

Eesti Rakenduslingvistika Ühing

Eesti Rakenduslingvistika Ühingu aastaraamat 4

Estonian Papers in Applied Linguistics 4

Toimetanud
Helle Metslang, Margit Langemets
ja Maria-Maren Sepper

Eesti Keele Sihtasutus
Tallinn 2008

SISUKORD

Statistilise keelemudeli adapteerimine eesti keele kõnetuvastuses <i>Statistical language model adaptation for Estonian speech recognition</i> Tanel Alumäe	5
Fraseologismide semantilise teisititõlgendamise võimalikkusest <i>On the semantic ambiguity of phraseologisms</i> Anneli Baran	17
Mida räägivad sõnasised pausid ja kordused sihtkeele käänamisprotsessist? <i>What do word-internal pauses, cut-offs, and retracings tell about inflectional process?</i> Annekatriin Kaivapalu	35
Kõnetehnoloogia vajab žanrilist lähenemist <i>Speech technology needs a genre-based approach</i> Krista Kerge, Hille Pajupuu, Kairi Tamuri, Heidi Meier	53
Sõnaraamatu kohtulingvistiline analüüs: Eesti pretsedent <i>Linguistic forensic analysis of a dictionary: an Estonian precedent</i> Margit Langemets, Piret Voll	67
Differences in adult L2 grammars of Lithuanian: the case of gender agreement <i>Erinevusi täiskasvanud õppijate leedu keele grammatikas, kes õpivad leedu keelt teise keelena – sooühildumine</i> Simona Mačiukaitė	87
Eesti verbi minevikuvormidest venekeelsete üliõpilaste kasutuses <i>Usage of Estonian past verb tenses by Russian-speaking students</i> Diana Maisla	105
In search of the best technique for vocabulary acquisition <i>Parimat sõnavara omandamise meetodit otsimas</i> Mohammad Mohseni-Far	121

Üleküsimine, ümbersõnastamine ja mittemõistmine infotelefonikõnedes <i>Clarification, reformulation and non-understanding in phone conversations</i> Siiri Pärkson	139
Ladina õigusterminite vormilisest ja sisulisest korrektsusest <i>Accuracy of the form and content of Latin legal terms</i> Merike Ristikivi	157
Konstruksioonipõhine keelemudel ja sõnaraamatumudel <i>Some lexicographic implications of a construction-based model of language</i> Heete Sahkai	171
Ühest katsest emotsiooniverbidega <i>Über ein experiment mit emotionsverben</i> Anni Silk	187
Teaduskeele seire Tallinna Ülikoolis: teadustöötajate hoiakud <i>Monitoring academic language at Tallinn University: academic staff attitudes</i> Anastassia Zabrodskaja	203
Kuidas me mõistame mõisteid? <i>How do we comprehend the concepts?</i> Ene Vainik, Toomas Kirt	225
Eesti-X-keele sõnastik ja grammatika <i>Estonian-X dictionary and grammar</i> Ülle Viks	247
Eesti silbisüsteemi struktuurist <i>A preliminary structural view of the Estonian syllable system</i> Leo Võhandu, Kairit Sirts, Eik Aab	263
POLEEMILIST Eesti keel ja teised keeled – Eesti jaoks <i>Estonian and other languages: for Estonia</i> Suliko Liiv	271

STATISTILISE KEELEMUDELI ADAPTEERIMINE EESTI KEELE KÕNETUVASTUSES

Tanel Alumäe

Ülevaade. Artiklis käsitletakse eesti keele suure sõnavaraga kõnetuvastuse statistilise keelemudeli adapteerimist. Adapteerimise lähteandmeteks on väike teemaspetsiifiliste lausete korpus. Adapteerimise käigus leitakse varjatud semantika analüüsi (LSA) abil suurest dokumendikorpusest antud teemale lähedaseimad tekstid. Saadud tekstide põhjal konstrueeritakse uus teemaspetsiifiline unigramm-mudel ning see kombineeritakse üldise N -gramm-mudeliga, mille tulemusena saadakse teemale adapteeritud N -gramm-mudel. Artiklis võrreldakse morfeeme, sõnu ja lemmasid adapteerimismudeli põhiühikutena.

Meetodit testitakse raadiouudistesalvestuste tuvastamisel. Tuvastuse esimeses faasis leitakse üldise keelemudeli abil igale uudisnupule tuvastushüpooteesid, mida kasutatakse keelemudeli adapteerimiseks. Tuvastuse teises faasis kasutatakse adapteeritud keelemudelit uute tuvastushüpooteeside saamiseks. Tulemused näitavad, et adapteerimisega saavutatakse oluline tuvastuskvaliteedi paranemine. Selgub, et morfeemidepõhisel adapteerimisel saavutatud 10-protsendiline vigade vähenemine on statistiliselt oluliselt suurem kui sõna- või lemapõhisel adapteerimisel saadud muutused. Artiklis analüüsitakse ka saadud tulemuste võimalikke põhjuseid.*

Võtmesõnad: kõnetuvastus, keelemudeli adapteerimine, LSA, lemmatiseerimine, morfeemid

1. Sissejuhatus

Suure sõnavaraga kõnetuvastuses kasutatakse statistilist keelemudelit sõnade aprioorse kontekstuaalse tõenäosuse hindamiseks. Tüüpiliselt kasutatakse keelemudelinähtena N -gramm-mudelit, kus eeldatakse, et sõna kontekstuaalne tõenäosus sõltub ainult temale eelnevast $N-1$ sõnast. Tavaliselt piirduakse trigramm-mudelitega

* Artikkel on valminud riikliku programmi "Eesti keele keeletehnoloogiline tugi (2006–2010)" rahalisel toel.

($N=3$). Statistiline keelemudel saadakse suure tekstikorpuse (kümned kuni sajad miljonid sõnad) analüüsi põhjal.

Statistilise keelemudeli adapteerimine on ülesanne, kus väikese olemasoleva teemaspetsiifilise korpuse põhjal kohandatakse üldist keelemudelit nõnda, et ta sobib paremini antud teemale. Kohandamise tulemusena peaks kõikide teemaga lähedalt seotud olevate sõnade ja sõnakombinatsioonide tõenäosused suurenema ning antud teemast semantiliselt kaugel olevate sõnade tõenäosused vähenema.

Viimastel aastatel on mitme keele kõnetuvastuse keelemudeli adapteerimiseks edukalt kasutatud nn varjatud semantika analüüsi (ingl *latent semantic analysis*, LSA) (vt nt Bellegarda 1998). Meetod kohandab järk-järgult keelemudelit vastavalt hiljuti tuvastatud sõnadele, kasutades sõnade dokumentides koosesinemise statistikat keelemudelis olevate sõnade unigramm-tõenäosuse ümberarvutamiseks.

See lähenemine aga ei pruugi olla otstarbekas flekteeruvate ja aglutinatiivsete keelte puhul. Sellistes keeltes on erinevate sõnavormide arv väga suur, mistõttu statistilises keelemudelis kasutatakse sõnade asemel morfeeme (Alumäe 2006) või tekstikorpuse statistika põhjal leitud morfeemilaadseid ühikuid (Siivola jt 2003). Et rakendada standardset LSA-põhist adapteerimismeetodit, peaks ka LSA-mudelis dokumentide esitamiseks kasutama morfeeme. Kuna LSA kasutab dokumentide esitamiseks järjestamata sõnade esinemissageduse ehk nn *bag-of-words* meetodit, tekkis autoril kahtlus, et dokumendi esitamine morfeemide esinemissagedusena annab dokumendi sisust vähem aimu, kui dokumendi esitamine sõnade või lemmade esinemissagedusena.

Selles artiklis tutvustatakse LSA-põhist statistilise keelemudeli adapteerimis-meetodit, kus LSA-mudelis kasutatavad ühikud ei pruugi kattuda keelemudeli ühikutega. See annab võimaluse kasutada semantiliste seoste modelleerimiseks sõnu, lemmasid, morfeeme või muid ühikuid. Treeningu käigus viiakse dokumendikorpuses olevad tekstid soovitud kujule, st leitakse neis olevate valitud ühikute sisaldus. Selle põhjal arvutatakse LSA sarnasusmudel. Kõnetuvastus koosneb siis kahest faasist: esimeses faasis leitakse igale tuvastatavale lausele N parimat lausekandidaati. Parimad lausekandidaadid viiakse LSA-mudeliga ühilduvale kujule (st näiteks lemmatiseeritakse) ning selle põhjal leitakse aktiivse teema kujutis LSA-ruumis. Seejärel leitakse antud teemale kõige lähedasemad treeningdokumendid ning nende põhjal kohandatakse esimeses faasis kasutatud üldist keelemudelit. Saadud adapteeritud keelemudeli abil arvutatakse esimeses faasis saadud lausekandidaatidele uued keelemudelipõhised skoorid ning selle põhjal leitakse uued parimad lausehüpoteesid.

2. LSA

Varjatud semantika analüüs (LSA) (Landauer 1998) on korpusepõhine matemaatiline meetod sõnade ja dokumentide semantilise sarnasuse arvutamiseks ja esitamiseks. LSA ülesandeks on M sõnast koosneva sõnahulga V ja N dokumendist koosnevad dokumendikorpuse T kujutamise vektorruumis, nii et iga sõna hulgas V ja iga dokument hulgas T oleks esitatavad selles ruumis. Selleks konstrueeritakse kõigepealt $M \times N$ maatriks W , mille iga elemendi W_{ij} väärtus on sõna w_i esinemisarvu kaal dokumendis d_j . Kasutatav kaalufunktsioon peaks arvesse võtma nii sõna

tähtsust antud dokumendis kui ka sõna semantilist informatiivsust. Üheks selliseks kaaluks (Bellegarda 1998), mida ka antud töös kasutatakse, on

$$(1) \quad W_{ij} = (1 - \varepsilon_i) \log_2 \left(1 + \frac{c_{ij}}{n_j} \right),$$

kus c_{ij} on sõna w_i esinemisarv dokumendis d_j , n_j on dokumendis d_j olevate sõnade arv ning ε_i on sõna w_i normaliseeritud entroopia korpuses T . Normaliseeritud entroopia vastab sõna w_i negatiivsele semantilisele informatiivsusele ning see arvutatakse järgmiselt:

$$(2) \quad \varepsilon_i = -\frac{1}{\log(N)} \sum \frac{c_{ij}}{t_j} \log \frac{c_{ij}}{t_j}.$$

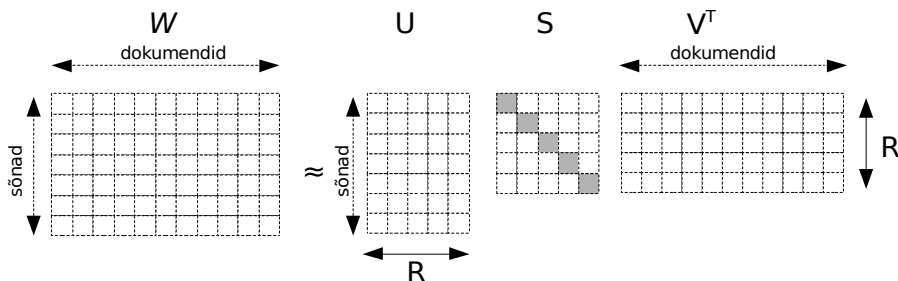
Siin $t_i = \sum c_j$ on sõna w_i esinemisarv korpuses T . Ühtlaselt ja paljudes dokumentides esinevate sõnade (näiteks side- ja asesõnad) entroopiaväärtus on lähedane ühele, ning suhteliselt vähestes dokumentides esinevate sõnade entroopiaväärtus lähedane nullile (nt allpoolkirjeldatud eksperimentides on sõna *ja* normaliseeritud entroopia 0,96 ning sõna *kvantarvuti* vastav väärtus 0,17). Sellise kaalufunktsiooni kasutamise tulemusena rõhutatakse maatriksis W rohkem selliseid sõnu, mille semantiline informatiivsus on suur (ning entroopia seega väike): on ju selge, et kui kahes dokumendis kasutatakse sagedasti sõna *ja*, siis ei pruugi nad olla sisult sarnased; kui aga mõlemas kasutatakse sagedasti sõna *kvantarvuti*, siis on dokumentide sisuline lähedus palju tõenäolisem.

Saadud maatriksi W veerud kirjeldavad korpuses olevaid dokumente skaleeritud sõnade esinemissagedustega. Igale dokumendile vastab M -mõõtmeline veeruvektor, kus M on tüüpiliselt 20 000–60 000. Kahe dokumendi võrdlemiseks tuleks siin võrrelda kahe dokumendivektori vastavaid elemente. Dokumendivektori suurest dimensionaalsusest tingituna on see üsna arvutusmahukas, samuti on olemas oht, et sisult sarnased dokumendid kasutavad pisut erinevat sõnavara (nt sünonüüme), mille tulemusena dokumentide kaugus vektorruumis tuleks ikkagi küllalt suur. Seetõttu on leitud (Deerwester jt 1990), et dimensionaalsuse vähendamiseks ning kompaktsete tunnuste esiletoomiseks on kasulik rakendada maatriksile W nn lõigatud singulaarlahutust (ingl *truncated singular value decomposition*, SVD).

Järguga R lõigatud singulaarlahutuse tulemusena lahutatakse esialgne maatriks W kolmeks faktoriks

$$(3) \quad W \approx \hat{W} = USV^T,$$

kus U on vasakpoolsete singulaarvektorite u_i maatriks ($M \times R$), S on R singulaarväärtusest koosnev diagonaalmaatriks ning V on parempoolsete singulaarvektorite v_j maatriks ($N \times R$). Maatriks \hat{W} on R järku parim lähendus algsele maatriksile W . Vektorid u_i esitavad sõnu w_i ja vektorid v_j dokumente d_j saadud LSA-ruumis (vt joonis 1).



Joonis 1. Singulaarlahutus sõna–dokument maatriksist

Kui enne singulaarlahutuse rakendamist esitati dokumente M -mõõtmeliste maatriksi veeruvektoritega, siis nüüd vastab igale dokumendile R -mõõtmeline vektor $v_i S$. Kuna R valitakse tüüpiliselt vahemikust 100...300 (autori eksperimentides $R=200$), siis saavutatakse sellise teisendusega oluline dimensionaalsuse vähenemine. Samuti on leitud, et lõigatud SVD leiab alge maatriksi W kõige olulisemad komponendid ning ignoreerib kõrgemat järku struktuure, mida võib pidada müüraks (Yang 1995). Teisisõnu, vektor $v_i S$ esitab kompaktselt dokumendi d_j sõnavara.

3. Keelemudeli adapteerimine

Keelemudeli adapteerimiseks kasutatakse üldise keelemudeli marginaalide adapteerimist vastavalt teemaspetsiifilistele lausetele semantiliselt lähedasemate dokumentide analüüsist saadud unigramm-tõenäosustele. Semantiliselt lähedasemad dokumendid leitakse eelnevalt kompileeritud LSA-mudeli põhjal.

3.1. Lähedaste dokumentide leidmine

Semantiliselt lähedaste dokumentide leidmiseks tuleb esmalt olemasolevatest teemaspetsiifilistest lausetest e adapteerimislausetest koosnev pseudodokument viia LSA-ruumiga ühilduvale kujule. Selleks tuleb kõigepealt leida lausetes esinevate erinevate ühikute (morfeemide, sõnade või lemmade) arv ning siis arvutada uue pseudodokumendi vektor \tilde{d}_p vastavalt valemile (1). Saadud pseudodokumendi kujutis LSA-ruumis on siis

$$(4) \quad \tilde{v}_p = \tilde{d}_p^T U S^{-1}.$$

Seejärel võib arvutada kauguse pseudodokumendi vektori \tilde{v}_p ja kõikidele korpuses olevatele dokumentidele vastavate vektorite v_i vahel, leides nurga vektorite $\tilde{v}_p S$ ja $v_i S$ vahel:

$$(5) \quad K(\tilde{v}_p, v_i) = \angle(\tilde{v}_p S, v_i S) = \arccos \frac{\tilde{v}_p S^2 v_i^T}{\|\tilde{v}_p S\| \cdot \|v_i S\|}$$

Kauguse põhjal järjestatakse kõik korpuses olevad dokumendid ning keelemudeli adapteerimiseks valitakse L antud teemale kõige lähedasemat dokumenti.

3.2. Kaalutud unigramm-mudeli arvutamine

Keelemudeli kohandamisel saadud dokumendikauguste teadmisel oleks kasulik, kui semantiliselt lähedasemad dokumendid annaksid adapteerimisel suuremat kaalu kui kaugemad dokumendid. Selleks arvutatakse L lähedaseimas dokumendis esineva i -nda keeleühiku kaalutud arv järgmiselt:

$$(6) \quad c_{adapt}(i) = \sum_{j \in C_L} \left(1 - \frac{K(\tilde{v}_p, v_j)}{\pi} \right) c_{ij} ,$$

kus C_L on L lähedaseimale dokumendile vastav indeksite hulk. Saadud murdosalistest keeleühiku esinemisarvudest koostatakse unigramm-keelemudel, kasutades nn loomulikku diskonteerimist (Ristad 1995). See diskonteerimismeetod valiti seetõttu, et teda on hõlbus kasutada murdarvuliste esinemisarvude puhul ning ta andis arendusfaasis häid tulemusi.

3.3. Kiire marginaalide adapteerimine

Kiire marginaalide adapteerimine (ingl *fast marginal adaptation*, FMA) (Kneser jt 1997) on meetod, millega saab üldise N -gramm keelemudeli kiiresti adapteerida teemaspetsiifilisele tekstile. Meetod kasutab algse keelemudelina segakorpusel treenitud keelemudelit. Adapteerimise käigus muudetakse algset keelemudelit nii, et selle marginaaljaotus oleks võrdne teemaspetsiifilistel tekstidel treenitud unigramm-mudeliga. Selgub, et see on samaväärne üldises N -gramm-mudelis olevate tõenäosuste skaleerimisega

$$(7) \quad P_{adapt}(w|h) = \frac{\alpha(w)P_{BG}(w|h)}{Z(h)} ,$$

kus $P_{adapt}(w|h)$ on adapteeritud sõna w kontekstuaalne tõenäosus, $P_{BG}(w|h)$ on sõna w esialgne kontekstuaalne tõenäosus ning $Z(h)$ on normaliseerimisfaktor, mis tagab tõenäosuste summeerumise üheks. Skaleerimisfaktor $\alpha(w)$ on ligikaudselt

$$(8) \quad \alpha(w) \approx \left(\frac{P_{adapt}(w)}{P_{BG}(w)} \right)^\beta ,$$

kus $P_{adapt}(w)$ on sõna w adapteerimisdokumentide põhjal leitud unigramm-tõenäosus, $P_{BG}(w)$ sõna w unigramm-tõenäosus üldise keelemudeli järgi, ning β valitud faktor vahemikus 0...1, mis eksperimentides seati võrdseks 0,5-ga. Skaleerimisfaktori ülesandeks on keeleühikute tõenäosuse suurendamine või vähendamine, vastavalt ühiku suhtelisele tõenäosusele adapteerimiskorpuses võrreldes üldise segakorpusega. Normaliseerimisfaktor on esitatav kujul

$$(9) \quad Z(h) = \sum_w \alpha(w)P_{BG}(w|h) .$$

4. Eksperimendid

4.1. Eksperimentaalse ülesande kirjeldus

Eksperimentaalseks ülesandeks võeti Eesti Raadio (ER) uudiste transkribeerimine. Kõnematerjal koosneb Vikerraadio täistunni lühiuudistest, mis on osa valmivast ER uudistekorpusest. Eksperimentideks valiti juhuslikult mõned olemasolevad uudistesaadete salvestused. Saadetes olev kõne transkribeeriti käsitsi ning segmenteeriti uudisnuppudeks ja lauseteks. Segmendid, mis ei sisaldanud kõnet vaid muud helimaterjali (nt ava- ja lõppsignatuurid), eemaldati. Meetodi testimiseks kasutati 21 minutit kõnematerjali, mis koosneb 44 uudisnupust ja 193 lausest. Erinevatele parameetritele parimate väärtuste leidmiseks kasutati 21 minutit kõnematerjali, mis koosneb 20 uudisnupust ja 101 lausest. Keskmine lausete arv ühes uudisnupus oli seega 4,6.

4.2. Akustilised mudelid

Kuna terve ER uudistekorpus ei ole veel transkribeeritud, ei saanud akustiliste mudelite treenimiseks kasutada uudistes olevat kõnet. Selle asemel kasutati mudelite treenimiseks eesti keele SpeechDat-tüüpi kõneandmebaasi (Meister jt 2003). Treenimiseks kasutati ainult andmebaasis olevaid aktsepteeritava kvaliteediga salvestusi, mille koguarv on 2969, kestusega 241 tundi. Erinevate kõneleajate arv on 1332. Kõne on digitaliseeritud 8 kHz sagedusel kasutades 8-bitist A-law-koodeeringut.

Akustiliste mudelite treenimiseks rakendati vaba lähtekoodiga SphinxTrain¹ tarkvarapaketti. Mudelite hulgas on 25 foneemmudelit, 5 erinevat müramudelit ning vaikuse mudel. Akustiliste tunnustena kasutati mel-sagedusskaalaga kepstrumkoefitsiente (ingl *mel frequency cepstrum coefficients*, MFCC), mis arvutati 130 Hz–3400 Hz laiusest sagedusribast. Sämplimisakna laius oli 25,6 ms ning sämplimisperiod 10 ms. Igast aknast arvutati 512-punktilise kiire Fourier' teisenduse (ingl *fast Fourier transform*, FFT) abil 31-ribaline filterpank, mida omakorda kasutati 13 kepstrumkoefitsiendi arvutamiseks. Akustiliste mudelitena kasutatakse kolme emiteeriva olekuga nn paremalt-vasakule topoloogiaga Markovi peitmudelite.² Väljundvektorid on 39-kohalised ning koosnevad 13 kepstrumkoefitsiendist ning neile vastavatest delta- ning delta-delta-koefitsientidest.³ Treenimise käigus saadakse seotud olekutega trifoone mudelid, millel on kokku 8000 olekut. Igat olekut modelleeritakse kaheksa Gaussi jaotuse summaga.

Enne tuvastamist adapteeriti SpeechDat-kõneandmebaasi põhjal treenitud mudeleid u 15 minuti ER uudistekorpuses oleva käsitsi transkribeeritud kõne põhjal.

Kõnetuvastussüsteemi häälussõnastik koostati automaatselt sõnade ortograafia põhjal, kasutades väikest arvu kontekstitundlikke reegleid ning käsitsi leitud reegleid tähtsamate võõrnimede ortograafia eestipärasele kujule viimiseks.

¹ Vt <http://cmusphinx.org> (01.01.2007).

² Paremt-vasakule topoloogiaga Markovi peitmudeliteks nimetatakse sellise topoloogiaga mudeleid, kus puuduvad nn *skip*-üleminekud, st mudeli suvalisest olekust saab liikuda ainult järgmisesse olekusse, või samasse olekusse tagasi. Sellise mudeli käivitamisel külastatakse alati kindlas järjekorras kõiki mudeli olekuid.

³ Delta- ja delta-delta-koefitsiendist võrduvad vastavate koefitsientide esimest ja teist järku tuletistega.

4.3. Keelemudel

Üldise statistilise keelemudeli treenimiseks kasutati järgnevaid osasid Tartu Ülikooli (TÜ) eesti keele segakorpusest (Kaalep, Muischnek 2005): ajalehed Postimees (33 miljonit sõna), Eesti Ekspress (7,5 miljonit sõna), Maaleht (4,3 miljonit sõna), ilukirjandus (4,2 miljonit sõna), ajakirjad Akadeemia (7 miljonit sõna) ning Kroonika (0,6 miljonit sõna), riigikogu stenogrammid (13 miljonit sõna). Lisaks sellele koguti Internetist Eesti Päevalehe artiklite korpuse (93 miljonit sõna) ja etv24.ee uudistekorpuse (4,8 miljonit sõna).

Keelemudeli komplekteerimiseks kasutati tarkvarapaketti SRILM (Stolcke 2002). Kõik tekstid töödeldi eelnevalt eesti keele morfoloogiaanalüsaatori ja -ühetaja (Kaalep, Vaino 2001) abil, mille põhjal segmenteeriti sõnad morfeemideks. Keelemudeli sõnavaraks on 60 000 morfeemi, mis valiti maksimaalse tõepära meetodil kõigi alamkorpuste segust TÜ tasakaalustatud korpuse põhjal. Kasutades saadud sõnavara, kompüleeriti kõikidele alamkorpustele oma trigramm-mudel. Trigramm-mudelite arvutamisel kasutati SRILM-is realiseeritud modifitseeritud Kneser-Ney diskonteerimist. Kõik trigramm-mudelid kompüleeriti lõpuks üheks üldiseks mudeliks, kasutades TÜ tasakaalustatud korpust optimaalsete kombineerimiskaalude leidmiseks.

Saadud keelemudeli entroopia-mõõde (ingl *perplexity*) kõnetuvastuseksperimentides kasutatavate uudislõikude transkriptsioonide suhtes on 128. Keelemudeli sõnavara katvus on 99,5%.

4.4. LSA-mudeli konstrueerimine

Semantiliste seoste modelleerimiseks konstrueeriti kolm LSA-mudelit: sõnade-, lemmade- ja morfeemidepõhine. Lähteandmetena kasutati ülaltoodud ajalehekorpuste ja etv24.ee artikliteks segmenteeritud tekste. Kokku on nendes korpustes u 500 000 artiklit. Lemmapõhise mudeli konstrueerimiseks asendati kõik artiklites olevad sõnad vastavate lemmadega, kasutades morfoloogiaanalüsaatorit. Morfeemidel põhineva mudeli konstrueerimiseks segmenteeriti sõnad morfeemideks. Sõna- ja lemmapõhise mudeli sõnavaraks võeti 60 000 kõige sagedasemat vastavat ühikut. Morfeemipõhise mudeli sõnavaraks võeti samad 60 000 morfeemi, mis olid eelnevalt valitud statistilise keelemudeli sõnavaraks. Sõnapõhise mudeli esialgses sõna-dokument maatriksis on u 84 miljonit nullist erinevat elementi. Lemmapõhise mudeli vastav väärtus on 79 miljonit ning morfeemipõhisel mudelil 117 miljonit.

Algsete maatriksite konstrueerimiseks ning dokumentide läheduse arvutamiseks kasutati autori loodud tarkvara. Maatriksite singulaarlahutuse arvutamiseks rakendati tarkvarapaketti PROPACK.⁴

⁴ [http://sun.stanford.edu/~rmunk/PROPACK/ \(01.01.2007\).](http://sun.stanford.edu/~rmunk/PROPACK/ (01.01.2007).)

4.5. Kõnetuvastuse meetodika

Kõnetuvastuseksperimentideks kasutati tarkvarapaketti CMU Sphinx 3.6.3.⁵ Tuvastuse esimeses faasis genereeriti igale lausungile 1000 parimat lausehüpoteesi. Iga uudisnupu parimatest lausehüpoteesidest konstrueeriti pseudodokumendid, mis projitseeriti LSA-ruumi. Igale uudisnupule leiti selle põhjal 1200 sarnaseimat dokumenti. Nende dokumentide põhjal konstrueeriti igale uudisnupule teemaspetsiifiline unigramm-mudel, mille abil adapteeriti üldises trigramm-mudelis olevaid kontekstuaalseid tõenäosusi. Seejärel arvutati igale esimeses faasis saadud lausehüpoteesile uued keelemudelipõhised skoorid. Need kombineeriti olemasolevate akustilise mudeli põhiste skooridega ning selle põhjal leiti igale lausungile uus parim lausehüpotees.

4.6. Tulemused

Kõnetuvastuse kvaliteedi hindamiseks kasutati tähevigade osakaalu (ingl *letter error rate*, LER), mis on defineeritud kui

$$(10) LER = \frac{S+D+I}{N} \cdot 100 ,$$

kus S on asendusviga, D kustutusviga, I vahelekirjutusviga ning N lauses olev tegelik tähtede arv. Sõnadevahelist tühikut käsitleti eraldi tähena, et võtta arvesse võimalikke kokku-lahku-kirjutamise vigu. Vigade osakaalu mõõdeti enne ja pärast adapteeritud keelemudeli rakendamist. Tulemused on toodud tabelis 1. Tähevigade järel on toodud adapteerimisega saavutatud suhteline vigade vähenemise protsent, võrrelduna ilma adapteerimiseta saadud tulemustega.

Tabel 1. Tähevigade osakaal enne ja pärast adapteerimist, koos suhtelise vigade arvu muutusega võrrelduna enne adapteerimist saadud tulemustega

Adapteerimine	Tähevigade osakaal	Suhteline vigade arvu muutus
–	7,1%	
sõnapõhine	6,7%	–6%
lemmapõhine	6,6%	–7%
morfeemipõhine	6,4%	–10%

Tulemuste statistiliseks kontrolliks rakendati Wilcoxon'i märgistatud astaktesti,⁶ mille abil analüüsiti tuvastuskvaliteeti erinevate uudislugude kaupa enne ja pärast adapteerimist. Osutus, et kõik adapteeritud mudelid parandasid statistiliselt oluliselt adapteerimata süsteemi tulemusi. Morfeemipõhine adapteerimine osutus oluliselt paremaks kui teised adapteerimisviisid, lemma- ning sõnapõhisel adapteerimisel olulist vahet ei olnud.

Tabelis 2 on toodud mõne uudisnupu esimene lause ning morfeemipõhise LSA-mudeli abil leitud sarnaseimate artiklite pealkirjad. Iga uudisnupu kohta on toodud neli talle lähedaseimat artiklit.

⁵ <http://cmusphinx.org> (01.01.2007).

⁶ Wilcoxon'i märgistatud astaktesti kasutatakse kahe valimi võrdlemiseks juhul, kui valimite jaotus erineb oluliselt normaaljaotusest. Test näitab, kas valimite vaheline erinevus on oluline või võib seda seletada juhusega.

Tabel 2. Mõned uudisnupud (toodud uudise esimene lause) ja neile lähedaste artiklite pealkirjad

Uudis	Lähedased artiklid
Seoses Tony Blairi visiidiga tõhustatakse tänasest alates Eesti piirikontrolli.	<ol style="list-style-type: none"> 1. Leedu peaminister sõidab visiidile Riiga. 2. Kallas sõidab Briti peaministriga kohtuma. 3. Balti siseministrid kohtuvad Saaremaal. 4. Balti riikide sisepiiride valvamiseks pole raha.
Laskesuusatamise maailmakarikasarja avaetapil Östersundis ei pääsenud Eesti teatenelik starti.	<ol style="list-style-type: none"> 1. Kuus hemoglobiinitasemega hädas olnud sportlast said loa võistelda. 2. Venemaa laskesuusatajad protestivad. 3. Teatesõite valitsesid Venemaa ja Norra. 4. Venelased kaebavad Lazutina ja Danilova karistused edasi.
Suurbritannia suurendab sõdurite arvu Afganistani missioonil.	<ol style="list-style-type: none"> 1. Poola saadab Afganistani 1000 lisasõdurit. 2. Suurpealetung Afganistanis. 3. Suurbritannia viib Afganistani veel 3300 sõdurit. 4. Afganistani hakkab stabiliseerima 3000 välismaist sõjaväelast.
Kümmekond hariduse ja terviseiga seotud organisatsiooni nõuab seksuaalkasvatuse muutmist kohustuslikuks õppeaineks.	<ol style="list-style-type: none"> 1. Tervisekaitseorganisatsioonid soovivad eraldi õppeainet seksuaalkasvatuse jaoks. 2. Arstid nõuavad koolidesse kohustuslikku seksuaalkasvatust. 3. Õpilased õpetavad eakaaslasti tervist hoidma. 4. Hariduse andmine on õpetaja isiklik mure.

5. Arutelu

Katsetulemused näitasid, et morfeemipõhine adapteerimine annab kõnetuvastuses paremaid tulemusi, kui sõna- või lemmapõhine adapteerimine. See on mõneti vastuolus autori esialgse hüpoteesiga, mis kahtles morfeemide võimes kanda semantilist sisu. Siiski, morfeemide hea tulemus on seletatav mitme asjaoluga.

Esiteks, morfeemipõhise adapteerimismudeli sõnavara saab seada identseks kõnetuvastuses kasutatava keelemudeli sõnavaraga. See tähendab, et morfeemipõhise adapteerimismudeli efektiivne katvus tuvastuse esimesest faasist saadud lausehüpoteeside suhtes on 100%.

Teiseks, viimasel ajal on mitmed korpuslingvistid tõdenud käändevormide olulisust võrreldes lemmadega või lisaks lemmadele. Näiteks František Čermáki (2007) arvates “leiutatakse lemmatiseerimise käigus uusi kujutletavaid maailmu, mida tegelikult tihti ei eksisteeri”. Lemmatiseerimise suurim oht on potentsiaalne liigne üldistamine ning sellest tulenev suur informatsioonikadu, mis ka antud lemmapõhise lähenemise käigus võib toimuda. Selle argumendiga on seletatav ka sõnapõhise mudeli suhteliselt väike allajäämine lemmapõhisele mudelile, kuigi sõnapõhise mudeli katvus on tunduvalt väiksem.

Kolmandaks, dokumendi esitus morfeemide kogumina võib anda dokumendi sisust rohkem informatsiooni, kui esialgu tundub. Näiteks raudteejaamast rääki-

vas artiklis on ilmselt kõrge morfeemide *raud, tee, jaam* ja tõenäoliselt ka *jaama* sagedus. Kuna LSA kasutab nn *bag-of-words* lähenemist, ei ole dokumendivektorit vaadates muidugi selge, mis järjestuses neid morfeeme kasutatakse. Sellest hoolimata – dokumendid, kus esinevad süstemaatiliselt koos nimetatud morfeemid, räägivad ilmselt siiski raudteejaamaga seotud teemal, kuigi teoreetiliselt saaks nendest morfeemidest ka teisi sõnu moodustada.

6. Kokkuvõte

Artikkel kirjeldas eksperimente eesti keele üldise statistilise keelemudeli automaatse adapteerimisega. Välja pakutud meetod kasutab keelemudelist iseseisvat LSA-põhist dokumentide vektoriruumi olemasolevale teemaspetsiifilisele tekstile semantiliselt lähedaste tekstide leidmiseks suurest dokumendikorpusest. Saadud dokumentide põhjal koostatakse uus teemaspetsiifiline unigramm-mudel ning see kombineeritakse üldise *N*-gramm-mudeliga, kasutades kiiret marginaalide adapteerimist. Eksperimentide käigus loodi kolm erinevat – sõnadel, lemmadel ja morfeemidel põhinev – LSA-mudelit ning võrreldi nende efektiivsust eesti keele suure sõnavaraga kõnetuvastuses.

Esialgseteks eksperimentide suhteliselt väikese hulga uudistesaadete salvestuste tuvastamisel näitasid, et kõigi kolme mudeliga saavutatakse oluline kõnetuvastuse kvaliteedi paranemine. Morfeemipõhise mudeliga saadud 10-protsendiline suhteline tähevigade vähenemine oli statistiliselt oluliselt suurem kui lemma- ja sõnapõhise mudeliga saadud muutused.

Tulevase töö ühe suunana on kavas suurendada eksperimentaalsete andmete mahtu. Samuti on kavas praegu kasutatav manuaalne uudiste segmenteerimine lauseteks ja uudisnuppudeks asendada automaatse segmenteerimissüsteemiga, mis looks realistlikuma rakendus-stsenaariumi.

Kuna uuringute käigus selgus, et morfeemid on sobivad ühikud keelemudeli adapteerimiseks, siis saab edaspidi eesti keele jaoks kasutada ka teisi levinud adapteerimismeetodeid, kus tüüpiliselt adapteerimismudeli sõnavara kattub keelemudeli sõnavaraga. Inglise keele puhul on häid tulemusi andnud PLSA (Federico 2002) ja PLDA (Tam, Schultz 2006). Eesti keele jaoks tuleks siis adapteerimisühikute kasutada sõnade asemel morfeeme, nagu ka *N*-gramm-mudelis.

Kirjandus

- Alumäe, Tanel 2006. Methods for Estonian Large Vocabulary Speech Recognition. Theses of Tallinn University of Technology. Thesis on informatics and system engineering 31. Tallinn: Tallinn University of Technology Press.
- Bellegarda, Jerome R. 1998. A multispan language modeling framework for large vocabulary speech recognition. – IEEE Transactions on Speech and Audio Processing 6 (5), 456–467.
- Deerwester, Scott; Dumais, Susan T.; Furnas George W.; Landauer, Thomas K.; Harshman, Richard 1990. Indexing by latent semantic analysis. – Journal of the American Society of Information Science 41 (6), 391–407.
- Federico, Marcello 2002. Language model adaptation through topic decomposition and MDI estimation. – Proceedings of ICASSP, Vol. 1. Orlando, 773–776.

- Kaalep, Heiki-Jaan; Muischnek, Kadri 2005. The corpora of Estonian at the University of Tartu: the current situation. – Proceedings of the Second Baltic Conference on Human Language Technologies. Tallinn, April 4-5, 2005. Tallinn: Tallinn University of Technology, 267–272.
- Kaalep, Heiki-Jaan; Vaino, Tarmo 2001. Complete morphological analysis in the linguist's toolbox. – Tõnu Seilenthal, Anu Nurk, Triinu Palo (eds). *Congressus Nonus Internationalis Fenno-Ugristarum, Pars V, Dissertationes sectionum: Linguistica. II.* Tartu, 9–16.
- Kneser, R.; Peters, J.; Klakow, D 1997. Language model adaptation using dynamic marginals. – Proceedings of Eurospeech, Vol. 4. Rhodos, 1971–1974.
- Landauer, Thomas K.; Foltz, Peter W.; Laham, Darrell 1998. Introduction to latent semantic analysis. – *Discourse Processes* 25, 259–284.
- Meister, Einar; Lasn, Jürgen; Meister, Lya 2003. Development of the Estonian SpeechDat-like database. – Proceedings of Eurospeech, Vol. 2. Genf, 1601–1604.
- Ristad, Eric Sven 1995. A natural law of succession. – Tech. Rep. TR-495-95, Computer Science Department, Princeton University.
- Siivola, Vesa; Hirsimäki, Teemu; Creutz, Mathias; Kurimo, Mikko 2003. Unlimited vocabulary speech recognition based on morphs discovered in an unsupervised manner. – Proceedings of Eurospeech. Genf, 2293–2296.
- Stolcke, Andreas 2002. SRILM – an extensible language modeling toolkit. – Proceedings of ICSLP, Vol. 2. Denver, 901–904.
- Čermák, František 2007. Some of the current problems of corpus and computational linguistics or fifteen commandments and general truths. – Plenary presentation. The 3rd Baltic Conference on Human Language Technologies. Kaunas, Lithuania, October 4-5, 2007.
- Tam, Yik-Cheung; Schultz, Tanja 2006. Unsupervised language model adaptation using latent semantic marginals. – Proceedings of Interspeech 2006. ICSLP. Pittsburgh, 2206–2209.
- Yang, Yiming 1995. Noise reduction in a statistical approach to text categorization. – Proceedings of SIGIR-95, 18th ACM International Conference on Research and Development in Information Retrieval. Seattle, 256–263.

Tanel Alumäe (Küberneetika Instituut) on lõpetanud Tallinna Tehnikaülikooli info- ja kommunikatsioonitehnoloogia erialal. Kaitses samas ülikoolis 2002. a magistr kraadi ja 2006. a doktorikraadi. Alates 2003. a töötab TTÜ Küberneetika Instituudi foneetika- ja kõnetehnoloogia laboris. Uurimisvaldkondadeks on kõnetuvastus, keeletehnoloogia, statistilised meetodid, masinõpe.
 tanel.alumae@phon.ioc.ee

STATISTICAL LANGUAGE MODEL ADAPTATION FOR ESTONIAN SPEECH RECOGNITION

Tanel Alumäe

Institute of Cybernetics at Tallinn University of Technology

This paper presents a statistical language model adaptation framework for Estonian large vocabulary speech recognition. Estonian is a highly inflected, agglutinative and compounding language. To reduce lexical variety, morphemes are used as basic units in a statistical language model. For language model adaptation, we use a small set of topic-specific sentences as an adaptation seed. Then, latent semantic analysis (LSA) is applied for finding semantically close texts from a large document corpus. The resulting adaptation corpus is used for compiling a topic-specific unigram language model for each story. The unigrams are combined with a background N -gram model using fast marginal adaptation, resulting in an adapted N -gram model. We compare words, lemmas and morphemes as basic units in the LSA model.

The method is tested on an Estonian broadcast news transcription task. In the first pass of the recognition, a general background language model is used for finding recognition hypotheses for all utterances. The hypotheses are then used as an adaptation seed to compile an adapted language model for each news story. In the second recognition pass, the adapted models are applied to find new recognition hypotheses. We observe a significant improvement in speech recognition quality after applying the adapted models. The 10% drop in letter error rate when using morpheme-based adaptation is significantly better than when using either word or lemma-based adaptation. The article also discusses some possible reasons behind this observation.

Keywords: speech recognition, language model adaptation, latent semantic analysis, fast marginal adaptation, morphemes, lemmatization

FRASEOLOGISMIDE SEMANTILISE TEISITITÕLGENDAMISE VÕIMALIKKUSEST

Anneli Baran

Ülevaade. Artiklis käsitletakse fraseoloogiat, leksikoloogiaga tihedalt seotud valdkonda. Enamasti räägitakse fraseologismide mõistmisega seonduvatest raskustest seoses mitteemakeelse kõneleja. Ent ankeetküsitlus kahe eestikeelse kooli vanemate klasside õpilaste fraseoloogiaalasest teadmusest osutas, et eeldatavalt üldtuntud väljendid – vähem või rohkem idiomaatilised fraseologismid ei pruugi seda olla keele noorema kõnelejaskonna jaoks. Küsitlus näitas ilmekalt, et fraseologismide mõtestamises-kasutamises on toimunud mõningad muutused, mis on eelkõige seotud kas väljendi tähenduse ülekantuse mittemõistmisega, st otsese tõlgendamisega või põhisõnast ajendatud assotsiatsioonide alusel teisititõlgendamisega.*

Võttesõnad: semantika, fraseoloogia, väljendid, ankeetküsitlus, eesti keel

1. Sissejuhatuseks

Fraseoloogiliste väljendite semantiline mitmepalgelisus on nähtus, mis fraseoloogiaga tegelejat kõige enam paelub. Fraseologismid ei kuulu iseenesest aga keele nende elementide hulka, mida ilmtingimata peaks tundma ja valdama (vrd sõnavaraalane vs. fraseoloogiaalane kompetents), samas on nad suurepäraselt vahendid keele ilmestamiseks. Tegeledes pikemat aega eesti kõnekäändude ja fraseologismide andmebaasis (EKFA) sisalduva materjaliga,¹ võib kergelt tekkida ettekujutus, et see aines on midagi, mis omab kohta ka tänapäeva aktiivses keelekasutuses. Tegelikult on tõenäolisem olukord, mida iseloomustab suurepäraselt tsitaat Delfi-blogist, kus arutletakse fraseologismide tõlkevastete üle:

Ja muidugi on eestikeelsed lahedad vasted ka olemas, võtame näiteks “nagu tuhat ja tuline” või “püksid sõeluvad püüli”. Oh jah, kahju kohe, et neil on selline “Eesti rahva ennemuistsete juttude” maik man. (Delfi.blog 2006)

* Artikkel on kirjutatud sihtfinantseeritava teadusteema 0032472s03 “Eesti ja teiste rahvaste folkloor: pärimus, identiteet ja globaliseerumine” raames.

¹ Tegemist on Eesti Kirjandusmuuseumi koosseisu kuuluva Eesti Rahvaluule Arhiivi laekunud materjaliga, mille digiteeritud versioon on kõigile huvilistele vabalt kättesaadav EKM-i folkloristika osakonna kodulehel aadressil <http://www.folklore.ee/justkui/> (24.09.2007).

Seega kaldutakse vanu kinnistunud väljendeid pidama millekski vanamoeliseks ning pigem välditakse nende kasutamist.

Fraseologismi mõistesse on kätketud ühe olulisema erijoonena, et tegu on n-ö keelilise valmismaterjaliga. Sellest lähtuvalt võib esitada terve rea küsimusi: kas fraseologismi tähendust tulebki mõista staatilisena; kas nn fraseologismi algne/algupärane tähendus on see tegelik ja ainuõige; kas fraseologismi tähendus võib varieeruda; kui kaugemale võib see varieeruvus ulatuda; kas tavapärasest hälbivas tähenduses mõistmine või kasutamine on valekasutus?

Soovides välja selgitada mõningate fraseoloogiliste väljendite tuntust ja tege-likku aktualisatsiooni, on kõige hõlpsam võimalus teha seda Interneti vahendu-sel – *on-line*-väljaannete kommentaarikülgedel, kodulehtedel, veebipäevikutes jm sisalduva põhjal. Kuigi vanasõnade- ja fraseologismidealase teadmuse uuringud eri keelte kõnelejaskonna seas on üldlevinud, puuduvad Eestis tänini sellelaadsed korralikud (täpsemalt laiaulatuslikud) küsitlused, ning fraseoloogide endi poolt läbiviidute kohta puuduvad andmed hoopiski. Pidades fraseoloogilisi keeleüksusi millekski, mis alluvad keelekasutuses nn automaatsele tuvastamisele ja töötlemisele, teisisõnu millekski, mis omavad kindlat kohta keelekasutaja mentaalses leksikonis, tuleb arvestada, et tegelikult on uuringud selles valdkonnas veel napid ja alles päris viimasel ajal hoogustumas.

Otsustasin küsitluse viia läbi nooremate keelekasutajate – gümnaasiumiõpi-laste – seas. Vanusemäära tingis vanasõnaurijate – peaaesjalikult saksakeelses keskkonnas tegutsevate – poolt kindlaks tehtud tõsiasi, et just vanuseks 16–18 on inimene omandanud vanasõnadealase teadmuse (Baur, Chlosta 1994: 3). Seega lähtutakse inimese keelilise ja kognitiivse võimekuse aspektist. Hõlbus oleks laiendada seda teesi ka fraseologismidele, sest fraseoloogiat laiemas tähenduses käsitledes kuuluvad siia ka vanasõnadega tihedas seoses olevad kõnekäänud (lähemalt vt Baran 1999 ja Õim 2004). Küsitlusega soovisingi saada muuhulgas vastust vanuselise aspekti paikapidavusele, kaldudes küll arvama, et pigem kujuneb mingitasemeline fraseoloogiaalane teadmus välja siiski varem, sõltudes rohkem iga konkreetse keelekasutaja taustast. V. Jenenšeki erinevate keelte peal tehtud uuringud näitavad, kuidas juba lasteaiaaegsed lapsed on suutelised fraseologisme mõistma ja kasutama (Jenenšek 2006: 138). Tegemist on lihtsalt keeleüksustega, mida omandatakse pidevalt, st see protsess ei lõppe mitte teataval vanuseastmel, vaid on dünaamiline. Seetõttu on selge, et nooremad keelekasutajad tunnevad tunduvalt vähem fraseologisme kui vanemad.² Samas tuleb ka meeles pidada, et fraseologismide tundmisel on omad erijooned.

Küsitluse eesmärgiks oli eelkõige proovida välja selgitada, kas kinnistunud keelendid omavad üldse kohta tänaste kooliõpilaste sõnavaras (leksikaalases teadmuses) ja kui omavad, siis kuidas on muutunud nende väljendite kasutus. See omakorda annaks aimu küsitletavate nn taustteadmistest ehk entsüklopeedilisest teadmusest. Fraseologismid on oma olemuselt selgelt mitmetähenduslikud. Küsi-mus ei ole aga üksnes keeleüksuse tajumises teljel mittekujundlik–piltlik/kujundlik. Kinnistunud väljendite kontiinum on astmeline (või siis radiaalne, st tuuma ja perifeeriga üksus), st nad on ka erineval määral motiveeritud ning sellest tingituna erinevalt mõistetavad/tõlgendatavad ning erinevate keelekasutajate jaoks erineval määral läbipaistvad. Eeldatavalt peaksid küll teatud fraseoloogilised väljendid

² Suurepäraseks näiteks väikesearvulise kasutajaskonnaga keele fraseoloogia alalt on Natalia Filatkina monograafia, kes on sarnaselt eesti keelele väikese kõnelejaskonnaga, nimelt letseburgi fraseoloogiat uurides läbi viinud ka põhjalikke küsitlusi, sh koolinoorte hulgas, ning on sedastanud, et fraseologismide tundmine ja tegelik kasutamine ei ole päris üksüheses suhtes, st neid fraseologisme, mida tuntakse, on märkimisväärselt rohkem kui on kombeks reaalselt kasutada (Filatkina 2005: 89).

(tuuma esindavad elik tuntumad) alluma hõlpsamalt intuiitiivsele tõlgendamisele kui teised.³

Nii Dmitri Dobrovolski (1997: 46 jj) kui Eckhard Roos (2001: 269 jj) on osutanud asjaolule, et nende väljendite, mille otsene tõlgendus on vastuolus kuulaja teadmusega, tõlgendus välistatakse. Sellised ütlused on tugeva piltlikkusega ning enamasti vastuolus tegelikkusega (või meie teadmistega sellest tegelikkusest) ning läheneb absurdile. Seega nad ei ole semantiliselt mitmetähenduslikud, kuna ei võimalda kahesugust tõlgendust.

Viimasel ajal kohtab üha sagedamini fraseoloogiaalastes sõnavõttudes ja kirjatükkides mõistet *esilduvus* (ingl *salience*), veelgi täpsemalt astmeline esilduvus (ingl *graded salience*), ning seda seoses fraseologismi semantikaga (vt Hallsteinsdóttir, Farø 2006). Selle kohaselt ei aktiveeru kinnistunud keelendi otsese tähenduse kõrval veel teinegi analoogne tegevus, milleks on fraseoloogilise tähenduse aktiveerimine mentaalses leksikonis, vaid esmalt aktiveeritakse teada olev ehk tuntud tähendus (ning seda sõltumata kontekstist). Ning alles seejärel, kui see tähendus ei sobi, asutakse konstrueerima teisi tähendusi. Salientsuse-teooria eestkõneleja Rachel Giora on rõhutanud idioomide puhul just asjaolu, et nende idiomaatiline tähendus ongi salientne tähendus ning aktiveeritakse kiiremini kui muud võimalikud tähendused (Giora 2003: 19).

Tegelikult ei ole sellises lähenemises midagi uut, sest juba aastaid varem on Raymond Gibbs oma psühholingvistiliste eksperimentide tulemusena sedastanud, et inimestel ei kulu idioomide, kaudse kõne, sarkasmi ja metafoori kujundliku tõlgenduse mõistmiseks/töötlemiseks mitte sugugi rohkem aega (ning mõningatel juhtudel isegi vähem) kui otsesest tähendusest arusaamiseks. Kujundliku keelekasutuse korral ei eelne tõlgendamisele mingit sellesama lausungi otsese tähenduse analüüsi etappi. R. Gibbs järeldas ka, et taolist käitlemist toetab kontekst, mis toob vastavad tähendused esile ehk muudab salientseks. Kusjuures konteksti all ei pea R. Gibbs silmas mitte üksnes kitsalt keelelist, n-ö tüüpilist kasutuskonteksti, vaid kontekstuaalset raamistikku entsüklopeedilise teadmuse näol (Gibbs 1989: 247 jj).

2. Küsitlus eesti koolinoorte seas

Selgitamaks välja just vanimate, n-ö põlisteks puhteestilisteks peetavate fraseoloogiliste väljendite tegelikku tuntuust, võtsin küsitlusse väga vanad väljendid – varaseimad fikseeringud 18. sajandi algusest (paari erandiga hilisemast ajast). Valikut raskendas EKFA-s sisalduva materjali rohkus ning ligi 170 000 teksti seast 35 väljendi selitamine ei ole n-ö kergete ülesannete killast. Valikul sai määravaks ühelt poolt väljendi ere kujundlikkus (subjektiivne kriteerium) ja teisalt populaarsus läbi aegade (objektiivne kriteerium), sellest peaks andma tunnistust tema esinemus nii EKFA-s kui ka fraseoloogiaasõnaraamatus (FS), kusjuures EKFA esinemuse puhul olid kriteeriumiks kõrge üleskirjutuste arv ja lai geograafiline esinemus mitte ainult 18.-19. sajandi, vaid ka 20. sajandi lõpul (v.a kaks väljendit: retooriline küsimus *Kas oled näinud kellaga siga?*, mis esineb peamiselt Lõuna-Eestis, ning väljend *üks silm vaatab tasku, teine taeva*, mis tuleb ette üksnes Saaremaal (laiemalt on tuntud

³ Kognitivistlikust vaatevinklist on igasugused väljendid vähem või rohkem läbipaistvad ning täiesti läbipaistmatuid väljendeid ei leidugi, sest alati on võimalik tõlgendamisel tugineda konkretiseerimis/kategoriseerimisvõimele (Lakoff 1987: 451).

üks silm vaatab kirikutorni, teine lihakarni ja üks silm vaatab silku, teine leiba)). Enamik küsitud väljendite varaseimatest allikatest, mis on esindatud EKFA-s, on ühtlasi eesti kirjakeele ajaloo uurimise olulisemad allikaid: A. T. Helle “Kurtzgefasste Anweisung zur Ehstnischen Sprache..” (1732), A. W. Hupeli “Ehstnische Sprachlehre für beide Hauptdialekte..” (1780), J. F. Helleri käsikirjaline (lõuna)eesti-saksa sõnaraamat (koostatud tõenäoliselt ajavahemikus 1823–1841), S. H. Vestringi eesti-saksa käsikirjaline sõnaraamat “Lexikon Esthónico Germanicum” (1720-ndad), F. J. Wiedemanni koguteos “Aus dem inneren und äusseren Leben der Ehsten” (1876), J. V. Jannseni koostatud Sannumetoja I–VII (1848–1860).

Miks aga peaksid vanad väljendid üldse veel aktiivses kasutuses olema? Siinkohal jagan seisukohta taani fraseoloog Ken Farøga, kes rõhutab tõsiasja, et fraseoloogilisi väljendeid ei kasutata mitte niivõrd kõnestrateegilistel eesmärkidel, vaid pigem seepärast, et nad sobivad kontekstiga või on mnemotehnilistel põhjustel teadliku kõneleja jaoks hõlpsasti haaratavad (Farø 2006: 60).

Küsitlavateks olid kahe gümnaasiumi õpilased vanuses 16–18, küsitluid kokku 147: Kuressaare Gümnaasiumis (KG) (X ja XII kl) 86: 33 noormeest (M) ja 53 neidu (N); Tartu Miina Härma Gümnaasiumis (MHG) 61 (X ja XII kl): 23 noormeest ja 36 neidu, 2 isikuandmeteta. Küsitlus oli anonüümne, oluliseks pidasin soolist ja vanuselist aspekti, mistõttu palusin küsitlavatel kindlasti märkida oma sugu ja vanus. Kaks Eestimaa eri otsas paiknevat kooli (kuhu tullakse ka mujalt maakonnast või ka – nagu Tartu kooli puhul – teistest lähematest maakondadest) sai valitud eesmärgiga välja selgitada murdetaustalt (kuivõrd selline tegur tänases keelesituatsioonis üldse rolli mängib) võimalikult erinevate keele kõnelejaskondade tegelikke erinevusi fraseoloogilise ainese tundmise osas. Seetõttu saigi küsitlusse lülitatud ka eelpool mainitud väljendid, mis kõnekäändude ja fraseologismide andmebaasi põhjal omavad peaasjalikult piirkondlikku levikut. Samuti pidasin küsimustikku koostades oluliseks mitte piirduda pelgalt suletud tüüpi jah/ei-küsimusega (tunnen / ei tunne), vaid jaatava vastuse korral soovisin ka lähemat selgitust fraseologismi tähenduse ja kasutamissituatsiooni kohta (avatud küsimus).

Küsitluses esitatud väljendite omaaegset tuntuust ja tegelikku tundmist illustreerib tabel 1, kus võib näha kõrvutatavalt kahesuguseid andmeid: väljendi järel teises veerus on sellesama väljendi koht Arvo Krikmanni poolt koostatud loendis “100 tuntumat eesti kõnekäändu” (Krikmann 1997) – aluseks üleskirjutuste hulk EKFA-s, kolmandas veerus sisaldub küsitlusega laekunud jaatavate vastuste üldarv ning sulgudes vastused kooliti.

Tabel 1. Küsitletud fraseologismide nimistu nende tuntuuse alusel

Fraseologism	Koht A. Krikmanni loendis	Jaatavate vastuste üldarv (KG + MHG)
<i>teeb sääsest elevandi</i>	52.	147 (86 + 60)
<i>ei lausu musta ega valget</i>	–	137 (79 + 58)
<i>süda kukkus saapasäärde</i>	95.	131 (80 + 51)
<i>elavad nagu koer ja kass</i>	99.	127 (75 + 52)
<i>nüüd on vesi ahjus</i>	13.	117 (70 + 47)
<i>vaatab läbi sõrmede</i>	–	114 (71 + 43)
<i>ei näe sõrmegi suhu pista</i>	–	109 (63 + 46)

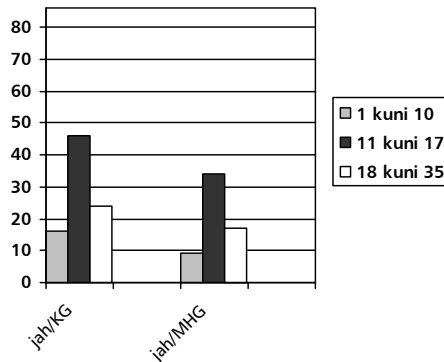
<i>käib nagu kass ümber palava pudru</i>	31.	102 (61 + 41)
<i>nagu juudi jõulupuu</i>	–	102 (55 + 47)
<i>hunt lambanahas</i>	–	100 (62 + 38)
<i>peenike peos</i>	86.	99 (58 + 41)
<i>üks jalg hauas, teine haua äärel</i>	11.	96 (57 + 39)
<i>elab peost suhu</i>	–	91 (54 + 37)
<i>see on hoopis teine tubakas</i>	–	83 (49 + 34)
<i>vorst vorsti vastu</i>	–	81 (42 + 39)
<i>ega sa klaassepa poeg ole</i>	39.	73 (46 + 27)
<i>kõrvatagused alles märjad</i>	84.	67 (35 + 42)
<i>must kass on vahelt läbi jooksnud</i>	83.	55 (35 + 20)
<i>pane või hambad varna</i>	18.	52 (34 + 18)
<i>segi nagu Kört-Pärtli särk</i>	–	42 (14 + 28)
<i>püksid sõeluvad püüli</i>	33.	41 (23 + 18)
<i>üks silm vaatab tasku, teine taeva</i>	38.	35 (23 + 12)
<i>mitu tuult jalge all</i>	–	32 (22 + 10)
<i>ega mu suu pole seinapragu</i>	60.	31 (24 + 7)
<i>tulid nagu tuld tooma</i>	52.	21 (18 + 3)
<i>küll tema teab, kus kivi all vähid on</i>	–	19 (12 + 7)
<i>koergi ei haugu takka</i>	–	14 (9 + 5)
<i>suurte kõrvadega supilontrus</i>	–	13 (9 + 4)
<i>lammas läheb vähi käest villa küsima</i>	21.	13 (9 + 4)
<i>mõni mõis või ahjutäis vorste</i>	–	12 (9 + 3)
<i>ei vea villast lõngagi katki</i>	–	12 (2 + 10)
<i>nagu hobuse unenägu</i>	–	11 (9 + 2)
<i>nüüd on konn mäta otsas</i>	9.	11 (7 + 4)
<i>võta silm näppu</i>	–	6 (2 + 4)
<i>kas oled näinud kellaga siga?</i>	–	0

Kuna küsitlus oli piiratud ajaliselt ühe koolitunniga, esitasid õpilased pigem tähenduste seletused kui kasutuse iseloomustused. Harvadel juhtudel täpsustati: *ise ma seda ei kasuta, kasutan harva, kasutan siis kui...* Sellised metakommunikatiivsed kommentaarid on eelkõige põhjendused mitte/harvakasutamise kohta.

Märkimisväärseid teadmiserinevusi koolide vahel välja ei tulnud, sh murdelise tagapõhjaga väljendite juures, mistõttu saab pigem rääkida ühistest joontest ning suundumustest. Samas tuleb arvestada, et tegemist on esialgse katsega selles vallas ning järeldustesse tuleb suhtuda teatava mõõndusega. Siiski saab välja tuua mõningad arvud, joonisel 1 on esitatud jaatavate vastuste suhe kahe kooli võrdluses.

Kuressaare Gümnaasiumi 86-st õpilasest vastasid jaatavalt – st tunnen küsitud fraseologismi – rohkem kui pooltele 24 õpilast, Miina Härma Gümnaasiumi 61-st 17 õpilast. *Jah*-vastuseid anti kõige enam 11–17 korral. Keskmiselt tunti 14 (KG-s

14,1, MHG-s 14,9) väljendit 35-st, mis teeb alla poole väljenditest. N-ö parimaid tundjaid (20–26 fraseologismi 35-st) oli KG-s 15 (12N + 3M), MHG-s 7 (5N + 2M), seega ühes napp kolmandik, teises veelgi vähem.



Joonis 1. Jaatavate vastuste suhe koolide võrdluses

Jaatavate-eitavate vastuste vahepeale jäid minu poolt küsimärgistatud vastused, mida analüüsingi lähemalt järgnevas peatükis. Selliste n-ö ärastpidiste tähendussetuste koguarv ei ole küll suur: MHG 2135 vastusest 102 ja KG 3010 vastusest 113. Keskmiselt teeb see mõlema kooli puhul 2,9–3,2 väljendit 35-st. Ent just sellised vastused huvitavad fraseoloogi iseäranis, sest näitavad, kuidas mõtestatakse väljendit juhul, kui seda tegelikult ei tunta. Huvitava iseärasusena kerkib esile tõsiasi, et enam andsid selliseid vastuseid noormehed, keda vastajate seas oli tütarlastest vähem.

Järgnevas võtan vaatluse alla just sellised tähendussetused, mis hälbivad mõnevõrra fraseologismi tavatähendusest. Fraseologismi põhitähendusena esitan fraseoloogiasõnaraamatus (FS) fikseeritu või seal mitteesinevuse või mõnevõrra erineva esinemise korral eesti kõnekäändude ja fraseologismide andmebaasis (EKFA) sisalduva.

3. Fraseologismide tähenduste seletused

3.1. Väljendi seletamine teise väljendiga

Küsitlusest ilmnes, et mõnede fraseologismide tähenduse seletamiseks kasutatakse teist väljendit ehk kujundlikku ütlemist. Tegemist on psühholingvistiliselt huvitava seigaga – ühe väljendiga seostub teine samas/lähedases tähenduses kasutatav väljend või värvikas üksiksõna. Sellist omalaadset koosinemist võib vaadelda kui mentaalset võrgustikku, mis seostub ühe ja sama kontseptuaalse/mõistelise valdkonnaga või milles üks väljend toimib mõjurina, kutsudes esile/aktiveerides muu selle nn mõjuriga mälus seonduva skriptina määratletava ilmingu.⁴ Dmitri Dobrovolski (1997: 36) on osutanud asjaolule, et väljendi töötlemist ei määragi mitte omandatud teadmised, vaid teatud kognitiivsed süvastruktuurid.⁵ Sageli võib selliste seletavate väljendite puhul täheldada ka struktuurilist sarnasust. Järgmise

⁴ Kognitivistikast pärinev mõiste *skript* (nt Wierzbicka 1999) on keskne huumoriteoorias (Victor Raskini skriptipõhine naljateooria), Eestis on selle (stiimul) sõna ümbritsevat semantilist infot märkiva termini kohta põhjalikult kirjutanud A. Krikmann monograafias "Koestler, Raskin, Attardo ja teised: lingvistiliste huumoriteooriate uuemaist areguist" (2004).

⁵ D. Dobrovolski peab siinkohal silmas varieerimist analoogia alusel, mitte normi järgimist.

viie väljendi puhul kasutati küsitluses kõige sagedamini tähenduse lahtiseletamisel teist samatähenduslikku väljendit.

- (1) **segi nagu Kört-Pärtli särk** 'kõik on pilla-palla laiali', 'suur segapudru', 'tohuvabohu' (FS 'segamini, sassis')
- (2) **ei lausu musta ega valget** 'vait kui sukk', 'ei ütle head ega halba', 'ei ütle ei nii ega naa' (FS 'mitte ühtegi sõna, mitte midagi ütleva')
- (3) **ei näe sõrmegi suhu pista** 'kottpime', 'pilkane pime', 'pime kui öö' (FS 'väga pime')
- (4) **üks jalg hauas, teine haua äärel** 'vaagub elu ja surma vahel', 'hinge heitmas', 'surmasuus', 'ühe jalaga hauas', 'kõigub elu ja surma vahel', 'muldvana', 'nagu noateral' (FS 'õige vana, kõrges eas, surmale lähedal')
- (5) **nüüd on vesi ahjus** 'pill on lahti', 'omadega plindris', 'karistus soolas', 'häda kaelas' (FS 'kellelgi on häda käes, keegi on täbarasse olukorda sattunud')

3.2. Väljendi mõistmine üksnes otseses tähenduses

Enim huvipakkuvateks pean fraseoloogina siiski küsitluses esitatud nn valesid seletusi. Nende puhul tuleb arvestada, et väljendi tegeliku mittetundmise korral (st vastates jaatavalt) võidi üritada seda tõlgendada ehk remotiveerida, kas siis otseselt või mingite assotsiatsioonide alusel. Fraseologismi puhul on tähtsaim semantiline aspekt, täpsemalt fraseoloogilise tähenduse ja otsese tähenduse vahekord. Kahe tähenduse vastukäivus, ühtesobimatus, vastuolu tingib väljendi (semantilise) idiomaatilisuse. Ilmneb, et seda idiomaatilisust ei pruugita enam tajuda elik ära tunda, st väljendid võivad keelekasutaja jaoks olla läbipaistmatud ning väljendi mõtestamine võetakse ette analüütiliselt. Seega on väljendi puhul esilduvaks tähenduseks otsene tähendus.

- (6) **vaatab läbi sõrmede** 'ei taha hästi vaadata', 'hirmus, hirm', 'piilub', 'kartlik' (FS 'millelegi taunitavale reageerimata jätma')
- (7) **püksid sõeluvad püüli** 'kulunud püksid', 'pükstes nii palju auke, katki', 'katkised püksid' (FS 'kartma')
- (8) **ega sa klaassepa poeg ole** 'ega sa õrnake pole', 'ettevaatlik', 'teeb kõike väga õrnalt ja ettevaatlikult', 'purunema', 'õrnake, pussy' (FS 'öeldakse valguse ees seisjale')
- (9) **kõrvatagused alles märjad** 'on ujumas käinud', 'alles sündinud', st 'kiirustamine' (FS 'noor ja kogenematu')
- (10) **ei näe sõrmegi suhu pista** 'ta on pime' (FS 'väga pime')
- (11) **peenike peos** 'midagi on peos' (FS 'raskes olukorras')

3.3. Väljendi tähendus on teisenenud

Järgneva neljateistkümne väljendi puhul joonistus samalaadsete vastuste arvukuse põhjal selgelt välja, et tähenduses on toimunud semantiline nihe. Need on väljendid, mille tähendus n-ö kumab läbi (osaliselt või täielikult motiveeritud), mistõttu tõlgendamisel tuginetakse salientsele otsesele tähendusele või assotsiatiivsetele seostele. Erinevalt eelmise rühma väljenditest võib siia paigutatuid vaadelda kui tavatähendusest mõnevõrra hälbivaid tähendusvariante (milledest mõned on tavatähendusele lähemal kui teised; selle vahemaa hindamine on aga loomulikult suuresti subjektiivne).

Võimaluse korral kaasan näitestikku ka küsitlusse võetud väljendiga samas või lähedases tähenduses kasutatud Interneti-näiteid, seda peamiselt eesmärgiga näidata selliste väljendite aktuaalsust keele väljenduslikumaks muutmise vahendina ka tänapäeval (keelenäited pärinevad viimasest viiest aastast). Nagu näidetest näha, võidakse fraseologismidega ringi käia üsnagi vabalt. Samas võiks oskust käidelda oma olemuselt klišeelist fraseologismi märglevalt kui näidet keele suurepärasest valdamisest.

(12) ***ei vea villast lõngagi katki***

FS –; EKFA 'laisk'

'nõrk', 'pole jõudu', 'ei ole tugev', 'õrn', 'ei ole võimeline kellelegi liiga tegema', 'saamatu millegi suhtes'

Kooliõpilaste seletused tuginevad otseselt puhtfüüsilisel omadusel – keegi on sedavõrd jõuetu, et ei ole võimeline isegi kaalult nii kerget asja nagu lõng katki tegema.

(13) ***must kass on vahelt läbi jooksnud***

FS 'kellegi vahel on lahkkelid, keegi on omavahel tülli või pahuksisse läinud'

Kooliõpilaste tähenduseseletuste aluseks on pigem uskumusest väljakasvanud nn argiarvamus – kui must kass üle tee läheb, siis tuleb õnnetus.⁶

'must kass toob paljude arvates halba õnne, kasutatakse kui tüllis on keegi', 'kõik läheb halvasti', 'on õnnetus juhtunud', 'kõik läheb mingil põhjusel untsu', 'häda ja tülli'

Tülli võibolla ei tulegi, aga must kass jookseb meie vahelt läbi küll. (Pere ja Kodu 2007)

(14) ***ega mu suu pole seinapragu***

FS 'keegi tahab ka midagi maitsta, keegi tunneb ka millestki maitsvast rõõmu'

Tähenduseseletuste aluseks on pigem suu kui kõnelemiselund. Sellest on omakorda visualiseerides tuletatud kahetised tähendused: võrdluses suuga on seinapragu avausena miski, mis on a) pidevalt lahti nii nagu inimesel, kes ei suuda suud kinni pidada, või siis hoopis vastupidiselt b) miski, kust heli väljatulek on välistatud/piiratud:

⁶ Vrd ka saksa fraseologismiga *da ist die schwarze Katze zwischen gekommen* 'will eine Sache gar nicht gedenken, so sagt man deshalb' (Röhrich 1973). [Tõlkes *seal on must kass vahele tulnud* 'öeldakse siis, kui asi ei taha kuidagi edeneda'.]

- a) 'lobamokk', 'mul on palju öelda', 'julgeb rääkida', 'pole vastusega kitsi', 'pole mõtet peale asja halvaksminekut mõliseda, kui enam midagi nagunii päästa ei anna';
- b) 'ega ta midagi välja ei räägi, usaldusväärne', 'ega ta jutukas pole', 'ega ma niisama tühja juttu ei aja', 'kõike ei räägita, rumalusi näiteks', 'seinapragu ei panda tähele ja seda ütleb tavaliselt inimene, kes tunneb, et temaga ei arvestata ega panda tähele'.

Korterite teema kohapealt peaks Atoneni suu küll seinapragu olema.
(Postimehe kommentaarikeskus 2004)

Imbetsill postita vähem, sinu suu on siin foorumis seinapragu, saa sellest lõpuks aru, sa seletad liiaga palju. (Freetime foorum 2007)

(15) **pane või hambad varna**

FS 'nälga jääma, nälgima'

Väljendit mõistetakse kui puhkama asumist või hoopiski millestki eksistentsiks vajalikust loobumist; viimasel juhul on tulemus traditsioonilise tähendusega võrreldes kaugemaleulatavam, tähistades pigem paratamatut lõppu.

'sure ära', 'tööd lõpetama, pooleli jätma midagi, surema', 'ehk lõpeta toiming, võrreldakse lõpu ehk surmaga', 'haige hakkab surema', 'puhkama hakkamise plaan', 'sure või maha', 'väga vana', 'süüa enam ei ole, järelikult ei ole hammastel ka muud teha kui puhata (varnas)', 'ehk lõpeta toiming (võrreldakse lõpu ehk surmaga)'

Kui teha paar visiiti kihvatoetri juurde, ongi kõik. Millest siis aga elada? Ega jää muud üle, kui hambad varna ja otse hauda.!!! (Kesknädal 2007)

(16) **nagu hobuse unenägu**

FS 'mõttetus, absurdus'; EKFA 'vilets asi; rumal või inetu või pikaldane inimene; segane asi'

Antud väljendi puhul torkab silma, et sõnaraamatu tähendus on tunduvalt üldistatum kui EKFA-s fikseeritu. Samas näitavad nii kooliõpilaste tõlgendused kui Interneti kasutusjuhud, et pigem on tähendus analoogne EKFA-s leiduvaga:

- a) 'suur uni', 'võid öelda inimese kohta, kes sulle kogu aeg järele käib', 'midagi mis on väga pikk, igav ning veniv', 'midagi väga pikka';
- b) 'mingi kirju situatsioon, tegevusterohke'.

Ja õhu teekond mootoriruumis on pikk nagu hobuse unenägu. (Saabnet 2006)

Paratamatult tekib aga küsimus, milleks siis haridus üldse on, kui ametite loomulik hierarhia on segi nagu hobuse unenägu. (Delfi 2007)

(17) **tulid nagu tuld tooma**

FS 'korraks, lühikeseks ajaks kellegi poolt läbi astuma, kedagi väga põgusalt külastama'

Seletuste aluseks on hoopiski tulega assotsieeruv tulekahju ning tuletoomise tegevuse all mõistetakse tule külge panemist, st hädaohtu kellegi poolt (a). Samas leidus üllataval kombel ka täiesti vastupidises, st positiivses tähenduses tõlgendusi (tuli kui valgus) (b):

- a) 'tüli norima', 'tulid tülitsema', 'riidu norima', 'tulid nagu pahandust/õnnetust tooma/tegema', 'tulid vihaselt',

Heameelt teeb seegi, et minister ei tulnud vaid tuld tooma ja poliitloengut pidama, vaid leidis aega paar tundi kohalikku elu vaadata. (EPL 2007);

- b) 'tulid head tooma', 'päästsid halvast'.

(18) **süda kukkus saapasäärde**

FS 'keegi hakkab kartma, lööb araks, kellelgi kaob julgus'

'süda murdus, kurb', 'inimene muutus õnnetuks', 'pettuma', 'langes suurde masendusse', 'hakkas häbi', 'šokiseisund', 'süda heldis'

Järjekordset massiivset ja argumenteerimata linnavaramüüki põhjendades ütles abilinnapea Aaviksoo, et toetas seda raske südamega. Uskuge mind, kõigile on näha, et süda on juba nii raske, et on vajunud saapasäärde. (Palts 2005)

Näide 18 on sellest, kuidas mõjus piltlik väljend võib omandada tähendusi, milles liigutakse mõisteliselt hoopis teistesse valdkondadesse (samas mitte ka ülemäära kaugelseisvatesse). Nii on siin tegu peamiselt inimese hingelise seisundi kirjeldamisega. Ka Internetist pärinev näide osutab ilmekalt, et süda võib liikuda allapoole vaid raskuse mõjul; raskelt mõjuv, rõhutud, löödud hoiak on aga seostatav pigem kurbusega. Tegemist on Lakoff-Johnsoni kõnepruugis orientatsioonimetafooriga: ÖNNELIK ON ÜLES, ÖNNETU/KURB ON ALLA (Lakoff, Johnson 1980: 15–16).

(19) **suurte kõrvadega supilontrus**

EKFA 'poisikeste omavaheline tõgamissõna'; EKSS 'kirumissõna'

Tähenduseseletustes lähtutakse välimusest ning mõtestamise järgmine arenemisjärk, milleks on üldistav edasiarendus, puudub:

- a) 'no lihtsalt kole', 'nii loll, et sarnaneb supikausile', 'tegemist on väga suure lontrusega', 'kõrvad välja veninud', 'kui väga suured kõrvad, et saaks suppi süüa';
- b) 'kuuleb kõike pealt, ka seda mida poleks vaja', 'kuuleb kogu aeg asju pealt, võib jutte levitada'.

Erinevalt kooliõpilaste antud seletustest, ei ole Interneti-näited mitte inimkesksed, vaid väljendit kasutatakse pigem looma, täpsemalt koera kohta. Ka slängisõnaraamatust (EESS) leiab väljendi märksõna *koer alt*.

Üks supilontrus loivab sulle padja peale (muideks nad just hommikul õppisid kõndimise ära!) ning sikutab end küünte abiga su otsaesisele magama. Nii me siis magasime terve öö otsa. (Ussipesa 2005)

Isiklikult kassid ei sümpatiseeri aga koeri armastan küll – endalgi üks spanjelist supilontrus. (Keskmaa 2004)

Supilontrus 'supisõbralik peni (eriti asjakohane ütlus, kui isend juhtub olema sündinud suurte laperdavate kõrvadega, mis iseenesest mõista alatasa supi sees tolgendavad)' (EESS)

(20) **hunt lambanahas**

FS 'malbusega oma kurje kavatsusi varjav inimene, vagurust teesklev kuri inimene'

Nii mitmegi seletuse taustaks on sümbolikeskne mõtlemine lambast kui arast loomast (a). Samas kui hundile omistatakse julgust, mis ei ole sugugi selle looma n-ö põhiomadus (pigem on selleks jultumus ehk häbematu julgus, mis on ka antud väljendi aluseks); võidakse ka – tõenäoliselt seda hästi tundvana – tavapärasest tähendusest originaalselt edasi arendada (b), sama kohtab ka Interneti-näidete puhul:

- a) 'hirm', 'arglik'; 'tegelikult on julge isik, aga näitab vastupidist', 'väljast nõrk, seest tugev', 'teeb, et on julge, aga tegelikult on arg', 'julge, aga pole mille üle uhkustada';
- b) 'mingi eriti tugev või lihtsalt kõva mees lööb araks', 'välimus tagasihoidlik, sisemus karm', 'näeb ilus välja, aga tegelikult sitt iseloom'.

Saatuslikuks saab hunt lambanahas, kus kiskja iseloom lõpuks peale jääb. Kui koalitsioonilepe ka vormistatakse, saab sellele saatuslikuks ikkagi Laari leppimatus temale antud positsiooniga. (PM kommentaarikeskus 2007)

Itaallastest rääkides pakuks sellist hunt lambanahas valem – võtad mittemidagiütleva kerega alfetta põhjale ehitatud sedaani ning varustada selle Alfa 24V 3,0 V6-ga (tulemus – imehea kaalujaotusega võimas ning üsna tuima välimusega isend). (Auto24 2006)

Eriti tore on see, et neile paigaldatakse ls1(corvette) mootorit, mõnus hunt lambanahas. (Streetrace 2007)

(21) **nüüd on konn mäta otsas**

EKFA 'uhke; suurustaja; vaese inimese kohta, kes on saanud teiselt varanduse, millega jõukalt elab; kui vihamees on saanud mõnele juhtivale kohale'

FS *nagu konn mättal* 'oma positsiooniga uhkeldav, suurustav inimene'

Küsitud väljend sellisel kujul fraseoloogiasõnaraamatus puudub, tõenäoliselt ongi rohkem tuntud võrdlusevormilisena (mida osutasid ka Interneti kasutusjuhud). Ka väljendi vähene tuntus kooliõpilaste seas on arvatavasti sellega seletatav. Nagu mõningatest õpilaste seletustest nähtub, võidakse väljendit mõista ka positiivses tähenduses – kõrgemale kohale (= positsioonile) jõudmine kui märk edenemisest, olukorra muutumine paremuse suunas (a). Sellisel juhul on äratuntav selle taga olev orientatsioonimetafoor HEA ON ÜLAL / HALB ON ALL. Teisalt aga antakse ka täiesti vastupidine seletus, mis sarnaselt fraseoloogiasõnaraamatu tähendusele on küll negatiivselt märgistatud, ent käib pigem olukorra kui inimese kohta (b):

- a) 'isik, kes tunneb ennast kuskil hästi ja turvaliselt', 'kõik on korras, valmis', 'eesmärk saavutatud';

Peale väheste visionääride, kes oma ambitsioone liigselt ei reklaami, on suurel osal meist kombeks elada omi asju ajades nagu konnad mättal, kes, tundes küll hästi oma lompi, on kogu maastiku suhtes üsna ükskõiksed. (Diplomaatia 2005)

On tavaline, et see, kes oma raha rahandusministrilt või Riigikogult on välja võidelnud, istub nagu konn mäta otsas ja peab seda raha enda omaks. (Kesknädal 2002)

- b) 'konn kuival', 'häda majas'.

(22) **mitu tuult jalge all**

FS 'kellelgi on palju tegemisi, ettevõtmisi, sündmusi ees'

Suure hulga tõlgenduste aluseks on pigem ettekujutus sihitust ja kiirest liikumisest (vrd *kadus kui tuul*) (a). Mõned ülekantud tähendused (b) näivad tuginevat assotsiatiivsetel seostel muude väljenditega, nagu *kust tuul, sealt meel*.⁷ Omaette tähendusena eristasin seletuse, milles tuuakse sisse muutusi märkiva *uute tuulte* n-ö edasiarendusena *paremad tuuled* (c). Sõnaühend *mitu tuult* on aga üsna tõenäoliselt juhtinud seletusi, mis seonduvad hoopiski väljendiga *mitu rauda tules* 'mitut võimalust alal hoidma' (d):

- a) 'põgenema', 'jookseb hästi kiiresti', 'väga kiire', 'ei suuda paigal olla, peab kogu aeg liikuma', 'reisib pidevalt', 'pidevas reisimises';
b) 'käib igal pool vastavalt sellele, kuidas tuju on', 'mitu mõtet, mida teha, aga ei ole veel otsustatud', 'palju mõtteid peas';
c) 'läheb sinna, kus paremad tuuled';
d) 'mitmes kohas käsi sees', 'on seotud mitme asjaga, nt mees semmib kahega korraga', 'käib mitme eidega'.

(23) **koergi ei haugu takka**

FS 'kellelgi pole juhtunuga asja, keegi ei päri juhtunu kohta, juhtunu ei lähe kellelegi korda'

Erinevalt tavatähendusest on kooliõpilaste seletustes sihtmärgiks pigem elus objekt, st kellegi vastu puudub huvi, keegi ei lähe korda sedavõrd, et isegi koer ei haugu ta peale.

'kellestki ei tunta puudust', 'eriti mõttetu mees'

Internetis leiduvad näited käivad tavapäraselt sündmuse kohta, mis on aset leidnud.

Prükkari surma üle ei haugu ka koer. (TOM 2002)

Kas too vald on omaette riik riigis, teevad, mis ise heaks arvavad, koergi ei haugu (politseist rääkimata). (Virumaa Teataja 2005)

(24) **elab peost suhu**

FS 'vaeselt, tagavaradeta, saadut kohe ära kasutades elama'

Väljendi tähendus avatakse otseselt, st keskne koht on toidul, mida kas suudetakse ise hankida (a) või saadakse seda teistelt (= teiste käest) (b). Antud väljendi seletuste puhul torkas silma ka teatav filosoofiline käsitlus (täna-homme, olevik-tulevik) (c), mis ülejäänud fraseologismide puhul niivõrd eredalt esile ei tõusnud. Üllataval kombel võib väljendil olla ka positiivne konnotatsioon – keegi ei ela mitte raisates, vaid hoopiski säästes (d):

- a) 'mis peos, see kohe söögiks', 'elab vaeselt, raha on ainult toidu jaoks';
- b) 'vaene inimene, vara pole, süüa saab, kui teised annavad', 'elab väga vaeselt tänu teiste annetustele', 'tehakse kõik ette-taha ära', 'elab teiste kulul', 'elab teiste almustest';
- c) 'elab üks päev korraga ja ei hooli homsest päevast', 'inimene ei planeeri oma tulevikku, vaid teeb nt juhutöid, et ära elada', 'palgapäevast palgapäevani';
- d) 'ei saa rahaga priisata', 'kokkuhoidlikult elab'.

Mis on vahepeal toimunud? Suurt midagi, elu nagu ikka: mitte peost suhu, aga peost peoni (soolaleivad, EV aastapäev, kokakursuste lõpetamine jne). (Wordpress.com 2007)

(25) **Kas kellaga siga nägid?**

Nagu juba sissejuhatuses sai märgitud, koosneb nimistu eeldatavalt üldtuntud väljenditest, siiski ühe erandiga, milleks on retoorilise küsimuse vormis väljend *Kas kellaga siga nägid?*, mis on suunatud inimesele, kes ei oska teretada. EKFA andmetel on selle väljendi kasutuse kõrgegaeg jäänud küll 19. sajandi lõppu ja 20. sajandi algupoolde, ent viimased fikseeringud pärinevad 1990. aastast. Vähemalt arhiivimaterjali (varaseim Wiedemannilt) põhjal võib järeldada, et tegu on peaaegselt mandrieestilise levikuga (v.a Jämaja, Mustjalas ja Karjas pelk *Oled sa siga näind?*) väljendiga. Paraku on aga jäänud mõnevõrra hämaraks väljendis sisalduva kujundi taust.⁸ Adudes, et selle väljendi populaarsus on vajunud minevikku, lülitasin selle ikkagi küsitlusse, lootes saada tõestust oletusele, et väljendit tänapäeval tõepoolest enam ei tunta. Seetõttu ei olnud üllatav tulemus, et väljend liigituski tundmatute kilda. Ent ilmnes huvitav erijoon – Tartu kooliõpilaste seas leidis neid, kes pakkusid omapoolse tõlgenduse. Tähendusversioonid lähtuvad ebahariliku kujundi nn sõnasõnalisest tõlgendamisest, st kellaga siga kui midagi ebaharilikku:

'midagi uskumatut näinu sarkastiliselt solvamine', 'miski võimatu kohta', 'oled sa midagi sellist enne näinud, oled imeasja näinud?', 'ehk paralleel "nagu sadul sea seljas" ehk sobimatu?', 'rikas inimene, aga halb, inimene, kes tahab väga moekas, teistest paremal paista'.

Samas osutab väljendi *nagu sadul sea seljas* kasutamine tähenduseseletusena assotsiatiivsetele seostele – sadul sea seljas kui midagi sobimatut või mõeldamatut, sama kellaga sea puhul.

⁸ Väljendi kasutamise ja kujunditekke kohta võib lähemalt lugeda Anne Hussari artiklist (2002).

Kokkuvõte

Kuigi fraseologismidele on iseloomulik eelkõige püsivus – millele osutavad juba nimetused püsiväljend, püsiütlus jmt – ei ole see omadus, mis välistaks igasuguse muutumise, ning seda nii leksikaalses kui semantilises plaanis. Teadupärast võivad fraseologismide kaks iseloomulikku tunnust – idiomaatilisus ja motiveeritus – olla erinevate väljendite juures eri kaaluga, ning keelekasutajale valmistavad tõelist raskust just kõrge idiomaatilisuseastmega ühendid. Gümnasistide hulgas katsena tehtud küsitlus näitas, et fraseologismi võidakse selle mittetundmise korral püüda siiski n-ö lahendada ning sõltuvalt väljendi motiveerituse astmest ei pruugi seletus olla sugugi nn tavatähendusest kaugemaleminev.

Selliste teistsuguste tõlgenduste esikolmiku osas on kokkulangevus vaid ühe väljendi osas, milleks oli *must kass on vahelt läbi jooksnud*, kaks ülejäänut on: KG-l *lõngaagi katki ei vea villast ja ega mu suu pole seinapragu*, MHG-l *nagu hobuse unenägu ja pane või hambad varna*. Suurima erinevusena torkas silma asjaolu, et Tartu kooliõpilased kaldusid tundmatu fraseologismi puhul esitama ka oletusi, omapoolseid tähendusversioone.

Küsitlust üldjoontes kokku võttes võib väita, et valitud traditsioonilisteks peetavad väljendid osutusid täiesti “elavateks” ka tänaste noorte seas, kuna esitatud väljenditest tunti kõiki (küll ühe erandiga). Kindlasti oleks tulemus olnud teistsugune, kui väljendid oleksid esitatud kontekstis. See oleks aga mõjunud pigem etteütlemissena, iseäranis kui sooviks on välja selgitada arusaamist fraseologismi ülekantusest.

Paljude väljendite puhul toimunud semantilised nihkumised – iseäranis huvi-pakkuvana mõlemas koolis ühtede ja samade väljendite puhul – tekitavad mitmeid küsimusi. Nimelt, kas võiks järeldada, et aja jooksul on toimunud tähenduse teisenemine? Ning kas seda tendentsi võiks üldistada kogu fraseoloogiaalase teadmuse kohta? Läbi viidud küsitluse põhjal sõandaks väita, et see teadmus on küll olemas, ent on mingil määral muutunud.

Siinkohal sobiks kasutada sotsiaalpsühholoogiast pärinevat terminit *mitmetähenduslikkuse tolerant*⁹ ning kohandada seda ka fraseologismide mõistmisele/mõtestamisele. Nagu küsitlusest ilmnes, käivitub väljendi mittetundmise korral ikkagi selles sisalduva kujundi otsene (või sõnasõnaline) tõlgendamine, ning kuigi selle tulemus on (äärmiselt) ebatõenäoline elik reaalsusele mittevastav, ei heideta seda kõrvale kui sobimatut. Siinkohal on oluline silmas pidada ka tõsiasja, et ei ole olemas üht ja ainust nn õiget tõlgendust, vaid neid võib olla mitu.¹⁰

Tõenäoliselt on mõistlik käsitleda fraseologismi tähendust avaramana kui seda on tema fikseeritud tavatähendus sõnaraamatus või andmebaasis. Selline arusaam annaks ka aluse suhtuda fraseoloogiliste väljendite mänglevasse kasutusse sallivalt. Hetkeseisuga on küll paraku rohkem küsimusi kui üheseid vastuseid, ent edasiseks uurimistegevuseks on sissejuhatas tehtud.

⁹ Selle mõistega tähistatakse võimet aktseptida erinevat, teistsugust, vasturääkivat. Omadust peetakse ühtlasi indiviidi kultuurilise kompetentsi eelduseks. Termin on väga levinud saksakeelses kultuuriruumis keeleõppe didaktika alal. Vt ka <http://de.wikipedia.org/wiki/Ambiguitätstoleranz> (24.09.2007).

¹⁰ Seda seisukohta on väljendanud fraseologismide puhul Harald Burger (2003: 56 jj), kes oma käsiraamatus peatub lähemalt mõistel *komponentide vaba tähendus* ning mõtestab seda kui fraseologismi koosseisu kuuluvate sõnade esinemist polüsemsetena vabas kasutuses.

Keelenäidete allikad

- Auto24 2006. http://www.auto24.ee/foorum/foorum.php?tid=72214&last_page=1 (24.04.2007).
- Delfi 2007. <http://www.delfi.ee/archive/article.php?id=15547934&categoryID> (24.09.2007).
- Delfi.blog 2006. <http://blog.delfi.ee/node/3256> (24.09.2007).
- Diplomaatia 2005. <http://www.diplomaatia.ee/est/artikkel.php?id=13> (24.09.2007).
- EESS = Loog, Mai. Esimene eesti slängi sõnaraamat. <http://www.eki.ee/dict/slang> (24.09.2007).
- EKFA = Baran, Anneli; Hussar, Anne; Õim, Asta; Õim, Katre. Eesti kõnekäändude ja fraseoloogismide andmebaas. <http://www.folklore.ee/justkui/index.html> (1.04.2007).
- EPL 2007. <http://www.epl.ee/?uudised=376311> (24.09.2007).
- Freetime 2007. <http://maasikad.ee/Helbreath/phpBB/viewtopic.php?t=24289&sid=bbc81280650131b74e464b196e15989c> (24.09.2007).
- Kesknädal 2002. <http://www.kesknadal.ee/index.php?aid=1769> (24.09.2007).
- Kesknädal 2007. <http://www.kesknadal.ee/?aid=741> (20.04.2007).
- Palts 2005. <http://www.palts.ee/index.php?id=1107> (20.04.2007).
- Pere ja Kodu 2007. <http://www.perekodu.ee/21415> (20.04.2007).
- Postimehe kommentaarikeskus 2004. http://www.postimees.ee/010704/kommentaarieskus.php?ARTIKKEL_ID=138208&TASK=KOMMENTAARID (20.04.2007).
- Postimehe kommentaarikeskus 2007. http://riigikogu.postimees.ee/online/riigikogu2007/kommentaarieskus.php?ARTIKKEL_ID=250998&TASK=KOMMENTAARID (24.09.2007).
- Saabnet 2006. <http://saabnet.ee/foorum/viewthread.php?fid=11&tid=2064&action=printable> (24.09.2007).
- Stretrace 2007. <http://www.stretrace.org/foorum/viewtopic.php?p=693150&sid=7908eace54d1482ce6170d13f43c6d8> (24.09.2007).
- TOM – täna otsustan mina 2002. https://www.eesti.ee/tom/ideas.py/show?idea_id=1467 (24.09.2007).
- Ussipesa 2005. http://ussipesa.blogspot.com/2005_05_01_archive.html (24.09.2007).
- Virumaa Teataja 2005. http://www.virumaateataja.ee/290605/kommentaarieskus.php?ARTIKKEL_ID=15025010&TASK=KOMMENTAARID (24.09.2007).
- Wordpress 2007. <http://mademoiselleparis.wordpress.com/2007/03/> (20.04.2007).

Kirjandus

- Baran, Anneli 1999. Kõnekäänud ja fraseoloogia – distinktsiooni võimalikkusest / vajalikkusest. – Lohetapja. Pro folkloristica VI. Tartu: Eesti Kirjandusmuuseum, 20–24.
- Baur, Rupprecht S.; Chlosta, Christoph 1994. Kennen Kinder heute noch Sprichwörter? Überlegungen zur Altersgrenze in Arbeiten zur empirischen Parömiologie. – C. Chlosta, P. Grzybek, E. Piirainen (Hrsg.). Sprachbilder zwischen Theorie und Praxis. Akten des Westfälischen Arbeitskreises "Phraseologie/Parömiologie" (1991/1992). Studien zur Phraseologie und Parömiologie 2. Bochum: Universitätsverlag Dr. N. Brockmeyer, 1–30.
- Burger, Harald 2003. Phraseologie. Ein Einführung am Beispiel des Deutschen. 2., überarbeitete Auflage. Grundlagen der Germanistik. Berlin: Erich Schmidt Verlag.
- Dobrovolski, Dmitri 1997. Idiome in kognitiver Perspektive. Zur Begriffsbestimmung. – M. Skog-Södersved (Hrsg.). Ethische Konzepte und mentale Kulturen 1. Auf der Suche nach ethischen Konzepten in Wortschatz und Phraseologie. Acta Universitatis Umensis, Umea Studies in the Humanities 139. Umeå, 17–57.
- EKSS = Eesti kirjakeele seletussõnaraamat. I–VII. Tallinn: Eesti Keele Sihtasutus, 1988–2007.

- Farø, Ken 2006. Ikonographie, Ikonizität und Ikonizismus: Drei Begriffe und ihre Bedeutung für die Phraseologieforschung. http://www.linguistik-online.de/27_06/faroe.html (24.09.2007).
- Farø, Ken; Hallsteinsdóttir, Erla 2006. Neue theoretische und methodische Ansätze in der Phraseologieforschung. http://www.linguistik-online.de/27_06/einleitung.html (24.09.2007).
- Filatkina, Natalia 2005. Phraseologie des Letzebuergeschen: Empirische Untersuchungen zu strukturellen, sematisch-pragmatischen und bildlichen Aspekte. Dissertation. Germanistische Bibliothek Bd. 22. Heidelberg: Universitätsverlag Winter.
- FS = Õim, Asta 2000. Fraseoloogiasõnaraamat. Tallinn: Eesti Keele Sihtasutus.
- Gibbs, Raymond W. Jr. 1989. Understanding and literal meaning. – *Cognitive Science* 13, 243–251.
- Giora, Rachel 2003. *On Our Mind: Salience, Context, and Figurative Language*. Oxford: Oxford University Press.
- Hussar, Anne 2002. Kas kellaga siga nägid? – *Oma Keel* 1, 15–20.
- Jenensšek, Vida 2006. Phraseologie und Fremdsprachenlernen. Zur Problematik einer angemessenen phraseodidaktischen Umsetzung. http://www.linguistik-online.de/27_06/jerensek.html (24.09.2007).
- Krikmann, Arvo 1997. 100 tuntumat eesti kõnekäändu. – A. Krikmann. *Sissevaateid folkloori lühivormidesse. 1: põhimõisteid, žanrisuhteid, üldprobleeme. Võrguteaviku URL* <http://www.folklore.ee/~kriku/LEX/LOENG4.HTM> (20.09.2007).
- Krikmann, Arvo (toim.) 2004. Koestler, Raskin, Attardo ja teised: lingvistiliste huumoriteooriate uuemaist arenguist. Reetor 4. Tartu: EKM, EKFK.
- Lakoff, Georg 1990. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago–London: University of Chicago Press.
- Lakoff, Georg; Johnson, Mark 2003. *Metaphors we Live by*. Chicago–London: University of Chicago Press.
- Roos, Eckhard 2001. *Idiom und Idiomatik. Ein sprachliches Phänomen im Lichte der kognitiven Linguistik und Gestalttheorie*. Aachen: Shaker.
- Röhrich, Lutz 1973. *Lexikon der sprichwörtlichen Redensarten*. Freiburg: Herder.
- Õim, Katre 2004. Fraseologism versus kõnekäänd. – M. Ereht, M-M. Sepper (toim.). *Emakeele Seltsi aastaraamat 50 (2003)*. Tallinn: Emakeele Selts, 129–142.
- Wierzbicka, Anna 1999. *Emotions Across Languages and Cultures: Diversity and Universals. Studies in Emotion and Social Interaction*. Cambridge: Cambridge University Press.
- Wikipedia. <http://de.wikipedia.org/wiki> (24.09.2007).

Anneli Baran (Eesti Kirjandusmuuseum) on lõpetanud Tartu Ülikooli eesti ja soome-ugri filoloogia erialal. Kaitses magistrikraadi Tartu Ülikoolis 1999. a eesti keele fraseoloogia alal, alates 2003. a doktorant eesti ja võrdleva rahvaluule osakonnas. On uurinud eesti fraseoloogismide semantikat ja struktuuri, fraseoloogia üldteooriat.
 anneli@folklore.ee

ON THE SEMANTIC AMBIGUITY OF PHRASEOLOGISMS

Anneli Baran

Estonian Literary Museum University of Tartu

Similarly to the constantly developing lexicon of a language, phraseology, which is closely interrelated with lexicology, is in constant progress. While an important characteristic of phraseologisms is their stability, indicated already by terms like ‘set phrases’, ‘fixed expressions’, etc. – this is not a quality that would rule out any kind of change, either on the lexical or the semantic level. It is generally known that the two characteristic features of phraseologisms – idiomaticity and motivation – may be of varying significance depending on the nature of a phrase, and a language user may find utterances of high level of idiomaticity particularly challenging. Problems with understanding phraseologisms are mostly discussed in relation with non-ethnic speakers. However, a questionnaire on the knowledge of phraseologisms, distributed among the secondary level students in two Estonian(-language) schools, revealed that phrases presumed to be commonly known – more or less idiomatic phraseologisms – may not be so for the younger generation of language users. The analysis of results shows that there have been certain changes in the conceptualisation and use of phraseologisms, and these changes are mostly associated with either the misunderstanding of the indirect meaning of a phrase, i.e., direct interpretation, or the misinterpretation based on associations of the phrase’s key word.

Keywords: semantics, phraseology, expressions, questionnaire, Estonian language

MIDA RÄÄGIVAD SÕNASISESED PAUSID JA KORDUSED SIHTKEELE KÄÄNAMIS-PROTSESSIST?

Annekatrin Kaivapalu

Ülevaade. Sõnavormide suulist produtseerimist iseloomustab kirjaliku keelekasutusega võrreldes eelkõige produtseerimisprotsessi kiirus ning aeglustusvõtete kasutus. Suulisele produtseerimisprotsessile on omased pausid ja kordused, mis peegeldavad muuhulgas selliseid psühholingvistilisi protsesse, mille kohta õppija kirjalik keelekasutus informatsiooni ei anna. Pauside esinemissagedus ja paiknemine sõnas viitavad sellele, kuidas keeleõppija morfoloogilise vormi struktuuri tajub. Kordused kaasnevad õppija enesemonitooringuga ning peegeldavad analoogiaprotsesse.

Artiklis käsitlen soome käändsõna mitmusevormide suulisel käänamisel esinevaid pause ja kordusi eesti- ja venekeelsete õppijate üksik-sõnatestides. Testi tulemustest selgus, et tüve- ja sufiksiallomorfide piirid on õppija keeletajus sageli ähmased. Sõnavormi morfoloogilise struktuuri tajumine ja produtseerimisviis sõltus suure osas muutesufiksi silbistruktuurist. Pauside ja korduste paiknemine andis tunnistust käänamisprotsessi võimalikust mittelineaarsusest. Samuti ilmnes, et käändeparadigma on õppijate jaoks psühholingvistiliselt reaalne mõiste. Pauside ja korduste kasutamise sagedus ei sõltunud kuigivõrd õppija emakeelest, vaid osutus pigem õppija individuaalseks eripäraks.

Võtmesõnad: psühholingvistika, morfoloogia, suuline produtseerimine, käänamistest, üksiksõnad, eesti- ja venekeelsed õppijad, soome keel võrreldena

1. Suulisest keelekasutusest ja selle erijoontest

Viimastel kümnenditel hoogustunud suulise kõne uurimine on nii eesti kui ka soome keeleteaduses keskendunud peaausjalikult emakeelsete kõnelejate keelekasutusele (nt Hokkanen 2001, Gerassimenko jt 2004, Mihkla jt 2005, Pajupuu, Kerge 2006). Keeleõppijate suulist keelekasutust on seevastu analüüsitud tunduvalt vähem (nt Kurhila 2002) ning keeleõppijate morfoloogiliste vormide suulist produtseerimist käsitlevad vaid üksikud uurimused (Kaivapalu 2005, 2008, Suni 2007).

Suulisele keelekasutusele on tunnuslik kirjaliku keelekasutusega võrreldes eelkõige lühem produtseerimisaeg. Kirjaliku keelekasutuse puhul on keelekasutajal aega mõelda ja valida erinevate võimaluste vahel, suuline keelekasutus aga nõuab mõtete kiiremat keelendamist. Seetõttu iseloomustavad spontaanset kõnet kordused, parandused, lõpetamata lausungid, venitused, partiklid, mitut liiki pausid, mis sageli ei lange kokku grammatiliste üksuste piiridega, varieeruv kõnetempo jne. Spontaanse kõne mittesoravuse põhjused võivad olla kognitiivsed (kõne planeerimise probleemidest tingitud), kommunikatiivsed (suhtlussituatsioonist tingitud) ja füsioloogilised (hingamine) (Meister, Lippus 2007).

Ka keeleõppija morfoloogiliste vormide suulisele produtseerimisele on omased pausid ja kordused, mille põhjused on esmajoonel kognitiivsed (Kaivapalu 2005, 2008) või kommunikatiivsed (Sunni 2008). Pausid ja kordused annavad tunnistust produtseerimiskäitusest ning peegeldavad muuhulgas selliseid psühholingvistilisi protsesse, mille kohta õppija kirjalik keelekasutus informatsiooni ei anna. Pauside ja korduste kasutussagedus ja paiknemine sõnavormis võimaldavad jälgida, kuidas õppija sõnavormi moodustab: mis talle raskusi valmistab, milliste morfeemivariantide vahel ta kõhkleb ja kuidas väljaöeldut parandab. Samuti viitavad sõnavormis esinevad pausid ja kordused sellele, kuidas õppija morfoloogilise vormi struktuuri tajub.

Käesolevas artiklis analüüsime eesti- ja venekeelsete soome keele õppijate suulisel käänamisel esinevaid pause ja kordusi ning püüan leida vastust järgmistele küsimustele:

- 1) kuidas õppija morfoloogilise vormi struktuuri tajub;
- 2) kas ja kuidas mõjutab õppija emakeel pauside ja korduste kasutamist;
- 3) kas käändeparadigma mõiste on õppija jaoks psühholingvistiliselt reaalne;
- 4) kas käänamisprotsess on alati lineaarne?

2. Ainestik ja informandid

Käesolev artikkel põhineb osal laiemast uurimisest, mille metodoloogiat olen varasemates artiklites (Kaivapalu 2005, 2006a, Kaivapalu, Martin 2007) detailselt kirjeldanud. Vaatluse all on kolm erinevat ainestikku:

- 1) suulise üksiksõnade käänamistesti abil kogutu;
- 2) kirjaliku üksiksõnade käänamistesti sooritamise ajal lindistatud õppijate kommentaarid oma tegevuse kohta;
- 3) retrospektiivsed intervjuud õppijatega pärast testi sooritamist.

Nii kirjalik kui ka suuline üksiksõnade käänamistest koosnes 30 soome käändsõnast. Testisõnade hulgas oli neli liitsõna (*talvirengas, musiikkiliike, kaupakeskus, matkalaukku*) ja neli tehissõna (*nirkas, veule, syhmä, sulkki*). Tehissõnad olid moodustatud "tavalistest" sõnadest esitähed asendamise teel. Testi lähtevormiks oli ainsuse nominatiiv ning testitavateks kääneteks mitmuse partitiiv, illatiiv ja elatiiv. Suulise testi ajal oli õpilastel ees kirjalik testiülesanne ja nende öeldud käändevormid lindistati. Testile järgnesid täpsustavad intervjuud õpilastega. Kirjaliku üksiksõnate testi sooritamise ajal palusin osal õpilastest n-õ valjusti mõelda ning kommenteerida oma tegevust. Lindistasin ka need kommentaarid ning intervjuueerisin seejärel õpilasi. Valjusti mõtlemise meetodit (sm *ääneenajattelu*, ingl *thinking aloud*) on peetud tõhusaimaks võimaluseks uurida põhjalikumalt keelematerjali produtseerimise kulgu ning saada selgust õppija mentaalsetest protsessidest (Block 1992, Swain, Lapkin 1995).

Informantideks olid eesti abiturientid Kärkla Ühisgümnaasiumist ja Tartu Forseliuse Gümnaasiumist ning vene abiturientid Peterburi 204. koolist. Mõlemad uurimuse sihtrühmad olid õppinud soome keelt võõrkeelena umbes 200 tundi ning ka kõikide muude taustategurite (vanus, sotsiaalne, haridus- ja kultuuritaust, motivatsioon) poolest ühesugused. Suulises testis osales kümme eestikeelset ja neli venekeelset soome keele õppijat. Eestikeelsed õppijad produtseerisid suulises testis kokku 778 ja venekeelsed õppijad 357 sõnavormi. Kirjaliku testi ajal lindistasin üheksa eesti- ja kolme venekeelse õppija kommentaarid. Kõiki suulises ja kirjalikus testis osalenud informante intervjuueerisin pärast testi sooritamist. Litereeritud keelenäidetes tähendab ? küsivat intonatsiooni, / lühemat, // pikemat pausi ja - katkestust produtseeritavas sõnavormis. Pauside lühemate või pikemate hulka liigitamise aluseks oli litereerija keeletaju. Eestikeelsete õppijate keelenäited on tähistatud lühendiga *ee* ja venekeelsete õppijate näited lühendiga *vn* vastava näite ees. Kui näide esines nii eesti kui ka venekeelsete õppijate testivastuste hulgas, ei ole lähtekeelt eraldi tähistatud, samuti juhul, kui õppija lähtekeel selgub näite kontekstist.

3. Kirjalik ja suuline üksiksõnade käänamistest võrdluses

Kirjalik üksiksõnatest on nii võõrkeeleõppes kui ka teise keele omandamisel traditsiooniline ning levinud meetod harjutamiseks ja õppija oskuste mõõtmiseks. Samuti on kirjalik üksiksõnatest olnud uurimistöös laialdaselt kasutusel ainekogumise efektiivse meetodina, mis võimaldab suhteliselt lühikese aja jooksul saada uurimisobjektiks vajalikus mahus sõnavorme, mida õppija spontaanses keelekasutuses esineb vähe, ei esine üldse või mida on võimalik vältida, asendades käänatava sõna teisega või jättes see käänamata (Martin 1992: 271–274).

Seevastu suulist üksiksõnade käänamistest teadaolevalt peaaegu ei ole kasutatud ei õppe- ega uurimistöös. Esimesed sellisuunalised katsed (Kaivapalu 2005) on aga andnud huvitavaid tulemusi ning rohkesti vihjeid edasiseks uurimiseks. Õppijate arvates raskendab suulist käänamist kirjalikuga võrreldes lühem produtseerimisaeg, mis jätab vähem võimalusi erinevate variantide kaalumiseks ning enesemonitooringuks.

A.K.: Kas kirjalikult on kergem [testi] teha kui suuliselt?
 P.: Eks vist mingi vahe ikka on. Kirjutades näeb, mis tuleb.
 A.K.: On nagu aega järele natukene mõelda ka?
 P.: Mh-h.

