

Eesti Rakenduslingvistika Ühing

Eesti Rakenduslingvistika Ühingu aastaraamat 3

Estonian Papers in Applied Linguistics 3

Toimetanud
Helle Metslang, Margit Langemets
ja Maria-Maren Sepper

Eesti Keele Sihtasutus
Tallinn 2007

EESTI RAKENDUSLINGVISTIKA ÜHINGU AASTARAAMAT 3

Estonian Papers in Applied Linguistics 3

Toimetajad / Editors: Helle Metslang (Helsinki/Tallinn), Margit Langemets (Tallinn),
Maria-Maren Sepper (Tallinn)
Kujundaja / Design: Piia Ruber

Toimetuskolleegium / Advisory board:

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

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

MLA Directory of Periodicals
MLA International Bibliography
Linguistics and Language Behavior Abstracts (LLBA)
Linguistic Bibliography Online (BL Online)
Directory of Open Access Journals, Lund University (DOAJ)
Eesti Rahvusraamatukogu digitaalne arhiiv (DIGAR)
Central and Eastern European Online Library (C.E.E.O.L)

Ühingu aadress / Contact information:

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

ery@eki.ee
www.rakenduslingvistika.ee

“Eesti Rakenduslingvistika Ühingu aastaraamat 3” väljaandmist on toetanud Eesti Keele Instituut ja Tallinna Ülikool.

Autoriõigus/Copyright: autorid ja Eesti Rakenduslingvistika Ühing

ISSN 1736-2563

SISUKORD

Kultuuri mõju eestlaste ja leedulaste ruumikasutusele <i>Cultural influence on personal space usage among Estonians and Lithuanians</i> Rene Altrov	7
<i>Variautor</i> – variisik või varjukirjanik? Variautor – <i>Strohmann oder Ghostwriter?</i> Anne Arold	35
Parental reactions to the morphologically correct and incorrect utterances of children <i>Vanema reaktsioonid lapse morfoloogiliselt korrektsetele ja ebakorreksetele lausungitele</i> Ingrida Balčiūnienė	47
Shift of use of capital (upper case) letters in english and Latvian marketing texts <i>Suurtähtede kasutuse muutus inglise ja läti turundustekstides</i> Eva Birzniece	57
Apologies in business communication <i>Vabandused töösuhtluses</i> Laura Čubajevaitė, Jūratė Ruzaitė	67
Suhtlusstrateegiad ja -taktikad müügivestlustes <i>Communicative strategies and tactics in telemarketing calls</i> Liina Eskor	83
Õppijakeel ja eesti vahekeele korpus <i>Learner language and Estonian Interlanguage Corpus</i> Pille Eslon, Helena Metslang	99
Machine learning techniques in dialogue act recognition <i>Masinõppe tehnikad dialoogiaktide tuvastamises</i> Mark Fišel	117

Kliendi soovide automaatne tuvastamine eestikeelsetes infodialoogides <i>Automatic recognition of customer requests in Estonian information dialogues</i>	135
Olga Gerassimenko, Tiit Hennoste, Riina Kasterpalu, Mare Koit, Andriela Rääbis, Krista Strandson, Maret Valdisoo, Evely Vutt	
Morphological modifications in Lithuanian child-directed speech <i>Leedu hoidjakeele morfoloogilised modifikatsioonid</i>	155
Laura Kamandulyte	
Võimu keel: asutus suhtleb kodanikuga <i>Language of power</i>	167
Reet Kasik	
Ametikeele oskuse mõõtmine: tekst ja ülesanne eesti keele kui teise keele kuulamistestis <i>Measuring specific purpose language ability: texts and tasks in Estonian as a second language listening tests</i>	177
Tiina Kikerpill, Ülle Türk	
Arvuti suhtluses <i>Interaction with the computer</i>	193
Mare Koit	
Rohkem kui sõnad <i>More than words</i>	211
Jaakko Lehtonen	
Kuidas hinnata suure panusega testide hindajaid <i>How to assess the raters of high-stakes tests</i>	221
Hille Pajupuu	
Keeleõppijate vältimisstrateegiatest eesti keele täis- ja osasihitise näitel <i>Learner strategies to avoid the use of the total and partial objects in Estonian</i>	235
Raili Pool	
Ladina keelest pärit terminite polüseemia ja sünonüümia eesti õiguskeeles <i>Polysemy and synonymy of Latin terms in Estonian legal language</i>	253
Merike Ristikivi	
Võimalusi helistaja identifitseerimiseks institutsionaalses telefonisuhtluses <i>Caller identification in institutional telephone conversation</i>	269
Andriela Rääbis	

The role of a child's gender in language acquisition <i>Lapse soo mõju keele omandamisel</i> Ineta Savickienė, Violeta Kalėdaitė	285
Õpetaja tagasiside õpilase vastusele algkoolitunnis <i>Teacher's feedback to students' answers in primary school lessons</i> Krista Strandson	299
Vene-eesti koodivahetuse korpus: kodeerimispõhimõtete väljatöötamine <i>Russian-Estonian code-switching corpus: elaboration of encoding principles</i> Anastassia Zabrodskaja	321
Emotsioonisõnad teises keskkonnas – millest kõnelevad loetelukatsete tulemused? <i>Emotion terms at work – what do the results of tasks of free listing reveal about?</i> Ene Vainik	339
Kakskeelsete leksikonide genereerimine paralleelkorpuse baasil <i>Generation of bilingual lexicons from a parallel corpus</i> Kaarel Veski	355
Pronoomenid lapsekeeles: mõnda <i>mina</i> ja <i>sina</i> omandamisest <i>Pronouns in child language: some fragments of a study of the acquisition of the pronouns mina and sina by Estonian children</i> Maigi Vija	373
RINGVAADE Tavakeelest ja tavamatemaatikast Leo Võhandu	385

KULTUURI MÕJU EESTLASTE JA LEEDULASTE RUUMIKASUTUSELE

Rene Altrov

Ülevaade. Artikkel tugineb uurimusele, mis vaatlleb kultuuri mõju ruumikasutusele erinevates suhtlusolukordades: kodus, töökeskkonnas ning väljaspool kodu ja tööd aega veetes. Uuringu eesmärk oli välja selgitada, millised on eestlaste ja leedulaste distantsikasutamise reeglid, kas need erinevad olenevalt suhtlussituatsioonist, kas eestlaste ruumikäitumine on ajas muutunud ning kas ilmneb erinevusi eestlaste ja leedulaste ruumikasutuses.

Andmeid koguti küsitluse teel, kus olid ette antud situatsioonikirjeldused nelja võimaliku käitumisvariandiga. Vastanute valikuid analüüsid tugineti G. Hofstede ja E. T. Halli kultuuriuringutele. Uuritavad olid eesti ja leedu rahvusest kontoritöötajad. Eestlaste kohta koguti andmeid kahel korral – 1998. ja 2006. aastal. Leedulaste uurimus viidi läbi 2006. aastal.

Uurimus näitas, et väärtused ja eelistused võivad erineda sõltuvalt suhtlusolukorrast ja -keskkonnast – peresuhtes võivad esile tulla ühed kultuurinormid, töösuhetes ja vaba aega veetes teised. Võrreldes eestlaste 1998. ja 2006. aasta küsitluse andmeid, iseloomustavad eestlaste ruumikäitumist valdavalt individualistlikule ja maskuliinsele kultuurile omased väärtused, kuid võib näha nihet maskuliinsetelt väärtustelt feminiinsete suunas. Leedulaste puhul ilmnas, et nende ruumikäitumine on mitmekesisem kui eestlaste oma, olenevalt suhtlussituatsioonist tulid esile ka mitmed vastandlikud väärtused.

Võtmesõnad: prokseeimika, ruumikäitumine, ruumikasutus, distants, eesti keel, leedu keel

Sissejuhatus

Kommunikatsiooni hulka kuulub muuhulgas ka ruumikäitumine, ehk see, kuidas kasutame ruumi enda ümber ja missugust distantsi vestluspartneriga hoiame. Ruumisuhted ja territooriumi piirid mõjutavad otseselt meie igapäevaelu, kuigi teadvustame seda endale alles siis, kui meie territooriumi piire rikutakse. Kontroll oma territooriumi piiride üle on aga üks rahulolu tingimusi, mistõttu inimesed tahavad seda kaitsta ja ka käituvad vastavalt (Gifford 1997: 96, Preston 2005: 151). Ruum, mida inimene enda ümber vajab, on kindlate piiridega ja peegeldab tema identiteeti, vanust, sugu ja isikupära. Samuti kajastab distantsihoidmine suhtlejate sotsiaalseid ja personaalseid vahekordi, tundeid, omandit ning võimupositsiooni (Adler, Elmhurst 1992: 81–82, Gifford 1997: 101–105, Gudykunst, Ting-Toomey 1988: 124).

Eri kultuurides on omad ruumireeglid. Kultuuriti varieerub, kuidas tajutakse avarust, ülerahvastatust, distantsinormide rikkumist ja nendega arvestamist (Gudykunst, Ting-Toomey 1988: 124).

Paul Prestoni väitel on ruumikasutus inimestevahelises suhtluses üks vähem uuritud valdkondi (2005: 151). Seetõttu on antropoloog Edward T. Halli 1960. aastate distantsiuurimused (vt Hall 1990) tänapäevalgi aluseks ruumi ja ruumisuhteid uurivatele teadlastele. E. T. Hallilt pärineb ka termin *prokseemika* (Hall 1990: 1), mis tähendab ruumi kommunikatsiooniliste funktsioonide uurimist ehk seda, kuidas inimesed liigendavad ruumi suhtlusdistantsideks; kuidas korraldada ruumi kodus, töökohas jne; kuidas planeerida linnu ning nendevahelisi maa-alasid.

Inimestevahelises suhtluses on E. T. Hall eristanud nelja peamist distantsitüüpi: intiimne, isiklik, sotsiaalne ja avalik ruumivöönd (Hall 1990: 116–125). Neid distantsse hoiab inimene sõltuvalt suhtlusolukorrast iseenda ja teiste vahel. See aga, millist distantsi mingis olukorras kasutatakse, sõltub suuresti kultuurist. Näiteks distants, mis ühes kultuuris on intiimse tähendusega, võib teises kultuuris kuuluda sotsiaalse suhtluse juurde (nt suudlemine) (Hall 1990: 118).

Ruumikasutus ja inimestevahelise distantsi suurus seostub otseselt kultuuriväärtustega (Samovar, Porter 2004: 190). Kõnesolevas uurimuses lähtuti Geert Hofstede (1980) kultuuridimensioonidest: individualism-kollektivism, maskuliinsus-feminiinsus, ebakindluse vältimise tarve ning võimudistants. Nende dimensioonidega on võimalik kirjeldada kõiki kultuure ja ühtlasi seletada ning prognoosida ka sarnasusi ja erinevusi kultuuridevahelises suhtluses (Gudykunst jt 1996: 21).

Individualismi-kollektivismi dimensioon iseloomustab rühma ja indiviidi vahelisi suhteid. Individualistlikes kultuurides on esmatähtsad üksikisiku huvid, soodustatakse eneseteostust ja eeldatakse, et inimesed hoolitsevad vaid iseenda ja oma lähedaste eest. Kollektivistlikes kultuurides seatakse esmatähtsaks rühma huvid. Rühm küll hoolib oma liikmetest, kuid vastutasuks nõuab lojaalsust (Gudykunst jt 1996: 22, Hofstede, Bond 1984: 419, Hofstede 2001: 226). Maskuliinsusdimensioon näitab maskuliinsete ja feminiinsete väärtuste kaalu ühiskonnas. Maskuliinsetes ühiskondades on soorollid kindlalt erinevad, tähtis on sõltumatus, ollakse orienteeritud rahale ja asjadele, tähtis on võim ja edasiliikumine. Feminiinsetes kultuurides pole soorollid fikseeritud, oluline on keskkond, elulaad ning suhted teiste inimestega (Gudykunst jt 1996: 47, Hofstede 2001: 297–299). Võimudistantsi dimensioon näitab suhteid erineva ühiskondliku ning varandusliku seisusega inimeste vahel:

kuivõrd väikese võimuga institutsioonide ja organisatsioonide liikmed aktsepteerivad võimu ebavõrdset jaotumist ja selle autoritaarset kasutamist (Hofstede, Bond 1984: 419, Hofstede 2001: 98). Ebakindluse vältimise dimensioon vaatlleb, kui ohtlikuks peab ühiskond ebakindlaid ja tundmatuid olukordi ning mil määral ta püüab neid kontrollida. Suure ebakindluse vältimise tarbega kultuurides arvatakse, et “kõik erinev on ohtlik”, väikese ebakindluse vältimise tarbega kultuurides valitseb veendumus, et “kõik erinev on imelik” (Gudykunst jt 1996: 42, Hofstede 2001: 161).

Kuigi G. Hofstede algne uurimus (1980) ei hõlmanud Ida-Euroopa riike, on hilisemate uuringute tulemusel esitatud andmeid ka Eesti (Hofstede 2001) ja Leedu (Mockaitis 2002) kohta. Eesti ja Leedu kultuuridimensioonide näitajad on tabelis 1.

Tabel 1. Kultuuridimensioonide indeksid Eesti ja Leedu kohta (Mockaitis 2005: 46)

	Individualism	Võimudistants	Ebakindluse tõrjumine	Maskuliinsus
Leedu	50	45	67	65
Eesti	60	40	60	30
Hofstede (2001) andmebaasi kõrgeim indeks	91	104	112	110
Hofstede (2001) andmebaasi madalaim indeks	6	11	8	5

Tabelis 1 esitatud andmete põhjal võib eeldada, et eestlastega võrreldes on leedulased märgatavalt kollektivistlikumad ja maskuliinsemad ning mõnevõrra suurema võimudistantsiga ja ebakindluse vältimise tarbega. Samas pole teada, kas eesti ja leedu erinevad väärtushinnangud tähendavad ühtlasi erinevusi vastavate maade vestluskäitumises. Käesolevas uurimuses soovitaksegi E. T. Halli ja G. Hofstede uurimustele toetudes teada saada:

- 1) milline on eestlaste ruumikäitumine ja kas see on kaheksa aasta jooksul muutunud;
- 2) milline on leedulaste ruumikäitumine võrreldes eestlastega.

Materjal ja meetod

Uuring koosnes kahest etapist. Eestlaste kohta viidi esimene ruumiuurimus läbi 1998. aastal (Altrov 1999). Tol korral oli küsitletavaid 40. Teine eestlaste ruumiuurimus viidi läbi 2006. aastal. Seekord oli küsitletavaid 30. Leedulasi küsitleti samuti 2006. aastal – vastanuid oli 27. Kõik uuringus osalejad olid kontoritöötajad, nende vanus oli 25–45 eluaastat ning neil oli vähemalt viieaastane töökogemus. Nii eesti kui ka leedu rahvusest küsitletavate seas oli naised ja mehed enam-vähem võrdselt. Naiste ja meeste tulemuste eraldi kajastamist ei peeta uuringus vajalikuks, kuna märkimisväärset erinevust vastustest ei ilmnenud.

Eestlaste ja leedulaste ruumikäitumist uuriti küsitluse teel. Küsitlus koosnes 12 valikvastusega situatsioonikirjeldusest. Küsitleva ülesanne oli neljast vastu-

sevariandist valida üks tema jaoks sobivaim. Situatsiooniküsimused olid jagatud kolme rühma: kodune keskkond, töökeskkond ning suhted väljaspool kodu ja tööd. Nii situatsioonikirjeldused kui ka nende vastusevariandid tuginesid E. T. Halli (1990) ja G. Hofstede (1980) kultuuriuuringutele, autori isiklikule kogemusele ja vaatlustele.

Küsitlusvorm töötati välja 1998. aastal. Küsitlust muuta ja kaasajastada ei peetud 1998. aasta ja 2006. aasta andmete võrreldavuse huvides otstarbekaks.

Küsitlus: situatsioonid, variandid ja tulemused

Küsitlus koosnes situatsiooniküsimustest, mis olid jagatud kolme rühma: sooviti teada, millised on eestlaste ja leedulaste ruumi- ja distantsieelistused kodus keskkonnas, töösuhetes ning väljaspool kodu ja tööd aega veetes ning millistest kultuuritunnustest antud suhtlusolukorrad johtuvad.

Koduga seotud küsimused

Kodust keskkonda uuriti situatsioonide kaupa. Sooviti teada saada, kas küsitletavad soovivad privaatsust ka oma pereliikmetega suheldes, kas peetakse tähtsaks eraldatust naabritest, millisel määral väärtustatakse elukeskkonda ja kas kodus eelistatakse kindlaid distantsinorme, millega rõhutatakse suhtlemise formaalsust, või on suhtlemine vaba ja inimesekeskne.

Esimene situatsioon

Esimese situatsiooniküsimusega uuriti, kui oluline on küsitletava jaoks privaatsus oma perekonnas, naabritest eraldatuse soov, või vastupidi, seotus kaasinimestega.

Ühel hetkel märkad, et Sinu pereliikmete arv on tublisti kasvanud. Otsustad oma väikese äärelinna korteri parema vastu vahetada. Mille soetad?

- A. *Väikese maja äärelinnas (püüame siiski ära mahtuda), kuid suure iluaiaga.*
- B. *Suure maja äärelinnas (kõik saavad oma toa), kuid aed on väike ja päikesevaene.*
- C. *Suurema äärelinna korteri (läheb pisut lähedamaks) ilusa põlispargi ääres.*
- D. *Samaväärse korteri, kuid kesklinnas. Leiad, et kesklinnas on paremad võimalused ajaveetmiseks väljaspool kodu: käia huvialaringides, kinos, teatris, olla sõprade seltsis kohvikutes/baarides.*

A-variandi valijatele on elukeskkond väga oluline. Kõnesolev variant võiks iseloomustada feminiinseid, kollektivistlikke kultuure. Feminiinsed ühiskonnad väärtustavad elukvaliteeti, hoolitakse teistest inimestest, suurt tähelepanu pööratakse elukeskkonnale (Hofstede 2001: 298, 306). Kollektivistlikke kultuure iseloomustavad

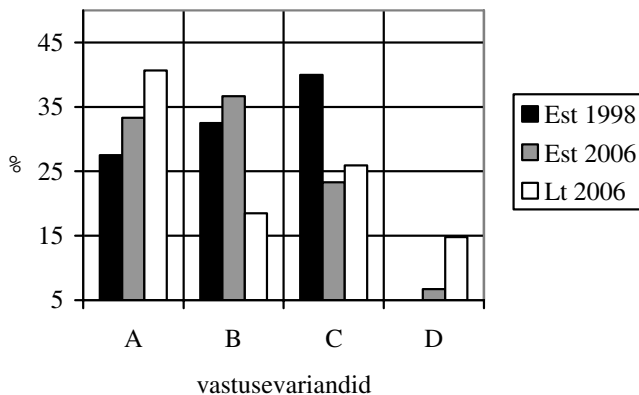
tugevad peresidemed (Hofstede 2001: 236), omaette olekul ei ole sellist tähtsust nagu individualistlikes kultuurides (Gudykunst, Ting-Toomey 1988: 123).

B-variandi valijatele on tähtis omaette olemise võimalus. Selline variant võiks iseloomustada maskuliinseid, individualistlikke, distantsi hoidvaid kultuure. Maskuliinsetes kultuurides rõhutatakse enda ego. Perekonnda peetakse küll oluliseks, kuid elukeskkonda ei väärtustata. Kui feminiinsetes kultuurides ollakse orienteeritud väikestele asjadele, siis maskuliinsetes kultuurides hinnatakse suuri silmapaistvaid asju (Hofstede 2001: 298–299, 306).

C-variandi valijatele on elukeskkond olulisem kui privaatsus. Selline variant võiks iseloomustada feminiinseid, kollektivistlikke kultuure. Feminiinsetes kultuurides on olulised suhted teiste inimestega, sh võõrastega (Hofstede 2001: 299). Samuti hinnatakse feminiinsetes kultuurides kõrgelt elukeskkonda ja -kvaliteeti (Hofstede 2001: 298–299). Kollektivistlikkusele viitab asjaolu, et puudub privaatsuse nõue (Gudykunst, Ting-Toomey 1988: 123).

D-variandi valijad korvavad kodust ruumikitsikust seltsieluga väljaspool kodu. Antud variant võib viidata individualistlikule ja maskuliinsele kultuurile. Individualistlikes ja maskuliinsetes kultuurides tähtsustatakse peamiselt iseenda ja oma pere vajadusi ning kodu ongi eeskätt oma perekonna jaoks (Hofstede 2001: 227, 306). Maskuliinsetele väärtustele viitab ka asjaolu, et eluaseme valikul peetakse oluliseks selle prestiižset asukohta (Hofstede 2001: 312).

Esimese situatsiooni vastused on joonisel 1.



Joonis 1. Esimese situatsiooniküsimuse vastused

1998. aastal toetusid eestlaste eelistused peamiselt feminiinsetele kollektivistlikele väärtustele (A + C): suhteliselt väike ruumivajadus, ilus elukeskkond. Ent esile tõuseb ka vastaspoolus, kuigi mitte domineerivana: suur maja, pereliikmete eraldumisvõimalus, ükskõiksus keskkonna suhtes (B), näidates individualistlike ja maskuliinsete väärtuste olemasolu kultuuris.

2006. aastaks polariseeruvad väärtushinnangud veelgi: suurenenud on nii suure maja eelistajate (B) kui ka keskkonna eelistajate hulk (A + C).

Võrreldes leedulasi eestlastega, on näha leedulaste elukeskkonna eelistus ning pühendumine inimsuhetele (A + C), seega tuginevad leedulaste eelistused feminiinsetele kollektivistlikele väärtustele.

Suhtlemisvõimalus linnakeskkonnas (D) jäi kõige vähem valitud variandiks nii eestlastel kui ka leedulastel.

Teine situatsioon

Teise situatsiooniga sooviti teada saada, kui suur on vastanute isikliku ruumi vajadus ja kas soovitakse eralduda võõrastest või ka oma pere liikmetest.

Ostsid endale just uue eluaseme. Mida seal esmajoones ette võtad?

- A. Jagad ruumid pereliikmete vahel ära.
- B. Paned akna ette läbipaistmatud kardinad, et naabrid sisse ei näeks.
- C. Lähed tutvud naabrite ning ümbruskonnaga.
- D. Lased osa vaheseinu maha lõhkuda, et tekiks avatud ruum.

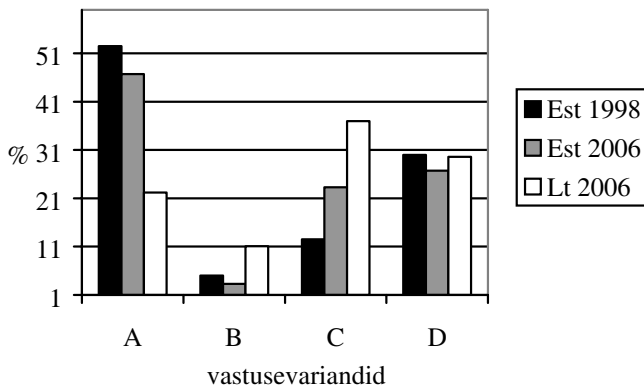
A-variandi valijaid iseloomustab tugev individualism. Soovitakse eralduda ka oma perekonnas. Individualistlikes kultuurides on privaatsusel suur tähtsus (Gudykunst, Ting-Toomey 1988: 121, 123).

B-variandi valijatele on võõrastest eraldumine olulisem kui distants omas perekonnas. Selline suhtumine võiks olla ebakindlust halvasti taluva kultuuri peegeldus, kuna ebakindlust halvasti taluvates kultuurides hoitakse võõraga distantsi (Hofstede 2001: 161, 180).

C-variandi valijatele on tähtsad elukeskkond ning inimsuhted. Selline variant võiks iseloomustada feminiinseid, kollektivistlikke ja ebakindlust hästi taluvaid kultuure. Feminiinsetes kultuurides on oluline elukeskkond, hoolitakse teistest inimestest ja inimsuhetest (Hofstede 2001: 298–299, 306). Kollektivistlikes kultuurides tähtsustatakse inimsuhteid, sotsiaalset võrgustikku ning naabruskonda (Hofstede 2001: 226–227). Ebakindlust taluvates kultuurides pööratakse tähelepanu suhetele ja tuntakse huvi võõra vastu (Hofstede 2001: 161, 170).

D-variandi valijaid iseloomustab kollektivistlikkus. Kollektivistlikus kultuuris on tähtsad suhted kaasinimestega, meie-tunne ning distantsihoidmist ei peeta oluliseks (Hofstede 2001: 226–227, 236, Gudykunst, Ting-Toomey 1988: 123).

Küsimusele antud vastused on esitatud joonisel 2.



Joonis 2. Teise situatsiooniküsimuse vastused

Eestlased on mõlemas küsitluses eelistanud A-varianti: jagada uude korterisse või majja elama asudes esmalt toad pereliikmete vahel ära. See on märk individualistlike väärtuste domineerimisest. Ent eestlastest eelistavad suhteliselt paljud ka avatud ruumi (D), mis soosib pereliikmete koosolemist ja omavahelist suhtlemist, kandes seega kollektivistlikke väärtusi.

Võrreldes esimese küsitlusega on märkimisväärselt tõusnud ka C-variandi eelistajate arv, kes uude elukohta asudes tutvuvad esmajoones sealse elukeskkonna ning naabritega. Nii nagu D-variant viitab see kollektivistlikele väärtustele, ent elukeskkonna tähtsustamine tõstab esile ka feminiinsed väärtused.

Kui C- ja D-variant kokku võtta, siis võib öelda, et individualistlike ja kollektivistlike väärtuste kandjaid on enam-vähem võrdselt, nagu ka eelmise situatsiooni puhul.

Võrreldes leedulasi eestlastega, on näha leedulaste ülekaalukas C- ja D-variandi eelistus. Seega iseloomustab leedulasi enim teistest inimestest ning keskkonnast hoolimine, samuti ka uudishimu võõra vastu.

Kõige vähem valituks osutus nii eestlaste kui ka leedulaste puhul B-variant – eraldumine kardinat taha. Sellest võib järeldada, et ebakindlamaid olukordi koduses ruumikasutuses talutakse mõlemas kultuuris suhteliselt hästi.

Kolmas situatsioon

Selle situatsiooniküsimusega sooviti teada saada, kas vastanud seavad ruumikasutuses iseenda ja teiste vahele kindlad piirid ning kuidas suhtutakse ruumikasutusreeglite rikkumisse.

Kuidas sisustad elutoa?

- A. *Kerge konstruktsiooniga hõlpsasti liigutatava mööbliga.*
- B. *Massiivse kalli mööbliga, mille sead paika aastateks.*
- C. *Nimeka disaineri loodud mööbliga.*
- D. *Kuidas juhtub, ja mis mööbel käepärast on.*

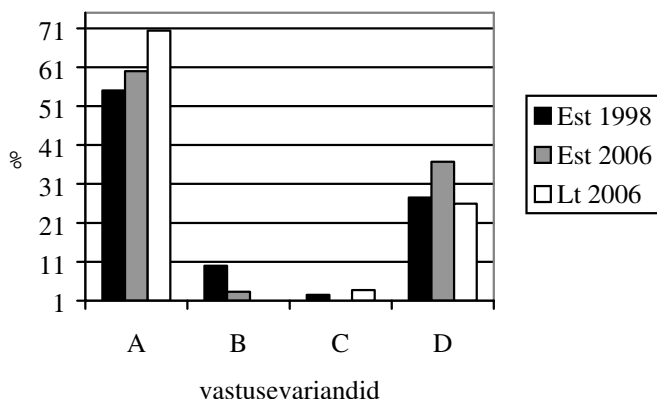
A-variandi valijatele ei ole stabiilsus kuigi oluline. Ruumisviibijatel lastakse distantsi reguleerida oma soovi ja tujude kohaselt. Selline tunnus on iseloomulik ebakindlust hästi taluvatele kultuuridele, kus käitumist määravad vähesed reeglid ning sallitakse ka muutusi ja ebaselgust ning ollakse orienteeritud inimsuhetele (Hofstede 2001: 161, 170). Antud vastusevariant viitab ka feminiinse kultuuri joontele, kuna hinnatakse inimestekesksust ja vastastikuseid suhteid (Gudykunst jt 1996: 47, Hofstede 2001: 299).

B-variandi valijad hindavad stabiilsust. Stabiilsusega soovitakse tõrjuda ebakindlust. Ruumi sisustaja kehtestab distantsireeglid (paneab toolide kaugused paika, valib laua suuruse jne). Ebakindlust halvasti taluvates kultuurides tekitavad ebakindlad olukorrad ohutunde, mistõttu neis on palju käitumist suunavaid reegleid ja teistsugune käitumine ei ole lubatud (Gudykunst jt 1996: 42, Hofstede 2001: 166, 169). Vastusevariant viitab ka maskuliinse kultuuri tunnustele, kuna maskuliinsetes kultuurides eelistatakse suuri asju ja pereliikmete erivajadustele tähelepanu ei pöörata (Hofstede 2001: 299, 306, 312).

C-variandi valijad ei kasuta distantsireguleerijana mööblit. Mööblil on nende jaoks muud funktsioonid. Ruumikorraldusest olulisem on mööbli väline hiilgus ja positsiooninäitamine. Iseloomulik on niisugune valik suure võimudistantsiga maskuliinsele kultuurile, kus pereliikmete vajadused jäetakse tagaplaanile ja tähtsad on raha ning asjad (Hofstede 2001: 299, 306).

D-variandi valijad ei kasuta või ei oska kasutada mööblit distantsireguleerijana. Mööblil pole muid funktsioone kui vaid olla mööbel.

Küsimusele antud vastused on esitatud joonisel 3.



Joonis 3. Kolmanda situatsiooniküsimuse vastused

Nii 1998. kui ka 2006. aasta eestlaste ning leedulaste vastustes domineeris kerge ja hõlpsasti liigutatava mööbli eelistus (A), mis näitab, et peresuhetes ei soovita ruumisviibijatele fikseeritud distantse peale suruda. Seega tulid esile feminiinsele ja eba-kindlust hästi taluvalle kultuurile omased jooned: kodus keskendutakse inimestele, tähtsustatakse vastastikuseid suhteid ja suhtlemise mugavust ning ollakse valmis muutusteks. Suhteliselt suur hulk vastanuid aga ei pööra mööblivalikule tähelepanu (D), mis näitab, et ei teadvustata mööbli võimalusi suhtluse reguleerijana.

Töoga seotud küsimused

Töösuhteid uuriti situatsioonide kaupa. Sooviti teada saada, kas töökeskkonnas tähtsustatakse füüsilist privaatsust või saadakse privaatsustunne muul viisil, näiteks pilgukontakti vältimisest ja rääkimise lõpetamisest. Ühtlasi sooviti välja selgitada, kas vastanud hindavad töökeskkonnas inimsuhteid või peavad nendest olulisemaks tööülesandeid. Samuti taheti teada saada, kuidas küsitletavad tõlgendavad puudutamist – kas puudutamine kuulub vestluskäitumise juurde või peetakse seda distantsinormide rikkumiseks.

Neljas situatsioon

Selle situatsiooniküsimusega uuriti, kas vastanud eelistavad visuaalset või füüsilist privaatsust, või eelistavad pigem olla teistega koos ning suhelda.

Suundud esimest päeva uuele töökohale. Tööruumi võid endale valida sää-
raste variantide seast:

- A. *Suur saal, mis on jaotatud lahtriteks.*
- B. *Kabinet neljale, kus igal töötajal on omaette privaatne lahter.*
- C. *Kabinet neljale, kus istutakse näod vastakuti.*
- D. *Privaatne kabinet.*

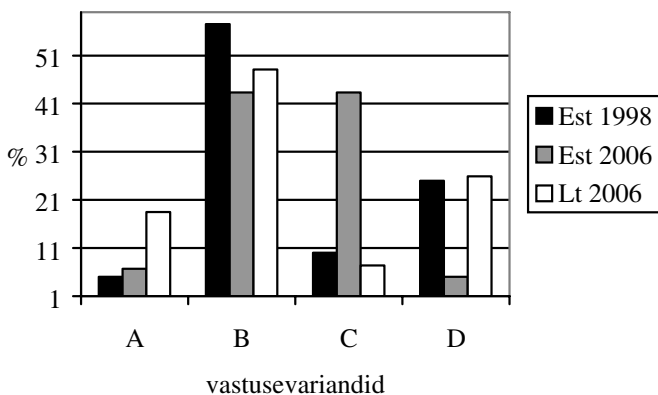
A-variandi valijad suhtlevad ise aktiivselt ning ootavad aktiivset suhtlemist ka teis-
telt. Selline variant iseloomustab eelkõige kollektivistlikke kultuure, kus rahavastiku
tihedus ei võimalda otsest eraldumist ja omaetteolemist. Sellises kultuuris on piir
privaatse ja sotsiaalse distantsivööndi vahel ähmane (Gudykunst, Ting-Toomey
1988: 123–128).

B-variandi valijad ei sea esmatähtsaks füüsilist privaatsust. Privaatsust on või-
malik saavutada ka visuaalsest kontaktist hoidumisega (Hall 1990: 133–134). Kuna
variandis kirjeldatud ruumikorralduse puhul on kolleegidel omavahel tõhusam
töötada, siis iseloomustab see variant maskuliinset kultuuri (ollakse orienteeritud
tööülesannetele ning tulemustele) (Hofstede: 2001: 298–299).

C-variandi valijaid iseloomustab soov üksteisega suhelda ning olla visuaalses
kontaktis. Kuna inimesed ja suhted on tähtsad, iseloomustab variant feminiinseid
kultuure (Hofstede 2001: 299). Kui soovitakse olla omaette, siis eralduakse ole-
masolevas ruumis, s.t lõpetatakse rääkimine (vt Hall 1990: 140, 159).

D-variandi valijad tajuvad ruumi kui isiklikku sfääri ning enda ego laiendust.
Selline valik iseloomustab sügavalt individualistlikku kultuuri (Hall 1990: 133,
Hofstede 2001: 227).

Küsimusele antud vastused on esitatud joonisel 4.



Joonis 4. Neljanda situatsiooniküsimuse vastused

1998. aastal eelistasid eestlased ülekaalukalt tööruumina kabinetti neljale, mis on
jaotatud privaatseteks lahtriteks (B). Sellise valiku puhul ei peeta esmaseks füüsilist
privaatsust, vaid piisab visuaalse kontakti vältimisest. Suhteliselt paljud vastanud
soovisid aga töökeskkonnas kolleegidest füüsiliselt eralduda ja eelistasid töötada
privaatses kabinetis (D). Seega tulid esile individualistlikule ja maskuliinsele

kultuurile omased tunnused, kus inimsuhetest olulisemaks peetakse eraldatust, tööülesandeid ja -tulemuslikkust.

2006. aastal eelistasid eestlased samuti töötada nelja kolleegiga ühes kabinetis, kus igal töötajal on oma privaatne lahter (B), ent samavõrra sooviti töötada nelja kaastöötajaga tööruumis, kus istutakse näod vastakuti (C). Seega, erinevalt varasemast küsitlusest on oluliseks muutunud ka vastastikune seotus ja suhted kaastöötajatega: näeme liikumist maskuliinsetelt väärtustelt feminiinsetele.

Võrreldes leedulaste ja eestlaste vastuseid, sarnanesid leedulaste valikud eestlaste vastustega 1998. aastal: kõige enam sooviti jagada tööruumi nelja kolleegiga, kusjuures igapäev oleks oma privaatne lahter (B). Samuti eelistasid paljud vastanud eralduda kaastöötajatest ning töötada privaatses kabinetis (D). Nii B- kui D-variant näitab, et töökeskkonnas peetakse inimsuhetest olulisemaks keskendumist tööülesannetele, see on aga maskuliinse kultuuri tunnus.

Viies situatsioon

Kultuure saab jagada avatud ja suletud ustega kultuurideks (vt Hall 1990: 134, 138–139). See tähendab, ühtedes kultuurides hoitakse ukse suletuna, teistes jällegi hoitakse ukse lahti. Ka on suletud või avatud ukse kultuuriti erinev tähendus, nt Saksamaal ei tähenda suletud uks, et inimene soovib olla segamatult omaette või on teoksil midagi, mida ei taheta teistele näidata, vaid suletud uksega säilitatakse ruumi terviklikkus ja luuakse töötamiseks vajalik keskkond (Hall 1990: 135–136). Põhjaameeriklastele seevastu on kinnine uks nõupidamiste, eravestluste ja tegevuste jaoks, mida ei taheta teistele näidata, ja suhtlema ollakse valmis siis, kui uks on avatud (Hall 1990: 138–139).

Viienda küsimusega sooviti teada, kas eestlased ja leedulased on suletud või avatud uste eelistajad.

Töötad läbikäidavas kontoriruumis. Tagaruumitöötajad tõmbavad oma tuppa minnes enda järel aegajalt ukse kinni. Kuidas suhtud ukse sulgemisse?

- A. *Arvad, et Sind püütakse kõrvale tõrjuda või tahetakse Sinu eest midagi varjata.*
- B. *Lükkad ukse lahti, et tagaruumitöötajatega seotust säilitada.*
- C. *Suletud ukse taga tunned end turvaliselt ja saad paremini tööle keskenduda. Kui tagaruumitöötajad jätavad ukse lahti, siis suled selle ise.*
- D. *Ei pööra sellistele pisasjadele tähelepanu.*

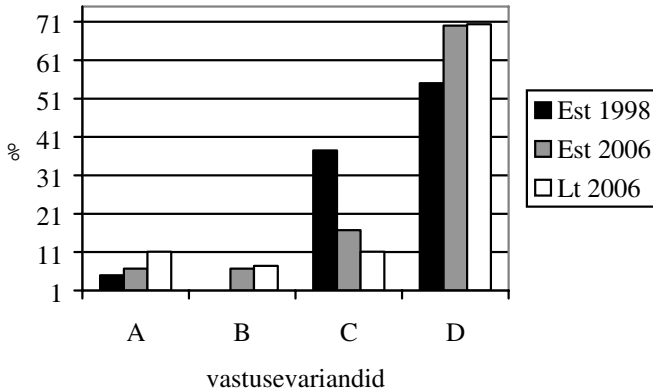
A-variant osutab avatud ustega kultuurile, ja ka ebakindlust vältivale kultuurile: suhtes ei sallita ebaselgust (Hofstede 2001: 161).

B-variant osutab avatud ustega kultuurile ja ka kollektivistlikule kultuurile: tahetakse olla seotud teiste inimestega (Hofstede 2001: 227).

C-variant osutab suletud ustega kultuurile ja ka individualistlikule kultuurile: soovitakse privaatsust ja austatakse üksiolemise õigust (Hall 1990: 133, Hofstede 2001: 227). Ukse sulgemisega luuakse distants iseenda ja väljaspool ruumi olijate vahel.

D-variandi valijate jaoks ei ole ukse sulgemine või avamine suhete reguleerimisvahend, ühtlasi näitab see, et ümbruses toimuvale ei pöörata suurt tähelepanu. Inimsuhete asemel ollakse orienteeritud tööülesannetele, mis on iseloomulik maskuliinsele kultuurile (Hofstede 2001: 298, 318).

Küsimusele antud vastused on esitatud joonisel 5.



Joonis 5. Viieendale situatsioonile antud vastused

Selle situatsiooni puhul olid eestlaste ja leedulaste küsitluse vastused sarnased: vastanud ei pööra tähelepanu, kas nende tööruumi uks on lahti või kinni (D). Seega ei saa kumbagi kultuuri liigitada ei suletud ega avatud ustega kultuuriks. D-variandi valik viitab maskuliinse kultuuri väärtustele, kuna ollakse orienteeritud töösuhetele. Samas võib D-variandi eelistus viidata ka asjaolule, et küsitletavad ei pruugi olla teadvustanud endale ukse kui suhtlusreguleerija funktsiooni.

Eestlaste kahanev vajadus ust sulgeda (vt C-variant) näitab, et privaatsust ei peeta enam nii oluliseks kui varasemas küsitluses.

Kuues situatsioon

Kuuenda situatsiooniküsimusega sooviti teada, milline tähtsus on füüsilise privaatsuse kõrval visuaalsel privaatsusel ja kas küsitletavad järgivad selles osas n-ö sakslaste mudelit, mille korral visuaalse kontaktiga ületatakse privaatse tsooni piir (vt Hall 1990: 133–134).

Jagad kabinetti kolleegiga. Sisse astub töökaaslase tuttav ja hakkab ilma Sind tähele panemata oma tuttavaga vestlema. Kuidas suhtud?

- Oled rahul, et Sind ei häirita – saad segamatult töötada.*
- Leiad, et Sinust on mööda mindud, tunned end puudutatuna.*
- Sul ükskõik, kes sisse astub või väljub ja kas Sinuga suheldakse või mitte.*
- Sekkid ka ise vestlusesse.*

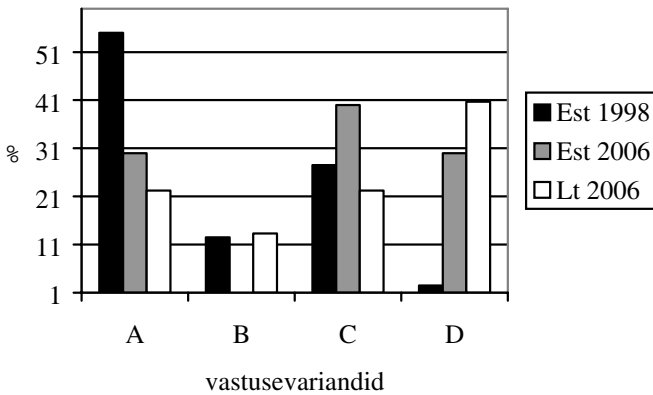
A-variandi korral aktsepteeritakse visuaalset privaatsust ning negatiivset viisakust: teine inimene jäetakse rahule, luuakse distants ja austatakse üksiolemisõigust. Tegemist on maskuliinse kultuuriga: tähtis on tulemus, edasiliikumine, ollakse orienteeritud tööülesannetele (Hofstede 2001: 298–299).

B-variandi valijad pööravad suurt tähelepanu inimsuhetele, vastandatakse “meie” ja “nemad”, visuaalset privaatsust ei tunnustata. Selline käitumine on iseloomulik kollektivistlikule kultuurile (Hofstede 2001: 226–227).

C-variant näitab, et suhteid teiste inimestega peetakse ebaoluliseks ja ümbruses toimuvale ei pöörata tähelepanu. Tegemist võib olla maskuliinse kultuuriga, kus eelistatakse keskenduda tööle, mitte suhtlemisele (Hofstede 2001: 298–299).

D-variandi valijate jaoks on oluline seotus ja suhted teiste inimestega. Privaatsust, sealhulgas visuaalset privaatsust ei vajata. Selline tunnus võib iseloomustada kollektivistlikke ja feminiinseid kultuure (Gudykunst, Ting-Toomey 1988: 123, Hofstede 2001: 299).

Küsimusele antud vastused on esitatud joonisel 6.



Joonis 6. Kuuendale situatsioonile antud vastused

Eestlased eelistasid 1998. aasta küsitluses A-varianti: nad on rahul, et töökeskkonnas neid ei häirita ja nad saavad tööle keskenduda. 2006. aastal oli A-variandi osakaal vastustes vähenenud ja eelistati C-varianti: teistega suhtlemisele ei pöörata tähelepanu. Nii A- kui C-variant iseloomustavad maskuliinset kultuuri, kus peatähelepanu on tööülesannetel ja nende tulemuslikkusel. Mõlemad variandid näitavad, et eestlased ei pea füüsilist eraldatust efektiivse töötamise jaoks ainuvõimalikuks lahenduseks, piisab ka visuaalsest kontaktist hoidumisest.

Erinevalt 1998. aastast oli 2006. aastal mh suur osa vastanuid, kelle jaoks olid tööülesannetest ja privaatsusest olulisemad suhted kaastöötajatega (D). Seega ilmnevad kultuuris ka feminiinseid ja kollektivistlikud tunnused.

Võrreldes eestlasi ja leedulasi, eelistasid leedulased vestlusesse sekkuda (D), mis tähendab, et suhteid ja vastastikust seotust kaastöötajatega peetakse oluliseks. Seega on tegemist feminiinsele ja kollektivistlikule kultuurile omaste tunnusojoontega. Samas oli suur hulk neid vastanuid, kes eestlastega sarnaselt ei pööra ümbritsevale tähelepanu ja on keskendunud tööle (A ja C), mis viitab maskuliinse kultuuri väärtustele.

Seitsmes situatsioon

Seitsmenda situatsiooniga sooviti teada saada, kas eestlased ja leedulased kasutavad sarnaselt inglastega distantsireguleerijana ka häälevaljust (vt Hall 1990: 142).

Kolleegiga kahekesi olles jutustad talle oma nädalalõpust, mille veetsid koos oma sõbra või sõbratariga. Äkki märkad, et ruumis on veel keegi. Kuidas reageerid?

- A. *Katkestad jutu ning lubad kolleegile, et räägid hiljem edasi.*
- B. *Tasandad häält ning jutustad loo lõpuni.*
- C. *Ei lase ennast kolmandast ruumisolijast häirida ning jätkad valjuhäälselt muljete jagamist.*
- D. *Kaasad vestlusesse ka kolmanda juuresolija.*

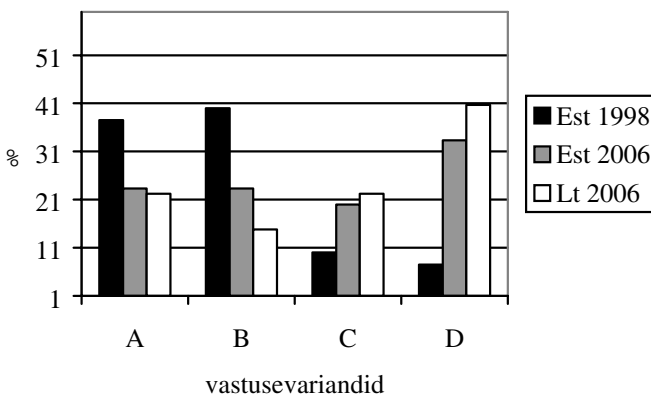
A-variandi valijaid iseloomustab vajadus säilitada privaatsus. Selline käitumine on omane eeskätt individualistlikule distantsi hoidvale kultuurile (Gudykunst, Ting-Toomey 1988: 123–124, Hofstede 2001: 227).

B-variandi valijad püüavad end teistest eraldada hääletugevust vähendades. Nii luuakse enda ümber privaatne tsoon ning antakse märku, et sissetungija pole oodatud (Hall 1990: 142). Enamasti on sellisel puhul tegemist individualistliku maskuliinsusesse kalduva kultuuriga (Gudykunst, Ting-Toomey 1988: 123–124).

C-variandi valijate jaoks on piir personaalse ning sotsiaalse ruumi vahel ähmane (vt Hall 1990: 154 jj). Selline valik näitab, et tegemist võib olla kollektivistliku feminiinse kultuuriga (Hofstede 2001: 227, 298). Samas võib hääle mittetasandamine tähendada ka avatuse demonstreerimist (vt Hall 1990: 142).

D-variandi valijad seavad esikohale inimsuhted. Selline käitumine on oodatav kollektivistlikus või siis feminiinses kultuuris, kus hinnatakse positiivset viisakust – teisele inimesele tähelepanu pööramist, tema kaasamist suhtlusesse (Hofstede 2001: 227, 298).

Küsimusele antud vastused on esitatud joonisel 7.



Joonis 7. Seitsmenda situatsiooniküsimuse vastused

Eestlaste 1998. aasta vastustes domineerisid A- ja B-variant, mis viitab asjaolule, et eestlased kasutavad distantsireguleerijana häälevaljust ning peavad eraldatust oluliseks. Selline valik osutab individualistliku ja maskuliinse kultuuri tunnustele. Samuti püütakse teist inimest mitte häirida ja soovitakse, et asjassepühendamatu vestlusesse ei sekkuks.

Eestlaste 2006. aasta vastustes on ülekaalus D-variandi valijad, mis vastupidiselt 1998. aasta tulemustele viitab, et inimsuhetele pööratakse ka töökeskkonnas rohkem tähelepanu. Taoline valik viitab feminiinse kultuuri väärtuste esiletõusmisele ja positiivse viisakuse ilmnemisele kultuuris (teist inimest ei jäeta suhtlusest kõrvale).

Võrreldes eestlaste ja leedulaste vastuseid, on leedulaste valik sarnane eestlaste 2006. aasta küsitluse tulemustega – kolmas ruumisviibija kaasatakse vestlusesse. Seega iseloomustavad leedulasi feminiinsed väärtused – teistest inimestest hoolimine.

Kaheksas situatsioon

Kaheksanda situatsiooniküsimusega sooviti teada saada, kuidas vastanud suhtuvad vestluspartneri hingeõhu tundmise – kas hingeõhu tundmine kuulub suhtluse juurde või peetakse seda ebameeldivaks ja ruumikitsikust tekitavaks (vt Hall 1990: 159–160, Samovar, Porter 2004: 186–187).

Rääkides on vestluspartner Sulle väga lähedale tulnud ja Sa tunned tema hingeõhku. Kuidas reageerid?

- A. *Oled väga häiritud: küsid, mida vestluspartner on söönud/joonud. Võimaluse korral pakud talle kurgupastille.*
- B. *Tunned end ebamugavalt, kuid püüad seda mitte välja näidata, taganed tasapisi.*
- C. *Sulle meeldib vestluspartneri hingeõhku tunda.*
- D. *Ei pane antud asjaolu tähelegi.*

A-variandi valijate jaoks on kindel distants vestluspartneriga väga oluline ning selle rikkumise suhtutakse resoluutselt otsest suhtlusstiili kasutades. Selline käitumine ja kõnestiil iseloomustab eeskätt individualistlikke maskuliinseid kultuure (vt Gudykunst, Ting-Toomey 1988: 124, Gudykunst jt 1995: 325).

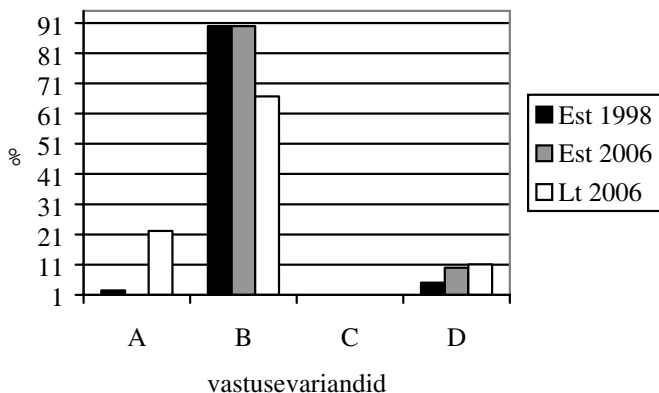
B-variandi valijad soovivad hoida distantsi, kuid samuti püütakse vältida konflikte, ebakindlaid olukordi talutakse A-variandiga võrreldes paremini. Selline valik võiks iseloomustada ebakindlust suhteliselt hästi taluvaid kultuure (emotsioonid hoitakse vaos) (Hofstede 2001: 160), kus ühtlasi aktsepteeritakse negatiivset viisakust.

C-variandi valijad pööravad suurt tähelepanu inimsuhetele, neile meeldib olla üksteise lähedal, piir intiimse ja personaalse tsooni vahel on ähmane (vt Hall 1990: 159–160). Enamasti iseloomustab selline valik kollektivistlikke ja feminiinseid kultuure (Gudykunst, Ting-Toomey 1988: 123–128, Samovar, Porter 2004: 186–187, 190).

D-variandi valijaid iseloomustab kas ebateadlikkus distantsireeglitest, või siis on distantsipiirid nende inimeste jaoks ähmased. Hingeõhu tundmist võetakse

kui loomulikku asja, mis ei häiri (vt Hall 1990: 157–160, Samovar, Porter 2004: 186–187). Selline valik võiks iseloomustada kollektivistlikke kultuure (Gudykunst, Ting-Toomey 1988: 123–128, Hall 1990: 53).

Küsimusele antud vastused on esitatud joonisel 8.



Joonis 8. Kaheksanda situatsiooniküsimuse vastused

B-variandi ülekaalukas valik (teise inimese hingeõhu tundmine on ebameeldiv, kuid seda püütakse mitte välja näidata), ja C-variandi mittevalimine (suhtluspartneri hingeõhu tundmine on meeldiv) eestlaste ja leedulaste seas näitab, et mõlemad kultuurid peavad distantsihoidmist sotsiaalses suhtluses vajalikuks, kuid distantsireegleid ei suruta inimesele peale otsesõnaliselt. Vältitakse konflikte ja hoidutakse teise inimese ärritamisest. Samas on leedulaste hulgas ka neid vastanuid, kes pidasid ruuminormide rikkumist vastumeelseks ja seda otsesõnu ka väljendaksid (vt A-variant).

Üheksas situatsioon

Kultuure võib jagada ka kontakt- ja mittekontaktkultuurideks. Kontaktkultuurides tulevad inimesed üksteisele lähedale ning vestluse käigus puudutavad teineteist. Mittekontaktkultuurides on inimese isiklik ruum suur (Gudykunst, Ting-Toomey 1988: 124–126). Üheksanda situatsiooniküsimusega sooviti teada, kuhu paigutuvad kontaktkultuuride-mittekontaktkultuuride dimensioonil eestlased ja leedulased.

Sinu ülemus on muhe mees. Tal on harjumus alluvatega vesteldes neile käsi õlale panna või siis seljale patsutada. Kuidas suhtud?

- Sulle meeldib, et Sind “omaks poisiks” peetakse ja võtad seda kui tunnustust.
- Püüad ülemusele mõista anda, et Sulle säärane semutsemine ei meeldi.
- Vastad samaga.
- Sulle meeldib ülemuse heatahtlik suhtumine, mitte patsutamine.

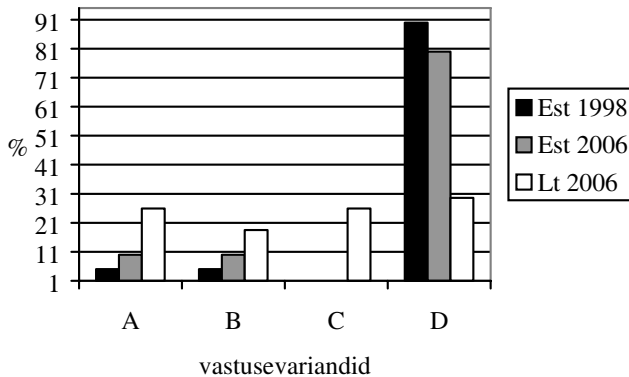
A-variandi korral talutakse füüsilist kontakti ning piir isikliku ja intiimse distant-sivööndi vahel on ähmane (vt Hall 1990: 154–159). Nii käitatakse enamasti suure võimudistantsiga kultuurides, kus ideaalne ülemus on kui hea isa, kelle soosingut tahetakse võita. Tunnistatakse ülemuse eesõigusi (vt Gudykunst, Ting-Toomey 1988: 124–126, Hofstede 2001: 98, 107), sh distantsireeglitest üleastumist.

