic Human Rights: Overcoming Linguistic Discrimination*. Berlin/New York: Mouton de Gruyter, 179–208.
- Rannut, Mart 2001. Eesti kirjakeelest ja selle kvaliteedist. – Sirje Mäearu, Peeter Päll (toim.). Keelekorralduse konverents 18.-19. novembril 1999, ettekanded. Tallinn: Eesti Keele Sihtasutus, 75–106.
- Rannut, Mart 2004. Language Planning in Estonia: Past and Present. Working Paper 16. CIEMEN/Mercator. <http://www.ciemer.org/mercator/pdf/wp16-def-ang.PDF> (1.03.2010).
- Rannut, Mart 2006. The role of education in reversing language shift in Estonia. – John Kirk (Ed.). *Belfast Studies in Language, Culture and Politics*. Belfast: Queen's University.
- Rannut, Mart; Rannut, Ülle 1995. Bilingualism – a step towards monolingualism or multilingualism? – Tove Skutnabb-Kangas (Ed.). *Multilingualism for All. European Studies on Multilingualism*, 4. Lisse: Swets and Zeitlinger, 183–198.
- Rannut, Ülle 2003. Muukeelsete õpilaste integreerimine eesti koolis. Õpik kõrgkoolile ning muukeelsete õpilastega töötavale aine- ja klassiõpetajale. Tallinn: TPÜ Kirjastus.
- Rannut, Ülle 2005. Keelekeskkonna mõju vene õpilaste eesti keele arengule ja integratsioonile Eestis. Tallinna Ülikool. Humanitaarteaduste dissertatsioonid, 14. Tallinn: TLÜ kirjastus.
- Rannut, Ülle; Rannut, Mart 2007. Tallinna õpilaste kodukeeled. – *Haridus*, 3–4, 7–10.
- Statistical Office of Estonia 2000 = Statistikaamet 2000. Statistika andmebaas: rahvaloendus 2000. Tallinn: EV Statistikaamet. <http://pub.stat.ee/px-web.2001/Database/Rahvaloendus/databasetree.asp> (20.02.2010).
- Vare, Silvi 2002a. Eesti keel vene koolis I. Eesti keele õpetamine vene üldhariduskoolis. <http://www.hm.ee/index.php?popup=download&id=3972> (20.02.2010).
- Vare, Silvi 2002b. Eesti keel vene koolis II. Eesti keele kasutamine vene üldhariduskoolis. <http://www.hm.ee/index.php?popup=download&id=3971> (20.02.2010).
- Vihalemm, Triin; Masso, Ann 2003. Identity dynamics of Russian-speakers of Estonia in transition period. – *Journal of Baltic Studies*, 34 (1), 92–116. doi:10.1080/01629770200000271

Mart Rannut (Tallinn University). His research interests cover: language policy and planning, linguistic imperialism, second language acquisition.
mart.rannut@tlu.ee

Ülle Rannut (Tallinn University). Her research interests cover several domains in sociolinguistics: language education policy, adaptation policies for immigrants and their children, second language acquisition issues, cross-cultural communication.
yrannut@tlu.ee

MITTE-EESTI ÕPILASTE VENESTUMINE TALLINNAS

Mart Rannut, Ülle Rannut

Tallinna Ülikool

Artiklis analüüsitakse Tallinna II–V klassi õpilaste kodukeeli, nende kvaliteeti ja kasutust. Võrreldakse kahte suurimat rühma (ukrainlased ja valgevenelased) ja kahte islami taustaga rühma (aserbaidžaanid ja tatarlased), et analüüsida keele säilimise ja kao faktoreid. Kuigi ukrainlased ja valgevenelased on arvuliselt palju suuremad, ei ole nad suutnud säilitada oma keelt nii edukalt kui aserbaidžaanid ja tatarlased. Andmete võrdlus näitab, et keele säilimist ja vastuseisu venestumisele on võimaldanud vene keelest erinevuse määr ja usuline erinevus vene õigeusust ning Nõukogude aja ateismist. Kodukeeli iseloomustab peamiselt suuline kasutus, samas kui kirjaoskus neis on tagasihoidlik või puudub täiesti. Keeli õpitakse tavaliselt mitteformaalselt, põhiliselt vanemate ja vanavanematega suheldes. Tulemuseks on teistest keeltest mõjutatud tugevalt varieeruv ning laenudega küllastunud kõnekeel, mis on kaugel vastavast kirjakeelest.

Võtmesõnad: hariduse keelepoliitika, rahvusrühmad, kodukeel, eesti keel, vene keel

LANGUAGE & AUTHORITARIANISM IN THE 20TH CENTURY: THE CASES OF ESTONIA AND CATALONIA

Delaney Michael Skerrett

Abstract. The 20th century saw the Soviet and Francoist regimes enforce their respective ideologies in Estonia and the Catalan-speaking territories in Spain. In both cases, the autochthonous language suffered under the stringent control of the mechanisms of censorship and repression. In fact, Soviet and Spanish leaders—representing both extremes of the political spectrum—tried to replace the use of the autochthonous language in many sociolinguistic domains with their own—Russian and (Castilian) Spanish—as these languages embodied the ideologies and the new orders that they wished to establish. This paper compares and contrasts the diverse methods of control over language carried out in Estonia and the Catalan-speaking areas of Spain in order to demonstrate that highly centralised multilingual states—whatever their political ideology—can make use of surprisingly similar means of control, ultimately depriving local linguistic communities of the ability to use and develop their own language.*

Keywords: diglossia, languages in contact, language policy, language planning, endangered languages, Estonian, Russian, Catalan, Castilian Spanish, Spanish

1. Introduction

The Soviet and Francoist regimes are two salient examples of attempts by 20th century authoritarian states to impose imperialistic linguistic policy as part of their political ideology. In both Soviet Estonia and the Catalan-speaking territories in Francoist Spain, the autochthonous language suffered under the stringent control of the mechanisms of censorship and repression. What is more, the leaders—from the extreme right in the Spanish case and the extreme left in the Soviet case—

* Acknowledgements. An earlier version of this paper was written in Spanish as a thesis at The University of Queensland. The thesis was supervised by Prof. Alfredo Martínez Expósito. I am extremely grateful for Prof. Martínez Expósito's insight and encouragement. I would also like to thank Miquel Strubell i Trueta for his support with that project.

attempted to replace the autochthonous variety in use in areas under their control in many domains with their own language: Russian in Estonia and Castilian Spanish in Spain. These languages embodied the ideologies and the new orders that they wished to establish. When authoritarianism came to an end and democracy was re-established, the sociolinguistic situation in these territories was one of diglossia, in which the vitality of Estonian and Catalan was threatened by the languages promoted by the Soviet Communists and the Spanish Fascists, respectively. This paper compares and contrasts the diverse methods of control over language carried out in Estonia and Catalonia in order to demonstrate that highly centralised multilingual states—whatever their political ideology—can make use of surprisingly similar means of control, ultimately depriving local linguistic communities of the ability to use and develop their own language.

2. Authoritarian repression: The Francoists and the Soviets

2.1. Catalan and the Francoist ideology

According to Hooper (1995: 413), Catalan suffered more under the Franco regime (1939–1975) than in any prior period of the history of the language, the fascists carrying out “a campaign [---] unparalleled in the region’s history”. Catalonia’s autonomy ended with the region’s occupation by Francoist forces (McRoberts 2001). In fact, the region itself was erased and it was divided into four provinces ruled directly by Madrid (Fishman 1991, May 2001). The leaders of Catalonia were killed or exiled (Keating 2001, Wright 1998). The Catalan language officially ceased to exist and was declared a “mere dialect” (Conversi 1997: 114, Fishman 1991: 297, Labanyi 1995, May, 2001: 242) of Castilian Spanish, “making use of a concept of ‘dialect’ belonging more to the 18th century than the 20th”¹ (Moreno Fernández 2005: 228). With regard to use of their language, Catalan-speakers were told “we are in Spain,”² “don’t bark like a dog,”³ and “speak in Christian”⁴ (Ager et al. 1995, Azevedo 2005, Fishman 1991: 297, Mar-Molinero 2000, May 2001: 242, Strubell i Trueta (personal communication 2006), Sánchez & Dueñas 2002). According to Moreno Fernández (2005: 228) “[i]f you are Spanish, speak the language of the Empire”⁵ was a catchphrase of the regime. This pressure to speak Castilian Spanish can be seen as part of Franco’s “glorious national Movement”⁶ which had as one of its goals the substitution of Catalan with Castilian Spanish, “the language of the Empire”⁷ (Conversi 1997, May 2001: 242; McRoberts 2001: 41, Strubell 1998: 156). For Franco, Castilian Spanish was a symbol of “la Patria” (“the Homeland”) and a singular Spanish national identity and the intention was thus to limit or indeed eliminate the other languages of the state (Mar-Molinero 2000: 84, Richards 1996, Shohamy 2006, Stewart 2002, Strubell i Trueta, Ros i Garcia 1984, Strubell 1998).

Conversi (1997: 109–110) maintains that “Francoism adopted the most radical politics of assimilation against non-Castilian cultures in modern Spanish history”, according to which it was the responsibility of everyone to speak Castilian Spanish

¹ “utilizando un concepto de ‘dialecto’ más propio del siglo XVIII que del XX”

² “estamos en España”

³ “no ladres como un perro”

⁴ “habla en cristiano”

⁵ “[s]i eres español, habla la lengua del Imperio”

⁶ “glorioso Movimiento nacional”

⁷ “la lengua del Imperio”

fluently, the language of “Hispanidad” (see also Richards 1996). As Wright (1998: 187, see also 2004) suggests, “[c]oercion and persecution were employed to achieve the goal of a strong, unitary state with a single language, a single history and a sense of a single tradition”. The—now famous—declaration of the right-wing intellectual and Minister of Finance, José Calvo Sotelo, “Spain: better red than dead”⁸ reflects the major importance of—indeed, the obsession with—unity on the part of the Francoists (Conversi 1997: 109–110, Strubell 1998: 156). Their ideology of the united nation, furthermore, was imposed repressively in areas, such as Catalonia, that had been more regionalist (as well as more socialist (Wright 2004), evident in the term “red-separatist”⁹ (Conversi 1997: 110)). Richards (1996: 150) maintains that the regime constructed a binary ideological division between “Spain” and “anti-Spain,” according to which the only legitimate identity was the *españolista*, the patriotic one. Franco appropriated the symbols of 15th century Spain, the period of Ferdinand and Isabella: patriotic Spaniards (Castilian Spanish-speakers) were thus taking part in a crusade on behalf of “Christian civilisation,” triumphing over “the malignant foreign powers” (Richards 1996: 150). The celebration of the discovery of the Americas as a Spanish victory over the *non-Spanish* and the *non-Christians* symbolised the Francoist victory over the separatist and non-Castilian peoples of Spain in general, or in other words, those “politically and morally poisoned elements” (Richards 1996: 158).

As was also the case under Soviet communism, praise was bestowed upon the social class considered to be the bearer of the ideological values of the new nation and, in this case, it was the Castilian farmers who possessed the true Spanish character (Richards 1996). Also similar to the Soviet case, myths were created so that people would believe in a new order—a “holy” empire and a common past—establishing the necessary conditions to believe in the positive future (Richards 1996). The communists were of course atheists but the communist ideology assumed a religious quality (Lauristin, Vihalemm 1993). Militarism was highly valued: Franco established parallels between his “glorious triumph” and the historic victories of Felipe II and el Cid (who, significantly, had conquered Valencia, a Spanish region where a variety of Catalan is spoken); victory parades re-enacted those of the heroes of the Middle Ages (like el Cid); school texts set up a direct link between Franco and these historic victors (Richards 1996). At the same time, nonetheless, lies were propagated to establish the belief that, ultimately, the regime was peaceful in nature: the message of the 1959 documentary film, *El camino de la paz (The Path of Peace)*, was that the Francoists, by way of the Civil War, had saved the Spanish people from the chaos of the previous times (Ellwood 1995).

The system of political propaganda, in both the Francoist and the Soviet cases, required the censorship and manipulation of the linguistic domains already in existence. In Spain, all books written in Catalan found in bookshops, public and private libraries, and publishing houses were destroyed (Hooper 1995, Labanyi 1995, Wright 1998). Pompeu Fabra’s “priceless collection” (Hooper 1995: 413) of books was burnt in the streets—Fabra was a respected engineer and linguist who had established the norms for the restandardisation of the Catalan in 1913 (George 2002, Mar-Molinero 1996, Moldoveanu, 2002, Strubell 1996). In total, hundreds of thousands of books were destroyed, particularly in Barcelona (Conversi 1997, Strubell 1998). Publications, posters, and advertising in Catalan were banned,

⁸ “España, antes roja que rota”

⁹ “rojo separatista”

and all existing material was removed by the Fascists as part of their “hunt for the separatist” (Conversi 1997: 111); organisations that did not comply with the new regulations were closed and individuals could be (and were) dismissed, arrested, jailed, or fined (Conversi 1997, Fishman 1991, Mar-Molinero 2000, Wright 1998, 2004). Private correspondence was censored and often destroyed (McRoberts 2001). Censorship of the postal services lasted until 1948 (Conversi 1997). Other symbols of the local culture, besides language, were also done away with: the hymn and the flag of Catalonia were declared illegal and patriotic monuments and statues were destroyed (Conversi 1997, McRoberts 2001).

The names of cities and towns were Castilianised; the street of the Virgin of Montserrat (the patron saint of Catalonia) was renamed the street of the Redeemer and the Library of Catalonia became the Central Library (Hooper 1995, Labanyi 1995, May 2001). The use of Catalan was banned in all public spheres: in administration, commerce (including on labels), the press, on the radio, and (later) on television, as well as in education and in the Church; even private conversations between people in Catalan were considered dangerous (Ager et al. 1995, Conversi 1997, Fishman 1991, Hooper 1995, Keating 2001, Labanyi 1995, May 2001, McRoberts 2001, Sánchez, Dueñas 2002, Spolsky 2004, Strubell i Trueta, Ros i Garcia 1984, Strubell 1996, 1998, Wright 2004). “Barbarisms”—foreign borrowings into Castilian Spanish—were declared illegal (Labanyi 1995). All films had to be dubbed into Castilian Spanish: “stripped of any trace of dialects”¹⁰ (Labanyi 1995, Mar-Molinero 2000, Moreno Fernández 2005: 228). In the Basque Country, tomb stones written in Basque had to be replaced by the family of the deceased (Labanyi 1995). In terms of the language use of the Church, officially the Catalan Catholic hierarchy was in agreement with Francoist language policy, a position which “was in contradiction with the pastoral practices of the bilingual areas”¹¹ of Spain, as the use of the autochthonous language “had been the norm since the 16th century”¹² (Moreno Fernández 2005: 228). Although anti-Catalan bishops were sent to Tarragona, Lleida, Barcelona, and Tortosa, local priests never fully supported the dictatorship and the Church, in fact, would become an important factor in the opposition movement (Conversi 1997).

2.2. Estonian and the Soviet ideology

According to Raun (1985), the control over language in Soviet Estonia (1939–1941 and 1944–1991) contrasted sharply with the relatively liberal approach of the inter-war independence period (1918–1940). Although prescriptive grammars and dictionaries were not a new phenomenon, “[m]ore so than in the past the various dictionaries of correct usage [were] looked upon as the final word in linguistic matters, and through its central control of all publication *the state [exerted] a powerful influence on the written language*” (Raun 1985: 22, my emphasis).

Thus literature, fine arts (art was to serve an educative, not aesthetic purpose), education, research, theatre, cinema, exhibitions, advertising, all publicly accessible printed matter (down to public transport timetables, confectionary wrappers, tickets, postcards, and calendars), and even statistics were all targets for the propaganda machine and censorship (Lauk 1999, Lauristin, Vihalemm 1993: 180): “every social activity acquired symbolic meaning related to communist ideology.

¹⁰ “prescindiéndose, en absoluto, de los dialectos”

¹¹ “entr[ó] en contradicción con la práctica lingüística de la acción pastoral en los territorios bilingües”

¹² “había sido palmario en Cataluña desde el siglo XVI”

Penetrating the minutest [sic] routines of everyday life, this ideology came to possess a religious character in Soviet society". The Soviet ideology thus formed the basis for control of all modes of expression: as all activities had meaning, these had to be regulated so that people would know how to behave in an appropriate manner, and the media in general, and journalism in particular, were to play a key role in educating the masses (Lauristin, Vihalemm 1993). As Lauk (1999: 19) states, "[t]he media were used for creating an alternative reality, 'an ideologically correct symbolic environment, filled with content designed to socialize the audience to the ideas and values of Communism'". Officially, however, there was no censorship in the Soviet Union, and the constitution of the Estonian SSR, for example, guaranteed freedoms of speech and press to all citizens, with the provision that these freedoms be used for "the consolidation and development of the socialist order" (Põhiseadus, as cited in Lauk 1999). It was not until 1986 that Gorbachev admitted in an interview with a French communist paper that censorship existed in the Soviet Union (Lauristin et al. 1993).

Myths were also created to lead people to believe in the superior status of the Russian language. Clachar (1998: 108; see also Hogan-Brun et al. 2007) states that, beginning in the mid-1950s, strong pressure was placed on the non-Russian nationalities of the Soviet Union to become bilingual, and emphasis on Russian as the language of "international communication" was stepped up from the late 1950s; Khrushchev, and later Brezhnev, could be seen to be distancing Soviet ideology from the earlier "nationalist in form, socialist in content" thesis, which held some hope for the promotion of the titular languages, and the Russian language emerged as "one of [the regime's] strongest hallmarks". Khrushchev introduced new Soviet ideological catch phrases such as "[Russian as] second national language", "jazyk mezhnatsional'nogo obshchenija" ("language of international communication"), "dobrovol'no prinjatyj" ("voluntarily adopted"), "obshchij leksicheskiy" ("common lexical fund"), and "vzaimoobogashchenie jazykov" ("mutual enrichment of languages") (Clachar 1998: 108), all to be achieved via "sblizhenie i slijanie natsii" ("convergence and fusion of peoples") (Knowles 1990: 150), with the result being "ravnopravie no ne vseгда ravnostennost" ("equal rights but not always an equal share of esteem") (Knowles 1990: 154).

In 1978, the USSR Council of Ministers sent an unpublished decree of "recommendations" to the education ministers of the individual republics; the aim of the instructions was to implement measures to increase the use of Russian in public affairs and in education (Misiunas, Taagepera 1993, Ozolins 2000). The decree gave absolute primacy, more explicitly than ever before, to the Russian language: as Clachar (1998: 109–110) describes, it "went beyond the lingua franca" and was declared "the only means of active participation in social life". It set out language policy in various stages to "place social and psychological pressure on Estonians [and] other nationalities of the Soviet Union [...] to acquire Russian". These stages ranged from an initial monolingualism in the national tongue all the way through to complete bilingualism and, finally, post-bilingualism, in which Russian was to effectively become the "new native language".

Although Estonian continued to be used in less strategic domains by native speakers and did not completely disappear from public use as Catalan did, Russian was nonetheless to pervade further into cultural life: its usage was increased

on radio and television, and even stipulated for increase in informal activities such as amateur theatre groups (Misiunas, Taagepera 1993). The increasing popularity of television of the time (a Latvian study found people were spending seven out of 20 weekly leisure hours watching television) only helped to further entrench Russification: in Estonia in 1975–76, 74% of broadcasting was in Russian (Misiunas, Taagepera 1993). By 1980, out of an average of 35.3 hours of television shown on various channels per day in Estonia, 29.3 hours were in Russian (83%) (Nørgaard et al. 1996). At the 1980 Estonian song festival, all official speeches were made, for the first time ever, in Russian (Misiunas, Taagepera 1993).

Censorship of existing literature played an important role in the Soviet system (Lauk 1999). Schools, bookstores, and libraries were instructed to remove pages from non-fiction and fiction books alike (or get rid of the books altogether) that were considered to be inconsistent with Soviet ideology (especially those by “nationalist” authors) (Misiunas, Taagepera 1993). 70,000 books were burnt in the Tartu University library alone (Clachar 1998); in the first year of Soviet occupation, 200,000 copies of Estonian books were destroyed; during the years 1944–48, almost half a million books and periodicals were burned; from 1947, school libraries were cleansed over a five-year period, destroying over 25,000 books annually: estimates range from ten to twenty million copies of printed matter in total were disposed of in this “cleaning up” of libraries and bookstores, a process which continued until the 1960s (Lauk 1999: 21). By that time, more than 86% of all the books as well as all the periodicals published in independent Estonia were included on the lists of prohibited literature (Lauk 1999). Lists of banned Soviet literature also appeared and these books, too, were destroyed; books in foreign languages (for example Chinese and Japanese) were also destroyed as the acceptability of their contents could not be verified, given that the censors could not read them (Lauk 1999).

As in Catalonia, the Soviet authorities made use of politico-ideological renaming. Proper names, particularly place and street names were often changed (Raag 2003)—the examples in Russia are well-known (Leningrad and Stalingrad)—and the historic capital of Estonia’s Saaremaa island was changed from Kuressaare to Kingissepa, after an Estonian communist (Taagepera 1993). Streets named after Soviet heroes abounded in Baltic cities. Ideological language was perhaps nowhere more evident than in the media, and it began with the very names of the publications: emphasising the communal nature of the Soviet people, their common future and commitment to communism.

2.3. Immigration and the development of diglossia

Russification in the sphere of language was significantly realised implicitly, within the framework of the massive state-controlled in-migration to Estonia (and Latvia), and, in all three Baltic states, the systematic process of deportation and terror. At least 700,000 people from the three countries were deported (Druvieta 1997). In addition to this figure, we should consider the more than 250,000 who fled to the West, those who were outright killed by the occupiers (also including Nazi Germany during the Second World War), and those removed through Estonian and Latvian territorial loss (initially to the Russian SSR and to then the Russian Federation) (Kallas 2002, Parming 1978), to come to a total population loss for the three small

Baltic states of approximately two million people: Misiunas and Taagepera (1993) estimate that in total Estonia lost 25% of its pre-World War Two population, Latvia 30%, and Lithuania 15%. Combined with the state-controlled in-migration (industrial workers, functionaries, Soviet military personnel as well as ex-convicts under the so-called 101 kilometre rule¹³), and (later) personal spontaneous migration to the Baltic, the result was a massive fall in the autochthonous proportion of the population: the native Estonian population fell from 88.2% in 1938 to 74.6% in 1959 to approximately 60% at independence (Lieven 1994). The city of Narva in north-east Estonia re-emerged after the war as exclusively Russian-speaking, due to a combination of migrant-staffed industrialisation and other immigration, and the fact that pre-war Estonian inhabitants were prohibited from returning to the area (Kallas 2002). These changes in the (ethno)linguistic makeup of the country, coupled with the privileged status given to Russian throughout the Soviet Union, meant that the language of everyday interethnic public communication in Estonia was Russian: a full set of public institutions was provided in Russian, while only a partial set was provided in Estonian (Adrey 2005, Clachar 1998, Nørgaard et al. 1996). Other non-native Russian-speaking immigrants to Estonia (chiefly Ukrainians and Belarusians) largely became Russian-speaking in the Soviet Estonian context as no public services were provided in their languages.

The sociolinguistic situation under authoritarianism—in both Estonia and Catalonia—can be classified as one of diglossia; as Hogun-Brun and colleagues (2007: 509) suggest, in the Estonian case, Russian enjoyed “an asymmetric hegemonic position”. Estonian, along with Latvian and Lithuanian, lost many vital and basic sociolinguistic functions within Soviet society, and the language hierarchy was clear: “in many everyday situations, the Balts [sic] were forced to speak Russian [---] The result was [---] the superior position of the Russian language” (Nørgaard et al. 1996: 178). As Adrey (2005: 458) states, “Russian was the exclusive dominant language in official spheres, i.e. state government, transport, industry, military, and in highly qualified employments and higher education, while Estonian and Latvian essentially channelled informal social communication”. There was and to a certain extent still is, furthermore, the prevailing mindset of many Russian-speakers that they had “the right to be monolingual no matter where they live and work [in the (former) Soviet Union]” (Karklins as cited in Druviete 1997, Ozolins 1999: 181). In a 2003 survey, 16% of ethnic Estonians surveyed stated that they had used Russian with sales or service assistants in the last month; 8% reported having done so with *government* employees (Vihalemm et al. 2004). Only 13% of Russian-speakers in Estonia claim not to be able to speak Estonian at all, yet 27% say they *never use it* (Vihalemm et al. 2004). Despite being *public* servants, a 2002 EU report found that there are still some police officers in Tallinn who “have virtually no command of Estonian” (Ozolins 2003: 27).

One of the main causes of sociolinguistic change—and the rise of diglossia—in Catalonia has also been immigration (especially that of 1959–72), and, just as in Estonia, the primary reason given for the immigration has been industrialisation and the lack of workers to implement it (Conversi 1997, Hooper 1995, May 2001), although some thinkers maintain that the Francoist and Soviet regimes used immigration as a premeditated tactic to dilute the autochthonous populations (Misiunas, Taagepera 1993, Wright 2004). In the 1950s and 1960s, more than a million

¹³ “Under this rule, a person released from the prison could not settle nearer than 100 km from large cities, Leningrad [St. Petersburg] being one of them (101 km to the west from Leningrad is Estonia)” (Kallas 2004: 18).

Castilian Spanish-speakers arrived in Catalonia (McRoberts 2001). If we consider that around 200,000 fled the Catalan-speaking territories (Conversi 1997), we can appreciate the sociolinguistic changes that took place. The proportion of Catalan-speakers declined from 90% in 1939 to 60% of the population in 1975 (May 2001). During the post-war “years of development”, Andalucians arrived to Catalonia in ever greater numbers, and although this was not the first mass immigration to the region, the Andalucians were the first who neither spoke Catalan nor had any historical relationship with Catalonia (Hooper 1995). As in Estonia, these immigrants had little incentive—or opportunity, for that matter—to learn the local language (May 2001). The Murcians, who had mostly arrived during the 1920s, had generally integrated into the Catalan society as Catalan had enjoyed widespread usage in the region and it was seen as a “language of social mobility” (Hooper 1995, May 2001: 243). An elderly Murcian who had given his eldest daughter the name Montserrat became a hero of the Catalan national cause (Hooper 1995).

The absence of the written use of Catalan (in other words, in formal domains and in the media), furthermore, created a situation in which many people could still speak the language but not read it or write in it; many more had difficulties reading the language as they were accustomed to an official written world which was completely Castilianised (Fishman 1991, Hooper 1995, Keating 2001, May 2001, Strubell i Trueta 1982, Strubell 1996). In 1979, among individuals born in Catalonia, only 19.2% were able to write in Catalan; the figure for people born outside of the region was a miniscule 1.7% (McRoberts 2001). Very few people were able to teach the language and civil servants had no experience whatsoever with its professional use (Strubell 1996). Furthermore, the home language of many Catalan-speakers was not Catalan: in cases in which Catalan-speakers were married to Castilian Spanish-speakers, the couple often used Castilian Spanish, given that both individuals could speak Castilian Spanish, but quite often the Castilian Spanish-speaker did not have a solid command of Catalan (Hooper 1995), this being particularly the case in Valencia (Strubell i Trueta 1982). Catalan, it should be noted, is not generally difficult for Castilian Spanish-speakers to learn (provided they have the motivation to do so) and many “immigrants”¹⁴ to the region have at least a passive knowledge of the language (Mar-Molinero 1994, 1995), as opposed to the Estonian case, where Estonian (being a Uralic language) and Russian (being an Indo-European language) are not even from the same family.

As is the case with Russian in the former Soviet Union, Castilian Spanish enjoys a privileged social status according to which many native speakers of the previously imposed language believe they have the right not to use Catalan, even if they know the language; Catalans and Estonians, furthermore, often feel obliged to use Castilian Spanish and Russian, respectively (Druvieta 2003, Lauristin, Heidmets 2002, McRoberts 2001, Strubell i Trueta 1982, Strubell 1996, Vihalemm et al. 2004). As Blas Arroyo (2005: 407) points out, in Catalan-speaking areas of Spain, and especially in the Eastern Fringe of Aragón, an area bordering Catalonia, Catalan can, “in switching to more formal domains as well as in conversation with strangers, be abruptly substituted by [Castilian] Spanish”¹⁵. In present-day Catalonia, a rather peculiar and perhaps surprising type of linguistic behaviour exists: two native-speakers of Catalan can speak to each other in Castilian Spanish without realising that the other speaks Catalan—a situation that has arisen due to

¹⁴ The usage of the term immigrant in the Catalan case under Franco refers to people that moved to the region from other parts of Spain, who were, to a large extent, monolingual Castilian Spanish-speaking.

¹⁵ “bruscamente sustituido por el español [sic] en el paso a los dominios formales, así como en la conversación con desconocidos”

the fact that the majority of Catalan-speakers speak unaccented Castilian Spanish, and they, in presuming that the interlocutor is a native Castilian Spanish-speaker, automatically choose that language (Strubell 2001). It should be made clear, nonetheless, that in neither of the cases presently discussed is it possible to speak of a low level of linguistic self-esteem among the autochthonous peoples, and it is precisely *because* of the solid language-based identities that these varieties were able to survive under authoritarianism (Conversi 1997, Doersam 1977). Sánchez and Dueñas (2002) assert that the linguistic minorities of Spain, in general, even achieved a certain prestige of their own as symbols of resistance to the authoritarian state. Blas Arroyo (2005: 408–409) asserts that Catalan “did not suffer significant linguistic loss despite the long Francoist interruption [---] either amongst native speakers or [---] the immigrants of Spanish-speaking origin”¹⁶. The fact that Catalan continued to be used during the dictatorship, albeit informally and unofficially, by people of higher socioeconomic groups has contributed to the association of the language with economic success, at least in Catalonia itself (Blas Arroyo 2005, Spolsky 2004). Nonetheless, as Conversi (1997: 117) asserts, “[i]n the 1950s there was a moment of real danger that Catalan would become a family language, extinct in the public sphere”. Similarly, Hooper (1995: 416–417) claims that when, “Franco died, the Catalan language had entered a crisis [which] [---] if allowed to continue would sooner or later finish it off”. Evald Mikkel, an Estonian political scientist (personal communication 2006) asserts that the Estonian language, in the 1980s, was verging on the “point of no return”: Estonian-speaking parents had begun to send their children to Russian-medium schooling in the belief that it would be in their future best interest. If independence had not arrived when it did, and this trend had indeed continued and become consolidated, Estonian may have very well entered into an extremely precarious process of linguistic endangerment. Similarly, during the Francoist era, Catalan-speaking parents in Valencia would often *speak to their children* in Castilian Spanish; in that part of Spain the local variety had come “to be seen as mark of rurality and of people without culture”¹⁷, which has, naturally, affected its intergenerational transmission (Blas Arroyo 2005: 409). Of those that speak the Valencian variety today, up to one-third do not use it in many public contexts such as in shops, hospitals, or in spare time activities; in a study on the use of Catalan in the Eastern Fringe of Aragón in 1995, more than 80% indicated that they do not use the language in domains related to religion or health (Blas Arroyo 2005, Boix i Fuster, Vila i Moreno 1998).

2.4. Languages in contact

Strubell (2001; see also Wright 2004: 208) asserts that Catalan continues to be threatened by the influence of Castilian Spanish, decades after the death of Franco: “the danger exists that the specific characteristics of Catalan—in terms of phonetics, grammar, syntax, and vocabulary—will gradually be lost and that Catalan will become, in effect, a mere dialect of another language (Castilian Spanish)”¹⁸ (Strubell 2001: 277). This danger is owed to the fact that—unlike

¹⁶ “sufrió una merma significativa de prestigio pese el largo paréntesis franquista [---] tanto entre los hablantes nativos como [---] entre los inmigrantes de origen hispanohablante”

¹⁷ “interpretarse como marca de ruralidad [---] y de gentes sin cultura”

¹⁸ “[e]xiste el peligro de que las características específicas del catalán – en cuanto a la fonética, la gramática y la sintaxis, y al léxico – vayan perdiéndose gradualmente, y de que el catalán se esté haciendo, en efecto, un mero dialecto de otra lengua (el castellano)”

countries such as the United States and Australia, in which immigrants have come from a variety of countries and thus have a variety of native languages—in Catalonia, the majority of immigrants are natively Castilian Spanish-speaking and the interference of Castilian Spanish is inevitable (Strubell 2001). In a very similar way, Russian-speakers make up the majority of immigrants in Estonia. The variety of Catalan spoken in Valencia has been influenced by Castilian Spanish to a much greater degree given the lesser presence of Catalan in the education system of that region, and the local colloquial language, in particular, includes hundreds of borrowings from Castilian Spanish (Strubell 2001). There is also an offensive term, *catanyol*, for “the process of hybridisation, through the influence of [Castilian] Spanish, in which Catalan is presently found, especially in the lowest sociolects”¹⁹ (Blas Arroyo 2005: 390), which attests to the quite significant interference of Castilian Spanish, at least among certain groups of Catalan speakers.

The form of Castilian Spanish spoken in Catalonia has also been influenced by Catalan; Blas Arroyo (2005: 386–387), for example, cites *iya vengo!* instead of *iya voy!* for the *I’m coming* of standard Castilian Spanish, and *sácate (la chaqueta)* instead of *quítate (la chaqueta)* for *take off (your jacket)*, both expressions showing Catalan interference. It is not clear that this type of expression is habitual among native Castilian Spanish-speakers in the region, although Blas Arroyo indeed later notes that the use of the partitive, which does not exist in standard Castilian Spanish, as in the phrase, “me gusta el disco, pero los *hay de mejores*,” “unlike other phenomena previously described, is not found among native speakers”²⁰ (Blas Arroyo 2005: 556; see also Moreno Fernández 2005). The use of possessive phrases instead of prepositional ones, such as *detrás mío* instead of *detrás de mí* (Stewart 1999: 114) is also attributed to Catalan, as are morphosyntactic peculiarities such as “*¿está la Julia?*” (*is the Julia there?*) and “*nadie no lo ha visto*” (“*nobody hasn’t seen it*”), and the lexical trait “*voy derecho a Barcelona*” for “*I’m going straight to Barcelona*” instead of *directo*, from the Catalan *dret* (Stewart 1999: 186). In terms of standard Castilian Spanish, it is now accepted—even by the country’s main broadsheet *El País*—to use the Catalan spelling *Generalitat*, whereas previously *Generalidad* was used (Stewart 1999).

The influence of Russian on Estonian vocabulary has been substantial, although it was only the case during the Soviet period—prior influence had been from German and Finnish (Erelt 2003, Lieven 1994, Raun 1985). German has, in fact, significantly shaped Estonian: nearly one-third of Estonian vocabulary is of Germanic-language origin (Raun 1985, Viires 2004). The Soviet Russian-language influence could be seen as more insidious, however, given the planned and deliberately political nature of the changes. Loan translation or calques were the most common form of creation of new lexical items, examples being *täitevkomitee* (‘executive committee’), *viisaastakuplaan* (‘five year plan’), and *lööktöö* (‘shock work’) (Raun 1985). Direct borrowings also occurred but with less frequency than calques, although we can see examples such as *kulak*, *narodnik* (‘a 19th century Russian socialist political term’), and *oblast* (‘a particular administrative area’) (Raun 1985); these were especially frequent where no metalinguistic equivalent existed (Knowles 1990, Vinay, Darbelnet 2004), meaning that the resultant Estonian-based term would be “either impossible or highly artificial [sic]” (Raun 1985: 24). There were occasions, however, where the official usage publication advocated the use of Estonian-based

¹⁹ “el proceso de hibridación en que se encuentra el catalán actual, por influencia del español [sic], especialmente entre los sociolectos más bajos”

²⁰ “a diferencia de otros fenómenos reseñados anteriormente, no se detecta entre los monolingües”

terminology when parallel forms existed: for example *läbipääsuluba* was to be used instead of *propusk* ('a pass') and *särgik* instead *maika* ('sleeveless shirt') (Raun 1985).

Raag (2003: 243) differentiates between the introduction of "fiction words" and "phenomena words" in Estonian. Fiction words were those that "project[ed] images devoid of any social basis in Soviet reality, and were used for propaganda reasons"; examples in Estonian include *avangard* ('vanguard'), *idealism* ('idealism'), *demokraatlik tsentralism* ('democratic centralism'), *parteisisene demokraatia* ('internal party democracy'), *helge tulevik* ('bright future'), *kommunismi ehitamine* ('the building of communism'), and *töövalve* ('watch of work'). On the other hand, phenomena words were those "denoting [novel but] real phenomena in the Soviet Union," such as the above terms for collective farms and the five-year plan, as well as *autahvel* ('honour board') (Raag 2003: 243).

There has, of course, been more "natural" influence between Russian and Estonian as well as the other titular languages of the Baltic states—mostly evidenced by the fact that Russian has adopted various direct borrowings and calques from Estonian, Latvian, and Lithuanian (Clachar 1998, Knowles 1990, 1999). Estonian Russian has seen the emergence of various linguistic forms, including lexical and morphosyntactic phenomena, apparently peculiar to that particular variety (Verschik 2004a, 2004b, 2005, 2006, 2007). Some of these, such as *stavit' v golovu*, (*to put on [a hat]*), are already, as Verschik (2007: 91) claims, "impossible in monolingual Russian" and "[f]or a monolingual speaker of Standard Russian [---] the meaning [---] remains opaque".

3. Conclusion

Llobera (1996) draws a parallel between the national resistance in the Baltic countries under the Soviet Union and that which occurred in Catalonia. Citing Smith (1996a, 1996b), the author asserts that "nationalism had never disappeared, it had just taken other cultural forms" (Llobera 1996: 192). Mar-Molinero (2000: 85) notes on the Catalan case, although we could easily substitute *Franco* with *Soviet*, that

[---] the regime's clever propaganda drive made one costly mistake: the rejection and even ridicule of the [local] languages became so closely associated with the ideology of the Franco regime that it also served as a point of reference around which to build the opposition to it. Language became a symbolic standard bearer for and against the regime.

As Raag (2003: 249) maintains, however, official Soviet language use possessed "a general militaristic flavour. A progressive Soviet citizen did not walk or stretch his [sic] legs, but headed for or marched in lines to sing and dance" and these "militaristic traits did not disappear completely from Estonian lexicon as long as the occupation lasted". For years, the Catalans were not permitted to speak their language in public for fear of being fined, jailed, or simply being called a dog or unchristian. We can suppose that, although it was widely understood to be propaganda, censorship and manipulation ultimately influenced the language, culture, and society of the Estonians and the Catalans. In Latvia, the "high level of linguistic tolerance" (Druviete

2003: 5) with Russian-speaking people *even if they speak Latvian* is proposed by Druviete (2003: 5) as part of the continuation of the “minority complex” formed during the Soviet period. This type of linguistic accommodation is certainly less frequent in Estonia but it still exists; in Catalonia it appears to be quite routine.

It seems to me, then, that the most serious sociolinguistic consequence of the Soviet and Francoist regimes is that related to the drastically reduced level of meaningful usage of the languages: severe restrictions in terms of *what* could be said and *how* it could be said impeded the development of cultural production in both Estonian and Catalan, thus impeding the languages themselves. In Catalonia and Estonia, the restriction of the usage of the autochthonous languages to particular domains did not allow these languages to develop along with the changing social conditions (May 2001). There was no corpus planning in Catalonia (Strubell 1996), while in Estonia it was restricted and prescriptive in nature (Hogan-Brun et al. 2007). And as Lauk (1999: 21) asserts, “Soviet censorship in Estonia clearly served the objective of Russification and destruction of Estonian national culture”; Francoism, as we have seen, had even more repressive goals towards Catalan-speakers.

Nevertheless, Estonian and Catalan survived in the face of authoritarianism and we can conclude that, in both cases, most people did not accept the officially sanctioned public usage of language and the related ideology. We can speak of *two sociolinguistic worlds* in Estonia, one official and Sovietised, the other unofficial and very much national in character (Vihalemm, Lauristin 1997). The Catalans also tenaciously maintained their cultural and linguistic worlds and, above all, “it must be recognised that [opposition to the dictatorship after some years] centered on [issues] of language and its recovery” (Conversi 1997: 137). The effect of highly centralised control of the national culture, and therefore also the language, however, should not be discounted. As Vihalemm and Lauristin (1997: 108) suggest—about the Baltic states but with clear relevance to post-Francoist Catalonia as well—“the unconscious adaptation to the Soviet cognitive and behavioural patterns that dominated in the spiritual environment of the Baltic people for decades [is] still quite large and [will remain] present for a long time [to come]”.

References

- Adrey, Jean-Bernard 2005. Minority language rights before and after the 2004 EU enlargement: The Copenhagen Criteria in the Baltic states. – *Journal of Multilingualism and Multicultural Development*, 26 (5), 453–468. doi:10.1080/01434630508668416
- Ager, Dennis; Hoffman, Charlotte; Sala, Rafael; Radnai, Zsófia; Wright, Sue; Yates, Alan 1995. The debate. – *Current Issues in Language and Society*, 2 (1), 91–108.
- Azevedo, Milton M. 2005. *Introducción a la lingüística española*. 2nd ed. Upper Saddle River, NJ: Pearson Education.
- Blas Arroyo, José Luis 2005. *Sociolingüística del español: Desarrollos y perspectivas en el estudio de la lengua española en contexto social*. Madrid: Ediciones Cátedra.
- Boix i Fuster, Emili; Vila i Moreno; F. Xavier 1998. *Sociolingüística de la llengua catalana*. Barcelona: Ariel.
- Clachar, Arlene 1998. Differential effects of linguistic imperialism on second language learning: Americanisation in Puerto Rico versus Russification in Estonia. – *International Journal of Bilingual Education and Bilingualism*, 1 (2), 100–118.
- Conversi, Daniele 1997. *The Basques, the Catalans and Spain*. Reno: University of Nevada Press.

- Doersam, Christopher 1977. Sovietization, culture, and religion. – Edward Allworth (Ed.). *Nationality Group Survival in Multi-ethnic States: Shifting Support Patterns in the Soviet Baltic Region*. New York: Praeger Publishers, 148–193.
- Druviete, Ina 1997. Linguistic human rights in the Baltic states. – *International Journal of Sociology of Language*, 127, 161–185. doi:10.1515/ijsl.1997.127.161
- Druviete, Ina 2003. *The Latvian Language in the 21st Century*. Riga: University of Latvia.
- Ellwood, Sheelagh 1995. The moving image of the Franco regime: *Noticiarios y documentales 1943–1975*. – Helen Graham, Jo Labanyi (Eds.). *Spanish Cultural Studies: An Introduction. The Struggle for Modernity*. Oxford: Oxford University Press, 201–203.
- Erelt, Mati 2003. Preface. – Mati Erelt (Ed.). *Estonian Language. Linguistica Uralica Supplementary Series*, 1. Tallinn: Estonian Academy Publishers.
- Fishman, Joshua A. 1991. *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*. Clevedon, UK: Multilingual Matters.
- George, David 2002. *Institute of Catalan Studies*. – Eamonn Rodgers (Ed.). *Encyclopedia of Contemporary Spanish Culture*. London: Routledge, 269–269.
- Hogan-Brun, Gabrielle; Ozolins, Uldis; Ramonienė, Meilutė; Rannut, Mart 2007. Language politics and practices in the Baltic states. – *Current Issues in Language Planning*, 8, 469–631. doi:10.1515/ijsl.1997.127.161
- Hooper, John 1995. *The New Spaniards*. London: Penguin Books.
- Kallas, Kristina 2002. *The formation of interethnic relations in Estonian SSR*. (Unpublished MA Thesis.) Central European University, Budapest.
- Keating, Michael 2001. *Nations Against the State: The New Politics of Nationalism in Quebec, Catalonia and Scotland*. 2nd ed. Houndmills, UK: Palgrave Macmillan.
- Knowles, Francis 1990. *Language Planning in the Soviet Baltic Republics: An Analysis of Demographic and Sociological Trends*. – Michael Kirkwood (Ed.). *Language Planning in the Soviet Union*. New York: St. Martin's Press, 145–173.
- Knowles, Francis 1999. Ethno-linguistic relations in contemporary Latvia: Mirror image of the previous dispensation? – *Current Issues in Language & Society*, 6 (1), 48–56.
- Labanyi, Jo 1995. *Censorship or the fear of mass culture*. – Helen Graham, Jo Labanyi (Eds.). *Spanish Cultural Studies: An Introduction. The Struggle for Modernity*. Oxford: Oxford University Press, 207–214.
- Lauk, Epp 1999. *Practice of Soviet censorship in the press: The case of Estonia*. – *Nordicom Review*, 2, 19–31.
- Lauristin, Marju; Heidmets, Mati 2002. *Introduction: The Russian minority in Estonia as a theoretical and political issue*. – Marju Lauristin, Mati Heidmets (Eds.). *The Challenge of the Russian Minority: Emerging Multicultural Democracy in Estonia*. Tartu: Tartu University Press, 19–30.
- Lauristin, Marju; Vihalemm, Peeter 1993. *The communist press in the Soviet Republics*. – Svennik Høyer, Epp Lauk, Peeter Vihalemm (Eds.). *Towards a Civic Society: The Baltic Media's Long Road to Freedom. Perspectives on History, Ethnicity and Journalism*. Tartu: Nota Baltica, 177–188.
- Lauristin, Marju; Vihalemm, Peeter; Brikše, Inta; Bulota, Jonas; Dūze, Dace; Šulmane, Ilze 1993. *The Period of Stagnation (1969–1986)*. – Svennik Høyer, Epp Lauk, Peeter Vihalemm (Eds.). *Towards a Civic Society: The Baltic Media's Long Road to Freedom. Perspectives on History, Ethnicity and Journalism*. Tartu: Nota Baltica, 211–220.
- Lieven, Anatol 1994. *The Baltic Revolution: Estonia, Latvia, Lithuania and the Path to Independence*. 2nd ed. New Haven: Yale University Press.
- Llobera, Josep R. 1996. *The role of commemorations in (ethno)nation-building. The case of Catalonia*. – Clare Mar-Moliner, Angel Smith (Eds.). *Nationalism and the Nation in the Iberian Peninsula*. Oxford: Berg, 191–206.
- Mar-Moliner, Clare 1994. *The politics of language: Spain's minority languages*. – *CLE Working Papers*, 3, 106–113.