Sama kinnitas eestikeelsete õppijate kirjaliku ja suulise üksiksõnatesti tulemuste võrdlus (Kaivapalu 2005: 205–208): kirjaliku üksiksõnatesti tulemused osutusid ootuspäraselt suulise testi tulemustest paremateks peaaegu kõikide testisõnade osas. Venekeelsed õppijad aga käänasid vastu ootusi kõiki testisõnu suuliselt paremini kui kirjalikult (Kaivapalu 2005: 231–234). Niisuguse tulemuse põhjusi on tõenäoliselt enam kui üks. Nähtavasti nõuab juba lähtekeelest tunduvalt erineva tähestiku ja ortograafia kasutamine täiendavaid jõupingutusi. Samuti võis tulemusi mõjutada venekeelsete õppijate eestikeelsetest õppijatest erinev käänamisstrateegiate kasutus. Kui eestikeelsed õppijad lähtusid käänamisel esmajoones lähte- või sihtkeele analoogiast, siis venekeelsete õppijate valdav strateegia oli käändevormide produtseerimine põhivormist õpitud reeglite abil. On võimalik, et sageli keeruliste, hierarhiliste ja mittelineaarsete käänamisreeglite rakendamine (lähemalt vt ptk 4.5.) raskendas venekeelsete õppijate kirjalikku produtseerimisprotsessi, kuna pikem mõtlemisaeg võib ka suurendada kõhklusit ja ebakindlust erinevate variantide valikul. Seevastu suulisel käänamisel ei ole alati aega reegleid kasutada ning analoogial põhinev produtseerimine pääseb vabamalt toimima (Kaivapalu 2006b: 82–83). Järgneva peatüki keskmes on suulise üksiksõnade käänamistesti tulemuste analüüs.

4. Aeglustamisvõtted suulises üksiksõnade käänamistestis

Vaatamata tervikuna parematele suulise testi käänamistulemustele kasutasid venekeelsed õppijad suulisel produtseerimisel pause ja kordusi tunduvalt rohkem kui eestikeelsed õppijad (vt tabel 1).

Tabel 1. Pauside ja korduste esinemissagedus eesti- ja venekeelsete soome keele õppijate suulises käänamistestis kordade arvuna ja protsendina kõikidest suuliselt produtseeritud sõnavormidest

	Eestikeelsed õppijad		Venekeelsed õppijad	
	Kordade arv	% kõigist sõnavormidest	Kordade arv	% kõigist sõnavormidest
Pausid silbipiiril	132	17,0	150	42,0
Pausid mujal sõnas	64	8,2	44	12,3
Kordused	119	15,3	84	23,5
Kokku	314	40,5	278	77,9

Nii sage aeglustamisvõtete kasutamine lubab oletada, et tegemist ei ole juhusliku nähtusega, vaid produtseerimisprotsessi kõneka peegeldusega, mille analüüs annab olulist informatsiooni ning võimaldab teha järeldusi mitmete keeleomandamise kesksete küsimuste kohta.

4.1. Pausid: kuidas õppija morfoloogilise vormi ehitust tajub

Nii eesti- kui ka venekeelsed soome keele õppijad kasutasid suulisel käänamisel kõige sagedamini pause, mis esinesid mitte ainult silbipiiril, vaid ka mujal sõnas. Silbipiiril esinevate pauside puhul on põhjust eeldada, et õppija lähtub esmajoones oma emakeelest ning silbitab nii, nagu ta on harjunud silbitama. Nii lähte- kui ka omandatava sihtkeele fonoloogia mõjutavad kahtlemata õppija produtseeritava morfoloogilise vormi ehituse taju (Niemi jt 2001: 197). Silbipiiril, kuid veelgi enam mujal sõnas esinevate pauside esinemissagedus ja paiknemine võivad anda vihjeid selle kohta, kas tüve-, tunnuse- ja lõpuallomorfide piirid on keeleõppija jaoks tõepoolest sama üheselt määratletavad kui õpitava keele süsteemi kirjeldustes.

Käesolevas uurimuses testitud soome keele mitmuse käänete muutesufiksud erinevad üksteisest sufiksi silbistruktuuri, produktiivsuse, kasutussageduse ja läbipaistvuse poolest.

Mitmuse partitiivis on kolm võimalikku sufiksit, mille vahel õppija peab käändevormi moodustamisel valima:

- 1) *-jA* (*kouluja, leluja*);
- 2) *-itA* (*vieraita, herneitä*);
- 3) *-iA* (*toisia, hyisiä*).

Neist kolmest võimalusest on *jA*-lõpuline mitmuse partitiiv soome keeles kõige üldisem ja produktiivsem. Samas on see sufiks ka läbipaistev, kuna moodustab silbi, milles on lihtne näha nii mitmuse kui ka käände markeerijat. *itA*-lõpulise mitmuse partitiivi puhul on samuti selgesti eristatavad mitmuse tunnus ja käändelõpp, kuid see sufiks ei ole tänapäeva soome keeles nii produktiivne kui *-jA*. Seevastu *iA*-lõpuline mitmuse partitiiv võib kergesti segi minna *i*-tüveliste käändsõnade ainsuse partitiiviga (nt *lasia*) (Niemi jt 2001: 196).

Mitmuse illatiivis on samuti kolm võimalikku sufiksivarianti:

- 1) *-ihin* (*pankkeihin, herkkuihin*);
- 2) *-isiin* (*kankaisiin, patsaisiin*);
- 3) *-iin* (*toisiin, veistoksiin*).

Kaks esimest, *ihin*- ja *isiin*-sufiksud, on silbistruktuuri ja läbipaistvuse osas võrreldavad partitiivi *itA*-sufiksiga. Käändtüüpides, kus *-ihin* ja *-isiin* vabalt vahelduvad (41 *vieras* ja 48 *hame*, PS 1990: XVI–XVII), eelistatakse tänapäeva soome keele kasutuses siiski *isiin*-lõpulist illatiivi (Itkonen 1979: 67). Reduplikatiivsete omadustega *iin*-sufiks (Karlsson 1983: 182) ei ole nii nagu ka partitiivi *-iA* sama läbipaistev kui eelmised.

Mitmuse elatiivi ainus võimalik sufiks on *-istA* (*kauppakeskuksista*).

Tüve- ja sufiksivariantide kombineerimise tulemus võib sihtkeeles olla nii normikohane (*tuot/tei/ta*) kui ka normivastane (*tuotte/ja*). Pausid produtseeritavas sõnas annavad tunnistust protsessi katke(sta)misest, teisisõnu õppija raskustest käändevormi moodustamisel. Pauside paiknemise põhjal võib teha oletusi muuhulgas selle kohta, kuidas õppija morfoloogilise vormi ehitust tajub.

Mitmuse partitiivi jA-sufiks produtseeriti peaaegu eranditeta omaette silbina (ee *kirkko//ja*, vn *kirk/ko/ja*; ee *matkalaukku/ja*, vn *matkalauk/ku/ja*; vn *oh/je/ja* pro *ohjeita*). Eestikeelsed õppijad produtseerisid *jA*-sufiksi kolmeteistkümnel ja venekeelsed õppijad kahel korral ka pikema häälikujärjendi osana (ee *talvi/*

renkkoja pro talvirenkaita; ee koh/toja pro kohteita; vn matkalauk/kuja). Käsitus *jA*-sufiksist kui sünkreetilise morfeemi taolisest tervikust (Niemi jt 2001: 196) seletab ka nii eesti- kui venekeelsete õppijate testivastustes ulatuslikult esinenud tüvevokaali vaheldust sõnavormides, kus seda normikohaselt tegelikult ei esine (nt *viereja pro vieraita; vilkkeja, vilkkoja pro vilkkaita; kangoja pro kangaita*). Tõenäoliselt tingib *jA*-sufiksi oletatav sünkreetilisus vajaduse tugevdada käändevormi mitmuslikkust täiendava tüvevokaali vahelduse abil, samas kui *itA*-sufiksi ees selliseid vaheldusi ei esine.

Mitmuse partitiivi -itA ning illatiivi -ihin ja -isiin seevastu jagunesid nii eesti- kui ka venekeelsete soome keele õppijate suulises produktsioonis kaheks nii, et mitmuse *i*-tunnus tüve osa ja *tA*-, *hin*- ja *siin*-lõpud produtseeriti eraldi (*tuot/tei/ta, pankkei//ta, pat/sai/siin, herkkui/hin*). Seega on põhjust oletada, et antud juhtudel tajuti mitmuse *i*-tunnust pigem (mitmuse)tüve osana kui iseseisva aglutineeritava morfeemina, samas kui partitiivi käändelõpud *-ta, -hin* ja *-siin* liideti iseseisvate morfeemidena tüvele. Reduplikatiivse *iin*-sufiksiga vorme esines nii eesti- kui ka venekeelsete õppijate testivastustes vaid üksikuid. Sel juhul olid need sõnavormid produtseeritud tervikuna (ee *ohjeiin pro ohjeisiin; ee musiikkiliikkiin pro musiikkiliikkeihin; vn tuotiin pro tuotteisiin*). Juhul kui sõna tüves esines *-s*, olid õppijad restruktureerinud enda jaoks *iin*-sufiksi *siin*-sufiksiks (*toi/siin*).

Enam varieeruvust esines **mitmuse partitiivi iA- ja mitmuse elatiivi stA-sufiks** suulisel käänamisel. *-iA* võis olla produtseeritud nii, et mitmuse *i*-tunnus jäi eelviimase silbi lõppu ja partitiivi lõpp moodustas eraldi silbi (ee *op/pa/si/a; vn kauppa/kesku/ksi/a; vn kerroksi/a*) või oli *iA*-sufiks pikema häälikujärjendi osa (*toi/sia; hyi//siä; ee veistoi/sia, vn veistok/sia; vn kauppakesku/ksia, ee kauppakesku/sia*). Esimesel juhul on põhjust oletada, et õppija tajus mitmuse *i*-tunnust pigem (mitmuse)tüve osana ja eristas enda jaoks selgelt käändelõpu; teisel juhul väljendas õppija jaoks vormi mitmuslikkust kas tüvi tervikuna või sõnavormi lõpphäälikute järjend, mille pikkus võis varieeruda. On võimalik, et eestikeelsete õppijate *si(a)*-järjendi omaette tervikuna produtseerimist soodustas lähtekeele *raskus-tüübi si-lõpuline* tüvevitmus *raskusi*.

Mitmuse elatiivivormide käänamisel järgiti sageli silbipiire: *tuot/teis/ta; vn laineis/ta, ee lainis/ta, ee lai/nois/ta pro laineista; her/neis/tä; vilkkais/ta, vilk/kais/ta; vierais/ta; ee talviren/gois/ta, vn talvirengais/ta pro talvirengaista*. Ilmselt soodustab käändelõpu *-stA* jagunemist kahe silbi vahele muuhulgas ka asjaolu, et *-s* on soome keeles inessiivi ja elatiivi, eesti keeles aga kõikide sisekohakäänete ühine komponent, millele järgneb konkreetse sisekohakäände spetsiifiline aines. Sisekohakäändeid moodustades on õppija jaoks lihtsam ja turvalisem produtseerida esialgu ainult kõikidesse sisekohakäänetes sobiv *-s* ja alles seejärel viia protsess lõpuni. Mitmuse *i*-tunnus võis elatiivis esineda iseseisvana (ee *vilkk/i/sta; ee nirkk/i/asta pro nirkkaista*), tüve osana (ee *hernei/s/tä, vn her/neis/tä; ee kankai/sta, vn kan/kai/sta; ee talvirenki/sta pro talvirengaista*), muutesufiksi (ee *tuotte//ista; ee talvireng/ista pro talvirengaista*) või pikema häälikujärjendi (vn *kerrok/sista; vn veis/to//ksista*) osana. Ulatuslikum variatsioon mitmuse elatiivivormide produtseerimisel võrreldes teiste testikäänetega on tõenäoliselt tingitud asjaolust, et vaid ühe võimaliku muutesufiksi olemasolul ei tarvitse õppija valida mitme variandi vahel ja seega ei ole sõnavormi struktuuri teadvustamine nii oluline kui enamate võimaluste korral.

Testitulemuste põhjal võib niisiis järeldada, et tüve-, tunnuse- ja lõpumorfeemide piirid ei ole õppija keeletajus sugugi nii selged ja üheselt määratletavad, kui keelekirjelduste ja keeleõpetuse põhjal võiks eeldada. Eriti varieeruv on mitmuse *i*-tunnuse positsioon: seda tajutakse kõige sagedamini (mitmuse)tüve osana, mõnikord muutesufiksi osana ja ainult harva iseseisva morfeemina. Kuna partitiivi *-ta*, illatiivi *-hin* ja *-siin* moodustavad omaette silbi, paigutub mitmuse *-i* produtseerimisel tüve viimase silbi lõppu ja on nagu tüve osa (vt ka Niemi 2001: 204–206). Sellist arusaama toetab eestikeelsete õppijate puhul lähtekeele tüvemitmus: eestikeelne õppija on harjunud sellega, et tüvi väljendab grammatilisi tähendusi, sealhulgas ka mitmust. Seevastu partitiivi *JA*-sufiksis moodustavad mitmuse tunnus ja käändelõpp omaette silbi ning seetõttu on tõenäoline, et sufiksit tajutakse pigem sünkreetilise morfeemina. On võimalik, et muutesufiksi silbistruktuur aitab muuhulgas kaasa sellele, kas mitmuse *i*-tunnus ühendatakse käänamisel tüve või käändelõpuga.

Muutesufiksitate produtseerimise mitmekesisus annab tunnistust sellest, et õppija tajub suhteliselt selgelt tüve algusosa ja silbikujulist käändelõppu. Seevastu tüve lõpuosa, mitmuse tunnuse ja teatud juhtudel ka käändelõpu piirid võivad õppija keeletajus olla vägagi ähmased ning grammatilisi tähendusi väljendavad muutesufiksidsid kas kogu tüve või tüveosa koosseisus. Morfeemiipiiride tajumise ähmasus seletaks osaliselt õppijate raskusi tüve- ja morfeemivariantide valikul: õppijal on raske aru saada, kus tüvi lõpeb ja tunnus algab, kus tunnus lõpeb ja käändelõpp algab ning milline sõnaosa (häälikujärjend) signaliseerib, et vorm on mitmuses. Seetõttu tugevdatakse vormi mitmuslikkust erineval viisil: tüvevokaali vahelduse abil seal, kus seda normikohaselt ei esine; mitme mitmuse tunnuse (sageli nii lähte- kui ka sihtkeelne) samaaegse kasutamisega, mitmuse *i*-tunnuse lisamisega ebamäärasesse paika tüve lõpuosas. Õppijad niisiis restruktureerivad enda jaoks produtseeritavat vormi (vt ka Martin 1995: 198). Morfeemiipiiride tajumise ähmasus suurendab ka tõenäosust, et käändevorm produtseeritakse analoogiale, mitte aga reeglitele toetudes (Kaivapalu 2006b: 82–83).

4.2. Kordused: õppija enesemonitooring

Kordused on levinumaid viise sobiva variandi otsimiseks (Muikku-Werner 2002: 16). Ei ole siis üllatav, et eesti- ja venekeelsed soome keele õppijad kasutasid sageli produtseeritava käändevormi katkestamist ja selle osalist või tervikuna kordamist. Kõige rohkem esines kordusi mitmuse partitiivivormide, vähem illatiivi ja elatiivi moodustamisel. Korduste rohke kasutamine partitiivis on ootuspärane: see oli testis esimene produtseeritav käändevorm ja õppijatele oli oluline leida sobiv produtseerimisviis ka järgnevate vormide jaoks. Arusaadavalt ei ole mitme tüve- ja sufiksivariandi ning nende kombinatsiooni vahel valimine kerge. Seetõttu jäeti käänamine sageli pooleli kontrollimaks, kas produtseeritud tüvevariant tundub sihtkeeles normipärane või normivastane ja alles seejärel viidi protsess lõpuni.

- (1) ee *talvirenkais- renkaisia*
ee *laino- lainoja*
vn *tuo- tuoteita*

Suurema kahtluse korral peatuti sõnas silbipiiril või mujal, jäeti produtseerimine veel kord pooleli või produtseeriti vorm esialgu osaliselt ja alles seejärel tervikuna.

- (2) ee *kan- kan/koja*
ee *nirk- nirk/koja*
ee *vilk- vilk/kaita vilkkaita*
ee *kang- kangoi/ta kangoita*
vn *hyine- hyi//sia*
vn *tuo- tuot/te/ja*
vn *hernei- herne/ja*

Selline järkjärguline produtseerimine annab õppijale enesemonitooringu võimaluse, võimaluse kontrollida, kuidas käändevorm “kõlab” ja vajaduse korral korrigeerida juba (osaliselt) produtseeritud vormi. Õige või väära, aga igal juhul õppija meelest sobiva lahenduse leidmise tähtsusest kohe esimeses testivormis annab tunnistust asjaolu, et võrreldes partitiiviga esines illatiivis kordusi 41 ehk 7,9% võrra vähem ja elatiivis vaid üksikuid (3). Siiski olid ka need kordused tingitud õppija kõhklumisest mitme võimaluse vahel (4).

- (3) ee *vie/rai- vie/rai/siin vieraisiin vieraisihin*
ee *nirk/koi/- nirkkoisiin nirk- nirkkois- nirkkaksiin*
ee *pats- patso/is- patsoi- patsoista*
ee *musiikkiliik/keis/ta liik- liikkei- liikeista*
ee *eevop/pai/sin oppai- oppai- oppaisin? oppai//siin*
vn *lainei- lai/nei/siin*
vn *koh/de- koh/de/si- siin*

- (4) vn *veu- velulei/sta*
vn *talvirenkai- ren- ren/kai/sta*

Õppijate kõhklused jagunesid laias laastus kahte rühma. Enamasti kõheldi kahe tüvevariandi vahel.

- (5) ee *hitaisi- hitaiseista*
ee *musiikkiliikk/ei- musiikki//liikkia*
ee *talvirenka- talviren/koja*
ee *talvirenkai- ta- talvi- rengoita*
vn *op- op//pai//ja op/pai/ta*
vn *sulkkija sulkkeja sulkkija sulkkeja*
vn *viera- vieroi- vier- vieroi/ta*

Kõhklusi põhjustas ka sufiksivariandi valik.

- (6) ee *vilkkai/sis- vilkkaksiin vilk/kai/si/hin*
ee *kohteja koh/tei/ta*
ee *musiikkiliik/kei/hin liik/kei/siin*
vn *vieras- viera- viera/ta vieria vieraja vieraita*

Kordustes peegelduvad õppijate sisemine ja väline enesemonitooring ning produtseeritu parandamine. Willem Levelt (1983: 56) on eristanud eneseparandustes

kolme etappi. Esimene etapp hõlmab produtseeritu pidevat monitooringut ning kohe, kui probleemi on märgatud, produtseerimisprotsess katkestatakse. Teist etappi iseloomustavad pausid, kõhklemised ja parandusväljendid (ingl *editing expressions*, sm *korjausilmaukset*). W. Levelt (1989) juhib tähelepanu ka asjaolule, et parandusväljendid erinevad vigade märkamise ja sobiva variandi otsimise korral ning esinevad peamiselt esimesel juhul.

Sama suundumus ilmnes eesti- ja venekeelsete soome keele õppijate suulisel käänamisel. Testisõnades *toinen* ja *kauppakeskus*, mille tüvevariantides esineb samuti *s*-häälik nagu elatiivi muutesufiksis, ajasid õppijad kergesti segamini partiivi- ja elatiivivormid. Kui viga märgati, parandati see pärast parandusväljendit.

- (7) *toisista ei toi- toisia*
kauppakeskuista ei kauppakes/kui/ta

Samamoodi toimiti teisteski testikäänetes.

- (8) *tuotteihin her -siin tuotteisiin*
her/nei//siin her herneihin

Vahel märgati alles kõikide testitavate käändevormide produtseerimise järel, et sõna oli käänatud ainsuses.

- (9) *tuote/ta : tuottesiin : tuottesta her tuo/tei/ta tuo/tei- tuotteihin oi -siin*
tuot/teis/ta
sulkkija : sulkkeihin : sulkkeista oi sulkkija sulkkeja sulkkeja sulk/kei/hin
sulk/kei/sta

Kui tegemist ei olnud vea märkamisega, vaid sobivama variandi otsimisega, kasutati erinevaid parandusväljendeid.

- (10) *kauppakesku- oota kauppakeskuissa*
kank/koi/hin oi kankkasiin kankkoihin
talviren//keista oi -aista
opaisi/s/sa ek-ee oi siis opaisis- ei opaisi/ssa
herkku- herkku/siin her- tak herkku/hin

Kordused kaasnevad seega õppija enesemonitooringuga ning peegeldavad analoogiaprotsesse: õppija peab otsustama, milliste kombinatsioonide vahel valida (MacWhinney 1978: 1–8).

4.3. Lähtekeele mõju aeglustamisvõtete kasutamisele

Eesti- ja venekeelsete soome keele õppijate suulise üksiksõnade käänamistesti tulemustest ilmnes, et suulisele produtseerimisele on omased teatud suundumused, mida järgitakse sõltumata lähtekeelest. Nii eesti- kui ka venekeelsed õppijad kasutasid samu produtseerimisviise ja muutesufikseid, mis sõltusid eelkõige silbi-struktuurist.

Oululisim erinevus eesti- ja venekeelsete õppijate suulises produtseerimisprotsessis on kvantitatiivne. Tabelis 2 on esitatud pauside ja korduste esinemiste

absoluutarv eesti- ja venekeelsete õppijate suulises käänamistestis. Iga informant moodustas kokku üheksakümmend käändevormi, kusjuures tuleb arvesse võtta, et pausid ja kordused võisid produtseeritavas sõnavormis ka kombineeruda.

Tabel 2. Pauside ja korduste esinemine eesti- ja venekeelsete soome keele õppijate¹ suulises käänamistestis

Informant	Pausid silbi- piiril	Pausid mujal kui silbipiiril	Kordused	Kokku
Eestikeelsed õppijad				
Mart	41	8	28	77
Paul	23	3	21	47
Katrin	17	24	1	42
Kaisa	5	13	17	35
Piret	19	4	10	33
Reet	17	4	11	32
Kersti	9	1	8	18
Ants	5	1	7	13
Tauno	1	4	5	10
Helle	2	3	2	7
Venekeelsed õppijad				
Anna	54	19	10	83
Leena	40	3	35	78
Mitja	31	9	30	70
Valja	26	17	7	50

Tabelist 2 ilmneb, et venekeelsed õppijad peatuvad produtseerimisprotsessis kaks korda sagedamini kui eestikeelsed õppijad. See asjaolu näib esmapilgul olevat vastuolus venekeelsete õppijate kirjalikust käänamistestist paremate suulise käänamistesti tulemustega. Nähtavasti aga ei takista produtseerimisprotsessi sagedasem katkestamine kõrgema normipäraste vastuste protsendi saavutamist, vaid isegi aitab sellele kaasa. Vähesed erinevused eesti- ja venekeelsete õppijate produtseerimisprotsessis johtuvad lisaks lähte- ja sihtkeele süsteemierinevustele vähemalt osaliselt ka käänamisstrateegiate erinevast jaotumisest. Eestikeelsed õppijad on altid kasutama nii lähte- kui ka sihtkeele analoogiat, samal ajal kui venekeelsed õppijad käänavad esmajoones reeglite abil ning sihtkeele analoogia võetakse appi alles siis, kui reeglite keerukus ületab õppija produtseerimisvõime piiri.

Samas osutus kasutatud pauside ja korduste hulk ning jagunemine mõlema õppijarühma puhul õppijakeskseks ilminguks. Nii eesti- kui ka venekeelsetel õppijatel olid suulisel käänamisel oma lemmiktoimetulekustrateegiad: osa õppijatest produtseerisid käändevorme peaaegu eranditeta tervikuna, suurem osa aga kasutas pause ja kordusi, kuid nende hulk ning jaotumine olid erinevad. Tähelepanu äratav vaid seik, et kõikidel venekeelsetel õppijatel esines käänamisel suhteliselt palju pause silbipiiril. On võimalik, et lühemate või pikemate pauside rohke kasutamine tuleneb lähte- ja sihtkeele erinevatest kvantiteedi- ja rõhusuhetest (Liekö

¹ Tabelis esinevad õppijate nimed on muudetud.

1992, Nenonen 2001): kahtluse korral on turvalisem produtseerida käändevorm silbikaupa, mitte aga pikemate sekvenssidena.

Eestikeelsete õppijate puhul annab lähedase sugulaskeele positiivsest mõjust või selle puudumisest tunnistust pauside ja korduste sagedasem kasutamine lähte- ja sihtkeeles eritüvelistes sõnades (tabel 3).

Tabel 3. Pauside ja korduste kasutamine eestikeelsete õppijate suulises käänamistestis testisõnärühmiti

Testisõnärühm	Pauside ja korduste arv	% kõigist sõnavormidest
STSK ²	81	38,9
ETSK	91	43,8
STEK	60	34,1
ETEK	82	45,1

Nii eesti ja soome keeles sama kui ka erineva käänamismalli korral on eritüvelistes testisõnades aeglustamisvõtteid kasutatud enam kui samatüvelistes sõnades. Seega on lähte- ja sihtkeele sõnatüvede sarnasus käänamismalli valikul täiendava vihjena suulisel produtseerimisel olulisem kui kirjalikul produtseerimisel. Õppija enesemonitooringu tähtsus niisiis kasvab, kui sõna tüvi on õppijale tundmatu ega ka lähtekeel anna vihjeid käänamismalli valikuks.

4.4. Käänamisprotsessi mittelineaarsus

Nagu eelnevas näidatud, on suurem osa kordustest seotud õppija kõhklustega sobiva tüve- ja/või sufiksivariandi ning nende kombinatsioonide valikul. Nii keelekirjeldustes kui ka keeleõppes on traditsiooniliselt *a priori* lähtunud oletusest, et sõnamuutmine on vasakult paremale kulgev lineaarne protsess: käänamisel ja pööramisel liidetakse muutumatule või vahelduvale tüvele tunnuseid ja lõppe. Siiski esineb üksikutes keelesüsteemi kirjeldustes vihjeid sõnamuutmise võimalikule mittelineaarsusele: mitmuse ja käände tunnused on sõltuvad teineteisest, tüve struktuurist ja morfofonoloogilistest muutustest (Karlsson 1983: 282). Järjekindlamalt esindab arusaama sõnamuutmise mittelineaarsusest Heikki Paunose (1983) morfoloogiamudel.

Veel enam kui keelesüsteemi kirjeldustes tõuseb küsimus sõnamuutmise lineaarsusest esile keeleõppijate käänamistulemusi analüüsid. Ilmselt on käänamine ja pööramine mingil sihtkeele omandamise etapil lineaarne protsess nendel, kellele seda on sellisena õpetatud. Iseküsimus on seejuures muidugi, kui pikk see etapp on ja kas lineaarne sõnamuutmine jääb keelekasutuses püsima. Eesti- ja venekeelsete soome keele õppijate suulistes käänamistestides esineb näiteid selle kohta, et tüvevariandi valik mõjutab sufiksivariandi valikut. Sageli põhjustab õppija kõhklemine mitme tüvevariandi vahel ka sufiksivariandi vahetuse.

- (11) *ee hitoja hi/taisia*
ee lakkia lak/ko/ja
ee hitoja hi/taisia
vn sulkkeihin sulkkiin

² Lühendite tähendused: STSK – eesti ja soome keeles sama tüvi, sama käänamine; ETSK – eri tüvi, sama käänamine; STEK – sama tüvi, erinev käänamine; ETEK – eri tüvi, erinev käänamine.

Samas võib protsess olla ka vastupidine. Kui juba produtseeritud osaline või ter-
vikkäänevorm tundub monitooringu tulemusena mingil põhjusel ebasobiv, tingib
sufiksivariandi vahetus tüvevariandi vahetuse.

(12) *ee vilkkai/sis- vilkkaisiin vilk/kai/si/hin*
ee talvirenk/ais/sa talvirengissa
ee nirk/koi/- nirkkoiisiin nirk- nirkkois- nirkkaiisiin
vn vie/rai- vie/rai/siin vieraisiin vieraisihin

Suulistest käänamistestidest ja intervjuudest õppijatega selgub, et sageli lähtutakse
mitmusevormide käänamisel just sufiksist, eriti aga käändelõpust, mis näiteks
elatiivis on ainsa võimaliku variandina eestikeelsetele õppijatele emakeele põhjal
tuttav. Alles seejärel proovitakse leida tüvevariant, sest viimane on eriti vahelduvate
tüvede puhul tunduvalt keerulisem.

(13) *ee tuotteista tuot- tuotteisista*
ee talviren//keista või -aista

Küsimus on siis pigem tüve- ja sufiksivariantide kombinatsioonides, mille vahel
õppija valib.

(14) *ee hyinei- hyi/nei/siin hyiniin*
ee musiikkilik/koja musiikkilik/kia
ee kankkoja ei kankkaisia kankkoja
ee kank/koi/hin või kankkaiisiin kankkoihin
vn vieras- viera- viera/ta vieria vieraja vieraita

Seejuures ei ole alati sugugi selge, kas tüvevariandi valik mõjutab sufiksivariandi
valikut või vastupidi (vt ka Martin 1995: 174–176). Tõenäoliselt on tegemist pigem
mitmesuunalise, omamoodi spiraalikujulise protsessiga, milles komponentide
mõju on vastastikune, nagu Urho Määttä (1998: 2) on tõdenud: “Keele struktuur
põhineb nii vormi- kui ka lauseõpetuse tasandil eritasandilistel ja -suunalistel
mõjuseostel.”

Käänamisprotsessi võimalik mittelineaarsus seostub tihedalt sõnamuutmis-
strateegiatega, eriti reeglite ja analoogia rolliga ning nendevahelise tööjaotusega.
Nagu eelnevalt tõdetud, viitab olukord, kus õppijal tuleb valida mitme tüve- ja
sufiksivariandi kombinatsiooni vahel, analoogiaprotsessidele. Seetõttu on kas lähte-
või sihtkeele analoogia kasutamine käänamisstrateegiana reeglitele toetumisest
efektiivsem juhul, kui tegemist on enam kui ühe tüve- ja/või sufiksivariandiga
(Kaivapalu 2006b: 83).

4.5. Kääneparadigma psühholingvistiline reaalsus

Ulatuslikult muutemorfoloogiat kasutavate keelte, nagu eesti ja soome keele puhul
on sõna muuteparadigma traditsiooniliselt keskne mõiste nii keelesüsteemi kirjel-
dustes (EKG I, ISK, Karlsson 1983, Paunonen 1983, Hint 1997) kui ka emakeeleõppes
(nt Hint 1983, Ehala 1997) isegi siis, kui mõistet ennast eksplitsiitselt ei rõhutata
(Õunapuu 2003). Seejuures aga teatakse väga vähe sellest, kui võrd reaalne on
muuteparadigma mõiste emakeelse keelekasutaja või keeleõppija keeletajus. Kas

sõnad rühmituvad mälus eelkõige tähenduse või pigem fonoloogilise kuju alusel? Kas sama sõna muutevormid kuuluvad kokku? Milline seos valitseb eri sõnade samade muutevormide vahel? Milline roll on süntaktilistel faktoritel, näiteks rekt-sioonil või kollokatsioonidel? Väheste selleteemaliste uurimuste põhjal (Paunonen 1983:61, Martin jt 2001: 74–75) on muuteparadigma osutunud kõige selgemaks sõnavormide ühendajaks: nii soomekeelsete keeikasutajate kui ka soome keele õppijate arvates valitseb sama sõna muutevormide vahel tunduvalt tihedam seos kui sõnavormi ja tuletise või sõnavormi ja sellele semantiliselt või fonoloogiliselt lähedase sõna vahel. Käesoleva uurimuse suulise testiosa tulemused kinnitasid muuteparadigma psühholingvistilist reaalsust vähemalt kahest aspektist.

Nii eesti- kui ka venekeelsed õppijad kasutasid sageli esimeses testikäandes, mitmuse partitiivis moodustatud tüvevarianti järgnevates testikäanetes aglutineerimisbaasina, millele liideti tunnuse- ja lõpuvariante. Seejuures võis tüvevariant olla saadud erinevaid käänamisstrateegiaid kasutades ega tarvitsenud olla ilmtingimata sihtkeeles normikohane.

- (15) *ee lainia : lainissa : lainis/ta*
ee hernei- herne/ita : hernei/s/sa : hernei/s/ta
ee herno- hernoja : hernoihin : hernoista
ee talvirenka- talviren/koja : talviren/koihin : talviren/gois/ta
ee tuot- tuot/te- tuotteita : tuot/tei/hin : tuotte//ista
vn vil/koi- vilk/koi/ta : vil/koi//hin : vilk/kois/ta
vn patse/ta : patse/hin : patse/sta
vn op/pa- op/pa/ta : op/pa/hin : op- oppa/sta

Kahtlemata ei saa eelnevate näidete puhul täiesti välistada testiefekti mõju, st eelneva testitava käändevormi mõju järgnevale vormile. Kui aga õppijad otsivad produtseerimisprotsessis abi sellistest käändevormidest, mida testis üldse ei küsita, on see kaalukas argument osutamaks, et sama testisõna erinevate käändevormide vahel valitseb õppija keeletajus seos.

- (16) *ee lelu lelu lelua palju on leluja*
ee lelu- lelun leloita
ee herkku herk/ku/a herkkui- kuita
vn opia opas opaita
vn kerro- kerros kerroi/ja kerroksi/a

- (17) *kohde kohteita kohte/siin. Ja seestütlev on ainsuses oleks mingi koh-teesta, vist siis mitmuses oleks kohteista.*

Seejuures ei olnud n-ö abivorm sageli sihtkeeles normikohane, vaid lühenenud ja lihtsustunud (18) nagu õppijakeelele üldiselt omane (MacWhinney 1987).

- (18) *oppaita. Sisseütlev on ainsuses vist tuleb oppeen, mitmuses oppiin, seestütlev on oppista.*

Samuti võis sihtkeeles normikohane abivorm põhjustada normivastase tulemuse. Algselt sihtkeeles normide kohaselt kulgenud produtseerimisprotsessis toimus pärast abivormi kasutamist muutus ning tulemuseks oli lühendatud ja lihtsustatud tüvega mitmuse seestütleva vorm.

(19) ee *kerros kerro/ksi/a kerroksiin kerros kerroseta kerroista*

Seega on alust oletada, et õppijate keeletajus on sihtkeele sama sõna eri käändevormid omavahel seotud ehk teisisõnu, õppijatel on oma arusaam käändeparadigma-sisestest seostest. See seostevõrk ei ole ilmingimata kõikidel õppijatel ühesugune, lõplikult välja kujunenud ega kõiki käändevorme hõlmav.

(20) ee *nirkai/sia : nirkai/sissa : nirkaista*
ee *kankai/si/a : kankai/ssa : kankai/sta*
ee *nirkkeja : nirkkoissa : nirkkoista*
ee *koh/toja : kohtoihin : kohteista*
ee *patsoja : patsissa : patsista*
vn *kohde- koh/de/ta : koh/de- koh/de/si- siin : kohtei- koh/tei/sta*
vn *veu- veulei- veu/lei/ta : veulei//hin : veu- velulei/sta*

Lähte- ja sihtkeeles samamoodi käänduvate sõnade puhul toimis abivormina lähtekeelne vorm juhul, kui õppija oli enda jaoks teadvustanud lähte- ja sihtkeele käänamise samasuse. Seejuures lähtekeelne abivorm võis, kuid ei tarvitsenud ka lähte- ja sihtkeele käänamismallide konvergentsi korral ilmingimata viia sihtkeeles normikohase tulemuseni (21); põhjuseks õppijate kartus moodustada lähtekeelega liiga sarnane vorm, mis võib osutada “petlikuks sõbraks” (Remes 1995: 7).

(21) *kangas, see on eesti keelega sarnane. kangaita kangastest kangai- kangaista ei kangaista*

kangas, eestikeelne sõna kango- kangoia kanga- kan/goi/sse kangas kangaista kangoista

Kui lähtuda sellest, et keele muutvorme tajutakse konkreetsete keelenditega lahutamatus seoses olevate kujundite (skeemide, mallide, paradigmade) abil (Bybee 1988, 2001: 19–27), siis on loomulik eeldada, et sihtkeele eri omandamisastmel olevate õppijate paradigmad erinevad üksteisest, koosnevad eri komponentidest ja nendevahelistest seostest. Õppijate produtseeritud paradigmadest ilmnes, kuidas käänetevahelisi suhteid tajutakse ja kuivõrd süstemaatiline või juhuslik käänamisprotsess on, ehk H. Paunose (1983: 61) sõnul, kuidas toimub keelesüsteemi sünd alati uue keelekõneleja keeletajus. Oli õppijaid, kelle paradigmad vastasid täielikult sihtkeele paradigmadele, kuid oli ka õppijaid, kes olid alles loomas uut grammatiliste kategooriate võrgustikku. Viimaste paradigmad sisaldasid lisaks sihtkeelsetele elementidele erinevaid sihtvormi seisukohalt normivastaseid tüve- ja sufiksivariantide kombinatsioone, lühendatud ja lihtsustatud tüvevariante ja lähtekeele käänamissüsteemi elemente. Käänamisprotsessi edukus sõltub seega õppija paradigma ja sihtkeele paradigma sarnasusest. Abivorm toimib seejuures käänetevaheliste seoste aktiveerijana ja tugevdajana.

Seega on vägagi tõenäoline, et käändeparadigma sõnavormide ühendajana on psühholingvistiliselt reaalne mõiste nii eesti- kui ka venekeelsete soome keele õppijate jaoks. Mõlema õppijarühma emakeeled ulatuslikult muutemorfoloogiat kasutavate keeltena toetavad sellist arusaama. Iseküsimus on muidugi see, kas see reaalsus on kujunenud kooli (emakeele)õpetuse tulemusena või mitte. Tõenäoli-

selt on kooliõpetus muuteparadigma psühholingvistilise reaalsuse kujunemisele vähemalt kaasa aidanud.

5. Kokkuvõte

Suuline üksiksõnade käänamistest osutus efektiivseks ja kasulikuks uurimismetodiks mitmeski mõttes. Suulisele produtseerimisele omane kiirus sundis õppijaid kasutama produtseerimisviise, mille (tõenäoliselt teadvustamata) lõppeesmärgiks oli protsessi aeglustamine. Käänamisprotsessis esinevate pauside ja korduste hulk, jaotumine ja paiknemine produtseeritavas sõnas andsid õppijate produtseerimisprotsessi kohta sellist teavet, mida kirjalik test ei võimalda. Pauside ja korduste paiknemise põhjal oli võimalik teha oletusi selle kohta, kuidas õppijad tajuvad tüve- ja sufiksivariantide piire ning nende omavahelisi suhteid. Kinnitust sai õpetamiskogemustele tuginev oletus, et tüve- ja sufiksiallomorfide piirid võivad õppija keeletajus olla vägagi ähmased. Kuna muutesufiksitate produtseerimisviis sõltus suuresti sufiksi silbistruktuurist, väljendas õppija jaoks sõnavormi mitmuslikkust sageli ebamäärane häälikujärjend, milles ühinesid nii tüve- kui ka sufikselemendid. Suuline käänamistest tõi kirjalikust testist selgemini esile ka käänamisprotsessi võimaliku mittelineaarsuse ning tõsiasi, et käändeparadigma on õppijate jaoks psühholingvistiliselt reaalne mõiste. Samuti selgus, et lähtekeel mõjutas aeglustamisvõtete kasutamist eelkõige kvantitatiivselt: venekeelsete õppijate suulises testis esines pause ja kordusi enam kui eestikeelsetel õppijatel. Samas ilmnesid mõlema õppijarühma sees pauside ja korduste kasutamisel märkimisväärsed individuaalsed iseärasused.

Suulise käänamistesti tulemuste tõlgendamisel ei tohi siiski unustada, et tegemist on morfoloogiliste vormide produtseerimisega testisituatsioonis, mitte aga õppija poolstruktureeritud või spontaanses keeekasutuses, mis pakuvad omakorda huvitavaid võimalusi edasiseks uurimiseks.

Kirjandus

- Block, E. L. 1992. See how they read: Comprehension monitoring of L1 and L2 readers. – *TESOL Quarterly* 26 (2), 319–343.
- Bybee, Joan 1988. Morphology as lexical organisation. – Michael Hammond, Michael Noonan (eds). *Theoretical Morphology: Approaches in Modern Linguistics*. San Diego: Academic Press, 119–141.
- Bybee, Joan 2001. *Phonology and Language Use*. Cambridge Studies in Linguistics 94. New York: Cambridge University Press.
- Ehala, Martin 1997. Eesti keele struktuur: õpik gümnaasiumile. Tallinn: Kännimees.
- EKG I = Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1995. Eesti keele grammatika. I. Morfoloogia. Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut.
- Gerassimenko, Olga; Hennoste, Tiit; Koit, Mare; Rääbis, Andriela; Strandson, Krista; Vutt, Evelyn 2004. Annotated dialogue corpus as a language resource: an experience of building the Estonian Dialogue Corpus. – *The Baltic Perspective*. 1st Baltic Conference on Human Language Technologies. Riga, Latvia, April 21–22. 150–155.
- Hint, Mati 1983. Eesti keele õpik IX klassile. Tallinn: Valgus.
- Hint, Mati 1997. Eesti keele astmevahelduse ja prosoodiasüsteemi tüpoloogilised probleemid. Tallinn / Helsinki: Eesti Keele Instituut.

- Hokkanen, Tapio 2001. Slips of the Tongue. Errors, Repairs, and a Model. *Studia Fennica Linguistica* 10. Helsinki: Finnish Literature Society.
- ISK = Hakulinen, Auli; Vilkuna, Maria; Korhonen, Riitta; Koivisto, Vesa; Heinonen, Tarja Riitta; Alho, Irja 2004. Iso suomen kielioppi. Toimituksia 950. Suomalaisen Kirjallisuuden Seura. Helsinki.
- Itkonen, Terho 1979. Retkiä nykysuomeen. *Tietolipas* 83. Helsinki: Suomalaisen kirjallisuuden seura.
- Kaivapalu, Annekatrin 2005. Lähdekieli kielenoppimisen apuna. *Jyväskylä Studies in Humanities* 44. Jyväskylä: Jyväskylän yliopisto.
- Kaivapalu, Annekatrin 2006a. Kontrastiivlingvistika ja vöörkeele õppimine. – Helle Metslang, Siiri Soidro (toim.). *Lähivertailuja* 16. Soome-eesti kontrastiivseminar Männikul 8.-9.05.2004. Tallinna Ülikooli eesti filoloogia osakonna toimetised 6. Tallinn: TLÜ Kirjastus, 71–100.
- Kaivapalu, Annekatrin 2006b. Reeglid ja analoogia vöörkeeleõppes soome mitmusevormide käänamise näitel. – Helle Metslang, Margit Langemets, Maria-Maren Sepper (toim.). *Eesti Rakenduslingvistika Ühingu aastaraamat 2. Eesti Rakenduslingvistika Ühing*. Tallinn: Eesti Keele Sihtasutus, 71–92.
- Kaivapalu, Annekatrin 2008. Lähtekeele mõju ja produtseerimiskontekst: eesti õpilased soome mitmusevormide käänamas. – Riho Grünthal, Helle Metslang (toim.). *Lähivertailuja* 18. Suomalais-Ugrilaisen Seuran Toimituksia. Helsinki: Société Finno-Ougrienne, 54–76.
- Kaivapalu, Annekatrin, Martin, Maisa 2007. Morphology in transition: The plural inflection of Finnish nouns by Estonian and Russian learners. – *Acta Linguistica Hungarica* 54 (2), 129–156.
- Karlsson, Fred 1983. Suomen kielen äänne- ja muotorakenne. Porvoo: WSOY.
- Kurhila, Salla 2002. Ei-syntyperäisten puhuma suomi: haaste kielentutkimukselle. – *Moni-Tori* 9:3, 59
- Levelt, Willem J. M. 1983. Monitoring and self-repair in speech. – *Cognition* 14, 41–104.
- Levelt, Willem J. M. 1989. *Speaking. From Intention to Articulation*. Cambridge, Massachusetts: MIT Press.
- Lieko, Anneli 1992. Suomen kielen fonetiikkaa ja fonologiaa ulkomaalaisille. Helsinki: Finn Lectura.
- MacWhinney, Brian 1978. *The Acquisition of Morphophonology*. Monographs of the Society for Research in Child Language Development 43. University of Chicago Press.
- MacWhinney, Brian 1987. The competition model. – Brian MacWhinney (ed.). *Mechanisms of Language Acquisition. The 20th Annual Carnegie Symposium on Cognition*. New Jersey: Erlbaum, 249–308.
- MacWhinney, Brian 1990. Psycholinguistics and foreign language acquisition. – Juhani Tommola (toim.). *Vieraan kielen ymmärtäminen ja tuottaminen*. AFinLA:n vuosikirja 1990. Suomen soveltavan kielitieteen yhdistyksen julkaisuja 48. Turku: AFinLA, 71–87.
- Martin, Maisa 1992. Taipuu, ei taivu... Suomen kielen taivutustesti ulkomaalaisille. Tuloksia ja pohdintoja. – Heikki Nyysönen, Leena Kuure (toim.). *Kielen ja kulttuurin omaksuminen*. AFinLA:n vuosikirja 1992. Suomen soveltavan kielitieteen yhdistyksen julkaisuja 50. Jyväskylä: AFinLA, 267–280.
- Martin, Maisa 1995. The Map and the Rope. Finnish Nominal Inflection as a Learning Target. *Studia Philologica Jyväskyläensia* 38. Jyväskylä: University of Jyväskylä.
- Martin, Maisa; Hihnala, Riikka; Sivonen, Ulla. 2001. Leksikko mielessä. Pieni koe sanojen järjestymisestä mentaaliosassa leksikossa. – Mirjaliisa Charles, Pirjo Hiidenmaa (toim.). *Tietotyön yhteiskunta – kielen valtakunta*. AFinLA:n vuosikirja 2001. Suomen soveltavan kielitieteen yhdistyksen julkaisuja 59. Jyväskylä: AFinLA, 61–80.

- Meister, Einar; Lippus, Pärtel 2007. Spontaanse kõne temporaalsest struktuurist. – Ettekanne Eesti Rakenduslingvistika Ühingu kevadkonverentsil “Keel ja leksikon” 25.4.2007, Tallinn.
- Mihkla, Meelis; Pajupuu, Hille; Kerge, Krista 2005. Statistical modelling of intonation and breaks for Estonian text-to-speech synthesizer. – Robert Vich (ed.). *Electronic Speech Signal Processing. Proceedings of the 16th Conference Joined with the 15th Czech-German Workshop “Speech Processing”*. Prague, September 26-28, 2005. Dresden: TUDpress, 91–98.
- Muikku-Werner, Pirkko 2002. Te käivitte Suomessa? Virolaisten ja saksalaisten oppijansuomen syntaksia. – Stiina Rantatalo, Helena Sulkala (toim.). *Tutkielmia oppijankielestä II. Oulun yliopiston suomen ja saamen kielen ja logopedian laitoksen julkaisuja 21*. Oulu: Oulun yliopisto, 7–43.
- Määttä, Urho 1998. Suomalaisesta kenttämorfologiasta. – Urho Määttä, Klaus Laalo (toim.). *Kirjoituksia muoto- ja merkitysopista. Folia fennistica & linguistica 21*. Tampere: Tampereen yliopiston suomen kielen ja yleisen kielitieteen laitos, 2–23.
- Nenonen, Sari 2001. Venäläiset suomenoppijat suomen kielen pitkien painottomien vokaalien havaitisijoina. – Mirjaliisa Charles, Pirjo Hiidenmaa (toim.). *Tietotyön yhteiskunta – kielen valtakunta. AFInLA:n vuosikirja 2001. Suomen soveltavan kielitieteen yhdistyksen julkaisuja 59*. Jyväskylä: AfInLA, 11–31.
- Niemi, Jussi; Heikkinen, Janne; Järvikivi, Juhani 2001. Miksei *kania* vaikka *munoja*? – Helena Sulkala, Leena Nissilä (toim.). *XXVII kielitieteen päivät Oulussa 19.-20.5.2000. Acta Universitatis Ouluensis B Humaniora 41*. Oulu: Oulun yliopisto, 191–197.
- Niemi, Sinikka 2001. Skeema-analogia, morfologia ja kielenomaksuminen. – Helena Sulkala, Leena Nissilä (toim.). *XXVII kielitieteen päivät Oulussa 19.-20.5.2000. Acta Universitatis Ouluensis B Humaniora 41*. Oulu: Oulun yliopisto, 198–207.
- Pajupuu, Hille; Kerge, Krista 2006. Hingav süntesaator ja pausid tekstiliigiti. – *Keel ja Kirjandus 3*, 202–210.
- Paunonen, Heikki 1983. Allomorfiin dynamiikkaa. – Auli Hakulinen, Pentti Leino (toim.). *Nykysuomen rakenne ja kehitys. 1. Näkökulmia kielen rakenteisiin. Tietolipas 93*. Helsinki: Suomalaisen kirjallisuuden seura, 57–85.
- PS = Suomen kielen perussanakirja. I–III. Kotimaisten kielten tutkimuskeskus. Helsinki: Edita, 2001.
- Remes, Hannu 1995. Suomen ja viron vertailevaa taivutustypologiaa. *Oulun yliopiston suomen ja saamen kielen ja logopedian laitoksen julkaisuja 2*. Oulu: Oulun Yliopisto.
- Swain, Merrill; Lapkin, Sharon 1995. Problems in output and the cognitive processes they generate: A step toward second language learning. – *Applied Linguistics 16 (3)*, 371–391.
- Suni, Minna 2007. Awareness of second language inflectional morphology: A case study on Finnish as second language. – *Acta Linguistica Hungarica 54 (2)*, 217–235.
- Suni, Minna 2008. Toista kieltä vuorovaikutuksessa: Kielellisten ressurssien jakaminen toisen kielen omaksumisen alkuvaiheessa. *Jyväskylä Studies in Humanities 94*. Jyväskylä: Jyväskylän yliopisto.
- Õunapuu, Toom 2003. Sõnavorm ja kõne. 7. klassi eesti keele õpik. Tallinn: Koolibri.

Annekatriin Kaivapalu (Tallinna Ülikool) on lõpetanud Tallinna Pedagoogilise Instituudi eesti keele ja kirjanduse erialal. Kaitstes samas ülikoolis magistriskraadi soome filoloogial 1999. a ja 2005. a doktoriskraadi teise keele omandamise alal (soome keel) Jyväskylän ülikoolis. Alates 2006. a töötab Tallinna Ülikoolis soome keele dotsendina. On uurinud kõrvutavalt eesti ja soome muutemorfoloogiat, eesti ja soome üliõpilaste arvamusi emakeelest, lähteelele mõju sihtkeele morfoloogia omandamisele, absoluutset ja suhtelist morfoloogilist kompleksust, eesti ja soome õppijakeele korpustega seonduvat. kaivapa@tlu.ee

WHAT DO WORD-INTERNAL PAUSES, CUT-OFFS, AND RETRACINGS TELL ABOUT INFLECTIONAL PROCESS?

Annekatrin Kaivapalu

Tallinn University

The oral processing of inflectional forms, in comparison with writing, is characterized by shorter processing times. The pauses, cut-offs, and retracings are a natural part of the oral producing process of native speakers and language learners alike. The paper argues that the occurrence, frequency and location of pauses, cut-offs and retracings within inflectional forms are not coincidental, but give valuable information about psycholinguistics processes. The argument is based on a study of morphological production of Finnish plural nominals by Estonian and Russian learners. The paper discusses the production in oral inflection tests of context-free words. The second data-type consists of the learners' comments on how they process the same words in a written test, and the retrospective interviews with the learners. Both data-types were used to explore the following questions:

1. How do the learners perceive the morphological formatives?
2. Is the notion of the inflectional paradigm psycholinguistically real for the learners?
3. Are the inflectional processes of the learners always linear?
4. How does the learners' first language influence the use of the pauses, cut-offs and retracings?

The study concludes that the borders of the morphological formatives are quite indistinct in the minds of the learners and that the perception and production of formatives depends mostly on the syllabic structure of the formative. The data also seems to confirm the assumption that the inflectional paradigms are psycholinguistically real in the learners' perceptions, but there are individual differences and variance by the L2 proficiency. The inflectional process is not necessarily always linear, and the interaction between stem and suffix allomorphs has an important role in processing inflectional forms. The use of the pauses, cut-offs and retracings is more determined by the personal characteristics of learners than by the first language influence.

Keywords: psycholinguistics, morphological processing, oral production, inflection test, context-free words, Finnish as a foreign language, Estonian and Russian learners

KÕNETEHNOLOOGIA VAJAB ŽANRILIST LÄHENEMIST

Krista Kerge, Hille Pajupuu,
Kairi Tamuri, Heidi Meier

Ülevaade. Oleme varasemas uurimuses näidanud, et ettelugemisel seostuvad pausid ja hingamine kui ette loetud teksti loomulikkuse tunnused tugevasti funktsionaalstiiliga (ajakirjandustekst, ilukirjandustekst). Tekstiuurimuse andmestik aitab muuta loomulikuks sünteeskõne, kuid ka tuvastada tekstiliiki ehk žanrit, eristada funktsionaalstiile ja autorite sõnastuslaadi ning muuta otstarbekamaks sõnastiku ja grammatika kasutuse teksti automaatses analüüsis. Artiklis osutatakse hingamise ja pausidega seostuvale ja teistele eesti teksti formaliseeritavate parameetrite uurimustele, mida keeletehnoloogia arenduses seni arvestatud ei ole. Varasema kontrollimiseks võrreldakse ilukirjandusteksti kahe žanri ettelugemise pause ja hingamist ajakirjandus uudise seniste andmetega. Tulemused kinnitavad, et pauside ja sissehingamise kestus ning nende seotus lause süntaktilise liigendusega erineb mitte ainult funktsionaalstiilide, vaid ka žanrite kaupa. Süntaktilist liigendust kannavad seejuures peamiselt kirjava-hemärgid. Ilukirjanduse puhul mängib pauside tekkes kaasa teksti lugejapoolne interpretatsioon. Tekst–kõne-sünteesis, mis on mõeldud tekstide ettelugemiseks, tuleks pauside ja hingamise genereerimisel lähtuda pikema ajakirjandusliku uudise kui muudeltki parameetritelt neutraalse tekstiliigi andmetest.*

Võtmesõnad: pausid, hingamine, funktsionaalstiil, žanr, uudised, ilukirjanduskeel, kõnesüntees, eesti keel

* See artikkel on valminud tänu Eesti Teadusfondi grandile nr 6742 ja riiklikule programmile "Eesti keele keeletehnoloogiline tugi".

Tekstiliikide parameetrid ja kõnetehnoloogia

Juba sünteeskõne intonatsiooni ja pauside modelleerimisega seoses (vt Mihkla jt 2005) tekkis huvi, kas ja kuidas sõltuvad ette loetud teksti paralingvistilised muutujad, näiteks hingamine ja pauseerimine keelekasutuse valdkonnast, millega seostuvad funktsionaalstiilid ehk allkeeled, ja žanrist, mis on kindla registri tekstiliigiline väljund. Et selline sõltuvus on olemas, sai kinnitust ette loetud teksti pauside ja hingamise detailsel analüüsil (vt Pajupuu, Kerge 2006).

Keeleroumi formaalselt mõõdetava leksikogrammatilise ja paralingvistilise varieeruvuse ideoloogiat ning seniste tekstiuurimuste andmestikku ei ole eesti kõnetehnoloogias kuigivõrd rakendatud, see võiks aga kasulikuks osutuda (nt Barnbrook 2007). Esimene kohalik katse nimetatud suunas oligi eelviidatud hingamise ja pauside määramine ajakirjandus- ja ilukirjandustekstis, kus vaatluse all olid pikemad raadiouudised ning ette loetud kriminaalromaan (Mihkla jt 2005, Pajupuu, Kerge 2006). Seda uurimust on ilukirjanduse poolel laiendatud armastusromaaniga (Tamuri 2007) ning võetud tulemused hiljem kokku (Kerge jt, ilmumas). Viimast uurimust tutvustame allpool üksikasjalikumalt.

Ettevalmistatud suulisele žanrile, nimelt loengule keskendub eesti suulise keele uurimist alustav kirjutus (Hennoste 1994); ettevalmistatud spontaanse kõne ajaorganisatsiooniline analüüs on toonud loengužanri taas päevakorradele (Meister, Lippus, ilmumas). Seda kõike on aga ilmselgelt vähe, kui võtta arvesse andmeid, mida kirjalike või kirjalikult ette valmistatud žanrite kohta on teada – just sellistele žanritele võiks kõnetehnoloogia arendus toetuda, seades sihiks keele, sealhulgas kirjakeele jätkusuutlikkuse, mille võti on Mart Rannuti (2003) järgi kõnetehnoloogia, k.a asjalik suuline arvutisuhtlus.

Tekstiuurimise andmestikku peaks peale kõnesünteesi vajama ka teksti automaatne analüüs jm keeletehnoloogilise rakenduse valdkonnad. Eesti teksti uurimusi, mis võiks seejuures pakkuda vähemalt tuge või ideid, on enamgi (vt üle vaadet allpool). Žanri tuvastamine aitaks kohati piirata süntaktilisel automaatanalüüsil või tõlkimisel kasutatavat leksikoni ja selle üksuste tõlgendusvõimalusi, seades ühtlasi tõenäosusjärjekorda grammatiliste vormide-mallide kasutuse ja tõlgenduse kindla allkeele ja sellele omase leksika kontekstis. Žanriomaseid lingvistilisi tunnuseid ei tule tekstiliigi tuvastuse juures arvesse võtta süsteemselt, piisaks kõige tugevamini korreleeruvate tunnuste koosinemuse modelleerimisest. Tekst–kõne-sünteesis võib omakorda toetuda keskmiste näitajatega ehk n-ö neutraalselt mõjuvatele žanritele (mis on odavam lahendus ja ilmselt vastuvõetav näiteks teksti sisu kiirest hõlmamisest huvitatud nägemispuudega inimestele). Süntesaatori kasutusala laienedes on võimalik rakendada žanri kiirtuvastust teksti põhjal ning programmeerida mitu lugemis- või asjaliku kõne mudelit, mis võtavad arvesse ette loetava või loodava teksti liiki seal, kus spontaansele kõnele toetumine mõjuks argikeelsena. See aga on ideaal n-ö pikemas plaanis, kui peaks tähtsustama suuline suhtlus arvuti ja tema kasutaja vahel. Rõhutatagu siinkohal, et spontaanse kõne analüüsiga ei ürita siinkirjutajad tegelda ega toetu niisiis ka sellekohastele uurimustele, mida Eestis on juba rohkesti.¹

Eesti tekstiliikide uuritud formaliseeritavad parameetrid

Nüüdiseesti allkeelte – nende nimi on kõnetehnoloogia ja korpuslingvistika ingliskeelses kirjanduses sageli *text-type* või *functional style* – süntaks muutub seniste uurimuste järgi ajas (vt Kerge 2002b: 44 jm, uustrukk Kerge 2003a [4]) ning erineb tugevasti ka süntaksplaanis (vt Kerge 2002a). Sajandivahetuse allkeelte parameetrite keskmisi esindab tekstilause pikkuse ja sisemise keerukuse, samuti teksti abstraktsust tõstvate *mine*-tarindite osatähtsuse poolest ajakirjandustekst. Näiteks kõigub eesti tekstilause pikkus u 9 ja 25 sõne vahel, ajakirjandusteksti lause on aga u 14 sõnet pikk. Kirjavahemärke ja sidendeid, mis viitavad lause tinglikule liigendatusastmele, on allkeeliti u 1–4, ajakirjanduses u 1,5. Regulaarset *mine*-nominalisatsiooni esineb teksti sõnede hulgas u 1–3,5%, ajakirjanduskeeles 1,8% (vt Kerge 2003a: 40–45).² Analoogilisi uurimusi on erakondade valimisplatvormide kohta (Rääk 2002), seaduste kohta (Rätsep 2003) ning autoristiili kohta žanriti (Kerge 2003b, Pöld 2007).

Lisaks on hulk muid lingvistilisi tunnuseid, mis eristavad tekstiliike ning ühtlasi iseloomustavad teksti laadi³ (s.o jutustavat, kirjeldavat, seletavat ja argumenteerivat teksti). Heidi Meieri (2003a) uurimuse järgi (vt ka Meier 2002, 2003b) on igale tekstiliigile omased tugevalt korreleeruvate tunnuste kimbud. Kuna tunnused on mehhaaniliselt tuvastatavad, teeb see võimalikuks tekstiliigi automaatse määramise.

H. Meier on allkeelte kaupa uurinud näiteks isikuliste asesõnade, lühendite, *mine*-tuletiste, 2–4- ja enam kui 14-tähele sõnade esinemissagedust; sõnaalgulise *g, b, d* ning võõrtähtede järgi hõlpsasti tuvastatavate võõrsõnade sagedust; rinnastavate ja alistavate sidesõnade esinemust ja sagedust; tekstiliikide sagedasimaid sõnavorme jpm, andes üsna hea ülevaate eesti tekstiruumi allkeelte parameetritest. Kõik parameetrid on ülevaatlikult esitatud ka tabelitena. Žanritest on uuritud esseed. (Täpsemalt Meier 2003a.)

Kogu see andmestik võiks kasulikuks osutada ka eesti kõnetehnoloogia rakendustes või pakkuda selles valdkonnas just žanrituvastuse ideid. Sama eesmärgi täidab nt Ülle Viksi ja Indrek Heina (2001) seadustekstianalüüs. Tuvastuskriteeriumide edasisel valikul pakuvad tuge ja eeskuju ka Douglas Biberi jt korpuseuuringud, mille põhjal tekstid oma retoorilise tüübi järgi moodustavad eelmainitud tekstilaadiga sarnaseid dimensioone (vt nt Biber 1995, Biber, Conrad 2001 vm).

Eestis on žanriti uuritud ka sõnaliigisuhteid. Seaduse andmeid kajastavad Ü. Viks ja I. Hein (2001). Teksti loomulikkusuuringute raames on sõnaliigisuhteid uurinud Krista Kerge, Hille Pajupuu ja Rene Altrov (2007), kes vaatlevad Frances Heyligheni ja Jean-Marc Dewaele (2002) meetodil, kui suur osatähtsus on nimisõnadel ning nimisõnaga seostuvana omadus- ja kaassõnadel *versus* muudel sõnaliikidel. Sõnaliigisuhete järgi on arvatud tekstiliigi formaalsusindeksid, mille järgi iga tekst paigutub formaalsuse–kontekstuaalsuse kontiinumis: nimisõnade ja nendega seotud sõnaliikide suur osatähtsus viitab teksti suuremale formaalsusele ja ühetimõistetavusele; koos muude sõnaliikide osatähtsusega kasvab kontekstuaalsus ja mitmetimõistetavus (vt Heylighen, Dewaele 2002). Uurimusega on eesti žanri-

² *mine*-tuletiste osatähtsusel on õigustekst esindatud äärmusliku näitajaga 6,8%, mida sama uurimuse järgi u 30% ulatuses võiks vältida kui asjatut nominaalstiili.

³ Inglisekeelses kirjanduses viidatakse teksti laadile terminiga *text-type*, saksakeelses *Textsort*, kuid termin *teksti laad* näib eesti keeles kõige paremini viitavat žanriülestele kvalitatiivsetele erijoonetele. *Tekstitüüp* kui üldsõna on olnud vabamas kasutuses juhul, kui ei ole põhjust kasutada täpse sisuga termineid (nt dialoog on kord suuline, kord kirjalik, kord intervjuužanr, kord kauplusedialoog, kord vaba vestlus). Samamoodi kasutatakse sõna *tekstitüüp* siinses artiklis.

test siiani hõlmatud kirjalik essee ning asjalik suuline monoloog ja dialoog (Kerge jt 2007), õigusaktid (Kerge 2006); hulk kirjavahetust: elektroonilised ametikirjad töötajate vahel ja asutusest välja (Rais 2007), ametikirjad asutuse sees, asutuste vahel ja asutuselt kodanikele, samuti poiste ja tüdrukute meilid (Kerge 2006); lisaks suuline hoidedialoog (Kerge 2006) ja parlamendikõned (Pajupuu 2007). Eksamiessede ning tigukirja- ja meilivahetuse taustal on nüüdseks uuritud ka isikustiili, täpsemalt Jaan Kaplinski esseistikat ja kirjavahetust (Pöld 2007). Žanriti erineb selgelt nii teksti tihedus (nimisõnade osatähtsus) kui ka kontekstuaalsus. Sellekohasedki andmed võiksid leida laiemat rakendust.

Loomulik kõneliigendus ja žanriuuringute tähtsustumine

Kõnesünteesi kontekstis on Eestis viimastel aastatel tõsiselt tegeldud lauseintonatsiooni loomulikustamisega, uurides selle reeglistamisvõimalusi (Mihkla jt 2003), ning jõutud statistiliste hinnangute eelistamiseni (Mihkla jt 2004). Ka pause ja pausieelseid pikendusi kui kõneliigenduselemente on üritatud mitmel meetodil modelleerida (Mihkla 2005, Mihkla, Kuusik 2005).

Loomulikus kõnes on samas iseenesestmõistetav ja enamasti ka kuulda hingamine. Täiesti ilmselt ei saa hingamispause näiteks sünteeskõnesse lisada juhuslikult, arvestada tuleb, et hingamispausid on seotud füsioloogiliste teguritega: ühest sissehingamisest teiseni peab kuluma niipalju aega, kuni sissehingatud õhuga rääkida saab. Seega ei tohiks sünteeskõne sissehingamissagedus erineda oluliselt inimese omast: kõne kiirust arvestades tuleks määrata hingamispausidele sobiv vahemaa, leida neile õige koht ning õige kestus. Taustal peetakse silmas, et tegemist on tekste ette lugeva ehk *Text-to-Speech* tüüpi (TTS) süntesaatoriga.

Mõte süntesaator hingama panna tuli kümnekond aastat tagasi Douglas H. Whalenilt (1994), kes leidis, et hingamise lisamine teeks inimestele teksti jälgimise kergemaks: ilma hingamiseta sünteeskõne tüütab kuulajat, mõjub üksluiselt ning ebaloomulikult. Ka Nick Campbell on kõnesüntesaatorite puudusena täheldanud võimetust anda edasi selliseid kõnes esinevaid helisid nagu naer või hingamine. N. Campbelli arvates tõstaks nende lisamine sünteeskõne vastuvõetavust oluliselt (Campbell 1998). Ometigi suur osa tänapäeva süntesaatoreid ei naera ega hinga.

Hingamist kõnes on uurinud François Grosjean ja Maryann Collins (1979) ja toonud välja kolm olulist aspekti:

- 1) kõneleja ühendab oma hingamismustri lause planeerimisega, st kohandab füsioloogilised vajadused lingvistiliste vajadustega, et tagada kõne sora-vus;
- 2) kõneleja kontrollib oma hingamist, st ta jaotab kõne regulaarsetesse hingamis-rühmadesse;
- 3) kõneleja hingamine on sõltuv süntaksist, st kõneleja hingab ainult siis, kui lausungi moodustajastruktuur seda lubab.

F. Grosjeani ja M. Collinsi (1979) järgi pole pausid määratud otseselt vajadusega hingata. Küll aga sõltub hingamine pausidest. Sageli teevad kõnelejad pikemaid pause süntaktiliselt olulistes kohtades, nt fraasipiiril, kus esinevad tavaliselt just

hingamispausid. Kuigi tegemist on suhteliselt ammuse uurimusega, pole rakenduslingvistikas (sh kõnesünteesis) seda teadmist ära kasutatud.

Kõnesünteesiga seoses on pauside olemust uurinud Chiu-Yu Tseng (2002) ja leidnud, et kõnes organiseeruvad pausid hierarhiliselt: kõige pikemad pausid markeerivad prosoodilise rühma piire.⁴ Prosoodiline rühm on võrdne vähemalt ühe hingamisrühmaga, kuid võib neid sisaldada rohkemgi. Hingamisrühma sees on omakorda lühemaid hingamiseta pause (Tseng, Chou 1999). Neist kahest uurimusest saab küll andmeid kirjavahemärkide kohale langenud pauside kestustest, ent ei saa teada, kuhu lugedes pause tegelikult tehakse. C.-Y. Tseng (2002) küll ütleb, et pausid võivad, kuid ei pruugi kokku langeda kirjavahemärkidega, ent ei käsitle seda teemat sügavamalt: me ei saa teada, kui palju pause langeb kokku kirjavahemärkidega, või millised neist on hingamispausid, millised mitte. Hingamispause ei uuri C.-Y. Tseng üldse lähemalt, näiteks ei saa me teada, kui suure osa hingamispausist moodustab sissehingamine, kui suure osa vaikus.

Pidades silmas mitut eesmärki, nt kõne loomulikkuse mõõtmist ja hindamist ning kõnesünteesi loomulikkuse tõstmist, püüdsime nendele küsimustele vastused leida, defineerides hingamis- ja muud pausid ning arvestades uurimise juures žanrilist eripära. Uurimuste rida puudutab kahe allkeele, ajakirjandus- ja ilukirjanduskeele ettelõetud teksti.

Pauside pikkust, sagedust ja asukohta eri tüüpi tekstides (spontaanne dialoog, professionaalne ja mitteprofessionaalne ettelõetud uudis) on rakenduslikul eesmärgil uurinud ka Sofia Gustafson-Čapková ja Beáta Megyesi (2001). Uurimusest järeldub, et pauside pikkus ja sagedus sõltub tekstitüübist: viimastest moodustub kontiinum, mille ühte suunda jääb dialoog (pausid pikemad, sagedamini esinevad) ja teise professionaalne uudiste ettelugemine (pausid lühemad, harvem esinevad).

Valdavalt tehakse pause olulistes kohtades, nt süntaktilistel piiridel ja semantiliselt tähtsate sõnade juures (viimased olid sõnumile aluseks). Monoloogides varieerivad kõnelejad tavaliselt pauside pikkust ja asukohta vastavalt informatsiooni struktuurile (Gustafson-Čapková, Megyesi 2001).

Gunnar Fant kolleegidega (2002, 2003) on uurinud pauside kestust ilukirjanduse ja uudiste lugemisel. Tulemused näitavad, et kõik uudiste pausid on ilukirjanduse omadest lühemad.

Kuigi kumbki vahetult nimetatud uurimustest ei tegele hingamispausidega, annavad need uurimused selge vihje vajadusele läheneda kõne uurimisele žanrist lähtuvalt.

Seni avaldatud selleteemalised eesti tulemused (Pajupuu, Kerge 2006) osutavad, et nii pausid kui ka hingamine on seotud teksti kompositsiooniga, s.o teksti osade ja neis sisalduvate lõikude ja lausete piiriga, ning tekstilause süntaktilise liigendusega, mida eesti keeles enamasti markeerivad kirjavahemärgid. Alustavalt võeti vaatluse alla kaks ettelõetud teksti liiki, pikemad ajakirjanduslikud uudised ning võrdluseks ilukirjandusest kriminaalromaan. Tulemused suunasid väitma, et sünteeskõne pause ja hingamist tuleks modelleerida ajakirjandusliku uudise andmetele toetudes (Pajupuu, Kerge samas). Et just ilukirjandusteksti andmed osutusid ootamatuks, suurendati uurimise järgmises etapis ilukirjandusliku materjali mahtu ja lisati armastusromaan (Tamuri 2007). Selle uurimuse tulemusi järgnevas ka tutvustatakse.

⁴ Prosoodiline rühm on kõneüksus, mis koosneb vähemalt ühest hingamisrühmast (sissehingamisest sissehingamiseni). Kirjalikus tekstis vastab sellele lõik (graafiliselt iseloomustab seda taane). Sama fenomen ilmneb ka kirjalikku teksti lugedes.

Pausid ja hingamine romaani allžanrites

Allpool on täpsustatult uuritud pauside paiknemist ja kestust etteloetud ilukirjandustekstis, täpsemini kahes romaani allžanris (armastusromaanis ja kriminaalromaanis) katkendid, vastavalt 19 lauset, 413 sõnet ja 39 lauset, 647 sõnet). Ilukirjandustekstid salvestati vaikselt ruumis Edirool R-09-ga loetuna professionaalsete näitlejate poolt (1 mees ja 1 naine, mõlemad lugesid samu tekste). Ajakirjandusliku uudise andmed pärinevad varasemast (pikemate uudiste terviklikud katkendid, 45 lauset, 666 sõnet), tekstid on Eesti Raadio professionaalsete uudisteluugejate studiosalvestised, reaalselt loetud uudised (1 naine ja 1 mees, mehe ja naise tekstid olid erinevad). Ideaalis võinuks materjal hõlmata rohkem žanreid ja rohkem ettelugejaid, ent pidades silmas oma eemärki – mõõta kõne loomulikkust ja pakkuda lähtematerjali sünteeskõne pauside genereerimiseks – pidasime piisavaks tüüpilistele žanritele keskendumist. Lähtusime seisukohast, et iga emakeelekõneleja loetud teksti võib pidada loomulikuks, selgitamaks välja pauseerimise üldisi tendentse.

Kõnes mõõdetud pausid rühmitasime kaheks: 1) hingamispausid; 2) hingamiseta pausid. Hingamispaus koosneb sissehingamisest, millele eelneb ja/või järgneb vaikus. Hingamiseta pausiks lugesime vähemalt 30 ms kestva vaikuse. Praat 4.5 (Boersma, Weenink 2006) kasutades mõõtsime mõlemas tekstiliigis pausi kestuse, hingamispausis lisaks ka sissehingamise kestuse. Hingamispauside keskmised kestused tekstiliigiti on esitatud tabelis 1. Otsisime ka pauside ja süntaksi seost, arvestades seda, et süntesaator tunneb ära ennekõike graafiliselt markeeritud lauseliigenduse (kirjavahemärgid, lõigupiir, kindel kirjavahemärgita sidend, nimi jne) ja et praktikas on näiteks kirjavahemärgiõpetust omakorda seotud intonatsiooni ja pauseerimisega.

Tabel 1. Hingamispausi ja sissehingamise keskmine kestus tekstilugejatel tekstiliigiti⁵

Meeshääl			Naishääl		
Paus t/sd (ms) ⁶	Sisse- hingamine t/sd (ms)	Sissehingamise osatähtsus pausi kestuses	Paus t/sd (ms)	Sisse- hingamine t/sd (ms)	Sissehingami- se osatähtsus pausi kestuses
AJAKIRJANDUSTEKST: UUDIS					
617/262	276/112	45%	597/256	269/99	45%
ILUKIRJANDUSTEKST: ARMASTUSROMAAN					
590/234	380/145	64%	606/240	374/174	62%
ILUKIRJANDUSTEKST: KRIMINAALROMAAN					
675/345	449/259	66%	780/482	362/222	46%

Ajakirjandustekstis ei erine hingamispauside keskmised kestused dimensioonil meeshääl–naishääl. Hingamise osatähtsus hingamispausis on 45%. Sellest sarnasusest joutuvalt käsitleme edasises ajakirjanduspauside analüüsis mees- ja naishäält koos.

Ka etteloetud armastusromaanis ei erine hingamispauside keskmised kestused dimensioonil meeshääl–naishääl: hingamispauside kestus on sarnane uudiste omadele, oluliselt erineb aga hingamise osatähtsus pausi kogukestusest, mis on üle 60%. Kui uudistes püütakse sisse hingata võimalikult kiiresti ja isegi kuuldamatult, siis ilukirjandust lugedes sellist vajadust pole, vastupidi, sissehingamine on sageli

⁵ Välja on jäetud prosoodilise rühma ehk lõigupiiri pausid kui eriliselt pikad. Nende keskmine kestus on antud tabelis 3.

⁶ t on pausi kogukestus või sissehingamise kestus, sd on standardhälve.

kuuldav ja pikk. Seega võime oletada, et hingamine on ilukirjandust lugedes mee-
leolu loomise teenistuses.

Kriminaalromaaani hingamispausid on armastusromaaani pausidest märkimis-
väärsetl pikemad, suurem on ka standardhälve. Võimalik, et erineva pikkusega
hingamispausidega lisatakse loetavasse põnevust. Selles žanris tuleb esile ka mees-
ja naislugeja erinevus: mees hingab sisse naisest oluliselt pikemalt (sissehingamise
osatahtsus vastavalt 66% ja 46%). Et armastusromaaani lugedes sellist erinevust välja
ei tulnud, siis ilmselt pole põhjus lugejate füsioloogilises erinevuses (nt naise väik-
semas kopsumahus), vaid pigem näitlejate erinevates põnevuse tekitamise võtetes:
meesnäitleja hingamispausi täidab suuresti sissehingamine, naisnäitleja oma aga
vaikus. Ilukirjandusžanriti esile tulnud lugejaerinevuste tõttu ja põhjusel, et nii mees
kui ka naine loevad samu tekste, käsitleme mees ja naishäält edaspidi lahus.

Pauside ja süntaksi seose analüüsi tulemused on esitatud tabelis 2.

Tabel 2. Pauside vastavus kirjavahemärkidele

Kirjavahemärk	AJAKIRJANDUSTEKST: UUDIS (meeshääl + naishääl)		
	Kirjavahemärkide arv tekstis	Pauside arv kirjavahemärgi kohal / neist hingamispause	
punkt lõikude vahel (kirjapildis taane)	14	14/14	
punkt lausete vahel	31	31/31	
koma	37	27/13	
mõttekriips	8	8/6	
koolon	6	6/5	
semikoolon	1	1/1	
paus kirjavahemärgita (ekstrapaus)		19/5	
Kirjavahemärk	ILUKIRJANDUSTEKST: ARMASTUSROMAAN		
	Kirjavahemärkide arv tekstis	Pauside arv kirjavahemärgi kohal / neist hingamispause	
		meeshääl	naishääl
punkt lõikude vahel (kirjapildis taane)	12	12/6	12/11
punkt lausete vahel	7	7/4	7/3
koma	45	16/8	20/12
mõttekriips	3	3/2	3/3
koolon	2	2/1	2/2
semikoolon	6	5/4	6/4
paus kirjavahemärgita (ekstrapaus)		49/22	9/0
Kirjavahemärk	ILUKIRJANDUSTEKST: KRIMINAALROMAAN		
	Kirjavahemärkide arv tekstis	Pauside arv kirjavahemärgi kohal / neist hingamispause	
		meeshääl	naishääl
punkt lõikude vahel (kirjapildis taane)	5	5/0	5/2
punkt lausete vahel	31	31/17	31/21
koma	58	12/7	31/11
mõttekriips	9	7/5	9/4
koolon	1	0	1/1
semikoolon	0	0	0
paus kirjavahemärgita (ekstrapaus)		41/24	23/7

Kirjavahemärkide seos pausidega on ilmne (vt tabel 2), üldjuhul tehakse nii ajakirjandusteksti kui ka ilukirjandusteksti lugedes kirjavahemärgi kohale paus. Kirjavahemärkidest tehakse kõige harvem paus koma kohal. Tekstiliikide erinevus on siiski märkimisväärne: ajakirjandustekstis tehakse paus 90% kõigist kirjavahemärkidest, seejuures 80% neist pausidest kasutatakse sissehingamiseks. Punkti ja lõigupiiri kohal tehakse ajakirjandustekstis alati paus ja hingatakse alati ka sisse, kuid ilukirjandustekstis ei ole ükski kirjavahemärk “kohustuslik” pausi ja sissehingamise koht. Isegi lõigupiir (mis ajalisel on väga pikk) võidakse ületada sissehingamata. Ilukirjandustekstis langeb paus kirjavahemärgiga kokku umbes pooltel juhtudel, erinevused ilmnevad nii žanriti kui ka ettelugejati.

Armastusromaani puhul teeb meesnäitleja pausi kõigist kirjavahemärkidest 60% kohal ning kasutab neist 55,5% sissehingamiseks. Naisnäitleja teeb pausi kõigist kirjavahemärkidest 67% kohal ja kasutab neist 70% sissehingamiseks. Kriminaalromaani puhul teeb meesnäitleja pausi kõigist kirjavahemärkidest 53% kohal ning kasutab neist 53% sissehingamiseks. Naisnäitleja teeb pausi kõigist kirjavahemärkidest 74% kohal ja kasutab neist 51% sissehingamiseks.

Ilukirjandust lugedes tehakse erinevalt ajakirjandustekstist palju nn ekstrapause (pausid mujal kui kirjavahemärkide kohal). Nende hulk ei sõltu niivõrd žanrist, kuivõrd teksti ettelugejast. Uudiste lugemisel moodustavad ekstrapausid 18% kõigist pausidest, meesnäitlejal armastusromaani lugedes 50,5% ja kriminaalromaani lugedes 40% kõigist pausidest, naisnäitleja teeb ekstrapause vähem: armastusromaanis 15% ja kriminaalromaanis 22% kõigist pausidest. Ajakirjandustekstis on ka ekstrapausid suhteliselt reeglipärased: neid tehakse sidesõnade *ja/ning*, nime- ja numbrite ees. Ilukirjandustekstis on ekstrapauside tekkekohad seevastu peamiselt sisulised: ilma teksti sisu mõistmata on nad reeglitamatud.

Tekstide lugemisel võib peamiseks žanrierinevuseks pidada niisiis seda, et uudiselugejad järgivad pause tehes küllaltki täpselt teksti vormiseiku (kirjavahemärke, numbritega kirjutatud arvsõnu, suurtähelisi nimesid), ilukirjandustekstide lugejad aga lähtuvad pigem teksti sisust kui vormist.

Ajakirjandusteksti võib seega tõenäoliselt võtta kui etteloetava teksti etaloni, millele toetada teistegi pragmaatilise orientatsiooniga ehk tarbetekstide esitus – ajakirjandus esindab eesti tekstiruumi keskmisi näitajaid (vt Kerge 2002a, 2003a) ja mõjub kõige neutraalsemalt, millise tahes žanri esitusega oleks tegemist.

Ajakirjandusteksti liigendust, punktuatsiooni, lugeja hingamist ja ekstrapause arvesse võttes moodustub kuus kestuselt erinevat pausirühma, mis – nagu Tseng'i ja tema kolleegi uurimusteski (Tseng 1999, Tseng, Chou 2002) – on hierarhilist laadi (vt tabel 3). Nende pausirühmade kestuste ligikaudseid väärtusi oleme soovitanud aluseks võtta neutraalsust taotleva sünteeskõne genereerimisel (vt Kerge, Pajupuu 2006).

Tabel 3. Pausirühmade keskmised kestused ajakirjandustekstis

Pausirühmad	Pausi keskmine kestus (ms)	Hingamise keskmine kestus (ms)
prosoodilise rühma paus (lõigupiir)	988	330
punktipaus	613	273
kooloni-, semikooloni-, mõttekriipsupaus	486	227
komapaus sissehingamisega	398	222
komapaus sissehingamiseta	170	–
ekstrapaus (<i>ja/ning</i> , nime, numbri ees)	69	–

Kokkuvõte

Uurimus näitab, et hingamispausid on vähemalt ettelugemise juures tugevasti süntaktiliselt tingitud siis, kui tegemist ei ole markantselt ekspressiivse tekstiga. Žanrilised erinevused on seniuuritus selgelt väiksemad kui erinevused allkeelte vahel: ilukirjandusproosa ettelugemises on oluline teksti sisuline interpretatsioon, mille nimel teinekord ohverdatakse isegi hingamine.

Spetsiaalsed ilukirjandusproosa lugemise reeglid tuleks seega kujundada veelgi kitsamalt aluselt kui žanr (siin romaan), see aga nõuaks eeltööna mahukat etteloetud kirjanduse korpust. Kuni sellist pole ja asjaomaseid reegleid kujundada ei saa, võib süntesaator kunstproosat ette lugeda samal moel nagu ajakirjandustekste, s.o ilma eriliste emotsioonideta. Seega, kuni emotsionaalse kõne korpus areneb (vt Altrov 2007), tuleks tekst–kõne süntesaatori pause modelleerides toetuda graafiliselt tuvastatavatele süntaktilise liigenduse elementidele. Nende tugevat seost pauserimisega näitab süntaktiliselt mitmeti keskmiseks osutunud tekstiliik, pikema neutraalse ajakirjandusliku uudise tekst.

Sellistel ja muudel artikli algupooles viidatud andmetel võib väita, et milline tahes rakenduslik keelekäsitelu nõuab lähenemist vähemalt allkeelte ja ideaalis žanrite kaupa. Lisaks sünteeskõnele võiks ka muudes keeletehnoloogilistes rakendustes aktiivsemalt kasutada tekstiuurimuse ülal viidatud andmeid ja pidada silmas, et keskmist eesti keelt ei ole olemas: on tema olukohased variandid, mis avalduvad keelekasutuse valdkonnale omastes žanrites, ja see nõuab žanrilist lähenemist keelele ka arvutuslingvistikas, nagu seda on võimaluste piires arvestatud kirjakeele korpuste kujundamisel.

Kirjandus

- Altrov, Rene 2007. Emotsionaalse kõne korpuse loomine eesti keele tekst–kõne sünteesi jaoks. Tekstimaterjali evalvatsioon *viha* näitel. Magistritöö. Tartu Ülikooli filoloogia-teaduskond. Eesti ja soome-ugri keeleteaduse osakond. http://dspace.utlib.ee/dspace/bitstream/10062/2739/1/altrov_rene.pdf (3.01.2008).
- Barnbrook, Geoff 2007. ... Uncovering the SECRET life of language: an introduction to text exploration tools. – Plenary presentation. The 3rd Baltic Conference on Human Language Technologies. Kaunas, Lithuania, October 4-5, 2007.
- Biber, Douglas 1995. Dimensions of Register Variation. A Cross-Linguistic Comparison. Cambridge: Cambridge University Press.
- Biber, Douglas; Conrad, Susan 2001. Register variation: A corpus approach. – Deborah Schiffrin, Deborah Tannen, Heidi Hamilton (eds). The Handbook of Discourse Analysis. Oxford: Blackwell, 175–96.
- Boersma, Paul; Weenink, David 2006. Praat: doing phonetics by computer (Version 4.5.08). Computer program. <http://www.praat.org/> (3.01.2007).
- 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.
- Fant, Gunnar; Kruckenberg, Ann; Gustafson, Kjell; Liljencrants, Johan 2002. A new approach to intonation analysis and synthesis of Swedish. – Proceedings. Speech Prosody 2002, Aix en Provence, France, 11-13 April 2002, 283–286. <http://aune.lpl.univ-aix.fr/sp2002/papers.htm> (3.01.2008).
- Fant, Gunnar; Kruckenberg, Anita; Ferreira, Joana B. 2003. Individual variations in pausing. A study of read speech. – PHONUM, Reports in Phonetics 9. Umeå University, 193–196. <http://www.ling.umu.se/fonetik2003/> (3.01.2008).
- Grosjean, François; Collins, Maryann 1979. Breathing, pausing and reading. – *Phonetica* 36, 98–114.
- Gustafson-Čapková, Sofia; Megyesi, Beata 2001. A Comparative study of pauses in dialogues and read speech. – Paul Dalsgaard, Børge Lindberg, Henrik Benner, Zheng-hua Tan (eds). EUROSPEECH 2001 Scandinavia, 7th European Conference on Speech Communication and Technology, 2nd INTERSPEECH Event, September 3–7. Aalborg, Denmark, 931–934.
- Hennoste, Tiit 1994. Prospektiivsed minimaalhesitatsioonid eesti keele suulises tekstis. – K. Pajusalu, V. Ylivakkuri (toim.). Lähivertailuja 7. Turun Yliopiston suomalaisen ja yleisen kielitiedeen laitoksen julkaisuja 44. Turku, 33–51.
- Heylighen, Frances; Dewaele, Jean-Marc 2002. Variation in the contextuality of language. An empirical measure. – *Foundations of Science* 7, 293–340.
- Kerge, Krista 2002a. Kirjakeele kasutusvaldkondade süntaktiline keerukus. – Reet Kasik (toim.). Tekstid ja taustad. Artikleid tekstianalüüsisist. Tartu Ülikooli eesti keele õppetooli toimetised 23. Tartu: Tartu Ülikool, 29–46.
- Kerge Krista 2002b. Aja- ja ilukirjandusteksti süntaktilise keerukuse dünaamika XX sajandil. TPÜ eesti filoloogia osakonna veebitoimetised. *Lingvistika* 1. Tallinn: TPÜ Kirjastus. <http://www.tlu.ee/fil/veebitoimetised/pdf/lingvistika1.pdf> (3.01.2007). Uustrükk Kerge 2003a [4].
- Kerge, Krista 2003a. Keele variatiivsus ja *mine*-tuletus allkeelte süntaktilise keerukuse tegurina. Tallinna Pedagoogikailikooli humanitaarteaduste dissertatsioonid 10. Tallinn: TPÜ Kirjastus.
- Kerge Krista 2003b. Autori stiil ja allkeele tekst. – Reet Kasik (toim.). Tekstid ja taustad II. Tekstianalüüsi vaatepunkte. Tartu Ülikooli eesti keele õppetooli toimetised 26. Tartu: Tartu Ülikool, 59–89.
- Kerge, Krista 2006. Kontekstuaalsus – on selles mingit seaduspära? – Ettekanne ja teesid. Tekstipäev 2006, 20. detsember, Tartu Ülikool.

- Kerge, Krista; Pajupuu, Hille; Altrov, Rene 2007. Tekst, kontekstuaalsus ja kultuur. – Keel ja Kirjandus 8, 624–637.
- Kerge, Krista; Pajupuu, Hille; Tamuri, Kairi. Ilmumas 2008. Where should TTS-synthesizer pause and breath? – The 3rd Baltic Conference on Human Language Technologies, 4-5 October. Kaunas: Vytautas Magnus University and Institute of Lithuanian Language.
- Meier, Heidi 2002. Olulisi aspekte tekstitüübivõrdluses. – Reet Kasik (toim.). Tekstid ja taustad. Artikleid tekstianalüüsist. Tartu Ülikooli eesti keele õppetooli toimetised 23. Tartu: Tartu Ülikooli Kirjastus, 101–114.
- Meier, Heidi. 2003a. Essee asend allkeelte tekstitüübivõrdluses. Magistritöö. Käsikiri Tallinna Ülikooli eesti filoloogia osakonnas. Tallinn: Tallinna Pedagoogikaülikool.
- Meier, Heidi 2003b. Essee allkeelte võrdluses. – Reet Kasik (toim.). Tekstid ja taustad II. Tekstianalüüsi vaatepunkte. Tartu: Tartu Ülikooli Kirjastus, 116–135.
- Meister, Einar; Lippus, Pärtel. Ilmumas 2008. On temporal organization of spontaneous Estonian: preliminary analyses results of lecture speech. – The 3rd Baltic Conference on Human Language Technologies, 4-5 October. Kaunas: Vytautas Magnus University and Institute of Lithuanian Language.
- Mihkla, Meelis 2005. Modelling pauses and boundary lengthenings in synthetic speech. – The 2nd Baltic Conference on Human Language Technologies, April 4–5, 2005. Proceedings. Tallinn, 305–310.
- Mihkla, Meelis; Pajupuu, Hille; Kerge, Krista 2003. Modelling and perception of the Estonian general questions with the *kas*-particle. – Proceedings of 15th ICPhS. Barcelona, 539–542.
- Mihkla, Meelis; Pajupuu, Hille; Kerge, Krista; Kuusik, Jüri 2004. Prosody modelling for Estonian text-to-speech synthesis. – The 1st Baltic Conference on Human Language Technologies. The Baltic Perspective, Riga, Latvia, April 21–22 2004. Riga, 127–131.
- Mihkla, Meelis; Kerge, Krista; Pajupuu, Hille 2005. Statistical modelling of intonation and breaks for Estonian text-to-speech synthesizer. – Robert Vich (ed.). Electronic Speech Signal Processing. Proceedings of the 16th Conference Joined with the 15th Czech-German Workshop “Speech Processing”. Prague, Sept. 26-28, 2005. Studentexte zur Sprachkommunikation 36. Dresden: TUDpress, 91–98.
- Mihkla, Meelis; Kuusik, Jüri 2005. Analysis and modelling of temporal characteristics of speech for Estonian text-to-speech synthesis. – *Linguistica Uralica* 2, 91–97.
- Pajupuu, Hille 2007. Sõnaliik ja kontekstuaalsuse variatsioonid. – Ettekanne ja teesid. VI rakenduslingvistika kevadkonverents “Keel ja leksikon”, 26.-27. aprill 2007, Tallinn. http://www.rakenduslingvistika.ee/word_doc/070405Teesid-2007-Keeljaleksikon.doc (3.01.2008).
- Pajupuu, Hille; Kerge, Krista 2006. Hingav süntesaator ja pausid tekstiliigiti. – Keel ja Kirjandus 3, 202–210.
- Pöld, Karin 2007. Kaplinski tekstide kontekstuaalsus ja süntaks. Seminaritöö. Käsikiri Tallinna Ülikooli eesti filoloogia osakonnas. Tallinn: Tallinna Ülikool.
- Rais, Kairi 2007. E-suhtlus ja selle õpetamine koolis (äriühingu näitel). Bakalaureusetöö. Käsikiri Tallinna Ülikooli eesti filoloogia osakonnas. Tallinn: Tallinna Ülikool.
- Rannut, Mart 2003. Eesti keele jätkusuutlikkusest. – Helle Metslang (koost.). Eesti kirjakeele kasutusvaldkondade seisundi uuringud. Tallinna Pedagoogikaülikooli eesti filoloogia osakonna toimetised 4. Tallinn: TPÜ Kirjastus, 214–235.
- Rätsep, Siret 2002. Haldusõiguse keel seaduskeele üldises raamis. Bakalaureusetöö. Käsikiri Tallinna Ülikooli eesti filoloogia osakonnas. Tallinn: Tallinna Pedagoogikaülikool.
- Rääk, Kristjan 2002. Poliitiliste programmide keelest. – Reet Kasik (toim.). Tekstid ja taustad. Artikleid tekstianalüüsist. Tartu Ülikooli eesti keele õppetooli toimetised 23. Tartu: Tartu Ülikooli Kirjastus, 115–127.

- Tamuri, Kairi 2007. Pausid ettelotud ilukirjandustekstis. Magistritöö. Käsikiri Tallinna Ülikooli eesti filoloogia osakonnas. Tallinn: Tallinna Ülikool.
- Tseng, Chiu-Yu 2002. The prosodic status of breaks in running speech: examination and evaluation. – Proceedings. Speech Prosody 2002, Aix en Provence, France, 11-13 April 2002, 667–670. <http://aune.lpl.univ-aix.fr/sp2002/papers.htm> (3.01.2008).
- Tseng, Chiu-Yu; Chou, Fu-Chiang 1999. A prosodic labeling system for Mandarin speech database. – Proceedings of ICPH99. San Francisco, California, 2379–2382.
- Viks, Ülle; Hein, Indrek 2001. Sõnavormide kasutus teadustekstides. – EKI. Publikatsioonid. <http://www.eki.ee/teemad/> (12.12.2007).
- Whalen, Douglas H. 1994. Computerized Speech Pauses for a Breath. – The Futurist, May/June, 28 (3), 7.

Krista Kerge (Tallinna Ülikool) uurimisvaldkonnad on keele variatiivsus, tekstianalüüs, rakenduslingvistika (L1 ja L2 omandamine, õigus- ja haldussuhtlus, kõne paralingvistiliste komponentide ja süntaksi seosed).
krista.kerge@tlu.ee

Hille Pajupuu (Eesti Keele Instituut) uurimisvaldkondadeks on kõneakustika, kultuuridevaheline kommunikatsioon, keeletestimine.
hille.pajupuu@eki.ee

Kairi Tamuri (Eesti Keele Instituut) uurimisvaldkonnaks on kõneakustika: tekstiliigid, emotsionaalne kõne, pausid ja hingamine.
kairi.tamuri@eki.ee

Heidi Meieri (vabakutseline uurija) uurimisvaldkond on tekstianalüüs.
heidi.meier@gmail.com

SPEECH TECHNOLOGY NEEDS A GENRE-BASED APPROACH

Krista Kerge, Hille Pajupuu, Kairi Tamuri, Heidi Meier

Tallinn University, Institute of the Estonian Language

In reading out a text, pauses and breathing as two of the naturalness parameters of the read-out text are closely related to its functional style (journalism, fiction). The results of text study may, apart from contributing to the naturalness of synthetic speech, help recognize the genre of the text, differentiate between functional styles as well as authors, and economize on the use of word lists and grammar in text automatic analysis. The article refers to some studies that relate to breathing and pauses as well as to some other formalizable parameters of Estonian texts, yet have not been considered in language technologies. To verify the above thesis the pauses and breathing in two genres of read-out texts of fiction are compared to the available data of a news text. The results prove that the duration of pauses and inhalations as well as their relation to the syntactic structure of the sentence does differ not only according to functional style but also according to genre. The syntactic structure is mainly indicated by punctuation marks. In fiction pausing partly depends on reader's interpretation. In text-to-speech synthesis for the effect of reading aloud breaks and breathing should be generated from the data of extended news texts as a parametrically neutral genre.

Keywords: pauses, breathing, functional style, genre, news, fiction, speech synthesis, Estonian

SÕNARAAMATU KOHTULINGVISTILINE ANALÜÜS: EESTI PRETSEDENT

Margit Langemets, Piret Voll

Ülevaade. 2003. aastal viidi Eestis esimest korda läbi kohtulingvistiline leksikograafiaekspertiis, mille käigus hinnati mitme sõnaraamatu sisu, lähtudes kriminaalsüüdistusest intellektuaalse omandi varguses. Artiklis kirjeldatakse Eesti Keele Instituudis läbi viidud leksikograafiaekspertiisi ja selle jaoks välja töötatud keeletehnoloogilist meetodit. Ekspertiisi tulemuste põhjal on välja pakutud piirmäär kakskeelsete sõnaraamatute originaalsuse üle otsustamiseks. Käsitlev juhtum on Eestis ainulaadne, kujutades säärasena kohtulingvistilist pretsedenti.

Võttesõnad: kohtulingvistika, leksikograafiaekspertiis, sõnaraamatud, inglise keel, eesti keel

1. Sissejuhatus: kohtulingvistikast ja sõnaraamatuplagiaadist

Eestis on kohtulingvistika pea tundmatu ala. Keeleekspertiisi on vajatud enamasti solvangut või laimu käsitlevate kohtuasjade juures. Kohtulingvistika tegelik töömaa on palju avaram, hõlmates nii õiguskeelt kui praktiliselt kõike, mis seondub keele ja kohtusüsteemi ning seadusandlusega (McMenamin 2002: 67). Enamasti uurivad kohtulingvistika eksperdid keelekasutust nii foneetilisel kui ka stilistilisel seoses laimujuhtumitega, nii on pea võimatuks osutunud leida leksikograafiapretsedenti. Autoriõigus ja plagiaatlus on kohtulingvistikas suhteliselt uued teemad üldse, mitte ainult Eestis, vaid ka mujal. Leksikograafiaalase seadusandluse puudumise üle on kurtunud näiteks Suurbritannia leksikograaf John Williams (1992: 566).

Iseenesest on sõnaraamatute pärast kohtuteed käidud juba 18. sajandil. Londonist on teada juhtum, milles plagiaatluses süüdistati Ann Fisheri sõnaraamatut “An Accurate New Spelling Dictionary and Expositor of the English Language”. Leiti, et teos võib olla kopeeritud John Enticki “The New Spelling Dictionary” (1765) pealt. Võrdlus näitas, et üks-ühele oli üle võetud 43,5%. Kaks aastat kestnud kohtuvaidluste tulemusel lubati sõnaraamat siiski 1773. aastal välja anda (Rodriguez-Alvarez

2006). Veel varasem, mitte küll kohtusse jõudnud näide on pärit 17. sajandist, kui Londonis ilmus Guy Miège'i "The Great French Dictionary. In Two Parts" (1688). Abel Boyer andis 1699. aastal välja "The Royal Dictionary", nimetades hea tava kohaselt ka kasutatud allikaid. Millegipärast "unustas" ta mainida G. Miège'i, kelle sõnaraamatust, nagu hiljem selgus, võeti kas otse või väikeste muudatustega üle umbes 70% märksõnadest koos vastetega (Cormier 2005).

Ka Eestis pole sõnaraamatuplagaadi kahtlus tundmatu asi. Arnold Kask (1955) on 18. sajandi sõnaraamatutest ülevaadet tehes leidnud, et nii Anton Thor Helle kui ka August Wilhelm Hupel on kasutanud suurel määral Salomo Heinrich Vestringi (umbes 1720-ndad) sõnaraamatu ainestikku, sealhulgas A. T. Helle (1732) käsiraamatus avaldatud sõnastik "toetub täiel määral" S. H. Vestringi käsikirjalisele sõnaraamatule (Kask 1955: 190). Käsiraamatu eessõnas on märgitud, et materjal on kogutud "ühe kristliku sõbra" ja läbi töötatud A. T. Helle poolt. "Kristliku sõbra" üle on arutlenud ka Arvo Krikmann (1986) ja Kristiina Ross (EEVA). A. Krikmanni arvates kinnitab A. Kase arvamust kõige tugevamini S. H. Vestringi alfabeet, mis on järjestatud väga veidral viisil ning mida sellisena kordab ka A. T. Helle "Anweisung" (Krikmann 1986: 84). K. Ross ei pea samuti A. T. Helle sõnastikku originaalmaterjaliks, sest see "tugineb S. H. Vestringi (märksa mahukamale) käsikirjalisele sõnastikule", ent kuna kohati on seal S. H. Vestringistki rikkalikumat materjali, siis "on oletatud, et Thor Helle kas lisas neile oma kogutud materjali või kasutas lisaks Vestringi käsikirjale veel kellegi materjale." (EEVA)

Järgnevas kirjeldatakse 21. sajandi algul Eestis läbi viidud leksikograafiaekspertiisi käiku. Alustatakse kohtuasjast ja autoriõiguse vastavatest kirjakohtadest, kirjeldatakse ekspertiisi lähteandmeid (sõnaraamatuid), ülesandeid ja Eesti Keele Instituudis välja töötatud meetodit. Seejärel analüüsitakse vaidlusalust materjali ette antud ekspertiisiülesannete kaupa. Lõpuks antakse ülevaade ekspertiisitulemuste rakendamisest Eesti õigussüsteemis, käsitletakse plagaadi piiritlemise probleemi ja pakutakse välja piirmäär kakskeelsete sõnaraamatute originaalsuse üle otsustamiseks.

2. Kohtuasi: sõnaraamatuvarguse süüdistus

2000. aastal kaebas kirjastus TEA kohtusse firma Festart, süüdistades viimast sõnaraamatuvarguses. Juhtumit menetleti kriminaalasjana seitse aastat, kuni 2007. aasta kevadeni. Festarti ja selle juhti Sergei Gluhhovit süüdistati selles, et nad aastatel 1999–2004 andsid CD-ROM-il välja inglise-eesti sõnaraamatu versioone, kasutades nende koostamisel TEA sõnaraamatute elektroonilisi põhjasid, omamata selleks kirjastuse luba. Vaidlustatud sõnaraamat pretendeerib originaalteose nimetusele, sest ei pakendil ega lisatekstides leidu ühtki viidet kasutatud allikatele. Muul juhtumil oleks tulnud ära näidata algse teose autori nimi, teose pealkiri (nimetus) ja allikas, kus teos on avaldatud, nagu sätestab autoriõiguse seadus (AÕS § 35 lg 3). Kohtuasja lahendamiseks sai Eesti Keele Instituut 2002. aastal Tallinna Politseiprefektuuri kriminaalosakonna majanduskuritegude talituselt ülesande läbi viia leksikograafiaekspertiis.¹

Olgu vahemärkusena öeldud, et kuni 2007. aasta lõpuni töötas Eestis kaks riiklikku ekspertiisiasutust,² mis aga kumbki nii spetsiifilist ekspertiisi ei tei-

¹ Ekspertiisi määras Tallinna Politseiprefektuuri kriminaalosakonna majanduskuritegude talituse uurija, politseivaneminspektor Carmen Kuus. Ekspertiisimäärus on koostatud Tallinnas 11.11.2002. Ekspertiisi teostaja valiti riigihanke korras.

² Kaks asutust – Eesti Kohtuarstlik Ekspertiisbüroo ja Kohtuekspertiisi ja Kriminallistika Keskus – liideti alates 1. jaanuarist 2008 üheks asutuseks nimega Eesti Kohtuekspertiisi Instituut (Riigi Teataja 24.08.2007).

nud.³ Lähimaks, st intellektuaalse sfääriga lõikuvaks võis ehk pidada dokumendi-, infotehnoloogia-, ka kujutise uuringut, millega eksperdid toona lähemalt tutvust ei teinud.

2.1. Autoriõigusest

Teoste hulgas, millele tekib autoriõigus, on autoriõiguse seaduses nimetatud väga mitmesuguseid teoseid: kirjalikke (ilukirjandus, publitsistika, erisugused teadustöö tulemused jms) ja suulisi teoseid (kõned, jutlused jms), muusika-, tantsu- ja kujutava kunsti teoseid, arvutiprogramme, standardeid, andmebaase jpm. Autoriõigus tekib ka tuletatud teosele, milleks on “teose tõlge, algse teose kohandus (adaptsioon), töötlus (arranjeering) ja teose muu töötlus” (AÕS § 4 lg 22). Põhiline kriteerium on, et teos oleks “autori enda intellektuaalse loomingu tulemus” (AÕS § 4 lg 2).

Ajal, mil arvutiajastu soodustab hõlpsat kopeerimist, diskuteeritakse ühtelugu autori- ja omanikuõiguse kontrolli ja vabakasutuse üle eri elualadel.⁴ Seaduses on sätestatud mõned intellektuaalse tegevuse tulemused, millele autoriõigus (seni) ei laiene, nagu rahvaloominguteosed, päevauudised, faktid, andmed, arvutiprogrammi elemendid jms (AÕS § 5). Sõnu *sõnastik* või *sõnaraamat* ei esine autoriõiguse seaduses kordagi.

Sõnaraamatule võiks autoriõigus tekkida kui originaalsele andmebaasile (“süsteemaatiliselt või meetodiliselt korrastatud iseseisvate teoste, andmete või muu materjali kogu”) või kui tuletatud teosele. Mille alusel aga eristada autoriõigusega kaitstavat (tuletatud) teost plagiaadist – selle kohta puuduvad autoriõiguses kaudsedki juhtnöörid. Kui arvutiprogramm on tervikuna kaitstud kui teos, aga selle elemendid mitte, siis kuidas on lood sõnaraamatuga? Süütuse presumptsiooni põhimõttel näeb seadus ette, et kui asja õiguskuulekuses kaheldakse, siis lasub tõendamiskohustus sellel, kes on vaidlustanud teose autoriõigusega kaitstuse (AÕS § 4), siinses kohtuasjas niisiis kirjastusel TEA.

2.2. Ekspertiisi lähteandmed ja ülesanded

Ekspertiisiakti⁵ sissejuhatavas osas on ekspertiisi lähteandmetena üles loetud kõik asjasse puutuvad sõnaraamatud, mis jagunevad kahte leeri. Ühel pool on vaidlustatud materjal ise ehk ekspertiisi objekt, milleks on kirjastuse Festart üllitatud inglise-eesti sõnaraamat CD-ROM-il ning seda toetav võrdlusmaterjal. Teisel pool on kannatanu võrdlusmaterjal (kakskeelsed inglise-eesti või eesti-inglise sõnaraamatud). Tabelis 1 on kirjas kõik kohtuasjas osalenud sõnaraamatud koos märkega nende staatuse kohta. (Edaspidi on artiklis kasutatud poolele osutavaid lühinimesid, nt FESTART-sõnaraamat, TEA teaduse ja tehnika sõnaraamat.)

³ 2008. a loodud asutuse põhikirjas (Riigi Teataja 20.12.2007) on üles loetud asutuses tehtavad uuringud, nt alkoholi-, DNA-, dokumendi-, hääle-, infotehnoloogia-, jälje-, kujutise-, käekirja-, narkootilise aine jpm uuringud (kokku 30 liiki), samuti ekspertiisiliigid: 1) kohtuarstlik, 2) kohtubioloogia-, 3) kohtukeemia-, 4) kohtusühhiatria- ja 5) kriminalistikaekspertiis.

⁴ Vt nt Evi Arujärve (2008) kirjutist “Kultuur – missioon ja äri”.

⁵ Ekspertiisiakt kriminaalasjas nr 00231000138 “Festart Dictionary English-Estonian. 70 000 Entries. CD-ROM Version 3.10”. Käsikiri Eesti Keele Instituudis. Leksikograafiaekspertiisi peaekspert: Piret Voll.