B-variandi valijad ei talu füüsilist kontakti ning ruumivööndite vahel tehakse vahet. Iseloomustada võiks selline suhtumine väikese võimudistantsiga individualistlike kultuure, kus ei kardeta oma suhtumist välja näidata (Hofstede 2001: 107, 116).

C-variandi valijad on orienteeritud inimsuhetele – lähedust ei peljata. Variant iseloomustab väikese võimudistantsiga feminiinseid kultuure. Ülemused ja alluvad on seal võrdsed (Hofstede 2001: 107–108, 298–299). William Gudykunsti ja Stella Ting-Toomey järgi (1988: 128) ei ole väikese võimudistantsi, väikese ebakindluse vältimise tarbega ning feminiinsetes kultuurides reeglid puudutamise suhtes nii ranged kui suure võimudistantsiga, suure ebakindluse vältimise tarbega ja maskuliinsetes kultuurides.

D-variandi valijad ei talu füüsilist kontakti ega distantsipiiride rikkumist. Iseloomulik on selline suhtumine individualistlikele maskuliinsetele kultuuridele (Gudykunst, Ting-Toomey 1988: 127–128). Võimudistants sellises kultuuris ei saa olla väga väike (üksteist ei patsutata), kuid ei saa olla ka väga suur (patsutamine ei meeldi, seda ei võeta kui soosingu osutamist).

Küsimusele antud vastused on esitatud joonisel 9.



Joonis 9. Üheksanda situatsiooniküsimuse vastused

Eestlased on mõlemas küsitluses eelistanud D-varianti: ülemuse heatahtlik suhtumine meeldib, kuid taunitakse patsutamist. See osutab asjaolule, et eestlaste puhul on tegemist mittekontaktikultuuriga: inimesed ei talu füüsilist kontakti ega distantsipiiride rikkumist. Samuti viitab puudutamisest hoidumine individualistliku ja maskuliinse kultuuri väärtustele.

Leedulaste küsitluses ei domineerinud antud situatsiooniküsimuse puhul selgelt ükski vastusevariant. Osa vastanuid eelistas sarnaselt eestlastega D-varianti, mille puhul puudutamisest hoidumine viitab individualistliku ja maskuliinse kultuuri väärtustele. Samas oli palju ka neid vastanuid, kes võtsid ülemuse õlalepatsutust kui

tunnustust (A), vastasid ise samaga (C) või kes püüdsid ülemusele mõista anda, et puudutamine on ebameeldiv (B). Kui A-variant osutab suure võimudistantsiga kultuurile, siis B- ja C-variant on vastupidiselt väikse võimudistantsiga kultuuri tunnus.

Kümnes situatsioon

E. T. Halli (1990: 122–123) järgi hoiavad inimesed mõjukate inimestega sotsiaalset distantsi (hoiavad üksteisest suhteliselt kaugelt). Kümnenenda situatsiooniküsimusega sooviti teada, kuidas käituvad kõnesolevas situatsioonis eestlased ja leedulased.

Oled suure osakonna juhataja. Täna on peadirektori juures igapäevane osakonnajuhatajate koosolek. Kuhu istud?

- A. *Arvestad väljakujunenud kohtadega ja istud n-ö oma kohale.*
- B. *Püüad istuda juhi kõrvale.*
- C. *Istud direktorist võimalikult kaugelt.*
- D. *Ei tee kohast numbrit, istud, kuhu juhtub.*

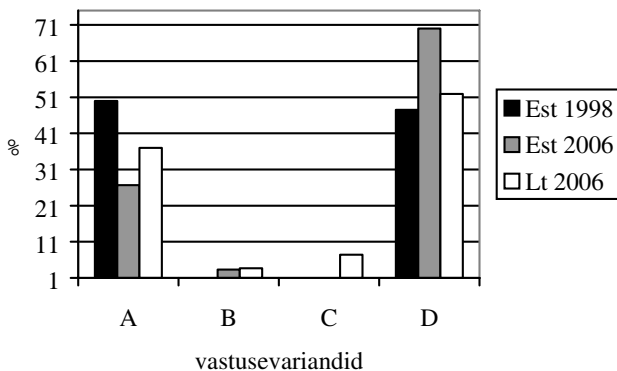
A-variandi puhul on koht ja inimene seotud. Selline valik iseloomustab ebakindlust halvasti taluvaid kultuure, kuna inimesi ahistab kõik ootamatu ja ebaselge (Hofstede 2001: 160–161).

B-variandi puhul võiks tegemist olla suure võimudistantsiga maskuliinse kultuuriga: kõrgemal positsioonil olevale isikule soovitakse meeldida, edukatele kuulub sümpaatiat (Hofstede 2001: 98, 107–108, 299, 318).

C-variandi valijaid iseloomustab lugupidav suhtumine ülemustesse. Iseloomulik on sellise variandi eelistamine suure võimudistantsiga kultuuris, kus ebavõrdsus on heakskiidetav ning alluvad ootavad käsked ja korraldusi (Hofstede 2001: 98, 107–108).

D-variandi puhul puudub koht ja inimese vahel seos. Selline variant osutab ebakindlusele hästi taluvale kultuurile (Hofstede 2001: 160–161).

Küsimusele antud vastused on esitatud joonisel 10.



Joonis 10. Kümnenenda situatsiooniküsimuse vastused

1998. aastal jagunesid eestlaste vastused vastandlike A- ja D-variandi vahel: arvestatakse väljakujunenud kohtadega või valitakse istekoht juhuslikult. Kui A-variant viitab ebakindlust halvasti taluvale kultuurile, siis D-variandi valijad taluvad ebakindlaid olukordi hästi.

2006. aastal on eestlaste seas vähenenud A-variandi osakaal ja samavõrd on suurenenud D-variandi eelistus. Seega tõusevad vastustes enam esile ebakindlust hästi taluva kultuuri tunnused. B-variandi vähene ja C-variandi mittevalimine viitab väikesele võimudistantsile ja feminiinsetele väärtustele – ühiskonnas hinnatakse võrdsust ja sõltumatust. Võim seondub vaid erinevate rollide ja ametikohtadega, kuid need ei ole jäävad: tänane alluv võib olla homme ülemus, ja vastupidi (vt Hofstede 1993: 42–75).

Võrreldes leedulasi eestlastega, sarnanes leedulaste valik eestlaste mõlema küsitluse tulemustega: kuigi leedulaste vastustes domineeris D-variant (istekohast ei tehta numbrit ja istutakse kuhu juhtub), oli suur hulk neid vastanuid, kes eelistasid A-varianti (soovivad istuda n-ö oma kohale). Kui D-variant viitab ebakindlust hästi taluva kultuuri väärtustele, siis A-variant on ebakindlust halvasti taluva kultuuri tunnus.

Vaba ajaga seotud küsimused

Suhteid väljaspool kodu ja tööd uuriti situatsioonide kaupa. Sooviti teada saada, kas vastanud suhtlevad avalikus kohas meelsasti võõraste inimestega või keskenduvad vaba aega veetes eeskätt suhetele oma sõprade ja lähedastega. Ühtlasi taheti välja selgitada, kuidas suhtuvad eestlased ja leedulased ebaselgetesse ruumilistesse olukordadesse – kas ruuminormide rikkumise korral kasutatakse otsesest või kaudset verbaalset stiili.

Üheteistkümnnes situatsioon

Selle situatsiooniküsimusega uuriti, kuidas eestlased ja leedulased käsitavad personaalset tsooni avalikus kohas ja kui oluline on vastanute jaoks privaatsus vaba aega veetes – kas võõras inimene lubatakse seltskonda ja haaratakse vestlusesse või tõrjutakse ta kõrvale, ning kuidas seda tehakse.

Veedad sõbraga aega populaarses kohvikus. Teie lauas on kaks vaba kohta, kuid üksikuid vabu kohti on mujalgi. Üks kohvikukülastajatest palub luba teie lauda istuda. Kuidas käitud?

- A. *Ütled, et soovite omaette olla.*
- B. *Lubad, kuid ise oled pahane, et just teie lauda tuldi. Õhtu on rikutud.*
- C. *Palud istuda ja kaasad ta peatselt vestlusesse.*
- D. *Lubad, kuid käitud hiljem temaga nagu õhuga.*

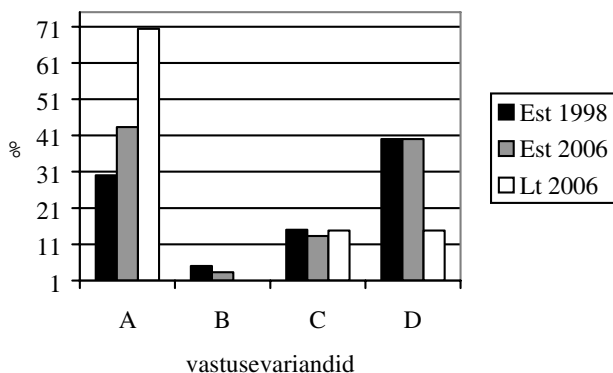
A-variandi valijad soovivad avalikus kohas privaatsust ning neid iseloomustab resoluutsus ja otsese verbaalse stiili kasutamine. Selline käitumine on iseloomulik individualistlikele maskuliinsetele kultuuridele (Salo-Lee jt 1996: 36–39).

B-variandi valijad kasutavad kaudset verbaalset stiili – inimestel on raske öelda ei. Enamasti on selline stiil omane kollektivistlikele kultuuridele (Gudykunst jt 1996: 29–34). Tehakse vahet sotsiaalse ja isikliku ruumivööndi vahel (vt Hall 1990: 154–159).

C-variandi valijad peavad väga tähtsaks inimsuhteid ning üksteisest hoolimist. Suheldakse ise meeleldi ja oodatakse seda ka teistelt. Variant iseloomustab kollektivistlike ja ka feminiinseid kultuure, kus hinnatakse positiivset viisakust (Hofstede 2001: 226–227, 298–299).

D-variandi valijad ei ole orienteeritud inimsuhtetele. Tegemist võib olla individualistliku ning maskuliinse kultuuriga, kus hinnatakse negatiivset viisakust (Hofstede 2001: 227, 299).

Küsimusele antud vastused on esitatud joonisel 11.



Joonis 11. Üheteistkümnenda situatsiooniküsimuse vastused

Selle situatsiooni puhul iseloomustasid eestlasi nii 1998. aastal kui 2006. aastal maskuliinsed ja individualistlikud väärtused – mõlemal juhul ei olnud eestlased valmis neile tundmatu inimesega suhtlema (A ja D). Samas, kui esimese küsitluse puhul aktsepteerisid vastanud negatiivset viisakust (võõras inimene lubatakse oma lauda, kuid teda vestlusesse ei kaasata), siis järgnevas uuringus oli nihe otsese suhtlusstiili kasutamise suunas (võõrale inimesele teatatakse otsekoheselt, et teda seltskonda ei soovita).

Leedulaste vastustes oli ülekaalus A-variandi valik, mis näitab, et eestlastest enam ollakse valmis oma personaalset ruumi ja privaatsust otsesõnu kaitsma. B-variandi (kaudne verbaalne stiil, raske anda eitavat vastust) mittevalimine ja D-variandi (hinnatakse teise inimese eraldumisõigust) vähene valik viitab, et leedulastele ei ole avalikus suhtluses omased kollektivistliku kultuuri tunnused ning nad hindavad võõrastega suheldes eestlastest märgatavalt vähem negatiivset viisakust. Seega toovad kõnesoleva küsimuse vastused leedulaste puhul esile individualistlikud maskuliinsed väärtused, kus domineerib otsene verbaalne stiil.

Kaheteistkümnnes situatsioon

Sooviti teada saada, kuidas eestlased ja leedulased suhtuvad ruumiliselt ebaselgetesse suhtlussituatsioonidesse, kus kaasinimese kavatsused on teadmata, ja kas nad

käituvad ruuminormide rikkumise korral aktiivselt ning agressiivselt või passiivselt ja taandudes (vt Gudykunst, Ting-Toomey 1988: 124).

Toidupoes leti juures seistes tajud, kuidas Sinu taga olev inimene Sulle järjest lähemale nihkub. Viimaks tunned kuklas juba tema hingeõhku! Kuidas reageerid?

- A. *Tunned hirmu ja püüad märkamatult eest ära liikuda.*
- B. *Vihastud ning nõuad, et selgatiikkija hoiaks “pikivahet”.*
- C. *Kuigi oled ärritunud, mõtled “Mina olin siin enne” ja ei lase ennast kõrvale tõrjuda.*
- D. *Ei pane antud asjaolu täheleegi.*

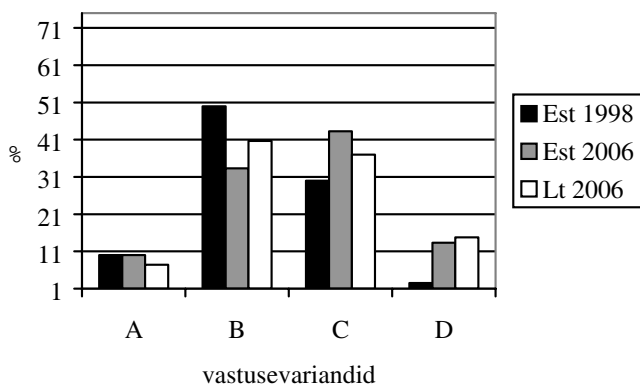
A-variandi valijad ei salli ebaselgust. Kõik võõrastav tekitab ohutunde. Distant-sipiiride rikkumisele reageeritakse passiivselt, taandudes. Iseloomulik on selline käitumine kollektivistlikele kultuuridele, kus ebakindluse vältimise tarve ei ole siiski kuigi suur. Sellises kultuuris ärritust otsesõnu ei väljendata (vt Gudykunst, Ting-Toomey 1988: 124, Gudykunst jt 1996: 42, Hofstede 2001: 160).

B-variandi valijad reageerivad ebaselgetes suhtlussituatsioonides resoluutselt otsest verbaalset stiili kasutades. Selline käitumine on iseloomulik individualistliku, maskuliinse kultuuritaustaga ja suure ebakindluse vältimise tarbega inimestele (Gudykunst, Ting-Toomey 1988: 124, Gudykunst jt 1995: 325, Hofstede 2001: 160).

C-variandi valijad ei kasuta ebaselgetes olukordades otsest verbaalset stiili, kuid samas väljendatakse kehakeelega oma suhtumist ning ei lasta end võõral inimesel kõrvale tõrjuda. Säärast käitumist võib oodata individualistlikes maskuliinsetes kultuurides (Gudykunst, Ting-Toomey 1988: 124, 128). Tõenäoliselt taluvad antud vastusevariandi eelistajad B-variandi valijatega võrreldes suhteliselt paremini ka ebakindlust, kuna ärritust ja viha otseselt ei väljendata (Hofstede 2001: 160).

D-variandi valijad ei hooli järjekordades seismisest: neile meeldib olla üksteise lähedal ja suhelda. Omane on selline valik kollektivistlikele kultuuridele (vt Hall 1990: 154–157, 162).

Küsimusele antud vastused on esitatud joonisel 12.



Joonis 12. Kaheteistkümnenda situatsiooniküsimuse vastused

Selle situatsiooni puhul olid eestlaste ja leedulaste tulemused suuremas osas sarnased, s.t kumbki kultuur ei talunud ruumiliselt ebaselgeid olukordi. A- ja D-variandi (omane kollektivistlikele kultuuridele) vähene valimine ning B- ja C-variandi (privaatus oluline) eelistamine viitab, et eestlasi ja leedulasi iseloomustavad individualistlikud ja maskuliinsed väärtused. Samas kasutavad leedulased ruuminormide rikkumise korral enam otsest verbaalset stiili (B), sarnanedes seega 1998. aastal vastanud eestlastega. 2006. aastal on eestlaste seas kasvanud C-variandi eelistajate hulk, millest võib järeldada, et eestlased on 8 aastaga muutunud ebakindlust paremini taluvaks. Leedulased taluvad võõrastega suheldes ebakindlust eestlastest halvemini.

Järeldused ja diskussioon

Majandusareng ja kiired muutused ühiskonnas võivad kultuuri muidu väga püsivaid väärtusi muuta (Inglehart, Baker 2000: 20). Eestlaste 1998. ja 2006. aasta küsitluse vahelise aega jäi nii Eesti ühinemine NATO-ga kui ka liitumine Euroopa Liiduga. Kahe küsitluse tulemuste võrdlus näitab väärtuste erinevust: eesti kultuuris on toimunud mitmed väärtusnihked. Leedulaste väärtusmuutustest töökeskkonnas aastatel 1993–1997 on ülevaate andnud Virginius Lepas (1999), praegustest väärtustest Audra I. Mockaitis (2005). Käesolev uurimus näitab, et väärtusmuutused Leedus jätkuvad. Leedu ruumikäitumine on eestlaste omast mitmekesisem, esile tulevad vastandlikud väärtused, muutuste suunda on keeruline määrata.

Uurimus näitab, et eelistused ja väärtused võivad erineda ka sõltuvalt keskkonnast: peresuhetes võivad esile tulla ühed kultuurinormid, töösuhetes ja vaba aega veetes teised.

Kodune keskkond

Kodust keskkonda puudutavate situatsiooniküsimuste puhul selgus, et 1998. aastal iseloomustasid eestlasi peresuhetes valdavalt kollektivistlikule ja feminiinsele kultuurile omased tunnused, kus eraldatusest olulisemaks peetakse suhteid teiste inimestega ning elukeskkonda eelistatakse privaatsusele (1. situatsiooni variandid A/C, 2. situatsiooni variant D). Ent esines ka vastaspoolus: ilusa keskkonna puudumist korvatakse ruumika majaga, kus igal pereliikmel oleks eraldumisvõimalus (1B), mis viitab individualistlike ja maskuliinsete väärtuste olemasolule kultuuris. 2006. aastal on väärtushinnangud veelgi enam jagunenud: suhteliselt palju vastanuid väärtustavad elukeskkonda ja tugevaid peresidemeid (1A/C, 2C/D), kuid suur on ka nende hulk, kelle jaoks on esmatähtis omaette olemise võimalus (1B, 2A). Seega, 1998. aasta ja 2006. aasta eestlaste küsitluste tulemusi võrreldes on koduses keskkonnas toimunud nihe kollektivistlikelt väärtustelt individualistlikele väärtustele. Kuna eestlaste materiaalne heaolu on tänaseks tunduvalt paranenud, võivad saadud tulemused kinnitada arvamust, et majanduslik areng suurendab kultuuris individualistlike väärtuste osakaalu: kui rahvuslik rikkus kasvab, on inimestel ressursse, mis võimaldavad neil lubada endale rohkem privaatsust ja individuaalseid otsuseid (vt Basabe jt 2002: 107, Schimmack jt 2005: 29).

Leedulaste puhul on kodusuhetes näha, et nende jaoks on oluline elukeskkond, hoolitakse teistest inimestest ja tähtsustatakse sotsiaalselt võrgustikku ning naabruskonda (1A/C, 2C/D). Seega kinnitavad saadud andmed A. Mockaitise (2005, vt ka tabel 1) uuringu tulemusi, mille kohaselt on leedulased eestlastest märgatavalt kollektivistlikumad. Näiteks, uude elukohta kolides jagavad eestlased kõigepealt toad pereliikmete vahel ära (2A), leedulased aga lähevad ja tutvuvad naabrite ja ümbruskonnaga (2C) ning lasevad osa vaheseinu maha lõhkuda, et tekiks avatud ruum (2D).

Vastuolu A. Mockaitise (2005) tulemustega tekkis aga maskuliinsuse-feminiinsuse dimensioonil. Nimelt on A. Mockaitise andmete põhjal leedu kultuur küllaltki maskuliinne, kuid käesoleva uurimuse põhjal võib oletada, et kodusuhetes iseloomustavad leedulasi pigem feminiinsed väärtused, nagu suhted kaasinimestega ja ilusa elukeskkonna eelistamine, mis korvab koduse ruumipuuduse.

Mõlema kultuuri kohta võib öelda, et oma pere keskel taluvad nii eestlased kui leedulased suhteliselt hästi ebakindlamaid olukordi. Pereliikmetele ei suruta peale fikseeritud distantse, vaid oluliseks peetakse suhtlemise mugavust – mööbli asukohta muudetakse vastavalt sellele, mida ruumis tehakse ja kellega seal suheldakse (3A).

Töökeskkond

Nii 1998. kui ka 2006. aasta eestlaste küsitlustest selgus, et töökeskkonnas suheldes on eestlastele omased individualistliku ja maskuliinse kultuuri tõekspidamised. Tööl keskendutakse eeskätt tööülesannetele ja -tulemustele, mistõttu soovitakse ka suuremat eraldatust. Samas ilmnes 2006. aasta küsitlusest, et privaatsust ei peeta enam nii oluliseks kui varem (5C). Privaatsust aitab hoida ka kaastöötajaga pilgukontakti vältimine ja hääle tasandamine või siis vaikimine (6A/C). Eestlane töötab meeleldi nelja kolleegiga ühes ruumis. Kui 1998. aastal soovisid eestlased, et ruum oleks eraldatud vaheseintega (4B), siis 2006. aastal sooviti lisaks sellele samavõrra istuda nelja kolleegiga näod vastakuti (4C).

Võrreldes 1998. aasta tulemusi 2006. aasta vastustega, ilmnes erinevus ka töökaaslastega suhtlemisel. Kui esimesest küsitlusest saadi teada, et eestlased ei tähtsusta töökeskkonnas suhteid kaastöötajatega, siis 2006. aastal tulid esile feminiinse kultuuri väärtused – ka töökeskkonnas pööratakse inimsuhetele rohkem tähelepanu (7D) ning teist inimest ei jäeta suhtlusest kõrvale. Seega on toimunud kultuuris muutus: suhetes kaastöötajatega näeme liikumist maskuliinsetelt väärtustelt feminiinsetele.

Eestlaste otsus valida tööruumiks neljane kabinet ning nende valmisolek töökaaslastega suhelda võib olla tingitud vajadusest olla töös efektiivsem. Nimelt on kolleegidel, kel on sarnased tööülesanded ja ühine eesmärk, ühes ruumis tõhusam töötada. Väike distantse kaastöötajate vahel muudab tööprotsessi efektiivsemaks, kuna kaastöötajad suhtlevad omavahel tihedamini (Kiesler, Cummings 2002, Nova 2005: 122 järgi). Mitteformaalne suhtlus on tööalase koostöö nurgakivi (Nova 2003: 10) ja arusaamine kolleegide tegevusest annab konteksti ka inimese enda tegevusele (Dourish, Bellotti 1992: 8).

Sarnaselt eestlaste tulemustele viitavad leedulaste andmed valdavalt individualistliku ja maskuliinse kultuuri väärtustele: tööl keskendutakse eeskätt töö-

ülesannetele ega lasta end kõrvalistest asjaoludest häirida. Seega on eestlastel ja leedulastel ühine soov töötada privaatsuses. Tööruumi valikul soovivad leedulased, nagu eestlased, töötada nelja kolleegiga ühes ruumis, igal oma privaatne lahter (4B). Samas on leedulaste seas suhteliselt palju neid vastanuid, kes eelistavad töötada omaette kabinetis (4D). Kuna leedulased on eestlastest kollektivistlikumad (Mockaitis 2005), võib privaatne kabinet olla ka prestiiži tunnus.

Eestlaste ja leedulaste erinevus tuli esile töökaaslastega suhtlemisel. Leedulaste puhul on suhetes kolleegidega valdav, et neid ei jäeta suhtlusest kõrvale (7D) ning võimalusel sekkutakse ka ise vestlusesse (6D). Märnatavalt vähem on leedulaste hulgas vastanuid, kes inimsuhetest enam soovivad pühenduda tööle ja meelsamini hoiduks töökaaslastega suhtlemisest (6A/C).

Küsitlustest ilmnes, et nii leedulased kui ka eestlased taluvad suhteliselt hästi ebakindlamaid olukordi – mõlemad kultuurid peavad distantsihoidmist sotsiaalses suhtluses vajalikuks, kuid seda otseselt teisele inimesele peale ei suruta ja suhtluskonflikte püütakse vältida (8B, 9D). Samas on leedulaste seas suur hulk vastanuid, kes on valmis vastumeelse ruumikäitumise korral seda otsesõnu ka väljendama (8A, 9B), eestlased aga pigem väldivad otsest ütlemist.

Nii eestlaste mõlemast küsitlusest kui ka leedulaste omast selgus, et ühe kultuuri piires võib olla erinevate kultuuriväärtuste kandjaid. Kui kultuuris puudub domineeriv joon, on suhtluskonfliktid kerged tekkima. On küllaltki tõenäoline, et eesti kultuuris põrkuvad töökeskkonnas maskuliinsed ja feminiinsed väärtused (vt situatsioon 4, situatsioon 6), ning väärtused, mis on omased ebakindlust hästi taluvatele ja ebakindlust halvasti taluvatele kultuuridele (situatsioon 8). Leedulaste puhul on pilt veelgi kirjum: väärtuskonfliktid võivad esile tulla võimudistantsi erinevuste pinnal (situatsioon 9), ebakindlate olukordade talumise erinevuse pinnal (situatsioon 9, 10) ning maskuliinsete-feminiinsete väärtuseelistuste pinnal (situatsioon 6).

Domineerivate väärtuste puudumine ühiskonnas näitab, et väärtused on muutumas. V. Lepeska (1999) täheldas väärtuste teisenemist Leedus ka ajavahemikul 1993–1997. Ta tõi esile kollektivismi tugevnemise, võimudistantsi suurenemise, maskuliinsetelt feminiinsetele väärtustele ülemineku. Suhteliselt muutumatuna püsis aga ebakindluse talumine.

Käesolevas artiklis käsitletud töökeskkonda puudutavad tulemused ei toeta kõiki V. Lepeska (1999) seisukohti. Kuigi leedulaste kollektivistlikkus on eestlaste omast suurem (Mockaitis 2005, vt ka tabel 1), ilmnevad 2006. aastal leedulastel sarnaselt eestlastega töökeskkonnas pigem individualistlikule kultuurile omased jooned. Niisugune tulemus on kooskõlas ka N. Basabe jt (2002) ja U. Scimmacki jt (2005) andmetega, mis näitavad, et majanduslik areng suurendab ühiskonnas individualistlikke väärtusi. Ja kuigi leedulased on A. Mockaitise (2005) järgi eestlastest oluliselt maskuliinsemad (vt ka tabel 1), toetab käesolev uurimus V. Lepeska (1999) tulemusi leedulaste feminiinsete väärtuste esiletõusust: leedulased pööravad eestlastest enam tähelepanu töökaaslastele. Selline käitumine iseloomustab aga just pigem feminiinseid kultuure. Võimudistants on leedulastel eestlastest mõnevõrra suurem (Mockaitis 2005, vt tabel 1) ning seda kinnitas ka käesolev uurimus, samas aga ei ole leedulased võimuküsimustes täielikus üksmeeles – on neid, kelle vastused viitavad suurele võimudistantsile ja neid, kelle omad väiksemale võimudistantsile.

Suhted väljaspool kodu ja tööd

1998. aasta ja 2006. aasta küsitlustest ilmnes, et väljaspool kodu- ja töökeskkonda suheldes on eestlaste ja leedulaste puhul olemas piir sotsiaalse ja isikliku ruumivõõndi vahel, kus suhteid võõrastega ei peeta tähtsaks ning neid oma seltskonnaga liituma ei oodata (vt situatsioon 11). Ilmselt kannavad eestlased ja leedulased väljaspool kodu ja tööd individualistlikule maskuliinsele kultuurile omaseid jooni, kus suhtlusnormid ja distantsipiirid on selgelt määratletud. Leedulaste ja eestlaste erinevus tuli ilmsiks ebaselgetes suhtlusolukordades verbaalse stiili kasutamises. Nimelt reageerivad leedulased ebamugavates ruumilistes olukordades valdavalt resoluutselt ning oma privaatsset ruumi otsesesõnu kaitstes (11A, 12B). Eestlastest annab suur osa mõista, et olukord on neile ebameeldiv, kuid oma viha ega ärritust keeleliselt ei väljenda (11D, 12C), kuigi oli ka neid vastanuid, kes sarnaselt leedulastega väljenduvad resoluutselt ja otsesest verbaalset stiili kasutades (11A, 12B). Saadud vastustest võib järeldada, et võrreldes leedulastega taluvad eestlased avalikus kohas ja võõrastega suheldes ebakindlaid olukordi paremini, mis kinnitab ka A. Mockaitise (2005, vt tabel 1) tulemusi leedulaste mõnevõrra halvemast ebakindluse talumisest. Vastuolu A. Mockaitise esitatud andmetega ilmnes aga individualismi-kollektivismi dimensioonil – erinevalt A. Mockaitise uurimusest ei ole käesoleva küsitluse tulemuste põhjal leedulastele avalikus suhtluses omased kollektivistliku kultuuri tunnused.

Kokkuvõte

Võrreldes 1998. ja 2006. aasta küsitluse tulemusi, iseloomustavad eestlasi koduses keskkonnas küll kollektivistlikud ja feminiinsed väärtused, ent ilmneb arengutendents individualistlike väärtuste suunas. Leedulaste 2006. aasta küsitlus näitas, et neile on koduses keskkonnas valdavalt omased kollektivistlikud ja feminiinsed väärtused. Eestlaste ja leedulaste ühisjooneks on suhteliselt hea ebakindluse taluvus.

Töösuhetes prevaleerivad eestlastel individualistlikud ja maskuliinsed tööks- pidamised, kuid toimunud on nihe maskuliinsetelt väärtustelt feminiinsemate suunas. Leedulaste küsitlustulemused viitavad samuti valdavalt individualistliku ja maskuliinse kultuuri väärtustele, ent nende ruumikäitumine tööl on eestlaste omast märgatavalt mitmekesisem. Eestlastest suurem suhtlemine töökaaslastega näitab leedulaste feminiinseid väärtuseid. Võimudistants on leedulastel eestlaste omast suurem.

Suheldes väljaspool kodu- ja töökeskkonda, on eestlastele ja leedulastele iseloomulikud individualistlikule ja maskuliinsele kultuurile omased väärtused. Võrreldes leedulastega taluvad eestlased avalikus kohas ja võõrastega suheldes ebakindlaid olukordi paremini.

Olenemata suhtlussituatsioonist iseloomustab leedulasi enamasti resoluutsus ja otsene verbaalne stiil, eestlased seevastu keeleliselt end nii otseselt ei väljenda.

Kirjandus

- Adler, Ronald B.; Elmhorst, Jeanne Marquardt 1992. *Communication at Work: Principles and Practices for Business and the Professions*. New York: McGraw-Hill.
- Altrov, Rene 1999. Eestlaste ruumikasutus suhtluses. Diplomitöö. Eesti Humanitaarinstituut. <http://sinine.ehi.ee/ehi/oppetool/lopetajad/altrov/altrov.html> (29.11.2006).
- Basabe, Nekane; Paez, Dario; Gonzales, Jose Luis; Rimé, Bernard; Diener, Ed 2002. Cultural dimensions, socioeconomic development, climate, and emotional hedonic level. – *Cognition and Emotion* 16 (1), 103–125.
- Dourish, Paul; Bellotti, Victoria 1992. Awareness and coordination in shared workspaces. – J. Turner, R. Kraut (Eds.). *Proceedings of CSCW'92: Conference on Computer Supported Cooperative Work*. ACM Press, 107–114.
- Hall, Edward T. [1966] 1990. *The Hidden Dimension*. Garden City, New York: Doubleday.
- Hofstede, Geert H. 1980. *Culture's Consequences: International Differences in Work-Related Values*. Newbury Park, CA: Sage.
- Hofstede, Geert H. 1993. Kultuurit ja organisatiot. Mielen ohjelmointi. Juva: WSOY.
- Hofstede, Geert H. 2001. *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. Thousand Oaks, London, New Delhi: Sage Publications.
- Hofstede, Geert H.; Bond, Michael H. 1984. Hofstede's culture dimensions. – *Journal of Cross-Cultural Psychology* 15, 417–433.
- Gifford, Robert 1997. *Environmental Psychology: Principles and Practices*. 2nd ed. Boston, London, Toronto, Sydney, Tokyo, Singapore: Allyn & Bacon.
- Gudykunst, William B.; Ting-Toomey, Stella 1988. *Culture and Interpersonal Communication*. SAGE Series in Interpersonal Communication. Newbury Park, CA: Sage Publications.
- Gudykunst, William B.; Ting-Toomey, Stella; Sudweeks, Stawart; Steward, Lea P. 1995. *Building Bridges: Interpersonal Skills for a Changing World*. Boston: Houghton Mifflin.
- Gudykunst, William B.; Ting-Toomey, Stella; Tsukasa, Nishida 1996. *Communication in Personal Relationships Across Cultures*. Thousand Oaks, London, New Delhi: Sage Publications.
- Inglehart, Ronald; Baker Wayne E. 2000. Modernization, cultural change, and the persistence of traditional values. – *American Sociological Review* 65, February, 19–51.
- Kiesler, Sara; Cummings, Jonathon N. 2002. What do we know about proximity and distance in work groups? A Legacy of research. – P. Hinds, S. Kiesler (Eds.). *Distributed Work*. Cambridge, MA: MIT Press, 57–80.
- Lepeska, Virginijus 1999. Changes in psychological characteristics of working adults in different stages of economic transition in Lithuania. – Open Society Institute, Center for Publishing Development. *Electronic Publishing Program*, 1–9.
- Mockaitis, Audra I. 2002. The national cultural dimensions of Lithuania. – *Ekonomika* 59, 67–77.
- Mockaitis, Audra I. 2005. A cross-cultural study of leadership attitudes in three Baltic Sea region countries. – *International Journal of Leadership Studies* 1.1, 44–63.
- Nova, Nicolas 2003. Socio-cognitive functions of space in collaborative settings: A literature review about space, cognition and collaboration. – CRAFT Research Report 1. École Polytechnique Fédérale de Lausanne September 2003, 1–35. (Vt ka http://icwww.epfl.ch/publications/documents/IC_TECH_REPORT_200481.pdf (25.01.2007).)
- Nova, Nicolas 2005. A review of how space affords socio-cognitive processes during collaboration. – *PsychNology Journal* 3 (2), 118–148. [http://www.psychology.org/File/PNJ3\(2\)/PSYCHNOLOGY_JOURNAL_3_2_NOVA.pdf](http://www.psychology.org/File/PNJ3(2)/PSYCHNOLOGY_JOURNAL_3_2_NOVA.pdf) (25.01.2007).
- Preston, Paul 2005. Proxemics in clinical and administrative settings. – *Journal of Healthcare Management* 50 (3), 151–154.

- Salo-Lee, Liisa; Malberg, Raija; Halinoja, Raimo 1996. Me ja muut. Kulttuurienvälinen viestintä. Helsinki: Yleisradio.
- Schimmack, Ulrich; Oishi, Shigehiro; Diener, Ed 2005. Individualism: A valid and important dimension of cultural differences between nations. – *Personality and Social Psychology Review* 9 (1), 17–31.
- Samovar, Larry A.; Porter, Richard 2004. *Communication Between Cultures*. Wadsworth Series in Speech Communication. Australia, Canada, Mexico, Singapore, Spain, United Kingdom, United States: Wadsworth/Thomson Learning.

Rene Altrov töötab Eesti Keele Instituudis, uurimisvaldkondadeks on kultuuridevaheline kommunikatsioon ja kõnekorpused.
rene.altrov@eki.ee

CULTURAL INFLUENCE ON PERSONAL SPACE USAGE AMONG ESTONIANS AND LITHUANIANS

Rene Altrov

Institute of the Estonian Language

The article is based on a study that examined cultural influences on personal space usage in different communicative situations at home, at work, and spending one's leisure out of home and office. The aim of the study was to find out the rules governing the use of personal distance among Estonians and Lithuanians, wondering (1) if there could be some differences due to the communicative situation, (2) if there have been any changes in the Estonians' personal space behaviour, and (3) if any differences could be observed in the usage of personal space by Estonians and Lithuanians.

The questionnaire provided situations with four options of behaviour. The response was analysed relying on cultural studies by G. Hofstede and E. T. Hall. The informant groups consisted of ethnic Estonian and ethnic Lithuanian office workers. The Estonian data were gathered on two occasions – in 1998 and 2006, the Lithuanians were questioned in 2006.

The results demonstrated that values and preferences may differ depending on the communicative situation and environment: some cultural norms may work in family, some others at work or leisure. The comparison of the Estonian data of 1998 and 2006 reveals that in most cases the personal space behaviour of Estonians can be associated with certain individualistic and masculine values, but a transition from masculine values to feminine ones can be sensed. As for Lithuanians, their personal space behaviour is more variegated showing up several opposite values, depending on the situation.

Keywords: proxemics, proximity, spatial behaviour, interpersonal distance, personal space, Estonian, Lithuanian

VARIAUTOR – VARIISIK VÕI VARJUKIRJANIK?

Anne Arold

Ülevaade. Kirjutis on ajendatud mõnedest mitmekeelsetes elektroonilistes ja trükitud erialasõnastikes esinevatest eksitavatest termini-kirjetest. Tähelepanu keskmes on autorsuse ja variautorsusega seotud mõistete rühm, eelkõige terminina tunnustatud nimisõna *variautor*, mida on võimalik mitmeti tõlgendada. Täiendsõnaga *vari* moodustatud sõnade rühma analüüsi abil näidatakse, mil viisil liitsõna aktuaalne tähendus sõltub põhisõna semantilisest struktuurist. Põhisõnana esinevate isikunimetuste juures saab seejuures määravaks asjaolu, kas isiku juures osutatakse konkreetsel juhul lihtsalt inimesele kui füüsilisele olendile või isikule kui teatava tegevuse läbivijale *resp.* omaduse kandjale. Liitsõna *variautor* ambivalentsus on tingitud sellest, et põhisõnas on ei tõuse kumbki nimetatud komponentidest piisavalt selgesti teise kõrval esile, mis võimaldab aktualiseerida täiendsõna eri tähendusvariante. Analüüsi tulemusena leiab kinnitust, et uue morfoloogiliselt kompleksse termini kasutuselevõtmisel on vaja täpselt määratleda selle baaselementide semantiline struktuur ning prognoosida nende kombineerimisel käivituv(ad) tõlgendusmehhanism(id), ennetamaks ohte morfosemantiliselt ambivalentsete sõnade tõlgendamisel ja tõlkimisel.

Võtmesõnad: leksikoloogia, leksikograafia, sõnamoodustus, terminoloogia, eesti keel, saksa keel, inglise keel

Leksikaalne polüseemia oskuskeeles

Polüseemia on keeles loomulik nähtus, mis ei pruugi emakeelses suhtluses tähelepanu pälvida, sest konkreetne situatsioon suunab kuulajat/lugejat tõlgendama teksti reeglina just nii, nagu kõneleja/kirjutaja on mõelnud. Samas võimaldab mõistete hägusus suhtluspartnerit ka sihilikult eksitada sel teel, et ambivalentset sõna kasutatakse kontekstis, kus selle tõlgendusvõimalusi on rohkem kui üks. Eesmärgiks võib sel juhul olla tahe väljendada humoorikalt (iironiliselt) või lihtsalt üllatada.

Halvemal juhul peitub sellise väljendusviisi taga soov oma tegelikke mõtteid varjata ja ühese tõlgendamisvõimaluse puudumist enda kasuks pöörata. Tõlkijale võib taotluslikult mitmeti mõistetav sõnakasutus tekstis saada lausa ületamatuks tõkkeks, ning seda mitte ainult ilukirjanduse vahendamisel. Georg Wilhelm Friedrich Hegeli filosoofiline õpetus pakub suurepärase näite selle kohta, kuidas ühe mitmetähendusliku verbi *aufheben* abil võib kokku võtta terve filosoofilise maailmakäsituse tuuma.¹ Hegeli tõlkijailt nõuab selle edasiandmine paratamatult omapoolsete kommentaaride lisamist viitega sõna *aufheben* tähendusvariantidele. Teiselt poolt võib aga lähtekeeles üheselt mõistetava termini sisu keelte tüpoloogilistest ja spetsiifilistest erisustest tingituna sihtkeelses tekstis kasutatavate vastete mitmeti mõistetavuse tõttu hägustuda. Järgnev kirjutis käsitleb mõningaid probleeme polüseemseid elemente sisaldavate terminite tõlgendamisel, tõlkimisel ja leksikograafilisel kirjeldamisel, keskendudes vormiliselt baasmorfeemile *vari(-)/varju-* ning mõisteliselt autorsuse eri vormidele ja nende tähistamisele.

Kes on variautor?

Ühiskonnas, milles kehtivaid suhteid püütakse üha täpsemalt lepingute, reeglite ja eeskirjadega korrastada, ei jää lahterdamistest puutumata ka vabad kunstid. Otse vastupidi – intellektuaalse omandi ja autorsusega seotud küsimused on hõivanud õigusnõustajate ja advokaatide töös arvestatava koha, ehkki autori defineerimine isikuna, “kes on loonud kirjandus-, kunsti-, teadus-, tehnilise vm teose /.../” (VL) ei valmista raskusi üheski keeles. Reeglina on autori nimi avalikkusele teada ja tema autoriõigused teose suhtes seadusega kaitstud. Jättes kõrvale kollektiivse autorsuse, kus kaasautorite panus töö valmimisse võib olla erinev, ja loomevarguse ehk plagiaadi, tuleb praktikas ette teisigi kõrvalekaldeid sellest pealtnäha lihtsast seosest autori ja teose vahel. Kõige üldisemalt saab neid määratleda kui juhtumeid, kus teose loomisega reaalselt seotud isiku(te) nimi/nimed ei ole nende endi tahtel avalikustatud või üheselt tuvastatavad. Valdavalt on siin tegemist olukorraga, kus autor avaldab oma teose kas anonüümselt või varjunime (e pseudonüümi) all. Varjunimi on harilikult fiktiivne, teiste sõnadega – sellenimelist isikut ei eksisteeri, kuivõrd autori taotluseks ei ole anda põhjust siduda oma teost ühegi reaalse isikuga. Tema autoriõigused ei ole sellisel juhul kuidagi ohustatud ning tema üksi on vastutav ka oma teose sisu eest. Keerulisem on lugu siis, kui tegelik looja püüab avalikkust suunata pidama töö autoriks mõnd teist reaalset isikut, kasutades teose avaldamisel viimase nime varjunimena. Siin tekib juba rida juriidilisi probleeme, mis hõlmavad nii materiaalseid kui eetilisi küsimusi: Kellel on õigus küsida honorari? Kes vastutab töö sisu eest? Kellele peaks suunama kriitika? Kuidas on kaitstud tegeliku autori õigused ja tagatud tema vastutus kirjutatu eest ning mil määral peaks teksti eest vastutama avalikkusele autorina esitletav (vari)isik, kellele tegelik autor on andnud üksnes võimaluse oma nime “laenamisega” nõustuda või mitte?

Tekstilooje alal leiab üha enam rakendamist veel kolmaski võimalus, mis inglise keeles on tuntud kui *ghostwriting* ning mis seisneb selles, et tekst väljendab küll selle autoriks peetava (avalikkusele tuntud) isiku mõtteid, kuid nende mõtete verbaliseerijaks on tolle isiku tellimuse anonüümseks jääv täitja, kelle “ametnimetus” on kujunenud *ghost writer*. See nimetus on kas otsese või tõlkelaenuna võetud

¹ Aufheben im Sinne von “Beseitigen”; Aufheben im Sinne von “Bewahren”; Aufheben im Sinne von “Hinaufheben” auf eine höhere Stufe bringen. Vt <http://www.philolex.de/hegel.htm> (20.08.2006).

kasutusele ka mitmetes teistes keeltes, sh eesti keeles.² Lisaks pakub Eesti Rahvusraamatukogu veebileheküljel paiknev raamatukogusõnastik (ERRS),³ mis sisaldab 4400 “Eesti Raamatukoguhoidjate Ühingu terminoloogiatoimkonnas läbiarutatud ja heakskiidetud terminit raamatukogunduse, bibliograafia, raamatuteaduse, info- tehnoloogia jm raamatukogunduse sidusaladelt”, eesti, inglise, vene, saksa, soome ja prantsuse keeles *ghost writer*’i eestikeelseks vasteks liitsõna *variautor*:

eesti	VARIAUTOR
seletus	isik, kelle nime all autor oma teose avaldab
	vt ka arvatav autor
	vt ka fiktiivautor
inglise	ghost writer
vene	тенево́й автор
saksa	Ghostwriter ⁴
soome	haamukirjoittaja ⁵ haamutekijä
prantsuse	auteur supposé (ERRS)

Ehkki “Eesti õigekeelsussõnaraamat ÕS 2006” (ÕS 2006) ega “Eesti kirjakeele seletussõnaraamat” (EKSS) sõna *variautor* hetkel veel ei sisalda, tundub see esmapilgul – võrreldes aeg-ajalt ka eesti keeles kasutatava toorlaenuga – olevat üsna õnnestunud vaste. Samas teeb aga nõutuks lisatud seletus: ‘isik, kelle nime all autor oma teose avaldab’. Nagu ülal märgitud, mõistetakse *ghost writer*’i all inglise keeles hoopiski **tundmatuks** jäävat autorit, kelle töö avaldatakse **võõra** nime all.⁶ Sama mõiste sisaldub ka enamikus teistes kirjes esitatud võõrkeelsetes vastetes. Venekeelse termini *тенево́й автор*⁷ kasutusnäited on seejuures mitmekesisemad.⁸ Eespool toodud tähenduse kõrval on seda väljendit pruugitud ka näiteks poliitiliste telgitaguste kirjeldamisel, tähistamaks märkamatuks jäävaid niiditõmbajaid. Prantsuskeelne vaste *auteur supposé*⁹, mis otseses tõlkes tähendab oletatavat autorit, annab *ghost writer*’i tähendust edasi vaid osaliselt, kuivõrd ei sisalda viidet sellele, et nimetatud isik pole mitte ainult arvatav, vaid ka tegelik autor (või õigemini: teksti tegelik kirjutaja).

ERRS-i kontaktaadressile saadetud lühimärkusele, milles sai juhitud tähelepanu sisulisele vastuolule termini seletuse ja selle võõrkeelsete vastete tähenduste vahel, saabus järgmine vastus:

Varjunime all teoste avaldamise korral ei esitata ju raamatu tiitliandmetes reeglina selle isiku nime, kes tegelikult teksti kirjutas (või arvatakse, et kirjutas). Ka on siin tegu mõiste tajumisega. Vrd ka ÕS: variisik – kelle nime

² Eesti keeles: *ghostwriter*, vt Mele Pesti. Esiteks kirjanikud. Aga naised ikka ka: Anna Gavalda: “.. Olete ühes intervjuus öelnud, et tahaksite olla ghostwriter (keegi kes kirjutab salaja tekste, mida teised oma nime all avaldavad). Kas see oli poos? – Absoluutselt tõsi! Ma ei taha olla avaliku elu tegelane, ma vihkan seda. Ma vihkann seda. Ma ju kuulan ja jälgin inimesi ja kirjutan neist – ma ei saa endale lubada anonüümsuse kaotust! ..” (Eesti Ekspress, 16.11.2005). <http://www.ekspress.ee/viewdoc/21EB5923C1B03C64C22570B4005E3683> (16.3.2006).

³ Vt <http://pang.nlib.ee/termin/> (17.11.2005), edaspidi: ERRS.

⁴ Ghost[w]riter [...raita], der; -s, - [enlgl. ghost-writer, eigtl. = Geisterschreiber] (bildungsspr.): Autor, der für eine andere Person, meist eine bekannte Persönlichkeit, bes. einen Politiker, schreibt u. nicht als Verfasser genannt wird ././ (GWDS) Ghostwriter [...], anonym er Autor, der im Auftrag und unter dem Namen anderer Personen Reden, Zeitungsartikel und Bücher schreibt. (Meyers GTL 1995).

⁵ Vt ka WOD <http://www.websters-online-dictionary.org/> (9.2.2006).

⁶ ghost writer – an unknown writer who produces copy to be signed by persons of prominence. (WOD).

⁷ литературный “раб”, vt WOD, <http://www.websters-online-dictionary.org/> (10.2.2006); автор-“призрак” (фактический автор, работающий на другое лицо), vt http://www.miripravo.ru/voc/voc_0.htm (16.2.2006);

безымянный автор; “негр” (автор, пишущий за другое лицо); автор-“призрак”; фактический автор, работающий на другое лицо, vt Автоматический словарь Мультитран. http://multitran.ru/c/m.exe?t=2996523_1_2 (10.2.2006).

⁸ Vt В. Танасийчук: Ещё одна загадка Эдгара По. <http://nauka.relis.ru/28/0509/28509080.htm> (25.2.2006); Вадим Шахович: Зуев М. Internet: Советы Бывалого Чайника. – Москва: Лаборатория базовых знаний, 1998. Первое сентября 72/1999 <http://ps.1september.ru/1999/72/5-4.htm> (25.2.2006).

⁹ Mitmete allikate põhjal *négre* (Dictionnaire International des Termes Littéraires. <http://www.ditl.info/lex/lexico.php> (10.2.2006); Langue française; <http://www.langue-fr.net/index.htm> (10.2.2006); WOD (10.2.2006); Wilpert 1989).

all keegi midagi teeb. Siit saab tuletada analoogia ka kõnesoleva terminiga, kas sõnade liitmise tulemusena saadud termini teine sõna on “isik” või “autor”, ei muuda siin sisu.¹⁰

Vastuolu *variautori* (= *ghost writer*) mõiste sisu ja selle kohta sõnastikus antud selgituse vahel ei lükka see kommentaar ometi ümber, ning seda kahel põhjusel. Esiteks samastatakse siin *varjunimi* ühtaegu *variisiku* ja *variautoriga*, osutades autori õigusele kasutada varjunime. Kummatigi ei sõltu varjunime kasutamise õiguse aktsepteerimisest või mitteaktsepteerimist midagi seoses variautorsusega. Üks autor võib kasutada ka mitmeid varjunimesid, jäädes seejuures ikkagi ainsaks isikuks, keda nende nimede all ilmunud teostega saab seostada. Variautorsuse korral on seevastu alati tegu **vähemalt kahe** isikuga, kusjuures ka *variisiku* ja *variautori* tõlgendamine ei ole sugugi *a priori* analoogne, nagu väidetakse viimase lausega, vaid siin on oma osa nii täiendsõna *vari* mitmetähenduslikkusel kui ka põhisõnade *isik* ja *autor* erineval semantilisel struktuuril. Kuna täiendsõnaga *vari* moodustatud liitsõnade tõlkimisel esineb vigu ka mõnedes teistes sõnaraamatutes (mõnedele neist osutatakse allpool) ning ka ülal tsiteeritud kirje on seni kohendamata, püüab käesolev kirjutis lahti seletada selliste sõnade tähenduse kujunemismehhanisme, et selliseid vigu oleks võimalik tulevikus ennetada.

Varju semantika liitsõna koosseisus

Nimisõna *vari* hõlmab mõisteid, mis on kõige üldisemalt lahti seletatavad kolme komponendi abil: (füüsiline) objekt, valgusvoog ja objekti taha jääv valgusvoo eest kaitstud ala. EKSS esitab märksõna *vari* (omastav: *varju*) all kümme tähendusvarianti, neist kolm sisaldavad omakorda vähem või rohkem eristatavaid varjundeid:

VARI

- 1 tume kujutis, mis tekib valgusvoo teele jääva keha taga
 - 1a midagi ähvardavat, ebameeldivat
 - 1b tume(dam) värvitoon v varjund; (tumedam) laik; (pildi) varju(s)tatud pind
- 2 (kellegi v millegi ebaselgelt nähtava kujutise kohta)
 - 2a hingelise seisundi peegeldus inimese näol
 - 2b (eitavates väljendites hrl täh.: ‘mitte kübetki, üldsegi mitte’)
- 3 päikese eest kaitstud ja seetõttu suhteliselt pime ja jahe koht, vilu
- 4 kaitse tuule, sademete, hädaohu vms eest; peavari, ulualune
 - 4a (ainsuse kohakäänetes:) ulualla, peitu; uluall, peidus; uluall, peidust
- 5 varjav rajatis, varjualune, ulualune
- 6 varjamine, varjav kate; varjav ese v vahend; sirm
- 7 (kellegi v millegi kohta, mis vaevu, väga nõrgalt meenutab võrdlusobjekti)
- 8 vaim, surnud inimese kehatu kuju
- 9 kurjus, kurjuse maailm
- 10 toit, kõhutäide
idiom. (*halba*) *varju heitma*.

Esitatud tähendused võib jagada kahte rühma olenevalt sellest, kas *vari* tähistab objekti taha jäävat valgustamata ala/kujutist (VARI₁) või valgusvoo teele jäävat, valgust varjavat objekti ennast (VARI₂). Eesti keeles on põhitähenduseks just esimene variant, mis omakorda sisaldab kaht alltähendust:

- a) VARI_{1a} – ‘valgusvoo teele jääva objekti taha tekkiv **kujutis**’ (~ *kuju, siluett*) (sks *Schatten*; ingl *shade, shadow*; vn *мень*);
- b) VARI_{1b} – ‘valguse eest varjatud **koht** või **ala**’ (~ *vilu, varjuline koht*) (sks *Schatten*; ingl *shade*; vn *мень*).

‘Valgusvoo teel asuv objekt’ on *varju* sekundaarne tähendus, mida kajastab ka EKSS, seades selle tähendusvariandi alles 6. kohale:

VARI₂ – ‘valgust (tuult vms) tõkestav objekt, varjav **kate**’ (~ *sirm*) (sks *Deckung, Schutz; Schirm*; ingl *screen, shade*; vn *завеса; прикрытие; защита; ширма; щит; экран*).

Variandid (4) ja (5) hõlmavad mõlemat aspekti, nii ‘varjavat objekti’ kui ‘varjatud ala’:

VARI_{1,2} – ‘kaitse tuule, sademete, hädaohu vms eest; varjav **rajatis, varju-alune**’ (~ *peavari, ulualune; varje*) (sks *Deckung; Obdach, Unterkunft*; ingl *shelter, lean-to; cover*; vn *укрытие, приют, кров*).

Variantide (7)–(10) juures on tegemist metafoorsete kasutusviisidega: VARI_{1a} sisaldub tähendusvariantides (7) ja (8), kus *varjuks* nimetatakse ähmasena, ebareaalsena näivat objekti ennast; VARI_{1b} sisaldub variandis (9) ning VARI₂ variandis (10).