- Mar-Molinero, Clare 1995. The politics of language: Spain's minority languages. – Helen Graham, Jo Labanyi (Eds.). *Spanish Cultural Studies: An Introduction. The Struggle for Modernity*. Oxford: Oxford University Press, 336–342.
- Mar-Molinero, Clare 1996. The role of language in Spanish nation-building. – Clare Mar-Molinero, Angel Smith (Eds.). *Nationalism and the Nation in the Iberian Peninsula: Competing and Conflicting Identities*. Oxford: Berg, 69–88.
- Mar-Molinero, Clare 2000. *The Politics of Language in the Spanish-speaking World: From Colonisation to Globalisation*. London: Routledge.
- May, Stephen 2001. *Language and Minority Rights: Ethnicity, Nationalism and the Politics of Language*. Harlow, UK: Pearson Education.
- McRoberts, Kenneth 2001. *Catalonia: Nation Building Without a State*. Don Mills, Canada: Oxford University Press.
- Misiunas, Romuald J.; Taagepera, Rein 1993. *The Baltic States: Years of Dependence 1940–1990*. Berkeley: University of California Press.
- Moldoveanu, Mihaíl 2002. Noucentisme. – Eamonn Rodgers (Ed.). *Encyclopedia of Contemporary Spanish Culture*. London: Routledge, 367–368.
- Moreno Fernández, Francisco 2005. *Historia social de las lenguas de España*. Barcelona: Ariel.
- Nørgaard, Ole; Hindsgaul, Dan; Johannsen, Lars; Willumsen, Helle 1996. *The Baltic States after Independence*. Cheltenham: Edward Elgar.
- Ozolins, Uldis 1999. Between Russian and European hegemony: Current language policy in the Baltic States. – *Current Issues in Language & Society*, 6 (1), 6–47.
- Ozolins, Uldis 2003. The impact of European Union accession upon language policy in the Baltic states. – *Language Policy*, 2, 217–238. doi:10.1023/A:1027320716791
- Parming, Tõnu 1978. Population changes and processes. – Tõnu Parming, Elmar Järvesoo (Eds.). *A Case Study of a Soviet Republic: The Estonian SSR*. Boulder, CO: Westview Press, 21–74.
- Raag, Virve 2003. Soviet Occupation and the Estonian Standard Language. – Anu Mai Kõll (Ed.). *The Baltic Countries Under Occupation: Soviet and Nazi Rule 1939–1991*. Stockholm: Department of History, Stockholm University, 241–250.
- Raun, Toivo U. 1985. Language development and policy in Estonia. – Isabelle T. Kreindler (Ed.). *Sociolinguistic Perspectives on Soviet National Languages: Their Past, Present and Future*. Berlin: Mouton de Gruyter, 13–35.
- Richards, Michael 1996. Constructing the nationalist state: Self sufficiency and regeneration in the early Franco years. – Clare Mar-Molinero, Angel Smith (Eds.). *Nationalism and the Nation in the Iberian Peninsula: Competing and Conflicting Identities*. Oxford: Berg, 149–167.
- Sánchez, Aquilino; Dueñas, María 2002. Language planning in the Spanish-speaking world. – *Current Issues in Language Planning*, 3 (3), 280–305. doi:10.1080/14664200208668042
- Shohamy, Elana 2006. *Language Policy: Hidden Agendas and New Approaches*. Milton Park, UK: Routledge.
- Smith, Graham 1996a. Introduction: The Baltic nations and national self-determination. – Graham Smith (Ed.). *The Baltic States: The National Self-determination of Estonia, Latvia and Lithuania*. New York: St. Martin's Press, 1–9.
- Smith, Graham 1996b. The resurgence of nationalism. – Graham Smith (Ed.). *The Baltic States: The National Self-determination of Estonia, Latvia and Lithuania*. New York: St. Martin's Press, 121–143.
- Spolsky, Bernard 2004. *Language Policy*. Cambridge: Cambridge University Press.
- Stewart, Miranda 1999. *The Spanish Language Today*. London: Routledge.
- Stewart, Miranda 2002. Language and national identity. – Eamonn Rodgers (Ed.). *Encyclopedia of Contemporary Spanish Culture*. London: Routledge, 293–294.

- Strubell i Trueta, Miquel 1982. Catalan sociolinguistics: A brief review of research. – *International Journal of the Sociology of Language*, 38, 70–84.
- Strubell i Trueta, Miquel; Ros i Garcia, María 1984. Introduction. – *International Journal of the Sociology of Language*, 47, 5–11.
- Strubell, Miquel 1996. Language planning and bilingual education in Catalonia. – *Journal of Multilingual and Multicultural Development*, 17 (2–4), 262–275. doi:10.1080/01434639608666279
- Strubell, Miquel 1998. Language, democracy and devolution in Catalonia. – *Current Issues in Language and Society*, 5 (3), 146–180.
- Strubell, Miquel 2001. Catalan a decade later. – Joshua A. Fishman (Ed.). *Can Threatened Languages be Saved? Reversing Language Shift, Revisited: A 21st Century Perspective*. Clevedon, UK: Multilingual Matters, 260–283.
- Taagepera, Rein 1993. *Estonia: Return to Independence*. Boulder, CO: Westview Press.
- Verschik, Anna 2004a. Aspects of Russian-Estonian codeswitching: Research perspectives. – *International Journal of Bilingualism*, 8 (4), 427–448. doi:10.1177/13670069040080040201
- Verschik, Anna 2004b. Estonian compound nouns and their equivalents in the local variety of Russian. – *Scando-Slavica*, 50 (1), 93–109. doi:10.1080/00806760410011132
- Verschik, Anna 2005. Russian-Estonian language contacts, linguistic creativity, and convergence: New rules in the making. – *Multilingua*, 24, 413–429. doi:10.1515/mult.2005.24.4.413
- Verschik, Anna 2006. Convergence in Estonia's Russian: Directional vs. static vs. separative verb. – *International Journal of Bilingualism*, 10 (4), 383–404. doi:10.1177/13670069060100040101
- Verschik, Anna 2007. Multiple language contact in Tallinn: Transfer B2 > A1 or B1 > A2? – *The International Journal of Bilingual Education and Bilingualism*, 10 (1), 80–103. doi:10.2167/beb372.0
- Vihalemm, Peeter; Lauristin, Marju 1997. Political control and ideological canonisation: The Estonian press during the Soviet period. – Eduard Mühle (Ed.). *Vom Instrument der Partei zur "Vierten Gewalt": Die Ostmitteleuropäische Presse als Zeithistorische Quelle*. Marburg, Germany: Verlag Herder-Institut, 103–109.
- Vihalemm, Triin; Masso, Anu; Vihalemm, Peeter 2004. *Eesti kujunev keeleruum*. – Veronika Kalmus, Marju Lauristin, Pille Pruulman-Vengerfeldt (toim.). *Eesti elavik 21. sajandi algul: ülevaade uurimuse Mina. Maailm. Meedia tulemustest*. Tartu: Tartu Ülikooli Kirjastus, 57–73.
- Viires, Ants 2004. *Old Estonian folklife*. (Translated by Mart Aru.) Tallinn: Ilo Publishing House.
- Vinay, Jean-Paul; Darbelnet, Jean 2004. *A Methodology for Translation*. (Translated by Juan C. Sager, M.-J. Hamel.) – Lawrence Venuti (Ed.). *The Translation Studies Reader*. 2nd ed. New York: Routledge, 128–137.
- Wright, Sue 1998. Catalonia: The geographical and historical context of the language question. – *Current Issues in Language and Society*, 5 (3), 181–189.
- Wright, Sue 2004. *Language Policy and Language Planning: From Nationalism to Globalisation*. Houndmills, UK: Palgrave Macmillan.

Delaney Michael Skerrett is a doctoral candidate in Applied Linguistics at The University of Queensland, researching Estonian post-Soviet language policy. He also lectures part-time in the Center for Baltic Studies and the Department of English at the University of Tartu. In 2009, he was invited to speak about the future of Estonian language policy and planning at the University of Barcelona.
d.skerrett@uq.edu.au

KEEL JA AUTORITAARSUS XX SAJANDIL: EESTI JA KATALOONIA JUHTUM

Delaney Michael Skerrett

Queenslandi Ülikool, Tartu Ülikool

XX sajand nägi sovetlikku ja frankistlikku režiimi teostamas oma ideoloogiat Eestis ja Hispaania katalaanikeelsesel territooriumil. Kummalgi juhul kannatas kohalik keel tsensoori- ja repressioonimehhanismide jäiga kontrolli all. Tegelikult püüdsid Nõukogude Liidu ja Hispaania liidrid, esindades poliitilise spektri äärmusi, vahetada kohalikku keelt välja omaenese, vastavalt vene ja (Kastiilia) hispaania keele vastu, sest need keeled kehastasid ideoloogiat ja uut korda, mida nad soovisid kehtestada.

Artiklis võrreldakse Eestis ja Hispaania katalaanikeelsetel aladel keele üle kehtestatud kontrolli mitmesuguseid meetodeid ning näidatakse, et väga tsentraliseeritud mitmekeelsed riigid – olgu mis tahes poliitilise ideoloogiaga – kasutavad üllatavalt sarnaseid kontrollivahendeid, röövides äärmusena kohalikult keelekogukonnalt võim(alus)e kasutada ja arendada oma keelt.

Võtmesõnad: diglossia, kontaktkeeled, keelepoliitika, keelekorraldus, ohustatud keeled, eesti keel, vene keel, katalaani keel, Kastiilia hispaania keel, hispaania keel

VAHENDAMATA JA VAHENDATUD NARRATIIV LASTE KÕNEARENGU HINDAMISEL

**Piret Soodla, Eve Kikas, Renate Pajusalu,
Aive Adamka, Sirli Parm**

Ülevaade. Artikkel käsitleb 162 eakohase kõnearenguga ja 27 alakõnega 6–7-aastase lapse kolmel viisil jutustatud suulisi narratiive – vahendamata narratiive ilma eelneva narratiivi mudelita, eelnevalt mudeldatud vahendamata narratiive ning vahendatud narratiive. Analüüs hõlmab narratiivide mikrostruktuuri (lausungite keskmist pikkust, grammatilist õigsust) ning makrostruktuuri (infoüksuste hulka, kompleksust). Eesmärgiks on selgitada, missugusel viisil koostatud narratiivid eristavad kõnearengu mahajäämust eakohasest arengust kõige paremini. Tulemustest selgub, et jutustamisoskuse hindamiseks narratiivi mikrostruktuuri tasandil sobivad kõigil kolmel viisil jutustatud narratiivid. Makrostruktuuri tasandil on parimaks hindamisvahendiks vahendatud narratiiv, millele järgneb eelnevalt mudeldatud vahendamata narratiiv.*

Võtmesõnad: keeleline areng, kõnepuuded, narratiiv, suuline keel, süntaks, morfoloogia, tekstistruktuur, eesti keel

1. Sissejuhatavat

Artikkel analüüsib üht kõnetegevuse olulist funktsiooni – narratiivi – ja selles peegelduvat kõnearengu taset. Mõistet *narratiiv* on keeleteaduses defineeritud erinevalt. Näiteks William Labovi järgi (1972: 359–360) on narratiiv keelekasutuse meetod, mille abil minevikus aset leidnud sündmusi edastatakse verbaalse lausejärjendina. Jerome Bruner (1985: 98–99) on aga narratiivi defineerinud kui ühte mõtlemise viisi, milles opereeritakse tegelaste kavatsuste ja ajaliselt järjestatud toimingutega, vastandudes seejuures teisele, loogilis-teaduslikule mõtlemisviisile. Hoolimata mõneti erinevatest definitsioonidest ollakse siiski ühel meelel, et narratiivne tekst jutustab tüüpiliselt tegelaste (kelleks võivad olla nii inimesed, loomad kui

* Artikkel on valminud Eesti Teadusfondi grantide nr 7388 ja 7492 toel.

ka väljamõeldud olendid) tegevustest ja/või tunnetest ja mõtetest ajalises mõõtmes. Käesolevas artiklis kasutame *narratiivi* sünonüümina ka terminit *jutustus*.

Narratiive peetakse rikkalikuks keelematerjaliks keeleteaduslikes, psühholoogia- ning kõnepatoloogiaalastes uurimustes. On leitud, et narratiivid eristavad hästi eri vanuses ja keeleliste võimetega lapsi (Liles 1987, Merritt, Liles 1987, Reuterskiöld Wagner jt 1999, Scott, Windsor 2000, Botting 2002, John jt 2003, Ukrainetz jt 2005). Paljudes uurimustes on osutatud erinevustele eakohase kõnearenguga (EK) ja alakõnega (AK)¹ laste iseseisvates ehk vahendamata narratiivides (nt Norbury, Bishop 2003, Pearce jt 2003, Fey jt 2004, Reilly jt 2004, Schneider jt 2006) ning ümberjutustustes ehk vahendatud narratiivides (nt Hayward jt 2007, Pankratz jt 2007). Tulemusi on aga raske üldistada keelte, uuritavate laste vanuse, uuringu protseduuri, hindamisvahendite ja uurimisobjektide erinevuste tõttu. Vahendamata ja vahendatud narratiive võrdlevaid uurimusi on senini läbi viidud vähe (Merritt, Liles 1987, 1989, Liles jt 1995, Reuterskiöld Wagner jt 1999, Botting 2002) ning tehtud järeldused on üsna vastuolulised. Puudub selgus, mil viisil koostatud jutustused eristavad eakohast ja mahajäänud kõnearengut kõige paremini. Logopeedilises töös on nimetatud teadmised aga olulise väärtusega.

Käesolevas uurimuses võrdleme 6–7-aastaste EK ja AK laste kolmel erineval viisil jutustatud narratiive (vahendamata narratiiv ilma eelneva narratiivi mudelita ja eelneva mudeliga ning vahendatud narratiiv). Eesmärgiks on selgitada, mis viisil koostatud narratiivid eristavad alakõnet eakohasest kõnearengust kõige paremini.

2. Narratiivi mikrostruktuur eakohase arenguga ja alakõnega lastel

Tekst on üles ehitatud hierarhiliselt kahel – mikro- ja makrostruktuuri – tasandil (van Dijk, Kintsch 1983). Nendest lähtuvalt analüüsitakse laste jutustavaid tekste ka kõnepatoloogiaalastes uurimustes (Hughes jt 1997: 111). Mikrostruktuur moodustub teksti väiksematest üksustest, s.t sõnadest, lausungitest ning lausungitevahelistest seostest. On leitud, et mikrostruktuuri näitajad eristavad hästi AK, iseäranis spetsiifilise kõnearengu puudega² lapsi eakohase kõnega lastest (Liles jt 1995, Scott, Windsor 2000). Järgnevas ülevaates käsitleme narratiivi lausungite keskmist pikkust, grammatilist õigsust ning lausungitevahelist sidusust, mis on vaatluse all ka töö uurimuslikus osas.

Pikki lausungeid peetakse üldjuhul iseloomulikuks süntaktiliselt keerukale väljendusele (Scott, Windsor 2000). Inglise keelel põhinevates uurimustes on leitud, et laste vanuse kasvades muutuvad pikemaks ka jutustustes kasutatud lausungid (Justice jt 2006, Tilstra, McMaster 2007), samuti eristab lausungite keskmine pikkus AK laste jutte EK laste omadest nii vahendamata (Scott, Windsor 2000, Bishop, Donlan 2005) kui ka vahendatud narratiivides (Hayward jt 2007).

Grammatikavigu jaotatakse süntaksivigadeks (nt obligatoorsete lauseliikmete väljajätt, vale sõnajärg) ja morfoloogiavigadeks (mingi tunnuse, lõpu või liite

¹ Alakõne ehk kõnearengu üldise mahajäämuse korral avalduvad semantikapuuded, vanusnormile ei vasta kõik keeleteasandid (sõnavara, grammatika, foneemid). Eristatakse kolme tasandit: esimene tasand (lalinkõnest kuni agrammatilise lauseni), teine tasand (agrammatilisest lausest kuni primitiivse tekstini), kolmas tasand (alakõne jääknähud, sh primitiivne tekst ja lause) (Karlep 1997: 11).

² Spetsiifilise kõnearengu puude (ingl *specific language impairment*) korral esineb oluline mahajäämus eakohasest kõnearengust. Tegemist on esmase kõnepuudega, s.t alakõne ei tulene vaimsest alaarengust, kuulmislangusest, neuroloogilistest haigustest, keskkonna mõjudest, "emotsionaalsetest häiretest" (millena enamasti peetakse silmas autistlikke häireid) ega artikulatsiooniaparaadi füüsilistest puuetest (Bishop 1997: 32–34).

äräjätmine, asendamine, lisamine). Eesti keele spetsiifikast tulenevalt lisanduvad morfofonoloogilised vead, mis avalduvad morfeemivariandi (käändelõpu, tunnuse, liite või tüve) vales valikus või moonutamises (Padrik 2006). Inglise keelel põhinevad uurimused on näidanud, et kõnearengu käigus grammatikavead vähenevad (Tilstra, McMaster 2007) ning need on sobivaks näitajaks EK ja AK laste jutustuste eristamisel (Scott, Windsor 2000, Norbury, Bishop 2003, Pearce jt 2003, Fey jt 2004). Kuna muutemorfoloogia omandamine on keelespetsiifiline protsess, olles mõjutatud nii sisendkeele sagedustest ja eelistustest kui ka morfooloogilistest, fonoloogilistest ning morfosüntaktilistest teguritest (Argus 2008: 41), ei saa teistes keeltes saadud tulemusi siiski otseselt üle võtta.

Lausungid on tekstis omavahel semantiliselt ja lingvistiliselt seotud. Leksikaalsetest vahenditest kasutatakse enam otsest kordust, sünonüüme, antonüüme, eri üldistusastmega sõnu (hüperonüüme), samatüvelisi tuletisi ja liitsõnu, temaatilisi sõnarühmi, ase-, määr- ja arvsõnu. Grammatilisteks vahenditeks on tegusõnade samad ajavormid, sõnajärg, siddendid, üldlaiendid ja sarnased lausekonstruktsioonid (Karlep 2003: 266). Sidusust peetakse üldjuhul teksti mikrostruktuuritasandi tunnuseks (Hughes jt 1997: 145), v.a lausetevahelisi siddendeid, mis suhestuvad enam makrostruktuuriga (Liles jt 1995). Leksikaalsetest vahenditest kasutavad eesti lapsed koolieelses eas ja algklassides valdavalt otsest kordust ja asesõnu, enamkasutatavateks siddenditeks on *ja*, *siis* ning *ja siis* (Pajuste 2007, Ruul 2009). Nii inglise kui ka eesti keeles on leitud, et arengu käigus õigesti kasutatud sidusvahendite sagedus tõuseb ja mitmekesistub (Peterson, Dodsworth 1991, Shapiro, Hudson 1991, Ruul 2009) ning AK laste jutud sisaldavad tavalastega võrreldes oluliselt vähem sidusvahendeid (Liles 1987, Ripich, Griffith 1988, Ruul 2009). Ümberjutustustes on leksikaalne sidusus (objektile viitamine) tugevam kui vahendamata tekstides, seda nii eakohase kui mahajäänud kõnearengu korral (Ripich, Griffith 1988, Pajuste 2007).

3. Narratiivi makrostruktuur eakohase arenguga ja alakõnega lastel

Makrostruktuuri all peetakse silmas teksti üldist sisulist ülesehitust. Makrostruktuur viitab otseselt teksti terviklikkusele. Et tekst oleks terviklik, peab jutustaja edastatava informatsiooni organiseerima ajaliselt ja põhjuslikult seotud järjendiks, mis on tähendusrikas nii kõneleja enda kui kuulajate jaoks (Shapiro, Hudson 1991). Narratiivi terviklikkusele aitavad kaasa mitmed struktuurielemendid, nt traditsioonilised jutustuse algus- ja lõpumarkerid, tegevuse aja ja koha ülevaade, tegelaste kirjeldused, tegevused, tegelastevaheline dialoog, eesmärgid, reaktsioonid, ettetulevad takistused, tegelase püüded neid ületada jne (Shapiro, Hudson 1991). Lisaks peetakse hea jutu tunnuseks jutustajapoolset vaatepunkti jutustatavale, s.t hinnanguid tegelaste, nende tegutsemise motiivide ja asetleidnud sündmuste kohta (Ukrainetz jt 2005).

Narratiivi makrostruktuuri analüüsiks on välja pakutud mitmeid mudeleid, millest levinuim on Nancy Steini ja Christine Glenni (1979) jutugrammatika mudel (ingl *Story Grammar Model*). Mudeli järgi koosnevad narratiivid teatud kindlat tüüpi informatsiooni sisaldavatest komponentidest, mis on omavahel seotud

ajaliselt ja/või põhjuslikult. Jutustuse makrostruktuur koosneb taustakirjeldusest ja episoodide süsteemist. Taustakirjelduses tutvustatakse peategelasi ning kirjeldatakse sotsiaalset, füüsilist või ajalist konteksti, milles sündmused hakkavad aset leidma. Jutustus võib koosneda vaid ühest, enamasti siiski mitmest omavahel seotud episoodist. Episood koosneb kuuest komponendist: 1) käivitav sündmus, mis tekitab tegelasel lahendamist vajava probleemi; 2) sisemine vastus – käivitavast sündmusest tingitud tegelase reaktsioon (tunded, mõtted); 3) sisemine plaan – tegelase plaan probleemi lahendamiseks või olukorra muutmiseks; 4) tegevus – tegelase toimingud plaani täitmisel; 5) tagajärg – tegevuse otsene tagajärg; 6) reaktsioon – tegelase reaktsioon tagajärjele. Struktuurilt kõige olulisemateks üksusteks peetakse käivitavat sündmust või sisemist vastust, eesmärgipärast tegevust ja tagajärge (Liles jt 1995). Vähemalt nimetatud kolme komponenti sisaldavat episoodi nimetatakse lõpetatuks. Ka ülejäänud komponendid annavad olulist informatsiooni, kuid pole narratiivi terviklikkuse aspektist hädavajalikud (Merritt, Liles 1987).

Kirjeldatud mudelil põhinevad paljud uurimused, kus on vaatluse all jutugrammatika komponentide esinemissagedus (nt Merritt, Liles 1987, John jt 2003) või lõpetatud ja lõpetamata episoodide osakaal narratiivis (nt Merritt, Liles, 1987, Ripich, Griffith 1988). Nimetatud meetodid ei ole aga osutunud efektiivseks alakõne diferentsiaaldiagnostikas (Liles jt 1995, Norbury, Bishop 2003). Phyllis Schneider jt (2006) soovivad narratiivide makrostruktuuri hindamisel arvesse võtta kõikides jutugrammatika komponentides sisalduvat informatsiooni, mis autorite hinnangul võimaldab paremini eristada AK ja EK laste jutustusi. Paraku on AK ja EK laste narratiive nimetatud aspektist vähe uuritud ning saadud tulemused on vastuolulised: on leitud, et nii vahendamata kui ka vahendatud jutustustes on olulise informatsiooni hulk AK lastel oluliselt väiksem kui EK korral (Reutersköld Wagner jt 1999, Schneider jt 2006), kuid on ka tulemusi, kus nimetatud erinevusi ei leitud (Norbury, Bishop 2003).

Siiski on hea narratiiv midagi enam kui keeleliselt korrektne, sidus ja sisult struktureeritud sündmuse edastamine. Et kuulaja tähelepanu saavutada ja oma mõtteid võimalikult täpselt edastada, kasutab oskuslik jutuvestja keelelisi ja mittekeelelisi vahendeid, nt hääliitsusi, hüüatusi, rõhutamist, kordamist, küsimusi kuulajale, liialdusi ja fantaasiat, hinnanguid jm (Hughes jt 1997:132). Nende nn ekspressiivsete elementide kasutamine kõnearengu käigus sageneb ja mitmekestub (Ukrainetz jt 2005). AK ja EK laste narratiivide võrdlusel pole aga üheste tulemusteni jõutud: erinevusi on leitud nii vahendamata (Reilly jt 2004) kui ka vahendatud jutustustes (Sleight, Printz 1985), kuid on ka uurimusi, kus erinevusi ekspressiivsete elementide kasutamisel ei leitud (Norbury, Bishop 2003).

4. Narratiivi analüüs kõnearengu hindamise meetodina

Kuigi alakõnega ja eakohase kõnearenguga laste narratiive võrdlevate uurimuste hulk on viimastel aastatel tõusnud, puudub siiani üksmeel sobivaima meetodi osas AK diagnoosimisel. Donna Merritt ja Betty Liles (1989) soovivad narratiivi makrostruktuuri hindamisel kasutada vahendatud jutustust, mille analüüs võimaldab jõuda usaldusväärsemate tulemusteni, võrreldes vahendamata jutustusega.

Lisaks on leitud, et erinevused EK ja AK laste rühmade vahel on oluliselt suuremad just vahendatud narratiivides (Merritt, Liles 1987, 1989). Samas eelistavad Schneider jt (2006) jutustamisoskuse hindamisel kasutada just vahendamata jutustust, mis peegeldab paremini sõltumatuid tekstiloomeskusi, kuna tuleb planeerida ja formuleerida täiesti iseseisvalt. Analüüsitulemuste usaldusväärsuse tagamiseks soovivad autorid kasutada pildiseeria alusel jutustamist.

Vahendamata jutustusi käsitlevates uurimustes on narratiive koostatud kahel viisil – ilma eelneva verbaalse näidise ehk mudelita (nt Liles 1987, Merritt, Liles 1987, 1989, Scott, Windsor 2000, Norbury, Bishop 2003, Pearce jt 2003, Reilly jt 2004) ning eelneva narratiivi mudeliga (näidisjutu esitamine, treeningjutu ühine koostamine enne testimist vm) (nt Reuterskiöld Wagner jt 1999, Fey jt 2004, Schneider jt 2006). Meile teadaolevalt ei ole aga nimetatud narratiivi tüüpe võrdlevaid uurimusi erineva kõnearenguga lasterühmades läbi viidud. Seega puudub senini üksmeel, missugust jutustamisviisi (vahendatud vs. vahendamata narratiiv, eelneva narratiivi mudeliga vs. eelneva mudelita jutustamine) eelistada laste jutustamisoskuste uurimises ning alakõne diagnostikas.

5. Eakohase kõnearenguga ja alakõnega eesti laste narratiivid: eesmärk ja hüpoteesid

Eesti laste jutustamisoskusi on viimastel aastatel uuritud vaid üliõpilastööde raames (nt Pajuste 2007, Adamka 2008, Ruul 2009) ning esialgu puudub terviklik ülevaade laste, sh AK laste oskustest. Seda oleks aga vaja laste kõnearengu hindamisel ning toetamisel. Uurimuses võrdleme 6–7-aastaste EK ja AK eesti laste narratiive, mis on koostatud kolmel viisil: vahendamata jutustus ilma eelneva narratiivi mudelita, vahendatud jutustus, vahendamata jutustus eelneva narratiivi mudeliga. Uurimuse teoreetiline olulisus seisneb AK laste tekstiloomeraskuste ja eri jutustamisviiside seostamises. Praktiline eesmärk on kirjeldada 6–7 aasta vanuste eesti laste jutustamisoskusi narratiivi mikro- ja makrostruktuuri tasandil ning selgitada, mil viisil jutustatud narratiivid eristavad alakõnet eakohasest arengust kõige paremini. Eesmärgist tulenevalt on püstitatud kolm hüpoteesi.

1. Võttes aluseks eelnevad uurimused, mis käsitlevad AK ja EK laste narratiivide mikrostruktuuri (nt Liles jt 1995, Reuterskiöld Wagner jt 1999, Scott, Windsor 2000, Botting 2002, Norbury, Bishop 2003, Pearce jt 2003, Fey jt 2004, Reilly jt 2004, Bishop, Donlan 2005, Peña jt 2006, Hayward jt 2007) ja makrostruktuuri (nt Reuterskiöld Wagner jt 1999, Schneider jt 2006), ning AK kolmanda tasandi diagnostilised kriteeriumid – primitiivne tekst ja lause (vt allmärkus nr 2), eeldasime, et kõigil kolmel viisil koostatud narratiivid on AK korral EK lastega võrreldes mikro- ja makrostruktuuri osas madalamal tasemel.
2. Teksti ümber jutustades, erinevalt iseseisvast tekstiloomest, ei väljenda kõneleja oma mõtet, vaid vahendab varem tajutud teksti. Selleks tuleb aga seda mõtestada ja mäletada. Teksti mõistmine, mis sõltub tajuja kognitiivsest arengust ja keelelistest oskustest, on AK korral eakohasest tasemest maha jäänud (vt Leonard 1998: 86). Mitmetes uuringutes on leidnud kinnitust ka AK laste mäluprotsesside eakohasest madalam arengutase

(vt Vance 2008). Toetudes teadmistele AK laste kognitiivsest ja keelelisest arengust, samuti vahendamata ja vahendatud jutustusi võrdlevate uurin-gute tulemustele (Merritt, Liles 1987, 1989), eeldasime, et vahendatud narratiivid eristavad EK ja AK laste rühmi paremini kui vahendamata narratiivid. Nimelt oletasime, et teksti reprodutseerimisel on EK lapsed AK lastest oluliselt edukamad teksti sisulisel ja keelelisel koostamisel ning erinevused jutustamisoskustes ilmnevad kõige selgemini just vahendatud narratiivides.

3. Eesmärgiga tutvustada lapsele jutu formaati, soovivad Schneider jt (2006) enne oskuse hindamist viia läbi harjutusjutt mingil muul teemal ehk kasutada mudeldamist. On leitud, et kuigi mudeldamisest on kasu nii eakohase kui mahajäänud kõnearenguga lastel, õpivad sellest enam EK lapsed ning mudeldamise mõju on oluliselt väiksem AK korral (Peña jt 2006). Käesolevas töös eeldasime, et vahendamata jutustus, millele eelnes näidisjutt, eristab EK ja AK laste rühmi paremini kui eelnevalt mudelda-mata vahendamata jutustus ning mudeldamisel on positiivne mõju eelkõige EK laste tekstiloomele narratiivi makrostruktuuri tasandil.

6. Meetod

6.1. Valim

Uuringus osales 189 eesti keelt emakeelena kõnelevat 6–7-aastast last, kellest 162 oli eakohaselt arenenud kõnega ning 27 alakõnega. EK lapsed käisid lasteaias koolieelsetes rühmades (6 maakonnast, 21 lasteaiast), lapsi testiti õppeaasta lõpus kevadel. EK rühma loeti lapsed, kelle kõnearengu taset pidasid lasteaiapäetajad keskmiseks või kõrgeks (küsitletud ankeedis), madalaks hinnatud kõnearengu tasemega lapsed jäeti valimist välja. AK lapsed käisid kõnepuuetega laste erikoolis, testimine toimus 1. klassi alguses sügisel. Kõik AK lapsed olid kooli suunatud rahvusvahelise haiguste klassifikatsiooni (RHK-10 1993) kliinilise diagnoosi alusel: 13 lapsel esines ekspressiivne kõnehäire (F80.1), 7 lapsel retseptiivne kõnehäire (F80.2), 7 lapsel segatüüpi spetsiifilised arenguhäired (F83)³, kõikidel lastel oli logopeed diagnoosinud alakõne (kolmas tasand). EK rühmas oli 72 (44%) poissi ja 90 (56%) tüdrukut, AK rühmas 16 (59%) poissi ja 11 (41%) tüdrukut. Enne uuringu läbiviimist paluti laste testimiseks lapsevanematelt kirjalikus vormis luba ning valimisse võeti vaid lapsed, kelle vanemad olid andnud selleks nõusoleku.

6.2. Vahendid

Jutustamisel kasutati kolme viieosalist pildiseeriat (kelgulugu, pallilugu, lumemem-melugu) ning kahte teksti (pallilugu, lumememmelugu).⁴ Pildiseeriatel kujutatud keskkond, tegelased ja tegevused olid eeldatavalt lastele tuttavad ning arusaadavad.

³ F80.1 korral on lapse väljendusoskus märkimisväärselt madalam ealisest normtasemest, kõnest arusaamine on seejuures normi piires. F80.2 korral on lapse arusaamine keelelistest konstruktsioonidest puudulik tema eale vastavast tasemest, ekspressiivne kõne on enamasti puudulik (RHK-10 1993: 223–226). F83 on "raskesti defineeritav, ebapiisava täpsusega määratletud häirete kategooria, mille puhul spetsiifilised keele-/kõne-, õpivulumuste ja/või motoorsete funktsioonide häired on segunenud" (RHK-10 1993: 239). Kõikidele nimetatud puuetele on iseloomulik alakõne ning mõningane kognitiivsete funktsioonide kahjustus, vaimse arengu tase on siiski eanormi piires.

⁴ Narratiivide süžee ning tekstid koostasid Merit Hallap, Marika Padrik, Piret Soodla ja Kati Sütt, pildid joonistas Jolana Laidma. Piltide ja tekstide näidiseid vt Pajuste (2007) ja Adamka (2008).

Narratiivide usaldusväärse kodeerimise eesmärgil püüti pildiseeriad ja tekstid koostada keeleliselt ja sisuliselt sarnase raskusastme ning struktuuriga. Pildiseeriatel kujutati vastas jutugrammatika mudelile (Stein, Glenn 1979), osutades mudeli sisukomponentidele (taust, käivitav sündmus, sisemine vastus, sisemine plaan, tegevus, tagajärg, reaktsioon). Pildiseeriatele vastavad tekstid sisaldasid samuti kõiki jutugrammatika komponente, tekstid olid sarnase pikkuse (palliloo 136, lumememmeloo 145 sõna), grammatilise ehituse (lausungite keskmine pikkus mõlemas jutustuses 6 sõna), infoüksuste hulga (palliloo 20, lumememmeloo 21 infoüksust) ning kompleksusega (palliloo 20, lumememmeloo 21 punkti). Lausungite, infoüksuste hulga ja narratiivi kompleksuse käsitlusi vt allosas 6.4.

6.3. Protseduur

Testimine viidi läbi individuaalselt. Iga laps jutustas kolm narratiivi, mis lindistati ning hiljem transkribeeriti. Esimese jutustuse aluseks oli kelguloo pildiseeria, teise ja kolmanda jutu aluseks olevaid seeriaid vahetati iga järgneva lapse korral: 98 last jutustasid teisena palliloo ja kolmandana lumememmeloo, 91 juhul oli pildiseeriate järjekord vastupidine.

1. Vahendamata narratiiv eelneva mudelita (edaspidi *1. narratiiv*). Testi läbiviija asetas lapse ette lauale kelguloo seeria esimese pildi ning palus ülejäänud pildid järjestada nii, et neist tekiks jutuke. Kui laps järjestas pildid valesti, pani testija need ise õigesse järjekorda. Lapsele anti piltide vaatamiseks aega ning paluti jutustada nende järgi üks lugu.
2. Vahendatud narratiiv (*2. narratiiv*). Testija asetas ükshaaval lapse ette pilte ning paralleelselt luges ette piltide kohta käiva jutukese (palli- või lumememmeloo). Peale jutu kuulamist paluti lapsel lugu ümber jutustada.
3. Vahendamata narratiiv eelneva mudeliga (*3. narratiiv*). Testimise protseduur oli sarnane 1. narratiivile. Lapsed jutustasid lumememmeloo- või palliloo (vastavalt sellele, kumba nad 2. narratiivina ei jutustanud). Narratiivi mudeliks oli 2. narratiivi aluseks olnud originaaltekst.

6.4. Narratiivide kodeerimine

Analüüs hõlmas narratiivide mikrostruktuuri – lausungite keskmist pikkust ja grammatilist õigust, ning makrostruktuuri – infoüksuste hulka ja kompleksust.

1. Lausungite keskmine pikkus (LP) on tekstis esinenud sõnade hulga ja lausungite hulga suhe. Lausungiteks peeti järgmisi süntaktilisi üksusi: 1) lihtlaused, 2) pealaused koos sellele alistuvate kõrvallausetega, 3) elliptilised lausungid, 4) väljendid, milles on puudu obligatoorsed lauseliikmed, kuid mis on eelnevast või järgnevast lausungist pauside abil eraldatud. Rinnastusseoses olevaid osalaused loeti eraldi lausungiteks. Lausungiteks jaotamisel lähtuti Diana Hughes'i jt (1997: 53) soovitustest, mõiste vastab ingliskeelsele terminile *communication unit*.
2. Grammatikavigade sagedus (GVS) on jutus esinenud vigade ja sõnade koguhulga suhe. Grammatikavigadeks loeti järgmisi eksimusi.

2.1. Süntaksivead: sõnajärjevead (*oli ilm ilus kevad*), obligatoorse lauseliikme puudumine (*Lili ei suutnud Martinit ___*), sidendivead (*Anna kartis selleks et vesi oli sügav*) ja sisulisse konteksti sobimatu sõnavorm (*see tüdruk sõidaks alla ja sõidaks* (pro sõitis) vastu puud).

2.2. Morfoloogiavead: süntaktilisse konteksti sobimatud sõnavormid, s.t vale morfoloogiline vorm (*nad hakkasid juba ülesse minna*) või sõnavormi moonutamine (*nad veeretasi. siis tõstsi.*). Morfoloogiavigade hulka lugesime ka morfofonoloogilised vead (*tooda pro tuua*).

Grammatikavigade hulka ei loetud 1) suulisele kõnele iseloomulikke nähtusi – katkestusi, kordusi, parandusi, lihtsustunud sõnavorme (*ehmund* pro *ehmunud*), sõnaalgulise *h*-i puudumist, 2) hääldusvigu (*tügruk* pro *tüdruk*), 3) leksikaalseid vigu (*solvunud* pro *kurb*) ning 4) pragmaatilisi vigu (nt eksimused sündmuste järjestamisel, arusaamatu sisu vm).

3. Infoüksuste hulk (IÜH) on olulise informatsiooni hulk narratiivis. Kodeerimisel võeti aluseks jutugrammatika mudel (Stein, Glenn 1979), s.t analüüsiti jutugrammatika komponentides sisalduvate infoüksuste hulka (Hayward, Schneider 2000 järgi). Näiteks taustakirjelduse osas arvestati punkte tege- laste, tegevuse aja ja koha ning muu olulise taustainformatsiooni nimeta- mise eest, tegevuse kategooria osas anti punkte tegelaste eesmärgipäraste üksiktoimingute eest jne. Sisult samade väidete eest anti vaid üks punkt. Lõpuks liideti saadud punktid kokku ning analüüsis kasutati jutugramma- tika komponentides sisalduvate infoüksuste koguhulka. Kodeerimisjuh- is vt Adamka (2008).
4. Narratiivi kompleksuse indeks (NKI) sisaldab järgmisi kategooriaid: tegelased, taustakirjeldus, käivitav sündmus, sisemised vastused, plaanid, tegevused, takistused, tagajärjed, dialoog, jutustajapoolsed hinnangud, jutu algust ja lõppu tähistavad vormelmarkerid ning aja- ja põhjusseoseid väl- jendavad keelendid (Petersen jt 2008 järgi). Lisaks nimetatud kategooria- tes sisalduva informatsiooni hulgale võetakse arvesse ka elementidevahelisi seoseid. Kodeerimisjuh- is vt lisa 1.

6.5. Kodeerimise usaldusväarsus

Analüüsitulemuste usaldusväarsuse eesmärgil kodeeris 50 (27%) lapse tekstid mitu kodeerijat. Grammatikavigade osas oli hindajatevaheline usaldusväarsus (väljen- datud Spearmani korrelatsioonikordaja *R* abil) 1., 2. ja 3. narratiivis vastavalt 0,57, 0,61 ja 0,57. Arvestades suulise keele suhteliselt lõtva struktureeritust ja väiksemat allumist mitmesugustele normingutele (võrreldes kirjaliku keelega), võib usal- dusväarsust vigade osas hinnata rahuldavaks. IÜH ja NKI osas olid kodeeringute kokkulangevused kõigis kolmes narratiivis head: Spearmani *R* oli vastavalt 0,85, 0,86 ja 0,84 (IÜH); 0,81, 0,88 ja 0,83 (NKI).

6.6. Statistilised analüüsid

Andmeanalüüsil kasutasime järgmisi statistilise analüüsi meetodeid: eri pildiseeriade alusel jutustatud narratiivide (2. ja 3. narratiivi) struktuurinäitajate (LP, GVS, IÜH, NKI) keskmiste väärtuste võrdlemisel t-testi, EK ja AK rühmade narratiivide struktuurinäitajate keskmiste väärtuste võrdlemisel korduvmõõtmistega dispersioonanalüüsi ANOVA, mille *post hoc* testina (erinevuste mustri väljatoomiseks) kasutasime Bonferroni testi. Tunnuste mõju suuruse hindamisel olid kasutusel Cohen'i d (t-testi ja Bonferroni testi korral) ja osaline eta ruut η_p^2 (ANOVA korral). Statistiliste testide puhul oli olulisuse nivoo $\alpha = 0,05$. Mõju suurust hindasime väikeseks ($d \geq 0,2$, $\eta_p^2 < 0,06$), keskmiseks ($d \geq 0,5$, $\eta_p^2 < 0,14$) või suureks ($d \geq 0,8$, $\eta_p^2 \geq 0,14$) (Vacha-Haase, Thompson 2004, Cohen jt 2007).

7. Tulemused

7.1. Pildiseeriade võrdlused

Eesmärgiga selgitada, kas kahe pildiseeria (palli- ja lumememmeseeria) alusel laste poolt koostatud narratiivid on analüüsitavate tunnuste osas sarnased ehk kas vastavaid narratiive saaks analüüsida koos, võrdlesime eri seeriade alusel koostatud EK laste jutustuste struktuurinäitajate keskmisi väärtusi (LP, GVS, IÜH, NKI) t-testi abil. LP, GVS ja IÜH osas palli- ja lumememmeseeria alusel koostatud narratiivide vahel statistiliselt olulisi erinevusi ei esinenud (nii 2. kui ka 3. narratiivis LP, GVS ja IÜH osas $p > 0,05$), erinevused ilmsesid vaid NKI-s (2. narratiiv: $t(160) = 2,48$, $p = 0,014$, $d = 0,39$; 3. narratiiv: $t(160) = 1,98$, $p = 0,049$, $d = 0,32$).

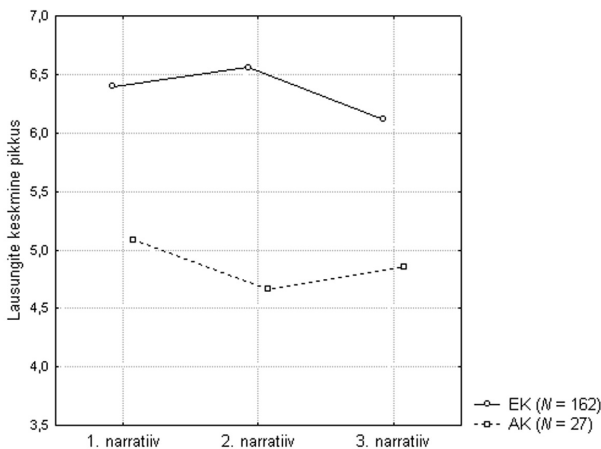
Kuna LP, GVS ja IÜH pooltest palli- ja lumememmejutud ei erinenud, analüüsisime eri seeriade alusel koostatud narratiive nende kategooriate osas edaspidi koos. NKI analüüsid viisime palli- ja lumememmeseeria järgi koostatud narratiivides läbi eraldi. Narratiivide struktuurinäitajate andmed EK ja AK rühmades on esitatud tabelis 1.