Tabel 1. Ekspertiisi lähteandmed (kohtuasjasse puutuvad sõnaraamatud)

Nr	Täiskirje ekspertiisiaktis	Lühinimi siinses artiklis	Materjali staatus
1	“FESTART DICTIONARY ENGLISH-ESTONIAN. 70 000 ENTRIES. CD-ROM Version 3.10” ISBN 9985-9063-4-9 ⁶	FESTART-sõnaraamat	vaidlustatud materjal
2	English-Latvian Dictionary / Angļu-latviešu vārdnīca. 3rd improved edition. Rīga: Jāņa Sēta Publishers & Printers Ltd, 1997. 1328 lk	FESTART inglise-lāti	võrdlusmaterjal
3	Teaduse ja tehnika seletav sõnaraamat I-II. Tallinn: TEA Kirjastus, 1997. 1667 lk	TEA teadus ja tehnika	võrdlusmaterjal
4	Teaduse ja tehnika seletav sõnaraamat. Eesti-inglise. Lisakõide. Tallinn: TEA Kirjastus, 1998. 544 lk	TEA teadus ja tehnika lisakõide	võrdlusmaterjal
5	PASSWORD. Inglise-eesti seletav sõnaraamat. Tallinn: TEA Kirjastus, 1995. 856 lk	TEA Password	võrdlusmaterjal
6	Eesti-inglise majandussõnastik. 2. trükk. Tallinn: TEA, 1998. 496 lk	TEA eesti-inglise majandus	võrdlusmaterjal
7	Inglise-eesti majandussõnaraamat. Tallinn: Keelefirma TEA, 1995. 496 lk	TEA inglise-eesti majandus	võrdlusmaterjal
8	Inglise-eesti ja eesti-inglise ärisõnastik. Tallinn: Keelefirma TEA, 1993. 560 lk	TEA inglise-eesti äri	võrdlusmaterjal
9	Festart; English-Estonian Dictionary of Economics 2.47 [elektrooniline teavik]	FESTART inglise-eesti majandus	täiendav võrdlusmaterjal
10	Silvet, Johannes 1989. Inglise-eesti sõnaraamat I-II. Kolmas, parandatud ja täiendatud trükk. Tallinn: Valgus	Silvet 1989	täiendav võrdlusmaterjal
11	Inglise-eesti sõnastik. Eesti Keele Instituudi vabavara Internetis ⁷	EKI sõnastik	täiendav võrdlusmaterjal
12	Festart. Keyword, English-Estonian Dictionary 2.44	FESTART Keyword	täiendav võrdlusmaterjal

Ekspertiisimääruses oli sõnastatud neli ülesannet (küsimust), millele ekspertiis pidi vastama.

1. Missugused tunnused viitavad võrdlusmaterjalina esitatud sõnastike ja/või nende elektrooniliste põhjade kasutamisele FESTART-sõnaraamatu koostamisel?
2. Kui palju on vaidlustatud materjalil (FESTART-sõnaraamatu) märksõnu või muid tunnuseid, mis kattuvad võrdlusmaterjalina esitatud sõnastike märksõnade või muude tunnustega?
3. Kuidas võis võrdlusmaterjalina esitatud sõnastike ja/või nende elektrooniliste põhjade alusel tehniliselt toimuda FESTART-sõnaraamatu märksõnade, vastete ja indeksi moodustamine?

⁶ Täpne märksõnade arv on 71 154.

⁷ Vt www.eki.ee/dict/inglise/ (20.11.2007).

4. Kas FESTART-sõnaraamat on originaalteos või on see tuletatud teos, mis ulatuses see teos kattub võrdlusmaterjalina esitatud teostega?

Järgnevas on neid küsimusi ühekaupa analüüsitud. Aga enne seda meetodist.

3. Meetod

Nagu juba eespool mainitud, ei olnud leksikograafiaekspertiisi tegemiseks otsest eeskuju leida. Peale selle tuli Tallinna Politseiprefektuuri kriminaalosakonna ettekirjutusena ekspertiisi läbi viia absoluutses ulatuses, st toetumist statistilisele valimile nad ei aktsepteerinud. Seega oli mõistlik, nagu ka nüüdisajale kohane, kasutada keeletehnoloogia abi, seda enam, et kogu materjal – nii vaidlustatud teos kui ka võrdlusmaterjal – loovutati ekspertide käsutusse elektroonilisel kujul.

Ekspertiisianalüüsi peamine meetod oli sõnaraamatute elektrooniliste versioonide läbiv võrdlus arvutiprogrammide abil. Aga küsimus oli: mida, millist üksust mõõta? Kui element ükski on liiga vähe, et sellele toetuda, siis elementide kombinatsioon võiks teataval määral peegeldada autori loomingut. Kakskeelses sõnaraamatus võiks vähimaks oluliste infoüksuste kombinatsiooniks pidada märksõna koos oma vaste(te)ga: ei saa olla kakskeelset sõnaraamatut, kus märksõnade juures pole tõlkevasteid (viiteartiklis asendab vastet viide) – need üksused on obligatoorsed igas tõlkesõnaraamatus. Vasteüksus võib koosneda ühest vastest (ühest sõnast või mitmesõnalisest väljendist, vt näide 1a). Sagedamini on antud mitu vastet, mis omavahel on eraldatud komaga (1b). Ei ole kuigi tavaline, et märksõnal ongi ainult üks vaste, eriti üldkeele sõnaraamatus, sest pole kaht keelt, mis üks-ühele kattuks.⁸ Väga tavaline on vastete loendis ka semikoolon, mis harilikult eraldab tähenduselt veidi kaugemaid vasteid (1b, 2).

- (1a) **post-paid** .. saatekuludeta [FESTART-sõnaraamat]
w pikkusühiku koormuse sümbol [FESTART-sõnaraamat]
- (1b) **voidable** .. annulleeritav, tühistatav; vaidlustatav [FESTART-sõnaraamat]
- (2) **micro-** väike, miljondik-; pisi-; miljondik, 0,0000001 m [FESTART-sõnaraamat]

Ekspertidid määratlesid analüüsi põhiüksuseks märksõna koos oma vaste(te)ga, lähtudes vasteloendis esinevast semikoolonist. Koma eraldab lähemaid vasteid ühekaupa ega sobi seega nii hästi rühma eraldajaks (märksõna-vaste-grupis oleks alati ainult üks vaste). Peale selle võib koma esineda vaste osana pikemates väljendites, murdarvudes jm (vt näide 2). Seega eraldab võrreldava märksõna-vaste-grupi (edaspidi MV-grupp) alati semikoolon, kusjuures märksõna korratakse iga grupi juures (vt tabel 2).

FESTART-sõnaraamatus, mis sisaldab 71 154 märksõna, on kokku 122 915 niisugust MV-gruppi. Ühe või teise TEA sõnastiku osakaalu FESTART-sõnaraamatus on iseloomustatud protsentides MV-gruppide kaupa. Välja toodud erinevuste või sarnasuste rõhutamiseks on (ekspertiisis esitatud) näidetes vastavad kohad alla joonitud, tabelitesse on lisatud selgituste tulp.

⁸ Spetsiaalses, nn kakskeelestatud sõnaraamatus võib olla põhimõtte esitada ainult üks vaste (seda tüüpi sõnaraamat on näiteks TEA Password), aga sealgi "segab mängu" leksikaalne mitmetähenduslikkus: üks vaste vahendab üht tähendust.

Tabel 2. Märksõna-vaste-gruppide (MV-grupid) näiteid

Sõnaartikkel	MV-grupid	Selgitus
micro- väike, miljondik-; pisi-; miljondik, 0,0000001 m	kokku 3 MV-gruppi: 1) micro- väike, miljondik-; 2) micro- pisi-; 3) micro- miljondik, 0,0000001 m	1. ja 3. MV-grupis on kaks vastet (eraldatud komaga), 2. MV-grupis on üks vaste
realise [verb] aru saama; täide viima; realiseerima	kokku 3 MV-gruppi: 1) realise aru saama; 2) realise täide viima; 3) realise realiseerima	kõik grupid koosnevad ühest (mitmesõnalisest) vastest

4. Esimene ülesanne: reetlikud tunnused

Tunnusena, mis võiks viidata sõnaraamatu ja/või nende elektrooniliste põhjade kasutamisele teise sõnaraamatu koostamisel, on ekspertiisil välja toodud järgmised elemendid ja/või seosed: a) korduvad märksõnad, b) märksõnavaliku kattuvus, c) MV-gruppide kattuvus, d) trükivead (õigekirjavead, küljenduselementide jäänused), e) haruldased vasted. Muud tehnilised detailid (nt muutunud tähestikjärjestus, lihtsustatud diakriitikutega tähed, eesti-inglise pöördindeks) on sõnastikuprogrammi eripärad ja siduvat informatsiooni ei kanna.

Märksõnavaliku või muude tunnuste kattuvust (seosed b ja c) on kirjeldatud järgmises, teist ülesannet käsitlevas 6. osas. Siin on lähemalt vaadeldud ülejäänuid reetlikke tunnuseid.

4.1. Korduvad märksõnad

FESTART-sõnaraamatu 71 154 märksõnast⁹ on korduvaid märksõnu 4585 (6,4%), st ligi 5000 juhtumil väljastatakse päringu tegemisel vastuseks mitu sama märksõnaga artiklit. See tõsiasi ei külvaks kahtlust, kui tegemist oleks seda märksõna *sisaldavate* väljenditega, mitte mitme identse märksõnaga. Näiteks “Eesti keele seletussõnaraamatus” tehtud märksõnapäring *must* toob vastuseks nii lihtsõnast märksõna enda (*must*) kui ka seda osist sisaldavad liitsõnad (*kärbsemust*, *mustjuur*, *mustsada*, *trükimust* jt) ja väljendid (*must lammas*, *must valgel*, *ei musta ega valget* jt). Sõnavarabaaside üks põhinõudeid on märgistada kirjeldusobjekt unikaalseks. Seega kui märksõna on kujult sama (nt homonüüm), on tal ikkagi üldjuhul mingi eristav tähis juures.

FESTART-sõnaraamatus on asi teisiti: lõviosa (u 90%) mainitud 4585-st artiklist on sellised, kus märksõna kordub muutumatul kujul. Väljastatud n-ö teise artikli vastena esineb siis enamasti majandustermin (vt tabel 3). See viib vägisi mõtted elektroonilisele kopeerimisele (majandussõnastikust?), kuivõrd tavalisel moel koostatud sõnaraamatus on sellised juhtumid enamasti ühe märksõna all, kas siis eri tähendusena (eraldatud tähendusnumbriga) või tähendusrühmana (semikooloniga).

⁹ Edaspidi on jälgimise hõlbustamiseks kasutatud ümardatud andmeid (70 000 märksõna), mis on kirjas ka toote ümbrisel.

Tabel 3. Korduvate märksõnadega artikleid FESTART-sõnaraamatus

Otsisõna	Väljastatav vastus	Selgitus
post-paid	post-paid [adjective] saatekuludeta post-paid [economics] postikulud tasutud, saatekuludeta, franko	märksõna kordub muutumatul kujul, kusjuures teise märksõna järel esineb majandustermini märgend economics
yuppie	yuppie .. [noun] jappi yuppie, yuppy [economics] juppi, jappi	väljastatakse kaks väga sarnast (inimese poolt toimetamata?) sõnaartiklit; märksõna kordub muutumatul kujul, kusjuures teise märksõna järel esineb majandustermini märgend economics

Ülejäänud juhtudel (10%) esineb korduv ois fraasi osana, mis on ka pabersõnastikes tavaline, kuigi seal on sellised juhtumid enamasti kasutusnäite või allmärksõna positsioonis (3), samas kui FESTART-sõnaraamatus on need tõstetud märksõna staatusse (4). Paljudes pabersõnastikes kuulub sõnaartiklisse märksõna tuletisi, mitmesõnalisi püsiühendeid ja väljendeid, samas kui arvutisõnastikes paigutatakse need kõik tavaliselt eraldi (st märksõna positsiooni). Kõrvalekaldena on siin paslik juhtida tähelepanu märksõnade arvule: sel moel eraldi esitatud märksõnad paisutavad sõnaraamatu mahtu, mida on muidugi soodus kasutada (arvuti)toodet reklaamides. Tegelikult võiks FESTART-sõnaraamatu “suurt” mahtu (70 000 märksõna) pidada võrdväärseks umbes keskmise suurusega (st 40–50 000 märksõnaga) pabersõnaraamatuga.

- (3) **voidable** [economics] annulleeritav, tühistatav; vaidlustatav • **voidable act** vaieldav toiming; **voidable title** vaieldav omandiõigus [võimalik artikkel pabersõnaraamatus]
- (4a) **voidable** [economics] annulleeritav, tühistatav; vaidlustatav [FESTART-sõnaraamat]
- (4b) **voidable act** [economics] vaieldav toiming [FESTART-sõnaraamat]
- (4c) **voidable title** [economics] vaieldav omandiõigus [FESTART-sõnaraamat]

4.2. Trükivead: õigekirja- ja poolitusvead, küljenduselementide jäänused

Leksikograafias (nagu muudeski kirjalikes tekstides) on samade trükivigade esinemine üks kindlmaid tõendeid selle kohta, et üks teos kopeerib teist: ilmselget viga ei ole lihtsalt võimalik teistmoodi paljundada, kui maha kirjutades: kas siis täiesti asjatundmatu inimesena või, mis veelgi hõlpsam, elektrooniliselt, kõike mehaaniliselt üle võttes. Ekspertiisis on välja toodud nii korduvaid õigekirja- ja poolitusvigu (vt tabel 4) kui ka küljenduselementide jäänuseid (vt näide 5). Näiteks 69-st ekspertiisi käigus leitud TEA Passwordi poolitusveast kordub 68 (98,5%) FESTART-sõnaraamatus.

Tabel 4. Kattuvad trükivead: õigekirja- ja poolitusvead

TEA inglise-eesti majandus	FESTART-sõnaraamat	Selgitus
BOOM buum, kõrgkonjunktuur; ..	boom [economics] buum, kõrgkonjunktuur; ..	õigekirjaviga: p.o <i>kõrgkonjunktuur</i>
quantitative CHANGE	quantitative change [economics] ..	õigekirjaviga: p.o <i>quantitative change</i>
SECOND bill of (exchange) sekundaveksel, ..	second bill of [economics] sekundaveksel	viga märksõnas: programmiga on välja jäetud kõik sulgudes osad
courtly [adjective] .. • peene-kombeline	courtly [adjective] õukondlik; peene-kombeline	poolitusviga

Küljendusvead on reeglina tekkinud uue materjali lisamisel ja/või millegi ümber-tõstmisel – see seletab ülearuseid tühikuid (5a), topelt-kirjavahemärke (5b), rea lõppu “unustatud” kirjavahemärke (5c) ja muid kirjavahemärkidega seonduvaid vigu (5d) (kõik näited FESTART-sõnaraamatust).

- (5a) **carriage free** [economics] franko, tasuta
- (5b) **letter of credit** [economics] akreditiiv, krediitkiri, rambuuri; maksekä-sund
- (5c) **compensate** [economics] korvama, hüvitama, tasuma; tasakaalustama, balansseerima,
- (5d) **appreciate** [economics] hindama; arvesse võtma; hinnas tõusma, hinda tõstma, väärtuses kasvama

Samalaadsele tehnilisele argumentile – veidrale tähestikjärjestusele – on osutatud juba 18. sajandil Helle-Vestringi sõnaraamatukriitikas, millest oli juttu siinse artikli sissejuhatuses.

4.3. Haruldased vasted

Kui suure kasutussagedusega sõnade vastete kattumine on mõnevõrra ootuspärane,¹⁰ siis haruldaste märksõnade ja kummaliste vastete kordumine mitmes sõnaraamatus seda kindlasti ei ole. Probleem on muidugi, kuidas kõik need haruldased vasted üles leida. Kuna see toona ei olnud keeletehnoloogiliselt lahendatav (põhimõtteliselt saaks mõõta iga eesti sõna sagedust tekstikorpuses – aga põhjendatud tulemuse saamiseks läheks vaja vähemalt 100 mln sõne suurust üldkeele korpust, mida Eestis pole veel praeguseni olemas), siis töid eksperdid n-ö käsitsi välja TEA sõnaraamatus silma hakanud mitmes mõttes erilisi vasteid.

Leksikograafide arusaamatul põhjusel korduvad need FESTART-sõnaraamatus, kuna tavaliselt selliseid asju uustrüki jaoks täiendatakse või parandatakse. Kui TEA teaduse ja tehnika sõnaraamatus, mille inglise aluspõhja (Chambers Ltd) on viimati uuendatud 1988. a, on esitatud (vananenud) märksõna ja vaste (6a), siis

¹⁰ Võib täiesti mõista Interneti-suhtlejate nõrdimust 2007. a kevadel (kui kohtuprotsess lõppes): “.. tegin ka eks-pertiisi: TEA sõnastik: *ajaleht* = *newspaper*; FESTART: *ajaleht* = *newspaper*. Oskad sa ette kujutada, ongi täpselt sama sõna pandud Festardil, järelikult maha tehtud! Piraat mis piraat! Kole ju ..” (Eesti Päevaleht Online, aprill 2007).

sama kordub FESTART-sõnaraamatus (6b), kus toimetajad ei ole kiiresti vananevaid infoteaduse jm termineid uuemate vastu välja vahetanud. Kui TEA teaduse ja tehnika sõnaraamatus on ära toodud spetsiifilisi termineid, mis üldkeele või väiksemates erialasõnastikes puuduvad (7a, 8a), siis on need olemas ka FESTART-sõnaraamatus (7b, 8b). Näidete 7 ja 8 puhul võib tähelepanu juhtida ka identses sõnastuses vastetele. Ebaharilik on ka allpool toodud vaste *tähtpäevuma* (10b). Võrdluseks: ekspertide kogemus sõnaraamatute koostamisel kinnitab, et isegi juhul, kui märksõna tähendusele on lubatud panna ainult üks vaste, esitavad eri tegijad selle vaste pisut isemoodi.

- (6a) **XNS** [Comp] Abbreviation for *Xerox network system*. .. • XNS [TEA teaduse ja tehnika]
- (6b) **XNS** [computers] XNS [FESTART-sõnaraamat]
- (7a) **VOS** [Image Tech] Abbreviation for *video on sound*. See **background video**. • helipealne video, taustvideo [TEA teaduse ja tehnika]
- (7b) **VOS** [photography] helipealne video, taustvideo [FESTART-sõnaraamat]
- (8a) **taungya** [Ecol] .. • agrometsanduse vorm troopilistel aladel [TEA teaduse ja tehnika]
- (8b) **taungya** agrometsanduse vorm troopilistel aladel [FESTART-sõnaraamat]

5. Teine ülesanne: kattuvate märksõnade või muude tunnuste määr

FESTART-sõnaraamatu materjali võrreldi ekspertiisis sisuliste alarühmade, valdkondade ja vastavate sõnaraamatute kaupa: 1) majandusterminid (võrdlusaluseks TEA inglise-eesti majandus), 2) tehnikaterminid (TEA teadus ja tehnika) ja 3) üldsõnavara (TEA Password ja Silvet 1989). Pärast iga alarühma analüüsimist eemaldati vastav osa (MV-grupid) algsest FESTART-sõnaraamatust. Iga järgnevat alarühma analüüsimise osa võrra vähendatud algmaterjalis.

5.1. Majandusterminid

Majandussõnavara osatähtsus on FESTART-sõnaraamatus (70 000 märksõna, 122 915 MV-gruppi) märkimisväärselt suurem kui muudes kakskeelsetes üldkeele sõnastikes: majandussõnad (märgendiga [*economics*] 'majandus') moodustavad FESTART-sõnaraamatu kogumahust ligi 40% (kokku 27 387 märksõna).

Majandusterminite osa eristub muudest sõnavarakihetidest mitme iseärasuse poolest: märksõnadel puudub hääldus (9-10); kogu FESTART-sõnaraamatus esinevad ainult selles rühmas mitmest sõnast koosnevad märksõnad (10); esineb trükivigu: veatüübid nagu "tühik koma ees" (9a), "koma järel puudub tühik" (9c), "kirjavahemärk rea lõpus" (9b) või "kaks kirjavahemärki üksteise järel" (10a) on selles rühmas ülekaalukas enamuses.

- (9a) **billion** [economics] biljon , miljard
- (9b) **bunco** , **bunko** [economics] pettus,
- (9c) **discuss** [economics] arvamusi vahetama, diskuteerima, polemiseerima, väitlema, vaidlema, arutama, arutlema
- (10a) **letter of credit** [economics] akreditiiv, krediitkiri, rambuuri, maksekä-sund
- (10b) **become due** [economics] saabuma, tähtpäevuma, aeguma; maksmisele kuuluma

Majandustermineid analüüsiti kahes etapis: esmalt võrreldi FESTART-sõnaraamatut teise sama kirjastuse sõnaraamatuga (FESTART inglise-eesti majandus, 27 500 märksõna), seejärel kontrolliti viimase võimalikku seotust TEA inglise-eesti majandussõnaraamatuga.

Kahe Festardi sõnaraamatu vigaste kohtade võrdlemisel ilmnes mitu asja. Esiteks, MV-gruppide kokkulangevus on sama hea kui täielik: võib väita, et FESTART-sõnaraamatu 27 891 MV-gruppi (23%) pärinevad sama kirjastuse majandussõnaraamatust (mis samuti pole originaalne, aga sellest allpool). Ülevõtmisel on automaatselt välja jäetud sulgudes olnud kommentaarid, täpsustused, keeleviited jms, seejärel tehtud korrektuur ja lisatud ka mõningaid uusi märksõnu. Tõenäoliselt on märgend [economics] lisatud lõpuks programmi abil ja ilma järgneva korrektuuri või toimetamiseta. Seda illustreerib märksõna *yeepy*, kus märgendi õige koht peaks olema eesti vaste *aktiivsed vanurid* ees, mitte ingliskeelse tähenduse seletuse keskel (11).

(11) **yeepy**, youthful [economics] energetic elderly people involved in everything aktiivsed vanurid [FESTART]

Teiseks tulevad esile tüüpiliste vigade tekkekohad, st kohad, kus midagi on muudetud, kas uue materjali elektroonilisel lisamisel, väljajätu kohal, sulgudes esitatud materjali ärajätmisel, küljenduse poolitusvihjete puhul jne (vt tabel 5).

Tabel 5. Festardi sõnaraamatute võrdlus: tüüpilisi veakohti

FESTART inglise-eesti majandus	FESTART-sõnaraamat	Vea tekke selgitus
APPRECIATE hindama; tänulik olema; arvesse võtma; hinnas tõusma, hinda tõstma	appreciate [economics] hindama; arvesse võtma; hinnas tõusma, hinda tõstma, väärtuses kasvama	vahela lisatud vaste ees
IRRETRIEVABLE korvamatu; lootusetu; tagasisaadamatu; taastamatu; heastamatu	irretrievable [economics] korvamatu, heastamatu; lootusetu; tagasisaadamatu; taastamatu	vahela lisatud vaste ees
REALITY reaalsus, tegelikkus	reality [economics] tegelikkus, reaalsus	kohavahetuse kohal
BILLION biljon (UK, Saksamaa), miljard (USA, Prantsusmaa)	billion [economics] biljon, miljard	sulgudes teksti väljajätu kohal
GROSS .. sum-maarne	gross [economics] .. sum-maarne ..	poolitusvihje kohal

Kolmandaks, kõigile FESTART inglise-eesti majanduse sõnaraamatust pärit märksõnadele on lisatud FESTART-sõnaraamatus erialamärgend [*economics*] (vt tabel 5, 2. veerg). Nii on tekkinud FESTART-sõnaraamatusse hulk “imelikke” artikleid, kus märksõnad korduvad (sellest oli ka eespool juttu), kusjuures ka erialamärgend on esitatud mõlemas artiklis (12a, 12b): sest kuskilt võetuna on materjalis juba olemas hääldusega ja märgendiga varustatud artikkel. Leidub artikleid, kus majanduse osa (nagu ka teised erialaterminid) on sattunud suisa üldsõnavaraosa keskele (13).

- (12a) **account** [‘ökaunt] [verb] lugema, arvama; punkte saavutama; tapma; [noun] arvestus; [*economics*] arve; seletus; jutustus [FESTART-sõnaraamat]
- (12b) **account** [*economics*] raamatupidamiskonto; konteerima, kirjendama; finantsaruanne; register; alaline klient; riigieelarve täitmise aruanne; börsitehingu arvutus; tasumata võlakohustus; aru/selgitusi andma; aruandlus, äriraamatud; vastutama, vastutust kandma [FESTART-sõnaraamat]
- (13) **principal** [‘prinsipl] [adjective] peamine; [noun] printsipaal; tööandja; direktor; [jurisprudence] peasüüdlane; [building] peatala; [*economics*] kapital; esinäitleja, esilaulja [FESTART-sõnaraamat]

Ekspertiisi käigus selgus, et FESTART inglise-eesti majandussõnaraamatus on võrdluses TEA inglise-eesti majandussõnaraamatuga omakorda täpseid MV-gruppide vastavusi 26 698 (97%). Elektroonilise küljenduspõhja kasutamisest annavad taas tunnistust poolitusvihjete täpne kordamine, kokkulangevad õigekirjavead jm. Näiteks leidub äärmiselt reetlik kirjaviga vastes äriraamatud, mis pärineb juba TEA inglise-eesti majanduse pabersõnastikust, esinedes nõnda igas kolmes allikas: esimesena TEA-s, seejärel kahes Festardi sõnaraamatus.

5.2. Tehnikaterminid

Pärast majandussõnavara (27 387 märksõna koos oma MV-gruppidega) eraldamist on vaadeldud allesjäänud osa (42 613 märksõna), milles omakorda eristub suur rühm – märksõnad ilma häälduseta, erialamärgenditega [*zoology*], [*chemistry*], [*physics*], [*building*], [*geology*] jne. Selliste erialamärgenditega sõnu oletasid eksperdid pärinevat TEA teaduse ja tehnika sõnaraamatust. Ilma häälduseta märksõnu, mis pole majandusterminid, on kogu FESTART-sõnaraamatus 25% (kokku 18 024 märksõna). Ühelt poolt kuuluvad siia hulka märksõnad, mille puhul sõnastikes hääldust sageli ei esitatagi (akronüümid, keemilised elemendid jms) ja märksõnad ilma igasuguse erialamärgendita. Teiselt poolt kattub osa erialatermineist tavakasutuses olevate sõnadega – siis on üldjuhul ka hääldus olemas (tõenäoliselt üle võetud TEA Passwordist, sellest on juttu allpool). TEA teaduse ja tehnika sõnaraamatuga toimiti tõenäoliselt järgneval viisil.

TEA sõnaraamatu elektrooniline versioon koosnes küljendusprogrammi Ventura failidest, millest mittevajaliku info ja kujunduselementide kustutamine võtab kogenud programmeerijal aega vaid mõne tunni (ekspertidel oli kasutada oma programmeerija, kes selle läbi proovis). Mittevajaliku info hulka kuulusid näiteks

inglisekeelsed seletused ja tähendusnumbrid. Sel viisil “puhastatud” sõnastikku jäid vaid inglise märksõnad, erialamärgendid ja eesti vasted, kokku 51 692 märksõna. Materjali võrdlus näitab, et sellest on otse üle võetud ligi 40% seal leiduvatest märksõnadest (kokku üle 20 000). Kogu FESTART-sõnaraamatust moodustab TEA teaduse ja tehnika sõnaraamatust pärit osa 16%. Ülevõtmine on võinud toimuda järgmiste üldreeglite alusel:

- 1) välja on jäetud mitmesõnalised inglise märksõnad (st märksõnad, mis sisaldavad tühikut);
- 2) kustutatud on kõik, mis oli sulgudes (meetod on tuttav juba varasemast, nii tekib näiteks vastete kordumine);
- 3) teisenduse käigus on kaotsi läinud kursiiv, fondivahetus jm, mis põhjustab sisulisi vigu: väärraid märksõnu ja seletusi; palju jälgi on endast maha jätnud numbrite töötl(emat)us;
- 4) kaotatud on tähendusliigendus: kustutatud on tähendusnumbrid, semi-koolon on asendatud komaga;
- 5) mitmest erialamärgendist on alles jäetud ainult esimene;
- 6) erialamärgendid on asendatud (osalt välja kirjutatud, osalt muudetud) (vt tabel 6).

Kõik loetletud asjaolud kinnitavad arvutiprogrammi, mitte inimesest koostaja kasutamist.

Tabel 6. Võrdlus TEA teaduse ja tehnika sõnaraamatuga: tüüpilisi veakohti

TEA teadus ja tehnika	FESTART-sõnaraamat	Selgitus
ageing [Eng] .. • .. 2. vananemine (plastide korral); 3. vananemine (ferromagneetikute korral)	ageing [physics] vananemine, vanandamine .. [physics] vananemine; vanandamine, vanandus	kustutatud on sulgudes osa: tulemuseks korduv erialamärgend, korduvad vasted
α-particle	α-particle	teisenduse käigus on kaotsi läinud erisümbolid: väär märksõna
sthene jõu ühik meeter-tonn-sekund-süsteemis, ekvivalentne 10 ³ N-ga	sthene jõu ühik meeter-tonn-sekund-süsteemis, ekvivalentne 103 N-ga	teisenduse käigus on kaotsi läinud ülaindeksid: väär seletus
albedo 1. albedo, peegeldusvõime; 2. neutronite peegeldumise võime pinnalt	albedo [physics] albedo, peegeldusvõime, neutronite peegeldumise võime pinnalt	kaotatud on tähendusliigendus, semi-koolon on asendatud komaga
mucilaginous [Bot, Zool] lima-	mucilaginous [botany] lima-	mitmest erialamärgendist on alles jäetud ainult esimene
[Aero] [Bot] [Behav] [ImageTech] [Space]	[aviation] [botany] (puudub) [photography] [space]	erialamärgendid on asendatud

folly .. • dekoratiivne pargirajatis <u>18. s.</u>	folly .. [architecture] dekoratiivne pargirajatis <u>1s.</u>	programm on (punktiga) numbrit pidanud tähendusnumbriks (8.) ja eemaldanud
grit .. • pulber kõvadest osakestest suurusega üle <u>76µm</u>	grit .. [technics] pulber kõvadest osakestest suurusega üle <u>76µm</u>	esimene sümbol (kreeka täht) on numברי järel kustutatud
singles .. • kivimitükk: katusekilt 12x8 tolli, süsi 1- <u>0,5 tolli</u> X-band .. • X-laineala: 1. sagedusriba 8...12 GHz; 2. radarite tööriba 8,8...10,68 MHz	singles .. [mineralogy] süsi 1- <u>0</u> X-band [technics] radarite tööriba <u>8</u>	programm on pidanud numbris esinevat koma vastete eraldajaks, mida inimtoimetaja kunagi ei teeks

Sellele, et FESTART-sõnaraamatu tegemisel võib olla kasutatud TEA teaduse ja tehnika sõnaraamatu elektroonilist põhja, osutavad taas TEA teksti vigased kohad (nt *madalatemperatuurne*, *pinuase nõlv* jt), mis parandati enne TEA teaduse ja tehnika lisaköite ilmumist (*madalatemperatuuriline*, *pinnasenõlv* jt). Samad vead on aga FESTART-sõnaraamatus täiesti olemas. Ekspertiisi käigus leiti ka selliseid kattuvaid kirjavigu, mida lisaköite tegemisel veel polnud märgatud (*harkelunieemaldus* pro *harkelundieemaldus* jt).

5.3. Üldsõnavara

Pärast kahest eelnevast allikast (TEA inglise-eesti majandus, TEA teadus ja tehnika) pärinevate MV-gruppide eemaldamist jäi alles umbes pool esialgsest mahust (35 298 märksõna, 48 000 MV-gruppi). Järgnevalt võrreldi seda materjali üldkeelt esindava TEA Passwordiga (23 629 märksõna, sh allmärksõnad – tuletised ja väljendid), kus iga tähenduse puhul on vasteid esitatud harilikult üks, harvemini kaks-kolm. Kuna eksperdid kahtlustasid TEA Passwordi ülevõtmist juba varem ilmunud kirjastuse Festart elektroonilisse sõnastikku (FESTART Keyword), siis lubati viimast kasutada täiendava võrdlusmaterjalina.

FESTART Keyword koosneb Festardi sõnastikele iseloomulikult hästi eristuvast majanduse- ja üldsõnavara osast. Lihtne analüüs annab alust oletada, et vähemalt kolmveerand FESTART Keywordi mittemajandussõnadest on üle võetud TEA Passwordi elektroonilisest põhjast. Kopeeritud on tervikartiklid, st peale märksõnade ja vastete ka hääldused, seletused ja näitelauseid. Eespool juba mainitud 69 TEA Passwordi poolitusveast 68 “esindatus” FESTART-sõnaraamatus kordub ka siin, ajaliselt varasemas Keywordis. Võrdluse käigus erinevaks tunnistatud artiklites on erinevused põhjustatud hääldusmärkide erinevast kodeeringust jt vormistuslikest erinevustest (näiteks punkt iga artikli ja iga tähenduse lõpus, vt tabel 7), mida kuigi tõsiselt ei saa võtta. Ekspersedid ei pidanud täpsemat kvantitatiivset analüüsi tarvilikuks, vaid pidasid tõendusmaterjaliks sisulisi (kvalitatiivseid) andmeid. Kuna TEA Passwordi korrektuur ja küljendus on olnud korrektne, siis leidub ka väga vähe trükivigu, millele toetuda.

Tabel 7. Festart Keywordi võrdlus TEA Passwordiga

TEA Password	FESTART Keyword	Selgitus
keep away to (cause to) remain at a distance: <i>Keep away -- it's dangerous!</i> • eemale hoidma	keep away to (cause to) remain at a distance: <i>Keep away - it's dangerous!</i> • eemale hoidma.	lisandus: punkt artikli lõpus
keep back 1 not to (allow to) move forward: <i>She kept the child back on the edge of the crowd;</i> <i>Every body keep back from the door!</i> • eemal hoidma 2 not to tell or make known: <i>I feel he's keeping the real story back for some reason</i> • enda teada hoidma	keep back 1 not to (allow to) move forward: <i>She kept the child back on the edge of the crowd;</i> <i>Every body keep back from the door!</i> • eemal hoidma. 2 not to tell or make known: <i>I feel he's keeping the real story back for some reason.</i> • enda teada hoidma.	lisandus: punkt iga tähenduse ja artikli lõpus; viga: p.o everybody (viga pärineb TEA Passwordist, kus rea lõpust puudub poolitusmärk, mida asendab tühik)
aggressively appreciatively comprehensively	aggressively appreciatively comprehensively	ilma häälduseta
mechanic .. • mehhaanik mechanical .. • mehhaaniline mechanically .. • mehhaaniliselt	mechanic mehaanik mechanical mehaaniline mechanically mehhaaniliselt	ebajärjekindlus h/hh tarvituses

Sellegipoolest on võimalik analüüsida (ja hinnata) kõigile sõnastikele omaseid parameetreid nagu sõnavalik, hääldused, vastete kokkulangevused, lisamärgendite kasutamine, artikli struktuur jm.

TEA Passwordis on lihtmärksõnu (vastandina mitmeosalistele) 18 647, neist 17 912 (96%) leidub FESTART-sõnaraamatus. Sõnastike sarnast tüüpi ja võrreldavat mahtu arvestades on see ootuspärane tulemus. MV-gruppe on TEA Passwordis kokku 31 391, millest 21 632-le (69%) leidub täpne paariline FESTART-sõnaraamatus. Samas puudub FESTART-sõnaraamatust umbes 8000 TEA Passwordi MV-gruppi (25% lihtmärksõnadest, nt *keenness* 'ind', *kidnap* 'inimest röövima' jt). Ebatäielikule kattumusele vaatamata võib FESTART-sõnaraamatust leida kokkulangevusi, mis endiselt seovad teda TEA Passwordi sisuga (haruldased vasted), nagu ka elektroonilise põhjaga (küljendusest pärit poolituskriipsud).

Üht osa TEA Passwordi märksõnadest – mis a) on ilma häälduse ja sõnaliigita, ja b) pole pärit TEA teaduse ja tehnika sõnaraamatust (kokku 1141) – võrreldi täiendavalt EKI sõnastikuga. Kui MV-gruppide kokkulangevus TEA ja FESTARDI vahel oli 97%, siis TEA ja EKI vahel ainult 53% (!). Kui TEA sõnaliigimärgendit sisaldavad MV-grupid kattusid FESTART-sõnaraamatuga 92% ulatuses, siis EKI-ga üksnes 24%. Need võrdlevad andmed võiksid olla kõnekas argument sõnaraamatu originaalsuse tõestamisel.

Kuna ekspertidele torkas pidevalt silma FESTART-sõnaraamatu heterogeensus, mis andis alust arvata, et sellegi allesjäänud materjali saamiseks on mehhaaniliselt ühendatud mitu allikat, siis võrreldi materjali ka Silvet 1989 sõnaraamatuga. Kahtlus, et ülejäänud materjal FESTART-sõnaraamatus pärineb Silvet 1989 sõnaraamatust, leidis kinnitust. Sellele viitab mitu asjaolu. Esiteks, peaaegu kõik eesti vasted järelejäanud FESTART-sõnaraamatu materjalis esinevad täpselt samas järjestuses Silvet 1989 vastete hulgas. Teiseks, leidub vasteid, mille tarvitus on ebaharilik, nüüdiskeeles vananenud ja mis ei esine ei tavakeeles ega üheski muus

sõnaraamatus, nt *adobe* – plonnohooone, *brief* – kohtule esitatav kaitse-kokkuvõte, *tight* – pinguline, *bridle* – pahandudes pead üles ning lõuga tagasi tõmbama jpt. Kolmandaks, sellise silmale märkamatu elemendi nagu (elektroonilises versioonis) tühik põhimärksõnaartikli lõpus leiame täpselt nendest kohtadest, kus Silvet 1989 on esitanud sulgudes kommentaare. Tühik – mis elektroonilisel kujul muutub “nähtavaks” – jäi artikli lõppu tõenäoliselt pärast seda, kui sulgudes olnud tekst kustutati (17). Sisuline viga tekkis samal põhjusel (sulgudes osad jäeti mehaaniliselt kõrvale): nõnda sugenes artiklisse kaks identset vastet (18b).

(17a) **adulterer** abielurikkuja (mees) [Silvet 1989]

(17b) **adulterer** abielurikkuja_ [FESTART-sõnaraamat]

(18a) **current** vool, voolus, (mere)hoovus, (elektri)vool [Silvet 1989]

(18b) **current** vool, voolus, hoovus, vool [FESTART-sõnaraamat]

Ekspertide hinnangul on Silvet 1989 vastete suurest hulgast (umbes 150 000 MV-paari) üle võetud umbes kolmandik. Vastete ülevõtmise kopeerimise teel välistab asjaolu, et 1999. aastal polnud sõnastiku elektroonilist varianti veel olemas: nii on osa kirjavigu tekkinud teksti sisestamisel (nt *silmamoondusrikke* pro *silmamoondustrikke*). Et FESTART on ka omalt poolt vasteid lisanud, sellele viitavad leitud vigased märksõnad või vasted, mis sellisena mujal ei esine (nt märksõna *thunderstruck* pro *thunderstruck*, vaste *äärmuste abinõude* pro *äärmuslike abinõude*).

5.4. Teise ülesande vastus

FESTART-sõnaraamatu majandussõnavara pärineb tõenäoliselt täielikult TEA inglise-eesti majandussõnaraamatust, olles peitva vaheastmena üle võetud varasemast FESTART inglise-eesti majandussõnaraamatu arvutiversioonist. Majandussõnavara osakaal kogu sõnaraamatus on 23% (vt tabel 9).

FESTART-sõnaraamatu tehnikasõnavara pärineb tõenäoliselt TEA teaduse ja tehnika sõnaraamatus, kust on üle võetud ligi 20 000 MV-gruppi. Tehnikasõnavara osakaal kogu sõnaraamatus on 16% (vt tabel 9).

FESTART-sõnaraamatu üldsõnavara pärineb kahest allikast: TEA Passwordist (4000 MV-gruppi) ja Silvet 1989-st (hinnanguliselt 69 000 MV-gruppi). Üldsõnavara osakaal kogu sõnaraamatus on 59% (3% + 56%) (vt tabel 9).

FESTART-sõnaraamatu originaalloomingu suuruseks on hinnanguliselt 2% (vt tabel 9).

Tabel 9. Eri sõnastike osakaal FESTART-sõnaraamatus

Sõnavararühm	Võrdlusobjekt	FESTART-sõnaraamatuga kattuvate MV-gruppide arv	Osakaal (%) FESTART-sõnaraamatus
Majandus-sõnavara	TEA inglise-eesti majandus	ligi 28 000	23%
Tehnika-sõnavara	TEA teadus ja tehnika	üle 20 000	16%
Üldsõnavara	TEA Password	4 000	3%
	Silvet 1989	(umbkaudne hinnang:) 69 000	56%
	Kirjastuse FESTART omalooming	(umbkaudne hinnang:) pisut üle 2 000	2%
FESTART-sõnaraamatus kokku		122 915	100%

6. Kolmas ülesanne: võimalik koostamisskeem

Eksperdid, kes lähtusid kirjastuse Festart väitest, et vaidlustatud FESTART-sõnaraamatu koostamine algas FESTART inglise-läti sõnastiku põhjal, pakkusid välja järgmise skeemi:

- 1) aluseks võeti märksõnastik ja hääldused FESTART inglise-läti sõnastikust, jättes välja mitmest sõnast koosnevad märksõnad;
- 2) jaotus sõnaliikideks võeti tõenäoliselt koos vastetega üle Silvet 1989 sõnaraamatust;
- 3) koostajad said tõenäoliselt ülesande lisada Silvet 1989 põhjal nii palju vasteid, kui nad tarvilikuks peavad;
- 4) nii valmis sõnastik, mille maht oli vähemalt 32 000 märksõna (koos häälduste, sõnaliikide ja erialamärgenditega). Sellele tehti pealiskaudne korrektuur;
- 5) saadud sõnastikku täiendati puuduvate märksõnade ja vastetega TEA Passwordist: programmiga nopiti välja ühesõnalised märksõnad ja nende vasted. Tekkisid esimesed ebaühtlused: osa märksõnu varustati nii häälduse kui ka sõnaliigiga, osa ainult sõnaliigiga, osa üksnes eesti vastega. Märksõnade arv kasvas 3000 võrra (kokku 35 000);
- 6) seejärel täiendati sõnastikku tõenäoliselt TEA teaduse ja tehnika sõnaraamatust pärit märksõnade ja vastetega (välja jäeti mitmest sõnast koosnevad märksõnad). Märksõnade arv kasvas üle 13 000 võrra (kokku üle 48 000). Erialamärgendite ühtlustamisel programmidega tekkis läbivaid vigu;
- 7) saadud sõnastikule lisati tõenäoliselt kogu FESTART inglise-eesti majandussõnaraamat (27 500 märksõna – mis omakorda oli üle võetud TEA inglise-eesti majandussõnaraamatu küljenduspõhjal). Kõik need lisatud märksõnad moodustasid eraldi üksuse (seetõttu esineb korduvaid märksõnu);
- 8) sõnastik sai oma lõpliku kuju: valminud FESTART-sõnaraamatus on 71 154 märksõna.

7. Neljas ülesanne: kas tegemist on originaal- või tuletatud teosega?

Plagiaatlus selle sõna laiemas tähenduses on sõnastike koostamisel vältimatu. Ühtki sõnastikku ei koostata nullist ja mis tahes uue sõnastiku autor saab materjali olemasolevatest sõnastikest. Kui selle tegevuse käigus a) viidatakse algallikatele, b) omandatakse litsentsiluba, ja c) luuakse loominguiliselt iseseisev ja algsest teosest sõltumatu teos – siis on (autoriõiguse sõnastuses) tegemist tuletatud teosega, mida plagiaatluseks ei nimetata.

FESTART-sõnaraamatut ei pea eksperdid originaalsõnastikuks, sest eksperitiisi tulemused näitavad, et selle materjal on mehaaniliselt kokku kuhjatud neljast sõnaraamatust. Kirjastuse Festart omapoolne, st originaalne panus on ainult 2%. Sellise mehaanilise ülevõtmise puhul ei tohiks autoriõiguste rikkumise üle otsustamisel saada määravaks asjaolud, et alusmaterjali ei võetud üle tervikuna (Silvet 1989 vastetest on kirjutatud ümber kolm viiendikku, TEA Passwordist viiendik, TEA teaduse ja tehnika sõnaraamatust umbes kolmandik) või et saadud sõnastik ei koosne enam ühest, vaid mitmest eri allikast. Juba algsõnastikele viitamata jätmine on käsitletav autoriõiguse rikkumisena, seda ka juhul, kui on leping, mis lubab nimetatud teoseid kasutada (AÕS § 35 lg 3).

Autorikaitse objektiks saab olla terviklik teos/andmekogu, sõnastike puhul on võimalik asuda seisukohale, et autorikaitse alla peaks kuuluma autori lisatud loominguiline panus, mis võib väljenduda märksõnavalikus, sõnastiku ülesehituses, uute vastete, definitsioonide, kasutusnäidete, kommentaaride, illustratsioonide jm lisamises vm viisil. Vaidlustatud sõnaraamat ei vasta ühelegi neist kriteeriumeist. Koostajate tegevus on mõnevõrra märgatav vaid majandussõnavara puhul.

Kõige selle põhjal on eksperdid FESTART-sõnaraamatu nimetanud tuletatud, mitte originaalteoseks. Autoriõiguse seaduse järgi on tuletatud teos „. teose tõlge, algse teose kohandus (adaptsioon), töötlus (arranzžering) ja teose muu töötlus.” (AÕS § 4 lg 21)

8. Lõpetuseks: oportuuniteedist

Kuigi ekspertiisi tõestas üsna üheselt vaidlustatud sõnaraamatu mehaanilise kompleerimise käigu, ilma selleks luba omamata, lõpetati kriminaalasi, mida Eesti õigussüsteemis oli menetletud seitse aastat, 2007. aastal oportuuniteedi alusel (Truu 2007). Mida selle all mõista? Aga seda, et riigi toimimises, nagu iga inimese elus, arvatakse mõned sündmused, teod või probleemid vähem oluliseks kui teised: oportuuniteediprintsiip on põhimõte, mis seadusena võimaldab otsustada riigi asjade otstarbekuse üle.

Eesti Vabariigi kriminaalkodeksisse lisandus oportuuniteedipunkt (KrMS § 168 täiendusena) uuesti 2004. aastal.¹¹ Oportuuniteedi põhiline mõte ongi olnud eelkõige vähendada politseiuurijate, prokuröride ja kohtute koormust, st lubada otsustada juba alamal astmel. Kriminaalmenetluse võib oportuuniteedile toetudes lõpetada, kui puudub avalik menetlushuvi ja kui süü ei ole suur (KrMS § 168, 202) või kui karistus on ebaotstarbekas (KrMS § 168, 203). Prokurör Endla Ülviste on

¹¹ Nõukogude kodeksis oli see samuti olemas, ja toimis väga hästi, nagu on kinnitanud Ain Seppik, kelle sõnatsi “seal tegi seda uurija”, mitte prokurör või kohus (väljavõte Riigikogu õiguskomisjoni istungi protokollist (23.10.2003).

küll väljendanud oma kergendust: “Te ei kujuta ette, kui palju on kohtutes neidsamu pasteetide, õllepudelite ja saiakeste lugusid” (Püüa 2004). Samas jääb kummitama võimalus “teise astme kuritegu niiöelda lämmatada”.¹² Ja kuidas ikkagi mõõta avalikkuse huvi? Kas lugeda kokku ajakirjanduses avaldatud kirjutised? Või asjasse segatud inimeste hulk? Või veel muud moodi? 2007. aastal avaldatud kokkuvõttest selgub, et lõviosa kriminaalmenetlustest ongi lõpetatud avaliku huvi puudumise tõttu (92% ehk 5239 kriminaalasja), vähem on toetunud ebaotstarbeka karistuse sättele (8% ehk 465 juhtu).¹³

Ehmatava ja kurvastava tõsiasi jana lisandus 2007. a sellesse pasteedi-õllepudeli-saiakeste loendisse ka loometöö vili – sõnaraamat –, ja samuti avaliku huvi puudumise tõttu. Päris karistuseeta vaidlustatud sõnaraamatu tegijad siiski ei jäänud: Harju maakohus otsustas, et S. Gluhhov peab riigi tuludesse tasuma 125 000 krooni ning tegema 100 tundi üldkasulikku tööd, firma Festart peab maksma riigi tuludesse 250 000 krooni, tasuma kriminaalmenetluse kulud (413 682 krooni) ja kannatanu TEA Kirjastuse kulud (216 889 krooni).

9. Ettepanek tuleviku jaoks

Eesti Keele Instituudis välja töötatud sõnastike kokkulangevust hindava märksõna-vaste-gruppide mõõtmise meetodi ja selle rakendamise põhjal kakskeelsete sõnaraamatute ekspertiisil tuleks kehtestada piirmäär, millest edaspidi lähtuda sõnaraamatu originaalsuse üle otsustamisel.

Tuletame veel meelde siinsest artiklist läbi käinud protsente ja fakte: 18. saj inglise Fisheri õigekirjasõnaraamat, kus üks-ühele oli teisest üle võetud 43,5%, lubati kohtuvaidluste järel siiski välja anda; teise, 17. saj inglise juhtumi puhul olevat üle võetud koguni 70% märksõnadest koos vastetega, aga too asi kohtuni ei jõudnud. FESTART inglise-eesti majandussõnaraamatu MV-gruppidest langes TEA majandussõnaraamatu materjaliga kokku 97%. Ühe kontrollrühma (1141 märksõna) osas oli FESTART-sõnaraamatu kokkulangevus TEA Passwordiga samuti 97%, samas EKI sõnastikuga 53%. Kui samas kontrollrühmas võeti arvesse kolmas üksus (sõnaliigimärgend), siis kattusid MV-grupid TEA-ga 92%, aga EKI sõnastikuga üksnes 24% (!) ulatuses. Korduva märksõna-osisega artiklitest (ligi 5000) on FESTART-sõnaraamatus 90% selliseid, kus märksõna kordub muutu-matul kujul. Ekspertiisi käigus leitud 69-st TEA Passwordi poolitusveast kordub 68 (98,5%) FESTART-sõnaraamatus.

Autorite ettepanek on kehtestada kahes sõnaraamatus kokkulangevate märksõna-vaste-gruppide piirmääraks umbes 60–75% (pöörates muidugi tähelepanu ka muudele “arvutinippidele”). Üle selle on tegemist plagiadiga, alla selle (tuletatud) originaalteosega.

¹² Märkt Raski sõnad Riigikogu õiguskomisjoni istungil (väljavõte protokollist 23.10.2003).

¹³ Justiitsministeeriumi kriminaalpoliitika osakonna ülevaade, ptk “Kriminaalmenetluse oportuniteediga lõpetamine (1.07.2004–31.03.2007)”, vt www.just.ee (6.10.2007).

Kirjandus

- Arujärvi, Evi 2008. Kultuur – missioon ja äri. – Eesti Päevaleht, 15.02.2008.
- AÕS = Autoriõiguse seadus. Elektrooniline Riigi Teataja, <https://www.riigiteataja.ee/ert/act.jsp?id=12791124> (20.02.2008).
- Cormier, Monique C.; Fernandez, Herberto 2005. From the Great French Dictionary to the Royal Dictionary. – *International Journal of Lexicography*, Vol. 18, No. 4, 479–507.
- EEVA = Ross, Kristiina. Anton Thor Helle (1683?–1748). Piiblitõlkija ja eesti kirja-keele arendaja, Jüri pastor. – Eesti vanema kirjanduse digitaalne tekstikogu. www.utlib.ee/ekollekt/eeva/ (20.11.2007).
- Helle, Anton Thor 1732. Kurtzgefaszte Anweisung Zur Ehstnischen Sprache, in welcher mitgetheilet werden I. Eine GRAMMATICA. II. Ein VOCABULARIUM. III. PROVERBIA. IV. AENIGMATA. V. COLLOQUIA. Tallinn: Eestimaa Konsistooriumi kirjastuskassa.
- Hupel, August Wilhelm 1780. Ehstnische Sprachlehre für beide Hauptdialekte, den revalschen und dörptschen, nebst einem vollständigen Wörterbuch. Riga-Leipzig: Hartknoch.
- Kask, Arnold 1955. S. H. Vestringi eesti-saksa sõnaraamat ja selle osa eesti leksikograafias. – Ajaloo-keeleteaduskonna töid. Tartu Riikliku Ülikooli toimetised 38. Tartu: Tartu Riiklik Ülikool, 181–192.
- Krikmann, Arvo 1986. Fraseoloogiline aines eesti vanemais grammatikates ja sõnastikes. Tallinn: Valgus.
- KrMS = Kriminaalmenetluse seadustik. Elektrooniline Riigi Teataja, <https://www.riigiteataja.ee/ert/act.jsp?id=12897094> (20.02.2008)
- Landau, Sidney I. 2001. *Dictionaries: The Art and Craft of Lexicography*. 2nd. ed. Cambridge University Press.
- McMenamin, Gerald R. 2002. *Forensic Linguistics. Advances in Forensic Linguistics*. CRC Press.
- Püüa, Marko 2004. Pasteedivaraste karistused leevenesid. – Postimees, 25.03.2004. Vt ka <http://www.postimees.ee/260304/esileht/siseuudised/129949.php> (5.03.2008).
- Rodriguez-Alvarez, Alicia; Rodriguez-Gil, Maria 2006. John Entick's and Ann Fisher's Dictionaries: An Eighteenth-century Case of (Cons)Piracy? – *International Journal of Lexicography*, Vol. 19, No. 3, 287–319.
- Truu, Aet 2007. Kohus lõpetas piraatkoopiaste valmistamise süüdistatava kriminaalaja oportuuniteedi alusel. – Äripäev, 04.04.2007. Vt ka <http://www.ap3.ee/Default2.aspx?ArticleID=5aa5daae-8403-4fco-a6b2-17848c1adfc3> (5.03.2008).
- Vestring, Salomo Heinrich 1998 [käsikiri ?1720-ndad]. *Lexicon Esthonicum Germanicum*. Trükki toimetanud Ellen Kaldjärvi. Tartu: Eesti Kirjandusmuuseum.
- Williams, John 1992. The question of plagiarism and breach of copyright in the dictionary-making process (with particular reference to the UK). – *EURALEX '92 Proceedings*. Tampere: Tampereen Yliopisto, 561–570.

Margit Langemets (Eesti Keele Instituut) on lõpetanud Tartu Ülikooli eesti keele ja kirjanduse erialal. Osales TEA-Festardi kohtuprotsessis tunnistajana. Uurimisvaldkonnad: leksikoloogia ja leksikograafia üldiselt, arvutileksikoloogia, leksikaalne semantika. margit@eki.ee

Piret Voll (Eesti Keele Instituut) on lõpetanud Tartu Ülikooli eesti keele ja kirjanduse erialal. Osales TEA-Festardi kohtuprotsessis peaekspertina. Uurimisvaldkonnad: leksikoloogia ja leksikograafia. piret.voll@eki.ee

LINGUISTIC FORENSIC ANALYSIS OF A DICTIONARY: AN ESTONIAN PRECEDENT

Margit Langemets, Piret Voll

Institute of the Estonian Language

In 2003 the first linguistic forensic assessment of lexicographic objects took place in Estonia. The case required expert assessment of the contents of several dictionaries involved in criminal accusation of theft of intellectual property. The article describes the process of lexicographic expert assessment carried out at the Institute of the Estonian Language, and the ad hoc language-technological method applied. Relying on the results, a critical measure has been elaborated to judge over the degree of originality of bilingual dictionaries. Being the first and only case of its kind in Estonia, it offers a linguistic-forensic precedent.

Keywords: forensic linguistics, lexicographic expert assessment, dictionaries, English, Estonian

DIFFERENCES IN ADULT L2 GRAMMARS OF LITHUANIAN: THE CASE OF GENDER AGREEMENT

Simona Mačiukaitė

Abstract. This article presents results obtained from an empirical investigation into the acquisition of gender agreement in Lithuanian as a second language (L2). A total of 59 L2 learners of Lithuanian and 10 Lithuanian native speakers, as controls, took part in this study. The participants came from diverse first language (L1) backgrounds (i.e., Russian, Ukrainian, Belorussian, Latvian, and English). The data were collected using a written translation task. The results of the study showed some statistically significant differences between English speaking learners of Lithuanian and non-native English speaking learners of Lithuanian. The findings support the Full Transfer Full Access hypothesis (Schwartz, Sprouse 1994, 1996) which argues for the interaction of L1 transfer and Universal Grammar in the development of adult L2 grammars.

Keywords: non-native language acquisition, adult learner, Universal Grammar, initial state grammar, Lithuanian

1. Introduction to the study

Languages exhibit variation in gender agreement. In Lithuanian, for instance, every noun is marked for gender and most modifiers preceding or following the noun must agree in gender with it, whereas in English, gender marking on nouns, as well as gender agreement within the noun phrase are absent. Consider representative examples given in (1).

- (1a) Sulting-as obuol-ys
 Juicy-masc apple-masc
 ‘A juicy apple’

- (1b) Sulting-a slyv-a
 Juicy-fem plum-fem
 ‘A juicy plum’

Since languages vary with respect to gender agreement, the question is, how does a second language (L2)¹ learner go about acquiring a new feature, not present in his/her language, and how s/he deals with the additional morphological complexity associated with the new feature (i.e., gender agreement). For example, Lithuanian belongs to a fusional type of language where a single morpheme marks number, case and gender. Consider representative examples given in (2).

- (2a) Ger-as vaik-as valg-o obuol-ius
 Good-sg.masc.nom child-sg.masc.nom eat-pres apple-pl.masc.acc
 ‘The good child is eating apples’
- (2b) Mokytoj-a pagyr-è ger-a vaik-a
 Teacher-sg.fem.nom praise-past good-sg.masc.acc child-sg.masc.acc
 ‘The teacher praised the good child’
- (2c) Ger-a mergait-è dav-è man knyg-a
 Good-sg.fem.nom girl-sg.fem.nom give-past I-dat book-sg.fem.acc
 ‘The good girl gave me the book’
- (2d) Mokytoj-a pagyr-è ger-a mergait-è
 Teacher-sg.fem.nom praise-past good-sg.fem.acc girl-sg.fem.acc
 ‘The teacher praised the good girl’
- (2e) Ger-i vaik-ai valg-o obuol-ius
 Good-pl.masc.nom children-pl.masc.nom eat-pres apple-pl.masc.acc
 ‘Good children are eating apples’
- (2f) Mokytoj-a pagyr-è ger-us vaik-us
 Teacher-sg.fem.nom praise-past good-pl.masc.acc children-pl.masc.acc
 ‘The teacher praised the good children’
- (2g) Ger-os mergait-ès valg-o obuol-ius
 Good-pl.fem.nom girl-pl.fem.nom eat-pres apple-pl.masc.acc
 ‘The good girls are eating apples’
- (2h) Mokytoj-a pagyr-è ger-as mergait-es
 Teacher-sg.fem.nom praise-past good-pl.fem.acc girl-pl.fem.acc
 ‘The teacher praised the good girls’

From examples in (2), we see that in Lithuanian, morphological inflections vary depending on the case, number and gender of the noun phrase. Specifically, examples in (2a), (2b), (2c) and (2d) illustrate the change in the morphological inflection of the noun phrase when the case marking for the phrase changes from nominative to accusative but gender and number stay the same (the noun phrases in (2a/2b) are marked for masculine gender/singular number and the noun phrases in (2c/2d) are marked for feminine gender/singular number). Examples given in (2e–2h) illustrate the changes in inflectional morphology when masculine and feminine

¹ In this paper, L2 is used to cover a term for non-native language acquisition of second, third, fourth, or even fifth language.

plural noun phrases (2e/2f and 2g/2h respectively) appear in two different positions, subject and object. In addition to this, even though there is always agreement between the noun and its modifier in all three features (number, case and gender), the morphological markings on the noun and the adjective do not always match (cf. examples in 2).

In Lithuanian, nouns and adjectives are declined according to the existing declensions (i.e., five nominal declensions and three adjectival). Consider representative examples of each declension given in (3) and (4).

- (3a) First Nominal Declension (masculine gender)
sod-as (nom.sg) – *sod-ai* (nom.pl) ‘garden’; *mėnul-is* (nom.sg) – *mėnul-iai* (nom.pl) ‘moon’; *pasiuntin-ys* (nom.sg) – *pasiuntin-iai* (nom.pl) ‘messenger’
- (3b) Second Nominal Declension (feminine gender)
sag-a (nom.sg) – *sag-os* (nom.pl) ‘button’; *kriauš-ė* (nom.sg) – *kriauš-ės* (nom.pl) ‘pear’
- (3c) Third Nominal Declension (masculine and feminine genders)
šal-is (nom.sg.fem) – *šal-ys* (nom.pl.fem) ‘country’; *žvėris* (nom.sg.masc) – *žvėrys* (nom.pl.masc) ‘animal’
- (3d) Fourth Nominal Declension (masculine gender)
dang-us (nom.sg) – *dang-ūs* (nom.pl) ‘sky’;
- (3e) Fifth Nominal Declension (masculine and feminine genders)
rud-uo (nom.sg.masc) – *ruden-ys* (nom.pl.masc) ‘autumn’; *ses-uo* (nom.sg.fem) – *seser-ys* (nom.pl.fem); *dukt-ė* (nom.sg.fem) – *dukt-ės* (nom.pl.fem)
- (4a) First Adjectival Declension
maž-as (nom.sg.masc) – *maž-i* (nom.pl.masc); *maž-a* (nom.sg.fem) – *maž-os* (nom.pl.fem) ‘little’; *geresn-is* (nom.sg.masc) – *geresn-i* (nom.pl.masc) ‘better’
- (4b) Second Adjectival Declension
stipr-us (nom.sg.masc) – *stipr-ūs* (nom.pl.masc); *stipr-i* (nom.sg.fem) – *stipr-ios* (nom.pl.fem) ‘strong’
- (4c) Third Adjectival Declension
nuolatin-is (nom.sg.masc) – *nuolatin-iai* (nom.pl.masc); *nuolatin-ė* (nom.sg.fem) – *nuolatin-ės* (nom.pl.fem) ‘permanent’; *geresn-ė* (nom.sg.fem) – *geresn-ės* (nom.pl.fem) ‘better’

From these examples, we see that to acquire Lithuanian nominal gender, a second language learner would have to memorize the inherent gender of each noun and would have to know the inflectional endings for each noun depending on the declension it belongs to, as well as the number and case of the noun. Now, to acquire a target-like gender agreement in Lithuanian, in addition to the above, the learner would have to learn the morphological endings of various modifiers, which also vary given the existing declensions and the case, number and gender of a given noun that

it would agree with. Besides the complexities of remembering all of these various inflections, the learner would also have to remember some of the ambiguous cases present in the Lithuanian nominal system.²

Given such complexity, the question is, how different is the learning of gender agreement between L2 learners from morphologically rich and morphologically poor first language backgrounds. Are these differences reflected in faulty gender agreement morphology or in the time it takes to acquire this grammatical feature? These are some of the questions posed for the study presented in this article.

To this day the work on gender agreement in the field of second language acquisition (SLA) has mainly focused on languages such as Spanish, French, and Dutch. Gender agreement is present in these languages, but when compared to Lithuanian gender agreement system, the systems do differ with Lithuanian showing by far more complexity. Thus, the hope is that the results of the present study will add more to the knowledge of how the acquisition of L2 gender agreement proceeds in languages that are morphologically very rich.

The organization of this paper is as follows. In the next section, the theoretical proposals for L2 acquisition are outlined. Then, a review of studies on the acquisition of gender agreement follows. Finally, the methods and the results of the study are presented and discussed.

2. Theoretical background of the study

In the field of SLA, several hypotheses have been put forth in order to explain how adult learners acquire grammatical features in L2 that are not present in their first language (L1) (for an in-depth overview of these hypotheses, see White 2003). One such hypothesis, namely The Failed Functional Features Hypothesis (FFFH) (Hawkins, Chan 1997) predicts that grammatical features that learner's L1 does not have but learner's L2 requires will not be fully acquired. The reason for this is that according to this view, adult learners no longer have direct access to the universal grammar (UG) which plays a central role in the acquisition of these features in L1. For adult L2 learners, UG can only be accessed indirectly, namely through L1 transfer. To illustrate how this works, let us take the languages represented in the current study sample. In the case of English learners acquiring Lithuanian gender agreement, the prediction of the FFFH will be that target-like acquisition of gender agreement will not take place because English language does not have this feature and it cannot be transferred to Lithuanian. Since UG is also unavailable to help these learners, the feature cannot be acquired in a target-like fashion. On the other hand, in the case of Ukrainian learners of Lithuanian, gender agreement will transfer to Lithuanian because gender agreement is present in Ukrainian and these learners will be able to acquire this feature in a target-like way.

Contrary to the FFFH, the Full Transfer Full Access (FTFA) hypothesis (Schwartz, Sprouse 1994, 1996) argues that UG is accessible to adult learners. The FTFA hypothesis makes it clear that L1 transfer is also involved in the building of L2 grammars. Specifically, the initial stages of L2 grammars are based on L1 grammars; however, UG comes into play when, for instance, L2 requires a feature that cannot be transferred from L1 because L1 does not possess it.

² As pointed out by a reviewer, the inflection for masculine 1st declension noun marked for singular number and nominative case and the inflection for feminine 3rd declension noun marked for singular number and nominative case is the same, namely *-is* (*mėnuš-is* 'moon' and *šal-is* 'country'). Also, pointed out by the same reviewer, the plural genitive ending is the same for all nouns (i.e., *ų*). In the current study, ambiguous cases were not used.

The essential difference between the FFFH and FTFA hypotheses is UG involvement proposed in the latter and the differences predicted for the end-state grammars of L2 learners. Using the same example of an adult English-speaking learner of Lithuanian, following the FTFA hypothesis, the prediction is that eventually gender agreement will be acquired by this learner. However, the initial state of an English-speaking learner would differ from the initial state of a Ukrainian-speaking learner since initially, L2 grammars will be built on L1 grammars. Since Ukrainian has gender agreement, this feature will transfer to Lithuanian and in the initial stages the learner will show his/her knowledge of this feature. On the other hand, the initial grammar of an English speaker will be based on English, which lacks the feature of gender agreement, and because of this, faulty gender agreement will surface more often in the L2 grammars of English speaking learners of Lithuanian, as opposed to Ukrainian-speaking learners. However, ultimately UG will aid English speaking learners of Lithuanian in the acquisition of gender agreement.

Lastly, the Full Access hypothesis (Flynn, Martohardjono 1994; Epstein et al. 1996, 1998), unlike the FFFH and the FTFA hypothesis, argues for full and direct UG access throughout the entire process of L2 acquisition. From this, it means that all adult learners despite their L1 backgrounds will proceed in a similar fashion when acquiring L2 grammars. Again, using the example of English and Ukrainian speakers of Lithuanian, both groups of learners will show similarities in the acquisition of gender agreement feature at all stages of their L2 grammar development since UG will guide this process.

To sum up the above discussion, the FFFH predicts that there will be differences in the grammars of the learners whose L1 have gender agreement and whose L1s do not have gender agreement when the L2 they are acquiring is Lithuanian (i.e., with gender agreement). The FTFA hypothesis predicts that initially the learners coming from languages with gender agreement and languages without it, will show differences, as L1 transfer will be the basis for their initial L2 grammars. However, this will change when learners pass through the initial stage. Finally, the Full Access view predicts that all learners, despite their L1 backgrounds, will show similarities in the acquisition of gender agreement since their initial state as well as the entire L2 development is driven by the direct access to UG.

3. Previous studies on gender agreement in SLA

The theoretical assumptions regarding how much influence L1 transfer or UG has on the acquisition of L2 grammatical features have been put to a test by a number of researchers in the field of SLA (see White 2003). A number of studies have also been conducted to examine the acquisition of gender agreement in order to determine which theoretical predictions (outlined in the previous section) hold true.

For instance, White, Valenzuela, Kozłowska-MacGregor and Leung (2004) investigated the acquisition of Spanish gender and number agreement using 48 French and 68 English native speakers of Spanish from three different proficiency levels (i.e., low, intermediate, and advanced). Both French and English languages have number agreement but vary with respect to gender agreement (i.e., English does not have it, whereas French does). The findings of the study showed that

only the proficiency level of a learner had an effect on the production of target-like gender agreement (i.e., learners with lower proficiency levels were less accurate in producing correct gender agreement in Spanish than were the learners with higher proficiency levels). Most importantly, the effect of the first language did not turn out to be significant in the acquisition of Spanish gender agreement, as would have been predicted by the FFFH. The authors concluded that gender agreement can be acquired by speakers whose first languages lack such features.

A few other interesting observations were made by White et al. (2004). They observed a difference in the production of number versus gender agreement. That is, number agreement was more accurately produced than gender agreement. And finally, a tendency among low proficiency learners was noticed to use a feminine noun with a modifier marked for masculine gender.

The opposite view, namely the FFFH, was pursued by Francheschina's (2001) study. In her study Francheschina examined the end-state grammar of an adult English speaker of Spanish, Martin, and investigated his knowledge of Spanish gender as reflected in the spontaneous production data. The findings of the study showed that in Martin's grammar there were more errors in gender agreement than in number agreement. Specifically, as given by Francheschina (2001: 237), in the subject's data there were a total of 239 (93%) gender agreement errors out of 257 and only 18 (7%) number agreement errors out of 257. Francheschina also found that Martin tended to substitute masculine gender for the required feminine or neuter. These findings led Francheschina to offer strong support for the FFFH.

Finally, the study conducted by Sabourin, Stowe, and de Haan (2006) examined the differences in the acquisition of Dutch inherent gender and gender agreement among English, German and Romance native speakers of Dutch as an L2. For the most part, this study was interested in finding out whether the same type of L1 transfer was used in the acquisition of inherent gender and gender agreement. In this study, L1 transfer was subdivided into two types, namely "surface transfer" and "deep transfer":

Surface transfer is represented by the direct transfer of morphologically similar gender realization between the L1 and L2. [---] Deep transfer, on the other hand, would be the transfer of the category gender (whether it is a congruent system or not) from the L1 to the L2. (Sabourin et al. 2006: 6)

The results of the study showed that all three groups, English, German, and Romance, showed excellent knowledge of inherent gender markings on Dutch nouns. However, when German and Romance groups were compared to the English-speaking group, the former outperformed the latter showing that transfer, more specifically "deep transfer", was involved in the assignment of inherent gender markings on nouns. Furthermore, when the results of German-speaking group were compared to the results of Romance group, Germans scored higher than Romance group showing the effect of "surface transfer" in the acquisition of inherent gender markings on Dutch nouns.

The findings regarding the acquisition of Dutch gender agreement argued for the significance of "deep transfer" since German and Romance groups did better than English speaking group in supplying correct Dutch gender agreement and no

differences of significant importance were observed among German and Romance groups.

Based on the theoretical proposals and the literature review presented above, the following research questions are formulated for our study: Does L1 play a role in the acquisition of Lithuanian gender agreement? Specifically, in the initial stages of L2 acquisition, do learners from L1s with gender agreement (i.e., Russian, Ukrainian, Latvian and Belorussian) outperform the English-speaking learners of Lithuanian?

By means of these research questions, the study sets out to find out which theoretical predictions presented above are true. If the findings show that all groups perform the same on Lithuanian gender agreement, we obtain support for the Full Access view. But if L1-related differences are observed, support for the FTFA view is obtained. Finally, if the English-speaking group performs significantly worse than other groups involved in the study, partial support is obtained for the FFFH. The reason why this would grant only partial support for the FFFH is because in the current study, only the learner's initial state, not the end state, is investigated.

4. Methodology

4.1. Subjects

The subjects in this study were 59 students from the *Introduction to Lithuanian I* class. Before taking this class, none of them had any exposure to Lithuanian. At the time of data collection, all students had 3 months of instruction in Lithuanian language. For the purposes of this study, the subjects were subdivided into 7 groups based on their native language. In the case of English speaking subjects, two groups were formed: first group (namely Group 6) consisted of students who had absolutely no prior knowledge of a language with gender agreement and second group (i.e., Group 7) consisted of students who have been exposed to a language with gender agreement from age 10 or later. Table 1 presents information about the subjects of the study.

Table 1. Subject information

Group number	Number of Subjects	L1(s)
1	17	Russian
2	10	Ukrainian/Russian
3	5	Latvian
4	4	Ukrainian
5	6	Russian/Belorussian
6	5	English
7	12	English

Besides the seven experimental groups, the study also had a control group which consisted of 10 native speakers of Lithuanian.

4.2. Data collection method

The data for this study were collected using a written translation task. The subjects were presented with an illustrated text in Lithuanian. They had to read the text and translate the English noun phrases (i.e., adjective + noun), verb phrases (*loves* ‘myli’) and prepositional phrases (*to the park* ‘į parką’) into Lithuanian. All subjects were fluent speakers of English. Since the subjects of this study were early beginners of Lithuanian, the noun phrases (NPs) that were used in the tasks were placed in the context which required nominative case only and all NPs were singular except for 3 cases that were plural and served as distracters of the study (e.g., *young sons* ‘jauni sūnūs’). Besides these three distracters, the study had a number of other distracters that took the form of prepositional phrases and verb phrases. When data were coded the distracters were not analyzed.

A representative example of the task is given in (5) and the task details are presented in Table 2.

- (5) Ana yra _____ (tall woman), o jos vyras Tomas yra _____ (short man).
‘Ana is a tall woman but her husband Tom is a short man’

Table 2. Task information

Number of Test Items	17
Number of Distracters	12
NPs with Grammatical Gender	5
NPs with Natural Gender	12
NPs with Masculine Gender	9
NPs with Feminine Gender	8

4.3. Analysis of data

In the analysis of the data the most conservative approach was taken in interpreting the responses given by the participants of this study. The process of the data analysis began with a transcription of the test items. Next, these items were coded using the following five categories:

- 1) Correct Masculine Agreement (CMA): (M+M)
žemas vyras ‘short (masc) man (masc)’
- 2) Correct Feminine Agreement (CFA): (F+F)
aukšta moteris ‘tall (fem) woman (fem)’
- 3) Incorrect Masculine Agreement (IMA): (*F+M)
**graži parkas* ‘beautiful (fem) park (masc)’
- 4) Incorrect Feminine Agreement (IFA): (*M+F)
**didelis kavinė* ‘big (masc) café (fem)’
- 5) Ambiguous

The distracters were not coded. The items that were coded were 9 NPs marked for masculine gender and 8 NPs marked for feminine gender. The ambiguous category

included the items that were unfinished, ambiguous, were missing the adjective or the noun, or were not produced at all. In addition to this, if the entire NP was changed from masculine into feminine (or vice versa) but had the target-like agreement, we counted such cases as correct feminine or masculine gender agreement respectively. For example, if the task asked to translate the NP such as ‘nice animal’, which would require a masculine gender agreement in Lithuanian, but the learner translated the phrase as ‘nice fox’, which requires feminine gender agreement, the phrase was counted as target-like if it had target-like feminine agreement or non-target like if the agreement between adjective and noun did not match. In addition to this, we counted as correct agreement if the noun and adjective agreed even though the noun was marked for the incorrect gender. Consider a representative example given in (6).

- (6) graž-i park-a
 Beautiful-fem park-fem
 ‘A beautiful park’

As we see from (6), ‘park’ is marked for feminine case in this learner’s L2 grammar, even though in Lithuanian ‘park’ must be masculine (*park-as*). However, since the adjective is marked for feminine gender too, there is feminine gender agreement between the noun and the adjective. Such cases were considered to have correct gender agreement.

5. Results

The results of the study are presented in tables given below. Each table shows the findings obtained from each group of subjects on their use of Lithuanian gender agreement. Consider the results presented in Tables 3–9.³

Looking at tables 3–9, we see that the category with the overall largest percentage was the ‘ambiguous’ category. All groups, except Latvian and Ukrainian native speakers (Tables 5 and 6, respectively), had the most responses that fell into the ‘ambiguous’ category. Looking within groups, the highest rates of ambiguous cases were produced by the two English speaking groups (41% and 46%). The next highest rate came from Russian-Belorussian speaking group (37%), Russian (34.3%), and Ukrainian-Russian (34%) groups. As already mentioned, the lowest rates were among Latvians and Ukrainians (24% and 20%, respectively). Even though the percentage rates for each group differ in this category, these differences were not found to be statistically significant, $F(6, 52) = 1.303, p < .272$. However, statistically significant differences were observed when English and non-native English speaking groups⁴ were compared, $F(1, 57) = 4.779, p < .033$.

Second, tables 3–9 show that correct feminine agreement (CFA) was produced in larger numbers than correct masculine agreement (CMA) even though the task contained more masculine noun phrases (i.e., there were 9 masculine and 8 feminine cases). This difference was seen in all groups except for the Latvians and Ukrainians.

³ The percentages given in the line called *Total* were calculated as follows: since there were 17 possible responses for each participant and the number of participants in each group varied, 17 possible responses were multiplied by the total number of subjects in a given group. For example, group 1 (Russian native speakers) had 17 subjects therefore 17 possible responses were multiplied by 17 subjects which yielded 289. Next, the number of responses for each category (i.e., CMA, CFA, IMA, IFA, and Ambiguous) was obtained and divided from 289 yielding a number listed in the *Total* line for each category.

⁴ Non-native English speaking groups refer to Groups 1 through 5.

Table 3. Results from Russian speakers of Lithuanian (n=17)

Subject number	CMA	CFA	IMA	IFA	Ambiguous
4	8	7	0	1	1
6	6	9	0	0	2
7	4	6	0	1	6
30	7	8	0	0	2
35	3	5	0	1	8
40	1	4	0	1	11
43	7	8	0	1	1
45	4	3	0	0	10
47	5	5	2	0	5
49	8	7	0	0	2
50	2	1	0	5	9
53	1	3	0	0	13
56	2	3	0	0	12
60	8	3	0	3	3
61	5	5	0	0	7
62	8	5	0	1	3
64	5	7	0	1	4
Total	84 (29%)	89 (31%)	2 (.7%)	15 (5%)	99 (34.3%)

Table 4. Results from Ukrainian/Russian speakers of Lithuanian (n=10)

Subject number	CMA	CFA	IMA	IFA	Ambiguous
1	9	7	0	1	0
3	5	3	0	3	6
5	0	7	1	0	9
28	6	8	0	0	3
33	6	7	1	0	3
34	1	6	1	1	8
42	0	3	2	0	12
44	4	4	1	1	7
57	6	6	0	0	5
65	7	4	0	1	5
Total	44 (26%)	55 (32%)	6 (4%)	7 (4%)	58 (34%)

Table 5. Results from Latvian speakers of Lithuanian (n=5)

Subject number	CMA	CFA	IMA	IFA	Ambiguous
36	8	8	0	0	1
41	6	5	0	0	6
55	6	4	0	1	6
63	7	7	0	0	3
66	4	7	1	1	4
Total	31 (36.%)	31 (36.5%)	1 (1%)	2 (2%)	20 (24%)

Table 6. Results from Ukrainian speakers of Lithuanian (n=4)

Subject number	CMA	CFA	IMA	IFA	Ambiguous
29	9	7	0	0	1
31	8	7	0	0	2
32	2	5	2	0	8
48	7	6	0	1	3
Total	26 (38%)	25 (37%)	2 (3%)	1 (2%)	14 (20%)

Table 7. Results from Russian-Belorussian speakers of Lithuanian (n=6)

Subject number	CMA	CFA	IMA	IFA	Ambiguous
2	1	4	0	1	11
37	8	7	0	0	2
38	3	6	0	1	7
51	4	7	0	1	5
52	3	4	0	2	8
58	4	7	0	1	5
Total	23 (23%)	35 (34%)	0 (0%)	6 (6%)	38 (37%)

Table 8. Results from English speakers of Lithuanian (n=5)

Subject number	CMA	CFA	IMA	IFA	Ambiguous
11	2	8	0	0	7
20	2	6	1	1	7
13	1	0	2	3	11
21	2	9	0	0	6
18	5	5	1	2	4
Total	12 (14%)	28 (33%)	4 (5%)	6 (7%)	35 (41%)

Table 9. Results from English speakers of Lithuanian (who also know a language with gender agreement) (n=12)

Subject number	CMA	CFA	IMA	IFA	Ambiguous
10	4	5	0	1	7
12	6	5	0	1	5
14	2	5	1	1	8
15	1	2	1	0	13
16	4	7	0	1	5
17	7	5	1	1	3
19	4	8	0	1	4
22	5	2	0	0	10
23	1	0	2	5	9
24	3	5	0	0	9
25	0	4	0	4	9
8	1	4	1	0	11
Total	38 (17%)	52 (25%)	6 (3%)	15 (7%)	93 (46%)

In the case of Latvian speakers of Lithuanian, the numbers for CFM and CMA were the same (36.5%), whereas in the case of Ukrainians, 38% went to CMA and 37% to CFA. The lowest numbers for CMA came from the two English-speaking groups (14% and 17%, respectively). In the case of CFA, the lowest number for this category came from the English-speaking group that had exposure to another language with gender agreement (25%). The first English-speaking group (without any exposure to a language with gender system) slightly outperformed the Russian and Ukrainian-Russian speaking groups on CFA (33%, 31%, and 32%, respectively).

Statistical significance between the groups was reached on CMA category,

$F(6, 52) = 2.173, p < .060$. When the differences between English and non-native English speaking learners of Lithuanian were statistically tested on CMA, the results were statistically significant as well, $F(1, 57) = 8.091, p < .006$. The results of group 7 on CMA when tested against all other groups, excluding Group 6, were found to be statistically significant, $F(1, 52) = 4.664, p < .036$. The results of group 6 when compared against the results of all other groups, were not found to be statistically significant, $F(1, 57) = 3.253, p < .077$. But when Group 6 was compared with all non-native English speaking groups (excluding Group 7), the results were found to be statistically significant, $F(1, 45) = 4.531, p < .039$.

As for the CFA category, no statistically significant differences were found among the groups, $F(6, 52) = .787, p < .584$. Also, when English speaking and non-native English speaking groups of learners were compared on CFA category, the results were not statistically significant, $F(1, 57) = 2.203, p < .143$. However, results were statistically significant when group 7 was compared against all other non-native English speaking groups, $F(1, 52) = 4.009, p < .050$. The same, however, was not observed for Group 6 when its production on CFA category was compared to all other groups, $F(1, 57) = .083, p < .775$.

In the case of IFA and IMA categories, there were more IFA cases than IMA. That is, masculine gender agreement was over generalized more often than feminine gender. However, in some groups the differences between IFA and IMA were minor. For instance, Ukrainian-Russian, Latvian and Ukrainian speaking subjects performed either exactly the same on both categories or with 1% difference between these two categories. The highest rate of IFA was 7% and it came from the two English-speaking groups; the highest rate of IMA was 5% and it came from the English-speaking group that did not have exposure to a language with gender agreement. However, the differences seen in percentages were not found to be statistically significant, $F(6, 52) = 1.649, p < .152$ (for IMA) and $F(6, 52) = .614, p < .718$ (for IFA) among all groups. As for English and non-native English speaking groups, only for IMA category the results were close to being statistically significant, $F(1, 57) = 3.303, p < .074$. No statistically significant differences were found between Group 6 and all other groups, $F(1, 57) = 2.735, p < .104$ (for IMA) and $F(1, 57) = .397, p < .531$ (for IFA). The same was true for Group 7 when it was compared to all non-native English speaking groups, $F(1, 52) = 1.439, p < .236$ (for IMA) and $F(1, 52) = 1.807, p < .185$ (for IFA).

In the case of the control group, there was no gender agreement errors found in their responses. Table 10 presents the results from the control group of the study.⁵

⁵ The responses obtained from the controls were coded as follows: CA (correct agreement), NN (null nominal) and Ambiguous. The null nominal is a grammatical option often used by native speakers of Lithuanian. In the null nominal construction, the noun is omitted from the noun phrase and the adjective or another modifier carries the necessary morphology for noun identification. In addition to this, the surrounding context provides the necessary information for the identification of the noun.

Table 10. Results from Lithuanian control group (n=10)

Subject	CA	NN	Ambiguous
C1	9	5	3
C2	17	0	0
C3	10	5	2
C4	17	0	0
C5	17	0	0
C6	15	0	2
C7	13	4	0
C8	17	0	0
C9	14	3	0
C10	12	4	1
Total	141 (83%)	21 (12%)	8 (5%)

As Table 10 shows, the majority of answers given by the controls, namely 95%, were target-like. 5% of ambiguous answers were not ungrammatical but instead lacked an adjective or had some other quantifier in place of an adjective.

Next, consider table 11 which presents information regarding the instances of faulty gender agreement found in the L2 grammars of Lithuanian learners. The groups are listed horizontally and the test cases (a total of 17) are listed vertically. Cases 1 through 5 involved nouns with grammatical gender, whereas cases 6 through 17 involved nouns with natural gender.

Table 11. Faulty agreement

Case number	G1	G2	G3	G4	G5	G6	G7	Total
1 (masc) little market 'mažas turgus'	0	1	0	0	0	0	1	2
2 (masc) beautiful park 'gražus parkas'	1	1	1	1	0	2	1	7
3 (masc) big university 'didelis universitetas'	1	3	0	1	0	0	0	5
4 (fem) big café 'didelė kavinė'	2	1	0	0	1	2	3	9
5 (fem) old shop 'sena parduotuvė'	7	4	1	1	3	2	6	24
6 (masc) quiet painter 'ramus dailininkas'	0	0	0	0	0	0	1	1
7 (fem) young woman 'jauna moteris'	1	0	0	0	0	1	2	4
8 (fem) beautiful woman 'graži moteris'	1	0	0	0	0	1	1	3
9 (fem) tall woman 'aukšta moteris'	0	0	1	0	0	0	1	2
10 (masc) short man 'žemas vyras'	0	0	0	0	0	0	0	0
11 (masc) good cook 'geras virėjas'	0	0	0	0	0	0	1	1
12 (masc) fat waiter 'storas padavėjas'	0	2	0	0	0	0	1	3
13 (fem) sweet waitress 'maloni padavėja'	2	0	0	0	0	0	0	2
14 (masc) sad waiter 'liūdnas padavėjas'	0	0	0	0	0	1	0	1
15 (masc) interesting man 'įdomus vyras'	0	0	0	0	0	0	0	0
16 (fem) young woman 'jauna moteris'	1	1	0	0	0	0	2	4
17 (fem) little daughter 'maža duktė'	1	0	0	0	2	1	1	5

As expected, table 11 shows that there were many more cases of faulty agreement among NPs with grammatical gender than among noun phrases with natural gender. Specifically, in the case of NPs with grammatical gender, there were 47 faulty agreements out of 295 possible (16%), whereas in the case of NPs with natural gender, there were only 26 faulty agreements out of 708 possible (4%) across all groups. ANOVA test did not show any statistically significant differences among all groups for these categories, $F(6, 52) = .679, p < .668$ (for the NPs with natural gender) and $F(6, 52) = .665, p < .678$ (for the NPs with grammatical gender).

Finally, the most common noun phrases marked with faulty gender marking were #5 (24 times) and #4 (9 times). These items are marked for feminine gender in Lithuanian. In order to account for these errors, table 12 provides information about the differences in gender morphology among the languages represented in the study sample for these two items, as well as for the other three noun phrases with grammatical gender.

Table 12. Grammatical gender markings

Item	Russian	Belo-russian	Latvian	Ukrainian	Lithuanian
'little market'	masc	masc	masc	masc	masc
'beautiful park'	masc	masc	masc	masc	masc
'big university'	masc	masc	fem	masc	masc
'big café'	masc	neut	fem	neut	fem
'old shop'	masc	masc	masc	fem	fem

As we see from table 12, the two noun phrases 'little market' and 'beautiful park' are marked with masculine gender in all languages, including Lithuanian, whereas the other three exhibit variation. One possible reason for why item #5 ('old shop') had the most errors in gender agreement is that in Lithuanian it is marked with feminine gender but in all other languages, except Ukrainian, 'shop' is masculine⁶.

The same explanation, namely L1 transfer, could account for the phrase #4 'big café' which is marked for feminine gender in Lithuanian and Latvian but masculine in Russian and neuter in Ukrainian and Belorussian. However, L1 transfer cannot account for the errors made by the English speaking subjects of this study since English does not have grammatical gender marking on nouns. An alternative explanation for the gender agreement errors found in the NPs #5 'old shop' and #4 'big café' is the L2 learner's tendency to overuse masculine gender. Similar findings have already been reported by White et al. (2004) and Francheschina (2001).

Finally, 6 cases were found where gender agreement was correct between the nominal and the adjective but the gender of the nominal was incorrectly supplied. In all these cases, feminine gender was overused. Consider representative examples provided in (7).

- (7a) *didel-ė universitet-ė*
 Big-fem university-fem
 'A big university'

⁶ A reviewer pointed out that Latvian and Lithuanian morphological systems are very similar and that it would be interesting to find out the errors that Latvians made in the case of 'old shop' which is marked with masculine gender in Latvian. In our study there were only 5 Latvians and only one made an error in gender agreement within the NP, 'old shop'. Specifically, the 'shop' was used with the correct feminine gender marking but the adjective 'old' was marked for masculine gender (i.e., *sen-as* (masc) *parduotuv-ė* (fem)).

- (7b) didel-ė turg-a
Big-fem market-fem
'A big market'
- (7c) graž-i park-a
Beautiful-fem park-fem
'A beautiful park'

Examples in (7) show the cases in which masculine gender is replaced by feminine gender but the agreement is correctly supplied in all these cases. These six cases came from Latvian (one case), Russian (one case) and English (four cases) native speaking learners of Lithuanian.

6. Discussion and conclusion

The research question of this study asked whether learners from L1s with gender agreement would outperform English-speaking learners of Lithuanian on gender agreement. The results of the study found statistically significant differences between English and non-native English speaking groups on Ambiguous and on CMA categories. In addition to this, the results from Group 7 were statistically significant on CFA category. Group 6, however, did not differ from other groups on CFA category. There were no statistically significant differences observed among groups on IMA and IFA categories.

Based on these findings, support for the FTFA hypothesis is obtained. Since learners from L1s with gender agreement performed better on some categories than English-speaking learners of Lithuanian, these differences could be attributed to L1 transfer. Recall that FTFA hypothesis argues that in the initial stages L1 transfer will be the basis for the development of L2 grammar. However, as the study shows, L1 transfer is not the only factor shaping the acquisition of gender agreement in the grammars of these learners. Since Group 6 did not differ from other groups on CFA category and both English-speaking groups did not differ from non-native English speaking groups on IMA and IFA categories, this strongly argues for the UG involvement in the development of L2 grammars of these learners.

If the predictions of the FFFH were borne out, the English-speaking group without knowledge of a language with gender agreement should have performed much worse than all other groups; however, on some of the categories, this group did not perform differently from non-native English speaking groups.

If the results of this study were to support the Full Access view, there should have been no differences among the groups since according to the Full Access view, UG directs the acquisition of L2 grammars without L1 interference. However, some differences were observed among the groups which suggest that L1 is playing a role in the initial stages of L2 grammars of these learners, as predicted by the FTFA hypothesis.⁷

⁷ We agree with a reviewer who pointed out that the differences between the groups tested in the current study were small and that many other factors "frequency of the input, word knowledge, morphophonological considerations of individual words" (directly taken from the reviewer's form) could have had an effect on the results of the study. We acknowledge the importance of these factors but leave the investigation of their effects on the acquisition of Lithuanian gender agreement for future research.

Abbreviations

acc – accusative
dat – dative
fem – feminine
masc – masculine
neut – neuter
nom – nominative
past – past
pl – plural
pres – present
sg – singular

References

- Epstein, Samuel David; Flynn, Suzanne; Martohardjono, Gita 1996. *Second Language Acquisition: Theoretical and experimental issues in contemporary research*. – *Brain and Behavioral Sciences* 19, 677–758.
- Epstein, Samuel David; Flynn, Suzanne; Martohardjono, Gita 1998. The strong continuity hypothesis: some evidence concerning functional categories in adult L2 acquisition. – Suzanne Flynn, Gita Martohardjono, Wayne O’Neil (eds). *The generative study of second language acquisition*. Mahwah, NJ: Lawrence Erlbaum, 61–77.
- Flynn, Suzanne; Martohardjono, Gita 1994. Mapping from the initial state to the final state: the separation of universal principles and language-specific principles. – Barbara Lust, Margarita Suner, John Whitman (eds). *Syntactic theory and first language acquisition: cross-linguistic perspectives*. Vol. 1: Heads, projections and learnability. Hillsdale, NJ: Lawrence Erlbaum, 319–335.
- Francheschina, Florencia 2001. Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. – *Second Language Research* 17, 213–247.
- Hawkins, Roger; Chan, Cecilia Yuet-hung 1997. The partial availability of Universal Grammar in Second Language Acquisition: The “failed functional features hypothesis.” – *Second Language Research* 13, 187–226.
- Sabourin, Laura; Stowe, Laurie A.; de Haan, Ger J. 2006. Transfer effects in learning a second language grammatical gender system. – *Second Language Research* 22, 1, 1–29.
- Schwartz, Bonnie D.; Sprouse, Rex A. 1994. Word order and nominative case in nonnative language acquisition. – Teun Hoekstra, Bonnie D. Schwartz (eds). *Language Acquisition studies in generative grammar*. Amsterdam: Benjamins, 317–368.
- Schwartz, Bonnie D.; Sprouse, Rex A. 1996. L2 cognitive states and the full transfer/full access model. – *Second Language Research* 12, 40–72.
- White, Lydia 2003. *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.
- White, Lydia; Valenzuela, Elena; Kozłowska-MacGregor, Martyna; Leung, Yan-Kit Ingrid 2004. Gender and number agreement in nonnative Spanish. – *Applied Psycholinguistics* 25, 105–133.

Simona Mačiukaitė (LCC International University, Lithuania) specializes in the fields of bilingualism/multilingualism, first and second language acquisition, research methodology in applied linguistics and generative grammar theory. Primary research areas include first language acquisition, morpho-syntactic aspects of second language acquisition and grammatical aspects of bilingual/multilingual code-switching.
smaciukaite@lcc.lt

ERINEVUSI TÄISKASVANUD ÕPPIJATE LEEDU KEELE GRAMMATIKAS, KES ÕPIVAD LEEDU KEELT TEISE KEELENA – SOOÜHILDUMINE

Simona Mačiukaitė

LCC Rahvusvaheline Ülikool, Leedu

Sooühildumine nimisõnafraasis ei ole omane kõigile keeltele. Kui teise keelena õpitavas keeles vastav nähtus eksisteerib, oleneb selle omandamise edukus sellest, kas nähtus on neile emakeelest tuttav või mitte. Näiteks on väidetud (Hawkins, Chan 1997), et täiskasvanud õppijad, kelle emakeeles sooühildumine puudub, ei suuda perfektselt ära õppida keelt, kus see esineb.

Käesolevas artiklis uuritakse, kuidas vene-, ukraina-, läti-, valgevene- ja ingliskeelsed õppijad omandavad leedu keele sooühilduvust. Katsealustel paluti teha kirjalik tõlge. Osalesid 59 muukeelset leedu keele õppijat ning 10 emakeelset leedulast.

Mõned statistiliselt olulised erinevused ilmned kahe inglise keelt oskava rühma vahel, kellest ühele oli inglise keel emakeeleks, teisele mitte. Täpsemalt oli tegemist ühilduvuserinevusega leedu meessugu sõnade tarvituses. Naissugu sõnade puhul jooksis veelahet teisiti: erinevus oli märgatav inglise keelt oskavate rühmade vahel, kellest üks oskas veel üht sooühildivat keelt, teine aga koosnes mitteinglastest.

Samas polnud naissugu sõnade puhul vahet mitteinglaste ja nende inglise keelt oskajate vahel, kes lisaks mõnda sooühildivat keelt ei osanud.

Saadud tulemuste põhjal loobutakse Hawkinsi ja Chani hüpoteesist Schwartzi ja Sprouse'i (1994, 1996) kasuks.

Märksõnad: teise keele omandamine, täiskasvanud õppija, universaalne grammatika, lähtegrammatika, leedu keel

EESTI VERBI MINEVIKUVORMIDEST VENEKEELSETE ÜLIÕPILASTE KASUTUSES

Diana Maisla

Ülevaade. Artiklis analüüsitakse Tartu Ülikooli Narva Kolledži ees-
tikeelsetel õppekavadel õppivate vene kodukeelega üliõpilaste oskust
kasutada eesti keele mineviku ajavorme. Uuritud on suulisi ja kirjalikke
keelenäiteid ning tehtud kokkuvõtte informantide metalingvistilistest
teadmistest mineviku kohta. Vaatluse all on ka, kuidas eesti keele kui
teise keele õppematerjalides on pakutud minevikuagade olemuse ja
kasutamisevaldkondade õppijatele tuge ja selgitusi. Lõpuks esitatakse
minevikuagade kujutavad skeemid, mida senini eesti keele kui teise
keele õpetuses kasutatud ei ole, kuid mis abistaks seletada ajavormide
(täendus)erinevusi.

Võtmesõnad: eesti keel teise keelena, teise keele õppimine ja oman-
damine, verbi ajavormid, eesti keel

1. Sissejuhatus

Mineviku ajavormide omandamise ja kasutamise seaduspärasusi on pikemat aega
uuritud germaani ja romaani keeltes. Keeleõppijad kasutavad algul verbi olevikus
ja täpsustavad aega määruste abil. Sündmustest jutustatakse kronoloogilises jär-
jekorras. Seejärel omandatakse morfoloogilised ajavormid. Õppijakeele ja esimese
keele vahe seisneb selles, et emakeele kõneleja oskab kasutada ka teisi vahendeid
ning võib sündmuste tegelikust järjekorrast kõrvale kalduda (Bardovi-Harlig
2000: 65). Näiteks lausetega *Istusime autosse. Enne seda olime kõigiga hüvasti
jätnud* on sündmusi kujutatud nende toimumisele vastupidises järjekorras. Sellist
võtet oskab kasutada enneminevikku valdav kõneleja, kuid mitte algaja keeleõppija.
Õppijate keelekasutusse ilmuvad liht-, täis- ja ennemineviku ning kestvate aegade
morfoloogilised väljendusvahendid teatud järjekorras. Tavaliselt on mineviku
omandamise suund lihtminevikult liitaegadele, kuid keeliti ja indiviiditi võib see
järjestus varieeruda, sest võõra keele grammatika omandamine sõltub ka sellega
kokkupuutumise või õpetamise järjekorrast (vt nt Bardovi-Harlig 2000).

Euroopa kultuuriruumis kujutatakse aega kolmemõõtmelises süsteemis koos ruumi ja liikumisega. Aeg voolab ühtlaselt, (inimese tajus küll vahel aeglaselt ja vahel kiiresti) ühes suunas – minevikust tulevikku. Aega on võimalik jaotada lõikudeks või perioodideks. Iga sündmuse, tegevuse või oleku saab asetada teatud ajapunkti või -vahemikku. Selline lähenemisviis võimaldab määratleda ajalisi suhteid kahe sündmuse vahel või sündmuse ja teatud ajahetke vahel (Evans 2005, Klein 1986). Keeles väljendab ajalisi suhteid grammatiline aeg abisõnade, määruste ja morfoloogia abil. Eraldi huvi pakub täisminevik, kuna selle ajavormi olemus ja kasutus võivad eri keeltes varieeruda või üldse puududa. Kuigi tegemist on minevikku väljendava ajavormiga, ei saa sellega samas lauses kasutada minevikku täpselt määratlevat määrust. Näiteks lause *Eile hommikul kell 10 olen reisilt tagasi jõudnud* ei ole eesti keeles õige, sest täpse aja määratluse *eile hommikul kell 10* tõttu peaks see olema lihtminevikus.¹ Mineviku ajavormide ja määruste omavaheline seos jääb selle artikli uurimisalast välja.

Eesti keele minevikuaegade õpetamise muudab keeruliseks asjaolu, et alati ei olegi kerge seletada, millisel juhul peaks valima lihtmineviku ja millal täismineviku. Mõlemad ajavormid võivad olla mõne tegevuse või olukorra kohta õiged. Näiteks ei ole ka emakeele kõneleja jaoks vahet lausetel:

Ma elasin kunagi selles majas.

Ma olen selles majas kunagi elanud.

Eestis algajate keeleõppijate minevikuvormide omandamisetappe uuritud ei ole, kuid edasijõudnute minevikukasutusele on viiteid Raili Pooli, Elle Vaimanni (2005) ja Anna Verschiku (2002, 2004) artiklites. A. Verschik on uurinud Narva Kolledži üliõpilaste keeleoskust ja leidnud, et mineviku kasutuses vigu ei esine. Silmas on peetud, et üliõpilased ei eksi olevikku ja minevikku eristades. R. Pool ja E. Vaimann kirjutavad eesti keelt kõrgtasemel valdavate üliõpilaste vigu analüüsid, et

.. on põhjust arvata, et õppijad on eesti keeles tehtud valikutel lähtunud emakeelest. Suurim rühm on siin ajavormivead, mille puhul vene keele ainus minevikuvorm on käesoleva uurimuse informantide vahekeeles üldistunud lihtminevikuna, jättes kõrvale täis- ja ennemineviku. (Pool, Vaimann 2005: 134.)

Teise keele õppijatel on raske omandada mõisteid, mida nende emakeeles ei ole. Kui eesti keeles eristatakse liht-, täis- ja enneminevikku, siis vene keeles on ajaloolise arengu tulemusena alles jäänud ainult üks mineviku ajavorm – lihtminevik. Osaliselt diferentseeritakse vene keele minevikku aspektikategooria abil, mis näitab, kas tegevus on ühe- või mitmekordne, lõpetatud või lõpetamata, pika- või lühiajaline (Külmoja jt 2003, Vaiss 2004). Kuidas keeleõppija mineviku ajavorme kasutab ja kuidas seletab eri vormide vahel valiku tegemise põhimõtteid, väljendab nende implitsiitset ja eksplitsiitset teadmisi keele kohta. Implitsiitne õppimine tähendab alateadlikku keele omandamist ja eksplitsiitne õppimine teadvustatud õppimist, mille käigus võetakse arvesse sõnastatud reegleid ning luuakse ja katsetatakse oma hüpoteese keele struktuuri kohta. Samamoodi on implitsiitsed teadmised keele kohta intuiitiivsed ja nende üle võib otsustada, kogudes keeleainet vabast keelekasutusest. Eksplitsiitsed teadmised ilmnevad kontrolltöodes ja testides, st kontrollitud

¹ Täisminevikust vt nt Macrory, Stone 2000, Pancheva, von Stechow 2003.

olukordades. Samuti väljendub see metalingvistiliste teadmistena, st võimega keelt kirjeldada ja põhjendada teatud vormi valikut (Ellis 2004, Ellis jt 2006).

2. Informandid ja andmekogumismeetod

Keelejuhtideks olid Tartu Ülikooli Narva Kolledži 90 üliõpilast, kes õpivad eesti keelsetel õppekavadel: noorsootöö, kohaliku omavalitsuse korralduse, koolieelse lasteasutuse ja humanitaarainete (eesti keel, inglise keel) õpetaja erialal. Pool informantidest on humanitaarainete õpetaja magistriõppekaval õppijad, kes on lisaks praktilisele keeleõppele läbinud teoreetilise eesti keele bakalaureuseõppe. Paljud neist juba töötavad valitud erialal. Enamikku üliõpilasi iseloomustab see, et nad on eesti keele selgeks õppinud ilma loomuliku keelekeskkonna toeta. Üliõpilased on pärit Ida-Virumaalt, suurem osa Narva linnast. Keelekeskkonnal, kui selle all mõelda elu- ja töökohas valdavalt kasutatavat keelt, ei pruugi olla olulist mõju keeleoskusele, sest iga õppija võib keelekeskkonna tekitada ka ise, valides sihtkeeles suhtlevaid sõpru, kasutades sihtkeelset massimeediat või osaledes sihtkeelses kultuurielus. Siiski on Ülle Rannut põhjalikus uurimuses tõestanud, et Ida-Virumaal segregatiivses keskkonnas on eesti keele oskus madalam, sest eestlastega suhtlemisvõimalused on väga piiratud (Rannut 2005). Narva Kolledži üliõpilased kasutavad eesti keelt õppetöös ning pedagoogilises tegevuses, näiteks koolipraktikal käies, kuid mitte igapäevase suhtluskeelena.

Informandid on sooritanud eesti keele kõrgtaseme eksami, mis põhineb Euroopa Nõukogu üldkeeleoskuse mudeli edasijõudnute taseme (ingl *vantage level*) keelefunktsioonide kirjeldusel. Kõrgtaseme sõnavara ja tarindite oskuse kohta on märgitud, et isik valdab praktiliselt kogu keele struktuuri ja sõnavara. Mõnikord võib esineda ebatäpsusi keeruliste ja harvade konstruktsioonide kasutamisel, üksikud vead on raskesti märgatavad. Kirjelduses märgitakse ka, et kõrgtaseme saavutanu kasutab harvu struktuure, mis pole tema emakeele väljendusvahenditega vastavuses (Loog, Kerge 1999: 17, 102). Ajakategooria väljendamine morfoloogiliste vahenditega kuulub just sellesse valdkonda ja on huvitav uurida, kui latusalt informandid seda paradigmat valdavad.

Keelenäited, st laused, milles esines või oleks pidanud esinema liitmineviku vorm, on kogutud korpuseks üliõpilaste suulistest ja kirjalikest ettekannetest, (praktika)tundides peetud vestlustest, esseedest, praktikapäevikutest ja muudest üliõpilastöödest, kuid mitte eesti keele grammatikatundides tehtud ülesannetest. Seega puuduvad näited harjutustest, mille sooritamisel on oluline keskenduda õige vormi valikule. Samuti on välditud referaate, milles on raske eristada iseseisvat ja refereeritud teksti. Kõrvale on jäetud ka teiste õppejõudude juhendamisel kirjutatud tööd, et välistada juhendaja korrigeeritud keelendeid. Suulistest näidetest osa pärineb lindistatud tekstidest, osa on üles märgitud üliõpilaste esinemistest ja vestlustest.

2005/06. õppeaastal on lindistatud neli tundi 27 üliõpilase jutustusi teemadel *kes ma olen, miks tulin õppima Narva Kolledžisse, mida tean valitud erialast, millised on mu tulevikuplaanid*. Enne intervjuu algust selgitati, et oodatakse võimalikult pikka jutustust eesti keeles, mis annaks informatsiooni igaihe keeleoskuse kohta ning aitaks kõiki üliõpilasi tundma õppida. Teemad olid kirjutatud tahvile ja üliõpilase kõneste sekkuti võimalikult vähe, peamiselt ainult julgustavate märkustega

sel juhul, kui kõneleja mõtted kippusid takerduma. Eesmärk oli koguda spontaanse keele näiteid suunates üliõpilasi rääkima olevikust, minevikust ja tulevikust. Eeldati, et jutustuses tuleb ette konstruktsioone, mis võimaldavad demonstreerida liitaegade kasutamise oskust, nt *Enne kui tulin Narva, olin õppinud Mainori kõrgkoolis*. Kuna liitaegade esinemissagedus on eesti keeles suhteliselt väike (vt Metslang 1997) ning intervjuueeritavad püüdsid end väljendada võimalikult lakooniliselt, kuid korrektselt, siis osutus andmete kogumine lindistamise abil ebaotstarbekaks.

2006/07. õppeaastal on vaatlusmeetodiga kogutud üliõpilaste suulist kõnet praktikatundidest (vaadeldud 30 akadeemilist tundi), loengutest-seminaridest, kus üliõpilased on vabalt vestelnud või esinenud pikemate sõnavõttudega, ja lõputööde kaitsmistelt. Üles on kirjutatud laused, kus liitmeneviku ajavorm esines või oleks pidanud esinema.

Kirjalik materjal on kogutud seotud tekstist, mitte üksiklausetest, seega on kontekstil oluline roll otsustamisel, kas kasutatud ajavormi pidada korrektseks või mitte. Suur osa materjalist on analüüsivatest töödest, kus on ette antud ülesanne ja/või loetelu, millest kirjutada. Seega on kirjutatud lausete sisu, kuid ilmselt osaliselt ka esinevad ajavormid sellest mõjutatud.

3. Eesti keele mineviku ajavormid

Keeleõppes esmalt defineeritakse grammatikakategooria, siis seletatakse, kuidas vorme moodustada, ning seejärel antakse kasutusjuhiseid. Käesoleva uurimuse informantidele vormimoodustus raskusi ei valmista, mistõttu on vaatluse all eri minevikuvormide tähendus ja kasutus esmalt grammatikakirjeldustes ja seejärel informantide keelekasutuses.

3.1. Mineviku ajavormid grammatikakirjanduses

Eesti keele akadeemilise grammatika järgi märgib imperfekt e lihtminevik definiitset minevikku, perfekt e täisminevik indefiniitset minevikku olevikulise vaatlushetke seisukohalt ja pluskvamperfekt e enneminevik indefiniitset minevikku minevikulise vaatlushetke seisukohalt (EKG I: 75–79, EKG II: 33).

“Eesti keele käsiraamatus” seletatakse täismineviku ja lihtmineviku vahetõrget järgmiselt:

Lihtminevik on lihtne, s.o üheplaaniline, nii vormilt kui ka sisult. Kõneleja kujutus kandub täielikult minevikku – ta ei seosta seda kuidagi olevikuga. --- Täisminevik seevastu on nii vormilt kui ka sisult kaheplaaniline. Temas on ühendatud oleviku seisund ja mineviku tegevus. --- Täisminevikku kasutatakse siis, kui eelnevast tegevusest tehakse kokkuvõtte käesoleva hetke seisukohast. (Erelt jt 2000: 346.)

Käsiraamatus “Keelehärm” (Metslang jt 2003) lisatakse eelnevale, et lihtminevik on jutustamise aeg, täisminevik minevikus toimunud sündmuste kirjeldamise või kõnealusel sündmusest kokkuvõtte tegemise aeg.

Ennemineviku kasutussoovitus “Eesti keele käsiraamatus” on järgmine:

Enneminevik teeb kokkuvõtte mingile minevikuhetkele eelnevast tegevusest, --- Kui pealauses on lihtminevik, kõrvallause aga väljendab pealause tegevusele eelnevat tegevust, tuleks kõrvallauses kasutada enneminevikku, mitte lihtminevikku. (Erelt jt 2000: 346.)