Liitsõnades esineb *vari* nii põhi- kui täiendsõnana. Mis puutub liitsõnadesse, kus *vari* on põhisõnaks (*õhtuvari, kuumavari, inimvari, naeratusevari* jne), siis nende tõlgendamine ei tekita harilikult probleeme, sest põhisõnas ja seega ka liitsõnas tervikuna sisalduv mõiste täidab tekstis referendina iseseisvat rolli, mille ühetähenduslik mõistmine on lauses eksplitsiitsete vahenditega tagatud. EKSS esitab sellise mudeli järgi moodustatud liitsõnad iga tähendusvariandi juures eraldi. Keerulisem on lugu liitsõnadega, milles mitmetähenduslik sõna on täiendsõnaks. Kuna kahest nimisõnast moodustatud liitsõna osade vaheline konkreetne sisuline seos ei ole vormi põhjal üheselt määratletav, sõltub suuresti põhisõna tähendusest, milline täiendsõna semantiline variant vastavas kombinatsioonis realiseerub.

Vaatleme lähemalt EKSS-is esitatud märksõnu, mille täiendkomponent on *vari*. Need liitsõnad, mida on kokku ca 40, võib jagada allrühmadesse vastavalt sellele, milline *varju* tähenduskomponent liitsõnas aktualiseerub ja mil määral on see seotud oma mõistelise vastaspoolega (varjav objekt varjatavaga ja vastupidi) ning kas referent on ise käsitletav varjuna või on sellega mingil muul viisil seotud.

Analüüsi tulemused võtab kokku tabel 1. Suurematest rühmadest on tabelisse kantud ainult kõige iseloomulikumad (tüüp)sõnad.

Analüüsitud liitsõnade puhul on valdavalt tegemist sõnadega, millega tähistatavad mõisted hõlmavad referendi kõrval ka teatavat teist objekti. VARI₁ ja VARI₂ on teineteisega onomasioloogiliselt seotud nagu paberilehe kaks poolt. Erandiks on ainult sõnad, mis sinse jaotuse järgi on paigutatud rühma (5) ja tähistavad (kuju-teldavaid iseseisvaid) varjusarnaseid objekte või nendega seonduvaid mõisteid.

Table 1. Täiendkomponendiga *vari-/varju-* moodustatud nimisõnade semantiline struktuur¹¹

O _r	1	2a	2b	3	4	5
N _{vari+n} [∈ VARI ₂]				<i>varjunimi</i>		
N _{vari+n} ⊂ VARI ₂	<i>varikatus</i> <i>varilaud</i> <i>varisein</i>	[O _x]	[O _x]	<i>variisik</i> <i>variettevõte</i> <i>varifirma</i>	O _x	
N _{vari+n} ⊂ VARI ₁	[O _x]	<i>varjupilt</i> <i>varjupool</i> <i>varjukülg</i>	<i>varivalitsus</i> <i>varikabinet</i>	O _x	? <i>variautor</i>	<i>varjukuju</i> <i>varjumees</i> <i>varjuriik</i>
N _{vari+n} [∈ VARI ₁]		<i>varjumäng</i> <i>varjuteater</i> <i>varjutaim</i> <i>varjuelu</i>	<i>vari-</i> <i>majandus</i>		? <i>varjukirjanik</i>	<i>varjupoks</i>
N _{vari+n} [∈ VARI _{1,2}]	<i>varjusadam</i> <i>varjukraav</i>					

O_r – referent;

N_{vari+n} – liitnimisõna täiendkomponendiga *vari-/varju-*;

O_x – tähistatavas mõistes sisalduv, referendiga lahutamatu seotud objekt;

[O_x] – referendiga mõisteliselt seotud, kuid sõna kasutuskonteksti seisukohalt irrelevantne objekt;

N ⊂ A – nimisõna, millega tähistatav mõiste kuulub mõistesse A;

N [∈ A] – nimisõna, millega tähistatav mõiste on semantiliselt seotud mõistega A.

Rühmadesse (1) ja (3) kuuluvad sõnad, mis tähistavad varjavaid objekte (N_{vari+n} ⊂ VARI₂) või mõisteid, milles varjamine on üheks signifikantseks tunnuseks (N_{vari+n} [∈ VARI₂]). Sealjuures ei ole rühma (1) mõistete (*varikatus*, *varilaud*) juures oluline, **mida/keda** need varjavad. Rühma (3) kuuluvad mõisted (*varjunimi*, *variisik*, *variettevõte*, *varifirma*) on aga varjatava objektiga lahutamatu seotud, see on nende sõnade kasutuskontekstis vähemalt sama oluline (või isegi olulisem) kui referent.

Analoogsed rühmad moodustuvad ka vastaspoolel. Rühma (2) kuuluvad nimisõnad tähistavad millegi varju või sellega seonduvat, kusjuures varjav objekt pole sõna mõistmise ega kasutamise seisukohalt oluline. Rühm (4), mille esindajaks võiks olla *variautor*, on vaadeldav rühma (3) peegelpildina. Kui *variisik*, *variettevõte* ja *varifirma* on avalikkusele teada ning tegutsevad (näiliselt) enda nimel, varjates tegelikke tellijaid, siis *variautor* kui tegija jääb ise reeglina tundmatuks, kujutades endast seega avalikkusele tuntud tellija varju.

Kui täiendkomponent *vari-/varju* on liidetud põhisõnale, mis tähistab füüsilist objekti, siis ei tekita liitsõna tõlgendamine erilisi probleeme: artefakte võib alati seostada teatava otstarbega, antud juhul (sihipärase) varjamisega, looduslike objektide (taimede ja loomade) puhul leiab rõhutamist neile loomuomane suhe varju (varjataluvus v -lembus). Abstraktseid mõisteid sisaldavad liitsõnad (*varjukuju*, *varjukülg*, *varjupool*, *varjukuju*, *varihing*, *varjumäng* jne) on seostatavad pigem teatud objekti varju (varjatud ala) kui varjava objektiga. *Varjusadam* ja *varjukraav* kui rajatised on seotud mõlema variandiga, sisaldades varju mõlemas tähenduses.

¹¹ Sümbolite valikul on toetutud Uno Mereste (2000: 179 jj) käsitlusele liitsõnaliste terminite semantilisest struktuurist ning Jan M. G Aarts ja Joseph P. Calberti monograafias (1979: 25 jj) kasutatud mõistet spetsifitseerivate semantiliste tunnuste (LPRIM) tähistusviisile.

Erilist tähelepanu väärivad liitsõnad, milles põhisõna tähistab kas inimest või tema tegevusega seotud mõisteid, nt *isik, mees, firma, ettevõte, kabinet, valitsus, riik, majandus*, teiste hulgas ka meid huvitav *autor*. Lähtuvalt *variator*i eelpool aktsepteeritud tähendusest (= *ghost-writer*) ei saa seda paigutada kokku ei *variisiku* ega *varivalitsusega*. Inimene kui reaalne füüsiline objekt võib nii varjata (*variisik*) kui ka millegi/kellegi varjus tegutseda (*variautor*), samuti võib ta luua teatud institutsioone millegi varjamiseks (*varifirma*) või varjus tegutsemiseks (*varikabinet*). Millest on tingitud ühe või teise tähendusvariandi aktualiseerumine?

On tuntud tõsiasi, et mitmetähendusliku sõna lähim kontekstipartner tõstab sõna potentsiaalsest tähendusest esile ühe konkreetse semeemi. Pisut vähem on rõhutatud seda, et ka selle semeemi juures on võimalik fokuseerida erinevaid seeme (vt Aarts, Calbert 1979: 48 jj; Arold 2000: 38 jj). Nii võib omadussõnaline täiend iseloomustada kas isikut kui füüsilist olendit või mõnd selle isiku olulist tunnust (isikuomadust, tegevusala jne). Kumb võimalus realiseerub, see sõltub nii nimisõna kui ka lisatava omadussõna semantilisest struktuurist: vrd: *suur poiss – suur laulja; hea mees, laps, kodanik – hea õpetaja, laulja, sekretär* jne. Mõnda fraasi, nagu näiteks *vana pillimees, vana õpetaja*, võib mõista kaheti: nende all võib mõelda nii lihtsalt eakat inimest kui ka pika muusiku- või pedagoogistaaziga isikut (hoolimata vanusest).

Teatavat analoogiat võib täheldada ka liitsõnade juures. Ehkki nii *isik* kui *autor* tähistavad inimest kui elusobjekti, mis võib toimida füüsilise (katte)varjuna, on *autori* juures semantiliselt oluline isiku **tegevus** loojana. Liitsõnas, kus täiendkomponendiks on *vari*, jääb *isik* sel juhul tagaplaanile ja *vari* on tõlgendatav kui varjatud keskkond, kus looja tegutseb.

Selgemini on see nüanss tajutav nn “läbipaistvate” isikunimetuste juures, milles referendi tegevus vm oluline tunnus on eksplitsiitselt väljendatud. Ehk ongi sellega seletatav sama mõiste tähistamiseks kasutatud alternatiivne väljend *varjukirjanik*, mida kasutab nt “ajaveebnik” Siim Teller oma kodulehel, kus ta värskeid lugemismuljeid jagades ühe raamatutegelase kohta muu hulgas kirjutab: “.. Tom on varjukirjanik e. stiilis see mees, kes tegelikult David Beckhami eluloo kirjutab, saab selle eest raha, aga nime raamatusse mitte.”¹²

“Varjude” roll ühiskonnas ja selle kajastused keeles

Kogu arutlus selle üle, millise varjuga ühel või teisel juhul tegu on, ei vääriks nii suurt tähelepanu, kui kirjeldatud probleem oleks eri keeltes enam-vähem ühesugune. Et aga mitmetes keeltes tehakse objekti *varju* ja objekti varjava *varju* vahel vahet ka leksikaalsel tasandil, sõltub tähendusvariandi õigest määramisest sobiva tõlkevaste leidmine. Nii on näiteks *varivalitsus* saksa keeles *Schattenregierung*¹³, *varjuelu – Schattendasein*, *varimajandus – Schattenwirtschaft*¹⁴, *varjuriik – Schattenreich*, *varjumees – Schattenmann*, *varjunimi* aga hoopiski *Deckname*, *varifirma – Scheinfirma*¹⁵ ja *variisik – Strohmännchen*¹⁶, mitte *Schattenperson*, nagu pakuvad uus “Eesti-saksa sõnaraamat” (ESS) ja “Eesti-saksa ärisõnastik” (ÄS), ega ka mitte *Hintermann* (ÄS). Kui nimisõna *Schattenperson* all mõistetakse saksa keeles pigem mingit kummituslikku kuju (vrd *Schattenmann*), siis *Hintermann*

¹² Vt http://www.siiimteller.com/archives/2004/07/mil_millington.html (22.03.2006).

¹³ Ingl *shadow cabinet* (EIS), vñ *теневое правительство*.

¹⁴ Ingl *underground economy, black economy* (EIS); vñ *теневая экономика*.

¹⁵ Ingl *front* (EIS), vñ *теневая фирма*.

¹⁶ Ingl *straw man, frontman* (EIS); vñ *подставное лицо*.

võib küll olla seotud variisikuga, seda aga ainult kui variisiku suunaja, varjatult tegutsev niiditõmbaja.

Nagu toodud näidetest selgub, võib eesti keeles sageli jääda selgusetuks, millist rolli üks või teine “vari” õieti täidab, kes selles “varjude maailmas” keda õieti varjab ja kuidas tuleks nimetada nende partnereid või mõistelisi vastaspooli.

Kõige lihtsam näib olevat lugu *varivalitsusega*, mis tegutseb küll realselt võimalikult oleval valitsuse varjus, kuid mitte selle otsese “varjuna”, vaid siiski iseseisvalt, juhindudes parajasti päevakorral olevatest probleemidest. Muutused valitsuse struktuuris ei pruugi kaasa tuua analoogilisi muutusi varivalitsuses.

Variisikutest, variettevõtetest ja varifirmadest saab rääkida seoses *varimajandusega*. Viimane eksisteerib suuresti tänu esimestele ja toimib nende kaudu. Sealjuures ei ole mõtet otsida varimajandusest firmade või ettevõtete analoogilisi struktuure. Allilm ei vaja legaalsete firmade koopiaid, seal aetakse äri oma seaduste järgi ning kehtivad omaette hierarhiad, mis kõik kokku moodustavad varimajanduse. Üks variisik võib tegutseda terve grupeeringu huvides, seda oma isikuandmetega varjates, nii nagu ka üks mafiooso võib varjuda mitme variettevõtte taha ja neid varjatult juhtida.

Variautori puhul on kahe isiku seos ühelt poolt kõige selgem ja otsesem: üks kirjutab (teise tellimusel), teist loetakse autoriks. Teiselt poolt jääb aga siin tihti ebaselgeks, kes on ikkagi **tegelik** autor. Kui tellimus esitatakse sel moel, nagu see toimus E. Vilde “Pisuhännas”, kus variautor Piibelehele anti vabad käed otsustamiseks, millest ja kuidas kirjutada, et insener Sanderile auhind välja teenida, siis pole kahtlust, et variautor = tegelik autor. Praktilises elus nii puhtaid “kaasusi” siiski peaaegu ei kohta. Variatorsuse spekter on väga lai, hõlmates erinevaid teenuseid, mida võib seostada konsultandi, nõuniku, erasekretäri vms tegevusalaga, kus kirjutaja saab töö tellijalt konkreetse ülesande, millest kirjutada, ja ka üldised juhised, kuidas seda teha. Seda ei peeta tänapäeval kummagi poole suhtes taunimisväärseks. Viimasel ajal on tõlke- ja toimetamisteenuste kõrvale tekkinud koguni omaette ametlikult tunnustatud (?) teenuseliik: *professional ghostwriting*, mille pakkujad avalikustavad (ilmselt siiski kokkuleppel tellijatega) koduleheküljel “oma” õnnestunud teoseid. Mil määral saab selliste tellimustööde puhul rääkida kirjutaja autorsusest, on mõlema osapoole sisetunde ja vastastikuse kokkuleppe küsimus. Eriti intrigeerivaks võib see küsimus kujuneda akadeemilise variautorsuse (ingl *academic ghostwriting*) alal.

Variautori abi on kasutanud eelkõige paljud poliitikud. USA presidentide Ricard M. Nixon ja Ronald Reagani kõnesid kirjutas väidetavalt Kenneth Khachigian¹⁷, George W. Bushi omasid Michael Gerson¹⁸. Neid kirjamehi nimetatakse ajakirjanduses ka lihtsalt *speechwriter*’iteks (kõnekirjutajateks), ehkki paralleelselt tuleb ette ka *ghost writer*. Selliseid näiteid võib leida ka lähemalt, eestikeelses ajakirjanduses küll enamasti vaid vihjamisi:

- (1) PR-mehed palkaksid Savile kõigepealt osavad kõnekirjutajad. Paavo Pettai (Idea) soovitaks Mihkel Mutti, aga kuuldavasti kasutab Mutti Tulviste. Mitmed pakuvad Lennart Merit ennast. Et siis ei kao head kõned ka tulevikus.¹⁹
- (2) Tuntud intellektuaal, presidendi kõnede kirjutaja, Tartu ülikooli professor, Edgar Savisaare tõlk ja Vikerraadio “Rahva Teenrite” sõltumatu ajakirjandusekspert Rein Veidemann käis Soomes.²⁰

¹⁷ Vt Markus Larsson. Iraagi probleem Bushile tähtsaim. EPLOnline. 30.1.2003. <http://www.epl.ee/?artikkel=226483>.

¹⁸ Vt Gregg Zoroya. He puts words in Bush's mouth. USA Today 4.10.2001. <http://www.usatoday.com/life/2001-04-11-bush-speechwriter.htm>.

¹⁹ Vt Madis Jürgen. See on ju see onu, kes eile Henri juures saunas olid Eesti Ekspress 10.5.2001. <http://www.ekspress.ee/Arhiiv/2001/19/Aosa/kuum10.html>.

- (3) “Kuna kirjanik Rein Veidemann on presidendi kõnede autor, siis taotleb ta autoriõigust, mis tähendab, et tekstis ei tohi muudatusi teha. Kuid president Rüütel tegi seda,” märkis anonüümseks jääda soovinud allikas.²¹

Näeme, et tipp-poliitikute võimalikest *ghost writer*’itest rääkides ei nimetata neid mitte *variautoriteks*, mis eelnevat arvesse võttes võiks olla omal kohal, vaid *kõnekirjutajateks* (*kõnede kirjutajateks*), kolmandas näites ka *kõnede autoriks*, kusjuures eelistatud on läbipaistva nimetuse kasutamine, mille sidumine täiendsõnaga *vari-/varju-* seostub pildiga, et kirjutamine toimub avalikkuse eest varjatult ehk varjus. *Autori* puhul sugeneb aga ambivalentsus just asjaolust, et ehkki kirjutamine/loomine on küll *autori* tähenduses oluline, ei tõuse see võõrsõnas piisavalt esile ja seepärast kiputakse *variautorit* tugevamalt seostama füüsilise isikuga, kes peaks justkui kedagi varjama (nagu *variisik*), mitte aga ise varjus tegutsema (mida kinnitab ka ülal tsiteeritud vastus minu märkusele). Teiseks põhjuseks pean siin aga seda, et ka kõige meisterlikum kõnekirjutaja formuleerib ikkagi eelkõige tellija mõtteid, seega on tema autorsus tegelikult küsitav.²²

Ääremärkuse korras võiks siinkohal ka küsida, kuidas tuleks nimetada selliste tekstide tellijat. Eesti keeles pole selle kohta eraldi sõna kasutusele võetud. Saksa keeles pakuvad mitmed sõnaraamatud ka selle isiku tähistamiseks inglise hõnguga sõna, nimelt *Mummy*.

Ghostwriter (engl. = Geisterschreiber), Schriftsteller, die beruflich im Auftrag und anonym unter dem Namen ihrer Auftraggeber (**mummies**), bes. Politiker, Filmschauspieler, Sportler, Industrieller u. a. Personen von öffentl. Interesse oder überlasteter Erfolgsschriftsteller (z. B. A. Dumas d. Ä.), meist nach genauer Absprache Bücher, bes. Memoiren, oder Reden verfassen, bes. In Amerika; in Frankreich *nègre* genannt. (Wilpert 1989)

Mum|my [mʌmi], der; -s, -s [engl. mummy = Mumie]: *Auftraggeber eines Ghostwriters*. (GWDS; FWB)

Mummy [...] *m.6 Auftraggeber eines Ghostwriters [zu engl. mum ‘schweigend’]*. (Wahrig 2002)

Ilmselt on siin aga tegu pseudolaenuga, nagu seda on *Handy*, *Dressmann* jmt, sest inglise keeles nimisõnal *mummy* niisugust tähendusvarianti ei ole – vähemalt ei ole seda võimalik leida üldkasutatavatest sõnaraamatutest ega ka otsingumootorite kaudu. Küllap ei ole selle järele ka erilist praktilist vajadust.

Kokkuvõte

Mõiste *variautor* ühetähenduslik määratlemine, s.t palgatud tekstikirjutaja eristamine tegelikku autorit varju jätvast variisikust, on vajalik nii juriidilisest kui lingvistilisest vaatepunktist. Autoriõiguste kaitse seisukohalt oleks oluline kokku leppida, kas ja kui võrd tuleks variautorsuse puhul kõnelda vastutuse jagamisest kahe isiku vahel. Ennekõike eeldaks see üksmeelt selles, keda õieti variautoriks nimetada. Ülal toodud argumentidest lähtuvalt võiks sõna *variautor* olla kasutatav kahes vastandlikus kontekstis:

²⁰ Vt Mart Soidro. Tühised lorad versus tõsised asjad. Sirp. 15.07.2005. <http://www.sirp.ee/arhiiv.php>.

²¹ Vt Tuuli Koch. Veidemannid tüüls Rüütli kantsleji juhiga. EPLOnline 5.3.2002. <http://www.epl.ee/artikkel.php?ID=197057>.

²² Ehkki ERSs seda ei kinnita, torkab eri allikate võrdlemisel silma, et ka teistes keeltes välditakse tellimustöö kirjutaja nimetamist autoriks.

- a) Ehkki raamat ilmus N nime all, ei kirjutanud N seda ise, vaid palkas seda tegema **variator**, kelleks oli tõenäoliselt X.
variator = ghost writer = varjukirjanik (tegelik kirjutaja/autor)
- b) Ehkki raamat ilmus N nime all, on tema kuuldavasti üksnes **variator**, tegelik autor on hoopis X, kes kasutas N nime pseudonüümina.
variator = variisik

Kuivõrd *variisik* hõlmab üldmõistena ka fiktiivset autorit, tähistades isikut, kelle nime all autori teos avaldatakse, leiab käesoleva kirjutise autor, et sõna *variator* sobiks juurutada esimeses (seega ERSS-is juba fikseeritud) tähenduses. Sõnastikes tuleks aga sellele anda õige seletus: *variator* – võõra (harilikult tuntud isiku) nime all avaldatud teksti (anonüümseks jääv) kirjutaja. Nõnda on seda juba tehtudki nt Tartu Ülikooli raamatukogu elektroonilise kataloogi märksõnastikus, kust võib leida kirje: *Variatorid, ghostwriters, ghostwriting; ghostwriter – faktiline autor, kes kirjutab mõne nimeka isiku kõned või tema nimel ilmuvad kirjutised*.²³ Rõhutamaks variatori rolli tellimuse täitjana, võiks seda vastavalt vajadusele asendada ka konkreetsemate nimetustega, nt *kõnekirjutaja, memuaarikirjutaja* (kes erinevalt oma mälestusi üles kirjutavast *memuaristist* paneb kirja kellegi teise memuaare) või Siim Telleri poolt kasutatud *varjukirjanik*. Viimati nimetatud varianti sobiks kasutada eelkõige seoses tellimustööna valminud ilukirjandusteosega.

Kirjandus

- Aarts, Jan M. G.; Joseph P. Calbert 1979. Metaphor and Non-Metaphor. The Semantics of Adjective-Noun Combinations. Linguistische Arbeiten 74. Tübingen: Niemeyer.
- Arold, Anne 2000. Kontrastive Analyse der Wortbildungsmuster im Deutschen und im Estnischen (am Beispiel der Aussehensadjektive). Dissertationes Philologiae Germanicae Universitatis Tartuensis 1. Tartu.
- EIS = Veldi, Enn 2002. Eesti-inglise sõnaraamat. Tallinn: Koolibri.
- 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.
- EKSS = Eesti kirjakeele seletussõnaraamat. 1988 –. Eesti Keele Instituut. Tallinn.
- ESS = Kann, Kallista; Kibbermann, Elisabeth; Kibbermann, Feliks; Kirotar, Salme 2003. Eesti-saksa sõnaraamat. 5., parandatud ja täiendatud trükk. Tallinn: Valgus.
- FWB = Duden. Fremdwörterbuch 1990. 5., neu bearbeitete und erweiterte Auflage. Dudenverlag.
- Mereste, Uno 2000. Terminoloogilise liitsõnamoodustuse semantilisi aluseid. – Oskuskeel ja seaduste keeleline rüü. Artikleid ja lühiaurimusi. Tallinn: Eesti Keele Sihtasutus, 175–217.
- Meyers GTL 1995 = Meyers großes Taschenlexikon in 24 Bänden. Hrsg. und bearb. von Meyers Lexikonredaktion. Mannheim, Leipzig, Wien, Zürich: BI Taschenbuchverlag.
- VL = Vääri, Eduard; Kleis, Richard; Silvet, Johannes 2000. Võõrsõnade leksikon. Kuues, uuendatud trükk. Parandanud ja täiendanud Eduard Vääri. Tallinn: Valgus.
- Wahrig, Gerhard 2002. Deutsches Wörterbuch. Neu herausgegeben von Renate Wahrig-Burfeind. Mit einem "Lexikon der deutschen Sprachlehre" Bertelsmann Lexikon Institut. 7. Auflage. Gütersloh/München: Bertelsmann.
- Wilpert, Gero von 1989. Sachwörterbuch der Literatur. 7., verb. u. erw. Aufl. Kröners Taschenausgabe, Bd. 231. Stuttgart: Kröner.

ÕS 2006 = Eesti õigekeelsussõnaraamat ÕS 2006. Tiiu Erelt (toim.), Tiiu Erelt, Tiina Leemets, Sirje Mäearu, Maire Raadik (koost.). Eesti Keele Instituut. Tallinn: Eesti Keele Sihtasutus.

ÄS = Hiimäe, Mari; Püssa, Katrin; Schulz, Anu 1994. Eesti-saksa ja saksa-eesti ärisõnastik. Tallinn: TEA Kirjastus.

Võrgumaterjalid

Автоматический словарь Мультитран. <http://www.multitran.ru/> (10.2.2006).

Dictionnaire International des Termes Littéraires. <http://www.ditl.info/lex/lexico.php> (10.2.2006).

ERRS = Eesti Rahvusraamatukogu raamatukogusõnastik. <http://pang.nlib.ee/termin/> (17.11.2005).

GWDS = Duden. Das große Wörterbuch der deutschen Sprache in 10 Bänden. 1999–2004. Aktualisierte Online-Ausgabe. Mannheim, Leipzig, Wien, Zürich: Dudenverlag. Bibliographisches Institut F. A. Brockhaus AG, Mannheim. <http://www.duden.de/index.php> (12.3.2007).

Jürgen, Madis. See on ju see onu, kes eile Henri juures saunas oli! Eesti Ekspress 10.5.2001. <http://www.ekspress.ee/Arhiiv/2001/19/Aosa/kuum10.html> (10.2.2006).

Koch, Tuuli. Veidemannid tülis Rüütli kantselei juhiga. EPLOnline 5.3.2002. <http://www.epl.ee/artikkel.php?ID=197057> (10.2.2006).

Langue française. Éditeur Luc Bentz. <http://www.langue-fr.net/index.htm> (10.2.2006).

Larsson, Markus. Iraagi probleem Bushile tähtsaim. EPLOnline. 30.1.2003. <http://www.epl.ee/?artikkel=226483> (10.2.2006).

Möller, Peter. Philolex. <http://www.philolex.de/hegel.htm> (20.8.2006).

Pesti, Mele. Esiteks kirjanikud. Aga naised ikka ka: Anna Gavalda. Eesti Ekspress 16.11.2005. <http://www.ekspress.ee/viewdoc/21EB5923C1B03C64C22570B4005E3683> (16.3.2006).

Soidro, Mart. Tühised lorad *versus* tõsised asjad. Sirp 15.7.2005. <http://www.sirp.ee/arhiiv.php> (10.2.2006).

Zoroja, Gregg. He puts words in Bush's mouth. USA Today 4.10.2001 <http://www.usatoday.com/life/2001-04-11-bush-speechwriter.htm> (16.3.2006).

Танасийчук, Виталий. Ещё одна загадка Эдгара По. <http://nauka.relis.ru/28/0509/28509080.htm> (25.2.2006).

Tarand, Kaarel. Kelle laulu nad laulavad? Eesti Päevaleht 12.7.2003. <http://www.epl.ee/artikkel/240274> (12.3.2006).

Tartu Ülikooli Raamatukogu elektroonilise kataloogi märksõnastik INGRID. <http://ingrid.utlib.ee/> (22.11.2006).

Teller, Siim. Ajaveeb Teller plöksib. http://www.siimteller.com/archives/2004/07/mil_millington.html (29.7.2004).

WOD = Webster's Online Dictionary with Multilingual Thesaurus Translation. <http://www.websters-online-dictionary.org/> (10.2.2006).

Wortschatz-Lexikon. http://wortschatz.uni-leipzig.de/index_js.html (09.2.2006).

Шахович, Вадим. Зуев М. Internet: Советы Бывалого Чайника. – М.: Лаборатория базовых знаний, 1998. Первое сентября 72/1999 <http://ps.1september.ru/1999/72/5-4.htm> (25.2.2006).

Anne Arold (Tartu Ülikool) on erialalt saksa filoloog, uurimisvaldkonnad: leksikoloogia, leksikograafia, kontrastiivne lingvistika (sõnamoodustus, valents).

Anne.Arold@ut.ee

VARIAUTOR – STROHMANN ODER GHOSTWRITER?

Anne Arold

Universität Tartu

Den Anlass für den Beitrag gaben einige fehlerhafte Eintragungen in elektronischen und Printwörterbüchern. Das Hauptaugenmerk wird auf die als Fachwort akzeptierte Personenbezeichnung *variator* gerichtet, die mehrere Deutungen ermöglicht. Die Analyse einer Gruppe von Zusammensetzungen mit dem polysemen Bestimmungswort *vari* zeigt, auf welche Weise die Deutung eines komplexen Wortes von der semantischen Struktur des jeweiligen Grundworts abhängt. Eine Personenbezeichnung kann sich als Grundwort einer Zusammensetzung entweder auf die Person als ein menschliches Wesen oder auf die Person als Handlungsträger bzw. Träger einer Eigenschaft beziehen. Wenn die beiden Möglichkeiten gleichermaßen aktualisiert werden können, kann das Kompositum mit dem polysemen Erstglied *vari* gelegentlich nach zwei einander ausschließenden Mustern interpretiert werden. Um Missverständnissen bei der Deutung und Übersetzung von morphosemantisch ambivalenten Termini vorzubeugen, sollte diese Besonderheit bei der Prägung eines neuen Terminus unbedingt berücksichtigt werden, indem man die semantische Struktur der jeweiligen Konstituenten genau ermittelt und die Deutungsvarianten ihrer Kombinationen prognostiziert.

Schlagwörter: Lexikologie, Lexikografie, Wortbildung, Terminologie, Estnisch, Deutsch, Englisch

PARENTAL REACTIONS TO THE MORPHOLOGICALLY CORRECT AND INCORRECT UTTERANCES OF CHILDREN

Ingrida Balčiūniene

Abstract. The acquisition of the first language is a long and complicated process. Besides incorrect pronunciation, one can observe various morphonological, grammatical, and lexical errors in the speech of young children. The goal of the present research was to analyse the parental reactions both to the correct and incorrect utterances of children and to discuss the dependence of parental reactions on the errors of the child's speech. The analysis is based on longitudinal data from the conversations between a Lithuanian girl (from the age of twenty months to thirty-two months) and her mother. The results of the study indicate that parents rather tend to react to the form than to the content of the incorrect utterances of children. On the other hand, the responses to the correct utterances are related more often to the content.

Keywords: child language, language acquisition, longitudinal corpus, conversation analysis, psycholinguistics, Lithuanian

1. Data and research methods

The acquisition of the first language is a long and complicated process. Besides incorrect pronunciation, one can observe various morphonological, grammatical, and lexical errors in the speech of young children. Parental reactions to the grammatically incorrect utterances of children may vary depending on the language type or the cultural and individual features of adult speech. However, such reactions differ from those to correct utterances. Moreover, these differences are universal and exist in most languages that were investigated. Previous research (Chouinard, Clark 2003) has discovered that parents tend to react to the errors in the speech of children by using special linguistic instruments (e.g. special questions such as "What did you say?" and "What are you trying to say?", remarks such as "I can't

understand you” and corrections such as “Not ..., but...” etc.). The main idea was that the differences between the parental reaction to correct and incorrect utterances can help the child in detecting and memorizing the errors of his or her speech. However, before proposing a hypothesis of the influence of the differences between parental reactions to correct and incorrect utterances on the development of the child’s grammar, one has to prove the statistical reliability of these differences.

The aim of the present study is to analyse and to calculate the parental reactions to both the correct and incorrect utterances of children and to discuss the dependence of parental reactions on the errors of the child’s speech. The analysis is based on the longitudinal data of Monika, a Lithuanian girl, and her mother. Their conversations were recorded by Monika’s mother (the author of this paper) in natural everyday situations: playing together, performing common tasks, and so on. The recordings started in March 2000 when Monika was one year and eight months old (age notation 1;8), and continued until she was two years and eight months old (2;8). The 15–20-minute conversations were recorded twice or three times per week. The total duration of recordings is 43 hours. The recordings were transcribed and partially coded according to the requirements of CHILDES (Child Data Exchange System, MacWhinney 1991). The longitudinal data of natural spontaneous conversations allows the researcher to analyse the development of children’s and child-directed speech and to discuss their relationship.

Monika was born in a middle-class family. During the period under study Monika did not attend any day-care centre and had no constant communication with other children; her language was mainly influenced by adult speech. Both her parents speak the north-eastern dialect of Lithuanian, but they use standard Lithuanian when talking to their daughter. Monika speaks standard Lithuanian, but also some phonetic and morphological features of the dialect were observed in her speech.

The study analyses the conversations between Monika and her mother that recorded when she was from one year and eight months to two years and eight months old¹.

2. Theoretical background

During the first phase of the research all the parental utterances were categorized as particular types of reactions. According to W. U. Dressler *et al.* (2006), all parental reactions can be divided into two main categories: **conversational** and **metadiscursive reactions**. Conversational parental reactions are related to the **content** of the child’s previous utterance, whereas metadiscursive reaction focuses on the **form**. According to this classification, parents can either react to *what was said by the child* (in the case of conversational reaction), or focus on *how it was said* (metadiscursive reaction). Both conversational and metadiscursive reactions can be divided into special types, such as *repetitions*, *expansions*, and *corrections* of the child’s utterances, and *questions*. In some cases parental responses are not related directly to the previous utterance of the child (e.g. the adult starts a new topic), but during the analysis presented in this paper all of responses were marked as conversational reactions (as part of the natural flow of conversation).

¹ All the conversations were recorded in 2000–2001 with the primary purpose of collecting data on Lithuanian child language. The parental speech of these conversations is analysed from 2004 only, so the mother’s language could not be influenced during the target period.

Table 1. Distribution of mother's reactions (1;8–2;8)

Number of reactions	1;8	1;9	1;10	1;11	2;0	2;1	2;2	2;3	2;4	2;5	2;6	2;7	2;8
Metadiscursive	133	266	221	160	86	146	24	92	72	116	22	50	60
Percentage of metadiscursive reactions	11%	24%	17%	16%	15%	14%	9%	9%	13%	7%	2%	6%	8%
Conversational	1102	841	1096	865	498	871	241	904	462	1464	904	793	702
Total	1235	1107	1317	1025	584	1017	265	996	534	1580	926	843	760

The analysis of the mother's coded reactions shows that the conversational reactions were more frequent during the period of analysis (see Table 1).

One can observe that metadiscursive reactions are more frequent at an earlier age of Monika (1;8–2;4) when the child is unable to pronounce some sounds (such as *r*, *l*, *s* etc.). Thus, the parent cannot comprehend many of her utterances and asks her to repeat or to explain them; starting with two years and four months metadiscursive reactions are rarely observed. The results of the analysis lead to the conclusion that during natural spontaneous conversation parents tend to react to the content of child's utterances more frequently than to the form.

3. Research findings

In order to test the hypothesis about the differences between the parental reaction to correct and incorrect children's grammar, all Monika's utterances were divided into correct and incorrect utterances, and each group was analysed separately.

The conversational parental reaction is more frequent both to the correct and incorrect utterances of the child, but one could observe a significant difference between the number of the metadiscursive reactions to the correct and incorrect utterances. The metadiscursive reactions account for 26% of reactions to grammatically incorrect utterances and for only 8% of all the reactions to grammatically correct utterances (see Figure 1).

3.1. Parental reactions to grammatically correct utterances of the child

When the child produces grammatically correct utterances, the parents react mostly (79–96%, see Figure 2) to the content and maintain the natural flow of conversation.

The most typical **conversational reactions** to correct utterances are expansions and questions. A conversational **expansion** can be defined as an expansion of the content of the child's previous utterance by the adult. In the earlier stage of Monika's speech acquisition (from one year and eight months to two years and one month), the mother expands most of her single-word utterances by (1) interpreting them according to the situation and the content of the conversation. Later (from two years and three months) the mother often expanded Monika's utterances by adding a particular modal verb (2) or an additional clause (3):

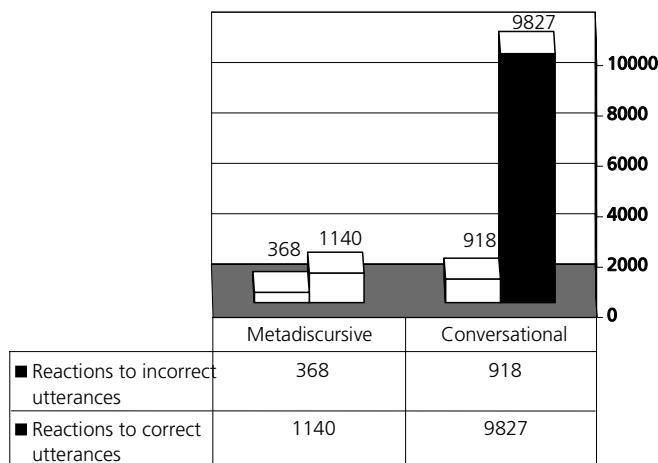


Figure 1. Distribution of the mother's reactions to grammatically correct and incorrect utterances of the child

- (1) CHILD: Teta.
'Aunt.'
MOTHER: Ir teta šneka.
'The aunt is also talking.' (1;8)
- (2) CHILD: Tėtukas vairuoja.
'Daddy is driving.'
MOTHER: Tėtukas moka vairuoti.
'Daddy is able to drive.' (2;1)
- (3) CHILD: Uždarytas.
'[window is] closed.'
MOTHER: Uždarytas tikrai, tėtukas nesusąls.
'[window is] really closed, daddy will not get cold.' (2;5)

The conversational **questions** are also very frequent reactions to the grammatically correct utterances of children. One can observe that these reactions are mostly used with the intention of maintaining conversation and are not used to elicit any information from the child (4):

- (4) Situation: Monika asks the mother to give her the cherries.
CHILD: Duok.
'Give [me].'
MOTHER: Ką mes dabar darysim? Šnekėsime apie vyšnią, ką? Šnekėsime?
Monika vyšnias valgė?
'What will we do now? Will we talk about the cherry? Will we talk?
Did Monika eat cherries?' (2;3)

Other conversational reactions, such as the **repetitions** and **corrections**, are not frequent in the mother's speech.

Metadiscursive reactions contain not more than 16% (the average being 8%) of all the mother's reactions to the child's grammatically correct utterances. They become rarer during the later period (from 2;4). **Questions** are the most typical

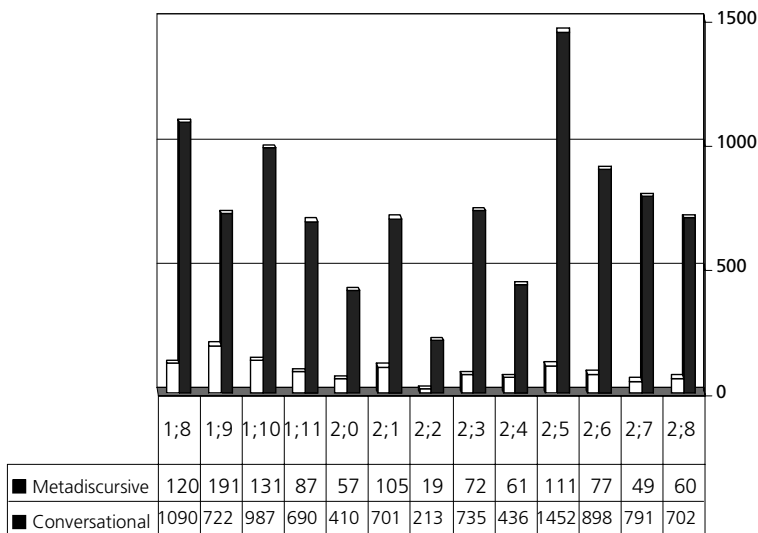


Figure 2. Distribution of the mother's reactions to grammatically correct utterances of the child

metadiscursive reactions to the grammatically correct utterances of the child. The mother uses them mostly when:

- 1) asking Monika to repeat (5) or to explain (6) the incomprehensible utterances:
 - (5) CHILD: Sėsk.
'Sit down.'
 - MOTHER: Ką sakai?
'What [do you] say?' (1;8)
 - (6) CHILD: Ant stalo pinigėlis.
'Coin [is] on the table.'
 - MOTHER: Kas ant stalo?
'What [is] on the table?' (2;3)
- 2) realizing that she has understood Monika's previous utterance correctly (7):
 - (7) CHILD: Guma.
'Gum.'
 - MOTHER: Guma, Monika, pasakei?
'Gum, Monika, [did you] say?' (1;9)

The **reformulations** constitute the second group of the most typical metadiscursive reactions of the mother to the grammatically correct utterances of the child. At an earlier stage (1;8–2;3) the mother reformulates Monika's utterances mostly by changing the morphological category of the child's word (8).

- (8) CHILD: Uždėsim tabiuką.
'[we] will put the tubby on.'
- MOTHER: Uždėk tabiuką.
'Put:IMP the tubby on.' (2;3)

Later on (from 2;4), when Monika is already able to use longer and more complicated constructions, the mother often reformulates the word order of the daughter's phrases (9).

- (9) CHILD: Galvytė gera.
'The head [is] good.'
MOTHER: Gera galvytė.
'Good head.' (2;5)

Lithuanian word order is not rigid, and both Monika's and her mother's utterances are grammatically correct. Thus, this type of parental reaction must be marked in Lithuanian as reformulation and not correction.

3.2. Parental reactions to grammatically incorrect utterances by the child

When the child produces a grammatically incorrect phrase, the mother often reacts (11–52%) to the form of the utterance (see Figure 3).

The most frequent **conversational reactions** to the grammatically incorrect utterances of children are **expansions** of the child's previous phrases (10) and **questions** (11).

- (10) CHILD: Mamytė, spausk.
'Mummy:NOM [=VOC]² push:IMP'
MOTHER: Taip, tik mamytė gali šitą spausti.
'Yes, only mummy can push this.' (2;1)
- (11) CHILD: Nukrapštyti.
'Scratch [= + modal verb].'
MOTHER: Ką tu ten nukrapštysi?
'What will you scratch?' (2;5)

Both of these reactions are related to the content but not to the form of the child's phrase. The mother does not stop the flow of natural conversation and does not demonstrate that Monika has produced an erroneous utterance. Conversational reaction to incorrect phrases can be related to parental intention to maintain conversation, to avoid disturbing the child and making her confused.

The frequency of **metadiscursive reactions** mostly depends on the type of the child's error. Particular errors, such as incorrectly used nominal forms, are followed by metadiscursive parental reactions more frequently (up to 64%), whereas other errors tend to be ignored more often. For example, metadiscursive reactions contain only 18% of all the parental reactions to incorrectly used verbal forms (Balčiūnienė 2006).

The most typical reactions to grammatically incorrect utterances are the **corrections** of the child's previous phrases. Parents can correct the child either directly (12) or implicitly (13).

- (12) MOTHER: Dėsim va tą baltą.
'[we] will put this white one on.'
CHILD: Šitas [= šitą].
'This:NOM [= ACC].'

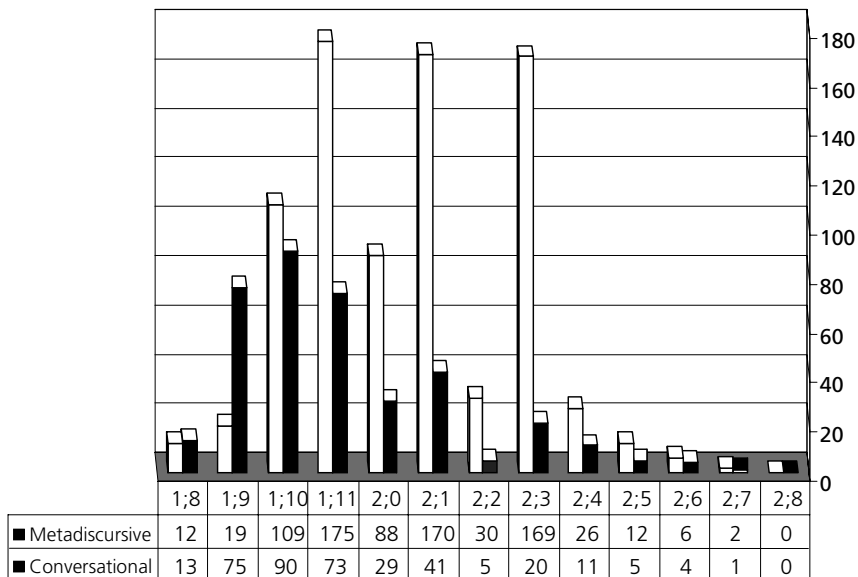


Figure 3. Distribution of the mother's reactions to grammatically incorrect utterances of the child

MOTHER: Šita.

'This:ACC.' (1;11)

(13) Sit: Monika is playing with coins.

CHILD: Sudėti.

'Put [= + modal verb].'

MOTHER: Į taupyklę reikia sudėti.

'[you] need to put [them] to the box.' (2;1)

The **questions** are also used frequently as metadiscursive reactions to children's incorrect utterances. The mother does not correct Monika's speech; by asking a metadiscursive question she directs the child's attention to the incorrect part of her previous utterance and expects Monika to detect the error herself (14), (15).

(14) CHILD: Šunis [= šuo].

'Dog.'

MOTHER: Šunis, pasakei?

'Dogs, did you say?' (1;10)

(15) CHILD: Pušinas [= tušinukas].

'Pen.'

MOTHER: O "pušinas" – tai tušinukas?

'Does 'pušinas' mean a pen?' (1;10)

Other metadiscursive reactions (**reformulations** and **expansions**) are rarely used as responses to the grammatically incorrect utterances of children.

4. Conclusions

In summary, the following conclusions can be made. In natural spontaneous conversations parents generally tend to react to the content of the child's previous utterances more frequently (74–92%) than to the form. The reliable difference between the distribution of the conversational and metadiscursive parental reactions was discovered while analysing them separately and comparing the parental reactions to the correct and incorrect utterances. When the child produces grammatically correct utterances, the parents usually react (79–96%) to the content of the utterance, whereas only 48–89% of incorrect utterances are followed by the parental reaction to their content. The most typical parental reactions both to the grammatically correct and incorrect utterances of children include conversational expansions and questions and metadiscursive questions. The metadiscursive corrections are used very frequently as a reaction to incorrect utterances.

References

- Balčiūnienė, Ingrida 2006. Parental reactions to the morphology of child utterances. – Paper presented in the workshop of the crosslinguistic project on pre- and protomorphology in language acquisition. Vienna, Vienna University. 2–4 February.
- MacWhinney, Brian 1991. *The CHILDES Project: Tools for Analyzing Talk*. Hillsdale: NJ Lawrence Erlbaum Associates.
- Chouinard, Michelle M.; Clark, Eve V. 2003. Adult reformulations of child errors as negative evidence. – *Journal of Child Language* 30, 637–669.
- Dressler, Wolfgang U.; Kilani-Schoch, Marianne; Balčiūnienė, Ingrida; Korecky-Kröll, Katharina; Laaha, Sabine 2006. The learnability of morphology is due to positive and negative evidence. Adult reactions to children's development of French, Lithuanian and German inflection. – Paper presented in the 12th International Morphology Meeting, workshop on acquisition and impairments of inflectional morphology. Budapest, Hungarian Academy of Sciences, 25–28 May.

Ingrida Balčiūnienė, Department of the Lithuanian language at Vytautas Magnus University (Kaunas, Lithuania), research interests: applied linguistics, language acquisition (mainly based on the longitudinal observation of child language development), conversation analysis.
i.balciuniene@hmf.vdu.lt

VANEMA REAKTSIOONID LAPSE MORFOLOOGILISELT KORREKTSETELE JA EBAKORREKTSETELE LAUSUNGITELE

Ingrida Balčiūnienė

Vytautas Magnuse nimeline Kaunase Ülikool

Esimese keele omandamine on kauakestev ja komplitseeritud protsess. Lisaks ebakorrektselle hääldusele esineb laste varases kõnes mitmesuguseid morfonoloogilisi, grammatilisi ja leksikaalseid vigu. Uurimuse eesmärk oli analüüsida nii kvalitatiivselt kui kvantitatiivselt vanema reaktsioone lapse korrektsetele ja ebakorrektsetele lausungitele ja uurida vanema reaktsiooni sõltuvust vea tüübist. Analüüs põhineb leedu tüdruku (1;8–2;8) ja tema ema vestluste longitudinaalsel uuringul. Uurimistulemused näitavad, et vanem reageerib lapse ebakorrektsel lausungi vormile sagedamini kui sisule, ja vastupidi: vastused korrektsetele lausungitele seostuvad enamasti lause sisuga.

Võtmesõnad: lastekeel, keele omandamine, longitudinaalne korpus, vestlusanalüüs, psühholingvistika, leedu keel

SHIFT OF USE OF CAPITAL (UPPER CASE) LETTERS IN ENGLISH AND LATVIAN MARKETING TEXTS

Eva Birzniece

Abstract. This paper is largely observational. It discusses shifts in capitalization practices in marketing texts in Latvian. Two of the observed shifts are changes in accepted Latvian orthography (in Latvian texts) and one is an extension of Latvian changes into English texts written in Latvia. One of the Latvian shifts may be due to external influence and one is from internal sources. These changes are considered as matters of orthographic change, cultural change, the introduction of new (or previously little utilized) communication channels and tools and demonstration of the evolution of new communication rules. The paper provides some hypotheses about what may cause this and makes some observations about what this reflects about culture and language and how these are applied to other languages.

Keywords: capitalization, orthography, language change, Latvian, English

Introduction

This paper discusses some unusual uses of capital letters in several situations in various marketing texts. Marketing texts (that is a wider and different notion from advertising texts) were chosen as they exhibit shifts in capitalization and this suggests a pattern and not just arbitrary changes of the existent norms. The first is a change from formally accepted Latvian orthography in the use of personal pronouns. The second is a reflection of the first in the use of capital letters in pronouns in English texts written in Latvia. The third usage change is taking English capitalization norms and using them in Latvian when they violate norms of Latvian orthography. These changes are observed in Latvian and English marketing texts written in Latvia.

Slipping out of the norm and then in some cases eventually changing it, are caused by new functions of communication and changing environment where several languages interplay. New functions of using capital letters are born of the need to assign new meaning and weight to specific words and concepts in marketing and advertising texts (that are rather new themselves in the Latvian language context). The changes of the language environment are mostly due to a closer interaction of the Latvian and English languages since the beginning of the 1990s. English is actively used in business communication. Sometimes the innovative use of the capital letter is overuse of an already existent norm in Latvian and sometimes it is adopted from an English convention. This process affects both English and Latvian usage as sometimes the new capitalization standards or the norms of Latvian are shifted to the English texts produced by authors whose first language is not English. This shift is based on the assumption that what works in one language must work in another as well disregarding the fact that any language is culturally laden and also norms of orthography carry meaning.

In this paper the terms *capital* and *upper case letters* are used interchangeably and are assumed to have the same meaning. This paper deals only with words where the first letter is capitalized as opposed to words with all upper case use.

1. Overuse of capitalization in personal pronouns in Latvian

A well-manifested change in the capitalization paradigm is overusing of capital letter for the second person plural pronoun *you*. The capitalized *you* is now used in practically all marketing texts – direct and public (Kotler 2006: 537), though its use in public marketing is not consistent with the norms of Latvian orthography.

In the Latvian language there are 2 forms of the 2nd person singular personal pronoun – the real singular *tu* and plural *jūs*. The plural form is used both for addressing several people and when addressing a single person with a need to show respect, formality and/or seniority. This norm is present in many European languages. The norm as suggested by standard textbooks of Latvian orthography is that second person plural pronoun is capitalized in personal writing (letters, invitations, etc.); thus, the use of capitalized *tu* and *jūs* in all case forms is expected only in a limited range of writing – personal letters, invitations, and applications. This is done to show respect and demonstrate formality to the person addressed and “to honour the addressee” (Romane 2000: 59). The authors of a manual of use of Latvian for written communication “Turi pa rokai, iemet aci” published in Canada in 1997, stress differentiation between capitalized and non-capitalized 2nd person pronoun (both singular and plural). They instruct that capitalized pronoun be applied only for a person in singular (Muižniece *et al.* 1997: 103). The same norm is recently reinforced by the authors of the manual of orthography where the authors specially indicate that personal pronouns *tu* and *jūs* and possessive pronouns *tavs*, *jūsu* (when they refer to a person in singular) in letters are written with a capital initial letter, for example, *sūtu Tev sveicienus; gaidu Jūsu atbildi; pateicos par Tavu sūtījumu* (Guļevska *et al.* 2002: 122–123). In practice, exactly when 2nd person pronoun is capitalized is a slightly elusive matter – each author chooses to use upper case letter when he/she assigns enough worth to the person addressed.

The capitalized 2nd person pronoun is a functional norm in many kinds of public (as opposed to direct) marketing and advertising writing. In contemporary marketing texts (bank brochures, leaflets, ads in printed media, web sites of companies and banks) this pronoun practically always appears in upper case when addressing a potential client, who is important but still anonymous (not personally known to the writer of the text). There seems to be no difference between texts that are addressed to many (and no one at the same time) and direct mailing which whenever possible has a concrete addressee.

(1a)

Uzņēmēja komplekts ir:

ekonomisks – *Jūs* ietaupīsiet līdzekļus, vienkopus iegādājoties pakalpojumus, kuru atsevišķa iegāde izmaksātu dārgāk;

efektīvs – lielāko daļu pakalpojumu *Jūs* varēsiet izmantot jau dažu stundu laikā pēc pieteikšanās... (Source: bank materials – a printed leaflet; *emphasis added*)

Translation into English:

Business services complex is:

economical – *You* will save money by getting a set of services that if purchased separately would cost more;

effective – *You* will be able to use most of these services within a few hours after applying... (Source: bank materials – a printed leaflet; *emphasis added*)

This text (1a) is from a bank brochure informing potential clients about the services of a service bundle for enterprise bank accounts.

The interplay between 2nd person singular and plural form is interesting as well. To gain more personalization, some marketing texts use singular *tu* instead of *jūs*. This is done to stress interest, personalization, friendship and closeness (e.g. – we like you, we are pals, look, what we've got for you here!).

(1b)

Drīzumā Valsts kases darbība pensiju 2. līmeņa pārvaldē tiks izbeigta. Ir pēdējais brīdis *Tev* pašam izvēlēties, kur uzkrāt savu pensiju. (Source: bank materials – a printed leaflet; *emphasis added*)

Translation into English:

State Treasury management of the second pension level is coming to an end. It is high time for *You* (2nd person singular) yourself to choose where to save your pension. (Source: bank materials – a printed leaflet; *emphasis added*).

This text (1b) is about pension plan availability and advantages provided by the new asset management company.

There is an interesting discourse in the society concerning these forms: many people perceive the plural form as a norm when talking to anyone unfamiliar, not to speak of people who are higher in the hierarchy or senior. It is curious that these forms – either singular or plural – carry very strong connotation. Even though usually the singular form is an expression of trust, friendship, liking, its use in

some contexts is viewed as rude, for example, two people who are not known to each other (one usually younger than the other) arguing in public transportation, where use of plural is traditionally socially expected. The same shift is possible also within the use of the plural form, which by definition implies respect. Still, by using *jūs* people very often show reservation, resentment, differentiation, sometimes also signalling power and status difference. Geert Hofstede defines power distance “as the degree of inequality among people which the population of a country considers as normal: from relatively equal (that is, small power distance) to extremely unequal (large power distance)” (Hofstede 1993: 89). There has been no research on power distance in Latvia, but work in Lithuania suggests that Lithuania has medium power distance (Baltrimiene 2005). Given the similarities between Lithuania and Latvia, probably Latvia has like power distance. It seems reasonable to assume that the strict differentiation in the language between singular and plural personal addresses indicates higher rather than lower power distance.