Tabel 1. Lausungite keskmine pikkus, grammatikavigade sagedus, infoüksuste hulk ja narratiivi kompleksuse indeks eakohase kõnearenguga ja alakõnega laste narratiivides: aritmeetilised keskmised ja standardhälbed (sulgudes)

Narratiivide struktuurinäitajad	1. narratiiv		2. narratiiv		3. narratiiv	
	EK	AK	EK	AK	EK	AK
Lausungite keskmine pikkus	6,40 (1,59)	5,09 (1,33)	6,56 (1,15)	4,66 (1,03)	6,11 (1,26)	4,85 (1,16)
Grammatikavigade sagedus	0,04 (0,04)	0,14 (0,11)	0,03 (0,03)	0,10 (0,08)	0,03 (0,03)	0,10 (0,07)
Infoüksuste hulk	7,04 (1,87)	5,70 (1,07)	10,04 (2,31)	6,85 (2,14)	7,88 (1,95)	6,07 (1,69)
Narratiivi kompleksuse indeks (rühm A)	7,81 (2,15)	6,17 (1,47)	14,09 (2,73)	9,50 (2,65)	9,98 (3,06)	6,25 (0,75)
Narratiivi kompleksuse indeks (rühm B)	7,99 (2,27)	6,73 (1,71)	12,92 (3,28)	9,93 (2,60)	10,80 (2,09)	9,20 (1,74)

7.2. Lausungite keskmine pikkus

EK lapsed koostasid jutustamisel lausungeid, mille sõnade arv ületas keskmiselt kuus sõna, AK lastel jäi LP viie sõna piiresse (vt joonis 1). Dispersioonanalüüsi tulemustest selgub, et kõnearengu taseme mõju oli statistiliselt oluline ($F(1,187) = 56,12$, $p < 0,001$, $\eta_p^2 = 0,23$), *post hoc* test näitas, et EK ja AK lastel erines LP kõikides jutustustes (1. narratiiv: $p = 0,003$, $d = 0,90$; 2. narratiiv: $p < 0,001$, $d = 1,74$; 3. narratiiv: $p = 0,005$, $d = 1,04$). Statistiliselt oluliseks ei osutunud jutustamisviisi mõju ($F(2,374) = 1,23$, $p > 0,05$, $\eta_p^2 = 0,01$) ega kõnearengu taseme ja jutustamisviisi koosmõju ($F(2, 374) = 2,26$, $p > 0,05$, $\eta_p^2 = 0,01$). Seega, LP oli AK rühma narratiivides oluliselt väiksem võrreldes EK lastega ning jutustamisviisi LP-le kummaski rühmas mõju ei avaldanud.



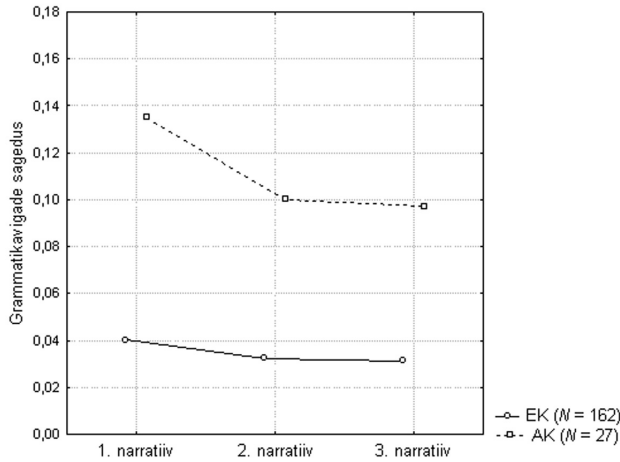
Joonis 1. Lausungite keskmine pikkus sõnades eakohase (EK) ja alakõnega (AK) laste narratiivides

7.3. Grammatikavigade sagedus

GVS erines lasterühmade ja eri viisil jutustatud narratiivide vahel mitmes mõttes (vt joonis 2). Statistiliselt oluline oli nii kõnearengu taseme mõju ($F(1,187) = 123,98$, $p < 0,001$, $\eta_p^2 = 0,40$), jutustamisviisi mõju ($F(2,374) = 11,70$, $p < 0,001$, $\eta_p^2 = 0,06$) kui ka kõnearengu taseme ja jutustamisviisi koosmõju ($F(2, 374) = 4,62$, $p = 0,011$, $\eta_p^2 = 0,02$). *Post hoc* test näitas, et GVS oli kõikides narratiivides AK rühmas suurem kui EK lastel ($p < 0,001$, $d > 0,8$). EK rühmas vigade sagedus kolme narratiivi lõikes ei muutunud ($p > 0,05$, $d < 0,28$), küll aga erines AK lastel vigade hulk 1. ja 2. narratiivi ($p = 0,007$, $d = 0,38$) ning 1. ja 3. narratiivi vahel ($p = 0,002$, $d = 0,42$).

Grammatikavigade liike analüüsisime vähendatud valimis, kaasates analüüsi 44 juhuslikult valitud EK lapse jutud ning kõigi 27 AK lapse jutud. Mõlemas rühmas olid ülekaalus süntaksivead (EK 64%, AK 69%), morfoloogiavead moodustasid umbes kolmandiku vigadest (EK 36%, AK 31%).

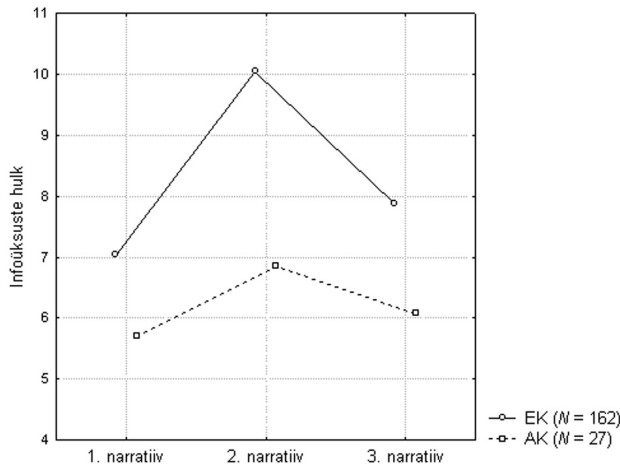
Tulemused näitasid seega, et GVS oli AK rühmas EK lastega võrreldes kõikides narratiivides oluliselt suurem, vigade jaotumine rühmades oli sarnane. AK rühmas langes GVS 2. ja 3. narratiivis oluliselt võrreldes 1. narratiiviga.



Joonis 2. Grammatikavigade sagedus eakohase (EK) ja alakõnega (AK) laste narratiivides

7.4. Infoüksuste hulk

IÜH osas ilmselt oluline kõnearengu taseme mõju ($F(1,187) = 44,34$, $p < 0,001$, $\eta_p^2 = 0,19$), jutustamisviisi mõju ($F(2,374) = 41,48$, $p < 0,001$, $\eta_p^2 = 0,18$) ning kõnearengu taseme ja jutustamisviisi koosmõju ($F(2,374) = 8,48$, $p < 0,001$, $\eta_p^2 = 0,04$). *Post hoc* test osutas EK ja AK laste narratiivide IÜH olulistele erinevustele 2. ja 3. narratiivis (vastavalt $p < 0,001$, $d = 1,43$ ja $p = 0,018$, $d = 0,99$), 1. narratiivi puhul ei ületanud erinevus olulisuse nivood ($p > 0,05$, $d = 0,88$). Jutustuste lõikes selgus, et EK lastel tõusis IÜH (1. narratiiviga võrreldes) oluliselt nii 2. kui ka 3. narratiivis (vastavalt $p < 0,001$, $d = 1,43$; $p < 0,001$, $d = 0,44$). Ka AK rühmas tõusis IÜH 2. ja 3. narratiivis, kuid mitte statistiliselt olulisel määral (vastavalt $p > 0,05$, $d = 0,68$ ja $p > 0,05$, $d = 0,26$).



Joonis 3. Infoüksuste hulk eakohase (EK) ja alakõnega (AK) laste narratiivides

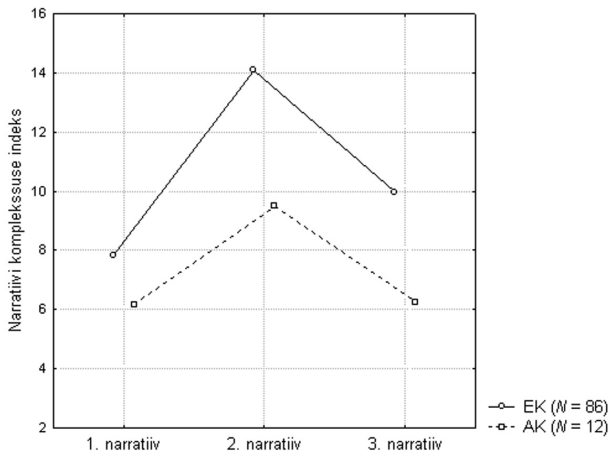
Analüüs näitas seega, et IÜH oli 2. ja 3. narratiivis AK rühmas oluliselt väiksem kui EK lastel ning jutustamisviisi statistiliselt oluline mõju ilmnes vaid EK laste puhul, kelle 2. ja 3. narratiivid olid 1. jutustusega võrreldes oluliselt inforikkamad.

7.5. Narratiivi kompleksuse indeks

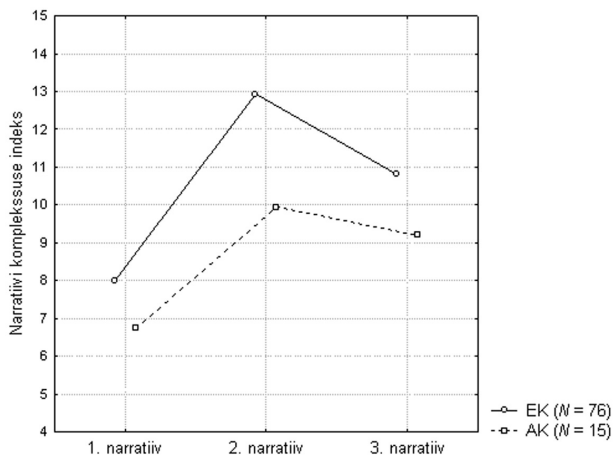
NKI osas viisime analüüsid läbi kahes rühmas: rühma A moodustasid lapsed, kes jutustasid teisea pallilugu ja kolmandana lumememmelugu ($N = 98$), rühmas B oli 2. ja 3. narratiivi aluseks olevate seeriade järjekord vastupidine ($N = 91$). Dispersioonanalüüsi tulemusel selgus, et statistiliselt oluline oli kõnearengu taseme mõju (rühmas A: $F(1,96) = 30,13, p < 0,001, \eta_p^2 = 0,24$; rühmas B: $F(1,89) = 13,93, p < 0,001, \eta_p^2 = 0,14$) ning jutustamisviisi mõju (A: $F(2,192) = 62,97, p < 0,001, \eta_p^2 = 0,40$; B: $F(2,178) = 48,34, p < 0,001, \eta_p^2 = 0,35$). Kõnearengu taseme ja jutustamisviisi koosmõju oli oluline rühmas A ($F(2,192) = 5,71, p = 0,004, \eta_p^2 = 0,06$), rühmas B ilmnes nõrk koosmõju tendents, jäädes siiski statistiliselt mitteoluliseks ($F(2,178) = 2,39, p = 0,095, \eta_p^2 = 0,03$).

Rühmas A erines NKI EK ja AK laste vahel statistiliselt olulisel määral 2. ja 3. narratiivis (vastavalt $p < 0,001, d = 1,71$ ja $p = 0,009, d = 1,67$), 1. narratiivis ei osutunud erinevused oluliseks ($p > 0,05, d = 0,90$). Rühmas B erines NKI EK ja AK laste vahel oluliselt vaid 2. narratiivis ($p = 0,021, d = 1,01$), 1. ja 3. narratiivi puhul ei olnud erinevused statistiliselt olulised (vastavalt $p > 0,05, d = 0,63$ ja $p > 0,05, d = 0,83$).

Kolme narratiivi võrdluse selgus, et EK lastel tõusis NKI oluliselt (1. narratiiviga võrreldes) nii 2. narratiivis (A: $p < 0,001, d = 2,56$; B: $p < 0,001, d = 1,75$) kui ka 3. narratiivis (A: $p < 0,001, d = 0,82$; B: $p < 0,001, d = 1,29$). AK lastel oli NKI 2. narratiivis samuti oluliselt kõrgem kui 1. narratiivis (A: $p = 0,002, d = 1,56$; B: $p < 0,001, d = 1,45$), 3. narratiivis olid A ja B rühmas tulemused erinevad, osutudes oluliseks vaid rühmas B (A: $p > 0,05, d = 0,07$; B: $p = 0,023, d = 1,43$).



Joonis 4. Narratiivi kompleksuse indeks eakohase (EK) ja alakõnega (AK) laste narratiivides (rühm A)



Joonis 5. Narratiivi kompleksuse indeks eakohase (EK) ja alakõnega (AK) laste narratiivides (rühm B)

Analüüs näitas seega, et NKI oli AK lastel EK lastega võrreldes oluliselt madalam 2. narratiivis (nii A kui B rühmas) ja 3. narratiivis (rühmas A). 1. narratiivis ei erinenud NKI EK ja AK rühmade vahel statistiliselt olulisel määral. Jutustamisviisi mõju ilmnes selgelt vaid EK lastel, kelle 2. ja 3. narratiivi NKI oli 1. narratiiviga võrreldes oluliselt suurem.

8. Arutelu ja kokkuvõte

Vaatluse all olid 6–7 aasta vanuste eakohase kõnearenguga ja alakõnega eesti laste suulised narratiivid: vahendamata jutustused ilma eelneva narratiivi mudelita, vahendatud jutustused ning vahendamata jutustused eelneva narratiivi mudeliga. Analüüsisime nimetatud kolmel viisil jutustatud narratiivide mikrostruktuuri – lausungite keskmist pikkust ja grammatikavigu, ning makrostruktuuri – jutugrammatika mudelil (Stein, Glenn 1979) põhineva olulise informatsiooni hulka ja narratiivi kompleksust, mis hõlmab jutustuse episoodilist struktuuri, jutustajapoolseid hinnanguid ning lausungeid siduvaid keelendeid (Petersen jt 2008 järgi).

Eeldasime, et kõigil kolmel viisil jutustatud narratiivid on AK lastel EK lastega võrreldes madalamal tasemel kõikide analüüsitavate näitajate osas. Hüpotees leidis kinnitust narratiivi mikrostruktuuri osas: AK laste LP oli oluliselt väiksem ning GV sagedus oluliselt kõrgem, võrreldes EK laste narratiividega. Tulemused on kooskõlas inglise keelel põhinevate uurimistulemustega lausungipikkuse (Scott, Windsor 2000, Bishop, Donlan 2005, Hayward jt 2007) ning grammatilise õigsuse kohta (Scott, Windsor 2000, Norbury, Bishop 2003, Pearce jt 2003, Fey jt 2004). Makrostruktuuri näitajate osas ilmnes samuti kõnearengu taseme mõju, mis siiski ei osutunud kõikide narratiivide osas statistiliselt oluliseks. Nimelt sisaldasid EK laste jutustused AK laste narratiividega võrreldes oluliselt rohkem infoüksusi vaid vahendatud narratiivis ja eelnevalt mudeldatud vahendamata narratiivis, eelnevalt mudeldamata vahendamata narratiivide IÜH erinevus ei olnud statistiliselt oluline. Ka narratiivi kompleksuse osas leidis hüpotees kinnitust vaid osaliselt,

s.t vahendatud narratiivide osas, mille NKI oli AK lastel oluliselt väiksem kui EK lastel. Eelnevalt mudeldamata narratiivide NKI ei olnud lasterühmade vahel oluliselt erinev, eelnevalt mudeldatud vahendamata narratiivide NKI erines oluliselt vaid rühmas A. Tulemused näitavad seega, et narratiivide mikrostruktuur (LP ja GVS) oli AK lastel EK lastega võrreldes oluliselt madalamal tasemel kõigil kolmel viisil jutustatud narratiivides. Narratiivide makrostruktuur (IÜH ja NKI) oli AK lastel oluliselt madalamal tasemel vahendatud jutustustes ning eelnevalt mudeldatud vahendamata jutustustes (osaliselt). Seega võib saadud tulemuste põhjal nõustuda Betty Lilesi jt (1995) ning Cheryl Scotti ja Jennifer Windsori (2000) väitega, et just lausetasandi näitajad eristavad hästi esmase alakõnega lapsi EK lastest, kuigi ka narratiivide makrostruktuur on AK korral eakohaselt arenenud lastega võrreldes primitiivsem.

Teise hüpoteesina eeldasime, et vahendatud narratiivid eristavad EK ja AK rühmi vahendamata jutustustest paremini. Mikrostruktuuri näitajate osas ei leidnud hüpotees kinnitust. Nimelt olid muutused LP osas kolme jutustuse lõikes väiksed ning ei osutunud statistiliselt oluliseks mitte kummaski lasterühmas. Grammatikavigade osas selgus, et EK rühmas oli GVS kolme narratiivi lõikes stabiilselt madal, AK rühmas oli see aga vahendatud ja eelnevalt mudeldatud vahendamata jutustuses oluliselt madalam, võrreldes eelnevalt mudeldamata jutustusega. Vigade vähenemine AK rühmas võib olla seletatav võimalusega, et 1. narratiivi puhul ei olnud AK lastele jutustamise ülesanne piisavalt arusaadav, mistõttu lapsed ei olnud motiveeritud koostama keeleliselt korrektset jutukest. Vahendatud jutustuse aluseks olev originaaltekst toimus lastele tõenäoliselt keelelise mudelina nii 2. kui 3. jutustuse koostamisel. Lisaks võib oletada, et vahendatud jutustuses reprodutseerisid AK lapsed osaliselt originaaltekstis esinenud grammatilisi vorme, mis tingis ka vigade vähenemise. Vigade langustrendi puudumine EK laste rühmas on seletatav niigi madala vigade hulgaga juba 1. narratiivis. Makrostruktuuri näitajate osas leidis hüpotees kinnitust: nii IÜH kui ka NKI erinevused olid EK ja AK rühmade vahel suurimad vahendatud jutustustes, näidates, et EK lastel tõusis vahendatud narratiivis informatsiooni hulk ja kompleksus AK lastega võrreldes oluliselt enam. Saadud tulemused on kooskõlas varasemate AK ja EK laste narratiivide makrostruktuuri uurimustega (Merritt, Liles 1987, 1989), toetades seisukohta, et vahendatud jutustus toob vahendamata jutustustega võrreldes paremini esile AK laste madalamad tekstiloomeskused.

Kolmandaks eeldasime, et eelnevalt mudeldatud vahendamata jutustus eristab EK ja AK rühmi paremini kui eelneva mudelita vahendamata jutustus: oletasime, et narratiivide makrostruktuur paraneb mudeldamise tulemusena enam EK lastel, lausetasandi näitajate osas me olulisi muutusi ei eeldanud. Hüpotees leidis kinnitust LP, IÜH ja NKI osas, GVS osas mitte. Nimelt selgus, et näidisjutu esitamisel ei olnud mõju LP-le kummaski kõnearengu taseme rühmas, grammatikavigade langustrend esines aga hoopis AK laste rühmas. Nagu eelnevalt põhjendatud, on üks võimalik seletus ülesande ebapiisav mõistmine 1. jutustuse koostamisel AK rühmas. IÜH tõusis EK laste jutustustes mudeldamise tagajärjel olulisel määral, AK rühmas aga mitte. Ka NKI oli EK laste eelnevalt mudeldatud vahendamata narratiivides oluliselt suurem, võrreldes eelnevalt mudeldamata jutustustega. AK rühmas jäid tulemused vastuoluliseks: rühmas B NKI tõusis, rühmas A jäi samaks. Hoolimata ebapiisavast statistilisest tõestusest NKI osas (mis võib olla tingitud väga väikestest AK alarüh-

madest), on tulemused siiski kooskõlas Elizabeth Peña jt (2006) uurimusega, kus leiti, et jutustamisele eelneval narratiivi formaadi õpetamisel on positiivne mõju eelkõige EK laste tekstiloomele makrostruktuuri tasandil ning AK lapsed õpivad mudeldamisest vähem. Eri viisil jutustatud narratiivide mikrostruktuuri analüüs näitas, et kuigi lausungite keskmist pikkust näidisjutu esitamine ei mõjutanud, oli sellel positiivne efekt lausungite grammatilisele õigsusele AK laste rühmas.

Käesoleva töö väärtus seisneb laste tekstiloomeoskuste ja eri jutustamisviiside seostamises, olles abiks logopeedidele alakõne diagnostika meetodite valikul. Tulemuste paikapidavust tuleks siiski edaspidi kontrollida suuremal valimil, seda eelkõige AK laste osas. Tulemuste alusel võib kinnitada, et narratiivid võimaldavad saada olulist informatsiooni laste kõnearengu taseme kohta. Jutustamisoskuse hindamiseks narratiivi mikrostruktuuri tasandil sobivad kõigil kolmel viisil jutustatud narratiivid, makrostruktuuri tasandil on sobivaim vahendatud narratiiv, millele järgneb eelnevalt mudeldatud vahendamata narratiiv. Võimalikult täpse ülevaate saamiseks lapse kõnearengu tasemest tuleks siiski kombineerida erinevaid jutustamisviise ning narratiivseid ülesandeid teiste hindamismeetoditega.

Viidatud kirjandus

- Adamka, Aive 2008. 6–8 aasta vanuste laste suuliste jutustuste makrostruktuur. Magistritöö. Tartu: Tartu Ülikool.
- Argus, Reili 2008. Eesti keele muutmorfoloogia omandamine. Tallinna Ülikooli humanitaarteaduste dissertatsioonid, 19. Tallinn: TLÜ Kirjastus.
- Bishop, Dorothy V. M.; Donlan, Chris 2005. The role of syntax in encoding and recall of pictorial narratives: Evidence from specific language impairment. – *British Journal of Developmental Psychology*, 23, 25–46. doi:10.1348/026151004X20685
- Bishop, Dorothy V. M. 1997. *Uncommon Understanding: Development and Disorders of Language Comprehension in Children*. Hove: Psychology Press.
- Botting, Nicola 2002. Narrative as a tool for the assessment of linguistic and pragmatic impairments. – *Child Language Teaching and Therapy*, 18, 1–21. doi:10.1191/0265659002ct2240a
- Bruner, Jerome 1985. Narrative and paradigmatic modes of thought. – Elliot Eisner (Ed.). *Learning and Teaching the Ways of Knowing*. Chicago: University of Chicago Press, 97–115.
- Cohen, Louis; Manion, Laurence; Morrison, Keith 2007. *Research Methods in Education*. London, New York: Routledge.
- Fey, Marc E.; Catts, Hugh W.; Proctor-Williams, Kerry; Tomblin, J. Bruce; Zhang, Xuyang 2004. Oral and written story composition skills of children with language impairment. – *Journal of Speech, Language, and Hearing Research*, 47, 1301–1318. doi:10.1044/1092-4388(2004/098)
- Hayward, Denyse V.; Gillam, Ronald B.; Lien, Phuong 2007. Retelling a script-based story: Do children with and without language impairments focus on script and story elements? – *American Journal of Speech-Language Pathology*, 16, 235–245. doi:10.1044/1058-0360(2007/028)
- Hayward, Denyse V.; Schneider, Phyllis 2000. Effectiveness of teaching story grammar knowledge to pre-school children with language impairment: An exploratory study. – *Child Language Teaching and Therapy*, 16, 255–284. doi:10.1191/026565900680410215
- Hughes, Diana; McGillivray, LaRae; Schmidek, Mark 1997. *Guide to Narrative Language: Procedures for Assessment*. Eau Claire, Wisconsin: Thinking Publications.

- John, Shonna F.; Lui, Mariko; Tannock, Rosemary 2003. Children's story retelling and comprehension using a new narrative resource. – *Canadian Journal of School Psychology*, 18, 91–113. doi:10.1177/082957350301800105
- Justice, Laura M.; Bowles, Ryan P.; Kaderavek, Joan N.; Ukrainetz, Teresa A.; Eisenberg, Sarita L.; Gillam, Ronald B. 2006. The index of narrative microstructure: A clinical tool for analyzing school-age children's narrative performances. – *American Journal of Speech-Language Pathology*, 15, 177–191. doi:10.1044/1058-0360(2006/017)
- Karlep, Karl 1997. Eesti vajab kõnepuuete ühtset logopeedia-klassifikatsiooni. – *Eripedagoogika: Logopeedia ja Emakeel*, 4–15.
- Karlep, Karl 2003. *Kõnearendus*. Tartu: Tartu Ülikooli Kirjastus.
- Labov, William 1972. *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.
- Leonard, Laurence B. 1998. *Children with Specific Language Impairment*. Cambridge, MA: MIT Press.
- Liles, Betty Z. 1987. Episode organization and cohesive conjunctives in narratives of children with and without language disorder. – *Journal of Speech and Hearing Research*, 30, 185–196.
- Liles, Betty Z.; Duffy, Robert J.; Merritt, Donna D.; Purcell, Sherry L. 1995. Measurement of narrative discourse ability in children with language disorders. – *Journal of Speech and Hearing Research*, 38, 415–425.
- Merritt, Donna D.; Liles, Betty Z. 1987. Story grammar ability in children with and without language disorder: Story generation, story retelling, and story comprehension. – *Journal of Speech and Hearing Research*, 30, 539–552.
- Merritt, Donna D.; Liles, Betty Z. 1989. Narrative analysis: Clinical applications of story generation and story retelling. – *Journal of Speech and Hearing Disorders*, 54, 429–438.
- Norbury, Courtenay F.; Bishop, Dorothy V. M. 2003. Narrative skills of children with communication impairments. – *International Journal of Language and Communication Disorders*, 38, 287–313. doi:10.1080/13682031000018133
- Padrik, Marika 2006. Milles seisneb kõnearengu puude spetsiifilisus? – *Eripedagoogika: Logopeedia ja Emakeel*, 26, 13–20.
- Pajuste, Maret 2007. *Lausungite keeleline seostamine 1. klassi eakohase arenguga ja õpiskustega õpilastel pildiseeria järgi jutustades*. Magistritöö. Tartu: Tartu Ülikool.
- Pankratz, Mary E.; Plante, Elena; Vance, Rebecca; Insalaco, Deborah M. 2007. The diagnostic and predictive validity of the Renfrew Bus Story. – *Language, Speech, and Hearing Services in Schools*, 38, 390–399. doi:10.1044/0161-1461(2007/040)
- Pearce, Wendy M.; McCormack, Paul F.; James, Deborah G. H. 2003. Exploring the boundaries of SLI: Findings from morphosyntactic and story grammar analyses. – *Clinical Linguistics and Phonetics*, 17, 325–334. doi:10.1080/0269920031000080019
- Peña, Elizabeth D.; Gillam, Ronald B.; Malek, Melynn; Ruiz-Felter, Roxanna; Resendiz, Maria; Fiestas, Christine; Sabel, Tracy 2006. Dynamic assessment of school-age children's narrative ability: An experimental investigation of classification accuracy. – *Journal of Speech, Language, and Hearing Research*, 49, 1037–1057. doi:10.1044/1092-4388(2006/074)
- Petersen, Douglas B.; Gillam, Sandra L.; Gillam, Ronald B. 2008. Emerging procedures in narrative assessment: The index of narrative complexity. – *Topics in Language Disorders*, 28, 115–130.
- Peterson, Carole; Dodsworth, Pamela 1991. A longitudinal analysis of young children's cohesion and noun specification in narratives. – *Journal of Child Language*, 18, 397–415. doi:10.1017/S0305000900011120
- Reilly, Judy; Losh, Molly; Bellugi, Ursula; Wulfeck, Beverly 2004. "Frog, where are you?" Narratives in children with specific language impairment, early focal brain injury,

- and Williams syndrome. – *Brain and Language*, 88, 229–247. doi:10.1016/S0093-934X(03)00101-9
- Reuterskiöld Wagner, Christina; Sahlén, Birgitta; Nettelblatt, Ulrika 1999. What's story? Narration and comprehension in Swedish preschool children with language impairment. – *Child Language Teaching and Therapy*, 15, 113–137.
- RHK-10 Psüühika- ja käitumishäired. Kliinilised kirjeldused ja diagnostilised juhised. Tõlge eesti keelde. Toim. Veiko Vasar. Tartu: TÜ Kirjastus.
- Ripich, Danielle N.; Griffith, Penny L. 1988. Narrative abilities of children with learning disabilities and nondisabled children: Story structure, cohesion, and propositions. – *Journal of Learning Disabilities*, 21, 165–173. doi:10.1177/002221948802100309
- Ruul, Riina 2009. 5–6-aastaste laste jutustuste sidusus ja selle hindamine. Magistritöö. Tartu: Tartu Ülikool.
- Schneider, Phyllis; Hayward, Denyse; Dubé, Rita V. 2006. Storytelling from pictures using the Edmonton Narrative Norms Instrument. – *Journal of Speech-Language Pathology and Audiology*, 30, 224–238.
- Scott, Cheryl M.; Windsor, Jennifer 2000. General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. – *Journal of Speech, Language, and Hearing Research*, 43, 324–339.
- Shapiro, Lauren R.; Hudson, Judith A. 1991. Tell me a make-believe story: Coherence and cohesion in young children Picture-elicited narratives. – *Developmental Psychology*, 27, 960–974. doi:10.1037/0012-1649.27.6.960
- Sleight, Christine C.; Printz, Philip M. 1985. Use of abstracts, orientations, and codas in narration by language-disordered and nondisordered children. – *Journal of Speech and Hearing Disorders*, 50, 361–371.
- Stein, Nancy L.; Glenn, Christine G. 1979. An analysis of story comprehension in elementary school children. – Roy O. Freedle (Ed.). *New Directions in Discourse Processing*, 2. Norwood, NJ: Ablex, 53–120.
- Tilstra, Janet; McMaster, Kristen 2007. Productivity, fluency, and grammaticality measures from narratives: Potential indicators of language proficiency? – *Communication Disorders Quarterly*, 29, 43–53. doi:10.1177/1525740108314866
- Ukrainetz, Teresa A.; Justice, Laura M.; Kaderavek, Joan N.; Eisenberg, Sarita L.; Gillam, Ronald B.; Harm, Heide M. 2005. The development of expressive elaboration in fictional narratives. – *Journal of Speech, Language, and Hearing Research*, 48, 1363–1377. doi:10.1044/1092-4388(2005/095)
- Vacha-Haase, Tammi; Thompson, Bruce 2004. How to estimate and interpret various effect sizes. – *Journal of Counseling Psychology*, 51, 473–481. doi:10.1037/0022-0167.51.4.473
- Vance, Maggie 2008. Short-term memory in children with developmental language disorder. – Courtenay F. Norbury, J. Bruce Tomblin, Dorothy V. M. Bishop (Eds.). *Understanding Developmental Language Disorders: From Theory to Practice*. New York: Psychology Press, 23–38.
- van Dijk, Teun A.; Kintsch, Walter 1983. *Strategies of Discourse Comprehension*. New York: Academic Press.

Lisa. Narratiivi kompleksuse indeksi kodeerimisjuhised
(Petersen jt 2008 järgi)

Narratiivi element	0 punkti	1 punkt	2 punkti	3 punkti
Tegelased	Peategelasi ei ole nimetatud või on kasutatud asesõnu.	Sisaldab vähemalt ühte peategelast, mis on märgitud mit-spetsiifiliselt.	Sisaldab ühte peategelast, kes on märgitud spetsiifilise nimega.	Sisaldab rohkem kui ühte peategelast, kes on märgitud spetsiifilise nimega.
Taustakirjeldus*	Ei ole viidatud ajale, kohale ega tegevusele.	Sisaldab ühte viidet ajale, kohale või tegevusele.	Sisaldab kahte või enam viidet ajale, kohale või tegevusele.	
Käivitav sündmus	Sündmust ega probleemi ei ole nimetatud.	Sisaldab vähemalt ühte otsest viidet sündmusele või probleemile, mis võiks kutsuda esile tegelase reaktsiooni, kuid puudub tegelase reaktsioon, mis oleks otseselt seotud selle sündmusega.	Sisaldab vähemalt ühte otsest viidet sündmusele või probleemile, mis kutsub esile tegelase reaktsiooni.	Kaks või enam eraldi viidatud sündmust või probleemi, mis kutsuvad esile tegelase reaktsiooni.
Sisemine vastus	Ei ole otsest väljendit tegelase psühholoogilise seisundi kohta.	Üks otsene väljend tegelase psühholoogilise seisundi kohta, mis ei ole aga põhjuslikult seotud käivitava sündmusega.	Üks või enam otsest väljendit tegelase psühholoogilise seisundi kohta, mis on põhjuslikult seotud käivitava sündmusega.	
Sisemine plaan	Ei ole otsest viidet tegelase plaanile tegutseda või lahendada olukorda.	Üks otsene väljend, kuidas tegelane võiks lahendada olukorda.	Kaks otsest väljendit, kuidas tegelane võiks käituda või lahendada olukorda.	Kolm või enam otsest väljendit, kuidas tegelane võiks käituda või lahendada olukorda.
Tegevus	Ei sisalda peategelas(t)e poolt tehtud tegevusi.	Sisaldab viiteid peategelas(t)e tegevustele, mis ei ole seotud käivitava sündmusega.	Sisaldab viiteid peategelas(t)e tegevustele, mis on otseselt seotud käivitava sündmusega.	
Takistus	Ei ole takistusi.	Üks takistus plaani või tegevuse teostumisel.	Kaks takistus plaani või tegevuse teostumisel.	

Narratiivi element	0 punkti	1 punkt	2 punkti	3 punkti
Tagajärg	Tegevuse tagajärke ei ole otseselt väljendatud.	Üks tagajärg.	Kaks tagajärge.	Kolm või enam tagajärge.
Jutustuse alguse ja lõpu markerid (nt <i>lõpp, oli kord, elas kord jne</i>)	Ei ole markereid	Üks marker.	Kaks või enam markerit.	
Ajamarkerid* (nt <i>kui, siis, peale seda, kohe jne</i>)	Ei ole ajamarkereid.	Üks ajamarker või ajamarkeri <i>siis</i> korduv kasutamine.	Kaks või enam ajamarkerit.	
Põhjusseoid markeerivad keelendid (nt <i>sellepärast, sest, sellepärast et jne</i>)	Ei ole põhjusseoid markeerivaid keelendeid.	Üks põhjusseost markeeriv keelend.	Kaks või enam põhjusseost markeerivat keelendit.	
Dialog	Ei ole dialoogi	Üks tegelane ütleb midagi.	Kaks või enam tegelast vestlevad.	
Jutustajapoolsed hinnangud	Ei ole hinnanguid.	Üks hinnang.	Kaks või enam hinnangut.	

* Märkus. Kodeerimissüsteemi on käesolevas uurimuses kohandatud.

Piret Soodla (Tallinna Ülikool, Tartu Ülikool) uurimistöö valdkonnad on suulise ja kirjaliku kõne areng, selle seosed akadeemilise edukusega ning vaimsete võimetega, kõne arengu mahajäämuse diagnostika.

piret.soodla@ut.ee

Eve Kikase (Tallinna Ülikool, Tartu Ülikool) uurimistöö valdkonnad on kooliõpetuse mõju mõtlemisele, laste ja täiskasvanute mõisted, erivajadustega laste võimed ja nende hindamine, laste agressiivsus ja selle areng.

eve.kikas@ut.ee

Renate Pajusalu (Tartu Ülikool) uurimisvaldkonnad on eesti keele semantika ja pragmaatika, esimese ja teise keele omandamise erinevad aspektid.

renate.pajusalu@ut.ee

Aive Adamka (Tartu Hiie Kool) töötab eesti keele ja kõneravi õpetajana, uurib tekstiloomeskuste arengut lastel.

Aive.Adamka@mail.ee

Sirli Parm (Tartu Ülikool) uurimisvaldkonnad on lapse keele omandamine, leksikaalne semantika.

sirli.parm@ut.ee

SELF-GENERATED AND RETOLD NARRATIVES AS A TOOL OF LANGUAGE ASSESSMENT

Piret Soodla, Eve Kikas, Renate Pajusalu, Aive Adamka, Sirli Parm

Tallinn University, University of Tartu, Tartu Hiie School

The article explores oral narratives of 6–7 year old typically developing (TD) children ($N = 162$) and children with language impairment (LI) ($N = 27$). Narratives were elicited by three methods: generated stories without previous narrative model and with previous narrative model, and retold stories. The analysis comprises narrative microstructure (mean length of utterances, grammatical errors) and macrostructure (quantity of information units, complexity). The purpose is to find out which method suits best for the identification of LI in language assessment. The results show that for microstructure level, all three kinds of stories distinguish children with LI from their peers with TD. For macrostructure level, retold stories enable better distinction between LI and TD as compared to self-generated stories, likewise generated stories with previous narrative model, as compared to stories without previous model.

Keywords: linguistic development, speech disorders, narrative, syntax, morphology, textual structure, Estonian language

KAS PAUSID KANNAVAD EMOTSIOONI?

Kairi Tamuri

Ülevaade. Pausidele on emotsionaalse kõne uurimustes suhteliselt vähe tähelepanu pööratud. Ilmselt seetõttu, et pausid ei ole esmased emotsiooni kandvad akustilised tunnused. Siiski on teada, et pausidel on emotsioonide eristamisel oma roll. Artiklis antakse vastus kahele küsimusele: kas eestikeelse etteloetud emotsionaalse teksti pausid sõltuvad oluliselt teksti emotsioonist ning kas teksti emotsioon on ära tuntav ainult pausierinevuste põhjal.*

Võtmesõnad: emotsioonide akustika, etteloetud tekst, pausigrupid, pauside arv, pauside asukoht, pauside iseloom, pauside kestus, kõne-süntees, eesti keel

1. Sissejuhatus

Emotsioonide lisamine sünteeskõnmesse muudab selle palju loomulikumaks ja ilmekamaks ning parandab mõistetavust. Et emotsioone tekst-kõnesünteesi jaoks modelleerida, on vaja üles leida need akustilised tunnused, mis emotsioone kannavad. Enamik kõnesünteesiga seotud uurimustest väidab, et igal emotsioonil on just talle omased akustilised parameetrid (vt Bachorowski 1999). Kirjanduse põhjal võib öelda, et tähtsateks emotsiooni kandvateks akustilisteks tunnusteks peetakse põhitooni ja selle kontuuri, kestust, hääle kvaliteeti, energiat, kõne kiirust, artikulatsiooni täpsust ja intensiivsust, kuid uuritud on ka emotsionaalse kõne pause (Murray, Arnott 2008, Yildirim jt 2004, ten Bosch 2003, Iida jt 2003, Schröder jt 2001, Montero jt 1998).

Uurimistulemused on näidanud, et pausid võivad emotsiooniti erineda nii kestuse kui ka esinemise sageduse poolest. Näiteks eristuvad ingliskeelses kõnes kõige selgemini kurbuselausete pausid. Nende esinemise sagedus on suurem ja võrreldes teiste emotsioonide pausidega on nad pikema kestusega. (Murray, Arnott 2008, Yildirim jt 2004, Schröder jt 2001)

Eestis kõnesünteesi kontekstis emotsioonide akustikat varem uuritud ei ole. Teiste keelte uurimistulemusi ei saa paraku automaatselt eesti keelde üle võtta, sest

* Uurimus on tehtud riikliku keeletehnoloogia programmi projekti "Eesti emotsionaalse kõne korpus", sihtfinantseeritava teadusteema SF0050023s09 ja Eesti Teadusfondi grandiprojekti ETF7998 raames.

emotsioonide akustiline väljendamine erineb kultuuriti: kõnelemine on kultuuriline tegevus ja igas ühiskonnas on emotsioonide väljendamisel oma reeglid (Burkhardt jt 2006, Douglas-Cowie jt 2003, Scherer 2003). Samuti eelistavad eri keelt kõnelevad inimesed ühte ja sama emotsiooni kuulda sünteeskõnes eri moods, s.t eri keelte puhul tajutakse sünteeskõne loomulikkust erinevalt (Burkhardt jt 2006).

2. Uurimisküsimused ja -materjal

Pausiuurimuse eesmärk oli leida vastus kahele küsimusele: 1) kas eestikeelse ettelõetud emotsionaalse teksti pausid sõltuvad oluliselt teksti emotsioonist¹ ning 2) kas teksti emotsioon on ära tuntav ainult pausierinevuste põhjal.

Esimesele uurimisküsimusele vastuse leidmiseks vaatlesin kolme põhiemotsiooni (kurbuse, viha ja rõõmu) pause. Uurisin pauside arvu, asukohta (sh vastavust kirjavahemärkidele²), iseloomu (hingamisega vs. hingamiseta pausid³) ning kestust. Uurimismaterjaliks olid sel puhul Eesti Keele Instituudi emotsionaalse kõne korpusel pärit ühe naiskeelejuhi ettelõetud ajakirjanduslikud tekstilõigud (vt Altrov 2008). Iga emotsiooni kohta valisin välja keskmiselt 16 lõiku, milles oli 120–123 lauset (vt tabel 1). Sõnesid oli lausetes keskmiselt 8–9, kirjavahemärke oli samuti enam-vähem võrdset. Pausiuurimuse tulemused esitan ptk 3.

Teise uurimisküsimuse materjalina kasutasin lõiku, mille laused olid katseisikud kuulamise põhjal määranud neutraalseks. Selle neutraalse lõigu pausi-parameetreid muutes viisin läbi kuulamistesti, et teada saada, kas teksti emotsioon on ära tuntav ainult pausierinevuste põhjal. Emotsiooni kuulamistesti tulemused esitan ptk 4.

Tabel 1. Pausiuurimuse materjal

	Kurbus	Rõõm	Viha
Lõike	15	16	18
Lauseid	120	120	123
Sõnesid	1053	966	1124
Kirjavahemärke	187	181	201

3. Kolme põhiemotsiooni (kurbuse, viha ja rõõmu) pausid

3.1. Pauside arv ja asukoht

Pause oli kurbuselõikudes 165, rõõmulõikudes 181 ja vihalõikudes 195 (vt tabel 2). Pauside arv olenes eeskätt kirjavahemärkide arvust, sest tavaliselt teeb lugeja pausi just kirjavahemärgi kohal (vt Tamuri 2007: 37, Pajupuu, Kerge 2006). Mujal kui kirjavahemärgi kohal tehtud pause nimetatakse lisapausideks. Näiteks on iseloomulik, et lisapausid kipuvad esinema teatud kohtades, milleks on päris- ja kohanimed, numbrid ning sidesõnad (Pajupuu, Kerge 2006).

¹ Selles töös ei võrreldud emotsioone neutraalse kõnega, kuna vastav võrdlusmaterjal emotsionaalse kõne korpusel puudub.

² Jagasin pausid teksti kirjavahemärkidele vastavuse järgi rühmadeks (nt punktipaus, komapaus jne).

³ Hingamisega paus koosneb sissehingamisest, millele eelneb ja/või järgneb vaikus. Hingamiseta paus on selline sõnade või lausete vaheline vaikus kõnes, mille kestus on vähemalt 30 ms. 30 ms sai valitud seepärast, et materjal oleks võrreldav varasemate eesti keelt puudutavate pausiuurimustega, kus pausi miinimumpikkuseks on võetud 30 ms (vt Pajupuu, Kerge 2006, Tamuri 2007: 14).

Tabel 2. Pauside arv ja asukoht

	Kurbus	Rõõm	Viha
Pauside arv	165	181	195
Pausid kirjavahemärkide kohal (% kõigist pausidest)	92%	92%	90%
Lisapausid (% kõigist pausidest)	8%	8%	10%

Kirjavahemärkide kohal tehtud pauside hulk emotsiooniti ei erinenud: kurbuse ja rõõmu puhul langesid 92% kõikidest pausidest kirjavahemärkide kohale, viha puhul oli see protsent 90.

Emotsiooniti ei erinenud oluliselt ka lisapausid, mis samuti siin uurimuses esinesid kõigil kolmel emotsioonil umbes pooltel kordadel sidesõnade *ja/ning* ees ja päris- või kohanime ees. Pooltel kordadel olid aga lisapausid juhuslikku laadi ning võisid sõltuda kas teksti sisust või lugejast endast (nt tema kopsumahust).

Analüüsi tulemuste põhjal võib öelda, et pauside arvu ja asukoha poolest kurbuse, rõõmu ja viha emotsioonid üksteisest märgatavalt ei erinenud ning kõigil kolmel langes suurem osa pause kirjavahemärkide kohale.