Enneminevikku kasutatakse eesti keeles suhteliselt harva võrreldes täismineviku ja eriti lihtminevikuga (vt Metslang 1997).

Minevikuaegadest rääkides on olulised kõnehetk, sündmuse toimumise hetk ja viitehetk. Enneminevik esineb eelkõige liitlausetes, millest üks osalause on lihtminevikus ja teine osalause viitab ajale enne seda, nt *Jõudsin teha kõike, mida olin enne planeerinud.*

Lihtmineviku ja täismineviku vahel valimine on keerulisem. Viidates Östen Dahlile ja Bernard Comriele, on Helle Metslang (1997: 102–105) eristanud eesti keele täisminevikul neli funktsiooni:

resultaadiperfekt – kirjeldatakse olukorda olevikus, mis on minevikus toimunud tegevuse resultaat (*olen lõpetanud gümnaasiumi*);

kogemusperfekt – räägitakse oleviku vaatepunktist minevikus toimunud või mitte toimunud tegevusest (*olen käinud / ei ole käinud Pariisis*);

püsiva/kestva olukorra perfekt – tegevus algas minevikus ja jätkub olevikus kõnemoendil (*olen Narvas elanud juba kolm aastat*);

kuuma uudise perfekt – lähiminevikus toimunud sündmuse erakordsuse rõhutamiseks (*Kanter on tulnud maailmameistriks*).

Kuna kuuma uudise perfekti kasutuskontekst on kitsas, siis ei saa eeldada, et täisminevik sellises funktsioonis õppijakeeles esineb.

Teise keele omandamisel õpitakse grammatikavorme seostama tähenduse ja kasutusega. Samuti tuleb õppida vahet tegema omavahel lähedaste, kuid semantiliselt erinevate vormide vahel. Nii liht-, täis- kui enneminevik väljendavad tegevust või olekut, mis on möödunud, kuid kolme vormi üksteisest eristamiseks on vaja mõista nende suhet kõnehetkega ning ajaga, millest räägitakse.

3.2. Mineviku liitajavormid informantide kasutuses

Kogutud keelenäited esindavad suhteliselt kitsast konteksti, sest Narva Kolledži üliõpilaste keel ja teemade ring, millest eesti keeles räägitakse või kirjutatakse, on piiratud peamiselt õppetöö vajadustega. Siiski esindavad koolipraktika tundides kogutud suulised näited erinevaid teemavaldkondi, mida eesti keele ainekava järgi (eriti gümnaasiumiklassides) käsitletakse – loodus, kultuur, majandus ja muu ühiskonnaelu.

Kokku on korpuses 129 eri verbi, neist kõige sagedamini esinevad (sulgudes kordade arv): *lõpetama* (16), *kasutama* (14), *tegema* (10), *õppima* (10), *olema* (9), *esitama* (8), *kirjutama* (8), *planeerima* (8), *koostama* (7), *saavutama* (7). Kõik need verbid esinevad üliõpilaste praktikapäevikutes ja kuuluvad tundide planeerimist ning läbiviimist kirjeldava sõnavara hulka. Verbi *lõpetama* ja *õppima* on kasutatud ka oma eluloost rääkides, nt *Olen lõpetanud X kooli. Olen õppinud X koolis.*

Minevikuvormide kasutuse uurimiseks on vastandatud liht- ja liitajavorme, st loendatud, kui paljudes lausetes liitmenevikku õigesti kasutatakse ja kui tihti see on asendatud lihtmenevikuga. Näitelauseid ja -fraase on kokku 287, millest

- täismenevikus on 84 suulisest kõnest ja 102 kirjalikust tekstist;
- ennemenevikus on 13 suulisest kõnest, 88 kirjalikust tekstist.

Eesmärk on teada saada, millised seaduspärasused esinevad keelejuhtide mineviku liitajavormide kasutuses. Tähelepanu keskmes ei ole vigade otsimine, kuid analüüsi huvides on keelenäited jagatud õigeteks ja vääradeks. Ajavormi kasutamise õigsuse üle otsustamisel lähtutakse ajavormi definitsioonidest, kasutusseletusest ja isiklikust keelevaistust.

Õigesti ja valesti kasutatud vormide kohta annab ülevaate tabel 1. Kuigi vene keeles puuduvad täis- ja ennemenevik grammatilise kategooriana, kasutatakse konstruktsioone, mis vastavad eesti keele liitmenevike umbisikulisele tegumoele. See asjaolu võib mõjutada umbisikuliste vormide kasutust. Tabelis 2 on andmed, milline on isikuliste ja umbisikuliste verbide arvuline suhe suuliste ja kirjalike ning täis- ja ennemenevikus olevate näidete vahel.

Tabel 1. Suuliste ja kirjalike keelenäidete arv

	Täismenevik				Ennemenevik			
	õiged	%	väärad	%	õiged	%	väärad	%
Suulised näited	44	52,4	40	47,6	7	53,8	6	46,2
Kirjalikud näited	75	73,5	27	26,5	38	43,2	50	56,8

Tabel 2. Keelenäidete arv isikulises ja umbisikulises tegumoes

	Täismenevik		Ennemenevik	
	Suulised / neist õiged	Kirjalikud / neist õiged	Suulised / neist õiged	Kirjalikud / neist õiged
Isikuline tegumood	64 / 26	46 / 22	5 / 3	13 / 7
Umbisikuline tegumood	20 / 18	56 / 53	8 / 4	75 / 46

Täismenevikuga õigete ja valede vormide osakaalu põhjal võib järeldada, et kirjutades on rohkem aega järele mõelda ning on võimalik kasutada ka vältimistehnikat, st kasutatakse lausekonstruktsioone, milles ollakse kindlad. Suulise keelekasutuse puhul on enesekontroll nõrgem, tihti tuleb kiiresti reageerida, sageli ei ole toetavat konteksti, mistõttu on ka valesid vorme rohkem. Täismenevikuga suulised näited on ülekaalukalt isikulises tegumoes. Ennemenevikuga suulisi näiteid on vähe, mis on kooskõlas asjaoluga, et seda ajavormi esinebki keeles suhteliselt harva. Kirjalike näidete suurem osakaal on seletatav üliõpilastööde eripäraga – suur hulk neist nõuab akadeemilise stiili järgimist, muuhulgas umbisikulist tegumoodi. Ennemenevikuga kirjalike näidete vead tulenevad peamiselt umbisiku ülemäärasest kasutamisest.

Järgnevalt on esitatud üksiklausetena esinenud näited, st ei ole räägitud pikemat seotud teksti ega ole olnud mingeid abistavaid vihjeid, mis oleks kõnelejat

aidanud õiget ajavormi valida. Siiski kuuluvad need näited vestluse või suhtlussituatsiooni konteksti, mis toetab ühe või teise minevikuvormi valikut.

- (1) Paneme [artikli] seinale neile, kes veel ei näinud ja ei lugenud. (P.o .. kes veel [eelnevate päevade jooksul] ei ole näinud ega (ole) lugenud.)
- (2) Iga laps lapsepõlves voolis plastiliinist midagi. (P.o Iga laps on lapsepõlves plastiliinist midagi voolinud.)
- (3) Mina ka olin Piusa koopas. (P.o Mina ka olen käinud Piusa koopas.)

Näited 1–3 vastavad sisult kogemusperfektile, kirjeldades ajaperioodi, mitte lühikest hetke. Esimeses näites on kontekst rääkijale ja kuulajale arusaadav. Jutt on artiklist, mille kõik oleksid pidanud juba varem läbi lugema, kuid ei ole lugenud. Teine näide on sissejuhatavat lause voolimise kasulikkust käsitlevas ettekandes. Kolmas näide on praktikatumnist, kus üliõpilane juhtis vestlust nahkhiirte eluviisi teemal.

Praktikatunde andvate üliõpilaste klassitoakõnes kordub küsimus:²

- (4) Kas te lõpetasite? või Kas saite valmis?

Enne uue ülesande andmist tahab õpetaja-praktikant teada, kas õpilastel on eelmine töö tehtud. Eestikeelne õpetaja küsib sellisel juhul: *Kas te olete lõpetanud?* või *Kas te olete ülesande/harjutuse valmis saanud?* Võrdluseks, olukorras, kus õpetaja küsib ühe või kahe õpilase käest: *Kas lõpetasid/te lobisemise?*, on see küsimus korrektne. Kui küsimus esitatakse rühmale, mille liikmed lõpetavad töö eri aegadel, siis on lõpetamine pikem periood ja eestlasele on loomulik kasutada täisminevikku.

Keeruline on liitaja *kas*-küsimuse moodustamine.

- (5) Kas te vaatasite vene televisiooni? (P.o Kas te olete vene televisiooni vaadanud?)
- (6) Kas teil oli kunagi stress? (P.o Kas teil on kunagi stress olnud?)
- (7) Kas teie klassis olid niisugused probleemid? (P.o Kas teie klassis on niisuguseid probleeme olnud?)

Näited 5–7 vastavad kogemusperfektile, konkreetsele ajale ei ole viidatud ja silmas peetakse ajaperioodi, mitte -hetke. Näite 5 puhul on võimalik ka olevikuline variant *Kas te vaatate vene televisiooni?*, sest üliõpilane tahtis teada, kas vestluspartner on üldiselt kursis sellega, mis vene televisioonis toimub (avaliku elu tegelased ropendavad teles), kuid kuna jutt jätkus konkreetsete näidetega minevikust, siis see asjaolu loob konteksti, milles õige on täisminevik. Lauset 6 kasutati kuulajate aktiveerimiseks stressiteemalise ettekande algul. Näites 7 juhtis üliõpilane praktikatumnis kiusamisteemalist vestlust ja tahtis teada, kas õpilased on kiusamisega kokku puutunud.

Üksikuid näiteid on täismineviku ülekasutusest, nt täisminevik lihtmineviku asemel (8–10).

- (8) Üks aasta tagasi ma olen lõpetanud Narva Pähklimäe gümnaasium. (P.o [Üks] aasta tagasi ma lõpetasin Narva Pähklimäe gümnaasiumi.)

² Kuigi kogumisel on lähtutud põhimõttest, et näited ei kordu, oleks siin pidanud tegema erandi ja loendama, kuna tunni juhtimisel kasutavad õpetajad ühesuguseid väljendeid.

(9) Nad ei olnudki abiellunud! (*P.o Nad ei abiellunudki!*)

(10) Kui Bert trammiga kooli sõitis, on ta ära söönud banaani. (*P.o Kui Bert trammiga kooli sõitis, söi ta ära banaani.*)

Näidetes 8-9 on lähtunud reeglist, et eluloost räägitakse suures osas täismineviku vormis ning jäetakse tähele panemata, et lauses on ajamäärus või üks kord toimuv sündmus, mis peaks suunama lihtmineviku kasutamisele. 9. lause lõpetas moodsa muinasjutu, kus prints ja tütarlaps kumbki oma teed läksid. 10. näide on Berti päevast jutustavast tekstist, tegevustel ei ole olevikuga seost ja sellepärast tuleb kasutada lihtminevikku.

Ennemineviku vormi kasutatakse peamiselt kirjalikes tekstides. Laused 11-12 on moodustatud reeglit järgides, st lauses on kaks minevikku, millest üks eelneb teisele.

(11) Kuulamisülesanne võttis palju aega, rohkem kui oli enne kavandatud.

(12) Kui lapsed olid õppinud lugema ja kirjutama, hakkasid nad uurima ladina keele grammatikat.

Õiged on ka lihtlaused 13-14, sest kirjeldatud tunnid olid juba lõppenud, kui üliõpilased neid kirjeldama asusid.

(13) Tund oli suunatud ainult grammatika õpetamisele ja kordamisele.

(14) Õpilased olid huvitatud tunni teemast.

Esimese keele mõju üle teise keele omandamisele ja kasutamisele keeleuurijate seisukohad lahknevad. Vahekeeleteooria (vt nt Richards, Schmidt 2002, Verschik 2004) toetab seisukohta, et keeleülekanne (ingl *transfer*) võib olla positiivne – kui emakeele struktuur on õpitava ehk sihtkeele struktuuriga nii sarnane, et teatud konstruktsioonide omandamine ja korrektne kasutamine ei põhjusta õppijale raskusi ega tekita vigu. Interferents ehk negatiivne ülekanne ilmneb näiteks sõnavara otsetõlkes ja emakeelsete grammatiliste konstruktsioonide kasutamisel sihtkeeles. Negatiivset ülekannet põhjustavad erinevused emakeele ja sihtkeele struktuuris, sealhulgas teatud kontseptsioonide puudumine emakeeles. Kathleen Bardovi-Harligi (2000: 182) järgi (inglise keele) mineviku omandamisel esimene keel märgatavat rolli ei mängi. Ta on toonud näiteid nelja eri emakeelega õpilaste tulemustest, kellest ühe emakeeles esineb inglise keelele sarnane aegade süsteem, kolmel mitte. Õppijate saavutused on enam seotud motivatsiooni ja isiksuse eripäradega kui esimese keele mõjuga. Tõenäoliselt on keelevigu võimalik seletada mitmete põhjustega, kuid kui informantide keeles ilmnevad teatud struktuurid kindlas kasutusvaldkonnas, milles nad vigu ei tee, ja samad struktuurid esinevad ka nende emakeeles, ning vastupidi, kui samu struktuure kasutatakse ka juhtudel, kus sihtkeeles on erinev struktuur, siis võib nähtust ülekandeks pidada.

Käsitlemisel olevates keelenäidetes võib umbisikulise tegumoe vormides tuvastada nii positiivset kui negatiivset ülekannet emakeelest. Näidetes 15 on positiivse ülekande näited eri üliõpilaste õpikut või teadusteksti analüüsivatest töödest.

(15) Raamat on adresseeritud pedagoogidele. Raamatus on vaadeldud paljud kasulikud teemad. Ka on toodud õpetajate sotsiaalse valmiduse uurimise

tulemused. Kõik lõigud on järjestikud ja seotud teineteistega. Suur osa on pühendatud järgmisele küsimusele .. Raamatus on ka toodud õpilaste isiksuse ja klassi tundmaõppimise meetodid. Väga põhjalikult on kirjutatud sellest, kuidas organiseerida .. tundi.

Kuna täismineviku umbisikulise tegumoe vorm langeb kokku vene keele struktuuriga, siis seda laadi lausetes vigu peaaegu ei esine.

Praktikaaruannetest kogutud negatiivse ülekande näidetes (16) on enne minevikuga liialdatud, sest kõik näitena esitatud laused oleksid lihtminevikus eestipärasemad.

- (16) Koos uue materjaliga oli käsitletud mõned eesti keele reeglid. Kõik vestlus oli esitatud kahes keeles, et suurendada õpilaste integreerumine. Tunni käigus olid kasutatud sellised õppevahendid: žestid, miimika, tahvel, makk. Iga antud ülesanne oli kontrollitud selleks, et tuua esile kuidas õpilased tulevad toime neile antud ülesannetega. Tunni lõpetamine oli korrektselt tehtud.

Suulises või kirjalikus elulookirjelduses ei valmista üliõpilastele raskusi isikulise tegumoe täismineviku kasutamine, sest seda õpitakse juba kesktaseme testi sooritamiseks: *olen sündinud Narvas, olen lõpetanud Kesklinna gümnaasiumi* jms. *Kas*-küsimusega lausete põhjal võiks arvata, et liitajaga küsilauseste moodustamine valmistab raskusi, kuid näiteid ei ole piisavalt palju, et nende alusel midagi järeldada. Akadeemilises keelekasutuses esineb palju umbisikulist tegumoodi ja selle kasutamine venekeelsetele õppijatele raskusi ei valmista.

3.3. Informantide teadmised minevikuaegade tähenduste kohta

Eri minevikuvormide kasutusoskus kirjeldab üliõpilaste implitsiitset teadmisi eesti keele mineviku väljendusvõimalustest. Järgnevalt on kokkuvõtte üldistest seisukohadest ja ilmekamate näidetest, kuidas informandid mõistavad erinevusi mineviku ajavormide vahel ning mida nad teavad eri vormide kasutamisest. Kirjalik küsitlus on läbi viidud humanitaarainete (eesti keel teise keelena) magistriõppe 18 üliõpilase seas. Neile anti ülesanne seletada, mis vahe on eesti keele minevikuaegadel ja mis juhtudel eri aegu kasutatakse.

Lihtminevikku kasutatakse

- kui räägitakse minevikust ja see on lõpetatud tegevus;
- minevikus sooritatud tegevuste diferentseerimiseks. Tegevus seejuures ei pruugi lõpetatud olla;
- hiljuti minevikus, eile toimunud tegevuse väljendamiseks. See tegevus võib veel korduda;
- kui tahetakse rõhutada fakti, kui tegevus on kindel (mitte üldine ega ebamäärane);
- [väljendamaks] mida mina tegin konkreetsel ajal või perioodil.

Täisminevik väljendab

- mineviku seost olevikuga ja kui tegevus on aegade kooskõlastamisega, väljendab eelnevat tegevust;
- et tegevus toimus, aga ei ole tähtis millal;
- et tegevus toimus mitte eile ja üleile, vaid pikka/mõnda aega tagasi (paar-kolm aastat);
- kindlasti lõpetatud tegevust;
- et tegevus on toimunud küllalt ammu või tehakse harva.

Enneminevik väljendab

- ammutoimuvat tegevust või eelnevat tegevust minevikus;
- enne toimunud tegevust, olevikuga seost ei ole. Seda aega kasutatakse tavaliselt ajaloo kirjutamisel;
- sündmust või tegevust, mis toimus näiteks 10 aastat tagasi;
- kaugel minevikus sooritatud ja lõpetatud tegevust, mis kestis kindla aja-perioodi;
- tegevust, mis oli sooritatud enne teist tegevust minevikus ja ei ole seotud praeguse hetkega.

Üliõpilased kirjutasid mineviku väljendamise võimalustest nii, nagu nad oleksid seletanud kellelegi, kes eesti keele reegleid ei tunne. Kõikide mineviku ajavormide kirjeldused on üldjoontes õiged, kuid kolme ajavormi kohta kirjutatud võrreldes ilmneb, et informantide arusaama järgi on lihtminevik hiljutine möödanik, täisminevik sellest varasem aeg ja enneminevik veelgi kaugem, lausa ajaloolisi sündmusi kujutav aeg.



Joonis 1. Grammatiliste aegade paiknemine ajajoonel keelejuhtide arusaamade järgi. (EM – enneminevik, TM – täisminevik, LM – lihtminevik, O – olevik, T – tulevik)

Ajavormide vastandusest selgub, et lihtminevikuga väljendatud tegevus ei pruugi lõpetatud olla, kuid tegevus täisminevikus on kindlasti lõpetatud. Siin on ilmselt aegade kasutamise reeglid vastupidi meelde jäänud või väljendatakse veel kord arusaama, et kuna täisminevikus toimus tegevus varem kui lihtminevikus, siis on see kindlasti lõpetatud. Enneminevikku peetakse haruldaseks ajavormiks, mis väljendab väga ammu toimunud sündmusi ja igapäevases suhtluses seda vaja ei lähe.

3.4. Minevikuajad õppekirjanduses

Narva Kolledžisse võetakse eestikeelsetele õppekavadele vastu keskmisest parema eesti kui teise keele oskusega noori. Keeleteadmised ja -oskuse on nad saanud üldhariduskoolist ning kolledži eesti keele tundide eesmärgiks on nende keele lihvimine ja täiendamine. Ka minevikuajadesse puutuv on õpitud juba üldhariduskoolis. Keeleoskus ja -teadmised võivad olla mõjutatud õpetamismeetoditest.

1990. aastatest alates on populaarne olnud kommunikatiivne keeleõppemeetod, mis vähemalt kümme aastat tagasi jõudis ka eesti keele õpetamisse. Meetod

rõhutab kõnelemisoskuse tähtsust ja kommunikatsiooni funktsionaalset printsiipi, st enese mõistetavaks tegemist. Grammatiline korrektsus ei ole primaarne, seda saab asendada suhtlust toetavate strateegiatega, nagu kehakeel, parafraseerimine, asendamine, abi otsimine jms. Süstemaatiline grammatika õpetamine, eriti keerulisemate grammatikakategooriate seletamine on teisejärguline.

Keeleõppemeetodi teoreetiline põhjendus on algselt pärit Noam Chomskylt (universaalse grammatika teooria), hiljem Stephen Krashenilt (monitori teooria) ja kinnitust saanud Rod Elliselt, kelle järgi eksplitsiitsel õpetamisel ei ole tihedat seost oskusega õpitud grammatikat kasutada. Aastatel 1978–1991 tehtud kuueteistkümne uurimuse tulemused mõne üksiku erandiga näitavad, et õpetamisel ei ole erilist mõju grammatilise korrektsuse saavutamisele pikema aja vältel (Ellis 1994: 624–626). Tunnis õpetatud reeglid ununevad või ei suudeta neid keelekeskkonnas rakendada.

Teine moodus grammatikat õpetada, sealhulgas ka Eesti koolides, on tähelepanu pööramine vormimoodustusele. Õppijad omandavad oskuse moodustada liht-, täis- ja enneminevikku kõikidest verbidest ning suudavad hästi täita vormimoodustusülesandeid, kuid õige vormi valik ja kasutamine tekitab raskusi.

Lisaks õppemeetoditele mõjutavad keeleõpet õppematerjalid, grammatika õppimist näidete ja seletustega reeglikogud, tabelid ja skeemid, harjutustikud ja õppetekstid, milles on kasutatud õpitavaid konstruktsioone. Õppekirjandus peab täitma teatud tingimusi, et see efektiivne oleks. Gee Macrory ja Valerie Stone on uurinud reeglite õpetamise mõju perfekti omandamisele ning tulnud järeldusele, et lihtsad ja lühikesed reeglid on efektiivsemad (Macrory, Stone 2000: 60). Eestis on Jaan Mikk sõnastanud lihtsa keele reeglid (vt nt Mikk 2001, Mikk, Elts 1999), mille mõte on selles, et õppetekstid, reeglid ja seletused peavad olema õppijatele selged ja kergesti arusaadavad, ainult sel juhul avaldavad need soovitud mõju. Lihtsas keeles kirjutatud tekstide järgi õppides saavutavad õpilased paremaid tulemusi kui keerulisi, kompleksseid, palju võõrsõnu sisaldavaid akadeemilises vormis tekste lugedes.

Nelja koolides ja keelekursustel kasutatava tänapäevase grammatikaallika (Siirak 1998, Valmis, Valmis 2003, Kallas 2001, Uuspõld, Valmet 2001) põhjal on püütud järgnevas välja selgitada, kuidas on mineviku ajavormide tähendusest ja kasutamisest kirjutatud venekeelsetele õppijatele ning kas õppekirjandusel on olnud mõju informantide keeleteadmistele. Grammatika seletused on kirjutatud vene keeles, kasutatud on eestikeelseid termineid, reeglite näitlikustamiseks on toodud näiteid. Siiski ei ole seletused väga lihtsad ega üheselt mõistetavad.

Lihtminevik

- väljendab fakti, et tegevus on toimunud minevikus;
- väljendab tegevust, mis eelnes antud kõnesituatsioonile;
- kirjeldab sellist minevikusündmust, mis on seotud konkreetse hetkega minevikus.

Viimane seletus on keeleõppe seisukohalt kõige lihtsam ja selgem, sest on täpsem kui esimene seletus ega ole nii teaduslik kui teine definitsioon.

Täisminevik

- märgib tegevust, mis on toimunud minevikus, kuid on seotud olevikuga ja võib jätkuda tulevikus;
- märgib tegevust, mis oli resultatiivne;
- märgib tegevust, mis toimus minevikus ja on kõnehetkeks lõppenud;
- märgib tegevust, mis toimus minevikus piisavalt ammu;

- väljendab sageli pikaajalist ja korduvat tegevust;
- võib täita imperfektiivse aspekti (vn *saveršennõi vid*) funktsiooni;
- on mineviku ajavorm, mida kasutatakse rääkides tegevustest, mis on toimunud elu jooksul ja täpne aeg ei ole tähtis.

Võõrkeele õppijale on viimane seletus kõige mõistetavam, sest see selgitab üht olulist lihtmineviku ja täismineviku erinevust.

Enneminevik

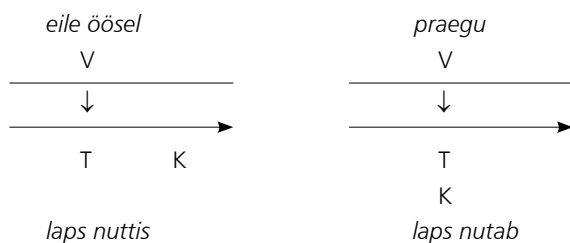
- märgib tegevust, mis eelnes teisele tegevusele minevikus;
- märgib tegevust, mis toimus varem, kui tegevus, mida väljendatakse lihtminevikuga;
- väljendab tegevust minevikus, mis eelnes kõnesituatsioonile, mis samuti toimus minevikus.

Nende definitsioonide peamiseks puuduseks on liigne keerukus. Üliõpilaste küsitluse vastustest võib järeldada, et nad on reeglite sõnastuse suures osas ära õppinud, kuid täpne mõte on jäänud selgusetuks. Mõni reegel on tekitanud segadust, näiteks väide, et täisminevik märgib tegevust, mis toimus minevikus piisavalt ammu. Üliõpilased on otsustanud, et “piisavalt” tähendab kahte või kolme aastat. Ennemineviku definitsioone lugedes on õppijad tähendusrõhu valesse kohta asetanud ja aru saanud, et enneminevik on toimunud enne teist tegevust minevikus, seega väga ammu.

4. Mis võiks olla teisiti?

Eesti keele kui teise keele õpetuses peaks minevikuaegade seletamisel toetuma klassikalistele skeemidele, mis määratlevad ajatähendusi kui suhet toimumisaja (T), viitehetke (V) ja kõnehetke (K) vahel (vt nt EKG I: 75–79, Metslang 1991), ja paigutama need kahele ajajoonele: ühele T ja K, teisele V. Nii on tehtud inglise keele grammatika õpikutes (vt nt Murphy 1997, Parrott 2004) ja soome keele grammatikas “Iso suomen kieloppi” (Hakulinen jt 2004: 1452–1454).

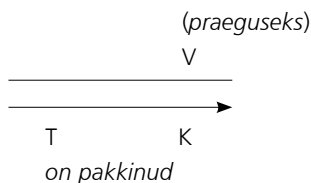
Kahe joone abil kujutatud **lihtmineviku ja oleviku** võrdlemisel saab selgeks, et oleviku puhul langevad kõnehetk, viitehetk ja toimumisaeg kokku (nt *praegu laps nutab*), kuid lihtmineviku viitehetk ja toimumisaeg on kõnehetkest varasemad (nt *eile öösel laps nuttis*).



Joonis 2. Lihtmineviku ja oleviku ajajooned

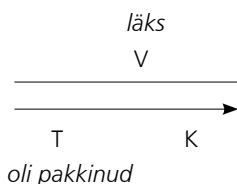
Täismineviku korral on viitehetk tavaliselt kõnehetkel, toimumisaeg aga on viitehetkest varasem: mineviku sündmusi vaadeldakse oleviku seisukohalt. Minevikusündmus ise on ajalisel ebamäärane: see võib olla toimunud ühe korra või rohkem,

ja võib ka viitehetkel jätkuda, nt *Sõber on oma asjad ära pakkinud*; *Sõber on alati oma asju korralikult pakkinud*; *Sõber on juba kaks tundi oma asju pakkinud*.



Joonis 3. Täismineviku ajajooned

Ennemineviku korral on toimumisaeg samamoodi viitehetkest varasem, viitehetk ise aga omakorda kõnehetkest varasem. Viitehetkeks on mingi minevikulise tegevuse või oleku aeg (nt *Sõber läks, kui oli pakkinud*).



Joonis 4. Ennemineviku ajajooned

Lisaks sõnalistele seletustele võimaldavad skeemid näitlikult kujutada ajalisi suhteid eri sündmuste, sh tegevuse/oleku ja kõnehetke vahel. Kaks paralleeljoont aitavad visuaalselt selgitada füüsilise aja ja grammatiliste ajavormide vahelist erinevust.

5. Kokkuvõte

Artiklis on uuritud kõrgtasemel eesti keelt valdavate Narva Kolledži eestikeelsetel õppekavadel õppivate üliõpilaste suulist ja kirjalikku keelekasutust. Kahe aasta jooksul kogutud andmete põhjal võib teha järelduse, et üliõpilased eristavad eesti verbi mineviku ajavorme, kuid teevad nende kasutamisel vigu. See kinnitab osaliselt R. Pooli ja E. Vaimanni uurimuse (2005: 134) ajavorme puudutavaid tulemusi – Narvas uuritud informantide lihtminevik ei ole küll üldistunud nii, et täis- ja enneminevik oleks kõrvale jäänud, kuid liitaegade kasutamisel tehtud vigade arv on suur. Erinevused kahe uurimuse vahel tulenevad ilmselt eri meetodite kasutamisest (R. Pool ja H. Vaimann on uurinud ainult kirjalikku keelekasutust esseedes) ja võimalik, et näidete hulgast ning kogumisperioodi pikkusest.

Üliõpilaste keelenäidete põhjal võib järeldada, et liitaegade moodustamisel valmistavad raskusi suulise esituse isikulise tegumoe vormid, sest täismineviku asemel kasutatakse palju lihtminevikku. Suulises spontaanses kõnes ei ole kõnelejal aega kasutatud ajavorme analüüsida. Enneminevikku suulises kõnes isikulisena peaaegu ei kasutata. Kirjalikus tekstis võib täismineviku umbisikulise tegumoe vormide kasutamisel märgata positiivset ülekannet emakeelest. Negatiivne ülekanne ilmneb ennemineviku umbisikuliste ajavormide ülemäärases kasutuses.

Õppematerjalide kirjelduste ja üliõpilaste sõnastatud arusaamade võrdlus mineviku ajavormide kasutamise kohta näitab, et informandid kordavad õppekirjanduse reegleid, kuid ei mõista sügavuti vahet eri ajavormide vahel. Kui uurida informantide teadmisi ajavormide kohta, siis ilmneb, et nende arusaama järgi kulgeb aeg ühel teljel. Olevikust tahapoole päris lähedale jääb lihtminevik, sellest kaugemale täisminevik ja väga kaugemale enneminevik. Eesti keele mineviku ajakategooriate graafiliseks kujutamiseks ei piisa ühest lineaarjoonest, sest liitjad väljendavad lisaks sündmuse toimumise ajale ka seost kõnehetke ja viitehetkega. Artiklis on välja pakutud lahendus, kuidas selgitada grammatiliste aegade erinevust kahe lineaarse ajajoone abil.

Edasist uurimist väärrib, millisel määral mõjutavad teadmised ajavormide kasutamise kohta praktilist keelekasutust. Selleks on vaja testida keeleoskust kontrollitud olukorras, kus informandid peaksid teadlikult ajavorme valima ning lisaks põhjendama tehtud valikuid.

Kirjandus

- Bardovi-Harlig, Kathleen 2000. Tense and Aspect in Second Language Acquisition: Form, Meaning and Use. Language Learning Monograph. Blackwell Publishers.
- EKG I = Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1995. Eesti keele grammatika. I. Morfoloogia, sõnamoodustus. Eesti Teaduste Akadeemia Eesti Keele Instituut. Tallinn.
- 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.
- Ellis, Rod 1994. The Study of Second Language Acquisition. Oxford Applied Linguistics. USA: Oxford University Press.
- Ellis, Rod 2004. The definition and measurement of L2 explicit knowledge. – Language Learning 54:2, 227–275.
- Ellis, Rod; Shawn, Loewen; Erlam, Rosemary 2006. Implicit and explicit corrective feedback and the acquisition of L2 grammar. – Studies in Second Language Acquisition 28, 339–368.
- Erelt, Mati; Erelt, Tiiu; Ross, Kristiina 2000. Eesti keele käsiraamat. Tallinn: Eesti Keele Sihtasutus.
- Evans, Vyvyan 2005. The Structure of Time. Human Cognitive Processing 12. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Hakulinen, Auli; Vilkkuna, Maria; Korhonen, Riitta; Koivisto, Vesa; Heinonen, Tarja Riitta; Alho, Irja 2004. Iso suomen kielioppi. Suomalaisen kirjallisuuden seuran toimituksia 950. Helsinki: Suomalaisen kirjallisuuden seura.
- Kallas, Roman 2001. Kõik on korras! Eesti keele õpik vene õppekeelelega gümnaasiumidele. Эстонский язык: практический курс и грамматика. Таллинн: J. Sarapu Kirjastus.
- Klein, Wolfgang 1986. Second Language Acquisition. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Külmoja jt = Кюльмоя, Ирина; Вайгла, Эда; Солль, Майе 2003. Краткий справочник по контрастивной грамматике эстонского и русского языков. Тарту: Tartu Ülikooli Kirjastus.
- Loog, Mai; Kerge, Krista 1999. Tuul tiibades. Eesti keele suhtluse kõrgtase. Tallinn: Riiklik Eksami- ja Kvalifikatsioonikeskus.
- Macrory, Gee; Stone, Valerie 2000. Pupil progress in the acquisition of the perfect tense in French: the relationship between knowledge and use. – Language Teaching Research 4, 1, 55–82.

- Metslang, Helle 1991. Ajast ja tema vormist verbitarindeis. – Keel ja Kirjandus 3, 163–68.
- Metslang, Helle 1997. On the use of the Estonian past tense forms through the last one 100 years. – Mati Ereht (ed.). Estonian: Typological studies II. Tartu Ülikooli eesti keele õppetooli toimetised 8. Tartu: University of Tartu, 98–145.
- Metslang, Helle; Krall, Ingrid; Pajusalu, Renate; Saarso, Kristi; Sõrmus, Elle; Vare, Silvi 2003. Keelehärm. Eesti keele probleemseid piirkondi. Käsiraamat. Tallinn: TPÜ Kirjastus.
- Mikk, Jaan; Elts, Jaanus 1999. A reading comprehension formula of reader and text characteristics. – Journal of Quantitative Linguistics 6, 3, 214–221.
- Mikk, Jaan 2001. Prior knowledge of text content and values of text characteristics. – Journal of Quantitative Linguistics 8, 1, 67–80.
- Murphy, Raymond 1997. Essential Grammar in Use: A Self-study Reference and Practice Book for Elementary Students of English with Answers. Grammar in Use Series. Cambridge: Cambridge University Press.
- Pancheva, Roumyana; Stechow, Arnim von 2003. On the present perfect puzzle. – K. Moulton, M. Wolf (eds). Proceedings of NELS 34, Vol. 2. Amherst: BookSurge Publishing, 469–484.
- Parrott, Martin 2004. Grammar for English Language Teachers. Cambridge: Cambridge University Press.
- Pool, Raili; Vaimann, Elle 2005. Vead kõrgtasemel eesti keele kõnelejate kirjalikus keelekasutuses. – Margit langemets (koost.), Maria-Maren Sepper (toim.). Eesti Rakenduslingvistika Ühingu aastaraamat 1. Tallinn: Eesti Keele Sihtasutus, 115–137.
- Rannut, Ülle 2005. Keelekeskkonna mõju vene õpilaste eesti keele omandamisele ja integratsioonile Eestis. Tallinna Ülikooli humanitaarteaduste dissertatsioonid 14. Tallinn: Tallinna Ülikooli Kirjastus.
- Richards, Jack C.; Schmidt, Richard 2002. Longman Dictionary of Language Teaching and Applied Linguistics. Pearson Education Limited.
- Siirak, Aino 1998. Eesti keele grammatika tabelites. Tallinn: Pangloss.
- Uuspõld, Valmet = Ууспõлд, Элен; Валмет, Айно 2001. Популярная грамматика эстонского языка. Eesti keele grammatiline vormistik: moodustamine ja kasutamine. Tallinn: Valgus.
- Valmis, Aavo; Valmis, Lembetar 2003. Lihtne eesti keele grammatika harjutuste ja võtmega. Эстонская грамматика для всех с упражнениями и ключом к ним. Tallinn: TEA.
- Vaiss, Natalia 2004. Verbi aspektikategooria eesti ja vene keeles. – Mart Rannut (koost.), Mart Rannut, Marju Kõivupuu, Tiit Päeva (toim.). Eesti keel: võõrkeelest teiseks keeleks. Tallinna Pedagoogikaülikooli eesti filoloogia osakonna toimetised 1. Tallinn: Tallinna Pedagoogikaülikooli kirjastus, 145–184.
- Verschik, Anna 2004. Mõnda vahekeelest ja kontrastiivsest analüüsist. – Mart Rannut (koost.), Mart Rannut, Marju Kõivupuu, Tiit Päeva (toim.). Eesti keel: võõrkeelest teiseks keeleks. Tallinna Pedagoogikaülikooli eesti filoloogia osakonna toimetised 1. Tallinn: Tallinna Pedagoogikaülikooli kirjastus, 132–144.
- Verschik, Anna 2002. Viga! Viga? Viga ... – Oma Keel 2, 76–82.

Diana Maisla (Tartu Ülikool) on lõpetanud Tartu Ülikooli eesti filoloogia erialal ja täitnud Tallinna Ülikooli pedagoogikamagistri õppekava. Töötab Tartu Ülikooli Narva Kolledžis ja õpib Tartu Ülikooli eesti keele ja tehnoloogia doktorikoolis. Uurimisvaldkond: eesti keele kui teise keele omandamine ja kasutus.
diana.maisla@ut.ee

USAGE OF ESTONIAN PAST VERB TENSES BY RUSSIAN-SPEAKING STUDENTS

Diana Maisla

University of Tartu

This article gives an overview on 1) how Russian-speaking students who are fluent in Estonian use temporal categories in Estonian; 2) what they know about the tense system of Estonian and 3) how the Estonian tense system is explained in grammar books written for schools with Russian as the language of instruction.

In Estonian past time is expressed by the simple past, the present perfect and the pluperfect, in Russian only by the simple past. The language data was collected from oral presentations and written texts produced by students during two years. Students distinguish between the three past tenses of Estonian but make a lot of mistakes. Features of both positive and negative transfer can be followed in students' presentations.

Students' opinions about the Estonian tense system repeat the rules given in grammar books. An unexpected finding was an understanding that the simple past represents a recent past, the present perfect expresses a time which passed at least three years ago and the pluperfect appears as a very distant past.

In the conclusion it is suggested to use the time lines in grammar sources to illustrate Estonian tenses. Two parallel lines enable to mark the time of speaking, the reference time and the time of action, thus improving students' understanding of the nature and differences between the simple past, the present perfect and the pluperfect.

Keywords: Estonian as a second language, second language learning and acquisition, verb tenses, Estonian

IN SEARCH OF THE BEST TECHNIQUE FOR VOCABULARY ACQUISITION

Mohammad Mohseni-Far

Abstract. The present study is intended to critically examine vocabulary learning/acquisition techniques within second/foreign language context. Accordingly, the purpose of this survey is to concentrate particularly on the variables connected with lexical knowledge and establish a fairly all-inclusive framework which comprises and expounds on the most significant strategies and relevant factors within the vocabulary acquisition context. At the outset, the study introduces four salient variables; learner, task and strategy serve as a general structure of inquiry (Flavell's cognitive model, 1992). Besides, the variable of context is introduced to enrich the examination process. The analysis specifically looks in depth at task-dependent strategies for this variable is much more substantial and yet in practice possesses more pedagogical implications. In lieu of seeking out superior strategies and techniques that bring about optimal outcomes, the investigator lays stress on the relative effectiveness of each individual strategy and technique. The organic idea of this exploration is that the most effectual and successful lexical development will occur in flexible and highly interactive syllabuses with a pedagogically well-reasoned balance between explicit and implicit activities.

Keywords: vocabulary learning/acquisition, lexical knowledge, lexical development, strategies and techniques, curriculum development

1. Introduction

Psychologists, linguists and language teachers have long been interested in vocabulary learning strategies. They have attempted to understand the role of the lexicon in language learning and communication. In particular, during the past fifteen years, the field of second language acquisition (SLA) has witnessed renewed attention to vocabulary learning and acquisition. There are many dimensions to

vocabulary acquisition, as reflected in the multitude of different areas of research being conducted on the topic.

In general, one way to study the task of vocabulary learning/acquisition is through the distinction between *knowing* a word and *using* a word. In operational terms, knowing a word may be seen as a continuum ranging from blurry recognition of its spelling or auditory pattern to (semantically, syntactically, stylistically) correct and contextually appropriate productive use. Retrieval of a word from the mental lexicon for productive use requires a higher degree of accessibility or, in other words, a more solid integration in various networks than is needed for receptive use (Groot 2000: 76). In other words, the purpose of vocabulary learning should include both *recalling* words and the ability to *apply* them automatically in a wide range of language contexts when the need arises. Vocabulary learning strategies, therefore, have to incorporate strategies for *recognizing* and *knowing* as well as *using* words. On the other hand, Henriksen (1999) draws attention to the fact that the acquisition of word meaning actually involves two interrelated processes: *item learning* (adding to the lexical store by creating extensional links, i.e. form-meaning mappings) and *system changing* (re-ordering/changing the lexical store via network building). The tendency in L2 vocabulary acquisition research has been to neglect the latter and focus on the former.

Another way to view vocabulary learning is to take it as a process of interconnected sub-tasks. When learners first encounter a new word, they might guess its meaning and usage from accessible tokens. Some learners might resort to consulting a dictionary. Others might make notes in the margins, between lines, or in a separate vocabulary notebook. Some learners will take advantage of simple rote repetition to commit the word to memory. Some would even attempt to use the word enthusiastically in a real context. Each of these task stages demands metacognitive decision, choice, and deployment of cognitive strategies for vocabulary learning. And each technique a learner puts to use will determine to a large degree how and how well a new word is learned/acquired.

This survey aims to profile a digest of recent research on vocabulary acquisition and to pinpoint areas that need further investigation. To this end, the paper centers on one particular area in depth, namely, vocabulary learning strategies. In so doing, it attempts to synthesize the most important and recent findings of research into vocabulary learning and acquisition, and to outline each of the main areas of research on the topic.

2. Critical survey

2.1. Preview

Mastery of vocabulary is an essential component of SLA. Effective L2 vocabulary acquisition is particularly important for foreign language learners who frequently acquire impoverished lexicons despite years of formal study (Hunt, Beglar 2005: 1). Today's language teachers and researchers have realized the important role of vocabulary in different pedagogical tasks. There is no doubt that virtually all L2 learners and their teachers are well aware of the fact that learning a L2 involves the learning of large numbers of words (Avila, Sadoski 1996, Laufer, Hulstijn 2001), yet

how to accomplish this task is often of considerable concern to them (Ott, Blake, Butler 1976: 37). How vocabulary is acquired and what the most efficient means are to promote effective acquisition are well established lines of enquiry in the field of SLA (De La Fuente 2002: 82).

The acquisition of a new lexical item is a complex process. Ellis, Tanaka and Yamazaki (1994: 457) argue that vocabulary acquisition involves discovering the frequency with which the item is used in speech and writing, its situational and functional uses, its syntactic behavior, its underlying form and the forms that can be derived from it, the network of associations between it and other items, its semantic features and, of course, the various meanings associated with the item. Due to this complexity, research on L2 vocabulary acquisition encompasses a number of different and diverse sub-areas. For example, how words are stored in the mental lexicon, automaticity of retrieval, the role of vocabulary in reading comprehension, the role of background knowledge, the implications of L2 vocabulary acquisition for L1 vocabulary research, how words are learned in context and other strategies for reading in L2 (Chun, Plass 1996: 184).

However significant lexical knowledge and vocabulary acquisition may be, the techniques and strategies recommended in this field remain challenging (Newton 2001: 30, Mohseni-Far 2006: 149). Although researchers and language teachers are becoming more and more convinced that vocabulary knowledge constitutes an essential part of competence in a L2, so far no comprehensive theories have been proposed that try to explain foreign language growth in terms of lexical development (Bogaards 2001: 321). In spite of the expansion in the amount of empirical research on vocabulary acquisition, consensus is lacking over issues such as the conceptualization of the process by which vocabulary acquisition occurs, the importance of context for acquiring vocabulary, and the extent to which learners build up specific strategies for vocabulary learning during their language acquisition.

In order to conduct the paper in a more systematic way, four prominent variables are taken into consideration; three from a cognitive model suggested by Flavell (1992), i.e., *learner*, *task*, *strategy*, as a general frame for the investigation as well as the variable of *context*. Initially, these four factors are introduced and defined, then in following chapters they are dealt with in more detail. Having taken these four variables into account, the paper also tries to shed significant light on major issues and controversies concerning each variable.

2.2. Cognitive variables and context

The four variables *learner*, *task*, *context*, and *strategy* are interconnected and work together to form the scaffold of learning/acquisition. The strategies a learner exercises and the effectiveness of these strategies very much depend on the learner him/herself (e.g., attitudes, motivation, prior knowledge, topic familiarity), the learning task at hand (e.g., type, complexity, difficulty, and generality), and the context or learning milieu (e.g., the learning culture, the influence of input and output opportunities). Therefore, an analysis of learning strategies will never be complete without identifying the learner, task and context configuration of the particular learning situation. Some strategies are more learner-dependent, some are more task-dependent, and others are more context-contingent.

The variable **learner** needs to be discussed from different perspectives. The learner brings to the language learning situation a wide spectrum of individual differences that will influence the learning rate and the ultimate learning result (Gu 2003: 2). The most widely reported learner factors include gender, language attitude, intelligence, prior knowledge, motivation, self-concept/image, personality, and cognitive and learning style. Motivation, emotion, and socio-cultural factors may affect the way in which people process information (Laufer, Hulstijn 2001: 7). The learner-dependent factors noted above to a large extent define how a learner handles a task.

The conception of the **learning task** incorporates materials being learned (such as the genre of a piece of reading) as well as the end the learner is attempting to achieve by using these materials (such as remembering, comprehending, or using language). *Task* can be defined as an activity or action which is carried out as a result of processing or understanding language. Laufer and Hulstijn (2001: 17) refine the meaning of task in the field of vocabulary acquisition in particular as “an activity in which meaning is primary”.

It is necessary to distinguish between two types of **context**. First, the learning context, which refers to the “learning environment and includes the teachers, the peers, the classroom climate or ethos, the family support, the social, cultural tradition of learning, the curriculum, and the availability of input and output opportunities” (Gu 2003: 2). Second, the language context, which refers to the textual or discursive place in which a particular word or structure can be found. Great importance has been attached to the latter by researchers and accordingly it will be subjected to fairly extensive and detailed investigation within this study. In support of the significance of context, Lawson and Hogben (1996: 106) suggest that from a psychological as well as a linguistic point of view, underlying the first guideline would be the need for vocabulary to be learned in context. Activities and reading materials that present words in meaningful contexts may contribute to vocabulary gains and are valued highly by students (Zimmerman 1997: 136).

A **learning strategy** (technique) covers a series of activities and efforts a learner makes to facilitate the completion of a learning task. Vocabulary learning strategies are any set of operations, steps, plans, routines used by the learner which affect this process. A technique is initiated when the learner examines the task, the situation, and what is available in his/her own mind. The learner then goes on to select, deploy, monitor, and weigh up the effectiveness of this activity, and decides if s/he needs to revise the plan and action. Theorists now place considerable stress on the importance of foreign language students’ developing autonomous learning strategies (Favretti, Silver, Tamburini, Gasser, quoted in Lawson, Hogben 1996: 106). Contextual guessing, skilful use of dictionaries, note-taking, paying attention to word formation and contextual encoding are some strategies normally applied by learners. Wenden (1987: 6) points out that language learning strategies have to consider different aspects of the language learning process. She identifies three areas in particular that language learning strategies refer to: (a) the actual *behaviour* of learners (what do learners do to learn an L2), (b) strategic knowledge (what learners know about the strategies they use), and (c) knowledge about aspects (other than strategies) of the L2 learning process, such as personal/motivational factors. These areas are summed up by Rubin (1987) as “what learners *do* to learn and *do to regulate* their [language] learning”. Language learning strategies are applicable to

a wide variety of language learning tasks, ranging from rather discrete and isolated tasks such as vocabulary, pronunciation and grammar to integrative tasks like oral communication and reading comprehension.

3. Vocabulary learning strategies: task-dependant perspective

Vocabulary learning strategies constitute a subclass of language learning strategies which in turn are a subclass of learning strategies in general. Fewer studies can be found on learner-related vocabulary learning strategies. Most of the empirical investigations on vocabulary learning strategies in a second language have focused on different sub-tasks of vocabulary learning. The emphasis of this study is placed upon this variable as well.

3.1. Inferring and vocabulary learning

The premise underlying this line of research is the belief that the vast majority of words learned in the L1 result from extensive and manifold exposures rather than direct instruction, and therefore successful vocabulary learning in a L2 should proceed in the same way. A number of topics have been put forward in the literature: guessing leading to vocabulary learning, inference from context, exposures needed to learn a word, and incidental (implicit) vocabulary learning and intentional (explicit) learning. Each of these is dealt with critically below in the chapter 3.3.

One of the strategies most often discussed in the literature is guessing word meaning from the clues made available by the context. Factors that affect the likelihood of success in inferencing include a context rich enough to provide adequate clues to guess a word's meaning (Celce-Murcia 2001: 290). Put another way, the unknown word to be guessed has to have plenty of comprehensible supporting context (Nation, Meara 2002: 44). In order to discuss this issue from a pedagogical standpoint, it is better to use the technical alternative "inferring" meaning from the context instead of "guessing".

In a sense, inference from context (IFC) is the mirror image of incidental acquisition on the strategy side. Nation has championed IFC as the "undoubtedly most important vocabulary learning strategy" (1990: 130), and recommends IFC especially for low-frequency words as their rarity does not repay the learning effort. One difficulty in this area of research is that there is no standard characterization of what is intended by context. In its broadest sense, of course, context may be said to comprise all the perceived phenomena that accompany the processing of a given stimulus, including the physical surroundings in which learning take place (Prince 1996: 479). Using the meanings of words together within the whole meaning of the sentence is the deepest level of processing and ensures the best memory (Cook 1991: 36). Crucial to inferring meaning from context is the degree to which context unveils word meaning. This degree is restricted by the learner's own background knowledge and constraints in the text itself. Hence, both teachers and learners must be aware that context functions to restrict meaning as well as to reveal it.

Some contexts do not provide a lot of information about a word, but others provide information that can take knowledge of the word forward. Lawson and Hogben (1996) found a lack of association between IFC and recall of word meaning, which they interpret as emphasizing the need to distinguish the use of IFC for the *generation* of new word meanings, and their *acquisition* for subsequent recall. Shu, Anderson and Zhang (1995: 79) suggest that significant learning from context is evident only when unfamiliar words appear repeatedly. When a context is rich enough, it will be sufficient to allow a learner to infer the full word meaning. On the other hand, if a context is too easily understood, then no vocabulary acquisition takes place. For example, at superior levels of proficiency, and when reading for comprehension – rather than to explicitly learn the meaning of unfamiliar words – the context may be easily understood without having to determine the meaning of every unfamiliar word (Pulido 2003: 241). Pulido (2004: 472) also states that words that may be easily guessed in the course of reading may not be better retained because of the lack of a need to allocate sufficient attention to the connection between the new word form and its meaning.

Aside from the richness of context, prior/background knowledge (topic familiarity) and the learner's vocabulary proficiency level are of great importance. There are many reports in the psychological and SLA literature that greater levels of background knowledge and expertise in a given subject matter result in the greater efficiency of attentional allocation during reading and enable richer analyses and textual interpretations, and, in turn, superior memory performance (Pulido 2003: 236). In simple terms, a learner is more successful at lexical inferencing when s/he is familiar with the topic. Since activation of appropriate knowledge structures stored in long-term memory is necessary to construct and integrate meaning successfully across discourse, it stands to reason that it will also have a strong bearing on the construction of meaning at the textual level. As a result, text recall is enhanced when learners possess and utilize the appropriate background knowledge.

The other significant issue is the learner's vocabulary proficiency level. The stronger vocabulary a learner has, the better s/he is able to understand the text that s/he interacts with (Chin 1999: 1). In an experiment, Prince (1996: 481) comes to the conclusion that advanced learners will make more efficient use of context than weaker learners, both during the study phase and during recall. To conclude, although IFC obviously has an important role to play in the overall scheme of vocabulary learning strategies (74% of Schmitt's (1997) intermediate level L2 learners of English used IFC, and 73% found it helpful), it cannot solve all reading comprehension, let alone acquisition, problems, and it needs to be supplemented by other vocabulary learning strategies.

3.2. Frequency

Frequency refers to the degree of probability of encountering a word in speech or print. It is obvious that the frequency of occurrence of the unknown words in the text has a significant impact on the retention of word meaning. The main reason given is that the reappearance of a word will strengthen the form-meaning connection in the learner's mental lexicon. In a more technical experiment by Hulstijn, Hollander and Greidanus (1996: 327) support was found for the hypothesis that frequency of

occurrence will foster incidental vocabulary learning more when advanced second L2 learners are given the meanings of unknown words through marginal glosses or when they look up meanings in a dictionary than when no external information concerning the meaning of unknown words is available.

It is a rather complicated task to mention the precise or relative number of words and accordingly exposures in a text required to learn a word. In addition, the number of exposures needed for mastery of a new word lies in many factors such as the salience of the word in the context, the richness of contextual clues, the learner's interest and his/her existing repertoire of vocabulary.

3.3. Incidental (implicit) vs. intentional (explicit) vocabulary learning

In implicit vocabulary learning learners are engaged in activities that do not focus attention on vocabulary. Incidental vocabulary learning is learning that occurs when the mind is focused elsewhere, such as on understanding a text or using language for communicative purposes (Celce-Murcia 2001: 289). Huckin and Coady (1999) point out that implicit learning cannot be totally incidental as at least *some* attention must be paid to the input by the learner. The implicit vocabulary learning hypothesis has its roots in Krashen's seminal *Input Hypothesis* (Krashen 1989), which maintains that the meanings of new words are acquired subconsciously as a result of repeated exposures in a range of contexts, where the conscious focus is not on form, but on the message.

The explicit vocabulary learning hypothesis holds that the employment of a range of vocabulary learning strategies can greatly facilitate and enhance vocabulary acquisition; in this view, learners are seen as active processors of information (Ellis 1995). From a pedagogically-oriented perspective, the goal of explicit teaching is "to lead the learner's attention", whereas the aim of an implicit focus on form is "to draw the learner's attention". Moreover, individual tasks can be located along an explicit or implicit continuum, and complex tasks may combine both explicit and implicit subtasks. From a cognitive psychological perspective, explicit learning can be contextualized as a conscious searching, building and testing of hypotheses and assimilating a rule following explicit instruction, whereas implicit learning is characterized by the automatic abstraction of the structural nature of the material arrived at from experience of instance (Hunt, Beglar 2005: 3).

The following appear to be essentials for successful incidental acquisition to occur (Schmitt and McCarthy 1997, Groot 2000):

- Adequate level of language proficiency (the ability to accurately decode the orthographic form of new words);
- A large L2 vocabulary (deducing the meaning of an unknown word requires a thorough understanding of the context, which in turn presupposes a large vocabulary);
- Strategic knowledge of the inferencing process;
- Context sufficiently rich in cues.

While definitions begin to clarify explicit and implicit learning, questions remain concerning what features of vocabulary and grammar are best learned explicitly or implicitly and how these processes occur. In addition to distinguishing between

explicit and implicit instruction, it is also necessary to consider which is more effective for promoting the acquisition of vocabulary knowledge. In fact, there is evidence in recent studies of second language learners that a combined approach is superior to a single learning method. Most researchers have recognized that a well-structured vocabulary programme needs a balanced approach that includes explicit teaching together with activities providing appropriate contexts for incidental learning (Celce-Murcia 2001: 286). Hunt and Beglar (2005: 3) also highlight the point that the most efficient learning involves a carefully selected combination of both explicit and implicit instruction and learning.

Finally, it should be mentioned that the most important explicit lexical instruction and learning strategies include lexis (word lists), using dictionaries and inferring from context, while the implicit approach primarily and basically involves learners in meaning-focused reading. In the preceding sections, inferring from context was discussed. Now, the role of the dictionary and then of meaning-focused reading will be examined.

3.3.1. The role of the dictionaries

Researchers are interested in investigating the part that dictionaries play in the learning of second language vocabulary. The dispute over the kind of dictionaries to be used in the foreign language classroom, and what dictionaries, if at all, should be used has always been an on-going one amongst language instructors and lexicographers. Presently, there is a prevalent view that L2 teachers should discourage learners from consulting dictionaries because extensive dictionary use can lead to word-for-word reading (Chin 1999: 3).

Three types of dictionaries are available: bilingual, monolingual, and bilingualized, and these can be found in either paper or electronic form. Both bilingual and monolingual dictionaries have their unique strong points and weaknesses for developing vocabulary knowledge.

Apart from the short and easy-to-understand definitions found in bilingual dictionaries, their strengths are: they can improve the reading comprehension of lower proficiency L2 learners, they assist vocabulary learning at all levels of proficiency (Hunt, Beglar 2005: 12). On the other hand, bilingual dictionaries 1) encourage translation, 2) foster one-to-one precise correspondence at word level between two languages, and 3) fail to describe adequately the syntactic behaviour of words (Gu 2003: 8).

In contrast, monolingual learners' dictionaries can be used to build and elaborate learner's vocabulary knowledge, using up-to-date and reliable sentence examples drawn from corpus data that provide information about meaning, grammar and usage (Hunt, Beglar 2005: 12). The monolingual entry can generally provide more detailed and precise information about idiomatic usage, common collocations, connotations, and register (Laufer, Hadar 1997: 189).

Since a combination of the good features of both types of dictionaries is attainable, there is considerable interest in the new bilingualized compromise dictionaries. A bilingualized entry typically includes: L2 definitions, L2 sentence information or L1 synonyms of the headword. These hybrid and fused dictionaries essentially provide translations in addition to retaining the good features of monolingual

dictionaries. Using bilingualized dictionaries is more efficient than using separate bilingual and monolingual dictionaries, and these dictionaries are more flexible because beginning and intermediate learners can rely on the L1 translation and advanced learners can concentrate more on the L2 part of the entry (Laufer, Hadar 1997, Gu 2003, Hunt, Beglar 2005).

The above types of dictionaries are also available in various electronic forms (software, pocket electronic and online dictionaries). It is obvious that electronic dictionaries (e-dictionaries) are easier and more convenient to use than printed dictionaries. They frequently permit the learners to search multiple sources (grammar, text usage as well as thesaurus), save and review words and definitions. Some e-dictionaries offer advanced searches, provide multimedia annotations, such as illustrations and video that assist in reading comprehension and vocabulary learning (Chun, Plass 1996: 185).

Regardless of the dictionary selected, learners need to learn how to use it effectively. In order to use a dictionary effectively, it is worth noting that more research is needed on what exactly learners do and how their dictionary strategies influence their learning results.

Research findings are inconclusive as regards the benefit of dictionary look-up for vocabulary acquisition/comprehension. The purpose of Hulstijn's (1993) study was to examine the connection between look-up behaviour and vocabulary knowledge on the one hand and inference ability on the other hand, taking into account the influence of task variables such as reading goal, word relevance and word inferability. The result was that learners did not look up all unfamiliar words, with look-up behaviour most strongly related to the perceived relevance of the word, but only modestly to readers' vocabulary knowledge. Interestingly, the ability to infer word meaning from context was not related to look-up behaviour at all, which suggests that a substantial proportion of good guessers may have been inclined to check their guesses subsequently. Finally, there seems to be no significant difference in lexical knowledge between 'maximalists' (subjects who looked up many words) and 'minimalists' (Hulstijn 1992, Chun and Plass 1996).

3.3.2. Meaning-focused reading

Reading is one of the most important ways that learners gain lexical knowledge incidentally. Recent developments in lexical semantics tell us a lot about vocabulary learning. Concentration on meaning-focused reading will result in incremental increases in vocabulary size, the elaboration of lexical knowledge and development of reading fluency. Zimmerman (1997: 123) believes that a considerable amount of word learning takes place incidentally through exposure to new words in meaningful contexts. Implicit learning through meaning-focused reading can occur incidentally as a result of learners' engaging in such activities as integrated task sets (a series of tasks requiring the use of multiple skills), narrow reading, rereading, timed and paced readings, intensive and extensive reading (Hunt, Beglar 2005: 15). Swanborn and Gloppe (2002: 98) also examine how reading texts for different purposes affects amounts of incidental word learning and then come to the conclusion that only the meaning of unknown words that are relevant for the reading purpose will be derived and recalled. Although these activities deserve greater attention, the

investigator intends to place emphasis upon extensive reading as the primary means for implicit learning for two reasons: its potentiality for triggering L2 learners and the fact that it can take advantage of the amount of meaningful input accessible to learners.

3.3.3. Extensive reading

In extensive reading, learners select and read large amounts of materials that appeal to them and are within their level of comprehension. The core of a meaning-focused input strand of a course is a well-organized, well-monitored, substantial extensive reading programme (Nation, Meara 2002: 40). Shu et al. (1995: 79) lay stress on the point that learning from context through extensive reading might be one of the important sources in L1 vocabulary growth. Intermediate and advanced L2 learners enlarge their vocabulary to a great extent through incidental learning during extensive reading (Hulstijn et al. 1996: 337).

For extensive reading to be efficient, learners must be frequently exposed to large amounts of comprehensible text. Since the knowledge gained from a single encounter with a lexical item is likely to be forgotten unless rapidly followed by another encounter, repeated contextualized exposures are necessary to consolidate and secure word meaning. Harmon (1999: 306) advocates the promotion of wide reading as an important vehicle for vocabulary development because learners gain words from context during reading. Teachers can promote consolidation through the use of post-reading activities in which learners first notice the target words by highlighting, underlining or circling them, and then process them by classifying, analyzing, or using the items productively (Hunt, Beglar 2005: 9).

3.4. Glosses and note-taking

Textual glosses have been used for a long time to facilitate L2 reading; other forms of glosses include pictorial (visual) and aural glosses, and various combinations thereof. In general, there exists a consensus among researchers that glosses facilitate reading comprehension and short-term vocabulary retention (Kost et al. 1999). As for textual glosses, there is a choice of language and form (single-choice or multiple-choice) glossing. Some studies have explored the effectiveness of the different options: Laufer and Shmueli (1997) found that glosses in L1 (Hebrew) led to better retention than L2 (English) glosses. Multiple-choice glossing, while encouraging deeper processing, suffers (in the printed form) from its failure to provide immediate feedback for learner errors. This problem can be attended to by using computers, however; Nagata (1999) found multiple-choice glossing with immediate feedback to be more effective than single-choice glossing.

After getting information about a lexical unit, learners may take notes, in the form of vocabulary notebooks, vocabulary cards, or simply notes in the margins or between the lines (marginal glosses). Note-taking is one of the basic strategies often recommended by researchers in the field of vocabulary learning. Vocabulary cards are invaluable in consolidating primary gains because of their ease of use and the number of communicative activities in which they can be used, as well as their

potentiality for increasing learner's enthusiasm. In a comparative empirical study Hulstijn et al. (1996: 336) conclude that the effect of marginal glosses will be greater than that of dictionary use because learners often do not make use of the dictionary. However, they then refine their conclusion in such a way that when readers do use a dictionary, the incidence of incidental vocabulary learning will be as good as, or even better than, when they are provided with marginal glosses.

3.5. Mnemonics: the role of memory

Among various other strategies frequently studied in the literature, one that requires a considerable amount of treatment and deep processing is mnemonics. Since vocabulary learning is essentially a memory matter, mnemonics should play a part in foreign language vocabulary learning as well. Mediation strategies – involving the new L2 word in some form of meaningful association – come in two varieties, *imagery mediation* and *semantic mediation*, and are examples of deep strategies. Imagery mediation in its plainest sense involves visualizing a mental picture or image of the L2 word in question; the most interesting variant of imagery mediation, however, is the *keyword method* (KW). KW has been mainly positively evaluated in the literature and can boast some impressive evidence of superiority over other direct or mnemonic strategies (Gu, Johnson 1996).

In the KW, the foreign word is remembered by being linked to a keyword, a similar sounding native word (the acoustic link), through an interactive image that involves both the foreign word and the native word (the imagery link). Avila and Sadoski (1996: 380) define the keyword method in two stages. First, the L1 word is associated with a familiar concrete word (keyword) based on acoustic similarities. The next stage is the production of an imaginal link between the target word and the keyword.

Since in this method the linkages from the vocabulary to a meaningful definition are provided, they will produce enhanced associative recall of definitions. As a result, the mnemonics links, particularly the imagery keyword technique, enable learners to memorize vocabulary more effectively. One of the frequently quoted claims of mnemonics is that people remember better over long periods of time if they have utilized mnemonic aids because they have a way to “get back at the word again” (Ott et al. 1976: 45).

Irrespective of the advantages of this method, there exist some limitations:

1. As the keyword is merely an approximation of the L2 form, ‘proper’ learning of correct L2 phonology and orthography is unnecessarily delayed (although Hulstijn (1997) points out that there are no theoretical reasons to assume KW plays an inhibiting role).
2. It depends on the nature of the words (KW can only be used for concrete nouns, rarely for abstract ones).
3. The mnemonic approach to vocabulary development emphasizes an unchanging one-to-one relationship between form and meaning. However, a key view in the applied linguist's conception of vocabulary is manifold meanings and multiple dimensions of meanings (referential, syntactic, pragmatic,

emotional, functional, literary, etc.). Especially in the case of polysemous words, it only helps establish *one* of the necessary meaning links.

4. Mnemonic devices are much less effective in productive vocabulary learning than in learning to comprehend the L2 because imagery association in the keyword technique allows retrieval of a keyword which is merely an approximation to the L2 form (Gu 2003: 15).

Although the applications of KW (and other mnemonic techniques, for that matter) are limited, its efficiency and value have been sufficiently proven. It is perhaps best seen as a helpful addition and complementary strategy, but not a substitute, for other vocabulary learning strategies.

3.6. Word-formation: focusing on form

Knowledge of lexical roots (etymological information and morphological origins) can assist in vocabulary development in that it helps learners predict or guess what a word means, elucidate why a word is spelt the way it is, and remember the word by knowing how its current meaning develops from its morphological roots. Learners should learn to identify morphemes which recur in a number of words and which can help them to identify at least part of the meaning, thus assisting them in guessing from context the meaning of apparently new items (Rivers 1981: 465). Contextual information and word morphology (e.g., word roots, affixes, and inflectional suffixes) are two major sources that readers use to interpret novel words (Mori 2003: 404).

Although most words can be decomposed into root words, prefixes, and suffixes, the degree to which these components specify the meaning of the whole varies widely. Shu et al. (1995: 80) introduce two major categories of words. First, the meaning of some words can easily be determined on the basis of the constituent morphemes or 'word parts' with little or no help from context (morphologically transparent words). Second, at the other extreme are words for which the components contribute almost nothing to the meaning (morphologically opaque words). In general, most words fall in between; although their meanings cannot be derived solely on the basis of word parts, their meanings are likely to be clear when they appear in even a moderately helpful context.

In this regard, two studies by Bogaards (2001: 321) underscore the importance of knowledge of form – but not of previously learned meaning – for the learning of new meanings for familiar forms. Then, based on his first experiment, he concludes that totally new single-word units are harder to learn and retain than multiword units of the same meaning but with a form that is made up of familiar words. In another technical study on the effects of semantic and structural elaboration on L2 lexical acquisition, Barcroft (2002: 323) provides evidence that increased semantic elaboration (evaluation of an item with regard to its meaning) can inhibit one's ability to encode the formal [phonemic/graphemic] properties of new words. Furthermore, in an interesting experiment Nagy, McClure and Mir (1997: 431) state that first language syntactic knowledge influences guesses about the meanings of unfamiliar words in a second language context. This effect is found among bilinguals who have experienced a variety of amounts and types of exposure to English.

A learner needs three skills in order to make use of affixation: breaking a new word into parts so that the affixes and roots are revealed; knowing the meaning of the parts; and being able to connect the meaning of the parts with the meaning of the word. Put another way, to make use of word parts, the learner needs to know the most useful word parts of English (20 or so high-frequency prefixes and suffixes are enough initially), needs to be able to recognize them in their various forms when they occur in words and needs to be able to relate the meanings of the parts to the meaning or definition (Nation, Meara 2002: 46).

4. Vocabulary learning strategies: learner-dependant view

From guessing at the first encounter, to possible dictionary use and note taking, to memorizing, encoding, and contextual activation, vocabulary acquisition is a dynamic process involving metacognitive choices and cognitive implementation of a whole spectrum of strategies. Whether and how a learner evaluates the task requirement and whether and how a cognitive strategy is utilized are often dependent more on the learner than on the task.

Researchers in the vocabulary learning strategy area have attempted to identify the ways in which “good” and “poor” learners approach lexical learning. Good learners are more aware of what they can learn about new words, pay more attention to collocation and spelling, and are more conscious of contextual learning. In contrast, poor learners are generally characterized by their apparent passivity in learning. Swanborn and Gloppe (2002: 99) suggest that poor learners are not as adept as good learners at adjusting their reading strategies to fit the reading purpose and accordingly learning more affectively from context. In other words, good learners are good comprehenders and learners who are better skilled at handling vocabulary acquisition tasks. In a similar study, Pulido (2003: 239) states that strong learners tend to have more efficient decoding skills and larger sight vocabularies than weak learners. So these outcomes in reading performance can be explained by individual differences and strategies applied in these areas. Gu and Johnson (1996: 668) also studied 850 university English as a foreign language students in China, and tried to establish how different vocabulary strategies were related to language learning outcomes. Both Pearson’s correlation and multiple regression analyses revealed that self-initiation, selective attention, and deliberate activation of newly learned words consistently predicted both vocabulary size and general proficiency. Other predictors of success included contextual learning, dictionary, and note-taking strategies.

5. Discussion

The factors and variables involved in vocabulary acquisition were critically discussed and analyzed in a rather comprehensive review of the relevant literature. The crucial notion developed here is that in lieu of seeking the best strategy that creates the best results, the relative efficiency of each individual strategy should be appreciated. In addition, the organic notion underlying this study is that the most effective and efficient lexical development will occur in multifaceted curriculums

that attain a pedagogically sound equilibrium between explicit and implicit activities for L2 learners at all levels of their development.

The following points outline the most important implications of the issues covered here.

1. The focus of research in the linguistics tradition regarding vocabulary acquisition has largely centered on vocabulary learning (what is learned/to be learned; product-oriented view) rather than acquisition (how is vocabulary learned, process-oriented perspective) (Crow 1986, Meara 1980).
2. From a psychological standpoint, memory strategies have received the major share of attention in vocabulary acquisition, probably for the reason that this learning/acquisition process has largely been taken as a memory problem.
3. Much of the stress on incidental vocabulary learning has zeroed in on how helpful incidental learning is and how much can be learned incidentally, often ignoring the fact that a lot can be learned intentionally and by design during reading with the help of strategies (e.g., inferring from context, dictionary use, note-taking, as well as intentional repetition).
4. Research on vocabulary learning strategies does indicate a tendency for good learners to be more active than poor ones, but the field would benefit from a clearer focus on precisely how learners learn lexical units and how their strategies are related to acquisition outcomes (Schmitt 2000).

Research efforts have largely been directed towards discovering and supporting the best strategy for vocabulary retention. In reality, learners tend to utilize a variety of strategies in combination. Therefore, as highlighted in this paper, it seems highly preferable to seek to design a multifaceted framework with a pedagogically sound equilibrium that takes advantage of the positive and appropriate strengths of each strategy in a due and sensible time and at relevant level.

6. Conclusion

The main purpose of this survey was to review the effects of learning strategies on L2 vocabulary acquisition during recent years. Vocabulary acquisition can be best conceived as a process in which L2 learners negotiate word meaning from a text level to a word level. This shift is necessary so that the learner can form a mental connection between the word form and his/her meaning premise (Mohseni-Far 2006: 162).

Word knowledge has linguistic, psycholinguistic, and sociolinguistic aspects. Lexical competence is far more than the ability to define a given number of words; it entails knowing a great deal about each word, including information about its general frequency of use, the syntactic and situational limitations on its use, its generalizability, its collocational probabilities, its core form, its derived forms, and its semantic features. The process by which learners acquire this information appears to take place gradually over a long period of time and is very complex and difficult to investigate. In conclusion, a lot of work has been done to find overall patterns of strategy use. However, the choice, use, and effectiveness of vocabulary learning strategies very much depend on the task (e.g., breadth vs. depth), the learner (e.g., cognitive and cultural styles of learning, motivation), and the context. Future

research, therefore, needs a more technical and closer approach that takes all the aforementioned features into account.

In operational terms, word knowledge includes the ability to recall meaning, infer meaning, comprehend a text, and communicate orally. No single approach can address all of these skills; when learners receive input about vocabulary only from reading or only from the use of lists, drills, or skill-building activities, they have not addressed the range of skills needed for word use. An effective approach to word learning should be multifaceted and comprehensive in what they require of the learner and rich in what they reveal about the target words. In a nutshell, the overall presumption must be that there is no simple answer to the key question of what the most efficient method/strategy of L2 word learning/acquisition is/should be. It depends very much on variables like degree of L1-L2 equivalence of the words to be learned, the intensity (both qualitative and quantitative) of processing, the age and cognitive level of the learner, the quantity and quality of rehearsal practice and variables discussed through this research. As a result, more experimentation systematically and analytically controlling these variables is required to collect data that will provide more insight into their relative importance and efficiency. Out of the meticulous examination of obtained results, a flexible, multi-purpose and eclectic skeleton and curriculum needs to be designed and developed so as to take best advantage of the strong points and effective potential of each strategy at its due time of utilization and which best fits with the existing overall characteristics of the pertinent task, context and learner. Therefore, a successful cognitive task performance (i.e. vocabulary acquisition) will be achieved within such a flexible, multiple-purpose and highly interactive framework.

References

- Avila, E.; Sadoski, M. 1996. Exploring new applications of the keyword method to acquire English vocabulary. – *Language Learning* 46, 379–395.
- Barcroft, J. 2002. Semantic and structural elaboration in L2 lexical acquisition. – *Language Learning* 52, 323–363.
- Bogaards, P. 2001. Lexical units and the learning of foreign language vocabulary. – *Studies in Second Language Acquisition* 23, 321–343.
- Celce-Murcia, M. 2001. *Teaching English as a Second or Foreign Language*. USA: Heinle & Heinle.
- Chin, C. 1999. The effects of three learning strategies on EFL vocabulary acquisition. – *The Korea TESOL Journal* 2, 1–29.
- Chun, D. M.; Plass, J. L. 1996. Effects of multimedia annotations on vocabulary acquisition. – *The Modern Language Journal* 80, 183–212.
- Cook, V. 1991. *Second Language Learning and Language Teaching*. London: Edward Arnold.
- Crow, J. T. 1986. Receptive vocabulary acquisition for reading comprehension. – *The Modern Language Journal* 70, 242–250.
- De la Fuente, M. J. 2002. Negotiation and oral acquisition of L2 vocabulary. – *Studies in Second Language Acquisition* 24, 81–112.
- Ellis, N. C. 1995. The psychology of foreign language vocabulary acquisition: Implications for CALL. – *Computer Assisted Language Learning* 8, 103–128.
- Ellis, R.; Tanaka, Y.; Yamazaki, A. 1994. Classroom interaction, comprehension, and the acquisition of L2 word meanings. – *Language Learning* 44, 449–491.

- Flavell, J. H. 1992. Metacognition and cognitive monitoring. – T. O. Nelson (ed.). *Metacognitive: Core Readings*. Boston: Allyn and Bacon, 3–8.
- Groot, P. J. M. 2000. Computer assisted second language vocabulary acquisition. – *Language Learning & Technology* 4, 60–81.
- Gu, Y. 2003. Vocabulary learning in a second language. – *TESL-EJ* 7, 1–26.
- Gu, Y.; Johnson, R. K. 1996. Vocabulary learning strategies and language learning outcomes. – *Language Learning* 46, 643–679.
- Harmon, J. M. 1999. Initial encounters with unfamiliar words in independent reading. – *R.T.E.* 33, 304–338.
- Henriksen, B. 1999. Three dimensions of vocabulary development. – *Studies in Second Language Acquisition* 21, 303–317.
- Huckin, T.; Coady, J. 1999. Incidental vocabulary acquisition in a second language: A review. – *Studies in Second Language Acquisition* 21, 181–193.
- Hulstijn, J. H. 1992. Retention of inferred and given word meanings: experiments in incidental vocabulary learning. – P. Arnaud, H. Béjoint (eds). *Vocabulary and Applied Linguistics*. London: Macmillan, 113–125.
- Hulstijn, J. H. 1993. When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. – *The Modern Language Journal* 77, 139–147.
- Hulstijn, J.H. 1997. Mnemonic methods in foreign language vocabulary learning: theoretical considerations and pedagogical implications. – J. Coady, T. Huckin (eds). *Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press, 203–224.
- Hulstijn, J. H.; Hollander, M.; Greidanus, T. 1996. Incidental vocabulary learning by advanced foreign language students: the influence of marginal glosses, dictionary use, and reoccurrence of unknown words. – *The Modern Language Journal* 80, 327–339.
- Hunt, A.; Beglar, D. 2005. A framework for developing EFL reading vocabulary. – *Reading in a Foreign Language* 17, 1–31.
- Kost, C. R.; Foss, P.; Lenzi, J. J. 1999. Textual and pictorial glosses: effectiveness on incidental vocabulary growth when reading in a foreign language. – *Foreign Language Annals* 32, 89–97.
- Krashen, S. D. 1989. We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis. – *The Modern Language Journal* 73, 440–464.
- Laufer, B.; Hadar, L. 1997. Assessing the effectiveness of monolingual, bilingual, and “bilingualized” dictionaries in the comprehension and production of new words. – *The Modern Language Journal* 81, 189–196.
- Laufer, B.; Hulstijn, J. 2001. Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. – *Applied Linguistics* 22, 1–26.
- Laufer, B.; Shmueli, K. 1997. Memorizing new words: Does teaching have anything to do with it? – *RELC Journal* 28, 89–108.
- Lawson, M. J.; Hogben, D. 1996. The vocabulary learning strategies of foreign language students. – *Language Learning* 46, 101–135.
- Meara, P. 1980. Vocabulary acquisition: a neglected aspect of language learning. – *Language Teaching and Linguistics Abstracts* 13, 221–246.
- Mori, Y. 2003. The roles of context and word morphology in learning new Kanji words. – *The Modern Language Journal* 87, 404–420.
- Mohseni-Far, M. 2006. Techniques utilized for vocabulary acquisition. – *Language Forum* 32, 147–165.
- Nagy, W.; McClure, E. F.; Mir, M. 1997. Linguistic transfer and the use of context by Spanish-English bilinguals. – *Applied Psycholinguistics* 18, 431–452.
- Nation, I. S. P. 1990. *Teaching and Learning Vocabulary*. Boston, MA: Heinle & Heinle.

- Nation, P.; Meara, P. 2002. Vocabulary. – N. Schmitt (ed.). *An Introduction to Applied Linguistics*. London: Arnold, 35–54.
- Nagata, N. 1999. The effectiveness of computer-assisted interactive glosses. – *Foreign Language Annals* 32, 469–479.
- Newton, J. 2001. Options for vocabulary learning through communication tasks. – *ELT* 55, 30–37.
- Ott, C. E.; Blake, R. S.; Butler, D. C. 1976. Implications of mental elaboration for the acquisition of foreign language vocabulary. – *IRAL*, XIV, 37–48.
- Prince, P. 1996. Second language vocabulary learning: the role of context versus translations as a function of proficiency. – *The Modern Language Journal* 80, 478–493.
- Pulido, D. 2003. Modeling the role of second language proficiency and topic familiarity in second language incidental vocabulary acquisition through reading. – *Language Learning* 53, 233–284.
- Pulido, D. 2004. The relationship between text comprehension and second language incidental vocabulary acquisition: a matter of topic familiarity. – *Language Learning* 54, 469–523.
- Rivers, W. M. 1981. *Teaching Foreign-Language Skills*. Chicago: The University of Chicago Press.
- Rubin, J. 1987. Learner strategies: theoretical assumptions, research history and typology. – A. Wenden, J. Rubin (eds). *Learner Strategies in Language Learning*. New York: Prentice Hall, 15–30.
- Schmitt, N. 1997. Vocabulary learning strategies. – N. Schmitt, M. McCarthy (eds). *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge : Cambridge University Press, 199–227.
- Schmitt, N. 2000. *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Shu, H.; Anderson, R. C.; Zhang, H. 1995. Incidental learning of word meanings while reading: a Chinese and American cross-cultural study. – *Reading Research Quarterly* 30, 79–95.
- Swanborn, M. S. L.; Gloppe, K. D. 2002. Impact of reading purpose on incidental word learning from context. – *Language Learning* 52, 95–117.
- Wenden, A. 1987. Conceptual background and utility. – A. Wenden, J. Rubin (eds). *Learner Strategies in Language Learning*. New York: Prentice Hall, 3–13.
- Zimmerman, C. B. 1997. Do reading and interactive vocabulary instruction make a difference? an empirical study. – *TESOL Quarterly* 31, 121–140.

Mohammad Mohseni-Far (State University of Shahid Chamran, Ahwaz, Iran) is currently M.A. student majoring in TEFL/Applied Linguistics. He is the member of Iranian National Foundation of Elite and also Young Researchers Club (YRC), Islamic Azad University, Ahwaz-Iran. His main areas of interest: English for Specific Purposes (ESP), psycholinguistics, systemic functional linguistics (SFL), vocabulary acquisition, syllabus design, discourse analysis, theories of second language learning and teaching.

Mmb_m2005@yahoo.com

Mohammad Mohseni-Far, M.A.
 Department of English,
 Faculty of Humanities & Letters,
 Shahid Chamran University,
 Ahwaz, Iran

PARIMAT SÕNAVARA OMANDAMISE MEETODIT OTSIMAS

Mohammad Mohseni-Far

Shahid Chamrani Riiklik Ülikool, Ahwaz, Iraan

Käesolev kriitiline analüüs on suunatud sõnavara õppimise ja omandamise meetoditele teise või võõrkeele kontekstis. Lähtudes sõnavara tundmisega seotud muutujatest püütakse luua võimalikult kõikehõlmav raamistik mahutamaks sõnavara omandamise tähtsaimaid strateegiaid ja tegureid. Neli olulist muutujat on *õppija*, *ülesanne* ja *strateegia* (need moodustavad uurimuse põhistruktuuri, vt. Flavelli kognitiivne mudel) ning *kontekst*. Tähelepanu keskmes on ülesandekesksed strateegiad, kuna ülesanne on oluliselt sisukam muutuja, samas sõltub see õpetajast enam kui muud. Püüdmatagi esile tõsta parimaid strateegiaid ja meetodeid optimaalse tulemuse saavutamiseks, rõhutatakse iga üksiku strateegia ja meetodi suhtelist efektiivsust. Eelnevast järeldub loomulikult, et edukaim sõnavara areng toimub paindlikus ning võimalikult interaktiivses õpikeskkonnas, kus valitseb pedagoogiliselt hästi põhjendatud tasakaal eksplitsiitsete ja implitsiitsete õpitoimingute vahel.

Märksõnad: sõnavara õppimine ja omandamine, sõnavara tundmine, sõnavara arendamine, strateegiad ja meetodid, õppekava arendamine

ÜLEKÜSIMINE, ÜMBERSÕNASTAMINE JA MITTEMÕISTMINE INFOTELEFONIKÕNEDES

Siiri Pärkson

Ülevaade. Igas inimestevahelises vestluses, sh dialoogis infotelefoni töötaja (infoandja) ja infotelefonile helistaja (kliendi) vahel võib esineda probleeme arusaamisega ning osalejate püüdu nende probleemidega toime tulla. Konversatsioonianalüüsi teoorias on kuulaja poolt esiletõstetud mõistmisprobleeme nimetatud partneri algatatud parandusteks. Eestikeelses suhtluses võib leida kolme erinevat tüüpi partneri algatatud parandusi ning neid nimetatakse vastavalt üleküsimine, ümbersõnastamine ja mittemõistmine. Artiklis uuritakse konversatsioonianalüüsi abil 126 eestikeelset infodialoogi. Ligi pooltes esineb vähemalt üks partneri algatatud parandus. Loomuliku suulise suhtluse mõistmiseks ja rakendamiseks (nt süsteemides, milles arvuti peab inimesega loomulikus keeles suhtlema) on vaja muuhulgas selgeks teha, mis täpselt suhtlusprobleemide tekkimist põhjustab ning milliseid keeliseid struktuure inimesed tegelikult kasutavad, kui nad signaliseerivad probleemi.

Võtmesõnad: arvutilingvistika, konversatsioonianalüüs, institutsionaalne suhtlus, suuline kõne, telefonisuhtlus, infodialoogid, üleküsimine, ümbersõnastamine, mittemõistmine, eesti keel

1. Sissejuhatus

Tavaliselt on igas vestluses suhtlus- või keeliseid probleeme, mida on vaja lahendada. Konversatsioonianalüüsis (vt ptk 2) nimetatakse sellist probleemide lahendamist paranduseks. Parandust saab algatada nii kõneleja ehk probleemi tootja kui ka vestluspartner ning paranduse saab läbi viia samuti kõneleja või vestluspartner. Käesolevas artiklis on vaatluse all vestluspartneri algatatud parandused.

Artikli teema on tihedalt seotud:

- eesti dialoogikorpuse (EDiK) arendamisega. Keskendudes dialoogides esinevatele probleemsetele kohtadele, saab arendada edasi dialoogiaktide analüüsi;
- eestikeelse infotelefoniteenuseid osutava dialoogsüsteemi loomisega.

Sissejuhatusel järgnevas peatükis tutvustatakse konversatsioonianalüüsi teooriat ja mõnesid artiklist arusaamiseks olulisi mõisteid. Seejärel antakse ülevaade töös kasutatavast materjalist ja uurimismeetodist. Järgneb analüüs ning iga partneri algatatud paranduse lähem vaatlus. Viimases peatükis esitatakse järeldused ja kokkuvõte.

2. Teooria

Salvestatud loomuliku interaktiivse keelekasutuse uurimist nimetavad Ian Hutchby ja Robin Wooffitt (1998: 14) **konversatsioonianalüüsiks** (ingl conversation analysis, lüh CA). CA tegeleb eeskätt vestluskäitumise regulaarsete muustritega. Eesmärgiks on anda struktuuriliste tunnuste süstemaatiline kirjeldus, mis toetub konkreetsete nähtuste analüüsile (Wooffitt 1990: 20).

Suuline vestlus kahe inimese vahel jaguneb **kõnevoorudeks** (ingl *turn*). Kõnevoor on ühe kõneleja jätkuv häälesolek. Tüüpiliselt on teineteisele järgnevatel kõnevoorudel erinevad autorid, kuid ühele kõnelejale võivad kuuluda ka kaks või enam üksteisele järgnevat kõnevooru (nt kui teine kõneleja ei võta talle pakutud kõnevooru, võib esimene kõneleja alustada taas oma kõnevooriga). Kõnevooru vorm, sisu ning pikkus pole ette määratud, vaid on varieerumisvabad (Hutchby, Wooffitt 1998: 47).

Kõnevoorud liigenduvad üksusteks, mida eesti keeles nimetatakse **lausungiteks** (ingl *utterance*) ning mille piirid on eelkõige intonatsioonilised, aga ka grammatilis-pragmaatilised (vt Hennoste jt 2002).

Dialoog jaguneb struktuurilt erinevateks väiksemateks ülesanneteks ehk dialoogiaktideks, millest mõned on kergesti lahendatavad (nt teretamisele tervitusega vastamine) ning teised on keerulisemad (nt küsimusele õige vastusega vastamine). EDiK-i märgendamisel on lähtutud muuhulgas põhimõttest, et iga lausung peab omama aktimärgendit, aga võib omada samaaegselt ka mitut märgendit. Kuna lausungi lõppu märgib eeskätt langev intonatsioon, mida märgendatakse punktiga, siis peab iga punktiga lõppev üksus saama aktimärgendi (Hennoste, Rääbis 2004).

EDiK-i märgendamisel on kasutatud Tartu Ülikoolis välja töötatud dialoogiaktide tüpoloogiat (Hennoste, Rääbis 2004)¹, mis põhineb konversatsioonianalüüsil ning mille järgi jagunevad partneri algatatud parandused (esiliikmed) üleküsimiseks, ümbersõnastamiseks, mittemõistmiseks ja muuks. Need mõisted on defineeritud järgmiselt (Hennoste, Rääbis 2004: 56, 58, 59).

- **Üleküsimine** (ingl *clarification*) on akt, mille puhul kuulaja kordab täpselt või väikeste modifikatsioonidega mõnda partneri lausungit, fraasi või sõna, et saada kinnitust selle kohta, et öeldu oli just selline (nt *kas te ütlesite nii?*).
- **Ümbersõnastamine** (ingl *reformulation*) on algatus, mille abil kuulaja pakub oma tõlgenduse (hüpooteesi, ümbersõnastuse, üldistuse, järelduse)

¹ Iga dialoogiakti nimetus (nt RIE: TERVITUS VÕI IL: TÄPSUSTAMINE) koosneb kahest osast: 1) akronüüm, mis sisaldab kas kaks või kolm tähte, millest esimesed kaks näitavad aktirühma (nt RI = rituaalid, IL = infolisa) ning kolmas, mida kasutatakse ainult naaberpaare moodustavate aktide puhul, näitab, kas on tegu naaberpaari esi- või järelliikmega (vastavalt E või J), 2) akti nimi (nt TERVITUS, TÄPSUSTAMINE).

Viidatud raamat sobib ka käesolevas artiklis esinevate mõistetega tutvumiseks.

varemõeldu kohta. Tema eesmärgiks on saada kinnitust, kas info mõte oli just selline (nt *kas te mõtlesite seda?*).

- **Mittemõistmine** (ingl *non-understanding*) on akt, millega kuulaja annab teada, et ta ei kuulnud mingit vooru või osa sellest, ei saanud millestki aru või on saadud info tema jaoks nii üllatav, et ta peab vajalikuks seda üle kontrollida.
- **Muu** (ingl *other*) hulka kuuluvad paranduse algatused, mis ei mahu toodud süsteemi.

CA järgi on verbaalse suhtluse oluline omadus organiseeritus järjenditena, sekventsidenä (ingl *sequence*). Sekventsiaalsus tähendab seda, kuidas üksteisele järgnevad kõnetoimingud liituvad üksteisega ja milliseid üksteisega seotud tegevuste järjendeid nad moodustavad (Hennoste, Rääbis 2004: 23). Sekventsisisüsteemi keskne mõiste on **naaberpaar** (ingl *adjacency pairs*). Naaberpaari esiliikme olemasolu nõuab kindlat järelliiget (nt hüvastijätt vastuhüvastijätku jne). Järelliikmed omakorda liigenduvad ootuspärasteks ning ootusevastasteks, mis realiseeruvad üldjuhul keeleliselt erineval kujul (Hennoste jt 2002).

Partneri algatatud parandused kuuluvad naaberpaare moodustavate aktide hulka ning tihti viitavad probleemlika asukohale, lokaliseerivad selle. Seega otsitakse partneri algatatud parandustes uut võimalust probleemkohast arusaamiseks ning samas pakutakse probleemi põhjustajale võimalus olukord ise lahendada. Võimalus eneseparandamiseks antakse isegi siis, kui teatakse viga ja osatakse seda parandada.

Kuna probleemne koht peab alati eelnema parandusele ning asetsema mõnes partneri algatatud parandusele eelnevas kõnevoorus, soovitab Marja-Leena Sorjonen kasutada terminit **probleemvoor**, mis on kõnevoor, kus probleemne koht asetseb (Sorjonen 1997).

3. Materjal ja meetod

Käesolev uurimismaterjal pärineb EDiK-ist, dialoogide analüüsimiseks rakendatakse konversatsioonianalüüsi meetodit. 2007. aasta novembri seisuga on EDiK-is 1061 teksti (178 000 sõnet), millest 945 on telefonikõned ning 116 silmast-silma vestlused.²

EDiK-ist on uurimistöösse valitud juhuslikult 126 dialoogi (8 356 sõna), kuigi on jälgitud, et partneri algatatud paranduste (üleküsimine, ümbersõnastamine, mittemõistmine) protsendiline osakaal oleks enam-vähem vastavuses korpuses esinemise suurusjärguga. Võrdluseks, EDiK-i dialoogides leidub 2007. aasta augusti seisuga kokku 583 (60%) üleküsimist, 208 (22%) ümbersõnastamist ning 174 (18%) mittemõistmist. Kõik valitud dialoogid on infodialoogid, st salvestused klientide telefonikõnedest infotelefonile. Seega on analüüsi all institutsionaalse suhtluse dialoogid.

Artiklis kasutatakse arvandmete esitamisel abivahendina Margus Treumuthi loodud dialoogikorpuse tööpinki.³ Lisaks kasutatakse kvantifitseerivaid väljendeid, nagu *tihti, üldiselt, korduvalt, suhteliselt harva, tavaliselt* jmt. Niisiis on

² <http://math.ut.ee/~koit/Dialoog/EDiC.html> (21.11.2007).

³ <http://math.ut.ee/~treumuth/> (08.08.2007).

kvantitatiivne informatsioon kohati esitatud suhteliselt ebamääraselt, kuid Paul ten Have'i (1999: 144) kinnitust mööda on konversatsioonianalüüsi rakendustes selliste kvantifitseerivate väljendite kasutamine tavaline.

4. Partneri algatatud parandused infodialoogides

Artikli uurimismaterjalis esineb partneri algatatud parandusi 60 dialoogis ja 66 dialoogis neid ei leidu. On tähtis mainida, et ühes dialoogis võib esineda rohkem kui üks partneri algatatud parandus (60-st dialoogist 31-s). Samuti võivad ühes dialoogis esineda koos erinevat liiki partneri algatatud parandused (60-st dialoogist 24-s). Erinevat liiki partneri algatatud parandused võivad olla mitte ainult ühes dialoogis, vaid ka lausa ühes kõnevoorus. 17-s dialoogis esineb ühte liiki partneri algatatud parandust rohkem kui korra. Analüüsitavates dialoogides esineb kokku 121 partneri algatatud parandust (vt tabel 1).

Tabel 1. Partneri algatatud paranduste esinemine analüüsitavates infodialoogides⁴

Partneri algatatud paranduse esiliige	Juhtumite arv	Juhtumite protsent	Dialoogide arv
Üleküsimine	72	60%	50
Ümbersõnastamine	25	21%	20
Mitterõistmine	21	17%	18
Muu (süsteemiseerimata aktid)	3	2%	3
Kokku	121	100%	91 (60)⁵

Infodialoogides on 1200 kliendi kõnevooru, infoandja kõnevoore esineb 1527 korda. Põhjus, miks infoandja voorude koguarv on suurem kui kliendi kõnevoorude arv, tuleneb ühelt poolt nn esimeste lausungite distributsiooni reeglist (Schegloff, 1968: 1076, viidatud ten Have 1999: 16 järgi), mille järgi vastaja räägib esimesena. Teisalt on infokõnedele omane, et klient annab märku, millal ta on valmis kõne lõpetama ning nii on infoandja sagedamini ka viimase lausungi ütleva.

Tabelist 2 võib näha, et üleküsimisi ning mitterõistmisi on algatanud rohkem klient ehk infotelefonile helistaja, seevastu ümbersõnastamist kasutavad enamasti infoandjad. Partneri algatatud parandusi on rohkem kliendil, kuid see tuleneb eelkõige üleküsimiste rohkusest.

Tabel 2. Partneri algatatud paranduste algatajad ja sagedused

	Klient	Protsent	Infoandja	Protsent
Üleküsimine	54	76%	18%	38%
Ümbersõnastamine	4	6%	21%	45%
Mitterõistmine	13	18%	8%	17%
Kokku	71	100%	47%	100%

⁴ Edaspidi kasutatakse lühendeid: PPE – partneri algatatud paranduse esiliige; ÜK – üleküsimine; ÜS – ümbersõnastamine; MM – mitterõistmine.

⁵ Ühes dialoogis võib esineda korraga erinevat liiki partneri algatatud parandusi. Seega, liitmistulemus tabelis on 91, kuid tegelikult esineb uurimismaterjalis partneri algatatud parandusi 60 dialoogis, nagu eespool on väidetud.

Kõikides analüüsitavates dialoogides on tegu partneri algatatud eneseparandustega. Niisiis võib väita, et eneseparandamise eelistamine ei ole inglise keele spetsiifiline, vaid kehtib samuti eestikeelsetes dialoogides. Seda väidet toetab Krista Strandsoni uurimistööde (Strandson 2000, 2001, 2002) raames läbiviidud argivestluste analüüs, mille järgi 144 vestluskaaslase algatatud parandusest 134 juhul jättis vestluskaaslane paranduse läbiviimise õiguse kõnelejale ning 10 juhul viis vestluskaaslane paranduse ise lõpuni.

4.1. Üleküsimine

Vormiliselt võib üleküsimised jaotada otsesteks üleküsimisteks ja modifitseeritud üleküsimisteks.

- **Otsene üleküsimine** (näide 1)⁶ on siis, kui kõnevoor jääb eelnevaga samaks.

- (1) 1. V: kuus? (.) [üks.] | KYJ: INFO ANDMINE || PPJ: LÄBIVIIMINE |
2. H: [kuus] üks. £ | KYE: VASTUST PAKKUV | | PPE: ÜLEKÜSIMINE |
3. V: [jah]. | KYJ: JAH | | PPJ: LÄBIVIIMINE |

Näitelõigus 1 kordab klient (rida 2) sõna-sõnalt infoandja (rida 1) kõnevooru.

- **Modifitseeritud üleküsimine** (näide 2) esineb siis, kui kõneleja kordab eelnevat kõnevooru ning teeb mõned grammatilised ja/või leksikaalsed muudatused. Kõneleja võib lisada mõne fraasi või, vastupidi, võib lühendada eelnevat kõnevooru.

- (2) 1. neli neli neli, | DIJ: INFO ANDMINE |
2. (0.5)
3. H: kolm `nelja jah | KYE: VASTUST PAKKUV | | PPE: ÜLEKÜSIMINE |
4. V: jah? | KYJ: JAH | | PPJ: LÄBIVIIMINE |

Kliendi üleküsimine leiab aset enamasti telefoni (või faksi) numbri kordamisel. Infoandja üleküsimine toimub, kui ta pole kindel, et kuulis, mõistis või mäletab korrektselt eelnevalt saadud olulist informatsiooni. Kuulmishäire või tõlgendus-raskusena tunduv võib olla ka infoandja poolt aja võitmiseks või planeerimiseks kasutatav vahend. Sel juhul ei pruugi üleküsimise vastus tegelikult rolli mängida, kui sealt just uut informatsiooni ilmsiks ei tule.

Üleküsimine esineb tihti siis, kui rääkija (uuritud dialoogides enamasti klient) usub või näib uskuvat, et ta ei kuule õigesti. Analüüsitud materjalis esineb ainult 3 juhtumit, kus klient kuuleb tõepoolest valesti. Nii väheste juhtumite arv võib viidata ka sellele, et üleküsimine on pigem tava kui vajadus. Kuid infotelefonile helistamise puhul on üleküsimine kindlasti ka funktsionaalne, sest üleküsimist kasutatakse enamasti kontrollimaks andmeid, mis peavad olema täpsed (aadressid, telefoninumbrid, nimed). Üleküsimise põhjused on esitatud tabelis 3.

Mõnigi kord põhjused veidi kattuvad ning seetõttu on üleküsimiste jaotus tinglik. Lisaks võivad mõned põhjused olla seotud mittekuulmise ja kinnituse otsimisega.

⁶ Näidetes on nummerdatud lausungid ning pausid kõnevoorude vahel ehk pausid, mis ei kuulu kummagi kõnevooru ega ühegi lausungi juurde. Näidetes kasutatud transkriptsioonimärgid on seletatud artikli lõpus.

Tabel 3. Üleküsimise põhjused

Põhjus	Juhtumite arv
(telefoni/faksi)number	38
aadress/asukoht	12
institutsiooni nimi/üldnimi	8
inimese nimi	4
info puudumine	4
ebaootuspärane või üllatav informatsioon	3
muu	3
Kokku	72

Loetelust on näha, et kõige enam üleküsimisi põhjustab (telefoni/faksi)number, mis iseenesest on infotelefoni dialoogidele väga omane. Aadressi all on mõeldud siin ja edaspidi tänava nime ning maja numbrit; asukoht on laialivalgavam mõiste, mis tähistab näiteks linna, linnajao, tänava nime või muud piirkonda (nt kesklinn, Lõuna-Eesti). Institutsiooni nimi tähistab ametlikku või kõneleja poolt arvatavat ametlikku nime. Üldnimi on nimetus, mis ei ole seotud konkreetse institutsiooniga (nt apteek, loomakliinik).

(Telefoni)numbrite teatamise korral esinevad järgmised reeglipärasused.

- Telefoninumbrit öeldakse osade kaupa (näites 3 read 4, 7, 9), üleküsimise akt toimub pärast viimast numbriosa ning korratakse ainult viimast numbriosa (3, rida 10). See on ka kõige sagedasem üleküsimise muster. Kuulaja omakorda annab numbriosade vahel tagasisidet (akti VR: NEUTRAALNE JÄTKAJA abil; 3, rida 6), et ta on kuuldel ja teine võib jätkata ja/või ta saab öeldust aru. Tagasiside võib väljenduda numbriosa kordamisena (VR: NEUTRAALNE JÄTKAJA; 3, rida 8) või lühikeste ühesõnaliste reageerimistena (*jaa?*; *jah*), mida on nimetatud ka minimaaltagasisideks (Hennoste, Rääbis 2004). Reeglipärane on seegi, et kuulaja jääb numbriosade vahel esimesena valitud tagasiside (numbriosa kordamine vs. lühikesed ühesõnalised reageerimised) juurde.

- (3) 1. H: £ öelge palun Tartu migratsiooni`ameti (.) selle telefoni`number. £
| DIE: SOOV |
2. V: jaa? | VR: NEUTRAALNE VASTUVÕTUTEADE |
3. üks=`hetk, | DIJ: EDASILÜKKAMINE |
4. (...) .hh `seitse `kolm `seitse, | DIJ: INFO ANDMINE |
5. (.)
6. H: £ jaa? £ | VR: NEUTRAALNE JÄTKAJA |
7. V: `seitse `viis, | DIJ: INFO ANDMINE |
8. H: £ jaa? £ | VR: NEUTRAALNE JÄTKAJA |
9. V: `kuus `null. | DIJ: INFO ANDMINE |
10. H: £ kuus `null.= £ | KYE: VASTUST PAKKUV | | PPE: ÜLEKÜSIMINE |
11. V: =jah= | KYJ: JAH | | PPJ: LÄBIVIIMINE |

- Telefoninumbrit öeldakse osade kaupa (näites 4 ridadel 1, 4, 6), üleküsimise akt toimub pärast viimast numbriosa ning korratakse kogu telefoninumbrit ehk siis järjekorras kõik kuulnud numbriosad (näites 4 real 8). Kuulaja annab tagasisidet ilma uut informatsiooni lisamata ega kõneõigust taotlemata. Tagasiside võib väljenduda numbriosa kordamisena (VR: NEUTRAALNE JÄTKAJA) või lühikeste ühesõnaliste reageerimistena *jaa?*; *jah* (VR: NEUTRAALNE JÄTKAJA). Üsna tavaline on ka neutraalne jätkaja *mhmh* (Hennoste, Rääbis 2004: 134), kuid analüüsitud dialoogides selline juhtum puudub.
- (4) 1. V: .hh `number, kinnitamata andmetel `seitse, (.) `kolm, `kuus | DIJ: INFO ANDMINE |
 2. (0.5)
 3. H: jaa? | VR: NEUTRAALNE JÄTKAJA |
 4. V: `kuus `seitse, | DIJ: INFO ANDMINE |
 5. H: jah | VR: NEUTRAALNE JÄTKAJA |
 6. V: seitse= `seitse. | DIJ: INFO ANDMINE |
 7. (0.8)
 8. H: `seitse `kolm `kuus `kuus `seitse `seitse `seitse. | KYE: VASTUST PAKKUV | | PPE: ÜLEKÜSIMINE |
 9. V: jah. | KYJ: JAH | | PPJ: LÄBIVIIMINE |
- Telefoninumbrit öeldakse osade kaupa, üleküsimise akt toimub pärast viimast numbriosa ning selles vóorus korratakse ainult viimast numbriosa (VR: NEUTRAALNE VASTUVÓTUTEADE) ja seejärel üleküsimisena kogu telefoninumbrit. Sellise skeemi korral kasutatakse neutraalse jätkajana eelmise numbrí kordamist. (Seda mustrit võib pidada ka eelmise punkti erandlikuks alaliigiks, sest selliseid juhtumeid esineb ainult 2.) Näites 5 toimub üleküsimine (rida 8) pärast info andmisi (read 2, 4, 6) ning pärast neutraalse jätkamise ja neutraalse vastuvótureate dialoogiakte (read 3, 5, 7).
- (5) 1. V: :h `Teguri kolkend `seitse. | DIJ: INFO ANDMINE |
 2. `ja: `number=on: > kinnitamata < `andmetel `seitse `kolm `kuus? | DIJ: INFO ANDMINE |
 3. H: kolm kuus? | VR: NEUTRAALNE JÄTKAJA |
 4. V: seitse öheksa? | DIJ: INFO ANDMINE |
 5. H: seitse öheksa? | VR: NEUTRAALNE JÄTKAJA |
 6. V: seitse seitse. | DIJ: INFO ANDMINE |
 7. H. seitse=seitse, | VR: NEUTRAALNE VASTUVÓTUTEADE |
 8. a `kordan seitse kolm `kuus seitse `öheksa seitse seitse. | KYE: VASTUST PAKKUV | | PPE: ÜLEKÜSIMINE |
 9. V: jah. | KYJ: JAH | | PPJ: LÄBIVIIMINE |
- Telefoninumber öeldakse välja ühes vóorus (vahel võivad olla lühikesed pausid) ning klient või infoandja kordab oma kõnevóorus kogu numbrit (üleküsimine). Näites 6 teavitatakse kogu telefoninumber ühes kõnevóorus (rida 4) ja pärast 0,5-sekundilist pausi (rida 5) toimub kliendi üleküsimine kogu numbrí kordamise näol (rida 6).

- (6) 1. V: jaa | VR: NEUTRAALNE VASTUVÕTUTEADE |
 2. üks=`hetk, | DIJ: EDASILÜKKAMINE |
 3. (...) haigekassal on kli`enditeeninduse `lühinumber. | DIJ: INFO ANDMINE |
 4. üks `kuus, (.) `kolm `kuus `kolm. | DIJ: INFO ANDMINE |
 5. (0.5)
 6. H: üks `kuus? `kolm kuus `kolm [jah] | KYE: VASTUST PAKKUV |
 | PPE: ÜLEKÜSIMINE |
 7. V: [jah.] | KYJ: JAH | | PPJ: LÄBIVIIMINE |

Ülalesitatud üleküsimise variandid kehtivad olenemata sellest, kas numbriosa koosneb 2, 3 või 4 numbrist. Siiski, 2 numbrist koosnemine on kõige tavalisem. Oluline on, et need mustrid kehtivad üksnes juhul, kui numbrütlemeise kestel pole partner algatanud ühtki parandust ning kogu number on saadud ladusalt lõpuni öelda.

EDiK-i tüpoloogiat järgides selgub, et antud artiklis analüüsitud dialoogides on kõik üleküsimised algatatud vastust pakuva küsimusena, mis eeldab vastust *jah* või *ei* (vt tabelit 4). Seejuures jaatav vastus ning nõustuv *ei* märgivad, et probleem on lahendatud. Samas esineb lisaks jaatusele või eitusele ka täpsustamist, mis võib järgneda jaatusele või eitusele, aga ka jaatuse/eituse asemel, st täpsustamisele ei eelne ega järgne eitust ega jaatust. See on ka põhjus, miks tabelis 4 on algatuste arv (72) väiksem kui parandusaktide arv (78) – ühele parandusalgatusele võib järgneda mitu parandust või täpsustust.

Tabel 4. Üleküsimivad parandusaktid (PPE: ÜLEKÜSIMINE)

Algatus	Algatuste arv	Parandusakt	Parandusaktide arv
KYE: VASTUST PAKKUV	72	KYJ: JAH	64
		IL: TÄPSUSTAMINE	10
		KYJ: EI	3
		KYJ: NÕUSTUV EI	1
Kokku	72		78

Enamus üleküsimistest moodustatakse vastust pakuva sõna või fraasi abil. Põhjusiks on, et enamasti küsitakse üle telefoninumber või osa sellest. Niisiis on nimeetatud üleküsimise vorm infotelefonikõnede spetsiifiline. Küsimuse märkimiseks kasutatakse palju sagedamini lausungi lõppu lisatud partiklit *jah* kui küsipartiklit *või/vä*. Üleküsimist algatab rohkem klient ning seda sõltumata vormist, milles (vastust pakkuv) küsimus on esitatud.

Juhul kui vastust pakkuv üleküsimine sisaldab tegusõna, mida infodialoogides esineb suhteliselt harva (16 juhtumit), siis peaaegu alati (14 juhtumit) kasutab paranduse algataja oleviku kindlat kõneviisi. Vaid ühel juhtumil kasutatakse lihtmineviku kindlat kõneviisi.