Though many people expect to be addressed in the plural form by a stranger (be it a person or a business), some marketers are courageous enough to use the singular form both in their oral and written materials.

When analysing the subject, a survey was distributed to learn respondents' preference regarding the use of *tu* vs. *jūs* and in case *jūs* was chosen, whether it should be capitalized. The age of the respondents is between 14 and 45; they are from secondary school students to people with completed higher education. 120 surveys were distributed via email and 60 were returned for a response rate of 50%. A significant majority (40 of 60) answered that they still expect to be addressed in the plural form. There was no age difference in the responses – both adults and teenagers answered that plural form is the appropriate one to be used in the marketing text. In comments to the survey, many respondents stressed that this is the way they address people and expect to be addressed by others whom they see for the first time or who are senior or perceived as respectable. 14 respondents (mostly men) said that it does not matter to them whether they are addressed by *tu* or *jūs*. Accordingly, only 6 respondents chose *tu* as a proper address form. In close parallel, most respondents (48 of 60) said that once *jūs* is used, it should also be capitalized. Several people first expressed their concerns of what is the correct/expected form. This implies that many people are rule and convention conscious and it is important for them to do the right thing. This is something marketers should pay attention to as well when deviating from the norm by adopting conventions from English: potential customers want to be treated the right way language wise as well.

Another trait of marketing texts in Latvian is shifting to more personalized, more dynamic expression. Thus the 3rd person usage has decreased in favour of 2nd person – instead of “the client, the customer” we see *you*. Though often this *you* is impersonal, it still gets treated as personal (by capitalizing it).

2. Bringing capitalization into English

The practice of using capitalized plural form of the personal pronoun is a culturally defined norm. These norms are very strong and very often almost unconscious; users perceive them as the only correct way there is. The use of one or the other form

is not only a rule of orthography but also culture expression. A project conducted by Radboud University (Nijmegen, the Netherlands) and Lessius Hogeschool in Antwerp (Flanders, Belgium) of intercultural business communication by email showed that even though the students from both countries share the same language, they demonstrate considerable cultural differences. Some of these are differences in power distance, masculinity, uncertainty avoidance, polychrony vs. monochrony, etc. Differences showed also in the use of 2nd person pronouns; a Dutch student in the first e-mail message addressed her Flemish counterpart using informal *you*, while the Flemish student uses the formal form. The authors of the project conclude that “this is not simply a question of superficial indications; it goes deep into the value-laden underground of one’s existence: something one sees only when confronted with it “on paper”” (Gerritsen, Verckens 2006: 55).

The variations in meanings and conventions are discussed in the article “Dear Friend“(?): Culture and Genre in American and Canadian Direct Marketing Letters” (Graves 1997) where the author compares direct marketing letters in USA and Canada. Both countries as in the case of Belgium and the Netherlands described above use the same language, but there are many historic, societal and cultural aspects that set the two countries apart and this difference is apparent in writing – both contents and form. On the contrary, we see mostly one-to-one alignment with the original Latvian text in many English marketing texts produced in Latvia.

In Latvian the use of capitalized *you* is brought over to marketing (and all kinds of informative) English texts created in Latvia by local authors. Even though capitalized *you* is used in some English advertisements produced in the UK, capitalized *you* in English is a violation of an English norm and appears faulty and confusing in an English text. Regardless of this and the fact that many of the texts are written or at least edited by people who have academic education in the English language (English philologists), this norm carried from the Latvian orthography prevails in the English texts. The overall quality of many texts suggests that this trend is not just ignorance and lack of knowledge, but an editorial policy of many companies and businesses.

Example (2a) is the text regarding the delivery of goods from an Internet shop in Latvia. Not only do the authors use capitalized *you*, but they also are not consistent about its use – in some places they use the upper case letter to begin the pronoun and in some other they do not. Comparing the English and the Latvian versions of the text, it is obvious that there are more *you* in the English translation: the Latvian text in the Example (2b) is basically 3rd person / impersonal (delivery of goods is possible, etc.) while in English a more personal voice is used by using *you*. There is the same possibility to use the 2nd person to express person neutral utterances by using 2nd person (Danielson, Porter 1990: 113–114). In this context capitalization is even more confusing – both in English and Latvian.

(2a)

1. Using the online shop ordering system.

This online shop system allows you to inspect the goods *You* have chosen, to enter delivery address, to chose the one of the delivery types. The delivery of the product is possible during 2-4 hours from the moment when *Your* order

is accepted in case if good is available in the warehouse. (Source: http://www.euroshop.lv/order_how/; *emphasis added*)

2. By free phone 800 06 00.

Tell the product *You* would like to buy, also tell desirable time and address of delivery. Our sales managers will consult *You* and register your order. The delivery of the product is possible during 2-4 hours from the moment when *Your* order is accepted in case if good is available in the warehouse. (Source: http://www.euroshop.lv/order_how/; *emphasis added*)

(2b)

1. Izmantojot Euroshop.lv iepirkšanās grozu.

Šī veikala sistēma *Jums* ļauj ielikt grozā vienu vai vairākas preces, apskatīties veiktos pirkumus, ievadīt līdz 5. piegādes adresēm, izvēlēties un norādīt piegāžu veidus. Preču piegāde iespējama 2-4 stundu laikā no pasūtījuma apstiprināšanas brīža, ja prece ir pieejama noliktavā. (Source: http://www.euroshop.lv/order_how/; *emphasis added*)

2. Izmantojot bezmaksas pasūtījumu tālruni 800 06 00.

Nosaucot savu vēlamu preci, piegādes laiku un vietu. Pārdošanas vadītāji *Jūs* konsultēs un reģistrēs *Jūsu* pasūtījumu. Preču piegāde iespējama 2-4 stundu laikā no pasūtījuma apstiprināšanas brīža, ja prece ir pieejama noliktavā. (Source: http://www.euroshop.lv/order_how/; *emphasis added*)

There is no consistency in the spelling of *you* with the capital letter as well (see the first instance of usage in 2a).

3. Specific vs. common nouns

Until relatively recently Russian was the major second language studied by Latvian speakers but now English is the most common second language (Bond *et al.* 2001). This suggests that English is the major intermediary language in Latvia, thus its impact on Latvian is inevitable. In English a common noun can be capitalized if it is specific as opposed to non-specific, e.g. god, God; project, Project; regulation(s), Regulation. The specified noun refers to the particular thing concerned vs. many others of the same category. This use of capitalized common nouns is a rather widespread practice in English. Example (3a) demonstrates this use.

(3a)

A description of The Visible Human Project® image data and how to obtain it (includes license agreement documents).

The Visible Human Project® FactSheet.

A sampler of images and animations from the *Project* (Source: http://www.nlm.nih.gov/research/visible/visible_human.html; *emphasis added*)

Some marketing and informative texts in Latvian have taken the English form of capitalising common nouns. This is done when referring to a specific project or event as opposed to any project or event. Common nouns in Latvian are capitalized

only in specific cases in titles and names of organisations and institutions. The rules are rather complex. Example (3b) demonstrates capitalization of a common noun *Pasākums* ‘event’ on a ticket to a concert, even though this does not add to the clarity of information – it is clear anyway that the ticket selling company assumes (limited) responsibility for and involvement only in the events it sells tickets for.

(3b)

SIA “Biļešu Paradīze” ir atbildīga tikai par biļešu pārdošanu. Nopirktās biļetes atpakaļ pieņem vai apmaina tikai *Pasākuma* atcelšanas un/vai pārcelšanas gadījumā, izņemot Patērētāju tiesību aizsardzības Likuma 34. panta 3. punktā noteiktos gadījumus. Par *Pasākuma* norisi pilnībā atbildīgs ir *Pasākuma* rīkotājs. Iekšējās kārtības noteikumus nosaka *Pasākuma* rīkotājs, saskaņā ar Civillikumu un sabiedrībā vispārpieņemtajām uzvedības normām. (Source: a ticket printed by “Biļešu Paradīze” ltd.; *emphasis added*)

Translation into English:

“Biļešu Paradīze” ltd. is only responsible for ticket sales. Tickets purchased are returned or exchanged only if the *Event* is cancelled or rescheduled, excluding reasons covered under paragraph 34, point 3 of the Consumer Protection Law. The organizer of the *Event* is completely responsible for *Event* procedures. The organizer of the *Event* is responsible to insure that the order and internal rules of the Event are in accordance with Civil Law and the accepted norms of society. (Source: a ticket printed by “Biļešu Paradīze” ltd.; *emphasis added*)

Spelling a common noun with a capital letter in the flow of the text is unaccustomed and confusing, as capitalization suggests a title or quality of a proper noun.

4. Capitalization in headings and titles

Another English use of upper case letters is in headings and titles. In English all words except short words (like *a, the, and, or, with*) are capitalized¹ while in Latvian only the first word of the heading and proper nouns are capitalized. Nowadays more and more samples of writing in Latvian have the English norm where every word gets capitalized.

(4)

Baltijas Nekustamā Īpašuma un Investīciju Konferenču
Reval Hotel Latvija, Rīga, 5.-6. Jūnijs (Source: newspaper publication of the real estate company Oberhaus)

Here every word of the title is written with upper case letters and in the additional information about the venue and place even the name of the month of June, which is never done in Latvian. This type of capitalization is more and more frequent in Latvian, as the authors evidently feel that English offers a uniform rule where nothing can go wrong and the possible reader is provided with the text of sufficient stress and importance.

¹ http://www.english.purdue.edu/handouts/print/grammar/g_caps.html (19.04.2006)

Conclusion

This paper presented a set of observations of orthography changes taking place in marketing texts produced in Latvian. One shows an internalization of English norms into Latvian and the other the absorption of Latvian norms in English language texts.

The use of capital (upper case) and small (lower case) letters forms the graphics of a text. Graphics of the text carries meaning just the same as the images, colours, layout and the actual words used in it. Voluntary changes of the expected graphics and/or orthography are unusual and confusing; they hinder perception of the text. The question of what these shifted and changed graphical representations really mean and communicate to the audience is still to be questioned and researched. Marketers should be aware that their texts may be perceived in a totally different way than they intend. The survey done with this study sustains the notion that Latvians expect what they believe is correct to be used and, by extension, they may not look kindly on deviations from the norm in marketing texts.

It is further suggested that these changes could easily be taken by readers (especially the young) as the norm and further replicated. It should be observed that because of its concrete existence on the printed page, written language has strength of permanence that spoken language may not have. So, spoken language allows more deviation.

We have to keep in mind that every norm is culturally laden and mechanical shifting of them from a language to language is at least counterproductive. As indicated in various sources, cultural differences and their representations can be disruptive (Gerritsen, Verckens 2006: 57). Therefore, people have to be aware and learn about intercultural differences and even more so people who advertise and market.

Language is a dynamic entity; it does change. Andrejs Veisbergs in his article "Where are you heading, Latvian?", suggests that there should be more flexibility in the way new tendencies and uses are elevated to the level of the norm (Veisbergs 2006: 2). He speaks mostly of vocabulary and spoken language; evidently changes in orthography take place over a longer period of time as their appearance in every individual case is an arbitrary decision. Even though some of the use decisions are arbitrary, we want to be aware of the process and the effects of the change.

References

- Baltrimiene, Renata 2005. Cultural Dimensions of Lithuania and its Relative Position in the Context With Other European Countries. http://www.klsmk.lt/gallery/_klsmk/mokslas/karjera/renata_baltrimiene.pdf (15.01.2007).
- Bond, Dzintra; Markus, Dace; Stockmal, Verna 2001. How do Latvians and Russians evaluate Latvian? Ohio University and the University of Latvia. <http://oak.cats.ohiou.edu/~bond/Latvians.htm> (15.01.2007).
- Danielson, Dorothy; Porter, Patricia 1990. Using English. Your Second Language. New Jersey: Prentice Hall Regents.
- Guļevska, D.; Miķelsons, A.; Porīte, T. 2002. Pareizrakstības un pareizrunas rokasgrāmata. Rīga: Avots.
- Graves, Roger 1997. "Dear Friend"(?): Culture and Genre in American and Canadian Direct Marketing Letters. – *The Journal of Business Communication* 34 (3), 235–252.

- Gerritsen, Marinel; Verckens, Jan Pieter 2006. Raising Students' Intercultural Awareness and Preparing Them for Intercultural Business (Communication) by e-mail. – Business Communication Quarterly 69 (1), 50–59.
- Hofstede, Geert 1993. Cultural constraints in management theories. – Academy of Management Executive 7 (1), 81–94.
- Kotler, Philip 2006. Marketing Management. New Jersey: Pearson Education.
- Muižniece, L.; Sināte, R.; Kronīte-Sīpola, S. 1997. Turi pa rokai, iemet aci. Kalamazū: LSC apgāds.
- Romane, A. 2000. Latviešu valodas rokasgrāmata. Rīga: Apgāds Zvaigzne ABC.
- Skujiņa, V. 1993. Valsts valodas prasmei lietvedības dokumentos. Rīga: SIA Biznesa komplekss. http://www.english.purdue.edu/handouts/print/grammar/g_caps.html (19.04.2006).
- Veisbergs, Andrejs 2006. Kurp ej, latviešu valoda? <http://www.leta.lv/lat/arhivs> (07.09.2006).

Eva Birzniece's (Business School at Riga Technical University) professional activity includes both teaching general English and academic writing. Her research interests are connected with Latvian and English literature and intercultural aspects of academic and bussines writing.
Eva.Birzniece@rbs.lv

SUURTÄHTEDE KASUTUSE MUUTUS INGLISE JA LÄTI TURUNDUSTEKSTIDES

Eva Birzniece

Riia Tehnikaülikool

Artikkel on suuresti ülevaatlik, arutledes kapitalisatsiooni tava muutuste üle Läti turundustekstides. Täheldatud nihetest on kaks seotud muutustega läti ortograafias (läti keelsetes tekstides) ja üks läti ortograafia laienemisega Lätis kirjutatud ingliskeelsetesse tekstidesse. Artiklis kirjeldatud läti muutustest on üks väliselt mõjustatud, teine on tingitud sisemistest muutustest. Muutusi on vaadeldud kui ortograafia, kultuurilise muutuse, uute (või varem vähetarvitatud) kommunikatsioonikanalite ja -vahendite käibelevõtu ning uute kommunikatsioonireeglite arengut näitlikustavat küsimust. Artiklis esitatakse mõned hüpoteesid, mis nähtust võiks olla põhjustanud, ning tehakse üldistusi selle kohta, mida see peegeldab kultuuri ja keele kohta ning kuidas on mõjutanud teisi keeli.

Võtmesõnad: kapitalisatsioon, ortograafia, keelemuutused, läti keel, inglise keel

APOLOGIES IN BUSINESS COMMUNICATION

Laura Čubajevaitė, Jūratė Ruzaitė

Abstract. The aim of the present paper is to examine the routine of apologizing in spoken business communication. Being face-threatening acts, apologies are of special importance in communication. Since apologies express sincerity and are remedial acts, apologizing helps to enhance mutual respect and to keep the relationship between colleagues stable, which is especially important in business communication. The present analysis focuses on four main expressions of explicit apologies, i.e. apologies with *sorry*, *I apologise*, *pardon* and *excuse me*. The data have been obtained from the sub-corpus of the British National Corpus, which includes transcripts of business communication and amounts to 1,321,844 words. The analysis accounts for the general frequency of different forms of apology in business communication as well as their frequency in relation to gender. The paper also focuses on the most typical patterns of apologies and argues that apologies range from neutral to highly emphatic or tentative. Finally, the present analysis investigates what is commonly apologised for in business settings.

Keywords: apologies, face-saving, politeness, gender, situational context, English

1. Introduction

The present paper discusses the routine of apologizing in spoken business communication. Apologies are of special importance in any communication since they have illocutionary force and thus are considered as speech acts; importantly, they are face-threatening acts. By admitting his/her fault, the speaker may lose face; therefore, to lessen the face-threat, special politeness strategies are required for an apology. When apologizing, most commonly speakers follow a highly predictable communicative routine to make the apology less personal and thus less face-threatening. On the other hand, apologizing is an act of expressing sincerity and a way of compensating for some damage caused to the hearer(s); thus apologizing

helps to enhance mutual respect and to keep the relationship between colleagues stable. In business communication, where stable partnership and mutual respect are especially significant, apologies play a special role.

The present analysis focuses on four main expressions of explicit apologies, i.e. apologies with *sorry*, *I apologise*, *pardon* and *excuse me*. Implicit apologies have not been taken into account since, as Robin Tolmach Lakoff (2001: 201) rightly observes, apologies in general (and especially the implicit ones) are “hard to identify, define, or categorize”. The analysis is primarily based on the corpus approach. The data have been obtained from the sub-corpus of the British National Corpus, which includes transcripts of business communication and amounts to 1,321,844 words. To make the data related to different genders comparable and to compensate for the uneven number of the words uttered by men and women in the corpus, relative frequencies (per 1 million words) will be calculated and will be presented in the analysis instead of absolute numbers.

The present analysis will focus on several aspects of explicit apologies. First of all, the general frequency of different forms of apologies will be examined. Their frequency will also be related to the social factor of gender to see whether male or female speakers apologise more extensively and whether they have any preference for different forms of apologies. Following the existing stereotypes, it can be hypothesized that women are more apologetic than men. Another aspect that will be studied is the usage patterns, which mainly include syntactic structures of different apologies. Finally, the paper will attempt to examine apologies on the pragmatic level by taking into account what is commonly apologised for in business settings. All these different aspects and levels of analysis will be related to the contextual factors, or situational features. The investigation will not take into account the phonological level since the nature of the available data does not allow us to do that. To limit the scope of the investigation, the turns following apologies are not analysed in greater detail either. This important aspect of the apologizing routine could be a focus of a separate study.

2. Theoretical preliminaries

The phenomenon of “the ritual organization of social encounters” was pointed to and discussed at length by Erving Goffman (1967: 45), who argues that human “interaction is organized on ritual principles” (ibid.). Conversational rituals or routines typically involve thanking, apologies, requests, offers and compliment responses. Apologies, just as some other conversational routines, have been studied rather extensively (see, for example, Owen 1983; Olshtain 1989; Vollmer, Olshtain 1989; Rintell, Mitchell 1989; Aijmer 1996; Lakoff 2001; Fahey 2005; Deutschmann 2006).

As Goffman (1967) observes, participants of a communicative act aim to maintain “ritual equilibrium” in communication, and apologizing is important since it helps to achieve this aim (Aijmer 1996). Similarly, George Leech observes that ‘apologizing ... can be regarded as an acknowledgement of an imbalance in the relation between s and h, and to some extent, as an attempt to restore the equilibrium’ (1983: 125). Since apologies have the function of restoring or reducing equilibrium, they are referred to as ‘remedial interchanges’ (Owen 1983).

Apologies are studied from different theoretical and methodological perspectives, such as semantics, speech act theory, sociolinguistic approach (see Aijmer 1996), the interdisciplinary approach of discourse analysis (Lakoff 2001) and corpus linguistics perspective (Deutschmann 2006). Being a culture-dependent phenomenon, conversational routines such as apologies are often studied in relation to cross-cultural variation (e.g. Wierzbicka 1991; Vollmer, Olshtain 1989; Blum-Kulka, House 1989; Wierzbicka 1991; Wolfston et al. 1989; Fahey 2005). A number of studies deal with the application of the pragmatics research and more specifically such speech acts as apologies to EFL (English as a Foreign Language) teaching. Such studies explore the importance of metapragmatic instruction on the speech act comprehension and production of non-native speakers (e.g. Rintell, Mitchell 1989; Bergman, Kasper 1993; Sbisà 1999; Eslami-Rasekh et al. 2004). Such studies clearly demonstrate practical applicability of pragmatic investigations of different speech acts.

Apologies are treated as ritualistic acts since speakers typically apologise in a relatively fixed way (Aijmer 1996: 80). Karin Aijmer (1996: 82) distinguishes two major sets of apologizing strategies, i.e. (1) explicit and implicit, (2) emotional and non-emotional strategies. Implicit apologies differ from the explicit ones as the former include no direct realization of an apology (e.g. *sorry*, *apologise*, *excuse*). If an apology is intensified (e.g. with the intensifier *very*), it is considered to be emotional.

Apologies are of special importance since they imply the speaker's guilt and thus are face-threatening (Olshtain 1989; Brown, Levinson 1994; Stenström 1994; Lakoff 2001). As Lakoff (2001: 201) points out, "apology, more than most speech acts, places psychological burden both on its maker and, less seriously, on its recipient". According to Elite Olshtain (1989: 156), when the speaker decides to apologise, s/he "is willing to humiliate himself or herself to some extent and to admit to fault and responsibility for X". Lakoff (2001) further notes that apologizing benefits the addressee, not the speakers. Therefore, apologizing is a face-saving act for the hearer (Olshtain 1989). Apologies are especially 'threatening' if the apologisee is a powerful person; the face loss is more serious then (Lakoff 2001). However, it is important to note that though apologies are always face-threatening, "not making a necessary apology may occasion more serious face loss in the long run" (Lakoff 2001: 211).

As some previous investigations suggest, the influence of context on the use of apologies is of high importance. However, apologies have not been studied extensively in situational contexts. Speech acts in general should be studied by applying context-sensitive approaches. Jacob I. Mey (2001: 219) suggests that linguists should study what he calls *situated* speech acts since, "[s]peech acts, in order to be effective have to be *situated*. That is to say, they both rely on, and actively create, the situation in which they are realised". For instance, Joanna Cutting (2000) in her analysis of casual conversations among students demonstrates that speech acts in general have a special role in different discourse communities. Similarly, Roberta Kevelson (1982), who focuses on legal speech acts, argues that speech acts should be of major concern to linguists in context-sensitive analyses. The importance of apologies is demonstrated in Brent Poole's (2001) study of apologies in non-synchronous computer mediated discourse among people from different cultural backgrounds.

Some previous investigations attempted to take into account the situational context by using the method of discourse-completion questionnaires (e.g. Olshtain 1989; Wolfson et al. 1989; Vollmer, Olshtain 1989). The drawback of such studies is that the apologizing behaviour of speakers is not observed in natural situations but the respondents have to provide probable apologizing forms for different situations described in the questionnaire. Nevertheless, the results of these studies strongly suggest that apologies are context-dependent. Besides, as Olshtain's (1989) research shows, the form of apologies depends on the speaker's status, that is, intensification of apologies rises with lower status.

Some previous research demonstrates that an analysis of apologies can yield very revealing results if this analysis considers situational contexts. For instance, Kevin Avruch and Zheng Wang (2005) study the importance of apology in the context of international negotiation between the U.S. and China. Their article examines the role of apologizing in relation to cultural and linguistic differences in the course of the negotiation. Aijmer's (1996) study of the distribution of apologies over different texts shows that different apologies are restricted to different texts. For instance, *sorry* is not found in public speeches, but it prevails in other text types (*sorry* is most frequent in conversations). The importance of register, genre and key is emphasized in Lakoff's (2001) interdisciplinary model as an important aspect that has to be taken into consideration in an analysis of apologies.

Social variation in the use of apology formulae is stressed by Lakoff (2001) and has been studied extensively by Mats Deutschmann (2006), who examined the use of the most frequent apologies in the spoken part of the British National Corpus. Social variation was observed in relation to the speaker's age and their class identity, which suggests that different apologies are important social markers. However, only minor differences between different genders were noticed.

Investigations of apologies in business settings, however, are not numerous. Apologies have been studied by David A. Hoffman (1998) in employment termination cases to show what constitutes an effective apology and what legal consequences it may have. Hoffman (1998) shows that appropriate apologies are effective in resolving disputes. How the forms of apologizing are attained by Thai business people has been examined by Ruja Pholsward (2003). José Camilo Davila (2004: 1) provides some theoretical considerations concerning the issues related to apologizing and forgiveness in the context of workplace relationships. In his study apologies and forgiveness are related to the degree of sincerity and severity of offence. Davila offers a theoretical model of how to examine three "hypothesized antecedents of forgiveness": offence severity, the content of the apology and the perceived sincerity of the offender's sincerity.

3. Discussion of the results

The present investigation has revealed that different forms of apologies are important in three main respects. First of all, apologies differ considerably in their frequency. There are also some minor differences in their frequency in men's and women's speech. Another important aspect is the usage of apologies, which are highly formulaic and follow largely predictable patterns. Finally, in relation to

apologies, it is important to take into account what is most commonly apologised for, since this is one of the important factors that can predetermine the pattern of an apology.

3.1. Frequency of different forms of apologies

The obtained data have revealed that different forms of apologies differ in their frequency. These differences are presented in Table 1, which provides not only general frequency of apologies but also their frequency in relation to gender.

Table 1. Frequency of different forms of apologies (per 1 million words)

	sorry	excuse me	pardon	apologise	Total
Women	510.05	34.9	14.6	27.67	587.22
Men	547.55	39.23	28.5	8.36	623.64
Total	1057.6	74.13	43.1	36.03	1210.86

Table 1 shows that apologies are rather frequent in business communication and total 1210.86 instances (see also Figure 1). The most noticeable differences can be observed between the total numbers of different forms of apologies. The data demonstrate that the most frequent apology is *sorry*, which occurs 1057.60 times per million. The other three apologies are considerably less frequent. *Excuse me* occurs 74.13 times; *pardon* occurs 43.1 times; *apologise* is even less frequent and occurs in 36.03 instances. Such a drastically higher frequency of *sorry*, as Figure 1 clearly demonstrates, can be explained by its usage peculiarities. It is the most neutral form of apologizing and thus it can be used in a much wider variety of situations. *Apologise*, being more formal than *sorry*, is of restricted usage since it may sound too emphatic. *Excuse me* is neutral in terms of its formality, but it is restricted mainly to the situations when the speaker wants to apologise for something embarrassing or rude that he/she will do. *Pardon* is mainly used when the speaker mishears and asks to repeat something.

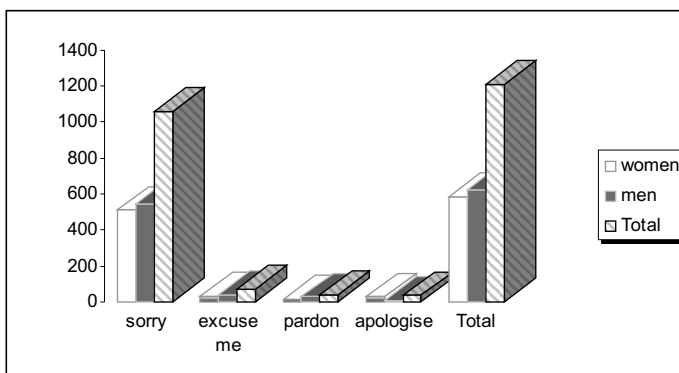


Figure 1. Distribution of different forms of apologies (frequency per 1 million words)

Different usage aspects of different forms of apologies will be further discussed in Section 3.2, which focuses on discourse functions of these different forms.

As far as gender differences are concerned, very little variation has been observed in the collected data. Men and women almost do not differ in the total number of apologies used in business communication (587.22 and 623.64 occurrences respectively). Though men use apologies more frequently, the difference is too slight to be important. Though we hypothesized that women might be more apologetic, this hypothesis has not been corroborated. Both men and women act as professionals in business communication, and their gender seems not to influence their verbal behaviour.

However, some differences between men and women do exist. These differences are related to the preference of certain forms of apologies. Table 1 shows that women clearly prefer the form *apologise* (27.67 occurrences), which in men's speech is three times less frequent (8.36 occurrences). Meanwhile, the frequency of the most neutral apology form *sorry* is higher in men's speech (547.55 occurrences) than in women's speech (510.05 occurrences). Male speakers use *pardon* almost twice as frequently as women (28.5 and 14.6 instances respectively). Men also demonstrate a slight preference for the form *excuse me*, but the difference in its distribution is just minor (34.9 occurrences in women's speech and 39.23 occurrences in men's speech). These results suggest that women prefer more formal apologies, whereas men are more inclined to use less formal forms.

3.2. Main usage patterns of different forms of apologies

The investigation has revealed several main usage patterns of apologies. First of all, in the majority of cases apologies are neutral and occur without any modifications. In addition, typically shortened forms prevail. For instance, *sorry* is preferred to the full form *I'm sorry*, *pardon* is preferred to *I beg your pardon* (for similar results in general English, see Aijmer 1996: 91). However, some apologies can be loaded and thus are either emphatic or tentative. There also occur double apologies, apologies with self-justification and apologies preceded by an interjection or a pause filler. All these patterns are discussed in greater detail further in the present section. To indicate the speaker's gender in each case, abbreviations M for male and F for female speakers will be used after each example.

3.2.1. Patterns of emphatic apologies

Emphatic apologies are based on the principle 'maximum sincerity and respect'. The degree of sincerity and respect is increased by using different linguistic means of intensification, examples of which are provided in (1)–(5).

- (1) Er, I must apologise again for what, er the way I addressed you, but of course with all these women equality er, movements going on, I never know whether its Mr, Mrs, or Ms. (M)
- (2) Alan, Alan I'm sorry to interrupt you, I do apologise. (M)

(3) A a again I am very very **sorry**. (M)

(4) I'm so **sorry**. (F)

(5) I'm awfully **sorry**... (M)

The examples above show that an apology is most commonly preceded by an intensifier (e.g. *very*, *so*, *awfully*). Another frequent intensifying device is the auxiliary *do* preceding the performative verb *apologise*, as in example (2). Finally, the performative verb *apologise* can be preceded by the modal verb *must*, which expresses strong obligation, to strengthen the illocutionary force of the apologizing utterance. In all these cases, the speaker maximally increases the threat to his/her own positive face, but by doing this, the speaker maximally reduces the threat to the hearer's face. Such techniques allow the speaker to preserve the stability of the relationship. Interestingly, in comparison to tentative apologies, which will be discussed below, emphatic apologies prevail in the collected data. This tendency suggests that speakers are more interested in maintaining a stable relationship and expressing positive attitudes to the interlocutors rather than saving their own face.

3.2.2. Compound apologies

Compound apologies, as referred to by Aijmer (1996), are closely related to emphatic apologies since the repetition of an apology strengthens its effect, as in examples (6)–(8).

(6) Right, first of all I'd like to **apologize** for the fact that Alan's report and my report especially the first half, are very similar. I'm **sorry** about that. (F)

(7) Oh **beg your pardon** I'm **sorry** I thought you said internal right. (M)

(8) And also **sorry**, **excuse me** like the other points there are you a smoker, no my wife does, she smokes twenty... (M)

Two apologies can follow one another in the same utterance, as in (7)–(8), or they can occur in two subsequent utterances, as in (6).

3.2.3. Patterns of tentative apologies

As has already been mentioned, tentative apologies are not as frequent as emphatic apologies. Tentative apologies contain mitigating devices that lessen the strength of the apology and make it less sincere, as can be seen in examples (9)–(11).

(9) I should perhaps **apologize on behalf of the hotel** for the temperature in the room this morning er I stayed here last night and woke up to find that not only was there no heat in the radiators, but there was no heat in the hot water. There wasn't any hot water. They had a major boiler breakdown last

night. Two boilers failed. Er so, so there's a distinct lack of central heating. Even more remarkable lack of hot water so (M)

(10) I'd like **apologize** for the room, it's er too cold really. (M)

(11) I I feel a bit **sorry** for for some of the counters because they can see certain on one palette for instance, I found there were three sorts of bags. (M)

In example (9), the apology is made tentative by means of modality. The utterance starts with the modal *should* and the adverb *perhaps*, both of which express lack of commitment (cf. *I must apologize* to see the contrast). In addition, the speaker does not make a personal apology, but apologises *on behalf of the hotel*, which allows the speaker to lessen personal responsibility and also face-threat. In example (10), the verb *apologise* is preceded by *'d like*, which weakens the force of the apology by making it less straightforward. In example (11), the mitigator *a bit* makes the apology less forceful. The examples above suggest that tentative apologies are less sincere than emphatic apologies and are based on the principle 'minimum face-threat' for the apologisee.

3.2.4. Apologies with a self-justification

In a great number of instances apologies are followed by a self-justification. Such self-justifications can be seen as special cases of face-saving and thus can be related to tentative apologies. Some examples of such extended apologies with self-justifying arguments are provided in (12)–(14).

(12) I'd like **apologise** for the room, it's er too cold really. I booked the conference not realising it was gonna be the control room by this time, and then this morning I turned up to find out we can't put the lights on without getting behind the bar, and we can't get behind the bar so that's why this er (M)

(13) I think what we have to understand, colleagues, is that this report actually is put to bed, print-wise, well in advance of this Congress. It doesn't in any way claim to be an absolutely up to the minute report of absolutely everything that we've done in the previously twelve months since last Congress, and quite frankly, it just could not be that. So, we do **apologize**, but hope you'll understand er, the delegate particularly, that we just cannot ensure that the report has got absolutely everything in it. (M)

(14) Erm I'm **sorry** about yesterday, but I just couldn't come. (F)

The examples above can be treated as cases of face-repair since the self-justifying explanations are the speakers' attempts to shield themselves. The speakers in examples (12)–(14) do apologise and admit their fault to demonstrate their sincerity, but simultaneously they disclaim their fault or try to lessen it by referring to some external circumstances that made their faults inevitable. As can be seen in examples (13) and (14), self-justifications are often introduced by the conjunction *but* (such constructions are called '*but*-prefaces' by Baker 1975, as cited in Lakoff 2001). Thus

apologies with a self-justification are well-balanced apologies that both express sincerity and respect to the speaker; they also minimize the face-threat.

3.2.5. Patterns with interjections and pause fillers

In a number of cases the interjection *oh* precedes the apology *sorry*, e.g. *Oh sorry*. See also example (15), which contains an interjection and a discourse marker in the utterance initial position.

(15) Oh, well, **sorry** about that er Marlene. (M)

Sometimes a pause filler precedes an apology:

(16) *Erm* I'm **sorry** about yesterday, but I just couldn't come. (F)

(17) *Er* let me **apologize** if the motion appears a little vague. (M)

The use of a pause filler in (16) and (17) allows the speaker to make a delay before uttering the apology.

Thus the investigation has shown that apologies are highly routinized and follow a largely predictable pattern. Apologies can have a different degree of force, depending on which several main types of apologies can be distinguished. First of all, apologies can be emphatic; in such a case, they primarily express the speaker's sincerity. Such apologies contain different intensifiers or double apologies. Tentative apologies, in contrast, are primarily face-oriented and contain different mitigating devices. Apologies with self-justifications are a balanced type of apologies since they both meet the requirement of sincerity and address the speaker's need for face-repair.

3.3. What is apologised for in business communication?

Another important aspect to take into account in relation to apologies is the faults that the speakers apologise for. On the basis of the data, two main reasons for apologies have been distinguished: (a) apologies for linguistic malfunctioning and (b) apologies for non-linguistic malfunctioning.

3.3.1. Apologies for linguistic malfunctioning (metalingual uses)

Most frequently apologies are made in cases of linguistic malfunctioning. Since in such instances the speaker comments on the discourse that is being produced, such uses of apologies will be called metalingual uses; see, for instance, example (18).

(18) However, the major problem, and you Ma'am have already touched upon this as well as the Chairman and I'm **sorry** to be repetitive but we do serve all yachtsmen, two and a half to three million of them whilst being financially supported by only sixty five thousand of them. (M)

In example (18), the speaker comments on what is being said and thus apologises for being repetitive. There are numerous other cases of linguistic malfunctioning, some of which are briefly discussed below.

- **Interruption**

To apologise for an interruption, different forms of apologies are used, as can be seen in (19)–(23).

(19) Robert Robert, **excuse me**, do you know what the er SSP allowance is?
(F)

(20) **Excuse me** a second I'll just do a quick (M)

(21) **Sorry** to interrupt you but there's nobody around at all and I've, yeah if I can just (F)

(22) Oh actually **sorry**, if I could just add a little bit to that. (M)

(23) Colleague, just one second, I do **apologise** for, for cutting in, but colleagues if you're, if it's your intention to leave Congress for a few moments please do it as quietly as possible. (M)

- **Disagreement (face-repair)**

In a number of cases an apology with *sorry* is used to mitigate disagreement, as in examples (24)–(26).

(24) I'm **sorry** if that's causing people a problem but that is basic human right. (F)

(25) I'm **sorry** Barbara, but time's a budget item with me these days, I can't wade through all this. (F)

(26) I'm **sorry** to say this, but the people that say money isn't everything, are the ones who've got it, aren't they? (M)

Usually, as the examples above suggest, in cases of disagreement an apology precedes a clause with *but*. Such uses of *sorry* can be treated as cases of face-repair. Since disagreement is an FTA, an apology in such instances functions as a mitigating device. By apologizing, the speaker admits the threat to the hearer's face and thus attempts to maintain the stability of the relationship.

- **Asking for repetition**

Apologies (usu. *sorry*) are also used in the speakers requests for repetition, as demonstrated in example (27).

(27) **Sorry** can you say again? (F)

- **Not being able to follow**

(28) **Sorry** I'm not, not with you. (F)

(29) **Sorry**, what did you say? (M)

- **Being too detailed**

(30) I'm sorry, I'm **sorry** I've been a bit steady getting there, but I think it's an important issue to tackle that... (M)

- **Self-correction**

Apologies in cases of self-correction are rather recurrent in the obtained data; see examples (31)–(34), where the corrected items have been underlined. As the examples suggest, the most frequent apology form in such cases is *sorry*.

(31) And divorce is or **sorry**, remarriage is not allowed for members of the Anglican Church. (F)

(32) We found that patients who were significantly younger er **sorry** we found that patients who were counseled for retrograde ejaculation were significantly younger (M)

(33) I had a hundred and twenty seven percent er **sorry** twenty seven percent above target er... (M)

(34) The Tories' reduction of public borrowing has been **excuse me** has seen many public services destroyed or disappeared forever. (M)

As examples (31)–(34) show, the routine of such apologies follows the pattern *corrected information + apology + correct information*.

It has to be noted that in requests for repetition, apologies for being unable to follow the speaker and apologies for self-correction, the apologizing pattern is usually very simple. A certain apology form without any elaboration is inserted in the utterance just to mark the speaker's awareness of some malfunctioning. With regard to the sincerity principle, such apologies are very neutral and routinized.

3.3.2. Apologies for non-linguistic malfunctioning

Apologies for non-linguistic malfunctioning are not as frequent as the ones for linguistic malfunctioning. Some examples of such apologies are provided below in (35)–(38).

(35) Now I'm **sorry** I haven't brought an acetate. (F)

(36) **Sorry**, I was dozing off over there. (F)

(37) I'm **sorry** I'm late. (M)

(38) I'm **sorry** I can't give you that. (M)

In the cases above, speakers apologise for their failure to perform something, as in (35), for some misbehaviour, as in (36) and (37), and refusal to do something, as in (38). As examples (35)–(38) show, all these cases of non-linguistic malfunctioning are just minor wrongdoings with a relatively low degree of the severity of the offence.

4. Summary and conclusion

The data have revealed that apologies are rather frequent in business communication; in the sub-corpus under investigation there occur 1210.86 cases of apologies. Apologies are slightly more frequent in men's than in women's speech, but this difference is too slight to be important (for similar results, see Deutschmann 2006). Interestingly, as the data show, there are some gender differences in the frequency of different forms of apology in men's and women's speech. The most neutral apology *sorry* is considerably more frequent in men's speech, whereas the most formal apologies with the performative verb *apologise* are more frequent in women's speech. This tendency suggests that women tend to use more formal forms of apology, at least in business settings. However, this tendency needs further testing on a bigger corpus.

Another observation in relation to frequency is that different types of apologies differ drastically in their frequency. The data show that most commonly apologies include *sorry*; apologies with the verb *apologise* are the least frequent ones. Similar results were obtained in Aijmer's (1996) corpus analysis of apologies, which were studied in general English without taking into account situational variation. Her data from the London-Lund Corpus show that apologies with *sorry* are the most frequent ones, whereas those with *apologise* are least frequent. The considerably higher frequency of *sorry* suggests, as Aijmer (1996: 84) observes, that *sorry* has become the most neutral and least marked routine. *Apologise*, according to her, is restricted mainly to formal situations (ibid.).

As for the linguistic patterns of apologies, they are highly routinized, or formulaic. Apologies can be of different degrees of strength, some being highly emphatic and others being rather tentative (cf. intensified and downgraded apologies in Olshtain 1989). Emphatic apologies (e.g. *I must apologise, I do apologise*) are based on the principle of 'maximum sincerity and respect', whereas tentative apologies (e.g. *oh well I apologise for that then*) are based on the principle of "minimum face-threat". However, in business communication apologies are most commonly very neutral and include the adverb *sorry* without any intensifiers or mitigators.

An interesting category of apologies includes those cases when an apology is followed by a *but*-explanation. Olshtain (1989: 158) refers to them as downgraded apologies, which raise "the question whether an apology was even necessary". Such apologies attempt to minimize the seriousness of the violation.

References

- Aijmer, Karin 1996. *Conversational Routines in English: Convention and Creativity*. London and New York: Longman.
- Avruch, Kevin; Wang, Zheng 2005. Culture, apology, and international negotiation: The case of the Sino-U.S. 'Spy Plane' crisis. – *International Negotiation* 10/2, 337–354.
- Baker, Charlotte L. 1975. This is just a first approximation, but... – T. J. San, J. L. Vance, R. E. Grossman (Eds.). *Papers from the Eleventh Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, 37–47.
- Bergman, Marc L.; Kasper, Gabriele 1993. Perception and performance in native and nonnative apology. – Gabriele Kasper, Shoshana Blum-Kulka (Eds.). *Interlanguage Pragmatics*. Oxford: Oxford University Press, 82–107.

- Blum-Kulka, Shoshana; House, Juliane 1989. Cross-cultural and situational variation in requesting behaviour. – Shoshana Blum-Kulka, Juliane House, Gabriel Kasper (Eds.). *Cross-Cultural Pragmatics. Requests and Apologies*. Norwood: Alex Publishing Corporation, 123–154.
- Brown, Penelope; Levinson, Stephen C. 1994. *Politeness. Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Cutting, Joanna 2000. *Analysing the Language of Discourse Communities*. Amsterdam: Elsevier.
- Davila, José Camilo 2004. Forgiveness as a function of offence severity, apology extensiveness, and perceived sincerity. – <http://www.agrh2004-esg.uqam.ca/pdf/Tome2/Davila.pdf> (20.07.2006).
- Deutschmann, Mats 2006. Social variation in the use of apology formulae in the British National Corpus. – Antoinette Renouf, Andrew Kehoe (Eds.). *The Changing Face of Corpus Linguistics. Language and Computers* 55. Amsterdam/New York: Rodopi, 205–221.
- Eslami-Rasekh, Zohreh; Eslami-Rasekh, Abbas; Fatahi, Azizollah 2004. The effect of explicit metapragmatic instruction on the speech act awareness of advanced EFL students. – *Teaching English as a Second Foreign Language. Electronic Journal* 8 (2). <http://www-writing.berkeley.edu/tesl-ej/ej30/a2.html> (15.07.2006).
- Goffman, Erving 1967. *Interaction Ritual: Essays in Face-to-Face Behaviour*. New York: Doubleday Anchor.
- Hahey, María Palma 2005. Speech acts as intercultural danger zones: A cross-cultural comparison of the speech act of apologizing in Irish and Chilean soap operas. – *Intercultural Communication* 8. <http://www.immi.se/intercultural/nr8/palma.htm> (19.07.2006).
- Hoffman, David A. 1998. The Use of Apology in Employment Cases. – <http://bostonlawcollaborative.com/documents/2005-07-apology-article.pdf> (19.07.2006).
- Kevelson, Roberta 1982. Language and legal speech acts: Decisions. – Robert Di Pietro (Ed.). *Linguistics and the Professions. Proceedings of the Second Annual Delaware Symposium on Language Studies*. Norwood: ALEX Publishing Corporation, 121–131.
- Lakoff, Robin Tolmach 2001. Nine ways of looking at apologies: The necessity for interdisciplinary theory and method in discourse analysis. – Deborah Schiffrin, Deborah Tannen, Heidi E. Hamilton (Eds.). *The Handbook of Discourse Analysis*. Oxford: Blackwell, 199–214.
- Leech, George 1983. *Principles of Pragmatics*. London, New York: Longman.
- Mey, Jacob L. 2001. *Pragmatics. An Introduction*. 2nd ed. Oxford: Blackwell Publishers.
- Olshtain, Elite 1989. Apologies across languages. – Shoshana Blum-Kulka, Juliane House, Gabriel Kasper (Eds.). *Cross-Cultural Pragmatics. Requests and Apologies*. Norwood: Alex Publishing Corporation, 155–173.
- Owen, Marion 1983. *Apologies and Remedial Interchanges*. Berlin, New York, Amsterdam: Mouton Publishers.
- Pholsward, Ruja 2003. Evaluation of language functions used by Thai business people. – http://www.utcc.ac.th/article_research/PDF/utccforum/EvaluationofLanguage.pdf (20.07.2006).
- Poole, Brent 2001. Apologies in computer-mediated communication. – *Pragmatic Matters. JALT Pragmatics SIG Newsletter* 2/3, 5–7.
- Rintell, Ellen M.; Mitchell, Candace J. 1989. Studying requests and apologies: An inquiry into method. – Shoshana Blum-Kulka, Juliane House, Gabriel Kasper (Eds.). *Cross-Cultural Pragmatics. Requests and Apologies*. Norwood: Alex Publishing Corporation, 248–272.
- Sbisà, Marina 1999. The room for negotiation in apologizing: Evidence from the Italian speech act of scusari. Paper read at the International Conference Pragm99 “Pragmat-

ics and Negotiation”, Tel Aviv and Jerusalem, 13-16 June. <http://www.univ.trieste.it/~dipfilo/sbisa/scuspap.html> (15.07.2006).

Stenström, Ana-Brita 1994. *An Introduction to Spoken Interaction*. London, New York: Longman.

Vollmer, Helmut J.; Olshtain, Elite 1989. The language of apologies in German. – Shoshana Blum-Kulka, Juliane House, Gabriel Kasper (Eds.). *Cross-Cultural Pragmatics. Requests and Apologies*. Norwood: Alex Publishing Corporation, 197–220.

Wierzbicka, Anna 1991. *Cross-Cultural Pragmatics. The Semantics of Human Interaction*. Berlin, New York: Mouton de Gruyter.

Wolfston, Nessa; Marmor, Thomas; Jones, Steve 1989. Problems in the comparison of speech acts across cultures. – Shoshana Blum-Kulka, Juliane House, Gabriel Kasper (Eds.). *Cross-Cultural Pragmatics. Requests and Apologies*. Norwood: Alex Publishing Corporation, 174–196.

Laura Čubajevaitė teaches in the Regional Studies Department, Vytautas Magnus University, Lithuania, Her research interests include foreign language teaching methodology and quality, and intercultural communication.
l.cubajevaitė@trs.vdu.lt.

Jūratė Ruzaitė teaches in the Department of English Philology, Vytautas Magnus University, Lithuania, Her research interests include discourse analysis (esp. spoken academic discourse), discourse and ideology, corpus linguistics, and vague language.
j.ruzaitė@hmf.vdu.lt.

VABANDUSED TÖÖSUHTLUSES

Laura Čubajevaitė, Jūratė Ruzaitė

Vytautas Magnuse nimeline Kaunase Ülikool

Artiklis analüüsitakse ingliskeelse suulise töösuhtluse vabandamistavasid. Vabandused kui mina-pilti ohustavad aktid on suhtluses eriti olulised. Kuivõrd vabandused väljendavad siirust ja on parandusaktid, aitab vabandamine parandada vastastikust lugupidamist ja säilitada stabiilseid suhteid kolleegide vahel, mis on töösuhtluses eriti oluline. Artikkel keskendub neljale eksplitsiitse vabandamise väljendile: *sorry*, *I apologise*, *pardon* ja *excuse me*. Andmed pärinevad korpuse British National Corpus allkorpusest, mis sisaldab 1 321 844 sõne mahus tööalaste vestluste litereeringuid.

Meeste kõnes esineb vabandusi veidi sagedamini kui naistel. Erinevusi leidub eri vabandusväljendite kasutamises. Kõige neutraalsem *sorry* on märgatavalt sagedasem meeste kõnes, formaalseim vabandus performatiivverbiga *apologise* aga naiste kõnes. Nähtavasti kalduvad naised formaalsema vabandamise poole, vähemalt töökeskkonnas. Suuresti erinevad eri vabandusväljendite üldised sagedused. Kõige sagedasem on *sorry*, kõige vähem esineb *apologise*.

Vabandusväljendite mallid on üpris kivistunud. Vabandused võivad olla eri tugevusega, mõned neist on väga emfaatilised, teised üsna ebakindlad. Emfaatilised vabandused põhinevad maksimaalsel siirusel ja lugupidamisel, ebakindlad aga mina-pildi minimaalsel ohustamisel. Omapärane kategooria on vabandused koos järgneva selgitava *but*-algulise lausega.

Võtmesõnad: vabandamine, mina-pildi säästmine, viisakus, sugu, situatsiooni-kontekst, inglise keel

SUHTLUSSTRATEEGIAD JA -TAKTIKAD MÜÜGIVESTLUSTES

Liina Eskor

Ülevaade. Artiklis antakse ülevaade Tartu Ülikooli Eesti dialoogikorpuse ühe osa, müügivestluste struktuurist ja müüja poolt kasutatavatest strateegiatest oma eesmärgi täitmisel. Tutvustatakse ka arutluspõhist suhtlustaktikate käsitlust ning vaadeldakse, millises ulatuses on võimalik seda müügivestluste peal rakendada.*

Võtmesõnad: korpuslingvistika, dialoogikorpus, korpuse märgendamine, müügivestlus, suhtlusstrateegia, arutluspõhine suhtlustaktika, eesti keel

1. Sissejuhatus

Käesolevas artiklis on vaatluse all müügivestlused ja neis esinevad suhtlusstrateegiad. Müügivestlused sisaldavad koolitusfirma töötajate kõnesid loodetavatele klientidele (s.t omavahel suhtlevad ametiisikud). Müügiagendi eesmärgiks on info kogumine kliendi kohta, kursuste vastu huvi tekitamine, ostusoovi korral argumenteerimine või kui see pole vajalik, siis ostuotsusele kaasa aitamine (pikemalt müügivestluste kohta vt Kasterpalu 2006). Artiklil on kaks eesmärki. Esiteks püütakse välja tuua analüüsitud müügivestluste nn laiemad strateegiad. Nende abil on võimalik leida müügivestluste sisemine struktuur, s.t vaadelda, millistele suhtluslikele pidepunktidele toetub müüja kliendis huvi tekitamiseks (peatükk 3). Teiseks ülesandeks on vaadelda, mil määral ja kuidas on võimalik rakendada müügivestlustes arutluspõhiseid suhtlustaktikaid (peatükk 4). Nimetatud ülesanded on eeskujuga ja innustust saanud Kristiina Jokineni suhtlusstrateegiate ideest konstruktiivse dialoogi juhtimise mudelis (peatükk 2.1) ning Tartu Ülikoolis arendatud arutluspõhisest suhtlusstrateegiate ja- taktikate käsitlusest (peatükk 2.2). Seni on eestikeelseid infodialooge (eraisikute kõned infotelefonile) analüüsitud K. Jokineni konstruktiivse dialoogi mudeli (Jokinen 1996a, 1996b) seisukohalt (Eskor 2004, 2005) ning vaadeldud ka suhtlusstrateegiate järjendeid ehk strateegiaid laiemas mõistes (Eskor 2006).

* Tööd on osaliselt toetanud Eesti Teadusfond (grant nr 5685).

Käeoleva uurimuse tarbeks analüüsiti algselt 16 müügivestlust (vestlused, mis olid selleks ajaks dialoogiaktidega¹ märgendatud ja ühestatud), kuid neist üheksas müügiteemalist vestlust ei arendatud (sest soovitud inimesega ei olnud võimalik rääkida). Seega pakkus sobivat analüüsimaterjali seitse müügivestlust. Kasutatud alamkorpuse suuruseks on 7792 sõna ehk 771 voozu.

Nii infodialoogid kui ka müügivestlused on osa Tartu Ülikooli Eesti dialoogikorpusest (EDiK)² ning näidetes on kasutatud vestlusanalüüsi transkriptsiooni. Korpus koosneb eri tüüpi institutsionaalsetest vestlustest ning on litereeritud ja dialoogiaktidega märgendatud. Korpuse loomise kaugemaks eesmärgiks on välja töötada kasutajaga loomulikus eesti keeles suhtlev dialoogsüsteem. Sellises süsteemis peaks arvuti oskama modelleerida inimese ajus toimuvaid protsesse: partneri jutu mõistmist, vastuse planeerimist ja ülesehitamist jms. Suhtlusstrateegiad võiksid aidata arvutil selles abiks olla.

2. Suhtlusstrateegia ja -taktika mõiste

2.1. Suhtlusstrateegiad konstruktiivse dialoogi juhtimise mudelis

Soome arvutilingvist Kristiina Jokinen on kasutanud suhtlusstrateegia mõistet konstruktiivse dialoogi juhtimise mudelis (Jokinen 1996a, 1996b). Mudel loodi kasutajale sobilike vastuste genereerimiseks infodialoogides ning selle lähtekohaks on kooperatiivse ja sidusa suhtluse aluseks olevad üldised suhtlusprintsiibid. Suhtlusprintsiibid on H. Paul Grice´i maksimumide täpsustused, s.t üldised põhimõtted, mille abil vestluses osalejad üksteise juttu mõistavad. P. Grice (1975) toob välja neli maksimumi: kvantiteedi-, kvaliteedi-, relevantsus- ja meetodimaksimum. K. Jokineni mudelis saab arvutiprogramm suhtlusprintsiiptide toel valida sobiva suhtlusstrateegia ja lahendada nii (kasutaja abiga) oma andmebaasi puudujääke. Sobiva strateegia valik tähendab valikut dialoogi jätkamise eri võimaluste vahel (nt küsida kasutajalt infot juurde, soovitada muid võimalusi, paluda küsimust korrata jm). K. Jokineni lähenemise puhul valitakse sobiv strateegia igale lausungile, arvestades sealjuures nelja faktorit: lausungi ootuspärasust (eelneva lausungiga võrreldes), lausungi teemat (kas on/ei ole vestluse senise teemaga seotud), initsiatiivi (kas vestluse initsiatiiv on kõnelejal või partneril) ning vestluse eesmärki (kas on käesoleva lausungi ajaks täidetud või mitte). Dialoog, kus info küsijal on laiemas mõistes üks eesmärk (nt bussi väljumisaja teadaasaamine), sisaldab endas tavaliselt ka mitut väiksemat eesmärki. Need tekitab vastaja oma täpsustavate küsimustega (küsimusele vastamiseks on tavaliselt tarvis üle küsida sihtjaama, nädalapäeva vms). Järgnevalt on toodud näide kõnest infotelefonile, kus on märgendatud lausungitele vastavad dialoogiaktid (Eesti dialoogiaktide tüpoloogia järgi, trükitähtedes, vt Hennoste, Rääbis 2004) ning suhtlusstrateegiad (Jokineni järgi, kaldkirjas, vt Jokinen 1996a). Helistajat tähistab *H* ja infotelefonil vastajat *V*.