Kui vaadata erinevaid kirjavahemärke, siis oli koma ainus, mille kohal lugeja pausi alati ei teinud (vt tabel 3). See kehtib kõigi kolme emotsiooni kohta. Komad ongi eesti keeles sellised kirjavahemärgid, mille kohal jäetakse paus tihtipeale tegemata. Ka mõned varasemad uurimused kinnitavad sama (vt Tamuri 2007, Pajupuu, Kerge 2006).

Tabel 3. Komad

	Kurbus	Rõõm	Viha
Komade arv	72	67	80
Komad, mille kohal paus tehti (% kõigist komadest)	51%	79%	67,5%

Uurimismaterjalis oli emotsiooniti komasid lausetes enam-vähem võrdselt: kurbuselausetes 72, rõõmulausetes 67 ja vihalausetes 80. Koma kohal tegi lugeja kõige vähem pause kurbuselausetes – umbes pooltel juhtudel. Märkimisväärselt rohkem tegi lugeja koma kohal pause aga viha- ja rõõmulausetes – vastavalt 67,5% ja 79% juhtudest.

3.2. Pauside iseloom

Pauside iseloomu all pean silmas hingamist. Jagasin pausid hingamisega vs. hingamiseta pausideks. Kui pausidele pööratakse kõnesünteesi uurimustes suhteliselt vähe tähelepanu, siis hingamine saab neis veelgi vähem tähelepanu. Samas on näiteks Nick Campbell (1998) väitnud, et hingamise lisamine tõstaks sünteeskõne loomulikkust oluliselt.

Uurisn emotsioonipause hingamise olemasolust lähtuvalt (vt tabel 4).

Tabel 4. Hingamispausid

	Kurbus	Rõõm	Viha
Hingamispauside arv	90	96	115
Hingamispausid (% kõigist pausidest)	54,5%	53%	59%
Hingamisega kirjavahemärgipausid (% kõigist kirjavahemärgipausidest)	57%	56%	61%
Hingamisega lisapausid (% kõigist lisapausidest)	31%	21%	55%

Kõigil kolmel emotsioonil oli hingamispause rohkem kui hingamiseta pause. Enim hingamispause oli vihalausetes, kuid see erinevus teistest emotsioonidest ei ole suur. Olulisem erinevus tuleb sisse lisapausides, kus viha puhul tegi lugeja teiste emotsioonidega võrreldes märgatavalt enam hingamispause.

3.3. Kirjavahemärkide ja hingamispauside vaheline seos

Kõige rohkem hingamispause tegi lugeja kurbuse ja rõõmu puhul lause lõpus (vastavalt 68% ja 67%), viha puhul aga koma kohal (67%) (vt tabel 5). Kõige vähem hingamispause tegi lugeja kurbuse- ja rõõmulõikudes koma kohal (vastavalt 27% ja 36%), vihalõikudes muude kirjavahemärkide (s.t v.a lauselõpp ja koma) kohal (56%).

Tabel 5. Hingamispauside osakaal kirjavahemärkide kohal

	Kurbus	Rõõm	Viha
Lauselõpp (% lauselõpupausidest)	68%	67%	58%
Koma (% komapausidest)	27%	36%	67%
Muud kirjavahemärgid (% muudest kirjavahemärgipausidest)	60%	40%	56%

3.4. Pauside kestus

Pauside kestus on seotud kõnetempoga. Uurimustest on selgunud, et pausid võivad emotsiooniti erineda ka kestuse poolest. Näiteks ingliskeelses kõnes eristuvad kõige selgemini kurbuselausete pausid, mis on võrreldes teiste emotsioonide pausidega pikemad (Murray, Arnott 2008, Yildirim jt 2004, ten Bosch 2003).

Mõõtsin programmiga Praat (Boersma, Weenink 2009) pauside kestused millisekundites ning jagasin pausid nende iseloomu ja asukoha järgi gruppidesse (vt tabel 6).

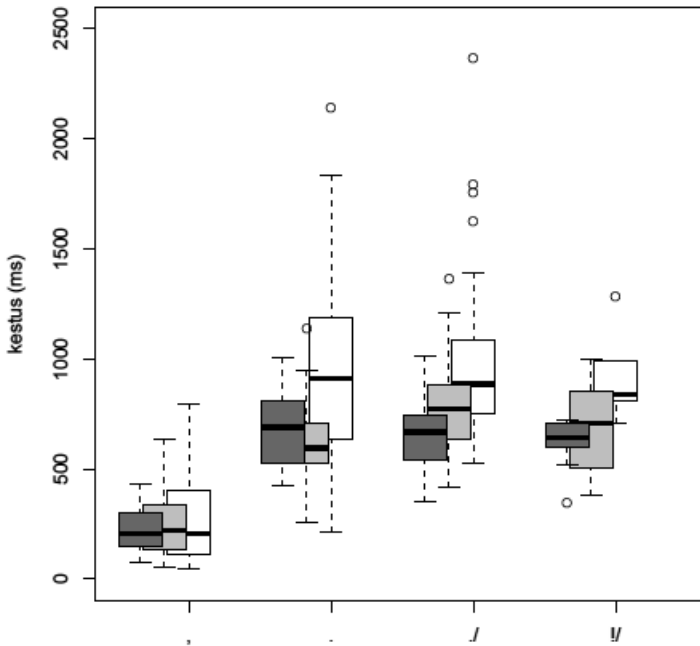
Koondtabelist on näha, et kestuste keskmiste järgi olid pikimad pausid kõigil kolmel emotsioonil punktipausid (kurbusel ja rõõmul hingamisega punktipausid, vihal hingamiseta punktipausid) ning kõige lühemad hingamiseta lisa- ja komapausid.⁴

Kui võrrelda emotsioonide pausigruppide kestusi omavahel, kooruvad välja mõned olulised erinevused (vt joonis 1). Joonisel 1 on võrdluseks välja toodud ka üks emotsiooniti mitte erinev pausirühm – komad.

⁴ Arvestatud pole nende pausirühmade andmeid, kus oli ainult üks esindaja.

Tabel 6. Pausirühmade keskmised kestused (millisekundites)

Märk	Emotsioon	N	Min	1. kvartil	Mediaan	Keskmine	3. kvartil	Maks
Lisapaus	kurbus	9	54	91	117	122,4	133	251
	rõõm	11	34	103,5	111	128,5	127	359
	viha	9	91	119	133	140,9	145	267
Hingamisega lisapaus	kurbus	4	300	366	436,5	417,5	488	497
	rõõm	3	365	416	467	448,3	490	513
	viha	11	289	326	399	385,6	412,5	555
Komapaus	kurbus	28	43	113,2	203	280,2	378,8	795
	rõõm	34	52	137,2	222	251,7	339	636
	viha	21	75	145	207	225	299	428
Hingamisega komapaus	kurbus	9	296	394	531	511	615	721
	rõõm	19	280	407	440	479,7	538,5	724
	viha	33	205	337	408	411,7	449	732
Koolonipaus	kurbus	0	–	–	–	–	–	–
	rõõm	2	139	180,8	222,5	222,5	264,2	306
	viha	1	364	364	364	364	364	364
Hingamisega koolonipaus	kurbus	1	438	438	438	438	438	438
	rõõm	1	777	777	777	777	777	777
	viha	2	296	366,5	437	437	507,5	578
Mõttekriip-supaus	kurbus	2	837	838,2	839,5	839,5	840,8	842
	rõõm	4	393	526,5	606,5	628	708	906
	viha	5	236	292	525	440,8	545	606
Hingamisega mõttekriip-supaus	kurbus	5	533	987	989	936	1033	1138
	rõõm	3	373	537,5	702	603,7	719	736
	viha	8	316	375	448	526,4	693,5	865
Mõtte-punktid	kurbus	5	407	412	786	742	972	1133
	rõõm	1	724	724	724	724	724	724
	viha	2	878	878,5	879	879	879,5	880
Hingamisega mõtte-punktid	kurbus	8	578	883,5	934	911,6	987	1085
	rõõm	1	474	474	474	474	474	474
	viha	0	–	–	–	–	–	–
Punktipaus	kurbus	30	211	637,8	913	944,4	1156	2136
	rõõm	25	257	524	594	637,8	710	1137
	viha	34	426	534	689	690,9	805,2	1006
Hingamisega punktipaus	kurbus	54	528	756,8	884,5	968,9	1076	2366
	rõõm	53	413	637	772	773,7	882	1364
	viha	43	353	542	667	664	746	1016
Hüüumärgi-paus	kurbus	0	–	–	–	–	–	–
	rõõm	7	256	445,5	540	570,9	708	893
	viha	6	384	626,5	635,5	623,7	655	802
Hingamisega hüüumärgi-paus	kurbus	6	709	813	835,5	910,8	954,8	1285
	rõõm	16	383	525,8	705	696,2	843	996
	viha	13	341	600	644	625,2	706	719
Küsimärgi-paus	kurbus	1	941	941	941	941	941	941
	rõõm	1	687	687	687	687	687	687
	viha	1	532	532	532	532	532	532
Hingamisega küsimärgi-paus	kurbus	2	479	601,8	724,5	724,5	847,2	970
	rõõm	0	–	–	–	–	–	–
	viha	1	623	623	623	623	623	623
Küsi-hüüu-märgipaus	kurbus	0	–	–	–	–	–	–
	rõõm	0	–	–	–	–	–	–
	viha	1	652	652	652	652	652	652
Hingamisega küsi-hüüu-märgipaus	kurbus	1	1260	1260	1260	1260	1260	1260
	rõõm	0	–	–	–	–	–	–
	viha	4	608	619,2	638	652,8	671,5	727



Joonis 1. Statistiliselt oluliselt erinevad pausirühmad (valge tähistab kurbust, helehall rõõmu ja tumehall viha; kaldkriips märgib hingamispause). Joonisel on esitatud iga pausirühma mediaan, esimene ja kolmas kvartiil ning miinimum ja maksimum

Emotsiooniti statistiliselt oluliselt erinevaks ($p < 0,05$)⁵ osutusid hingamiseta punktipaused, hingamisega punktipaused ja hingamisega hüüumärgipaused (arvutatud kestuse keskmiste põhjal).

- Hingamiseta punktipaused
 1. kurbus (944 ms) vs. rõõm (638 ms) ($p = 0,01$)
 2. kurbus (944 ms) vs. viha (691 ms) ($p = 0,006$)
- Hingamisega punktipaused
 1. kurbus (969 ms) vs. rõõm (774 ms) ($p = 0,000$)
 2. kurbus (969 ms) vs. viha (664 ms) ($p = 0,000$)
 3. rõõm (774 ms) vs. viha (664 ms) ($p = 0,003$)
- Hingamisega hüüumärgipaused
 1. kurbus (911 ms) vs. viha (625 ms) ($p = 0,017$)

Pauside modelleerimisel tasub neid erinevusi arvesse võtta.

Ülejäänud pausigrupid kestuselt emotsiooniti oluliselt ei erinenud.

Mõõtmistulemustest selgus, et kõige pikemad paused olid kurbusel, lühimad aga vihal (kestuste keskmine). See läheb kokku teiste keelte emotsionaalse kõne uurimuste tulemustega.

⁵ Kasutatud on Student's t-Testi. Püstitati 0-hüpootees: kahe emotsiooni keskmiste vahe on null ($p > 0,05$).

4. Emotsiooni kuulamistest

Pausiurimuse tulemustele tuginedes koostasin kuulamistesti, et teada saada, kas teksti emotsioon on ära tuntav ainult pausierinevuste põhjal.

Kuulamistesti aluseks oli (neutraalne) neljalausealine lõik, mille lausetest olid katseisikud (kuulamise põhjal) hinnanud neutraalseks kolm (lõigu esimese, kolmanda ja neljanda lause), ühe (lõigu teise lause) aga rõõmuemotsiooni kandvaks. Puhast, ainult neutraalsetest lausetest koosnevat lõiku korpusest leida ei õnnestunud. Programmi Sound Forge Pro 9 abil muutsin pauside kestust, asukohta ning iseloomu. Nii moodustusid kaheksa erineva pausimustriga lõiku (2 x neutraalne, viha, kurbus ja rõõm). Neli neist olid lihtsama pausimustriga, s.t neutraalses lõigus on muudetud ainult lauselõpupauside kestusi vastavalt pausiurimuse tulemustele (vt näide 1). Lauselõpupausid valisin välja seepärast, et nende erinevus oli emotsiooniti statistiliselt oluline. Teised neli olid keerulisema struktuuriga (vt näide 2). Neis muutsin nii pauside kestust, iseloomu kui ka asukohta vastavalt pausiurimuse tulemustele. Neutraalsetes lõikudes ei muutnud ma midagi.

- (1) Neutraalne: Klaasipõhja valatakse seda vahest kolm sentimeetrit [./] (456 ms) Ainult tirtsukese rohkem, kui enesest lugupidav sommeljee mekkimiseks pakub [./] (640 ms) Seitsekümmend krooni, palun [.] (495 ms) Aitäh.

Viha: Klaasipõhja valatakse seda vahest kolm sentimeetrit [./] 664 ms Ainult tirtsukese rohkem, kui enesest lugupidav sommeljee mekkimiseks pakub [./] 664 ms Seitsekümmend krooni, palun [.] 691 ms Aitäh.

Rõõm: Klaasipõhja valatakse seda vahest kolm sentimeetrit [./] 774 ms Ainult tirtsukese rohkem, kui enesest lugupidav sommeljee mekkimiseks pakub [./] 774 ms Seitsekümmend krooni, palun [.] 638 ms Aitäh.

Kurbus: Klaasipõhja valatakse seda vahest kolm sentimeetrit [./] 969 ms Ainult tirtsukese rohkem, kui enesest lugupidav sommeljee mekkimiseks pakub [./] 969 ms Seitsekümmend krooni, palun [.] 944 ms Aitäh.

- (2) Neutraalne: Klaasipõhja valatakse seda vahest kolm sentimeetrit [./] (456 ms) Ainult tirtsukese rohkem, kui enesest lugupidav sommeljee mekkimiseks pakub [./] (640 ms) Seitsekümmend krooni, palun [.] (495 ms) Aitäh.

Viha: Klaasipõhja valatakse seda vahest [] 141 ms kolm sentimeetrit [.] 691 ms Ainult tirtsukese rohkem [./] 412 ms kui enesest lugupidav sommeljee mekkimiseks pakub [./] 664 ms Seitsekümmend krooni, palun [.] 691 ms Aitäh.

Rõõm: Klaasipõhja valatakse seda vahest kolm sentimeetrit [./] 774 ms Ainult tirtsukese rohkem [.] 252 ms kui enesest lugupidav sommeljee mekkimiseks pakub [.] 638 ms Seitsekümmend krooni, palun [./] 774 ms Aitäh.

Kurbus: Klaasipõhja valatakse seda vahest kolm sentimeetrit [.] 944 ms Ainult tirtsukese rohkem, kui enesest lugupidav sommeljee mekkimiseks pakub [./] 969 ms Seitsekümmend krooni, palun [./] 969 ms Aitäh.

Pärast testi koostamist palusin kümnel katseisikul heli põhjal määrata, mis emotsioon lõigus kõlab – kas viha, rõõm, kurbus või on lõik ilma erilise emotsioonita, s.t neutraalne. Testi tulemustest pidi selguma, kas lause emotsiooni saab muuta ainult ühe akustilise parameetri abil, s.t pausiga manipuleerides.

Lõigud, kus oli muudetud ainult lauselõpupauside kestusi (vt näide 1), kõlasid kuulajale endiselt neutraalsena (vt tabel 7). Sama tulemuse sain lausete puhul, kus oli muudetud nii pauside kestust, iseloomu kui ka asukohta (vt näide 2 ja vt tabel 8).

Seega ei suutnud pausid üksinda lõigu emotsiooni muuta.

Tabel 7. Lõigud, kus oli muudetud ainult lauselõpupauside kestus

Vastusevariandid				
Lõigu emotsioon	Rõõm	Viha	Kurbus	Neutraalne
Rõõm	0%	10%	0%	90%
Viha	10%	10%	10%	70%
Kurbus	0%	20%	10%	70%
Neutraalne	0%	20%	10%	70%

Tabel 8. Lõigud, kus oli muudetud nii pauside kestus, iseloom kui ka asukoht

Vastusevariandid				
Lõigu emotsioon	Rõõm	Viha	Kurbus	Neutraalne
Rõõm	10%	10%	10%	70%
Viha	10%	10%	0%	80%
Kurbus	10%	20%	0%	70%
Neutraalne	10%	10%	0%	80%

5. Kokkuvõte

Pausiuurimus näitas, et eestikeelses emotsionaalses kõnes võivad pausid emotsiooniti erineda nii oma asukoha, iseloomu kui ka kestuse poolest. Kuigi analüüsist selgus, et pauside vahel oli siin-seal emotsiooniti olulisi erinevusi, näitas kuulamistest, et sellest üksinda siiski ei piisa, et kuulajad suudaksid nende põhjal emotsioone üksteisest eristada. Samas ei saa ka väita, et pausidel pole emotsioonide eristamisel üldse mingit tähtsust. Pausid võivad muutuda oluliseks koos teiste emotsioone eristavate akustiliste tunnustega. Kas see nii ka on, seda näitab edasine emotsiooniakustika uurimine.

Pauside arvu, asukoha ja iseloomu analüüsi tulemusi kokku võttes saab esitada mõned soovitusel pauside modelleerimiseks sünteeskõne jaoks.

Pauside asukoht – kõigil kolmel emotsioonil tasub pausid panna kirjavahe- märkide kohale. Erand tuleb sisse ainult koma juures – rõõmu- ja vihalausetes võib pausi panna ka koma kohale, kurbuselausetes pigem mitte. Kuna lisapauside osakaal oli emotsiooniti väike (8–10%) ning sõltus pooltel juhtudel kas teksti sisust või lugejast endast, siis lisapause modelleerida ei ole oluline.

Pauside iseloom – kõigil kolmel emotsioonil võib vähemalt pooled pausid modelleerida hingamispausideks. Pausid tuleb teha mõistliku vahemaa tagant,

intervalli võib arvutada näiteks deklinatsiooni oletatava kestuse järgi. Kui viha puhul ei ole oluline, millise kirjavahemärgi kohale hingamine panna, siis rääkimises sobib hingamine pigem lauselõppu ning kurbuselauluses tuleb jälgida, et hingamine (ja paus) ei satuks koma kohale.

Viidatud kirjandus

- Altrov, Rene 2008. Eesti emotsionaalse kõne korpus: teoreetilised toetuspunktid. – Keel ja Kirjandus, 4, 261–271.
- Bachorowski, Jo-Anne 1999. Vocal expression and perception of emotion. – *Current Directions in Psychological Science*, 8 (2), 53–57. doi:10.1111/1467-8721.00013
- Boersma, Paul; Weenink, David 2009. Praat: doing phonetics by computer (Version 5.0.46). <http://www.praat.org/> (19.01.2009).
- Burkhardt, Felix; Audibert, Nicolas, Malatesta, Lori; Türk, Oytun; Arslan, Levent M.; Auberge, Veronique 2006. Emotional prosody – does culture make a difference? – SP-2006, paper 207.
- Campbell, Nick 1998. Where is the information in speech? (And to what extent can it be modelled in synthesis?). – *Proceedings of the Third ESCA/COCOSDA Workshop on Speech Synthesis*. Australia: Jenolan Caves, 17–20.
- Douglas-Cowie, Ellen; Campbell, Nick; Cowie, Roddy; Roach, Peter 2003. Emotional speech: Towards a new generation of databases. – *Speech Communication*, 40 (1), 33–60. doi:10.1016/S0167-6393(02)00070-5
- Eesti emotsionaalse kõne korpus. <http://193.40.113.40:5000> (29.09.2009).
- Iida, Akemi; Campbell, Nick; Higuchi, Fumito; Yasumura, Michiaki 2003. A corpus-based speech synthesis system with emotion. – *Speech Communication*, 40 (1), 161–187. doi:10.1016/S0167-6393(02)00081-X
- Montero, Juan Manuel; Gutierrez-Arriola, Juana M.; Palazuelos, Sira; Enriquez, Emilia; Aguilera, Santiago; Pardo, José Manuel 1998. Emotional speech synthesis: from speech database to TTS. – *ICSLP-1998*, paper 1037.
- Murray, Iain R.; Arnott, John L. 2008. Applying an analysis of acted vocal emotions to improve the simulation of synthetic speech. – *Computer Speech and Language*, 22, 107–129. doi:10.1016/j.csl.2007.06.001
- Pajupuu, Hille; Kerge, Krista 2006. Hingav süntesaator ja pausid tekstiliigiti. – Keel ja Kirjandus, 3, 202–210.
- Scherer, Klaus R. 2003. Vocal communication of emotion: A review of research paradigms. – *Speech Communication*, 40 (1), 227–256. doi:10.1016/S0167-6393(02)00084-5
- Schröder, Marc; Cowie, Roddy; Douglas-Cowie, Ellen; Westerdijk, Machiel; Gielen, Stan 2001. Acoustic correlates of emotion dimensions in view of speech synthesis. – *EUROSPEECH-2001*, 87–90.
- Tamuri, Kairi 2007. Pausid ettelõetud ilukirjandustekstis. Magistritöö. Käsikiri Tallinna Ülikooli eesti keele ja kultuuri instituudis.
- ten Bosch, Louis 2003. Emotion, speech and the ASR framework. – *Speech Communication*, 40 (1), 213–225. doi:10.1016/S0167-6393(02)00083-3
- Yildirim, Serdar; Bulut, Murtaza; Lee, Chul Min; Kazemzadeh, Abe; Deng, Zhigang; Lee, Sungbok; Narayanan, Shrikanth; Busso, Carlos 2004. An acoustic study of emotions expressed in speech. – *INTERSPEECH-2004*, 2193–2196.

Kairi Tamuri (Eesti Keele Instituut) uurimisvaldkond on emotsionaalse kõne akustika.
kairi.tamuri@eki.ee

DO THE PAUSES IN READ TEXT CARRY EMOTION?

Kairi Tamuri

Institute of the Estonian Language

Pauses have received relatively little attention among the researchers of emotional speech. It may be caused by the fact that pauses are not considered to belong to the primary acoustic parameters rendering emotions. However, they do help to distinguish between emotions (Murray, Arnott 2008, Yildirim et al. 2004, ten Bosch 2003, Iida et al. 2003, Schröder et al. 2001, Montero et al. 1998).

This research looks for an answer to two questions: 1) if the pauses in an emotional text read out in Estonian differ considerably depending on the emotion of the text; 2) if the emotion contained in the text can be recognised by looking at the pause differences only.

Although the research proved that some pauses differ by emotion, the perception test shows that this parameter alone is not sufficient to distinguish between emotions.

Keywords: acoustics of emotions, read out text, pause groups, number of pauses, location of pauses, character of pauses, duration of pauses, speech synthesis, Estonian

STATISTILISED MEETODID MURDEKORPUSE ÜHENDVERBIDE TUVASTAMISEL

Kristel Uiboaed

Ülevaade. Sõnadevahelise seose tugevuse mõõtmise statistikuid kasutatakse arvutilingvistikas püsiühendite tuvastamisel. Statistikud võimaldavad korpusel kahele sõnale arvutada nendevahelise seose tugevuse väärtuse, mille põhjal võib otsustada, kas tegemist on püsiühendiga või mitte. Statistikute kasutamise eelis on, et arvesse ei võeta ainult sõnade koosinemise, vaid ka ühendit moodustavate sõnade eraldiesinemise sagedusi. Artiklis teen katse rakendada statistikuid Eesti murrete korpusel kaheliikmeliste ühendverbide automaatsel tuvastamisel. Katsetatud on kolme murderühma peal eraldi nelja statistikut: t-skoori, vastastikuse informatsiooni väärtust MI, hii-ruut statistikut ning log-tõepära funktsiooni.*

Võtmesõnad: arvutilingvistika, korpuslingvistika, murdeuurimine, meetodid ja vahendid, statistika, eesti keel

1. Sissejuhatus

Keelekorpuste levik on teinud võimalikuks erinevate keelenähtuste uurimise suuremahulise materjali põhjal. See eeldab omakorda erinevate statistiliste meetodite kasutamist lingvistiliseks andmetöötuseks ning vajaduse korral uute rakenduste väljatöötamist. Oleks mõeldamatu, et inimene suudaks nii suurt andmestikku läbi vaadata ning seda analüüsida. Käesolev töö on esmane suurem katse rakendada arvutilingvistikas laialt levinud statistilisi meetodeid ka Eesti murrete korpusel¹ peal. Ühendverbide automaatseks tuvastamiseks Eesti murrete korpusel kasutatakse sõnadevahelise seose tugevuse mõõtmise statistikuid (ingl *association measures*) – statistilisi valemeid, mille abil on võimalik arvutada igale korpusel olevale sõnaühendile seose tugevuse väärtus (*association score*), mis osutab sõnadevahelise statistilise seose tugevusele. (Evert 2008)

* Artikli valmimist on toetanud riikliku programmi "Eesti keel ja kultuurimälu 2009" projekt EKKM09-111 ja riikliku programmi "Eesti keele keeletehnoloogiline tugi 2006" projekt EKKTT06-14 ning Eesti Teadusfondi grant nr 7464.

Täna anonüümseid retsensente asjatundlike ettepanekute ja kommentaaride eest.

¹ Eesti murrete korpus www.murre.ut.ee (30.08.2009).

Erinevad statistikud põhinevad erinevatel teooriatel, seega erinevate statistikute abil arvatud sõnadevahelise seose tugevuse väärtused pole omavahel võrreldavad, kuna väärtuste leidmise alused pole ühtsed. Statistikud võivad põhineda lihtsalt sageduste kokkulugemisel, informatsiooniteoorial² või neil võib olla statistiline alus (Evert 2008, Krenn, Evert 2001).

Sõnadevahelise seose tugevuse mõõtmise statistikuid kasutatakse arvutileksikograafias, eelkõige püsiühendite tuvastamisel, ning paljudes arvutilingvistilistes rakendustes, näiteks masintõlkes. Statistikuid on võimalik rakendada muudegi grammatiliste konstruktsioonide ja lingvistiliste nähtuste uurimisel. Näiteks on statistikute abil võimalik uurida verbi ja objekti suhteid, verbide transitiivseid ja intransitiivseid omadusi (Stefanowitsch, Gries 2009).

Artikli eesmärk on tutvustada statistiliste meetodite tööpõhimõtteid ning rakendada neid murdekorpuse morfoloogiliselt märgendatud ja automaatselt osalausestatud tekstide peal. Esiteks tutvustan statistikute kasutamiseks vajalikke mõisteid ja andmeid, kirjeldan kolme võimalikku materjali liigendamiseviisi ning näitan, kuidas arvulised väärtused valemitesse saadakse. Seejärel teen ülevaate kasutatud materjalist. Viimases osas kirjeldan nelja väljavalitud sõnadevahelise seose tugevuse mõõtmiseks kasutatavat statistikut, mida olen testinud murdekorpuse ühendverbide peal sihiga välja selgitada parimad statistikud murdekorpuse ühendverbide tuvastamiseks ning murdekorpuse materjaliga töötamiseks laiemalt edaspidi. Statistika terminoloogias lähtun Rahvusvahelise Statistika Instituudi kodulehel olevast terminipangast (ISI).

2. Statistiliste meetodite tööpõhimõtetest

2.1. Mis andmeid on vaja sõnadevahelise seose tugevuse mõõtmise statistikute kasutamiseks?

Sõnadevahelise seose tugevuse väärtused on kvantitatiivselt mõõdetavad ja nad iseloomustavad sõnadevahelise seose tugevust. Neid leitakse erinevate statistiliste valemite abil. Väärtuse tõlgendamise põhimõte on lihtne: kõrged väärtused osutavad sõnadevahelisele tugevale seosele, väiksed ja negatiivsed väärtused sellele, et sõnad pigem välistavad koosesinemisel üksteist.

Selleks, et püsiühendeid sõnadevahelise seose tugevuse väärtuse järgi leida, peaks kõigepealt piiritlema, mida sõnade koosesinemise all mõeldakse. Kas soovitakse leida ainult n-õ tõelisi püsiühendeid või lihtsalt seada sõnadevahelise seose tugevuse mõõdud mingile skaalale, tõmbamata seejuures kindlat piiri püsiühendite ja mitte-püsiühendite vahele. (Evert 2009)

Teiseks peab kindlaks tegema, kuidas püsiühendeid grupeerida: kas ollakse huvitatud ainult väga tugeva seosega sõnaühenditest, mida peetakse täiesti iseseisvateks leksikaalseteks või grammatilisteks üksusteks, või soovitakse ainult kindlaks teha, milline on vaadeldava sõna levinum kontekst. N-õ tõeliste püsiühendite kindlakstegemisel määrab väärtuste piiri uurija ise, otsustades, kui suur peab olema sõnadevahelise seose tugevuse väärtus, et neid veel koosinevaks pidada. (Evert 2009)

Kõige lihtsam viis korpusest püsiühendeid leida on lihtsalt nad kokku lugeda – leida sõnade koosinemise arv. Sellise lihtsa meetodiga võib saavutada juba üsna head tulemused ja eriti hästi töötab see läbipaistvate ning kindla konstruktsiooniga püsiühendite peal (*kõnet pidama, tantsu lööma*). Sageli võivad püsiühendi eri osad paikneda lauses üsna vabalt ning sellistel juhtudel jääb sageduste kokkulugemisest väheks, siin on vajalik seose tugevuse üle otsustamiseks keerulisemaid meetodeid. (Manning, Schütze 2003: 153–158)

Sageli jääb sõnadevahelise seose üle otsustamisel lihtsalt nende koosinemiskordade kokkulugemisest väheks. Kui moodustada korpuse bigrammide ehk sõnapaaride sagedusloend, siis ilmselt oleks seal üsna kõrgel kohal ühend *ja ka*. See aga ei tähenda veel, et tegemist on püsiühendiga. Mõlemad sõnad esinevad korpuses sageli ning seega on loomulik, et ka nende koosinemist tuleb tihti ette. Lisaks lihtsale koosinemise sagedusele tuleb arvesse võtta ühendit moodustavate sõnade eraldiesinemise sagedusi – marginaal- ehk ääresagedusi.³ Samuti peab arvestama korpuse suurust ehk valimimahtu, kust püsiühendeid leitakse. Seega saame statistiliste meetodite kasutamiseks sageduse arvutamise andmestiku, mis moodustub järgmistest näitajatest: O – sõnade koosinemise sagedus korpuses ehk valimis, f_1 ja f_2 – vastavalt esimese ja teise sõna eraldiesinemise sagedus, N – valimimaht (valimimahust tuleb täpsemalt juttu edaspidi). (Evert 2009)

Lisaks on statistiliste meetodite kasutamiseks vajalik leida koosinemise teoreetiline sagedus E (ingl *expected frequency*), mis näitab, kui suur on sõnade koosinemise oodatav tõenäosus, eeldusel et sõnad esinevad tekstis üksteisest sõltumatult. Üks võimalus on arvutada teoreetiline sagedus järgmise valemi põhjal: $E = f_1 * f_2 / N$ ehk sõnade eraldiesinemise sageduste korrutis tuleb jagada valimi suurusega. (Evert 2009)

Keerulisemate statistiliste meetodite jaoks on eelnevalt vaja koostada kahe-mõõtmeline sagedustabel (*contingency tabel*), mis arvutatakse sõnade eraldi- ja koosinemissageduste ning valimimahu põhjal. Tabel 1 esitab kahemõõtmelise sagedustabeli ning tehted, kuidas leitakse väärtused tabeli eri lahtrite jaoks.

Tabel 1. Kahemõõtmeline sagedustabel (Evert 2009)

O_{11}	O	$f_1 - O$	O_{12}
O_{21}	$f_2 - O$	$N - f_1 - f_2 + O$	O_{22}

Oletame, et meil on korpus 48 705 osalausega ($N = 48\,705$). Korpuses esinevad sõnad *vastu* ja *võtma* samas osalauses 78 korda ($O = 78$). *Vastu* ja *võtma* eraldiesinemise sagedused on vastavalt 458 ja 670 ehk $f_1 = 458$ ja $f_2 = 670$. Sellistel tingimustel saab koostada tabelis 2 esitatud kahemõõtmelise sagedustabeli.

Tabel 2. Kahemõõtmeline sagedustabel (*vastu võtma*)

O_{11}	78	$458 - 78 = 380$	O_{12}
O_{21}	$670 - 78 = 592$	$48\,705 - 458 - 670 + 78 = 47\,655$	O_{22}

³ Statistiline termin sõnade eraldiesinemissageduste kohta selles kontekstis on marginaal- ehk ääresagedused. Selles artiklis kasutan siiski terminit *sõnade eraldiesinemissagedus*, mis on intuiitiivselt arusaadavam.

2.2. Kolm lähenemisviisi materjali liigendamiseks

Selles osas tutvustan kolme lähenemisviisi, millest võib lähtuda püsiühendite kandidaatpaaride moodustamisel. Näidetena olen siin kasutanud kirjakeelset materjali,⁴ sest kirjakeeles on laused kirjavahemärgistatud ning laused ja osalaused selgesti eraldatavad, mistõttu on kirjeldatavaid põhimõtteid lihtsam mõista.

1. **Kindlas naabruses koosinemine** (ingl *surface cooccurrence*) on kõige levinum lähenemine sõnadevahelise seose tugevuse mõõtmiseks. Selle järgi peetakse kahte sõna koosinevaks, kui nad esinevad sageli koos teatavas kauguses ehk samas aknas (*collocational span*), mida arvestatakse kahe sõna vahele jäävate sõnade järgi. Akna suuruse määrab uurija, tavaliselt jääb see kolme ja viie sõna vahele. Akna suurust võib märkida kujul L_5 ja R_5 , mis tähendab, et uuritava sõna vasakule ning paremale poole arvestatakse viis sõna ehk aken on sümmeetriline. Aken võib olla ka asümmeetriline, näiteks L_2 ja R_4 , kui uuritava sõna vasakule poole jääb kaks ning paremale neli sõna. Samuti peab otsustama, kas arvestatakse ainult sõnu või ka muid märke (kirjavahemärke, numbreid). Tuleb määratleda, kuidas käsitletakse mitmesõnalisi ühendeid ning kas koosinevad sõnad võivad ületada lausepiire. (Evert 2009)

Näide (1) illustreerib kindlas naabruses koosinemist. Vaadeldavateks sõnadeks on *juttu* ja *ajama*. Uuritav sõnavorm on *juttu*. Akna suurus on L_3 ja R_3 ehk konteksti arvestatakse mõlemalt poolt kolm sõna. Osalause piire pole arvesse võetud, kirjavahemärke akna suuruse määramisel ei arvestata.

- (1) Kiusu pärast, oma täieliku ükskõiksuse pärast tegin sellest siiski **juttu**, puudutasin keelatud asju. Viieaastaselt oli poiss leitanud viisi, kuidas linnas **juttu ajama** jäänud õdesid kähku liikuma saada. Ta võis vabalt ka kogu selle aja vaikselt uksepeal nende **juttu** pealt kuulata ja seejärel tuppa astudes teatrit tegema hakata. Aga seal, emaga kõõgis **juttu ajades**, olevat tulnud talle just Alfredi tublidus meelde. Need mehed aga, kes niisugust hinnapoliitikat **ajavad**, oma nägu kail ei näita, nemad istuvad kaldal uhketes büroomajades ning nende juurde iga soovijat ei lastagi.

Juttu ajama esineb samas aknas kaks korda ehk $O = 2$. Sõnade *juttu* ja *ajama* eraldi-esinemise sagedused on vastavalt neli ja kolm ehk $f_1 = 4$ ja $f_2 = 3$. Valimimaht ehk N antud näite puhul $N = 84$ (kirjavahemärke pole arvestatud). Sõnade arvu järgi arvestatakse valimimahtu ainult kindlas naabruses koosinemise lähenemisest lähtudes.

2. **Tekstilise koosinemise** (*textual cooccurrence*) puhul peetakse sõnu koosinevaks, kui nad esinevad koos samas tekstiüksuses, tavaliselt lauses, osalauses või lausungis. Kindlas naabruses koosinemise lähenemisele heidetakse ette akna suuruse suvalisust, mis võib tulemust vales suunas mõjutada just vaba sõnajärjega keeltes, kus omavahel tugevas seoses olevad sõnad võivad üksteisest paikneda kaugel. Seega sobib sellistes keeltes paremini aknaks kogu lause või osalause ehk püsiühendite statistilisel tuvastamisel tuleb aluseks võtta tekstilise koosinemise lähenemine. See lähenemine suudab tuvastada ka nõrgemas seoses olevaid püsiühendeid. (Evert 2009)

Tabel 3 illustreerib tekstilist koosinemist. Siin võetakse arvesse osalause piire ning vaadeldakse sõnade *peale* ja *käima* koosinemist samas osalauses.

⁴ Eesti kirjakeele korpus: <http://www.cl.ut.ee/korpused/kasutajaliides/> (30.08.2009).

Tabel 3. Tekstiline koosinemine

Osalause	Info koosinemise kohta
Ta oli ise riskialdis ja hasartne,	— —
temal käis kõik ikka kogu panga peale .	peale käima
Selle vastu polnud mul kõige väiksematki huvi,	— —
aga et ma olin Aimest kuulda saades kindlalt otsustanud Liisuga jälle söber olla,	— —
siis käisin talle tormiliselt peale :	peale käima
“Tule ikki!	— —
Joost oleks varanduse ära andnud,	— —
aga õpilane ei tahtnud vist,	— —
või kui tahtiski,	— —
siis ei julgenud peale käia ,	peale käima
oli võtsilt viisakas, korralik.	— —
Juhangi läks selle peale ägedaks.	peale
Eile õhtul käisin auhinda vastu võtmas.	käima

Ühend *peale käima* esineb koos samas osalause kolm korda, seega $O = 3$. Sõnade eraldiesinemissagedused arvestatakse selle järgi, mitmes osalause kumbki sõna esineb, antud näites $f_1 = 4$ ja $f_2 = 4$. Osalauseid on 13 ehk $N = 13$.

3. **Süntaktilise koosinemise** (*syntactic cooccurrence*) lähenemine on kõige suuremate piirangutega, mille kohaselt sõnu peetakse koosinevateks ainult juhul, kui nende vahel on kindel süntaktiline seos, näiteks verb koos subjekti või objektiga. Süntaktilise seose seisukohalt erinevaid püsiühendeid vaadeldakse eraldi. Süntaktilise koosinemise lähenemine on sobilik olukordades, kus koosinevad sõnad asuvad teineteisest kaugel. See lähenemine võetakse sageli aluseks püsiühendite tuvastamisel, kuna paljud leksikaliseerunud ühendid esinevad koos tavaliselt kindlas süntaktilises seoses (näiteks adverbi ja verbi ühend võib koos moodustada öeldise). (Evert 2009)

Süntaktilise koosinemise lähenemist illustreerib tabel 4. Oletame, et soovime teada saada, kas *tantsu lööma* on püsiühend.

Tabel 4. Süntaktiline koosinemine

Lause	Info koosinemise kohta
Igatsedes vaatas ^[pred] Grethel noori ^[obj] , kes jõululaule ^[obj] lausid ^[pred] ja lõbusalt tantsu ^[obj] lõid ^[pred] .	vaatas noori lausid jõululaule lõid tantsu
Koerad lõid ^[pred] tantsu ^[obj] .	lõid tantsu
Ta sulges ^[pred] silmad ^[obj] ja alustas ^[pred] tantsu ^[obj] .	sulges silmad alustas tantsu
Ta lõi ^[pred] mind ^[obj] selgeks.	lõi mind
Kas te lõite ^[pred] mätaste vahel siis tantsu ^[obj] ?	lõite tantsu
Ta avas ^[pred] ukse ^[obj] .	avas ukse
Tantsisime ^[pred] kolm tantsu ^[obj] .	tantsisime tantsu
Nad piinasid ^[pred] tantsu ajal jalgu ^[obj] .	piinasid jalgu

Siin leitakse vajalikud andmed natuke erineval viisil. Oletame, et näites kasutatud väike korpus on süntaktiliselt märgendatud, s.t iga lauseliige on tähistatud vastava märgendiga. Antud juhul uurime ainult predikaadi ja objekti kombinatsioone, seega tuleb esmalt tuvastada kõik sellises seoses olevad sõnaühendid, mis koos moodustavad valimi. Tabelis 4 on 11 predikaadi ja objekti ühendit, seega $N = 11$. Eraldame uuritava ühendi *tantsu lööma*, näites on neid kolm, seega $O = 3$. Sõnade eraldiesinemissagedused on vastavalt $f_1 = 5$ ning $f_2 = 4$. Tegelikult esineb sõna *tantsu* tekstis 6 korda. Viimases lauses aga pole ta objekti rollis ning kuna antud juhul on huvi keskmises vaid objekti ja predikaadi ühendid, siis muus funktsioonis esinevaid vaatlusaluseid sõnavorme sageduste kokkulugemisel ei arvestata.

Kindlas naabruses koosinemise lähenemine on osutunud kasulikuks leksikograafias ja korpuslingvistikas. Tekstilise koosinemise lähenemist on kergem rakendada ning see on tõhusam kõikvõimalike juhuslike väljendite välistamisel. Süntaktilise koosinemise lähenemine on tõhusam kindla konstruktsiooniga püsiühendite leidmisel, aga seab eelduseks korpuse süntaktilise märgendatuse.

3. Materjal

Käesoleva töö materjaliks on Eesti murrete korpuse morfoloogiliselt märgendatud tekstid (sisuga 01.06.2009) ning nendest leitud ühendverbid. Tekstid on eelnevalt lausestatud, kasutades Kaili Müürisepa loodud suulise kõne süntaksi analüsaatorit, mida on kohandatud murdetekstide lausestatmise jaoks (selle kohta vt Lindström, Müürisep 2009). Lausestatud tekstide põhjal moodustasin kaheliikmelised kombinatsioonid kõigi osalause sõnadega, neist otsisin automaatselt välja adverbi ning verbi ühendid. Vaatluse all on ainult kaheliikmelised ühendid. Välja on jäetud verbi *olema* sisaldavad ühendid, samuti ei ole ma kaasanud modaalverbe (*saama, võima, tohtima, pidama*) sisaldavaid ühendeid, kuna need käituvad sageli abiverbina. Kui sõnakombinatsioonid moodustada osalause sees kõigi verbide ning adverbidega, moodustuks ühendverbi kandidaat ka abiverbi funktsioonis esineva verbivormiga, nt osalause *töö sai ära tehtud* moodustuksid ühendid *ära tegema* kui *ära saama*, s.t toimuks teatav ülegenereerimine.⁵

Saadud verbi ja adverbi ühendite peal on lisaks rakendatud adverbide stopp-sõnade loendit, s.t eemaldatud on ühendid, mille adverbiline komponent reeglina ühendverbi ei moodusta, nt *hullusti, jõudsalt, korrapäraselt* jt (täpsemalt vt Uiboaed 2008). Samuti olen analüüsist välja jätnud ühendid, mis esinevad samas osalause vähem kui kolm korda.

Kõik katsed on tehtud kolmes osas, arvestades kolme erinevat murderühma. Põhjaeesti murderühma moodustavad idamurre, keskmurre, läänemurre ja saarte murre (kokku 279 126 sõna). Lõunaeesti murderühm hõlmab Mulgi, Tartu, Setu ja Võru murde (147 666 sõna). Madalaima esindatusega (75 144 sõna) rannikumurde murderühma moodustavad Alutaguse ning rannamurre.

Leidmata jäänud ühendverbide väljaselgitamiseks olen kasutanud testkorpust, mis sisaldab 10% kogu korpuse tekstidest. Rannikumurdes jääb testkorpuse põhjal leidmata 2% olemasolevatest ühendverbidest, lõuna- ja põhjaeesti rühmas on osakaalud mõlemad 16%. Korpuses olemasolevad ühendverbid jäävad leidmata peamiselt lausestaja vigade tõttu, sest ühendverbe leitakse ainult ühe osalause

⁵ Uiboaed (2010) kirjeldatud katse, mis põhineb samal materjalil, erineb siinsest selle poolest, et seal pole ühtegi statistikut kasutatud: vastavate ühendverbide sagedused on lihtsalt kokku loetud (mitu korda nad ühes osalause koos esinevad). Seega pole arvesse võetud ka sõnade eraldiesinemise sagedusi. Samuti olid seal osalausestatamise alused teised (automaatsel osalausestatamisel ei kasutatud süntaksianalüsaatorit).

piirides (piirid on seatud automaatselt). Kui lausestaja on osalause piiri seadnud valesti, nii et ühendverbi eri osad jäävad kahte erinevasse osalause, siis sellised ühendverbid jäävad tuvastamata.