Tabel 5. Üleküsimist algatava dialoogiakti lingvistiline struktuur⁷

Dialoogiakt	Lingvistiline struktuur	Juhtumite arv
KYE: VASTUST PAKKUV	sõna/fraas	46
	sõna/fraas + <i>jah</i>	9
	lause + <i>jah</i>	6
	lause	5
	see + lause + <i>jah</i>	3
	lause + <i>või/vä</i>	2
	sõna/fraas + <i>vä</i>	1
Kokku		72

4.2. Ümbersõnastamine

Kui telefoninumbri üleküsimine on infodialoogides isegi ootuspärane, siis (telefoni)numbrite ümbersõnastamine on pigem erandlik ning võib esineda juhtudel, kus järjest on mitu ühesugust numbrit (vt näidet 7) või kui infot antakse mitme (telefoni)numbri kohta, mille algus on sama ja lõpp muutuv.

- (7) 1. H: te:re | RIJ: VASTUTERVITUS |
 2. {-} `Tartus kurgu-nina-kõrva kliiniku registra`tuuri `telefoni` numbrit. | DIE: SOOV |
 3. V: jah? | VR: NEUTRAALNE VASTUVÕTUTEADE |
 4. üks= `hetk, | DIJ: EDASILÜKKAMINE |
 5. (2.0)
 6. V: `registra`tuur on: seitse (.) kolm üks? | DIJ: INFO ANDMINE |
 7. (0.8)
 8. H: jaa? | VR: NEUTRAALNE JÄTKAJA |
 9. V: üheksa seitse, (0.8) seitse seitse. | DIJ: INFO ANDMINE |
 10. (1.0)
 11. H: > kolm seitset on lõpus=sis, < | KYE: VASTUST PAKKUV | | PPE: ÜMBERSÕNASTAMINE |
 12. V: jah. | KYJ: JAH | | PPJ: LÄBIVIIMINE |

Näites 7 soovib klient kliiniku registratuuri telefoninumbrit (rida 2). Infoandja alustab numbrit ütlemist real 6 ja lõpetab real 9. Sellele järgneb kliendi algatatud numbriosa ümbersõnastamine real 11.

Ümbersõnastamine võib näidata, millises osas on ümbersõnastaja teadmistes või arusaamises lünk, mis vajab täitmist.

- Institutsiooni nimi põhjustab ümbersõnastamist iseäranis juhul, kui nimi on konsonantidest koonev lühend (nt PFT) või kuulajal tekib õigekirja suhtes kahtlus – siis tuleb vajalikke komponente täpsustada ning see viib ümbersõnastamiseni. Tõenäoliselt nõuavad ümbersõnastamist (ja ka teisi partneri parandusalgatusi) sagedamini nimed, kus on eeldatav inglise keele mõju.

⁷ Fraasiks on loetud kahe, kolme või enama numbrit loetelu. Lause tunnuseks on verbi olemasolu.

- (8) 1. V: .hh kas=ee `Farma on `efiga või pee=`hashsh. | KYE: ALTERNATIIV |
 | VTE: VASTUSE TINGIMUSTE TÄPSUSTAMINE |
 2. H: ee `ei, `Farma=nagu (.) nagu: | KYJ: MUU | | VTJ: VASTUSE
 TINGIMUSTE TÄPSUSTAMINE |
 3. V: `ef `aa `er `em `aa.= | KYE: VASTUST PAKKUV | | PPE: ÜM-
 BERSÕNASTAMINE |
 4. H: =just | KYJ: JAH | | PPJ: LÄBIVIIMINE |

Näitelõigus 8 on infoandja ümbersõnastamise (rida 3) põhjuseks firma nime õigekiri (read 1 ja 2).

- Kui info andmisel on olemas asukoht või aadress, siis peab soov sisaldama ka institutsiooni või selle üldnime. (Teoreetiliselt on võimalik variant, kus institutsiooni nime/üldnime ei esine, nt *Õelge, mis firma asub sel aadressil*, kuid uuritava materjali hulgas selline juhtum puudub.) Mõnikord on aga üldnime asemel kasutatud tegevust (nt *ujula, veekeskus – ujuma*). Sellistel juhtudel võib infoandja tegevuse sõnastada ümber institutsiooni üldnimeks.

- (9) 1. H: te:re. | RIJ: VASTUTERVITUS |
 2. paluksin teada `Tartus on võimalik `Mõisavahe tänaval `soolaravi teha.
 | DIE: SOOV |
 3. t=ega te ei=oska seal kontaktnaa- `numbrit anda. | KYE: JUTUSTAV
 KAS |
 4. V: `soolakambrid jah? | KYE: VASTUST PAKKUV | | PPE: ÜMBER-
 SÕNASTAMINE |
 5. H: jah | KYJ: JAH | | PPJ: LÄBIVIIMINE |

Näites 9 soovib klient teada, kas Mõisavahe tänaval on võimalik soolaravi teha (rida 2). Infoandja sõnastab tegevuse *soolaravi teha* ümber üldnimeks *soolakamber* (rida 4). Ümbersõnastamine on seotud helistaja öeldud andmete viimisega andmebaasis kasutatavale kujule (nt *soolaravi* → *soolakambrid*).

Infoandja ümbersõnastamised ongi enamasti seotud otsesemalt või kaudsemalt andmebaasiga (andmebaasi struktuur või andmete õigekiri). Infoandja muudab sisusõnu (nt *soolaravi* → *soolakamber*; *Kaa salong* → *Kaa arvutisalong*) või lisab teavet (nt *Elmar* → *raadio Elmar*). Kliendi ümbersõnastamised on seotud mõnes eelnevas voorus kuulnud teabe kordamise ehk faktide täpsustamisega.

Ilmneb, et ümbersõnastamist kasutades soovib infoandja aidata nii klienti kui ennast. Ühelt poolt annab ümbersõnastamine teavet, kas ümbersõnastatud kõnevooru on õigesti mõistetud. Teisalt arendab see vestlust ning viib lähemale suhtluseesmärgi saavutamisele. Ümbersõnastamine on niisiis oluline lüli kõnevooru mõistmise ning dialoogi arengu vahel.

Erinevalt teistest partneri algatatud parandustest algatab ümbersõnastamise enamasti infoandja. Selle põhjuseks on infotelefoni vestluste omapära, kus infotöötaja peab andmebaasist leidma vajaliku info ja kasutab ümbersõnastamist kinnituse saamiseks, et ta asub otsima õiget (vajalikku) teavet. Infoandja ümbersõnastamine on lühem kui eelnev kliendi kõnevoor, sest infoandja teeb eelmisest kõnevoorst kokkuvõtte, korrastab ja kohandab eelmise kõnevooru andmed andmebaasi struktuurile vastavaks.

Mõlemad, nii infoandja kui klient, algatavad ümbersõnastamise enamasti vastust pakkuva küsimusena. Enamjaolt järgneb sellele järgmises kõnevoorus jaatus (nõustuv *ei* täidab seejuures sama rolli mis jaatamine). Vahel lisatakse ka täpsustavat infot, harvem vastatakse ümbersõnastamisele eitusega või informatsiooni andmisega. Tähelepanu väärib, et parandus iseseisvalt puht täpsustamisena ei esine, st täpsustamisele peab eelnema kinnitamine või ümberlukkamine (siin erineb ümbersõnastamine üleküsimisest).

Tabel 6. Ümbersõnastavad aktid (PPE: ÜMBERSÕNASTAMINE)

Algatus	Arv	Parandus	Arv	Protsent
KYE: VASTUST PAKKUV	24	KYJ: JAH	19	61,3%
		IL: TÄPSUSTAMINE	6	19,4%
		KYJ: EI	3	9,7%
		KYJ: INFO ANDMINE	1	3,2%
		KYJ: MUU	1	3,2%
KYE: SULETUD KAS	1	KYJ: NÕUSTUV EI	1	3,2%
Kokku	25		31	100%

Sarnaselt üleküsimisele on ka ümbersõnastamise korral võimalik, et ühele parandusalgatusele järgneb mitu paranduse läbiviimist. Tähtis on siiski mainida, et võimalik on ka vastupidine, st esineb mitu partneri parandusalgatust (nt ümbersõnastamist), millele järgneb üks läbiviimine, kuigi sellised juhtumid on harvad (analüüsitud materjalis ainult 1 juhtum).

Ümbersõnastamise moodustamiseks on mitmeid võimalusi (vt tabelit 7), mis on struktuurilt sarnased üleküsimise moodustamisele. Enamikel juhtumitel kasutatakse oleviku kindlat kõneviisi, vaid ühel juhtumil kasutatakse lihtmineviku vormi.

Tabel 7. Ümbersõnastamist algatavate dialoogiaktide lingvistiline struktuur⁸

Dialoogiakt	Lingvistiline struktuur	Juhtumite arv
KYE: VASTUST PAKKUV	sõna/fraas	5
	sõna/fraas + <i>jah</i>	7
	sõna/fraas + <i>või</i>	2
	lause	5
	lause + <i>jah</i>	2
	(<i>tähendab</i>) fraas + <i>jah</i>	1
	(<i>tähendab</i>) fraas/lause + <i>või</i>	2
KYE: SULETUD KAS	fraas + <i>jah</i>	1
Kokku		25

⁸ Kuna juhtumeid on niivõrd vähe, pole protsente arvatud.

4.3. Mittemõistmine

Mittemõistmise põhjuseks on, et infoandja või klient ei kuule või ei mäleta olulist osa suhtluspartneri eelmisest kõnevoorst selle infokülluse tõttu. Mittemõistmist võib põhjustada lause, sõna või hääldamisest tekkiv mitmetähenduslikkus. Teisalt ei ole mittemõistmiseks vaja mitmeti mõistetavat lausungit, piisab, kui üks kõneleja leiab, et teine ütles midagi ebatõest või mittesobilikku. Lisaks kasutatakse mittemõistmist mittekuulmise märkimiseks.

Tabel 8. Mittemõistmise algatamise sisulised põhjused

Dialogiakt	Põhjus	Juhtumite arv
PPE: MITTEMÕISTMINE	telefoni/faksi number	9
	institutsiooni nimi/üldnimi	7
	aadress/asukoht	3
	mittekuulmine/informatsiooni üleküllus	2
Kokku		21

Tabelis 8 esitatud teabele võib lisada, et mittekuulmine või informatsiooni üleküllus võib olla seotud ka muude esitatud põhjustega, kuid on eraldi välja toodud üksnes siis, kui juhtumit ei saa teiste tabelis 8 nimetatute hulka paigutada.

Kliendi jaoks on peamine probleemiallikas telefoni (või faksi) number. Seejuures esineb mittemõistmist siis, kui infoandja ütleb eelnevas kõnevoorus korraga (ilma tagasisidet ootamata) kogu numbrikombinatsiooni või enam kui kolm numbrit.

Probleemvoorule osutamise ja partneri algatatud paranduse viisid on järgmised.

- Infoandja või klient kordab eelmise voo lausungist osa, millest ta aru sai, viidates sellega, et soovib teada ülejäänut ehk mõistmatuks jäänud lausungi osa.
- (10) 1. V: üks üheksa kuus üks. | DIJ: INFO ANDMINE |
 2. (1.2)
 3. H: £ üks `üheksa? £ | KYE: AVATUD | | PPE: MITTEMÕISTMINE |
 4. (.)
 5. V: kuus? (.) [üks.] | KYJ: INFO ANDMINE || PPJ: LÄBIVIIMINE |

Dialogikatkes 10 järgneb infoandja telefoninumbri lausumisele (rida 1) pärast 1,2-sekundilist pausi (rida 2) kliendipoolne parandusalgatus (rida 3). Klient kordab tõusva intonatsiooni (märgend ?) ja aktsendiga (märgend £...£) esimest kahte numbrit, andes sellega mõista, et järgnevaid numbreid ta kas ei suutnud meelde jätta (nt info üleküllus) või ei saanud aru (nt ei kuulnud). Pärast mikropausi (rida 4) viib infoandja läbi paranduse (rida 5), andes kliendile puuduva teabe.

- Kui probleemiallikat peitev lausung koosneb mitmest erinevast n-ö elemendist (nt ettevõtte nimi + aadress; asukoht + ettevõtte nimi), korratakse seda probleemvoo elementi, mida mõisteti. Sellisel juhul oodatakse ainult selle elemendi piires puuduva info kordamist.

- (11) 1. H: £ [ee] tere. | RIJ: VASTUTERVITUS |
 2. (.) `tahtsin küsida `Tartus=e (.) `Kalda tie `kolmkümmend (.) `pleki `ukse `koda (.) kas te `saate mu anda. | KYE: JUTUSTAV KAS |
 3. (1.2) on sellised (.) {teil või} £ | KYE: SULETUD KAS |
 4. (4.2)
 5. V: ja `aadress oli `Kalda `tee? | KYE: AVATUD | | PPE: MITTEMÕISTMINE |
 6. (.)
 7. H: £ `kolmkend kui ma õieti `mäletan. £ | KYJ: INFO ANDMINE | | PPJ: LÄBIVIIMINE |

Näites 11 esitab klient pärast vastutervitust (rida 1) jutustava *kas*-küsimuse abil kolmest elemendist koosneva päringu (rida 2). Need elemendid on asukoht (*Tartu*), aadress (*Kalda tee 30*) ja institutsiooni üldnimi (*pleki-ukse-koda*). Infoandja küsimus käib ainult ühe elemendi, aadressi kohta (rida 5). Ta lokaliseerib oma mittemõistmise avatud küsimusega: *ja `aadress oli `Kalda `tee?*, oodates aadressist arusaamatuks jäänud osa kordamist antud elemendis. Klient teeb ootuspärase paranduse (rida 7), kus sarnaselt eelmisele näitele öeldakse üksnes elemendist puuduv osa.

- Probleemsele kohale ei viidata, vaid küsitakse otse. Sellistel juhtumitel on probleemiallikas väga selgelt piiritletud.

- (12) 1. H: tere. | RIJ: VASTUTERVITUS |
 2. (0.8) öelge (.) kus on Tartus ee `Kaa sa `long. | KYE: AVATUD |
 3. (1.5)
 4. V: e=`kuidas=se sa `long oli | KYE: AVATUD | | PPE: MITTEMÕISTMINE |
 5. H: `Kaa. | KYJ: INFO ANDMINE | | PPJ: LÄBIVIIMINE |
 6. Kaa sal[ong.] | KYJ: INFO ANDMINE | | PPJ: LÄBIVIIMINE |

Näites 12 koosneb kliendi avatud küsimus (rida 2) kahest elemendist: asukohast (*Tartu*) ja institutsiooni nimest (*Kaa salong*). Infoandja algatab mittemõistmise institutsiooni nime suhtes (rida 4), tehes seda antud kontekstis ja suulise kõne vaba vormi arvestades otsese küsimusega: *kuidas see salong oli?* Kuna dialoogis eelneb ainult üks kõnevoor, siis ei saa kahtlust olla, millele asesõna *see* osutab. Klient kordab infot kahes järgnevas lausungis (rida 5, 6). Esimeses lausungis parandab ta väga täpselt (kitsalt) mittemõistmist põhjustanud osa (*`Kaa.*), teises lausungis kordab tervet probleemset elementi (*`Kaa sal[ong.]*).

Mittemõistmine moodustatakse enamasti avatud küsimusena, kuid vahel ka soovina (vt tabelit 9). Paranduse läbiviimised on vormistatud informatsiooni andmisena. Seda nii juhul, kui mittemõistmise algatus on vormistatud avatud küsimusena, kui juhtudel, mil mittemõistmine on algatatud soovi näol.

Sarnaselt teistele partneri algatatud parandustele, võib ka mittemõistmise korral järgneda ühele parandusalgatusele mitu läbiviimise akti. Ilmselt on see inimsuhtlusele omane, et ei parandata ainult väärnimõistetut, vaid vahel jagatakse ka lisateavet või teavet mõne nüansi kohta, kus teine võiks sarnaselt eksida.

Pole ühte ja ainsat kindlat viisi, millega rääkijad väljendaksid mittemõistmist. Kuna vaatluse all on suhteliselt väike arv mittemõistmise juhtumeid, on keeleliste seaduspärasuste märkamine raskendatud.

Tabel 9. Mittemõistmise puhul kasutatavad parandusaktid (PPE: MITTEMÕISTMINE)

Algatus	Arv	Protsent	Parandus	Arv	Protsent
KYE: AVATUD KÜSIMUS	19	90%	KYJ: INFO ANDMINE	23	88%
DIE: SOOV	2	10%	DIJ: INFO ANDMINE	3	12%
Kokku	21	100%		26	100%

Nii nagu teised partneri algatatud parandused, käib ka mittemõistmine kas vahetult eelnenud või mõne varem eelnenud kõnevooru kohta. Erinevalt üleküsimisest ja ümbersõnastamisest kasutatakse mineviku vormi enam kui oleviku vormi (vastavalt 7 ja 3 juhtumit), samuti ei lisata lausungi ega kõnevooru lõppu abisõnu *jah*, *või*, *vä* vmt. Mittemõistmist väljendatakse suhteliselt lühidalt, st sõna, fraasi või lühikese lausega:

- (üineem +) küsisõna + sõna/fraas + *oli* (3 juhtumit);
- sõna (5 juhtumit), millest 3 on avatud tüüpi parandusalgatused (vt tabel 10);
- fraas (6 juhtumit);
- lause (7 juhtumit).

Mitmed fraasid ja laused on vormistatud nn lükkküsimusena ehk lähikontekstis (partneri lausungis) leidub kõneleja fraasis/lauses puuduv osa.

Avatud tüüpi (ingl *open class*) järgmise kõnevooru parandusalgatus (ingl *next turn repair initiation*) leiab kasutamist ainult viimase abinõuna, mil vajatakse kõnevooru täielikku uuesti esitamist, sest selline avatud liiki süntaktiline esitus eeldavalt vihjab või osutab, et kõneleja ei ole kuulnud eelmist kõnevooru või piisavalt aru saanud, et korrata midagigi sellest (Clark, Schaefer 1987: 29). Seega ei lokaliseeri avatud tüüpi parandusalgatus probleemi keset ning laieneb kogu eelnenud kõnevoorule. Tabelis 10 on loetletud käesoleva artikli raames uuritud mittemõistmistes leiduvad avatud tüüpi parandusalgatused. (Teistes partneri algatatud parandustes neid ei leidu.)

Tabel 10. Avatud tüüpi parandusalgatused

Parandusalgatuse vahend	Autori tulemused	K. Strandsoni tulemused
<i>ah</i>	2	0
<i>kuidas</i>	1	9
<i>mis asi</i>	1	0
<i>hakkame otsast peale</i>	1	0
Kokku	5	9

Infodialoogides esines 4 erinevat avatud tüüpi mittemõistmise parandusalgatust (*ah*; *kuidas*; *mis asi*; *hakkame otsast peale*). Materjali vähesed näited kinnitasid Leelo Keevalliku (1999: 49) väidet, et küsiva intonatsiooniga *ah* moodustab pea alati omaette kõnevooru ja tähendab 'mis sa ütlesid?'.

Krista Strandsoni (2002: 24) ametlikes telefonivestlustes aset leidnud avatud liiki parandusalgatuste loetelu on pikem (*mh*; *misasja*; *mida (misse)*; *kuidas*;

vabandust; oi ma ei kuule, kuidas palun; oi mitte ei kuule), kuid peale avatud paranduse *kuidas* ei esine samuti ühtegi teist parandusalgatuse vahendit üle 2 korra. K. Strandsoni ning käesoleva uurimistöö tulemusi võrreldes selgub tõsi-asi, et avatud tüüpi parandusalgatuste loetelu on suhteliselt lai, sest peale sõna *kuidas* kattuvaid avatud tüüpi parandusalgatusi ei esinenud. Siiski võib nõustuda K. Strandsoni (2002: 35) kokkuvõtva väitega, et “vestluskaaslase avatud tüüpi reformuleeringu algatus osutab vestluskaaslase kuulumisprobleemile, arusaamisraskusele või sellele, et eelmine kõnevoor oli vestluskaaslase jaoks mingil põhjusel ebaootuspärane.”

5. Järeldused ja kokkuvõte

Suhtlusprobleeme täiesti vältida pole võimalik. Need on üks osa loomulikust inimes-tevahelisest suhtlemisest. Inimesega loomulikus suulises keeles suhtleva dialoogsüsteemi edukaks töötamiseks on vaja täpselt teada, mis tekitab suhtlusprobleeme, kuidas probleeme vältida ja kuidas tekkinud probleemidega edukalt toime tulla. Just seetõttu on vajalik uurida inimestevahelisi vestlusi. Saades aru inimestevahelise suhtluse iseärasustest, on võimalik seda masinale modelleerida.

Käesolevas artiklis on vaatluse all 126 inimestevahelist suulist infodialoogi eesti dialoogikorpusest. Neist 60 dialoogis leidub üks või enam üleküsimise, ümbersõnastamise, mittemõistmise või muud (ehk süstematiseerimata) dialoogiakti, kokku 121 partneri algatatud parandust. Üleküsimine esineb 72, ümbersõnastamine 25 ja mittemõistmine 21 korda.

Fakt, et peaaegu pooltes dialoogides esineb üks või mitu partneri algatatud parandust, näitab et tegu on tõsise probleemiga dialoogsüsteemi loomise seisukohalt vaadates.

Kõigil partneri algatatud parandustel on omad tunnused, mille poolest nad üksteisele sarnanevad või erinevad. Paljud tunnused on infotelefonikõnede spetsiifilised, nt on iseloomulik numברי üleküsimine. Samuti on välja kujunenud mustrid, mil moel infoandja numbrit ütleb ja kuidas klient numbrit arusaamist väljendab. Samas on numברי ümbersõnastamine harv nähtus. Ümbersõnastamine on tihti seotud kliendi öeldud andmete viimisega infoandja andmebaasis kasutatavale kujule. Nendeks andmeteks on näiteks institutsiooni nimi/üldnimi ja aadress/asukoht. Mittemõistmist kasutatakse märkimaks, mida partneri öeldust mõistetakse ja mis osa ei mõisteta. Mittemõistmise põhjuseks võib olla näiteks mittekuulmine. Analüüsitud dialoogides põhjustab mittemõistmist enam telefoninumber ja institutsiooni nimi või üldnimi.

Materjalis esinevad kindlad keelelised struktuurid, kuidas üleküsimine, ümbersõnastamine ja mittemõistmine moodustatakse. Seega võib öelda, et kuigi inimes-tevaheline kõne on spontaanne ja raskesti etteaimatav, panevad institutsioonikõne, infodialoog ja telefonivestlus inimesed rollidesse, mis mõjutavad nende suhtlemist nii, et tekivad teatavad mustrid, kuidas suhtlusprobleemide tekkimisest partnerit teavitatakse.

Partneri algatatud parandused vajavad edaspidi veel uurimist. Nii on näiteks vaja kontrollida, kas käesolevas artiklis esitatud materjalist ilmnenu seaduspärasused, on kehtivad (ja samas vahekorras) ka suurema hulga dialoogide puhul.

Mark Fišeli dialoogikorpuse tööpink, mida on kasutatud käesolevas töös abivahendina, sisaldab ka automaatset dialoogiaktide märgendajat. Märgendaja töötab statistiliste printsiipide põhjal, mistõttu aktimärgendid ei pruugi alati õiged olla. Seega hetkel töötava märgendaja väljund ei sobi üldiseks kasutamiseks enne, kui inimene-ekspert pole tulemusi üle vaadanud. Edasises uurimistöös tuleb leida reegleid, mis selle automaatse märgendaja täpsust suurendaksid, sh ka partneri algatatud paranduste märgendamisel.

Transkriptsioonimärgid

H – klient

V – infoandja

(.) – mikropaus (0,2 sekundit või lühem)

(...) – mikropausist pikem paus

(1,2) – pausi pikkus sekundites

. – langev intonatsioon

`vaadake – terve sõna rõhutus või intonatsiooni tõus

sa`long – ebatavaline rõhuline silp

>...< (sissepoole osutavad nooled) – kiirendatud löik

<...> (väljapoole osutavad nooled) – aeglustatud löik

... – muust kõnest vaiksem löik

£...£ – aktsenti osutavad märgid

@...@ – hääletooni või hääle kvaliteedi muutumine (nt kellegi kõne jälgendamine)

.hhh – häälekas sissehingamine (üks h-täht vastab 0,1 sekundile)

[– pealerääkimise algus

] – pealerääkimise lõpp

= – kaks iseseisvat üksust on hääldatud kokku

{või} – loogiliste sulgude sees on halvasti kuulnud tekstilöik või kõneleja nimi

{--} – ebaselgeks jäänud sõna või kõneleja

{---} – pikem ebaselgeks jäänud löik

Kirjandus

Clark, Herbert H.; Schaefer, Edward F. 1987. Collaborating on contributions to conversations. – Language and Cognitive Processes 12, 19–41.

Ten Have, Paul 1999. Doing Conversation Analysis: A Practical Guide. London: Sage.

Hennoste, Tiit; Rääbis, Andriela 2004. Dialoogiaktid eesti infodialoogides: tüpoloogია ja analüüs. Tartu Ülikooli arvutiteaduse instituut. Tartu: Tartu Ülikooli Kirjastus.

Hennoste, Tiit; Gerassimenko, Olga; Kasterpalu, Riina; Koit, Mare; Rääbis, Andriela; Strandson, Krista; Valdisoo, Maret 2005. Questions in Estonian information dialogues: form and functions. – V. Matousek, P. Mautner (eds). Text, Speech and Dialogue. 6th International Conference TSD 2005. Berlin: Springer-Verlag, 420–427.

Hutchby, Ian; Wooffitt, Robin 1998. Conversation Analysis: Principles, Practices, and Applications. Blackwell Publishing Inc.

Keevallik, Leelo 1999. Informatsioonikäsitluse partikkel *ahah* telefonivestluses. – Emakeele Seltsi aastaraamat 43. Tartu: ETA Emakeele Selts, 34–56.

Schegloff, Emanuel 1968. Sequencing in conversational sequence. – Gail Jefferson (ed.). Lectures on Conversation. Vol. I. Oxford: Basil Blackwell, 3–11.

Sorjonen, Marja-Leena 1997. Korjausjäsenys. – L. Tainio (toim.). Keskusteluanalyysin perusteet. Tampere: Vastapaino, 111–137.

- Strandson, Krista 2000. Teise vooru reformuleeringud eesti vestlustes. Bakalaureusetöö. Käsikiri Tartu Ülikooli filosoofiateaduskonna eesti keele osakonnas. Tartu: Tartu Ülikool.
- Strandson, Krista 2001. Kuidas vestluskaaslane parandusprotsessi algatab? – Keel ja Kirjandus 6, 394–409.
- Strandson, Krista. 2002. Vestluskaaslaste algatatud reformuleeringud eesti vestlustes: reformuleeringualgatuse vahendeid ja põhjuseid. Magistritöö. Käsikiri Tartu Ülikooli filosoofiateaduskonna eesti keele osakonnas. Tartu: Tartu Ülikool.
- Wooffitt, Robin 1990. On the analysis of interaction: an introduction to conversation analysis. – Paul Luff, Nigel Gilbert, David Frohlich (eds). Computers and Conversation. Computers and People Series. London: Academic Press, 7–38.

Võrgumaterjalid

Eesti Dialoogikorpuse veebileht. <http://math.ut.ee/~koit/Dialoog/EDiC.html> (21.11.2007).

Eesti Dialoogikorpuse tööpink. <http://math.ut.ee/~treumuth/> (08.08.2007).

Siiri Pärkson on lõpetanud Tartu Ülikooli eesti kirjanduse kirjanduse erialal. Kaitses 2007. a humanitaarteaduste magistrikaadi arvutilingvistika erialal. On uurinud loovust ja suulises suhtluses kasutatavaid partneri algatatud parandusi.
siiri_parkson@hotmail.com

CLARIFICATION, REFORMULATION AND NON-UNDERSTANDING IN PHONE CONVERSATIONS

Siiri Pärkson

University of Tartu

It is impossible to completely avoid communication problems; they are a part of natural interpersonal communication. For a dialogue system to successfully work with people in natural spoken language, it is important to know what causes these problems, how to avoid them and how to deal with the problems that have appeared. Because of this it is important to study interpersonal conversations. Understanding of the peculiarities of natural interpersonal communication helps us to model artificial dialogue systems.

There are 126 dialogues from the Estonian dialogue corpus under investigation in the present paper. In 60 of these, there are one or many clarifications, reformulations, non-understandings or other (unclassified) dialogue acts, a total of 121 partner-initiated repairs. Clarifications appear 72 times, reformulations 25 times and non-understandings 21 times.

Each type of partner-initiated repair has a set of features which makes us possible to compare them against each other. Many of these features are specific to information telephone calls. Thus, for example, clarification of numbers is very typical to information dialogues. Also, there are certain patterns in which way a number is spoken by information provider and how the client expresses his/her understanding in it. At the same time, reformulation of numbers happens very rarely. Instead, it is usually connected with conversion of information coming from clients to representation used by a database. Examples of such information are institutional name/general name and address/location. Non-understanding is used to indicate which part of information is understood and which part is not. In the analysed dialogues, telephone numbers, institutional names or general names introduce most non-understandings.

There are certain language structures which are used to make up clarification, reformulation and non-understanding repairs. Hence it can be said that even though a human speech communication is very spontaneous and hard to predict, factors like institutional speech, information dialogue and telephone call put people into roles which introduce a strong impact on the flow their communication. Moreover, certain patterns appear to inform a partner about problems of communication.

Keywords: computational linguistics, conversation analysis, interaction in institutional settings, spoken language, telephone conversation, information dialogues, clarification, reformulation, non-understanding, Estonian

LADINA ÕIGUSTERMINITE VORMILISEST JA SISULISEST KORREKTSUSEST

Merike Ristikivi

Ülevaade. Õigusteadus on oma olemuselt tekstiteadus, milles väga suur tähtsus on keelelistel väljendusvahenditel. Juriidilise jõu saavad sõnad ja väljendid siin vahetult keele kaudu. Käesoleva artikli uurimismaterjal pärineb Eesti ainsa õigusteadusliku ajakirja *Juridica* 1993–2007 aastakäigu artiklitest. Uurimuses käsitletav periood on olnud Eesti õigusele murranguline aeg: tagasi on pöördutud lääneliku õigusruumi, mis lähtub ladinakeelsest baasist. Õigusreformiga on kaasnenud ka muudatused Eesti juristide terminikasutuses, lülitumine Euroopa õiguskultuuri on märgatavalt suurendanud ladina terminite kasutamist. Samas kaasnevad võõrkeelendite tarvitamisega mitmed praktilised probleemid. Artikkel keskendub ladina õigusterminite ortograafiale ja selle eripärale ning morfoloogiale ja selle suhtele eesti-keelse lausega.

Võtmesõnad: õiguskeel, terminoloogia, keelekontaktid, keelevead, ortograafia, morfosüntaks, ladina keel, eesti keel

Sissejuhatus

Õiguskeele ja -terminoloogia analüüsid viidatakse sageli juriidiliste tekstide raskepärasusele ja keerukale sõnastusele ning rõhutatakse tarvidust koostada materjalid selges, kergesti jälgitavas, ühemõttelises ja arusaadavas keeles (Mellinkoff 1963, Tamm 1997, Oksaar 1999, Asprey 2003, Kukk 2003). Selge ja mõistetav tekst ei tohi samas olla ebapädev ja ülelihtsustatud. Juriidilised terminid kannavad ja vormistavad oma eriala mõistestikku, nad on aine objekti ja sisu põhilised väljenduselemendid (Erelt 1982: 26, Kull 2000: 95) ning sellistena õigusteaduse keele lahutamatu osa. Selle teemaringi keskmes on ka ladina keelest pärit terminid, mida õiguskeeles kasutatakse tsitaatsõnadena. Võib küsida, kas tänapäevases õiguskeeles ladina terminid pigem täpsustavad mõtet ja abistavad lugejat või muudavad teksti mõistetamatuks?

Henn Saari (1980: 741) märgib, et keeltevahelise terminoloogilise ühtsuse moodustab ühtne sisuplaan, ühtne mõistevõrgustik. Mandri-Euroopa õigus on arenenud ladina keele baasil ning lähtub Rooma õigusel põhinevast mõistete süsteemist.¹ Ladina keel on olnud tihedas ajaloolises seoses Euroopa õiguse arenguga, peamine hulk õigusteadusealasest kirjandusest, aga ka õigusaktid on koostatud möödunud sajandite jooksul ladina keeles (Ristikivi 2005: 199–200). Ladina terminite kasutamist soodustab samuti nende lühidus ja keeleline ökonoomsus.² Tihti ei ole võimalik nappi ladina terminite sõna-sõnalt tõlkida ning vasted võivad ulatuda mitme lauseni.³ Täpselt formuleeritud ladina terminid muudavad hõlpsamaks ka juristide rahvusvahelise kommunikatsiooni ning võimaldavad efektiivselt suhelda erialaliselt keele- ja kultuuripiiridest sõltumata (Benke, Meissel 1997: 10). Samuti osutatakse, et ladina väljendite kasutamine hõlbustab Euroopa kohtusüsteemi ühendamist ning muudab juriidilise kirjanduse rahvusvaheliselt mõistetavaks (Knütel 1994: 251).

Seetõttu ei oleks ladina terminite täielik eemalejätmine õigustekstist põhjendatud ja otstarbekas. Selgus ja arusaadavus kujunevad probleemiks siis, kui termineid ei kasutata õigustekstis korrektselt. Selles kontekstis annab ladinakeelsete terminite kasutamine ja sellega kaasnev probleemistik võimaluse hinnata juristide ja õiguskultuuri taset, sealhulgas ka õigusharidust. Millise ettevalmistuse alusel Eesti juristid ladina termineid kasutavad? Artiklis uuritavat perioodi 1993–2007 – Eesti Vabariigi astumist Euroopa Liitu ja sellele eelnevat 15 aastat – iseloomustab Eesti õigustegevuses väljumine nõukogulikust õigusest ja liitumine lääneliku õigusruumiga, rahvusvaheliste õigusaktide uurimine ning kohandamine oma riigi seadustega. Selle aja jooksul on koostatud ja avaldatud uued õpikud pea kõigi õigusharude kohta, mis tutvustavad juuraüliõpilastele Euroopa õigustraditsioone ning koos sellega erialase terminikasutuse põhimõtteid. Põhiline osa ladina juriidilisest terminoloogiast omandatakse õigusteaduse õppekavasse kuuluvate juriidilise ladina keele kursuse ja Rooma eraõiguse aluste kursuse käigus. Teisalt kuulub peamine osa Eesti ainsa õigusteadusliku ajakirja *Juridica* autoritest põlvkonda, kui klassikalisi keeli uuriv eriala Eesti ülikoolide ametlikes õppekavades puudus.⁴ Rooma õigust küll õpetati, kuid selle valdkonna uurimistraditsioon oli katkenud. Seega on paljude praeguste õigustekstide autorite teadmised ladina keele grammatikareeglitest napid ja juhuslikud. Seda tausta arvestades on arusaadav, et ühelt poolt on uus aeg toonud kaasa euroopaliku väljenduslaadi ja ladina terminite rohkuse, teiselt poolt aga ei ole praktiline ettevalmistus nende korrektseks kasutamiseks alati piisav.⁵ Nii kaasnevad võõrterminite kasutamisega eksimused nii ortograafia ja morfoloogia reeglite vastu kui ka süntaktilisel ja semantilisel tasandil.

¹ Ladina termineid kasutatakse suurel hulgal ka angloameerika õigussüsteemis, ehkki selle õigusmõisted ja sõnavara ei ole otseselt välja kujunenud Rooma õiguse alusel, vaid arenenud enamasti iseseisvalt kohtupraktika alusel (Mattila 2002: 208–210). Poleusemiast nende kahe õigusühema terminite vahel vt Ristikivi 2007.

² Tavaliselt esinevad ladina juriidilised terminid üksiksõnadena – lihttõlvisõnadena ja liitsõnadena, nt *usus* 'kasutamine, kasutusõigus', *usufructus* 'kasutusvaldus', või sõnaühenditena, nt *bona fides* 'hea usk, heauskus'. H. Saari (1980: 741) järgi on termin nomineeriv keeleüksus, kuid see ei pea olema nimisõna. Nii on ladina juriidilisteks terminiteks substantiivide kõrval ka verbid, omadus-, ase-, arv- ja määrsõnad, nt *non liquet* 'asi ei ole selge, pole tõendatud', *bilateralis* 'kahekülgne', *ad hoc* 'selleks, kindlaks juhtumiks', *privatim* 'eraviisiliselt', *bis* 'kaks korda, topelt' jne. Reeglina on teised sõnaliigid peale nimisõnade üldkeele varamusse kuuluv aines, mis erialases kontekstis on saanud spetsiifilise tähenduse.

³ Näitena võib tuua termini *ius alluvionis*, mida eesti keelde tõlgitakse 'juurdekasvuõigus veekogu uhte puhul'. Väljend tähendab, et maatüki omanikul on õigus ka mere või jõe poolt tema krundile kantud pinnasele (Adomeit jt 2005: 147).

⁴ Klassikalise filoloogia eriala suleti Tartu Ülikoolis 1954. ning taasavati 1990. aastal (Lill 2003: 11).

⁵ Ehkki terminikasutus peegeldab õigushariduse taset, tuleb ajakirja puhul tähele panna toimetaja rolli. Ajakirjal *Juridica* on tegevtoimetaja kõrval ametis ka keeletoimetaja, kuid ei ole selge, kas nende ülesandeks on ka kõigi artiklites esinevate võõrterminite (ladina väljendite kõrval leidub ka inglise-, prantsuse- ja saksaakeelseid termineid) õigekirja kontrollimine.

Eestis on ladina väljendite õigekirja küsimusi põgusalt kajastanud Anne Lill (1996), Erki Silvet (1997) ja Peeter Kask (Kask, Lill 1997). Ka allakirjutaja on tähelepanu juhtinud üldisematele probleemidele õigusterminite kasutamisel (Ristikivi 2003, 2005, 2007a, 2007b). Põhjalikumat analüüsi varem sellel teemal ilmunud ei ole. Käesolev artikkel püüab latinisti pilguga vaadelda õigusteaduse tekstides esile tõusvaid küsimusi ja leida neile vastus: millised ladina terminid valmistavad eesti juristidele kasutusraskust, milline on ladina tsitaatide vorm eesti lauses ning õigusteksti kvaliteet.

Uurimismaterjal

Õiguse keelekasutuses eristatakse õigusteaduse oskuskeelt ja seaduskeelt (Oksaar 1999: 204, Mereste 2000: 412–415, Kerge 2002: 37–38). Õigusteaduse seadus- ehk õigusaktide keelt kasutatakse ühiskonnale üldtähtsate probleemide lahendamisel ja teadvustamisel. Õigusteaduse oskuskeeles käsitletakse teaduslikult õigusprobleeme, kirjeldatakse ja analüüsitakse seaduskeeles väljendatut (Mereste 2000: 415). Käesoleva artikli uurimisobjekt on õigusteaduse oskuskeel, uurimismaterjaliks on õigusteaduslikud analüüsid ja kommentaarid ajakirjas *Juridica*.⁶ Seadustekste ning ladina õigusterminite problemaatikat Eesti seadustes otseselt uuritud ei ole. Seetõttu on õiguskeele mõiste piiritletud siin õigusteaduse keelega.

Juridica on Tartu Ülikooli õigusteaduskonna ajakiri, millest on nüüdseks välja kasvanud üle-eestiline juristide ajakiri. Ajakirjas on avaldatud enamik Eesti õigusteaduslikest publikatsioonidest (Varul 2002: 209). *Juridica* kui lingvistilise uurimisobjekti muudab representatiivseks kirjutajate lai ring. Aastate jooksul on ajakirjas artikleid avaldanud lisaks õigusteadlastele ja tegevjuristidele mitmete eluvaldkondade esindajad. 2001. aastaks, s.t üheksa ilmunisaasta jooksul oli *Juridica* autorite arv kasvanud 360 kirjutajani (Pruks 2001: 56). Teemade ringis on viieteistkümne aasta jooksul (1993–2007) olnud Eesti avaliku ja eraõiguse kõrval rahvusvaheline õigus ja Euroopa Liidu õigus, aga ka õigusteooria, õigusajalugu, õigusfilosoofia. Õigusajaloo- ja Rooma õiguse alaseid artikleid, mis juba oma loomu poolest sisaldavad rohkem ladina keelest pärit ainet, on uuritava materjali hulgas siiski küllalt vähe. *Juridica* numbrid on kas tervikuna pühendatud kindlale teemale või esitatakse läbilõige erinevatest juuravaldkondadest: käsitletakse õigusloome hetkeprobleeme ning antakse teoreetilis-praktiline kommentaar seaduste rakendamisel kerkinud küsimustele, sealjuures on kommenteerijateks enamasti seaduste autorid ise.

Uurimus hõlmab ajakirja *Juridica* aastakäike 1993–2007, kokku 146 numbrit, 968 artiklit ja 5879 lehekülge.⁷ Uurimusest on välja jäänud alates 1996. aastast kord aastas ilmuv ingliskeelne *Juridica International*, kuna eesmärgiks on uurida eelkõige eesti õiguskeelt.

Artiklites on 784 erinevat ladinakeelset terminit. Üldse esineb ladina termineid 4756 korral, seega keskmiselt 30 väljendit ühe ajakirja numbriga kohta ning 4–5 väljendit artikli kohta. Lehekülgede arvu jagamine väljendite arvuga näitab, et ladina keelt kasutatakse keskmiselt *Juridica* igal teisel leheküljel. Ladina terminite kasutamine on aastate jooksul järjepidevalt tõusnud: 1993. aastal esines artiklites kokku 66 ladina terminit, s.t keskmiselt 1,08 terminit ühe artikli kohta. Võrdlusena

⁶ Vt ka www.juridica.ee. 1994. aastal liitus *Juridica* ajakirjaga Eesti Jurist (1990–1994), mis vahendas peamiselt õigusinformatsiooni ja oli välja kasvanud ENSV Justiitsministeeriumi bülletäänist Nõukogude Õigus (1967–1989).

⁷ Esimesel, 1993. aastal ilmus 6 numbrit, aastakäigud 1994–2007 sisaldavad igaüks 10 numbrit.

seitse aastat hiljem, 2000. aastal kasutati juba 252 terminit, s.t keskmiselt 2,9 ühe artikli kohta. Viimastel aastatel on artiklites esinenud ladina termineid juba 450 korral, s.t keskmiselt 5,3 ühe artikli kohta.⁸

Ladina terminite kasutamine sõltub peamiselt õigusharu ajaloolisest taustast. Nii näiteks on ladina terminid pea alati õigusteooria, õigusfilosoofia, kriminaalõiguse, rahvusvahelise õiguse, lepinguõiguse ja pärimisõiguse artiklites. Kõigi nimetatud valdkondade terminoloogia on suuremas osas kujunenud antiik- või keskajal. Vähesel määral esineb ladina termineid tööõiguse, perekonnaõiguse ja äriõiguse artiklites. Ladina keele nappust võib siin seletada asjaoluga, et nimetatud õigusvaldkondade kujunemine ja uurimine on toimunud peamiselt 20. sajandil. Seega side Rooma õigusega, millest on käibesse läinud enamik ladina õigustermineid, neil puudub või on kaudne.

Kõige sagedamini *Juridicas* esinevad ladinakeelsed väljendid on esitatud tabelis 1. Väljendid nagu *expressis verbis*, *de facto* ja *de iure* esinevad artiklites mõnikord üldises ja neutraalses tähenduses, samas on autorid neid rakendanud ka kitsamas juriidilises mõttes. Peamiselt kasutatakse *Juridicas* ladina keelest pärit materjali aga termini tähenduses normatiivsete argumentidena ning see kannab spetsiifilist juriidilist informatsiooni.

Tabel 1. Ajakirjas *Juridica* perioodil 1993–2007 kõige sagedamini esinevad ladinakeelsed terminid

Ladinakeelne termin	Tähendus ⁹	Esinemisarv
<i>corpus iuris</i>	'õiguskogumik'	243
<i>lex mercatoria</i>	'kaubandusõigus'	145
<i>de lege ferenda</i>	'välja antava seaduse järgi'	128
<i>expressis verbis</i>	'sõnaselgelt, rõhutatult'	127
<i>culpa in contrahendo</i>	'lepingueelne vastutus'	83
<i>lex fori</i>	'kohtu asukohamaa seadus'	71
<i>de facto</i>	'faktiliselt'	68
<i>forum</i>	'kohus'	64
<i>de lege lata</i>	'kehtiva õiguse kohaselt'	63
<i>pacta sunt servanda</i>	'lepingutest tuleb kinni pidada'	47
<i>erga omnes</i>	'kõikide suhtes'	46
<i>lex specialis derogat generali</i>	'eriseadus tühistab üldise (seaduse)'	42
<i>ius cogens</i>	'imperatiivne õigus, peremptoorne norm'	39
<i>nullum crimen nulla poena sine lege</i>	'ei ole süütegu, ei ole karistust ilma seaduseta'	38
<i>praeter legem</i>	'seaduse kõrval'	32
<i>de iure</i>	'juriidiliselt'	28
<i>obiter dictum</i>	'muuhulgas öeldu; juhuslik märkus'	28
<i>in dubio pro reo</i>	'kahtluse korral süüdistatava kasuks'	25
<i>bona fides</i>	'hea usk, heauskus'	22
<i>ne bis in idem</i>	'ärge olgu kaht (karistus) sama asja eest'	19

⁸ Ladina keele ja Rooma õiguse vastu on huvi kasvanud ka mujal Ida-Euroopas ning Venemaal (Mattila 2006: 138). Näiteks võib tuua W. Wolodkiewiczzi uurimuse, mille kohaselt on Poola kohtumaterjalides ladina õigusterminite kasutamine viimastel kümnenditel märgatavalt suurenenud (Wolodkiewicz, Krzynyówek 2001).

⁹ Siin ja edaspidi autori tõlge.

Ladina terminid ja õigusteksti kvaliteet

Õiguskeele uurijad on tõdenud, et õigusteksti keeleline kvaliteet ei vasta hea keele normidele, juriidilise teksti laused on sageli väga keerulise struktuuriga ning lugeja peab teksti mõistmiseks tugevalt keskenduma (Kukk 2003: 100–102). Samuti märgitakse, et õigustekst koormab lugeja mälu pikkade täiendfraasidega või üksikesse paigutatud lausetega, mida pole võimalik mõista enne, kui jõutakse põhisonani. Kui mõistmiseks olulist teavet sisaldub palju täiendites, kõrvallauses ja lauseühendites, on terviklause mõttest raske aru saada (Kasik 1995: 19). Õigusteksti kvaliteeti mõjutab ka juriidiliste võõrterminite kasutamine ebakorrektses seoses ja tähenduses.

Eeltoodud arvandmete alusel võib märkida, et juristidele meeldib ladina keelt kasutada. Materjali uurimisel torkab aga silma, et autorid jätavad enamasti tekstis kasutatud ladinakeelse väljendi juurde märkimata eestikeelse vaste. Õiguskeele uurijate sõnul seisneb oskuskeele problemaatilisus muuhulgas selles, et väljendi õigeks kasutamiseks on vajalikud teadmised asja ja situatsiooni kohta ning oskussõnade kasutamisel on vaja arvestada ka vastuvõtja arusaamisega erialakeelest (Oksaar 1999: 202). Ehkki *Juridica* on teadusväljaanne, peetakse sihtrühmaks nii Eesti juristkonda kui laiemat huviliste ringi (Varul 1993: 2). Tihti võib aga *Juridica* artiklitest leida harva esinevaid või spetsiifilist informatsiooni sisaldavaid ladinakeelseid väljendeid, mis ilma juriidilise tausta või kommentaarideta ei pruugi olla täielikult arusaadavad. Sageli ei piisa ka heast ladina keele tundmisest, et neid väljendeid õigesti mõista. Seega võib tõlke puudumisel kergesti kujuneda olukord, et erialakaugem lugeja tõlgendab väljendeid ja termineid teisiti kui vastava ala spetsialist. Ladinakeelne termin võib põhjustada mõistmisraskusi ja arusaamatust ka juristidele endile. Tänapäeva õigusteadus on sedavõrd spetsialiseerunud, et tsiviilõiguse tundja ei tarvitse teada, mida üks või teine termin tähistab nt karistusõiguses. Võib isegi juhtuda, et terve lause jääb ebaselgeks, kui mõnda ladina terminis leiduvat sõna ei mõisteta (nt 1–3).

- (1) “Seaduslikkuse ehk legaliteedipõhimõte toetub EIÕK (Euroopa Inimõiguste Konventsioon) artiklile 7 ning tähendab analoogiakeeldu *malam partem* ja tagasiulatuva jõu keeldu koos nõudega *lex mitior*.” (Sootak 1999: 178)
- (2) “Õiguse rakendamisel ei saa kohustuslikuna kehtida põhimõte *singularia non sunt extendenda*.” (Luts 1996: 350)
- (3) “Samuti mõisteti juba tol ajal *jus gentium*’i all selliseid rahvusvahelisi käitumishorme, nagu *rerum repetitio* ja sellest keeldumise korral sõjakuu-lutamist.” (Uibopuu 1998: 189)

Näidetes 1–3 esinenud ladina terminitest on tuntum ainult *jus gentium* (= *ius gentium*) ‘rahvaste õigus’ (3). Termineid *malam partem* ‘halbemata osa, halvemas osas’, *lex mitior* ‘leebem seadus’ (1), *singularia non sunt extendenda* ‘üksiksätteid ei tohi laiendada’ (2), *rerum repetitio* ‘asjade tagasinõudmine’ (3) on ajakirja *Juridica* kümne aastakäigu numbrites kasutatud ainult üksikutel kordadel. Ilma tõlketa võib lausete 1–3 tähendus jääda ebaselgeks. Muidugi pole termini laiema tuntusega arvestamine (ja seega tõlkevaste ärajätmine) alati üheselt määratletav.

Siiski võiks kahtluse korral ja parema mõistetavuse huvides omakeelse vaste pigem lisada kui ära jätta.

Tõlkevaste ärajätmisega samaväärset arusaamatust võivad põhjustada ka laused, milles ladina termin on tõlgitud, kuid esitatud eestikeelne vaste ei ole täpne (nt 4).

- (4) “*Ratio legis est anima legis* – seaduse mõte on seaduse asi.” (Narits 1994: 229)

Ladina *anima* tähendab ‘hing, olemus’, seega oleks õigem tõlkida ‘seaduse mõte on seaduse hing/olemus’. Näites 5 on ekslik vaste antud tuntud õigusteaduslikule printsibile, mille korrektne tõlge eesti keeles on ‘hea usk, heausksus’.

- (5) “*Bona fides*, eesti keeles truudus ja usk, tuleneb otseselt õigusideest.” (Uibopuu 1998: 189)

Näites 5 on tsitaadi autor vaste esitamisel lähtunud heausksust tähistavast saksa õigusterminist *Treu und Glauben*.¹⁰ Seega ei ole lauses 5 otsest viga, kuid eksitab terminoloogiliselt. Samuti osutab selline juhtum asjaolule, et Eesti juristide erialakeelt mõjutavad tugevalt võõrkeelsed materjalid.

Esmärgipärasele selgusetulelusele ei aita samuti kaasa teksti ilma tõlkevasteta paigutatud ja sisult sobimatu ladina termin. Näites 6 on ladina terminit kasutatud direktiivi tõlkes.

- (6) “Müüjaks või tarnijaks loetakse a contrario iga füüsilist või juriidilist isikut, kes tegutseb eesmärkidel, mis on seotud tema põhitegevuse, ettevõtluse või ametiga.” (Ginter 2001: 502)

Ladina termin *a contrario* tähendab ‘vastaspoolelt; vastupidiselt’ ning lause mõte jääb praegusel kujul ebaselgeks. Selguse toovad direktiivi originaaltekst (6a) ja eestikeelne justiitsministeeriumi ametlik tõlge (6b).

- (6a) “(c) ‘seller or supplier’ means any natural or legal person who, in contracts covered by this Directive, is acting for purposes relating to his trade, business or profession, whether publicly owned or privately owned.”¹¹

- (6b) “c) müüja või teenuste osutaja – füüsiline isik, kes käesoleva direktiiviga hõlmatavate lepingute raames toimib eesmärkidel, mis on seotud tema kaubandus-, majandus- või kutsetegevusega, olenemata sellest, kas ettevõtte on era- või riigiomanduses.”¹²

Täpsem oluks selles kontekstis kasutada ladina terminit *ex contractu* ‘lepingust tulenevalt, lepingu järgi’, mis oleks kooskõlas nii eestikeelse teksti sisu kui ka direktiivi ingliskeelse originaali mõttega.

Ladina väljendite sobimatule kasutamisele on viidatud erialakirjanduses ka varem. Ajakirjas *Õiguskeel* on ilmunud artikkel meditsiinitermini *restitutio ad integrum* ‘terviku juurde viima; meditsiinis: terveks tegema’ ekslikust kasutamisest juriidilises kontekstis, kus korrektsem oleks õigusväljend *restitutio in integrum* ‘taastamine tervikuna, endise olukorra ennistamine’ (Kask, Lill 1997).

¹⁰ Vt nt Meyer 2002: 52.

¹¹ Vt <http://eur-lex.europa.eu/Notice.do?val=294482:cs&lang=en&list=294482:cs,&pos=1&page=1&nbl=1&pgs=10&hwords=&checktexte=checkbox&visu=#texte> (22.02.2008). 31993L0013. Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts; Official Journal L 095, 21/04/1993 p. 0029–0034.

¹² Vt <http://eur-lex.europa.eu/Notice.do?mode=dbl&lang=en&lng1=en,et&lng2=bg,cs,da,de,el,en,es,et,fi,fr,hu,it,lt,lv,mt,nl,pl,pt,ro,sk,sl,sv&val=294482:cs&page=1&hwords=> (22.02.2008).

Oskussõnadena ei ole põhjust ladina termineid vältida, kui selle taga on püüe eristada ja määratleda täpselt mingi nähtuse olemust. Terminite kasutamine hõlbustab erialainimeste omavahelist suhtlemist ja mõistmist. Ent arvestades Juridica sihtrühma, kelleks juristide kõrval on ka teiste erialade esindajad, oleks otstarbekas lisada võõrterminile ka omakeelne ja korrektne vaste. Korrektnete terminikasutuse ja tõlke esitamise eelduseks omakorda on, et autor ise tunneb täpselt termini tähendust.

Ortograafiavead

Ladina terminite ortograafia- ehk õigekirjavigu esineb Juridica artiklites sageli. Ortograafilised variandid ja vead võivad omakorda muuta morfoloogilist vormi. Eriti palju leidub ladina terminite käändevormi- ja ühildumisvigu, aga ka lihtsalt tähevigu. Läbi aastate korduvad vead ja ebasoovitavad konstruktsioonid omakorda osutavad, et õiguse keelekasutus omandatakse eelkõige olemasolevate tekstide kaudu ja toel. Juridica andmete põhjal valmistab probleeme nii ladina terminite kokku- ja lahkukirjutamine, kui ka vokaali või konsonandi õigekiri. Siinjuures ei saa aga väita, et vead puudutaksid eelkõige vähe kasutatavaid ja mittetuntud termineid. Vääräl kujul võib artiklitest leida ka väljendeid, mis arvandmete järgi esinevad sageli ja peaksid kuuluma juristide üldsõnavara hulka.

Ladina õigusväljendite puhul on levinud ja lubatud sõnade ja prepositsioonide kokku- ja lahkukirjutamisel tekivad variandid, nt *iuris prudentia* ja *iurisprudencia* 'õigusteadus', *sub poena* ja *subpoena* 'karistuse ähvardusel' jpt (Ristikivi 2007a: 263). Verbi ja käändsõna ühend, nt Juridica artiklist leitud *respondeatsuperior*, tuleb aga kindlasti lahku kirjutada: *respondeat superior* 'kõrgem (isik) vastutagu'. Lahku tuleb kirjutada ka ühe sõnana esinev ladina eituspartikkel, arvsõna ja eessõna, Juridicas nt *nebisin idem pro ne bis in idem* 'mitte kaks korda sama asja eest'.

Vokaalide puhul aetakse tihti segi sarnase kirja- ja tähed *a* ja *o*, kuid esineb ka muid tähevigu, nt *ultima ratio pro ultima ratio* 'viimne vahend'; *culpa levissima pro culpa levissima* 'kergeim hooletus'; *sine qua non pro sine qua non* 'ilma milleta ei';¹³ *sensus verborum est anima legis pro sensus verborum est anima legis* 'sõnade mõte on seaduse hing'; *vacatio legis pro vacatio legis* 'seaduse puudumine'.¹⁴ Viimases näites muutub eksimuse puhul ka termini tähendus: *vacatio* tähendab eesti keeles 'kutsumine'. Konteksti järgi otsustades ei olnud siiski artikli autori kavatsus terminit selles muutunud tähenduses kasutada.

Mitmetel juhtudel põhjustab autori ebakindlust väljendi kirja- ja tähed *i* osas ja eksimist õigekirjas õigustermini hääldus. Alates 19. sajandi lõpust on ladina keele hääldamisel lähtutud kahest suunast.¹⁵ Nn klassikaline suund järgib antiikajal levinud hääldust, mille puhul sõnade hääldus ei erine reeglina kirja- ja tähed (*ti* [ti], *c* [k]). Teine, traditsiooniline suund lähtub hilisladina perioodil umbes 4. sajandiks välja kujunenud hääldusest, mis sai aluseks ka ladina keelest arenenud romaani keeltele. Selle järgi erineb teatud häälikute ja häälikuühendite hääldus kirja- ja tähed (nt *c* [ts], *ti* [tsi]). Eesti kultuuriruumis soovitatakse klassikalist hääldust juhul, kui õpitakse ladina keelt antiikaja ja antiikautorite tekstide tundmaõppimiseks. Kui keskendutakse aga juura või meditsiini terminoloogiale, on otstarbekam järgida traditsioonilist hääldust, mille järgi on parem jälgida ka seoseid ladina keele ja

¹³ Lühenenud vorm terminist *conditio sine qua non* 'tingimus, ilma milleta ei; vältimatult vajalik nõue, ilma mille järgimiseta ei saabu soovitatav õiguslik tagajärg' (Adomeit jt 2005: 63).

¹⁴ Ajavahemik seaduse väljakuulutamise selle kehtima hakkamiseni (Adomeit jt 2005: 258).

¹⁵ Eri põhimõte on angloameerika õigussüsteemi juristidel, kes hääldavad ladina keelt inglise keele hääldusreeglite järgi (Kelly 1988: 196–200). Siiski ei ole selline hääldus Eesti juristide seas veel levinud ega avalda mõju ladina terminite kirja- ja tähedile.

võõrsõnade vahel (*obligatio* – obligatsioon, *ius civile* – tsiviilõigus) (Adomeit jt 2005: 13).

Juridica artiklites esinenud konsonantide vead puudutavadki eelkõige tähti, mille hääldus kirjapildist erineb ning sarnaneb mõne teise tähe hääldusega. Häälduselt on sarnased ja valmistavad rohkem raskust *c* ja *s*, *q* ja *g*, *q* ja *c*, *t* ja *d*, nt *ius est ars boni et aequi* pro *ius est ars boni et aequi* 'õigus on teadus heast ja õiglasest'; *lucrum sessum* pro *lucrum cessans* 'saamata jäänud tulu'; *numerantur sententiae*, *non ponderander* pro *numerantur sententiae*, *non ponderantur* 'hääli loendatakse, mitte ei kaaluta'; *qui prodest* pro *cui prodest* 'kellele on (see) kasulik'. Viimases väljendis esineva vea tulemusel muutub ka väljendi tähendus: *qui prodest* tuleks tõlkida eesti keelde 'kes või mis on kasulik', asesõna on siin nominatiivis; originaalväljendis on asesõna aga daativis.

Tsitaatsõnade ja -väljendite kasutamismeel on kirjutada neid nagu lähtekeeles (Erelt 1999: 16). Juridica algusaastate artiklitest võib aga leida ekslikke vorme, milles autorid on termini tähistamisel lähtunud ladina keele häälduspõhimõtetest ning märkinud terminisse muutuse, kus hääldus erineb kirjapildist: *c* [ts] ja *x* [ks]. Näiteks *ekspressis verbis* pro *expressis verbis* 'sõnaselgelt'; *lex spetsialis* pro *lex specialis* 'eriseadus'; *sine periculo sotsiali* pro *sine periculo sociali* 'ilma ühiskonnaohtlikkusest'. Samas esinevad need terminid jällegi arvandmetes väga tihti (vt tabel 1). Seega ei saa seostada termini tuntuust ja korrektset kasutust.

Soo- ja käändevormid ning eessõnad

Lisaks õigekirjavigadele leidub Juridica artiklites eksimusi ladina õigusterminite käände- ja soovormides ning eessõnade kasutamisel. Väljendite vead osutavad tüüpilistele raskustele, mis tekivad ladina keele kasutamisel eesti keeletausta mõjul. Näiteks tekitab probleeme soovormide moodustamine ja ühildamine, sest soo kategooria eesti keeles puudub.

Ladina keeles ühilduvad soos nimi- ja omadussõna. Ladina sõna *mos* 'tava, komme' on meessoost ning ühildub omadussõna meessoovormiga. Juridica artiklist leitud väljendis on aga ekslikult kasutatud omadussõna naissoovormi: *bonae mores* pro *boni mores* 'head tavad'. Palju eksitakse termini *ius* 'õigus' kasutamisel. See sõna on ladina keeles kesksõost ning eeldab kesksõost omadussõnalise täiendi kasutamist. Mitmetes artiklites seevastu on kasutatud omadussõnadel mees- ja naissoo lõppe: *ius naturalis* pro *ius naturale* 'loomuõigus'; *ius commutativa* pro *ius commutativum* 'võrdsustav õigus';¹⁶ *ius distributiva* pro *ius distributivum* 'jaotav õigus'.

Nimi- ja omadussõnad ühilduvad ladina keeles omavahel alati ka käändes. Väljendis *mala fide* 'pahausksest' peavad olema põhisõna ja täiend mõlemad ablatiivi käändes, artiklist aga võib leida variandi *mala fidem*, milles adjektiiv on ablatiivi ja substantiiv akusatiivi lõpuga. Probleeme tekitab ka väljend *lex posterior derogat legi priori* 'hilisem seadus tühistab varasema seaduse', milles *lex posterior* ühildub nominatiivis ja *legi priori* ainsuse daativis. Artiklist leitud variandis *lex posteriori derogat leges priori* on eksitud mitme grammatikareegli vastu: ühildumist ei toimu *lex posteriori* nominatiivi ja daativi vahel; *leges priori* ei ühildu ka arvus: siin on kõrvuti mitmuse nominatiiv ja ainsuse daativ. Mitmel korral esineb Juridica

¹⁶ Õigusteaduses on rohkem tuntud väljendid *iustitia commutativa* 'võrdsustav õiglus' ja *iustitia distributiva* 'jaotav õiglus'.

artiklites termin *stricto sensu* 'otseses mõttes' ebaõigel kujul *strictu sensu*, milles omadussõnal on nimisõna käändelõpp. Selliste ühildumisvigade vältimine ning ladina terminite õigekirja kontrollimine on eriti olulised põhjusel, et ajakirja *Juridica* kasutatakse ka õppematerjalina. Halba eeskju annab õppejõu artiklis esinev väär vorm. Näiteks vale käändelõpuga termin *strictu sensu* esines esmalt õppejõu artiklis ja seejärel *Juridica* bakalaureusenumbris.¹⁷

Eksimusi võib leida samuti ladina eessõna ja sellega ühilduva käände kasutamisel. Prepositsioon *ad* 'juurde, poole' nõuab enda järele akusatiivi käänat, artiklis esineb aga ablatiiv: *poena absoluta ad effectum* pro *poena absoluta ad effectum* 'täielik karistus tulemuse saavutamiseks'. Prepositsiooni *ex* '-st' kasutatakse koos ablatiiviga, artiklis esineb see aga koos genitiivi käändega: *ex iniuriæ ius non oritur* pro *ex iniuriæ ius non oritur* 'õigusrikkumisest ei sünni õigust'.

Ladina tsitaatide vormiline sobivus eesti lauses

Ladina terminite kasutamisel osutub kõige üldisemaks probleemiks tsitaatsõna ühildumine eestikeelse tekstiga. Suurelt osalt sünteetilises ladina keeles tähistatakse grammatilisi vahakordi sufiksitena. Ladina väljendi äratuntavust ja arusaadavust võib seega mõjutada isegi termini ainsuse ja mitmuse vaheldumine.

Üldine on termini äratuntavuse probleem ladina III deklinatsiooni sõnade juures. Selles käändkonnas erineb genitiivi tüvi, millest moodustatakse ülejäänud käänded, sageli nominatiivis lähtevormist. Ainsuse nominatiiv on siin enamasti tüve lühenenud ja muundunud vorm. Enamasti on ladina õigusterminid kasutusel ainsuse või mitmuse nominatiivis, millest viimane on moodustatud genitiivi tüvest ja võib vähem või rohkem ainsuse nominatiivist erineda, nt *ius–iura* (tüvi: *iur-*) 'õigus'; *opus–opera* 'töö' (tüvi: *oper-*); *lex–leges* 'seadus' (tüvi: *leg-*); *mos–mores* 'tava' (tüvi: *mor-*). Kui selline sõna esineb väljendis mitmusevormis, ei jõuta tõlkimisel tihti õige lähtesõnani. Näiteks termini *vis* 'võim' mitmus on *vires*, mida tõlkimisel aetakse sageli ekslikult segi sõnaga *vir*, mitmuses *virī* 'mees'.

Ainsuse või mitmuse küsimus tõuseb üles ladina II deklinatsiooni *um*-lõpuliste kesksõost sõnade puhul, mille mitmuse lõpp on *-a*, nt *pactum*, mitm *pacta* 'leping'; *argumentum*, mitm *argumenta* 'tõendus'; *obiter dictum*, mitm *obiter dicta* 'muu hulgas öeldu; juhuslik märkus'. Selline mitmuse lõpp langeb kokku I deklinatsiooni sõnade ainsuse lõpuga, vrd *culpa*, mitm *culpæ* 'süü, hooletus'. Seetõttu tõlgitakse sageli sõna *digestum*, mitm *digesta* 'korrastamine' eesti keelde mitmuse ühildumisveaga *Digestad* (pro *Digestid*),¹⁸ kus koos esinevad nii ladina kui eesti mitmuse lõpp. Eraldi saab välja tuua eksimustüübi, kus termini nominatiivis lähtevormi asemel on kirjutatud selle genitiivtüvi. Näiteks esineb sõna *natio* 'rahvas; rahvus' asemel tüvi *nation* (7).

- (7) "Kannavad ju ka kreekakeelne *ethnos* ja ladinakeelne *nation* sama tähendust." (Kerikmäe 1995: 330)

Sarnaselt esineb tüvi *iur* akusatiivis oleva objekti *ius*¹⁹ asemel (8).

¹⁷ Vrd Uibopuu 1997: 237 ja Parmas 2002: 324. Alates 1996. aastast ilmub igal aastal *Juridica* teemanumber, milles on avaldatud kokkuvõtteid juurailiõpilaste bakalaureuse- ja magistratöödest.

¹⁸ *Digestid* (ehk *Pandektid*) on Rooma õiguse üks tähtsamaid allikaid, keiser Justinianuse korraldusel aastatel 530–533 koostatud 50 raamatust koosnev väljavõtete kogum Rooma klassikalise ajastu juriidilisest kirjandusest (Adomeit jt 2005: 85).

¹⁹ *Ius* 'õigus' on kesksõost sõna, millel nominatiivi ja objektkäände akusatiivi vormid kattuvad.

- (8) “Tööjaotust kohtuniku ja protsessiosaliste vahel näitab tabavalt sentents – *Narra mihi factum, ego tibi narrabo iur* – näita mulle asjaolud ja mina näitan sulle õigust.” (Pärsimägi 1996: 392)

Sellised vead tekivad enamasti põhjusel, et paljud ladina sõnad (eriti III deklinatsioonist) esinevad võõrsõnadega ja võõrkeeltes just oma tüve alusel. Ka sõnad *natio* ja *ius* on nii eesti kui teistes keeltes tuntud oma tüvede *nation*- ja *iur*- kaudu (vrd ee *natsioonaalne, juriidiline*; ingl *national, juridical*; sks *national, juristisch*; pr *national, juridique*).

Tsitaatsõnade ja -väljendite puhul rõhutavad keelekorraldajad, et need on puhtvõõrkeelsed ning neid kirjutatakse nagu lähtekeeles. Tsitaatsõnad oma põhisõnadega ei ühildu, need vormistatakse muust tekstist eristamiseks teises kirjas, näiteks kursiivis. Liitsõnas ühendatakse tsitaatsõna sidekriipsu abil ning käänatakse ülakoma abil (Erelt 1999: 16–17, Mäearu 2005: 171). Juridica artiklite põhjal valmistab autoritele raskust ladina termini kahe vormi – nominatiivse lähte- ja määrusliku ablatiivivormi – õige sobitamine eestikeelsesesse teksti. Näiteks esinevad ladina terminid *ultima ratio* ‘viimne vahend’ (9) ning *lex artis* ‘(eriala)kunsti reegel, oskusuue’ (10) tekstis oma lähtevormis, nominatiivsena ning ei ühildu grammatiliselt ülejäänud lausega.

- (9) “Abinõud on halduslikud ning neid kohaldatakse nagu likvideerimistki *ultima ratio*.” (Hirsch 1996: 20)
- (10) “Üldtunnustatud on seisukoht, et arst peab ravi läbi viima arstikunsti reeglite kohaselt ehk *lex artis*.” (Nõmper 2000: 443)

Sobivam oleks lauses 9 ablatiivi vorm: *ultima ratione* ‘viimse vahendina, äärmisel vajadusel’ või lisada eestikeelne täiendav sõna vastavas vormis, nt *ultima ratio*-põhimõttel. Näites 10 oleks olnud korrektsem kasutada kas termini ablatiivivormi: *lege artis* ‘oskusuuete kohaselt’ või paigutada ladina termin lauses ettepoole koos eesti käändelõpuga: .. *arstikunsti reeglite ehk lex artis’e kohaselt*. Leidub ka vastupidine juhtum, kus ladina termin *lege artis* esineb ablatiivi käändes määrusena, kuid nominatiivne lähtevorm *lex artis* sobiks ja ühilduks kontekstiga täpsemini. Näites on eksitud ka ladina tsitaatsõna ning eesti käändelõpu kasutamisel, need kaks osist tuleb eraldada ülakomaga: *lege artis’e*.

- (11) “.. aastatuhandete vältel asendus see aukartusega arsti *lege artise* vastu ning arsti ja patsiendi usaldusliku vahekorraga.” (Sootak 1995: 331)

Näites 12 on ladina termin *contra bonos mores* ‘heade tavade vastaselt’ poolik, ladina eessõna *contra* asemel on kasutatud eesti tagasõna *vastaselt*. Autor on lisanud ladina terminile ka ebatäpse tõlke: *bonos mores* on akusatiivi käändes ‘häid kombeid’, mis omakorda ei ühildu ülejäänud lausega.

- (12) “Kohus otsustas, et Inglise kohtuotsuses kasutatud mõiste “hästi käituda”, see on mitte käituda *bonos mores* (ld. k. head kombed) vastaselt, mida Inglise õiguses on defineeritud kui käitumist, mis on “enamiku kaasini-meste arvates pigem vale kui õige”, on ilmselt ebaselge ega anna piisavat juhust selleks, kuidas tulevikus käituda.” (Maruste 2001: 18)

Kokkuvõte

Uurimus ladina õigusterminite vormilisest ja sisulisest korrektsusest õigusteadusliku ajakirja *Juridica* 1993–2007 aastakäikude põhjal toob välja praktilised probleemid, mis kaasnevad võõrkeelendite tarvitamisega. Ladina keeles tähistatakse grammatilisi vahakordi sufiksitena. Ladina termini äratuntavust ja arusaadavust võib mõjutada seega termini ainsuse ja mitmuse vaheldumine, samuti termini kasutamine käändevormis või prepositsiooniga. Mõistmisprobleeme põhjustavad ka võõrtermini tõlke puudumine, samuti vale tõlke lisamine ladina terminile ning sobimatu väljendi kasutamine.

Ortograafiavigade puhul jäävad silma ladina keele häälduspõhimõtetest tulevad eksimused (s pro c, g pro q, c pro q, ks pro x, ts pro c). Eksitakse ka sarnase kirja- ja tähtedega (a ja o) kirjutamisel ning tsitaatsõna korrektsel tähistamisel. Ladina terminite kasutamisel osutub kõige üldisemaks probleemiks aga tsitaatsõna sisulise sobitumine ja keeleline ühildumine eestikeelse tekstiga. Kõik need probleemid mõjutavad ladina terminite eestikeelses kontekstis kasutamisel ka lause süntaktilist ja semantilist struktuuri. Korrektsel väljendusviisi pole võimalik saavutada, mõistmata ladina keelest pärit termini lingvistilist eripära.

Kuidas sellised eksimused grammatikareeglite vastu tekivad? Uuritavat perioodi iseloomustab Eesti õigustegevuses rahvusvaheliste õigusaktide uurimine ning kohandamine oma riigi seadustega. Tagasipöördumine Euroopa õigusruumi on mõjutanud ka õiguskeele kasutust. Eesti juristid tuginevad oma kirjatöödes sageli võõrkeelsetele, enamasti saksa- ja ingliskeelsetele materjalidele. Võib oletada, et mitmed õigekirja- ja vormivead on tekkinud tsitaatide ümberkirjutamise või puuduliku kvaliteediga koopiade uurimise käigus. Samuti selgitaks võõrkeelsete materjalide kasutamine artiklites vigu ladina termini ühildumisel eestikeelse tekstiga. Võõrkeelte mõju eesti õigustekstile on siinsest artiklist kõrvale jäänud uurimisülesanne.

Termini tuntust ja korrektsust ei saa seostada. *Juridica* materjali põhjal ei leia kinnitust, et vead puudutaksid eelkõige vähe kasutatavaid ja mittetuntud termineid. Vääril kujul esineb artiklites ka väljendeid, mis arvandmete järgi esinevad sageli ja peaksid kuuluma juristide üldsõnavara hulka.

Ladina keele terminite korrektseks kasutamiseks tuleb nii õppetöös kui juuraajakirjade artiklites juhtida enam tähelepanu eesti ja ladina keele struktuuri erinevustele (nt soovormid ja eessõnad) ning ühildumisprobleemidele. Samuti kujundada ja suurendada juristide harjumust kasutada keeleabimaterjale (sõnastikud ja grammatikakäsiraamatud) ning juhtida seeläbi tähelepanu ladina terminite tähendusele ja õigekirjale ning hõlbustada nii tööd erialaste materjalidega.

Kirjandus

- Adomeit, Klaus; Ristikivi, Merike; Siimets-Gross, Hesi 2005. Ladina-eesti õigussõnastik. Peeter Kask (toim.). Tallinn: Eesti Keele Sihtasutus.
- Asprey, Michèle M. 2003. Plain Language for Lawyers. 3rd ed. Annandale: The Federation Press.
- Benke, Nikolaus; Meissel, Franz Stefan 1997. Juristenlatein. Lateinische Fachausdrücke und Redewendungen der Juristensprache, übersetzt und erläutert. Wien: Juridica.
- Erelt, Tiiu 1982. Eesti oskuskeel. Tallinn: Valgus.
- Erelt, Tiiu 1999. Eesti ortograafia. 3. täiend. tr. Tallinn: Eesti Keele Sihtasutus.
- Kask, Peeter; Lill, Anne 1997. Õigustermini *restitutio in integrum* ja väljendi *restitutio ad integrum kasutamine*. – Õiguskeel 3, 12–14.
- Kasik, Reet 1995. Õiguskeele lauseehitus. – Õiguskeel 5, 16–21.
- Kelly, Henry A. 1988. Lawyers' Latin: *Loquenda ut Vulgus?* – Journal of Legal Education 38, 195–207.
- Kerge, Krista 2002. Kirjakeele kasutusvaldkondade süntaktiline keerukus. – Reet Kasik (toim.). Tekstid ja taustad. Artikleid tekstianalüüsisist. Tartu Ülikooli eesti keele õppe-
tooli toimetised 23. Tartu: Tartu Ülikooli Kirjastus, 29–46.
- Knütel, Rolf 1994. Rechtseinheit in Europa und römisches Recht. – Zeitschrift für europäisches Recht, 244–276.
- Kukk, Inga 2003. Õiguse ja halduse keel. – Maire Raadik (toim.). Eesti kirjakeele kasutusvaldkondade seisundi uuringud. Tallinna Pedagoogikaülikooli eesti filoloogia osakonna toimetised 4. Tallinn: TPÜ Kirjastus, 96–117.
- Kull, Rein 2000. Kirjakeel, oskuskeel ja üldkeel. Tallinn: Eesti Keele Sihtasutus.
- Lill, Anne 1996. Ladina keel õiguskeele alusena. – Õiguskeel 4, 9–13.
- Lill, Anne 2003. Kakssada aastat klassikalist filoloogiat Eestis: kahe alguse lugu. – Ivo Volt (toim.). Kakssada aastat klassikalist filoloogiat Eestis. Morgensterni Seltsi toimetised I. Tartu: Tartu Ülikooli Kirjastus, 9–18.
- Mattila, Heikki E. S. 2002. Vertaileva oikeuslingvistiikka. Helsinki: Kauppakaari Lakimiesliiton Kustannus.
- Mattila, Heikki E. S. 2006. Comparative Legal Linguistics. Aldershot: Ashgate.
- Mellinkoff, David 1963. The Language of the Law. Boston, Toronto: Little, Brown and Co.
- Mereste, Uno 2000. Oskuskeel ja seaduste keeleline rüü. Tallinn: Eesti Keele Sihtasutus.
- Meyer, Dieter 2002. Juristische Fremdwörter, Fachausdrücke und Abkürzungen. 11. Aufl. Neuwied, Kriftel: Lutherhand.
- Mäearu, Sirje 2005. Keelekasutuse keerdkäigud. Õigekiri. – Aime Vettik, Erki Silvet (koost.). Keel ja õigus. Tallinn: Kirjastus Juura, 169–185.
- Oksaar, Els 1999. Kõnekeel, erialakeel, õiguskeel. – Juridica 4, 199–207.
- Pruks, Peep 2001. Law Journals Juridica and Juridica International. Publishing Legal Literature. – Faculty of Law, University of Tartu. Tartu: Sihtasutus Iuridicum.
- Ristikivi, Merike 2003. Ladina keel ajakirjas Juridica 1993–2002. – Juridica 10, 727–732.
- Ristikivi, Merike 2005. Latin: The Common Legal Language of Europe? – Juridica International 10, 199–202.
- Ristikivi, Merike 2007. Ladina keelest pärit terminite polüseemia ja sünonüümia eesti õiguskeeles. – Helle Metslang, Margit Langemets, Maria-Maren Sepper (toim.). Eesti Rakenduslingvistika Ühingu aastaraamat 3. Estonian Papers in Applied Linguistics 3. Tallinn: Eesti Keele Sihtasutus, 253–267.
- Ristikivi, Merike 2007. Lexica iuridica in Juridica: Latin Terms as a Reflection of Europeanisation of Estonian Legal Culture. – Juridica International 12, 173–179.
- Saari, Henn 1980. Omasõna ja võõrsõna paarid eesti oskussõnavaras (1). – Keel ja Kirjandus 12, 737–743.
- Silvet, Erki 1997. Ladinakeelne väljend õiguskeeles. – Õiguskeel 2, 16–19.

Tamm, Kaljo 1997. Põhiseaduse terminoloogiast. – Õiguskeel 1, 7–12.
Wolodkiewicz, Witold; Krzynówek, Jerzy 2001. Lacinskie paremie w europejskiej kulturze prawnej i orzecznictwie sądów. Warszawa: Liber.

Materjali allikad

- Ginter, Carri 2001. Tüüptingimustest VÕS ja direktiivis 93/13/EMÜ. – *Juridica* 7, 501–513.
- Hirsch, Hans Joachim 1996. Juriidiliste isikute kriminaalvastutusest. – *Juridica* 1, 17–24.
- Juridica*. Tartu Ülikooli õigusteaduskonna ajakiri. 1993–. Peatoimetaja Paul Varul. Tartu: Sihtasutus Iuridicum.
- Kerikmäe, Tanel 1995. Mis on vähemus? – *Juridica* 8, 330.
- Luts, Marju 1996. Lünga vastu tõlgendamise või analoogiaga? (Diskussioonist juriidilises meetodiõpetuses). – *Juridica* 7, 348–352.
- Maruste, Rait 2001. Sõnavabadus ja selle piirid. – *Juridica* 1, 15–23.
- Narits, Raul 1994. Tõlgendamine: teadus või seadus? – *Juridica* 9, 228–230.
- Nõmper, Ants 2000. Stigma abordi kaasus. – *Juridica* 7, 435–450.
- Parmas, Andres 2002. Inimsusevastased kuriteod rahvusvaheliste kriminaalkohtute käsitluses. – *Juridica* 5, 321–329.
- Pärsimägi, Andra 1996. Kohus see on kohtunik... – *Juridica* 8, 391–396.
- Sootak, Jaan 1995. Meditsiiniõiguse eetilistest lähtekohtadest. – *Juridica* 8, 331–332.
- Sootak, Jaan 1999. *Corpus Iuris* ja Eesti kriminaalõigus. – *Juridica* 4, 176–184.
- Uibopuu, Henn-Jüri 1997. Üldised inimõigused Eesti 1920., 1937. ja 1992. aasta põhiseadustes. – *Juridica* 5, 235–241.
- Uibopuu, Henn-Jüri 1998. Eesti põhiseadus, rahvusvahelised suhted ja rahvusvaheline õigus. – *Juridica* 4, 187–196.
- Uibopuu, Henn-Jüri 2005. *Ius cogens*: tänapäeva rahvusvahelise õiguse üldtunnustatud norm või heatahtlik soovitus. – *Juridica* 4, 273–286.
- Varul, Paul 1993. Austatud *Juridica* esimese numbri lugeja! (toimetajaveerg). – *Juridica* 1, 2.
- Varul, Paul 2002. Austatud lugeja! (toimetajaveerg). – *Juridica* 4, 209.

Merike Ristikivi (Tartu Ülikool) on erialalt klassikaline filoloog, uurimisteemaks on ladina keele mõjutused eesti õiguskeeles.
merike.ristikivi@ut.ee

ACCURACY OF THE FORM AND CONTENT OF LATIN LEGAL TERMS

Merike Ristikivi

University of Tartu

The article discusses the orthography of Latin legal terms and its peculiarities, and Latin morphology and its relationship with Estonian sentence. Law is a field where linguistic means of expression are of utmost importance. This discipline operates directly through language; a word or expression acquires juridical power in it. The material analysed comprises the terms collected from the law review *Juridica* issued from 1993 to 2007. The time period looked at in the given research has been a decisive era in the development of Estonian law: it is characterised by turning back to the Western legal environment which largely depends on the Latin language. The legal reform in Estonia has been accompanied by changes in the usage of terms by Estonian lawyers. The integration of Estonian legal language into European legal culture is reflected by a relatively high increase in the usage of terms in Latin. At the same time, there occur several problems when Latin terms are employed.

In Latin, a synthetic language, grammatical relationships are represented by applying suffixes and inflectional endings. As a result, the recognition and understanding of a Latin term may be affected by distinguishing the singular and the plural form, as well as using of a term in different case forms or with various prepositions. Difficulties in understanding can also be caused by the use of a less known Latin expression or the improper use of an expression in a sentence.

Mistakes frequently appear in the spelling of Latin terms, as well as in the agreement between case forms and gender forms and in translation of Latin terms. The most common problem when using Latin terms, however, is adapting the foreign words to the context and incorporating them into the Estonian sentences. Ordinarily mistakes occur in the usage of two forms – the basic form in the nominative case and the adverbial in the ablative – in the proper context.

Keywords: legal language, terminology, language mistakes, orthography, syntax-morphology interaction, Latin, Estonian

KONSTRUKTSIOONIPÕHINE KEELEMUDEL JA SÕNARAAMATUMUDEL

Heete Sahkai

Ülevaade. Artiklis arutletakse võimaluse üle kajastada sõnaraamatus nähtusi, mis on olnud ajendiks konstruktsioonipõhise keelemudeli väljatöötamisele ning mis on problemaatilised nii leksikaalsel projektsioonil ja kompositsioonilisel semantikal põhinevate keeleteooriate kui traditsioonilise sõnaraamatumudeli jaoks. Artiklis vaadeldakse kolme tüüpi nähtusi: esiteks idiosünkraatilises morfosüntaktilises mallis kinnistunud sõnavormide erandlikku grammatilist käitumist, teiseks sõna esinemist mingis tähenduses ainult kindlates konstruktsioonides ja kolmandaks sõnavormide produktiivset esinemist eripärase vormi ja tähendusega konstruktsioonis. Kahe esimese nähtuse käsitlemiseks sõnaraamatus pakutakse välja konstruktsiooniploki lisamine sõnaartikli struktuuri, kolmanda nähtuse kirjeldamiseks tuleb aga leksikoni täiendada “konstruktikoniga”.*

Võtmesõnad: leksikograafia, konstruktsioonigrammatika, süntaks, eesti keel

1. Sissejuhatus

Sõnaraamatu ülesehituse aluseks on meie ettekujutus keele ülesehitusest. Artiklis arutlen selle üle, millised järeldused võiksid tuleneda leksikograafia jaoks konstruktsioonipõhisest keelemudelist ja selle väljatöötamise ajendiks olnud nähtustest, mis on problemaatilised nii paljude alternatiivsete keeleteooriate jaoks kui ka traditsioonilise sõnaraamatumudeli seisukohast.

Nimelt on konstruktsioonipõhine lähenemine kasvanud välja tähelepanekust, et idiomaatiliste ja erandlike nähtuste osakaal keeles on senistes teooriates eeldatust oluliselt suurem ning neid ei saa esitada üksnes idioomide ja erandite lõpliku loeteluna, vaid nad võivad olla ka osaliselt produktiivsed ja reeglipärased (Fillmore jt 1988). Ehk teiste sõnadega, erandliku vormiga keeleväljendid, mille struktuur

* Artikkel põhineb VI rakenduslingvistika konverentsil peetud ettekandel “Keelemudel ja sõnaraamatumudel” (Sahkai 2007) ja selle kirjutamist on toetanud Eesti Teadusfondi grant nr 5970. Olen tänulik Margit Langemetsale ja Ülle Viksile, kes nõustusid artikli käsikirja lugema ja kommenteerima.

ei taandu neid moodustavate leksikaalsete üksuste kombinatoorsetele omadustele ja üldistele grammatilistele seaduspäradele, võivad olla moodustatud osaliselt produktiivse malli järgi, mis tähendab, et grammatilised seaduspärad võivad olla väga erineva üldisus- ja produktiivsuseastmega. Ning teiselt poolt, idiomaatilist või mittekompositsioonilist tähendust ei pruugi kanda mitte ainult konkreetset, kinnistunud sõnaühendid, vaid ka produktiivselt moodustatud väljendid, ehk tähendus võib olla seotud ka (osaliselt) produktiivsete morfosüntaktiliste mallidega.

Selliste osaliselt produktiivsete ja tähendust kandvate idiosünkraatiliste mallide representeerimiseks välja töötatud teooriates (nt Goldberg 1995, Kay, Fillmore 1999, Croft 2001) ei jagata keelelist teadmist leksikoniks ja grammatikaks, vaid leksikaalseid ja grammatilisi üksusi käsitletakse ühe kontinuumina, mis ulatub sõnadest ja idioomidest osaliselt või täielikult skemaatiliste ja produktiivsete morfosüntaktiliste mallideni, millel võib samuti olla oma tähendus ja pragmaatiline funktsioon. See tähendab, et kõiki keelelise teadmise üksusi kujutatakse ühes formaadis, nn konstruktsioonidena, millel on kindel vorm, kombinatoorsed omadused ja (potentsiaalselt) tähendus ning pragmaatilised omadused. Seega kui enamikus teooriates kujuneb lause tähendus kompositsiooniliselt seda moodustavate leksikaalsete üksuste tähendustest, siis konstruktsioonipõhises mudelis võivad lause tähendusse oma panuse anda ka morfosüntaktilised mallid, nii et sõna tähendust kontekstis ei määra üksnes tema leksikaalne tähendus, vaid ka konstruktsioon, mille koosseisus ta esineb. Samuti ei ole lause vorm üksnes seda moodustavate sõnade kombinatoorsete omaduste peegeldus, vaid võib olla määratud ka seda moodustavate konstruktsioonide poolt, millel on samuti oma kombinatoorsed omadused: mitte üksnes sõnad ei "vali" oma süntaktilist ümbrust, vaid ka morfosüntaktilised mallid valivad sõnu, mis võivad nendes esineda. Ning et sellised mallid on väga erineva üldisusastmega, ei lange nende poolt valitud sõnade klassid alati kokku traditsiooniliste sõnaliikidega. Kui sõna on sellises idiosünkraatilises konstruktsioonis kinnistunud, võib selles esinemist pidada küll tema leksikaalseks omaduseks, ent tema süntaktilist käitumist ei saa sel juhul kirjeldada sõnaliigi terminites ning tema tähendust selles konstruktsioonis ei saa kirjeldada tema kontekstivälise tähendusena. Kui aga sõna esineb konstruktsioonis produktiivselt, ei saa tema süntaktilist käitumist ei taandada tema enda kombinatoorsetele omadustele ega kirjeldada sõnaliigi terminites.

Konstruktsioonipõhine lähenemine seab seega kahtluse alla kaks traditsioonilise sõnaraamatumudeli (ja paljude keeleteooriate) eeldustest: et sõnad jagunevad oma distributsiooni alusel suhteliselt väikeseks hulgaks sõnaliikideks, mis alluvad ühesugustele grammatilistele seaduspäradele, ning et sõnad on autonoomsed üksused, mille tähenduse lauses määrab nende inherentne leksikaalne tähendus ning mille morfosüntaktilise käitumise määravad nende endi kombinatoorsed omadused nagu sõnaliik ja reksioonimall. Sellest tuleneb, et praegune sõnaraamatumudel jääb hätta juhtumitega, kus sõnade käitumist ei saa kirjeldada sõnaliigimärgenditega ning kus sõna tähendus ja/või süntaks on konstruktsioonispetsiifiline ehk seotud konkreetse morfosüntaktilise kontekstiga.

Leksikograafidele ei ole selliste nähtuste olemasolu muidugi uudis, kuid sõnaraamatumudeli aluseks olevast keeleteooriast sõltub see, kuidas sellistesse nähtustesse suhtuda. Seni on eeldatud (ja ka valitsevad keeleteooriad ennustavad), et selliste nähtuste näol on tegemist pigem üksikute marginaalsete eranditega. Konstruktsioonipõhine keelemudel aga ennustab, et tegemist on piisavalt ulatuslike,

süsteematailiste ja kesksete nähtusega keeles, et nendega sõnaraamatu struktuuri kujundades arvestada ning neid sõnaraamatut koostades sõnade tähendusest ja käitumisest süsteematailiselt otsida.

Varem on idiosünkraatiliste ja idiomaatiliste konstruktsioonide leksikograafilisest käsitlemisest rääkinud konverentsiettekandes Charles Fillmore (2006) seoses täis-tekstide märgendamise kui uue menetluse kasutuselevõtmisega FrameNeti¹ projekti raames. C. Fillmore'i nägemuse järgi peaks selle menetluse tulemusena sündima elektrooniline ressurss, mis koosneb kolmest komponendist: inglise keele leksikonist, kus on muuhulgas kirjeldatud sõnade esinemist erandlikes konstruktsioonides, wiki-põhimõttel täiendatavast inglise keele "konstruktikonist", kus konstruktsioonide kirjeldused on seotud neis esineda võivate sõnade kirjetega leksikonis (välja arvatud muidugi juhul, kui konstruktsioon on täiesti skemaatiline ega ole seotud konkreetsete sõnadega), ja märgendatud tekstid, kus on annoteeritud ka konstruktsioonid.

Artiklis püüan näidata, et konstruktsioonide kirjeldamine on vajalik ka traditsioonilises sõnaraamatus, ning pakun selleks välja spetsiaalse konstruktsiooniploki lisamise sõnaraamatu struktuuri (peatükk 2.1.). Konstruktsioonide kirjeldamise vajalikkus oleneb muidugi sõnaraamatu tüübist ja funktsioonist – võib eeldada, et ennekõike on see vajalik keeleõppijatele mõeldud sõnastikes. See, milline täpselt peaks olema sellise konstruktsiooniploki ülesehitus ja koht sõnartikli struktuuris, jääb siiski veel edasise uurimise objektiks.

Illustreerin konstruktsiooniploki võimalusi ja piire kolme näitega. Artikli 2. osas kirjeldan põgusalt idiosünkraatilistes süntaktilistes mallides kinnistunud parameetersõnu, mille erandlikku süntaktilist käitumist konstruktsiooniplokk võimaldaks kirjeldada. 3. osas vaatlen sõna *minema* tähendusploki 'alustama, hakkama, algama' EKSS-is ning püüan näidata, et tegemist on konstruktsioonispetsiifilise tähendusega, mille esinemiskontekste võiks samuti kirjeldada konstruktsiooniplokkides. 4. osas toon näite juhtumist, kus sõnade idiosünkraatilist süntaksit ja tähendust ei ole võimalik ka konstruktsiooniplokis kirjeldada, sest nad omandavad selle produktiivses konstruktsioonis esinedes. Selliste nähtuste kirjeldamiseks tuleks leksikoni täiendada "konstruktikoniga". 5. osas teen artiklist kokkuvõtte.

2. Idiosünkraatiline grammatika

Idiosünkraatilises konstruktsioonis kinnistunud sõnu, mille süntaktilist käitumist ei saa kirjeldada sõnaliigimärgendiga, sobivad näitlikustama parameetersõnadena toimivad nimisõnavormid. Tegemist on küllalt suure rühma sõnadega, millest paljud on sagedased ning mis asuvad leksikaalsete ja grammatiliste sõnade vahepeal, sest nad ei täida nimisõnadele omast referentsiaalset funktsiooni, kuid ei moodusta ka päris suletud klassi. Seega on tegemist funktsionaalselt oluliste sõnadega, millest sagedasemate käitumist võib olla vajalik kirjeldada eriti keeleõppijatele suunatud sõnaraamatutes.

Näiteks võib tuua konstruktsiooni, mis koosneb partitiivis parameetersõnast ja selle täiendist (nt *ilusat värvi*). Täiendiks võib olla kas adjektiiv- või genitiivatribuut,² kuid mitte mõlemad korraga, samuti mitte relatiivlause. Seega ei käitu partitiivsed parameetersõnad reeglipäraste nimisõnadena, kuid neid ei saa paigutada üheselt ka mingisse muusse sõnaliiki. Sõnaraamatutes on neid küll

¹ <http://framenet.icsi.berkeley.edu/> (17.12.2007).

² Seetõttu võib rääkida ka kahest erinevast mallist, [adjektiivatribuut + partitiivne N] ja [genitiivatribuut + partitiivne N]. Sellise lähenemise kasuks räägib see, et mõned partitiivsed parameetersõnad ei kombineeru vabalt mõlema atribuudiga, vaid eelistavad neist ühte, nt *nägu* eelistab genitiivatribuuti: *vrđ Ta on venna nägu*, aga **Ta on vistrikulist nägu*; mõned adjektiivid on siiski võimalikud, nt *imelikku nägu*.

seostatud kaassõnadega, sest nimisõna- ja kaassõnafraasid täidavad lauses osaliselt samu funktsioone ning parameetersõnad nõuavad lisaks erinevalt enamikust nimisõnadest laiendit, meenutades sellega kaassõnu, kuid nende laiendamisvõimalused ei lange sageli kokku kaassõnade omadega. Kirjeldatud konstruktsiooni(de)s võivad esineda näiteks järgmised sõnad: *laadi, moodi, stiili, liiki, tüüpi, sorti, klassi, tõugu, rassi, värvi, tooni, karva, mõõtu, kasvu, kaalu, nägu...* Samas järgmised sõnad, mis on mingis mõttes sarnase tähendusega, mallis esineda ei saa, ehk tegemist ei ole päris produktiivse konstruktsiooniga: *kategooriat, rahvust, hõimu, kuju, vormi, maitset, lõhna, välimust, ilmet...*

Selliseid parameetersõnade malle on eesti keeles terve rida, igauks veidi erineva vormi, distributsiooni ja produktiivsusega; morfoloogiliselt on neis esindatud enamus semantilisi käändeid, aga ka partitiiv, nagu eelpool illustreeritud. Sageli esinevad sama sõna vormid mitmes sellises konstruktsioonis. Näiteks sõnast *mood* on kolm parameetersõna, mis kõik käituvad erinevalt (vt tabel 1): *moodi* kombineerub adjektiivatribuudiga (A) ja genitiivatribuudiga (G) ning võib lauses toimida täiendi (Atr), öeldistäite (Pred) ja määrusena (Adv); *moel* võib kombineeruda adjektiiv- ja genitiivatribuudiga ja ka mõlemaga korraga (A+G), samuti relatiivlausega (Rel; nt *Ta käitus moel, mis jättis ebaviisaka mulje*), kuid saab lauses olla ainult määrus; *moega* võimaldab samu täiendeid nagu *moel*, välja arvatud relatiivlause, ent võib lisaks määrusele olla ka täiend ja öeldistäide. Ehkki laias laastus võib rääkida näiteks partitiivsest mallist, adessiivsest mallist ja komitatiivsest mallist, ei ole kõik samas käändes parameetersõnad siiski sama distributsiooniga: näiteks *viisil* tundub käituvat samamoodi nagu *moel*; *kombel* aga mitte (vt tabel 1). Seega võib parameetersõnade süntaksit sageli käsitleda üksikute sõnade idiosünkraatilise käitumisena, mõnel juhul aga võib rääkida ka osaliselt produktiivsetest konstruktsioonidest.

Tabel 1. Parameetersõnade süntakiline variatiivsus

	A	G	A+G	Rel	Atr	Pred	Adv
<i>moodi</i>	+	+	-	-	+	+	+
<i>moega</i>	+	+	+	-	+	+	+
<i>moel</i>	+	+	+	+	-	-	+
<i>kombel</i>	+	+	-	-	-	-	+

Sõnaraamatu koostaja ette seavad sellised sõnad kaks probleemi: esiteks ei ole sõnavormi käitumine ennustatav algvormi põhjal, ehkki tähenduslik seos on olemas, ning teiseks ei ole see kirjeldatav sõnaliigimärgendiga. Morfoloogiliselt reeglipärased käändevormid on seega süntakiliselt paradigmast eraldunud, kuid on säilitanud semantilise seose lähtesõnaga ning pole üle läinud ühtegi teise selgelt defineeritavasse sõnaliiki.

2.1. Konstruktsioonid sõnaraamatus

Sõnaraamatutes on selliste sõnade puhul kasutatud üldjoontes kolme tüüpi lähenemist. Esimene lähenemine on parameetersõnade esitamine vastavate nimisõnade reeglipäraste vormidena, st sõnavormi eripärast süntaksit parameetersõnana ei ole kirjeldatud (nt *laadi, liiki, tüüpi* EKSS-is) ega mõnel juhul ka näidetega illust-

reeritud (nt *tooni* koos genitiivtribuudiga EKSS-is). Selline lähenemine kajastab parameetersõna semantilist seost lähtesõnaga, kuid mitte selle iseseisvat staatust ja süntaksit.

Teise võimalusena on parameetersõnad esitatud omaette märksõnadena ja nende käitumist on kirjeldatud erinevate sõnaliigimärgendite kombinatsioonidega, näiteks VVS-is on omaette märksõnad *moodi Adv A Ks*, *laadi A*, *sorti A*, *karva Adv A*, kuid mitte näiteks *tüüpi*, *liiki*, *tooni*, *värvi*, *mõõtu*, *kasvu*, *kaalu*, *nägu*, *tõugu*. Selline lähenemine kajastab parameetersõna iseseisvumist paradigmat, kuid nagu öeldud, ei ole sõnaliigimärgendid nende käitumise kirjeldamiseks piisavad.

Kolmas kasutatud lahendus on parameetersõnade eripärase käitumise informaalne kirjeldamine lähtesõna artiklis eraldi tähendusploki alguses. Seega on sellist lahendust kasutatud eelkõige siis, kui erandliku käitumisega kaasneb tähenduserinevus, nt EKSS-is on osaliselt kirjeldatud sõnu *moodi*, *moel* ja *moega* (nd 1), kuid mitte sõnu *laadi*, *liiki*, *tüüpi*. Kirjeldused ei ole siiski täielikud: mainitud on käändevorme, adjektiivtribuuti ja postpositsioonilaadsust (ehk genitiivtribuuti), kuid mitte distributsiooni lauses ega ka nende kolme erinevusi, sest sõnaartikli liigendus põhineb tähendusel.

(1) mood

..

4. (hrl. partitiivis, adessiivis v. komitatiivis koos adjektiivse täiendiga:) viis, komme. ..

5. [gen.] hrl. substantiivi järel postpositsiooni laadis (väljendab tegelikult mingit sarnasusvõrdlust: kellegi, millegi sarnane v. sarnaselt). ..

6. (hrl. partitiivis:) liik, laad. *Sõpru on mitut moodi*. * *Sealt kostis juba selle uut moodi pilli, lõõtsmoonika hääli*.. A. Kalmus.

Kuna kõigil neil lahendustel on nii tugevusi kui nõrkusi, tundub kõige eelistavam nende kolme kombinatsioon, mis võimaldaks kajastada nii parameetersõna semantilist seost lähtesõnaga (ja tagada tema leitavuse lähtesõna artiklist) kui tema iseseisvat staatust, ja ühtlasi kirjeldada tema erandlikku grammatikat. Sellise lahenduse näiteks sobib sõna *kombel* esitus EKSS-is (nagu eelpool toodud näidetest näha, ei ole sellist esitusviisi kasutatud siiski süstemaatiliselt).

(2) kombel vt. komme.

komme, *kombe*¹⁹ s.

..

4. kombel (mingil) viisil, moel, moodi. **a.** koos eelneva (ühilduva) omadus- v. asesõnaga (mõnikord on asendatav vastavast omadus- v. asesõnast tuletatud adverbiga). .. || koos sõnadega *kole*, *hirmus*, *armetu*, *hale* jms. väljendab millegi rohkust v. intensiivsust. .. **b.** [gen.] hrl. nimisõna järel postpositsiooni laadis (väljendab tegelikult võrdlust eelneva nimisõnaga: *mehe kombel* 'nagu mees' jne.). .. **c.** kõnek. koos määrsõnaga. ..

Kombel on antud omaette märksõnana viitega artiklile *komme*, kus ta on rasvaselt esile tõstetud omaette tähendusplokis ja esitatud otse adessiivivormis, mitte lihtsalt sulgudes kirjeldusega "hrl. adessiivis", nii et see on kasutaja jaoks kergelt leitav; lisaks on allplokides kirjeldatud informaalset erinevaid konstruktsioone, milles *kombel* esineb.

Seega on tuntud vajadust kirjeldada sõnavormide iseseisvat ja eripärast käitumist ka muude vahenditega peale sõnaliigimärgendite, kuid sellised kirjeldused ei ole täielikud ega süstemaatilised ning nende jaoks ei ole loodud sõnaraamatu struktuuris omaette kohta. EKSS-is on eripärase vormi ja käitumise kirjeldus seotud märksõna tähenduse seletusega, ent nagu toodud näidetestki ilmneb, ei saa märksõna vormide eripärast käitumist alati siduda märksõna iseseisva tähendusega: nt sõna *kombel* puhul on tähendus omistatud märksõna kindlale vormile, mitte märksõnale endale, ning teiselt poolt sõnade nagu *liik*, *tüüp* või *laad* vormide tähendus parameetersõna konstruktsioonis ei eristu piisavalt märksõnade leksi-kaalsest tähendusest, et anda alust omaette tähendusploki loomiseks. Samuti ei võimalda selline lahendus kirjeldada juhtumeid, kus sõna esineb samas tähenduses erinevates konstruktsioonides, nt artikli järgmises osas vaatlen verbi *minema* esinemist tähenduses 'alustama, hakkama, algama' ja püüan näidata, et see tähendus on konstruktsioonispetsiifiline, kuid esineb mitmes konstruktsioonis.

Kui tähenduse seletuse juures esitatud grammatikakirjeldus on allutatud tähendusel põhinevale liigendusele, ehk täpsustatud on märksõna ühe tähenduse kasutuskontekste, siis sõnaliigimärgenditel põhinev grammatikakirjeldus on tähenduse suhtes primaarne – nt sõna *all* tähendust postpositsioonina ja adverbina on seletatud ühtemoodi, kuid sõnaliigierinevuse tõttu on nad esitatud erinevates struktuuriüksustes. Kuid ka sõnaliigi väli on mõeldud märksõna, mitte selle vormide kirjeldamiseks ega sobi seetõttu selliste konstruktsioonide kirjeldamiseks, milles märksõna võib esineda ainult kindlas vormis. Teiste sõnadega, oleks eksitav esitada samas numeratsioonis vaheldumisi informatsiooni märksõnaks oleva lekseemi ja selle üksikute vormide kohta.

Veel üks võimalus konstruktsioonide kirjeldamiseks oleks fraseologismiplokk – on ju idiosünkraatilise vormi ja/või mittekompositsioonilise tähendusega konstruktsioone nimetatudki "formaalseteks idioomideks" (Fillmore jt 1988: 505–506). Ka selliste konstruktsioonide näol on tegemist sõnast suuremate üksustega, kus sõna esineb kindlas vormis ja/või tähenduses ning millel endal on oma kindel vorm ja tähendus. Kuid nagu C. Fillmore jt sealsamas märgivad, võivad idioomid põhineda konstruktsioonidel, ehk vajalikuks võib osutuda konstruktsiooni illustreerimine idioomiga.³ Seega võiks konstruktsioonide kirjeldamiseks ette näha omaette ploki, mille ülesehitus oleks sarnane fraseologismiplokiga (näiteks EKSS-is on fraseologismiplokis koht seletuse, märgendite ja näidete jaoks, vt 1. kd lk 8 p 26), kuid mis võiks allüksusena sisaldada ka fraseologismiplokki. Samuti on võimalik, et juhtudel, kus sõna või sõnavorm esineb samas tähenduses mitmes konstruktsioonis (nagu *minema* järgmises peatükis illustreeritud konstruktsioonides), on vajalik eraldi kirjeldada sõna(vormi) tähendust konstruktsioonis ja konstruktsiooni kui terviku tähendust; ei ole ka välistatud, et ühel konstruktsioonil võib olla mitu tähendust, märgendit vms. Seega kui fraseologismid on EKSS-is esitatud alati sõnartikli lõpus, siis konstruktsioone kui vähem idiomaatilisi üksusi võib olla otstarbekas kirjeldada ka tähendusplokkide sees.

Konstruktsiooniploki kasutuselevõtt tähendaks, et sõnartiklisse toodaks märksõna sõnaliikide, tähenduste ja märksõna sisaldavate fraseologismide kõrval sisse veel üks sõltumatu liigendusalus: märksõna (vorme) sisaldavad erandliku ehituse ja/või mittekompositsioonilise tähendusega konstruktsioonid. See võimaldaks näiteks luua omaette ploki parameetersõnade *liiki* või *tüüpi* süntaktilise käitumise kirjeldamiseks

³ Idioomide ja idiomaatiliste konstruktsioonide tihed seos tuleb välja ka järgmises peatükis, kus käsitletakse *minema* esinemist algamist või alustamist väljendavas tähenduses nii püsiühendites kui osaliselt produktiivsetes mallides.

hoolimata sellest, et nad ei anna alust omaette tähendusloki loomiseks. Sõnaliigimärgendi tähendus konstruktsiooniploki lisamisest ei muutuks: kuna ka praegu on sõnaartiklis olemas täiendavad grammatilise käitumise kirjeldamise vahendid, mõistetakse sõnaliigimärgendit järelikult juba praegu vaikumisi kehtiva üldistusena, millest tulenevad implikatsioonid saab tühistada täpsustavate kirjeldustega. Samuti säiliks võimalus anda grammatilist infot märksõna üksikute tähenduste kohta.

Kuidas täpselt oleks otstarbekas konstruktsioone sõnaraamatus kirjeldada, on siiski keeruline küsimus ja jääb käesolevas artiklis lahtiseks. Artikli eesmärk on eelkõige põhjendada konstruktsioonide kirjeldamise vajalikkust omaette struktuuriüksuses.

Võib tekkida küsimus konstruktsioonide kirjeldamise teostatavusest: peaks ju sõnaraamatu koostaja sellisel juhul lisaks tähendusele analüüsima küllalt detailselt ka sõna süntaktilist käitumist. Sellise analüüsi teostatavuse kasuks räägib siiski see, et nagu eeltoodud näidetest näha, on seda ka seni osaliselt tehtud. Ka on see vajalik ainult kindlat tüüpi sõnaraamatutes, eelkõige õppesõnastikes, ning eelkõige sagedasemate sõnade puhul. Lisaks peaks konstruktsioonide kirjeldus jääma kasutajasõbralikkuse huvides informaalseks nagu senigi ega peaks olema põhjalikum kui tavapärane semantiline ja grammatiline analüüs.