(1)

H: kas ´telefoninumbriid mis algavad numbritega viis ´kaks, (.) kas need on ´Rakvere või: ´Pajusti omad. KYE: ALTERNATIIV, TVE: PAKKUMINE *finish/start*

¹ Dialoogiaktid on tegevused, mida inimene kõne abil teeb. Aktid jagunevad kahte suurde rühma: naaberpaariaktid (nt küsimus-vastus) ja üksikaktid (ei vaja partneri tagasisidet, nt jätkaja *mhmh*). Iga akt kannab nime, mis sisaldab akti rühmanime (nime alguses oleva akronüümi kaks esimest tähte, nt IL = infolisa), märgat selle kohta, kas ta on naaberpaariakt või mitte, kas esi- või järelliige (DIJ = direktiivi järelliige) ning akti konkreetne nimi, mis annab akti semantilise/funktsionaalse sisu (DIE: SOOV, KY: INFO PUUDUMINE). Vt ka Hennoste, Rääbis 2004.

² Vt ka <http://math.ut.ee/~koit/Dialoog/EDiC.html> (20.09.2006).

V: {-} (.) tähendab (.) vaatan? KYJ: EDASILÜKKAMINE, TVJ: VASTU-
VÕTMINE *notrelated*

(22.0) no 'Pajustis akkavad numbrid viis 'seitse=aga: (0.5) Rakvere::s (.) ma
nüüd ei 'näe=et viis 'kaks päris oleksid=aga KYJ: ALTERNATIIV: EITAV
follow-up-old

H: ahah VR: NEUTRAALNE INFO OSUTAMINE UUEKS *somethingelse*
aga võibolla see siin on mingi 'muu koha number lihtsalt. SEE: ARVAMUS
new dialogue

Helistaja küsimus infotelefoni töötajale on alternatiivküsimus, samuti on see esitatud pakkumisena vestluse teema vahetamiseks (esitatud näide on lõik vestluse keskelt). Sellistele tingimustele vastab strateegia *finish/start* ('eelmine lõpetatud, alusta uuega'), s.t selline küsimus on oodatud, vestluse teemaga seotud, küsimuse esitamise ajaks on vestluse varasemad eesmärgid täidetud ning initsiatiiv on kõnelejal (s.t helistajal). Vastaja on teema vahetamisega nõus, kuid info otsimiseks vajab ta pisut aega, seetõttu lükkab ta vastamise edasi. Sellise lausungi saab märgendada strateegiaga *notrelated* ('teemaga sidumata'): lausung ei ole partnerile ootuspärane (ta ootab vastust), see ei ole vestluse teemaga otseselt seotud, vestluse eesmärk on lausungi ajaks täitmata (vastust ei ole ju veel antud) ning initsiatiiv on vastajal. Sellele järgnebki vastus. Kuna helistaja küsimus oli esitatud alternatiivküsimusena, on nüüd tegemist alternatiivvastusega. Kumbki pakutud alternatiiv pole õige, aga kolmandat varianti siiski ei pakuta. Strateegiaks on siin *follow-up-old* ('jätka eelmisega'): vastus kui selline on oodatud, see on teemaga seotud, vestluse eesmärk on selle lausungi ajaks veel täitmata (kuigi just parajasti täitmisel) ning initsiatiiv on endiselt vastajal. Helistaja järgmine lausung (*ahah*) on vabatahtlik reaktsioon, saadud info on uus. Sobiv strateegia on siin *somethingelse* ('muu'), s.t antud lausung ei ole vestluse kontekstis otseselt ootuspärane (kuna see on lihtsalt helistaja reaktsioon, millele mingit reageeringut ei oodata), samas on ta eelneva lausungiga seotud (reageering eelnevale) ning ka vestluse eesmärk on selleks hetkeks juba täidetud. Initsiatiiv on partneril (s.t tegemist ei ole küsimusega). Järgneb helistaja seisukohavõtt, täpsemalt arvamus, mis on tähistatud strateegiaga *new dialogue* (uus dialoog). See ei ole otseselt oodatud, samas on teemaga seotud ning ka vestluse eesmärk on täidetud. Helistaja on nüüd kõneleja rollis (s.t ta ootab seisukohavõtule infotöötaja reaktsiooni).

2.2. Suhtlusstrateegiad ja –taktikad arutluspõhises dialoogimudelil

Tartu Ülikoolis on loodud dialoogimudel, mis kaasab arutluse mudeli (Koit 1996, Öim 1996, Koit, Öim 2003). Selle kohaselt reguleerivad inimese (praegusel juhul kliendi) arutlust selle üle, kas teha või jätta tegemata mingi tegevus D, kolme liiki faktorid: sisemised ehk subjektiivsed faktorid *soov ja vajadus* ning väline faktor *kohustus* (Koit 1996: 99–114, Koit, Öim 2003: 721–735). Subjekt soovib teha tegevust D, kui D meeldivad küljed ületavad ebameeldivad. Subjektil on vaja teha D, kui D tegemise kasu on suurem kui kahju. Subjekt on sunnitud/kohustatud tegema D, kui D tegemata jätmise toob kaasa karistuse. Püüdes jõuda selgusele, kas teha D või mitte, lähtub inimene kõigepealt soovist (s.t kas tegevuse meeldivad küljed

ületavad ebameeldivad). Kui see nii on, siis kontrollitakse vahendite olemasolu. Kui vahendid on olemas, kaalutakse tegevuse D muid positiivseid ja negatiivseid külgi (kasulikkust ja kahjulikkust). Kui D on keelatud tegevus, kaalutakse ka tegemisele järgnevat karistust. Kui head küljed ületavad summaarselt halvad, siis on arutluse tulemuseks otsus teha tegevus D, vastupidisel juhul otsus jätta D tegemata. Sellises mudelis on tarvis kokku liita nimetatud aspektide (meeldivus, ebameeldivus, kahjulikkus jt) kaalud, s.t kaaludel peaksid olema arvulised väärtused. Kuigi reaalsuses seda muidugi ei tehta, võib siiski kõik omadussõnad, mida nt meeldivate ja ebameeldivate tegevuste kirjeldamisel kasutatakse, teatud skaalale paigutada ning saada niimoodi vastava aspekti kaalu arvulise väärtuse (Koit, Õim 2003: 726).

Arutleva subjekti motivatsioonisfääri mudeliks on **kaalude vektor** $k = (k(\text{ressursid}), k(\text{meeldiv}), k(\text{ebameeldiv}), k(\text{kasulik}), k(\text{kahjulik}), k(\text{kohustuslik}), k(\text{keelatud}), k(\text{karistus}_p), k(\text{karistus}_{\text{mitte-D}}))$, kus D on tegevus; $k(\text{ressursid}) = 1$, kui subjekt omab kõik vajalikud ressursid tegevuse D tegemiseks (vastasel korral 0); $k(\text{kohustuslik}) = 1$, kui D on subjekti jaoks kohustuslik (vastasel korral 0); $k(\text{keelatud}) = 1$, kui D tegemine on subjekti jaoks keelatud (vastasel korral 0). Ülejäänud kaalude väärtusteks on mittenegatiivsed täisarvud.

Suhtlemise reeglid võtab arutluspõhises dialoogimudelis kokku suhtlusstrateegia. See on algoritm, mida osaleja kasutab oma suhtluseesmärgi saavutamiseks. Strateegia rakendamiseks võib sõltuvalt olukorrast kasutada erinevaid suhtlustaktikaid, nt partneri ahvatlemine, veenmine, ähvardamine. Lähtudes üle-eelmises loigus sõnastatud tegevusi reguleerivatest faktoritest, rõhutab vestluses osaleja (müügivestlustes müüja) ahvatlemise puhul partnerile tegevuse meeldivust, veenmise puhul kasulikkust ja ähvardamise puhul tegevuse tegematajätmisele järgnevat karistust. Suhtlustaktika valik sõltub väga paljudest asjaoludest, kõige üldisemalt saab need jagada kaheks: müüja subjektiivsed aspektid (meeleolu, isikuomadused) ja objektiivsed aspektid. Siin peetakse silmas suhtlust kui sotsiaalset nähtust ning konkreetsemalt selle erinevaid parameetreid, nt kas suhtlus on koostöine või vaenutsev, isiklik või isikupäratu (tavaliselt ametikõnelus), milline on suhtluskaugus (kas suhtlevad sõbrad või ülemus ja alluv), milline on suhtlemise toon (sõbralik või ebasõbralik) (Koit, Õim 2003: 728).

Müügidialooge analüüsidest ilmnevad konkreetset olukorrad, kus nimetatud taktikaid rakendatakse. Näitena võib tuua veenmise taktika: autor A, partner B, tegevus D: A tahab, et B teeks D (Koit 1996: 104).

1) Kui B vastus on “ei”, “ $k(\text{kasulik})$ väike”, “ $k(\text{kohustuslik})=0$ ”, “ $k(\text{meeldiv})$ väike”,

siis esitada vastuargument kaalu $k(\text{kasulik})$ suurendamiseks.

2) Kui B vastus on “ $k(\text{ressursid})=0$ ”,

siis näidata ressurside olemasolu või nende hankimise võimalust.

3) Kui B vastus on “ $k(\text{kahjulik})$ suur”,

siis esitada argument kaalu $k(\text{kahjulik})$ vähendamiseks.

4) Kui B vastus on “ $k(\text{kohustuslik})=1$ ja $k(\text{karistus})$ mitte-D väike”,

siis esitada argument kaalu $k(\text{karistus mitte-D})$ suurendamiseks.

5) Kui B vastus on “ $k(\text{keelatud})$ ja $k(\text{karistus D})$ suur”,

siis esitada argument kaalu $k(\text{karistus D})$ vähendamiseks.

6) Kui B vastus on “ $k(\text{ebameeldiv})$ suur”,

siis esitada argument kaalu $k(\text{ebameeldiv})$ vähendamiseks.

3. Müügistluste struktuur suhtlusstrateegiate kaudu

Käesolev peatükk annab ülevaate müügistluste struktuurist. Kuigi K. Jokineni suhtlusstrateegiad on loodud kasutamiseks infodialoogides ja müügistluste jaoks need eriti hästi ei sobi (kuna siin ei ole suhtluse eesmärgiks info andmine ja saamine, vaid koolituskursuste tutvustamine ja/või müük), püütakse siin K. Jokinenist innustust saades leida müügistlustest strateegiaid (kui dialoogiaktide järjendeid), mis kulgevad läbi mitme vooru ning annavad kokkuvõttes ülevaate vestluste struktuurist. Siin pakutud müügistluste struktuur strateegiate järjendina on selline:

- sissejuhatus ja täpsustamine;
- osavõtlikkus;
- veenmine;
- lähenemine teisest küljest;
- nõustumine kliendiga;
- kontakt tulevikuks.

Järgnevalt on dialoogide osad esitatud strateegiana (strateegia nimi kaldkirjas) ning ka dialoogiaktide järjendina. Sulgudes olevad aktid võivad konkreetsest vestlusest puududa.

3.1. Sissejuhatus ja täpsustamine

Sissejuhatusel algab müügistluste sisuline osa (koolituse pakkumine). Sellele eelnevaid vestluse algusrituaale, identifitseerimist ja viisakusküsimusi on kirjeldanud R. Kasterpalu (2006). Üldjuhul on sissejuhatuses müüja küsimus (nn taustauuring) firma suuruse, koolitusvajaduse, tulevikuplaanide jms kohta (2). Sissejuhatus võib olla esitatud ka väitena (3).

(2)

M: kuidas see juhatus (.) otsus nüüd `otsustus. | KYE: AVATUD |

K: ee `noh me `rääkisime sellest `pikalt hh | KYJ: INFO ANDMINE |

M: [jaa?] | VR: NEUTRAALNE JÄTKAJA |

K: [rääkisime] sellest pikalt jah=aga me ei jõudnudki `kokkuleppele et=kas
meil: `on vaja või `ei. | KYJ: INFO PUUDUMINE |

Aktijärjendina³:

M: KYE: AVATUD/ JUTUSTAV KAS/ SULETUD KAS

K: KYJ: JAH/ INFO ANDMINE/ INFO PUUDUMINE/ KEELDUMINE/ EI

(3)

M: .hh me oleme teiega sin `rääkinud `koolituse `teemadel. | SEE: VÄIDE |

K: [jaa?] | SEJ: NÕUSTUMINE |

M: [mitmeidki] kordi. | IL: TÄPSUSTAMINE |

Aktijärjendina:

M: SEE: VÄIDE

K: SEJ: NÕUSTUMINE

(M: IL: TÄPSUSTAMINE)

³ Aktijärjendid siin ja edaspidi kirjeldavad kõiki sarnaseid situatsioone, mitte ainult eelnenud näidet. Sulgudes esitatud aktid võivad puududa.

Sissejuhatusega samasuguse struktuuri ja sisuga vestlust leidub ka dialoogi keskel, siis on see tähistatud müüja strateegiaga *täpsustamine*:

(4)

M: kuidas teil muidu plaanis `on kas jääbki (.) `kolm inimest müügi peale või või tulevikus ka (.) `laiendada (.) peaks müük [`suurenema] | KYE: AVATUD |

K: [tändab] see `laieneb ainult võibolla siis `oma:: `jae` müügi: `võrgu:: [ee] | KYJ: INFO ANDMINE |

3.2. Osavõtlikkus

Osavõtlikkust on eristatud kolmel korral. Müüja näitab omaalgatuslikult, et ta mõistab klienti. Osavõtlikkust võib väljendada vastust pakkuva küsimusega (5) või ka arvamuse avaldamisega (6).

(5)

M: no siis ma kujutan ette=et praegu on tõesti `väga=väga kiire `aeg vist. | KYE: VASTUST PAKKUV |

K: mhmh. | KYJ: JAH |

(.) tuleb tõdeda. | KYJ: JAH |

M: \$ jah. \$ | VR: NEUTRAALNE VASTUVÕTUTEADE |

Aktijärjendina:

M: KYE: VASTUST PAKKUV

K: KYJ: JAH

(M: VR: NEUTRAALNE VASTUVÕTUTEADE)

(6)

M: .hh st ma kujutan ette kui nüüd ee (.) pood `lahti läheb või tähendab mitte `kaubamaja vabandust läheb lahti nüüd (.) detsembri `kuus ja=ja (.) mt ja see `jõulud ja see `kiire aeg `mööda läheb, (0.5) .hh ja hakkab see `kevad pihta ja hakkab selline nüüd juba väike ru `tiin tekkima=et mis on see `töö=ja=ja mis on vaja seal `töö juures `teha, [.hh] | SEE: ARVAMUS |

K: [tähendap,] (0.5) me `tahaksime `tegelikult kui me (.) mingit `koolitust `tahaksime siis `enne`kõike tahaksime me `müüjatele seda `teha. | SEJ: MUU |

Aktijärjendina:

M: SEE: ARVAMUS

K: SEJ: NÕUSTUMINE/ MUU

(M: VR: NEUTRAALNE VASTUVÕTUTEADE)

3.3. Veenmine

Veenmist võib vaadelda argumentide vahetamisena. Järgnevalt müüja poolt kasutatud argumente.

- **Koolituse kasulikkus kliendile.** Näiteks: koolitus aitab rutiini peletada, koolituse abil saab firma teavet ja oskusi, mille abil oma maine konkurentidest kõrgemale tõsta. Koolitus on lödvestav, aitab kogemusi vahetada.
- **Sobivate koolituste läbiviimise kord ja sisu.** Näiteks: klient on enda arvates sobiliku kursuse välja valinud, kuid tema usutlemisel avastab müüja, et kliendile sobiks hoopis teine kursus, ja annab sellest teada.
- **Koolituse sisu ja toimumisaja muutmine vastavalt kliendi vajadustele.** Näiteks pakub müüja kursust, mis kliendi arvates on mõeldud algajaile ja seega talle ei sobi. Müüja teatab, et teemasid võib muuta ja täiendada.
- **Koolitusel saadud teadmisi saab uutele töötajatele edasi anda,** nii et juba katseajal oleks uutest inimestest müügitöös rohkem abi.

Kuigi vestluste lõpuks soovib müüja teada konkreetset aega, millal uuesti kliendile helistada, selgitab ta, et konkreetse koolituse valik on väga oluline ning selleks on tarvis aega. Kuna müüja ei soovigi nende vestluste käigus ühtegi konkreetset kokkulepet sõlmida, võib järeldada, et ta täitis oma eesmärgi, s.t sai kursusi tutvustada ja ennast meelde tuletada. Müüja saavutas kõigis vestlustes kokkuleppe, et võtab kliendiga mõne aja pärast jälle ühendust. Ükski klient ei keeldunud sellest.

Veenmine dialoogiaktide järjendina tähendab müüja poolt tüüpiliselt aktide DIE: PAKKUMINE või DIE: ETTEPANEK kasutamist (7). Lisaks sellele võib müüja kasutada ka erinevaid seisukohavõtte, nt SEE: VÄIDE (8), SEE: ARVAMUS (9), samuti info andmist üksikaktina (10).

(7)

M: .h et=et kui nüd (.) ee `teie kontigenti `vaadata sis võibola (.) võibola seal kui `teemasid natukene (0.2) e seles mõtes täiendada `muuta et=et võtta midagi `sellist asja `sisse mida nüd `algajatele `päris EI (.) ole mõtet `pakkuda. | DIE: PAKKUMINE |

K: jah. | DIJ: NÕUSTUMINE |

M: .h midagi lisamüügi=[mm] `lisamüügi sarnast `kaupluses | DIE: PAKKUMINE |

K: [jah.] | DIJ: NÕUSTUMINE |

Aktijärjendina:

M: DIE: PAKKUMINE/ ETTEPANEK

(K: VR: NEUTRAALNE JÄTKAJA)

(M: IL: SELETAMINE)

K: DIJ: NÕUSTUMINE/ PIIRATUD NÕUSTUMINE

(M: IL: SELETAMINE/ TÄPSUSTAMINE)

(8)

M: ja ja `samas siis kui tuleb see `uus inimene kes hakkab ak `tiivse müügiga tegelema (.) näiteks .h sis `sellisel puhul tal on .h `praegusel inimesel `on `kogemus nendest `toodetest ja kogemus `firmast=et e siis ta saab anda neid

`teadmisi edasi ka .h ka `ueele inimesele et see juba `katseajal saaks .h saaks
`tõesti sellist ee `eesmärgipärast `müügitööd teha. | SEE: VÄIDE |
K: mhmh? | SEJ: PIIRATUD NÕUSTUMINE |

Aktijärjendina:

SEE: VÄIDE

SEJ: PIIRATUD NÕUSTUMINE/ NÕUSTUMINE/ MUU

(9)

M: .hh aga (0.8) mt nojah `seles mõts=et `kui ta läheb ka edasi ütleme sinna
`kauplusesse võetakse uus `müügi `mees siis .hh ma usun=et `kindlasti `need
`teadmised mis tal on sellisest `müügist ja kliendi `mõjutamisest kuluvad
`marjaks `ära ka sealsamas `kaupluses. | SEE: ARVAMUS |

K: ee ma ei vaidle `vastu. | SEJ: NÕUSTUMINE |

Aktijärjendina:

M: SEE: ARVAMUS

K: SEJ: NÕUSTUMINE

(10)

M: .h aga=aga=jah (0.5) seles mõttes `küll=et kui `jah vanad olijad mõtlevad
`tõesti et=et see on=ee `sellisele: ütleme `ueele(.) `ueele `inimesele `rohkem
mõeldud sis sis seda `irmu ei tohiks .h ja=uh `olla iseenesest et=et se=on
> selline küllaltki < ja kui teda on vel nädala `vahetusele panna ja kokku
viia mingisuguse firma `sisese .h sellise mõnusa `õhtu veetmisega=et=siss
ta on: noh pigem `lõdvestav ja=ja (.) sad `oma kogemusi: `välja rääkida
`töökogemusi=aga teistelt (.) teistelt `uusi kogemusi=ja .h teatud `situatsioone
`arutada ja leida [võbola] | YA: INFO ANDMINE |

3.4. Lähenemine teisest küljest

Kui klient ei ole koolitusest huvitatud (või ta näiteks ei oska öelda, kas tema firma võiks tulevikus koolitust tellida), püüab müüja läheneda teisest küljest ja vestlust edasi arendada. Selleks kasutab ta akte KYE: AVATUD (11, rida 2) või KYE: VASTUST PAKKUV (12, rida 2).

(11)

1. K: `teema on: `lahtine. | IL: KOKKUVÕTMINE |

(0.8)

2. M: jaa:| VR: NEUTRAALNE VASTUVÕTUTEADE |

.h ee kuidas o:lid `lood nende `referentsidega: kas [ee] | KYE: AVATUD |

3. K: [eõ]õ [`sain] `kätte selle: mm:: (1.0) `noh selle `muu=te=es=er selle
[`Akutrelli] | KYJ: INFO ANDMINE |

(12)

1. K: ma nüüd hheh hetkel jälle `öelda et mul on juba `plaanid `tehtud ma e:i
söendaks paraku .hh `järje `kordselt. | KYJ: INFO PUUDUMINE |

2. M: * mhmh. * | VR: NEUTRAALNE VASTUVÕTUTEADE |

.hh no õõ (.) võobla=t tõesti üks asi on sellised `kirjalikud `plaanid õõ

`graafikutena `paika `panna=aga .h aga=ilmselt sellised `mõtted `plaanide tasandil `peas on ol- on siiski [`olemas.] | KYE: VASTUST PAKKUV |
3. K: [on=on.] | KYJ: JAH |

Aktijärjendina:

M: KYE: AVATUD/ VASTUST PAKKUV

K: KYJ: INFO ANDMINE/ JAH

3.5. Nõustumine kliendiga

Kui klient avaldab millegi kohta arvamust (tihti aktidega SEE: VÄIDE või DIE: ETTEPANEK), siis müüja alati nõustub kliendiga, kasutades akte SEJ: NÕUSTUMINE, (13); DIJ: NÕUSTUMINE (14). Analüüsitud dialoogides ei vaielnud müüja kunagi kliendile vastu.

(13)

K: ja ainult ei ole jutt (.) müügist (.) sest `müük on `müük on üks osa. |
SEE: VÄIDE |

M: ja müügi- on loomu- loomu `likult üks [osa.] | SEJ: NÕUSTUMINE |

Aktijärjendina:

K: SEE: VÄIDE

M: SEJ: NÕUSTUMINE

(14)

K: [et] leida mingi::d siuksed pool `pühad või ma ei tea mis sin tulemas `on {üldse} | DIE: ETTEPANEK |

M: jah. | DIJ: NÕUSTUMINE |

no seda on võimalik iseenesest ka: ka noh `Tiritamme poolt ka alati `riigipühadel `teha et see ei ole [nagu] ee piirav | DIJ: NÕUSTUMINE |

Aktijärjendina:

K: DIE: ETTEPANEK

M: DIJ: NÕUSTUMINE

3.6. Kontakt tulevikuks

Kontakt tulevikuks tähistab vestluse lõpuosa, kus üks osalejatest pakub välja uue suhtluse aja. Strateegia võib algatada nii müüja kui klient. Müüja algatust esines kolmes vestluses, nt (15) ja (16), kliendi algatust neljas vestluses, nt (17). Strateegia on vormistatud küsimuse ja vastuse järjendina.

(15)

M: .h ee ja: ja `millal oleks `minul otstarbekas uuesti ühendust `võtta sel teemal | KYE: AVATUD |

et äkki san ka `ise `aidata kuidagi `kaasa teie otsuse kujunemisele. | DIE: PAKKUMINE |

(o.8) kas [midagi lähemalt] | YA: MUU |

K: [teeks äkki] ee `selle kuu: ütleme pealt kahe`kümmendal. | KYJ: INFO
ANDMINE | | DIE: ETTEPANEK |
M: mt (.) no `teeme `nii? | DIJ: NÕUSTUMINE |

Aktijärjendina:

M: KYE: AVATUD

K: KYJ: INFO PUUDUMINE + DIE: PAKKUMINE

M: DIJ: NÕUSTUMINE

(16)

M: mhmh. | VR: NEUTRAALNE VASTUVÕTUTEADE |

.hh nii=et võtan teiega ühendust kolmeteistkümnendal `jaanuaril? | DIE:
PAKKUMINE |

ja räägime juba sis kokku `saamisest, (.) millal [võiks] | DIE: ETTEPANEK |

K: [täitsa] `nõus. | DIJ: NÕUSTUMINE |

Aktijärjendina:

M: DIE: ETTEPANEK

K: DIJ: NÕUSTUMINE

(17)

K: `lepiksime `nimodi `kokku et=ee `kui ma nüüd olen ära `otsustanud
[`kas] ma | DIE: PAKKUMINE |

M: [* mhmh *] | VR: NEUTRAALNE JÄTKAJA |

K: tulen `ise või saadan `müügimehe [tähendab] ee ma `helistan teile `tagasi.
| DIE: PAKKUMINE |

M: [* mhmh *] | VR: NEUTRAALNE JÄTKAJA |

(0.2)

M: mt jaa. | DIJ: NÕUSTUMINE |

Aktijärjendina:

K: DIE: PAKKUMINE

M: VR: NEUTRAALNE JÄTKAJA

M: DIJ: NÕUSTUMINE

Käesolevas peatükis käsitletud nn laiema strateegiad kirjeldavad müügivestluste üldist struktuuri. Strateegiad võib jagada nende esinemiskoha järgi kaheks: strateegiad, mis võivad esineda ükskõik millises vestluse osas (täpsustamine, osavõtlikkus, veenmine, nõustumine kliendiga ja lähenemine teisest küljest), ning strateegiad, millel on kindel asukoht (sissejuhatus – vestluse alguses, kontakt tulevikuks – vestluse lõpus). Strateegiaid saab eristada ka nende esinemissageduse järgi: iga vestluse juurde kuuluvad sissejuhatus, veenmine, nõustumine kliendiga ja kontakt tulevikuks. Osavõtlikkust ja lähenemist teisest küljest igast müügivestlusest ei leia.

4. Suhtlustaktikad müügivestlustes

Järgnevalt vaadeldakse, milliseid suhtlustaktikaid (vt eespool 2.2) on võimalik leida müügivestlustest. Kuna müügivestlustes osalejad peavad jääma anonüümseks, siis ei saa käesolevas artiklis kahjuks avaldada neljast kõnevoorst pikemaid näiteid.

Seega ei ole võimalik müüja suhtlustaktikate kasutamist täiel määral demonstree-rida, kuid loodetavasti annavad ka lühemad dialoogilõigud asjast ülevaate.

Näidetes (18) ja (19) veenab müüja klienti koolitusel osalema. Klient esitab vastuväiteid ning müüja püüab teda veenda koolituse vajalikkuses. Klient väidab, et tema poe töötajad ei ole nõus koolitusele minema, kuna nad on juba vanad profid (rida 1), s.t esitab argumenti *k(kasulik) väike*. Selle peale püüab müüja suurendada *k(kasulik)* kaalu argumentidega *inimeste tundma õppimine, vaheldus ja huvitavad ülesanded* (rida 4).

(18)

1. K: ütsid=nimodi=et=ee (.) kui sa (.) mind (.) `müüjate `kursustele tahad `saata sis ma võin sulle öelda et ma: sinna ei `lähe eksju=noh. | YA: INFO ANDMINE |

2. M: a`haa. | VR: NEUTRAALNE INFO OSUTAMINE UUEKS |

3. K: hehe | YA: MUU |

4. M: \$ tegelikult \$ j(h)ah, se on küllaltki `huvitav `lähenemine nad nad ise saaksid `küllaltki `palju sealt `juurde. | SEE: VÄIDE |
tändap=se=on=noh .h et kui kogu `firmaga seal `olla=ja=ja: sab ää noh teises `situatsiooniski oma `oma inimesi rohkem `tundma ja=ja se on `vaheldus=ja seal on `huvitavaid `ülesanded mida nad lahendada peavad `koos=ja .h | SEE: VÄIDE |

Näites (19) kasutab müüja ahvatlemise taktikat (rida 1), suurendades *k(meeldiv)* kaalu argumentidega *mõnus õhtu, lõdvestus, kogemuste vahetamine*. Sellele klient vastuargumente ei esita.

(19)

1. M: .h aga=aga=jah (0.5) seles mõttes `küll=et kui `jah vanad olivad mõtlevad `tõesti et=et see on=ee `sellisele: ütleme `uuele (.) `uuele `inimesele `rohkem mõeldud sis sis seda `irmu ei tohiks .h ja=uh `olla iseenesest et=et se=on > selline küllaltki < ja kui teda on vel nädala `vahetusele panna ja kokku viia mingisuguse firma `sisese .h sellise mõnusa `õhtu veetmisega=et=siss ta on: noh pigem `lõdvestav ja=ja (.) sad `oma kogemusi: `välja rääkida `töökogemusi=aga teistelt (.) teistelt `uusi kogemusi=ja .h teatud `situatsioone `arutada ja leida [võbola] | YA: INFO ANDMINE |

2. K: [mhmh] | VR: NEUTRAALNE JÄTKAJA |

3. M: paremad | YA: INFO ANDMINE |

4. K: [{ah soo te}] | YA: PRAAK |

Näites (20) esitab müüja argumentid *hea suhtumine klienti, maine tõstmine, kursus aitab eesmärki saavutada* kaalu *k(kasulik)* suurendamiseks, s.t ta näitab siin samuti koolituse tellimise kasulikke aspekte (read 1 ja 3).

(20)

1. M: (.) ja ja kui nüüd `seal panna=nüd=ee `paika need ee (.) `inimeste=õõ `inimese `suhtumine teie `kaubamajasse=et et=see=on < `väga: (.) `hea `suhtumisega .hh `klienti > selline `virma=sis ma kujutan ette=et `see loob ka väga jah `tõepoolest `väga `head .hh `võimalused õõ (.) tõsta see `maine oopis `kõrgemale kui nüd teie `konkurendid Pärnus. | SEE: ARVAMUS |
.hh | YA: MUU |

2. K: just nimelt. | SEJ: NÕUSTUMINE |
 (0.5) see ongi `eesmärk. | SEJ: NÕUSTUMINE |
 3. M: jah? | VR: NEUTRAALNE VASTUVÕTUTEADE |
 et see=on `kindlasti `suur `abi sellest > `kursustest oleks `just nimelt selle
 (0.8) `eesmärgi saavutamisel, < kindla `sti üks võibolla isegi `peamisi |
 SEE: ARVAMUS |

Lõpetuseks veel näiteid, kuidas müüja kasutab erinevaid argumente rõhutamaks kliendile tegevusest saadavat kasu (iga näite juures on märgitud müüja konkreetne argument).

- (21)
 argument: *koolitusel saadavad teadmised on igas olukorras kasulikud*
 M: .hh aga (0.8) mt nojah `seles mõts=et `kui ta läheb ka edasi ütleme sinna
 `kauplusesse võetakse uus `müügi `mees siis .hh ma usun=et `kindlasti `need
 `teadmised mis tal on sellisest `müügist ja kliendi `mõjutamisest kuluvad
 `marjaks `ära ka sealsamas `kaupluses. | SEE: ARVAMUS |

- (22)
 argument: *võimalus saada toetusraha*
 M: sis ma kujutan `ette ku me `hakkame nüüd=nüüd `talvel kohe sellest
 `rääkima ja=ja asju `plaane tegema sis me võime seal saada POOL summast
 sellest mis `tegelikult maksma läheb. | YA: INFO ANDMINE |

- (23)
 argument: *koolituse sisu on võimalik kliendile sobivamaks muuta*
 M: .h et=et kui nüüd (.) ee `teie kontingenti `vaadata sis võibolla (.) võibolla
 seal kui `teemasid natukene (0.2) e seles mõtes täiendada `muuta et=et
 võtta midagi `sellist asja `sisse mida nüüd `algajatele `päris EI (.) ole mõtet
 `pakkuda. | DIE: PAKKUMINE |

- (24)
 argument: *koolitatud töötaja saab uusi töötajaid koolitada*
 M: ja ja `samas siis kui tuleb see `uus inimene kes hakkab ak `tiivse müügiga
 tegelema (.) näiteks .h sis `sellisel puhul tal on .h `praegusel inimesel `on
 `kogemus nendest `toodetest ja kogemus `firmast=et e siis ta saab anda neid
 `teadmisi edasi ka .h ka `uuele inimesele et see juba `katseajal saaks .h saaks
 `tõesti sellist ee `eesmärgipäras `müügitööd teha. | SEE: VÄIDE |

- (25)
 argument: *head kursused*
 M: (0.8) sest=et jah `juhtimine on ju `Tiritammes päris `põhjalikult
 `esindatud seal `koolituse osas ja seal on kah `erinevad (.) .hh `erinevad
 `võimalused. | YA: INFO ANDMINE |

Algne ettekujutus ja ootus, et müügistlustes leidub lisaks veenmisele ka vähemalt samapalju näiteid ahvatlemise taktika kasutamise kohta, ei täitunud. Ahvatlemist, s.t tegevuse meeldivuse rõhutamist, kasutas müüja väga vähe. Muidugi võib mõne näite juures tekkida küsimus, et kas tegemist on veenmise või ahvatlemisega, sest millegi kasulikkuse näitamine võib muuta selle ka meeldivaks. Nii et veenmine ja ahvatlemine ei ole täiesti eraldiseisvad ega teineteist välistavad.

5. Kokkuvõtteks

Artikli alguses seati ülesandeks kahe küsimuse lahendamine. Kõigepealt kavandati vaadelda müügivestluste struktuuri, et näha, millistele suhtluslikele pidepunktidele ehk strateegiatele toetub müüja kliendis huvi tekitamiseks. Müügivestluste analüüsi järel võib vestlustes esile tuua kuus strateegiat: sissejuhatus ja täpsustamine, osavõtlikkus, veenmine, lähenemine teisest küljest, nõustumine kliendiga, kontakt tulevikuks. Neist igaühes kasutab müüja mitmesuguseid võtteid kliendi tähelepanu püüdmiseks. Strateegiad jagunesid esinemiskoha järgi kaheks: strateegiad, mis võivad esineda ükskõik millises vestluse osas (täpsustamine, osavõtlikkus, veenmine, nõustumine kliendiga ja lähenemine teisest küljest), ning strateegiad, millel on kindel asukoht (sissejuhatus vestluse alguses, kontakt tulevikuks vestluse lõpus). Esinemissageduse järgi jagunesid strateegiad samuti kaheks: iga vestluse juurde kuuluvad sissejuhatus, veenmine, nõustumine kliendiga ja kontakt tulevikuks. Osavõtlikkust ja lähenemist teisest küljest igast müügivestlusest ei leia. Nimetatud strateegiate abil püüdis müüja koolituste vastu huvi tekitada, et klient ikka sooviks neist osa võtta. Müüja ülesandeks on märgata iga kliendi n-ö huvisähvatust ning sellest kinni haarata. Müüja veenab klienti koolituse vajalikkuses, esitades selleks erinevaid argumente. Ta ei vaidle kunagi kliendile vastu, on abivalmis ja vastutulelik, nagu müüja olema peab.

Teiseks analüüsiti, mil määral ja kuidas on müügivestlustes võimalik rakendada arutluspõhiseid suhtlustaktikaid. Müügivestlustele oli iseloomulik, et kliendid ei olnud eriti huvitatud koolituse tellimisest. Neil puudus selleks soov ja enda arvates ka vajadus. Kliendi otsust vestluse lõpuks mõjutas aga müüja. Oluliseks sai, kuidas müüja suutis näidata, et kliendil vajadus koolituse järele ikkagi on. Vajaduse väljaselgitamisel sai klienti juba edasi mõjutada ja loota, et kliendis tekib seepeale ka soov. Suhtlustaktikate seisukohalt oli analüüsitud vestlustes müüja põhiliseks mõjutusvahendiks kliendi veenmine, s.t tegevuse kasulikkuse rõhutamine. Veenmist saab vaadelda argumenteerimisena: müüja esitab kliendis huvi tekitamiseks erinevaid argumente. Müüja argumendid rõhuvad enamasti just tegevuse kasulikkusele. Samas näitas analüüs, et nii mõnegi argumendi puhul ei ole alati võimalik öelda, kas rõhutatakse tegevuse kasulikkust või kasutatakse hoopis ahvatlemise taktikat ja näidatakse tegevuse meeldivaid külgi. Kuna esitatud analüüsi aluseks on ainult seitse dialoogi, ei pruugi siin esitatu kehtida kõikide müügivestluste puhul. Praeguste väidete kinnitamiseks või ümberlükkamiseks peaks nimetatud vestlustaktikate esinemist kontrollima kindlasti kogu müügivestluste korpus.

Kirjandus

- Eskor, Liina 2004. Dialoogiaktid ja suhtlusstrateegiad: Eesti dialoogikorpuse analüüs. Magistritöö Tartu Ülikooli üldkeeleteaduse õppetoolis. Tartu.
- Eskor, Liina 2005. Dialoogiaktid ja suhtlusstrateegiad: Eesti dialoogikorpuse analüüs. – Keel ja Kirjandus 9, 711–727.
- Eskor, Liina 2006. Suhtlusstrateegiad infodialoogides. – Mare Koit, Renate Pajusalu, Haldur Õim (toim.). Keel ja arvuti. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 6. Tartu: Tartu Ülikooli Kirjastus, 183–195.
- Grice, Paul 1975. Logic and Conversation. – D. Davidson and G. Harman (Eds.). The Logic of Grammar. Encino, CA: Dickenson, 64–75.

- Hennoste, Tiit; Rääbis, Andriela 2004. Dialoogiaktid Eesti infodialoogides: tüpologia ja analüüs. Tartu Ülikooli arvutiteaduse instituut. Tartu: Tartu Ülikooli Kirjastus.
- Jokinen, Kristiina 1996a. Cooperative response planning in CDM. Reasoning about communicative strategies. – Proceedings of the 11th Twente Workshop on Language Technology: Dialogue Management in Natural Language Processing Systems. Twente, The Netherlands, 159–168. <http://www.mlab.uiah.fi/~kjokinen/papers/199606TWEN.PDF> (06.12.2006).
- Jokinen, Kristiina 1996b. Goal formulation based on communicative principles. – Proceedings of the International Conference on Computational Linguistics. COLING-96. Copenhagen, Denmark, 598–603. <http://www.mlab.uiah.fi/~kjokinen/papers/199608COLI.PDF> (06.12.2006).
- Kasterpalu, Riina 2006. Algusrituaalid müügikõnedes. – Mare Koit, Renate Pajusalu, Haldur Õim (toim.). Keel ja arvuti. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 6. Tartu: Tartu Ülikooli Kirjastus, 156–169.
- Koit, Mare 1996. Implementing a dialogue model on the computer. – Haldur Õim (Ed.). Estonian in the Changing World. University of Tartu, Departement of General Linguistics. Tartu: University of Tartu, 99–114.
- Koit, Mare; Õim, Haldur 2003. Eestikeelse dialoogi modelleerimine. – Keel ja Kirjandus 10, 721–735.
- Õim, Haldur 1996. Naive theories and communicative competence: Reasoning in communication. – Haldur Õim (Ed.). Estonian in the Changing World. University of Tartu, Departement of General Linguistics. Tartu: University of Tartu, 211–231.

Liina Eskor (Tartu Ülikool) on erialalt arvutilingvist, tegeleb suulise keele uurimisega.
 liina.eskor@ut.ee

COMMUNICATIVE STRATEGIES AND TACTICS IN TELEMARKETING CALLS

Liina Eskor

University of Tartu

In telemarketing calls communication takes place between business representatives on a win-win basis, which means that the caller is offering a competence (supposedly) needed by the receiver. The calls may have been addressed to top leaders of bigger Estonian companies. The original material for the present study consisted of 16 calls. However, only seven of them provided appropriate for the study as nine calls ended before ever reaching the marketing stage (for absence of the right person on the receiver end). The subcorpus analysed contained 7792 words used in 771 turns.

The calls are part of the Estonian dialogue corpus (EDiK) maintained at the University of Tartu. The transcription system used is typical for conversation analysis. The article points out some broader conversational strategies, which last through several turns. Following those strategies helps to find out the internal structure of the calls, i.e. to disclose how exactly the caller attempts to arouse interest at the other end of the line. As was proved by the dialogues analysed typical telemarketing calls consist of the following phases: introduction and specification, kind attention, persuasion, approach from another aspect, agreeing with the client, establishing contacts for the future.

Keywords: corpus linguistics, dialogue corpus, corpus annotation, telemarketing calls, communicative strategy, communicative tactics based on reasoning, Estonian

ÕPPIJAKEEL JA EESTI VAHEKEELE KORPUS

Pille Eslon, Helena Metslang

Ülevaade. Eesti keele õpetamine võõrkeelena on kutsunud esile vajaduse minna üle käsitsi tehtud väikeuuringutelt õppija vahekeele korpusepõhisele ja osaliselt automatiseeritud uurimisele, kuna suurte andmehulkade analüüs annab objektiivsema pildi eri tasemel õppijate arengust ja vajadustest teel sihtkeele omandamisele. Kirjaliku vahekeele korpuse alusel saab uurida ning arendada kirjutamis- ja lugemisoskust; korpusuuringute tulemusi võib kasutada uute ainekavade, õppevahendite, koolisõnastike ja käsiraamatute koostamisel. Käesolev artikkel tutvustab Tallinna Ülikooli üld- ja rakenduslingvistika õppetooli juures alustatud vene emakeelega õppijate eesti vahekeele korpuse loomise tööd eesmärgiga toetada teise keele omandamise (ingl second language acquisition) teaduslikku uurimist ja eesti keele kui teise keele õpetamise edasist arendamist.*

Võtmesõnad: korpuslingvistika, vahekeel, õppijakeele korpus, veaanalüüs, pedagoogiline grammatika

Sissejuhatavalt

Tänapäeva keeleuuringuid eristab varasematest võimalus teha järeldotsi suurte andmekogude alusel, milleks kasutatakse erinevat tüüpi elektroonilisel töödeldavaid korpuseid. Sama suund on iseloomulik ka võõrkeele omandamise teooria ja keeleõpetusmetoodika arengule. Sel eesmärgil on koostatud õppijakeele (ingl *learner corpus*) ehk vahekeele korpuse (*interlanguage corpus*), keeleõppeks sobivate tekstide korpuse (*corpora for learners*). Täna on korpusepõhine vahekeele- ja veaanalüüs seotud uute õppematerjalide loomise (õpikud, grammatikad, sõnastikud) ja arvutipõhise keeleõppega (*computer assisted language learning*), samuti tõlkija- ja õpetajakoolitusega. Et kiirendada ja arendada õppurite keelematerjali kogumise, kirjeldamise ja uurimise kõiki etappe, on viimastel aastakümnetel loodud ja kasutusele võetud mitmesuguseid universaalsemaid ja kitsama suunitlusega korpusedarkvara lahendusi. Nii minnakse näiteks õppijakeele korpuse käsitsimärgendamisele üle

* EVKK loomine on seotud järgmistest programmide ja sihtfinantseeritavate teemadega: sihtfinantseeritav projekt nr 0132493s03 "Eesti keelekeskkonna arengu analüüs, modelleerimine ja juhtimine" (2003–2007); riiklik programm "Eesti keel ja rahvuslik mälu" (2004–2008), grant R 05/01 "Koodivahetuse, vahe- ja lastekeele korpuse töötlemine ja haldamine"; ETF-i grant nr 6151 "Koodivahetuse, eesti vahekeele ning lastekeele andmekorpuse koostamine ja üldkirjeldus" (2005–2008).

Autorid tänavad Lars Borinit Göteborgi ülikoolist heade soovitude ja informatsiooni eest.

poolautomaatsele. Automatiseeritud meetodikal põhinevad uurimused võimaldavad tänu muutunud materjalimahtudele saada usaldusväärsemaid tulemusi ja leida uut laadi informatsiooni vahekeele haruldasemate nähtuste kohta.

Käesolev artikkel tutvustab erinevat tüüpi õppijakeele korpusi, mis maailmas on loodud, ning Eesti vahekeele korpust (EVKK) nende taustal. Pikemalt peatutakse EVKK veaklassifikatsioonil ning kirjutatakse sellest, missugused teaduslikud ja rakenduslikud perspektiivid avanevad eesti keele kui teise keele uurimisel seoses korpuse kasutamisega.

Õppijakeel ehk vahekeel

Vahekeel kujuneb seoste alusel, mida õppija loob emakeele (K1) ja õpitava võõrkeele (K2) vahel (Selinker 1969, 1972, 1992, Corder 1981 jt), mõnikord ka juba omandatud võõrkeel(t)e, emakeele ning õpitava keele vahel (Michiels 1999, Cenoz jt 2001, Hufeisen, Neuner 2003 jt). Vahekeelest on samuti räägitud kui sihtkeele variandist, mille kõrvalekaldumist normist mõjutavad vähemalt viis psühholingvistilist protsessi: 1) ülekanne K1 või K2, K3 jne põhjal; 2) ülekanne, mille õppija teeb õpetamise alusel; 3) õpistrateegiatest tulenev vigane keelekasutus; 4) suhtlusstrateegiatest lähtuv kõrvalekalle normist; 5) üldistamine ning reegli kasutamine selleks sobimatus kontekstis. Seega on vahekeel dünaamiline, tal puudub stabiilsus, tema areng on järk-järguline lähenemine sihtkeele normipärasele kasutamisele. Võib ka kujuneda olukord, kus vahekeel edasi ei arene. Ilmnevad fossiliseerumise ehk vea kinnistumise ning tagasimineku nähud: teatud suhtlussituatsioonides pöördub õppija vahekeele eelnevate etappide juurde ning hakkab tegema vigu, mida ta varem ei teinud (Tönshoff 1995: 5–6).

Õppijakeele korpustest maailmas

Õppijakeele korpused on maailmas uuemaid, kuid kiiresti levivaid korpuse tüüpe. Vanemaid neist on 1980. aastatel loodud Euroopa Teadusfondi teise keele andmebank (European Science Foundation Second Language Data Bank)¹. Laiemalt hakati seda tüüpi korpusi rajama 1990. aastate algul, mitte-inglise sihtkeele korpusi veelgi hiljem. Järgnevalt esitame olulisemad jooned, mis iseloomustavad erineva ülesehituse ja funktsiooniga õppijakeele korpusi.

Õppijakeele korpusi on loodud mitmesugustel eesmärkidel teoreetilisest uurimistööst kuni keeleõppe rakendusliku pooleni: näiteks eri tasemel keeleõppijate vahekeele tekstide alusel saab uurida teatud grammatilise struktuuri omandamist, õpikute ja sõnastike vastavust keeleõppijate tegelikele vajadustele (vt ka Kitsnik 2006: 94–99), õppijakeele korpust saab rakendada, koostades korpusepõhiseid lihtsamaid ja keerulisemaid keeleõppesüsteeme. Osa korpuseid on ette nähtud ennekõike vaid korpusepõhiseks uurimis- ja õppetööks, teised on planeeritud olema aluseks elektroonilistele keeleõppesüsteemidele. Arvutipõhise keeleõppe toetuseks on loodud mitmeid õppijakeele korpusi kasutavaid rakendusi: vigade identifitseerimine, viitamine grammatikareeglile või sõnastikule ja õpikeskkonnad. Õppijakeele korpuste koostamise põhimõtted ja analüüsivõimalused on mitmekesised.

Muudeks olulisteks õppijakeele korpuste eristajateks on veel emakeelekõnele-
jate kontrollkorpuse olemasolu, allikmaterjali keel, kirjalike tekstide või suulise kõne
kogumine, register ja tekstiliik. Mugava meetodina on viimasel ajal populaarsust
võitnud elektrooniliste keeletestide korpuseks koondamine (vt Cobb 2003). Teiseks
levinud tekstitüübiks on õppurite esseed. Samas leitakse, et õppijakeele alusuuring-
ute tarvis on oluline kasutada mitte ainult kirjalikku, vaid ka suulist allikmaterjali,
kuna võrreldes kirjalike tekstidega on suuline kõne vabam metalingvistilisest mõjust
ning peegeldab õppija vahekeele arengut ja selle taga peituvaid mentaalseid prot-
sesse paremini (vt Myles 2005: 375). Kirjaliku keele korpused toetavad eelkõige
kirjutamis- ja lugemisoskuse uurimist ning arendamist. Leidub märgendamata ja
märgendatud õppijakeele korpusi, viimased erinevad omakorda märgendamiseks
valitud info ning märgendamise automaatsuse astme poolest.

Õppijakeele korpuste märgendamisel on praegusajal kasutusel ilmselt sama
palju märgendamissüsteeme, kui palju on võõrkeele omandamise uurijaid. Sellele
vaatamata võib täheldada selget liikumist standardiseerimise suunas. Parim näide
on siinkohal CHILDES-i korpustevõrgustik², mida on arendatud 1980. aastate
algusest saadik. Korpustevõrgustiku juurde kuuluva tarkvarapaketi märgendus-
moodul on põhjalik, võimalusterohke ning tugeva tehnilise toega. Seda kasutatakse
lapsekeele (emakeele) uurijate seas standardina, mida pidevalt uuendatakse ja
täpsustatakse. Seetõttu kasutab tänaseks enamik suulise kõne õppijakeele korpusi
CHILDES-i korpustevõrgustikku, mille märgendus on mugandatud CHILDES-i TEI
standardile.³ Korpuste loomiseks ja töötlemiseks on analoogselt CHILDES-iga väga
võimalusterohke ka Esseni ülikoolis Raymond Hickey (2003) poolt koostatud 27
programmist koosnev üldine tarkvarapakett. WordSmithi korpusetarkvara leiab
näiteks konkordantse, sagedasemaid sõnu tekstis (ingl *keywords*) jm.⁴

Maailmas on loodud nii avalikke huve kui kommertseesmärke teenivaid õppi-
jakeele korpusi. Sellest sõltub ka korpuse ligipääsetavus, lisaks oleneb juurdepää-
suvõimalus ka korpuse tehnilise teostuse valikust. Oluliseks eristavaks jooneks on
ka erinevate korpuste mastaapsus (“suured” inglise sihtkeelele ja “väikesed” mitte-
inglise sihtkeelele põhinevad õppijakeele korpused).

Maailma suurimad õppijakeele korpused on enamasti inglise sihtkeelelega:
Cambridge Learner Corpus⁵ (20 miljonit sõnet), Longman Learners' Corpus⁶ (10
miljonit sõnet) ning International Corpus of Learners' English⁷. Viimane sisaldab
2 miljonit sõnet ja on samuti esimesi kirjaliku õppijakeele korpusi maailmas, selle
on loonud 1990. aastate esimesel poolel Louvaini korpuslingvistika keskus eesotsas
tuntud korpuslingvistide Sylviane Grangeri ja Fanny Meunieriga (vt Granger 2002).
Väikesemahuliste korpuste (40 000–600 000 sõnet) hulka kuuluvad nt Antwerp
Corpus of Institutional Discourse⁸, prantsuse sihtkeelelega FRIDA⁹ ning Corpus of
English by Japanese Learners¹⁰. Järgnevalt tutvustame neid korpusi lähemalt.

Cambridge Learner Corpus (CLC) ja Longman Learners' Corpus (LLC) – Cambridge'i ja Longmani õppijakeelekorpuste aluseks on kommerts-
eesmärkidel loodud kirjastuste andmekogud, mille põhjal arendatakse keeleõppe

² CHILDES koosneb mahukast vestlustekstide andmebaasist, märgenduspõhimõtete süsteemist CHAT ning tarkvara-
pakettist CLAN, mis on mõeldud korpuse tekstide transkribeerimiseks, märgendamiseks ja analüüsiks. CHILDES-is on
konkordantsileidjad, sagedusloendurid, sõnaliigimärgendajad paljudele keeltele, mida on üsna lihtne kohandada oma
uuringuvajadustega. CHILDES-i aina laiemat levikut soodustab vaba ligipääs. <http://chilides.psy.cmu.edu/> (17.08.2006).

³ TEI (Text Encoding Initiative) on rahvusvaheline juhendite kogu, mis võimaldab keeleuurijatel jt vormistada ning an-
noteerida erinevaid tekste. <http://tei-c.org/> (17.08.2006).

⁴ <http://www.lexically.net/downloads/version4/html/index.html> (27.08.2006)

⁵ http://www.cambridge.org/elt/corpus/learner_corpus.htm (17.08.2006)

⁶ <http://www.longman.com/dictionaries/corpus/learners.html> (17.08.2006)

⁷ <http://www.ceclfltr.ucl.ac.be/Cecl-Projects/lc/icle.htm> (17.08.2006)

⁸ <http://www.contragram.ugent.be/newsle12.html#Acid> (17.08.2006)

⁹ <http://www.latl.unige.ch/freetext/en/description.html> (17.08.2006)

¹⁰ <http://leo.meikai.ac.jp/~tono/paper/crg.pdf> (07.09.2006)

materjale. LLC ja CLC sisaldavad erineva emakeele ja keeletasemega õppijate vahekeeli. Vaba juurdepääs korpustele kahjuks puudub.

Saja erineva emakeelega õppurite kirjalikku inglise keelt peegeldav CLC koosneb tuntud standardsetest *online*-keeletestidest, mis jaotuvad allkorpusteks vastavalt õppija keeletasemele ja testi liigile. Märghendatavaks infoks on valitud grammatikavead. Korpuse töötlemise muudab mugavaks võimalusterohke statistikamoodul. Statistikat saab piiritleda õppija keeletaseme või testitüüpide järgi.

CLC-i andmeid kasutatakse rakendusuuringutes Cambridge'i ülikooli kirjastuse tarvis, et anda välja õppijate vajadustele paremini vastavaid õpikuid, keeletestideks valmistumise käsiraamatuid, sõnastikke ja grammatikaid.

International Corpus of Learners' English (ICLE) – rahvusvaheline inglise õppijakeelte korpus koosneb 19 erineva emakeelega üliõpilaste 500–1000-sõnelistest kirjalikest esseedest kogumahuga 2 miljonit sõnet. Allkorpuste mahuks on 200 000 sõnet. Sinna kuuluvad näiteks soome, poola, saksa, hiina, prantsuse ja leedu emakeelega kõnelejate korpused, lisaks inglise keelt emakeelena kõnelejate tekstidest koostatud testkorpus. Õppijate keelekasutust võimaldavad väga täpselt uurida erineva taustaga keelejuhtide ankeedid.

Korpuses on märghendatud ortograafiavead, leksikaalsed, leksikaalgrammatilised, grammatilised ja fraseoloogia vead. ICLE korpusele on loodud automaatne sõnaliigimärghendaja-lemmatiseerija.