4. Sõnadevahelise seose tugevuse mõõtmise statistikud

Selles peatükis näitan katsete abil, kuidas sõnadevahelise seose tugevuse mõõtmise statistikuid töötavad ning milliseid otsuseid nende põhjal teha võiks. Lisaks väga lihtsatele meetoditele (nt sageduste kokkulugemine) on olemas ka terve hulk informatsiooniteoorial ja statistilise olulisuse testimisel põhinevaid statistikuid. Mina olen tutvustamiseks välja valinud neli: kaks lihtsamat statistikut – 1) t-skoori (ingl *t-score*) ja 2) vastastikuse informatsiooni (MI) mõõdiku (*Mutual Information*) – ning kaks statistilist statistikut – 3) hii-ruut-statistiku (*chi-squared statistic*) ja 4) log-tõepära funktsiooni (*log-likelihood*).

Tavaliselt kasutatakse statistikuid kandidaatandmete seast tugevama seosega ühendite väljavalmimiseks. Kõigepealt leitakse mingi hulk ühendeid (mõne osas 2.2 kirjeldatud meetodi abil) ning nende jaoks arvutatakse statistilise seose tugevuse väärtused. Seejärel järjestatakse loend ümber ning koosinevateks loetakse ainult teatav hulk ühendeid sagedusloendi esimesest osast (väärtuse, millest alates ühendeid enam koosinevateks ei loeta, peab uurija määrama ise). (Krenn, Evert 2001)

Arvutuste tegemisel lähtusin tekstilise koosinemise põhimõttest – aknaks määrasin osalause, seega valimimaht on osalause arvu. Statistikute väärtused on arvutatud järgneva andmestiku abil: osalause arvu, ühendit moodustavate sõnade sagedused korpuses, sõnade koosinemise arv samas osalause ning nende koosinemise oodatav sagedus. Kõik katsed on tehtud kolmes osas, arvestades kolme murderühma, seega iga murderühma sagedusandmed on erinevad. Iga statistiku kirjelduse juures olevad pingeread on järjestatud vastava statistiku väärtuste alusel, s.t eespool on ühendid mida statistik loeb tugevasti kokkukuuluvateks.

Statistikute töö kontrollimiseks vaatasin läbi sada kõrgema väärtusega ühendverbi igas murdrühmas ja iga statistiku kaupa eraldi. Saja kõrgema väärtusega ühendi hulgas on korpuses ühendverbidena mitteesinevaid ühendeid ühe ja kolme ühendi vahel ning enamjaolt on need ühendid sattunud loendisse märgendusvigade tõttu. Seega võib statistikute tööd pidada tõhusaks.

4.1. t-skoor

t-skoor arvutatakse järgmise valemi põhjal:

$$t\text{-skoor} = \frac{O - E}{\sqrt{O}},$$

kus O on osalause hulk, kus sõnad koos esinevad, ja E on teoreetiline sagedus.

Tabel 5 esitab kahanevalt 50 kõrgeima t-skoori väärtusega ühendit.

Tabel 5. 50 kõrgeima t-skoori väärtusega ühendit eri murderühmades (k-sag – ühendi koosesinemissagedus samas osalauses, t-sk – t-skoori väärtus)

PÕHJA	k-sag	t-sk	LÕUNA	k-sag	t-sk	RANNIKU	k-sag	t-sk
peale_panema	97	8,59	ära_koolema	56	7,03	ära_võtma	30	4,42
välja_tulema	92	8,11	peale_panema	37	5,52	maha_laskma	19	4,19
sisse_panema	91	7,82	ära_surema	35	5,44	peale_hakkama	20	4,03
ära_võtma	103	7,50	sisse_panema	31	4,83	välja_vedama	20	4,03
ära_surema	63	7,39	ära_tapma	22	4,21	sisse_laskma	16	3,80
ära_kaduma	51	6,55	välja_tulema	22	3,75	tagasi_tulema	18	3,65
ära_viima	68	6,37	maha_jääma	16	3,70	sisse_panema	17	3,55
valmis_tegema	42	5,61	üles_ajama	15	3,65	ära_surema	16	3,53
üles_panema	50	5,55	tagasi_tulema	16	3,54	välja_tulema	24	3,30
kinni_siduma	31	5,47	üles_tulema	20	3,53	maha_lööma	12	3,29
ära_tooma	57	5,43	vastu_tulema	14	3,37	ringi_käima	13	3,29
alla_panema	39	5,42	kinni_hoidma	12	3,36	ära_lahutama	11	3,14
lahti_võtma	34	5,38	kinni_panema	18	3,36	ära_tulema	34	3,12
välja_võtma	43	5,34	ära_minema	65	3,34	läbi_käima	14	3,09
maha_laskma	29	4,99	ära_kuivatama	12	3,28	ära_viima	15	3,04
läbi_käima	37	4,95	kinni_võtma	15	3,27	kinni_panema	13	2,94
ära_tapma	29	4,88	ära_müüma	13	3,19	vastu_võtma	10	2,93
ette_panema	31	4,84	ühes_võtma	12	3,17	välja_minema	22	2,83
ära_kuivatama	26	4,63	kinni_kõitma	10	3,12	välja_võtma	13	2,81
kinni_panema	38	4,58	ette_panema	11	3,06	vastu_tulema	11	2,72
tagasi_tulema	25	4,39	üles_panema	16	3,01	peale_panema	11	2,67
vastu_võtma	19	4,11	ära_uppuma	10	2,96	lahti_tegema	11	2,65
sisse_minema	43	4,10	vastu_võtma	10	2,91	üles_võtma	10	2,65
kokku_vedama	18	4,08	välja_võtma	13	2,90	alla_vedama	8	2,59
ära_sööma	32	4,00	sisse_tulema	16	2,90	üles_tõstma	7	2,58
sisse_laskma	22	3,97	ära_pesema	12	2,90	ette_panema	8	2,51
pealt_võtma	17	3,84	manu_panema	10	2,80	valmis_tegema	8	2,48
üles_tõusma	15	3,80	ära_võtma	33	2,77	edasi_minema	10	2,48
maha_kukkuma	14	3,67	välja_ajama	9	2,70	üles_panema	11	2,47
välja_minema	41	3,67	ära_katkuma	10	2,67	ära_kaduma	7	2,46
läbi_pistma	14	3,67	ära_viima	17	2,59	maha_kukkuma	6	2,43
ära_kuivama	17	3,63	maha_tapma	7	2,53	lahti_päästma	6	2,43
peale_hakkama	25	3,62	ära_peksma	12	2,51	kinni_siduma	6	2,41
sisse_ajama	19	3,59	ära_keetma	10	2,49	üles_tõusma	6	2,41
ära_müüma	17	3,53	ära_eksima	7	2,48	mööda_minema	7	2,38
segamini_minema	16	3,53	ära_lõikama	11	2,46	kokku_ajama	6	2,35
ära_murdma	15	3,50	sisse_lööma	7	2,45	läbi_ajama	6	2,32

PÕHJA	k-sag	t-sk	LÕUNA	k-sag	t-sk	RANNIKU	k-sag	t-sk
maha_panema	29	3,46	ära_kaduma	8	2,45	külge_panema	6	2,31
ära_kaotama	13	3,44	katki_lööma	6	2,41	välja_paistma	6	2,30
ülal_istuma	12	3,43	välja_minema	15	2,41	ära_vajuma	6	2,28
maha_võtma	20	3,43	üle_minema	8	2,41	tagasi_minema	10	2,27
ära_lõhkuma	14	3,43	ühes_käima	7	2,37	välja_viima	8	2,26
välja_ajama	18	3,39	taga_ajama	6	2,37	püsti_tõstma	5	2,23
välja_laskma	18	3,38	ära_hõõruma	7	2,29	kokku_vedama	6	2,22
ära_tulema	77	3,34	ära_kuivama	7	2,29	sisse_lööma	6	2,19
üles_ajama	15	3,31	sisse_laskma	7	2,28	kokku_korjama	5	2,19
ära_lõikama	19	3,27	peale_valama	5	2,20	maha_jätma	5	2,18
üles_võtma	19	3,26	mööda_minema	6	2,20	üles_leidma	5	2,17
kokku_panema	22	3,25	maha_sadama	5	2,19	ära_rookima	5	2,15
ära_niitma	17	3,25	valla_laskma	5	2,19	ära_korjama	6	2,13

t-skoor järjestab ühendverbid kõigis murderühmades suhteliselt hästi ehk sagedamad ühendverbid saavad kõrgema t-skoori väärtuse. Kuigi analüüsisist on välja jäetud *saama*-ühendid, võis t-skoori väärtuste põhjal näha, et vastavad ühendid saavad suhteliselt madala väärtuse, mis vastab tegelikkusele, sest tegemist ei ole ühendverbide vaid adverbi ja abiverbi funktsioonis esineva verbivormi ühenditega. Eriti hästi oli seda näha põhjaeesti murderühmas, kus kasutatakse palju *saama*-passiivi. Sama kehtib paljude *pidama*-ühendite kohta nii põhja- kui lõunaeesti murderühmas.

Kuigi Stefan Evert on oma doktoritöös t-skoori kritiseerinud, pidades seda täiesti sobimatuks statistikuks sõnadevahelise seose mõõtmiseks (Evert 2004: 82–83), on ta teinud katseid, kus t-skoor on andnud häid tulemusi (Krenn, Evert 2001). Sama näitab ka murdematerjal.

t-skoori väärtuste põhjal tuleb selgelt välja statistikute kasutamise suurim eelis lihtsalt sageduste kokkulugemise ees – statistik võimaldab arvesse võtta sõnade eraldiesinemise sagedusi. Sellest tulenevalt saavad madalama väärtuse sagedastest sõnadest moodustunud ühendid, mis ei moodusta reeglina ühendverbi (*välja pidama, alla pidama, ära hakkama*). t-skoori tulemuste parandamiseks oleks vaja meetodit, mis välistaks sagedatest sõnadest moodustunud ühendverbide mitte väljajäämist vaatluse alt. Nii nagu on juhtunud lõunaeesti murderühmas, kus madala väärtuse on saanud mõned sagedatest sõnadest moodustunud ühendverbid (*ära tegema, sisse saama*). Üks võimalus selleks on kõrge koosinemissagedusega ja madalate t-skoori väärtustega ühendid käsitsi läbi vaadata, mis küll praeguse murdekorpuse mahu juures oleks mõeldav, kuid kindlasti mitte efektiivselt meetod.

Murdematerjali t-skoori väärtuse piiri, millest väiksema väärtusega ühendeid enam kokkukuuluvateks ei loeta, pole kerge seada. Osaliselt on see kindlasti tingitud murdekorpuse materjali vähesusest, sest üldjuhul töötavad statistikud paremini väga suuremahuliste andmekogude peal. Nende katsete põhjal võiks ranniku- ja põhjaeesti murderühmas seada t-skoori piirväärtuseks 1 ning lõunaeesti murderühmas 0,5, s.t kui t-skoor on väiksem ühest või 0,5-st, siis ühendit enam ühendverbiks ei loeta.

4.2. Vastatikuse informatsiooni väärtus (MI)

Vastastikuse informatsiooni väärtus MI leitakse valemiga:

$$MI = \log_2 \frac{O}{E},$$

kus O on osalause hulk, kus sõnad koos esinevad, ja E on teoreetiline sagedus.

Tabel 6 esitab kahanevalt 50 kõrgeima MI väärtusega ühendit.

MI eelistab harvaesinevaid ühendeid, mille eri osad on samuti korpusel madala esinemissagedusega (Evert 2009). Seda kinnitavad ka minu katsed. Everti järgi osutab negatiivne MI väärtus antikollokatsioonidele ehk sõnadele, mis pigem välistavad teineteise läheduses esinemise.

Põhjaeesti murderühmas said kõrgema negatiivse väärtuse pigem sageli koosesinevad sõnaühendid (*ära tegema, ära ütleva, ära käima, ära panema*) ning kontrolli käigus selgub, et tõepoolest pole tegemist sagedaste ühendverbidega. Kõrgem koosinemissagedus tuleneb asjaolust, et süntaksianalüsaator pole hakkama saanud korrektse osalausestamisega, mistõttu on üheks osalauseks loetud mitu ning sellest tulenevalt on ühendverbi kandidaate ülegenereeritud. Statistiku ise ei saa arvesse võtta, kas tegelikkuses on tegemist ühe või mitme osalausega. Lõunaeesti murderühmas saavad samalaadsed ühendid samuti kõrge negatiivse väärtuse, kuid antud murderühmas on tegemist siiski ühendverbidega, mis tähendab, et statistiku väärtus ei peegelda tegelikkust. Rannikumurde murderühmas sarnaselt põhjaeesti murderühmaga on küll paljud mitte-ühendverbid saanud madala väärtuse, kuid kõrge koha pingereas on saanud näiteks *välja varastama, pealt jääma, maha lõikama, välja lõppema*, mis ei esine tekstis ühendverbina. Seega on MI käitumine eri murderühmades mõnevõrra erinev.

Tabel 6. 50 kõrgeima MI väärtusega ühendit eri murderühmades (k-sag – ühendi koosinemissagedus samas osaluses, MI – MI väärtus)

PÕHJA	k-sag	MI	LÕUNA	k-sag	MI	RANNIKU	k-sag	MI
segamini_vaduma	3	9,16	mööda_sõitma	4	7,85	püsti_tõstma	5	8,13
ümbert_siduma	3	8,57	peale_siputama	3	7,51	maha_kukkuma	6	7,30
katki_närjama	3	7,65	ümber_mähkima	3	7,22	lahti_päästma	6	6,69
alla_neelama	3	7,41	maha_murduma	3	6,96	üles_lendama	3	6,18
kinni_sõlmima	3	7,24	kinni_haarama	3	6,95	tagasi_langema	3	6,12
ilma_jätma	3	7,07	ökva_tõmbama	3	6,80	kinni_siduma	6	5,99
kinni_katma	7	6,87	kinni_kängima	4	6,78	üles_tõusma	6	5,86
ülal_istuma	12	6,87	üles_atma	4	6,49	kokku_korjama	5	5,48
seest_kiskuma	3	6,81	üles_tõusma	4	6,49	katki_lõikama	3	5,44
alla_vajuma	3	6,74	perra_jätma	3	6,48	ära_triivima	3	5,35
ringi_pöörama	3	6,63	kinni_katma	4	6,37	ära_täitma	3	5,35
hiljaks_jääma	4	6,55	kinni_kõitma	10	6,23	ringi_keerama	4	5,34
üles_pooma	4	6,51	maha_kargama	4	6,06	üles_tõstma	7	5,29

PÕHJA	k-sag	MI	LÕUNA	k-sag	MI	RANNIKU	k-sag	MI
lahti_kaevama	3	6,24	katki_lööma	6	6,01	maha_jätma	5	5,26
üleväl_istuma	3	6,19	peale_valama	5	5,92	üles_leidma	5	5,18
kinni_kargama	5	6,10	maha_röökima	4	5,79	edasi_lükkama	3	5,15
üles_paisutama	4	5,93	maha_sadama	5	5,61	kinni_hoidma	4	5,08
peksa_andma	6	5,89	valla_laskma	5	5,56	maha_müüma	4	5,00
edekohe_andma	3	5,89	kinni_mähkima	4	5,46	ära_selgima	4	4,76
ümber_pöörama	5	5,80	välja_kiskuma	5	5,39	ära_rookima	5	4,67
kinni_siduma	31	5,76	ilma_jääma	4	5,26	maha_laskma	19	4,65
üles_tõusma	15	5,76	kinni_hoidma	12	5,07	kokku_ajama	6	4,59
maha_vajuma	3	5,73	alla_lööma	5	5,03	kaasa_tooma	3	4,57
maha_kukkuma	14	5,71	tähele_panema	3	4,96	taga_vahtima	3	4,43
katki_lõikama	9	5,69	maha_visrama	5	4,95	taga_ajama	4	4,39
katki_raiuma	3	5,68	taga_ajama	6	4,93	maha_lööma	12	4,35
edasi_lükkama	3	5,68	peale_keema	4	4,92	ära_kaotama	4	4,35
läbi_pistma	14	5,63	maha_tapma	7	4,51	järele_jääma	3	4,33
välja_rapsama	3	5,63	pakku_minema	3	4,50	sisse_laskma	16	4,30
lõhki_lööma	4	5,55	ära_kustuma	4	4,49	ümber_lööma	4	4,27
lõhki_lõikama	3	5,48	ära_puherduma	4	4,49	läbi_ajama	6	4,24
kokku_laduma	4	5,38	ära_kuivatama	12	4,27	ära_lahutama	11	4,22
ilma_jääma	6	5,33	ära_hääduma	5	4,22	edasi_ajama	4	4,22
tagasi_lükkama	4	5,30	ära_jahtuma	5	4,22	külge_panema	6	4,16
sisse_soolama	3	5,26	ära_külmuma	5	4,22	välja_varastama	3	4,14
kinni_tarima	3	5,24	ära_rehkma	5	4,22	pealt_jääma	3	4,11
järgi_andma	7	5,19	ära_survima	5	4,22	välja_paistma	6	4,04
põiki_laskma	3	5,18	ära_vassima	5	4,22	tagasi_keerama	3	3,99
välja_noppima	8	5,04	vahele_Panema	3	4,22	ära_üppuma	4	3,89
järele_jääma	11	5,01	üles_visrama	3	4,13	ära_vajuma	6	3,85
valmis_õmblema	3	4,98	üles_ajama	15	4,11	ära_kaduma	7	3,83
välja_kannatama	3	4,95	ära_vahetama	3	4,07	vastu_võtma	10	3,77
maha_jätma	11	4,90	mant_tulema	4	4,05	eemale_minema	4	3,76
püsti_seisma	3	4,87	ära_koolema	56	4,05	maha_lõikama	3	3,72
peale_laduma	5	4,82	läbi_ajama	5	4,05	välja_lõppema	3	3,70
kokku_lükkama	5	4,81	täis_sööma	4	4,03	peitu_minema	3	3,60
kokku_siduma	11	4,80	üles_käänama	3	4,02	sisse_ajama	5	3,59
ümber_keerama	3	4,78	ära_üppuma	10	4,00	maha_surema	5	3,58
vargil_käima	4	4,75	ära_leotama	5	4,00	alla_vedama	8	3,55
sisse_puurima	3	4,74	ära_väsima	5	4,00	ringi_käima	13	3,29

Selline tulemus on tugevasti mõjutatud asjaolust, et ühendit moodustavad sõnad on korpuses väga sagedad ning MI töö põhimõttele vastavalt peavadki nad madala väärtuse saama. MI väärtust mõjutab tugevasti ainult sõnade eraldiesinemise sagedus, nende koosesinemise sagedus samas osalauses ei paista tulemusele olulisest mõju avaldavat, nagu võib näha rannikumurde murderühmas. Rannikumurde murderühmas oli MI töö eriti tõhus, mis puudutas modaalverbi ühendeid – MI väärtuste viimase osa moodustasid suures osas just ainult adverbide ja modaalverbi ühendid, mis aitab välistada selliste ühendite arvestamist ühendverbide hulka.

Kuna MI on loodud madala esinemissagedusega ühendite leidmiseks, siis võimendab statistika selliste ühendite väärtusi üle, mistõttu satuvad nad pingereas kõrgetele kohtadele. Selle võimenduse tasakaalustamiseks on pakutud erinevaid lahendusi. Üks võimalus on kogu valem läbi korrutada sõnade koosesinemissageduse väärtusega (Evert 2009). Seda tehes muutub MI väärtuste loend märgatavalt – kõrgema MI väärtuse saavad sagedamini esinevad ühendverbid.

Teise võimalusena on välja pakutud ühendite koosesinemise sagedusele piiri seadmist, kõige efektiivsemaks on peetud kümmet, ehk sõnad, mis esinevad samas osalauses vähem kui kümme korda, jäetakse vaatluse alt välja (Evert 2009). Rakendades sama tehnikat Eesti murdematerjali peal, moodustub MI väärtuste pingerida n-õ väga tugevatest ühendverbidest, mis on kindlasti lähenemisviisi eelis. Antud materjali peal ei oleks siiski see lahendus, kuna suurem osa ühendeid esinebki korpuses samas osalauses vähem kui kümme korda, näiteks rannikumurde murderühmas jääks selle tulemusena alles vaid vähem kui 50 ühendit.

MI väärtuse üle, millest alates ühendit enam ühendverbiks ei peeta, pole kerge otsustada, Eesti murdekorpuse puhul oleks see ilmselt võimatu. Sellise väärtuse seadmine tuleks kõne alla, kui jätta välja näiteks vähem kui kümme korda esinevad ühendid ning järelejäänud ühenditele seada teatud piirväärtus. Nagu eelnevalt mainitud, poleks see siiski murdekorpust silmas pidades selle vähesuse tõttu ainumõeldav lahendus.

Eesti murdematerjali peal võiks MI tulemuse parandamiseks kombineerida erinevaid väljapakutud lahendusi. Näiteks võib esialgse kandidaatandmestiku põhjal leida madala esinemissagedusega ühendverbid, seades piiri, millest vaatluse alla jäävad ainult eespool olevad ühendid. Seejärel võiks järjestada esialgse loendi uuesti, jättes välja vähem kui kümme korda samas osalauses esinevad ühendid. Saadud loendile võiks rakendada samuti piirväärtuse, mis välistaks kandidaatide seast sagedasti samas osalauses koosesinevad mitte-ühendverbid.

Krenn ja Evert (2001) väidavad, et MI ei sobi üldse näiteks väljendverbide ja muude kujundlike püsiühendite leidmiseks. Murdekorpuse ühendverbide leidmisel on MI mõningate kitsenduse seadmisel suhteliselt efektiivne.

4.3. Hii-ruut-statistik

Hii-ruudu arvutamiseks kasutatakse kahemõõtmelist sagedustabelit (vt eespool) ja arvutatakse valemi

$$hii\text{-ruut} = \sum_{ij} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

abil, kus O on osalausetes hulk, kus sõnad koos esinevad, ja E on teoreetiline sagedus.

Kuna hii-ruut on n-ö kahepoolne mõõdik (s.t annab suure positiivse väärtuse nii tugevale seosele – sõnadele, mis kindlasti moodustavad püsiühendi kui ka tugevale “tõukumisele” – sõnadele, mis “väldivad” üksteise naabrust), jäädes kogu aeg positiivseks (ruudu väärtus on alati positiivne), peaks tulemuse läbi korrutama (–1)-ga, juhul kui $O < E$ (teoreetiline koosinemissagedus on suurem tegelikust koosinemissagedusest, s.t sõnad esinevad üksteise naabruses harvem, kui võib oodata nende eraldiesinemise sageduste põhjal), saades hii-ruudust ühepoolse mõõdiku, mis eristab nii positiivseid kui negatiivseid seoseid (Evert 2009).

Tabel 7 esitab kahanevalt 50 kõrgeima hii-ruudu väärtusega ühendit (väärtused on läbi korrutatud (–1)-ga, juhul kui $O < E$).

Hii-ruut-statistik eelistab sarnaselt MI-ga madala esinemissagedusega sõnu ja ühendeid, samas antakse kõrged väärtused n-ö väga tugevatele ühendverbidele. Hii-ruudu väärtuste järjestus sarnaneb MI väärtuste loendiga. Madala väärtuse põhjaesti murderühmas saavad sageli koosineda sõnad, mis üldjuhul ei ole ühendverbid (*välja pidama, välja saama, välja hakkama*).

Tabel 7. 50 kõrgeima hii-ruut statistiku väärtusega ühendit eri murderühmades (k-sag – ühendi koosinemissagedus samas osalauses, hii-ruut – hii-ruut statistiku väärtus)

PÕHJA	k-sag	hii-ruut	LÕUNA	k-sag	hii-ruut	RANNIKU	k-sag	hii-ruut
segamini_vaduma	3	1707,98	mööda_sõitma	4	913,86	püsti_tõstma	5	1391,25
kinni_siduma	31	1638,07	ära_koolema	56	858,84	maha_kukkuma	6	939,95
ülal_istuma	12	1384,91	kinni_kõitma	10	736,05	lahti_päästma	6	610,66
ümbert_siduma	3	1138,12	peale_siputama	3	543,91	maha_laskma	19	447,34
ära_surema	63	823,39	ümber_mähkima	3	440,82	kinni_siduma	6	371,02
kinni_katma	7	812,33	kinni_kängima	4	435,17	üles_tõusma	6	339,68
üles_tõusma	15	787,28	ära_surema	35	387,63	sisse_laskma	16	288,22
maha_kukkuma	14	708,43	kinni_hoidma	12	382,88	üles_tõstma	7	262,42
läbi_pistma	14	671,67	katki_lööma	6	377,78	maha_lööma	12	224,82
peale_panema	97	609,76	maha_murduma	3	371,09	kokku_korjama	5	214,02
katki_närjama	3	595,87	kinni_haarama	3	367,89	üles_lendama	3	213,65
ära_kaduma	51	538,64	üles_atma	4	353,46	tagasi_langema	3	203,66
alla_neelama	3	507,85	üles_tõusma	4	353,46	ära_lahutama	11	188,83
katki_lõikama	9	450,56	peale_panema	37	344,77	maha_jätma	5	183,11
kinni_sõlmima	3	449,24	õkva_tõmbama	3	329,49	üles_leidma	5	173,28
välja_tulema	92	448,49	kinni_katma	4	324,42	peale_hakkama	20	170,21
kokku_vedama	18	439,72	peale_valama	5	295,72	välja_vedama	20	167,42
ilma_jätma	3	399,20	perra_jätma	3	262,41	ringi_keerama	4	154,45

PÕHJA	k-sag	hii-ruut	LÕUNA	k-sag	hii-ruut	RANNIKU	k-sag	hii-ruut
lahti_võtma	34	385,92	maha_kargama	4	260,25	kokku_ajama	6	133,90
hiljaks_jääma	4	372,01	maha_sadama	5	236,67	ringi_käima	13	130,54
üles_pooma	4	359,69	üles_ajama	15	233,87	kinni_hoidma	4	128,28
sisse_panema	91	358,14	valla_laskma	5	228,31	katki_lõikama	3	124,43
peksa_andma	6	347,47	ära_kuivatama	12	217,27	maha_müüma	4	120,98
maha_laskma	29	346,67	maha_röökima	4	215,58	ära_rookima	5	120,47
järele_jääma	11	337,66	välja_kiskuma	5	201,29	vastu_võtma	10	119,93
kinni_kargama	5	334,98	maha_jääma	16	183,30	ära_triivima	3	119,18
seest_kiskuma	3	331,07	sisse_Panema	31	183,12	ära_täitma	3	119,18
alla_vajuma	3	315,19	ära_tapma	22	179,70	ära_surema	16	108,50
maha_jätma	11	310,54	taga_ajama	6	172,31	ära_võtma	30	105,34
vastu_võtma	19	301,97	kinni_mähkima	4	169,39	ära_selgima	4	103,31
ringi_pöörama	3	291,95	alla_lööma	5	154,04	läbi_ajama	6	102,67
kokku_siduma	11	287,27	maha_tapma	7	147,61	edasi_lükkama	3	100,78
ümber_pöörama	5	269,52	ära_uppuma	10	147,59	tagasi_tulema	18	99,62
ära_tapma	29	263,55	ilma_jääma	4	146,75	külge_Panema	6	98,58
ära_kaotama	13	261,44	maha_viskama	5	145,46	sisse_Panema	17	93,25
välja_noppima	8	251,09	ette_Panema	11	122,62	ära_kaduma	7	88,55
valmis_tegema	42	248,17	vastu_tulema	14	120,12	välja_paistma	6	88,17
järgi_andma	7	243,65	ühes_võtma	12	119,05	alla_vedama	8	79,48
ära_kuivatama	26	241,40	peale_keema	4	114,42	taga_ajama	4	76,77
üles_paisutama	4	237,17	tagasi_tulema	16	111,74	ära_vajuma	6	76,62
alla_Panema	39	232,90	vastu_võtma	10	108,97	ära_kaotama	4	75,53
ilma_jääma	6	232,41	ära_eksima	7	100,95	ümber_lööma	4	69,78
ära_võtma	103	229,70	ära_müüma	13	91,28	edasi_ajama	4	67,20
lahti_kaevama	3	221,43	tähele_Panema	3	90,50	kaasa_tooma	3	66,14
pealt_võtma	17	219,66	ära_hääduma	5	87,70	mööda_minema	7	60,70
üleval_istuma	3	213,94	ära_jahtuma	5	87,70	taga_vahtima	3	59,02
tähele_Panema	11	212,68	ära_külmuma	5	87,70	ette_Panema	8	58,26
ära_sulatama	10	211,90	ära_rehkma	5	87,70	läbi_käima	14	57,26
ette_Panema	31	188,77	ära_survima	5	87,70	järele_jääma	3	55,12
ära_viima	68	188,17	ära_vassima	5	87,70	ära_uppuma	4	52,84

Sarnaselt MI-ga ei piisaks ka hii-ruudu puhul kollokatiivsuse ja mitte-kollokatiivsuse piirväärtuse seadmisest, kuna nii oleks vaatluse alla jäävate ühendite loend suhteliselt kaootiline. Nii nagu MI puhulgi paraneks tulemus, kui välja jätta ühendid, mis esinevad koos samas osalauses vähem kui kümme korda. Sellisel juhul jällegi jääksid välja harvaesinevad ühendverbid ning kuna murdekorpuse materjal on suhteliselt väike, on paljude ühendverbide esinemissagedus väiksem kui kümme. Hii-ruut oleks murdematerjali peal tulemuslikum, kui välja jätta madala esinemissagedusega sõnadest koosnevad ühendid, sest erinevalt MI-st ei “lükka” ta pingerea lõppu kõrge esinemissagedusega ühendeid. Madala sagedusega ühendite tuvastamisel töötab paremini MI.

4.4. Log-tõepära funktsioon

Log-tõepära funktsiooni väärtuse leidmiseks kasutatakse samuti kahemõõtmelist sagedustabelit (vt eespool) ning arvutatakse valemi

$$\text{log-tõepära} = 2 \sum_{ij} O_{ij} \log \frac{O_{ij}}{E_{ij}}$$

abil, kus O on osalausete hulk, kus sõnad koos esinevad, ja E on teoreetiline sagedus.

Nii nagu hii-ruut on ka log-tõepära funktsioon kahepoolne mõõdik ning selle ühepoolseks muutmiseks tuleb statistiku väärtus läbi korrutada (-1) -ga, juhul kui $O < E$.

Tabel 8 esitab kahanevalt 50 kõrgeima log-tõepära funktsiooni väärtusega ühendit (väärtused on läbi korrutatud (-1) -ga, juhul kui $O < E$).

Kirjanduse põhjal võis eeldada, et parima tulemuse annab log-tõepära funktsiooni kasutamine ning Eesti murdematerjalil tehtud katsed kinnitasid seda. Log-tõepära funktsioon on loodud just arvutilingvistiliste rakenduste jaoks ja on leidnud palju kasutust selles valdkonnas. Seda peetakse sobivaks paljude erinevate ülesannete lahendamiseks (Evert 2009, Krenn, Evert 2001). Kõigis murderühmades peegeldab log-tõepära väärtuste loend kõige enam tegelikkust ning Everti järgi korreleeruvad log-tõepära funktsiooni väärtused kõige rohkem inimeste poolt antud hinnangutega. Samuti võis kõigis kolmes murderühmas näha, et log-tõepära annab pigem väiksemad väärtused *saama*-ühenditele, mis reeglina on abiverbi ja adverbi ühendid (eriti just põhjaeesti murderühmas). See osutab taas statistikute kasutamise eelisele lihtsalt ühendite sageduste kokkulugemise ees. Samuti ei võimenda log-tõepära funktsioon üle madala sagedusega andmestikku.

Log-tõepära funktsiooni väärtust, millest alates ühendeid enam koosinevateks ei loeta, pole murdematerjali peal kerge määrata. Nagu eelnevalt mainitud, on üheks põhjuseks korpuse väiksus. Tehtud katsete põhjal võiks väärtuse piir, millest alates sõnu enam koosinevateks ei loeta, olla põhjaeesti murderühmas umbes 13 miljardi kandis, lõunaeesti murderühmas 4,5 miljardi kandis ning kirderanniku murderühmas umbes 2 miljardit.

Table 8. 50 kõrgeima log-tõepära funktsiooni väärtusega ühendit eri murderühmades (k-sag – ühendi koosinemissagedus samas osalauses, log – log-tõepära väärtus)

PÕHJA	k-sag	log	LÕUNA	k-sag	log	RANNIKU	k-sag	log
peale_ panema	97	2,15E+11	ära_koolema	56	7,21E+10	maha_laskma	19	1,25E+10
ära_surema	63	2E+11	ära_surema	35	3,6E+10	ära_võtma	30	9,68E+09
välja_tulema	92	1,77E+11	peale_ panema	37	3,29E+10	sisse_laskma	16	9,54E+09
sisse_panema	91	1,61E+11	sisse_panema	31	2,24E+10	peale_ hakkama	20	9,42E+09
ära_kaduma	51	1,43E+11	ära_tapma	22	1,91E+10	välja_vedama	20	9,12E+09
ära_võtma	103	1,38E+11	ära_ kuivatama	12	1,81E+10	maha_ kukkuma	6	7,75E+09
kinni_siduma	31	1,31E+11	kinni_kõitma	10	1,61E+10	tagasi_tulema	18	7,32E+09
ära_viima	68	1,01E+11	üles_ajama	15	1,48E+10	maha_lööma	12	7,16E+09
valmis_ tegema	42	9,03E+10	kinni_hoidma	12	1,48E+10	ära_lahutama	11	6,98E+09
lahti_võtma	34	8,95E+10	maha_jääma	16	1,43E+10	ringi_käima	13	6,88E+09
alla_panema	39	8,34E+10	tagasi_tulema	16	1,25E+10	ära_surema	16	6,8E+09
üles_panema	50	7,75E+10	ära_uppuma	10	1,24E+10	sisse_panema	17	6,65E+09
maha_laskma	29	7,53E+10	välja_tulema	22	1,23E+10	püsti_tõstma	5	5,99E+09
ära_tapma	29	7,51E+10	vastu_tulema	14	1,2E+10	lahti_päästma	6	5,86E+09
ära_tooma	57	7,13E+10	ära_minema	65	1,16E+10	vastu_võtma	10	5,21E+09
välja_võtma	43	7,1E+10	üles_tulema	20	1,07E+10	välja_tulema	24	5,19E+09
ära_ kuivatama	26	6,8E+10	ette_panema	11	1,05E+10	üles_tõstma	7	5,15E+09
ette_panema	31	6,68E+10	ära_müüma	13	1,05E+10	kinni_siduma	6	5,13E+09
üles_tõusma	15	6,29E+10	ühes_võtma	12	1,04E+10	üles_tõusma	6	5,06E+09
läbi_käima	37	6,17E+10	kinni_panema	18	9,7E+09	ära_tulema	34	4,96E+09
ära_kaotama	13	5,84E+10	kinni_võtma	15	9,56E+09	läbi_käima	14	4,87E+09
kokku_ vedama	18	5,82E+10	vastu_võtma	10	8,96E+09	ära_viima	15	4,45E+09
ülal_istuma	12	5,81E+10	katki_lööma	6	8,8E+09	kinni_panema	13	4,27E+09
maha_ kukkuma	14	5,77E+10	ära_eksima	7	8,5E+09	välja_minema	22	3,9E+09
vastu_võtma	19	5,71E+10	ära_pesema	12	7,78E+09	ära_kaduma	7	3,86E+09
läbi_pistma	14	5,66E+10	manu_ panema	10	7,73E+09	külge_ panema	6	3,81E+09
tagasi_tulema	25	5,46E+10	mööda_ sõitma	4	7,71E+09	alla_vedama	8	3,81E+09
tähele_ panema	11	5,33E+10	üles_panema	16	7,62E+09	vastu_tulema	11	3,77E+09
kinni_panema	38	5,12E+10	kinni_ kängima	4	7,51E+09	ära_rookima	5	3,77E+09
ära_sulatama	10	4,87E+10	maha_tapma	7	7,5E+09	välja_võtma	13	3,75E+09
pealt_võtma	17	4,69E+10	peale_valama	5	7,34E+09	kokku_ korjama	5	3,71E+09
sisse_minema	43	4,08E+10	ära_häduma	5	7,25E+09	kokku_ajama	6	3,7E+09

PÕHJA	k-sag	log	LÕUNA	k-sag	log	RANNIKU	k-sag	log
kinni_katma	7	4,07E+10	ära_jahtuma	5	7,25E+09	mööda_minema	7	3,62E+09
sisse_laskma	22	4E+10	ära_külmuma	5	7,25E+09	maha_jätma	5	3,59E+09
järele_jääma	11	3,96E+10	ära_rehkma	5	7,25E+09	üles_leidma	5	3,57E+09
ära_lõhkuma	14	3,86E+10	ära_survima	5	7,25E+09	ette_panema	8	3,52E+09
ära_murdma	15	3,85E+10	ära_vassima	5	7,25E+09	üles_võtma	10	3,49E+09
ära_sööma	32	3,85E+10	ära_võtma	33	7,23E+09	peale_Panema	11	3,48E+09
ära_kuivama	17	3,79E+10	välja_võtma	13	7,17E+09	lahti_tegema	11	3,46E+09
maha_jätma	11	3,73E+10	taga_ajama	6	7,13E+09	valmis_tegema	8	3,4E+09
segamini_minema	16	3,64E+10	sisse_tulema	16	7E+09	läbi_ajama	6	3,4E+09
kokku_siduma	11	3,59E+10	välja_ajama	9	7E+09	ära_vajuma	6	3,32E+09
katki_lõikama	9	3,52E+10	valla_laskma	5	6,95E+09	välja_paistma	6	3,3E+09
ära_tulema	77	3,51E+10	üles_atma	4	6,92E+09	ära_selgima	4	3,14E+09
ära_lõppema	10	3,44E+10	üles_tõusma	4	6,92E+09	edasi_minema	10	3,04E+09
ära_müüma	17	3,39E+10	maha_sadama	5	6,89E+09	üles_Panema	11	2,89E+09
välja_minema	41	3,36E+10	ära_katkuma	10	6,68E+09	ringi_keerama	4	2,87E+09
sisse_ajama	19	3,18E+10	kinni_katma	4	6,62E+09	kinni_hoidma	4	2,76E+09
peale_hakkama	25	3,1E+10	välja_kiskuma	5	6,59E+09	üles_lendama	3	2,74E+09
välja_noppima	8	2,95E+10	ära_leotama	5	6,18E+09	kokku_vedama	6	2,71E+09

5. Kokkuvõte

Sõnadevahelise seose tugevuse mõõtmise statistikuid kasutatakse mitmetes arvuti-lingvistilistes rakendustes sõnadevahelise seose tugevuse määramiseks. Statistike kasutamise eeliseks lihtsa sageduse kokkulugemise ees on asjaolu, et statistikud võtavad arvesse ka ühendit moodustavate sõnade eraldiesinemise sagedused. Lisaks sõnade eraldiesinemise sagedustele on statistikute väärtuste arvutamiseks vajalik leida nende koosinemise sagedus, oodatav koosinemise sagedus ning valimimaht. Sõnadevahelise seose tugevuse mõõtmisel võib lähtuda erinevatest alustest. Kindlas naabruses koosinemisel arvestatakse vaadeldava sõna konteksti ehk akent, mille määrab uurija.

Käesolevas töös rakendasin Eesti murdematerjalil nelja statistikut: t-skoori, vastastikuse informatsiooni väärtust MI-d, hii-ruut statistikut ning log-tõepära funktsiooni (statistikute arvulised väärtused pole omavahel võrreldavad). Katsematerjaliks olid murdekorpuse morfoloogiliselt märgendatud ja lausestatud tekstid. Katsed tegin kolme murderühma peal eraldi.

Katsetulemuste pingeridade põhjal võib öelda, et omavahel sarnasemaid tulemusi annavad hii-ruut ning MI ja t-skoor ning log-tõepära funktsioon. Hii-ruut ja

MI eelistavad selgelt madala esinemissagedusega ühendeid, mille komponentide esinemissagedused korpuses on samuti madalad.

Võrreldes igas murderühmas nelja statistiku sadat kõrgeima väärtusega ühendit, võib näha, et ranniku- ja lõunaeesti murderühmas on väga sarnased ka hii-ruudu ning log-tõepära väärtused. Statistikutevahelised erinevused on väikseimad rannikumurde murderühmas ning suurimad põhjaeesti murderühmas. Üks põhjus on kindlasti materjali erinev hulk korpuses – põhjaeesti materjal on tunduvalt paremini esindatud. Omavahel kõige vähem kattuvad kõigis murderühmades MI ja t-skoori ning MI ja log-tõepära funktsiooni väärtused.

Ühtegi statistikut ei saa pidada teistest ühemõtteliselt paremaks.⁶ Erinevad statistikud sobivad erinevat tüüpi ülesannete lahendamiseks. Statistiku sobivus sõltub kollokatiivse seose tüübist, korpuse suurusest, valdkonnast, keelest jmt. Näiteks oli antud katsete puhul üsna määrav lausestaja töö efektiivsus, sest mida korrektsemalt on osalused eraldatud, seda adekvaatsem on kandidaatandmestik. Samuti võib statistikute töös näha erinevusi eri murderühmade vahel. Näiteks sobib MI paremini põhjaeesti murderühmale rakendades, sest annab madalad väärtused sagedastest sõnadest moodustunud mitte-ühendverbidele. Samuti toimib statistik ka lõunaeesti murderühmas, kuid siin esinesid sagedastest sõnadest moodustunud ühendid sageli ühendverbina, mis MI tööpõhimõttest tingituna saavad madalad väärtused.

Nii MI kui ka hii-ruut annavad kõrgemaid väärtusi madala sagedusega andmes-tikule, kuid sellise andmestiku jaoks paistab paremini töötavat siiski MI. Madala sagedusega ühendverbide eraldamiseks võiks MI-d kombineerida log-tõepära funktsiooniga. Log-tõepära funktsioon töötab selgesti kõige efektiivsemalt kõigis murderühmades, kuid MI eraldab väga hästi just madala esinemissagedusega ühendverbe. Seega ei oleks mõistlik valida ühte kindlat statistikut kogu materjali jaoks. Murrete ühendverbide leidmiseks sobib selgelt kõige paremini log-tõepära funktsioon, kuid parema tulemuse saamiseks peaks kasutama ka MI-d. Ranniku- ja lõunaeesti murderühmas võib madala esinemissagedusega ühendite leidmisel rakendada ka hii-ruut-statistikut.

Viidatud kirjandus

- Eesti kirjakeele korpus. <http://www.clut.ee/korpused/kasutajaliides/> (30.08.2009).
- Eesti murrete korpus. www.murre.ut.ee (30.08.2009).
- Evert, Stefan 2004. *The Statistics of Word Cooccurrences: Word Pairs and Collocations*. Dissertation, Institut für maschinell Sprachverarbeitung, University of Stuttgart. <http://elib.uni-stuttgart.de/opus/volltexte/2005/2371/pdf/Evert2005phd.pdf> (30.08.2009).
- Evert, Stefan 2008. *Association Measures*. <http://www.collocations.de/AM/> (01.09.2009).
- Evert, Stefan 2009. *Corpora and collocations*. – A. Lüdeling, M. Kytö (Eds.). *Corpus Linguistics. An International Handbook*. Berlin: Mouton de Gruyter, 1212–1249.
- Stefanowitsch, Anatol; Gries, Stefan Th. 2009. *Corpora and Grammar*. – A. Lüdeling, M. Kytö (Eds.). *Corpus Linguistics. An International Handbook*. Berlin: Mouton de Gruyter, 933–952.
- ISI – International Statistical Institute (The ISI Glossary of Statistical Terms). <http://isi.cbs.nl/glossary/> (01.09.2009).

⁶ Krenn ja Evert (2001) on muuhulgas väljendanud ka arvamust, et ükski meetod ei anna oluliselt paremaid tulemusi kui tavaline sageduse arvutamine (vt ka Uiboed 2010).

- Krenn, Brigitte; Evert, Stefan 2001. Can we do better than frequency? A case study on extracting PP-verb collocations. – Proceedings of the ACL Workshop on Collocations. Toulouse, France, 39–46.
- Lindström, Liina; Müürisep, Kaili 2009. Parsing Corpus of Estonian Dialects. – Eckhard Bick, Kristin Hagen, Kaili Müürisep, Trond Trosterud (Eds.). Proceedings of the NODALIDA 2009 workshop. Constraint Grammar and Robust Parsing. May 14, 2009. NEALT Proceedings Series, 8. Odense: Northern European Association for Language Technology, 22–29.
- Manning, Christopher D.; Schütze, Hinrich 2003. Foundations of Natural Language Processing. London: The MIT Press.
- Quantum Computing and Quantum Information Processing. National Laboratory for Scientific Computing. <http://virtualo1.lncc.br/~giralddi/qc/itheory.html> (01.09.2009).
- Uiboaed, Kristel 2008. Ühendverbid Eesti murrete korpuses. Magistritöö. Tartu Ülikool, eesti ja üldkeeleteaduse instituut.
- Uiboaed, Kristel 2010. Ühendverbid eesti murretes. – Keel ja Kirjandus, 1, 17–36.