Samuti võib küsida, kas selliste idiosünkraatiliste nähtuste osakaal keeles on piisavalt suur, et nendega sõnaraamatu struktuuris arvestada. Konstruktsioonipõhine keelemudel ennustab, et see on vajalik, kuid seni on siiski tegemist hüpoteesiga, mida tuleb alles kontrollida. See, milline keelemudel on sõnaraamatumudeli aluseks, ongi tähtis seetõttu, et andmestikus märgatakse eelkõige seda, mida oodatakse seal olevat. Kui lähtutakse eeldusest, et sõnad käituvad valdavalt reeglipäraselt ning nende tähendus kontekstis tuleneb nende leksikaalsest tähendusest, siis ei märgata nende võimalikke erandlikke omadusi ega konstruktsioonispetsiifilisi tähendusi. Alles teadlik lähtumine konstruktsioonipõhisest keelemudelist või hüpoteesist, et sellise mudeli ajendiks olevad nähtused on keeles seni eeldatust sagedasemad, teeb võimalikuks nende nähtuste märkamise ja nende ulatuse kindlakstegemise.

Mingi ettekujutuse idiosünkraatiliste nähtuste osakaalust annab see, et juba näiteks toodud parameetersõnad moodustavad üsna suure rühma ning on lisaks suhteliselt sagedased ja ka teataval määral funktsioonisõnad. Kindlasti pole parameetersõnad ainuke selline rühm. Üheks näiteks sellisest sõnaklassist eesti keeles, mida ei saa kirjeldada traditsiooniliste sõnaliigimärgenditega, võib tuua seisundisõnad, millel on nii adverbide kui adjektiivide jooni (vt Tiits 1982). C. Fillmore nimetab eelpool viidatud ettekandes selliste sõnade hulgas ka ajaväljendeid ja arvsõnu. Ühe universaalse grammatiseerumisnähtusena, millega kaasnevad samalaadsed kirjeldusprobleemid, võib nimetada nimisõnade grammatiseerumist nimisõnalises kvantorifraasis (vt Koptjevskaja-Tamm (2001), Traugott (käsikiri)). Näiteks eesti keeles on hulganimisõnadel mitmeid arvsõnade omadusi (EKG II: 144–145, Koptjevskaja-Tamm 2001: 536–540) ning nad võivad olla sihitisena muutumatud (Erelt 2000: 94–95).

Kokkuvõtteks, paradigmast iseseisvunud sõnavormide eripärase süntaksi kirjeldamiseks võiks lisada sõnaraamatu struktuuri fraseologismiplokiga sarnase konstruktsiooniploki ning lisaks esitada sõnavormi märksõnaviitena. Järgmises osas vaatlen konstruktsiooniploki kasutamist konstruktsioonispetsiifilise tähenduse konteksti(de) kirjeldamiseks.

3. Konstruksioonispetsiifiline tähendus

Artikli selles osas illustreerin juhtumit, kus sõna kannab teatud tähendust ainult kindlate konstruksioonide koosseisus, st tegemist on konstruksioonispetsiifilise tähendusega. Niisuguse tähenduse esitamine sõnaraamatus sõna leksikaalse, kontekstivälise tähendusena võib olla eksitav taas eelkõige keeleõppijatele mõeldud sõnastikus. Seetõttu võib osutada vajalikuks sellise tähenduse esinemiskontekstide kirjeldamine, milleks sobib samuti konstruksiooniplokk.

Näiteks konstruksioonispetsiifilisest tähendusest võib tuua sõnaartikli *minema* 6. tähendusploki 'alustama, hakkama, algama' EKSS-is.

(3) **minema**

..
6. protsessi v. tegevust alustama; hakkama, algama, puhkema. *Mootor ei tahtnud käima minna. Usu elektrijaam läheb käiku. Lind läks lendu. Maja läks välgust põlema. Ahi läks hõlpsasti küdema, kütte. Kohvivesi läks keema. Haav läks mädanema. Mahl on käärima, pealt hallitama läinud. Selline kuuldus, kumu on liikuma, liikvele läinud. Istutasin kadaka, aga see läinud kasvama. Koerad läksid purelema. Eks katsu minna rahva vastu! Läheb nüüd sulane peremehega maid jagama! Läksid omavahel vaidlema, nägelema, tülli, riidu, raksu, käsitsi kokku, kaklema, kiskuma. Läks vaidluseks, ütlemiseks, nägelemiseks, tüliks. Nüüd läheb löömaks, mölluks, madinaks, raginaks. Hommikul läheb sõiduks. Traktoristidel kippus palgapäeviti viinavõtmiseks minema. *Kas siis mõni kaupmees teil kahjuga kauplema läheb, hurjuh! O. Luts. | **läks!** las läheb, las minna. Tähele panna! Valmis olla! Läks! *.. (valab Filosoofile veiniklaasi konjakit täis): Läks! E. Rannet.*

See plokk jaguneb tegelikult erinevateks konstruksioonideks, millest osa on kinnistunud ühendid, osa teatud määral produktiivsed, kuid mitte kirjeldatavad verbi *minema* reksioonina. Väljaspool neid kontekste ehk vabalt kombineerudes *minema* sellist tähendust ei kanna.

Esimene tüüp vaatlusaluses tähendusploki illustreeritud konstruksioone on püsiühendi laadsed juhtumid: *käiku / lendu / kütte / liikvele / kasvama / tülli / riidu / raksu / käsitsi kokku minema*. Üleminek sarnase tähendusega püsiühendite rühma ja mingi tähendusrühma piires produktiivse malli vahel on muidugi sujuv, kuid pigem püsiühenditest võib rääkida siis, kui ühendis esinev laiendsõna ei ole ka lähedaste sünonüümidega vabalt asendatav. Näiteks ritta *tülli / riidu / raksu minema* võib lisada ka ühendid *naginasse / raginasse minema*, kuid mitte enam väljendeid *#sõnavahetusse / sõimlusse / rüdllemisse / nägelusse minema*.

Teiseks on tähendusploki illustreeritud mõned produktiivsed konstruksioonid. Esimene neist koosneb verbist *minema* ja füüsilist või verbaalset konflikti väljendava verbi *ma*-infinitiivist: *vaidlema / nägelema / kaklema / kiskuma minema*. Ka siin ei ole tegemist vaba ühendiga, sest *Poisid läksid õues kaklema* tähendab küll 'poisid hakkasid kaklema', kuid *Poisid läksid sööma/mängima* ei tähenda sama mis 'poisid hakkasid sööma/mängima'. Samas on mall produktiivne, st tegemist ei ole üksikute väljendverbidega, vaid selles võivad esineda kõik antud tähendusrühma kuuluvad verbid, mille hulk on avatud, sest sellesse tähendusrühma lisandub

pidevalt ekspressiivseid, kujundlikke või slängisõnu. Konstruksiooni produktiivset kasutust illustreerivad järgmised Google'i otsingumootoriga leitud näited.

(4) Poisid läksid Gerly pärast kätsima :) ! (myass.ee/My%20Web%20Sites/joulupyhad.htm 17.12.2007)

.. mängijad ja treenerid läksid alaliiduga lubatud tasude pärast kraaklema. (etvSPORT.ee/index.php?0539507 17.12.2007)

Keskerakond ja sotsid läksid valimiskomisjoni ukse taga kähmlema. (www.postimees.ee/150404/online_uudised/131607.php 17.12.2007)

Järgmine konstruksioon koosneb *minema* ainsuse 3. isiku vormist ja tegevusnimisõna translatiivist: *vaidluseks / ütlemiseks / nägelmiseks / tüliks / löömaks / mölluks / madinaks / raginaks / sõiduks / viinavõtmiseks minema*. Eriti mine-teonimega, kuid ka teiste tegevusnimisõnadega on see mall üsna produktiivne, nagu näitavad Internetist leitud laused näites 5, kuid ka seda malli ei saa kirjeldada *minema* reksioonimallina, sest *minema* peab selles esinema kindlas vormis.

(5) Peale sööki läks köögi ja koridori tapeetimiseks. (www.misjon.ee/index.php?id=17 17.12.2007)

Pidutsemiseks läheb kella kümnest. (www.ehh.ee/index.php?main_id=8&news_id=154 17.12.2007)

Läks jubedaks kisaks ja sagimiseks. (velo.clubbers.ee/foorum/viewtopic.php?start=175&t=9923 17.12.2007)

Igatahes mai algul läheb tööks ja mõne nädalaga peaks kõik korda saama. (vana.www.sakala.ajaleht.ee/rubriigid.html?number=603 17.12.2007)

Varsti läheb reisiks läbi Praha Dubaisse. (trip.ee/node/17052 17.12.2007)

Veel üks vaadeldavas tähendusplokis illustreeritud mall koosneb verbist *minema* ja *ma*-infinitivis seisundi-, protsessi- või liikumisverbist: *käima / liikuma / põlema / küdema / keema / mädanema / käärima / hallitama minema*. Selles mallis esineda võivate verbide hulk on küll suur, kuid piiratud kindlate tähendusrühmadega. Näiteks tundub, et *minema* ei esine tähenduses 'hakkama' ühenduses psühholoogilist seisundit väljendavate verbidega, vrd #*Läksin kurvastama / rõõmustama*, aga ka mitmete muude protsessi- ja seisundiverbide tähendusrühmadega, vrd *Maja läks põlema*, aga #*Maja läks lagunema, vajuma; Vesi läks keema*, aga #*Vesi läks jahutama, lainetama, purskama, tõusma; Haav läks mädanema*, aga #*Haav läks valutama, veritsema, sügelema, paranema, armistuma*.

Kirjeldatud tähendusplokis on veel kinnistunud vormel *Läks!* ja üks pragmaatilise funktsiooniga lausemall, mida illustreerivad näitelauseid *Läheb nüüd sulane peremehega maid jagama!* ja *Kas siis mõni kaupmees teil kahjuga kauplema läheb, hurjuh!* Kui muuta need laused neutraalseteks väitlauseteks (*Sulane läks peremehega maid jagama, Kaupmees läks kahjuga kauplema*), ei tähenda *minema* enam 'hakkama', ehk see tähendus on antud juhul seotud konkreetse lausemalliga.

Viimaks on plokis veel näitelause *Eks katsu minna rahva vastu!*, kus *minema* ei tundu tähendavat 'alustama' või 'hakkama' ning mis on tähendusplokki sattunud võib-olla sellepärast, et *vastu minema* on siin sünonüümne ühendverbiga *vastu hakkama*.⁴

⁴ Üks osaliselt produktiivne konstruksioon, mis kirjeldatud tähendusplokist puudub, on *minema* 3. isiku vorm kombinatsioonis teonime adessiiviga, nt *Maja läheb lammutamisele* (vt Sahlkai 2005).

Seega võib öelda, et kõik näited, mis illustreerivad verbi *minema* kasutamist tähenduses 'alustama, algama, hakkama', illustreerivad kas konkreetseid püsiühendeid või osaliselt produktiivseid idiomaatilisi malle, mida ei saa esitada *minema* reksioonina, nt kujul [*minema* + *ma*-infinitiiv], sest nad on piiratud kindlate tähendusrühmadega või verbi kindla vormiga. Seega on verbil *minema* tähendus 'alustama, algama, hakkama' ainult kindlates konstruktsioonides. Selle tähenduse esitamine *minema* leksikaalse tähendusena võib olla eksitav eelkõige eesti keele õppijatele suunatud sõnastikes, sest tegemist on tavalise tähendusülekandega ja paljudes keeltes on verb *minema* inhoatiivse või tuleviku abiverbina täielikult grammatiseerunud. Ka siin võiksid lahenduseks olla konstruktsiooniplokid, kus kirjeldatakse põgusalt neid kontekste, milles verb vastavat tähendust kannab. Omaette lahendust nõuavad aga püsiühendid, mis on rohkem käsitletud teema (Muischnek 2006, Veldi 2006).

Küsimus, kas eeltoodud näide illustreerib üksikjuhtu või üldisemat nähtust, nõuab muidugi uurimist, kuid asjaolu, et minemisverbi grammatiseerumine aspektuaalseks või temporaalseks abiverbiks on universaalne nähtus (vt nt Heine, Kuteva 2002: 161–163), viitab pigem sellele, et tegu ei ole ainulaadse näitega. Võib oletada, et nii nagu 2. osas toodud näide kehtib parameetersõnade kohta laiemalt, nii võib ka see näide kehtida üldisemalt verbide kohta, mis võivad abiverbistuda, sest grammatiseerumine algab alati kindlas kontekstis (Heine, Kuteva 2002: 2). Selliste verbide kandidaatideks on tuumverbid (*olema, tegema, tulema, minema, käima, seisma, istuma, saama, võtma, andma, panema, viima, tooma, pidama, võima, hakkama, ajama, laskma, jääma* ning kognitiiv- ja suhtlusverbid – vt Pajusalu jt 2004: 11–12). Tundub, et konstruktsioonispetsiifilisi tähendusi kalduvad omandama ka verbid, mis moodustavad sageli püsiühendeid (*tegema, minema, andma, saama, panema, võtma, ajama, jääma, pidama, lööma* – vt Muischnek 2004: 576), millest omakorda võivad välja kasvada poolproduktiivsed mallid (näiteks reeglipärast *mine*-teonime sisaldavatele poolproduktiivsetele mallidele võivad vastata sarnase tähendusega verbi-nimisõna püsiühendid ning ka verbiühenditel võib olla paralleelseid verbi-nimisõna püsiühendeid, vrd nt *kaklema minema* ja *tüllil minema*).

Kokkuvõtteks, väljapakutud konstruktsiooniplokki saaks kasutada ka juhul, kui sõnal on konstruktsioonispetsiifilisi tähendusi, mille konteksti kirjeldamine on vajalik. Järgmises peatükis vaatlen aga juhtumit, mis illustreerib konstruktsioonide kirjeldamise piire.

4. Produktiivne konstruktsioon

Võimalusel kirjeldada konstruktsioone sõnaraamatus on ka oma piirid: sõna konstruktsioonispetsiifilist süntaksit ja tähendust saab sõnaraamatus kirjeldada üksnes siis, kui sõna on eripärase vormi ja/või tähendusega konstruktsioonis kinnistunud, nii nagu üksikud parameetersõnad oma mallides või verb *minema* artikli 3. osas kirjeldatud konstruktsioonides; ehk teistpidi öeldes, sõnaraamatus saab kirjeldada ainult konstruktsioone, mis sisaldavad kindlaid sõnu või sõnavorme. Kui aga sõna omandab konstruktsioonispetsiifilise tähenduse või süntaksi tänu kombineerumisele produktiivse konstruktsiooniga, milles esineda võivate sõnade hulk on avatud, ei saa sõna eripärast käitumist ja/või tähendust selles konstruktsioonis sõnaraamatus

kuidagi kirjeldada. Teiste sõnadega, sõna esinemist selles konstruktsioonis ei saa esitada tema leksikaalse omadusena, sest selle tingivad konstruktsiooni, mitte sõna kombinatoorsed omadused. See, et tähendust ei saa kirjeldada ega grammatikat annoteerida ainult sõnade loendi kujul ehk sõnaraamatus, on ootuspärane, kui keeleline teadmine ei jagune leksikoniks ja (suurelt osalt sõnade kombinatoorsetele omadustele taanduvaks) grammatikaks, vaid moodustab ühe erineva üldisusastmega sümboolsete üksuste võrgustiku. Seega tuleb konstruktsioonipõhisest perspektiivist lähtudes täiendada sõnaraamatut ja seal sõnaliigimärgenditega indekseeritavat traditsioonilist grammatikat n-ö "konstruktkoniga" sissejuhatuses refereeritud C. Fillmore'i ettekande mõttes (Fillmore 2006).⁵

Näiteks poolproduktiivsest konstruktsioonist, mida ei saa sõnaliigimärgendiga indekseerida ega konstruktsiooniplokis kirjeldada, võib tuua adessiivse viisi- ja põhjusemääruse, mida illustreerivad järgmised näited (lähemalt vt Sakhai 2006).

- (6) Eesti Kirjameeste Seltsi otsusel mindi 1872 üle uue kirjaviisi kasutamisele. (http://www.miksike.ee/docs/referaadid/rahvuslik_liikumine.htm 17.12.2007)

See on lugu kõrgel kohal asjamehest, kes FBI sundimisel pidi üle kahe aasta suurfirma hämaraid asju nuhkima. (<http://movies.ee/uudis/996/17.12.2007>)

Konstruktsioon toimib lauses määrusena ja koosneb adessiivis tegevusnimisõnast ja selle tegevussubjekti vormistavast genitiivsest nimisõnafaasist. Sellist määrust sisaldavad laused koosnevad seega kahest propositsioonist, kusjuures põhipropositsioon (esimeses näitelause *1872 mindi üle uue kirjaviisi kasutamisele*) on tõlgendatav sekundaarse predikaadina toimiva tegevusnimisõna semantilise argumentina. Konstruktsiooni põhjana esinevad sõnad omandavad selles tunnuseid, mida ei saa pidada nende leksikaalseteks omadusteks: teonime semantilisi argumente ei realiseerita laienditena nimisõnafaasist (vrd *?Eesti Kirjameeste Seltsi otsusel hakata looma ühist kirjakeelt mindi 1872 üle uue kirjaviisi kasutamisele.*), mitteprotsessuaalsed teonimed omandavad protsessuaalse tähenduse ja tavaliselt objektatribuudiga kombineeruvad teonimed saavad kombineeruda ainult subjektatribuudiga: nt *soosing* on tavaliselt mitteprotsessuaalse tähendusega ja kombineerub subjektatribuudiga ning *soosimine* protsessuaalse tähendusega ja kombineerub objektatribuudiga,⁶ antud konstruktsioonis aga on mõlemad protsessuaalse tähendusega ja kombineeruvad subjektatribuudiga:

- (7) .. linnavõimude soosingul on avalikud huvid tagaplaanile surutud .. (www.parnupostimees.ee/101106/kommentaarikeskus.php?ARTIKKEL_ID=10069109&TASK=KOMMENTAARID 17.12.2007)
.. tema soosimisel vohab firmas le travail clandestin ehk ebaseaduslik töö. (<http://arhiiv2.postimees.ee:8080/leht/99/02/06/valised.htm> 17.12.2007)

Et konstruktsioonis esineda võivate nimisõnade hulk ei ole piiratud konkreetsete sõnadega, vaid üksnes kindlate tähendusrühmadega, ei saa neid konstruktsioonipetsiifilisi omadusi esitada sõnaraamatus üksikute sõnade omadustena. Näiteks võib Internetiotsinguga leida sellest konstruktsioonist järgmisi sõnavorme, mis

⁵ Ülle Viks juhtis mu tähelepanu sellele, et eesti keele kohta on üks konstruktkonilaadne ressurss juba olemas: Huno Rätsepa "Eesti keele lihtlause tüübid" (Rätsep 1978).

⁶ Protsessuaalsete ja mitteprotsessuaalsete teonimede genitiivatribuudi subjektilisusest ja objektilisusest vt Kasik 1968: 133.

sageduse järgi otsustades ei ole selles konstruktsioonis kinnistunud ja esindavad seega konstruktsiooni produktiivset kasutust: (*midagi tehti kellegi*) *finantstoetusel / inspiratsioonil / kaasaulamisel / kaasamängimisel / kureerimisel / moosimisel / niiditõmbamisel / promomisel / sobitusel / tagantsundimisel / tagantorkimisel...*

Paljud sõnavormid on selles konstruktsioonis siiski kinnistunud ning selliste vormide erandlikku süntaksit ja/või tähendust saab sõnaraamatus kirjeldada konstruktsiooniplokis; näiteks kinnistunud vormidest võib tuua sõnad *abil, ettepanekul, käsul, loal, nõusolekul, osavõtul, saatel, soovitusel, toel...*

Selliste kinnistunud juhtude praegune esitus sõnaraamatus kajastab traditsioonilist eeldust, mille järgi sõna tähendus kontekstis tuleneb tema leksikaalsest tähendusest ning sõna süntaktilise käitumise määravad tema leksikaalsed kombinatoorsed omadused: konstruktsioonispetsiifilist protsessuaalset tähendust on kirjeldatud sõna leksikaalse tähendusena ning erandlikku süntaksit on käsitletud sõnavormi leksikaalse kategooriamuutuse (kaassõnastumise) tulemusena. Sellist lähenemist võib illustreerida vaatlusaluses konstruktsioonis kinnistunud vormide *toimetusel* ja *kirjastusel* esitusega EKSS-is. Sõna *kirjastus* artikli teine tähenduslokk illustreerib konkreetse sõnavormi konstruktsioonispetsiifilise tähenduse omistamist märksõnale, mis võib taas olla eksitav keeleõppija jaoks.⁷

(8) **kirjastus**, -e 5 s.

1. mitmesuguseid trükiseid kirjastav ettevõte. ..

2. kirjastamine. *Romaan ilmus "Postimehe", autori enda kirjastusel.*

Artikkel *toimetus* illustreerib sõna eripärase käitumise kirjeldamist kategooriamuutusena.

(9) **toimetus**, -e 5 s.

1. toimetamine, talitus, tegemine, toiming, töö. ..

2. **toimetusel** (postpositsioonile lähedasena:) kellegi poolt toimetatuna (4. täh.). *J. V. Jannseni toimetusel ilmunud ajaleht. Kooli almanahh anti välja õpilaste endi toimetusel.*

3. ajalehte v. ajakirja väljaandev asutus v. organ, samuti selle v. kirjastuse, raadio mingi osakond; selle tööruum(id). ..

Kaassõnadeks ei saa selles konstruktsioonis kinnistunud sõnu siiski pidada nende tähenduse tõttu, avatud sõnarühma kuulumise tõttu ega võimaluse tõttu kombineeruda adjektiivatribuudiga, nt (10).

(10) Samalt autorilt ja A. L.-i tehnilisel toimetusel ilmus 2000. aastal "Inimese füsioloogia". (http://biomedicum.ut.ee/arfs/ajalugu/tartu_ulikooli_fusioologia_instituut_1981-2006/oppetoo/17.12.2007)

Toodud näited illustreerivad veel kord selliste juhtumite senise käsitlemise variatiivsust; seda näitlikustab ka see, et sünonüümsed vormid *toimetamisel* ja *kirjastamisel* ei kajastu sõnaraamatus üldse, sest reeglipäraseid *mine*-teonimesid üldise reeglina omaette märksõnadena ei esitata.

Ka see on empiirilise uurimise küsimus, kui palju on keeles selliseid produktiivseid konstruktsioone, mida ei saa kirjeldada sõnaliikide terminites. Näiteks toodud konstruktsiooni võib pidada suhteliselt marginaalseks, samas näitab Maria-Maren Sepper (2006: 54–56), et see konstruktsioon on muutunud viimastel aastakümne-

⁷ Samamoodi on esitatud vorm *väljaandel*, selle vahega, et protsessuaalne tähendus ei ole esitatud mitte omaette tähendusena, vaid tähendusvarjundina.

tel üheks peamiseks vahendatud väite vormistamise viisiks ajakirjanduskeeles (nt lauses *Tõnissoni ütlusel tuleb vend tagasi*, Sepper 2006: 49). Peamiselt kasutatakse selleks konstruktsioonis kinnistunud vorme, mida saab sõnaraamatus kirjeldada (nt *sõnul, teatel, väitel*), kuid on ka vorme, mida on selles mallis kasutatud ilmselt produktiivselt (nt olen Internetiotsinguga leidnud sellest konstruktsioonist järgmisi vorme: *ennustusel, hoiatusel, järeldusel, kartusel, lubadusel, meenutusel, mäletamisel, rõhutusel, seletusel, selgitusel, veendumusel, viiteil, viitel*). Need sõnad moodustavad aga ainult ühe konstruktsioonis esineda võivatest tähendusrühmadest (vt Sahlkai 2006).

Eelkirjeldatud konstruktsiooniga sarnased on häält väljendavate *-n*-deverbaalide adessiivivormid (nt *mürinal*), millest paljud on esitatud nt VVS-is adverbidena (VVS: 786–787). Tegemist on siiski produktiivse konstruktsiooniga, milles esinevad vormid ei käitu reeglipäraste nimisõnadena, kuid neid ei saa kirjeldada ka adverbidena, sest nad võivad kombineeruda genitiiv- ja adjektiivatribuudiga, nt (11).

(11) Kui igipedajad metsadest traktorite mürinal lahkuvad .. (www.loodusajakiri.ee/eesti_loodus/EL/vanaweb/9805/kotkad.html 17.12.2007)

Pole aimugi, mis toimub Eestis ja kogu maailmas just nüüdsama, sel hetkel, kui masin monotoonsel mürinal seda Päevalehe soovitud arvamustlugu trükib. (www.epl.ee/artikkel_275118.html 17.12.2007)

Samalaadseks nii sõnaliikide kui vormi- ja sõnamoodustuse piire ähmastavaks nähtuseks eesti keeles võib pidada nimisõnade nn seisundivorme, millest paljud on samuti kinnistunud, kuid mida saab moodustada ka produktiivselt; teiste sõnadega, nimisõnast saab moodustada produktiivselt seisundisõnu, mis langevad vormiliselt osalt kokku ühe käändevormiga, kuid mis ei käitu nagu nimisõnad ja mida ei saa üheselt kirjeldada ka kui adverbe või adjektiive (vt Tiits 1982). Näiteks võib produktiivselt moodustatud seisundivormideks pidada sõnu *rihmas* ja *suukorvis* järgmises lauses (vrd ?*Koer olgu lühikeses rihmas*).

(12) .. looduses olgu koer rihmas ja suukorvis. (www.rm.k.ee/pages.php3/010303 17.12.2007)

Kokkuvõtteks, produktiivses konstruktsioonis kinnistunud sõnade või sõnavormide käitumist ja tähendust on võimalik taas kirjeldada konstruktsiooniplokis, sellises konstruktsioonis produktiivselt esinevate sõnade idiosünkraatilist käitumist ja/või tähendust saab aga kirjeldada ainult “konstruktsioonis”.

5. Kokkuvõte

Vaatlesin artiklis kolme tüüpi nähtusi, mis on problemaatilised nii leksikaalsel projektsioonil ja kompositsioonilisel semantikal põhinevate keeleteooriate jaoks kui traditsioonilise sõnaraamatumudeli jaoks. Nendeks olid esiteks idiosünkraatilises morfosüntaktilises mallis kinnistunud sõnavormide erandlik grammatiline käitumine, teiseks sõna esinemine mingis tähenduses ainult kindlates konstruktsioonides ja kolmandaks sõnavormide produktiivne esinemine eripärase vormi ja tähendusega konstruktsioonis. Selliste nähtuste kajastamine sõnaraamatus on problemaatiline, sest idiosünkraatilist grammatilist käitumist ei saa kirjeldada või

indekseerida traditsioonilise sõnaliigimärgendiga ning konstruktsioonispetsiifilist tähendust ja grammatikat ei saa esitada märksõna leksikaalsete omadustena. Konstruktsioonipõhine keelemudel aga ennustab, et selliseid nähtusi on keeles rohkem kui seni eeldatud, mistõttu nende kirjeldamine on vajalik eriti keeleõppijatele suunatud sõnastikes. Pakkusin välja, et kahte esimest nähtust, sõnavormide eripärast käitumist ja konstruktsioonispetsiifilise tähenduse esinemiskontekste, võimaldaks sõnaraamatus kirjeldada fraseologismiploki taolise konstruktsiooniploki lisamine sõnaartikli struktuuri. Lähtumine eeldusest, et idiosünkraatilisust ja idiomaatilisust on sõna(vormi)de käitumises ja tähenduses seni arvatust rohkem, tähendab muidugi ka seda, et sõnaraamatu koostaja peab seda andmestikust süstemaatiliselt otsima ja andmestikku senisest põhjalikumalt analüüsima. Sellise analüüsi teostatavuse kasuks räägib siiski see, et nagu artiklis toodud sõnaraamatunäidetest ilmneb, on seda ka seni osaliselt tehtud; samuti ei pea selline analüüs olema põhjalikum tavapärasest semantilisest ja grammatilisest analüüsist ning konstruktsioonide kirjeldus ei pea olema ammendav ega formaalne.

Kolmandat tüüpi nähtust, produktiivses konstruktsioonis omandatud erandlikku käitumist ja konstruktsioonispetsiifilist tähendust, ei saa aga kirjeldada ka konstruktsiooniplokis. See on ootuspärane, kui keeleline teadmine ei jagune leksikoniks ja grammatikaks, vaid moodustab ühe erineva üldisusastmega sümboolsete üksuste võrgustiku. Seetõttu tuleb sõnaraamatut ja seal sõnaliigimärgenditega indekseeritavat traditsioonilist grammatikat paratamatult täiendada n-ö “konstruktioniga”, kus on võimalik kirjeldada ka produktiivsete konstruktsioonide vormi, tähendust ja kombinatoorseid omadusi.

Eelpool kirjeldatud nähtuste osakaal keeles ja seega nende kirjeldamisega seotud probleemi ulatus tuleb küll alles empiirilisel kindlaks teha, kuid artiklis püüdsin näidata, et ka juba olemasolevate andmete põhjal võib öelda, et tegemist ei ole üksiknähtustega. Selliste nähtuste avastamise ja märkamise tingimuseks on lähtumine keelemudelist, mis eeldab nende olemasolu. Seetõttu on ka sõnaraamatu koostamisel tähtis, millisest keelemudelist lähtutakse.

Kirjandus

- Croft, William 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- EKG II = Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1993. *Eesti keele grammatika. II. Süntaks*. Kiri. Tallinn: Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut.
- EKSS = Eesti kirjakeele seletussõnaraamat. I–VII. Tallinn: Eesti Keele Instituut, 1988–2007.
- Erelt, Mati 2000. Sihitis. – Tiiu Erelt, Mati Erelt, Maire Raadik, Tiina Leemets, Sirje Mäearu. *Keelenõuanne soovitab 2*. Tiiu Erelt, Maire Raadik (koost., toim.). Eesti Keele Instituut. Tallinn: Eesti Keele Sihtasutus, 91–97. Vt ka <http://www.eki.ee/keeleabi/artiklid2/sihitis.html> (17.12.2007).
- Fillmore, Charles J. 2006. The articulation of lexicon and constructicon. – *Plenaaretkanne 4. rahvusvahelisel konstruktsioonigrammatika konverentsil Tokyo ülikoolis* 1.–3.09.2006.
- Fillmore, Charles J.; Kay, Paul; O'Connor, Mary Catherine 1988. Regularity and idiomacity in grammatical constructions. – *Language* 64, 501–538.
- Goldberg, Adele E. 1995. *Constructions: a Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

- Heine, Bernd; Kuteva, Tania 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Kasik, Reet 1968. Omastavalise täiendi subjektilisusest ja objektilisusest *mine*-konstruktsioonis. – *Keel ja struktuur* 2. Tartu: Tartu Riiklik Ülikool, 125–139.
- Kay, Paul; Fillmore, Charles J. 1999. Grammatical constructions and linguistic generalizations: The *What's X doing Y?* construction. – *Language* 75, 1–33.
- Koptjevskaja-Tamm, Maria 2001. “A piece of the cake” and “a cup of tea”: Partitive and pseudo-partitive nominal constructions in the Circum-Baltic Languages. – Östen Dahl, Maria Koptjevskaja-Tamm (eds). *The circum-Baltic languages: typology and contact*. Vol. 2. Grammar and Typology. *Studies in Language Companion Series* 55. Amsterdam, Philadelphia: Benjamins, 523–568.
- Muischnek, Kadri 2004. Verbi ja noomeni püsiühenditest eesti keeles. – *Keel ja Kirjandus* 8, 574–589.
- Muischnek, Kadri 2006. Verbi ja noomeni püsiühendid eesti keeles. *Dissertationes philologiae estonicae Universitatis Tartuensis* 17. Tartu: Tartu Ülikooli Kirjastus.
- Pajusalu, Renate; Tragel, Ilona; Veismann, Ann; Vija, Maigi 2004. Tuumsõnade semantikat ja pragmaatikat. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 5. Tartu: Tartu Ülikooli Kirjastus.
- Rätsep, Huno 1978. Eesti keele lihtlausete tüübid. *Emakeele Seltsi toimetised* 12. Tallinn: Valgus.
- Sahkai, Heete 2005. Teonimi perifrastilises verbivormis. – *Keel ja Kirjandus* 10, 790–807.
- Sahkai, Heete 2006. Konstruktsioonipõhise keelekirjelduse võimalustest adessiivse viisi- ja põhjumsääruse näitel. – *Keel ja Kirjandus* 10, 816–831.
- Sahkai, Heete 2007. Keelemudel ja sõnaraamatumudel. – Ettekanne VI rakenduslingvistika konverentsil “Keel ja leksikon”, 26.-27. aprillil 2007 Tallinnas.
- Sepper, Maria-Maren 2006. Indirektaal eesti 19. sajandi lõpu ja 20. sajandi aja- ja ilukirjanduskeeles. *Magistritöö*. Käsikiri Tallinna Ülikoolis. Tallinn: Tallinna Ülikool.
- Tiits, Mai 1982. Seisundiadverbidest. – *Keel ja Kirjandus* 1, 17–21.
- Traugott, Elizabeth Closs (käsikiri). *Constructions, emergent constructions, and the notion of ‘newness’*. <http://dialoguematters.stanford.edu/twiki-dm/bin/viewfile/Main/WorkshopThree?rev=1.2;filename=TraugottEmergCxns.doc> (17.12.2007).
- Veldi, Enn 2006. Uusi võimalusi ühendteguõnade kajastamiseks eesti-inglise sõnaraamatutes. – Ettekanne V rakenduslingvistika kevadkonverentsil “Keel töises suhtluses”, 20.-21.aprill 2006, Tallinn.
- VVS = Viks, Ülle 1992. Väike vormisõnastik II. *Sõnastik ja lisad*. Tallinn: Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut.

Heete Sahkai (Eesti Keele Instituut) on lõpetanud Eesti Humanitaarinstituudi romaani filoloogia erialal. Kaitses magistrkraadi Tartu Ülikoolis 2002. a prantsuse filoloogia alal. Alates 2002. a töötab Eesti Keele Instituudis ja on Tallinna Ülikooli doktorant eesti filoloogia erialal. Osaleb “Suure eesti-prantsuse sõnaraamatu” koostamises. On uurinud eesti keele süntaksit. Heete.Sahkai@eki.ee

SOME LEXICOGRAPHIC IMPLICATIONS OF A CONSTRUCTION-BASED MODEL OF LANGUAGE

Heete Sahkai

Institute of the Estonian Language

Construction-based theories of language are the result of the acknowledgement that idiosyncratic and idiomatic phenomena constitute an important part of linguistic knowledge and are often subject to partial regularities. Such phenomena pose problems for many alternative linguistic theories based on lexical projection and compositional semantics, as well as for certain important assumptions of lexicography. One of these is the assumption that distributional patterns divide words into a relatively small set of parts of speech so that the grammatical behaviour of words can be indexed in a dictionary with part of speech tags. Another assumption is that words are relatively autonomous entities and thus that their meaning and behaviour in context is determined by their inherent semantic and combinatorial properties.

The paper explores three types of phenomena that pose problems for current lexicographic practices based on these assumptions. The first of these is the exceptional syntactic behaviour of words or word forms that have become entrenched in an idiosyncratic pattern which cannot be described in terms of parts of speech, the second is construction-specific meaning, and the third is idiosyncratic behaviour and/or construction-specific meaning acquired by words or word forms in a productive construction.

It is proposed that the occurrence of the head word or some of its forms in a construction in which it displays idiosyncratic behaviour and/or construction-specific meaning should be described in a separate structural unit in the dictionary article, similarly to its occurrence in a phraseological unit. It is argued that omitting information about constructions or conflating it with information about the inherent semantic and grammatical properties of the head word would be misleading, especially in learners' dictionaries. The possibility to describe constructions in a dictionary is however limited, as is illustrated by the third phenomenon examined in the paper: productive constructions that contain no fixed lexical material can only be represented in an eventual "constructicon" complementing the lexicon and the traditional grammar.

Keywords: lexicography, Construction Grammar, syntax, Estonian

ÜHEST KATSEST EMOTSIOONIVERBIDEGA

Anni Silk

Ülevaade. Eesti- ja saksakeelsete emotsiooniverbide võrdlevaks semantilis-süntaktiliseks uurimuseks oli kõigepealt vaja koostada loend nimetatud verbide kohta. Seejuures tekkis paratamatult küsimus, mis need emotsiooniverbid õigupoolest on ehk kus on piir emotsiooni- ja mitte-emotsiooniverbide vahel. Ajendatuna arvukatest emotsioonimõistetest ning mitmesugustest emotsioonisõnade nimekirjadest, mille koostajad on eeldatavasti lähtunud erinevatest põhimõtetest, otsustasin küsitleda eesti ja saksa keelt emakeelena kõnelejaid, lastes neil otsustada n-õ piiripealsete sõnade kuulumise emotsiooniverbide hulka. Artiklis antakse ülevaade emotsiooniverbide kindlakstegemisest, kirjeldatakse tehtud katset lähemalt ning püütakse analüüsida saadud tulemusi. Katse tulemused olid mõneti ootuspärased, hõlmates nii üpris üksmeelseid valikuid kui ka vastuolusid näiteks samatüveliste või -sisuliste sõnade hindamisel. Näib, et katseisikud lähtusid sõnade liigitamisel suhteliselt laiaast emotsiooniverbide määratlusest, mis hõlmab muu hulgas ka hulgaliselt hinnangulisi kõneaktiverbe.*

Võtmesõnad: leksikaalne semantika, emotsioonisõnad, eesti keel, saksa keel

Uurimisvaldkond ja eesmärk

Vajadus koostada emotsiooniverbide nimekiri tekkis seoses mahukama uurimusega, mille eesmärgiks on võrdlevalt uurida emotsiooniverbide semantika ja süntaksi vahelisi seoseid eesti ja saksa keeles. Et uurimusse taheti kaasata võimalikult palju kummaski keeles olemasolevaid emotsiooniverbe, ei sobinud nende kogumiseks välimeetod, nagu seda on kasutanud nt Ene Vainik (2001, 2004), kuna küsitluse käigus ei pruugi katseisikule kõik sõnad meenuda ning osa nimetatud sõnadest kuulub emotsiooniverbide semantilise välja perifeeriasse vaid põhjuslikkuse või assotsiatiivseoste kaudu (Vainik 2002a: 74, 2002b: 548) või jääb sellest rangelt võttes hoopis välja (Vainik 2002b: 548). Kui eesmärk on saada võimalikult täielik pilt

* Artikkel on valminud ETF-i grandri nr 7149 osalisel toel.

keeles olemasolevatest emotsiooniverbidest, ei ole väga tõenäoline, et keskmisele keelekasutajale meenuvad küsitluse käigus ka harvem kasutatavad tundevarjundeid tähistavad sõnad, ehkki ta neid küll teab ja tunneb. Sellised sõnad on aga samuti emotsioonisõnavara oluline osa.

‘Keeleteaduslikuks uurimuseks ei sobi hästi ka psühholoog Jüri Alliku rakendatud meetod, millega valiti emotsioonisõnade nimekirja “enamik eesti keeles sagedamini kasutatavaid ja üldarusaadavaid tundmusi või emotsioone väljendavaid sõnu” (Allik 1996: 8), kusjuures loetelu pidi tingimata sisaldama 60-sõnalise inglise keele emotsioonide nimekirja tõlkevasteid. Kui tahetakse uurida emotsioonisõnu ühes keeles, ei ole mõttekas võtta sõnade nimekirja üle teisest keelest, sest sageli ei ole ühe keele emotsioonisõnale teises keeles täpset vastet ja nii võib tõlkimisel tekkida segadus. Näiteks võib ingliskeelse sõna *disgust* tõlkida eesti keelde nii ‘vastikus, jälkus, tülgastus, tülkus, läilus’ kui ka ‘kibe pettumus’, eesti *kahjurõõmule* ei ole täpset vastet gruusia keeles ja jaapanikeelsele *amae*-le ei ole teistes keeltes üldse vastet.

Nimetatud põhjustel otsustasin nii eesti kui ka saksa keele emotsiooniverbid koguda kirjalikest allikatest, nagu seda on teinud ka nt Wilhelm Janke ja Günter Debus (vt Debus 1988), Marcos Herrera Burstein (1997) ning Rolf Kailuweit (2005), kes koostasid nimekirjad peamiselt sõnaraamatutes leiduva keeleainese põhjal. Et uurijad ei ole üheselt defineerinud emotsiooni ja emotsioonisõna, siis sõnastasin enne verbide kogumist olemasolevate käsitluste ning nimekirjade põhjal enda jaoks emotsiooniverbide definitsiooni, mida üritasin nimekirju koostades silmas pidada.

Definitsiooni alusel kogutud verbide seas oli aga terve hulk selliseid, mille puhul ei olnud kerge otsustada, kas ikka on tegemist emotsiooniverbiga. Üks võimalus on sel juhul lasta katseisikutel hinnata, kas või millisel määral sellised sõnad uuritavasse semantilisse välja kuuluvad. Nii toimisingi ebaselge staatusega verbidega – katseisikud said verbide loetelu palvega märkida selles ära sõnad, mis nende arvates on emotsiooniverbid. Kuna puudub objektiivne kriteerium selle kohta, mille alusel mingit sõna emotsioonisõnaks pidada (Tischer 1993: 54), siis otsustasin analoogselt L. Schmidt-Atzertiga (1980) lugeda emotsiooniverbideks sõnad, mida pidasid nendeks vähemalt pooled katseisikutest.

Järgnevalt kirjeldatakse mõnda varasemat uurimust emotsioonisõnade kindlakstegemise kohta, antakse ülevaade tehtud katsest, mille eesmärgiks oli teha kindlaks, milliseid emotsiooniverbide semantilise välja perifeeriasse kuuluvaid verbe peavad emotsiooniverbideks eestlased ning milliseid sakslased, ning võrreldakse saadud tulemusi ja arutletakse nende võimalike põhjuste üle. Eelkõige pidid tehtud katse tulemusena selgunud verbid olema aluseks eesti ja saksa keele emotsiooniverbide võrdleval semantilis-süntaktilisele analüüsile, ent selgus ka muud huvitavat informatsiooni emotsiooniverbide semantilise välja liigendamise kohta kummaski keeles.

Emotsioonisõnad ja nende kindlakstegemine

Sarnaselt emotsioonidega, mida on uuritud ja üritatud defineerida juba antiikajast peale ning mille olemuses ei ole siiani üksmeelele jõutud,¹ puudub tänaseni ka emotsioonisõnade kindel definitsioon. Sageli on emotsioonisõnadega seotud uurimuste keskmes olnud emotsioonisõnade kategoriseerimine (eri keeltes) (nt Marx 1982, 1997, Schmidt-Atzert 1987, Russell 1991, Allik 1996), hoopis harvem on emotsioonisõnu uuritud nende keelelistest aspektidest (nt Burstein 1997, Siironen 2001, Toluk 2002, Kailuweit 2005). Ehkki enamasti ei olegi täpselt defineeritud, mis emotsioonisõnad on, on vahel täpsustatud nende funktsiooni – nii on emotsioonisõnadeks peetud sõnu, mis on keeles emotsioonide või tundmuste tähistamiseks (Õim 1971, Schmidt-Atzert 1987), väljendamiseks (Allik 1996) või iseloomustamiseks (Jäger, Plum 1988), on seotud tajuga ning väljendavad hinnangulisi ja mentaalseid seisundeid ning protsesse (Fries 2007). Lisaks emotsioonisõnade erinevale määratlemisele ning uuritavate sõnade liigilise kuuluvuse² ja hulga varieeruvusele on ka emotsioonisõnu ennast nimetatud erinevalt – *affective lexicon* 'afektiivne leksikon' (Ortony jt 1987), *Gefühlswörter* 'tundesõnad' (Fries 2000, 2002, ka Toluk 2005), *psychische Wörter* 'psüühilised sõnad' (Fries 2007), *psych-verbs* 'psüühilised verbid' (Klein, Kutscher 2005) või *psychological verbs* 'psühholoogilised verbid' (Vanhoe 2002, Caluianu 2007).

Olgugi et enamik uurijaid on üksmeelel selles, et emotsioonisõnade semantiline väli on ilma selgete ja üheselt kindlaks määratud piirideta (vt nt ülevaateid Debus 1988, Jäger, Plum 1988) ning et emotsioonisõnade kindlakstegemiseks ei olegi võimalik leida mingeid kriteeriume, mille alusel otsustada, kas sõna on emotsioonisõna või ei ole (Fehr, Russell 1984), on Andrew Ortony, Gerald Clore ja Mark A. Foss (1987) väitnud, et peab olema võimalik tuua välja need reeglid, mille alusel saab määrata, kas sõna kuulub emotsioonisõnade hulka või ei kuulu. Analüüsinud u 500 sõna semantiliselt, jõuavad nad järeldusele, et emotsioonisõnade kõige paremateks esindajateks on sõnad, mis kirjeldavad inimese hinges toimuvaid selge seisundimääratlusega protsesse ja mida saab seostada tunnetest lähtuva hinnanguga mingile objektile, sündmusele või inimesele (Ortony jt 1987: 358). Otsustamisel soovivad nad kasutada reeglit, mille kohaselt tõelist emotsioonisõna tajutakse sellena nii lauses "Ma olen X" kui ka lauses "Ma tunnen ennast X-ina". Nii tajutakse sõna *angry* 'vihane' mõlemas lauses emotsioonisõnana, aga nt *abandoned* 'mahajäetud' omandab emotsionaalse varjundi alles teises lauses ja sedagi tänu verbile *tundma*. Seega ei ole *mahajäetud* tõeline emotsioonisõna (vt Ortony jt 1987: 346–347).

A. Ortony jt (1987) jagavad kogu emotsioonidega seotud sõnavara ehk afektiivse leksikoni kolme alagruppi: sõnad, 1) mis kirjeldavad väliseid olusid (nt *abandoned* 'mahajäetud', *beloved* 'armastatud'), 2) millega tähistatakse füüsiliselt tajutavaid seisundeid (nt *sleepy* 'unine', *hungry* 'näljane'), ning 3) mis kirjeldavad inimese hingeelu ja psüühilisi või mentaalseid seisundeid. Viimase grupi jagavad nad omakorda pigem kognitsiooniga seotud (*interested* 'huvitatud'), pigem afektiivsusega

¹ Erialases kirjanduses leidub kümneid erinevaid emotsioonimääratlusi, alates füsioloogilistest ja neuronaalsetest kuni kognitiivsete ja käitumuslikeni. Kõige sagedamini on emotsioone iseloomustavate tunnustena nimetatud nendega kaasnevat näoilmet, mingile objektile suunatust ja sellele kognitiivse hinnangu andmist, võimatust neid kehas lokaliseerida ning mingit välist ärritajat nende esilekutsujana. Tihti peale ei esine need tunnused kõik koos, vaid mingi kombinatsioonina, ent üldiselt on emotsioon siiski kompleksne psühhofüsioloogiline protsess, mis tekib vastuseks mingile välisele stiimulile ja mis annab tõuke olla valmis mingiks käitumuslikuks muutuseks (Silk 2006, vt ka Mandl, Euler 1983: 5–11, Ulrich 1989: 31–38, Merten 2003: 13–17, Mitmannsgruber 2003: 15–27).

² Erinevates emotsioonisõnade nimekirjades on emotsioonisõnadena käsitletud kas ainult adjektive või ainult substantiive, nii adjektive kui ka substantiive või sõnu erinevatest sõnaliikidest, st nii substantiive, adjektive, verbe kui ka adverbe.

seotud (*glad* 'rõõmus', *scared* 'hirmunud') ning nii kognitiivse kui afektiivse komponendiga sõnadeks (*apathetic* 'ükskõikne', *discouraged* 'heitunud'). Seejuures olevat just teine grupp ehtsad, n-ö kesksed emotsioonisõnad. Lisaks kognitiivsele ja afektiivsele komponendile eristavad A. Ortony jt mentaalsete seisundite kohta käivates sõnades veel käitumuslikku komponenti, mis esineb koos kognitiivse või afektiivse komponendiga.

Ehkki selline sõnade jagamine komponentide alusel näib üsna loogiline ja selgepiiriline, ei ole siiski alati lihtne kindlaks teha, milline komponent mingis sõnas domineerib. Autorid isegi on tunnistanud, et toetuda tuleb pigem intuitsioonile (Ortony jt 1987: 352). Samuti jääb arusaamatuks, miks liigitatakse näiteks *aggravated* 'ärritatud' ja *delighted* 'vaimustatud' afektiivsete, aga *alarmed* 'ärevuses' ja *admiration* 'vaimustus, imetus' afektiiv-kognitiivset seisundit väljendavate sõnade hulka või miks ei peeta sõnu *baffled* 'hämmeldunud' ja *amazed* 'hämmastunud' afektiivse, vaid kognitiivse komponendiga sõnadeks. Samuti tundub, et tähenduselt sarnased sõnad on kord ühes, kord teises grupis, nt *stunned* 'vapustatud' loetakse kognitiivse domineeriva, *shaken* 'vapustatud' afektiivse domineeriva komponendiga sõnaks.

Näiteid emotsioonisõnade määramise vastuolulisusest leiab mujaltki. Näiteks erinevates uurimustes käsitletakse emotsioonisõnadena erinevatesse sõnaliikidesse kuuluvaid sõnu. Valdavalt on emotsioonisõnadena uuritud kas emotsioonimõistega seonduvaid substantiive (nt Marx 1982, Schmidt-Atzert 1987, 1997, Goozen, Frijda 1993, 1997) või adjektiive (nt Allik 1996) või on uuritud mõlema sõnaliigi esindajaid segamini (nt Õim 1971, Ortony jt 1987, Russell 1991, Frijda jt 1995). Harvem on uurimisobjektiks olnud vaid emotsiooniverbid (nt Siirainen 2001, Toluk 2002, Kailuweit 2005, Caluianu 2007). E. Vainiku eesti keele emotsioonisõnavara uurimus (2001, 2004) hõlmab keelendeid sõnaliigist sõltumata, kusjuures domineerivad substantiivid (73%), adjektiive ja verbe on enam-vähem võrdses koguses (Vainik 2002: 252). Küllap on arusaam substantiividest ja adjektiividest kui tüüpilisematest emotsioonisõnadest mõnes mõttes ka mõistetav – moodustavad valdavalt just substantiivid keelekategoriate n-ö põhitasandi (vt nt Kövecses 1995: 4), mis keelt õppides esimesena omandatakse (vt nt Argus 2004: 37) ning mis loetelukatsetes sageli esimesena meenuvad (Vainik 2002), adjektiivi abil aga samastatakse ennast emotsiooniga (nt *Ma olen kade*) (Allik 1996: 8).

Teiseks võib nimetada märkimisväärseid erinevusi erinevates uurimustes kasutatud emotsioonisõnade loeteludes. See, et põhiemotsioonideks peetav *rõõm* või *kurbus* ja nende vasted teistes sõnaliikides esinevad kõigis emotsioonisõnade nimekirjades (Schmidt-Atzert, Ströhm 1983: 131), on üsna loomulik, ent küsitavused tekivad siis, kui üks ja sama sõna (või selle tuletised) kord on, siis jälle ei ole loetud emotsioonisõnaks. Kui võrrelda näiteks A. Ortony jt (1987) koostatud inglisis-, Jüri Alliku (1996) ja Ivi Toluki (2002) eesti-, Mari Siiraineni (2001) soome- ning Rolf Kailuweiti (2005) prantsus- ja itaaliakeelset emotsioonisõnade loetelu, siis lisaks nende erinevale pikkusele ning sellele, et kolm viimast koosnevad verbidest, jääb mulje, et just verbid ja nendest saadud tuletised (partitsiibid, kesksõnad) on need sõnad, mis tekitavad vaidlusi oma kuulumise üle emotsioonisõnade hulka. Nii peab vaid R. Kailuweit sks verbe *trauen*, *vertrauen* 'usaldama' emotsioonisõnaks, A. Ortony jt liigitavad selle ingliskeelse vaste *intimate* kognitiivsete mitteafektiivsete meeelseisundite hulka ning ülejäänud kolmes loetelus seda sõna ei esine. *Alan-*

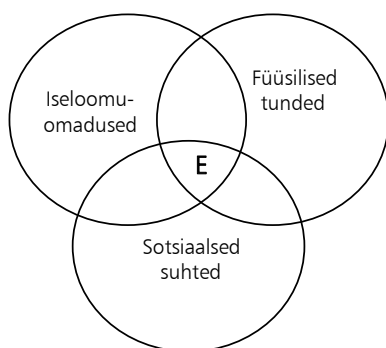
dama ja *alandatud* toovad oma loetelus ära nii J. Allik, I. Toluk kui ka A. Ortony jt (kusjuures ingl *submissive* 'alandlik, alistuv' ei ole afektiivne seisund ja seega ka mitte n-ö õige emotsioonisõna), R. Kailuweit mainib *alandama* kui mitte-emotsioonisõna ning M. Siironeni loetelust selle vastet ei leia. Vastuolusid on veelgi: R. Kailuweit peab emotsiooniverbiks pr *se flatter* 'lootust hellitama', A. Ortony jt liigitavad ingl *expectant* ja *hopeful* '(l)ootusrikas' kognitiivsete mitteafektiivsete meeleseisundit väljendavate sõnade hulka; nii R. Kailuweit kui ka M. Siironen välistavad emotsioonisõnadena pr *désapprouver* 'hukka mõistma', pr *reprimander* 'taunima' ning sm *paheksua* 'pahaks panema', sama meelt tundub olevat I. Toluk, samas kui A. Ortony jt nimetavad ingl verbe *disapprove-of* 'pahaks panema' ning sama tüüpi *approve-of* 'heaks kiitma' afektiivse komponendiga sõnade loendis. Üldse paistab nimetatud uurimuste aluseks olnud loetelusid võrreldes silma, et J. Alliku emotsioonisõnade hulka kuuluvad ka mitmed n-ö iseloomu- või A. Ortony jt liigituse järgi välist olukorda või füüsilist seisundit kirjeldavad sõnad (nt *vaenulik, tülinoriv, hooletu, laisk, haiglane, unine, uljas, petetud, mahajäetud*). A. Ortony jt kesksete emotsioonisõnade hulgast leiab aga mitmeid üsna ootamatuid sõnu, mida enamasti on peetud pigem kognitiivseteks sõnadeks (nt *forgive* 'andestama', *grateful* 'tänulik', *hope* 'lootma', *want* 'tahtma').

Nimetatud viies uurimuses toodud sõnaloendeid võrreldes võib öelda, et M. Siironen on lähtunud kõige kitsamast emotsiooniverbi määratlusest (loendi tegusõnad on valdavalt seotud asjaomase nimisõnalise emotsiooninimetusega). Ehkki ka I. Toluk olevat lähtunud samadest kriteeriumitest nagu M. Siironengi, sisaldab tema sõnaloend pigem kognitiivse komponendiga verbe (nt *kahtlema, kõhklema, ebalema*), mida ta seostab tundmusega (Toluk 2002: 14), ning ka mitmeid pigem iseloomu kirjeldavate sõnade hulka kuuluvaid verbe (nt *juhmuma, kalgistama, arastama, julgema, kurjuma, tigetsema, tigetama*). Iseloomusõnu esineb suhteliselt palju ka J. Alliku emotsioonisõnade nimekirjas (nt *tragi, laisk, vaenulik, jõuline, vapper, haiglane*).

Ehkki nendest emotsioonisõnade loenditest kolm on koostatud lingvistilise (eelkõige sõnade süntaktilis-semantilise käitumise, aga ka sõnamoodustuse) uurimuse jaoks (R. Kailuweit, M. Siironen, I. Toluk), J. Alliku oma psühholoogias rakendatava positiivsete ja negatiivsete emotsioonide mõõtmise skaala loomiseks ning A. Ortony jt olidki oma uurimuses keskendunud emotsioonisõnade kindlakstegemisele, siis erinev eesmärk ei tundu olevat loetelu koostamist oluliselt mõjutanud. Teistest eristub eespool nimetatud põhjustel vaid pisut M. Siironeni loetelu.

Loetletud näited peaksid olema ilmekaks tõestuseks, et emotsioonisõnad ei moodusta mingit kindlaksmääratud homogeenset hulka, mida saab ühe reegli abil kindlaks teha ning mis on alati ja ainult emotsioonisõnad. Pigem tuleb nõustuda nendega, kes soovivad ka emotsioonisõnade ja -mõistete puhul lähtuda prototüübi-teooriast (Fehr, Russell 1984, Kövecses 1995: 9), mille kohaselt võrreldakse objekti ja tema omadusi vastava kategooria tüüpilise esindaja ehk prototüübiga. Erinevate kategooriate piirid on hägusad ja üleminekud sujuvad, seega saab ka sõnade kuulumist emotsioonisõnade hulka pidada kas rohkem või vähem tõenäoliseks, mitte tõeseks või valeks. Emotsioonisõnad moodustavad pigem võrgustiku, mille liikmed (ülem-, alam- ja põhitasandi mõisted; erinevatesse sõnaliikidesse kuuluvad sõnad) on omavahel seotud mitmesuguste suhete kaudu. Emotsioonisõnade semantiline väli ei ole homogeenne ega teistest semantilistest väljadest rangelt eraldatud, vaid

tema n-õ tuum on kompaktne, äärealad aga kattuvad teiste, semantiliselt lähedal seisvate väljade või kognitiivsete valdkondadega (iseloomuomadusi, füüsilisi tundeid ning sotsiaalseid suhteid kajastavad semantilised väljad, vt joonis 1). Tuumiku moodustavad n-õ prototüüpsed emotsioonisõnad, emotsioonide põhinimetustega seonduvad sõnad, nagu *viha, kurvastama, rõõmus* (neid sõnu mainitakse ka loetelukatsetes sagedamini ja esimestena, Vainik 2001b: 67); semantilise välja äärealal ehk perifeerias paiknevad sõnad, mis võivad kuuluda samal ajal nii emotsiooni- kui ka taju-, meeleolu- või iseloomusõnavarasse ning mille kindla kuuluvuse üle on seega ka raskem otsustada (nt *üleannetu, hooletu, uljus, hullus, julgema, sütitama, valutama*).



Joonis 1. Emotsioonisõnade semantiline väli (E) teiste semantiliselt lähedaste väljade suhtes. (Joonise koostamisel on võetud aluseks emotsioonikategooria paiknemine Vainik 2004: 39)

Eesti- ja saksakeelsete emotsiooniverbide nimekirja koostama hakates tuli ikka ja jälle esitada küsimus, mis need emotsiooniverbid on ehk kus jookseb piir emotsiooniverbide ja mitte-emotsiooniverbide vahel. Seni kohatud emotsioonisõnade nimekirjades (Ortony jt 1987, Schmidt-Atzert 1987, Tischer 1993, Frijda jt 1995, Allik 1996, Siironen 2001, Toluk 2002, Kailuweit 2005) esinenud verbe³ võrreldes jagasin emotsiooniverbid enda jaoks kõigepealt kaheks. Emotsiooniverbid kitsamas mõttes tähistavad mingi emotsionaalse seisundi või emotsiooni esilekutsumist, tekkimist või kogemist ning need moodustavad tuumosa ja ka enamuse (nt *jahmuma, rõõmustama, vihkama, üllatuma*). Emotsiooniverbid laiemas mõttes (perifeeria) hõlmavad tundeeluga üldisemalt seotud verbe, mis ei ole otseselt seotud ühegi kindla emotsiooni või tundeelamusega (nt *tundma, tunnetama, aimama*) või väljendavad pigem meeleolu, tuju või suhtumist (nt *meeldima, sallima, mornitsema, kiuslema*). Kuna aga ka meeleolu on seotud tundeeluga, ainult et emotsioonidest eristab seda pikem kestus, väiksem intensiivsus ja puuduv objekt (Schmidt-Atzert 1981: 30, Mitmannsgruber 2003: 30–32), siis lugesin ka need sõnad emotsiooniverbide hulka kuuluvaks.⁴

Ehkki teaduskirjanduses puudub tänaseni ühtne arusaam emotsioonide olemusest ning emotsioonisõnugi ei ole üheselt defineeritud, pidasin siiski mõistlikuks enda uurimuse jaoks emotsiooniverbid täpselt määratleda. Püüdes hõlmata emotsiooniverbe nii kitsamas kui ka laiemas mõttes, defineerisin emotsiooniverbid

³ Frijda jt (1995) käsitlevad nii substantiive, adjektiive, adverbe kui verbe, Schmidt-Atzert (1987) ja Tischer (1993) õigupoolest küll ainult emotsioone tähistavaid substantiive. Substantiivid transformeeriti (nagu Schmidt-Atzert (1981) toimis teiste sõnaliikidega, et saada vaid substantiive) aga verbideks, mis tähistasid omakorda vähem emotsioonidele ja rohkem käitumisele viitavate tegevusi (nt *Dankbarkeit* 'tänuликkus' – *tänama*, *Stolz* 'uhkus' – *uhkustama*) või emotsionaalsete protsesside läbielamist ja põhjustamist (nt *Sehnsucht* 'igatsus' – *igatsema*, *Frustration* 'frustratsioon' – *frustreerima*). Just viimased on emotsiooniverbid (Kailuweit 2005).

⁴ Üldise arusaama järgi moodustavad meeleolud emotsioonidele justkui üldise tausta (Mandl, Euler 1983: 6), nii et mingi meeleolu olemasolu soodustab teatud emotsiooni tekkimist, nt agressiivses meeleolus inimene läheb kergemini raevu kui mingis muus meeleolus (vt nt Leising 2003: 31).

kui verbid, mis kirjeldavad või tähistavad mingeid emotsionaalseid läbielamisi või seisundeid, nende esilekutsumist või tundmahakkamist. Teatud mõttes sisaldab see määratlus liigitust, mida kasutab M. H. Burstein (1997) verbide puhul, mis kirjeldavad emotsionaalseid sündmusi (sks *emotionale Ereignisse*) – ta käsitleb verbe, mis kirjeldavad seisundeid (nt *kurb olema*), üleminekuid ühest emotsionaalsest seisundist teise (nt *kurvaks muutuma, kurbuma*) ning tinglikult öeldes “tegevusi” (sks *Handlungen*, ingl *acts*), mis kutsuvad esile emotsiooni (nt *kedagi kurvaks tegema, kurvastama*). Sarnaselt on tundeverbid määratlenud I. Toluk: “verbid, mis oma tähenduselt väljendasid emotsionaalseid psüühilisi protsesse või tundmusi, samuti emotsionaalse seisundi muutust.” (Toluk 2002: 14) Emotsiooniverbi asemel kasutab I. Toluk küll terminit *tundeverb*, ehkki ta ei käsitle verbe, mis väljendavad füüsilist tunnet. Et *tunne* võib nii eesti kui ka saksa keeles lisaks emotsionaalsetele läbielamistele ja tundmustele tähistada ka füüsilist aistingut (nt *nälja-, valu-, väsimustunne*) (Silk 2006, vt ka Fries 2003), siis on käesolevas artiklis eelistatud terminit *emotsiooniverb*, mis tundub ühemõttelisem.

Materjal ja katse

Eespool nimetatud emotsiooniverbide määratlust silmas pidades kogusin kättesaadud allikatest, nagu emotsiooniverbiloenditest, emotsioonisõnade nimekirjadest, psühholoogilistes testides, ÕS 1999-st, EKSS-ist (1988–2005), saksakeelsetest seletussõnaraamatutest (Duden 1996, DWDS), sünonüümi- (Õim 1991, Duden 1997, Görner, Kempcke 1978) ja pöördõnastikest (Mater 1970) 367 saksa- ja 340 eestikeelset emotsiooniverbi. Loendist jäid välja murdesõnad ning sõnad, mille juures on sõnaraamatus märgend *vananenud* või *vananev*. Kogutud sõnade hulgas oli 58 saksa- ja 102 eestikeelset verbi mille liigitumine emotsiooniverbiks oli küsitav. Sellised verbid tähistasid kas üldisi kognitiivseid kogemusi või nende põhjustamist (nt sks *beleben* ‘elavdama’, *demütigen* ‘alandama’, *sich gedulden* ‘kannatama, kannatlik olema’, *locken* ‘meelitama’, *schönreden* ‘meelitama’; ee *uskuma, veenduma, raatsima*), tuju (nt sks *gnatzen* ‘tusatsema’, *maulen* ‘mossitama’; ee, *mossitama, tigetsema, süngestama*) või olid seotud ootuste ja soovidega (nt sks *begehren* ‘ihaldama; soovima, tahtma’, *erhoffen* ‘ootama, lootma’, *geliüsten* ‘himustama, ihaldama’, *harren* ‘(pikisilmi) ootama’, *herbeiwünschen* ‘(kellegi v millegi tulekut) soovima’, *hoffen* ‘lootma’, *schwärmen* ‘unistama’; ee *soovima, tahtma, ootama*).

Kuna varasemad uurimused näitasid, et ühtseid kriteeriume emotsiooniverbide määramiseks ei ole ning et tihtipeale ongi nende määramine üsna subjektiivne, tundus mõistlikuna korraldada katse ning lasta emotsiooniverbid määrata kindlaks katseisikutel. Nn ebakindla staatusega verbide nimekiri jagati 76-le saksa keelt (2006. a) ja 47-le eesti keelt (2005. a) emakeelena kõnelejale. Saksakeelsete verbide hindajateks olid peamiselt Kieli Christian Albrechtsi nimelise ülikooli germanistika eriala üliõpilased vanuses 21–32 aastat, lisaks 13 erineval erialal (nt ametnik, õpetaja, majandus, müügitöötaja, kunstnik) töötavat kõrgharidusega inimest Nürnbergis vanuses 31–52 aastat. Eestikeelsed verbid anti hinnata 37-le kõrgharidusega (nt filoloogia, raamatupidamine, õigusteadus, logistika, keemia, bioloogia, geoloogia, stomatoloogia) ja 10-le kesk- või keskeriharidusega (nt raamatupidamine, mehhaanika, kaubandus) eestlasele vanuses 25–71 aastat.

Katseisikutele anti ülesandeks verbide hulgast välja valida ja alla joonida need, mis nende arvates on emotsiooniverbid, kusjuures rõhutati, et ei ole õigeid ega valesid vastuseid, vaid tahetakse teada üksnes subjektiivset arvamust. Ajapiirangut küsitluslehe täitmisel ei olnud. Eelnevalt tehtud prooviküsitlus näitas, et vaatamata sellele, et oli rõhutatud kõigi vastuste aktsepteeritavust, teatasid mitmed katseisikud, et nemad ei oska vastata, kui ei tea, mis on emotsiooniverb. Et esialgsel verbide valikul kirjalikest allikatest oli lähtunud emotsiooniverbi määratlusest, sai küsitlusele lisatud ka emotsiooniverbi definitsioon. Juurde märgiti, et see võib olla otsustamisel abiks. Oleks ülesandeks olnud lihtsalt katseisiku arvates sobivate sõnade väljavalimine, oleks katse tulemustest saanud ehk leida ka viiteid sellele, kuidas peegelduvad emotsioonid kahe keele maailmapildis. Katse eesmärgiks oli siiski saada enam-vähem sarnastel alustel koostatud nimekirjad, lisaks lasti katseisikutel valida mitte kõigi võimalike sõnade, vaid nende verbide vahel, mille kuulumine emotsiooniverbide hulka ei olnud autorile selge; esialgne valik oli juba tehtud. Ka emakeelsete katseisikute valik ei olnud tingitud mitte soovist saada teavet vastava keele maailmapildi kohta, vaid eeldusest, et emakeeles leiduvad sõnad on katseisikule oma tähenduse ja kasutuse poolest tuttavad.

See, kas ja kui palju kasutasid katseisikud otsustamisel definitsiooni abi, ei ole siiski teada, sest ülesande juhises ei palutud sellest täpselt lähtuda. Võib siiski oletada, et tulemustes peegeldub eelkõige see, mismoodi tõlgendasid katseisikud emotsiooniverbe definitsioonist lähtuvalt (juhul kui nende enda arusaam emotsiooniverbist ei ühtinud küsitluslehel tooduga).

Katse tulemuste tõlgendamisel võtsin eeskujuks L. Schmidt-Atzerti tehtud uuringu (1980), mille käigus lasti katseisikutel hinnata etteantud nimisõnade kuuluvust emotsioonisõnade hulka viiepalliskaalal (4 – kindlasti emotsioon, 0 – kindlasti ei ole emotsioon). Vältimaks seda, et suure osa etteantud sõnadest paigutavad katseisikud skaala keskele (2 – ei tea), nagu juhtus L. Schmidt-Atzerti katses, lasin verbe hinnata vaid sellest seisukohast, kas tegemist on emotsiooniverbiga või ei.⁵ Emotsioonisõnade hulka kuuluvaks luges L. Schmidt-Atzert sõnad, millele katseisikud olid omistanud vähemalt väärtuse 3 (st pigem emotsioon). Sellest lähtuvalt otsustasin, et verbe, mille joonisid emotsiooniverbina alla vähemalt pooled katseisikutest, peetakse suurema tõenäosusega emotsiooniverbiks. Oluline on siinkohal rõhutada tõenäosust, sest väita kindlalt, et väljavalitud sõnad ongi emotsioonisõnad, ei saa. Tegemist on vaid selle katse tulemustega ning mõni teine katse oleks andnud ehk hoopis teise tulemuse. Väljavalitud sõnade kuulumise üle päris lõplikku nimekirja edasiseks analüüsimiseks otsustasin siiski pärast võrdlemist teiste loeteludega.

Tulemused

102-st eestikeelsest verbist pakkusid katseisikud emotsiooniverbiks umbes pooled – 46. Seega jäi eestikeelsesse verbiloendisse kokku 284 verbi (kuna esialgses nimekirjas oli 340 verbi ja 56 verbi puhul oli ülekaalus arvamus, et pigem ei ole tegemist emotsiooniverbiga). Nn valituks osutunud verbide seas oli võimalik leida teatud seaduspärasusi. Näiteks peeti üsna üksmeelselt emotsiooniverbiks halvustamise (*alavääristama, halvustama, häbistama, laitma, parastama, pahaks panema,*

⁵ Siinkohal tuleks mainida ka seda, et L. Schmidt-Atzerti (1980) korraldatud ja käesolevas artiklis kirjeldatud uurimise eesmärgid olid samuti pisut erinevad, millest oli tingitud ka katseisikutele esitatud ülesande erinev püstitus. L. Schmidt-Atzert soovis teada, millise tõenäosusega peetakse mingit sõna emotsioonisõnaks, sinne artikli eesmärgiks on leida mingigi piir emotsioonisõnade ja mitte-emotsioonisõnade vahel. Lisaks käsitles L. Schmidt-Atzert emotsioonisõnadena vaid nimisõnu, emotsioonisõnavara tüüpilisemaid esindajaid (Vainik 2002b: 542), mitte verbe, nagu siinses uurimuses.

parastama, vaenutsema, vägivallatsema), tujutsemise (*jonnima, kapriisitsema, tujutsema, kiuslema*) ja lõbutsemisega seotud verbe (*lustima, lõbutsema, vallatlema*), ei peetud aga soovide (*ootama, tahtma*), uskumise (*uskuma, veendum*) ja kannustamisega seotud (*õhutama, ärgitama, kannustama*) verbe ning üksikuid verbe *huvituma, trotsima, morjendama*.

Samas võis valikute seas täheldada ka teatud vastuolusid, mis annab ehk taas tunnistust sellest, et nimetatud verbid ei klassifitseeru üheselt ühtegi kindlasse leksikaal-semantiliselt välja. Nimelt moodustus terve hulk transitiivseid-intransitiivseid verbipaare (semantiliselt kas kausatiividest ja inhoatiividest või seisundiverbidest)⁶, mille üht liiget pidasid katseisikud emotsiooniverbiks (alljärgnevates sõnapaarides nimetatud esimesena), teist aga mitte (*ergutama-ergastama, erkuma; lõögastuma-lõögastama; piinama, piinlema-piinutama; rahulduma-rahuldama; süngestuma-süngestama*), kusjuures emotsiooniverbiks peetud verb tundub olevat pigem intransitiivne verbivorm.

58-st saksakeelsest verbist valisid katseisikud emotsiooniverbiks 45, nii et lõpuks jäi saksakeelses verbide nimekirja 354 verbi. Ka sakslased pidasid sarnaselt eestlastega emotsiooniverbiks halvustamisega seotud tegusõnu (*demütigen* 'alandama', *entwürdig* 'alandama, häbistama', *herabwürdig* 'alandama, alavääristama') ega pidanud uskumisega (*schwören* 'kindel olema, kindlalt uskuma', *sichergehen* 'veendum', *überzeugen* 'veenma') ning (I)ootusega seotud verbe (*sich getrösten* 'lootma jääma', *gewärtigen* 'ootama, lootma'). Samuti moodustusid paarid või grupid sarnase tähendusega (aga erinevalt eesti keelest erineva tüvega) verbidest, millest üks tundus kuuluvat emotsiooniverbide hulka (alljärgnevates sõnapaarides nimetatud esimesena), teine mitte (*schmeicheln* 'meelitama' – *schönreden, schöntun* 'meelitama'; *demütigen, herabsetzen, entwürdig* 'alandama' – *herabsetzen* 'alavääristama'; *sich abgrämen* 'murest otsa jääma' – *sich abhärmen* 'murest otsa jääma'; *sich verfinstern, sich verdüstern* 'süngestuma' – *sich umdüstern* 'süngestuma'; *aufstacheln, aufmuntern* 'reibastama; ergutama' – *stacheln* 'õhutama, ergutama', *anregen* 'ergutama', *beleben* 'elavdama, ergutama'; *verführen* 'kiusatusse viima, ahvatlema' – *locken* 'peibutama, ahvatlema', *verlocken* 'ahvatlema, kiusatusse viima'; *sich entspannen* 'lõögastuma' – *sich abspannen* 'lõdvestuma').

Lisaks juba nimetatud sarnasustele eestlaste ja sakslaste sõnavalikus (alandamise, (I)ootuse ja uskumisega seotud sõnad) esines veelgi verbe, mille puhul kummagi keele rääkijad tegid ühesuguse otsuse. Emotsiooniverbiks peeti mõlemas keeles verbe (sulgudes saksakeelne vaste, mis esines saksakeelses nimekirjas) *unistama* (*träumen*), *lõögastuma* (*sich entspannen*), *andestama* (*vergeben*), *süngestuma* (*sich verdüstern, sich verfinstern*); emotsiooniverbiks ei peetud sõandama (*sich getrauen*), roiduma (*ermatten*), peibutama (*locken, verlocken*). Ilmnes aga ka erinevusi: kui sakslased pidasid emotsiooniverbiks transitiivseid verbe *aufmuntern* 'reibastama', *ermütigen* 'julgestama', *aufstacheln* 'virgutama, ergutama' (kausatiivid) ning intransitiivset verbi *vertrauen* 'usaldama' (seisundiverb), siis eestlaste meelest nende sõnade vasted või neile tähenduselt sarnased sõnad *ergastama, reibastama; julgema; ärgitama; uskuma* emotsioonisõnade hulka ei kuulunud. *Rahulduma* oli eestlaste arvates emotsiooniverb, sakslaste arvates (*vorliebnehmen*) mitte, nagu ka eesti *rahuldama*.

⁶ Kausatiivid on alati transitiivsed (*Tema sõnad jahmatasid mind*), inhoatiivid intransitiivsed (*Kati kokkus seda kuuldes*), seisundiverbid võivad olla nii transitiivsed (*Ma armastan sind*) kui ka intransitiivsed (*Teet närvitses kohutavalt*). Vt täpsemalt I. Toluk (2002).

Arutelu ja kokkuvõte

Lahknevused samatüveliste või -täenduslike sõnade liigitamises ühte või teise gruppi ei iseloomusta vaid seda katset, sama näitas ka lühike ülevaade erinevate autorite koostatud emotsioonisõnade nimekirjadest. Kõigi kirjeldatud katsete tulemusi kõrvutades võib öelda, et katses osalenud isikud on sõnade liigitamisel lähtunud ilmselt emotsioonisõnade laiemast määratlusest, mille kohaselt liigituvad emotsioonisõnade alla ka mitmed kindlat tegevust tähistavad sõnad. Nende all võib esile tuua mitmeid verbe, mis oma olemuselt ei ole otseselt seotud emotsioonidega, vaid nimetavad pigem kõneakte, millega taotletakse kas positiivset või negatiivset muutust teise inimese emotsionaalses seisundis (nt ee *alavääristama*, *alandama*, *hääbistama*; sks *demütigen*, *herabwürdigigen*, *entwürdigigen* 'alandama'), või milles väljendub kõneleja / emotsiooni kogeja hinnang situatsioonile või enda emotsionaalsele seisundile (nt *pahaks panema*, *halvustama*, *laitma*). *Vaenutsema* on näide suhtumist kirjeldavatest verbidest ning *vägivallatsema* tähistab negatiivseid emotsioone väljendavat tegevust. Ka E. Vainik on teinud sarnase tähelepaneku, et hinnangulise suhtluse puhul on emotsionaalse mõjutamise komponent olemuslik ning hinnanguliste kõneaktide seostumine emotsiooniverbidega on ootuspärane (Vainik 2007).

Selles mõttes sarnanevad tehtud katse tulemused pigem J. Alliku ja A. Ortony jt kui lingvistide (R. Kailuweit, M. Siironen, I. Toluk) nimekirjas leiduvaga. Emotsiooniverbideks ei peetud aga kummaski keeles (l)ootusega seotud verbe, mille on emotsiooniverbide hulka lugenud nii R. Kailuweit (2005) kui ka N. Fries (2000). Uskumuste ja veendumustega seotud sõnade pidamist mitte-emotsioonisõnadeks võib seletada ilmselt sellega, et uskumine ja veendumine (nagu ka ootamine ja lootmine) seostuvad pigem mõistuse kui tunnetega, neis domineerib afektiivse komponendi asemel kognitiivne. Kognitiivsus tundubki olevat see aspekt, mis teeb otsustamise keeruliseks. Tekivad küsimused, kas kognitiivne komponent sõna tähenduses on tugevam kui afektiivne ja kas sõna võib sellegipoolest lugeda emotsioonisõnaks. Kognitiivse komponendiga verbidest peeti emotsiooniverbiks nt ee *unistama*, *andestama*, *rahulduma*; sks *träumen* 'unistama', *vergeben* 'andestama', *vertrauen* 'usaldama'.

Lõbu ja lõbutsemist tähistavate verbide (*lustima*, *lõbutsema*, *vallatlema*) pidamine emotsiooniverbiks on iseenesest loogiline – seostub ju omadussõna *lõbus* (sks *lustig*) mitte iseloomu, vaid tuju või meeleoluga, tujud ja meeleolud on aga emotsioonidega tihedalt seotud (Schmidt-Atzert 1981: 30, Euler, Mandl 1983: 5–6, Fuchs-Lévy, Greulich-Janssen 1991; Mitmannsgruber 2003: 30–32). Ehkki lõbu kuulub mitmes emotsioonisõnade nimekirjas rõõmu alla (nt Schmidt-Atzert, Ströhm 1983), esineb ta ka omaette emotsioonikateooriana (nt Schmidt-Atzert 1987). Mainitud sõnad väljendavad lisaks pillerkaaritamisele ja vallatuste tegemisele tõepoolest ka vastava tunde või omaduse omamist ("mul on lõbus") või enda samastamist sellega ("ta on vallatu").

Vastupidiselt sakslastele ei ole eestlastest katseisikud pidanud emotsiooniverbideks *õhutama*, *ärgitama*, *kannustama* ja *reibastama*, vastavasisuline sõnarühm on puudu ka J. Alliku ning I. Toluki sõnaloendist. Sakslased seevastu on transitiivseid verbe (inhoatiive) *anspornen* 'kannustama', *aufmuntern* 'reibastama; ergutama, virgutama', *aufstacheln* 'virgutama, ergutama; tagant kihutama' näinud

pigem emotsioonisõnana, samuti esineb *innustama*, *sütitama* või *ärgitama* nii R. Kailuweitil, M. Siironeni kui ka A. Ortonyl sõnaloendis (pr *inspirer*, *s'inspirer*; sm *innostaa*, *innostua*; ingl *eager*, *enthusiastic*). Nimetatud sõnadki kannavad endas inimestevahelise suhtlemise seda poolt, mis hõlmab suhtluspartneri mõjutamist ning suunamist. Nähtavasti ei ole nende sõnade semantikas olulisel kohal mitte niivõrd emotsionaalse seisundi muutmine (*ärgitamise* tulemusel ei muutu ärgitav ärksaks, nagu ei saa ka öelda, mill(is)eks muutub kannustatav või õhutatav), kuivõrd partneri mõjutamine nii, et see hakkaks midagi tegema.

Erinevalt paljudest emotsioonisõnade nimekirjadest, k.a ülalmainitud, ei peetud tehtud katses emotsioonisõnaks ka verbe *huvitama-huvituma*. *Huvi* tundub olevat aga üks neid emotsioonikategooriad, mille kuulumises emotsioonisõnade semantiliselt välja ei ole ühel meel ka kõik emotsiooniuurijad – nii on huviga seotud sõnu käsitletud oma uurimustes ameeriklased (Averill, Izard, Kotsch), mitte aga rootslased (Ekman, Lundberg) ja sakslased (Schmidt-Atzert, Ströhm) ning need sõnad olevat emotsioonide diferentsiaalskaalal saanud väga madala hinnangu (Schmidt-Atzert, Ströhm 1983: 131).

Silma paistab seegi, et kumbagi keelt kõnelevad katseisikud on otsustanud pigem mitme konkreetset tegevust (nt füüsiline kontakt, hinnanguline verbaalne kommunikatsioon, liikumine ruumis) tähistava verbi (nt ee *laitma*, *vägivallatsema*, *parastama*, *vallatlema*; sks *herabwürdigen* 'alandama', *verführen* 'kiusatusse viima') kui ülekantud tähenduses emotsioonisõnana kasutusel oleva verbi (nt *liigutama*; *berücken* 'võluma, lummama', *entzaubern* 'lummusest vabastama') "kasuks". See võib viidata sellele, et ehkki emotsioonisõnad on sarnaselt emotsioonidele üsna abstraktsed, püütakse neidki endale pisut konkreetsemana ette kujutada. E. Vainiku tehtud vaba loetlemise ülesandes on nimetatud emotsioonidega seonduvate tegusõnade hulgas samuti rohkelt selliseid sõnu, mis väljendavad emotsioone kehakeele, näoilme või mingi tegevuse kaudu (nutmine, naermine, ohkamine), seega konkreetset ja nähtavalt (Vainik 2002a, 2007).

Tehtud katse tõestas veel kord, et emotsioonisõnade kui kategooria ja semantiliselt välja piirid on hägusad ning iga keelekasutaja jaoks individuaalsed. Ei ole siis imestada, et nii nagu tänaseni ei ole suudetud üheselt defineerida emotsioone, nii ei ole ka suudetud kokku leppida ühtses emotsioonisõnade definitsioonis. Eesti ja saksa keelt emakeelena kõnelevate katseisikute küsitluse tulemused näitavad, et emotsiooniverbide mõistet tajutakse suhteliselt ebamääraselt, see ei hõlma vaid konkreetsete emotsioonidega seotud või emotsionaalseid seisundeid kirjeldavaid verbe, vaid ka näiteks sõnu, mis tähistavaid tegevusi, mille eesmärgiks on suhtluspartneri emotsionaalset seisundit verbaalselt mõjutada või väljendada hinnangut situatsioonile või partnerile. Keeltevahelistest erinevustest olulisemaks osutusid siiski keelesisesed ning individuaalsed lahknevused – tuli ette ka seda, et tähenduselt sarnaste verbide puhul (nt *ergutama-innustama-ärgitama-õhutama*) valiti emotsiooniverbide hulka vaid üks või mõned, aga mitte kõik.

Nähtavasti võibki erinevate katseisikutega ja küsitlute hulgaga saada isesuguse tulemuse, kuivõrd emotsioonisõnade näol on tegemist kategooriaga, mida võib tajuda, tõlgendada ja liigitada äärmiselt subjektiivselt.

Kirjandus

- Allik, Jüri 1996. Eesti keele emotsioone väljendava sõnavara tähendus. – Keel ja Kirjandus 1, 6–12; 2, 92–99.
- Argus, Reili 2004. Verbi eelkäijatest esimeste kolmeliikmeliste miniparadigmadeni eesti ja soome keele võrdlusel. – Helle Metslang (koost.), Maria-Maren Sepper, Jane Lepasaar (toim.). Toimiv keel II. Töid rakenduslingvistika alalt. Tallinna Pedagoogikauilikooli eesti filoloogia osakonna toimetised 3. Tallinna Pedagoogikauilikool, Eesti Keele Instituut. Tallinn: TPÜ kirjastus, 37–52.
- Burstein, Marcos Herrera 1997. Sprachliches Erfassen von Emotionen im Spanischen: ein Beitrag zur kognitiven Semantik. Inaugural-Dissertation zur Erlangung der Doktorwürde der Philosophischen Fakultät der Albert-Ludwigs-Universität zu Freiburg i. Br.
- Debus, Günter 1988. Psychologie der Gefühlswörter. Empirisch-experimentelle Untersuchungsansätze und -ergebnisse. – Ludwig Jäger (Hrsg.). Zur historischen Semantik des deutschen Gefühlswortschatzes: Aspekte, Probleme und Beispiele seiner lexikographischen Erfassung. Aachen, Alamo.
- Duden = Duden Deutsches Universalwörterbuch 1996. 3., neu bearb. und erw. Aufl. Mannheim, Leipzig, Wien, Zürich: Dudenverlag.
- Duden sinn- und sachverwandte Wörter: Synonymwörterbuch der deutschen Sprache 1997. Herausgegeben und bearbeitet von Wolfgang Müller. Der Duden in 12 Bänden; Bd.8. Mannheim: Dudenverlag.
- EKSS = Eesti kirjakeele seletussõnaraamat I–VII. 1988–2007. Eesti Keele Instituut. Tallinn: Eesti Keele Sihtasutus.
- Fehr, B; Russell, James A. 1984. Concept of emotion viewed from a prototype perspective. – Journal of Experimental Psychology: General, 113, 464–484.
- Fries, Norbert 2000. Sprache, Gefühle, Emotionen und Emotionale Szenen. Vortrag zum 65. Geburtstag von Inger Rosengren. Lund.
- Fries, Norbert 2003. Gefühlswortschatz im GWDS. – Herbert Ernst Wiegand (Hrsg.). Untersuchungen zur kommerziellen Lexikographie der deutschen Gegenwartssprache I. Lexicographica, Series Maior Supplementbände zum Internationalen Jahrbuch für Lexikographie 113. Tübingen: Niemeyer.
- Frijda, Nico; Markam, Suprapti; Sato, Kaori; Wiers, Reinout 1995. Emotions and emotion words. – James A. Russell, José-Miguel Fernánandez-Dols, Antony S. R. Manstead, J. C. Wellenkamp (eds). Everyday Conceptions of Emotions. An Introduction to the Psychology, Antropology and Linguistics of Emotion. NATO ASI Series; Series D: Behavioral and Social Sciences 81. Dordrecht, Boston, London: Kluwer Academic Publishers, 121–143.
- Fuchs-Lévy, Irmgard; Greulich-Janssen, Gisela 1991. Über das Wesen und die Dynamik des Gefühls. Eine Studie tiefenpsychologisch und philosophisch orientierter Gefühlstheorien in ihrer Bedeutung für die Psychotherapie. Dissertationen im Arbeitskreis für Tiefenpsychologie, Gruppendynamik und Gruppentherapie Berlin, herausgegeben von Prof. Dr. Josef Rattner. Berlin: Verlag für Tiefenpsychologie.
- Goozen, Stephanie van; Frijda, Nico H. 1993. Emotion words used in six European countries. – European Journal of Social Psychology, Vol. 23, 89–95.
- Görner, Herbert; Kempcke, Günter (Hrsg.) 1978. Synonymwörterbuch: sinnverwandte Ausdrücke der deutschen Sprache. Leipzig: Bibliographisches Institut.
- Jäger, Ludwig; Plum, Sabine 1988. Historisches Wörterbuch des deutschen Gefühlswortschatzes. Theoretische und methodologische Probleme. – Ludwig Jäger (Hrsg.). Zur historischen Semantik des deutschen Gefühlswortschatzes: Aspekte, Probleme und Beispiele seiner lexikographischen Erfassung. Aachen: Alamo.
- Kailuweit, Rolf 2005. Linking: Syntax und Semantik französischer und italienischer Gefühlswörter. Linguistische Arbeiten 493. Tübingen: Max Niemeyer Verlag.

- Kövecses, Zoltán 1995. Language and emotion concepts. – James A. Russell, José-Miguel Feranández-Dols, Antony S. R. Manstead, J. C. Wellenkamp (eds). *Everyday Conceptions of Emotions. An Introduction to the Psychology, Anthropology and Linguistics of Emotion*. NATO ASI Series; Series D: Behavioral and Social Sciences 81. Dordrecht, Boston, London: Kluwer Academic Publishers,
- Leising, Daniel 2003. Veränderungen interpersonal-affektiver Schemata im Verlauf psycholanalytischer Langzeitbehandlungen. Doktorarbeit. Universität Heidelberg.
- Mandl, Heinz; Euler, Harald A. Euler 1983. Begriffsbestimmungen. – Harald A. Euler, Heinz Mandl (Hrsg.). *Emotionspsychologie: ein Handbuch in Schlüsselbegriffen*. U-&-S-Psychologie. München, Wien, Baltimore: Urban & Schwarzberg.
- Marx, Wolfgang 1982. Das Wortfeld der Gefühlsbegriffe. – *Zeitschrift für Experimentelle Psychologie*, Bd XXIX, H. 1, 137–146.
- Marx, Wolfgang 1997. Semantische Dimensionen des Wortfeldes der Gefühlsbegriffe. – *Zeitschrift für Experimentelle Psychologie*, Bd XLIV, H. 3, 478–494.
- Mater, Erich 1970. *Rückläufiges Wörterbuch der deutschen Gegenwartssprache*. Leipzig: VEB Verlag Enzyklopädie.
- Merten, Jörg 2003. *Einführung in die Emotionspsychologie*. Verlag W. Kohlhammer.
- Mitmannsgruber, Horst 2003. *Kognition und Emotion. Die Regulation von Gefühlen im Alltag und bei psychischen Störungen*. Bern, Göttingen, Toronto, Seattle: Verlag Hans Huber.
- Ortony, Andrew; Clore, Gerald L.; Foss, Mark A. 1987. The Referential Structure of the Affective Lexicon. – *Cognitive Science* 11, 341–364.
- Russell, James A. 1991. Culture and the Categorization of Emotions. – *Psychological Bulletin* 110, 3, 426–450.
- Schmidt-Atzert, Lothar 1980. Die verbale Kommunikation von Emotionen: Eine Bedingungsanalyse unter besonderer Berücksichtigung physiologischer Prozesse. Dissertation. Justus-Liebig-Universität Gießen.
- Schmidt-Atzert, Lothar 1981. *Emotionspsychologie*. Kohlhammer-Studienbücher: Psychologie. Stuttgart, Berlin, Köln, Mainz: Kohlhammer.
- Schmidt-Atzert, Lothar 1987. Zur umgangssprachlichen Ähnlichkeit von Emotionswörtern. – *Psychologische Beiträge* 29, 140–163.
- Schmidt-Atzert, Lothar; Ströhm, Walter 1983. Ein Beitrag zur Taxonomie der Emotionswörter. – *Psychologische Beiträge* 25, 126–141.
- Siiroinen, Mari 2001. *Kuka pelkää, ketä pelottaa: nykysuomen tunneverbien kielioppien ja semantiikkaa*. Suomalaisen kirjallisuuden seuran toimituksia 844. Helsinki: Suomalaisen kirjallisuuden seura.
- Silk, Anni 2006. *Gefühl und Emotion in der Sprache*. – Anne Arold, Dieter Cherubim, Dagmar Neuendorff, Henrik Nikula (Hrsg.). *Deutsch am Rande Europas*. Humaniora: Germanistica I. Tartu: Tartu University Press.
- Tischer, Bernd 1993. *Die vokale Kommunikation von Gefühlen*. Fortschritte der psychologischen Forschung 18. Weinheim: Psychologie-Verlags-Union.
- Toluk, Ivi 2002. *Eesti keele tundevid. Bakalaureusetöö*. Tartu Ülikool, eesti keele õppetool, käsikiri. Tartu: Tartu Ülikool
- Ulrich, Dieter 1989. *Das Gefühl. Eine Einführung in die Emotionspsychologie*. 2., durchgesehene und ergänzte Auflage. München: Psychologie Verlags Union.
- Vainik, Ene 2001. *Eestlaste emotsioonisõnavara*. Magistritöö. Tartu Ülikool, eesti keele õppetool, käsikiri. Tartu: Tartu Ülikool
- Vainik, Ene 2002a. Kas eestlased on “kuumaverelised”? Eestlaste rahvalikust emotsioonikategorias. – Mati Ereit (peatoim.), Evi Ross, Asta Õim (toim.). *Emakeele Seltsi aastaraamat 47 (2001)*. Tallinn: Emakeele Selts, 63–86.
- Vainik, Ene 2002b. Millest on tehtud eestlaste emotsioonisõnavara? – *Keel ja Kirjandus* 8, 537–553.

- Vainik, Ene 2004. *Lexical Knowledge of Emotions: the Structure, Variability and Semantics of the Estonian Emotion Vocabulary*. *Dissertationes linguisticae universitatis Tartuensis* 5. Tartu: Tartu University Press.
- Vainik, Ene 2007. Emotsioonisõnad töises keskkonnas – millest kõnelevad loetelukatsete tulemused? – Helle Metslang, Margit Langemets, Maria-Maren Sepper (toim.). *Eesti Rakenduslingvistika Ühingu aastaraamat 3. Estonian Papers in Applied Linguistics* 3. Tallinn: Eesti Keele Sihtasutus, 339–353.
- Õim, Asta 1991. *Sünoniüümisõnastik*. Tallinn: Autori kulu ja kirjadega.
- Õim, Haldur 1971. Isiku mõistega seotud sõnarühmade semantiline struktuur eesti keeles. – *Keele modelleerimise probleeme* 4. Tartu Ülikooli toimetised 278. Tartu, 5–260.

Võrgumaterjalid

- Caluianu, Daniela 2007. *Psychological Verbs and the Transitivity in Romanian, English and Japanese*. <http://www3.aa.tufs.ac.jp/~ritsuko/project/calu9810.html> (27.03.2007)
- DWDS = Das Digitale Wörterbuch der deutschen Sprache des 20. Jh. <http://www.dwds.de/> (10.09.2006)
- Fries, Norbert 2002. *Emotionen. Linguistische und kognitionswissenschaftliche Aspekte ihrer Explikation*. <http://www2.rz.hu-berlin.de/inside/linguistik/institut/syntax/download.htm#Materialien> (29.04.2005)
- Fries, Norbert 2007. *Die Kodierung von Emotionen und Texten*. – *Journal of Literary Theory* 1 (2). <http://www2.hu-berlin.de/linguistik/institut/syntax/docs/fries2007a.pdf> (15.06.2007)
- Klein, Katarina; Kutscher, Silvia 2005. *Lexical Economy and Case Selection of Psych-Verbs in German*. <http://www.linguistics.ruhr-uni-bochum.de/~klein/papers/LexEconPsych.pdf> (27.03.2007)
- Vanhoe, Henk 2002. *Aspects of the syntax of psychological verbs in Spanish. A lexical functional analysis*. – Miriam Butt, Tracy Holloway King (eds). *Proceedings of the LFG02 Conference*. National Technical University of Athens. CSLI Publications. <http://csli-publications.stanford.edu/LFG/7/lfg02vanhoe.html> (27.03.2007)

Anni Silk (Tartu Ülikool) on lõpetanud Tartu Ülikooli saksa filoloogia erialal. Kaitses samas ülikoolis 2002. a magistriraadi võrdlevast leksikaalsest polüseemiast. Alates 2003. a töötab kontrastiivse lingvistika teadurina. Uurimisvaldkonnad: polüseemia, leksikaalne semantika, kontrastiivne lingvistika, valents, emotsiooniverbid.
Anni.Silk@ut.ee

ÜBER EIN EXPERIMENT MIT EMOTIONSVERBEN

Anni Silk

Universität Tartu

Beim Zusammenstellen der estnischen und deutschen Emotionsverblisten entstand immer wieder die Frage, was die Emotionswörter (in diesem Fall Emotionsverben) eigentlich sind. Verschiedene Wortlisten in der Literatur scheinen nach völlig verschiedenen Kriterien zusammengestellt zu sein. Einige Autoren (z. B. Mari Siirainen) gehen anscheinend von einer recht begrenzten Emotionswortbestimmung aus, nach der Emotionsverben fast ausschließlich mit bestimmten Emotionen in Zusammenhang stehen, nach anderen (z. B. Jüri Allik, Andrew Ortony, Gerald L. Core und Mark A. Foss) umfassen diese auch einige weitere kognitive Wörter bzw. Wörter, die eher Charaktereigenschaften bezeichnen.

Um festzustellen, welche Verben estnische und deutsche Muttersprachler als Emotionsverben empfinden, wurden 76 deutsche und 47 estnische Versuchspersonen gebeten, über die Zugehörigkeit der vorgegebenen (und vorher als sog. Grenzfälle bezeichnenden) Verben zum Wortfeld der Emotionsverben zu entscheiden. Die Ergebnisse des Experiments zeigen, dass sowohl Deutsche als auch Esten im Allgemeinen eher konkrete Tätigkeitsverben (wie z. B. *laitma* 'tadeln', *schmeicheln*) als Verben mit übertragener Bedeutung (wie z. B. *berücken*, *entzaubern*) zu Emotionswörtern zählen und Emotionsverben als ein ziemlich weites Wortfeld betrachten, zu dem auch mehrere (bewertende) Sprechaktverben (z. B. *demütigen*, *entwürdigen*, *herabwürdigen*). Besonders schwankend schienen die Entscheidungen bei Verben mit der kognitiven Komponente zu sein – ein Teil von ihnen wurde für Emotionsverben gehalten, ein anderer dagegen nicht.

Es sollte noch bemerkt werden, dass die Unterschiede zwischen Individuen und innerhalb einer Sprache wesentlich größer waren als die zwischen Deutsch und Estnisch. Das beweist nochmals, dass Emotionswörter ein Gebiet mit sehr subjektiven Grenzen und Kriterien darstellt.

Schlagwörter: lexikale Semantik, Emotionswörter, Estnisch, Deutsch

TEADUSKEELE SEIRE TALLINNA ÜLIKOO LIS: TEADUSTÖÖTAJATE HOIAKUD

Anastassia Zabrodskaja

Ülevaade. 2006. aastal alustati Tallinna Ülikoolis seirega “Eesti teaduskeele seisund TLÜ-s”. Artikli alguses tutvustatakse inglise keele olukorda teaduskeelena ja kirjeldatakse eesti teaduskeelt kaitsvaid riiklikke ja TLÜ-siseseid arengudokumente. Järgmisena keskendutakse professorite, dotsentide ja (vanem)teadurite teaduskeelealastele probleemidele ja vajadustele, käsitledes publitseerimise ja õpetamisega seotud seiku. Seejärel analüüsitakse teadustöötajate hoiakuid eesti teaduskeele suhtes ja koolitusvajadust nende enese arvamuse järgi.*

Võtmesõnad: teaduskeel, inglise keel, eesti keel

1. Inglise keele pealetung teadusmaailmas

Nüüdismaailma peamine teaduskeel on inglise keel, mille globaalset levikut on põhjalikult uuritud paljudes publikatsioonides (Cheshire 1991, Crystal 1995, 1997, McArthur 1998, Graddol 2000, Ammon 2001). David Graddol (2000: 9) nimetab inglise keelt teaduse ja tehnoloogia rahvusvaheliseks valuutaks, sest see on uurijate ühine keel, milles avaldatakse teadustulemusi, peetakse ettekandeid ning kaitstakse väitekirju. Wayt W. Gibbs (1995) kirjeldab, kuidas mõned rahvuskeelsed ajakirjad muutusid ingliskeelseks: alustati ingliskeelse ülevaate nõudmisega, siis hakati artikleid tõlkima inglise keelde ja lõpuks hakati ajakirja kaastöödeks vastu võtma ainult ingliskeelseid kirjutisi. Tähelepanuväärset mõju on avaldanud inglise keel kunagi tugevale ja iseseisvale saksakeelsele teadusmaailmale. Wolfgang Viereck (1996) toob näite saksakeelse ajakirja Zeitschrift für Tierpsychologie kohta: 1950. aastal olid kõik kaastööd saksakeelsed, ent 1984. aastal oli 95% vastuvõetud artikleist juba ingliskeelsed. Kahe aasta pärast sai ajakirja ingliskeelse nime Ethology. D. Graddol (2000: 9) rõhutab, et 76% Saksamaa loodusteadlastest nimetab inglise keelt oma põhiliseks töökeeleks – füüsika, keemia ja bioloogia eriala teadlastest vastavalt 98%, 83% ja 81%. Harald Haarman ja Eugene Holman (2001: 231) märgivad, et

* Artikkel on valminud Euroopa Sotsiaalfondi toel.

1950-ndatel muutus Soome teaduslik ja akadeemiline elu ingliskeelseks selle tõttu, et 1953. aastal alustati rahvusvaheliste vahetusprogrammidega, mis andsid Soome õppejõududele võimaluse õpetada ja oma uuringuid läbi viia USA ülikoolides.

Mida rohkem selliseid ümberlülitumisi inglise teaduskeelele on olnud, seda loomulikum tundub uurijatele ingliskeelseid töid lugeda ja avaldada. Ulrich Ammoni (2003: 244) statistilistest andmetest saab suurepärase ülevaate anglo-saksa bibliograafiliste andmebaaside kohta. Inglise keel domineerib selgelt kõigis teadusharudes, nt loodusteadustes (bioloogia, keemia, füüsika, meditsiin, matemaatika) on 87,2% publikatsioonidest ingliskeelsed, sotsiaalteadustes (sotsioloogia, ajalugu, filosoofia) on see näitaja 82,5%.

Teadlaste töö on igale ühiskonnale äärmiselt oluline mitmel tasandil. Globaalsete arengute taustal muutub Eesti teadustegevus vaieldamatult aina ingliskeelsemaks. On õige aeg mõelda, kuidas eesti keel areneb teaduskeelena ning mida teha, et üliõpilane ja õppejõud kasutaksid teadustöid koostades laitmatut eesti keelt.

2. Eesti teaduskeele olulisus

See, millises eesti keeles õppejõud oma ainet käsitleb, on mitmeti oluline mõjur. Kui haritlane väljendub ladusalt ja täpselt, on üliõpilasel iga päev hea eeskuju ning ainegi saab hõlpsamini selgeks. Kui õppejõud suudab juhtida üliõpilase tähelepanu ka kirjalike tööde keelele, on kasutegur mitmekordne. Seesama kehtib veel enam muukeelsetest koolidest tulnud üliõpilaste kohta. Head eesti (eriala)keelt valdavad noored õpetajad viivad nõnda täpsema ja selgema keele oma aine kaudu (ka muukeelsetesse) üldhariduskoolidesse. Uurimistulemused teeb rahvale arusaadavaks selge ja täpne emakeel.

Arenenud eesti teaduskeel kui oluline osa oskuskeelest on eestikeelse teadmusühiskonna jätkusuutliku arengu peamisi eeldusi ning kajastub mitmes arengudokumendis. Eesti kõrgharidusstrateegia aastateks 2006–2015 (EKHS 2006) seab üheks eesmärgiks “tagada eestikeelse haridus- ja kultuuriruumi areng. Tuleb kindlustada eestikeelse kõrgharidusliku õppe jätkumine ja areng Euroopa avatud haridusruumis. Eesmärgiks on, et kõigis õppesuundades saab eestikeelset kõrgharidust kõigil õppeastmetel.” Samas riiklikuks dokumendis rõhutatakse ka erialase eesti keele loengute tähtsust: “Erialase eesti keele oskuse parendamiseks lülitatakse õppekavadesse erialase eesti keele õpe vähemalt 1 AP mahus. Arendatakse eestikeelset õpitarkvara, toetatakse eestikeelsete õppevahendite koostamist.”

Üks “Eesti keele arendamise strateegia 2004–2010” (EKAS 2004) eesmärke on “tagada eestikeelne kõrgharidus (katta kõik erialad oskussõnastike ning eestikeelse õppekirjandusega, säilitada olulises osas eestikeelne õpe ja avaldada tähtsamad teadustulemused ka eesti keeles), vältides mis tahes teadusala täielikku võõrkeelsetumist; tagada kõrgkoolilõpetajate eesti keele oskuse kõrge tase.”

Tallinna Ülikooli (TLÜ) põhikirja § 6 järgi on “ülikooli missioon toetada Eesti jätkusuutlikku arengut haritlaste ettevalmistamise, uurimistöö ja selle tulemuste rakendamise, ühiskondliku mõttevahetuse ja akadeemilise partnerluse edendamise kaudu.”

TLÜ arengukava 2006–2010 eesmärke on “hooldiseda Eesti intellektuaalse keskkonna sidususe eest ühiskonnaga, avada see keskkond senisest enam Euroopas

ja mujal maailmas toimuvatele dialoogidele ja debattidele ning toetada nendega Eesti ühiskonna sees käivat arutelu ja kriitilist eneserefleksiooni, aidates sellega kaasa Eesti vaimse iseseisvuse säilimisele.” TLÜ arengukava 2004–2007 järgi tuleb tagada õppekavadega piisav erialase eesti keele oskus, panustades seeläbi eesti keele kui teaduskeele arengusse. Erialase eesti keele oskuse parandamiseks võimaldatakse tudengitele valikainete raames eesti keele õpet vähemalt 1 ainepunkti mahus.

Eesti ülikoolides on üleilmastumine viinud selleni, et õppejõude valides hinnatakse koha taotlejat nimelt võõrkeelsete üllitiste loetelu järgi, arvestamata eestikeelseid – või paremal juhul pidades neid teisejärguliseks. Seda on rõhutanud ka akadeemik Anto Raukas (2002: 77).

Hulk inimesi on aastaid tegutsenud selle nimel, et eesti keel jääks täisväärtuslikuks, sh suudaks väljendada teadusmõtet. Kümnekond aastat tagasi on tehtud ettepanek kiiresti käivitada eesti haritlaste emakeelekoolitus, mis keele ja mõtlemise seotuse tõttu on igati kasulik mis tahes erialal (Nemvalts 1999). On selge, et kehvast eesti keeles mõtlemine ei vii täisväärtusliku teadustulemuseni.

Igal üliõpilasel tuleb analüüsida teaduskirjandust, sõnastada oma uurimisprobleem, töö eesmärgid, hüpotees, ülesanded. Üliõpilane peab olema võimeline valida kirjutatava teadustöö keelt ja stiili. Praktika käiguski peaks üliõpilane jälgima mõtete väljendamise selgust ja täpsust, keelekasutuse korrektsust.

Et saavutada kõigi üliõpilaste eesti erialakeele väga hea tase, on vaja kõigi eriala õppejõudude ja teadurite teadlikku suhtumist nii iseenda kui ka üliõpilaste eesti teaduskeele kasutusse. Seepärast on tarvis esmalt pakkuda eesti teaduskeele täiendõpet ülikooli akadeemilistele töötajatele. See on saanud võimalikuks osana Euroopa Sotsiaalfondi toetatavast teadus-arendustööst “Uus eesti keele ja kakskeelse õppe süsteem Tallinna Ülikoolis” (UUEKS)¹ 2006–2008, millel on kolm põhieesmärki.

1. Parandada muukeelsete üliõpilaste õpitulemusi eesti keele tugiprogrammi abil.
2. Tagada praktilise eesti erialakeele oskus kõikidele bakalaureuseõppe üliõpilastele ning toetada õppejõude ja teadureid teaduskeele kasutamisel.
3. Käivitada aineõpetajate täiendõppeprogramm, mis vastab venekeelse kooli eestikeelsele õppele ülemineku ning kultuuriderohkes koolis töötavate õpetajate vajadustele.

UUEKS-i laiem ülesanne on tagada eesti teaduskeele arendamine eri valdkondades.

3. Seire metodoloogia

Selgitamaks välja eesti teaduskeele olukorda, on TLÜ-s käivitatud teaduskeele seire sihiga välja selgitada:

- a) teadustöötajate ja üliõpilaste hoiakud eestikeelse teaduskeele suhtes;
- b) teadustöötaja teaduskeelealased probleemid ja vajadused (peamised raskused eestikeelsete loengute/õppematerjalide jms ettevalmistamisel);
- c) doktorantide teadustegevuse kogemused ning teaduskeelealased probleemid ja vajadused;

¹ www.tlu.ee/uueks (27.02.2008).

- d) eesti ja inglise keele kasutamisel teaduskeelena TLÜ-s;
- e) erialase eestikeelse kirjanduse olemasolu ja kättesaadavus;
- f) erialase ingliskeelse kirjanduse olemasolu ja kättesaadavus.