Selle korpuse põhjal on viidud läbi teadusuuringuid paljudes riikides ja väga mitmekesistel teemadel (õppijakeele süntaks, sõnavara ja sõnasagedused, diskursuse ning stiili omapära, erinevate emakeeltega õppijate vahekeelte võrdlus).

Antwerp Corpus of Institutional Discourse (ACID) – Antwerpeni institutsionaalse diskursuse korpus loodi 1990. aastate kesksajaks. ACID peegeldab edasijõudnute inglise keele kasutamist ametialases suhtluses, keskendudes võõrkeelse ning emakeelse, kirjaliku ning suulise diskursuse pragmaatiliste aspektide uurimisele ja võrdlemisele. Korpuses on 900 kirjalikku teksti, mis pärinevad äri- ja mittetulusühingute ning akadeemiliste asutuste ametialasest suhtlusest. Korpus lubab võrrelda õppijakeelt emakeelekõnelejate omaga, samuti hollandlaste ja inglaste emakeelset kirjalikku ärisuhtluse keelt (vt Pelsmaekers 1997). Vaba ligipääs korpusele puudub.

Nii suulise kui kirjaliku diskursuse märghendamiseks on kohandatud CHILDES-i CHAT/CLAN-märghendussüsteem, mis võimaldab ka teha suurte tekstilõikude kvantitatiivset analüüsi. Korpuses on märghendatud kolme liiki andmeid: situatsiooniinfo, viisakusstrateegiate kasutamine ning semantiline-süntaktiline info (Pelsmaekers 1997). Situatsiooniinfo all märghendatakse alluvussuhted, tähistatakse ametikirja eesmärk konkreetses suhtlusprotsessis (algatus, küsimus jne), määratakse kommunikatsiooni domeen (nt äripakkumine) ja nn geograafiline info (nt asutusesisene vs. asutustevaheline kiri). Viisakusstrateegiad on märghendite moodulis olulisimad, rõhuasetus on ebamugavast situatsioonist väljatulemise strateegiatel. Süntaktiliste ja semantiliste märghendite moodul on vähem põhjalik ning samuti seotud viisakusstrateegiatega: kõneviis, tegumood, subjekti tüüp, põhiverbi tüüp, kõrvallause tüüp, modaalsed abisõnad jms.

FRIDA ja FreeText. FRIDA on prantsuse õppijakeelekorpus, FreeText on selle alusel loodud arvutipõhine keeleõppesüsteem. FRIDA koosneb 450 000 sõnest, mille moodustavad erinevate emakeeltega prantsuse keele õppijate kirjalikud

tekstid. Tekste püüti koguda võimalikult autentsetest keelekasutussituatsioonidest, kus õppijad käituvad loomunguliselt (seminarid, interaktiivsed õppetegevused, tekstimõistmisülesanded ning kontrollitud ja vaba tekstiloomes ülesanded). FRIDA korpuses on vead märgendatud ning iga vea tähtsust on hinnatud. Vaba ligipääs korpusele ja keeleõppesüsteemile puudub.

FreeText valmis 2003. aastal ülikoolide ja äri sektori koostöös Euroopa Liidu rahvusvahelise projektina. Osalesid Manchesteri ülikool, Genova ülikool ning prantsuse firma Softissimo. FreeText tugineb uuematele teise keele omandamise teooriatele ning on mõeldud kesk- ja kõrgtasemel prantsuse keele õppijate kommunikatiivsete oskuste parandamiseks. FreeTexti veadiagnoosimise süsteemi on treenitud FRIDA märgendatud osa peal, mille maht on 300 000 sõnet. FreeTextis töötavad kõrvuti automaatse veaotsinguga ka teised loomuliku keele automaattöötlusvahendid: süntaksianalüsaator, lausestruktuuri äratundja, lause reformuleerija ja tõlkija, kõnesüntesaator. Veadiagnoosimissüsteemis on eraldi välja toodud ortograafia-, grammatika- ja semantikavead. FreeText on esirinnas arvutipõhise keeleõppe arendamisel: siia kuuluvad interaktiivne õpikeskkond, mis võimaldab õppida prantsuse keeles koostama viit tüüpi kirjalikke tekste ning mida õpetaja saab vastavalt vajadusele kohandada; grammatikakontrollija-speller; grammatika käsiraamat, mille konkreetsetele lõikudele tehakse õpikeskkonnas viiteid; sõnasetikud.

FreeTexti automaatse veadiagnoosimise süsteemi aluseks on mõjustus- ja sidususteoorial põhinevad tehnikad (vt Vandeventer Faltin 2003). Analüsaator vaatleb õppijakeele kõiki tasandeid, parandab testülesannete vead, annab tagasisidet ning viiteid konkreetsetele grammatikareeglitele e-käsiraamatus. Grammatikakäsiraamat on koostatud veastatistika alusel ning see täieneb pidevalt õppuritele antava tagasiside põhjal. FreeTexti loomine on andnud olulise panuse keeleõppeteooria ning õpetuspraktika arengusse, avardunud on arusaamad vahekeele korpuste koostamise ja arendamise võimalustest, veadiagnostikast ning keele automaattöötlusvahendite kasutamisest vigade leidmisel ja märgendamisel.

Corpus of English by Japanese Learners (CEJL) – jaapani-inglise õppijakeele korpuse maht on ligi miljon sõnet ja sisaldab jaapanlaste kirjaliku ning suulise keelekasutuse näiteid. CEJL koosneb neljast allkorpusest: 1) ingliskeelsed kirjalikud tekstid (12–19-aastaste noorte jutustav ja argumenteeriv proosa); 2) ingliskeelne suuline materjal (argumenteerimine ja koomiksi kirjeldus); 3) jaapanikeelsed kirjalikud tekstid (jutustav ja argumenteeriv proosa); 4) inglise keele kui võõrkeele õpikutekstide korpus (kontrollkorpus). Vaba juurdepääs korpusele puudub.

CEJL-i alusel on uuritud vahekeele arenemise mustreid ning võrdlevalt ka emakeele ja õppijakeele korpuseid. Korpuse töötlemisel on kasutatud erinevat tarkvara, näiteks WordSmith ja MonoConc Pro. Märgendamisel on kolm aspekti: sõnaliigid, vead ja teksti metainfo. Veamärgendusel eristatakse nii lingvistilist klassifikatsiooni (nt vealiigitus keeletasandite järgi – fonoloogiast teksti ja diskursuse vigadeni; sõnaliikide vead; eri tasandi moodustajate vead; vealiigitus grammatikakategooriate alusel – aeg, tegumood, loendatavus, transitiivsus, referendi ja keeleüksuse vahelise suhte iseloom, mille hulka on arvatud ka kollokatsioon) kui vealiigitust sihtkeele modifitseerimisviisi alusel (nt ärajätt ja lisamine). Ärajätt on siin erinevalt ellipsist ja nullelementidest mittegrammatiline nähtus: **He'll pass his exam and I'll ø too* 'Ta läbib eksami ja ma ø ka', lisamine on keelereegli järgimine vales situatsioonis.

Veel eristab CEJL-i märgendus sihtkeele modifitseerimisviisi järgi ebakorrektselt vormimoodustust regulariseerimise, irregulariseerimise või topeltmarkeerimise tõttu, õigete vormide vale kasutust ja järjestamist. Segunemisvigadeks peetakse niisuguseid nähtusi nagu ühtesulamine, kontaminatsioon, ristassotsiatsioon ja hübriidiseerimine: **according to Erica's opinion* *'Erica arvates järgi' – õiged on: *according to Erica / in Erica's opinion* 'Erica arvates / Erica arvamuse järgi'.

Teksti metainfo sisaldab teavet ülesannete tüüpidest (loominguline tekst, suuline kõne, paaristöö, jutustamine), õppeasutusest, kust tekst on pärit, läbitud keeleõppe mahust, tööle antud hindest.

Mitte-inglise sihtkeele korpustest

Inglise õppijakeele korpused on loodud tunduvalt enam kui muudel sihtkeeltele põhinevaid, kuid oma võimaluste ja mahu poolest on mitte-inglise sihtkeele korpused (ingl *non-English learner corpora*) sageli huvipakkuvad. Pikaajalised kogemused mitte-inglise õppijakeele korpuste arendamisel on näiteks Skandinaavias, kus juba aastatel 1973–1980 Björn Hammarberg jt löid ligi 112 000 sõnelise rootsi sihtkeele **SSM korpuse**¹¹ (Svenska som Målspråk), mis koosneb eri emakeelte ja keeletasemetega õppijate rootsikeelsetest esseedest.

Stockholmi ülikooli lingvistikateaduskonnas on välja töötatud **Andraspråkets StrukturUtveckling korpuse** (ASU)¹² – õppijatelt erinevatel õppeetappidel kogutud sihtkeele struktuuri arengu korpuse ehk nn longituudkorpuse. ASU sisaldab kirjalikku ja suulist materjali rootsi kui kolmanda keele õppijatelt, kajastades nende keelekasutuse ajalist arengut. Korpuse kogumaht on 490 000 sõnet, sisaldab ka kontrollkorpust. ASU on loodud erilaadsete võrdlevate uuringute tarvis ning õppijakeele veaanalüsaatorite arendamiseks.

Rootsis on erinevad õppijakeele korpused koondatud **SVenska Andraspråks-TEexter** ehk **SVANTE korpusesse**¹³ (Göteborgi ülikool), mis kuulub üldisesse rootsi keelepanka (Språkbanken). SVANTE korpuse tekstid on lemmatiseeritud ning sõnaliikide osas märgendatud, korpust täiendatakse pidevalt.

Alates 2001. aastast kasutatakse SVANTE korpust Stockholmi kuningliku tehnoloogiasüsteemi ja Stockholmi ülikooli koostööprojektis **CrossCheck**¹⁴ (Svensk grammatikkontroll för andraspråksskribenter, rootsi õppijakeele grammatikakontrollija), mille raames tehakse ulatuslikku arendustööd rootsi keele kui teise keele elektrooniliste õppevahendite koostamisel. CrossChecki olulisemaid tulemusi on veadiagnoosimise süsteem (vt Bigert 2005), mis arvestab sõna kasutuskonteksti iseärasustega.

Rootsis on koostatud ka prantsuse, saksa ja hispaania sihtkeele õppijakeele korpused. Soomes pole seni mitte-inglise õppijakeele elektroonilistele korpustele rõhku pandud, eelistatud on tõlke- ja paralleelkorpuste arendamist, mille uurimisest saadud tulemusi on kasutatud nii tõlkijakoolituses kui keeleõppe eesmärkidel. Huvi õppijakeele korpuste loomise vastu on ilmnunud alles viimasel ajal: III Virsu konverentsil Joensuu tutvustas Jarmo Jantunen (2007) laiahaardelist rahvusvahelist projekti soome õppijakeele korpuse loomisest Oulu ülikoolis.

Norra õppijakeele korpuse **Language learner corpus of Norwegian as a second language**¹⁵ koosneb immigrandide keeletestide arhiivist. Korpused on nii

¹¹ Infot korpuse kohta vt <http://www.ling.su.se/DaLi/research/xcheckpres.html> (01.02.07).

¹² <http://www.ling.su.se/staff/ham/projects.html> (07.01.2007)

¹³ <http://www.svenska.gu.se/~svelb/svante/> (07.01.2007)

¹⁴ <http://www.csc.kth.se/tcs/projects/xcheck/index-en.html> (07.01.2007)

¹⁵ <http://decentius.aksis.uib.no/corpus/askdemo-home.xml> (01.02.2007)

sõnaliigi-, süntaktiline kui vealiigimärgendus. Veaanalüsaator lisab veamärgendi, annab hinnangu vea raskusastme kohta ja teeb veaparanduse. Teistest Euroopas kõneldavate keelte õppijakeele korpustest võib esile tuua rahvusvahelise itaalia õppijakeele korpuse VALICO¹⁶ ning FLLOC-andmebaasi¹⁷ (koondab kuut prantsuse sihtkeelega suulist õppijakeele korpust).

Eesti vahekeele korpuse üldtutvustus

Tallinna Ülikooli üld- ja rakenduslingvistika õppetooli Eesti vahekeele korpus¹⁸ kuulub väikeste mitte-inglise õppijakeele korpuste alla, sisaldab vene emakeelega eesti keele õppijate kirjalikke töid, on koostatud ja elektrooniliselt töödeldud kindlal uurimistö eesmärgil. Korpuse olulisimaks töövahendiks on konkordantsiprogramm, mis annab erinevate kriteeriumide alusel paindlikud otsinguvõimalused. Et luua konkordantsiprogramm, tuli eelnevalt defineerida vea mõiste, keelevigade liigitamise alus ning sellest lähtuvalt jõuda mitmemõõtelise lingvistilise veaklassifikatsioonini (vt EVKK mitmemõõtelisest lingvistilisest veaklassifikatsioonist).

EVKK on kõigile veebis kättesaadav ükskeelne avatud korpus, millesse võib tekste jätkuvalt lisada. 2006. aasta lõpus on EVKK osaliselt märgendatud tekstikogu, mis sisaldab osaliselt tasakaalustatud alg-, kesk- ja kõrgetaseme tekste erineva sotsiaalse taustaga keeleoskajatelt kogumahu 459 731 sõnet (märgendatud 75 749 sõnet, millest 8123 olid vigase keelekasutuse näited). Korpuse alusel saab uurida nii korrektset kui vigast õppijakeelt. Korpust saab kasutada tekstiarhiivina (ilma märgenduseeta) või märgendatud variandina, kus keeleviga on välja toodud ja lingvistiliselt tõlgendatud. Tekstikogust saab leida samaliigilisi vigu, näha neid kitsamas kontekstis (süntagma), ± 1–2 lauses või konkreetsetes terviktekstis. Jälgida saab ka mingi sõna õigeid ja vigaseid esinemisnäiteid ning sama vealiiki ühes tekstis, ühes tekstitüübis või kõikides korpuse tekstides. Samuti on võimalik keelevigade infot seostada õppuri sotsiaalse päritolu, staatuse, hariduse, soo, vanuse ja emakeelega ning keelevaldamise tasemega (sarnaselt Euroopa Nõukogu keeleoskustasemete süsteemiga A-, B- ja C-tase).

Järgnevalt anname ülevaate EVKK teksti- ja informandivalikust, samuti korpuse väljatöötamise etappidest.

EVKK metainfo kohta annab jooksvalt teavet korpuse statistikalehekülj. Näiteks autori keeletaseme seisukohalt sisaldab korpus hetkel kõige enam kesktaseme tekste (52,4%), järgnevad alg- (22,5%) ja kõrgetaseme (25,1%) tekstid. Autorite sooline kuuluvus (64,7% naised, 35,3% mehed) ja haridustase (keskharidus 67,4%, põhiharidus 18,7% ja kõrgharidus 13,4%) ei ole tasakaalus. Disproportsioon on nähtav samuti tekstide autorite vanuses (52,9% informantidest on kuni 18-aastased) ja sotsiaalses taustas (49,5% üliõpilasi, 27% õpilasi, 16% teenistujaid jne). Autorite valikul on teadlikult piiratud kahe Eestimaa regiooniga, kus on kõige rohkem venekeelset elanikkonda: 59,9% korpuse tekstide autoritest on pärit Tallinnast või Harjumaalt ja 29,1% Ida-Virumaalt. Samas tuleb märkida, et informantide hulka, kes on end kirja pannud Tallinna/Harjumaal elanikuna, võib olla sattunud ka Ida-Virumaalt ja teistest Eesti piirkondadest pärit inimesi (Tallinnas õppimas või täiendkoolitusel).

Teksti tüüpidest sisaldab EVKK kõige rohkem abivahendeid kasutamata kirjutatud esseid (46%), vastuseid küsimusele (17,1%) ja kirju (8,6%). Eelistatud

¹⁶ www.bmanuel.org/projects/br-HOME.html (29.12.2006)

¹⁷ <http://www.flloc.soton.ac.uk> (4.01.2007)

¹⁸ <http://evkk.tlu.ee> (26.02.2007)

on loominguulist laadi töid. Et aga õppetöös tuleb ette ka muid kirjaliku kontrolli vorme (referaat, tõlge, grammatikaharjutus, muu – 10,7%), siis ei olnud põhjust neid õppijakeele tekste korpusest välja jätta.

Metainfo statistika põhjal saab teha mõningaid järeldusi korpuse tasakaalustatuse kohta. Samas aga ei tasu tasakaalustatuse eesmärki korpuse tegemisel absolu-tiseerida. Võõrkeeli omandatakse enamasti noores eas, rahule jäädakse põhiliselt alg- ja kesktasemega, naised õpivad meestest meelsamini jne.

EVKK arendamine on toimunud etapiti. Esimene etapp oli tarkvara valik sea-tud eesmärkidest lähtuvalt. Sobivat vabavara EVKK töötlemiseks mitmemõõtelise lingvistilise veaklassifikatsiooni alusel (vt Eslon 2006b: 15–17, 19–20) ei olnud võimalik leida. Seetõttu tuli luua originaalne tehniline teostus (insener Vahur Rebas), mille aluseks võeti Tallinna Ülikooli haridustehnoloogia keskuse tarkvara ja avatuse põhimõte.

Järgmise etapina asuti korpust käsitsi märgendama. Töö käigus selgus, et nii tehnoloogilist poolt kui esialgset lingvistilist veaklassifikatsiooni tuleb järk-järgult täpsustada. Selles tööfaasis ei saa kasutada olemasolevaid eesti keele tehnoloogilisi ressursse (vt ülevaade Muischnek jt 2003). Näiteks morfoanalüsaator analüüsib grammatiliselt õigeid vorme, kuid vigade leidmise ja märgendamise eesmärgil on see kasutu.

Korpuse loojad on otsustanud avada 500 000 sõnest koosneva märgendatud korpuse, milles sisalduvat materjali saab kasutada nii teadusliku uurimistö ees-märgil kui õpetajakoolituses ja praktilises keeleõppes. Kasutajal on võimalik luua korpuse oma töökeskkonda ning arendada veamärgendust ja tekstivalikut vastavalt kitsale uurimisteamale.

Hetkel pole korpuse kodulehel avalikkusele veel kättesaadavad päring, statis-tika, veaklassifikatsioon ning märgendamata ning märgendatud tekstid, küll aga saab EVKK ja õppijakeele korpustega seotud infot. Korpuse kõiki mooduleid on kasutatud Tallinna Ülikooli eesti filoloogia osakonna üliõpilaste ning magistrantide õppetöös ja uurimistevõttes.

EVKK mitmemõõtelisest lingvistilisest veaklassifikatsioonist

EVKK-s on keeleviga mõistetud grammatikareeglile või kommunikatiivsele ees-märgile mittevastava keelekasutusena, mille hulka ei kuulu väsimusest ja hooletusest põhjustatud eksimused ning keelevääratused. Samuti ei vaadelda vigadena keele-lisest kreaativsusest põhjustatud variatsioone, mille eesmärk on erinevate keelesi-seste võimaluste mänguline kasutamine mõtte väljendusrikkamaks edastamiseks. Veaklassifikatsioonist on välja jäetud ka keele kultuurilised ja psühholoogilised aspektid.

Veaklassifikatsiooni loomisel on arvestatud keelendi õigsust edastatava tähenduse seisukohast (semantika), keelendi vormilist õigsust (grammatika) ning kasutuse sobivust (pragmaatika). Tegu on lingvistilise veaklassifikatsiooniga, milles lähtutakse keelest kui paradigmaatiliste ja süntagmaatiliste seoste süsteemist (vt tabel 1). Paradigmaatilised seosed on hierarhilised, iga alamtaseme keelend sisaldub igas kõrgemas tasandis ning iga kõrgema tasandi keelend hõlmab alamtasemelisi (grafeem, morfeem, sõna, sõnaühend, lause, tekst) ja vastupidi. Süntagmaatilised seosed on lineaarsed, neis kombineerub leksika-, morfoloogia- (leksikaalgramma-

tika, morfonoloogia, morfoloogia, morfosüntaks) ja lausetasandi (süntaks, registri valik ja pragmaatika) reeglistik. Minimaalne märgendusüksus on sõne, maksimaalne lause. Lausest suuremate üksuste märgendamine (nt märgendid “teksti sidususvahendite kasutamine” ning “teksti sõnastuse ja registrivaliku vastavus autori pragmaatilisele eesmärgile”) on taandatud miinimumini.

Keelesüsteemi süntagmaatilise ja paradigmaatilise telje ristumisel tekib kahe-mõõteline lingvistiline veaklassifikatsioon (vt Hufheisen 1991, Michiels 1999), mille põhjal saab eristada 18 veaklassi (vt tabel 1).

Tabel 1. Keelevigade lingvistiline klassifikatsioon

Süntagmaatika Paradigmaatika	Semantika	Grammatika	Pragmaatika
Tekst	1	2	3
Lause	4	5	6
Sõnaühend	7	8	9
Sõna	10	11	12
Morfeem	13	14	15
Grafeem	16	17	18

Iga veaklass sisaldab erineva hulga konkreetseid liike ja alamliike, mis on omakorda vastastikku paradigmaatilistes ja süntagmaatilistes seostes. Vealiike ja alamliike on EVKK veaklassifikatsioonis kokku 170. Ülemliidide hulka kuuluvad leksikaalsed ja sõnatuletusvead; leksikaalgrammatilised, morfonoloogilised, morfoloogilised, morfosüntaktilised, süntaktilised ja kommunikatiivsed vead. Iga ülemligi vead jagunevad alamliikideks. Näiteks leksikaalgrammatiliste vigade alamliikidena on välja toodud tegevuse piiritletus/piiritlematus, transitiivsus/intransitiivsus, tegevuslaadi väljendavad verbisufiks, analüütilised verbid, afiksaaladverbide kasutamine, noomeni tuletusliidete kasutamine. Iga alamliik võib omakorda jaguneda: näiteks tegevuse piiritletuse/piiritlematuse all on välja toodud omastavaline/nimetavaline ja osastavaline sihitis; ainult osasihitist võimaldavad verbid; ainult täissihitist võimaldavad verbid; konteksti vahendid, mis määravad tegevuse piiritletuse või piiritlematuse; kvantiteedisõnade kasutamine tegevuse piiritlemise vahendina jne. Üldisele veaklassifikatsioonile on lisatud ka märgend “Proovi kätt”, millega märgendatakse neid vigu, mida erinevatel põhjustel pole olnud võimalik selgelt määratleda.

Veaklasside, -liikide ja alamliikide väljatoomine ning alamliikide jagunemine veelgi kitsamateks vea ilminguteks annab võimaluse kirjeldada viga mitmemõõteliselt, rääkida veast mitte ainult kui tagajärjest, vaid välja tuua ka vea tekkepõhjust (vt Eslon 2006b). Need kaks poolt – põhjus ja tagajärg – on vigade interpreteerimisel omavahel lahutamatu seotud. Ühe või teise märkimata jätmine tekitab ebaadekvaatseid tõlgendusi, mis omakorda võivad üle kanduda õppeprotsessi, veateraapiasse ja -diagnostikasse, luues nähtumusel põhinevaid formaalseid või suisa valesid seoseid. Seetõttu toetub EVKK märgendussüsteem veaklassi, -liigi ja alamliikide vahelistele seostele. Vea tekkepõhjuste väljatoomisel arvestatakse ka

K1 ja K2 vaheliste sümmeetria, asümmeetria ja analoogiaseoste olemasoluga (vt Eslon 2006a: 22). Näiteks:

- (1) kui sõna *viskuma* asemel on ekslikult kasutatud sõna *viskama*, siis võib viga tõlgendada mitme veaklassi omavaheliste seoste alusel: veaklass 10 (semantika kokkupuutepunkt sõna tasandiga), veaklass 11 (grammatika kokkupuutepunkt sõna tasandiga), veaklass 13 (semantika kokkupuutepunkt morfeemi tasandiga), veaklass 14 (grammatika kokkupuutepunkt morfeemi tasandiga), veaklass 16 (semantika kokkupuutepunkt grafeemi tasandiga) ja veaklass 17 (grammatika kokkupuutepunkt grafeemi tasandiga). Teisisõnu: semantika- ja grammatikaviga on eristatav grafeemi tasandil (16 ja 17), kuid põhjustatud eksimusest morfeemi tasandil, kuna ilmselt ei ole õppija omandanud *u*-refleksiivi semantikat (13) ja grammatikat (14), mille tagajärjeks on kahe erineva sõna semantiline (10) ja grammatiline (11) äravahetamine lauses.
- (2) Ainsuse osastava käände ekslik lõpp *-k* sõnas *arvutik* (*Ma õpen vell arvutik*) võib olla interpreteeritud kui grammatikaviga, mis avaldub grafeemi tasandil (veaklass 17 – grammatika kokkupuutepunkt grafeemi tasandiga, grammatikaviga grafeemi tasandil) ning on põhjustatud oskamatusest moodustada ainsuse osastavat käännet (veaklass 14 – grammatika kokkupuutepunkt morfeemi tasandiga, grammatikaviga morfeemi tasandil). Viga on võimalik mõtestada ka häälduspärase kirjaviisina, mille võib olla tinginud õppija metakeeleline seisukoht – “kirjutan nagu kuulen”. Selgelt on see nähtav ka näiteks sõna *prägu* (õige: *praegu*) kasutamisel (*Ja prägu ma kirjutan sulle kirja eesti keelest tunni*). Tegelikult kajastub häälduspärases kirjaviisis vene ja eesti häälikusüsteemide erinevus, mis põhjustab nihkeid häälikute tajumisel ning ilmneb häälduses ja õigekirjas.
- (3) Lauses *Oman kõrgharidust ja neljaaastane töökogemus sissetuleva turismi alal* on allakriipsutatud sõnaühend saanud korpuses mitu märgendit. Kõigepealt on esile toodud sõnaühendi grammatikaviga (veaklass 8 – grammatika ja sõnaühendi kokkupuutepunkt): ainsuse nimetava käände asemel peab kasutama ainsuse osastavat (*Oman .. nelja-aastast töökogemust pro neljaaastane töökogemus*). Seejärel on märgendaja näinud lause tasandi grammatikaviga (veaklass 5 – grammatika ja lause kokkupuutepunkt): liidetud on normaallause ja omajalause, milles sisalduvatel predikaatverbidel on erinev argumendistruktuur (*kes omab mida?, kuid kellel on mis?*). Õppijakeeles on aga toimunud üldistamine, mille tulemusel nimetatud kaks struktuuri on ühendatud normaallause malli alusel. Järelikult saab välja tuua süntaksivea alamliigi – lauseliikmete ärajätmine (*Oman kõrgharidust ja mul on nelja-aastane töökogemus*). Teise interpretatsiooni kohaselt võib siin näha kahe objektnoomeni ühildumisviga arvus ja käändes (*Oman kõrgharidust ja nelja-aastast töökogemust*). Analüüsitav viga on märgendatud ka verbireksiooni veana (veaklass 8), objektnoomeni käände veana (veaklass 14).

Toodud näited kinnitavad ammuteada seisukohta vea interpretatiivsest iseloomust (Corder 1981) – tavapäraselt saab üks keeleviga ikka mitu märgendit ning see ei johtu veaklassifikatsiooni puudulikkusest, vaid on vea olemuse väljendus. Kui loomuliku keele süsteemi lingvistilise kirjeldamise puhul räägitakse interpretatiivsetest

kategooriatest ja keelendite keerulisest funktsionaalsest potentsiaalist (vt Бондарко 1996), siis sama kehtib ka vahekeele süsteemi ja selles ilmnevate vigade suhtes – interpretatiivsus on veale iseloomulik tunnus. Järelikult peab vea olemusega adekvaatne klassifikatsioon olema mitmemõõteline: iga sügavama tasandi viga on iga kõrgema tasandi konkretiseering vähemalt kolmes aspektis (semantika, grammatika, pragmaatika). Hetkel on EVKK vealiikide määramise aluseks ligi 300 tunnust. Niisugust veaklassifikatsiooni võib visualiseerida vealiikide vastastikuste sõltuvuste puuna, tänu millele võime näha vigade hierarhiat ning kirjeldada viga ja selle tekkepõhjusi tunduvalt avaramalt, samas ka täpsemalt kui seni on tehtud (nt Stemberger 1985: 33–39, Lähdemäki 1995, Pool, Vaimann 2005).

Lisaks võimalusele õppijakeelt süsteemselt uurida saab mitmemõõtelise veaklassifikatsiooni alusel parema ülevaate ka sellisest tähtsast mõjurist sihtkeele omandamisel nagu lähte- ja sihtkeele sisesed keelendite ja kategooriate vahelised üleminekulad ehk leksikaalgrammatiline perifeeria, milles ilmnevad keelendite sekundaarsed funktsioonid ning millele tugineb keeleline kreatiivsus. Just kategooriate üleminekuladel ja keelenditevahelistes seostes realiseerub nende funktsionaalne potentsiaal, võimalus olla kasutatud samades või erinevates sekundaarsetes funktsioonides ning samades või erinevates kontekstitüüpides. Seetõttu ongi oluline uurida ühelt poolt K1 ja K2 grammatilist perifeeriat ning teisalt otsida kahe erineva keelesüsteemi vahelisi seoseid, mis on aluseks emakeele positiivsele mõjule teise keele omandamisel (vt Kaivapalu 2005: 23–38).

Samas on ka vahekeele analüüs näidanud, et “vead tekivad eelkõige sel juhul, kui õpitava ja emakeele keelendeid kasutatakse sekundaarsetes funktsioonides” (vt Гак 2001: 18). Järelikult aitab seoste ning üleminekulade leidmine erinevate veaklasside, liikide, alamliikide jne vahel meil paremini mõista nii õppijakeele olemust kui vigade tekkepõhjusi. Seoste ilmnmisel peaks olema võimalik teatud vealiikide järgi prognoosida teiste nendega enamseotud vigade esinemise tõenäosust.

Järgnevalt mõned näited sellest, millist osa etendavad lähte- ja sihtkeele vahelised seosed, samuti mõlema keele süsteemi perifeersed nähtused vene emakeelega eesti keele õppijate vigade tõlgendamisel. Näited (4–8) on leitud EVKK märgendatud korpuseosast päringumooduli abil.

- (4) Otsingule “põhitähenduse viga” andis päring hetkel näiteid 26 tekstist, milles see vealiik on märgendatud. Näidete hulgast valisime sõna *värske* ebakorrekse kasutuse (*On vaja, et meie riietus on värske ja puhas*), mis on seotud selle sõna sekundaarse tähendusega, eelkõige õppija emakeeles. Vea täpsustusena on lisatud märgend “sõnade semantiline ühildamatus”. Põhjus: vene keeles on sõna *свежий* ‘värske’ sekundaarseid tähendusi ‘äsja pestud ja triigitud, puhas’ (rõivastest kõnelemise kontekstis: *на нём свежая рубашка / рубашка не первой свежести = tal on seljas puhas / mitte eriti puhas särk*). Eesti keeles see sekundaarne tähendus puudub – *riietus* ei saa olla *värske*, kuna need kaks sõna on omavahel semantiliselt sobimatud. Küll aga võib riietusese olla *värskelt triigitud*, s.t *äsja*. Näide kuulub veaklassi 7 (semantika kokkupuutepunkt sõnaühendi tasandiga), vealiik – leksikaalne viga, alamliik – väljendustava vastu eksimine, alamliigi jaotused – sõnade semantiline ühildamatus ja kollokatsioon. Toodud näite alusel saab väita, et K1 ja K2 üleminekulaks ning samas neid keeli siduvaks nähtuseks on

adjektiivide *värske* ja *свежий* semantiline ühildamatus/ühildatus substantiividega *riietus* ja *одежда*.

- (5) Otsingule “paronüümi kasutamine”, mis kuulub leksikaalsete vigade alamliiki “põhitähenduse viga”, leidis päring vastuseks 30 teksti. Analüüsime tekstikatket *Ilmuvad igasuguseid kirjanduserühmid näiteks: “Tarapita” – mis oli pooleli poliitiline, ja pooleli kirjanduslik ühend* (õige: *ühendus*), *veel üks kirjanduse rühm “Siuri” (Maria Under, Tuglas, Alle ja teised.)*. Õppijakeeles on segi läinud sihtkeele häälduselt lähedased, kuid vormilt ja tähenduselt erinevad sõnad *ühend* ja *ühendus*, mille tulemusena tekib paronüümia (veaklass 10 – semantika kokkupuutepunkt sõna tasandiga). Samas kannab see juhtum ka märgendit “afiksaalse sõnatuletuse viga (nimisõnaliite *-us* kasutamine)” – veaklass 11 (grammatika kokkupuutepunkt sõna tasandiga) ja veaklass 14 (grammatika kokkupuutepunkt morfeemi tasandiga). Vea tekkepõhjus peitub sihtkeele sõnade *ühend* ja *ühendus* tähendus- ja vormierinevuste mittenägemises, mille tulemusel õppija ei adu piiranguid nende semantilises ühildatuses/ühildamatuses (vrd *keemiline ühend* = *химическое соединение*, kuid *kirjanduslik ühendus* = *литературное объединение*). Positiivne ülekanne oleks võimalik seoste loomisel lähtekeelega, kus *ühend* ja *ühendus* vasteks on nii vormilt kui põhitähenduse poolest omavahel erinevad sõnad *соединение* ja *объединение*.
- (6) Otsingule “samasse mõistepesasse kuuluva sõna kasutamine” leidis päring vastuseks 8 teksti, mille hulgast valisime järgmise näite: *Minu arvates see oleks tore reis. Võib* (õige: *saab*) *tutvuda uute inimestega ja veeta aega rõõmsalt ja kasulikult*. Lauses esile toodud viga kuulub klassi 10: liik – leksikaalne viga, alamliik – põhitähenduse viga, mille jaotuseks on “viga samasse mõistepesasse kuuluva sõna kasutamisel”. Täpsustusena on lisatud märgend “viga modaalsõnade kasutamisel”. Teise interpretatsiooni kohaselt võib antud juhtumit siduda veaga sõna tähendusvarjundite eristamisel (klass 10, liik – leksikaalne viga, alamliik – tähendusvarjundi viga). Põhjus peitub ühelt poolt kahe sõna semantika osalises kattumises, teiselt poolt erinevustes kontekstuaalsetes piirangutes. Modaalverbi *võima* tähendusväljas saab eristada vähemalt nelja võrdväärset komponenti: 1. ‘(hüpoteetiliselt) võimalik olema’, 2. ‘reaalne olema, realselt võimalik olema’, 3. ‘lubatud olema’, 4. ‘suuteline olema, oskama’. Verbi *saama* põhitähendus on aga mittemodaalne – ‘omandama, kätte saama’. Modaalverbide klassiga on ta seotud sekundaarse tähenduskomponendi ‘reaalne olema, realselt võimalik olema’ vahendusel (vrd *Siin on väga kätärrikas. Lähme ärklikorrusele, seal saame puhata*), millele lisandub ajatähenduse varjund – lähitulevikulisus. Kuna analüüsitava tekstikatke autor kirjutab reaalsest võimalusest tutvuda lähitulevikus uute inimestega ning veeta aega rõõmsalt ja kasulikult, siis lause mõtet arvestades tuleks sõna *võib* asemel kasutada *saab* – see, millest räägitakse, saab olema ja saab tulema. Tegu on veaga, mitte aga sünonüümsusest tuleneva variatiivsusega. Sellest annab järeldada, et modaalverbide *võib* ja *saab* kontekstiseotus tuleneb keelesisestest varjatud protsessidest, mille aluseks on keelendite sekundaarsed tähendused ja funktsioonid keelesüsteemi leksikaalgrammatilises

perifeerias. Vene emakeelega õppijale ei ole eesti keele perifeersete nähtuste alla kuuluv erisus sugugi kergesti arusaadav, kuna vene avaratähenduslikul modaalverbil *мочь* 'võima' on väga rikas semantiline potentsiaal, mis ei pane verbi kasutamisele kontekstuaalseid piiranguid. Sama laieneb ka modaal sõnale *можно* 'võib'.

Lisame näiteid teistelt keeletasanditelt.

- (7) Morfosüntaksi all on ühe alamliigina välja toodud viga "ma-infinitiivi kasutamine seoses faasiverbiga": *Seal me hakkame elada ja ööbime* (õige: *hakkame elama ja ööbima*). Teise interpretatsiooni kohaselt on siin morfoloogiaviga: alamliik – indikatiivi oleviku ajavormi kasutamine (*Seal me elame ja ööbime*). Analüüsitava näite puhul on oluline, kuidas teksti autor mõtestab sündmuse ajalist kulgu. Sõlmküsümuseks saab tulevikulisuse semantika väljendamine. Eesti keele analüütilise verbikonstruktsiooniga *hakkame elama* ja vene keeles sellele vastava imperfektiivse aspekti analüütilise tuleviku vorm *будем жить* edastavad meile arusaama kaugemas tulevikus toimuma hakkavast tegevusest. Sel puhul on eesti tegevusverb alati *ma-infinitiivi* vormis ja vene tegevusverb imperfektiivses aspektis. Sisuliselt on siin tegu keeltevahelise üks-ühese vastavusega. Sama nähtus on jälgitav ka eesti keele sünteetilise verbi oleviku vormi *vaatame* ja vene keele perfektiivse aspekti oleviku-tuleviku vormi *посмотрим* vahel, sest mõlemad väljendavad lähitulevikulisust, tegevuse vahetut järgnemist kõnecomendile, nt *vaatame, mis teha annab = посмотрим, что можно сделать; astun kohe tema juurest läbi = сразу же зайду к нему* jne. Olenevalt lause semantilisest, samuti ajadeiktiliste sõnade ning muude rõhumarkerite olemasolust kontekstis võib ajatähendusele lisanduda ka modaalne tähendusvarjund – kõnecomendile järgnev tegevus on reaalselt võimalik (*Nad lähevad ju otavahel tilli = Они же поспорятся*). Kuna eesti ajavormidel ja vene aja-aspektivormidel on tulevikulisuse semantika edastamisel ühesugune funktsionaalne potentsiaal, siis kõneleb see üks-ühesest seosest nende vahel. Seetõttu on vene emakeelega õppijal kaugema ja lähitulevikulisuse väljendamisel võimalus toetuda keeltevahelisele korrelatiivsusele, mis samas aitab realiseerida K1 keelekompetentsis sisalduvaid teadmisi, luues tingimused positiivse ülekande tekkimiseks.
- (8) Otsing "moodustajate süntaktilised funktsioonid" tõi päringus välja 32 teksti; üks selle vealiigi alamliikidest on "sihitise viga", nt *Me ka kasutame pesuvahendid* (õige: *pesuvahendeid*). Eesti keeles sõltub objektnoomeni käändevalik mitmest tegurist: näiteks objekti määratlematusest (umbmäärane kogus), kirjeldatava situatsiooni imperfektiivsusest/perfektiivsusest.

Objekti määratlematus/määratletus ning selle olulisus aspektisituatsiooni mõestamisel on vene lähtekeelega õppija jaoks erandlik nähtus, kuna tema emakeeles sõltub sihite kasutamine transitivverbi reksioonistruktuurist, on seotud käandekategooriaga ning alles seejärel räägitakse objekti olemusest (elus/eluta, haaratud tegevusest osaliselt/täielikult, otsene/kaudne suunaobjekt). Tegum on erinevustega keelemeeles, mis kajastub ka grammatikakirjeldustes ja terminoloogias. Katsed

ühitada eesti ja vene keele objektuuringutes otse- ja kaudsihitist on seetõttu esile kutsunud vastuväiteid: eesti keeles puudub kaudobjekti süntaktiline omapära (vt Erelt 2002). Seega on eesti keele kui teise keele õppes mõttekas mitte arutada selle üle, kas objekt on määramata/määratud, vaid toetuda objektinoomeni käändevaliku õpetamisel mõlema keele jaoks universaalsetele momentidele nagu kirjeldatava situatsiooni imperfektiivsus/perfektiivsus.

Kuna näitelause (8) semantika viitab tegevuse pidevusele ja seega imperfektiivsele iseloomule, siis kasutatakse objektinoomeni reeglina osastavas käandes. Sama viga on märgendatud ka leksikaalgrammatilisena, alamliik – tegevuse piiritletuse/piiritlematuse väljendamine. Selgituseks: kui subjekt midagi pidevalt kasutab, siis on tegu piiritlemata tegevusega, mida võib mõtestada ka subjektile omase harjumusena (meil on harjumus või komme kasutada pesuvahendeid). Kui aga millegi kasutamine on subjekti jaoks ammendunud, siis kirjeldatakse tegevust piiritletuna. Selle väljendamiseks sobivad eesti keeles ühendverbid, antud juhul *ära kasutama*. Siinkohal võib näha selgepiirilist korrelatsiooni K1 ja K2 vahel: eesti keeles on tegevuse piiritlematuse/piiritletuse ja imperfektiivsuse/perfektiivsuse edastamisvahend objektinoomeni käändevorm (partitiiv vastandatuna nominatiivile/genitiivile); vene keeles aga vastandus imperfektiivne/perfektiivne aspekt. Peale selle on eesti keele objektinoomeni käände kasutamine seotud ka sünteetilise/analüütilise verbivormi valikuga. Neist esimene kirjeldab piiritlemata tegevust – *kasutan, kasutasin pesupulbrit; olen, olin kasutanud pesupulbrit*, teine aga tegevuse piiritletust – *kasutan, kasutasin pesupulbri ära; olen, olin pesupulbri ära kasutanud*. Võrdluses vene keelega ilmneb vastavus eesti keele sünteetilise verbi ja vene keele imperfektiivse aspekti ning eesti analüütilise verbi ja vene perfektiivse aspekti vahel. Järelikult on lause aspektuaalse semantika interpreteerimine K2-s tihedalt seotud K1 keeleteadmistega ja universaalset laadi keeltevaheliste korrelatsioonidega.

Kokkuvõtvalt: mitmemõõtelise veaklassifikatsiooni alusel saab uurida sihtkeele reeglite vale kasutamist ja ekslikke ülekandeid õppijakeeles süsteemselt, tuues välja teise keele omandamiseks olulised universaalsed nähtused, mis tingivad vajaduse uutal alustel rajanevate ainekavade, õpikute, õppematerjalide loomiseks. Selles osas on mõttekas võrrelda ka õppijakeelt ja sõnasagedust loomuliku keelekasutuse ja sõnasagedusega.

EVKK ja pedagoogiline grammatika

Tavapäraselt on pedagoogilise grammatika loomisel tuginetud kontrastiivgrammatikale, mille ideoloogia on seotud teesiga K1 interferentsi otsustavast mõjust K2 omandamisel. Nii on rõhutatult esile toodud K1 negatiivne mõju K2 omandamisele. Selle teooria seavad kahtluse alla paljud vahekeele uuringud, samuti eelnenud veaanalüüsi näited (vt EVKK mitmemõõtelisest lingvistilisest veaklassifikatsioonist).

Vahekeelt võib vaadelda seoste võrgustikuna lähte- ja sihtkeele vahel, mida õppija ei loo kontakteeruvate keelte kontrastsuse põhjal, vaid vajadusest K2-st aru saada, otsides seejuures kokkulangevusi, lähedasi ja analoogseid nähtusi. Ning kui midagi jääb seletamata, seosed eelneva ja järgneva vahel on loomata, siis tekitab õppija neid ise – lihtsatest assotsiatsioonidel põhinevatest ülekannetest abstraktsete

analoogiateni (nt Ehala 2000: 28). Seetõttu peaks pedagoogilise grammatika loomisel tingimata arvestama vahekeele veaanalüüsi tulemusi: missugused vead ja kui sageli ette tulevad (vea esinemissagedus), kuidas vead on omavahel seotud (seoste statistika), kuidas õppijakeele vead on seotud sihtkeele autentse sõna- ning vormikasutuse ning -sagedusega, sõnaühendite moodustamise ja muu sarnasega. Alles seejärel tasub asuda otsima K2 omandamise tegelikke raskuspunkte. Seejuures on muidugi oluline, et vahekeele uurimine oleks laiapõhjaline ning tugineks erinevate lähtekeeltega õppijate sihtkeele kasutusnäidetele nii kirjalikus kui suulisel kõnes. Niisugust seisukohta on varemalt väljendanud nt F. Myles, kes leiab, et õppijakeelt saab analüüsida kahel viisil: 1) seades kitsamaid eesmärgi ja võrreldes õppijate vahekeelt sihtkeele korrektse grammatikaga ning 2) seades avaramaid eesmärgi ja uurides õppija vahekeelt kui iseseisvat terviklikku lingvistilist süsteemi, milleks aga on vaja võrrelda erinevaid emakeeli rääkivate õppijate keelekasutust liikumisel sama sihtkeele poole (Myles 2005: 378, vt ka Granger 1998).

Võib minna ka teist teed ja võrrelda vahekeele veaanalüüsi tulemusi K1 ja K2 vaheliste seoste grammatikaga ehk korrelatsioonigrammatikaga (mille kohta vt Eslon 2006a: 24, Eslon 2006b: 17–21). Vastupidiselt kontrastiivgrammatikale ei tugine korrelatsioonigrammatika keelekontrastidele, vaid püüab näha tüpoloogiliselt erinevate keelte vahel üks-üheseid vastavusi ning samalaadsust, mille tulemusena rakendub lähtekeele analoogiapõhine positiivse ülekande mehhanism. Just seetõttu ei saa keeleõppes välja jätta K1 ja K2 süsteemide kõrvutatavat kirjeldust, kuid erinevalt kontrastiivgrammatikast tuleb see viia teistlaadi teoreetilis-metodoloogilisele alusele – põhjuslike seoste tuvastamisele lähte- ja sihtkeele vahel (vt Eslon 2006a: 17, 19–20). Pedagoogilise grammatika koostamise eesmärgil on mõttekas vahekeele veaanalüüsi tulemusi võrrelda K1 ja K2 korrelatsioonigrammatikaga ning leida vahekeelt ja korrelatsioonigrammatikat ühendavaid momente. Tegelikult tähendab see keeleteooria, rakenduslingvistika ja ka ainedidaktika integreerimist keeleõppe eesmärgil (vt ka Hudson 2004).

Rääkides seni koostamata eesti keele kui võõrkeele pedagoogilisest grammatikast, on käesoleva artikli autoritel eelnevalt olnud mõte seostada selle loomine EVKK-põhise vahekeele analüüsi tulemuste ja korrelatsioonigrammatikaga. EVKK veaanalüüsiga on seotud empiiriline suund pedagoogilise grammatika loomisel, korrelatsioonigrammatika väljatöötamisega aga teoreetiline suund. Muidugi oleks ideaalne luua eesti-x-keele korrelatsioonigrammatika, mis teatud mõttes oleks universaalse grammatika näide. Esialgu jääb see ilmselgelt üle jõu käivaks ülesandeks, kuna eeldab ühe võimaliku väljundina keelilise käitumise modelleerimist erinevates keeltes ning suhtlussituatsioonides. Esialgne töö piirneb vene emakeelega õppijate vahekeeletekstides sisalduvate vealiikide, alamliikide jne määramise ning nendevaheliste seoste väljaselgitamisega, mis sagedusstatistika rakendamisel aitab jõuda vigade tekkepõhjusteni. Paralleelselt toimub uurimistöö eesti-vene korrelatsioonigrammatika koostamiseks. Aluseks on võetud funktsionaalsemantilised väljad (vt Eslon 2006a: 20–21).

Pedagoogilisel grammatikal on võõrkeeleõppes keskne koht: grammatika teeb valiku keeleainesest, mis ainekavas järjestatakse keelendite sageduse, produktiivsuse ja omandamise raskuse alusel. Pedagoogilise grammatika põhjal on võimalik üles ehitada süsteemne keeleõpe, koostada selle läbiviimiseks vajalik ainekava, kirjutada õpikud, luua sõnastikud ja muud õppeprotsessi toetavad vahendid.

Vahekeele korpuse alusel tehtud veastatistika ning erinevatesse veaklassidesse ja -liikidesse kuuluvate vigade omavaheliste seoste võrgustiku kindlakstegemine aitab välja töötada eesti keele õppeks vajaliku reeglite süsteemi.

Kirjandus

- Bigert, Johnny 2005. Automatic and Unsupervised Methods in Natural Language Processing. Doctoral Thesis. Stockholm. <http://www.nada.kth.se/~johnny/docs/thesis.pdf> (9.01.2007).
- Бондарко Александр Владимирович 1996. Проблемы грамматической семантики и русской аспектологии. Санкт-Петербург: Изд-во Санкт-Петербургского ун-та.
- Cenoz, Jasone; Hufeisen, Britta; Jessner, Ulrike (Eds.) 2001. Looking Beyond Second Language Acquisition. Studies in Tri- and Multilingualism 6. Tertiärsprachen und Mehrsprachigkeit Bd. 6. Tübingen: Stauffenberg Verlag.
- Cobb, Thomas 2003. Analyzing late interlanguage with learner corpora: Quebec replications of three European studies. – Canadian Modern Language Review 59 (3), 393–423.
- Corder, Pit 1981. Error Analysis and Interlanguage. London: Oxford University Press.
- Ehala, Martin 2000. Second language learners' impact on the structure of Estonian. – Kiira Allikmets (Ed.). Languages at Universities Today and Tomorrow. Proceedings of the Methodology Conference of the Language Centre, 19-20 May 2000. Tartu, 20–32.
- Erelt, Mati 2002. Hierarhiatset tüpoloogias. – Renate Pajusalu, Ilona Tragel, Tiit Hennoste, Haldur Õim (toim.). Teoreetiline keeleteadus Eestis. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 4. Tartu: TÜ Kirjastus, 34–40.
- Eslon, Pille 2006a. Analoogiast keelte kõrvutamisel. – Keel ja Kirjandus 1, 15–24.
- Eslon, Pille 2006b. Eesti vahekeele korpusest korrelatsioonigrammatikani. – Helle Metslang, Margit Langemets (toim.), Maria-Maren Sepper (keeletoiim.). Eesti Rakenduslingvistika Ühingu aastaraamat 2. Estonian Papers in Applied Linguistics 2. Eesti Rakenduslingvistika Ühing. Tallinn: Eesti Keele Sihtasutus, 11–24.
- Гак Владимир Григорьевич 2001. Семасиологический функциональный подход и типология функций. – Теоретические проблемы функциональной грамматики. Материалы Всероссийской научной конференции, Санкт-Петербург, 26–28 сентября 2001 г. Отв. ред. Александр Владимирович Бондарко. Санкт-Петербург: Наука, 17–19.
- Granger, Sylviane 1998. The computerized learner corpus: A versatile new source of data for SLA research. – Sylviane Granger (Ed.). Learner English on Computer. London, New York: Longman, 3–18.
- Granger, Sylviane 2002. A Bird's-eye view of learner corpus research. – Sylviane Granger, Joseph Hung, Stephanie Petch-Tyson (Eds.). Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching. Amsterdam/Philadelphia: John Benjamins, 3–33.
- Hickey, Raymond 2003. Corpus Presenter. Software for Language Analysis (With a Manual and A Corpus of Irish English as Sample Data) + CD-ROM. Amsterdam, Philadelphia: John Benjamins. <http://www.uni-essen.de/~lan300/HICKEY.htm> (27.08.2006).
- Hudson, Richard 2004. Why education needs linguistics (and vice versa). – Journal of Linguistics 40 (11), 105–130.
- Hufeisen, Britta 1991. Englisch als erste und Deutsch als zweite Fremdsprache. Empirische Untersuchung zur fremdsprachlichen Interaction. Frankfurt/M etc: P. Lang.
- Hufeisen, Britta; Neuner, Gerhart (Eds.) 2003. Mehrsprachigkeitkonzept – Tertiärsprachen – Deutsch nach Englisch. European Centre for Modern Languages. Strasbourg: Council of Europe Publishing. <http://www.ecml.at/documents/pub112G2003.pdf> (9.03.2007).

- Jantunen, Jarmo 2007. Oppijansuomen piirteitä korpusvetoisesti. – III kansainvälinen Virsu-konferenssi 19.-20. tammikuuta 2007 Joensuussa. <http://www.joensuu.fi/suomi/virsu/virsuabstraktit07.html> (6.02.2007).
- Kaivapalu, Annekatrin 2005. Lähdekieli kielenoppimisen apuna. Maisa Martin, Pekka Olsbo, Marja-Leena Tynkkynen (toim.). Jyväskylä Studies in Humanities 44. Jyväskylä: Jyväskylän yliopisto.
- Lähdemäki, Eeva 1995. Mikä meni pieleen? Ruotsinkielisten virheet suomen ainekirjoituksessa. Fennistica 11. Åbo: Åbo Akademis tryckeri.
- Michiels, B. 1999. Die Rolle der Niederländischkenntnisse bei Französischsprachigen Lernern von Deutsch als L3: Eine empirische Untersuchung. – Zeitschrift für Interkulturellen Fremdsprachenunterricht 3 (3). http://www.spz.tu-darmstadt.de/projekt_ejournal/jg-03-3/beitrag/mich1.htm (6.02.2007).
- Muischnek, Kadri; Orav, Heili; Kaalep, Heiki-Jaan; Õim, Haldur 2003. Eesti keele tehnoloogilised ressursid ja vahendid. Arvutikorpused, arvutisõnastikud, keeletehnoloogiline tarkvara. Urve Talvik (toim.). Tallinn: Eesti Keele Sihtasutus.
- Myles, Florence 2005. Interlanguage corpora and second language acquisition research. – Second Language Research 21 (4), 373–391.
- Pelsmaekers, Katja 1997. Dutch, English and English L2 business language in contrast: Working with the ACID corpus. – Contragram (Quarterly newsletter of the Contrastive Grammar Research Group of the University of Gent) 12. <http://www.contragram.ugent.be/newsle12.html#Acid> (6.02.2007).
- Pool, Raili; Vaimann, Elle 2005. Vead kõrgtasemel eesti keele kõnelejate kirjalikus keelekasutuses. – Margit Langemets (koost.), Maria-Maren Sepper (toim.). Eesti Rakenduslingvistika Ühingu aastaraamat 1. Estonian Papers in Applied Linguistics 1. Eesti Rakenduslingvistika Ühing, Eesti Keele Instituut. Tallinn: Eesti Keele Sihtasutus, 115–137.
- Selinker, Larry 1969. Language transfer. – General Linguistics 9 (2), 67–92.
- Selinker, Larry 1972. Interlanguage. – International Review of Applied Linguistics 10, 209–231.
- Selinker, Larry 1992. Rediscovering Interlanguage. New York: Longman.
- Stemberger, Joseph P. 1985. The Lexicon in a Model of Language Production. New York, London: Garland Publishing Inc.
- Tönshoff, Wolfgang 1995. Ausgewählte Forschungsergebnisse und Denkanstöße für die Unterrichtspraxis. – Fremdsprache Deutsch (Zeitschrift für Praxis des Deutschunterrichts). Sondernummer 1995: Fremdsprachlertheorie, 4–15. http://www.edition-deutsch.de/fremdsprache/pdf/fd_s095.pdf (05.10.2005).
- Vandeventer Faltin, Anne 2003. Syntactic Error Diagnosis in the Context of Computer Assisted Language Learning. PhD Thesis. Genova University. <http://www.unige.ch/cyberdocuments/theses2003/VandeventerA/these.pdf> (07.09.2006).