Kristel Uiboaed (Tartu Ülikool). Uurimisvaldkonnad on korpuslingvistika, eesti murded, statistilised meetodid keeleteaduses.
kristel.uiboaed@ut.ee

STATISTICAL METHODS FOR PHRASAL VERB DETECTION IN ESTONIAN DIALECTS

Kristel Uiboaed

University of Tartu

The aim of this study was to assess different statistical methods of automatic collocations extraction from the corpus. To extract the collocations, association measures (AM) were applied and the association scores (AS) for the collocation candidates found in the corpus were calculated. An AS indicates the collocational strength between two words. An advantage of the AMs is the fact that in addition to the co-occurrence frequency, the marginal frequencies of collocating words are also taken into account. To calculate the AS, the following data is needed: co-occurrence frequency, marginal frequencies of collocating words, expected frequency and the sample size.

There are different approaches to applying AMs: words can be considered collocational only if they appear in the same collocational span, in one text unit (clause, sentence, utterance), or if they carry together some syntactic function. This paper attempts to apply AMs for phrasal verb detection from the Corpus of Estonian Dialects (CED). Texts of CED were morphologically tagged and parsed. Combinations of adverbs and verbs were extracted and AS was calculated for every collocation candidate. Experiments were run on three different dialect groups applying four different association scores: t-score, Mutual Information, chi-squared test and log-likelihood.

The results indicate that log-likelihood and t-score outperform MI and chi-squared test. The outcomes of different measures vary the most in the Northern dialect group. The best measure for dialect data in general is log-likelihood. However, MI and chi-squared test work well with low frequency data. In the Northern dialect group the best AM for low-frequency phrasal verb detection is MI, however, in the North-Eastern and Southern groups chi-square test works well for the same purpose. To achieve better results different scores should be combined.

Keywords: computational linguistics, corpus linguistics, dialectology, methods and tools, statistics, Estonian

KUIDAS ÕPETADA KÖNESÜNTEESAATORILE EMPAATIAT? EMOTSIOONI AUTOMAATSE TUVASTUSE VÕIMALUSTEST EESTIKEELSE KIRJALIKUS LAUSES SISALDUVA INFO PÕHJAL

Ene Vainik

Ülevaade. Kõnesünteesi loomulikustamise üheks eelduseks on arvuti võime tunda tekstist ära emotsionaalseid lauseid. Artiklis tutvustatakse eestikeelse kirjutatud teksti automaatse emotsioonituvastaja (AET) loomise taustaks olevaid teoreetilisi seisukohti ja kirjandust, millest võib ülesande lahendamisel abi olla. Seejärel antakse ülevaade inimeste poolt piisava kindlusega ära tuntud emotsioonilauseste mitmetest parameetritest ja hinnatakse nende rakendatavust arvuti tulevases emotsioonituvastuses. Osutub, et automaatse emotsioonituvastaja loomisel ei ole võimalik otse üle võtta välismaiseid eeskujusid, kuna emotsioonide väljendamine on mõningal määral kultuurisidus ja nende väljendus kirjalikus tekstis ripub ära konkreetse keele eripärast. Tuvastuseks vajalikud keeleressursid – emotsioonidega seotud eesti keele sõnade ja konstruktsioonide loendid – tuleb AET jaoks algusest peale ise luua.*

Võtmesõnad: emotsioonid, kirjalik tekst, kõnesüntees, tundmusanalüüs, afektirehkendus, eesti keel

1. Sissejuhatus

Empaatia tähendab võimet mõista teise inimese tundeid ja emotsioone ning loomulikku inimestevahelist suhtlust me ilma selleta ette ei kujuta (vt nt Goleman 2006). Ka inimese ja arvuti suhtluse arendamine on jõudnud niikaugemale, et mingil määral oodatakse arvutitki inimese emotsioonidega “samal lainel olemist”. Nii on näiteks eesti keeletehnoloogia arendamise raames kerkinud ülesanne muuta kõnesüntesaatori poolt (re)produtseeritavate lausungite kõlapilti loomulikumaks,

* Selle artikli valmimisel on tulnud finantstugi kahest allikast: sihtfinantseeritav teema nr SF0050023s09 ja Eesti Teadusfondi grant nr 7149.

reaalsele inimsuhtlusele lähedasemaks (Altrov, Pajupuu 2007). Muude meetmete (nagu nt hingamispauside, vt Kerge jt 2008) kõrval plaanitakse süntesaatorit õpetada kõnes võimalikult inimlähedaste akustiliste vahenditega väljendama tekstiga kaasnevaid emotsioone.

Selleks, et süntesaatori poolt ette loetud tekst mõjuks võimalikult loomulikuna, peab kõlapilt olema vastavuses teksti autori poolt kogetud emotsiooniga (Iida jt 2003, Ververidis, Kotropoulos 2006). Selle eesmärgi saavutamise juures on kaks olulist lõiku: 1) modelleerida emotsioonide väljendamisega kõlalised parameetrid ja 2) õpetada arvutit kirjaliku teksti põhjal aru saama, millist emotsionaalset registrit peaks süntesaator igal antud juhul rakendama. Käesolev artikkel tegeleb neist teisena mainitud ülesande võimalike lahenduste tutvustamisega ja püüab hinnata nende perspektiivikut eesti keele spetsiifikat arvestades.

Üks on selge – selleks, et arvutit saaks “õpetada” kirjalikus tekstis sisalduvat emotsiooni tuvastama, peame olema veendunud, et emotsioonidel on kirjalikus tekstis tõepoolest mingid märgilised kandjad olemas. Kuni ei ole saanud üldiseks tavaks lisada kirjutatud lausesse spetsiaalseid emotsioonidele viitavaid märke (nt emotikone)¹, tuleb eeldada, et emotsiooni tuvastamiseks piisab eestikeelses kirjalikus lauses sisalduvatest muudest kirjamärkidest ja et tavainimesed tulevad emotsiooni tuvastamisega kirjutatud lauses sellele infole tuginedes üldjuhul suuremate probleemideta toime. Kuivõrd need eeldused aga paika peavad, sellest allpool lähemalt.

Artiklis tutvustatakse esmalt taustaks olevaid teoreetilisi seisukohti emotsioonide keeles väljendumisest ning antakse ülevaade kirjandusest, millest võib eesti keele kirjaliku lause emotsioonituvastaja (edaspidi AET) loomisel abi olla. Empiirilises osas esitatakse nn andmekaeve (ingl *data mining*) tulemused, mille käigus analüüsiti 361 eestikeelset lauset, milles inimesed olid emotsiooni olemasolu või puudumise piisava üksmeelega kindlaks teinud. Andmekaeve eesmärgiks oli välja selgitada kõikvõimalikud emotsioonide tunnuslikud parameetrid kirjutatud tekstilauses ja hinnata nende rakendatavust arvuti tulevases emotsioonituvastuses. Artikli viimasel osas vaetakse tulemusi ja arutletakse AET loomise perspektiivide üle.

2. Kuidas saavad emotsioonid verbaalses suhtluses väljenduda?

Selleks, et õpetada arvutit emotsioone tekstist ära tundma, peaksime esmalt teadma, kuidas need emotsioonid seal põhimõtteliselt saavad esindatud olla. Kas on üldse enesestmõistetav, et iga lausega kaasneb ilmtingimata emotsioon?

Üldiselt arvatakse, et mingisugune emotsionaalne seisund või meeleolu saadab inimest kogu aeg. Vastavalt sellele, mis oludega tuleb kokku puutuda, kerkivad lühiaegselt esile intensiivsemad reaktsioonid ning nende situatsioonikohast tõlgendamist kogetaksegi emotsioonide ja tunnetena (Damasio 1999). Kui me vaikimisi eeldame, et emotsionaalne foon (oma teatud fluktuatsioonidega) saadab inimest kogu aeg, siis on täiesti loomulik eeldada, et see ei kao ära ka kõneprotsessi ajal, et see toimib keelelise kommunikatsiooniga samaaegselt. Selles mõttes võib tõesti emotsioon saata põhimõtteliselt iga inimliku autoriga lausungit. Samas on selge, et emotsioonide kõnes “kohaloleku” määr varieerub suuresti nii isikuti kui ka eri kultuurides ning žanrides (Wilce 2009, Biber, Finegan 1989).

¹ Ettepaneku nn multialfabeedi laialdaseks kasutuselevõtuks on teinud Valdur Mikita (2008).

Teoreetilises plaanis on emotsioone ja keelelist suhtlemist peetud kuuluvaks interpersonaalse kommunikatsiooni eri liikidesse. Emotsioon on loomult spontaanne, kasutades nn analoogkoodi, ja keel on sümboliline, kasutades digitaalset koodi (Buck 1989, Foolen 1997, Watzlawick jt 1967). Kuivõrd aju-uuringud näitavad kummagi kommunikatsiooniliigiga seotud aktiivsete keskuste paiknemist aju eri osades, on emotsioone ja keelelist suhtlemist peetud teineteisest sõltumatuteks nähtusteks. Teised uurijad väidavad siiski, et emotsioonid ja keel on küll kaks paralleelset suhtlusvahendit, kuid et nad on siiski vastasmõjulised (Reilley, Seibert 2003). See seisukoht sobib kokku ka keeleteaduses traditsiooniliselt juurdunud tõekspidamisega, et keelele on muude funktsioonide kõrval omane ka ekspressiivne funktsioon (Foolen 1997 ja tema viide Bühler 1934).

Kognitiivne lingvist Ad Foolen oletab, et kõnelejate emotsioonid ei väljendu mitte ainult kontseptualiseerituna (referentsiaalses plaanis/ideatsiooniliselt, nt öeldes, et keegi on *kurb*), vaid et emotsioonidel on kõnes avaldumiseks ka otsetee (1997). See otsetee tähendab seda, et põhimõtteliselt digitaalne ja analoogsõnum põimuvad. Seda seisukohta toetavad valdavalt psühholoogide tööd, milles uuritakse, kuidas inimesed tuvastavad emotsioone kõne prosoodiast lähtuvalt (nt Scherer jt 2008). Kuid emotsioonide ja kõne vastasmõju ei piirdu kindlasti üksnes prosoodiaga. Emotsioonide verbaalse väljendumise otsetee võib tähendada ka seda, et digitaalse koodi (keele) elementide seas leidub selliseidki, mis vahendavad emotsionaalset sõnumit pigem analoogse kodeerimise võtteid kasutades, näiteks ikooniliselt. On näiteks esitatud hüpotees, et afekti või hinnangu väljendamiseks sobivad eelkõige need keelendid, mis väljendavad kogemuse, omaduse või tegevuse intensiivsust ja/või selle kestust ning korduvust (Downes 2000).

Emotsioonide kõnes väljendamine ja nende äratundmine näib olevat midagi baasilist, mis omandatakse varases lapseeas koos keele õppimise ja sotsiaalse kohanimisega. Antropoloogid on leidnud, et afekti verbaalne väljendamine hakkab peale üksiksõna tasandist ja kõik grammatilised vormid positiivse ja negatiivse afekti väljendamiseks omandatakse juba enne viiendat eluaastat (Ochs 1988, viidatud Wilce 2009: 44 kaudu).

Nii lingvistid kui ka antropoloogid on ühel nõul, et emotsioonidel ei ole kõnes mingit neile ainuomast kindlat väljendusviisi – jälgi emotsioonidest võib leida pea kõigil nn traditsioonilistel keeletasanditel (vt Foolen 1997 ja tema viited; Wilce 2009). Eesti keele peale mõeldes: foneetilises plaanis on juba Andrus Saareste (1927) juhtinud tähelepanu häälikute labialiseerimisele beebidega suheldes või siis sõnaalgulise hääliku liigsele aspireerimisele (nn järsk alge) peente proudade erutatud kõnes (nt *Khuidas nii?*, *Vhäega hilus!*, *See on ju phöörane!*). Leksikonis avalduvad emotsioonid näiteks emotsiooninimetustena (nt Vainik 2002), piltlike kirjeldustena (nt Vainik, Velt 2006), vande- ja kirumissõnadena (Voll 2002); emotsionaalsele laetusele osutavad terved sõnaklassid nagu hüüdsõnad või rõhumäärsõnad (Veski 1982). Morfoloogias avaldub emotsionaalne suhtumine nt deminutiivsete ja pejoratiivsete sufiksrite, kliitikute kasutamise läbi; süntaksi valda kuulub nt spetsiaalsete lausetüüpide nagu nn kogeja- või eksklamatiivlause olemasolu; toimuda võivad muutused tavapärases sõnajärjes ja infostruktuuris. Ülalootletuga pole emotsioonide eestikeelses kõnes avaldumise võimaluste ring arvatavasti kaugeltki ammendatud. Mil määral ja kuidas täpselt avalduvad emotsioonid aga eestikeelses kõnes selle kirja pandud, prosoodiast “vabastatud” kujul, vajab samuti alles lähemat uurimist.

3. Kas me saame teiste kogemusi üle võtta?

Kui nn põhiemotsioonid on kultuuriülesed ja universaalsed (Ekman 1982), ning ühised meil koguni teiste imetajatega (Darwin 1998 [1872], viidatud Wilce 2009 vahendusel), siis emotsionaalsete reaktsioonide tõlgendamine ja kogemine on kultuuriliselt mingil määral varieeruv (Wilce 2009) ja annab aluse kategooriate (emotsioonimõistete) loomiseks, mis eri keelte ja kultuuride lõikes sugugi üks-üheselt ei kattu (Kövecses 2006, Wierzbicka 1999). Seetõttu ei saa kindlasti eeldada, et kusagil teisel maal loodud emotsioonituvastaja suudaks tuvastada eestlaste emotsioone. Veel enam, kui eesmärgiks on luua emotsioonituvastaja kirjaliku teksti tunnuste põhjal, siis see on loomulikult seotud konkreetse keelega ja selle kirjaviisiga, mis juba iseenesest välistab teiste keelte ja kultuuride kogemuse automaatse ülevõtmise. Võtta saab aga eeskujuna ja seda on kindlasti mõtet teha, kuna emotsioonide kirjutatud keeles väljendumise põhiprintsiibid võivad olla sarnased.

Keeleteaduses on tekstide hinnangulisuse ja afektiga tegeldud peamiselt tekstilingvistika vallas (nt Biber, Finegan 1989, Hunston, Thompson 2001). Rakenduslikku suunda, mille raames tegeldakse meetodite väljatöötamisega selleks, et oleks võimalik automaatselt tuvastada tekstides sisalduvat hoiakut, hinnangut või emotsiooni nimetatakse *tundmusanalüüsiks* (ingl *sentiment analysis*). Tehisintellekti loomise paradigma raames tegeldakse sarnaste asjadega ka *afektirehkenduse* nime all (*affective computing*) (Picard 1997, Yan jt 2008)². Suuremad edusammud on mõistagi keskustes, kus ülesandele on võimalik pühendada suuremaid ressursse (nii inim-, rahalisi kui juba täies mahus toimivaid keele- ja teadmusressursse). Eelisaasendis on inglise keel, mille suuremahulised märgendatud korpused, tesaurused, WordNet, FrameNet, ConceptNet ja spetsiaalsete afektisõnastike olemasolu loomulikult hõlbustavad uute täppsihitusega rakenduste loomist ja testimist. Ka vene keele puhul teatatakse toimivatest lahendustest³. Praeguse seisuga on üldine pilt kirju ning kirjutatud teksti emotsiooni määramise alased uuringud varieeruvad mitmes suhtes, mida alljärgnevalt põgusalt kirjeldatakse.

Esiteks võib tekstiline sisend ehk analüüsi objekt mahult varieeruda. Osa lähenemistest analüüsib ja rehkendab terviktekstide emotsionaalset hoiakut (Shanahan jt 2006, Bracewell jt 2006), osa keskendub lõigule (Iida jt 2003), kõige fokuseeritumad lähenemised üritavad määrata lause, sh isegi osalause emotsiooni (nt Osherenko 2008, Wu jt 2006, Al jt 2007, Wilson jt 2005, Liu jt 2003). Kuivõrd eesti keele kõnesünteesis on akustilise modelleerimise põhiühikuks lause (Mihkla 2007), siis AET loomisel on kasulikumad kogemused just fokuseeritud lähenemistest.

Teiseks varieeruvad lähenemised selles osas, mida õigupoolest soovitakse määrata: kas konkreetset emotsiooni (nt hirm, viha, rõõm, kurbus) või emotsioone üldisemalt iseloomustavaid parameetreid. Esimese lähenemisviisi puhul eeldatakse, et emotsioonid on kategooriana nn loomulik liik (ingl *natural kind*), mille liikmed erinevad üksteisest situatsioonist ja nimetustest olenemata (Hacking 1999, viidatud Wilce 2009 kaudu). Sellest eeldusest järeldub, et eristatavad peavad olema ka üksikute konkreetsete emotsioonide (nt viha, kurbuse, rõõmu) märgilised väljendused tekstis. Teise lähenemisviisi – nn dimensionaalse emotsioonimudeli – puhul ei otsita tekstist mitte niivõrd viiteid konkreetsetele emotsioonidele, kuivõrd keskendutakse emotsioone iseloomustavate üldisemate parameetrite (nagu valents, intensiivsus) tuvastamisele (Obrenovic jt 2005). Rohkem, kauem ja seni suurema

² Kirjandust tasub ostida ka otsingusõnadega nagu *affect sensing, assessing of sentiment ja parsing of affect*.

³ Väidetavalt toimiv teksti emotsionaalsuse ja muude psühholoogiliste parameetrite kaardistamise süsteem VAAL 2000 on loodud vene, ukraina ja inglise keele jaoks, vt <http://www.vaal.ru/proekt/vaal2000.php> (25.02.2010), (vt ka Beljanin 2000: 218).

eduga on tegutsetud teksti autori üldise hinnangu ja arvamuse polaarsuse (positiivne vs. negatiivne) ning intensiivsuse (aktivatsiooni) automaatse tuvastamise alal (nt Osherenko 2008, Osherenko, André 2007, Al jt 2007, Wilson jt 2005). Kui eesmärgiks on seatud konkreetsete emotsioonide tuvastamine, siis on nende arv tavaliselt piiratud vaid nn põhiemotsioonidega (nt Liu jt 2003) või kasutatakse nn vahepealse üldisuseastmega kategooriad (nt *õnnelik, õnnetu, neutraalne*) (Wu jt 2006, Bracewell jt 2006). On ka mõned lähenemised, mis hõlmavad suuremat emotsioonide ringi ja seovad neid üldiste klassimääratlustega (*positiivne, negatiivne, neutraalne*) (Grefenstette jt 2006, Bednarek 2008, Mathieu 2006). On autoreid, kes peavad teksti hinnangulisuse määramist ja emotsiooni tuvastamist eri ülesanneteks, mida saab lahendada jupikaupa (Al jt 2007) ning ka autoreid, kes peavad emotsioone hinnangutega lahutamatu seotuks (Bednarek 2009). Selles osas ei ole eesti keele loodava emotsioonituvastaja puhul erilist valikut – AET peab olema suuteline tuvastama konkreetseid emotsioone (esialgu tundma teksti põhjal ära kurbust, viha, rõõmu ja neutraalsust). Seda seetõttu, et eesti kõne emotsionaalsete aspektide akustilise modelleerimise ja sünteesimise aluseks on võetud just kategoriaalne ja mitte dimensionaalne emotsioonimudel ning sellele vastavalt on loodud ka emotsionaalse kõne korpus (vt Altrov 2008).

Lähenemised erinevad ka selles suhtes, kas eesmärgiks seatakse nn kõne autori emotsiooni/hinnangu tuvastamine (nt Biber, Finegan 1989, Hunston, Thompson 2001) või üritatakse tekstist leida vihjed emotsioonidest referentsiaalse kõnelemise kohta, mis võib olla ka nn mitte-autori emotsioon (nt Bednarek 2008). Näiteks hüüatused ja hüüdlauseid, samuti piltlik keelekasutus seostuvad kahtlemata keele ekspressiivse funktsiooniga (Foolen 1997), mispuhul emotsiooni pigem väljendatakse kui kirjeldatakse. Seevastu otseste emotsiooni- ja tundenimetuste (nagu *rõõm, viha, kurbus, nõrdimus, kadedus* jne) kasutamine (eriti kellegi teise) emotsioonide reflekteerimiseks kujutab endast põhimõtteliselt teistsugust tegevust ja seda seostatakse keele deskriptiivse funktsiooniga⁴. Kuigi nii mõnedki emotsioonituvastuse süsteemid ennast selle eristusega ei vaeva, oleks eesti keele süntesaatori otstarvet silmas pidades ilmselt oluline aru saada, millal lause otseselt väljendab mingit emotsiooni (ja järelikult tuleb mõnd spetsiifilist akustilist registrit rakendada) ja millal lause täiesti emotsioonivabalt kirjeldab või mainib kellegi kolmanda isiku emotsiooni (mispuhul eri registrit rakendada poleks tarvis). Lauset *Oh, sa reo!* (ekspressiivne lausung) võiks süntesaator ju öelda tunderõhuliselt, kuid deskriptiivset lausungit *Naabrimees oli eile väga vihane ja karjus oma naise peale* poleks süntesaatoril vaja karjuda, kuigi see sisaldab endas otseseid vihjeid emotsioonile (*vihane, karjus*).

Julgelt võib öelda, et enamik lähenemisi kasutab emotsiooni/hinnangu tuvastamiseks tekstist leitavaid leksikaalseid tunnuseid. Kasutatakse nn emotsionaalseid võtmesõnu (ingl *emotional keywords*) või nn leksikaalset külgetõmbavust (*lexical affinity*), mis teatud lekseemidel, nagu nt *vähk* või *majanduskriis*, emotsioonide suhtes on (vt ülevaadet ja kriitikat Wu jt 2006: 166, Liu jt 2003: 126). Selleks kasutatakse kas juba olemasolevaid afektisõnastikke (nt Whissell 1989, viidatud Osherenko 2007 vahendusel), kirjeldatakse nende loomist (nt Grefenstette jt 2006, Mathieu 2006) või üritatakse kasutada ka üldotstarbelistes sõnastikes sisalduvat infot (Osherenko 2007). Mõningal juhul kombineeritakse otsest emotsioonisõnavara nn lokaalse grammatikaga, mis tähendab teatud kinnistunud leksikaal-

⁴ Hea ülevaate selle dihhotoomia esile toomisest eri autorite poolt leiab allikast (Bednarek 2008: 10).

grammatiliste mallidega arvestamist (Hunston, Sinclair 2001, Bednarek 2008). On tehtud katseid tuvastada emotsiooni mitte ainult otsese, vaid ka piltliku, metafoorse keelekasutuse põhjal (nt Zhang 2008, 2009). Kriitikana leksikaalse lähenemise vastu on välja toodud, et mitte kõik leksikast lähtuvad süsteemid ei ole lause tasandil afekti kalkuleerimise puhul võimelised arvestatama eituse valentsi muutva mõjuga (Liu jt 2003). Näiteks võib süsteem lause *Täna polnud just õnnelikem päev minu elus* ainult eksplitsiitse emotsioonisõnaga *õnnelikem* arvestades selle eksikombel “õnnelikuks” klassifitseerida. Teiseks ei tunne üksnes leksikonile toetuvad süsteemid emotsiooni ära nn “ridade vahelt”, juhul kui lauses puudub otsene emotsioonile viitav lekseem (nt jääb tuvastamata nõrdimus lause *Vanemad tulevad ööseks koju üksnes magama* puhul). Sellel põhjusel on katsetatud ka süsteeme, mis arvestavad emotsiooni määramisel üldiste maailmateadmistega, toetudes selleks suurtele kodifitseeritud teadmusbaasidele nagu ConceptNet ja Open Mind Commonsense (Liu jt 2003).

Loodav AET hakkab ka kindlasti arvestama leksikaalse tasandi infoga, samuti peab ta olema valmis toime tulema eitavast kõnest sugeneva võimaliku tähendusmuutusega lause tasandil. Selleks läheb tarvis ühest küljest toetavat keeleressurssi – emotsioonituvastuses relevantsete sõnade loendit, ning teiseks algoritmi, mis arvestuslike emotsionaalsete “väärtuste” koosmõju lause tasandil kalkuleeriks. Eesti keele aglutineeriv-flekteerivat tüpoloogilist eripära arvestades võib osutada otstarbekaks ka morfoloogiliste tunnuste ja lõppude koosinemise muustritega arvestamine. Raskeks läheb kahtlemata arvutile “üldiste maailmateadmiste” õpetamine, mille põhjal ta saaks teha järeldusi emotsioonide kohta lausetes, kus neid otse mainitud ei ole.

4. Mida saaks arvuti õppida inimestelt?

Järgnevalt kirjeldan andmekaeve tulemusi, mis on saadud katsetest, kus eesti keelt emakeelena oskavad inimesed on lausete emotsioone üksnes lause kirja pildi põhjal määranud. Eesti keele emotsionaalse kõne korpuse⁵ loomise raames on välja otsitud sobiva pikkusega emotsionaalsed tekstilõigud (loe täpsemalt Altrov 2008), tükeldatud need lauseteks, lausete ja lõikude järjekord tahtsi segi aetud ning lastud katseisikutel otsustada iga lause puhul n-ö kontekstivabalt, kas lausega kaasnev emotsioon on *kurbus, viha, rõõm, neutraalne* või *ei oska öelda*. Selliseid teste, kus tuli lause emotsiooni üle otsustada ilma heli kuulmata, oli kokku kuus. Andmete analüüsi algusajaks (25.03.09) oli iga testi sooritanud keskmiselt 18,2 inimest. Testide sooritajaid oli kokku 55 (31 naist ja 24 meest, keskmine vanus 33,3 aastat, STDEV 11,51). Mõni inimene osales ka mitmes testis, kuid mitte keegi neist ei osalenud sama korpuse kuulamiskatsetes.

⁵ Eesti emotsionaalse kõne korpus ja sellega seotud testid on liginetavad aadressilt [http://193.40.113.40:5000/\(2.03.3010\)](http://193.40.113.40:5000/(2.03.3010)). Kirja pildi järgi emotsioonide tuvastamise katsed peituvad sildi taga EMOTSIOONITEST (ilma helita)1–6.

4.1. Emotsiooni äratundmine üksiku lause põhjal

Kõigepealt peab ütleva, et üksnes kirjalikul kujul ja ilma kontekstita esitatud lausete puhul ei olnud katseisikutel ilmselt sugugi lihtne emotsiooni või selle puudumist (neutraalne) määrata. Seda näitab määrangute suhteliselt suur hajuvus. Tabelist 1 nähtub, et kokku ainult 15 lause puhul oli kõikidel katseisikutel täiesti üksmeelne arvamus selle kategoriseerimise kohta. See moodustab kõigest 2,61% nendest 574 lausest, mida inimestel lasti hinnata. Samas peab ütleva, et lauseid, mida kõik oleks üksmeelselt emotsioonide suhtes määramatuks pidanud, ei olnud sootuks – iga lause kohta oli võimalik võtta vähemalt seisukoht, kas see on neutraalne või sisaldab mingit emotsiooni. Selline tulemus näib kõnelevat sellest, et iga lause võib potentsiaalselt mingi emotsiooniga seostuda, kuid et tugevaid ja seega hästi äratuntavaid emotsioone kogetakse (ja väljendatakse keeleliselt) suhteliselt harva. Teisest küljest näitab selline tulemus, et kirjalikus tekstis ei pruugi olla sugugi piisavalt märgilist informatsiooni emotsioonide tõsikindlaks tuvastamiseks. Viimane järeldus ei ole AET loomise eesmärki silmas pidades just kuigi lootustandev. Jääb ka kolmas võimalus, et informatsioon emotsioonide kohta on siiski lausetes olemas, kuid katseisikud mingil põhjusel ei suuda seda kasutada, näiteks nende isikliku madala empaatia tõttu⁶. Igatahes saab teha järelduse, et kui ka lõigu tasandil on emotsioon üldiselt tuvastatav (Iida jt 2003), siis mitte iga lause neis lõikudes pole eraldivõetuna võrdse väljendusjõuga.

Lapsed saab paigutada kontinuumile nende emotsionaalse väljendusjõu selguse järgi, mida näitab hinnangute üksmeelse määr. Kõige üksmeelsemalt (100%) tunti viha ära näiteks järgmistes lausetes: *Ei saa nii, et töökoht tuhnib su eraelus! Iga-sugustel peksta ma ennast ka ei lase! Praegune tants alkoholipiirangute ümber muutub juba labaseks ja naeruväärseks tsirkuseks!* Kurbus tunti üksmeelselt ära näiteks järgmistes lausetes: *Ma ei tea, kuidas edasi... Mul ei ole enam kodu. Tean, et tal on tööl jamad, kuid ikka jääb tunne, et äkki olen mina süüdi või et ta tahab meie juurest ära minna... Rõõmsaks peeti üksmeelselt lauset: *Fantastiline staadion ja fantastiline publik!* Neutraalsus väljendus kõige selgemini lausetes: *Ma ei tunne Kadri Kukke isiklikult ja Seitsekümmend krooni, palun.* Kontinuumi teises otsas on laused, mille osas katseisikud kõige rohkem kahtlesid, näiteks *Siis on see ilus kannatus* (kas rõõm või neutraalsus?), *Eh, sõbrad, nii see ikka ei käi!* (viha või neutraalsus?) ja *Korraldus on ikka väga "kõva"* (rõõm või viha?). Iseloomulik on, et need laused sisaldavad vastuolulist infot. Tõenäoliselt kujuneb ambivalentsete tunnete nagu nostalgia, kahjurõõmu, ironia jm mitmeplaaniiliste tundmuste tuvastamine pähkliks ka arvutile.*

Tabel 1. Erineva üksmeelsuse määraga ära tuntud emotsioonid

	100%	75%	50%
viha	7	69	178
kurbus	5	34	87
neutraalne	2	16	80
rõõm	1	51	86
ei oska öelda	0	0	4
Kokku	15 = 2,61%	170 = 29,62%	435 = 75,78%

⁶ 55-st katseisikust sooritasid empaatiat mõõtvat testi 20 ja said seal keskmiselt täiesti normikohase empaatiamaära: 42,35 (STDEV 8,30) (vrd Baron-Cohen, Wheelwright 2004).

4.2. Võimalikke tunnuseid, mis aitavad emotsioone ja neutraalsust lauses eristada

Viisteist 100-protsendilise kindlusega määratud lauset on väga vähe selleks, et selgitada, millised märgilised tunnused võiksid arvutil aidata emotsioone kirjalikus tekstis ära tunda. Seetõttu on edasises sõnavara ja muude karakteristikute analüüsis võetud arvesse laused, mille puhul vähemalt 51% vastanutest on määranud emotsiooni/neutraalsuse ühtemoodi. Laused, mida pole osatud määrata, on jäetud välja; välja on jäetud ka eri testides kordunud laused (kokku 78). Arvesse tulevate lausete arv on 361 ja see moodustab 63% katses testitud lausetest. Uuritavate hulgas on 129 viha lauset, 84 rõõmlauset, 76 kurbuselauset ja 72 neutraalset lauset. Kuna eri emotsioonidega lausete arv ei ole võrdne, on järgnevas analüüsis karakteristikud andmete võrreldavuse mõttes normeeritud (arv 100 lause kohta). Et lisaks leksikale võib olla informatiivne ka nt emotsionaalsete lausete suhteline lühidus (Saareste 1927), lauselõpumärk (Laur 1987), sõnaliigiline koosseis vm tunnused mistahes keeletasanditelt, siis vaadeldaksegi järgnevas nn andmekaeve korras mitmesuguste tunnuste distributsiooni emotsiooniti võrreldes samade tunnuste esinemisega neutraalsetes lausetes.

Lõpumärgistus on üks käegakatsutavamaid asju, mis näib mingil määral peegeldavat lause emotsiooni. Umbes kolmandik lausetest (115) sisaldas mõnd tavapärasest punktist erinevat lõpumärki⁷ – kas hüüumärki, küsimärki, hüüu- ja küsimärgi kombinatsiooni või kolme punkti. Tabelis 2 on näidatud, kuidas eri emotsiooniga laused lõpumärgistuse alusel eristuvad neutraalsest lausest, mis kaldub valdavalt lõppema punktiga. AET-i jaoks võiks sõnastada tendentsi: kui lause sisaldab hüüu- või küsimärki või kolme punkti, on emotsiooni tõenäosus suur. Küsi- ja hüüumärk (nii eraldi kui kombinatsioonis) ennustavad viha võimalust, üksik hüüumärk rõõmu võimalust ning kolm punkti kurbuse võimalust. See tulemus on hästi kooskõlas Mall Lauri uurimusega, kes leidis samuti, et nii hüüu- kui küsimärk ennustavad suure tõenäosusega mitmesuguseid aktiivseid emotsioone ja punkt seostub kas neutraalsuse või kurbusega (1987). Samas ei saa unustada, et praeguses uurimuses ligi kaks kolmandikku lausetest lõppes punktiga ja nendest vaid umbes kolmandik (28%) olid hinnatud tõesti neutraalseks. Ülejäänute osas oli kõigi kolme emotsiooni (viha, rõõm, kurbus) osakaal võrdne. Järelikult ei saa AET lauselõpu punkti põhjal veel selle neutraalsust järeldada ja lahti jääb võimalus, et tegemist võib siiski olla emotsionaalse lausega.

Tabel 2. Lauselõpumärkide distributsioon emotsiooniti (arv 100 lause kohta)

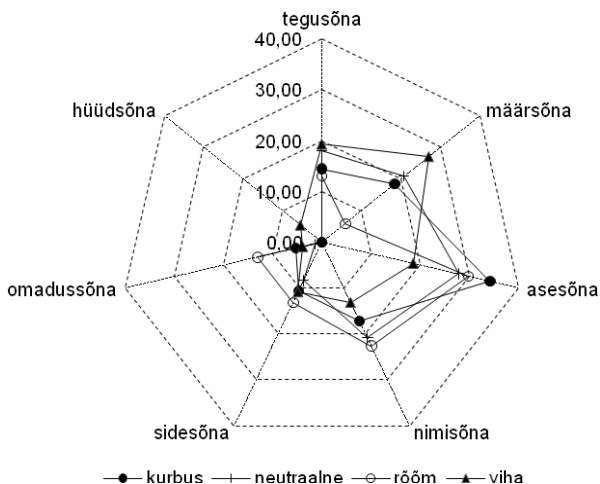
	neutraalne	rõõm	kurbus	viha
küsimärk	2,78		1,32	10,08
hüüumärk		34,52	7,89	28,68
kolm punkti			13,16	0,78
küsi- ja hüüumärk			2,63	8,53
punkt	97,2	65,5	73,68	50,4

⁷ Nende märkide esinemust ei vaadatud mitte ainult lause absoluutses lõpus, vaid ka näiteks lausesse sisestatud otsese kõne jm juhtudel.

Eitav kõne on samuti tunnus, mis võiks oluline olla. Näib loogiline arvata, et negatsioon on seotud negatiivsusega mitte üksnes lause sisus, vaid ka tundetoonis. Kui vaadata adverbide *ei* ja *mitte* ning verbi *olema* vormide *pole*, *polegi*, *polnud* esinemust uuritavas materjalis, siis selgus, et kõige rohkem seostuvad nad kurbusega. Üle poole 100 kurbuselausest (57,8) sisaldas mõnda neist sõnadest. Neutraalsetest lausetest oli eitav kolmandik (33,3%) ja vihalausestest 40%. Rõõmsatest lausetest sisaldas neid sõnu aga vaid 15,48%. Seega osutubki, et eitavat kõnet on rohkem lausetes, milles tunti ära negatiivsed emotsioonid (viha, kurbus). Positiivsete emotsioonidega (rõõm) laused pole küll kõik jaatavas kõnes, kuid enamikul juhtudest (84%) on see nii. Valdavalt on jaatavas kõnes ka neutraalsed laused (66,6%). Sellised statistilised tendentsid iseloomustavad emotsioonilauseid üldiselt ja ei ole eraldi võetuna kindlasti rakendatavad ühe konkreetse lause üle otsustamisel. AET jaoks võiks siiski sõnastada tendentsi: eitava kõne olemasolul on suurem negatiivsete emotsioonide tõenäosus.

Lause pikkus on järgmine asi, mille poolest emotsionaalsed laused võiks neutraalsest eristuda. Saareste on emotsionaalsete lausete puhul maininud nende tunnusest lühidust ja alistuvate kõrvallauseste puudumist (1927). Materjali analüüs osutab, et sellesuunalist tendentsi neutraalsete lausete keskmise pikkuse (9,64 sõnet) ja emotsionaalsete lausete keskmiste pikkuste vahel polnud võimalik tuvastada (rõõm 11,3; viha, 11,2; kurbus 9,8). Emotsionaalsed laused olid neutraalsetest keskmiselt isegi veidi pikemad. Lause keskmine pikkus sõnedes ei ole kindlasti mingi kriteerium AET-le, kuna keskmise sisse mahub pikkuse varieerumine väga suurtes piirides (alates ühesõnelistest lausungitest, nagu *Absurd!*, *Aitab!*, *Miks?*, kuni 23-sõnelise pikima lauseni antud valimis). Kõige rohkem varieerub rõõmsa lause pikkus (STDev 5,18 sõnet). Kui lausete pikkust mõõta aga tähekohtades, siis torkab silma, et kurbuselaused on teistest emotsioonidest (viha, rõõm) keskmisel tervelt 15,5 tähekohta lühemad, järelikult iseloomustab neid koosnemine suhteliselt lühematest sõnedest. Keskmine sõne pikkus kurbuse lauses ongi teistest lühem – 4,83 tähemärki (kõigis teistes lausetüüpides stabiilselt 5,5). Siit võiks vahest küll AET jaoks sõnastada tendentsi: kui lauses on suhteliselt rohkem lühikesi sõnesid (kuni 4 tähemärki), siis on põhjust kahtlustada kurbust.

Lause esimese sõna liik on veel selline pisut pealiskaudsem tunnus, mille järgi arvutil võiks olla lihtne orienteeruda. Ligikaudu neljandik uuritud lausetest (26%) algas mingit sorti asesõnaga, 19% määrsõnaga, 18% nimisõnaga, 17% tegusõnaga, 11% sidesõnaga ning 2% hüüdsõnaga. Joonisel 1 esitatud radiaaldiagramm näitab lausete esimeste sõnade sõnaliigilist kuuluvust emotsiooniti võrrelduna neutraalse lausega. Osutub, et tegusõna lause alustajana ei ole kuigi informatiivne, küll aga näiteks määrsõna. Määrsõnaga algaval lausel on suurem tõenäosus olla viha lause ja suhteliselt väiksem tõenäosus olla rõõmu lause. Asesõna lause alguses ennustab kõige rohkem kurbust (eriti kui see on 1. isiku asesõna) ja kõige vähem viha. Omadussõna lause alustajana esineb mingil määral vaid rõõmu lausetes. Sidesõna esineb lause alguses pigem emotsionaalsetel kui neutraalsetel lausetel. Olgu osutatud, et tendentsid on küll olemas, kuid kindlasti pole need kindlad reeglid. On ilmne, et lause esimese sõna liigiline kuuluvus võib olla abistavaks, kuid kindlasti mitte otsustavaks kriteeriumiks emotsioonide *vs.* neutraalsuse tuvastamisel.



Joonis 1. Lause esisõnade liigiline distributsioon

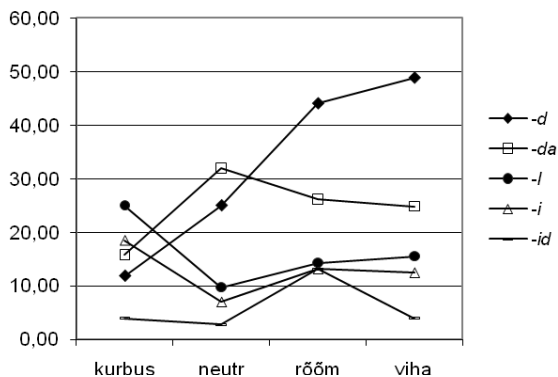
Kõikide lauses sisalduvate sõnade liigi määramine on kahtlemata juba ressursimahukam tegevus. Kui eeldada, et neutraalses lauses on sõnade liigiline jaotumus optimaalne, siis silma torkavad mõned iseloomulikud kõrvalekalded sellest. Nii kurbuse kui ka viha lausetes esineb rohkem määrsõnu (mõlemas keskmiselt 200) kui neutraalsetes (ca 160), seevastu rõõmlausetes on määrsõnu neutraalsest vähem (keskmiselt 116). Adverbide märksa väiksem esinemus rõõmu lauses võib osaliselt seletuda eelpool sedastatud tendentsiga, et nad on valdavalt jaatavas kõnes (s.t ei sisalda sagedasi adverbe *ei* ja *mitte*). Kui kõigis teistes lausetüüpides on nimisõnu keskmiselt pisut üle 200, siis kurbuse lauses vähem (159).

AET jaoks võiks sõnastada tendentsi: kui lauses on kaks või rohkem adverbi, tuleks kahtlustada viha või kurbust. Kui adverbe on lauses üks või vähem, siis on tõenäolisem neutraalsuse või rõõmu võimalus. Teine tendents oleks: kui lauses on nimisõnu üks või vähem, on suurem kurbuse võimalus.

Mitmesuguste morfoloogiliste markerite⁸ iseloomulik distributsioon tuleb samuti kõne alla ühe võimaliku kriteeriumina, mis võiks arvutil aidata erineva emotsiooniga lauseid ära tunda. Mõnevõrra üllatuslikuna tuli välja, et kõikides analüüsitud lausetüüpides oli eristatavate morfoloogiliste markeritega sõnavormide arvu suhe ilma markeriteta sõnavormide arvu 2 : 3. See tähendab, et lauses iga tunnust või lõppu omava sõna kohta tuleb poolteist sõna, mis on kas muutumatud või o-morfeemiga. Kurbuselausetes ilmnes pisut suurem tendents kasutada markerita sõnu, mis on üheks seletuseks eelpool sedastatud tõsiasi, et kurbuse laused olid teistest mõnevõrra lühemad mitte niivõrd sõnede kui tähekohtade arvu poolest.

Kui eeldada, et neutraalsetes lausetes on kõikvõimalikke markereid optimaalselt, siis torkab silma mõningate markerite suhteline vähesus või rohkus emotsioonidega lausetes. Joonis 3 kujutab suuremaid esiletuleku erinevusi näidanud markerite distributsiooni. Silma torkab markeri *-d* suur esinemus rõõmu ja viha lausetes ja vähesem esinemus kurbuse lausetes. Valdavalt on tegemist mitmuse nominatiivi tunnusega ja jääb mulje, et mingil määral selle kasutamine korreleerub tundeelamuse aktiivsuse või intensiivsusega.

⁸ Laused analüüsiti Filosoofi morfanalüsaatoriga ja ühestati käsitsi. Markeriteks nimetatakse siin analüsaatori poolt eristatavaid tunnuseid ja lõppe.



Joonis 3. Morfoloogiliste markerite jaotumus emotsiooniti (arv 100 lause kohta)

Selline tulemus on kooskõlas teoreetilises osas esitletud hüpoteesiga, et afekti või hinnangu väljendamiseks sobivad need keeleelemendid, mis väljendavad kogemuse omaduse või tegevuse intensiivsust ja/või selle kestust, korduvust (Downes 2000). Veel torkab silma, et kurbuse lausetes on kõiki markereid võrreldes neutraalsega vähem, välja arvatud *-l* (adessiivi lõpp) ja *-i* (mitmus või minevik). Rõõmlauseid eristab markeri *-id* (mitmuse osastav või minevik) suhteline rohkus. Pelgalt morfoloogiliste markerite äratundmise ja nende statistilise jaotuvuse põhjal on väga raske sõnastada AET-le juhust selle kohta, mida järeldada näiteks konkreetses lauses ette tuleva mitmuse nominatiivi tunnuse *-d* põhjal. On siiski ligi kaks korda suurem tõenäosus, et lause viitab mõnele aktiivsemale emotsioonile (rõõm, viha) ja poole väiksem tõenäosus, et ta viitab kurbusele, kuid välistada ei saa ka neutraalsust.