Seire raames küsitleti (ajavahemikus 01.09.2006–10.12.2006) kokku 400 inimest: 200 üliõpilast (bakalaureuse- ja magistriõpe), 65 doktoranti ning 135 teadustöötajat (professorid, dotsendid, (vanem)teadurid). Küsitluse jaoks töötati välja kolm ankeeti: doktorandi, teadustöötaja ning BA- ja MA-üliõpilase ankeet.²

Peamised uurimisküsimused olid:

- a) kui palju ja mis keeles on ilmunud TLÜ kirjastuse kaudu teadustöid?
- b) kui suure osa nendest moodustavad doktoritööd?
- c) kui suur on võõrkeelse doktoritöö puhul eestikeelse resümee osa?
- d) mis keeles avaldab oma uurimistöö tulemusi TLÜ teadustöötaja/doktorant üldse?
- e) kui suur on teises keeles kirjutatu eestikeelse resümee osa?
- f) kuidas varieerub avaldamiskeel teadusalati?
- g) kuidas on publitseerimiskeel viimasel aastakümnel muutunud?

Kolme küsimustiku koostamisel lähtuti järgmistest põhimõtetest.

- a) Sihtrühmapõhisus – küsitletavate sihtrühmade eripärast tulenevalt koostati eri sihtrühmade jaoks erinevad küsimustikud.
- b) Küsimuste võrdlev esitus – 40, 21 ja 49 küsimust (vastavalt üliõpilase, doktorandi ja teadustöötaja) ankeedi kohta, erinedes sihtrühmiti eri teemade käsitlemise detailsuse astmelt. Ankeedid olid üles ehitatud viisil, mis võimaldas võrrelda eri vastajarühmade antud vastuseid küsimuste lõikes. Samuti oli vastajatel võimalus lisada oma vastusevariante, kommentaare ja selgitusi ankeetide täitmisel. Ankeetimine oli anonüümne. Terviklikuma ülevaate saamiseks kombineeriti uuringus kvantitatiivset meetodit kvalitatiivsega.

Andmete võrdlemisel ja kirjeldamisel on kasutatud protsentjaotust, vastused töödeldi programmiga MS Excel. Andmed on esitatud joonistena. Pikemad sõnalised vastused üldistati analüüsi käigus kategooriatesse ja kodeeriti.

Klasteranalüüsi abil jaotati küsitluse andmed gruppidesse. Vastustevaheliste seoste uurimiseks ja analüüsimiseks kasutati statistilise andmetöötluse paketti SPSS 14.0. Arvuliste näitajate ehk muutujate seoste iseloomustamiseks kasutati korrelatsioonikoefitsienti r , mis näitab, kui tugev on seos muutujate vahel. Selle abil uuriti, kuidas on valimi sotsiodemograafilised taustaandmed seotud valitud vastustega. Korrelatsioonikoefitsient r võib jääda vahemikku -1 kuni 1 . Mida lähemal 1 -le, seda suurem positiivne korrelatsioon, st mida rohkem üht, seda rohkem ka teist asja, mida lähemal -1 -le, seda tugevam pöördkorrelatsioon, st mida rohkem üht, seda vähem teist.

- a) $0,1$ kuni $0,3$ – nõrk korrelatsioon;
- b) $0,3$ kuni $0,5$ – keskmine korrelatsioon;
- c) rohkem kui $0,5$ – tugev korrelatsioon.

Reliabiliteetanalüüsiga (Cronbachi alfa) aga selgitati välja muutujate vaheline sarnasus. *p* tähistab usaldusnivood. See peab olema 0,05 või vähem, et tulemust saaks pidada statistiliselt oluliseks, st see ei ole juhuslik erinevus, vaid tegelik. *p* väljendab veavõimalust – 0,05 on 5%-line eksimisvõimalus ja 0,01 on 1%-line eksimisvõimalus.

Artiklis käsitletakse teaduskondade kaupa ainult teadustöötajate suulise küsitluse teel saadud andmeid nende õpetamiskogemuse ja avaldamise kohta, esitatakse hoiakuid eesti teaduskeele kohta ning analüüsitakse arusaama teaduskeelekoolituse vajalikkuse suhtes. Uuringuaruanne tervikuna valmib 2008. aasta maikuu lõpuks. Peale teaduskeele seire tutvustatakse selles ka venekeelsete üliõpilaste ja erialase eesti keele teadustöötajate arvamusi, kogemusi ning hoiakuid erialase eesti keele loengute kohta.³

4. Teadustöötajate küsitlus

Teadustöötajatele mõeldud küsimused olid jagatud kolme rühma: a) vastajast üldiselt (sugu, vanus, teaduskond, eriala, keelteoskus), b) TLÜ töökogemustest üldiselt (publitseerimine, loengute pidamine, teaduskeele koolituse vajadus) ja c) eestikeelsete loengute probleemistiku väljaselgitamine (õppematerjalide keel, venekeelsete üliõpilaste eesti keele oskuse tase). Eraldi esitati informantidele 12 väidet eesti teaduskeele kohta, millega tuli kas nõustuda või mitte nõustuda. (Teadustöötajate ankeeti vt lisa 1–3.)

Õppeosakonna andmeil töötas TLÜ-s 2006. aastal umbes 450 teadustöötajat, nendest küsitleti 135, kelle vastused olid edaspidiseks töötlemiseks ja analüüsiks kõlblikud. Intervjuud viisid läbi TLÜ eesti filoloogia üliõpilased, keda oli eelnevalt instrueeritud intervjuu läbiviimise tehnika kohta. Nii moodustas valim 30% üldkogumist. Et teaduskondade suurus varieerub oluliselt, erineb ka seal töötavate professorite, dotsentide ja (vanem)teadurite arv (vt tabel 1). Võrreldes TLÜ teadustöötajate hulka ja vastanute jaotust teaduskonniti võib tõdeda, et koguti enam-vähem representatiivsed andmed.

Tabel 1. Vastanud teadustöötajate arv teaduskonniti

Teaduskond	Vastanute arv
Ajaloo Instituut	2
Eesti Humanitaarinstituut	7
filoloogiateaduskond	33
kasvatusteaduste teaduskond	17
kehakultuuri teaduskond	7
kunstide teaduskond	12
matemaatika-loodusteaduskond	34
sotsiaalteaduskond	23
Kokku	135

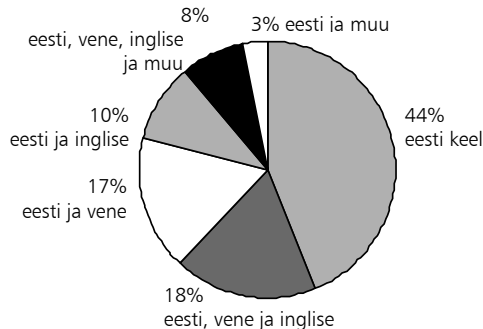
³ Aruanne avaldatakse UUEKS-i kodulehel.

5. Uuringu tulemused

5.1. Eesti keeles publitseerimisest ja õpetamisest

Üle poole (59%) vastustest laekus koos teaduskraadi, ameti, soo, vanuse, keeleoskuse ja muude sotsiodemograafiliste andmetega. Suhteliselt tagasihoidlikult oli kirjeldatud osakonda ja tööstaaži TLÜ-s (vaid 42%). Põhjuseks toodi ankeedi anonüümsuse printsiip, arvati et nende andmete põhjal saab küsitleja hiljem respondendid vajaduse korral tuvastada.

Igapäevaselt kasutab mitut keelt üle poole küsitletud TLÜ teadustöötajatest (vt joonis 1).

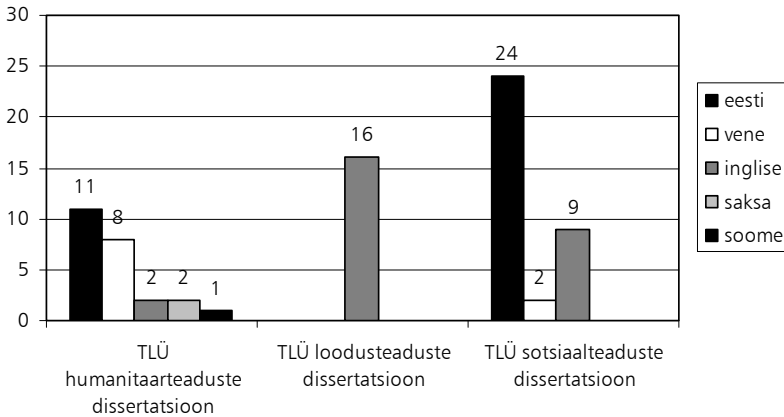


Joonis 1. TLÜ teadustöötajate igapäevane keelekasutus

Igapäevase keeleoskuse seosed teiste näitajatega on järgmised: teadlased, kes on lõpetanud ülikooli välismaal ($r = 0,24$, $p = 0,05$), kasutavad igapäevaselt rohkem võõrkeeli. Samuti on nõrk korrelatsioon ($r = 0,29$, $p = 0,05$) võõrkeelekasutuse ja teaduskraadi vahel: doktorikraadiga teadlased kasutavad võõrkeeli rohkem kui kandidaadi- või magistriskraadiga teadlased. Ilmnes veel, et mida rohkem võõrkeeli teadustöötaja iga päev kasutab, seda vähem on talle olulised terminoloogiasõnas- tikud ($r = 0,28$, $p = 0,05$).

TLÜ teadustöötaja seisab valiku ees, mis keeles oma teadmisi kirjalikult edasi anda. Üle kolmveerandi (78%) vastanutest tunnistas, et neil tuleb oma teaduslikke artikleid inglise keeles kirjutada. Pea kõigis matemaatika-loodusteaduskonna vastustes väideti, et sel alal ei olegi võimalik eesti keeles avaldada. Loodusteaduste eesti teaduskeele tulevik on küsitav. Doktoritööd näitavad teaduspublitseerimise tulevikusuundumusi. Seetõttu soovib siinkirjutaja esile tõsta, mis keeles ja kui palju kirjutatakse TLÜ-s doktoritöid (vt joonis 2).

Pooled vastanutest kinnitavad, et nad on avaldanud oma teaduspublikatsioone ka eesti, vene, saksa või soome keeles. Kui humanitaarvaldkonna teadustöötajad tunnistasid, et neil on võimalik publitseerida eelretsenseeritavates eestikeelse- tes ajakirjades ja nimetavad selliseid ajakirju, nagu Keel ja Kirjandus, Emakeele Seltsi aastaraamat, Eesti Rakenduslingvistika Ühingu aastaraamat, Mäetagused, Tuna, siis matemaatika-loodusteaduskonna esindajate sõnul ei ole nende alal teaduslikke eelretsenseeritavaid eestikeelseid ajakirju, sotsioloogia erialal on need võimalused üsnagi piiratud (Akadeemia on mingis osas eelretsenseeritav) ning ka kasvatusteadustes puuduvad vastavad ajakirjad.

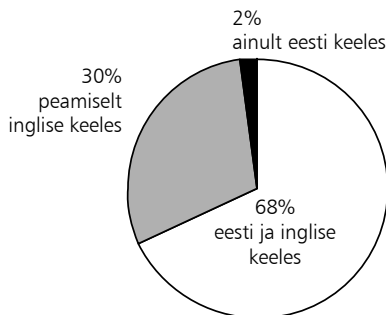


Joonis 2. TLÜ doktoritööde keel teadusalati (seisuga 27.02.2008)

Eestis on võrdlemisi head võimalused inglise keeles avaldamiseks olemas, nt humanitaarvaldkonnas Trames, Acta Historica Tallinnensia ja Linguistica Uralica, loodusteadlastel Proceedings of Estonian Academy of Sciences. Biology. Ecology; Proceedings of Estonian Academy of Sciences. Chemistry; Proceedings of Estonian Academy of Sciences. Engineering; Proceedings of Estonian Academy of Sciences. Geology; Proceedings of Estonian Academy of Sciences. Physics. Mathematics.

Rohkem kui kolmandik teaduspublikatsiooni küsimustele vastanutest (vt lisa 1, küsimused 9-10) on avaldanud viimase kahe aasta jooksul keskmiselt 2,4 publikatsiooni rahvusvaheliselt tunnustatud ajakirjades ja 4,1 artiklit muudes ajakirjades. Samas on levinud arvamus, et ka ajakirjanduslikku aktiivsust tuleks atesteerimisel arvesse võtta, nt Õpetajate Lehes või Hariduses ilmuvaid kasvatus- ja keeleteaduslikke üllitisi. Peab kahjuks tõdema, et ajakirjade kohta (vt lisa 1, küsimused 7-10) vastati üsna lohakalt, jättes ajakirjade nimetused välja kirjutamata.

68% vastanuist leiab, et publitseerida tuleb eesti ja inglise keeles. Samas arvab peaaegu kolmandik (30%), et publitseerima peaks peamiselt inglise keeles. Ainult eesti keeles publitseerimist pooldab vaid 2% laekunud vastustest (vt joonis 3).



Joonis 3. Mis keeltes tuleks publitseerida teadustöid

Tulemused näitavad (positiivne korrelatsioon $r = 0,65$, olulisusnivoo $p = 0,01$), et publitseerima peab eeskätt inglise keeles ja vaid rahvusvahelistel publikatsioonidel

on väärtus. Ehk siis need, kes väärtustavad rahvusvahelist publitseerimist, tunnistavad, et seda tuleb teha eeskätt inglise keeles. Samuti pidasid rahvusvahelise publitseerimise väärtustajad põhjendatuks seda, et eestikeelseid publikatsioone ei väärtustata ($r = 0,24$, $p = 0,05$).

Üldiselt ilmnas, et avaldatud rahvusvaheliste publikatsioonide arv oli keskmises korrelatsioonis arvamusega, et avaldada tuleb eeskätt inglise keeles ($r = 0,33$, $p = 0,01$).

Huvitava joonena ilmnas, et need, kes avaldasid kohaliku levikuga ajakirjades, tunnetasid vähem erialakeeleõppe taseme nõrkust ($r = 0,29$, $p = 0,05$) kui need, kes avaldasid kohalikes teadusajakirjades vähe. Seose tingib ilmselt teadusvaldkondlik erinevus: neis valdkondades, kus avaldatakse rohkem inglise keeles (loodusteadused), tunnetatakse teravamalt puudujääke erialase keele õppes. Erialadel, kus emakeelne publitseerimine on laiemalt levinud, on ka erialakeele õpetamise tase parem.

Vaieldamatult on oluline publitseerida teadusartikleid inglise keeles. Eesti teadlane peab oskama inglise keelt, et esineda rahvusvahelistel konverentsidel, pidada loenguid välisriikides ning lugeda teaduskirjandust, kuid tal on vaja juhendada üliõpilasi, propageerida oma uurimisvaldkonda, pidada ka eestikeelseid loengusarju ning koostada emakeelseid (kõrgkooli)õpikuid. Inglisekeelsete doktoritööde maht paneb mõtlema, kas edaspidi suudetakse ja tahetakse arendada eesti loodus- ja/või sotsiaalteaduste terminoloogiat.

EKAS (2004: 21) näeb ette erialakeeleõpikute koostamist ja oskuskeele arendamist nende abil. Seire käigus aga selgus, et 48% TLÜ teadustöötajatest ei pea eestikeelsete kõrgkooliõpikute kirjutamist võimalikuks ajanappuse tõttu. Kolmandik (34%) vastanutest leiab, et vajalikud on heatasemelised eestikeelsed õppematerjalid või koondkonspektid, sest õpikute kirjutamine on aeganõudev protsess ning paljude erialade puhul (eeskätt informaatika-, matemaatika- ja loodusteadustes) osutub õpik vananenuks juba enne trükki minekut. 17% arvab, et eestikeelse terminoloogia kestmist soodustavad ka veebipõhised kursused. Ühtlasi oleks kergendaks see ka ainekursuste ettevalmistamist ja läbiviimist.

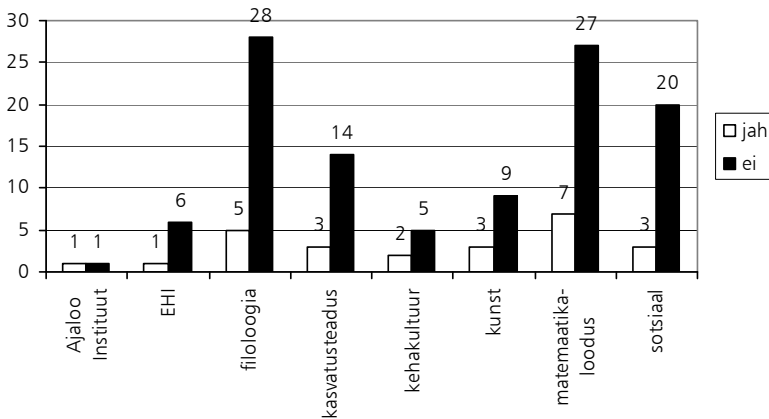
72% kõigist respondentidest arvab, et venekeelsete üliõpilaste kehv eesti erialakeele oskus pole nende tööd raskendanud. Samas peab 63% vajalikuks venekeelsetele üliõpilastele erialase eesti keele õpetamist kogu õpingute aja jooksul. Levinud on arvamus, et loodus- ja sotsiaalteadustes on eesti erialakeele õpetamine nõrk.

TLÜ teadustöötajad ei eelista selgelt võõrkeelse õpetuse suurendamist: enam-vähem võrdselt on esindatud positiivne ja negatiivne hoiak, ning pooled ei osanud sellele küsimusele üldse vastata. 70% vastanulist kasutab eesti- ja võõrkeelset õppekirjandust vahekorras 33-67%.

Järgnevalt vaadeldakse õppejõudude ankeetimise tulemusi küsimuste ja väidete kaupa, esiteks uuritakse teaduskeele koolitusvajadust, teiseks teaduskeele positsiooni.

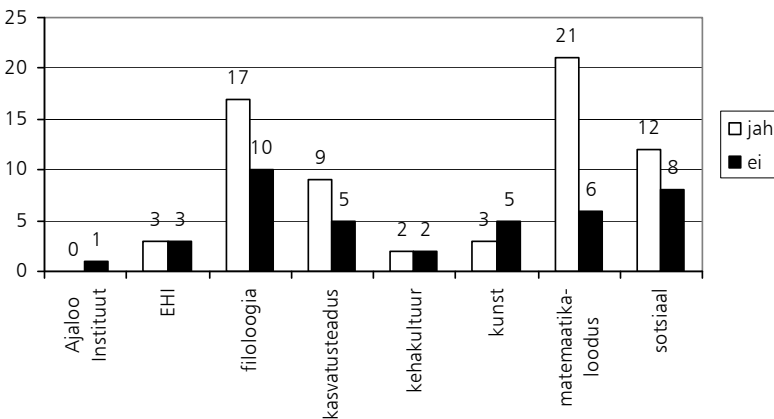
5.2. Teaduskeele koolitusvajadusest

Kas olete käinud kursustel, mis aitasid kaasa erialaste teadustekstidega töötamisel? (Lisa 1, küsimus 12.) Jaatavalt vastas 25 informanti (vt joonis 4), nendest ainult 17 täpsustas, mille poolest olid need kursused vajalikud. Nende vastuste põhjal peeti kasulikuks võimalust teadusteksti iseärasustega tutvuda, arendada oskust mõista teadusteksti loogikat, lihvida tõlkeoskust ja retsensioonide jms teadustekstide kirjutamise kunsti õppida. Toodi välja ka keeleõppe aspekt: lisaks tööle teadustekstiga õpitakse ka keelt, toetudes teaduskeele uuendamisele ja kirjakeelsele eneseväljendusele.



Joonis 4. Teaduskeele koolitustel osalus

Neilt, kes pole teaduskeele koolitustel käinud, küsiti, **kas soovitakse, et Tallinna Ülikool korraldaks vastavat koolitust** (lisa 1, küsimus 14). Joonisel 5 näeme, et koolitusvajadus on kõrge: filoloogia- ja sotsiaalteaduskonnas tunnevad eesti teaduskeele koolituse vastu huvi üle poole vastanud teadustöötajatest. Matemaatika-loodusteaduskonnas on huvilisi lausa kolm neljandikku. Probleemina toodi asjaolu, et koolitusel osalemiseks on raske aega leida.



Joonis 5. Eesti teaduskeele koolitusel osalemise soov

Kes peab need kursused läbi viima? (Lisa 1, küsimus 15a.) Võimalik oli valida kolme vastusevariandi vahel – oma ala tuntud ekspert, keeleteadlane või keelekorraldaja – või lisada oma vastus. Teaduskonniti jagunesid vastused järgnevalt: 30 matemaatika-loodus- ja sotsiaalteaduskonnas töötavat informanti arvas, et koolitama peaksid oma ala tuntud eksperdid; 12 filoloogiateaduskonna esindajat eelistasid keeleteadlasi; 3 kunstiteaduskonna töötaja arvates peaksid koolitust läbi viima keelekorraldajad.

Mida tuleb kursuste raames õpetada? ja Milliseid oskuseid soovite kursuste jooksul eriti arendada? (Lisa 1, küsimus 15b-c.)⁴ Teadustöötajate sõnul tuleks kindlasti õpetada uuemat terminoloogiat, mida olemasolevad erialasõnastikud ei kajasta, teadustekstile omast keelekasutust kõigil keeletesanditel, hea teadusteksti loomist, eriti aga arendada selle koostamise loogikat. Rõhutati, et teaduskeele kasutamist on väga raske õpetada, sest erialakeele oskused peavad põimuma peente keeleliste nüanssidega ning et eriti oleks eesti teaduskeele koolitus vajalik noorte teadustöötajate jaoks. Kursuste korraldajatelt oodati seda, et osalejad saaksid jõukohaseid ülesandeid ja et neid saaks ka praktiliselt harjutada.

5.3. Teaduskeele positsioonist

Asjaolu, et õppejõu akadeemilise edukuse näitajaks ei peeta praegu eesti keeles avaldatud teadusartikleid ja/või monograafiat, on õige. (Lisa 3, küsimus 2.) Väitega oli nõus 67 ja polnud nõus 68 teadustöötajat. Rõhutatakse, et palju sõltub väljaande kvaliteedist.

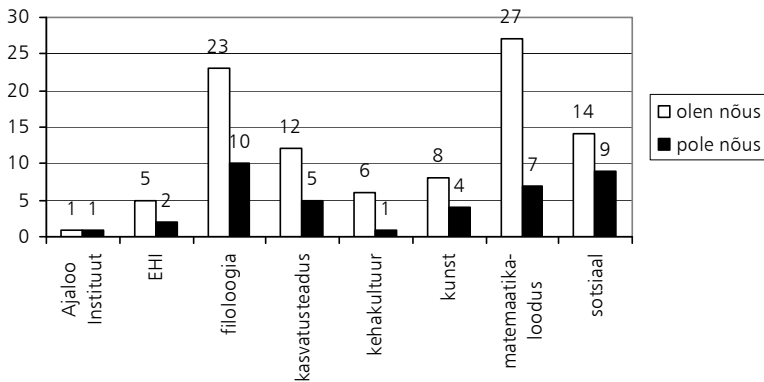
Nõue, et teadusartiklil on õige väärtus ainult sel juhul, kui see on avaldatud rahvusvaheliselt tunnustatud ajakirjas, on täiesti õigustatud. (Lisa 3, küsimus 3.) Jaatavalt on vastanud 63 ja eitavalt 72 teadustöötajat. Umbes pool vastanutest arvab, et teadustöötaja peab publitseerima nii eesti kui ka võõrkeeles, 27% arvab, et eeskätt inglise keeles ning 8% leiab, et piisab ainult eesti keeles avaldamisest.

Saadud vastustest võib järeldada, et TLÜ teadustöötajate arvates tähtsustatakse Eesti akadeemilistes ringkondades eesti keelt tõesti vähe (vt joonis 6). Selle väitega oli nõus üle poole iga struktuuriüksuse töötajatest. Eriti sageli olid jaatava vastuse valinud matemaatika-loodus- ja filoloogiateaduskonna töötajad – vastavalt 27 ja 23 korda.

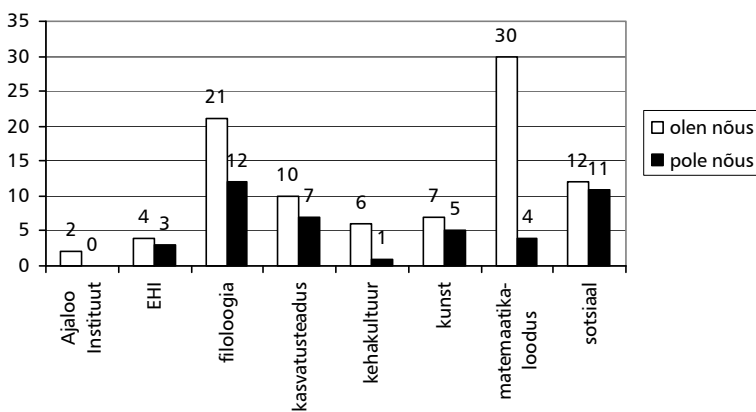
Ametlik terministandardimine aitab minu alal uute terminite loomisel. (Lisa 3, küsimus 8.) Väitega oli nõus kokku 92 teadustöötajat (vt joonis 7). Eriti kõrged näitajad olid jällegi matemaatika-loodus- ning filoloogiateaduskonnas – vastavalt 30 ja 21.

Sõnaraamatud, mis sisaldavad minu eriala uusimaid termineid ja nende definitsioone, on mulle vajalikud. (Lisa 3, küsimus 9.) Teistest rohkem olid väitega nõus matemaatika-loodusteaduste esindajad (vt joonis 8).

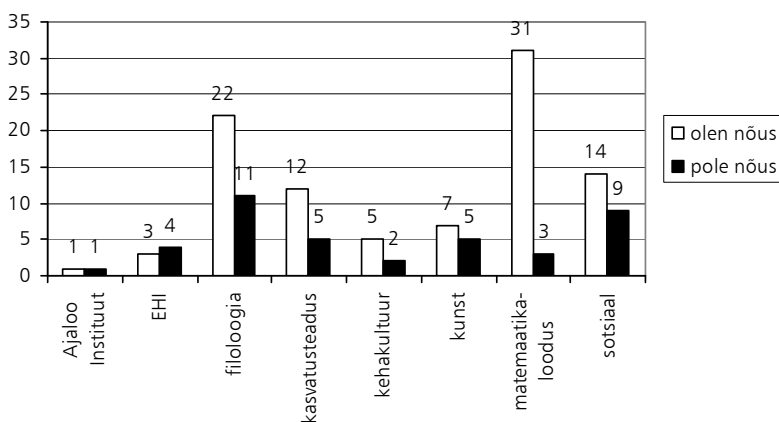
⁴ Kuna neile küsimustele vastas vähe intervjueeritavaid, ei tooda siin erinevusi teaduskondade lõikes välja.



Joonis 6. Teadustöötajate nõusolek väitega *Eesti keelt tähtsustatakse vähe Eesti akadeemilistes ringkondades*



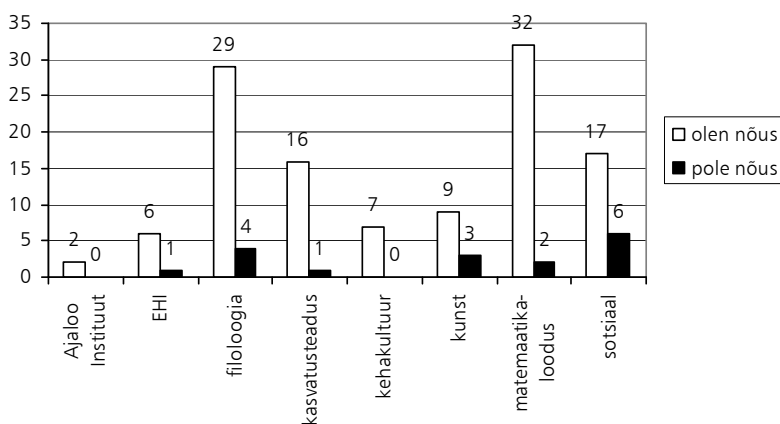
Joonis 7. Teadustöötajate nõusolek väitega *Ametlik terministandardimine aitab minu alal uute terminite loomisel*



Joonis 8. Teadustöötajate nõusolek väitega *Sõnaraamatud, mis sisaldavad minu eriala uusimaid termineid ja nende definitsioone, on mulle vajalikud*

Tõenäoliselt kasvab vajadus kasutada uusimaid termineid sisaldavat sõnaraamatut eestikeelsete loengute ettevalmistamiseks, eestikeelsete üllitiste või eestikeelsete ettekannete koostamiseks, sest teadustöötaja eesmärgiks on ka selge ja üldaktsepteeritava terminoloogia arendamine. See on aga suhteliselt keeruline ülesanne: autorid kirjeldavad uusi või võrdlemisi sarnaseid nähtusi erinevaid termineid kasutades. Et mitte tuua eesti teaduskeelde segadust, mis valitseb inglise keeles formuleeritud teooriates, on jätkuvalt vaja uute terminite loomist, mida oleks ohutu kasutada, ning olemasolevate terminite semantilise tähenduse laiendamist. Samas võib kommentaaridest välja lugeda, et palju oleneb ka eriala levikust, populaarsusest, tähtsusest ühiskonna seisukohalt. Toodi näide, et kui tegu on mõne Eestis vähe levinud teadusega (nagu arheoloogia), siis on vähe tõenäoline, et vastav sõnaraamat üldse koostatakse. 38% vastajatest kommenteeris, et kuigi antud sõnaraamatud oleks vajalikud, pole nende erialade terminisõnaraamatuid olemaski.

Loengumaterjali ettevalmistamisel ei õnnestu vältida inglise laenterminite kasutamist. (Lisa 3, küsimus 10.) Jooniselt 9 on näha, et jaatavaid vastuseid on kokku 118 ja mõtlemisainet see kindlasti annab. Siinkohal võib arutleda selle üle, kui võrd õigustatult kasutatakse laentermineid. Kahjuks ei ole võimalik kontrollida, kas sage laenterminite kasutamine on tingitud tõesti ainult eestikeelsete vastete puudumisest. Võib tõdeda, et tihti ei tegele vastava ala terminoloogia arendamisega ei keeleteadlased ega -korraldajad. Seetõttu võivad nõnda moodustatud toorlaenud olla liiga pikad või kohmakad, mis raskendab nende kasutamist ja tähendusest arusaamist. Huno Rätsep (2002: 176) on toonitanud, et uusi sünonüüme on eestikeelsel teadusel vaja just sarnaste mõistete eristamiseks ja täpsustamiseks. Ladinatüvelised terminid sobituvad eesti oskuskeelde hästi, kuid ei paista läbi ja seepärast on võõr-oma-terminipaar ikka parim. Omavaste otsimisel on väga hea toetuda korraga lähtekeele seletussõnaraamatule ja lähte-sihtkeele kakskeelsele sõnaraamatule.



Joonis 9. Teadustöötajate nõusolek väitega *Loengumaterjali ettevalmistamisel ei õnnestu vältida inglise laenterminite kasutamist*

Kui kõrvutada kõiki vastuseid, ilmneb, et teadustöötajad tajuvad eesti teaduskeele arendamise vajadust teravalt. Kõigile mõistetavalt oma ideid esitav eesti haritlaskond suudab tõhusamalt edendada Eesti arengut. Keele selguse ja täpsuse teadlik

viljelemine kandub sedakaudu igäiheni. Eriti oluline on täpsus teaduskeeles, tagades ka õpetlaste omavahelise parema mõistmise. Seega on eesti oskuskeeles, sh teaduskeeles seire, uurimine, ülikooli- ja täiendusõpe Eesti ühiskonna strateegiliste huvide seisukohalt esmatähtis.

6. Järeldusi

Teadustöötaja ankeedi abil koguti andmeid eesti keeles õpetamise, avaldamise ning eesti teaduskeelde suhtumise kohta. Püüti luua pilt eesti teaduskeeles kasutamise kohta TLÜ-s ning koguda kvantitatiivset infot eri teadusvaldkondades peetud ainekursuste kohta.

Eesti keelest rääkides rõhutatakse alati, et on olemas eestikeelne kõrgharidus ja teadus. Uurimistulemused kinnitavad, et TLÜ teadustöötajate arvates suhtutakse eesti teaduskeelde üleolevalt. Omaette probleem on erinev eesti- ja ingliskeelsete (teadus)publikatsioonide väärtus teadusaruandluse jaoks. Kahjuks ei ole Eestis mitmel (kitsamal) erialal eelretsenseerivaid ajakirju. Mitme eriala teadustöötajad tunnistavad, et ingliskeelsetele terminitele on üha rohkem vaja arusaadavaid eestikeelseid vasteid leiutada, seda nii loengute ettevalmistamise kui ka ainekursuste pidamise ajal. Teadustekstide koostamine on oskus, mis vajab järjekindlat lihvimist.

Kui enamik teadustöötajaid leiab, et eesti keel teaduskeelena on alahinnatud ja eestikeelsed terminid on ülivajalikud ning naljalt ei õnnestu vältida inglise laenterminite kasutamist, siis on vajadus eesti teaduskeele koolituse järele tõesti kõrge. Ühtlasi aitab see niihästi teaduskeelt lodusamaks ja täpsemaks muuta kui ka suhtumist parandada.

Eestikeelsed kõrgkooliõpikud on vajalikud seetõttu, et neid koostades korrustub teaduskeel ja mõistestik ning luuakse uusi termineid. Eesti teaduse eri valdkondadesse kuuluvaid selgeid termineid vajavad nii teadlased, õppejõud kui ka üliõpilased. Kindlasti tuleks tihendada koostööd eesti teaduskeelekorralduse ja ametliku terministandardimise vahel. On aeg muuta eesti teaduskeele uurimine ja õpetamine süsteemseks, kogu Eesti kõrgharidust ja teadust hõlmavaks pidevaks protsessiks.

Kirjandus

- Ammon, Ulrich (ed.) 2001. *The Dominance of English as a Language of Science*. Berlin: Mouton de Gruyter.
- Ammon, Ulrich 2003. *The international status of the German language*. – Jacques Maurais, Michael A. Morris (eds). *Languages in a Globalising World*. Cambridge: Cambridge University Press, 231–249.
- Cheshire, Jenny (ed.) 1991. *English Around the World: Sociolinguistic Perspectives*. Cambridge: Cambridge University Press.
- Crystal, David 1995. *The Cambridge Encyclopedia of the English Language*. Cambridge: Cambridge University Press.
- Crystal, David 1997. *English as a Global Language*. Cambridge: Cambridge University Press.
- EKAS 2004 = Eesti keele arendamise strateegia 2004–2010. Haridus- ja Teadusministeerium. Eesti Keelenõukogu. Tartu, 2004. <http://www.eki.ee/keelenoukogu/> (27.02.2008).

- EKHS 2006 = Eesti kõrgharidusstrateegia aastateks 2006–2015. <http://www.hm.ee/index.php?popup=download&id=4778> (27.02.2008).
- Gibbs, W. Wayt 1995. Lost science in the third world. – *Scientific American*, August, 76–83.
- Graddol, David 2000. *The Future of English?* London: The British Council. <http://www.britishcouncil.org/learning-elt-future.pdf> (27.02.2008).
- Haarman, Harald; Holman, Eugene 2001. The impact of English as a language of science in Finland and its role for the transition to network society. – Ulrich Ammon (ed.). *The Dominance of English as a Language of Science*. Berlin: Mouton de Gruyter, 229–260.
- McArthur, Tom 1998. *The English Languages*. Cambridge: Cambridge University Press.
- Nemvalts, Peep 1999. Pigem oma kui võõras! (“Muutuv keel”). – *Keel ja Kirjandus* 11, 756–763.
- Raukas, Anto 2002. Eesti Entsüklopeedia eesti teaduskeele hoidjana. – Tõnu Tender, Elsa Pajumaa (toim.). *Eesti keel teaduskeelena ja Euroopa Liit*. Tallinn: Tallinna Teadlaste Maja väljaanne, 73–83.
- Rätsep, Huno 2002. Sõnaloo raamat. Tartu: Ilmamaa.
- Tallinna Ülikooli arengukava aastateks 2004–2007. <http://www.unit.ee/dokid/arengukava.htm> (27.02.2008).
- Tallinna Ülikooli arengukava aastateks 2006–2010. <http://www.tlu.ee/files/arts/712/Areng6179dd412159910b6aed1cc6faac9fad.pdf> (27.02.2008).
- Tallinna Ülikooli põhikiri 2006. <http://www.tlu.ee/files/arts/4088/TLY%20p594b11e-a91ac59974270daea3e280331.pdf> (27.02.2008).
- UUEKS = Projekt “Uus eesti keele ja kakskeelse õppe süsteem Tallinna Ülikoolis” (2006–2008). www.tlu.ee/uueks (27.02.2008).
- Viereck, Wolfgang 1996. English in Europe: its nativisation and use as a lingua franca, with special reference to German-speaking community. – Roland Hartmann (ed.). *The English Language in Europe*. Oxford: Intellect, 16–24.

Anastassia Zabrodsckaja (Tallinna Ülikool) on lõpetanud Tartu Ülikooli Narva Kolledži eesti keel teise keelena erialal. Kaitsnud magistrikraadi Tallinna Ülikoolis 2005. a eesti filoloogia erialal. Alates 2005. a on Tallinna Ülikooli lingvistika doktorant ning töötab teadurina eesti keele ja kultuuri instituudis. On uurinud keelekümblust, laste kakskeelsust, vene-eesti koodivahetust ja venekeelsete üliõpilaste vene ning eesti keele kasutust, tegeleb teaduskeele seirega.
 anastaza@tlu.ee

Lisa 1. Üldküsimused.

1. Teaduskond:
 - a. Osakond:
 - b. Õppetool:
 - c. Amet:
 - d. Sugu: M N
 - e. Vanus: alla 25 26–29 30–34 35–39 40–44
 45–49 50–54 55–59 üle 60
2. Kraad ja selle omandamise aasta:
 - f. mis ülikooli olete lõpetanud:
 - i. mis oli(d) õppekeel(ed):
3. Keelteoskus:
 - g. emakeel (esimene omandatud keel):
 - h. mis keeli kasutate igapäevases suhtluses?
 - i. mis keel(t)es oskate erialast vestlust pidada?
 - j. mis keel(t)es oskate erialast kirjandust lugeda?
 - k. mis keel(t)es oskate erialaseid artikleid kirjutada?
 - l. mis keel(t)es oskate ettekannet pidada?
 - m. mis keelt/keeli sooviksite juurde õppida?
 - ii. mis eesmärgil?
 - iii. kus?
 1. välismaal
 2. TLÜ kursustel
 3. muul viisil (kuidas täpsemalt?)
4. Kui kaua olete Tallinna Ülikoolis töötanud:
 - n. alla 1 aasta
 - o. 1 aasta
 - p. 2–4 aastat
 - q. 5–9 aastat
 - r. 10–14 aastat
 - s. 15–19 aastat
 - t. 20–24 aastat
 - u. 25–29 aastat
 - v. 30–34 aastat
 - w. üle 35 aasta
5. Mis keel(t)es olete oma teaduspublikatsioone avaldanud?
6. Mis keel(t)es tuleb Teil teaduslikke artikleid kirjutada?
7. Kas Teil on võimalik publitseerida eelretsenseeritavates eestikeelsetes ajakirjades? Kas võite neid ajakirju nimetada?
8. Kas Teil on võimalik publitseerida Eestis väljaantavates eelretsenseeritavates ingliskeelsetes ajakirjades? Kas võite neid ajakirju nimetada?

9. Mitu teaduspublikatsiooni ja millistes rahvusvaheliselt tunnustatud ajakirja(de)s olete avaldanud viimase paari aasta jooksul?

- | | |
|--------|------------|
| x. 0 | aa. 7–9 |
| y. 1–3 | bb. üle 10 |
| z. 4–6 | |

10. Mitu teaduspublikatsiooni ja millistes muu(de)s ajakirja(de)s olete avaldanud viimase paari aasta jooksul?

- | | |
|---------|------------|
| cc. 0 | ff. 7–9 |
| dd. 1–3 | gg. üle 10 |
| ee. 4–6 | |

11. Missuguseid loengukursusi ja mis keeles olete õpetanud (loetlege vähemalt 5 kursust, märkides ka õppekeelt/-keeli).

Kui loengu ajal kasutati mitut keelt, täpsustage palun, miks (näiteks õppejõud tõlkis vene keelde, et venekeelsed üliõpilased saaksid loengu sisust paremini aru; õppejõud kasutas ingliskeelset terminoloogiat vajaliku erialase eestikeelse terminoloogia puudumise tõttu jms):

- a. ainekursus:
 - õppekeel(ed):
- b. ainekursus:
 - õppekeel(ed):
- c. ainekursus:
 - õppekeel(ed):
- d. ainekursus:
 - õppekeel(ed):
- c. ainekursus:
 - õppekeel(ed):

12. Kas olete käinud kursustel, mis aitasid kaasa erialaste teadustekstidega töötamisel?

- a. jah
- b. ei

13. Kui jah, mille poolest olid need kursused kasulikud ja vajalikud?

14. Kui ei, kas sooviksite, et Tallinna Ülikool korraldaks vastavat koolitust?

- c. jah
- d. ei

15. Kui jah, vastake palun täpsemalt:

- e. kes peab need kursused läbi viima?
 - i. oma ala tuntud eksperdid
 - ii. keeleteadlased
 - iii. keelekorraldajad
 - iv. keegi teine (täpsustada)
- f. mida tuleb kursuste raames õpetada?
- g. milliseid oskuseid soovite kursuste jooksul eriti arendada?
- h. mida veel ootaksite kursuste korraldajatelt?

Lisa 2.

Palun teid vastata alljärgnevale küsimustele kõigi ainete kohta, mida andsite eelmisel semestril:

Tuleb märkida, mitu kursust õppejõud luges:

1. Kursuse nimetus:
2. Kursus oli:
 - a. kohustuslik aine
 - b. kõrvalaine
 - c. vabaaine
 - d. muu (täpsustada)
3. Kursuse kuulajad olid:
 - a. BA-üliõpilased (vana kava järgi)
 - b. MA-üliõpilased (vana kava järgi)
 - c. BA-üliõpilased (uue kava järgi)
 - d. MA-üliõpilased (uue kava järgi)
 - e. PhD-üliõpilased
4. Üliõpilased olid:
 - a. eesti emakeelega
 - b. vene emakeelega
 - c. muu emakeelega (v.a vene)
 - d. segarühm (*a* ja *b* koos)
 - e. segarühm (*a* ja *c* koos)
 - f. segarühm (*a*, *b* ja *c* koos)
 - g. segarühm (*b* ja *c* koos)
5. Kui oskate, siis nimetage palun, mis eriala(de)lt olid need üliõpilased:
6. Hinnake erineva eesti keele tasemega üliõpilaste koostöö produktiivsust:
 - a. oli produktiivne
 - b. ei olnud produktiivne
 - c. sõltus loenguvormist
 - d. sõltus üliõpilastest endast
7. Venekeelsete üliõpilaste kehv eesti erialakeele oskus on õppejõu tööd raskendanud:
 - hh. jah
 - ii. ei
8. Antud kursust oleks pidanud lugema eesti- ja venekeelsetele üliõpilastele eraldi:
 - jj. jah
 - kk. ei

9. Kas antud aines on olemas eestikeelne kõrgkooliõpik?
 ll. jah
 i. mina olen selle õpiku autor
 ii. mina olen selle õpiku üks autoritest
 mm. ei
10. Kas Te peate eestikeelse õpiku kirjutamist antud valdkonnas vajalikuks? Miks?
 nn. jah
 oo. ei
11. Kas kavatsete eestikeelset õpikut ise kirjutada? Miks?
 pp. jah
 qq. ei
12. Kas olete kursuse raames üliõpilastele õppematerjale koostanud?
 rr. jah – vastake ka järgmistele küsimustele.
 ss. ei – jätkake vastamist küsimusest 18.
13. Mis keel(t)es need õppematerjalid olid?
14. Miks oli vajalik koostada õppematerjalid just selles keeles/neis keeltes?
15. Kas õppematerjalid olid mõeldud:
 tt. kõikidele üliõpilastele
 uu. venekeelsetele üliõpilastele
 vv. võõrkeelsetele üliõpilastele (v.a vene)
16. Kas materjali koostamisel puutusite kokku vajaliku eestikeelse terminoloogia puudumisega?
 ww. jah
 xx. ei
17. Kui jah, siis mida Teil tuli teha?
18. Kas ainekursuse pidamisel tuli Teil eestikeelseid termineid leiutada?
 yy. jah
 zz. ei
19. Mis keeles oli kohustuslik kirjandus ja miks?
 a. ainult eesti keeles
 b. eesti keeles
 c. hispaania keeles
 d. inglise keeles
 e. itaalia keeles
 f. ladina keeles
 g. prantsuse keeles
 h. rootsi keeles
 i. saksa keeles
 j. soome keeles
 k. vene keeles
 l. muu(de)s keel(t)es (täpsustage palun)

20. Mis keeles oli soovitatav kirjandus ja miks?

- a. ainult eesti keeles
- b. eesti keeles
- c. hispaania keeles
- d. inglise keeles
- e. itaalia keeles
- f. ladina keeles
- g. prantsuse keeles
- h. rootsi keeles
- i. saksa keeles
- j. soome keeles
- k. vene keeles
- l. muu(de)s keel(t)es (täpsustage palun)

21. Andke vastus protsentides, mil määral tuli kasutada eesti- ja võõrkeelset kirjandust:

22. Mis oleks kursuse ettevalmistamist, läbiviimist kergendanud?

Lisa 3.

Olen või pole nõus järgmiste väidetega:

1. Eesti keelt tähtsustatakse vähe Eesti akadeemilistes ringkondades.
 - a. olen nõus
 - b. pole nõus
2. Asjaolu, et õppejõu akadeemilise edukuse näitajaks ei peeta praegu eesti keeles avaldatud teadusartikleid ja/või monograafiat, on õige.
 - a. olen nõus
 - b. pole nõus
3. Nõue, et teadusartiklil on õige väärtus ainult sel juhul, kui see on avaldatud rahvusvaheliselt tunnustatud ajakirjas, on täiesti õigustatud.
 - a. olen nõus
 - b. pole nõus
4. Võõrkeelse õpetuse osatähtsust Tallinna Ülikoolis peaks:
 - a. suurendama. Miks?
 - b. vähendama. Miks?
5. Minu valdkonnas on eesti erialakeele õpetamine nõrk.
 - a. olen nõus
 - b. pole nõus
6. Õppejõud peab publitseerima:
 - a. ainult eesti keeles.
 - olen nõus
 - pole nõus
 - b. nii eesti kui ka võõrkeeles.
 - olen nõus
 - pole nõus
 - c. eeskätt inglise keeles.
 - olen nõus
 - pole nõus
7. Mis keeles peaks Teie erialal publitseerima?
8. Ametlik terministandardimine aitab minu alal uute terminite loomisel.
 - a. olen nõus
 - b. pole nõus
9. Sõnaraamatud, mis sisaldavad minu eriala uusimaid termineid ja nende definitsioone, on mulle vajalikud.
 - a. olen nõus
 - b. pole nõus

10. Loengumaterjali ettevalmistamisel ei õnnestu vältida inglise laenterminite kasutamist.

a. olen nõus

b. pole nõus

11. Muukeelsetele üliõpilastele peab õpetama erialast eesti keelt ainult esimese aasta jooksul.

a. olen nõus

b. pole nõus

12. Muukeelsetele üliõpilastele peab õpetama erialast eesti keelt kogu õpingute aja jooksul.

a. olen nõus

b. pole nõus

Kas Teil on veel mingeid kommentaare/soovitusi? Kõik ettepanekud ja asjakohane kriitika ka antud ankeedi kohta on teretulnud.

MONITORING ACADEMIC LANGUAGE AT TALLINN UNIVERSITY: ACADEMIC STAFF ATTITUDES

Anastassia Zabrodskaia

Tallinn University

The world's predominant academic language is English, and unquestionably, Estonian scientific activities are continually moving towards the increasing use of English. It is important that university students and staff have the opportunity to publish their research in good quality Estonian. The article deals with measures taken at Tallinn University (TLU) to promote research and academic writing in Estonian.

The project "The New System of Estonian and Bilingual Teaching in Tallinn University"⁵, with the goal of assuring the development of academic Estonian in all fields of research, was started in 2006. As part of the project, the status of academic Estonian is being monitored.

The paper is comprised of four chapters. Firstly, the state of English in scientific world is outlined. Also, documents that try to retain a prominent place for Estonian in university curricula are described. Given that, the methodology of the first stage of monitoring is explained. Then the results are discussed. The paper examines the attitudes of university academic worker's towards academic Estonian and how great the need of training concerning the use of the academic Estonian language is. The main research questions are: Where does TLU staff publish their research results? Are there possibilities to publish in the Estonian language? What kinds of problems and needs related to the preparation of lecture/study materials in Estonian do professors have?

In general, those scholars who have graduated from universities abroad use more foreign languages in their everyday life. Academics with doctoral degrees use foreign languages more than PhD candidates or master's degrees. It was found that the more foreign languages are used, the less the need for terminology dictionaries is felt. Those who value publications on the international level, acknowledge that this should be done first and foremost in English. Also, these respondents agree with the opinion that the publications written in Estonian are not considered as relevant as international level publications. The results show that the number of international publications was in correlation with the opinion that a scholar has to publish primarily in English. It should be noted that those who publish mainly in journals with local distribution, recognize the weak level of teaching of Estonian for specific purposes in their field of study more rarely than those who publish less in Estonian and more internationally. This also correlates with different scholarly domains. For instance, the data show that academics in the natural sciences, where publications are mostly in English, are in greater need of courses on academic Estonian. Quality terminology dictionaries are needed for preparation of courses, articles or reports held in Estonian. And it is considered appropriate, that compiling university textbooks in the Estonian language should be paid greater attention.

Keywords: academic language, English, Estonian

KUIDAS ME MÕISTAME MÕISTEID?

Ene Vainik, Toomas Kirt

Ülevaade. Artiklis antakse ülevaade erinevatest teoreetilistest lähenemistest mõistete olemusele ning nende vastasmõjust lingvistilise semantikaga. Käsitlemist leiavad klassikaline, prototüübi ja nn teooria teooria, lisaks võrgumudel, ruumiline, dünaamiline ja evolutsiooniline lähenemine. Ühe ja ainsa õige mõisteteooria väljaselgitamise asemel jõutakse järeldusele, et erinevad teoreetilised lähenemised on mõistete juures tähtsustanud eri aspekte olenevalt seatud eesmärgist. Artiklis näidatakse mitmesuguste lingvistikas käibel olevate semantiliste esituste seost vastavate mõisteteooriatega.*

Võtmesõnad: lingvistika, semantika, epistemoloogia, kognitiivteadus

1. Sissejuhatus

Õpetlaste arutelu mõistete olemuse üle on saanud alguse muistsest Kreekast ja kestab tänapäevani, ühtne arusaam mõistete olemusest aga puudub tänini.¹ Kui varem kuulusid mõisted abstraktse mõtlemise vormidena pigem loogika ja filosoofia² pärusmaale siis tänapäeval on mõiste korruga paljude distsipliinide huviorbiidis. Nüüd tunnevad mõiste vastu huvi ka (neuro)psühholoogia, lingvistika, informaatika, matemaatika ja loomulikult eelkõige kognitiivteadus (ingl *cognitive science*) kui interdistsiplinaarne uurimisala, mis seab tunnetusprotsesside seaduspärasuste avastamise ja kirjeldamise oma esmaseks uurimisülesandeks.

Lingviste sunnib mõiste olemuse üle mõtisklema ühest küljest teoreetiline huvi ning soov mõista inimkeele aluseks olevaid tunnetusmehhanisme; teisest küljest on vajadus aga ka puhtpraktiline. Näiteks terminitöös on tarvis seada kindel termin vastavusse kindla mõistega ja tagada terminitõlgete adekvaatsus tähistatava sisu alusel (vt nt Tavast 2004). Mõistepõhiselt käib ka tesaurusetüüpi sõnastike³ ning elektroonsete ressursside, nagu WordNet, koostamine (Fellbaum 1999: 7). Semantilise komponendi kirjeldusteta ei pääseta ka tehisintellektis ja

* Artikkel on valminud ETF-i grandri nr 7149 toel.

¹ On tunnuslik, et püüd mõistest teadaolevat kokku võtta on päädinud mitte integreeritud vastusena küsimusele *mis on mõiste?*, vaid asjassepuutuvate olulisemate kirjutiste mahuka kollektisioonina (vt Laurence, Margolis 1999).

² Filosoofidest on mõistete üle arutlenud nt Platon, Aristoteles, Kant, Locke.

³ Sedalaadi töö pioneeriks eesti keele osas on A. Saareste "Eesti keele mõisteline sõnaraamat" (1958).

keele modelleerimises. Mõiste ja tähenduse vahekorra osas aga üksmeel puudub: mõned autorid samastavad mõiste ja leksikaalse tähenduse (nt Cruse 2000), mõned peavad mõistet tähendusest avaramaks (nt Aitchison 2003), mõned räägivad mõiste tähendusest (Osgood jt 1975), mõned keeleüksuste mõistelisest tähendusest (Leech 1990 [1974]); kognitiivse lingvistika paradigmas tuuakse tähenduse avamiseks sisse mõistestamine (ingl *conceptualisation*) (Langacker 1987, 2002) jne.

Üldise arusaama kohaselt on nn leksikaalsed mõisted vastavuses sõnade ehk nende sümboolse esitusega (vt Laurence, Margolis 1999). Iga kord, kui me millegi kirjeldamiseks või määratlemiseks kasutame sõnu, toetume ka mõistetele. Paradoks on selles, et mõisteid defineerides oleme olukorras, kus kasutame ühtede sümbolite sisu edasiandmiseks teisi sümboleid. Tekib suletud süsteem ning küsimus, kuidas siiski sõnad kui mõiste sümboolid tähenduse omandavad, jääb vastuseta (Harnad 1999).

Sõnade ja mõistete vahelise seose uurimist on peetud lausa mädasooks – alaks, mis on täis keerukust ja võhiklikkust (Aitchison 2003: 43). Ka käesoleva artikli maht ei luba meil üritada avada mõiste erinevate mõistmisviiside ja mõiste ning tähenduse vahekorra kogu sügavust ja mitmepalgelisust. Kuivõrd on ilmne, et keel (sõnadel) on mõistete mõistmises täita oluline roll, seame järgnevas eesmärgiks tutvustada erinevaid mõiste mõistmise viise ja nende vastasmõju lingvistilise semantikaga. Tulemusena peaks lugeja suutma orienteeruda erinevate lähene-misviiside rägastikus ja saama ülevaate sellest, kuhu ollakse tänapäeval mõistete olemuse seletamisel jõutud.

2. Naiivne arusaam

Kõige igapäevasemas mõttes tuleb *mõiste* mõiste käibele praktilisest vajadusest võtta kuidagi kokku sõnadetagune reaalsus (ld *conceptum* < *concupere* = *com* 'kokku' + *capere* 'võtma'). Näiteks ühiskonna reguleerimisel õigusaktidega on ülioluline kokku leppida, et sõnu kasutatakse järjepidevalt samade reaalsuskogumite tähistamiseks. Mõistete täpne maht ja kirjeldamine on väga oluline ka teaduses. Mistahes teadusliku käsitluse juurde kuulub selle põhimõistete määratlemine ja defineerimine. Seega oleks mõistete üks aspekte justkui nende piiritletus ja defineeritavus. Neilt eeldatakse ka vähemalt mõningat püsivust ja mentaalset manipuleeritavust, mis panebki meid mõistetest tavaliselt mõtlema kui iseseisvatest entiteetidest või asjadest.

Kuigi teadaolevalt ei ole neid asju seni keegi oma silmaga näinud,⁴ ei näi see tekitavat kahtlusi nende olemasolus. Me oleme rahul sellega, et mõisted toimetavad omasoodu, mõjutades seda, kuidas me asjadest aru saame. Kui me igapäevaselt mõtleme, arutleme, järeldame, aga ka unistame ja fantaseerime ning oma kaaslastega seda kõike jagame, siis ise sellest teadlikud olemata toetume paljudele ühistele mõistetele. Mõistete jagatus, kollektiivsus tagab selle, et me suudame teha end arusaadavaks teistele ning mõista teiste poolt väljendatut. Keele õppimisega paralleelselt me omandame ka vastavas kultuuris käibel olevad mõisted. Seega näib, et mõisted nagu neid vahendav keelgi on olemas enne meid ja me lihtsalt õpime neid kasutama.

Mida olemuslikku on eesti keel oma kasutajate jaoks sõnasse *mõiste* talletanud? Substantiiv *mõiste* on *e*-tuletis verbist *mõistma*, millel EKSS-i järgi on 4 põhilist

⁴ Küll aga kultiveerivad mõned keeled arusaama mõistete kasutatavusest, nt soome keeles on mõiste vasteks *käsi-te*, ka meie räägime mõistmisega paralleelselt *käsitamisest*. Inglise keelt kõnelevad rahvad suudavad aga mõisteid *haarata* (vt Jackendoff 1999).

tähendust: (i) 'taipama, aru saama, selgusele jõudma', nt *Ma mõistan, mida sa tunnend*; (ii) 'mõistatama', nt *Mõista, mõista mis see on?*; (iii) 'oskama', nt *Kas sa saladust pidada mõistad?*; (iv) 'hinnangut andma, otsust langetama, määrama', nt *Kurjategija mõisteti süüdi*. Seega võib mõiste olla seotud nii taipamise, äraarvamise, oskamise kui otsustamisega, mis kõik on seotud inimese vaimsete võimete rakendamisega.

Ehkki *e*-liitega tuletatakse regulaarselt teosubstantiive (nt *teke = tekkimine*) ja protsesside üksikjuhtumeid (nt *hüpe* vrd *hüppamine*), kuulub sõna *mõiste* "muude deverbaalide" rubriiki, mille tähendus võib olla nt tegevusvahend, saadus, tulemus (EKG I: 504). Seega võib *mõiste* olla mõistmise protsessis nii vahendiks, saaduseks kui tulemuseks. Tegu on kindlaskujunenud leksikoniüksusega, kus liite tähendust on sõnatähendusest raske eristada.

Verbi *mõistma* etümoloogia osutab, et tegu on vana soome-ugri sõnaga, millel on vasteid olemas paljudes sugulaskeeltes, nt on soome sõna *muistaa* 'mäletada' sama algupäraga (Mägiste 1982–1983). Algseks tüvekujuks on rekonstrueeritud **muja-* (Häkkinen 2004). Otsides sugulaskeelte toel selle verbi võimalikku algsemat ja konkreetsemat tähendust, jõuame lüüdi ja vepsa verbini *mujada*, tähendustega 'maitsta, katsuda, haista'. Seega on vaimsete tegevuste tähendused eesti ja soome verbidel *mõistma* ja *muistaa* tagasi viidavad tajuprotsesse nimetavale tähendusele, mis on hästi kooskõlas tendentsidega ka paljudes indo-euroopa keeltes (Sweetser 1990).

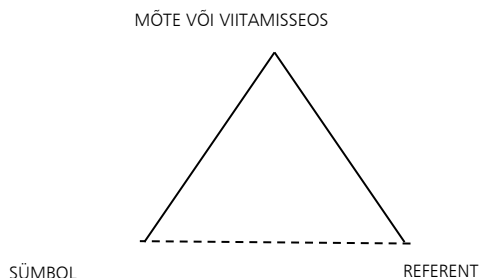
3. Mõiste mitmest vaatenurgast

Kõik teadusalad, mis mõistega tegelevad, on ühel meelel selles, et mõisted on häda- vajalikud selliste komplekssete organismide nagu inimene edukaks toimetulekuks. On leitud, et meie kogemus oleks organiseerimatu kaos ja me ei suudaks sellest õppida, kui me ei lahterdaks seda stabiilsetesse kategooriatesse. Iga uus kogemus oleks kordumatu ja eelnevast ei oleks kasu selle interpreteerimisel. Jagatud kategooriaid peetakse ka suhtluse eelduseks (Cruse 2000: 127).

On spetsiifikast lähtuvalt defineeritakse mõistet eri teadusaladel siiski mõnevõrra erinevalt. Näiteks entsüklopeedia annab loogikast lähtuva definitsiooni: mõiste on "abstraktse mõtlemise vorm, mis peegeldab tegelikkuse esemeid ja nähtusi nende oluliste tunnuste, seoste ja suhete kaudu [---] Mõiste peegeldab tegelikkust adekvaatsemalt kui taju või kujutus" (EE). Psühholoogias on mõistet defineeritud kui filtrit, mille läbi kogemust sõelutakse (Russell, Lemay 2000: 491), aga ka kui teadmiste hulka, mida vaikumisi kasutatakse kõrgemale tunnetusele aluseks olevates kognitiivsetes protsessides (Machery 2007: 20). Lingvistikas on defineeritud mõisteid kui salvestatud teadmiste kimpe, mis esindavad meie kogemuses esinevate sündmuste, entiteetide, situatsioonide jm liigendatust (Cruse 2000: 127).

Mida on mõistritel kui abstraktse mõtlemise vormidel, teadmiste kimpudel või tajufiltritel tegemist keeleteadusega? Lingvistiline perspektiiv mõistele saab alguse keelest, st vaevalt satuks lingvistide huviorbiiti mõiste ilma selle keelelise tähistuseta. Pigem on ikka vastupidi: lingvistika läheb liikvele keele uurimisest, ja kuivõrd keelemärk on tähistaja ja tähistatava ühtsus, siis tuleb kirjeldada ka tähistatavate tähendust. Siin tulebki mõnikord mängu mõiste.

Denotatsioonide võimalikult täpses kirjeldamises on nähtud sõnasemantika ülesannet (Karlsson 2002: 238). Kui tegu ei ole just budistliku keelekäsitlusega,⁵ on keeleline tähendus määratletud positivistlikult: miski (tähistaja e keeleline sümbol) on vastavuses millegi muuga (tähistatav e referent). Näiteks sõnaga *tool* oleks justkui vastavuses tähistatav mööbliese. Asjad tähistatavaga pole siiski nii lihtsad: klassikalisesena on käibel Charles Ogdeni ja Ivor Richardsi (1989 [1923]) semiootiline kolmnurk (vt joonis 1), mis reaalse maailma objekti kui võimaliku referendi kõrval seab sümboliga vastavusse ka mõtte või viite (ingl *thought or reference*). Viimane on sümboli suhe mõiste või ideega, st mingi kujutisega inimmees (Karlsson 2002). Pidev joon tähendab joonisel 1 otsesest ja katkendjoont kaudset vastavust. Meie tooli näites seega tähistaks sõna *tool* otseselt mõtet toolist või viitamisest tooli mõistega ja alles selle viitamisest kaudu ka vastava esemega. Siin on *mõiste* käibel abistava terminina tähenduse olemuse avamisel.



Joonis 1. C. Ogdeni ja I. Richardsi semiootiline kolmnurk (1989 [1923])

Samuti semantika klassikuks peetav John Lyons (1977: 96) asetab mõiste aga juba otse semiootilise kolmnurga tipuks (ingl *signification*), vahetades sellel kohal välja originaalis olnud mõtte või viitamisest. Referendi asemel kasutab ta terminit *significatum*. Sõna *tool* on otsesest mõistega *TOOL* ja viimane on otsesest tooli kui esemega. Üldistus, et keeleline sümbol kui tähistaja on vastavuses referendi e tähistatavaga mõiste kaudu, on kujunenud tunnustatud tõdemuseks. Sellist vaadet keele ja maailma seoste, mis määratleb väljendi või sõna tähenduse mõistena, on hakatud nimetama kontseptualismiks (Lyons 1977: 110).

Kuigi kirjanduses arutletakse jätkuvalt selle üle, kas üldse on olemas mingi abstraktne mõistete tasand, mis seisab eraldi sõnade tähendusest, või on sõnade tähendused mõistetega samastatavad (Aitchison 2003: 43), tundub, et üldiselt usutakse tänapäeval mingite keeleüksuste vastavate mentaalsete representatsioonide olemasolu inimese ajus. Käibel on terminid nagu *mentaalne leksikon* (Aitchison 2003) või *mentaalkeel* (*mentalese*, Pinker 2000). Järgnevalt tutvustamegi erinevaid lähenemisi mõistetele kui võimalikele tegelikkuse representatsioonidele.

3.1. Klassikaline lähenemine

Läbi ajaloo on arvatud, et kategooriad on kindlalt piiritletavad ja et leksikaalsed mõisted on määratletavad nende kasutamiseks vajalike ja piisavate tingimuste teadmise kaudu. Näiteks mõistet *VANAPOISS* me saaks kasutada, kui kõnealune isik

⁵ Budistlikus keeleteoorias lähenetakse keele tähistusvõimele hoopis negatsiooni kaudu: näiteks sõnaga *laud* välistatakse mõeldavate asjade hulgast kõik, mis on olemuselt *mittelaud* (vt Thampi 2005).

vastaks kõigile vajalikele tingimustele: VALLALINE, MEES ja TÄISKASVANUD. Samas on need tingimused ka piisavad: igauks, kes vastab neile tingimustele, kuulub automaatselt kategooriasse VANAPOISS (vt lähemalt Laurence, Margolis 1999). Mõistetest mõeldakse kui abstraktsetest asjadest, millel on mahulise kasti omadused: vaimus mõeldud ese saab olla kas selle sees või väljas, mitte aga olla kuidagi serva peal ripakil. Kategooria sees on liikmed võrdse staatusega.

Selline selgeid piire ja kindlaid liikmesustingimusi rõhutatv lähenemine mõistetele sobis hästi 1970-ndatel strukturalismist välja kasvanud objektivistlikule semantikakoolkonnale lingvistikas, mille raames peeti tähendusi entiteetideks, mida sai formaliseerida ja ka täpselt kirjeldada binaarse opositsioonis olevate tunnusepaaride abil (Karlsson 2002). Lisaks kontrastiprintsiibile oli oluline struktuuriprintsiip – tähendusi usuti koosnevat väiksematest üksustest, millest räägiti kui semantilistest komponentidest, tähenduse kirjeldust nimetati komponentanalüüsiks. Näiteks sõna *mees* tähendus oli määratletud kui +INIMENE, +TÄISKASVANUD, +MEESSOOST, sõna *naine* tähendus erineb ühe komponendi, –MEESSOOST, poolest; *tüdruk* aga on kirjeldatav kui +INIMENE, –TÄISKASVANUD, –MEESSOOST. Samu binaarseid tunnusekomplekte kasutavad mõisted moodustasid suuremaid üksusi e semantilisi välju, näiteks kuuluvad MEES, NAIN, POISS ja TÜDRUK välja ‘INIMSUGU’ (vt nt Leech 1990 [1974]: 89–122).

Klassikalisele mõistekäsitlusele toetub ka terminiarendus, mille raames näiteks Eesti terministandardis⁶ on mõiste määratletud kui teadmusüksus, mille moodustab *tunnuste ühene kombinatsioon*. Eri staatusega on *olulised tunnused* (nagu soo- ja liigitunnused), mis on mõiste mõistmiseks möödapääsmatud, ning *eristavad* tunnused, mis aitavad sugulasmõisteid üksteisest lahus pidada. Näiteks *tooli* mõiste võiks defineerida olulise tunnusega “mööbliese ühe inimese istumiseks” ja eristavaks tunnuseks võiks olla “leeni olemasolu”, mis eristab seda sugulasmõistetest *taburet* ja *järi*. Nende tunnustega on määratletud selle mõiste sisu e intensioon. Selle mõiste ekstensiooni e mahu moodustavad kõik esemed, mis sellesse kategooriasse kuuluvad, kusjuures kogu maailma kirjususest hoolimata ei ole ükski tool põhimõtteliselt teistest “toolim”.

Mõlemas tutvustatud lingvistilises rakenduses on tunda klassikalisele mõisteteooriale omast täpsust ja kategooriate jäikust. Mõiste rakendamise *tingimuste* asemel on lingvistilises tähendusekirjelduses vaid *semantilised komponendid* või *tunnused*. Mõistete psühholoogiline realistlikkus ei ole arutlusobjektiks.

3.2. Prototüübiteooria

Prototüübiteooria on osa üldisemast kognitiivsest pöördest inimtunnetuse mõistmisel, mis leidis aset eelmise sajandi viimasel veerandil, kui järjest enam hakati kilbile tõstma kategooriate kogemuslikkust ja psühholoogilist usutavust. 1970-ndatel korraldatud empiirilised katsed sõnatähenduste uurimiseks osutasid, et kategooriad, millega inimesed realselt otsustusi langetades opereerivad, on piiridelt hägused ja sisemiselt organiseeritud esindavate liikmete ehk prototüüpide ümber. Kategooria liikmelisus osutus astme küsimuseks, mida ei otsustata mingite kindlate tunnuste olemasolu või puudumise alusel, vaid sarnasuse põhjal tüüpilise esindajaga

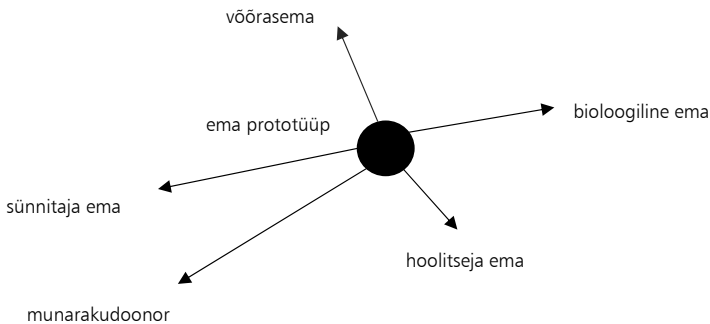
⁶ Autorid tänavad Arvi Tavastit võimaluse eest tutvuda terministandardi EVS-ISO 1087-1:2002 projektiga.

(Rosch 1978). Näiteks selle üle, mis mööbliesemega on tegemist meie istumise all, otsustame selle eseme sarnasuse põhjal tüüpilise tooliga. Ka prototüübiteoorias räägitakse siiski kategooriate *tunnustest*, kuid need ei ole vajalikud ja piisavad, vaid *tõenäolised* ja *olulised* (Rusell, Lemay 2000). Mida lähemal kategooria keskmeele e prototüübile paikneb kõnealune mõeldav ese, seda enam ühiseid tunnuseid tal sellega arvatavasti on. Ükski tunnus eraldivõetuna siiski kuidagi kriitiline ei ole. Näiteks võib ka istumiskõlbmatu lagunenu tool olla siiski kategooria *TOOL* liige, kuigi suhteliselt kauge tüüpilisest esindajast; samuti võib tooliks teatud tingimustel klassifitseerida istumiseks sobiva puunoti.

Kategooriatest hakati “kastisüsteemi” asemel mõtlema ruumiliselt, tsentri ja radiaalselt eemalduva perifeeria terminites. Võimalikuks osutusid kategooriatevahelised üleminekuvalad ning mõeldava eseme kuulumine korraga mitmesse kategooriasse. Mõistete prototüübiteooria on saanud väga laia kõlapinna lingvistikas (vt nt Erelt 1984), osalt kindlasti ka selle tõttu, et lingvistilised kategooriad ise alluvad prototüübi lähenemisele palju hõlpsamini kui “kastisüsteemile”.

Lingvistilise semantika jaoks tõi see kaasa, et sõna tähendus ei olnud enam hõlpsasti defineeritav ega kirjeldatav semantiliste tunnuste abil. Mingi väljendi tähenduse kirjeldamine tähendas esitada nii selle tüüpilisemad kui ka ebatüüpilisemad kasutused, mis kõik koos moodustavad nn kimpesituse (ingl *cluster*) või radiaalse kategooria (vt joonis 2).

Prototüübiteooria sobib hästi seletama kategooriate hajusust ja on psühholoogiliselt usutav. Selles osas, kas kategooriat esindav prototüüp on olemuselt pigem abstraktne ja ideaalne skeem või konkreetne eksemplar, puudub teoreetikutel üksmeel (vt Russel, Lemay 2000: 494).



Joonis 2. Mõiste *EMA* radiaalne esitus (Lakoff 1987 järgi)

3.3. Teooria teooria⁷ ning teadmiste sidusus

Mõistete seondumine omavahel ja üldisemate tunnetusprotsessidega on pannud aluse nn teooriakesksele lähenemisele (Murphy, Medin 1999). Väidetakse, et ainuüksi sarnasusest kas mingile ideaalsele prototüübile või antud klassi parimale esindajale ei piisa mõistete sidususe seletamiseks. Mõisted on hoopis vaadeldavad lülidena mõistmisprotsessis, mis tähendab vaikimisi teooriate loomist selle kohta, kuidas on nähtused omavahel seotud. Suhe üldisemasse teoriasse otsustab, mil-

⁷ Kasutame siin terminit *teooria teooria* S. Laurence'i ja E. Margolise (1999: 43) eeskujul pisut laiemas tähenduses, kui seda on sarnase nime all tuntuks saanud tendents luua teooriaid teiste inimeste meeleisundite kohta – ingl nn *theory theory of mind*.

lised omadused on olulised ja millised mitte. Näiteks KANA, PORGAND ja MANDEL pole kuidagi sarnased; omaduseks, mis seob neid kategooriaga TOIT, on nende söödavus, mis omakorda on osa naiivsest toitumisteooriast (näide Russell, Lemay 2000: 495). On põhjust arvata, et näiteks emotsioonimõistete, aga ka isiksuseomaduste jagunemine positiivseteks ja negatiivseteks tagab vastava rahvaliku teooria sidususe üldise inimest käsitleva naiivteooriaga, mille aluseks on hea ja kurja polaarne vastandus (vt Vainik 2002, Orav 2006).

Lingvistilises semantikas on mõistete sidusust tähtsustav lähenemine toonud kaasa tähenduste kirjeldamise suhete kaudu üldisemate teadmisstruktuuridega, mida on nimetatud küll *freimiks* (Fillmore 1985), *skriptiks* (nt Wierzbicka 1999), *kultuuriliseks mudeliks* (D'Andrade 1987), *kognitiivseks stsenaariumiks* (nt Lakoff, Kövecses 1987) või *idealiseeritud kognitiivseks mudeliks* (Lakoff 1987).

Freimi kui teadmiste struktuuri püütakse kirjeldada propositsioonidest koosneva kompaktsel semantilisel esituse abil, mida kutsutakse samuti *freimiks*. Selline esitus võib kirjeldada ka üksikmõistet, kuid enamasti kirjeldab mitut omavahel seotud mõistet. Näiteks JULGE ja ARG kui põhimõtteliselt sama situatsioonitüübiga seotud vastandlikke omadusi hõlmavad mõisted on Heili Oraval kirjeldatud ühe ja sama JULGUSE freimi abil (vt joonis 3), mille elementidest vaid mõne vastandlikkus eristab kahte mõistet (Orav 2006: 82–83). Propositsioonides figureerivad üldistatud tähendusega freimielemendid, mis defineeritakse iga freimiesituse komplekti tarbeks selle eesmärgi silmaspidavalt. Näiteks kirjeldab H. Orav eesti keele isiksuseomaduste sõnavara (ligi 1200 mõistet) ära, kasutades 15 defineeritud freimielementi.

<p>JULGUS hinnatavaks on Agent ja tüüpiliselt Agendi TEGEVUS</p> <p>Agent on JULGE, kui :</p> <ul style="list-style-type: none"> a) Agent hindab situatsiooni b) Agent tahab Eesmärki c) Agent teab Ohtu d) selle vaatamata Teeb Agent midagi Eesmärgi saavutamiseks e) Tagajärjeks on tüüpiliselt Agendi Eesmärgi saavutamine <p>Väärtus: Hindajapoolne hinnang pigem Positiivne.</p> <p>Agent on ARG, kui:</p> <ul style="list-style-type: none"> a) Agent hindab situatsiooni b) Agent teab ja kardab Ohtu c) Agent ei julge midagi teha <p>Väärtus: Hindajapoolne hinnang Negatiivne</p>

Joonis 3. JULGUSE freim (Orav 2006)

Skriptid on teadmiste struktuuri esitused, mis kirjeldavad sündmuste ajalise ja/või põhjusliku järgnevusega seotud mõistestruktuure. Anna Wierzbicka on leidnud selle vormi olevat hästi sobiva emotsioonimõistete sisu avamiseks (1999). Ka skriptid koosnevad propositsioonijärgenditest, mis omakorda koosnevad piiratud arvust nn semantiliste metakeele primitiividest (vt nt Wierzbicka 1996).

Õnn (X tundis õnne)

- (a) X tundis midagi (sest X mõtles midagi)
- (b) Mõnikord inimene mõtleb:
- (c) "minuga juhtus väga häid asju"
- (d) ma tahtsin, et sellised asjad juhtuks
- (e) midagi muud ma tahta ei oska"
- (f) kui see isik mõtleb, tunneb ta midagi väga head
- (g) X tundis midagi selletaolist (sest X mõtles midagi selletaolist)

Joonis 4. Mõiste ÕNN skript (Wierzbicka 1999)

Stsenaarium e kognitiivne mudel on mõnevõrra vabam propositsioonijärjendeist koosnev mõiste kirjeldus, mis metakeelena kasutab tavalise keele piiramatuid vahendeid. On näiteks väidetud, et teada, mida tähendab inglisekeelne sõna *anger* 'viha', on sama, mis teada mõiste VIHA prototüüpset stsenaariumit e kognitiivset mudelit (Fehr, Russell 1984).

- (1) Solvamine. Kahju põhjustaja solvab vastaspoole Mina. Solvav sündmus pahandab solvatu Mina. Kahju põhjustaja on süüdi. Rünaku intensiivsus kaalub üles karistuse intensiivsuse (mida selles faasis ei ole) ja tekitab seega tasakaalutuse. Rünnak põhjustab viha tekkimise.
- (2) Viha eksisteerimine. Sisaldab ka füsioloogilisi protsesse, mis on vastava emotsiooniga seotud; ekspressiivne komponent. Eksisteeriv viha avaldab Minale füsioloogilist mõju (kuumus, rõhumine, ärevus, erutus). Viha rakendab Mina suhtes jõudu, et üritada kätte maksta.
- (3) Püüe viha kontrollida. Mina avaldab vastumõju, et viha kontrollida.
- (4) Kontrolli kadumine. Viha intensiivsus ületab piirid. Viha valdab Mina, kes ilmutab vihast käitumist (otsustusvõime kadumine, agressiivne tegevus). Need nimetatud käitumisilmingud kahjustavad mina. Selles faasis on kahju põhjustanu ohus.
- (5) Karistus, kättemaks. Mina sooritab pahategija suhtes kättemaksuakti. Kättemaksu intensiivsus tasakaalustab rünaku intensiivsuse. Viha intensiivsus langeb nulli. Viha lakkab eksisteerimast.

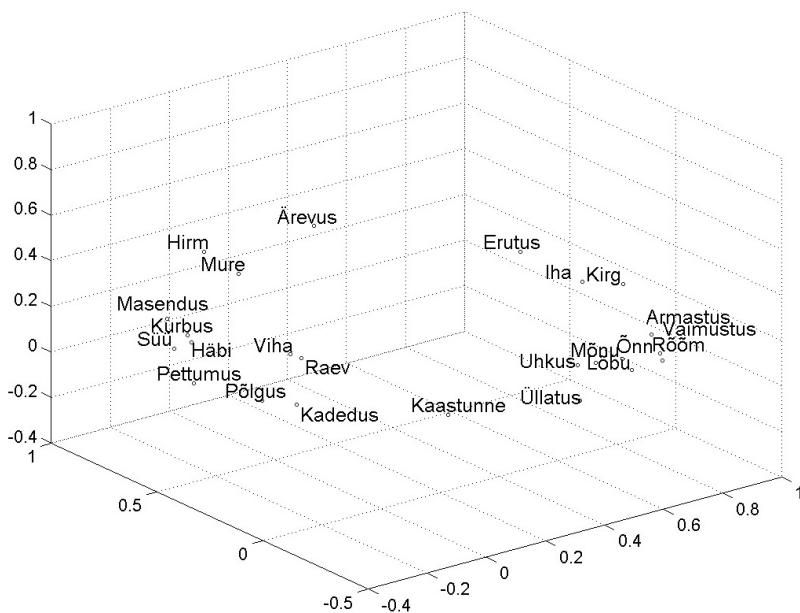
Joonis 5. VIHA stsenaarium ameerika inglise keeles (Lakoff, Kövecses 1987)

Termin *idealiseeritud kognitiivne mudel* sobib üldnimetuseks komplekssetele struktureeritud teadmistesüsteemidele. Need mudelid on idealiseeritud, kuna võtavad kokku seda teadmist, mis on kogemuseilene, samas aga ka selles mõttes, et võivad sisaldada kultuurilisi paleusi. Kui näiteks mõiste ABIELU idealiseeritud kognitiivne mudel oleks pelgalt statistiline, peaks see sisaldama konflikte ja vähemalt 50%-lise tõenäosusega ka lahutust. Tegelikuses näib aga ABIELU mõistesse kuuluvat soovmõtlemine "kui nad surnud ei ole, elavad nad õnnelikult tänase päevani".

Laiemaid teadmisstruktuure arvestavad lähenemised mõistetele üritavad seletada kategooriate radiaalset struktuuri ning mõistete prototüüpseusega seotud omadusi kui teadmiste üldise idealiseeritud mudelesituse kõrvalprodukte (Lakoff 1987: 68). Idealiseeritud kognitiivse mudeli rolli nähakse pidevalt toimivate mõis-testusprotsesside ja mõtte vahetu meediumi – nn mentaalsete ruumide – struktureerijana (Lakoff 1987: 281).

3.4. Ruumiline lähenemine

Mõte sõnade tähendust ruumiterminites kirjeldada ei olnud siiski uus ja läheb ajas isegi kaugemale tagasi kui prototüübiteooria. Charles Osgood, George Suci ja Percy Tannenbaum (1975 [1957]) käsitasid tähendust psühholoogilise fenomenina biheivioristlikus stiimuli ja reaktsiooni vaimus. Nende jaoks oli tähendus määratletud kui vahendusprotsess representatsioonisüsteemis, mis on ühteaegu märgi vastuvõtmise paratamatuks tagajärjeks ja märgile reageerimise vältimatuks eelduseks (Osgood jt 1975: 9). Tähendusprotsessi mõõtmiseks pakkusid nad välja semantilise diferentsiaali meetodi, mis tähendab sõnade/mõistete hindamist 7-astmelistel skaaladel, mille otstes on polaarseste tähendustega omadussõnad (nt *külm-soe, kaugel-lähedane*). Selle meetodi abil erinevaid mõisteid hinnates jõudsid nad järeldusele, et kõik mõeldavad semantilised eristused taanduvad kolmele statistilisele faktorile, millele nad andsid nimeks hinnang (ingl *evaluation*), jõud (*potency*) ja aktiivsus (*activity*).⁸ Mingi mõiste tähendus indiviidi jaoks on operantselt määratletud vastavate faktorilaadungite jaotumusest kolmel dimensioonil; mõiste tähendus kultuuris kujutab endast vastavate keskmiste faktorilaadungite jaotumust. Nad nimetavad neid faktoreid semantilise ruumi kolmeks mõõtmeks. Lisaks ühe konkreetse mõiste paigutuse selgitamisele lubab see meetod kujutada ka tervet mõistete hulka ja nende vastastikuseid (ruumilisi) suhteid (Osgood jt 1975: 87–89). Vaata näitena joonist 6, mis kujutab 24 eestikeelse emotsioonimõiste paigutust kolme statistilise faktoriga määratud semantilis ruumis. Mõisteid kirjeldavad andmed on kogutud 100 keelejuhilt semantiliselle diferentsiaalile väga lähedase meetodiga (vt Vainik 2004).

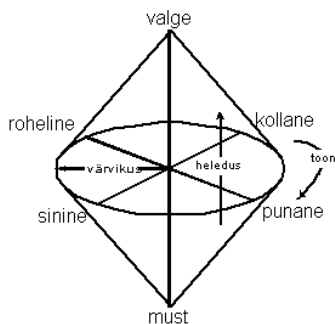


Joonis 6. 24 emotsioonimõistet kolme faktoriga määratud semantilis ruumis

⁸ Tähelepanuta on jäetud võimalus, et need kolm faktorit kirjeldavad pigem polaarseste adjektiivide mõõtmisvõimet kui nähtusi, mida nad "mõõdavad".

Kognitiivse lingvistika raames kõneldakse ühelt poolt keele semantilisest ruumist kui piiramatu ulatusega mõistepotentsiaalide väljast (Langacker 1987, eesti keeles vt Vainik 1992) ja teiselt poolt rõhutatakse selle organiseeritust valdkondadena. Viimastest osa on nn abstraktsed valdkonnad, mis koosnevad ise teistest hõlmavamatest mõistetest. Näiteks mõistel SÕRM on mõtet üksnes seoses hõlmavama mõistega KÄSI. SÕRM omakorda on valdkonnaks mõistetele KÜÜS ja SÕRMENUKK. Nn põhivaldkonnad (ingl *basic domains*) nagu AEG, RUUM, AINE, KVANTITEET, MUUTUS ON sellised, millele teised mõisted on taandatavad, kuid mis ise enam millekski ei taandu (Cruse 2000). Neist vähemalt AEG ja RUUM on iseloomustatavad otseselt ruumiliste dimensioonide abil. Enamus mõisteid on määratletud mitte ühe, vaid mitmest valdkonnast moodustuva valdkonnamaatriksi suhtes (vt nt Cruse 2000: 140, Langacker 1987, Evans, Green 2006: 230–247).

Kognitiivteadustes on mõistete ruumilise paigutuse ideed edasi arendanud Peter Gärdenfors (1997, 2000). Tema eesmärgiks on ühe geomeetrilise mudeli abil ära kirjeldada keele, meele ja maailma vahekorrad. Nn mõisteruumide teooria kohaselt on inimese käsutuses kolm põhimõtteliselt eristatavat representatsioonide tasandit. Välismaailmaga kõige rohkem on seotud nn eelmõisteline tase (ingl *subconceptual level*), kus mõisteid esindavad tajumodaalsustele vastavad kvaliteedidimensioonid. Kõige abstraktsem ja välise reaalsusega nõrgemalt seotud on sümboolsete representatsioonide tasand, kus mõisteid esindavad mingis keeles kasutatavad sümbolid, nt sõnad. Abstraktsuse astmelt vahepealne on nende kahe tasandi vahele jääv mõistelise representatsiooni tasand. Mõisted on kirjeldatavad alade või punktidenäruumis, mis on määratud eelmõistelise tasandi kvaliteedidimensioonidega (vt näiteks kolme dimensiooniga määratud värvimõistete paigutust jooniselt 7).



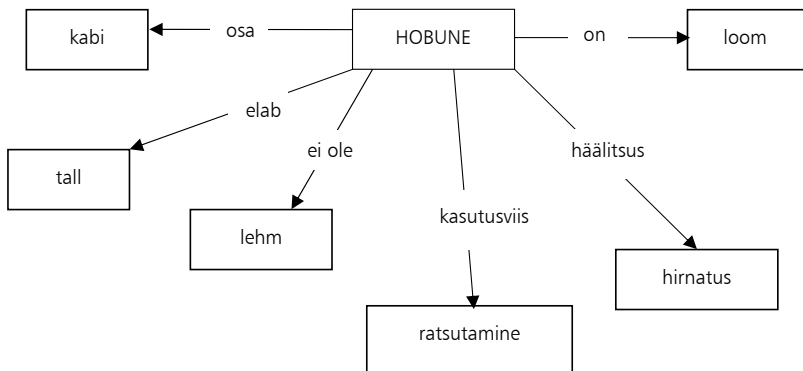
Joonis 7. Värvimõistete paiknemine mõisteruumis (Gärdenfors 2000)

Katseid mõistetega seotud tunnetusprotsesse ruumiterminites kirjeldada on tehtud teisigi. Näiteks mentaalsete ruumide teooria tegeleb põhiliselt tähenduse konstrueerimise uurimisega, sh eriti valdkondade vahelise ülekande (ingl *mapping*) probleemistikuga ja kognitiivsete segude (ingl *blend*) näol uute mentaalsete ruumide moodustamisega (Fauconnier 1994).

3.5. Võrgumudel

Kuna kognitiivses lingvistikas kuulub mõiste kirjeldusse alati selle taustsüsteem, olgu siis abstraktse valdkonna või dimensioonide näol, siis on selge, et keeleline semantika ei saa enam olla autonoomne ning et mingi tähenduse täpne analüüs tähendab tunnetuse kogu arengu arvessevõttu (Langacker 2002). Viimane on aga seotud sellega, kuidas usutakse inimteadmised organiseeritud olevat. Ruumilise paigutuse ideega konkureerib teadmiste võrgumudel, millel rajaneb nn entsüklopeedilise semantika käsitlus. Selle kohaselt on sõna sissepääsuks teadmiste võrgustikku, mis koosneb üksustest ja neid kaartena siduvatest suhetest. Näiteks sõna tähenduspoolusele kuuluvad antud indiviidil kõik selle sõlmega seostuvad teadmised, mälestused ja kogemused (Langacker 1987: 155).

Mõisteid kujutletakse seega sõlmedena komplekses multidimensionaalses võrgustikus, kus nad on määratletud eelkõige seoste kaudu teiste mõistetega. Näiteks sõna *hobune* tähendus on mõiste HOBUNE ja see sisaldab mitmeid ning eri sorti seoseid teiste mõistetega. A. Cruselt (2000: 128) leiame järgneva mõiste HOBUNE esituse (vt joonis 8). Seosed võrgustikus kuuluvad teatud tüüpidesse (osa, liik, kasutusviis jne) ja on erineva tugevusega. Leksikaalne üksus aktiveerib mõiste ja sellega seotud mõisted teadmiste võrgustikus vastavalt seoste tugevusele ja esilduvusele. Võib oletada, et HOBUNE puhul on kategoriseeriv suhe “on” mõistega LOOM esildavam kui näiteks negatiivne kategorisatsiooniosustus “ei ole” mõiste LEHM suhtes. Kategoriseerimise prototüübiloomus toimib ka siis, kui küsimus on semantiliste seoste tugevuses või esilduvuses.



Joonis 8. Mõiste HOBUNE võrgustik (Cruse 2000)

On loomulik, et selliselt mõistetud mõiste ei ole täiesti muutumatu, vaid arenev vastavalt sellele, kuidas indiviidi teadmiste võrgustikku lisatakse uusi seoseid või sõlmi. Teadmiste võrgumudelile annab jumet asjaolu, et kui mõistetel on mingi neurofüsioloogiline korrelaat, siis seda tuleb näha pigem neuronite koaktivatsioonimustrina närvivõrgustikus, kui mingi selgelt lokaliseeritava või piiritletava üksusena inimese ajus. Teatavate närvühenduste hõlpsus võrreldes teistega taandub nende kasutussagedusele ning sellega seotud neuronite aktivatsioonilävede madalusele (Pinker 1997).

3.6. Dünaamiline lähenemine

Mõistetest on võimalik mõelda ka mitte ainult kui suhteliselt stabiilsetest asjadest, vaid ka kui vaimsetest protsessidest, mis transformeerivad kogemuse toormaterjali käepärasteks üksusteks. Kuna välismaailma stiimulid on lõpmatult mitmekesised, siis tajus, mälus, mõtlemises, probleemide lahendamises ja teistes psühholoogilistes protsessides on ökonoomsem neid andmeid käsitleda teatavate tervikpakettidena, mis ühendavad omavahel piisavalt sarnanevaid stiimuleid. Mõistetest saab seega mõelda kui kognitiivse ökonoomia teenistuses olevatest protsessidest, kontseptuaaliseeringutest, mis on dünaamilised, hajusate piiridega ja individuaalsed (Russell, Lemay 2000: 492).

Sarnast mõtlemist võib tuvastada kognitiivses lingvistikas, kus deklareeriti tähenduse samastamist mõistestusega (ingl *conceptualisation*), ja hakati semantika esmaülesannet nägema mõtete ja mõistete struktuuri analüüsimises ning kirjeldamises (Langacker 1987, 2002, Evans, Green 2006). Mõistestamist käsitatakse seejuures üsna avaralt: see hõlmab nii uusi mõttemoodustisi (ingl *conceptions*) kui ka kinnistunud mõisteid; sensoorset, kinesteetilist ja emotiivset kogemust, vahetu sotsiaalse, füüsilise ja lingvistilise konteksti tuvastamist jne. Mõistestamine seisneb kognitiivses infotötluses; kirjelduse ülesandena nähakse seda mentaalset kogemust moodustavate kognitiivsete sündmuste iseloomustamist (Langacker 2002). Teisal täpsustatakse, et kontseptualisatsiooni saab käsitleda kahel tasandil: (i) fenomenoloogilisel tasandil kui vaimset kogemust; (ii) kognitiivsete sündmuste tasandil kui näiteks neuroloogilist aktiivsust. Vaimse kogemuse protsess tähendabki mingit neuronaalse aktiivsuse ilmnemist ja vaimset kogemust on loomulik selle kaudu seletada, või vähemalt püüda tabada neile omast kognitiivsete sündmuste arhitektuuri (Langacker 2002: 149).

Lähenemises, mida on nimetatud ka uus-empiritsismiks (Machery 2007), on seda ülesannet võetud tõsiselt ja kujundatud käsitus mõistetest kui situatsiooniliste kontseptuaaliseeringute kogumitest (Feldmann Barrett 2006). Mõisted ei ole mitte üksnes dünaamilised, vaid ka multimodaalsed, st erinevate tajukanalite kaudu seotud situatsioonidega, millest nad võrsuvad.

Näiteks mõiste *VIHA* ei ole mitte semantilises mälus paiknev klassikaliselt organiseeritud kogum infot, ega ka mitte prototüübi ümber organiseeritud tavaarusaam sellest emotsioonist. Mõiste *VIHA* on pigem representatsioonide või kontseptuaalsete teadmiste pakkide kogum, mis kõik üksteisest erinevad. Iga kontseptuaaliseering on situatsioonisidus selles mõttes, et vastab antud isiku vajadustele antud olukorras. Kõik need juhtumid, kui kogetud tuumafekti on *viha* sõnaga tähistatud, annavad oma panuse *VIHA* mõiste kujunemisse. Selliselt käsitatud mõiste on heterogeenne, varieerub ühel isikul vastavalt olukorra nõuetele, rääkimata varieerumisest isikute või kultuuride lõikes.

Mõisted kui situatsioonisidusad kontseptuaaliseeringud ei ole sõltumatud sensorimotoorsetest sündmustest ega talletatud mingit sorti propositsioonidena. Selle asemel käsitatakse neid tajusümbolitena (ingl *perceptual symbols*) ehk varem antud kategooria esinemisega kaasnenud sensorimotoorsete seisundite osaliste taaslavastustena, nn simulatsioonidena (Feldmann Barrett 2006, ja tema viited Barsalou jt 2003). On võimalik luua lõpmata palju erinevaid simulatsioone, mis sõltuvad hetke eesmärkidest ja situatsioonist, ning seetõttu on mõisteline esitus äärmiselt

kontekstipõhine ja dünaamiline (Barsalou 2003). Igas olukorras aktiveeritakse teadmistest vaid olukorrale vastav osa. Näiteks otsustab mingi konkreetse situatsiooni aspekt, kas relevantne on see viha mõisteesindus, mis sisaldab karjumist, ründamist, hammaste krigistamist või nutmist (Feldmann Barrett 2006: 33).

Sarnased ideed on jõudnud ka lingvistikasse. Näiteks on uuritud tegevusverbide, nagu inglise *grasp* 'haarama', neuroloogilist esindatust ajus ja leitud, et samad taju ja liikumisega seotud närvikeskused/neuronid aktiveeruvad nii reaalse liigutuse, selle visuaalse taju korral kui ka vastava mõiste meenutamisel. Osutus, et millegi sooritamine ning selle pelk kujutlemine kasutavad sama neuromaalist substraati. Need tulemused lubavad autoritel väita, et kontseptuaalne teadmine on kehapõhine (ingl *embodied*), rajanedes meie sensorimotoorsel süsteemil (Gallese, Lakoff 2005). Kuivõrd viimane on multimodaalne (hõlmates nägemise, kuulmise, puudutuse, liikumise jne), on seda ka keel, mis tänu inimese ajus leiduvatele peegelneuronitele on võimeline juba olemasoleva tajumodaalsuste süsteemi uut moodi kasutusele võtma.

Vittorio Gallese ja George Lakoff kõnelevad kehamõistetest (ingl *embodied concept*) kui skeemidest, mis sugenevad meile sotsiaalsest ja füüsilisest interaktsioonist maailmaga vastavalt meie keha ja aju olemuslikele parameetritele (Gallese, Lakoff 2005: 12). Ka nn põhitasandi kategooriate (Rosch 1978) esilduvuse ühe põhjusena toovad nad välja, et meil on sensorimotoorsed programmid, suhestumaks näiteks pigem *TOOLI* ja *LAUAGA* kui *MÖÖBLIGA* üldiselt. Piltlikult öeldes: seda mis on tool, teab kõige paremini tagumik.

Kokkuvõttes väidab lähenemine, mida siin dünaamiliseks nimetasime, et mõisted/tähendused ei ole ei stabiilsed ega invariantid, et nad on loomult multimodaalsed ja sugenevad inimese interaktsioonist keskkonnaga. Kognitiivne ökonoomia ei avaldu ainult kogemuse paljususe kokkuvõtmises väiksesse arvu mõistetes, vaid ka tajuesinduste kaasatuses mõistelisse representatsiooni.

3.7. Evolutsiooniline lähenemine

Kui juba on jõutud mõistete/tähenduste protsessiloomuse tõdemiseni, siis on loomulik ka küsida, milleks meile sellised protsessid? Mõistete eksistentsiks vajaliku, kuid samas ka väga energianõudliku aju väljakujunemine evolutsiooni käigus on pidanud andma selle omanikele eeliseid kas siis loodusliku või sugulise valiku survet silmas pidades.

Mõistete moodustamise (ingl *concept formation*) ja sümbolite põhinemise (ingl *symbol grounding*) uurimine on kaks suunda kognitiivteadustes, mis neile küsimustele püüavad vastata. Silmas peetakse ka seda, et kirjeldatavaid protsesse peaks olema võimalik arvutil modelleerida. Näiteks nähakse mõistete moodustamise analoogiat andmete klassifitseerimises ning sarnase käitumisega alamklasside leidmisele suunatud andmeanalüüsi ja -kaevandamise meetodites⁹ (Witten, Frank 2005).

Mõiste moodustamiseks nimetatakse protsessi, mille käigus isik õpib korraldama teatud kogemust üldisteks reegliteks või klassideks (Britannica). Tegevuses tõstab isik üles konkreetse eseme, mõtetes kasutab ta aga klasse, teades eseme omadusi üldiselt. Selline üldistatud teadmine võimaldab toime tulla uutes situat-

⁹ Näiteks leiab P. Gärdenfors (2000), et iseorganiseeruvate kaartide meetod (Kohonen 2000) võimaldab leida mõisteruumi varjatud dimensioonid ning tuua välja andmetes peituvat mõistete struktuuri. Formaalne mõistete analüüs (Ganter, Wille 1998) püüab leida mõistelist kuuluvust omaduste ja objektide vaheliste seoste kaudu.

sioonides. Sümbolite põhinemist seletatakse kolmetasandilise struktuuri kaudu, milles madalaim tase on mõiste moodustamiseks vajaliku sensoorse informatsiooni projitseerimine ikooniliseks esituseks, järgnev tase on teatud omadusi üldistav kategooriaalne esitus esinevatest mustritest ning kõrgemaks tasemeks on mustri sidumine sümbolitega (Harnad 1990).

Vastust küsimusele, miks sellised protsessid aset leiavad, nähakse alalhoiuis-tinktis.¹⁰ Tähenduse omistamist nähakse kogu elusloodusele omase fenomenina (Hoffmeyer 1996). Füüsilise keskkonna projektsioon mõisteliseks esituseks ei toimu üks-üheselt, vaid on piiratud organismi vajadustega edukaks toimetulekuks antud keskkonnas. Klassifitseerimise täpsus peab olema vastavuses olulisusega organismi eksistentsi seisukohast ning peab võimaldama maksimaalset teavet keskkonna kohta minimaalse kognitiivse kuluga (Rosch 1978).

Ainult klassifitseerimine ei kirjelda aga asjade kogu tähendust ning seepärast on mõned uurijad leidnud, et lisaks sündmust või objekti füüsiliselt kirjeldavatele dimensioonidele on mõistete moodustamisel oluline ka väärtuse (ingl *value*) dimensioon (Zlatev 2001). Asjaomases kirjanduses nimetatakse seda ka *tähenduseks* (ingl *meaning*) ning väidetakse, et teadmine objekti või sündmuse tähendusest (= väärtusest) sugeneb läbi tagasisideaehelate (Cisek 1999). Organismi heaolu soodustavad objektid või sündmused (nagu toit) hinnatakse headeks ja seda kahjustavad nähtused (nagu oht) halvadeks. Selline eristamine annab tähenduse nii tajule – kas tajutav objekt või sündmus võimaldab saavutada soovitud olekut või annab märku võimalikust ebasoovitavast olekust – kui ka tajude põhjal moodustatud mõistetele.¹¹ Viimaseid salvestatakse mälus kui hulka neuronite klikke, mis kodeerivad erinevaid tunnuseid/omadusi üldistest spetsiifilisteni ning on leitud, et sealhulgas talletatakse ka sündmusega kaasnenud emotsionaalset kogemust (Lin, Osan, Tsien 2006). Emotsionaalne info annabki mõistetele väärtuse.

Edukus olelusvõitluses ei sõltu ainult katse ja eksituse teel saadavast tagasisidest, oluline on hinnata ka seda, mida objektid või sündmused meil lubavad saavutada. Lapsed õppivad juurde uusi mõisteid, sõnu, objekte ja sündmusi mitte niivõrd nende omaduste kaudu, vaid vastavalt sellele, mida need võimaldavad neil peale hakata (Gibson 1979). Väidetakse ka, et kategooriad, mida me maailma tunnetamiseks kasutame, on eesmärgistatud, ning et kategooria kõige tüüpilisemad näited on need, mis suurendavad eesmärgi saavutamist, mitte ilmingimata need, mis on sagedased (Feldmann Barrett 2006).

Ei saaks öelda, et evolutsioonilise kasulikkuse argumendid oleks kuigi laialdaselt lingvistilisse semantikasse jõudnud. Katseid keelenähtusi olelusvõitlusega seletada võib siiski leida nt E. Vainiku (2002) ja H. Orava (2006) töödest, kes on uurinud emotsioonide ja isiksuseomaduste sõnavara semantikat ning leidnud sealt läbiva jaotuse heaks ja halvaks. Evolutsiooniline lähenemine sobib hästi mõningaid nähtusi seletama, kuid empiirilisi tõendeid selle toimimisest on raske saada, kasvõi juba sellepärast, et evolutsioon on oma olemuselt väga pikaajaline protsess. Lahenduseks võiks olla mõistemoodustuse modelleerimine arvutil. Näiteks on spetsiaalse programmi loomisega simuleeritud algselt teadvusetute agentide evolutsiooni “toitvate” ja “toitmatute” objektide keskkonnas, kus nood mitme põlvkonna möödudes omandasid teadvuse alged, st võime klassifitseerida keskkonnas olevaid objekte ja nähtusi ning luua nendest elementaarse mõistelise esituse (vt Kirt 2007).

¹⁰ Nn NASA määratluse järgi on elu ennast alalhoidev keemiline süsteem, mis järgib darvinistliku evolutsiooni põhimõtteid (Shapiro 2007). Elu aluseks peetakse autokatalüütilisi tsükleid (Kauffman 1993). Elusate süsteemide peamine omadus on tagada eksistentsiks vajalikud tingimused ning seetõttu hoida teatud parameetreid soovitud vahemikus.

¹¹ On tähelepanuväärne, et ka C. Osgood, G. Suci ja P. Tannenbaum (1975) jõudsid semantilise diferentsiaaliga mõisteid mõõtes tulemusele, et *hinnangu* dimensioon seletas kaks korda nii palju vastuste variatiivsust kui *jõud* või *aktiivsus*, viimaste laadung ületas kaks korda mistahes järgnevat faktorit.

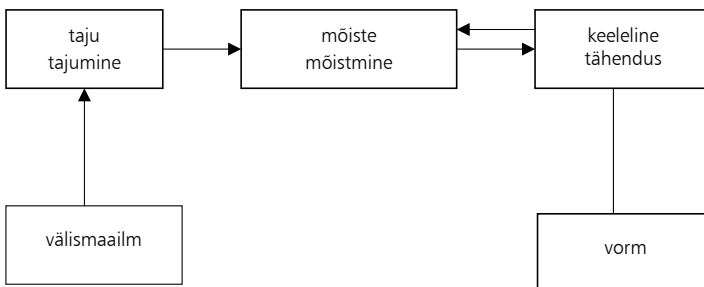
4. Kokkuvõte ja arutelu

Mõiste mõistmise ja seletusviiside mitmekülgsus, mida eelnevas tutvustasime, annab aimu nendest jõupingutustest, mida inimkond on sellest nähtusest aru saamiseks teinud. Meie eesmärgiks ei olnud välja selgitada ühte ja “õiget” mõistete mõistmise viisi. Tahtlikult hoidusime erinevatele lähenemistele hinnanguid andmast, keskendudes selle asemel mõisteteooriate vastasmõjudele lingvistilise semantikaga, mille jaoks on pea igas lähenemises olnud midagi kasulikku ja edasiviivat. Osutub, et tähenduste kirjeldamine, seletamine ja modelleerimine on erinevad eesmärgid, millest lähtuvalt on ka lingvistikas katsetatud erinevaid semantilisi esitusi.

Keelelise tähenduse uurija võiks teha enda jaoks teoreetilise valiku lähtuvalt eesmärgist. Kui eesmärgiks on tähenduste täpne kirjeldamine ja fikseerimine (nagu sõnaraamatutöös, terminiarenduses või arvulingvistilistes rakendustes), tuleks usutavasti valida klassikalisel mõisteteoorial põhinev komponentanalüüs või ka teooria teorial põhinev kirjeldusatribuutika piiratud arvu selgelt määratletud primitiivide abil. Tesauruse tüüpi rakendusele sobib tähenduste võrgumudel oma defineeritud seosetüüpidega. Kui eesmärgiks on tähendusloomega seotu psühholoogiline seletamine, sobib prototüübiteooria. Kui semantikast liigutakse edasi mentaalsete representatsioonide arhitektuuri kirjeldamisele, võiks kaaluda ruumilist või dünaamilist lähenemist. Evolutsioonilise vaatenurga rakendatavus lingvistilises semantikas alles selgub. Huvitatule annab see igatahes võimaluse keelest leitavaid fenomene siduda kõikjal eluslooduses iseenesest aset leidvate protsessidega.

Mõiste uurimist raskendab asjaolu, et ta on ise mõiste ja et enamus mõisteid kirjeldavatest teooriatest on paratamatult sunnitud nende olemust määratlema teistele mõistetele toetudes. Katsed kirjeldada mõisteid/tähendusi mingit sorti semantiliste komponentide, primitiivide, freimielementide, kvaliteedidimensioonide, tajumodaalsuste, sensorimootorsete esinduste või muu ontoloogiliselt “primitiivsema” kirjeldusatribuutika varal kujutavad endast kõik katseid sellest nõiaringist pääseda. Järjest nutikamate semantiliste kirjelduste väljamõtlemine on seotud ühelt poolt sooviga muuta silmale nähtamatud representatsioonid nähtavaks ja manipuleeritavaks ning teiselt poolt sooviga mõisteid ja mõistmisprotsessi tervikuna mõista ning modelleerida.

Mida me mõistest praegu rohkem mõistame kui artikli alguses? Kindlasti seda, et tegu on kompleksse nähtusega inимтunnetuses, mida saab käsitleda eri vaatenurkadest, tuues esile selle erinevaid olemuslikke tunnusjooni ja käitumist. Joonis 9 esitab kokkuvõtlikult lüüsisid, mis erinevates mõisteteooriates oluliseks on osutunud.



Joonis 9. Representatsioonitasandite koondesitus (Evans, Green 2006: 7)

Paremal näeme keelemärki kui tähenduse ja vormi vastavust, mis mõjutab mõistmist ja mõisteid. Vasakul näeme aga välismaailma, mis läbi tajuprotsesside meieni jõuab ning mõistete kujunemisele aluse paneb. Joonisel on välja toodud keelelise tähenduse ja mõistete omavaheline vastasmõju, viimaseid on käsitatud kui tajudest tulenevaid representatsioone, mis on seotud keelelisele sümbolile omase tähendusega (Evans, Green 2006). Jooniselt puudub aga mõistete (tagasi)mõju tajule ja muudele tunnetusprotsessidele, mille olemasolu mõningad uuringud on samuti näidanud (nt Halberstadt, Niedenthal 2001).

Mõisted on kahepalgelised, kuuludes ühest küljest tähendusena sõnade (st keele) juurde ning olles seega kollektiivsed ja jagatavad; teisest küljest võtavad mõisted kokku kogemust ja teadmisi, mida igauks on maailma nähtuste kohta individuaalselt oma teadmistesse talletanud. Neid kahte põhimõttelist mõistete tüüpi on nimetatud erinevalt, nt objektiivseks *vs.* vaimumõisteks (ingl *mental concept*) (Lyons 1977) või eksternaliseeritud *vs.* internaliseeritud mõisteteks (Jackendoff 1999). Kahe põhimõtteliselt erineva lähenemise võimalusest on tingitud vähemalt osa teoreetilisest möödarääkimisest eri alade kirjandustes. Keeleteaduses on tegeldud rohkem mõistete kollektiivsusest ja suhtluslikkusest tuleneva loomusega; psühholoogias ja kognitiivteadustes on keskendutud enam mõistmisprotsessidele indiviidis.