Pille Eslon (Tallinna Ülikool) on uurinud vene keele funktsionaalset grammatikat, üld- ja kõrvutava keeleteaduse probleeme, modaalsust, aspektuaalsust, infinitiivi.
pille.eslon@tlu.ee

Helena Metslang töötab Integratsiooni Sihtasutuses, teaduslikeks huvialadeks on süntaks, eriti lauseliikmetevahelised üleminekuvalad, ning õppijakeele korpused.
helena.metslang@meis.ee

LEARNER LANGUAGE AND ESTONIAN INTERLANGUAGE CORPUS

Pille Eslon, Helena Metslang

Tallinn University / Non-Estonian's Integration Foundation

The current article introduces the design, aims and future applications of the Estonian Interlanguage Corpus (EIC) in the background of other learner language corpora in Europe and Asia. Estonian Interlanguage Corpus is an annotated monitor corpus of learner language texts, being created at the chair of General and Applied Linguistics at Tallinn University. It comprises written Estonian language texts, mainly by Russian speakers. The objective of EIC is to support the development of pedagogy and research and of acquisition of Estonian as a second language.

EIC is a free online corpus which will by the end of its first stage of development have 500 000 strings of error-annotated texts plus a larger volume of non-annotated texts. The most important tool of EIC is the concordancer which gives flexible search options by various criteria. The specially created error classification has 170 tags and is multidimensional (there are the syntagmatic and paradigmatic axes). To the highest level of tags there belong lexical errors, derivational errors, lexicogrammatical, morphophonological, morphological errors, morphosyntactic, syntactic and communication errors and unspecified errors (ambiguous, doubtful errors).

The results of EIC error analysis are a good basis for writing a systematic pedagogical grammar for Estonian as a second language. Analyzing errors statistically and determining the network of links between the errors will help to describe better the system of rules necessary for learning Estonian as a second language. After creating the pedagogical grammar it is possible to build on it syllabi, courses, textbooks, dictionaries and other means to support the process of learning Estonian as a second language.

Keywords: corpus linguistics, interlanguage, learner language corpus, pedagogical grammar

MACHINE LEARNING TECHNIQUES IN DIALOGUE ACT RECOGNITION

Mark Fišel

Abstract. This report addresses dialogue acts, their existing applications and techniques of automatically recognizing them, in Estonia as well as elsewhere. Three main applications are described: in dialogue systems to determine the intention of the speaker, in dialogue systems with machine translation to resolve ambiguities in the possible translation variants and in speech recognition to reduce word recognition error rate.

Several recognition techniques are described on the surface level: how they work and how they are trained. A summary of the corresponding representation methods is provided for each technique. The paper also includes examples of applying the techniques to dialogue act recognition.

The author comes to the conclusion that using the current evaluation metric it is impossible to compare dialogue act recognition techniques when these are applied to different dialogue act tag sets. Dialogue acts remain an open research area, with space and need for developing new recognition techniques and methods of evaluation.*

Keywords: conversation analysis, computational linguistics, dialogue act, machine learning, Bayes classifier, hidden Markov model, neural network, decision tree

1. Introduction

Dialogue acts (DAs) have derived from speech acts introduced by Austin (Traum 1999). He introduced the notion of utterances as actions which change the state of the environment, dialogue participants, etc. (in contrast to the notion of utterances as expressions which can be evaluated to true or false). Austin distinguished three types of actions performed by utterances: locutionary (action of saying something –

* This work was supported by the Estonian Science Foundation, grant no. 5685.

shaping the utterance, pronouncing it and using it to refer to real world objects), perlocutionary (action performed by saying something – achieved effects, special to the particular situation – e.g. persuading or surprising) and illocutionary (action performed in saying something – e.g. informing, requesting, asking, answering, warning, apologizing etc.). The latter type has later on been mostly worked with in subsequent research.

Searle extends Austin's work on illocutionary acts (Traum 1999). His main contribution was the attempt to define necessary and sufficient conditions for the act to be performed. These were presented as game definition rules: conditions of normal input/output (necessary for one to express himself and others to understand), propositional content (content restrictions), environment, sincerity (alignment of actual attitudes with the ones expressed in the utterance), etc. Searle also presented several dimensions along which the speech acts can vary, and proposed a speech act taxonomy based on the dimensions.

Dialogue acts were first introduced by Bunt (1994). Although his definition almost repeats the definition of speech acts (DAs – functional units used by the speaker to change the context), the exact notion behind DAs changes from author to author. DAs are also called dialogue moves, utterance types, utterance classes and (returning to the origin) speech acts. DA recognition has also several names, such as recognition, classification, tagging, etc.

Applications. The most common application of automatically recognized dialogue acts are dialogue systems. DAs are used to recognize the intention of the speaker, which helps to determine the necessary response dialogue act (Prasad, Walker 2002), (Fernandez et al. 2005), or even works as shallow parsing – i.e. recognizing DAs is equivalent to understanding the utterance on a more general level (Lendvai et al. 2003).

Another frequent usage of DAs is machine translation in dialogue systems. A correctly recognized DA can help resolve ambiguities in translating utterances. The grammatical form of an utterance doesn't always coincide with the meant intention (e.g. "could you close the window" is not a question, but a request). In addition, different languages have grammatically different polite forms meaning the same thing.

A typical example is the VerbMobil project (Reithinger, Maier 1995, Reithinger et al. 1996, Küssner 1997). There are also other examples (Lee et al. 1997, Levin et al. 2003, etc.).

DAs can be recognized from prosodic and other speech wave features. This can be used to reduce the word recognition error rate in speech recognition (Wright 1998, Wright et al. 1999, Grau et al. 2004, Alshawi 2003).

Hagen and Popowich (2000) describe a dialogue system which is based on a grammar of dialogue acts, which is used to determine the behavior of the whole system and also constrains the possibilities in DA recognition.

Another application of automatic DA recognition is conversational analysis: for instance, it enables selecting subcorpora with specific DA tags, which might be of interest to the researcher.

In Estonia research on DA recognition is done mostly in the University of Tartu. The application is the first one mentioned: recognized DAs are used for determining the speaker's intention and for choosing the proper response DA.

Paper structure. The paper is structured as follows. The 2nd section is devoted to DA taxonomies. It describes the ones used in the projects referred to in this paper. The 3rd section reviews several DA recognition techniques, giving a brief introduction to the technique and describing how it is commonly applied to DA recognition. The 4th section gives an overview of research in DA recognition done in Estonia. The 5th section concludes the paper with a discussion.

2. Dialogue act taxonomies

Although this report is mainly devoted to DA recognition techniques, it is necessary to introduce the main DA taxonomies (or DA tag sets) that are used in the projects, serving as examples for DA recognition.

A DA taxonomy must compromise between two factors. First, the definitions of DA tags must be clear enough in order to be easily separable. If they are not, agreement between human taggers¹ will be low. On the other hand it is efficient to define a reusable taxonomy, which is general enough to be applicable to many different problems.

There seems to be little agreement on how exactly to achieve the compromise. As it can be seen from the projects referred to in this report, many of them prefer using self-defined DA taxonomies. Others use one or another existing taxonomy. The most popular taxonomy, initially designed to be universal, is DAMSL. Other taxonomies have been developed for some corpora, like CallHome or VerbMobil, and have later on gained popularity.

The acronym DAMSL stands for Dialogue Act Markup in Several Layers (Allen, Core 1997). The dialogues are annotated on four different levels, which are the communicative status, the information level, forward-looking function and backward-looking function. It is important to note that some utterances might lack a tag on some levels; for instance an utterance can have the backward-looking function specified, and the forward-looking function missing.

The communicative status isn't marked for most of the utterances. It specifies whether the utterance was uninterpretable, abandoned or was a self talk. The latter indicates that the speaker is not intending the information he talks about for other dialogue participants.

The information level annotation provides an abstract characterization of the content of the utterance. It includes four categories: task fulfilling, task management, communication management and "other level". Task fulfilling means that the utterance is directed at fulfilling the general task of the dialogue, like asking for the time of a flight. Task management indicates an attempt of coordinating the activity of the two speakers; for instance a proposal of switching the current problem or topic. Communication management represents conventional phrases that maintain contact, perception and understanding. These include greetings, closings, acknowledgements, stalling for time (e.g. "let me see"), speech repairs and misunderstandings. The other level category indicates utterances that do not fit into the first three categories: like jokes or small talk.

The forward-looking function corresponds to Austin's illocutionary act, specifying which action the utterance performs. This level includes such categories as

¹ The agreement is most commonly measured with the kappa-statistic (Siegel and Castellan 1988), which equals the percent of utterances which were assigned the same DA tag by all human taggers.

statement, exclamation, information request, etc. The backward looking function level describes the responsive aspect of the utterances, for instance acceptance, rejecting, misunderstanding, full answer, etc.

Another widely used taxonomy based on DAMSL, was designed for the Switchboard corpus (Godfrey et al. 1992) and is called SWBD-DAMSL (Jurafsky et al. 1997). 80% of the tags are present in both the original and the modified version, some tags are added (for instance “non-verbal” category in the communicative status) and some DAMSL categories are subdivided (for instance “statement-non-opinion” and “statement-opinion”).

The taxonomy designed for the Spanish dialogue corpus CallHome (Levin et al. 1999) includes in its original version 232 DA tags. These represent the combination of a general category (like statement/question) and the more specific descriptions, e.g. whether the utterance describes the emotional state of the speaker. The original set is frequently collapsed by uniting similar DA tags into one. One of such collapsed variants is the CallHome37, which has all the statements and the back-channels collapsed into a single category, and includes 37 tags. In its turn CallHome37 is collapsed into CallHome10, which includes 8 most general categories (statement, question, answer), a tag for abandoned sentences and a tag for noise.

Among the DA taxonomies designed for single projects, there is the VerbMobil taxonomy (Reithinger, Maier 1995). It also has a single layer of annotation and includes 33 DA tags. These describe the utterance from the point of the performed action; for example *GREET*, *GIVE REASON*, *REJECT*, *FEEDBACK POSITIVE*, etc.

Another taxonomy designed for a single project is the one of the Estonian Dialogue Corpus (EDiC). The taxonomy has two layers of annotation. The upper layer specifies the general typology of the utterance: ritual, questions/answers, directive, additional information, repair, etc. There are twelve types in total, out of which 7 have paired tags (one for initiation and one for response) and 5 have a single tag. The lower layer describes the utterance in greater detail. Each general type has several subtypes on this level: for instance rituals have greeting, thanking, apologizing, etc.; questions/answers have wh-questions, open and closed yes/no questions, refusal to answer, yes/no answers; and so on. In total there are 126 DA tags on the level of detailed annotation. The EDiC DA taxonomy as well as the corpus itself are comprehensively described in (Gerassimenko et al. 2004).

It can be seen that taxonomies vary greatly in size and design, which produces a number of advantages and disadvantages for automatic DA recognition.

3. DA recognition techniques

This chapter describes several techniques successfully applied to DA recognition. The techniques are rarely applied to the textual or speech wave representation of the utterance, because the connection between these and the utterance dialogue act is too complicated for the contemporary techniques to grasp. Therefore both the utterance and the dialogue act have to be encoded in order to be used as the technique’s input and output respectively.

A common way to encode an utterance is to describe it and/or its words with features. By content these can be linguistic (sentence structure class, word

morphology, parts-of-speech, etc.), prosodic (intonation changes, speech melody contour classes, etc.), keyword-based (wh-word, some content words, presence binary indicators, etc.), statistical (WEBSOM vectors significance vectors – see sct. 4) etc. The three most common feature data types are binary, numerical and nominal. Numerical features are usually encoded as scalars. Binary true/false values are replaced with 1/0 and also encoded as scalar values. Nominal features are most commonly encoded as vectors, composed of binary indicators, each corresponding to a possible feature value. Thus the component, corresponding to the current value, equals 1, and the other components equal 0. Since scalars can be viewed as 1-component vectors, the encoded features can be concatenated to compose an input vector. The output (the dialogue act) is encoded as a nominal feature.

The less common encoding methods, specific to the technique, will be described in the following subsections, devoted to the corresponding techniques.

The two most commonly used evaluation methods are the percentage of correctly classified utterances and the confusion matrix (the rows of the latter correspond to the original DA tags and the columns – to the hypotheses of the model; thus each cell shows how many utterances of some DA have been classified as another DA tag). This report will focus on the former technique, since the latter is specific to its DA taxonomy, and two confusion matrices of different taxonomies cannot be compared to each other directly.

3.1. N-gram approach

One of the simplest techniques used in DA recognition is the n-grams. It is based on the assumption that the current dialogue act is explicitly determined by k preceding dialogue acts (the Markov assumption). Therefore the candidate for the n -th dialogue act is chosen by the principle

$$c_n = \operatorname{argmax}_c P(c|c_{n-1}, \dots, c_{n-k+1})$$

The conditional probabilities are extracted from tagged corpora by counting all existing DA sequences. These are simply the number of occurrences of the sequence (c_{n-k+1}, \dots, c_n) in the training corpora, divided by the number of occurrences of a shorter sequence, $(c_{n-k+1}, \dots, c_{n-1})$. The most common values for k are 2 and 3 (in which case the technique is called, correspondingly, bigrams and trigrams). Using larger values only makes sense when longer dependencies are known to exist in the data.

Larger values heavily increase the sparse data effect. In order to lessen the effect smoothing is sometimes used. One of the most commonly used smoothing techniques is the deleted interpolation. The idea is to use n-gram probabilities of lower order in case the higher order n-gram is absent. The lower-order n-grams are penalized with smaller weights in order to provide privilege for higher-order n-grams.

The VerbMobil project is the most basic example of using n-grams (including deleted interpolation) for dialogue act recognition (Reithinger, Maier 1995). Another example is (Lee et al. 1997), where dialogue acts are used for resolving ambiguities in translations from Korean. Their method is based on the assumption that the

conditional n-gram probability of the current utterance u_n is approximated by the following product:

$$P(u_n|u_{n-1}, \dots, u_{n-k+1}) \approx P(u_n|c_n)P(c_n|c_{n-1}, \dots, c_{n-k+1})$$

3.2. Hidden Markov Models

In case of Hidden Markov Models (HMMs) a process is modeled as two parallel sequences of states, out of which one is observable and the other one – hidden. The states of the hidden sequence comprise a stochastic FSA, with the transition probabilities specified. The two sequences are aligned, and the hidden states are said to “generate” the observable ones, whereas each observable state can be generated by each hidden state with a preset probability. The last parameter, describing a model instance, is the vector of the initial probabilities of the hidden states.

Several assumptions are made about the nature of the dialogue process, the first one being the Markov assumption. The strongest one is that transition and generation probabilities remain constant throughout the process.

All three parameters that describe the model (the initial, transition and generation probabilities) are estimated from the training corpus. The estimation is commonly done with the forward-backward algorithm; applying the model (i.e. finding the most probable hidden state sequence given the observable states) is done with the Viterbi algorithm. A thorough description of HMMs can be found in (Rabiner, Juang 1986).

When modeling dialogues and DAs with HMMs, utterances (or the features that represent them) are the observable outputs and DA tags are the hidden ones. Wright (1998) uses intonation events (intonation fall/rise/etc.) as the features, achieving the precision of 72%. In (Ries 1999) HMMs are applied unconventionally: the input is composed of lattices of words and segments, taken from a speech recognizer. The Viterbi algorithm is used to derive the probabilities of all possible DAs of the utterance, which are then used as input of a multilayer perceptron (see section 3.5), along with several prosodic features. The best achieved precision is 76%.

3.3. Bayes classifiers

The Bayes classifiers get their name from Bayes theorem, which is used in the derivation of the technique. The main idea is the same as with n-gram technique – to maximize the probability of the DA tag c , but here instead of preceding acts, arbitrary features f_1, \dots, f_n are allowed to describe the utterance:

$$\hat{c} = \underset{c}{\operatorname{argmax}} P(c|f_1, \dots, f_n)$$

Unlike the n-grams, in this case estimating the probability directly from the training data would require it to be of enormous size to avoid the sparse data effect. Instead, some independency assumptions are made, which make the estimation easier. In the extreme case all features are considered to be independent and the DA tag probability is estimated by

$$P(c) \prod_i P(f_i|c)$$

This case is called the naive Bayes classifier.

It is necessary to note however, that the features are rarely fully independent in the strict sense. Instead of assuming them to be, the dependencies can be specified with a directed acyclic graph. For example, with a graph given on figure 1 the DA tag probability is estimated by

$$P(f_i|c) \cdot P(c|f_j, f_k) \cdot P(f_k|f_j) \cdot P(f_j)$$

This case is referred to as Bayesian networks. The dependency network (or the graph) can be composed manually or automatically. The latter way can be achieved by regarding the statistical dependencies between features in the training corpus and ignoring the ones with dependency strength below some quota.

Grau et al. (2004) use the naive Bayes classifier with the bag-of-words method: the feature set is composed of binary features, each indicating the presence or absence of a specific word. They also use a modified version of the classifier (uniform naive Bayes classifier) which neglects the DA probability $P(c)$ in equation 1. They achieve a result of 66% on the DAMSL-switchboard corpus and DA taxonomy. Ivanovic (2005) describes application of the same modification to a subset of the DAMSL tag set (12 tags), which results in 80% precision. Levin et al. (2003) also use the bag-of-words, but they use binary grammatical features instead of word features. They achieve a precision of 51% using the NESPOLE corpus (Costantini et al. 2002).

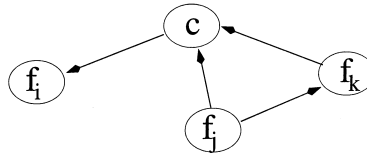


Figure 1. An example of a Bayesian network

Keizer et al. (2002) describe 2 experiments of DA recognition with Bayesian networks. In the first they use three features: sentence type (declarative, imperative, etc., a total of 10 types), subject type (1st/2nd/3rd person) and sentence punctuation (ends with a question or exclamation mark, a full stop, no punctuation, etc.). The network is composed manually and is quite small and simple. They achieve 44% precision with the Schisma corpus (slightly modified DAMSL tag set).

They further apply the technique to their own small dialogue corpus annotated with the same tag set. The used features describe utterances on the surface level: keyword-based features, features indicating whether the sentence starts with some predefined sequence, whether the sentence is/isn't a wh-question, etc.; a total of 13. The network is generated automatically in an iterative manner: first a small part of the corpus was tagged with DAs manually, and then a network was generated based on that part. After training the network had an average precision of 69%. The resulting network was applied to a bigger part of the corpus, and a new network was generated based on that bigger part. Training the new network resulted in 83% precision.

3.5. Multilayer perceptrons

Multilayer perceptrons (MLP) is one of the most frequently used neural network techniques, in general as well as in DA recognition. Neural networks are designed analogically to the human brain. They are capable of learning complex non-linear dependencies between input and output.

A neural network is composed of simple computational units (called neurons), which are organized into a network. The neurons have input and output connections, which are used to connect them. These connections are weighed, which means that the signal is multiplied by the connection weight – e.g. if the weight is 0, the connection doesn't let the signal through, or if it is 1, the signal is strong.

Neurons have output values which are computed based on their inputs, which are other neuron output values, weighed by the incoming connections. The network output complexity is achieved by combining the simple functions, that each neuron implements.

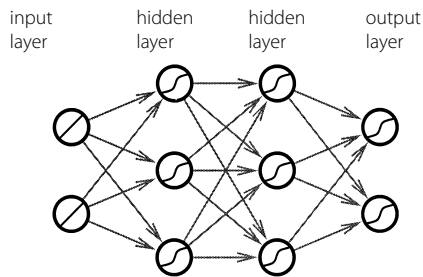


Figure 2. An example of a multilayer perceptron

MLP neurons are grouped into a linearly ordered set of layers, whereas the first layer is assigned the role of the input layer, and the last one – of the output layer. Connections are only allowed in the increasing direction. The function which neurons implement can be different, but must be monotonic and differentiable; it is usually chosen to be a sigmoid. Figure 2 shows a multilayer perceptron with 2 hidden layers.

The most common training algorithm for MLP is error back-propagation. Its description, as well as other algorithms can be found in (Haykin 1999).

Wright (1998) describes training an MLP with one hidden layer on the DCIEM corpus (Bard et al. 1996) (annotated with 12 different DA tags). Using only prosodic features extracted from the speech waves (a total of 54 binary features) she achieves 70% precision.

In Sanchis and Castro (2002) bag-of-words is used to form the MLP input. Before the method is applied all words are replaced with 3 types of categories: words in their basic forms (verbs in infinitive, nouns in singular, adjectives in singular and without gender, etc.), task-specific categories (departure and arrival cities, train types, etc.) and general categories (city names, days of week, months, numbers, etc.). Furthermore the infrequent words are ignored. In total the vocabulary size is 138 categories. The tag set includes 16 tags; the resulting MLP accuracy is 92%.

Levin et al. (2003) apply MLP to the same input as the naive Bayes classifier (binary grammatical feature bag-of-words), achieving the precision of 72% for English and 68% for German.

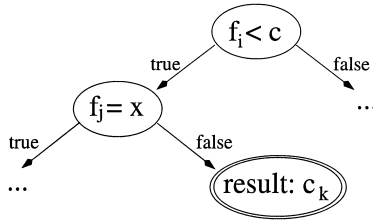


Figure 3. An example of a decision tree

3.6. Decision trees

Decision trees (or classification and regression trees, CART) have the advantage of being a white-box type method, meaning that it is easy for a human to interpret a trained model (relatively to the black-box models, like HMM or MLP).

Each node in a decision tree contains a set of conditions, describing which arc to continue traversing. Each decision is based on one of the parameters, describing the input sample (utterance in case of DA recognition). Each leaf is assigned a DA tag. Recognition is then performed by starting from the tree root and moving along the arcs according to the decisions. When a leaf is reached, the utterance is assigned the DA tag which corresponds to the leaf. A decision tree example can be found on figure 3.

Training decision trees stands for automatically composing them. A common algorithm is the Ross Quinlan ID3 algorithm. The main idea is to select into the tree root the parameter which causes the entropy of the divided subsets to be minimal; the entropy is defined as

$$H = - \sum_i p_i \log_2(p_i)$$

where p_i is the number of utterances, assigned the tag i according to the new tree. The algorithm therefore aims at composing as small a tree as possible.

Wright (1998) applies decision trees to the same data as MLP (binary features extracted from speech waves). Levin et al. (2003) also use the same data as with MLP and naive Bayes classifier (binary grammatical feature bag-of-words). Both methods result in 70% precision.

3.7. Transformation-based learning

Transformation-based learning (TBL) was introduced by Brill (1993). It is based on a set of rules, which are applied consecutively to the data, changing some tags into other ones. The rules are controlled by preset templates; the most common ones are of the type “if current tag is A, it is preceded by tag B and/or the word C is present in one of the preceding N utterances, change the current tag to D”.

Rules are composed in a supervised manner. Having a marked training corpus, all possible rules are generated from the templates, after which the rules are selected iteratively: the rule bringing the biggest improvement to the precision is selected on each iteration. The process is continued until one of the stopping criteria is met;

the most common is that no improvement is brought by applying any rule.

Since the total amount of all possible rules can be huge, it is computationally expensive to test each rule, especially since most of them bring precision degradation. One way to improve the situation is to use the Monte-Carlo pruning method. According to that, a fixed number of rules is selected by random; only these rules are later tested. Although this might exclude the very best rule from the selected set, it is highly probable that the set will contain a rule which will still bring a lot of improvement, even if not the maximum.

Samuel et al. (1998) describe applying transformation-based learning with the Monte-Carlo optimization to the VerbMobil corpus and DA tag set. Besides the usual TBL features (i.e. the neighboring utterances and DAs) they use the speaker, punctuation information, word and dialogue act cue² presence indicators. They achieve a precision of 75%.

Samuel et al. (1999) describe using dialogue act cues only, also applying TBL to the VerbMobil corpus. Although the focus of the paper is the automatic selection of the DA cues, they describe the DA recognition experiments, which result in a model with 72% precision.

3.8. Memory-based learning

Memory-based learning (MBL) is a relatively simple supervised classification technique. It exploits the idea that handling new data can be done by comparing it to earlier experience, instead of extracting rules from it and applying them. Therefore in MBL the training samples are stored into memory and the new unseen samples are compared to them.

There are several methods of comparing the stored and the unknown new samples. The simplest and the mostly used one is the k -nearest-neighbors classification. This method consists of defining a distance measure between the samples, and finding k stored samples which have the smallest distance to the new, unclassified sample. It is assumed that the “nearest” samples are similar to the new sample and therefore their classification can be extrapolated on it. The new sample is therefore assigned the class which dominates among the “neighbors”.

A common way to store the training samples is to replace them with n features, and to save the classification info. The simplest distance metric for this representation is the overlap metric. It simply sums the individual feature distances of the two samples. If a feature is scalar or vector, Euclidean distance is used, normalized with the maximum value for the given feature. In case of nominal features the individual distance is 1 if the values differ and 0 if they coincide. The combination of k -NN classification with the overlap metric is called the IB1 algorithm. Other metrics as well as a comprehensive description of memory-based learning can be found in the Tilburg Memory-Based Learner (TiMBL) reference guide (Daelemans et al. 2004).

Levin et al. (2003) apply TiMBL to the corpora of the NESPOLE machine translation project. They used the IB1 algorithm with 1 neighbor. The features are the same as in their MLP, decision tree and naive Bayes experiments (binary grammatical feature bag-of-words). They achieve a precision of 70%.

² Word sequences that are often met in the dialogue.

Lendvai et al. (2003) also use IB1. The features they use include prosodic features extracted from speech waves, bag-of-words vectors based on the speech recognition guesses about the pronounced words and context features (preceding tags etc.). The accuracy of the trained learner is 73.5%.

Fernandez et al. (2005) use features extracted from morphological and PoS information (such as the presence of a wh-word, repeated words in the utterance and its predecessor, etc.). They use the modified value difference metric (Daelemans et al. 2004), claiming that it performs better than the default (overlap) metric. The technique results in 87% precision.

3.9. Less frequently used techniques

Küssner (1997) applies a rule-based approach with the FLEX++ system to DA recognition. The used features include prosodic, semantic, syntactical as well as the DA guess of some unmentioned statistical DA tagger. Since the papers focus is not explicitly DA recognition, no experiment results are reported. In Lendvai et al. (2003) another rule-based approach is used, which is the RIPPER algorithm. The data representation is the same as in their experiment with memory-based learning. The achieved precision is 60%.

Serafin et al. (2003) and Serafin, Eugenio (2004) describe using latent semantic analysis (LSA) for DA recognition. The key-point idea is taking the word-document matrix (each component equals the number of times a word occurs in a document), compressing it with Singular Value Decomposition (SVD) and later comparing compressed new utterances to the matrix using Euclidean distance. The best reported result is 79%.

Tanaka and Yokoo (1999) describe alternative approach where DA tagging is integrated with discourse segmentation. The evaluation method is labeled bracket matching (Nagata 1994), the authors report the highest precision to be 75%.

4. DA recognition research in Estonia

In Estonia systematic research in the area of dialogue systems is done at the University of Tartu. The project incorporating the research aims at building a natural language interface to a database of transport timetables. The application of DA recognition is therefore the one that is common for dialogue systems: to recognize the intention behind the speaker's utterance and generate a proper response.

The peculiarity of DA recognition in this case lies in the high granularity of the DA taxonomy. The relative incidence of several DAs is low, which means both that their definition is very specific and that it's hard for a machine learning technique to not skip them in the learning process.

Fishel (2005) and Fišel, Kikas (2006) describe adapting several text classification preprocessing methods to DA recognition. WEBSOM-style preprocessing (Honkela et al. 1997), where the features describe the whole utterance, is tested with multilayer perceptrons. The same preprocessing method is adapted for processing the utterance one word at a time with simple recurrent networks: this way words are

represented with either random vectors (with each component ranging between -1 and 1) or random binary vectors (where each component is either -1 or 1). A different approach is coding the words with significance vectors (Wermter 2000). Finally, complete morphological description of each word as a single feature is tested.

All the experiments with neural networks were done with the upper layer of the DA taxonomy (see section 2). The best accuracy, showed by recurrent networks with significance vector coding, is 58%; the accuracy of other techniques ranges between 40% and 55%. The networks succeed at classifying DA tags represented in many utterances, but completely ignore the less frequent ones.

Another applied technique are decision trees; the experiments are described in Kikas (2005) and Fišel, Kikas (2006). Several types of input features are used in this case. First of them are the keyword-based features; intonation marks are also exploited. Secondly, the neighboring DAs are also included as a hint to the classifier. Finally the most frequent bi- and trigrams are used. The experiments are focused on the detailed level of the EDiC DA taxonomy; the described task is therefore to distinguish between 126 different tags.

The resulting accuracy is 45%. The decision trees appear to be capable of classifying the most frequent DA tags as well as neural networks; they also succeed in recognizing other DAs which are less frequent. However, DA tags represented in less than 1% of utterances are also ignored.

Currently a new algorithm of DA recognition is being developed. The general idea is to find regular expressions that best match utterances of definite DAs. These expressions can then be combined into a single decision tree, where each token of the expression serves as a tree node, which has two child-nodes: one for matching utterances and one for not matching ones. No experimental results have been published so far.

The latest experiments with the EDiC taxonomy detailed level exploit a Bayes-like classifier. The used features are the 2 preceding DA tags (equivalent to trigrams), the number of words k and the utterance words. In order to take into consideration the different number of words in every utterance, word probabilities are averaged with a geometrical mean:

$$\hat{c} = \operatorname{argmax}_c P(c|c_{n-1}, c_{n-2}) \cdot P(c|k) \cdot \left(\prod_{i=1}^k P(w_i|c) \right)^{\frac{1}{k}}$$

The resulting accuracy is 61%, which is noticeably close to the EDiC taxonomy kappa-statistic: whereas the statistic was measured on 45 dialogues and equalled 73%, the current results were measured with 10-fold cross-validation on 776 dialogues. However, in order to compare the two directly, the results should be calculated on the same dialogue sets.

As shown by this section, the statistical methods tested so far did not succeed to fully solve the taxonomy. On the other hand, fully describing it with manually composed rules seems not to be feasible given the complexity of the taxonomy considering the available resources. The author believes that further improving the DA recognition for the full EDiC taxonomy would require combining rule-based methods with statistical ones. For instance, specific errors of the Bayes classifier can be corrected by adding manually deduced rules.

5. Discussion

Several DA recognition techniques have been reviewed; a summary of the referred projects can be found in tables 1, 2 and 3 in appendix A.

The reported precisions range from 50% to 95%, whereas in several cases the precisions of applying the same technique to different corpora and/or tag sets differ greatly (e.g. the precision of naive Bayes is in different cases 50%, 66% and 82%). Therefore the technique choice doesn't seem to be the most important one.

It is possible to deduce filigree influences of exact combinations of recognition techniques and representation methods on the precision, but according to the author's opinion it would not make any sense. The hit-count precision does not take into consideration the differences of the DA taxonomies (size, definition clarity, design structure). Although it's a convenient and simple value, it can't be a basis of a conclusion that one technique or data representation is better than the other, unless the used DA taxonomy was the same.

In order to compare DA recognition techniques based on the results of experiments, applying them to different DA taxonomies, an evaluation metric has to be developed, which would take into consideration the DA tag set parameters, possibly also the size of training/testing corpora, training quality etc.

The conclusion is that since none of the techniques appear to be ideal or fully robust, and proper evaluation is absent, dialogue acts are an open research area, requiring a lot of work concerning recognition, representation and evaluation methods.

References

- Allen, James; Core, Mark 1997. Draft of DAMSL: Dialog Act Markup in Several Layers, Technical Report, Discourse Research Initiative.
- Alshawi, Hiyan 2003. Effective Utterance Classification with Unsupervised Phonotactic Models. – Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, Vol. 1. Edmonton, Canada, 1–7.
- Bard, Ellen; Sotillo, Cathy; Anderson, Anne; Taylor, Martin 1996. The DCIEM Map Task Corpus: Spontaneous Dialogue Under Sleep Deprivation and Drug Treatment. – Proceedings of ICSLP'96. Philadelphia, PA, USA, 1958–1961.
- Brill, Eric 1993. A Corpus-Based Approach to Language Learning. PhD thesis. Department of Computer and Information Science, University of Pennsylvania.
- Bunt, Harry C. 1994. Context and Dialogue Control. – THINK 3, 19–31.
- Daelemans, Walter; Zavrel, Jakob; van der Sloot, Ko; van den Bosch, Antal 2004. TiMBL: Tilburg Memory-Based Learner Reference Guide. Technical Report ILK 04-02. Tilburg University and University of Antwerp.
- Fernandez, Raquel; Ginzburg, Jonathan; Lappin, Shalom 2005. Using Machine Learning for Non-Sentential Utterance Classification. – Proceedings of the 6th SIGdial Workshop on Discourse and Dialogue. Lisbon, Portugal, 77–86.
- Fishel, Mark 2005. Dialogue Act Recognition in Estonian Dialogues Using Artificial Neural Networks. – Proceedings of the 2nd Baltic Conference on Human Language Technologies. Tallinn, 231–235.
- Fišel, Mark; Kikas, Taavet 2006. Dialoogiaktide automaatne tuvastamine. – Mare Koit, Renate Pajusalu, Haldur Õim (toim.). Keel ja arvuti. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 6. Tartu: Tartu Ülikooli Kirjastus, 233–245.

- Gerassimenko, Olga; Hennoste, Tiit; Koit, Mare; Rääbis, Andriela; Strandson, Krista; Valdisoo, Maret; Vutt, Evely 2004. Annotated Dialogue Corpus as a Language Resource: An Experience of Building the Estonian Dialogue Corpus. – Proceedings of the 1st Baltic Conference on Human Language Technologies. Riga, Latvia.
- Godfrey, John J.; Holliman, Edward; McDaniel, Jane 1992. SWITCHBOARD: Telephone Speech Corpus for Research and Development. – Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. San Francisco, CA, USA, 517–520,
- Grau, Sergio; Sanchis, Emilio; Castro, Maria Jose; Vilar, David 2004. Dialogue Act Classification Using a Bayesian Approach. – Proceedings of the 9th International Conference Speech and Computer, 495–499.
- Hagen, Eli; Popowich, Fred 2000. Flexible Speech Act Based Dialogue Management. – Proceedings of the 1st SIGdial Workshop on Discourse and Dialogue. Hong Kong, 131–140.
- Haykin, Simon 1999. Neural Networks: A Comprehensive Foundation. 2nd ed. Prentice-Hall, Inc.
- Honkela, Timo; Kaski, Samuel; Lagus, Krista; Kohonen, Teuvo 1997. WEBSOM – Self-Organizing Maps of Document Collections. – Proceedings of Workshop on Self-Organizing Maps. Helsinki, 310–315.
- Ivanovic, Edward 2005. Dialogue Act Tagging for Instant Messaging Chat Sessions. – Proceedings of the ACL Student Research Workshop. Ann Arbor, Michigan, 79–84.
- Jurafsky, Daniel; Shriberg, Elizabeth; Biasca, Debra 1997. Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual, draft 13, technical report 97-01. Institute of Cognitive Science, University of Colorado.
- Keizer, S.; op den Akker, R.; Nijholt, A. 2002. Dialogue Act Recognition with Bayesian Networks for Dutch Dialogues. – 3rd SIGdial Workshop on Discourse and Dialogue. Philadelphia, USA, 88–94.
- Kikas, Taavet 2005. Dialoogiaktide tuvastamine eestikeelsetes dialoogides otsustuspuude abil. Bakalaureusetöö. Käsikiri Tartu Ülikooli arvutiteaduse instituudis.
- Küssner, Uwe 1997. Applying DL in Automatic Dialogue Interpreting. – Proceedings of the International Workshop on Description Logics. Yvette, France, 54–58.
- Lee, Jjae-won; Kim, Gil Chang; Seo, Jungyun 1997. A Dialogue Analysis Model with Statistical Speech Act Processing for Dialogue Machine Translation. – Proceedings of the Spoken Language Translations EACL'97 Workshop. Budapest, Hungary, 10–15.
- Lendvai, Pirooska; van den Bosch, Antal; Krahmer, Emiel 2003. Machine Learning for Shallow Interpretation of User Utterances in Spoken Dialogue Systems. – Proceedings of the EACL-03 Workshop on Dialogue Systems: Interaction, Adaptation and Styles of Management. Budapest, Hungary, 69–78.
- Levin, Lori; Langley, Chad; Lavie, Alon; Gates, Donna; Wallace, Dorcas; Peterson, Kay 2003. Domain Specific Speech Acts for Spoken Language Translation. – Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue. Sapporo, Japan.
- Levin, Lori; Ries, Klaus; Thyme-Gobbel, Ann; Levie, Alon 1999. Tagging of Speech Acts and Dialogue Games in Spanish Call Home. – Proceedings of the ACL Workshop “Towards Standards and Tools for Discourse Tagging”. Somerset, NJ, USA, 42–47.
- Nagata, Masaaki 1994. A Stochastic Japanese Morphological Analyzer Using a Forward-DP and Backward-A* N-best search algorithm. – Proceedings of the 15th conference on Computational linguistics, Vol. 1. Kyoto, Japan, 201–207.
- Prasad, Rashmi; Walker, Marilyn 2002. Training a Dialogue Act Tagger for Human-Human and Human-Computer Travel Dialogues. – Proceedings of the 3rd SIGdial Workshop on Discourse and Dialogue. Philadelphia, Pennsylvania, 162–173.
- Rabiner, Lawrence; Juang, Bbiing-Hwang 1986. An Introduction to Hidden Markov Models. – ASSP Magazine, IEEE 3 (1), 4–16.

- Reithinger, Norbert; Engel, Ralf; Kipp, Michael; Klesen, Martin 1996. Predicting Dialogue Acts for a Speech-To-Speech Translation System. – Proceedings of the 4th International Conference on Spoken Language Processing, Vol. 2. Philadelphia, Pennsylvania, 654–657.
- Reithinger, Norbert.; Maier, Elisabeth 1995. Utilizing Statistical Dialogue Act Processing in VERBMOBIL. – Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics. Cambridge, Massachusetts, 116–121.
- Ries, Klaus 1999. HMM and Neural Network Based Speech Act Detection. – IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 1. Phoenix, Arizona, 497–500.
- Samuel, Ken; Carberry, Sandra; Vijay-Shanker, K. 1998. Dialogue Act Tagging with Transformation-Based Learning. – Proceedings of the 17th International Conference on Computational linguistics, Vol. 2. Montreal, Quebec, Canada, 1150–1156.
- Samuel, Ken; Carberry, Sandra; Vijay-Shanker, K. 1999. Automatically Selecting Useful Phrases for Dialogue Act Tagging. – Proceedings of the 4th Conference of the Pacific Association for Computational Linguistics. Waterloo, Ontario, Canada.
- Sanchis, Emilio; Castro, Maria Jose 2002. Dialogue Act Connectionist Detection in a Spoken Dialogue System. – Proceedings of the 2nd International Conference on Hybrid Intelligent Systems. Santiago de Chile, Chile, 644–651.
- Serafin, Riccardo; Eugenio, Barbara D. 2004. FLSA: Extending Latent Semantic Analysis with Features for Dialogue Act Classification. – Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics. Barcelona, Spain, 692–699.
- Serafin, Riccardo; Eugenio, Barbara. D.; Glass, Michael 2003. Latent Semantic Analysis for Dialogue Act Classification. – Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, Vol. 2. Edmonton, Canada, 94–96.
- Siegel, Sidney; Castellan, John N. 1988. Nonparametric Statistics for the Behavioral Sciences. 2nd ed. New York: McGraw-Hill.
- Tanaka, Hideki; Yokoo, Akio 1999. An Efficient Statistical Speech Act Type Tagging System for Speech Translation Systems. – Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics. College Park, MD, USA, 381–388.
- Traum, David R. 1999. Speech Acts for Dialogue Agents. – Anand Rao, Michael Wooldridge (Ed.). Foundations and Theories of Rational Agents. Kluwer Academic Publishers, 173–206.
- Wermter, Stefan 2000. Neural Network Agents for Learning Semantic Text Classification. – Information Retrieval 3 (2), 87–103.
- Wright, Helen 1998. Automatic Utterance Type Detection Using Suprasegmental Features. – Proceedings of the 5th International Conference on Spoken Language Processing, Vol. 4. Sydney, Australia, 1403.
- Wright, Helen; Poesio, Massimo; Isard, Stephen 1999. Using High Level Dialogue Information for Dialogue Act Recognition Using Prosodic Features. – Proceedings of an ESCA Tutorial and Research Workshop on Dialogue and Prosody. Eindhoven, The Netherlands, 139–143.

Appendix: Summary of DA Recognition Projects

This appendix includes the tables summarizing the DA recognition methods, data representation methods and results of projects, referred to in the DA recognition techniques section.

Table 1. DA recognition in dialogue systems

Project reference	DA taxonomy	Recognition technique	Data representation	Best result
(Keizer et al. 2002)	DAMSL	Bayes	linguistic, punctuation	44%
(Ivanovic 2005)	SWBD-DAMSL	Bayes	BoW, DA n-grams	82%
(Samuel et al.1999)	VM	TBL	DA cues	72%
(Samuel et al.1998)	VM	TBL with Monte Carlo	speaker, punctuation, DA cues	75%
(Ries 1999)	CallHome	HMM+MLP	prosodic and word features	76%
(Serafin and Eugenio, 2004)	CallHome	FLSA	W*D matrix	79%
(Prasad and Walker, 2002)	DARPA	RIPPER	context and word features	99%
(Fishel and Kikas, 2006)	EDIC	Bayes	words, n-grams	61%
		CART	keywords, n-grams	45%
(Lendvai et al. 2003)	self-defined	TIMBL	prosodic, BoW	74%
		RIPPER	(same)	60%
(Fernandez et al. 2005)	self-defined	SLIPPER	linguistic and antecedent	87%
		TIMBL	(same)	87%
		MaxEnt	(same)	87%
(Sanchis and Castro, 2002)	self-defined	MLP	reduced lexicon BoW	92%

Table 2. DA recognition in machine translated dialogue systems

Project reference	DA taxonomy	Recognition technique	Data representation	Best result
(Küssner 1997)	VM	FLEX++	prosody, syntax, semantics	
(Reithinger and Maier 1995)	VM	n-grams	DA n-grams	81%
(Reithinger et al. 1996)	VM	n-grams	DAs, grammar	76%
(Levin et al. 2003)	NESPOLE	TIMBL	grammatical	70%
		CART	(same)	70%
		MLP	(same)	70%
		Bayes	(same)	50%
(Lee et al. 1997)	self-defined	n-grams	syntax patterns, DA n-grams	79%
(Tanaka and Yokoo 1999)	self-defined	self-defined probabilistic method	morpheme sequences	75%

Table 3. DA recognition in speech recognition

Project reference	DA taxonomy	Recognition technique	Data representation	Best result
(Wright et al. 1999)	DCIEM	CART	prosody features	69%
(Wright 1998)	DCIEM	CART	intonation events	71%
		HMM	prosodic features	72%
		MLP	prosodic features	70%
(Grau et al., 2004)	SWBD-DAMSL	Bayes	n-grams	66%
(Alshawi 2003)	AT&T	HMM	phoneme sequences	95%

Mark Fišel (Tartu Ülikool) on erialalt informaatik. Uurimisvaldkonnad: masinõpe, masintõlge, dialoogisüsteemid.
fishel@ut.ee

MASINÕPPE TEHNIKAD DIALOOGIAKTIDE TUVASTAMISES

Mark Fišel

Tartu Ülikool

Artiklis käsitletakse dialoogiakte, nende rakendusi ja automaatse tuvastamise tehnikaid – nii Eestis kui ka mujal maailmas. Kirjeldatakse kolme põhilist rakendust: dialoogisüsteemides selleks, et tuvastada rääkija kavatsusi, masintõlkega dialoogisüsteemides selleks, et lahendada mitmesust võimalikkude tõlkevariantide vahel, ning kõnetuvastuses selleks, et vähendada sõnade tuvastuse vigade arvu.

Kirjeldatakse mitmeid tuvastamistehnikaid. Kõige sagedamini kasutatavad neist on Markovi peitmuldel, mitmekihiline tajur, naiivne Bayesi klassifitseerija, Bayesi võrgud, otsustuspuud, mälu põhine ning transformeerimispõhine õppimine. Iga tehnikat kirjeldatakse üldisel tasemel: töötamise printsiipe ning seda, kuidas toimub treenimine. Artikkel sisaldab mitmeid näiteid selle kohta, kuidas neid tehnikaid saab rakendada dialoogiaktide tuvastamiseks. Eraldi pööratakse tähelepanu dialoogiaktide tuvastamise uurimisele Eestis.

Autor jõuab järeldusele, et tuvastamistehnika valik ei olegi dialoogiaktide puhul kõige olulisem. Mõjuvaim faktor on hoopis dialoogiaktide süsteemi karakteristikud (suurus, aktide definitsioonide selgus jms); samas aga enim kasutatavad hindamiskriteeriumid ei võta neid karakteristikuid arvesse. Seega on võimatu võrrelda tuvastamise meetodeid selliste eksperimentide alusel, kus neid rakendatakse erinevatele dialoogiaktide süsteemidele.

Dialoogiaktid on avatud uurimisala, kus on võimalust ja vajadust arendada uusi tuvastamistehnikaid ning hindamise kriteeriume.

Võtmesõnad: vestlusanalüüs, arvutuslingvistika, dialoogiakt, masinõpe, Bayesi klassifitseerija, Markovi peitmuldel, neurovõrk, otsustuspuu

KLIENDI SOOVIDE AUTOMAATNE TUVASTAMINE EESTIKEELSETES INFODIALOGIDES

**Olga Gerassimenko, Tiit Hennoste, Riina Kasterpalu,
Mare Koit, Andriela Rääbis, Krista Strandson,
Maret Valdisoo, Evely Vutt**

Ülevaade. Käsiloleva uurimuse eesmärk on leida keelelisi märguandeid, mida arvuti saaks kasutada dialoogiaktide tuvastamisel. Artiklis piirdume kliendi soovidega kui dialoogi modelleerimisel kesksete aktidega. Analüüsimiseks valisime 144 institutsionaalset telefonikõnet Tartu Ülikooli Eesti dialoogikorpusest. Need jagunevad nelja situatsioonirühma: infotelefon, reisibüroo, polikliiniku registratuur, takso tellimine. Analüüsitud dialoogides esitavad kliendid kaht liiki soove: infosoovid (helistaja vajab kindlat infot, nt telefoninumbrit) ja ametniku tegevust taotlevad soovid (reserveerida aeg arsti vastuvõtule, saata takso). Analüüs näitab, et kliendi soovides esineb piiratud hulk verbivorme ja kindlad verbid on süstemaatiliselt seotud kindlate kontekstitingimustega. Dialoogide struktuuri kui dialoogiaktide järgnevuse esitame Markovi ahelana, mis aitab arvutil kasutaja aktidele adekvaatselt reageerida.*

Võtmesõnad: dialoogikorpus, dialoogiaktide tuvastamine, vestlusanalüüs, eesti keel

1. Sissejuhatus

Suhtlemisel loovad ja tõlgendavad osalejad dialoogiakte. Selleks et modelleerida suhtlust arvutil, tuleb analüüsida, kuidas dialoogiakte keeleliselt väljendatakse ja milliseid keelelisi tunnuseid saaks kasutada arvuti kui vestluspartner lausungites esitatud dialoogiaktide tõlgendamiseks.

Tuntakse kahte dialoogiaktide automaatse interpreteerimise arvutimudelite klassi (Jurafsky, Martin 2000).

* Tööd on osaliselt toetanud Eesti Teadusfond (grant nr 5685).

Esimene klass on märguandepõhine ehk tõenäosuslik. Siin eeldatakse, et kuulaja kasutab erinevaid leksikaalseid, kollokatsioonilisi, süntaktilisi, prosoodilisi või vestluse struktuuriga antud märguandeid, mis aitavad tal otsustada, kuidas kõneleja akti tõlgendada. Selle lähenemisviisi aluseks on nn mikrogrammatikate idee (Goodwin 1996), mille kohaselt saab iga dialoogiaktiga siduda omaette väikese grammatika. Selliseid märguandeid on sagedasemate dialoogiaktide puhul laialdaselt kirjeldatud (Jurafsky jt 1998). Märguandepõhised mudelid lähenevad akti tõlgendamisele kui klassifitseerimisülesandele ja lahendavad seda statistiliste klassifitseerijate treenimisega korpustel, kus on märgendatud dialoogiaktid.

Teine mudelite klass kasutab järeldamist. Kui analüüsida nt lauset *Kas te oskate mulle öelda X?*, siis selle otsene tõlgendus oleks küsimus: *Kas te olete võimeline mulle üttelema X?* Kõneleja tegeliku soovi *Ütelge mulle X* tuletab kuulaja otsese küsimuse töötlemise tulemusel. Järeldamismudelid põhinevad arvamusloogikal ja rakendavad loogilist tuletust, et arutleda kõneleja kavatsuste üle.

Selles artiklis keskendume märguandepõhisele mudelile. Analüüsime eestikeelsetes ametikõnedes esinevad soove, kus kõneleja eesmärk on kas saada kuulajalt teatavat infot või saavutada, et kuulaja teeks teatava tegevuse. Seega piirdume siin võimalike soovide ühe alamhulgaga.

Teine piirang on see, et me vaatleme ainult kliendi soove. See on arusaadav, sest meie kaugem eesmärk on luua selline dialoogsüsteem, mis suudaks täita infoametniku rolli ja seega interpreteerida kasutaja akte ning nendele adekvaatselt reageerida.

Meie otseseks eesmärgiks on leida kliendi soovide tuvastamiseks sobivad leksikaalsed ja süntaktilised märguanded ning nende kasutamise tingimused. Käesolevas artiklis keskendume kliendi soovides predikaadi funktsioonis esinevatele verbidetele.

2. Mis on soov?

Maailmas on juba mitmeid aastaid läbi viidud erinevaid projekte korpuste märgendamiseks ja ühiste standardite loomiseks. On välja töötatud mitmeid dialoogiaktide märgenduskeeme, millest tuntuim on DAMSL (Dialog Act Markup in Several Layers, vt Allen jt 1997).

Eesti dialoogikorpuse märgendamiseks oleme Tartu Ülikoolis välja töötanud oma dialoogiaktide tüpologia, mille aluseks on vestlusanalüüsi põhimõtted (Hennoste, Rääbis 2004). See erineb mõnevõrra märgenduskeemist DAMSL, sest ta lähtub eestikeelsete suuliste dialoogide analüüsist ja vastab täpselt meie eesmärgile – uurida eestikeelset suulist suhtlust. Kui skeemis DAMSL esineb akt *request*, mis võib tähendada nii küsimust kui ka infosoovi (s.t direktiivi), siis meie tüpologia teeb vahet küsimuste ja direktiivide vahel. Küsimustel on eesti keeles selged formaalsed tunnused (küsisõnad, intonatsioon, sõnajärg), mida saab kasutada märguannetena nende automaatsel tuvastamisel. Direktiividel aga selliseid tunnuseid ei ole. Seega meie tüpologia kohaselt on *Kas te oskate mulle öelda X?* küsimus, aga *Ütelge mulle X* või *Tahan teada X* direktiiv (soov).

Meie tüpoloogias on üldse kolme liiki direktiive: soovid, pakkumised ja ettepanekud. Soov väljendab kõneleja kavatsust saada teatavat infot või kutsuda esile adressaadi teatav tegevus (H – helistaja, klient, V – vastaja, ametnik)¹:

¹ Näidetes kasutame vestlusanalüüsi transkriptsiooni. Vt ka Hennoste, Rääbis 2004, Strandson (käesolevas kogumikus), Rääbis (käesolevas kogumikus), <http://math.ut.ee/~koit/Dialoog/TranscriptionMarks.rtf> (26.03.2007).

H: → tere, hh' ütle mulle palun Parkeri kabinet number.

V: mt üheksakend neli neli null seitse.

Ettepanekud ja pakkumised (mingi tegevuse tegemiseks) on teineteisega sarnased, mõlemal juhul on sobivad reaktsioonid nõustumine või keeldumine. Seejuures on ettepaneku puhul tegevuse tegijaks kuulaja (*helistage hiljem*), pakkumise puhul aga kõneleja ise (*ma saadan teile programmi*) (Hennoste, Rääbis 2004).

Soovide esitamiseks saab kasutada erinevaid verbe, sõltuvalt sellest, missugust reaktsiooni oodatakse kuulajalt (*tõuse püsti, võta leiba* jne). Käesolevas artiklis, nagu eespool öeldud, piirdume kliendi soovidega, kus

- 1) klient vajab infot (nt telefoninumbrit) või
- 2) klient ootab, et ametnik teeks teatava tegevuse (nt saadaks talle takso).

Seejuures kaasneb tegevuse sooritamisega alati ka info andmine: ametnik informeerib klienti, et ta on tegevuse lõpetanud (*jaa, takso tuleb teile*) või et ta ei ole suuteline seda tegema.

3. Empiiriline materjal ja meetod

Meie analüüs põhineb Eesti dialoogikorpusel EDiK.² Korpus sisaldab umbes 800 suulist dialoogi, sh üle 700 telefonikõne. Korpuses on märgendatud dialoogiaktid. Iga dialoogi on teineteisest sõltumatult märgendanud kaks isikut (kokku on märgendajaid olnud 10) ja seejärel on märgendused ühtlustatud.

Käesoleva artikli jaoks valisime korpusest juhuslikult 144 telefonikõnet (kokku 19 938 tekstisõna). Kõnedes on esindatud neli suhtlussituatsiooni: 1) infotelefon (küsitakse telefoninumbreid, aadresse jms), 2) takso tellimine, 3) kõned reisibüroosse ja 4) kõned polikliinikusse. Kokku leidus dialoogides 130 kliendi soovi (tabel 1), millest üks jäi edasisest analüüsist välja, sest ta oli esitatud vene keeles.

Tabel 1. Kasutatud korpus

Dialoogi tüüp	Dialooge	Teksti-sõnu	Kliendi soove
Infotelefon	60	4384	56
Reisibüroo	36	12104	33
Polikliinik	26	2422	22
Takso	22	1028	19
Kokku	144	19938	130

Analüüsimiseks kasutame kahte meetodit. Esmalt püüame kliendi soovides leida tüüpilisi märguandeid lihtsa statistika abil. Eesti dialoogikorpuse analüüsimiseks on välja töötatud tarkvara, mis muuhulgas võimaldab hõlpsasti loendada mitmesuguseid dialoogides esinevaid nähtusi: sõnu või sõnajärgendeid, transkriptsiooni-elemente, dialoogiakte (Treumuth 2005).