4.3. Sõnavara

Uuritud 361 lauses kasutati kokku 3134 sõnet, neist 1545 erinevat. Keskmiselt tuli iga sõne andmebaasis seega esile 2 korda, kuid reaalsus on see, et suurem osa (76%) sõnesid esines vaid ühekordselt. Sõnesid, mis tulid materjalis esile 2 või enam korda, oli 369. Materjal on ilmselgelt liiga väike, et teha paikapidavaid statistilisi järeldusi igale emotsioonile iseloomuliku korduva sõnavara kohta. Eriti arvestades asjaolu, et sagedusloetelude tipus esinevad sõnad, mis on eesti keeles üldiselt sagedama esiletulekuga ja seega vaevalt on kõnekad lause emotsionaalse eriväärtuse seisukohast. Kõnekam on see osa sõnavarast, mis tuleb esile ainult teatud kindla emotsiooniga lausetes ja mitte mujal. Nagu kirjanduse ülevaatest selgus, tasub teiste uurijate kogemusi arvestades pöörata eritähelepanu ühest küljest nn emotsionaalsetele võtmesõnadele (milleks on otsesed emotsiooni- ja tundenimetused, nagu *viha*, *rõõm*, *kurbus*, vt lähemalt Vainik 2002) ja teisest küljest nn leksikaalse külgetõmbega sõnadele (sõnad, millega antud kultuuris kaasneb kas positiivne või negatiivne emotsioon või konnotatsioon, nagu *vähk* või *majanduskriis*).

4.3.1. Võtmesõnad

Sõnu, mida annaks kvalifitseerida otseste emotsiooninimetuste hulka, tuli materjalis esile 55. Eeldades, et ühes lauses ei olnud mitut emotsiooninimetust, teeb see emotsioonidele otse (mõisteliselt) viitavate lausete osakaaluks kõigest umbes 14%. Tegelikult on aga selliste lausete osakaal veel väiksem, kuna siiski leidis lauseid, mis kasutasid mitut otsest emotsioonisõna, nt *Kui enne suhtusin sinusse heatahtlikult ja usaldusega, siis sinu avaliku valetamise tõttu suhtun sinusse nüüd umbusu ja ükskõiksusega. Mind väsitab moodsa kohustus olla ikka ja alati roosa ning rõõmus*. Mõlemaid näiteid peeti vihalauseteks.

Lauses sisalduva emotsiooninimetuse põhjal ei saa kummati mitmelgi juhul üks-üheselt järeldada selle lause enda emotsiooni. Näiteks mainiti kurbuselauses ka positiivseid emotsioonisõnu nagu *armastama* ja *hoolima*, tõsi küll, eitavas kõnes. Kurbust ennast mainiti aga ainult ühes lauses. Lisaks mainiti ka pettumust, viha, süüd, nõrdimust, löödud olemist ning mitmel eri viisil hirmu (*kartma*, *õudne*, verb *julgema* eitavas kõneviisis). Rõõmulauses mainiti samuti rõõmu vaid ühel korral, see-eest korduvalt mainiti õnne (*õnnelik*) ja armastust (*armas*, *armastama*). Mainiti ka meeldimist, tänulikkust, rahulolu ja üllatust. Vihalausetes mainiti viha kahel korral, selle kõrval aga ka pettumust, imestust, muretsemist, umbusku, väsimust, ärritust, õudu. Positiivsetest emotsioonisõnadest tulid vihalausetes esile rõõm ja usaldus (vt ülal esitatud näiteid).

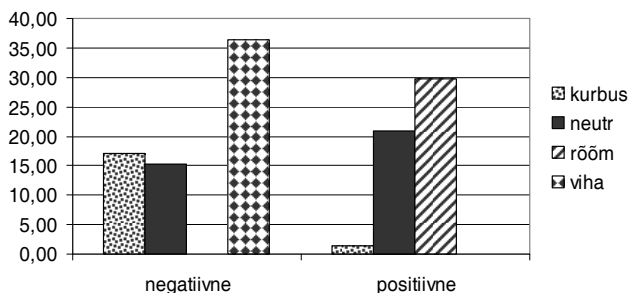
Tähelepanuväärne on, et otsest emotsiooninimetust sisaldanud lausete hulgas oli väga vähe neutraalseid lauseid (kokku 3). See tähendab seda, et otsene emotsiooninimetuse ennustab suure tõenäosusega lause neutraalsest erinevat kvaliteeti. Kahtlemata on see hea uudis AET-le, kuna saab sõnastada selge tendentsi: kui lause sisaldab mõnda otsestest emotsiooninimetustest, on neutraalsus vähe tõenäoline. Seega peaks AET osaks kindlasti olema eestikeelsete emotsiooni- ja tundenimetuste loend. Sellest loendist üksi jääb aga kindlasti väheks, kuna tervelt 83% emotsioonilausetest, mis katseisikute poolt olid piisava kindlusega määratud, ei sisaldanud mingit sellekohast otsest viidet. Suurem osa teksti põhjal määratud emotsioonidest väljendus kuidagi teisiti kui otseste emotsiooninimetusega.

Veel üks võimalik nn võtmesõnade rühm on emotsioonidega seotud semantiliselt, keelekasutaja tavateadmiste kaudu. Need on sõnad, mis kirjeldavad või nimetavad emotsioonide nn väljendusliigutusi (nt *naeratus*, *pisarad*, *näkku irvitama*, *aplaus*, *ovatsioonid*), füüsilist väljaelamist (*peksma*, *möllama*) või verbaalset väljaelamist (*hüüatama*, *käratama*, *sõimama*). Selliseid sõnu sisaldavate deskriptiivsete lausete osakaal oli materjalis suurim rõõmu puhul (10,7 lauset saja kohta), viha ja kurbuse puhul vähe (vastavalt 3,8 ja 2,6). Ka selliste väljendite loend peaks kuuluma AET koosseisu.

Võtmesõnadeks võiks olla ka kinnistunud piltlikud emotsioonikirjeldused, kuid neid tuli uuritavas materjalis esile väga vähe (nt *Hing kargas täis!* (viha); *Ja Mu hing läks katki...*, *Ma läksin liimist lahti...* (kurbus)), kuid kahtlemata tuleks needki AET jaoks loenditena esitada. Mida aga esitada ei saa, ja mis jääbki ilmselt peavalu valmistama, on uudisloominguline piltlikkus, nt *Nagu oleks Eestisse jäänud üks suur auk, kuhu on koos Juhaniga langenud suur osa meie inimlikkusest* (kurbus).

4.3.2. Leksikaalne külgetõmme

Üks strateegia, mida katses osalenud isikud nähtavasti kasutasid emotsioonide omistamisel lausetele, kus puudusid otsesed emotsiooninimetused või vihjed emotsioonide väljendustegevusele, oli emotsiooni järeldamine lauses sisalduva hinnangulisuse põhjal. Nagu kirjanduses osutatud, on hinnangut mitmesugust, ning osa sellest on tihedasti seotud lausele omistatava emotsiooniga (Bednarek 2009). Joonisel 4 on näha hinnangut sisaldavate lausete distributsioon aineses. Kuivõrd hinnangut esines ka neutraalsetes lausetes, ei ole ka see üksi võetuna piisavaks emotsioone eristavaks kriteeriumiks. AET-l on kindlasti oluline arvestada ka hinnangu valentsi, kuna siin ilmneb kindlasuunaline tendents, mis on näha ka joonisel 4: negatiivset hinnangut sisaldav lause ei ole väga tõenäoliselt rõõmus ning positiivset hinnangut sisaldav lause ei ole suure tõenäosusega vihane ega kurb.



Joonis 4. Hinnangu distributsioon lausetüüpides (arv 100 lause kohta)

Leksika, mille järgi hinnangut tekstis potentsiaalselt võiks ära tunda, on mitmekesine ja vajab kahtlemata veel põhjalikumat läbiuurimist. Olgu siinkohal vaid loetletud ja näitlikustatud mõningad ainekust esinenud võimalused. Näiteks negatiivset hinnangut implitseerivad adjektiivid nagu *haige, halb, kalk, lootusetu, arusaamatu, keeruline, raske, segane, kahtlane, kättesaamatu, küündimatu, labane, laisk, müüdav, naeruväärne, oskamatu, paks, süüdimatu, rumal, tobe, õudne*. Positiivset hinnangut implitseerivad *hea, suurepärane, loomulik, perekne, rahulik, suursugune, tasakaalukas, vajalik, huvitav, fantastiline, õige, elujõuline, geniaalne, ligitõmbav, professionaalne, täiuslik, tore, uhius*. Substantiividest kalduvad endaga siduma negatiivset hinnanguväärtust nt *hilp, punt, lapsikus, kiusamine, valetaja, vastutamatus, mõttelaiskus, kriis, majanduskrahh* ja lisaks rodu poliitikaga seonduvat sõnavara: *minister, president, poliitik, riigiasutus, riik, valitsus, alkoholiäärang, hinnatõus, kommunismikuritegu*. Positiivne hinnanguväärtus võib kaasnedda sõnadega nagu *arukus, soojus, avarus, sõber, sõbranna, lapsepõlvesõber, soojendamise, privileeg, armumine, kingitus*. Esile tulnud verbidest sisaldavad negatiivset hinnangulist komponenti nt *liialdama, katkema, korrumpeeruma, juhtuma, kulutama, passima, peksma, trügima, tuhnama, valetama*; positiivset võitma, *tasakaalustama, mõtestama, mõistma, väärima, meeldima*.

Tundub mõistlik, et AET oskaks arvesse võtta lause hinnangulisusest sugenevat võimalust emotsionaalseks interpretatsiooniks. Selleks tuleb moodustada

hinnanguliste kõrval- ja põhitähendustega sõnade loend, milleks mõningat abi pakub kindlasti Piret Volli eesti keele konnotatiivseid aspekte käsitlev magistritöö (2002). Lisaks tuleb hinnanguliste adjektiivide ja verbide leidmiseks süstemaatiliselt läbi töötada sõnastikud, ning võimalik, et korraldada ka katseid. See on kindlasti üks AET loomise ressursinõudlikumaid etappe. Lihtne ei saa lause hinnanguväärtuse määramise ja sellest emotsiooni järeldamise õpetamine arvutile küll sellegipoolest kindlasti olema. Esiteks juba eituse võimaliku mõju tõttu ning teiseks ka selle tõttu, et komplekssete lausete eri osades võidakse kõrvutada või vastandada eri hinnanguväärtustega nähtusi, mille mehaaniline kalkuleerimine ei pruugi anda adekvaatset tulemust.

4.4. Ekspressiivsus

Üks raskesti mõõdetav tunnus, mis emotsionaalseid lauseid neutraalsetest ilmselt eristab ja mille järgi katseisikud arvatavasti intuiitiivselt otsuseid langetasid, on ekspressiivsus ehk väljenduslikkus. Ekspressiivse lausungi puhul on aru saada, et see väljendab lausungi autori enda samal hetkel kogetavat emotsiooni. Ekspressiivset lausungit võib iseloomustada lühidus, hüüdelisus, retoorilised küsimused, direktiivsus, ütte kasutamine, polaarsed adjektiivid, intensiivistavad adverbid, kordus, hüüdsõnad, vande- või kirumissõnad, kinnisvormelid, deminutiivid, pejoratiivid, piltlik keelekasutus (Foolen 1997). Näiteid ainekust esinenud röömsatest ekspressiividest: *Suur tänu kõikidele häälekatele fännidele! Kõik ei ole veel kadunud, sõbrad! Ja milline publik! Ilus! Hea Kadri, ma soovin, et Sul elus hästi läheks! Kurbust väljendavad nt See oli nii suur kaotus! Ma ei tunne, et minust midagi sõltuks. Ma ei tea, kuidas edasi... ning viha: Absurd! Sotti ei saa! Paluks mitte ühtegi piüksu kriisist ja katastroofist! No saage aru! No mis sa tahad nendest naistest saada?! Nii ju ei lähe! Nagu öeldakse, ei au ei häbi! Kus on nüüd politsei?! Aitab!*

Saja lause kohta oli ekspressiivseid kurbuse puhul umbes 9, rõõmu puhul 8,3 ja viha puhul 18,6. Seega ei kujuta ekspressiivsus endast kaugeltki emotsioonilause vältimatut tunnust. Teisest küljest on ekspressiivsete lausungite äratundmine süntesaatori jaoks üks kõige olulisemaid ülesandeid, kuna just siin on vaja rakendada tavalisest erinevat intonatsiooniregistrit. Arvatavasti on suhteliselt raske arvutile selgeks teha kõiki ekspressiivsuse kriteeriume, kuna selles vallas on vahendite repertuaar väga mitmekesine. Loetletud tunnustest on lühidus, hüüu- ja küsimärgid kergesti leitavad, samuti ilmselt ütte. Interjektsioonid moodustavad suletud sõnaklassi ja neidki saab AET-le nimekirjana ette anda. Samuti on mõeldav koostada piltlike ja kinnisväljendite loendid. Samas on ebaselge, kuidas õpetada eristama arvutit retoorilist küsimust harilikust või retoorilist direktiivi tavalisest käsust-keelust.

5. Kokkuvõte ja arutelu

Kuidas siis õpetada arvutile empaatiat, kuidas õpetada teda üksnes kirjaliku lause põhjal selle võimalikku emotsiooni ära tundma? Üks on selge – arvuti ei ole inimene, ta ei saa ennast samastada lause autori või protagonistiga ja ta ei loe midagi oma multimodaalsele kogemusele tuginedes välja n-ö “ridade vahelt” (vrd Marmolejo-Ramos jt 2009). Arvutil tuleb lähtuda palju pealiskaudsematest tunnustest, ta saab otsida ja tuvastada ainult seda, mis on kirjas digitaalsetel märgiridadel.

Käesolevas artiklis sai läbi arutatud tähtsamad teoreetilised seisukohad ja mujal tehtud tekstianalüüsimeetodite sobivus eeskujudena. Mida on sellest ülevaatest kõrva taha panna arvutile empaatiat õpetamise tarbeks? Põhimõtteliselt on emotsiooni kaasnemine lausega alati võimalik, kuivõrd lause kujutab endast lausuja eneseväljendust. Samas ei saa öelda, et see oleks nii alati ja enesestmõistetavalt. Normaalne, vaimselt terve inimene kogeb ja väljendab (tugevaid) emotsioone vastavalt situatsioonile ainult aeg-ajalt ja sedasama peaks tegema ka “empaatiline” kõnesüntesaator. Arvutigi peaks esiteks oskama eristada neutraalset teksti emotsionaalsest ja teiseks tegema vahet olulisematel kirjas kajastuvatel emotsioonidel. Teiseks on teksti ja selles sisalduvate lausete emotsionaalsuse määr kindlasti žanrist. Douglas Biber ja Edward Finegan leidsid ulatusliku tekstikorpuse (500 teksti) läbisõelumise tagajärjel, et ainult 3% uuritud ingliskeelsetest tekstidest sisaldasid märkimisväärselt afektile viitavat sõnavara (adverbid, adjektiivid ja verbid). Kõige enam viidati seejuures emotsioonidele isiklikes kirjades (Biber, Finegan 1989). Käesoleva artikli autori meelest ei saa seetõttu eeldada, et mistahes eestikeelset teksti žanrist olenemata iseloomustaks leksikaalselt markeeritud emotsionaalsus. Usutavasti on ka eesti keele puhul emotsioonide väljendamine mõnes tekstiliigis ootuspärasem kui teistes. Emotsioonituvastus ja žanrianalüüs võiksid tulevikus üksteist vastastikku hõlbustada.

Kindlasti on kasulik olla kursis meetoditega, mida teiste keelte puhul on emotsioonide või hinnangute tuvastamisel rakendatud, kuid otse ülevõtmiseks sobivat eeskju kirjanusest siiski praegu veel ei leidunud. Seda eeskätt selle tõttu, et eesti tulevase emotsioonituvastaja eesmärk ei ole mingi üldotstarbeline teksti meeoleu või valentsi analüüs, vaid et ta peab suutma määrata konkreetse lause emotsiooni ja seda üsna täpselt piiritletud kategooriate terminites. Teiseks on selge, et mis puudutab emotsioonidega seotud sõnade ja võimalike konstruktsioonide loendeid, siis tuleb need eesti keele jaoks algusest peale ise luua vastavalt sellele sõnavarale ja lokaalsele grammatikale, mida eestikeelsetes emotsioonilausestes kasutatakse.

Läbi sai kaalutud ka inimeste poolt piisava üksmeelega ära tuntud emotsioonilause mitmete formaalsete ja sisulisemate tunnuste kasutatavus automaatses emotsioonituvastuses. Artikli lisana esitatakse veel kord kokkuvõtlikult need tendentsid, mis andmekaeva tulemusena õnnestus empiirilise materjali analüüsil välja selgitada. Osa tendentsidest väljendavad positiivseid seoseid nähtuste vahel (kui A, siis tõenäoliselt B), osa negatiivseid (kui A, siis pigem mitte B). Ükski neist tendentsidest ei ole siiski reegli kindlusega, nad saavad anda üksnes osalise panuse tulevase AET töösse. Ilmselt peabki AET hakkama rehkendama mitmete võimalike lauses leitavate mingi emotsiooni tõenäosust suurendavate ja vähendavate tunnuste koosmõju.

AET loomise seisukohast on kindlasti tähtis ka kasutatavate kriteeriumide ressursinõudlikkus. Mida sisulisem kriteerium, seda ressursinõudlikum. Suhteliselt lihtne on panna arvuti vaatama lauselõpumärke või arvutama lause pikkust sõnedes või tähekohtades. Tunduvalt töömahukam seevastu on morfoloogiline analüüs (koos ühestamisega) ning leksika võrdlemine emotsioonide suhtes relevantsete sõnade sõnastikuga, mida praeguse seisuga eesti keele jaoks ei ole veel olemaski. Eesti keele afektisõnastik tuleb selle keeletehnoloogilise rakenduse tarbeks alles spetsiaalselt luua. Ideed selle kohta, kust peale hakata ja kuidas kasutada olemasolevaid eesti keele ressursse ning tehtud uuringuid, on olemas, kuid selle kõige järeleproovimine ning kirjeldamine nõuab palju tööd ning on kahtlemata juba uue artikli teema. AET võiks rakendada lausest otsitavate tunnuste hierarhiat: esmalt otsida “odava-maid” tunnuseid ja alles siis, kui neist ei piisa, vaagida ressursinõudlike tunnuste olemasolu või puudumist.

Nagu välja tuli, kujutas emotsiooni määramine ilma kontekstita esitatud kirjaliku lause puhul endast küllaltki rasket ülesannet ka katses osalenud inimestele, kelle arvamused kippusid lahknema. Autor ei ole praeguse seisuga eriti optimistlik selle suhtes, et AET tuleks inimesest “targem”. Ilmselt nii mõnegi lause puhul ongi võimalik mitmene interpretatsioon ja päris paljude tekstis ette tulevate lausete kohta ei piisa üldse tõenäosuslikke tunnuseid mingi konkreetse emotsiooni määramiseks. On täitsa võimalik, et üksikus kirjalikus tekstilauses (eriti kui see on kontekstisidus) ei sisaldugi taustaks olevate emotsioonide kohta piisavalt märgilist infot.

Simone Schnall (2005) väidab tuginedes Sperberi ja Wilsoni relevantsus-teooriale (1995), et kommunikatsioonis ei toimigi põhimõtteliselt nn “kodeerimis-mudel”, vaid et oluline on jagada ühist reaalsust ning teha järeldusi suhtleja kavatsuste kohta esitatud asitõendite põhjal. Kui seda seisukohta aktsepteerida, siis ei peakski AET otsima tekstist niivõrd konkreetsele emotsioonile omast koodi kui olema võimeline tegema tõenäosuslikke järeldusi asitõendite põhjal, milleks on igas konkreetses tuvastamisele tulevas lauses sisalduv info. “Empaatiline arvuti” peaks mingil määral jagama ka eestlaste ühist reaalsust (näiteks sõnastiku näol, mis talletab “üldiselt positiivsed” ja “üldiselt negatiivsed” asjad, omadused jne).

Kokkuvõttes ei ole arvutile empaatia õpetamise ülesanne võib-olla siiski päris lootusetu, kuivõrd arvuti ei pea ju tundeid päriselt jagama, vaid looma lihtsalt pii-savalt usutava mulje võimest olla emotsionaalselt ettelooetava teksti autori “nahas”. Kõnesüntees, mille vajadustest on AET ellu kutsutud, võiks seada eesmärgiks mitte eksida lause hinnangulise valentsi ja intensiivsuse osas ja jätta emotsionaalse tõlgenduse tegemise (konkreetsete kategooriate mõttes) ja/või mittetegemise rahumeeli kuulaja teha. Kui kõnesünteesi vajadustest lähtuv emotsioonituvastaja on kord loodud, võib selle potentsiaalne rakendus olla tulevikus laiemgi. Võimalikuks saab hinnata mistahes tekstide emotsiooni ja see võib kasulikuks osutada näiteks elektroonse kirjavahetuse, interneti e-postituste, kommentaaride jm tekstide automaatsel käsitlemisel.

Viidatud kirjandus

Al, Mostafa; Shaikh, Masum; Prendinger, Helmut; Mitsuru, Ishizuka 2007. Assessing senti-ment of text by semantic dependency and contextual valence analysis. – A. Paiva, R. Prada, R. W. Picard (Eds.). *Affective Computing and Intelligent Interaction*. Second International Conference, ACII 2007 Lisbon, Portugal, September 12-14, 2007. Pro-

- ceedings. *Lecture Notes in Computer Science*, 4738. Berlin, Heidelberg: Springer-Verlag, 191–202. doi:10.1007/978-3-540-74889-2_18
- Altrov, Rene 2008. Eesti emotsionaalse kõne korpus: teoreetilised toetuspunktid. – *Keel ja Kirjandus*, 4, 261–272.
- Altrov, Rene; Pajupuu, Hille 2008. The Estonian emotional speech corpus: Release 1. – *The Third Baltic Conference on Human Language Technologies*. Vilnius: Vytauto Didžiojo Universitetas; Lietuvių kalbos institutas, 9–15.
- Baron-Cohen, Simon; Wheelwright, Sally 2004. The empathy quotient: An investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. – *Journal of Autism and Developmental Disorders*, 34 (2), 163–175. doi:10.1023/B:JADD.0000022607.19833.00
- Bednarek, Monika 2008. *Emotion Talk across Corpora*. Houndmills, New York: Palgrave Macmillan.
- Bednarek, Monika 2009. Dimensions of evaluation. Cognitive and linguistic perspectives. – *Pragmatics and Cognition*, 17 (1), 146–175.
- Beljanin, Valeri 2000. *Osnovõ psiholingvističeskoj diagnostiki: modeli mira v literature*. Moskva: Trivola.
- Biber, Douglas; Finegan, Edward 1989. Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. – *Text*, 9 (1), 93–124. doi:10.1515/text.1.1989.9.1.93
- Bracewell, David B.; Minato, Junko; Ren, Fuji; Kuroiwa, Shingo 2006. Determining the emotion of news articles. – D-S. Huang, K. Li, G. W. Irwin (Eds.). *Computational Intelligence. International Conference on Intelligent Computing, ICIC 2006, Kunming, China, August 16-19, 2006. Proceedings, Part II, Vol. 2*. Berlin, Heidelberg, New York: Springer, 918–923. doi:10.1007/11816171_115
- Buck, Ross 1989. On the semiotics of emotion: Spontaneous and symbolic communication. – Walter A. Koch (Ed.). *For a Semiotics of Emotion*. Bochum: Studienverlag Dr. Norbert Brockmeyer, 1–33.
- Bühler, Karl 1934. *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Stuttgart: Fischer.
- ConceptNet. <http://conceptnet.media.mit.edu> (07.12.09).
- Damasio, Antonio R. 1999. *The Feeling of What Happens. Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace & Company.
- Darwin, Charles 1998. *The Expression of Emotion in Man and Animals. Introduction and Commentary* by P. Ekman. Chicago: Chicago University Press.
- Downes, William 2000. The language of felt experience: Emotional, evaluative and intuitive. – *Language and Literature*, 9 (2), 99–121. doi:10.1177/096394700000900201
- Eesti emotsionaalse kõne korpus. <http://193.40.113.40:5000/> (07.12.09).
- Foolen, Ad 1997. The expressive function of language: Towards a cognitive semantic approach. – S. Niemeier, R. Dirven (Eds.). *The Language of Emotions. Conceptualization, Expression and Theoretical Foundation*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 15–32.
- FrameNet. <http://framenet.icsi.berkeley.edu/> (07.12.09).
- Goleman, Daniel 2006. *Social Intelligence. The New Science of Human Relationships*. London: Hutchinson.
- Grefenstette, Gregory; Qu, Yan; Evans, David A.; Shanahan, James G. 2006. Validating the coverage of lexical resources for affect analysis and automatically classifying new words along semantic axes. – James G. Shanahan, Yan Qu, Janyce M. Wiebe (Eds.). *Computing Attitude and Affect in Text: Theory and Applications*. Dordrecht: Springer, 93–107. doi:10.1007/1-4020-4102-0_9
- Hacking, Ian 1999. *The Social Construction of What?* Cambridge MA: Harvard University Press.

- Hunston, Susan; Thompson, Geoff 2001. *Evaluation in Text. Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press.
- Hunston, Susan; Sinclair, John 2001. A local grammar of evaluation. – S. Hunston, G. Thompson (Eds.). *Evaluation in Text. Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press, 74–101.
- Iida, Akemi; Campbell, Nick; Higuchi, Fumito; Yasumura, Michiaki 2003. A corpus-based speech synthesis system with emotion. – *Speech Communication*, 40 (1–2), 161–187. doi:10.1016/S0167-6393(02)00081-X
- Kerge, Krista; Pajupuu, Hille; Tamuri, Kairi 2008. Where should TTS-synthesizer pause and breath? – *The Third Baltic Conference on Human Language Technologies*. Vilnius: Vytauto Didžiojo Universitetas; Lietuvių kalbos institutas, 143–149.
- Kövecses, Zoltan 2006. *Language, Mind and Culture. A Practical Introduction*. Oxford: Oxford University Press.
- Laur, Mall 1987. *Eesti kõne emotsionaalseid aspekte*. Käsikirjaline magistritöö Tallinna Pedagoogikakõrgkoolis.
- Liu, Hugo; Lieberman, Henry; Selker, Ted 2003. A model of textual affect sensing using real-world knowledge. – *Proceedings of the 8th International Conference on Intelligent User Interfaces*, Miami, Florida, USA, 125–132.
- Marmolejo-Ramos, Fernando; de Juan, Maria R. E.; Gygas, Pascal; Madden, Carol; Roa, Santiago M. 2009. Reading between the lines. The activation of background knowledge during text comprehension. – *Pragmatics and Cognition*, 17 (1), 77–107.
- Mathieu, Yvette Y. 2006. A computational semantic lexicon of French verbs of emotion. – James G. Shanahan, Yan Qu, Janyce M. Wiebe (Eds.). *Computing Attitude and Affect in Text: Theory and Applications*. Dordrecht: Springer, 109–124. doi:10.1007/1-4020-4102-0_10
- Mihkla, Meelis 2007. *Kõne ajalise struktuuri modelleerimine eestikeelsele tekst-kõne sünteesile – Modelling the temporal structure of speech for the Estonian text-to-speech synthesis*. *Dissertationes linguisticae Universitatis Tartuensis*, 8. Tartu: Tartu Ülikooli Kirjastus.
- Mikita, Valdur 2008. *Metsik lingvistika*. Tallinn: Grenader.
- Obrenovic, Zeljko; Garay, Nestor; López, Juan Miguel; Fajardo, Inmaculada; Cearreta, Idoia 2005. An ontology for description of emotional cues. – Jianhua Tao, Rosalind W. Picard, Tieniu Tan (Eds.). *Affective Computing and Intelligent Interaction*. First International Conference, ACII 2005, Beijing, China, October 22-24, 2005. *Proceedings. Lecture Notes in Computer Science*, 3784. Berlin Heidelberg: Springer, 505–512. doi:10.1007/11573548_65
- Ochs, Elinor 1988. *Culture and Language Development: Language Acquisition and Language Socialisation in a Samoan Village*. Cambridge: Cambridge University Press.
- Osherenko, Alexander 2008. Towards semantic affect sensing in sentences. – *Proceedings of AISB 2008*, 1–4.
- Osherenko, Alexander; André, Elisabeth 2007. Lexical affect sensing: Are affect dictionaries necessary to analyze affect? – A. Paiva, R. Prada, R. W. Picard (Eds.). *Affective Computing and Intelligent Interaction. Second International Conference, ACII 2007 Lisbon, Portugal, September 12-14, 2007. Proceedings. Lecture Notes in Computer Science*, 4738. Berlin, Heidelberg: Springer-Verlag, 230–241. doi:10.1007/978-3-540-74889-2_21
- Picard, Rosalind W. 1997. *Affective Computing*. Mass.: The MIT Press.
- Reilly, Judy; Seibert, Laura 2003. Language and emotion. – R. J. Davidson, K. R. Scherer, H. H. Goldsmith (Eds.). *Handbook of Affective Sciences*. New York: Academic Press, 535–559.
- Saareste, Andrus 1927. *Tundmused tegurina keele arengus*. – *Eesti Keel*, 7, 161–184.

- Scherer, Klaus R.; Johnstone, Tom; Klasmeyer, Gundrun 2008. Vocal expression of emotion. – M. Lewis, J. M. Haviland-Jones, L. Feldman Barrett (Eds.). *Handbook of Emotions*. Third ed. New York: Guilford Press, 433–456.
- Schnall, Simone 2005. The pragmatics of emotion language. – *Psychological Inquiry*, 16, 28–31.
- Shanahan, James G.; Qu, Yan; Wiebe, Janyce M. (Eds.) 2006. *Computing Attitude and Affect in Text: Theory and Applications*. Dordrecht: Springer. doi:10.1007/1-4020-4102-0
- Sperber, Deidre; Wilson, David 1995. *Relevance: Communication and Cognition*. 2nd edition. Oxford: Blackwell.
- Zhang, Li 2009. Exploration of affect sensing from speech and metaphorical text. – M. Chang, Rita Kuo, Kinshuk, Gwo-Dong Chen, Michitaka Hirose (Eds.). *Learning by Playing. Game-based Education System Design and Development*. 4th International Conference on E-Learning and Games, Edutainment 2009, Banff, Canada, August 9-11, 2009. *Proceedings. Lecture Notes in Computer Science*, 5670. Berlin/Heidelberg: Springer-Verlag, 251–262. doi:10.1007/978-3-642-03364-3_31
- VAAL 2000. <http://www.vaal.ru/proekt/vaal2000.php> (07.12.09).
- Vainik, Ene 2002. Millest on tehtud eestlaste emotsioonisõnavara? – *Keel ja Kirjandus*, 8, 537–553.
- Vainik, Ene; Velt, Anneli 2006. Viha metafoorid ja kognitiivne mudel eesti keeles. – *Keel ja Kirjandus*, 2, 104–122.
- Ververidis, Dimitrios; Kotropoulos, Constantine 2006. Emotional speech recognition: Resources, features, and methods. – *Speech Communication*, 48 (9), 1162–1181. doi:10.1016/j.specom.2006.04.003
- Veski, Asta 1982. *Eesti keele grammatika II*. Tartu.
- Voll, Piret 2002. Tähenduse konnotatiivne aspekt ja selle kajastamine ükskeelses sõnaraamatus. Käsikirjaline magistritöö Tartu Ülikoolis.
- Watzlawick, Paul; Bavelas, Janet B.; Jackson, Don 1967. *Pragmatics of Human Communication. A Study of Interactional Patterns, Pathologies, and Paradoxes*. New York: W. W. Norton & Company.
- Whissell, Cynthia M. 1989. The dictionary of affect in language. – R. Plutchik, H. Kellerman (Eds.). *Emotion: Theory, Research, and Experience*. New York: Academic Press, 113–131.
- Wierzbicka, A. 1999. *Emotions across Languages and Cultures. Diversity and Universals*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511521256
- Wilce, James M. 2009. *Language and Emotion*. New York: Cambridge University Press.
- Wilson, Theresa; Wiebe, Janyce; Hoffmann, Paul 2005. Recognizing contextual polarity in phrase-level sentiment analysis. – *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, Vancouver, British Columbia, Canada, 347–354.
- WordNet. <http://wordnet.princeton.edu/> (07.12.09).
- Wu, Chung-Hsien; Chuang, Ze-Jing; Lin, Yu-Chung 2006. Emotion recognition from text using semantic labels and separable mixture models. – *ACM Transactions on Asian Language Information Processing*, 5 (2), 165–182. doi:10.1145/1165255.1165259
- Yan, Jiajun; Bracewell, David B.; Ren, Fuji; Kuroiwa, Shingo 2008. The creation of a Chinese emotion ontology based on HowNet. – *Engineering Letters*, 16 (1), 166–171.

Ene Vainiku (Eesti Keele Instituut) uurimisobjektiks on eesti keele emotsioonisõnavara, selle struktuur, semantika ja varieerumine. Huvipakkuvateks teemadeks on ka psühholingvistika, kognitiivne semantika ja keele ning emotsioonide interaktsioon laiemalt.
Ene.Vainik@eki.ee

Lisa. Tunnuste ja emotsioonide seostumise tendentsid

Kategooria	Tunnus (selgitus)	Emotsioon
lõpumärgistus	(lõpumärk ei ole:) üksik punkt	vaevu neutraalne
	(lõpumärk on:) ? / ! / ?!	pigem viha
	(lõpumärk on:) !	pigem rõõm
	((lõpu)märk on:) ...	pigem kurbus
eitav kõne	(lauses sisaldub:) ei/mitte/pole/polnud/polegi	pigem negatiivne
jaatav kõne	(lauses ei sisaldu:) ei/mitte/pole/polnud/polegi	pigem neutraalne/rõõm
sõnede pikkus tähemärkides	lühikeste sõnede (kuni 4 tähemärki) osakaal lauses suurem kui pikkadel	pigem kurbus
lause esimese sõna liik	adverb	pigem viha vaevu rõõm
	pronoomen (eriti I isik)	pigem kurbus vaevu viha
	adjektiiv	pigem rõõm
	sidesõna	vaevu neutraalne
sõnaliigid läbivalt	adverbe kaks või rohkem	pigem negatiivne
	adverbe üks või vähem	pigem neutraalne/rõõm
	substantiive üks või vähem	pigem kurbus
morfoloogiline marker	-d	pigem viha/rõõm
nn võtmesõna	otsene emotsiooninimetus	vaevu neutraalne
	emotsioonide väljenduse kirjeldus	vaevu neutraalne
nn leksikaalne külgetõmme	negatiivne hinnang	vaevu rõõm
	positiivne hinnang	vaevu viha/kurbus

HOW TO TEACH EMPATHY TO A SPEECH SYNTHESIZER? ON THE POSSIBILITY OF IDENTIFYING EMOTION SOLELY FROM WRITTEN ESTONIAN SENTENCES

Ene Vainik

Institute of the Estonian Language

There is a growing need for more naturalness in synthetic Estonian speech. One of the measures to be taken is to teach the computer to apply differentiated acoustic registers according to the different emotions (joy, sadness, anger) or neutrality of the text. The present article deals with how to detect the quality of emotion vs. neutrality relying solely on the information present in the Estonian written text.

The first part of the article provides a theoretical overview of how, in principle, emotions can be expressed in speech. Besides the ideational/referential mode (i.e. the use of a literal emotion term, like *sadness*) there is a variety of linguistic cues of expressiveness to be detected. As a result of an overview of the literature on sentiment analysis and affective computing it is stated that the approaches vary in many respects. First, there is a variation in what is considered to be the unit of analysis in the first place (text, passage, sentence, or clause). Secondly, what exactly is being looked for: particular emotions in terms of specific categories (e.g. fear, joy, sadness, anger etc.) or more abstract dimensions (e.g. valence, intensity). The third aspect to be noticed is whether the authorial or also non-authorial affect is taken into account. Most of the approaches exploit lexical features and compare the items present in the text with an affective lexicon.

In the empirical part of the study the results of a statistical analysis of Estonian sentences (altogether 361) are presented. The sentences were retrieved from the Estonian Emotional Speech Corpus, where 55 test subjects had evaluated their emotion (joy, sadness, anger) or neutrality. To put the results very briefly is to say that identifying the emotion of a sentence entirely out of context must have been a pretty demanding task. There was not much congruence in the evaluations. Looking for the features which could work for the computer as cues for automatic emotion detection ended up in a list of probabilistic tendencies. Features such as punctuation, length of sentence (in characters), part of speech of the first word, negation etc. were used. Also the main lexical means and some strategies of attributing emotion according to the evaluative value of the sentence were described in some detail.

The article ends with a conclusion that enabling a speech synthesizer to imitate human emotions can be compared with modeling human empathy. This is not an easy task. None of the systems created for other languages can be simply adapted for Estonian. This is firstly because emotions and their expressions are culture-specific to some extent and, secondly, because the Emotion Detector should rely on an Estonian affect dictionary, which does not exist, yet. Consequently, there is still a lot to do in the field.

Keywords: emotions, written text, speech synthesis, sentiment analysis, affective computing, Estonian

CONCEPTUAL CO-PRESENCE OF MOTION AND EMOTION IN THE ESTONIAN TERMS OF PERSONALITY

Ene Vainik, Toomas Kirt, Heili Orav

Abstract. The purpose of the study was to find out whether the conceptual connection of emotion and motion is holding in the domain of personality traits. In a quantitative study 40 Estonian terms of personality traits were investigated and a conclusion was driven that, indeed, the qualities of emotionality and motion are perceived as co-present characteristics. In further data analysis some visualized measures were applied in order to get further insights into the hidden structure of the data. The self-organizing map (SOM) analysis revealed an additional dimension of axiological evaluations present in the semantics and the SOM meta-analysis technique revealed groups of near synonymous words as well as gave overview of the more general structure common in the two data sets. The latter was tentatively explained by the person's habitual level of activation.*

Keywords: motion, emotions, terms of personality traits, semantics, self-organizing map, Estonian

1. Introduction

We consider the vocabulary of personality traits as a potential meeting point of emotion and motion in the human conceptual representation of other peoples' behaviour. It has been claimed that conceptual knowledge influences our experience of other people (Feldman Barrett 2006: 28). The general idea is that our knowledge of people and situations automatically and effortlessly shapes what we “see” people doing and gives rise to our explanations for that behaviour. One aspect of such conventional knowledge is the idea that a person's emotion and physical mobility are somehow linked. Several observations back up our assumption that those three phenomena – motion, emotions and personality traits – could be conceptually linked.

* This study was supported by grant no. 7149 of the Estonian Science Foundation and grant no. MJD22 financed with the support of the European Social Fund.

First, there is a conceptual connection between emotions and motion. Both psychologists and some cognitive linguists share an agreement that emotions are linked to a tendency of motion either towards or away from something (see Gibbs 2006: 243 and his references).¹ Motion as a more embodied and experience related cognitive domain serves as a source to conceptualize emotional phenomena metaphorically. There are such conceptual metaphors mentioned in the literature as A CAUSED CHANGE OF STATE (EMOTION) IS MOTION CAUSED BY A FORCE; CONTROL OVER AN EMOTIONAL ACT IS CONTROL OVER MOTION; EMOTIONAL RESPONSES ARE OTHER-PROPELLED MOTIONS (Kövecses 2000: 58–59).

Second, there is a connection between tendencies to move and certain personality traits. Motivational psychology differentiates between two important motivation styles which it formulates directly as motion tendencies – a tendency to approach the goal and a tendency to avoid (withdraw from) the anti-goal (Carver et al. 2000). Both emotional disposition and motivation are identified as significant components of the structure of personality (Nõlvak, Valk 2003).

Third, there is an area of intersection between vocabularies of personality traits and emotions. Studies into psychology and linguistics alike have shown that concerning the content, vocabularies describing personal qualities and emotions are partly overlapping (Plutchik 1980, Szelid, Geeraerts 2008, Vainik, Orav 2005).

Our research question is whether the conceptual interrelatedness of motion and emotion also holds in the domain of personality traits. Our intuition is that it does – i.e. we set a hypothesis that emotions and motion are co-present in the conceptualizations of personality traits. We have already approached this research question in a previous qualitative study that enabled to find out the main ways in which the domain of physical motion is exploited in order to conceptualize emotion-related aspects of character terms (Vainik, Orav 2009). In that previous study we applied the framework of cognitive metaphor theory (Lakoff, Johnson 1980, 1999) and systematic detection of basic embodied image-schemas (Johnson 1987, Peña Cervel 2003). However, this previous analysis was based mostly on the linguist's insights into the lexicon and its underlying mechanisms and did not say much about the co-presence of emotionality and mobility in the character terms for an ordinary speaker of Estonian.

Our theoretical framework of concepts is based on the Peter Gärdenfors' (2000) theory of conceptual spaces which distinguishes three non-competing levels of representations. In this model the most abstract level is the symbolic level on which an observation is described by means of some symbol system (e.g., language). The second level is the conceptual level on which observations are located as points or an area in the conceptual space. The least abstract level is the sub-conceptual level on which the observations are characterized by inputs from sensory receptors which form the quality dimensions of the conceptual space. The quality dimensions thus form a (perceptual) basis for conceptual representation which in turn can be communicated by means of the respective linguistic units.

When applied to the field of personality traits, this would mean that on the most abstract level we have terms of personality traits, on the conceptual level we have concepts of personality traits and on the sub-conceptual level we have discrete (e.g., perceptual) qualities of personality. While asking people to assess some qualities

¹ This is implied even by the fact that the English terminology – the terms *emotion* and *motion* – are derived from the same Latin root *movere*, 'to move'. The term *motivation* originates from the same Latin root (The Concise Oxford Dictionary).

of a person characterized by a specific term we are using an access on the most abstract level (a term of personality trait). At the same time we are addressing the perceptual qualities of the assessed person (the subconceptual level). As a result the intermediate, conceptual level of representation is revealed which would otherwise not be accessible as such. While a personality trait term evokes co-presence of some qualities (e.g., emotionality and mobility) in the inquiry addressed to the subconceptual perception it would most probably also mean the co-presence of those qualities on the conceptual level i.e. their co-conceptualization.

In addition, the theory offers a geometrical model of assessing concept similarity based on spatial relations. As Gärdenfors (2000) has suggested, there is an analogy between the conceptual spaces and the self-organizing map (Kohonen 2000). During the self-organizing process the points in high-dimensional space are mapped onto a two-dimensional output map that can be identified as a conceptual space. The self-organizing map is one way of modeling how the geometric structure of concepts can be created.

In the present study we focus on detection whether the two qualities of personality – emotionality and mobility – are truly perceived and conceptualized as co-present in the case of character traits' terms. In order to do so we used character terms as verbal stimuli and gathered assessments of emotionality and mobility given by laypersons. We subjected the quantitative data to statistical analysis and in the next phase of the study we applied the methods of data reduction, which give the results in the form of visual layouts. Both traditional statistical analysis and visualized measures taken together revealed that, indeed, the qualities of emotionality and motion are perceived as co-present characteristics. The visualized measures gave some further insights into the data, too. The self-organizing map (SOM) analysis revealed an additional dimension of axiological evaluations present in the semantics and the SOM meta-analysis technique revealed groups of nearly synonymous words as well as gave an overview of the more general structure common in the two sets of data. The latter was tentatively explained by the person's habitual level of activation.

2. Materials and procedures

2.1. Selection of the terms

The initial data pool of linguistic expressions for our analysis was gathered by Orav (2006) using the field method of tasks of free listing (Corbett, Davies 1997). In a multiple-test series, the participants ($N = 100$; age 14–90; the number of men and women was balanced) had to list words in Estonian that describe a person's internal characteristics.² A total number of about 1270 words were elicited (Orav 2006).

The next step was to select the terms which to address to in a quantitative study. The number of the verbal stimuli could not exceed 50 because the purpose was to use ordinary speakers of Estonian on a voluntary basis and they had to fill in two independent questionnaires. Therefore the number of terms under investigation was limited to 40 and in addition 10 controls were selected.

² The test subjects were illustratively explained the difference between words describing external characteristics and internal qualities (i.e. personality). External traits are connected mostly with people's looks (as *thin*, *blonde*, *big*), but internal are those characteristics which we can register only by observing someone's behaviour.

As the next step, two of the authors (Heili Orav and Ene Vainik) assessed the total list of words independently, relying on their intuition and introspection.³ Although the evaluation criteria were left open deliberately, the result was a great concurrency in terms of words associated with motion tendencies (87%). Even though the total number of words making reference to emotions was nearly twice as high, there was less concurrence in the independent assessments (70%). While selecting the terms both researchers had mutually independently used two criteria as a basis for introspect decisions: a) the meaning of the word as a whole suggested motion or emotion (e.g., *kiire* ‘fast’, *emotsionaalne* ‘emotional’), or b) the word or phrase was clearly derived from some certain motion verb (e.g. *heitlik* ‘fickle, lit. throw+ADJ’) or emotion term, (e.g. *rõõmsameelne* ‘joyful, lit. joy-SG+GEN+mind+ADJ’).

For further testing we selected those character terms that both authors had independently assessed as motion-related and at least one of the authors had assessed as emotion-related. The controls were selected so that both authors agreed that 5 of them had to do only with motion tendencies (*sportlik* ‘sporty’, *kiire* ‘fast’, *laisk* ‘lazy’, *virik* ‘diligent’, *edasipüüdlik* ‘ambitious, lit. forward+pursue+ADJ’) and 5 of them only with emotions with no reference to motion whatsoever (*kade* ‘envious’, *õel* ‘mean’, *rõõmsameelne* ‘joyous’, *tundlik* ‘sensitive’, *närviline* ‘nervous’).