Ajalises perspektiivis on mõisteuurijate huvi liikunud üha enam mõistete määratlemise ja defineerimise probleemistikult nende psühholoogilise realismi ja neuropsühholoogiliste korrelaatide leidmisele. Suhteliselt pikalt on oldud üksmeelel, et mõisted on loomult amodaalsed, st sõltumatud mingi konkreetse tajukanali spetsiifikast (Machery 2006). Neid on peetud stabiilseks ideograafiliselt (st samal isikul on nad alati samased) ja nomoteetiliselt (st kõigil isikutele samas kultuuris on asjast enam-vähem samane ettekujutus) (Feldmann Barrett 2006: 32). Nagu näitasime dünaamiliste ja evolutsiooniliste mõisteteooriate tutvustuses, on tänapäevaks nende seisukohtade üksmeel kõikuma löönud.

Huvitav on veel tähele panna tutvustatud teoreetiliste mõistekäsitluste katuvusi artikli algul esitatud naiivsete arusaamadega. Esiteks on see olnud vajadus elunähtusi kokku võtta ja nendega manipuleerida (nt *käsitada* või *haarata*), mis on sundinud mõistetest mõtlema “klassikalisel” viisil kui mahulistest kategooriatest ja objektidest. Teisalt on põnev, et eestikeelse sõna *mõiste* etümoloogia viib tagasi maitsmis-, katsumis- ja haistmisaistingutele, olles sellega kooskõlas vana hea empiritsismi, aga ka vägagi tänapäevase mõistete multimodaalsuse teooriaga. Dünaamilise vaatenurgaga haakub ka eestikeelse sõna *mõiste* vormiline kuulumine *e*-deverbaalide klassi, mis laseb sellega tähistatavat tõlgendada mõistmisprotsessi lülina. Verbi *mõistma* hinnangut sisaldav tähendus (nt *õigeks vs. hukka*) osutab aga paralleelsusele evolutsioonilise lähenemisega ja selles välja toodud positiivse ja negatiivse tagasiside funktsiooniga. Mõistmise ja mõistete tähtsus pole seega mitte ainult klassifitseerivas teadmises nähtuste omadustest, vaid ka automaatses hinnangus, mida mõistja selles olukorras ette võiks võtta. Ehkki meie oma artiklis hinnangutest tahtlikult hoidusime, andsime meiegi vihjeid, mida ühe või teise mõiste mõistmise viisiga saab lingvistikas peale hakata.

Kui püüda lõpetuseks vastata pealkirjas esitatud küsimusele, siis tuleb välja, et mõisteid saabki mõista mitmel eri viisil. Olenemata vaatenurkade paljususest ollakse üksmeelel selles, et mõisted ja mõistmine on kesksel kohal meie arusaamises keele, maailma ja inimmeele vastastikustest seostest.

Kirjandus

- Aitchison, Jean 2003. *Words in the Mind: An Introduction to the Mental Lexicon*. 3rd ed. Blackwell Publishing.
- Barsalou, Lawrence W. 2003. Situated simulation in the human conceptual system. – *Language and Cognitive Process* 18 (5–6), 513–562.
- Barsalou, Lawrence W.; Simmons, W. Kyle; Barbey, Aron K.; Wilson, Christine D. 2003. Grounding conceptual knowledge in modality-specific systems. – *Trends in Cognitive Sciences* 7, 84–91.
- Britannica. Concept formation. *Encyclopædia Britannica*. <http://www.britannica.com/eb/article-70206> (18.06.2007).
- Cisek, Paul 1999. Beyond the computer metaphor: Behaviour as interaction. – *Journal of Consciousness Studies* 6 (11–12), 125–142.
- Cruse, Alan D. 2000. *Meaning in Language. An Introduction to Semantics and Pragmatics*. Oxford Textbooks in Linguistics. Oxford: Oxford University Press.
- D'Andrade, R. G. 1987. A folk model of the mind. – Dorothy Holland, Naomi Quinn (eds). *Cultural Models in Language and Thought*. Cambridge: Cambridge University Press, 112–148.
- EE = Eesti Entsüklopeedia. 1992. Tallinn: Valgus.
- EKG I = Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1995. *Eesti keele grammatika. I. Morfoloogia*. Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut.
- EKSS = Eesti kirjakeele seletussõnaraamat. I–VII. Tallinn: Eesti Keele Instituut, 1988–2007.
- Erelt, Mati 1984. Prototüübid ja keeleteadus. – *Keel ja Kirjandus* 2, 86–97.
- Evans, Vyvyan; Green, Melanie 2006. *Cognitive Linguistics. An Introduction*. London: Lawrence Erlbaum Associates Publishers.
- EVS-ISO 1087-1 = Terminoloogiatoõ. Sõnastik. Osa 1: Teooria ja rakendus. Tallinn: Eesti Standardikeskus, 2002.
- Fauconnier, Gilles 1994. *Mental Spaces. Aspects of Meaning Construction in Natural Language*. New York: Cambridge University Press.
- Fehr, Beverley; Russell, James. A. 1984. Concept of emotion viewed from a prototype perspective. – *Journal of Experimental Psychology* 113 (3), 464–486.
- Feldmann Barrett, Lisa 2006. Solving the emotion paradox: Categorization and the experience of emotion. – *Personality and Social Psychology Review* 10 (1), 20–46.
- Fellbaum, Christiane (ed.) 1999. *Wordnet. An Electronic Lexical Database. Language, Speech, and Communication*. London: The MIT Press.
- Fillmore, Charles 1985. Frames and the semantics of understanding. – *Quaderni di Semantica* 6.2, 222–254.
- Gallese, Vittorio; Lakoff, George 2005. The brain's concepts: The role of the sensory-motor system in conceptual knowledge. – *Cognitive Neuropsychology* 22, 455–479.
- Ganter, Bernhard; Wille, Rudolf 1998. *Formal Concept Analysis: Mathematical Foundations*. Berlin: Springer-Verlag.
- Gibson, James J. 1979. *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum.
- Gärdenfors, Peter 1997. Meanings as conceptual structures. – M. Carrier, P. Machamer (eds). *Mindscapes: Philosophy, Science, and the Mind. The Pittsburgh-Konstanz Series in the Philosophy and History of Science*. Pittsburgh: Pittsburgh University Press, 61–86.
- Gärdenfors, Peter 2000. *Conceptual Spaces. The Geometry of Thought*. London: The MIT Press.
- Halberstadt, Jamin P.; Niedenthal, Paula M. 2001. Effects of emotion concepts on perceptual memory for emotional expressions. – *Journal of Personality and Social Psychology* 81 (4), 587–598.

- Harnad, Stevan 1990. The symbol grounding problem. – *Physica D* 42, 335–346.
- Hoffmeyer, Jesper 1996. *Signs of Meaning in the Universe*. Translated by B. J. Haveland. *Advances in Semiotics*. Bloomington: Indiana University Press.
- Häkkinen, Kaisa 2004. *Nykysuomen etymologinen sanakirja*. Juva: WSOY.
- Jackendoff, Ray 1999. What is a concept that a person can grasp it? – E. Margolis, S. Laurence (eds). *Concepts: Core Readings*. Cambridge, MA: MIT Press, 306–333.
- Karlssoon, Fred 2002. *Üldkeeleteadus. Tõlkinud ja kohandanud R. Pajusalu, J. Valge, I. Tragel*. Tallinn: Eesti Keele Sihtasutus.
- Kauffman, Stuart 1993. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford: Oxford University Press.
- Kirt, Toomas 2007. An agent based simulation for testing the emergence of meaning. – *System and Information Science Notes* 2 (1), 70–73.
- Kohonen, Teuvo 2000. *Self-Organizing Maps*. 3rd ed. *Springer Series in Information Sciences*. Berlin: Springer-Verlag.
- Lakoff, George 1987. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: The University of Chicago Press.
- Lakoff, George; Kövecses, Zoltan 1987. The cognitive model of anger inherent in American English. – D. Holland, N. Quinn (eds). *Cultural Models in Language and Thought*. Cambridge: Cambridge University Press, 195–221.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar I. Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 2002. *Concept, Image, and Symbol. The Cognitive Basis of Grammar. Cognitive Linguistic Research*. Berlin: Mouton de Gruyter.
- Laurence, Stephen; Margolis, Eric. 1999. *Concepts and Cognitive Science*. – E. Margolis, S. Laurence (eds). *Concepts: Core Readings*. Cambridge, MA: MIT Press, 3–81.
- Leech, Geoffrey 1990 [1974]. *Semantics*. Harmondsworth: Penguin Books.
- Lin, Longnian; Osan, Remus; Tsien, Joe Z. 2006. Organizing principles of real-time memory encoding: Neural clique assemblies and universal neural codes. – *Trends in Neurosciences* 29 (1), 48–57.
- Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Macherey, Edouard 2007. Concept empiricism: A methodological critique. – *Cognition* 104, 19–46.
- Murphy, Gregory; Medin, Douglas 1999. The role of theories in conceptual coherence. – E. Margolis, S. Laurence (eds). *Concepts: Core Readings*. Cambridge, MA: MIT Press, 425–458.
- Mägiste, Julius 1982–1983. *Estnische etymologische Wörterbuch*. 1–12. Helsinki: Finnisch-Ugrische Gesellschaft.
- Ogden, Charles K., Richards, Ivor A. 1989 [1923]. *The Meaning of Meaning: A Study of the Influence of Language Upon Thought and of the Science of Symbolism*. London: A Harvest/HBJ Book; New York: Harcourt Brace.
- 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.
- Pinker, Steven 1997. *How the Mind Works*. London: Penguin Books.
- Pinker, Steven 2000. *The Language Instinct*. London: Penguin Books.
- Rosch, Eleanor 1978. Principles of categorization. – E. Rosch, B. B. Lloyds (eds). *Cognition and Categorization*. Hillsdale, NJ: Erlbaum, 27–48.
- Russell, James A.; Lemay, Ghyslaine 2000. Emotion concepts. – M. Lewis, J. Haviland-Jones (eds). *Handbook of Emotions*. 2nd ed. New York: The Guilford Press, 491–503.
- Saareste, Andrus 1958. *Eesti keele mõisteline sõnaraamat I*. Stockholm: Vaba Eesti.
- Shapiro, Robert 2007. A simpler origin for life. – *Scientific American* 296 (6), 25–31.

- Sweetser, Eve 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge Studies in Linguistics 54. Cambridge: Cambridge University Press.
- 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: Lund University Cognitive Studies, 153–162.
- Tavast, Arvi 2004. Mõiste mõiste: tähendusteooriate mõju oskuskeelepraktikale. – Rahvusvaheline terminoloogiakonverents “Eesti oskuskeel 2003” 10. ja 11. oktoobril 2003. Ettekanded. Eesti Terminologia Ühing, Eesti Keele Instituut, Emakeele Selts. Tallinn: Eesti Keele Sihtasutus, 7–18.
- Thampi, Mohan, G. B. 2005. Budistlik keeleteooria. – *Akadeemia* 6, 1161–1174.
- Vainik, Ene 1992. Genereerib inimene, mitte grammatika. – *Keel ja Kirjandus* 3, 156–160, 4, 216–223.
- Vainik, Ene 2002. Kas eestlased on “kuumaverelised”? Eestlaste rahvalikust emotsioonikategooriast. – M. Ereht, E. Ross, A. Õim (toim.). *Emakeele Seltsi aastaraamat 47 (2001)*. Tallinn: Emakeele Selts, 63–86.
- Vainik, Ene 2004. *Lexical Knowledge of Emotions: The Structure, Variability and Semantics of the Estonian Emotion Vocabulary*. *Dissertationes Linguisticae Universitatis Tartuensis* 5. Tartu: Tartu University Press.
- Wierzbicka, Anna 1996. *Semantics. Primes and Universals*. Oxford: Oxford University Press.
- Wierzbicka, Anna 1999. *Emotions Across Languages and Cultures: Diversity and Universals*. *Studies in Emotion and Social Interaction*. Cambridge: Cambridge University Press.
- Witten, Ian; Frank, Eibe 2005. *Data Mining: Practical Machine Learning Tools And Techniques*. 2nd ed. Morgan Kaufmann Series in Data Management Systems. San Francisco, CA: Morgan Kaufmann.

Ene Vainik (Eesti Keele Instituut) on lõpetanud Tartu Ülikooli eesti keele ja kirjanduse erialal. Täiendanud ennast praktilise psühholoogia alal Professionaalse Psühholoogia Erakoolis. Magistri- (2001) ja doktorikraadi (2004) on kaitsnud üldkeeleteaduses Tartu Ülikooli juures. 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 Kirt (Küberneetika Instituut) kaitses 1999. a Tallinna Tehnikaülikoolis tehnikamagistri kraadi ja 2007. a doktorikraadi. Töötab TTÜ Küberneetika Instituudi foneetika ja kõnetehnoloogia laboratooriumis erakorralise teadurina. Uurimisteemadeks iseorganiseerumine, andmeanalüüs ja (tehis)intellekt. Toomas.Kirt@mail.ee

HOW DO WE COMPREHEND THE CONCEPTS?

Ene Vainik, Toomas Kirt

Institute of the Estonian Language,
Institute of Cybernetics at Tallinn Technical University

Scholars differ over the meaning of *concept*. Having first belonged to the field of logic and philosophy, concepts are now of interest also for (neuro)linguists, psychologists, cognitive scientists and and specialists in informatics. The aim of the article is to provide an overview of different theoretical approaches to concept and to demonstrate the relations of those approaches to certain tendencies in linguistic semantics.

In the *classical approach* it is held important that concepts should be defined in terms of necessary and sufficient conditions. In linguistics the approach based on very precise differentia has become the core of semantic component analysis, while its application has been central, e.g., in the field of terminology. *Prototype theory*, largely drawing on psychological reliability, explains concepts as radial categories organized around the most representative members or category prototypes. For linguistics this approach implied that identification of the meanings of a word requires consideration of its typical as well as less typical uses alongside their different statuses. The *theory theory* aims at explaining the essence of concepts by means of certain more general knowledge structures (theories) that the concepts are part of. This approach has been used a lot for different semantic representations and description models of semantic content, such as *frames*, *scripts* or *scenarios*.

There have also been attempts to describe concepts in *spatial terms*, using dimensional representations. Cognitive linguistics, as well as cognitive sciences in general, are rather prone to construct some kind of semantic or conceptual spaces. Another type of concept explication is a *network model* of knowledge, which serves as basis for the so-called encyclopaedic approach to semantics. Here a concept is described by its relations with other concepts in the network of knowledge. The main linguistic applications are thesaurus dictionaries and electronic resources like WordNet. The *dynamic approach* views concepts as individual and variable situation-related processes going on in human brain, which have to do with processing and storing sensory data. It has been found that basically same parts of the neural substrate are involved in representing concepts (e. g. such as is represented by the verb *grasp*) as well as their underlying events (some act of *grasping*). Concepts are said to be simulations or reenactments of sensorimotor representations of events. That approach lays the basis for the so-called interactive semantics, arguing that human (incl. bodily) interaction with the external world is responsible for the development of the *embodied concepts* or *schemes*. *Evolutionary approach* sets out to reveal the mechanism of symbol grounding and concept formation, trying to explain it by ubiquitous loops of positive and negative feedback. It is vital for an organism, on the one hand, to tell hazardous situations from beneficial ones, and on the other hand, to assess what are the *affordances* of the situations.

Instead of aiming at the one and only theory of *what the concepts really are* the article concludes that different theoretical approaches focus on different aspects of concepts. Description, explication and modelling of linguistic concepts are three different objectives, each thus preferring a different semantic representation.

Keywords: linguistics, semantics, epistemology, cognitive science

EESTI-X-KEELE SÖNASTIK JA GRAMMATIKA

Ülle Viks

Ülevaade. Artikkel käsitleb probleeme, mis on seotud grammatilise info esitamisega kakskeelses sõnaraamatus. Lähemalt tutvustatakse Eesti-X-keele sõnastikku (EXS), tema kohta sõnastike tüpoloogias ja sellest tulenevaid nõudeid. EXS on universaalne leksikaal-grammatiline andmebaas, millest on võimalik genereerida suvalise kakskeelse sõnaraamatu elektrooniline põhi, kus eesti lähtekeel on standardkujul olemas. Grammatiline info esitatakse EXS-is struktureeritult, hõlmates erinevaid keelekirjelduse tasandeid fonoloogiast semantikani; keeleandmete valik ja esitusviis tuginevad sõnastiku ja grammatika vahekorra analüüsile. Vormistusel peab EXS kui metatasandi andmebaas olema paindlik ning sõltumatu tulemsõnastiku sihtkeelest, tüübist ja kujundusest. Lõpuks tutvustatakse lühidalt töövahendeid, mille abil EXS-i luua ja rakendada: sõnastike haldussüsteemi EELex (leksikograafi töökeskkond) ja grammatilist kirjegeneraatorit.*

Võtmesõnad: kakskeelne leksikograafia, grammatiline informatsioon, keeletöötlus, XML, sõnastike haldussüsteem, eesti keel

1. Eesti-X-keele sõnastik

“Eesti-X-keele sõnastik” on keeletehnoloogia projekti “Leksikograafi töökeskkond” allprojekt (2006–2010). Eesti Keele Instituudis (EKI) algatatud Eesti-X-keele sõnastik (EXS) on kavandatud kui universaalne leksikaal-grammatiline andmebaas, mis sobiks aluseks eelkõige uute eesti lähtekeeleka kakskeelsete sõnaraamatute koostamisel (inimese jaoks), aga ühtlasi peaks ta olema kasutatav mitmesugustes keeletehnoloogia rakendustes (arvuti jaoks).

EXS pole valmis sõnastik, vaid pigem metatasandi nähtus. See on paindlik polüfunktsionaalne andmebaas, millest on võimalik genereerida põhimõtteliselt ükskõik millise kakskeelse (eesti lähtekeeleka) sõnaraamatu elektrooniline põhi, kus eesti pool on standardkujul juba olemas. Selles on ära tehtud esialgne sõnavalik,

* Tööd on toetanud riikliku programmi “Eesti keele keeletehnoloogiline tugi” projekt “Leksikograafi töökeskkond” ja Eesti Teadusfondi grant nr 5969.

esitatud grammatilised andmed, näidatud sõna tähendusliigendus koos vajalike seletustega, lisatud sõna kasutusinfo, olulised sõnaühendid jne.

EXS-i olemuse ja koostamisaluste mõistmiseks püüan määratleda, milline on tema koht sõnastike väga kirevas maailmas.

1.1. Sõnastikutüpoloogiast

Sõnaraamat (laiemas tähenduses sõnastik) on teatavasti korrastatud sõnavarakogu, mis koosneb kindla struktuuriga sõnaartiklitest. Sõnaartiklit alustab märksõna, millele järgnevad teatud skeemi kohaselt mitmesugused andmed märksõna kohta, nt hääldus, sõnaliik, reksioon, stiilmärgend, seletus, sünonüümid, kasutusnäited, vasted teises keeles jne.

Sõnaartikli ülesehitus, andmete hulk ja esitusviis sõltuvad sõnastiku tüübist. Sõnaraamatuid võib liigitada paljude eri tunnuste järgi, kõige üldisemad neist on:

- **keelte hulk**, mis jagab sõnastikud ükskeelseteks ja kaks- või mitmekeelseteks;
- **sihtgrupp**, mille alusel eristuvad üldsõnastikud (laiale kasutajaskonnale) ja mitmesugused erisõnastikud (kitsale ringile) – igal kasutajagrupil on oma vajadused ja nõuded sõnaraamatule, millest sõltub eelkõige märksõnavalik, aga ka info valik, maht ja esitusviis sõnaartiklis;
- **otstarve**, mis liigitab sõnastikud funktsiooni järgi: nad kirjeldavad keelenähtusi, annavad õigekeelsuse soovitusi (normatiivne), abistavad keeleõppijat (õppesõnastik) või tõlkijat (tõlkesõnastik), toetavad teksti mõistmist (passiivne sõnastik) või moodustamist (aktiivne sõnastik) jne – ka otstarve mõjutab märksõnavalikut ning sõnaartikli koostist ja sisu;
- **maht**, mis seostub nii märksõnade arvuga kui ka artiklis esitatava info hulgaga – tavaliselt eristatakse kolme mahuklassi märksõnade arvu järgi: väike (10 000–20 000), keskmine (30 000–50 000), suur (60 000 ja rohkem) –, kuid mahu hindamisel tuleb arvesse ka sõnaartikli suurus.

Tunnused võivad kombineeruda, andes mitmeid eri- ja alaliike. Kakskeelsete (v mitmekeelsete) sõnastike ülekaalukas põhitüüp on **üldsõnastik**. See on universaalne, mõeldud eelkõige tõlkimiseks ühest keelest teise. Peamised liigitusalused on:

- sihtgrupp, nt tõlkija, keeleõppija, turist, spetsialist, teadlane, lingvist, keeletehnoloog jne;
- otstarve: kirjeldav e tõlkesõnastik, õppesõnastik, vestmik, formaalne leksikon jne; otstarbega seostub ka jaotus passiivseks v aktiivseks sõnastikuks (vt lähemalt 1.2.);
- maht: suurem osa tõlke- ja õppesõnaraamatuid on keskmise mahuga, suure näiteks võiks olla soome-eeesti suursõnaraamat (SUVI: 90 000 märksõna, 2 köidet), tüüpilised väikesed on nt turistsõnastikud.

Kakskeelseid **erisõnastikke** leidub suhteliselt vähe, erinevalt ükskeelsetest, mille seas on mitut liiki erisõnastikke.¹ Kakskeelsete erisõnastike liigitusalustest võiks välja tuua:

¹ Ükskeelsete erisõnastike liigitusalused on spetsiifilised, olulisemad neist on:

- märksõnatüüp, nt lühendid, võõrsõnad, (fraseoloogilised) väljendid, kohanimed jne;
- infotüüp e mille kohta sõnaartiklis teavet antakse, nt ortograafia, hääldus, morfoloogia, tuletus, valents/reksioon, sünonüümid/antonüümid, riimid, etümoloogia jne;
- tekstitüüp e allkeel, nt släng, erialane tekst (terminid), murded, vana kirjakeel, piibel, autoritekid jne;
- makrostruktuur e sõnastiku organiseerimise viis, nt pöörsõnastik, konkordants, mõisteline sõnastik, tuletuspe-rede sõnastik jne.

- märksõnatüüp, nt fraseoloogilised väljendid, käibefraasid;
- tekstitüüp, nt släng, erialane tekst (terminoloogiasõnastikud on sagedamini mitmekeelsed).

1.2. Aktiivne ja passiivne sõnaraamat

Kakskeelsete sõnastike üks olemuslikumaid liigitusi on jaotus passiivseks ja aktiivseks, mis seostub kasutaja teadvuses toimuvate erisuunaliste keeletöötlusprotsessidega: analüüsiga ja sünteesiga.

Passiivse sõnastiku eesmärgiks on toetada kasutajat võõrkeelse teksti mõistmisel e analüüsil. Selleks et tundmatu sõna sõnastikust üles leida (ja selle tähendust vaadata), peab kasutaja võõrkeelses tekstis tegema kõigepealt morfoloogilise analüüsi, st ta peab suutma kindlaks teha, milline on lauses oleva muutevormi algvorm, sest sõnastikes on märksõnad algvormis. Passiivse sõnastiku puhul on sõnastiku lähtekeel (märksõnade keel L1) kasutaja jaoks võõrkeel ning sõnastiku sihtkeel (vastete keel L2) on kasutaja emakeel.

Aktiivse sõnastiku eesmärgiks on toetada kasutajat võõrkeelse teksti loomisel e sünteesil. Selleks, et oma lauset grammatiliselt õigesti kokku panna, peab kasutaja suutma moodustada (sünteesida) sõnastikus esitatud algvormist kõik lauses vajalikud muutevormid. Aktiivse sõnastiku puhul on sõnastiku lähtekeel (märksõnade keel L1) kasutaja emakeel ning sõnastiku sihtkeel (vastete keel L2) on võõrkeel.

Neist eesmärkidest tulenevad erinevused sõnaraamatute ülesehituses e makro- ja mikrostruktuuris, eelkõige puudutab see andmete valikut ja esitusviisi sõnaartiklis. Makro- ja mikrostruktuuri erinevusi aktiivses ja passiivses sõnastikus on käsitlenud nt Soome keeleteadlane Marja Järventausta (2004) seoses õppesõnastikega (soome ja saksa keele näitel). Võtsin tema poolt esile toodud põhipunktid kokku tabelisse 1.

Tabel 1. Sõnastikuinfo passiivses ja aktiivses sõnastikus

	passiivne sõnastik (analüüs)	aktiivne sõnastik (süntees)
märksõnavalik	palju sõnu, sh haruldased	keskne sõnavara
andmete valik	märksõna tähendus	märksõna kasutusinfo
metakeel	sihtkeel (L2)	lähtekeel (L1)
grammatiline info	märksõna juures	vaste juures
vormid märksõnana	jah: ebareeglipärased	ei

Järgnevad kommentaarid on tabeli 1 vaba tõlgendus.

1. Sõnavalik passiivses sõnastikus peab olema suur, et see kataks võimalikult suure osa loetavates tekstides leiduvast sõnavarast. Aktiivne sõnastik seevastu võib piirduda ainult olulise sõnavaraga, mis võimaldab kasutajal ennast väljendada. Mida paremini kasutaja keelt valdab, seda suuremat sõnavara (ja sõnaraamatut) ta oma tekstide loomiseks vajab.
2. Andmete valik passiivse sõnastiku sõnaartiklis peab arvestama eelkõige märksõna tähendusi, et kasutaja saaks sõnastikust teada, mida tekstis esinev sõna tähendab. Aktiivne sõnastik peaks keskenduma märksõna kasutusele

- erinevates kontekstides (nii grammatilistes kui ka semantilistes), et kasutaja suudaks valida sõnastikus pakutud variantidest tema lausesse sobivaima.
3. Metakeel (seletuste ja märgendite keel) peaks olema kasutaja emakeel, et ta seletustest ja kommentaaridest saaks paremini kätte kogu sisu. Passiivses sõnastikus on emakeel sihtkeele (L2) rollis, aktiivses aga lähtekeele (L1) rollis.
 4. Grammatiline info peaks passiivses sõnastikus olema esitatud märksõna juures, et võõras keeles oleva sõna vormidest ja grammatilisest käitumisest tekiks ülevaade. Aktiivses sõnastikus vajavad grammatilist infot aga vasted, et oleks võimalik moodustada vajalikud vormid ja kasutada neid lauses vastavalt sihtkeele (L2) reeglitele.
 5. Sõna (ebareeglipärased) muutevormid märksõna positsioonis (koos viitega algvormile) on passiivses sõnastikus vajalikud selleks, et kasutaja saaks võimaluse siduda kõik tekstis esinevad ühe sõna muutevormid vastava märksõnaga, olgu nad omavahel nii erinevad kui tahes. Eesti keeles, mida iseloomustab suur tüvemuutuste hulk (vrd *sugu : soo, uba : oa, saag : sae, raba : rappa, siduma : seob, ütleva : öelda, minema : läinud* jne), on see väga oluline. Aktiivses sõnastikus pole vormide esitamine märksõnana vajalik, sest märksõna on kasutaja emakeeles.

M. Järventausta (2004: 27) sõnul on uurijad üsna üksmeelsed selles, et ideaaljuhul peaks iga keelepaari jaoks olema neli sõnastikku: kummagi suuna jaoks nii aktiivne kui ka passiivne, sõltuvalt kasutaja emakeelest.² Praktikas on sellist luksust muidugi raske leida, enamasti üritatakse kummagi suuna sõnastikku kokku koondada nii passiivse kui aktiivse sõnastiku omadused. Sedasama üritame ka meie Eesti-X-keele sõnastikus, aga me püüame seda teha teadlikult ja kindla eesmärgiga.

1.3. Tüpoloogilised nõuded EXS-ile

Sõnastikutüpoloogia seisukohalt on EXS-i andmebaas **kakskeelne üldsõnastik**, st universaalne sõnastik, mis on mõeldud eelkõige tõlkimise abivahendiks. Et aga EXS on metatasandi andmebaas, siis peab ta mingil moel hõlmama ka nende sõnastikutüüpide omadusi, mille aluseks ta saab. Selleks peab ta vastama järgmistele nõuetele.

1. EXS-i andmebaasi **maht** on maksimaalne, nii märksõnade arvu kui ka info hulga osas, kuid sellest peab saama genereerida väiksemaid tulemsõnastikke.
2. EXS on otseselt suunatud kahele **sihtgrupile**: a) leksikograafidele erinevate kakskeelsete sõnaraamatute koostamiseks, b) keeletehnoloogidele mitmesuguste sõnastikurakenduste loomiseks (sh masintõlge). Kaudselt – leksikograafide vahendusel – võivad EXS-i sihtgruppideks saada kõik tavapärased kakskeelse sõnaraamatu kasutajad (e lõppkasutajad), nt tõlkijad, keeleõppijad, turistid jt.
3. EXS-i **otstarbe** tunnussõnad on: kirjeldav, passiivne (ja aktiivne), formaalne.

Keelekirjeldus EXS-is peab olema põhjalik ja hästi struktureeritud, et sealt saaks teha erinevaid väljavõtteid, arvestades konkreetse tulemsõnastiku vajadusi. EXS on eelkõige passiivne sõnastik, mis siiski ei välista selle hilisemat kasutamist aktiivse sõnastikuna. Vastavalt otseste sihtgruppide vajadustele pööratakse EXS-is formalismidele suuremat tähelepanu, kui see tavasõnastikes kombeks on. Sõnastiku formaalsus ei tähenda sisu kaotamist, vaid sisu esitamist kujul, mis on kättesaadav eksplitsiitselt, tuginemata kasutaja oletatavatele taustateadmistele.

2. Grammatilise info esitus EXS-is

2.1. Keelekirjeldustasandid

EXS-i universaalset iseloomu ja võõrkeelset kasutajat arvestades on seal kavas esitada grammatilist infot võimalikult paljudelt keelekirjeldustasanditelt: 1) morfofonoloogiast välde ja ebareeglipärane rõhk, 2) morfoloogiast põhjalik vormimoodustuse kirjeldus ja jõudumööda andmed ka vormikasutuse kohta, 3) tuletusest kindlasti liitsõna-liigendus ja komparatsioon, võib-olla ka tuletustüübid, 4) (morfo)süntaksist sõnaliik ja arv, rektsioon ja püsiühendid, 5) semantikast võimaluse korral semantilised rollid, edaspidi ehk ka semantilised tüübid.

Kõige põhjalikumalt on välja töötatud morfoloogia esitussüsteem, mille aluseks on avatud morfoloogiamudel (Viks 2000a) ja sellele vastav morfoloogiline klassifikatsioon (VVS ja Viks 2001). See erineb mõnevõrra ÕS2006 tüübistikust (vt 2.3.), kuid sobib kakskeelse sõnastiku jaoks paremini, sest tema aluseks on formaalne mudel. VVS-i tüübistik on testitud ja leidnud praktilise rakenduse eesti keele morfoloogiataravaras (nt EKI automaatne morfoloogiasüsteem³, Filosoofi morfoloogiline analüsaator (Kaalep 1996) ja sellel tuginev eesti keele speller) ning mitmes kakskeelses sõnaraamatus, nt EVS, NEEN, Viks 2003 jt. Morfoloogiasüsteemiga koos on lahendatud ka morfofonoloogia probleemid, sest need on omavahel vastastikuselt sõltuvuses. Süntaktilise info kohta on ilmunud ülevaade (Langemets jt 2005). Palju konkreetseid eeltöid on tehtud tuletusinfo osas, mis suures osas toetub töös olevale sõnapere materjalidele⁴ (Vare 2002).

Grammatilise info esitamise üksikasjadeni ma siin ei lähe, vaid peatun sõnastiku ja grammatika vahekorral, millest otseselt sõltub sõnastiku keeleline funktsionaalsus.

2.2. Sõnastiku ja grammatika vahekord

Milleks üldse ja kui palju vajab sõnaraamat grammatilist infot, kui selle jaoks on olemas spetsiaalsed grammatikaraamatud? Inimene haarab sõnastiku järele tavaliselt siis, kui ta tahab teada sõna tähendust, õigekirja või vastet teises keeles. Kuid sellest ei piisa aktiivse keelekasutuse puhul – vaja on ka teada, kuidas konkreetset sõna ühel või teisel puhul pruukida: kuidas teda hääldada, kuidas käänata-pöörata, millisesse positsiooni teda lauses saab paigutada, milliste teiste sõnadega või muutevormidega sobib teda kokku panna jne, jne. Seda tarkust on raske grammatikast üles leida, kuidagi ta on seal olemas.

³ <http://www.eki.ee/tarkvara/> (28.09.2007).

⁴ Kasutatav EKI sisevõrgus: <http://www.eki.ee/dict/sp/> (28.09.2007).

Sõnastik ja grammatika sisaldavad suures osas sama grammatilist infot, kuid nad on organiseeritud erinevalt ja ideaalis peaksid teineteist vastastikku täiendama. Grammatika esitab keelenähtusi üldistatult, tema struktuuri põhiüksus on grammatiline seaduspärasus (keelereeglite kujul), mida illustreeritakse üksikute näidetega (hrl sõnadega), milles see või teine grammatiline omadus avaldub. Ükski grammatika ei saa loetleda kõiki iga reegli alla kuuluvaid sõnu. Sõnastik seob needsamad keelenähtused konkreetsete sõnadega: sõnastiku struktuuri põhiüksus on sõna, mille juures sõnaartiklis esitatakse temale iseloomulikud grammatilised omadused.

Kuid ka sõnastikus ei saa grammatiline info olla ammendav, nt puudub seal ruum näitamaks sõna kõiki muutevorme või sõna võimalust esineda lauses teatud lauseliikmena jne. Seetõttu esitatakse grammatiline sõnastikuinfo tavaliselt üldistatud kujul, nt klassitähisena või reegliviidana, mis juhib kasutaja reeglite endi juurde. Üldisema kategooria näide võiks olla sõnaliik: sõnad kirjes varustatakse sõnaliigimärgenditega (nt *ehtne Adj*), millele vastavad reeglid (põhimõtted sõnade üldise morfoloogilise ja süntaktilise ja semantilise käitumise kohta) tuleb leida grammatikaraamatust. Morfoloogiliste reegliviitadena kasutatakse eesti leksikograafilises traditsioonis sõna juures tüübinumbreid (nt *krae S 26*), need juhivad sõnaraamatu lisaosas esitatud tüübikirjelduste (e reeglite) juurde, mille alusel saab vajalikud vormid moodustada (vt Viks 2001). Samal kombel saab esitada sõnastikus ka regulaarse sõnamoodustuse reeglid, millele viitavad sõnaartiklis tuletustüüpide tähised.

Kui mõni sõna käitub teisiti, kui üldine grammatikareegel ette näeb, siis sellekohane info täiendab sõnaartiklis esitatud klassitähist. Nt sõna *eht* puhul tuleb adjektiivimärgendile lisada märkus, et see sõna ei käändu (erinevalt tüüpilistest omadussõnadest). Erandliku morfoloogiaga sõna kirjes tuuakse lisaks tüübinumbri-
rile ära konkreetsed erandlikud muutevormid (nt *pea S 26; pl p päid, adt pähe*).

Grammatilise info valik ja esitusviis sõnaartiklis sõltuvad suurel määral sõnastiku tüübist: sellest, milline on sõnastiku sihtgrupp ja otstarve, kas sõnastik on suur või väike, üks- või kakskeelne. Ükskeelne sõnastik lähtub emakeelse kasutaja vajadustest, kes tavaliselt otsib sõnastikust lisa oma olemasolevatele teadmistele. Muidugi võivad ükskeelset sõnastikku kasutada ka muukeelsed inimesed, aga nende jaoks on ta mingis mõttes liiane (info on kohati liiga põhjalik ja detailne), aga mingis mõttes ka puudulik, sest kõike pole otse välja kirjutatud. Ükskeelsesse sõnastikku on vaikimisi sisse kodeeritud eeldus, et kasutaja valdab keelt ja seega ei pea seal kõike eksplitsiitselt näitama – seda nii sisulise poole esitamisel kui ka sõnaartikli vormistamisel (nt vormide või näidete lühendamise puhul). Seevastu kakskeelse sõnastiku adressaadilt ei saa head teise keele oskust nõuda (selle saavutamiseks aitab sõnastik alles kaasa). Selles suhtes on võõrkeelne kasutaja sarnane arvutiga – mõlema jaoks kehtib eksplitsiitsuse ja formaalsuse nõue. Oluline on, et sõnastiku kasutaja saaks talle vajaliku info ja selle rakendamise juhtnöörid kätte ammendavalt ja ühemõttelise töökindlusega.

2.3. Ükskeelne kõrvalepõige

Toon kõrvalepõikena ühe näite ükskeelsest ÕS2006-st, kus morfoloogiline info on esitatud küll põhjalikult, kuid üksnes sellist kasutajat silmas pidades, kes eesti keelt valdab ja otsib sõnastikust tuge peamiselt kahtluse korral. Võõrkeelse kasutaja jaoks pole vormide moodustusalgoritm piisavalt selge ning on mõnikord lausa vasturääkiv.

Näiteks noomeni üks keerulisemaid vorme – pluurali partitiiv – on sõnaartiklis harva esitatud ja kasutaja peab selle ise tüübikirjelduse põhjal moodustama. Tüübid 20, 21 ja 22 erinevadki omavahel ainult mitmusevokaali poolest. Tüübikirjeldused sõnaraamatu lisas (ÕS2006: 20) on sellised (siin lühendatult):

20. mitmuse osastavas *-e*
paks, paksu, .paksu, .paksu *ja* paksusse, .paksude, .pakse *ja* .paksusid, .paksudesse *ja* paksesse
21. mitmuse osastavas *-i*
külm, külma, .külma, .külma *ja* külmasse, .külmade, .külmi *ja* .külmasid, .külmadesse *ja* külmisses
täht, tähe, .tähte, .tähte *ja* tähesse, .tähtede, .tähti *ja* .tähtesid, .tähtedesse *ja* tähisses
22. mitmuse osastavas *-u*
õrn, õrna, .õrna, .õrna *ja* õrnasse, .õrnade, .õrnu *ja* .õrnasid, .õrnadesse *ja* õrnusse
leib, leiva, .leiba, .leiba *ja* leivasse, .leibade, .leibu *ja* .leibasid, .leibadesse *ja* leivusse

Neisse tüüpidesse kuuluvate sõnade kirjetes pluurali partitiivi vormi pole, õige vormi saamiseks tuleb uurida tüübikirjeldust ja toimida analoogia põhjal tüüpsõna eeskujul. See on kasutajale ebamugav, aga siiski võimalik. Nt

haug <20: haugi, .haugi>	→ .hauge <i>ja</i> .haugisid	(<i>nagu</i> paks)
halg <20: halu, .halgu>	→ .halge <i>ja</i> .halgusid	(<i>nagu</i> paks)
õng <21: õnge, .õnge>	→ .õngi <i>ja</i> .õngesid	(<i>nagu</i> täht)
pung <21: punga, .punga>	→ .pungi <i>ja</i> .pungasid	(<i>nagu</i> külm)
lõug <22: lõua, .lõuga>	→ .lõugu <i>ja</i> .lõugasid	(<i>nagu</i> leib)

ÕS2006-s on aga ka tüüpe, mis sisaldavad erineva mitmusevokaaliga sõnu, kusjuures tüübikirjeldus näitab ainult üht vokaali, nt 24 (ainus tüüpsõna *pood*). Tüübikirjeldus:

24. *pood'*, *poe*, .*poodi*, .*poodi ja* .*poesse*, .*poodide*, .*pood'e ja* .*poodisid*, .*poodidesse*

Sel juhul on pluurali partitiiv kõigi sõnade kirjetes antud ja seda EI TOHI moodustada tüübikirjelduse järgi.

laug 1. <24: lau, .laugu; .lauge <i>ja</i> .laugusid>	(<i>vrđ</i> halg <20>)
saag <24: sae, .saagi; .saage <i>ja</i> .saagisid>	(<i>vrđ</i> haug <20>)
roog <24: roa, .rooga; .roogi <i>ja</i> .roogasid>	(<i>vrđ</i> pung <21>)
lõög <24: lõa, .lõõga; .lõõgu <i>ja</i> .lõõgasid>	(<i>vrđ</i> lõug <22>)

Eelnevast järeldeb (kirjutamata) algoritm kasutaja jaoks:

- 1) Kui pluraali partitiiv on sõnastikukirjes olemas, võta sealt.
- 2) Kui kirjes pluraali partitiiv puudub, moodusta see tüübikirjelduses oleva tüüpsõna eeskujul.

Segadus tekib neis tüüpides, kus tüübikirjelduses on pluraali partitiivi vorm antud, aga kõigist selle tüübinumbriga sõnadest teda ei moodustata või ei tarvitata (kusujuures eri sõnadel on partitiiv ka erineva vokaaliga), nt tüüp 16: tüüpsõnad *rida* ja *sõda*. Tüübikirjeldus:

16. rida, rea, rida, .ritta ja .reasse, ridade, ridu ja ridasid, ridadesse
sõda, sõja, sõda, .sõtta ja sõjasse, sõdade, sõdu ja sõdasid, sõdadesse

Algoritmi järgides juhtub nii, et sõna *tuba* saab oma pluraali partitiivi vormi *tube* kirjest kätte:

tuba <16: toa, tuba, .tuppa ja .toasse; tube ja tubasid>

Aga sõnade *koda*, *õde*, *tegu* puhul satub kasutaja eksiteele, sest sõnaartiklis pluraali partitiivi vormi pole ja järelikult tuleks see teha tüübikirjelduse järgi, aga tüübikirjeldus näitab ainult *u*-lõpulist vormi koos *sid*-lõpulisega. Nii võib keeleoskamatu kasutaja moodustada vigased muutevormid:

- | | | |
|--|--------------------|------------|
| koda <16: koja, koda, .kotta ja kojasse> | → *kodu ja kodasid | (vrd sõda) |
| õde <16: õe, õde> | → *õdu ja õdesid | |
| tegu <16: teo, tegu, .teosse ja .tekku> | → *tegu ja tegusid | |

Kui ükskeelne sõnastik ehk võib endale lubada mõningaid lünki keelekirjelduses, lootes kasutaja keeleoskusele, siis kakskeelses sõnastikus peab grammatika esitus olema täpsem ja algoritmilisem. Seetõttu pole EXS-i grammatika esitamisel võetud eeskujuks kirjakeele normi alust ÕS2006.

3. EXS kui universaalne andmebaas

EXS-ile seatud põhinõudest (alus erinevatele kakskeelsetele sõnaraamatutele) lähtudes postuleeritakse kolm sõltumatusenõuet: EXS peab olema sõltumatu tulemsõnastiku sihtkeelest, sõnastiku tüübist (mahust, sihtgrupist, otstarbest) ja kujundusest.

3.1. Sõltumatus sihtkeelest

EXS-i põhinõue on sõltumatus koostatava kakskeelse sõnaraamatu sihtkeelest – ta peab sobima aluseks paljude eesti-X keelepaaride sõnaraamatutele. See tähendab, et ta peab olema eelkõige eesti keele andmebaas, mis kirjeldab eesti sõnade leksikaal-grammatilisi omadusi sel kombel, et nendega saaks kõrvuti panna mistahes teise keele sõnade leksikaal-grammatilised omadused.

Eelkõige puudutab see sõnade tähenduse avamist, ja eriti **tähendusliigen-****dust**, mis peaks olema hästi detailne ja lähtuma eesti keelest. Kui eesti sõna mitu

eri tähendust saavad sihtkeeles sama vaste, siis need võetakse kokku alles tulemsõnastiku (X-keele sõnaraamatu) koostamisel, sest mõne teise keele vasted võivad eesti sõna eri tähenduste vahel teisiti jaotuda.

Teine oluline valdkond on **grammatiline info**, mis peab olema esitatud formaalselt ja algoritmiliselt, st arusaadaval kujul inimesele, kes eesti keelt piisavalt ei oska ja kelle emakeel ei paku tuge eesti grammatilise süsteemi mõistmisel. Universaalsete keelenähtuste esitamisel on mõttekas järgida üldist leksikograafist traditsiooni (nt sõnaliigimärgendid sõnastikukirjes märksõna juures ja tähendusnumbrite juures). Suuremat tähelepanu nõuavad grammatilised nähtused, mis on eesti keelele spetsiifilised, nagu välted, morfoloogia, rektsioon jne. Keelespetsiifilist infot saab esitada kas a) tervikuna sõnaartiklis, nt välde, rektsioon, erandlikud keelendid, või b) täiendades kirjade infot vastava grammatikaosaga, nt tüübnumber ja mõned vormid kirjes, kogu tüübikirjeldus ja muutevormide moodustusalgoritm sõnastiku lisas. Kui mõni väiksem tulemsõnastik ei vaja nii põhjalikku grammatilist infot, siis seal on võimalik osa andmeelemente esitamata jätta.

Tulemsõnastiku sihtkeelest sõltub otseselt **metakeel** (st seletuste ja märgendite keel), mis peaks olema lõppkasutaja emakeel. Kuigi EXS-i enda metakeel on eesti keel, peab ta arvestama teistega: tulemsõnastiku jaoks peab olema loodud võimalus vahetada teatud elementide keelt või lisada teises keeles olev üksus.

3.2. Sõltumatus sõnastiku tüübist

EXS on sõltumatu tema baasil koostatava kakskeelse tulemsõnastiku mahust ja otstarbest ning lõppkasutaja sihtgrupist. See tähendab, et ühest andmebaasist peab saama automaatselt genereerida erineva märksõnade arvuga ja erineva andmevalikuga standardseid sõnastikupõhju, mida konkreetse tulemsõnastiku koostaja saab oma töö käigus modifitseerida.

3.2.1. Märksõnade arvu reguleerimiseks on esialgu plaanis ette valmistada kolm erimahulist standardit. Seda saaks teha näiteks nii, et andmebaas sisaldab maksimaalse hulga märksõnu (ja kirjeid), kusjuures iga märksõna saab külge tunnuse, millisesse mahukategooriasse ta kuulub: kas suurde, keskmisse või väiksesse sõnastikku. Otsustava tähtsusega parameetrid mahukategooria määramisel on kuuluvus põhisõnavarasse, sagedus tekstides ja mõisteline kuuluvus (nt kui sageduse põhjal on ühte mahukategooriasse valitud mõni nädalapäeva nimetus, siis peab vaatama, et kaasa tuleks kõik nädalapäevad).

Selleks, et sõnu saaks paindlikult erimahulistest sõnastikes esitada ja et vajalikku infot kaotsi ei läheks, on oluline, et iga leksikaalne üksus oleks andmebaasis eraldi täiskirjena oma märksõnaga, mitte paigutatud mõne teise märksõna alla. Peamised kriitilised üksused on:

- **homograafilised** sõnakujud, nt *hall* (3 homonüümi);
- **liitüksused**: tuletised, liitsõnad, sõnaühendid (mille tähendus ei tulene komponentide tähendustest);
- **sõna muutevormid** leksikaalse üksusena (oma tähendusega), nt *õues*, *tasuta*, *elagu*, *avatud*;
- **üksuste variandid**: ortograafilised, häälduslikud, morfoloogilised jne.

Oma lõpliku koha sõnastiku struktuuris saavad sellised üksused tulemsõnastikus, ja neid lahendusi võib olla mitmeid. Vastavalt mahukategooria tähisele saab EXS-i märksõna teatud suurusega sõnastikupõhjas kas märksõnaks või artikli mingi allstruktuuri põhiüksuseks (allmärksõnaks) või mingi (näite)loendi osaks või jääb hoopis välja. Tulemsõnastiku toimetajal on võimalus sõnastiku ülesehitust ka ise kujundada.

3.2.2. Mahuvalikuga kuulub kokku **andmevalik**: suurem sõnastik esitab hari-likult ka põhjalikumat infot märksõna kohta. Põhiliselt sõltub andmevalik aga tulemsõnastiku sihtgrupist ja otstarbest. Sobiva valiku tegemiseks on mõeldav kasutada filtrite süsteemi. Näiteks nii, et andmebaas sisaldab märksõna kohta maksimaalset infot, kuid iga konkreetse tulemsõnastiku jaoks valmistatakse ette filter, mis osa infot välja praagib. Filtriga määratakse kirje koostis ja ülesehitus: milline on andmeelementide valik ja asukoht sõnaartikli struktuuriskeemis. Välja võivad jääda nt hääldus, etümoloogia, reksioon; morfoloogia võib olla antud erineva põhjalikkusega jne.

Enamasti on grammatilised andmed üksteisest küllaltki sõltumatud, ja ka nende valik on suhteliselt vaba. Eriolukorras on morfoloogiline info, mis on tihedalt seotud muutevormide analüüsi ja sünteesi algoritmidega – muidu pole kindlustatud ei märksõna leidmine ega sõna kasutamine lauses. Välja on töötatud kolm morfoloogiakirjete standardit, mis sisaldavad mitmeid väiksemaid sisulisi elemente, mis peaksid filtrisse kuuluma ühe tervikliku kompleksina. Neid on eri sõnastikes ka juba kasutatud:

- 1) vormipõhine kirje (suur) – sõna kõik põhivormid ja muuttüüp (nt EVS):
 tuba <tuba t`oa tuba t`uppa, tuba[de tuba[sid & tub/e S 18>
 lastetuba <+tuba t`oa tuba t`uppa, tuba[de tuba[sid & tub/e S 18>
 andma <`and[ma `and[a anna[b `an[tud, `and[is `and[ke V 34*>
- 2) tüvepõhine kirje (keskmine) – kõik erinevad tüvevariandid ja muuttüüp (nt NEEN eesti pool):
 tuba [toa, *adt.* tuppa 18e]
 laste-tuba [toa, *adt.* tuppa 18e]
 and**ma** [annab, *ips.* antud 34] ..
- 3) tüvemuutuste-põhine kirje (väike) – olulisi tüvemuutusi näitavad vormid ilma muuttüübita (nt ENNET eesti pool):
 tuba <toa, *adt* tuppa>
 laste-tuba[~]
 an|**dma** <-nab -tud>

Kaks esimest – vormi- ja tüvepõhine kirje – annavad mõlemad piisava info selleks, et oleks võimalik nii morfoloogiline analüüs kui ka süntees, kuid tüvepõhine kirje eeldab lõppkasutajalt suuremat valmidust teha tööd tüübikirjeldustega. Vormipõhine kirje on kasutajale mugavam ja sobib paremini nõrgema keeleoskuse puhul, sest kirjes on kõik sõnavormide moodustamiseks vajalikud põhivormid välja kirjutatud. Kolmas – tüvemuutuste-põhine kirje – näitab ainult põhilisi tüvevariante ja annab enamasti võimaluse märksõna leidmiseks (lemmatiseerimiseks), aga täielik morfoloogiline analüüs ja vormide süntees jäävad selle info põhjal kättesaamatuks.

3.2.3. Sõnastiku otstarbega ja kasutaja emakeelega on seotud tema **aktiivsuse ja passiivsuse** määratlus. Milline peaks olema EXS sellest aspektist? Kuna andmebaasi suund on E–X (eesti on lähtekeel), siis järelikult: X-keelsele kasutajale peaks see olema passiivne ja võimaldama E-keelsest tekstist arusaamist (E-analüüsi); E-keelsele kasutajale peaks see olema aktiivne ja võimaldama X-keelse teksti loomist (X-sünteesi). Et aga tegemist on pooliku sõnastikuga E–(X), kus eesti keele andmed on põhjalikult esitatud, samas kui X-keele asemel on ainult tühjad struktuuripositsioonid, siis langeb teine variant ära: X-keele sünteesi toe peab tagama iga konkreetse keelepaari jaoks koostatav sõnaraamat ise. EXS on järelikult orienteeritud X-keelsele kasutajale ja on passiivne sõnastik (võimaldab E-analüüsi).

Kaugemas perspektiivis hoiame sihikul siiski ka sõnastiku pööramise võimalust, nii et sama andmebaasi saaks kasutada X–E sõnastike loomisel. (X)–E puhul on eesti sihtkeeleks ning järelikult: E-keelsele lõppkasutajale peaks see olema passiivne ja võimaldama X-keelsest tekstist arusaamist (X-analüüsi); X-keelsele kasutajale peaks see olema aktiivne ja võimaldama E-keelse teksti loomist (E-sünteesi). Meie poolikus sõnastikus jääb sellest paarist alles teine variant, mis on samuti orienteeritud X-keelsele kasutajale, kuid on aktiivne sõnastik (võimaldab E-sünteesi).

Suurim muutus sõnastiku pööramisel on see, et endine märksõna liigub vaste positsiooni ja koos sellega ka märksõna juurde kuulunud info, kusjuures algselt passiivse sõnastiku lähtekeele nõuete kohaselt esitatud andmed peavad hakkama sobima aktiivse sõnastiku sihtkeele nõuetega (vrd tabel 1). Aktiivse sõnastiku jaoks vajalik grammatiline (sünteesi)info on EXS-is olemas, samuti ka sõna kasutamisega seotud andmed, nt valdkonna- ja stiilimärgendid, rektsioon jms, tarvis on kogu see andmestik transformeerida vastainfole sobivasse vormingusse. Sõna tähendusega seotud teave vajab põhjalikumalt revideerimist – see on vaste juures esitamiseks liiane.

Sõnastiku pööramine saab võimalikuks tänu EXS-i andmebaasi universaalsusele: kuna keeleandmete valik eesti märksõna juures on maksimaalne ja andmed ise põhjalikult struktureeritud ning struktuuritähistega varustatud, siis saab nende positsiooni ja staatust sõnastiku struktuuris vastavalt vajadusele muuta.

Kokkuvõttes peaks loodav EXS-i andmebaas olema orienteeritud X-keelsele lõppkasutajale, kusjuures andmete esitusviis peab võimaldama genereerida sellest passiivsete (suund E–X) ja edaspidi ka aktiivsete (suund X–E) tulemsõnastike põhjasid.

3.3. Sõltumatus kujundusest

Lõpuks peab EXS olema sõltumatu ka konkreetse tulemsõnaraamatu kujundusest (vormistusest). Selle (nagu üldse sõnastikuandmete töötlemise) tagatiseks on andmete esitamine XML (ingl *Extended Markup Language*) vormingus (vt ka 4. ptk). Kogu koostamis- ja toimetamistöö toimub XML märgendusega andmebaasis, kus iga sõnaartikli element on varustatud XML sildiga, mis näitab üksnes elemendi sisulist staatust. Kuidas üks või teine sõnaartikli komponent küljendatud tekstis välja näeb (fondid, kirjastiilid, taanded jms), määratakse eraldi reeglitega, mis pole seotud sisulise koostamistööga. Iga tulemsõnastiku toimetaja saab need reeglid ise määrata ja oma sõnastikku sel viisil kujundada.

Teine kujundusega seotud probleem on teksti lühendamine, mida tavasõnastikes kasutatakse paberi kokkuhoiu huvides küllalt palju. Muutvormidel asendatakse korduv (märksõnaga identne) sõnaosa sidekriipsuga, nt ÕS2006

kajut <g: -i> (*loe*: kajuti) ..

alfa.beet <2Q: -beedi, -.beeti> (*loe*: alfabeedi, alfa.beeti) ..

Näidetes asendab märksõna või tema osa sageli tilde, nt EVS:

algupära|ne .. ~ne kirjandus (*loe*: algupärane kirjandus) ..; ~sed tekstid (*loe*: algupärased tekstid) ..

algus .. julge ~ (*loe*: julge algus) ..; .. ~te ~ (*loe*: alguste algus) ..

ÕS2006 teeb lühendusi ka liitsõnaloendi piires, nii et lühendusmärgi asemele tuleb panna mitte märksõna, vaid loendi esimene või viimane sõna, nt

komme .. Elu_u, laua_ukomme. Eksi_u, ime_u, õnne_ukombel. Kombe_ulõtv,
v_uvastane, v_uvabadus, v_uõpetus

EXS-i andmebaasis pole ruumi kokkuhoid oluline, tähtsam on funktsionaalsuse tagamine. Seetõttu on kõik keelendid lühendamata välja kirjutatud, nii et neid saab andmebaasis otsida ja töö käigus ümber paigutada ilma vormistuse pärast muretsemata. Kõik lühendamised toimuvad automaatselt vastavate reeglite alusel küljendusvaate genereerimise ajal koos muu kujundusega. See tagab ka lühenduste töökindluse, nii et võõrkeelne kasutaja (v arvuti) saab lihtsa formaalse algoritmi alusel lühendatud kohad taastada.⁵

3.4. Ülaltoodud sõltumatusnõuete (sihtkeelest, sõnastikutüübist, kujundusest) täitmine EXS-i andmebaasis annab uue X-sõnaraamatu toimetajale üsna vabad käed. Ta saab valida sobiva suurusega sõnastikupõhja ja häälestada selle vastavalt oma tulemsõnaraamatu vajadustele nii andmete hulga ja struktuuri kui ka kujunduse ja muu vormistuse osas.

4. EKI sõnastike haldussüsteem ja kirjegeraator

EXS-i universaalsuse ja paindlikkuse kindlustamisel mängib suurt rolli EKI keeletehnoloogide poolt loodud töövahend – sõnastike haldussüsteem EELex –, mis kujutab endast leksikograafi elektroonilist töökeskkonda ning toetab ja abistab sõnaraamatu koostajat, võimaldades ka rühmatööd veebis (lähemalt vt Langemets jt 2006, Loopmann 2007). Seda kasutatakse nii EXS-i enda koostamisel kui ka kakskeelsete tulemsõnaraamatute tegemiseks.

EELex-i tööpõhimõtted toetuvad Nancy Ide juhtimisel sõnastike jaoks välja töötatud mudelile (Ide jt 2000), mille alustala on sõnastiku vormi ja sisu sõltumatus ning vormingute variatiivsus. Sõnastikku käsitletakse kui XML vormingus kodeeritud dokumenti, mida saab transformeerimiskeele XSL (ingl *Extensible Stylesheet Language*) abil teisendada muudeks XML dokumentideks, mille ele-

⁵ Üldjuhul kehtib sõnastikes lühendamisel kokkulepe, et muutvormides või näidetes asendab sidekriips v tilde tervet märksõna (nt EVS: algus .. ~te ~). Kui märksõnast kasutatakse ainult algusosa, siis tähistatakse lühendamiskoht mingi märgiga (nt EVS: algupära|ne .. ~sed). ÕS2006-s pole aga muutvormide lühendamisel formaalset algoritmi järgitud ja on jäetud lühendamiskohad märksõnas tähistamata. Seetõttu on osa ÕS2006 lühendusi arusaadavad ainult emakeelsele kasutajale, kes sõnavorme teab. Vrd nt

lihas <13: -e> ja **lihane** <13: -se>

kirmas <13: -e> ja **ainumas** <8: -ma, -mat>

.lööke[ne] <12 ja 13: -kese> ja **.maakene** <12 ja 13: -se>

järguline <12: -lise> ja **järgutine** <12: -se>

püra.miid <20: -miidi, -.miidi> ja **püra.miidjas** <8: -ja, -jat>

mentide konfiguratsioon (struktuur) võib olla hoopis teistsugune. See mudel annab võimaluse sõnastike struktuurielemente soovitud moel esitada, valida ja omavahel kombineerida, samuti ka sõnastikest infot otsida ja sobival kujul väljastada.

Kõik sõnastikubaasid EELex-is on XML vormingus (lähemalt vt Loopmann jt 2006: 248 jj), mis on kujunenud leksikograafias üldtunnustatud standardiks. XML on standardiseeritud andmekirjelduskeel, mis võimaldab andmeid struktureerida sisu järgi, sõltumata nende vormistusest. Iga sisuline sõnaartikli element (märksõna, tähendus, stiilimärgend, tüübinumber jne) on märgistatud tähisega (XML sildiga) ja tal on oma kindel koht sõnastiku struktuuriskeemis. Kuigi kogu sõnastiku koostamis- ja toimetamistöö toimub XML vormingu baasil, pakutakse kasutajale võimalust töötada ka visuaalselt mugavamate vaadetega, nt täita struktuurilahtreid sõnastikutabelis või teha väiksemaid parandusi küljendusvaates. XML vormingule toetub ka tulemsõnastike genereerimine EXS-i andmebaasist.

Haldussüsteemiga on plaanis siduda EKI grammatiline kirjegeneraator, mis dialoogis sõnaraamatu koostajaga lisab sõnaartiklisse morfoloogilise info, vastavalt valitud kirjestandardile. Kirjegeneraatori varasemat versiooni on juba mitme sõnastiku puhul kasutatud (vt Viks 2000b). Praegu käib selle edasiarendamine, et temast saaks igale leksikograafiale käepärane töövahend. Kirjegeneraatorit ei hakata kasutama kõigi märksõnade puhul, sest EXS-i andmebaas on valmiskujul salvestatud teatud hulk standardseid kirjeid. Kui aga koostaja lisab sõnaraamatusse mõne uue sõna, mida andmebaas pole, siis käivitub dialoogis töötav kirjegeneraator. Dialoogiliidese abil lahendatakse muuhulgas ka probleemid, mis pole arvuti poolt üheselt lahendatavad, nt kirjete jaotumine homonüümsete märksõnade vahel. Sõnastiku koostajal on võimalus teha sõna sisust sõltuvaid valikuid ja vajaduse korral korrigeerida generaatori väljundit.

Kirjandus

- Ide, Nancy; Kilgarriff, Adam; Romary, Laurent 2000. A formal model of dictionary structure and content. – U. Heid, S. Evert, E. Lehmann, C. Rohrer (eds). Proceedings of the ninth EURALEX International Congress, EURALEX 2000. Stuttgart: Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, 113–126.
- Järventausta, Marja 2004. Sanakirjat suomenoppijan avuna. – H. Sulkala, H. Laanekask (toim.). VIRSU II. Suomi ja viro kohdekielinä. Lähivertailuja 15. Suomen ja saamen kielen ja logopedian laitoksen julkaisuja 24. Oulu: Oulun Yliopisto, 24–44.
- Kaalep, Heiki-Jaan 1996. ESTMORF: A morphological analyzer for Estonian. – H. Öim (ed.). Estonian in the Changing World. Tartu: University of Tartu, 43–98.
- Langemets, Margit; Loopmann, Andres; Viks, Ülle 2006. The IEL dictionary management system of Estonian. – G-M. de Schryver (ed.). DWS 2006: Proceedings of the Fourth International Workshop on Dictionary Writing Systems. Pre-EURALEX workshop, Turin, 5th September. Turin: University of Turin, 11–16.
- Langemets, Margit; Mägedi, Merike; Viks, Ülle 2005. Süntaktiline info sõnastikus: probleeme ja väljavaateid. – M. Langemets (koost.), M-M. Sepper (toim.). Eesti Rakenduslingvistika Ühingu aastaraamat 1 (2004). Tallinn: Eesti Keele Sihtasutus, 71–98.
- Loopmann, Andres 2007. Sõnastike haldussüsteem EELex. Magistritöö. Käsikiri Eesti Keele Instituudis. Tartu: Tartu Ülikool.
- Loopmann, Andres; Sein, Kati; Viks, Ülle 2006. Sõnastike haldussüsteem Eesti Keele Instituudis. – M. Koit, R. Pajusalu, H. Öim (toim.). Keel ja arvuti. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 6. Tartu: Tartu Ülikooli Kirjastus, 246–258.

- Vare, Silvi 2002. Eesti keele sõnaperede sõnaraamatu koostamise põhimõtetest. – Keel ja Kirjandus 12, 855–866.
- Veldi, Enn 2000. Estonian and English: A lexicographer's point of view. I. – M. Ereht (ed.). Estonian: Typological Studies IV. Publications of the Department of Estonian of the University of Tartu 14. Tartu: Tartu Ülikooli Kirjastus, 173–197.
- Viks, Ülle 2000a. Eesti keele avatud morfoloogiamudel. – T. Hennoste (toim.). Arvutuslingvistikalt inimesele. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 1. Tartu: Tartu Ülikooli Kirjastus, 9–36. http://www.eki.ee/teemad/avatud_mrf.html (28.09.2007).
- Viks, Ülle 2000b. Kuidas tekib sõnastikukirjese grammatika. – Keel ja Kirjandus 7, 486–495. <http://www.eki.ee/teemad/kirjegeneraator.html> (28.09.2007).
- Viks, Ülle 2001. Muuttüübid eesti sõnastikes. – M. Langemets (toim.). Leksikograafiaseminar: Sõna tänapäeva maailmas. Leksikografinen seminaari: Sanat nykymaailmassa. Eesti Keele Instituudi toimetised 9. Tallinn: Eesti Keele Sihtasutus, 157–185. <http://www.eki.ee/teemad/tyybijutt.html> (28.09.2007).
- Viks, Ülle 2003. Viron muoto-oppia / Eesti morfoloogia. Viron vastineiden taivutussanasto / Eesti vastete morfoloogiasõnastik. – Soome-eesti suursõnaraamat 2. Tallinn: Eesti Keele Instituut; Helsinki: Kotimaisten kielten tutkimuskeskus, 929–1301.

Sõnaraamatud ja võrgumaterjalid

- ENNET = Farbregd, Turid; Lepp, Hille; Viks, Ülle 1999. Estisk Lommeordbok. Eesti-norra norra-eesti. Oslo: Kunnskapsforlaget.
- EVS = Eesti-vene sõnaraamat. 1997–2006 (1–4). Tallinn: Eesti Keele Sihtasutus.
- NEEN = Farbregd, Turid; Kangur, Sigrid; Viks, Ülle 2006. Norra-eesti eesti-norra sõnaraamat. 2. trükk. Tallinn: Eesti Keele Sihtasutus.
- SUVI = Soome-eesti suursõnaraamat. 1-2. Anu Haak, Paul Kokla, Külli Kuusk, Helga Laanpere (koost.), Valdek Pall (toim.). Tallinn: Eesti Keele Sihtasutus, 2003.
- VVS = Viks, Ülle 1992. Väike vormisõnastik I, II. Tallinn: Keele ja Kirjanduse Instituut.
- ÕS2006 = Eesti õigekeelsussõnaraamat ÕS 2006. Tiiu Ereht (toim.), Tiiu Ereht, Tiina Leemets, Sirje Mäearu, Maire Raadik (koost.). Eesti Keele Instituut. Tallinn: Eesti Keele Sihtasutus.
- EKI automaatne morfoloogiasüsteem. <http://www.eki.ee/tarkvara/> (28.09.2007).

Ülle Viks (Eesti Keele Instituut) on lõpetanud Tartu Riikliku Ülikooli eesti keele ja kirjanduse erialal. Kaitses samas ülikoolis 1978. a kandidaadiväitekirja eesti verbide morfoloogilisest klassifikatsioonist ja 1994. a doktoriväitekirja eesti keele klassifikatoorsest morfoloogiast. Uurimisvaldkonnad: avatud morfoloogiamudel ja selle realisatsioon arvutis (sõnavormide reeglipõhine süntees ja analüüs), grammatilise informatsiooni esitamine sõnastikes.
ylle@eki.ee

ESTONIAN-X DICTIONARY AND GRAMMAR

Ülle Viks

Institute of the Estonian Language

Problems connected with representation of grammatical information in a bilingual dictionary are discussed. The focus lies on the Estonian-X dictionary (EXS), its place in dictionary typology and consequent requirements. The EXS is a universal lexicogrammatical database enabling to generate the electronic basis for an arbitrary bilingual dictionary, providing Estonian as the source language ready-made in a standard format. It is designed to present grammatical information as well structured as possible, tapping various levels of language description from phonology to semantics; data selection and presentation are based on relations of dictionary vs. grammar. The EXS being a metalevel database its format should be flexible enough to be independent of the target language, type and design of the output dictionary. A brief introduction is made of the available tools to develop and implement the EXS: the dictionary management system EELex (lexicographer's working environment) and a grammatical entry generator.

Keywords: bilingual lexicography, grammatical information, language processing, XML, dictionary management system, Estonian

EESTI SILBISÜSTEEMI STRUKTUURIST

Leo Võhandu, Kairit Sirts, Eik Aab

Ülevaade. Artiklis uurime eesti keele silpide sagedusmaatriksit, mis on loodud eesti kirjakeele korpuse baasil, kasutades ilukirjandustekste aastatest 1988–1998. Vaatluse all on 1000 kõige sagedamini esinevat silpi. Eksperimentaalsete uurimismeetoditena kasutasime Hamiltoni tee ning sotsiaalse võrgustiku klastrite leidmist eesmärgiga uurida silbisüsteemi sisemist varjatud struktuuri. Hamiltoni tee leidmiseks vajaliku programmi loomiseks kasutasime programmeerimiskeelt J-6.01. Programm ühendab kõik silbid järjestikusse maksimaalse ühenduste summaga jadasse ilma silpe kordamata. Teise meetodina leidsime sotsiaalse võrgustiku klastrid 100 ja 1000 kõige sagedamini esineva silbi jaoks. Selles lühikeses artiklis esitame graafiliselt ainult 100 silbi võrgustiku ilma täpsemate selgitusteta selle leidmise matemaatilise meetodi või esituse interpretatsiooni kohta.

Võtmesõnad: arvutilingvistika, silbitamine, silbiseostus, graafesisus, Hamiltoni tee, silpide sotsiaalne võrgustik, eesti keel

1. Sissejuhatus

Raalingvistika areng näitab üha veenvamalt, et mida rohkem me suudame tekstikäsitlust automatiseerida, seda kasulikumad ja põnevamad on tulemused. Seejuures on oluline, et me peame üritama võimalikult vähe kasutada üldlingvistilisi eelteadmisi. Miks? Nõnda toimides me jälgime teaduse üldist arengujoont. Formaalsete meetodite õitseage sai alguse reaalse maailma suhteliselt lihtsalt kirjeldatavate või modelleeritavate objektsüsteemide uurimisest (nt mehaanika, optika). Tekkinud probleemide lahendamiseks loodi tasapisi rida meetodeid, mis töötasid nn polünoomiaalses ajas. Samade meetodite ülekanne hägusast või diskreetset modelleerimist nõudvatele objektsüsteemidele ei õnnestunud aga kuigi hästi. Peamiseks põhjuseks oli see, et hägusad süsteemid on põhiliselt iseorganiseeruvad, mitte kindlate reeglite järgi konstrueeritavad. Piltlikult öeldes, Loojal ei olnud olemas formaalset süsteemi (matemaatikat), mille abil ta oleks saanud deterministlikku maailma ehitada.

Kellelgi meist ei ole valemit otsa ees ega südamesse peidetud, mille kohaselt me peaksime käituma.

Iseorganiseeruvuse idee võimaldas käsitleda protsesse ja nähtusi, milles nn avatud süsteemi sisemise organisatsiooni keerukus kasvab, ilma et seda välismõjud otseselt juhiksid või haldaksid. Termin *iseorganiseeruvus* võttis ilmsi kasutusele 1947. a psühhiaater ja insener W. Ross Ashby. Laiemalt tuli see sõna käibele 1948. a Norbert Wieneri küberneetikaraamatu kaudu (eesti keeles Wiener 1961). Asjast huvitusid ka filosoofid ja nii avaldas Noam Chomsky ning Hilary Putnami õpetaja Nelson Goodman 1951. a raamatu “The Structure of Appearance” (“Nähtumuse struktuur”). Matemaatika ju sellega tegelebki, et leida nähtumust kirjeldavaid või peidetud struktuure. Et selline lähenemine viljakas on, näitab kasvõi akadeemik Jaan Einasto grupi töö tumeda aine uurimisel astronoomias (Näha pole ju midagi!).

Nendest ideedest lähtudes üritamegi tungida eestikeelse teksti kui nähtuse peidetud struktuuri, seejuures võimalikult iseorganiseeruvuse ideed kasutades. Loomulikult jääb see ekskursioon pinnapealseks, aga ehk leidub järeltegitajaid ja ümbermõtlejaid.

2. Andmestik

Võtsime Tartu Ülikoolis loodud kirjakeele korpuse 1988–1998 ilukirjandusosa (vt ka Hennoste jt 2001).¹ Kasutades Eesti Keele Instituudi silbitajat² ja J-keelt³ koostasime kõigi silbipaaride jaoks järgnevuste sagedustabeli. Enne silbitamist teisendasime teksti 8-bitilisse UTF-formaati.⁴ Täpsemaks analüüsimiseks võtsime 1000 sagedasemat silpi pluss tühiku iseseisva üksusena. Saadud sagedustabeli veerus on reas esitatud silbile järgnenud silp. Töötlustarkvara loomiseks kasutasime J-keele versiooni 6.01. Et kogu sagedustabeli esitamiseks selles lühikäsitluses ei ole ruumi, siis esitame tervest 1001x1001 sagedustabelist vaid selle vasakpoolse ülemise 10x10 nurga (kõige tihedama osa tabelist).

Tabel 1. Silpide sagedustabel

	<i>le</i>	<i>ta</i>	<i>ma</i>	<i>ja</i>	<i>se</i>	<i>da</i>	<i>o</i>	<i>li</i>	<i>ga</i>	<i>te</i>
<i>le</i>	535	545	1024	98	24	125	96	72	707	191
<i>ta</i>	37	146	1438	582	62	1599	0	76	1750	129
<i>ma</i>	719	1179	37	604	115	145	0	214	608	207
<i>ja</i>	343	174	119	23	17	71	0	176	262	330
<i>se</i>	1419	179	254	23	15	2713	18	113	1392	167
<i>da</i>	598	346	407	93	23	402	0	22	93	93
<i>o</i>	2709	3	4261	21	10	41	0	8661	16	13
<i>li</i>	147	266	190	58	858	76	33	56	106	138
<i>ga</i>	92	274	181	38	26	86	0	17	37	20
<i>te</i>	1011	112	2628	1	8	1272	28	77	779	52

Tabelit 1 vaadeldes on näha, et esineb rida loomulikke järgnevuspaare (paarile järgnev arv näitab paari esinemissagedust):

¹ Vt <http://www.cl.ut.ee/korpused/baaskorpus/> (10.10.2007).

² Vt <http://www.eki.ee/tarkvara/silbitus/> (10.10.2007).

³ J-keel = Jsoftware, vt <http://www.jsoftware.com/stable.htm> (10.10.2007).

⁴ Vt <http://www.chmaas.handshake.de/delphi/freeware/xvi32/xvi32.htm> (10.10.2007).

o-li 8661, *se-da* 2713, *o-le* 2709, *te-ma* 2628, *ta-ga* 1750, *ta-da* 1599, *ta-ma* 1438, *se-le* 1419 jne.

Mida taoliste tabelitega teha? (Eriti, kui need tabelid nii suured on!)

Andmeanalüüsi üldteoorias on selliste maatriksite korrastamiseks mitmeid meetodeid. Enamasti baseeruvad need meetodid mingite seoste või mõõtude maksimeerimisel.

3. Silpide järjestamisest. Hamiltoni tee

Kõige esimene meetod on mõtteliselt lihtne ja üritab silbid niimoodi järjestada, et nad kõik oleksid üksteisele järgnevad ja iga silbipaari koosesinemiste arvude summa üle kogu uue järjestuse oleks maksimaalne. Seda kõiki silpe ühendavat optimaalset sidusat teed nimetatakse *Hamiltoni teeks*. Graafiteoorias ja diskreetse optimeerimise valdkonnas on Hamiltoni tee leidmine hästituntud ülesanne ja kuulub nn NP-keerukate probleemide hulka. Mida see tähendab? N objektist koosneval süsteemil võib kõiki objekte paigutada $N!$ erinevasse järjestusse. Kui väikese objektide arvu korral on nende järjestuste leidmine lihtne töö (nt kolme objekti a, b, c korral on järjestusi 6: abc, acb, bac, bca, cab, cba), siis juba $N=10$ korral on järjestusi $10!=3628800$. Kui oletada, et iga järjestuse ja selle naaberseoste sageduste summa leidmiseks kulub ainult 1 sekund, siis kümne objekti kõigi Hamiltoni teede leidmiseks ja hindamiseks kuluks ei rohkem ega vähem kui 42 ööpäeva!

Selgub, et näiteks $N=20$ puhul tuleks kõikide variantide läbimise korral tööd teha $7,7 \cdot 10^{10}$ aastat. Tööaeg on nii suur, et tulemus ei huvitakski enam kedagi. (Vahemärkusena lisame, et J. Einasto andmetel kõrvetab Päike meid kõiki juba umbes miljardi aasta pärast olematusse ja Maa ise hävib ca 5 miljardi aasta pärast.) Nüüd peaks selge olema, et 1000 objekti korral pole mingit lootustki täpse lahendi leidmiseks. Arvutiteaduses on rida ligikaudseid meetodeid, mis päris korralikke lähendeid annavad.

Me kasutame ühte J-keelset programmi, mis võimaldab Hamiltoni teed leida ka üpris suurte maatriksite korral. Selle keele plussiks on ülilühikesed programmid (mis ei tarvitse küll kergesti arusaadavad olla). Kui meie silpide sagedustabeli nimeks on *mm*, siis vastav Hamiltoni tee programm, mis stardib *i*-ndast silbist, läbib kordusteta kõik ülejäänud silbid ja trükib selle jada ka välja, ei ole kuigi pikk:

```
imax=. i.>./
ntg=. dyad define
imax o({:x)}({:x}){y
)
hp=.]`($:@(],ntg&mm))@.((#mm)&~:@#)
```

Näiteks hp 6 annab tulemuseks silbijada

o li se da le ma ta ga ja te

Toome näitena veel mõned silbijadad:

Ja te ma ta ga se da o li
Te ma ja le o li se da ta ga

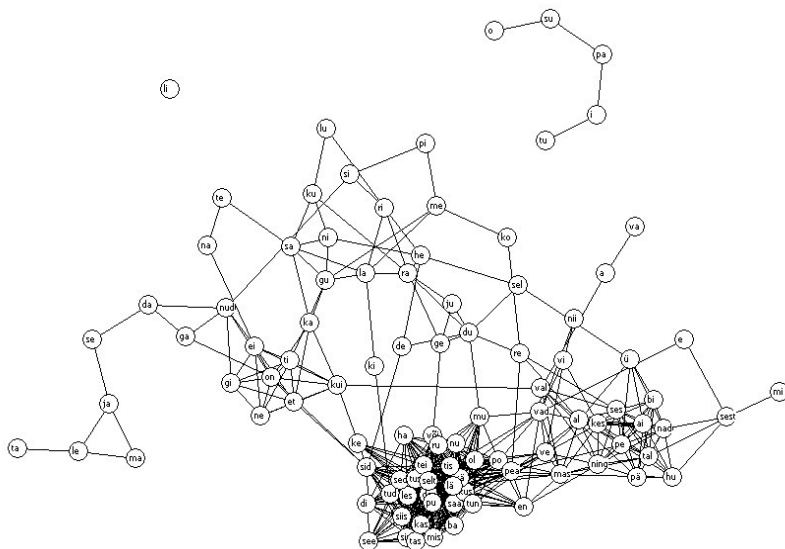
Näeme, et silpide seostus on vägagi suupärane. Huvitav on neid “jorusid” võrrelda üle 40 aasta tagasi Leo Võhandu informatsiooniteooria seminaris Tartu Ülikoolis Priit Järve (1965: 97–103) poolt saadud tulemustega eesti keele fraaside entroopiamudelite kohta. Kahe- ja kolmetäheliste ühendite puhul saadud fraasid olid järgmised:

*Siserdamega mibul ta lid neijud ka onndagine
Tas peanud lestemate kõnniigla kaltsed ase*

Ilmne on, et kahtäheliste silpide ühenduvus on parem 3-grammidest saadud tulemusest. Oleks põnev, kui mõni üliõpilane otsiks P. Järve artikli välja ja üritaks silpidega teha sama testi piisavalt mahuka tekstikorpuse peal.

4. Silbid süsteemina

Eelmises jaotuses vaadeldud Hamiltoni tee kõrval on veel terve rida meetodeid, mis silbisüsteemi sisemist peidetud struktuuri avada aitaksid. Me vaatleme siin vaid ühte Eik Aabi poolt loodud silpide sotsiaalvõrgulise sõprusmodeli tulemuspilti. Ideeks on valida antud silbile lähim naaber mitte paaritise koosinemise sageduse, vaid 100- või 1000-silbilise koosinemise alusel. Nii saadud sarnasusmõõt näitab, kui võrd sarnaselt kaks vaadeldavat silpi käituvad ülejäänud sagedaste silpide suhtes. Teisisõnu, me otsime silpe, mille käitumismuster silpide kogumis on “globaalselt” analoogiline. Iga silp otsib teiste silpide hulgast endale “iidolit”. Iga silbi jaoks leitakse efektiivsus $e=R/E$, kus R on sõprade arv ja E on nendele kokku kuluv energia (kauguste summa). Iidoliks valitakse sõprade hulgast need, kellel on suurem efektiivsus (ehk tihedus tema ümber on suurem). Algoritmi täielik kirjeldamine ei oleks siinkohal sobilik. Küll aga esitame 100 sagedasema silbi graafilise pildi.



Joonis 1. Silpide sotsiaalne võrgustik

5. Mis on sellest kõigest kasu?

Silbistruktuuri uurimine on alles algstaadiumis, seepärast on praegu raske kõiki võimalikke rakenduskohti välja tuua. Nimetame vaid mõned:

- Markovi peitmudelite õppematerjali minimeerimine;
- kõrge katteväärtusega lihttekstide moodustamine. Kasulik oleks see näiteks meie keele õpetamisel võõramaalastele;
- keel kui “Small World” mudelite klassi kuuluv süsteem;
- vanade tekstide silbistruktuuri uurimisel saab infot nii fonotaktika kui silbitaktika muutumise trendidest;
- teades sagedusseostuste süsteemi nihkeid, saab ennustada edasisi muutusi (nn iteratiivne ennustamine) jne jne.

6. Mida edasi teha?

Juba mainitud teemade uurimise kõrval tasub ilmselt tõsiselt tähelepanu osutada ka morfeemidele. Soomes loodi huvitav süsteem Morfessor morfeemide automaatseks eraldamiseks, et vältida kõnetuvastuses sõnastikuväliste sõnade suurt hulka (Siivola jt 2003). Soome keele jaoks töötas see süsteem efektiivselt. Türgi ja suahiili keele puhul ei olnud tulemused nii head.

Eks me teksti saa tükeldada just nende kolme struktuurielemendi – sõnade, silpide ja morfeemide abil. Milline nendest meie käsitluse jaoks sobivaim on, seda näitab tulevik.

Kirjandus

- Hennoste, Tiit; Kaalep, Heiki-Jaan; Muischnek, Kadri; Paldre, Leho; Vaino, Tarmo 2001. The Tartu University Corpus of Estonian Literary Language. – Tõnu Seilenthal, Anu Nurk, Triinu Palo (eds). Congressus Nonus Fenno-Ugristarum, Pars IV, Dissertationes sectionum: Linguistica. I. Tartu, 337–344.
- Järve, Priit 1965. Eesti keele kirjapildi entroopiast ja liiasusest. – Keel ja struktuur 1. Töid struktuuraalse ja matemaatilise lingvistika alalt. Tartu, 97–103.
- Siivola, Vesa; Hirsimäki, Teemu; Creutz, Mathias; Kurimo, Mikko 2003. Unlimited vocabulary speech recognition based on morphs discovered in an unsupervised manner. – Proceedings of the 8th European Conference on Speech Communication and Technology (EUROSPEECH 2003). Geneva, Switzerland, 2293–2296.
- Wiener, Norbert 1961. Küberneetika. Tallinn: Valgus.

Korpused ja tarkvara

- Eesti kirjakeele korpused 1890–1990. <http://www.cl.ut.ee/korpused/baaskorpus/> (10.10.2007).
- Eesti Keele Instituudi tarkvara. Silbitus. <http://www.eki.ee/tarkvara/silbitus/> (10.10.2007).
- J-keel = Jsoftware, <http://www.jsoftware.com/stable.htm> (10.10.2007).
- Freeware Hex Editor XV132. Version 2.51. <http://www.chmaas.handshake.de/delphi/freeware/xvi32/xvi32.htm> (10.10.2007).

Leo Võhandu on lõpetanud Tartu Ülikooli matemaatikaosakonna, praegu Tallinna Tehnikaülikooli emeriitprofessor. Uurimisvaldkonnad: andmeanalüüs, keerukate andmekogumite peidetud struktuuri avamine, graafiteooria.
leovoo@hotmail.ee

Kairit Sirts (Tallinna Tehnikaülikool) on informaatika eriala magistrant. Uurib eestikeelse silbisüsteemi peidetud formaalset struktuuri.
karambula@hotmail.ee

Eik Aab (Tallinna Tehnikaülikool) on informaatika eriala doktorant, teemaks iseorganiseeruvate sotsiaalsüsteemide uurimismeetodid.
eik.aab@innofusion.ee

A PRELIMINARY STRUCTURAL VIEW OF THE ESTONIAN SYLLABLE SYSTEM

Leo Võhandu, Kairit Sirts, Eik Aab

Tallinn University of Technology

Using the Corpus of Estonian Literary Language and specifically the selection of fiction texts from years 1988–1998 we have studied the frequency table of Estonian syllables (1000 most frequent syllables). As an experimental study we have used Hamilton Path (HP) and Social Network Clusters to study the inner structure of syllable system. For HP we have created and represented a program in J-6.01 which connects all syllables into a sequential path with the maximal connections sum and without a syllable repetition. As another method we have created SNC networks for 100 and 1000 most frequent syllables. In this short article we have presented graphically only the 100 syllable network without explaining exactly either the mathematical method or networks interpretation.

Keywords: computational linguistics, syllabification, syllable association, graph representation, Hamilton path, syllable social network, Estonian

EESTI KEEL JA TEISED KEELED – EESTI JAOKS

Suliko Liiv

Ülevaade. Ühinemine Euroopa Liidu ja NATO-ga muutis oluliselt eesti- ja võõrkeelte õpetamise-valdamise senist paradigmat Eestis. Uus paradigma eeldab võimalikult laiemat eelnevat avalikku arutelu: tuleks läbi vaielda erinevad visioonid ning sündivate kompromisside alusel sõnastada pikaajaline eesti ja võõrkeelte õpetamise ning kasutamise strateegia Eestis.

Artikli eesmärk on elavdada uueneva keelepoliitika paradigma arutelu arvestades, et haridus- ja teadusministri lauale on mitmeaastase töö tulemusena jõudnud Eesti võõrkeelte strateegia esialgne kirjalik versioon.

Võtmesõnad: mitmekeelsus, võõrkeelte strateegia, keelepoliitika, eesti keel

Sissejuhatus

Arvestades väikese riigi väheseid ressursse, on otstarbekas teadvustada Eesti reaalne keelepoliitiline olukord: 1880-ndail oli venestamine, 1920-30. aastatele oli tunnuslik eesti keele säilitamine-arendamine ning saksa keele mõjude vähendamine. 20. sajandi teine pool kujunes vene keele mõjudega toimetulemiseks. Kahe eelneva kogemusega varustatult on mõistlik mitte jääda ajaloo objektina ootama sündmuste spontaanset kulgu, vaid tegutseda võimalikult aktiivselt oma emakeele edasikestmise kindlustamiseks: analüüsida praegust hetkeolukorda ning arutada võimalikult kiiresti süsteemselt läbi eesti keele jätkusuutlikkus; defineerida riigikeele tulevikuvision ühitatult suureneva erinevate võõrkeelte survega sellele ning kaaluda selle visiooni võimalikult ideaalilähedasemaks saavutamiseks vajalikud meetmed.

Euroopa Liidu tööturu avanemine esitab Eesti haridussüsteemi keeleõppele uue taseme- ja väljundinõude – haridussüsteemi väljundiks olgu Eesti elanike suutlikkus õppida ja töötada edukalt nii Eestis kui ka mujal maailmas. Üleilmas-

tumise visioonis peitub ilmselt eestlaste jaoks seni selgelt sõnastamata nõue: et väikeriigina ja Euroopa Liidu liikmesriigina olla üleilmselt konkurentsivõimeline, tuleb meil osata laitmatult võõrkeeli, sh inglise keelt, jagada euroopalikke hoiakuid ning väärtustada maailma euroopalikult. 2007. a oleme oma kooliharidusega sarnaspoole liikumas: inglise keelt õpib 83,7%, vene keelt 40,9%, saksa keelt 17,8% ja prantsuse keelt 3,0% üldhariduskoolide õpilastest (EHIS). Võõrkeelte õppimise soodsat mõju isiksuse arengule on tõestanud mitmed uuringud (Byram, Fleming 1998, Clyne, Rossi-Hunt, Isaakidis 2004, Cook 2002, Williams 2004). Samas oleks hädavajalik ka teisesuunalise mõju võimas tugevdamine: eesti keele ja kultuuri võimalikult aktiivsem tutvustamine võõrkeelte vahendusel.

Peaaegu kümme aastat tagasi juhtis käesoleva artikli autor tähelepanu vajadusele töötada välja võõrkeelepoliitika arvestades Eesti Euroopa Liidu liikmeks saamisega (Liiv 1998: 639–640). Taas on käes aeg läbi vaadata möödunud aastate kogemused ning ajakohastada muutunud oludest johtuvalt keelepoliitikat.

Eesti keel rahvusliku identiteedi alustalana

Keel on peamine sotsialisatsiooni vahend. Samal ajal kui laps õpib rääkima, kasvab temast ühiskonna liige. Põhiliselt lapsepõlves omandatavate oskuste-hoiakute eiramine keelepoliitikas loob soodsa pinnase sotsiaalsetele konfliktidele.¹

Tänini on Eestis valminud kaks keelevaldkonna strateegilist arengukava.² Need jätkuvalt olulised dokumendid kannavad endas paratamatult eelmise sajandi viimase aastakümne Eesti vaimse õhustiku hõngust ja eesmärke: taasiseseisvumine, desovetiseerimine ja Euroopa Liitu pürgimine. Nende üks emotsionaalsemaid taustategureid oli püüd neutraliseerida naaberriigi tugevnenud survet muuta Eesti vene keele eelistusega kakskeelseks nõukogude inimeste piirkonnaks. Samas on eestlastel peamiseks identiteedi alustalaks enam kui sajandi vältel olnud just oma maa keel, mistõttu Teise maailmasõja järgseil aastail keskendusid siinsed poliitilised pinged valdavalt keele, kultuuri ning majanduse paradigmadele. Seda aega ise mitte üleelanuile võib tunduda artefaktina, et Eestis praktiliselt puudusid 20. sajandil üleilmselt nii sagedased religioossed ja rassilised (sealhulgas antisemiitlikud³) pinged.

Nüüd, kus oleme nii Euroopa Liidus⁴ kui ka NATO-s, tuleks uusi tulevikuparadigmasid eeldades ning võimalike vigade ja konfliktide vältimiseks lähitulevikus koostada, põhjalikult läbi arutada ning ilmingimata riiklikul tasandil vastu võtta pikaajalise toimega keeledokument, kus oleks eraldi poliitilise konsensuse alusel määratletud võõrkeelepoliitika ülesanded ning nende elluviimise vahendid 21. sajandi Eestis. Lähtealusena võiks kasutada 2003. aasta algul, enne Euroopa Liiduga ühinemist koostatud uurimust “Eesti elavik 21. sajandi algul” (Vihalemm jt 2004), mis annab ülevaate Eesti elanike võõrkeelte oskusest, kasutamisest ja hoiakutest rahvus- ja vanuserühmade lõikes.

¹ Vt nt Rannut jt 2003: 92.

² Mitte-eestlaste loimimisele suunatud “Muukeelse elanikkonna keeleõppestrateegia” (1998) ning eesti keele arendamisele ning kaitsele pühendatud “Eesti keele arendamise strateegia 2004–2010” (2004), kus loetletakse ka võõrkeeli, millised oluliselt mõjustavad eesti keelt olevikus ja lähitulevikus. Artikkel on kirjutatud ja esitatud enne “Eesti võõrkeele strateegia aastani 2015” töövariandi esitamist haridus- ja teadusministrile detsembris 2007.

³ Neil sõjajärgseil aastail ilmesid mitmes NSVL venekeelsetes kõrgkoolides ja teadusasutustes antisemiitlikud meeleolud, mis tõi paljusid juudi rahvusest üliõpilasi õppima Eesti kõrgkoolidesse, ka asusid Eestisse elama mitmed silmapaistvad juudi rahvusest intellektuaalid. Autori märkus.

⁴ Teoreetiliselt on eesti keel Euroopa Liidu üheks võrdseks töökeeleks, praktiliselt peame aga tõdema, et meil on neid laialdasi õigusi raske oma õlgadel kanda ja seetõttu peame paratamatult minema kompromissile, valides Euroopa Liidus omi asju ajades praktilise käibekeelena mõne suurema (ilmselt inglise) keele.

Riikliku võõrkeelte strateegia vajalikkus Eestis

Sellega tõstatub rida prognoositavaid probleeme, mille ennetamisele on juba praegu vaja mõelda.

Tunnustades aksioomina eesti keelt ainsa riigikeelena Eestis, annaks riikliku võõrkeelte strateegia selge sõnastus Eestile ülimalt vajaliku varuaja ennetamiseks ja minimeerimaks mitmeid siinmail uue immigratsioonilaine tekkel tõenäoselt pingestuvaid eri keelte, kultuuride, religioonide ja rasside vahelisi pingeid. Taolisi pingeid ei ole suutnud vältida isegi eriti immigrandisõbralik Austraalia.

Kuivõrd Eesti on väike hõredalt asustatud riik,⁵ kus loomulik iive madal ning tööjõupuudus süveneb, tuleb eeldada: 21. sajandil koos Eesti elatustaseme, sh sotsiaalturbe jõudsa paranemisega tõuseb immigratsioonisurve Eestisse. Sellega tõstatub rida prognoositavaid probleeme, mille ennetamisele on juba praegu vaja mõelda.

1. Kõikidele Eesti alaliste elanikele kogu 21. sajandi kestel oleks keeleoskuse prerogatiiviks võimekus suhelda eesti keeles vabalt töö- ja olmetasandil. Kuidas seda motiveerida, milliste vahendite-võtetega tagada?
2. Maksimaalselt palju elanikest omandaks piisaval tasemel uue kõnekeele (tõenäoliselt eelistataks inglise keelt, pole aga välistatud ka prantsuse ja saksa keele oskuse laienemine).
3. Mis on motivatsiooniks ja tagatiseks, et 20. sajandi teisel poolel siia saabunud immigrantide järglased omandaksid koolihariduse käigus riigikeele, rahvusvahelise suhtlemise keelena inglise keele ning piisaval tasemel ka oma esivanemate keele? Seejuures tuleb nendele samaaegselt keeleõppega kindlasti põhjalikumalt õpetada uue asukohamaa ajalugu, kultuuri ning selgitada, et uue asukohamaa kodakondsuse omandamine tähendab lisaks õigustele ka otsest kohustust olla lojaalne selle maa põhiseadusele ning sellele tuginevatele seadustele. Juhul, kui seda läbimõeldult ei tehta, sugenevad kindlasti pinged, millega kaasneb immigrantide teise põlvkonna pidev nõudmine anda ka nende emakeelele Eestis ametliku keele staatus. Seega tuleks riigikeele õpetamine ilmtingimata ühitada uue asukohamaa ajaloo, kultuuri tutvustamise ning sellele riigile vääramatu lojaalsuse kasvatamisega.
4. Teise maailmasõja järgsed immigrantid, kes olid/on küll valdavalt venekeelsed, kuid kuulusid NSVL-is elanud ca 120 eri rahvuse või rahvusgrupi hulka (näiteks sel perioodil Eestisse elama asunud neenetsid, vt Viikberg 1999) seisavad sellisel juhul loomulikult küsimuse ees: kas nad on muutumas nelikeelseks.

See toob usutavasti kaasa paljude väikerahvuste järglaste siirdumise olmekeeles seniselt vene keelelt riigikeelele. Viimast protsessi on ilmselt võimalik abistada, luues neile omakeelseid kultuuriklubisid (umbes nii nagu Kanadas ja USA-s tegutsesid 20. sajandi teises pooles eesti pagulasklubid).

5. Pärast Eesti taasiseseisvumist saabunud uusimmigrantid, kelle puhul on tegemist juba eri rasside ja religioonide alanud sisserändega. Oluline strateegiline probleem on, et nad ei jääks Eestisse elama asudes pelgalt ingliskeelseks ega liituks olmetasandil 20. sajandi teisel poolel immigrerunud venekeelse elanikkonnaga.

⁵ Eestis 32 in/km², Taanis 126 in/km², Hollandis 481 in/km², Indias 394 in/km², Hiinas 138 in/km² (EE).

Seega aitab pikaajaline võrkeelte strateegia juba praegu läbi mõelda uusimmigrantide laine Eesti ühiskonda sulandamise strateegia: kas võtame eri maadest ja eri religioonidest saabuvad ühiskonda vastu kompaktselt koos nende kodumaal tavaks saanud kommetega, kärpides sellega eestlaste seniseid kultuuritraditsioone ja pannes aluse Eesti praeguse riigikeele õiguste laiendamisele immigrantide nõudel varem-hiljem ka inglise ning vene keelele? Sellisel puhul peame eelnevalt põhjalikult läbi arutama, kuidas vältida põliselanikel tekkivat traditsioonilise kultuuri ning elulaadi laienevat ahistamistunnet üha lisanduvate uute võorkultuuride ja meile võõraste, kuid oma nõudmisi aktiivselt esitavate uute religioonide voogudes.

Ka toimiks võrkeelte omandamise pikaajaline strateegia tulevases multikultuurseks muutuvast Eestis immigrantide eri põlvkondade ning põhielanike vaheliste pingete olulise mõjurina (vt ka Bertriveau jt 2001: 279–288).

Võib arvata, et kolmandast maailmast hakkab meile saabuma ennekõike madala haridustaseme ja teistsuguse ellusuhtumisega uusimmigrante, kellele peame kindlasti selgeks õpetama meie riigikeele ning inglise, saksa, prantsuse või hispaania keele. Täni on aga isegi see, kas Eesti muukeelsele elanikule on riigikeel võõrkeel või mitte, kõikides nüanssides läbi vaidlemata ning defineerimata.

Ülalesitatust tulenevalt vajab eesti keele kui immigrantide jaoks vähemalt teise keele – uue asukohamaa riigikeele – õpetamine Eesti muukeelsele elanikkonnale (sh uusimmigrantidele) nende eesti ühiskonda lõimumise soodustamiseks tulevikule mõeldes juba praegu hoolikat läbimõtlemit. Muuhulgas on vaja ka eesti keele, kultuuri ning sinise põlisrahva traditsioonide võimalikult aktiivsemat tutvustamist vähemalt Euroopa Liidus, miks mitte ka kogu maailmas. Seejuures tuleb aga poliitilise korrektsuse tagamiseks vältida vanade ja uusimmigrantide religioosete märkide ja põhiraahvuse kultuuriliste tõekspidamiste konflikti. Pole ilmselt vaja põhjalikumalt selgitada: parema võõrkeelteoskuse ja kasutuskogemusega kaasneb indiviidi suurem kultuuriline avatus, tolerantsus ja informeeritus, erinevate kultuuride läheduse ning isiklike liikumisvõimaluste paranemise tunnetamine. Seeläbi soodustab parem võõrkeelteoskus ja keeleõppevõimaluste aktiivsem kasutamine teiste faktorite kõrval kultuuriliselt salliva ning avatud isiksuse kujunemist. (VKS: 8.)

Haridus- ja teadusminister Mailis Reps algatas 2005. aastal võõrkeelte strateegia koostamise ning kinnitas seda ettevalmistava komisjoni koosseisu ning komisjoni sisemise töökorra. Komisjoni põhiülesanneteks määrati senise võõrkeelte poliitika analüüsimine ning võõrkeelte strateegia väljatöötamine.

Valminud käsikirjas “Eesti võõrkeelte strateegia aastani 2015” ongi vaatluse all võõrkeelte õppe küsimused üld-, kutse- ja kõrghariduses, samuti täiskasvanuhariduses ning erinevates mitteformaalse hariduse vormides. “Elukestva õppe süsteemi kuulub nii formaalharidus kui ka sellele järgnevad õpingud. Käsitlemist leiavad ka keeleoskuse tunnustamise ja kvaliteedi tagamise küsimused. Võõrkeelte oskuse ja kasutamisega seotud küsimusi vaadeldakse ühiskonna sotsiaalse sidususe, üksikisikute mobiilsuse ning tööhõive seisukohalt. Samuti käsitletakse võõrkeelte temaatikat seoses riigi funktsioonide täitmise ning rahvusvahelise suhtluse eri tasanditega. Võõrkeelte strateegias analüüsitakse Eesti senist võõrkeelteõppe korraldust, õiguslikku seisundit, võõrkeelte alaseid uuringuid ning sõnastatakse võõrkeelte poliitika põhimõtted ning visioon aastaks 2015.” (VKS: 5.)

Võõrkeelte strateegias lähtutakse ennekõike keelenõudlusest ja tuginetakse Eesti põhiseaduses väljendatud aadetele.

Võõrkeelte strateegia sihtrühmad

Eri inimrühmadel läheb keeli vaja erineval määral, millel põhineb võõrkeelte strateegia sihtrühmade eristamine. Inimeste rühmitamine mistahes moel on paratamatult suhteline, rühmade määratlemine vaieldav, teinekord liigselt hinnanguline. Siin kasutatud rühmitus kajastab artikli autori nägemust võimalikest sihtrühmadest.

1. Eesti eliit (diplomaadid, Eesti Vabariigi Euroopa Liidu ametnikud, poliitikud, tippteadlased, kultuuritegelased jt), kes riigikeele kõrval võiks vabalt kõnes ja kirjas osata veel näiteks viit-kuut Euroopa Liidus enim kõneldavat keelt ja kelle töö on esindada Eesti Vabariiki Euroopa Liidu eri foorumitel. Tinglikult siis rühm riigikeel pluss viis-kuus Euroopa Liidus enim kõneldavat keelt. Eesti osaleb paljudes rahvusvahelistes organisatsioonides, mille eelduseks on riigiametnike, asjatundjate, tõlkide ja tõlkijate hea võõrkeelteoskus. Euroopa Liidu õiguse rakendamine nõuab näiteks head võõrkeelteoskust paljudelt õigus- ja teistelt erialaspetsialistidelt (VKS: 8). Hetkel on nende järgi suur puudus, selliste inimeste absoluutarv Eestis võiks hetkel olla umbes 10 000–14 000 inimest (umbes 1% elanikest).
2. Eesti kõrgtasemel haritlased (teadlased, pedagoogid, arstid jt), kes lisaks riigikeelele oskaksid vabalt kõnes ja kirjas kuni viit Euroopa Liidu või meie lähemate naabrite keelt. Inglise keele hädavajalik oskus on sel tasemel vajalik erialase kirjanduse lugemisel ning teadusartiklite kirjutamisel, samuti välismaa kolleegidega suhtlemisel. Nende inimeste absoluutarv Eestis võiks olla umbes 20 000–25 000 (umbes 2% elanikest).
3. Ülejäänud kõrgharidusega inimesed, kes lisaks riigikeelele oskaksid vabalt kõnes ja kirjas inglise keelt ning lisaks vene ja soome keelt Põhja-Eestis, Lõuna-Eestis näiteks läti keelt (umbes 5% elanikest).
4. Gümnaasiumiharidusega inimesed, kes lisaks riigikeelele oskaksid vabalt kõnes ja kirjas inglise keelt ning mitte-eestlased ka oma esivanemate keelt.
5. Põhikooliharidusega inimesed, kes lisaks riigikeelele oskaksid elementaarselt kõnelda inglise keeles.

Eri põhjuseil (tervis, motivatsioonipuudus, eakad pagulased ja immigrandid, jms) tuleb ilmselt leppida, et umbkaudu 6% ehk umbes 85 000 21. sajandi keskpaiga Eesti elanikku kõneleb vaid üht keelt. Seejuures on needki ilmselt kümnekond väga erinevat keelt, mis toob täiendavaid sotsiaalseid probleeme nende lõimumiseks Eesti ühiskonda ja kultuuri.

Kokkuvõte

Kogu elanikkonna piisav võõrkeelte valdamise tase ja selle suhestamine riigikeelega on Eesti tuleviku, kodurahu ja majandusedukuse jaoks strateegiliselt sedavõrd oluline, et eeldaks kindlasti riiklikku süsteemset korraldatust. Põhjalikult läbiarutatud võõrkeelte strateegia ühitatuna eesti keele jätkusuutlikkuse edendamiseks Eestis aitaks ühemõtteliselt kaasa, et kõik Eesti elanikud (rahvusest olenemata) muutuksid üha enam mitmekeelsemateks, säilitades seejuures ometi eesti keele riigikeelena.

Keelepoliitika paradigma on Eesti liitmisel Euroopa Liidu ning NATO-ga oluliselt teisenenud, mistõttu tuleb spetsialistide poolt juba ettevalmistatud esialgne strateegia versioon kiiremas korras riigi uusi olusid arvesse võttes läbi vaielda ja sõnastada see uue pikemaajalise keelestrateegiana.

Kui vaid suudaksime ühendada kõik jõud Eestis, kes oskaksid osaleda ja tahaksid jõuda tulemuseni ning kui ei tekiks eesti inimesele omaseid poliitilisi takistusi võõrkeele strateegia kinnitamisele ja läbiviimisele.

Kirjandus

- Bertriveau, Adrian; Luik, Jüri; Tiido, Harri 2001. Globaliseerumine, integratsioon Euroopa struktuuridesse ja selle mõju rahvuslikule identiteedile ning kultuurile. – A. Bertriveau (koost.), L. Listra (toim.). Eesti identiteet ja iseseisvus. Tallinn: Avita, 279–288.
- Byram, Michael; Fleming, Michael (eds.) 1998. Language Learning in Intercultural Perspective: Approaches through Drama and Ethnography. Cambridge Language Teaching Library. Cambridge: Cambridge University Press.
- Clyne, Michael; Rossi-Hunt, Claudia; Isaakidis, Tina 2004. Learning a Community Language as a third Language. – The International Journal of Multilingualism 1 (1), 33–52.
- Cook, Vivian (ed.) 2002. Portraits of the L2 User. Clevedon: Multilingual Matters. European Day of Languages (26.9.01.). Press Centre.
- EE = Eesti Entsüklopeedia 2007, 15 kd.
- EHIS = www.ehis.ee (7.03.2008)
- Eesti keele arendamise strateegia 2004–2010. Tartu: Haridus- ja Teadusministeerium, 2004.
- Viikberg, Jüri (koost. ja toim.) 1999. Eesti rahvaste raamat: rahvusvähemused, -rühmad ja -killud. Tallinn: Eesti Entsüklopeediakirjastus.
- Rannut, Mart; Rannut, Ülle; Verschik, Anna 2003. Keel, võim, ühiskond: sotsiolingvistika ja keelepoliitika õpik. Tallinn: Tallinna Pedagoogikaülikooli Kirjastus.
- Liiv, Suliko 1998. Eesti võõrkeeleturnust. – Keel ja Kirjandus 9, 639–640.
- Muukeelse elanikkonna keeleõppestrateegia 1998. Tallinn: Eesti Keelestrateegia Keskus.
- Vihalemm, Triin; Masso, Anu; Vihalemm, Peeter 2004. Eesti kujunev keeleruum. – V. Kalmus, M. Lauristin, P. Pruulmann-Vengerfeldt (toim.). Eesti elavik 21. sajandi algul: ülevaade uurimuse *Mina. Maailm. Meedia* tulemustest. Tartu: Tartu Ülikooli Kirjastus, 57–74.
- VKS = Eesti võõrkeelte startegia aastani 2015. (Käsikiri Haridus- ja Teadusministeeriumis.)
- Williams, Kevin 2004. Modern foreign languages. – J. White (ed). Rethinking the School Curriculum: Values, Aims and Purposes. London: Routledge/Falmer, 117–127.

Suliko Liiv (Tallinna Ülikool) on lõpetanud Tartu Ülikooli germaani-romaani filoloogi, inglise keele õpetaja erialal. Alates 1993. a töötab Tallinna Ülikooli inglise keele professorina ja inglise keele õppetooli juhatajana, on täitnud ka filoloogiateaduskonna dekaani ja germaani-romaani filoloogia osakonna juhataja kohuseid. Alates 2008. a on Tallinna Ülikooli germaani-romaani keelte ja kultuuride instituudi direktor. Uurimisvaldkonnad: kontrastiivuuringud, kultuuridevaheline suhtluspädevus, keelepoliitika, võõrkeelte õpetamise meetodika.

liiv@tlu.ee

ESTONIAN AND OTHER LANGUAGES: FOR ESTONIA

Suliko Liiv

Tallinn University

The language situation in Estonia as a member of multicultural and multilingual European Union has changed in comparison with last century and Estonia needs a clear strategy in order to teach foreign languages to its citizens working here in Estonia as well as in other European Union states.

The current article gives an overview of the needs to teach and learn foreign languages in Estonia. The article also raises problems and offers solutions for Estonian foreign language strategy plan. A well-discussed foreign languages strategy together with promoting the sustainability of Estonian in Estonia would unequivocally lead to a situation where all residents of Estonia would become increasingly multilingual and the Estonian language would still be preserved as the official language.

Keywords: multilingualism, foreign language strategy, language policy, Estonian

Trükitud 150 eksemplari