2.2. The experiment

The data was gathered in April 2008 during two weeks via electronic media (by means of e-formular <http://www.eformular.com>). The experiment consisted of two questionnaires directly addressing the two qualities under investigation (see Appendix 1). The questionnaires had to be completed in serial order—at first the questionnaire about emotionality and next about mobility. There was one question that occurred in both questionnaires: “*To what extent is a person characterised as X (a specific character term) also emotional/mobile?*” The questions were followed by the 50 words (40 + 10). At the end of both questionnaires there was a place to fill in the data about the age and gender.

Every word was evaluated twice by every person on a 5 point *Likert*-scale (very little, somewhat, fairly, above average, very much emotional/mobile). A possibility was also left open to skip a word in the case the person would think that the given quality (emotionality/mobility) had nothing to do with the specific presented character term. This possibility was used very little and unsystematically.

2.3. Participants

The inquiry was carried out by electronic media and it was out of our reach to direct the selection of participants in terms of age, gender or educational background. The appeal to fill in those two questionnaires was spread out in the Internet through a network of family members and friends and their contacts. Apparently, it was filled in by persons who were cooperative enough and happened to have time. It took approximately 45 minutes altogether to fill in both questionnaires.

³ For information about introspection as a method in linguistics (see e.g. Huomo 2008) and cross-references. Assessment of word lists by two independent experts in order to find words suitable for some semantic category has been widely used in psychology (see e.g. Aavik, Allik 2002).

The eventual number of participants was $N = 161$ ($F = 142$, $M = 19$), age averaged at 32,8 ($SD = 12,0$).

2.4. Handling of the data

The verbal assessments on *Likert-scale* were transformed into numerical values ranging down from 5 (*very much*) to 1 (*very little*). It was carefully detected whether the rows of the two tables matched the persons (relying on the data about age and gender and computer address given by e-formular).

In the following analysis the data was first analyzed statistically to find out the average levels and standard deviations of emotionality and mobility ratings associated with our stimulus terms. In order to detect systematic intercorrelations between ratings of emotionality and mobility the analysis of bivariate correlations was carried out (two-tailed T-test). As the next step the visual methods of data-analysis were applied in order to find out some further structure possibly hidden in the distribution of the numeric data.

2.5. Self-organizing maps

The self-organizing map (SOM) (Kohonen 2000) is a widely used method to visualize multidimensional data. The SOM can be used to project multidimensional data into a two-dimensional map while preserving proximity relationships as well as possible. The data items located close to each other in multidimensional space will appear as nearby units in a two-dimensional map.

The learning algorithm of the SOM has a few basic steps which are reiterated a number of times. Prior to learning the output map is initialized randomly. Thereafter one random input unit is chosen and compared with all output nodes on the map to find the closest unit on the output grid. Next, this best match or winning node and its neighborhood are changed to become closer to the input unit, using the following equation:

$$\text{Weight(New)} = \text{Weight(Old)} + \text{Learning rate} \times \text{Neighbourhood function} \times (\text{Input value} - \text{Weight(Old)})$$

During the learning process the learning rate and the neighbourhood function shrinks. As a result of this process the output becomes ordered and all the output vectors are valued so that the total distance from the input vectors is minimized.

The distance between each pair of map units is represented on a topological map called the Unified Distance Matrix (U-matrix) by their location and also with colour coding. A bright shade corresponds to a small distance between two map units and a dark shade represents a bigger difference between the map units. The points on the output map that lie in the light area belong to the same group or cluster, while the dark area shows the borders between the clusters.

2.6. Method of comparing the self-organising maps

As the SOM represents data on a two-dimensional topological map another question arises: how to measure similarity of two or more maps describing similar data? This study proposes a methodology of SOM meta-analysis (Kirt, Vainik, Võhandu 2007) that uses the local neighbourhood as the basis to measure the similarity between maps. The method expects the neighbourhood relations to remain stable even when the overall orientation of the map changes. The proposed methodology to measure similarity between the self-organizing maps consists of two main stages.

First, to analyze general organization the resulting map is visually examined and clusters and their borders are identified, in addition the general orientation and locations of data items are also identified. Thereafter the matrix of neighbourhood relations is formed. Neighbourhood assessment is based on the location of the best matching units (BMU – a point on the map that is the closest to the input data vector) on the self-organizing map. Two data items are neighbours if they are marked to locate on the same node or in the neighbouring nodes depending on the neighbourhood range. The second part of the similarity measurement consists of establishing how much the two neighbouring matrixes are identical to what is a maximum isomorphic subset. Here we can perform a new meta-level analysis and reorder and visualize the isomorphic sub-graph to see commonly shared information between two maps. For output the Graphviz⁴ software has been used.

The next stage of similarity analysis is the calculation and assessment of similarity coefficients. The similarity measurements coefficients have typically values between 0 and 1. Value 1 indicates that the two objects are completely similar and value 0 indicates that the objects are not similar at all. Two similarity coefficients – Simple Matching Coefficient (SMC) and Jaccard Coefficient (J) (see Tan, Steinbach, Kumar 2005) – have been used. The SMC rates positive and negative similarity equally and can be used if positive and negative values have equal weight. Jaccard Coefficient (J) is used if the negative and positive matches have different weights (are asymmetric). Jaccard Coefficient ignores negative matches and can be used if the variables have many 0 values.

3. Results

The results in terms of mean values, standard deviations and correlations between evaluations given to the same word in the two tasks are given in Appendix 2. Observations of the mean values of assessed emotionality and mobility showed that there was covariation in the rates: one part of the terms rated high on both emotionality and mobility and the other part rather low on both. There were no terms rating high on one scale and low on the other. The highest ratings of both emotionality and mobility were given to terms *hüperaktiivne* ‘hyperactive’, *pidurdamatu* ‘rampant’, *pöörane* ‘frenetic’ and the lower ratings on both qualities to *apaatne* ‘apathetic’, *flegmaatiline* ‘phlegmatic’, *loid* ‘listless’, *passiivne* ‘passive’, *tagasihoidlik* ‘modest’, *uimane* ‘sluggish’. There are, of course, variations in the levels of emotionality and mobility co-present in the meanings of the words. For

example in the meanings of *aktiivne* ‘active’, *elav* ‘lively’, *energiline* ‘energetic’, *püsimate* ‘restless’, *reibas* ‘cheery’, the rate of mobility prevailed while in the case of *ründav* ‘offensive’, *temperamentne* ‘spirited’, *äkiline* ‘quick-tempered’, the aspect of emotionality prevailed.

The two-tailed T-test showed that for 45% of the words there were significant positive correlations ($0,46 > r > 0,3$; $p < 0,0001$) between the given rates of emotionality and mobility, on 32% of the words there were a little less significant and weaker positive correlations ($0,27 > r > 0,18$; $p < 0,001$) and in 23% of the cases there was still significant ($p < 0,05$) positive correlation $0,17 > r > 0,09$.

A conclusion can be drawn that in the case of studied 40 Estonian terms both emotionality and mobility were either present or non-present in their meanings in a correlated way. One can say that these two characteristics were co-conceptualized. Thus, the results of the statistic analysis are in accordance with our intuition that mobility and emotionality are perceived and conceptualized as co-present human qualities in the meanings of Estonian character terms.

As an addition to the test-words 10 terms of personality trait were added to the questionnaires as controls that were supposed to have to do only with one of the two characteristics. Surprisingly, the control words shared the similar pattern of significant positive correlations (see Appendix 3). Only in case of two of the controls *tundlik* ‘sensitive’ and *rõõmsameelne* ‘joyous’ the correlations were too weak $r < 0,06$ and nonsignificant. There are two explanations: a) the selection of control words failed; b) in the case of any term of personality trait it is possible to assess the presence or absence of both emotionality and mobility as compared to some default norm of these qualities. We assume that both characteristics have to do with the person’s habitual level of activation. A person with the habitual level of activation lower than the “norm” is characterized by terms implying low emotionality and low mobility, a person with habitual activation level higher than the “norm” is characterized by terms implying both high emotionality and mobility.

The traditional statistical analysis revealed the significant positive correlation between emotionality and mobility on 90% of the cases. The strength of correlations did not exceed 0,45, though, remaining just moderate on most of the cases. In connection with those observations some further questions arise. Is the moderate positive correlation all about the measurable interrelations of emotion and motion in the understanding of personality traits? How are the terms of personality traits similar or dissimilar in their ability to reflect both the person’s emotionality and mobility? Is there possibly some kind of hidden structure in the data-matrix that does not show up in the traditional statistical analysis? In an attempt to answer those questions we turned to some of the visual measures of data analysis.

3.1. The SOM-s of emotionality and mobility

The next step of our study was to subject the data of assessed emotionality and mobility to the process of self-organization. The respective topological layouts are presented on Figure 1 and Figure 2.⁵ We used the SOM Toolbox (Alhoniemi et al. 2005) for projecting the abstract and numeric properties of data items into a visible and comprehensible code of their relative spatial locations and colour. There are two

⁵ Next to the SOM-s we represent the layouts of concepts in form of Table 1 and Table 2 in order to give the English glosses of the Estonian terms.

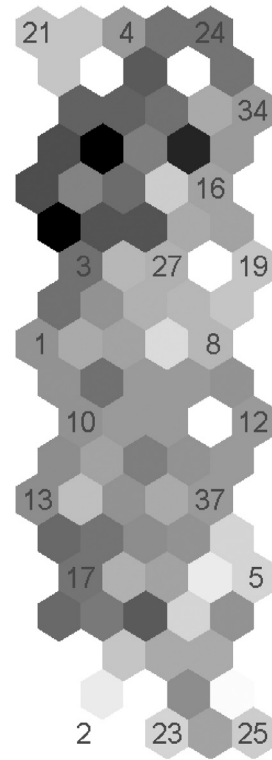
principles to keep in mind while interpreting the results: closeness of the map units and a brighter shade of colour on the map also means similarity of data items.

The very first glance at the Figures tells that they are similar in respect of their extended layouts in one dimension. This is of no surprise because both questionnaires measured the level of just one quality, emotionality or mobility, respectively. The concepts which rated higher on those qualities tend to “shrink” to the bottom of the graph and concepts where those qualities were low have “risen” to the uppermost parts of the graphs. However, the overall shape of the graphs is not purely scalar, apparently due to the significant variations of the individual assessments. The rationale of the emergent horizontal dimension remains a question of further interpretation, though.

Table 1. BMU-s of the SOM of emotionality

21 apathetic phlegmatic sluggish listless passive	4 stable modest	24 submissive
		34 altruistic lickspittle
		16 outgoing
3 mooncalf	27 inconsistent	19 dynamic spunky selfish
1 intransigent		8 lusty for travelling eager
10 blaming		12 cheery active
13 fickle harsh		37 intrusive
17 unstable		5 energetic intensive restless lively
2 rampant offensive quick-tempered frenetic tameless gushy hyperactive	23 spirited	25 excitable spontaneous

Figure 1. SOM of emotionality



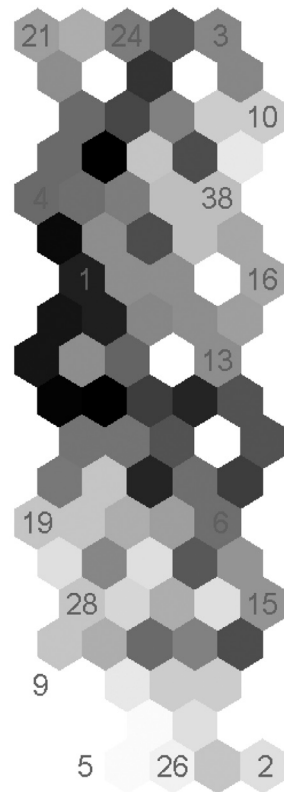
A possibility has been pointed out by Krzeszowski (1997) that the dimension of axiological evaluations is showing up. Interestingly, the location of positive vs. negative concepts is reversed on both graphs and such a distinction is detectable only in the bottom part of the graphs where the emotionality and mobility ratings were the highest. The terms of personality traits referring to very low emotionality and/or mobility are valued negatively, anyway.

Another thing to look at on the maps is the local neighbourhood relations of concepts. Similar data items are located close to each other and different data items are located apart. This principle enables the clusters of concepts to emerge. In both maps there are small clusters in the uppermost left corner that are separated by relatively darker shade of the colour. These are the concepts where the measured qualities occurred at a very low level (*apathetic, listless, passive, sluggish, submissive, modest*). On Figure 2 there is another darker border of clusters noticeable in the middle section of the graph.

Table 2. BMU-s of the SOM of mobility

21 apathetic phlegmatic sluggish listless passive	24 submissive	3 mooncalf modest
		10 blaming altruistic
4 stable		38 lickspittle selfish
1 intransigent		16 outgoing inconsistent
		13 fickle unstable harsh
19 dynamic excitable spontaneous		6 offensive quick-tempered intrusive
28 spunky eager		15 gushy
9 intensive spirited		
5 energetic lusty for travelling cheery active lively	26 hyperactive	2 rampant frenetic tameless restless

Figure 2. SOM of mobility



It can be concluded that in the case of the rated mobility the terms of character traits tend to gather into three clusters: very low mobility, intermediate level or normative level of mobility and a cluster of high mobility (in the bottom). In the case of emotionality ratings the concepts are divided into two: a cluster of very low or missing emotionality and a cluster where the level of emotionality gradually decreases from the very high level to intermediate or normative levels.

The third thing to look at on the maps is the local neighbourhood relations of map units. The closer the visual neighbourhood the higher is the rate of similarity of the data. Concepts located in the same cell of the SOM-map are the most similar to each other (in respect of given emotionality and mobility ratings). In Tables 1 and 2 the coincident terms have been separated by cell borders. We can see groups of seemingly near-synonymous concepts on both maps which reveals that as regards to the levels of emotionality and mobility there are no very sharp distinctions made between those concepts. Distinctions are made between those concepts that are not neighbours and do not fall into the same cluster.

In order to answer the question how similar or dissimilar are the local neighbourhood relations on both maps the method of pure visual tackling may not suffice, though. In the next section we use a metaanalytical method of comparing two (or more) SOM maps that share the range of units but may differ in their numerical descriptions and therefore in the topological layouts and local neighbourhood relations describing their similarity.

3.2. Isomorphic subgraph of the SOM-s of emotionality and mobility

The aim of the meta-analysis of the SOM was mainly to systematically compare the neighbourhood relations present on both maps and identify similarities between two conceptual spaces. There are 1600 potential neighbourhood relations possible between the pairs of our 40 emotion terms. Only a part of the neighbourhood relations is realized on both maps.

To look for the isomorphic subgraph means to identify those neighbourhood relations that are present on both SOM-s. The neighbourhood matrixes of emotionality and mobility were compared and a matrix of common neighbours was found and visualized. The summary of the neighbourhood relations between the two tasks is given in Table 3. The number of relations was measured with a different range of neighbourhood, starting from 1 to 2. An increase in the neighbourhood range also causes an increase in the number of relevant neighbourhood relations. The SMC coefficient is decreasing if the neighbourhood is increasing because the growing range increases the possibility of relations between the words that actually do not belong to the same neighbourhood. The SMC can be used as an indicator of stability. The SMC coefficient can also be used if the samples are exclusive like in the case of the lexical study. Jaccard Coefficient is increasing if the neighbourhood range is widening because it characterizes the positive matches and there is a tendency to have more positive matches if the number of neighbourhood relations increases. In both cases the number of positive matches remains rather low and therefore the value of coefficients is also low.

Table 3. Neighbourhood similarity of the SOM of terms of personality

Neig. Range	Emotionality	Mobility	Similar neig.	Total neig.	SMC	Jaccard coef.
1	206	198	76	328	0,8425	0,23171
2	406	426	216	616	0,75	0,35065

In case the neighbourhood range is provisionally set on 1, several separate fragments of conceptual networks are formed (Figure 3). The general structure of the data does not appear as a connected system. There are groups of near synonyms that share the estimated level of correlated emotionality and mobility. For example the group of low emotionality/mobility: *loid* 'listless', *apaatne* 'apathetic', *passiivne* 'passive', *uimane* 'sluggish', *flegmaatiline* 'phlegmatic', or the group with high level of correlated emotionality and mobility: *pöörane* 'frenetic', *hüperaktiivne* 'hyperactive', *taltsutamatu* 'tameless', *pidurdamatu* 'rampant'. There are also groups that share the intermediate level of correlated emotionality and mobility ratings. In addition, there are some concepts visible that are not connected to other concepts like *püsimate* 'restless', *stabiilne* 'stable', *temperamentne* 'spirited', *hõljuja* 'mooncalf' etc. These polymorphic units show up in different neighbourhood surroundings on our two maps.

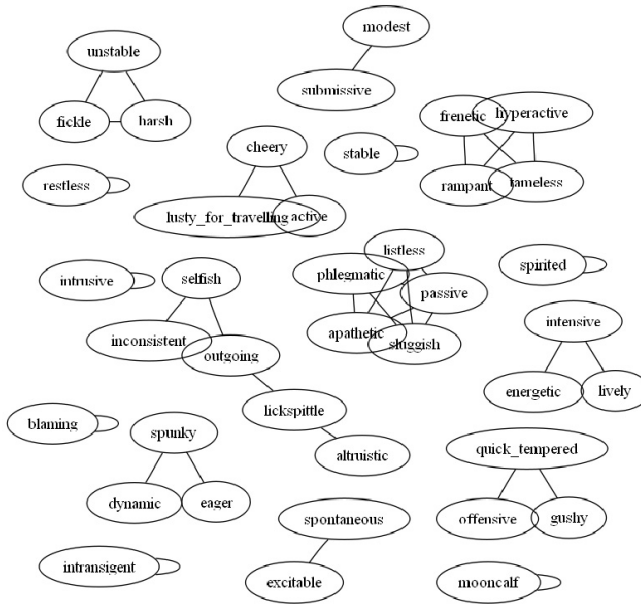


Figure 3. Visualized isomorphic subgraph of terms of personality (emotionality vs. mobility). Neighbourhood range 1

When the measured neighbourhood range is set on 2, the graph becomes connected (Figure 4). An increase in the neighbourhood range also causes an increase in the number of relevant neighbourhood relations. Therefore the graph is not the best means for studying the particular range of neighbours for any single concept. What it is good for is the general overview of the common structure hidden in the two sets of data. It can be speculated that it represents the common structure or a backbone of the data which in our case looks kind of symmetrical. There are at least two almost

exclusive groups of data visible on the left side and the right side that had the weakest neighbourhood relations. These are the subgroups of low and high emotionality/mobility accordingly and only the concept *järeleandmatu* 'intransigent' is situated in the middle having the least to do with both of the groups.

In the section of the results of the traditional statistical analysis we hypothesised about the person's habitual level of activation as a more general property that can cause the levels of emotionality and mobility to covariate. The diagonal dimension of the graph on Figure 4 represents such a more general dimension. Interestingly enough, the term *järeleandmatu* 'intransigent' represents the intermediate or normative level of activation. It is also visible on the graph that the subgroup of high activation is more differentiated than the group of low activation. The terms in the upper part of the high activation subgroup highlight the person's physical behaviour (mobility) which is evaluated positively while the lower part the of the high activation subgroup highlights lack of emotional regulation skills which is mostly evaluated negatively.

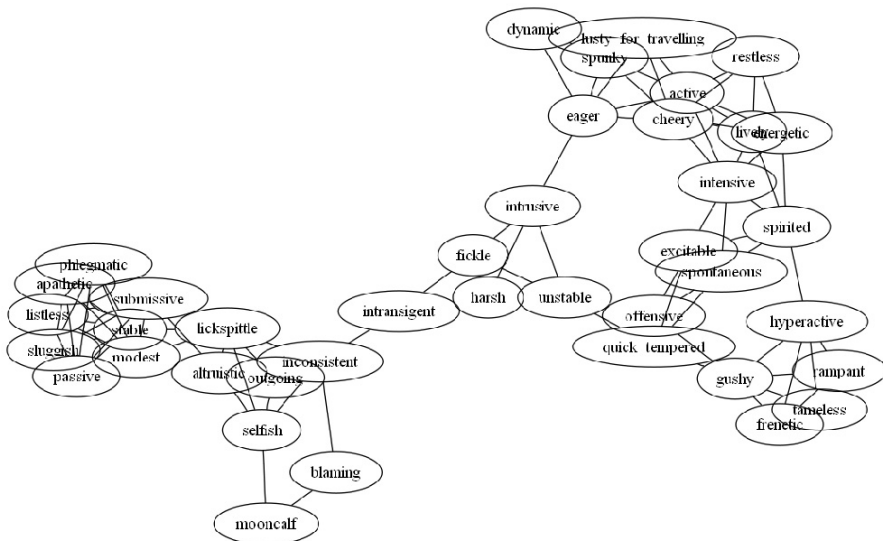


Figure 4. Visualized isomorphic subgraph of terms of personality (emotionality vs. mobility). Neighbourhood range 2

4. Summary and discussion

On the basis of the observations of other authors (Carver et al. 2000, Gibbs 2006, Kövecses 2000, Plutchik 1980) as well as our own previous semantic study (Vainik, Orav 2009) we hypothesized that the categories of emotions and physical motion could be conceptually linked and that this connection could be detected in the semantics of the terms of personality. In order to find out whether this conceptual link means co-presence of emotionality and mobility for Estonian laypersons 40 stimulus terms were selected and two inquiries measuring these qualities were carried out.

As a result we had two sets of data measuring the conceptual co-presence of motion and emotion in the Estonian terms of personality. The traditional statistical analysis gave us the mean values and standard deviations of the given mobility and emotionality ratings; it allowed establishing the intercorrelations and their significance. It occurred that both emotionality and mobility were either present or non-present in the meanings of 40 words in a correlated way. The two-tailed T-test showed significant but rather weak positive correlations between ratings of emotionality and mobility. Our main hypothesis was thus proven – emotionality and mobility occurred as positively correlated throughout almost the whole range of stimulus terms according to the opinions of the Estonian laypersons. None of the correlations was strong, though.

As the traditional statistical analysis was unable to reveal any hidden patterns possibly inherent in the data, the two data sets were further visualized by the method of the Self-Organizing Maps. The SOM-s show kind of extended layouts with the clusters of words of high/low emotionality and mobility respectively. In addition, a dimension of axiological evaluation is recognizable as the horizontal dimension of the layouts. Presence of those dimensions reminds two of the Osgood's three universal factors of *activity*, *potency* and *evaluation* that are claimed to measure the semantic space universally (Osgood, Suci, Tannenbaum 1975). Some other researchers of personality traits (in English) have also ended up in a two-dimensional solution of activation–evaluation circumplex (Plutchik 1980). As regards to the emerging dimension of evaluation it is interesting to pay attention to Orav's (2006) indication that the Estonian personality vocabulary always includes either positive or negative evaluation. The same tendency has been pointed out by Vainik (2004) in the case of emotion vocabulary. Thus, understanding and conceptualising other people's traits and emotions always brings along their approval or disapproval. The importance of implicit evaluation in concept formation has been stressed also by other authors (e.g., Zlatev 2001).

The visualizing techniques enabled to “translate” the abstract and numeric properties of data items into a visible and comprehensible code of their relative spatial locations and colour (SOM) or graphic connection (the meta-analytic tool). Visualization allows getting the first glimpse on the data quickly due to the pictorial superiority effect (Medina 2008⁶: 233). The limitations of visualization techniques are that sometimes some *a priori* knowledge is needed to understand what the visualized structure means and how to interpret the results. The most appropriate visualization technique that reveals the peculiarities of the data in the best possible way is not always predictable beforehand and some experimentation might be needed. It is interesting to note that other researchers of emotion and personality terms have used a method called correspondence analysis to get some visual clusters of the data units under investigation (Szelid, Geeraerts 2008).

In order to compare the two topological layouts systematically we used a similarity measurement technique that enabled us to analyze the two sets simultaneously. The meta-analysis of two SOM-maps showed 1) the clusters of nearly-synonymous words with the most correlated emotionality and mobility ratings, 2) the hidden common structure of the two sets of data. The latter supports our intuition about the person's habitual level of activation as the general property explaining the co-variation of mobility and emotionality. We assume that the structural

⁶ This is a popular book based on peer-reviewed scientific studies. The references for what the author calls a brain rule of visual superiority are retrievable from the site http://www.brainrules.net/pdf/references_vision.pdf (2.03.2010).

elements which were found to be shared between our two SOM-s (Figure 3 and 4) have the greatest chance of belonging to the conceptual level of representation that is assumed to be on the intermediate position in the hierarchy of abstractness (Gärdenfors 2000).

The three levels of symbolic, conceptual and sub-conceptual representation of personality traits are compatible also with the levels proposed by some students of emotion knowledge (Niedentahl, Setterlund, Jones 1994). The theorists of symbol grounding also claim that the particular symbol system is used to carry our conceptual knowledge and that it is grounded into the surrounding world through the sensory-motor interaction (Harnad 1990). Other authors, however, say that such grounding has some difficulties in representing abstract concepts (Barsalou 2008) and that sensory-motor and conceptual processing are based on same system (Barsalou 2003). Therefore the conceptual knowledge is claimed to be embodied and mapped within a sensory-motor system (Gallese, Lakoff 2005).

In the present experimental study into the Estonian layperson's intuitions we were able to detect to what extent the categories of emotionality and mobility are conceptually co-present in the meanings of a set of Estonian terms of personality traits. The finding of their positive correlation is fully in accordance with the insights of the cognitive psychologists (see Gibbs 2006: 243) as well as cognitive semanticists (see Kövecses 2000: 58–59). Understanding the conceptual components of character trait terms is important for the general theory of person perception. In a sense, person perception is the study of the theory of mind – how the categorization process allows people to attribute mental states to identify, explain, and predict behaviour (Feldman Barrett 2006: 28).

Abbreviations

A.SBST	agentive substantive
ADJ	adjective
CAR.ADJ	caritive adjective
GEN	genitive
INF	infinitive
PRS.PTCP	present participle
SG	singular

References

- Aavik, Toivo; Allik, Jüri 2002. The structure of Estonian personal values: A lexical approach. – *European Journal of Personality*, 16, 221–235. doi:10.1002/per.439
- Alhoniemi, Esa; Himberg, Johan; Parhankangas, Juha; Vesanto, Juha 2005. SOM Toolbox (Version 2.0). [Computer software and manual.] <http://www.cis.hut.fi/projects/somtoolbox/> (11.11.2005).
- Barsalou, Lawrence W. 2003. Situated simulation in the human conceptual system. – *Language and Cognitive Process*, 18 (5–6), 513–562. doi:10.1080/01690960344000026
- Barsalou, Lawrence W. 2008. Grounded cognition. – *Annual Review of Psychology*, 59, 617–645. doi:10.1146/annurev.psych.59.103006.093639
- Carver, Charles S.; Sutton, Steven K.; Scheier, Michael F. 2000. Action, emotion, and personality: Emerging conceptual integration. – *Personality and Social Psychology Bulletin*, 26 (6), 741–751. doi:10.1177/0146167200268008

- Corbett, Greville G.; Davies, Ian R. L. 1997. Establishing basic color terms: Measures and techniques. – C. L. Hardin, L. Maffi (Eds.). *Color Categories in Thought and Language*. Cambridge: Cambridge University Press, 197–223.
- Feldman Barrett, Lisa 2006. Solving the emotion paradox: Categorization and the experience of emotion. – *Personality and Social Psychology Review*, 10 (1), 20–46.
- Gallese, Vittorio; Lakoff, George 2005. The brain's concepts: The role of the sensory-motor system in conceptual knowledge. – *Cognitive Neuropsychology*, 22, 455–479. doi:10.1080/02643290442000310
- Gibbs, Raymond W. Jr. 2006. *Embodiment and Cognitive Science*. Cambridge: Cambridge University Press.
- Gärdenfors, Peter 2000. *Conceptual Spaces. The Geometry of Thought*. London: The MIT Press.
- Harnad, Stevan 1990. The symbol grounding problem. – *Physica D*, 42, 335–346. doi:10.1016/0167-2789(90)90087-6
- Huumo, Tuomas 2008. Joko lingvistin nojatuoli joutaisi kaatopaikalle? Introspektio-lingvistiikan asemasta kognitiivisessa kielitiedessä. – *Emakeele Seltsi aastaraamat*, 53 (2007), 163–180.
- Johnson, Mark 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago, IL: University of Chicago Press.
- Kirt, Toomas; Vainik, Ene; Vöhandu, Leo 2007. A method for comparing self-organizing maps: Case studies of banking and linguistic data. – Y. Ioannidis, B. Novikov, B. Rachev (Eds.). *Proceedings of Eleventh East-European Conference on Advances in Databases and Information Systems*. Varna, Bulgaria: Technical University of Varna, 107–115.
- Krzyszowski, Tomasz P. 1997. *Angels and Devils in Hell. Elements of Axiology in semantics*. Warszawa: Wydaietwo Energeia.
- Kohonen, Teuvo 2000. *Self-organizing Maps*. 3rd ed. Springer Series in Information Sciences Series, 30. Berlin: Springer.
- Kövecses, Zoltan 2000. *Metaphor and Emotion. Language, Culture and Body in Human Feeling*. Studies in Emotion and Social Interaction. Cambridge: Cambridge University Press.
- Lakoff, George; Johnson, Mark 1980. *Metaphors We Live By*. Chicago, London: The University of Chicago Press.
- Lakoff, George; Johnson, Mark 1999. *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. New York: Basic Books.
- Medina, John 2008. *Brain Rules*. Seattle: Pear Press.
- Niedenthal, Paula M.; Setterlund, Marc B.; Jones, Douglas E. 1994. Emotional organization of perceptual memory. – P. M. Niedenthal, S. Kitayama (Eds.). *The Hearts Eye. Emotional Influences in Perception and Attention*. New York: Academic Press, 87–113.
- Nõlvak, Aire; Valk, Raivo 2003. Isiksus ja emotsioonid. – J. Allik, A. Realo, K. Konstabel (toim.). *Isiksusepsühholoogia*. Tartu: Tartu Ülikooli Kirjastus, 171–191.
- Orav, Heili 2006. *Isiksuseomaduste sõnavara semantika eesti keeles*. Dissertationes linguisticae Universitatis Tartuensis, 6. Tartu: Tartu Ülikooli Kirjastus.
- Osgood, Charles E.; Suci, George J.; Tannenbaum, Percy H. 1975 [1957]. *The Measurement of Meaning*. Urbana–Chicago: University of Illinois Press.
- Peña Cervel, Sandra M. 2003. *Topology and Cognition: What Image-Schemas Reveal about the Metaphorical Language of Emotions*. LINCOS Studies in Cognitive Linguistics. Munich: Lincom Europa.
- Plutchik, Robert 1980. *Emotion: A Psychoevolutionary Synthesis*. New York: Harper & Row Publishers.
- Szelid, Veronika; Geeraerts, Dirk 2008. Usage-based dialectology. Emotion concepts in the Southern Csango dialect. – *Annual Review of Cognitive Linguistics*, 6, 23–49.
- The Concise Oxford Dictionary 1999. The tenth edition. Oxford: Oxford University Press.

- Tan, Pang-Ning; Steinbach, Michael; Kumar, Vipin 2005. Introduction to Data Mining. Boston: Addison Wesley.
- Vainik, Ene 2004. Lexical Knowledge of Emotions: The Structure, Variability and Semantics of the Estonian Emotion Vocabulary. *Dissertationes linguisticae Universitatis Tartuensis*, 5. Tartu: Tartu Ülikooli kirjastus.
- Vainik, Ene; Orav, Heili 2005. Tee tööd ja näe vaeva, ... aga ikka oled vihane. – Keel ja Kirjandus, 4, 257–277.
- Vainik, Ene; Orav, Heili 2009. Edasipüüdlik, aga tagasihoidlik. Eesti keele kehapõhised kujutelmad isikuomadustest. – Keel ja Kirjandus, 11, 830–840.
- Zlatev, Jordan 2001. A hierarchy of meaning systems based on value. – C. Balkenius, J. Zlatev, H. Kozima, K. Dautenhahn, C. Breazeal (Eds.). *Proceedings of the First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*. Lund University Cognitive Studies, 85. Lund: LUCS, 153–162.

Ene Vainiku (Eesti Keele Instituut) uurimisobjektiks on eesti keele emotsioonisõnavara, selle struktuur, semantika ja varieerumine. Huvipakkuvateks teemadeks on ka psühholingvistika, kognitiivne semantika ja keele ning emotsioonide interaktsioon laiemalt.
Ene.Vainik@eki.ee

Toomas Kirti (Tartu Ülikooli kognitiivse psühholoogia labor) peamised uurimisvaldkonnad on andmeanalüüs, neurovõrgud, iseorganiseerumine, (tehis)intellekt.
Toomas.Kirt@mail.ee

Heili Orava (Tartu Ülikooli eesti ja soome-ugri keeleteaduse instituut) peamised uurimisvaldkonnad on (leksikaalne, kognitiivne, arvuti-) semantika.
Heili.Orav@ut.ee

Appendix 1. Copies of the headers of the 2 e-formulars

ISELOOM JA EMOTSIOONID - Mozilla Firefox

Prindi... Lehekülje sätted... Lehekülg: 1 kokku 3 Mõõtkava: Mahuta lehele Pustpaigutus

ISELOOM JA EMOTSIOONID http://www.eforumar.com/eforumar_eeivaade.php?vorm=emote

ISELOOM JA EMOTSIOONID

Järgnevalt esitame teile rea väljendeid, mida on kasutatud inimese iseloomustamiseks. Palume hinnata, kui võrd on antud SÕNAGA iseloomustatud inimene teie arvates ühtlasi EMOTSIOONALNE?

Valida on variantide vahel:

- väga vähe
- mõnevõrra
- üsnagi
- üle keskmise
- ülimalt

Palume spontaanselt hinnata kõiki sõnu. Vastuseks võib jätta üksnes need sõnad, mille puhul on EMOTSIOONALAASE määr teie arvates tõesti ennustamatu.

Õigeid ja valesid vastuseid ei ole olemas, meid huvitavad vaid teie hinnangud. Ärge palun unustage vajutamast lõpuks ka nuppu "saadan"

NB! Küsitlus on anonüümne ja uurib üksnes sõnade omadusi, mitte teie ega kellegi teise isikut. Küsitluse tulemused avaldatakse teaduspublikatsioonides (ETF grant nr 7149).

*sa in aru juhendist ja olen valmis alustama

jah ei

LAISK

väga vähe mõnevõrra üsnagi üle keskmise ülimalt

JÄRELEANDMATU

väga vähe mõnevõrra üsnagi üle keskmise ülimalt

SPORTLIK

väga vähe mõnevõrra üsnagi üle keskmise ülimalt

RÕÕMSAMEELNE

ISELOOM JA LIIKUVUS - Mozilla Firefox

Prindi... Lehekülje sätted... Lehekülg: 1 kokku 3 Mõõtkava: Mahuta lehele Pustpaigutus

ISELOOM JA LIIKUVUS http://www.eforumar.com/eforumar_eeivaade.php?vorm=liik

ISELOOM JA LIIKUVUS

Järgnevalt esitame teile rea väljendeid, mida on kasutatud inimese iseloomustamiseks. Palume hinnata, kui võrd on antud SÕNAGA iseloomustatud inimene teie arvates ühtlasi LIIKUV?

Valida on variantide vahel:

- väga vähe
- mõnevõrra
- üsnagi
- üle keskmise
- ülimalt

Palume spontaanselt hinnata kõiki sõnu. Vastuseks võib jätta üksnes need sõnad, mille puhul on LIIKUVUSE määr teie arvates tõesti ennustamatu.

Õigeid ja valesid vastuseid ei ole olemas, meid huvitavad vaid teie hinnangud. Ärge palun unustage vajutamast lõpuks ka nuppu "saadan"

NB! Küsitlus on anonüümne ja uurib üksnes sõnade omadusi, mitte teie ega kellegi teise isikut. Küsitluse tulemused avaldatakse teaduspublikatsioonides (ETF grant nr 7149).

*sa in aru juhendist ja olen valmis alustama

jah ei

LAISK

väga vähe mõnevõrra üsnagi üle keskmise ülimalt

JÄRELEANDMATU

väga vähe mõnevõrra üsnagi üle keskmise ülimalt

SPORTLIK

väga vähe mõnevõrra üsnagi üle keskmise ülimalt

RÕÕMSAMEELNE

Both formulars contained 50 (40 + 10) character terms presented in a serial order and followed directly by the 5-point Likert-scale. Choosing the rate of emotionality/mobility was optional. At the end of both questionnaires there was a place to fill in the data about the age and gender.

Appendix 2. Data gathered about the 40 Estonian terms

Estonian term of character trait	English translations	Glosses of the figurative expressions	E_mean	E_StDev	L_mean	L_StDev	Correlation_E&L
flegmaatiline	phlegmatic		1,715	0,982	1,428	0,723	0,464***
pugeja	lickspittle	hole up+A.SBST	2,4	0,972	2,322	0,93	0,425***
omakasu-püüdmatu	altruistic	self+interest-SG+GEN+pursue+CAR.ADJ	2,128	0,892	2,216	0,895	0,417***
hasartne	excitable		3,869	0,975	3,629	0,841	0,409***
spontaanne	spontaneous		3,827	0,989	3,541	0,939	0,407***
loid	listless		1,478	0,805	1,258	0,673	0,384***
temperamentne	spirited		4,201	0,842	3,937	0,871	0,357***
vastutulelik	outgoing	against+come+ADJ	2,586	0,962	2,617	0,858	0,353***
omakasu-püüdlük	selfish	self+interest-SG+GEN+pursue+ADJ	2,629	1,064	2,404	1,04	0,347***
ebajärjekindel	inconsistent	non+track-SG+GEN+firm	2,74	1,068	2,352	0,978	0,338***
ründav	offensive	attack+PRS.PTCP	4,182	0,937	3,44	1,05	0,336***
passiivne	passive		1,548	0,737	1,39	0,67	0,326***
apaatne	apathetic		1,75	1,178	1,216	0,491	0,324***
taltsutamatu	tameless	tame-INF+CAR.ADJ	4,219	0,966	4,043	0,951	0,324***
pöörane	frenetic	turn+ADJ	4,437	0,81	4,143	0,893	0,324***
uimane	sluggish		1,422	0,73	1,264	0,623	0,317***
dünaamiline	dynamic		2,817	1,04	3,652	0,998	0,307***
hüperaktiivne	hyperactive		4,164	0,999	4,708	0,648	0,306***
pidurdamatu	rampant	restrain-INF+CAR.ADJ	4,248	0,975	4,081	0,942	0,271**
pealetükkiv	intrusive	onto+strive+PRS.PTCP	3,462	1,066	4,248	0,975	0,27**
püsimatu	restless	stay+INF+CAR.ADJ	3,522	1,018	4,275	0,894	0,267**
järsk	harsh	steep	3,335	1,172	2,686	0,941	0,259**
heitlik	fickle	throw+ADJ	3,563	1,039	2,704	0,991	0,25**
järeleandmatu	intransigent	after +give+CAR.ADJ	3,075	1,154	2,383	0,976	0,235**
hõljuja	mooncalf	hover+A.SBST	2,76	1,191	2,081	0,798	0,233**
stabiilne	stable		2,013	0,901	2,21	0,822	0,231**
pealehakkaja	eager	upon+start+A.SBST	3,006	0,884	3,785	0,744	0,23**
äkiline	quick-tempered	sudden+ADJ	4,255	0,91	3,503	1,004	0,222**
tagasihoidlik	modest	backwards+hold+ADJ	1,942	0,868	1,782	0,622	0,218**
ebastabiilne	unstable		3,8	1,017	2,73	1,039	0,203**
elav	lively		3,671	0,842	4,107	0,628	0,185**
etteheitev	blaming	ahead+throw-PTCP	3,161	1,089	2,125	0,924	0,168*
reisimishimuline	lusty for travelling		2,987	1,072	4,219	0,722	0,167*
intensiivne	intenseive		3,648	1,044	3,931	0,776	0,163*
tragi	spunky		2,679	0,883	3,762	0,72	0,153*
aktiivne	active		3,182	0,961	4,422	0,598	0,117*
energiline	energetic		3,562	0,998	4,344	0,643	0,108*
ülevoolav	gushy	over+flow+PRS.PTCP	4,352	0,93	3,717	0,995	0,106*
allaheitlik	submissive	down+throw+ADJ	2,155	1,058	1,551	0,592	0,103*
reibas	cheery		3,258	0,917	4,163	0,611	0,092*

Note: Boldfaced are the mean values that are equal or greater than 4, highlighted by italics are the mean values less than or equal to 2. The significance of correlations: *** $p < 0,0001$; ** $p < 0,001$; * $p < 0,05$.

Appendix 3. Data gathered about the 10 control terms

Estonian term of character trait	English	Gloss	E_ mean	E_ StDev	L_ mean	L_ StDev	Correlation_ E&L
närviline	nervous		4,082	0,962	3,481	0,974	0,335***
edasipüüdlik	ambitious	forward+ pursue+ADJ	2,576	0,878	3,497	0,900	0,328***
õel	mean		2,875	1,083	2,062	0,878	0,283***
laisk	lazy		1,650	0,790	1,294	0,658	0,210**
virg	diligent		2,500	0,939	3,857	0,805	0,190**
kade	envious		2,801	1,153	1,822	0,748	0,142*
sportlik	sporty		2,550	1,094	4,609	0,634	0,109*
kiire	fast		2,866	1,055	4,331	0,687	0,101*
tundlik	sensitive		4,182	0,928	2,422	0,907	0,051
rõõsameelne	joyous		3,975	0,908	3,875	0,687	0,045

Note: Boldfaced are the mean values that are equal or greater than 4, highlighted by italics are the mean values less than or equal to 2. The significance of correlations: *** $p < 0,0001$; ** $p < 0,001$; * $p < 0,05$.

LIIKUMISE JA EMOTSIOONIDE KAAS-KONTSEPTUALISEERIMINE ISIKUOMADUSTE SÕNAVARAS

Ene Vainik, Toomas Kirt, Heili Orav

Eesti Keele Instituut, Tartu Ülikool, Tartu Ülikool

Käesoleva artikli eesmärgiks oli kontrollida, kas hüpotees liikumise ja emotsioonide kaas-kontseptualiseerimisest isikuomaduste sõnavaras vastab tõe. Uuringu keelematerjalina kasutasime Orava poolt 2003. aastal läbi viidud isikuomaduste loetelukatsete tulemusi (1272 keelendit), millest sõltumatute eksperthinnangute baasil eraldati potentsiaalselt nii liikumisega kui emotsioonidega seotud väljendid (106). Uurijate intuitsioone emotsioonide ja liikumise ko-kontseptualiseerimise esinemise ja määra kohta kontrolliti küsitluste abil, milles osales 162 inimest (e-formular). Uurimise all oli 40 testsõna (nt *apaatne*, *pealetükkiv*, *temperamentne*, *hüperaktiivne*), mille kõigi puhul tuli esile positiivne korrelatsioon emotsionaalsus- ning liikuvushinnangute vahel. Selline empiiriline tulemus näitab, et keelekõnelejad tajuvad neid kahte karakteristikut inimese iseloomus koosinevatena.

Lisaks traditsioonilisele analüüsile kasutasime oma uuringus ka iseorganiseeruvate kaartide (ingl *self-organizing map*, SOM) meetodit, et tuua välja andmete varjatud struktuur ning muuta tulemused visualiseerimise läbi kergesti haaratavaks.

Iseorganiseeruvate kaartide meetod kujutab endast edasisidega neurovõrku, mis läbi iseorganiseerumise protsessi projitseerib paljumõõtmelise sisendandmesitiku kahemõõtmelisele väljundkaardile, paigutades sarnased sisendid lähestikku. Kasutades kaartide lokaalse naabruse sarnasuse eeldust saime hinnata erinevate kaartide sarnasust, mõõtes nende lokaalsete naabrussuhete vastavust. Artiklis esitatud kahe SOM-kaardi meta-analüüs näitab 1) lähisünonüümsete sõnade “kobaraid” koos kõige rohkem korreleerunud emotsioonide ning liikumise näitajatega 2) kahe andmehulga varjatud üldist struktuuri. Tulemustest nähtuvad kõrge/madala emotsionaalsuse ja mobiilsusega iseloomusõnade hulgad ning ka sõnade semantikkasse peidetud aksioloogilised hinnangud – nii positiivsed kui ka negatiivsed. See tõestab veelkord, et iseloomusõna ja emotsioone tajutakse alati kas meeldivate või ebameeldivatena.

Võtmesõnad: liikumine, emotsioonid, isikuomadused, semantika, iseorganiseeruvad kaardid, eesti keel