Teiseks, eesti keeles kasutatakse soovide väljendamiseks erinevaid verbivorme (tabel 2). Meie eesmärk on leida, missugustel tingimustel erinevaid variante kasutatakse. Selleks viime läbi sõnavormide kasutuse mikroanalüüsi vestlusanalüüsi

² <http://math.ut.ee/~koit/Dialoog/EDiK.html> (12.09.2006)

meetodil (vestlusanalüüsi kohta vt nt Hutchby, Wooffitt 1998). Infotelefoni- ja reisibürookõnedes ootavad kliendi soovid vastuseks info andmist (nt *ma paluks filo 'soofiateaduskonna`dekanaadi`numbrit.*) ning polikliiniku- ja taksokõnedes ametniku tegevust (*ma palun`taksot`Ringtee`kuuskend kaheksa`bee.*). Kui klient polikliinikusse helistades või taksot tellides vajab infot, siis vormistab ta oma kavatsuse alati küsimusena, mitte aga soovina (*kas teil`nihukest taksot`ka on kuhu`viis inimest peale mahuks.*).

4. Verbivormide analüüs

Kliendi soovid võib jaotada kahte klassi: sellised, mis sisaldavad verbi (predikaadi funktsioonis), ja sellised, mis verbi ei sisalda (tabel 2). Kliendi soovides esineb üksnes piiratud kogus verbe. Kõige sagedasemad on *soovima* (s.h laiendiga variandid *teada soovima, küsida soovima*) – 35 juhtu ning *paluma* – 24 juhtu (kokku 46% soovidest). Lisaks nendele esinevad veel verbid *tahtma* (ja selle variandid) – 15 korda ning *ütleva* – 8 korda, kokku 23 korda (18%). Nimetatud neli verbi moodustavad 64% kõigist juhtudest. Veel neli verbi esinevad 4-6 korda (kokku 19 korda ehk 15%). Ülejäänud verbe on soovides kasutatud 1-2 korda.

Eesti keeles on võimalik soovi väljendada nii käskiva kõneviisiga (*istu/istuge*), kindla kõneviisiga (*ma tahan/me tahame*) kui ka tingiva kõneviisiga (*ma tahaks/me tahaksime*) (vt EKG II: 174, 178). Vormide arv aga, milles eri verbid kliendi soovides predikaadi funktsioonis esinevad, on piiratud. Vaatame neid lähemalt.

Verbi **soovima** tähendus sisaldab soovi, palvet. Lisaks sellele sisaldab ta ka viisakust (EKSS). Meie korpuses esinevates soovides kasutatakse seda verbi ainsuse esimeses isikus ning kas tingivas või kindlas kõneviisis (*ma sooviks/soovin*). Küsimustes võib seda verbi kasutada teises isikus. Käskiva kõneviisi kasutamine on võimalik käsu korral – et partner ise sooviks midagi (*soovi X*), mida infodialogides mõistagi ei esine.

Tabel 2. Kliendi soovide arv korpuses (erinevate verbivormidega ja ilma verbita)

Verb	Kõneviis			Kokku
	kindel	tingiv	käskiv	
soovima	2	24		26
teada soovima		8		8
küsida soovima		1		1
tahtma	3	7		10
teada tahtma		4		4
küsida tahtma		1		1
paluma	12	12		24
ütleva			8	8
võtma		6		6
vaja olema		5		5
huvitama	1	3		4
huvitatud olema	2			2

Verb	Kõneviis			Kokku
	kindel	tingiv	käskiv	
panema	1	1	2	4
andma			2	2
küsima	1	1		2
oskama	2			2
saama	1			1
olema pakkuda	1			1
vaatama			1	1
vabandama			1	1
Verbiga kokku	26	73	14	113
Verbita				16
Kokku				129

Ühendit *teada soovima* oleme automaatse tuvastuse vajadustest lähtudes käsitlenud verbi *soovima* variandina ning samamoodi toimunud ka teiste infiniitsete verbilaienditega konstruktsioonide puhul. Siin on lisatud verb *teada*, mis osutab soovile saada teadmisi. *Teada soovima* esineb soovides niisamuti nagu *soovima* (meie korpuses ainult tingivas kõneviisis): *hh e sooviksin teada: natuke `Austria suusareisi kohta*.

Analoogiliselt on verbi *soovima* variant *küsida soovima*. Meie korpuses on selle kohta üksainus näide ja tingivas kõneviisis (*sooviksin küsida Petseri `kloostri kohta*).

Verb ***tahtma*** on verbi *soovima* sünonüüm, kuid tal on tugevam tähendus: pingsalt, vältimatult vajama (EKSS). Soovides esineb *tahtma* ainult tingivas kõneviisis. Kindel kõneviis väljendaks käsku, aga see on institutsionaalsetes dialoogides ebatavaline. Tingiv kõneviis pehmendab verbiga *tahtma* väljendatud soovi (EKG II: 35, Quirk jt 1972).

Teada tahtma on *tahtma* variant ja teda kasutatakse samal viisil. Meie korpuse näited on ainult tingivas kõneviisis: *ma=taks] teada `Ee Kaubamaja ilusa `longi telefoninumbrit*.

Küsida tahtma on samuti *tahtma* variant ja tedagi kasutatakse samamoodi. Meie korpuse ainus näide on tingivas kõneviisis: *tahaks=küsida seda: Alek `sandri auto=ja=kaatrikeskuse `telefoninumbrit*.

Verb ***paluma*** on eesti keeles tähelepanuväärne. Kõigepealt kasutatakse teda ainsuse esimeses isikus viisakussõnana *palun*, teiseks võib teda kasutada tavalise verbina (EKSS). Tema tähendus on sarnane verbi *soovima* tähendusega, kuid *paluma* on pehmem.

Soovides kasutatakse verbi *paluma* ainult kindla või tingiva kõneviisi esimeses isikus (*palun/paluksin*), teise isiku korral (*sa palud/paluksid*) ei oleks tegu sooviga. Käskivas kõneviisis saab seda verbi kasutada ainult kuulajale käsu edastamiseks, et too endale midagi paluks (*palu nüüd...*).

Verbi *paluma* kasutuses on selged erinevused olenevalt sellest, kas tegu on telefoninumbri või muu info soovimisega. Klient küsib telefoninumbreid enamasti nii, et *paluma* on lausungis ainus verb (*palun haigla*), harva lisab ta sõna *number* (*palun haigla number*). Selliseid info tüüpi osutusega lausungeid ei kasutata aadresside soovimisel või takso tellimisel (*palun taksot*). Infotelefonile helistades

tahavad kliendid sageli teada just telefoninumbreid, seetõttu on ametniku jaoks loomulik, et klient oma soovi vormistamisel info tüüpi ei väljenda.

Verbide *soovima* ja *paluma* vahel on erinevus: kui telefoninumbrit pärides kasutatakse verbi *soovima*, siis lisatakse alati nimisõna *number*. Seega võib teha järelduse, et verbi *palun* kasutatakse sellistes situatsioonides enamasti vormelina.

Verbi **ütleva** semantika käskivas kõneviisis (*ütlev/ütelge*) kirjutab ette tema tõlgendamise infosoovina, mitte aga tegevussoovina. Tingiva või kindla kõneviisi ainsuse esimeses isikus ei saa seda verbi soovi väljendamiseks kasutada (*ma ütles/ma ütlen*).

Ranget käsku võib väljendada kindla kõneviisi teise isiku abil (*sa ütled, kuhu sa asjad peitsid*), kuid meie korpuses sellised näited puuduvad.

Tingiva kõneviisi teist isikut võib soovides kasutada üksnes koos küsiva partikliga *kas* (*kas te ütlete...*), kuid selliseid lausungeid tõlgendame me küsimustena (mis aga pole käesoleva artikli teema).

Võtma on verb, mis semantiliselt pole seotud sooviga. *Ma võtaks/ma võtan* näitab, et kõneleja soovib midagi saada. Samal ajal näitab verbi tähendus seda, et soovitud saab võtta teatud kohast ja et partner on kohustatud selle kõneleajale andma. Seetõttu esineb *võtma* infosoovides ainult tingivas kõneviisis, et soovi pehmenada. Seda verbi ei saa infosoovides kasutada käskiva või tingiva kõneviisi teises isikus (*võta/võtke, sa võtad/võtaks*). Käskiva kõneviisiga soov ei oleks infosoov, vaid väljendaks käsku adressaadile teha midagi (nt *võta raha*). Kindlas kõneviisis saab seda verbi kasutada ainult ühendverbi koosseisus, väljendamaks ranget käsku, mitte aga (info)soovi (*sa võtad selle üles!*).

Adverbi ja verbi kombinatsioon **vaja olema** osutab vajadusele millegi või kellegi järele (EKSS). Soovides saab seda verbi kasutada ainult kindla või tingiva kõneviisi esimeses isikus (*hh mul oleks vaja sõita aprillis Budapesti*). Meie korpuses on kõik näited tingivas kõneviisis. Siin ei kasutata kindlat kõneviisi, sest see mõjuks kui nõudmine, kohustamine (mul on vaja, seepärast sa pead andma).

Huvitama on verb, mis ei näita kõneleja konkreetset soovi, vaid ainult huvi millegi vastu (EKSS). Selles funktsioonis saab teda kasutada kindla või tingiva kõneviisi esimeses isikus (*min:d uvitavad turismireisid Lõuna=Eestis*).

Olema huvitatud on verbi *huvitama* variant ja teda kasutatakse samal viisil.

Kumbagi vormi ei kasutata konkreetsetes infosoovides, vaid ainult sissejuhatavates lausungites, mis alustavad teemat, näitamaks, millest jutt tuleb (*hh olen uvitatud reisidest Skandi'naaviamaadesse=h.; ma olen huvitatud=õ selle: ses informatsioonist=et > ku palju võiks < maksta sõit Inglismaale*).

Verbi **panema** tähendus on eesti keeles lai ja hägune (EKSS). Kõigepealt võib ta tähendada füüsilist tegevust (*pane maha*). Infosoovides on *panema* ühendverbi *kinni panema* (siin tähenduses 'reserveerima') osa. Teda kasutatakse sellistes dialoogides, kus kliendil on vaja reserveerida arsti vastuvõtuaega (*aga paneme :: 'edasi ka 'ravi juba 'kinni*). Seda ühendverbi kasutatakse mõnikord (sõltuvalt kontekstist) ilma adverbina. Meie infosoovides kasutatakse *panema* käskiva kõneviisi teises isikus (*pane/pange*), tingivas kõneviisis *paneks* ja kindla kõneviisi mitmuse esimeses isikus *paneme*. Kindla kõneviisi ainsuse teise isiku kasutamine oleks tõlgendatav kui range käsk (*sa pane mulle selle numbri!*). Meie andmetes selliseid näiteid ei esine.

Ülejäänud verbid esinevad soovides igauks ainult 1-2 korda.

Verbil **andma** on eesti keeles mitmesuguseid üldisi tähendusi (EKSS). Soovides saab teda kasutada ainult käskivas kõneviisis (*andke `laua oma `ka*). Küsimustes võib seda verbi kasutada tingiva kõneviisi teises isikus (*kas sa annaks*).

Oskama ja **saama** väljendavad mõlemad muu hulgas subjekti võimelisust tegevust sooritada (EKSS). Infosoovides kasutatakse mõlemat modaalses funktsioonis konstruktsioonides *oskama öelda, saama öelda, oskama soovitada*. Kõik sellised konstruktsioonid sisaldavad lisaks ka sõna *äkki* või *võib-olla*, mis pehmen-davad soovi ja annavad ametnikule võimaluse infot mitte anda või mitte omada (*et võibolla oskate midagi `soovitada. // äkki te oskate `aadressi öelda*).

Vaatama ja **andma** esinevad käskivas kõneviisis. Kindla kõneviisi esimeses isikus väljendavad nad kõneleja enda tegevust (*ma vaatan/annan*). Infotelefo-nile helistades väljendab *vaadake* käsku ametnikule otsida midagi oma andme-baasist.

Olema pakkuda eeldab, et partner pakuks midagi. Sellist konstruktsiooni ei saa kasutada konkreetse info soovimiseks, vaid ainult üldises soovis, kus kõne-leja ei tea kindlalt, mida vastuseks oodata (*aga võibolla teil on: `teisi jaanuaris variante ka `pakkuda*).

Küsima ja **vabandama** ei ole meie infosoovides predikaadi funktsioonis kasutatust leidnud (vt ptk 5).

5. Eri verbivormide kasutamise tingimused

Eri verbe kasutatakse dialoogides erinevatel tingimustel. Tüüpilisel infosoovil on meie korpuses järgmised tunnused.

- 1) Suhtlus leiab aset võõraste vahel. Meie korpuses on vähe tuttavatevahelisi kõnesid (nt arst räägib polikliiniku registratuuri töötajaga). See mõjutab verbivormide esinemust.
- 2) Soovi esitab klient suhtluse algul. Enamasti järgneb soov vahetult rituaalsele sissejuhatusele (kutsungi vastuvõtmine, tutvustus, tervitus):

V: info´ telefon=Kersti=tere

H: → tere ma palun: ´juuksuriäri Mag´ noolia.

Vahel kasutatakse sissejuhatava osa lõpus ja soovi esitamise eel veel ühte lausungit, mis näitab, et järgneb soov. Mõnikord esitab klient küsimuse, et otsustada, kas planeeritud soovi on võimalik väljendada (kas ametnik suudab anda oodatavat infot, nt eraisikute telefoninumbreid jms):

H: .hh tere. .hh ee (.) ´kas te tegelete ´lennupiletitega ´ka.(.)

V: ikka.

H: → .hh ee mul oleks vaja sõita aprillis ´Budapesti.

Lisaks esmastele soovidele esineb ka soove, mis tekivad suhtluse käigus.

- 3) Tüüpilise soovi kolmas tingimus on, et klient on selle valmis mõelnud enne helistamist. Seega on soov helistamise põhjuseks.

Tüüpilised soovid saab jaotada kahte rühma. Esiteks, konkreetset soovid, kus kõneleja eesmärk on kas saada teatavat täpset infot (telefoninumber) või algatada

ametniku teatav tegevus (takso saatmine). Teiseks, üldised soovid, mis üksnes juhatavad sisse teema. Sellisele soovile järgneb küsimus-vastus-alamdialoog, kus kogutakse infot.

Vaatame erinevate verbide kasutamise tingimusi kliendi soovides (vt ka tabelid 3–6).

Verbivormi **sooviks** kasutatakse erinevates dialoogitüüpides nii konkreetsete kui ka üldiste tüüpiliste soovide väljendamiseks.

Kindel kõneviis *ma soovin* esineb meie korpuses kahel korral ja mõlemal juhul alustab pooleli jäävat lausungit. Seetõttu võib seda tõlgendada kui viga, mille kõneleja ise parandab (*ma soovin=e mingi `Tar `Furgo on siin kuskil `Ringteel mis laenutab fur`goone.*)

sooviks teada ja *sooviks küsida* kasutatakse ainult üldiste soovide väljendamiseks (*.hh=em sooviksin: teada natuke sellest `Soome reisist aasta `vahetusel.*). Sellised soovid esinevad reisibüroodialoogides.

Tabel 3. Verbide esinemine taksodialoogides

Verb	Kõneviis			Kokku
	Kindel	Tingiv	Käskiv	
soovima		8		8
tahtma	1			1
paluma	6	2		8
Verbita				3
Kokku	7	10	0	20

Tahaks kasutatakse kõigis dialoogitüüpides ja enamasti konkreetsete soovide väljendamiseks, mis on helistamise tegelikuks põhjuseks (meie andmetes on sellest ainult kaks erandit).

Vormi *ma tahan* kasutatakse ainult konkreetsete soovide väljendamiseks, mis on helistamise tegelikud põhjused, kuid üksnes tuttavate suhtluses (*tead mina tahan ka ühte kaarti ja ühte arvet tänase kuupäevaga.*)

Verbiühendeid *tahaks teada* ja *tahaks küsida* kasutatakse erineval viisil. Kõik kasutused esinevad võõrastevahelises suhtluses: *ma=taks teada `Ee Kaubamaja ilusa `longi telefoninumbrit.*

Nii **paluks** kui ka **palun** kasutatakse peaaegu eranditult takso tellimiseks või infotelefonile helistamisel (reisibüroo- ja polikliinikudialoogides on kummaski üksainus näide). Kõik kasutused väljendavad konkreetseid soovide.

Verbivormi **öelge** kasutatakse peaaegu eranditult tüüpilistes soovides ning ta väljendab infotelefonidialoogides konkreetseid soovide. Ainult ühel korral on seda verbi kasutatud hilisemas soovis eelmise soovi täpsustamiseks (*vot `öelge=mulle=säl: mingit `info `numbrit või .hh.*)

Võtaks esineb ebatüüpilistes konkreetsetes soovides. Meie andmetes kasutatakse seda vormi kahes rollis. Esiteks, kokkuvõtva otsuse tegemiseks, kui kõneleja teab, et soovitud info leitakse (*.hh ee no ma igaks juhaks võtaks selle `interneti aadressi.*). Teiseks võib *võtaks* osutada, et kõneleja vajab rohkem infot (*võtaks `veel mõne.*)

Oleks vaja esineb erinevates dialoogitüüpides ainult tüüpilistes soovides ja väljendab kas konkreetset või üldist soovi (*oleks vaja* `Pärnust `Arko´rok nõude `laenus telefoninumbrit; *hh mul oleks vaja* `sõita aprillis `Budapesti).

Tabel 4. Verbide esinemine registratuuridialoogides

Verb	Kõneviis			Kokku
	Kindel	Tingiv	Käskiv	
soovima		6		6
tahtma	2	4		6
paluma		1		1
vaja olema		1		1
panema	1	1	2	4
Verbita				3
Kokku	3	13	2	21

Tabel 5. Verbide esinemine infotelefonidialoogides

Verb	Kõneviis			Kokku
	Kindel	Tingiv	Käskiv	
soovima	2	5		7
tahtma (sh teada -, küsida -)		5		5
paluma	6	8		14
ütleva			8	8
võtma		5		5
vaja olema		1		1
andma			2	2
küsima	1			1
oskama	1			1
saama	1			1
vaatama			1	1
vabandama			1	1
Verbita				7
Kokku	11	24	12	54

Mind huvitavad / mind huvitaks / ma olen huvitatud esinevad ainult reisibüroodialoogides ja väljendavad peaaegu alati suhtluse alguses üldisi soove (*min:d uvitavad turismireisid* `Lõuna=Eestis). Suhtluse käigus tekkinud soovi kohta on üksainus näide, mis samuti väljendab üldist soovi (*hh vot neid m-m* `vaatamisväärsusi ja `neid nagu `vitakski).

Panema (kinni) kasutatakse ainult registratuurikõnedes ja ainult konkreetsete ebatüüpiliste soovide väljendamiseks, mis on tekkinud suhtluse käigus. Meie andmetes esineb kolm vormi.

Vorm *paneme* osutab koostööle ja on tüüpiline meditsiinidialoogides (*just, just. aga paneme::* `edasi ka `ravi juba `kinni).

Vormi *paneks* kasutatakse samuti võõrastevahelises suhtluses (*hh`paneks ühe`raviaja ka`kinni=et niigu`nii ma pean`tulema*). Ilma isikule viitava lõputa tingiva kõneviisi kasutamine võib osutada kõneleja ebakindlusele. Samuti on see viisakusvahend, mis vähendab survet soovi täitjale.

Vormi *pane* kasutatakse tuttavate suhtluses (*e no pane üksteist=e`neliküm-mend vast siis=e,`Kutsar`Maire tuleb siia`tagasi*).

Aja reserveerimise soove väljendatakse tüüpiliselt verbivormi *ma sooviks* abil.

Tabel 6. Verbide esinemine reisibüroodialoogides

Verb	Kõneviis			Kokku
	Kindel	Tingiv	Käskiv	
soovima (sh teada -, küsida -)		14		14
tahtma (sh teada -)		3		3
paluma		1		1
võtma		1		1
vaja olema		3		3
huvitama	1	3		4
huvitatud olema	2			2
küsima		1		1
oskama	1			1
olema pakkuda	1			1
Verbita				3
Kokku	5	26	0	34

Andke leiab kasutust infotelefonikõnedes dialoogi kestel tekkivate konkreetsete soovide väljendamiseks (*andke`laua oma`ka*).

Oskate soovitada /oskate öelda /saate öelda esinevad dialoogi käigus tekkivates ebatüüpilistes soovides, kus klient ei oska soovi ise esitada või ei ole kindel, et infoandja saab vastata (*et võibolla oskate midagi soovitada*).

Vaadake kasutatakse ainult dialoogi käigus tekkivate konkreetsete soovide väljendamiseks (*hh`aga`vaadake sellist=nagu`Pee`Ef`Tee*).

On pakkuda esineb üldises soovis (*aga võibolla teil on: `teisi jaanuaris variante ka`pakkuda*).

Küsin/küsiksin kasutatakse ainult lausungi selles osas, mis annab eelinfot, näitamaks, et järgneb soov. Lausungi sooviosa ei sisalda predikaati (*sellist`asja küsin`Pärnust`lapsehoidjat*).

Vabandage võib kasutada predikaadina käskiva kõneviisi korral või vormelina, mis eesti keeles väljendab viisakust. Meie andmetes kasutatakse teda vabandava vormelina. Lausungis on soovi sisaldav osa esitatud kui deklaratiivne lause kindlas kõneviisis (*vabandage`Tartus on mingi`söögi`koht nimega`Maailm*).

Need harvad verbid esinevad infotelefonidialoogides (*vaadake, tahaks küsida, saate, andke, vabandage*), reisibüroodialoogides (*sooviks teada, olen huvitatud, on pakkuda, sooviks küsida, huvitab/huvitaks*) või mõlemas nimetatud dialoogitüübis (*võtaks, tahaks teada, küsin/küsiks, oskate*).

6. Verbita soovid

Kuusteist soovi ei sisalda verbi. Sellised soovid saab jaotada kolme rühma. Esimese rühma moodustavad 9 noomenifraasi (`teisipäeva `päraslõuna pärast `kolme), teise rühma 6 dialoogipartiklit *jah/jaa/mhmh* ja kolmandasse rühma kuulub fraas *no nii, nii*.

Rolli alusel võib needsamad soovid jaotada kaheks. Esimese rühma moodustavad 13 soovi, mis on reaktsioonid ametniku eelmisele voorule. Ametnik on midagi pakkunud ja kliendi nõusolek on samaaegselt ka soov.

V: ahah, ma vaatan ´äkki mõni ´sidepunkt. (2.5) näiteks siin on ´Kuressaare sidepunkt ja (0.8) ja: ´liikuv ´postkontor. (.)

H: → no nii, [nii?]

V: [.h e] neli viis kaks?

Ülejäänud kolm soovi, mis on vormistatud noomenifraasidena, esitavad tüüpilisi konkreetseid soove infotelefoni- ja taksodialoogides:

V: `Maria=Takso tere.

H: → õõ `Lossi: `kolmteist.

Niisiis on kuueteistkümnest verbita soovist kolmteist (81%) dialoogi käigus tekkinud ebatüüpilised soovid.

7. Testimine

Viisime läbi kaks eksperimenti, moodustades selleks Tartu Ülikooli Eesti dialoogikorpusel baasil testkorpusel, kuhu võtsime 505 dialoogi (kokku 57 585 tekstisõna). Testkorpuses leidis 466 kliendi soovi.

Esimese eksperimendi eesmärgiks oli kontrollida, kui paljud meie analüüsitud korpuses esinevad tingivas kõneviisis verbid on kliendi infosoovides predikaatideks. Püstitasime hüpoteesi, et kui arvuti leiab lausungis sellise tingivas kõneviisis verbivormi, siis see lausung on tõenäoliselt soov. Tulemused on esitatud tabelis 7. Nagu näha, eristavad mõned märguanded soove väga hästi (*sooviks(in)*, *paluks(in)*). Keskmine soovide tuvastusprotsent oli analüüsitud korpuses 81 ja testkorpuses 78. Muidugi on soovide tuvastamiseks vaja leida lisaks tingivas kõneviisis verbidele ka muid märguandeid.

Teises eksperimendis testisime nelja verbivormi, mis esinesid tüüpilistes konkreetsetes soovides: *öelge*, *palun*, *paluks*, *sooviks*. Eesmärk oli teada saada, kui paljud neid verbivorme sisaldavad kliendi lausungid on soovid. Vaatlesime testkorpuses ainult neid kliendi soove, mis vahetult järgnesid rituaalsele algusele. Selliseid soove oli 270. Verb *öelge* esines 70 korda (neist 63 korda soovides), *palun* 110 korda (105 soovides), *paluks* 50 korda (48 soovides), *sooviks* 61 korda (54 soovides). Seega võib järeldada, et verbivormid on päris head märguanded kliendi esimeste konkreetsete soovide tuvastamiseks.

Muidugi võivad needsamad verbid esineda ka küsimustes, kuid me saame eeldada, et arvuti tuvastab küsimused enne soove, kasutades selleks muid, küsimustele spetsiifilisi keelelisi märguandeid.

Tabel 7. Tingivas kõneviisid verbide esinemiste arv analüüsitud korpuses ja testkorpuses

	Analüüsitud korpuses	Soovides	Testkorpuses	Soovides
sooviksin	11	7	27	25
sooviks	18	17	29	27
tahaks	11	7	6	2
paluksin	2	2	18	8
paluks	11	11	31	29
võtaks	9	6	24	24
oleks vaja	5	5	8	2
huvitaks/ uvitaks	6	3	13	7
paneks	1	1	9	7
küsisin	1	1	1	0
tahaks teada	5	4	2	1
tahaks küsida	1	1	2	1
sooviks teada	1	1	3	0
sooviksin teada	8	7	1	1
sooviksin küsida	1	1	6	6
Kokku	91	74/81%	180	140/78%

8. Arutelu

Nagu korpuse analüüs näitab, esinevad infosoovides enamasti kindlad verbid kindlates vormides.

Esiteks võib sellised verbivormid jaotada kahte rühma.

Esimese rühma verbide puhul kasutatakse käskivat kõneviisi (*üttele, pane, anna, vaata*; 13 soovi). Neidsamu verbe saab kasutada kavatsuse väljendamiseks tingivas kõneviisis ainult *kas*-küsimustes (*kas te (ei) vaataks/annaks/paneks/ütleks*).

Teise rühma kuuluvad verbivormid tingiva või kindla kõneviisi esimeses isikus. Sellisel viisil on formuleeritud 99 soovi. 73 neist sisaldavad verbi tingivas kõneviisis ning ülejäänud 26 kindlas kõneviisis.

Tingiv kõneviis ongi üldiselt seotud sooviga, lisades viisakust. Eesti keeles on tingival kõneviisil teatavasti täpne morfoloogiline tunnus *-ks-* (*-ksi-*), mida saab automaatsel tuvastamisel kasutada (EKG I: 235).

Kindel kõneviis on deklaratiivse akti universaalne vorm (Quirk jt 1972). Kõigil analüüsitud juhtudel on ta paralleelvariandiks tingiva kõneviisi kõrval. Ainsaks erandiks on ühend *huvitatud olema*, mis esineb meie andmetes kaks korda.

Mõningaid kindla kõneviisi vorme saab küll kasutada infosoovides, kuid mitte institutsionaalsete dialoogide puhul, sest nad väljendavad ranget ja ebaviisakat käsku (*sa ütled*).

Teiseks, soovid võib jaotada kahte rühma: tüüpilised ja ebatüüpilised. Tüüpiline soov rahuldab järgmisi tingimusi:

- suhtlus leiab aset võõraste vahel;
- soov esitatakse suhtluse alguses;
- autor on selle valmis mõelnud enne suhtluse algust ja sellisena on see helistamise tegelik põhjus.

Tüüpilised soovid saab jaotada kahte liiki: 1) soovid, kus kõneleja taotleb konkreetset infot või tegevust, 2) üldised soovid, mis juhatavad sisse teema.

Meie analüüs näitab, et soovides kasutatakse kindlaid verbe kindlatel tingimustel:

sooviks, paluks, palun, öelge tüüpilistes konkreetsetes soovides (üheainsa erandiga),

sooviks teada, sooviks küsida; mind huvitab /huvitaks /ma olen huvitatud ainult tüüpilistes üldistes soovides (meie andmetes esinesid sellised soovid üksnes reisibüroodialoogides).

Mõlemat liiki soove saab formuleerida vormi *oleks vaja* abil.

Vorme *võtaks, paneks /pane/paneme, andke, vaadake* kasutatakse ainult konkreetsetes soovides, mis tekivad dialoogi käigus.

Seevastu vorme *oskate soovitada, oskate öelda, saate öelda, on pakkuda* kasutatakse üldistes soovides, mis tekivad dialoogi käigus.

Tüüpiliste ja ebatüüpiliste soovide erinevus tuleb selgelt esile registratuuridialoogides. Tüüpiliste soovide esitamiseks kasutatakse enamasti vormi *ma sooviks(in)*. Klient kasutab vormi *tahaks* sellisel juhul, kui pakutud aeg on ebasobiv. Kui suhtlus leiab aset tuttavate vahel (arst ja registratuuritöötaja), siis kasutatakse soovide väljendamiseks vorme *ma tahan, tahaks*.

Verbi *panema* kõik kasutused on spetsiifilised ja ebatüüpilised (klient helistab tagasi või formuleerib soovi, mis on tekkinud dialoogi käigus).

Kaks kasutust (lisaks veel *palun*) ei ole soovlause predikaadid: *küsiksin/küsin* on eelteade, *vabandage* on viisakusvormel.

Kolmandaks, nagu nägime, vormistatakse tüüpilised soovid enamasti tingivas või käskivas kõneviisis. Ainsad erandid on *palun*, mille tähendus sisaldab viisakust, ja *mind huvitab/ma olen huvitatud*, mis rõhutab kõneleja huvi.

Verbi mittesisaldavad lausungid väljendavad enamasti ebatüüpilisi konkreetseid soove.

9. Dialoogi modelleerimine arvutil

Kui arvuti on tuvastanud predikaadi funktsioonis esineva verbi alusel kliendi soovi, siis verbi laiendid aitavad tal mõista selle soovi sisu (*mul oleks vaja sõita aprillis Budapesti; sooviks taksot Puurmanni viisteist.; ma palun: juuksuriäri Mag'noolia* jne), et soovi täita. Muidugi ei ole otstarbekas, et arvuti täidaks korraga erinevate ametnike rolle (infotelefonil vastaja, taksodispetšer jne), ühe ainevaldkonnaga piirdumine võimaldab oluliselt lihtsustada kliendi soovide prognoosimist ja tõlgendamist.

Enamasti ei saa ametnik kliendi soovi kohe täita, vaid peab esitama täpsustavaid küsimusi, s.t algatama alamdialoogi. Selliseid alamdialooge võib algatada ka klient³:

H: sooviks taksot `Puurmanni `viisteist. | DIE: SOOV |
(0.5)

V: → ja `kelle `nimele. | KYE: AVATUD | | VTE: VASTUSE TINGIMUSTE TÄPSUSTAMINE |

³ Dialoogiaktide märgendid vt nt <http://math.ut.ee/~koit/Dialoog/EDiC.html> (12.09.2006).

H: Ülle? | KYJ: INFO ANDMINE | | VTJ: VASTUSE TINGIMUSTE TÄP-
 SUSTAMINE |
 (.)
 V: `Ülle `nimele. | KYE: VASTUST PAKKUV | | PPE: ÜLEKÜSIMINE |
 H: jah. | KYJ: JAH | | PPJ: LÄBIVIIMINE |
 V: ja `tuleb teile `auto. | DIJ: MUU |

Arvuti peab suhtlejana kasutama teadmisi dialoogi struktuuri kohta – dialoogiaktide võimalikke järgnevusi –, mis aitab arvutil nii kasutaja akte tuvastada kui ka nendele adekvaatselt reageerida.

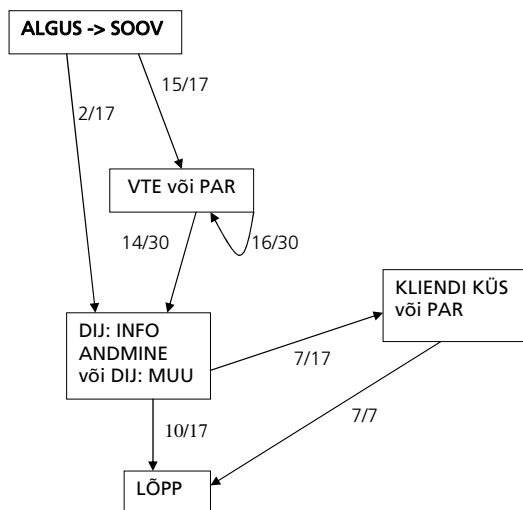
Üks võimalus dialoogi struktuuri esitada on nn Markovi ahel – graaf, mille sõlmedele vastavad dialoogiaktid ja servadele võimalikud üleminekud ühelt aktilt teisele koos üleminekutõenäosustega.

Uurisime oma analüüsitavas korpuses kliendi sooviga algavate dialoogide struktuuri ja moodustasime vastavad Markovi ahelad (joonised 1–4). Sooviga algavaid taksodialooge on 17, registratuuridialooge 15, infotelefonidialooge 48 ja reisibüroodialooge 23. Ülejäänud 41 dialoogis esitas klient pärast rituaalset sissejuhatust küsimuse. Lihtsuse huvides on joonistel ära jäetud rituaalne algus ja lõpp ning alamdialoogide struktuuri ei ole eraldi välja toodud.

Joonistel on kasutatud Eesti dialoogiaktide tüpologia aktinimesid ja järgmisi lühendeid:

- VTE – vastuse tingimuste täpsustamine (küsimus – vastus),
- PAR – parandussekvents (paranduse algatus – läbiviimine – [hindamine]),
- KÜS – (saadud infot) täpsustav küsimus ja vastus,
- DIJ – direktiivi (antud juhul soovi) järelliige.

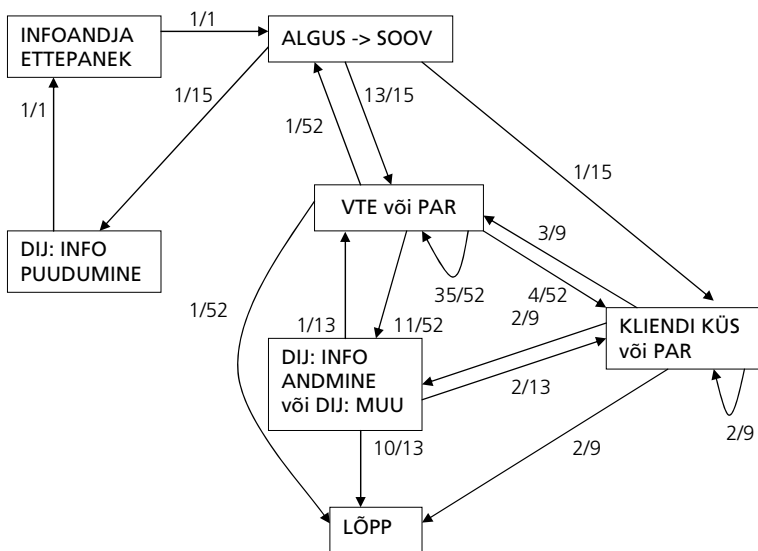
Nagu joonistelt näha, on taksokõned suhteliselt lihtsa ülesehitusega. Ülejäänud dialoogitüüpides on hargnemisvõimalusi ja erinevaid üleminekuid tunduvalt rohkem. Eriti keeruline on reisibüroodialoogide struktuur, need on ka tunduvalt pikemad kui muud dialoogid (vrd tabel 1).



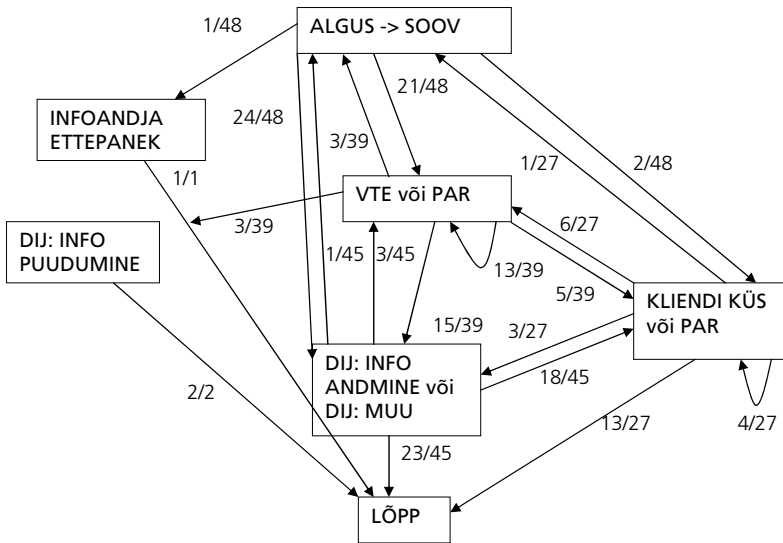
Joonis 1. Taksodialoogide mudel

Järgmises näites on paksus kirjas toodud see osa taksodialoogist, mida kajastab joonis 1: kliendi soovile järgneb ametniku vastuse tingimuste täpsustamine ja kliendi vastus (joonisel tähistab seda alamdialogi VTE), misjärel ametnik teatab soovi täitmisest (DIJ: MUU):

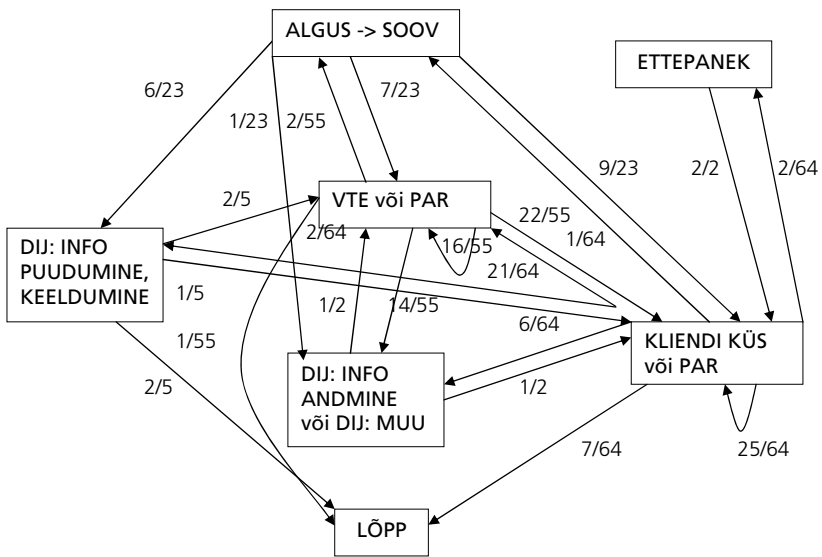
((kutsung)) | RIE: KUTSUNG |
 V: `Maria takso | RIJ: KUTSUNGI VASTUVÕTMINE | | RY: TUTVUS-
 TUS |
 tere? | RIE: TERVITUS |
 H: tervist, | RIJ: VASTUTERVITUS |
sooviks taksot `Puurmanni `viisteist. | DIE: SOOV |
(0.5)
V: ja `kelle `nimele. | KYE: AVATUD | | VTE: VASTUSE TINGI-
MUSTE TÄPSUSTAMINE |
H: Ülle? | KYJ: INFO ANDMINE | | VTJ: VASTUSE TINGIMUSTE
TÄPSUSTAMINE |
(.)
V: `Ülle `nimele. | KYE: VASTUST PAKKUV | | PPE: ÜLEKÜ-
SIMINE |
H: jah. | KYJ: JAH | | PPJ: LÄBIVIIMINE |
 V: ja `tuleb teile `auto. | DIJ: MUU |
(.)
 H: ai[täh.] | RIE: TÄNAN |
 V: [`kõi]ke=ead teile. | RIE: SOOVIMINE |
(.)
 H: `näge`mist | RIE: HÜVASTIJÄTT |



Joonis 2. Registratuurialoogide mudel



Joonis 3. Infotelefonidialogide mudel



Joonis 4. Reisibüroodialoogi mudel

Analüüsitud dialoogide tüüpilise struktuuri võib esitada regulaaravaldisega:

kliendi soov (ametniku algatatud täpsustus või parandus)*
ametniku vastus (kliendi algatatud täpsustus või parandus)*.

Tärn (*) tähistab siin sulgudes toodud alamdialoogi esinemist null või enam korda, s.t kliendi soovile võib järgneda (aga ka mitte järgneda) ametniku algatatud alamdialooge, seejärel annab ametnik soovitud info ja pärast vastuse saamist võib klient omakorda algatada täpsustavaid alamdialooge.

Sellisele muustrile vastavad kõik sooviga algavad taksodialoogid, 85% infotelefon- ja 80% registratuurikõnedest, kuid ainult 60% reisibüroodialoogidest. Põhiline muustrist kõrvalekalle seisneb selles, et ühe osaleja algatatud alamdialoogi järel (või ka sees) võib partner omakorda algatada alamdialooge, s.t initsiatiiv liigub ühelt osalejalt teisele, küsija ja vastaja rollid vahelduvad. Näiteks esitab ametnik enne infoandmist täpsustava küsimuse (algatades sellega alamdialoogi) ja klient jätkab kohe pärast vastamist omakorda täpsustava küsimusega:

V: mis `kuupäev teid `uvitab. | KYE: AVATUD | | VTE: VASTUSE TINGIMUSTE TÄPSUSTAMINE |

H: ee praegu `kuupäeva võibola täpselt ei `oskagi `öelda, | KYJ: INFO PUUDUMINE | | VTJ: VASTUSE TINGIMUSTE TÄPSUSTAMINE |

H: .hh aga sooviks teada=et mitu `inimest on ültse `rühmas. | KYE: AVATUD |

H: kui suured grupid teil `on. | KYE: AVATUD |

V: `Varmõnsheik on `lennuki `täis sada= `seitsekend `inimest. | KYJ: INFO ANDMINE |

Või esitab klient pärast info saamist uue küsimuse ja ametnik täpsustab omakorda, enne kui vastab sellele küsimusele:

H: hh aga kas sellist varianti on `ka=et=ee näiteks kahekümne= `öeksandal minna. | KYE: JUTUSTAV KAS |

V: nii | VR: NEUTRAALNE VASTUVÕTUTEADE |

V: ja siis mitu päeva `olla. | KYE: AVATUD | | VTE: VASTUSE TINGIMUSTE TÄPSUSTAMINE |

H: noh ma=i=tea. | KYJ: EDASILÜKKAMINE | | VTJ: VASTUSE TINGIMUSTE TÄPSUSTAMINE |

H: viis= `kuus | KYJ: INFO ANDMINE | | | VTJ: VASTUSE TINGIMUSTE TÄPSUSTAMINE |

V: äm=ma usun `küll et on selline variant ka `võimalik. | KYJ: JAH |

Koostatud Markovi ahelad vajavad muidugi täpsustamist, enne kui neid dialoogsüsteemis saab kasutada.

10. Kokkuvõte ja edasine töö

Valisime analüüsimiseks 144 institutsionaalset telefonikõnet Tartu Ülikooli Eesti dialoogikorpuselt. Need jagunevad nelja situatsioonirühma: infotelefon, reisibüroo, polikliiniku registratuur, takso tellimine.

Meie analüüsitud dialoogides esitavad kliendid kaht liiki soove. Ühed on infosoovid (helistaja vajab kindlat infot, nt telefoninumbrit), teised on ametniku tegevust taotlevad soovid (reserveerida aeg arsti vastuvõtule, saata takso). Sellest seisukohast jagunevad dialoogid kahte rühma. 1) Infotelefoni- ja reisibüroodialoogides on helistaja soovid eranditult infosoovid. 2) Takso- ja registratuurialoogides esinevad soovid on eranditult tegevust ootavad soovid.

Sealjuures on oluline, et ka tegevuse soovile vastatakse ikkagi keeleliselt, infoga, mis ütleb, et soov täidetakse või mis tingimustel saab seda soovi täita.

Analüüsi tulemusel saab teha järgmised järeldused.

- Soovid vormistatakse väheste verbide abil.
- Keskne on tingiv kõneviis ning selle sees teatud verbid kindlates vormides (tabel 2).
- Kindlates situatsioonides on omad vormelistunud konstruktsioonid: info-telefonis *palun/paluks, öelge, sooviks*; takso tellimisel *palun, sooviks*, registratuuris *sooviks*.
- Teatud situatsioonides juhatab soov sisse pikema jutu, nagu reisibürookõnedes.
- Teatud sõnavormid on kasutusel vestluse keskel, kokkuvõttena või uue teemaosa algatajana.

Seega näitas meie analüüs, et kliendi soovid vormistatakse eesti institutsionaalsetes dialoogides piiratud arvu verbivormide abil, mis esinevad lausungis kui predikaadid. Verbi võimaliku kasutuse määrab tema semantika.

Lisaks analüüsisime kliendi sooviga algavate dialoogide struktuuri, mille esitasime Markovi ahelatenä.

Küsimusi ja nende tuvastamise märguandeid on analüüsitud artiklis (Hennoste jt 2005) ning rituaale on käsitletud töös (Hennoste, Rääbis 2004).

Edaspidi kavandame analüüsida rohkem dialooge, et täpsustada verbivormide kasutustingimusi. Samuti tuleb moodustada detailsed Markovi ahelad, mille alusel saaks dialoogsüsteem kasutaja dialoogiakte tõlgendada ja neile adekvaatselt reageerida.

Kirjandus

- Allen, James; Core Mark 1997. Draft of DAMSL: Dialog Act Markup in Several Layers. <http://www.cs.rochester.edu/research/cisd/resources/damsl/RevisedManual/RevisedManual.html> (4.07.2006).
- 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.
- EKSS = Eesti kirjakeele seletussõnaraamat. 1988 –. Eesti Keele Instituut. Tallinn.
- Goodwin, Charles 1996. Transparent vision. – Elinor Ochs, Emanuel A. Schegloff, Sandra A. Thompson (Eds.). Interaction and Grammar. Cambridge: Cambridge University Press, 370–404
- Hennoste, Tiit; Rääbis, Andriela 2004. Dialoogiaktid eesti infodialoogides: tüpoloogia ja analüüs. Tartu Ülikooli arvutiteaduse instituut. 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. – Václav Matousek, Pavel Mautner (Eds.). Text, Speech and Dialogue. 6th International Conference TSD 2005. Springer, 420–427.
- Hutchby, Ian; Wooffitt, Robin 1998. Conversation Analysis. Principles, Practices and Applications. Cambridge: Polity Press.
- Jurafsky, Daniel; Martin, James H. 2000. Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Upper Saddle River (NY): Prentice Hall.
- Jurafsky, Daniel; Shriberg, Elizabeth E.; Fox, Barbara; Curl, Traci 1998. Lexical, prosodic, and syntactic cues for dialog acts. – Proceedings of ACL/COLING–98 Workshop on Discourse Relations and Discourse Markers. ACL, 114–120.
- Quirk, Randolph; Greenbaum, Sidney; Leech, Geoffrey; Svartvik Jan 1972. A Grammar of Contemporary English. London, New York: Longman.
- Treumuth, Margus 2005. A software tool for the Estonian Dialogue Corpus. – Proceedings of the Second Baltic Conference on Human Language Technologies. Tallinn, 341–346.

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

Tiit Hennoste (Helsingi Ülikool, Tartu Ülikool) uurimisvaldkondadeks on kommunikatsiooniteooriad, suuline keel jm. tiit.hennoste@ut.ee

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

Mare Koidu (Tartu Ülikool) erialaks on keeletehnoloogia; on uurinud dialoogi modelleerimist arvutil. mare.koit@ut.ee

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

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

Maret Valdisoo (Tartu Ülikool) on uurinud dialoogi modelleerimist arvutil. maret@ut.ee

Evelyn Vutt (Tartu Ülikool) on uurinud dialoogi modelleerimist arvutil. nurm@ut.ee

AUTOMATIC RECOGNITION OF CUSTOMER REQUESTS IN ESTONIAN INFORMATION DIALOGUES

**Olga Gerassimenko, Tiit Hennoste, Riina Kasterpalu,
Mare Koit, Andriela Rääbis, Krista Strandson, Maret Valdisoo,
Evely Vutt**

The goal of the paper is to find out linguistic cues which can be used for automatic recognition of dialogue acts by the computer. We limit us with customer requests here which are the most important acts for modelling the dialogue on the computer. We have chosen 114 dialogues from the Estonian Dialogue Corpus of the University of Tartu. Four situational groups are represented in the dialogues: directory inquiries (phone numbers, addresses, etc.), calls for taxi, calls to travel agencies, calls to outpatients' offices. There were 130 customer requests in our analysed corpus. Customers ask either for information or for an action of the operator. As the analysis demonstrates, only certain modes and persons of certain verbs can be used in order to express requests. Certain verbs are systematically related to certain contextual conditions. We represent the structure of a dialogue as a sequence of dialogue acts using Markov chain. It helps the computer to recognise the user acts and adequately react to them.

Keywords: dialogue corpus, dialogue act recognition, conversation analysis, Estonian

MORPHOLOGICAL MODIFICATIONS IN LITHUANIAN CHILD-DIRECTED SPEECH

Laura Kamandulyte

Abstract. Child-directed speech (CDS) is very important for the language development of a child. When talking to children adults modify their language by simplifying it and adapting it to the language acquisition stage of the child. Words and constructions that are used in CDS are acquired earlier by the child.

This paper analyses the morphological modifications of Lithuanian CDS.

The analysis proves that the most frequent morphological modifications of CDS include: 1) the higher frequency of verb tokens over noun tokens, 2) the low frequency of indefinite pronouns, demonstrative pronouns, and adjectives, and 3) the high frequency of diminutives and interjections.

Prevalence of the denominative parts of speech has been influenced by the cognitive function; the abundance of verbs is conditioned by short utterances. The high frequency of diminutives is related to pragmatic function, intimacy, and emotions.

Keywords: language acquisition, longitudinal corpus, premorphology, protomorphology, modularized morphology, register of CDS, input, Lithuanian

1. Introduction

Many studies emphasize that CDS differs from adult speech by phonetic (high tone, clear articulation), morphological (abundance of nouns, especially diminutives, and verbs), syntactic (limited vocabulary, short sentences, large number of interrogative sentences) and pragmatic features (many questions and references) (Cruttenden 1994, Dressler, Barbaresi 1994, Ferguson 1977, Harkness 1976, Snow 1994).

CDS of the Lithuanian language is a new topic, and there has been little research in this field. Phonetic, lexical, and morphological features of Lithuanian CDS were briefly surveyed by Wójcik (1994). The researchers who have analysed acquisition of the Lithuanian language have also discussed the impact of adult speech on CDS. Savickienė (1999, 2000, 2003) described the features of CDS in regard to noun usage, whereas Wójcik (2000) described the usage of verbs in speech directed to children. There have been no other studies of CDS in Lithuania.

In 1977 Ferguson (1977: 232–235) noted three main functions of CDS:

- 1) the function of communication and self-expression. The structure of CDS is in large part a response to the need for improved communication when one of the participants has only a limited ability to use language normally.
- 2) the function of language teaching. The structure of CDS reveals some simplifications, and the complexity is gradually restored as the child's language progresses.
- 3) the function of socialization. CDS offers a means of identifying the social roles which the child must learn as requiring differential behaviour in his society.

According to Ferguson (1977), and due to these three functions, various CDS modifications appear, differentiating it from usual adult speech. One of the major modifications is considered to be the allocation of parts of speech differing from that of adult speech.

The goal of this paper is to analyse the characteristic morphological features of Lithuanian CDS and to discuss the relation between the morphological modifications in CDS and the child's age.

2. Method

Studies in child language and CDS are ambivalent – on the one hand there is psychology and, on the other, linguistics. It is typical of psycholinguistic research, such as psychological sciences, to analyse the individual data. In this study the method of longitudinal research is applied.

The study focuses on naturally recorded conversations of a mother and her child Rūta. The conversations were recorded by the mother, a graduate in Lithuanian philology, 3–4 times per week. She was instructed to possibly cover all kinds of everyday situations. Data collection started when the daughter was 1;3¹ and was continued until the daughter was 3;5. The mother recorded 35 hours of naturally occurring conversations at home and in other places (e.g. while on a visit and taking a walk) in situations of free play and everyday situations (e.g. eating, washing, and dressing). The data were transcribed in CHAT format and coded morphologically using the program CHILDES (MacWhinney 2000).

The analysis was carried out according to the child's different stages of language acquisition – premorphologic, protomorphologic, and modularized morphology (Dressler et al. 1995–1996; Dressler 1997). To illustrate these stages, the following ages were chosen: the premorphologic stage is represented by age 1;7, protomorphologic by age 1;9, and modularized morphology by age 2;5.²

¹ The numerals show the child's age (e.g. 1;3 – 1 year 3 months).

² The stages of acquisition were determined by I. Savickienė (1999).

During the premorphologic stage, the imitation of utterances and mechanical repetitions are typical of children's speech. The rules of morphology are applied by analogy, and the basics of vocabulary are evident during protomorphology. The stage of modularized morphology of children's speech already shows the acquisition of morphological categories, composite sentences, and the use of various vocabularies.

The results of CDS were compared with the data of the analysis of conversations between the same informant and another adult. Five hours of conversations between the informant and the other adult (a friend of the informant) cover different conversational themes (everyday situations, work, and research). The data were transcribed and coded morphologically using the program CHILDES (MacWhinney 2000).

The CDS is comparable with ADS (adult-directed speech) because conversations were recorded in both cases in everyday situations, and in both cases the informant was communicating very intimately (with a daughter and a good friend). The results of CDS were compared with the results of typical ADS (when the ability of an addressee to use language is unlimited) because the modifications of CDS can be set when the language is compared with the typical language, which does not have any special modifications. Otherwise, one might argue that the feature of language may just be a peculiarity of the language of a certain speaker and does not have to be a pervasive feature of the CDS register (Remick 1976, Kempe et al., forthcoming).

Child language and CDS became the subject of research interest in Lithuania only a few years ago; therefore, there are not many recorded and transcribed data of child language and CDS. Also, the results can not be compared with the longitudinal data of the language of other informants. However, the studies on other languages show that the method of individual longitudinal data research is reliable for psycholinguistic studies.

3. Quantitative analysis of morphological modifications

The hypothesis of this study was that since special functions are typical of the register of CDS, they should be reflected by more or less abundant use of particular parts of speech.

3.1. Inflected parts of speech

All the denominative parts of speech (noun, adjective, verb, numeral, pronoun, adverb) are inflected in Lithuanian. They denominate things, phenomena, features, or actions. Because of these functions, denominative parts of speech are dominant in speech. It is natural that this tendency prevails in CDS because the usage of denominative parts of speech has been influenced by the cognitive function. The mother names the different things, actions, features and gives the child a new knowledge about the world. However, the distribution of denominative parts of speech in the speech of the mother differs according to the child's stage of language acquisition (see Table 1).

Table 1. Distribution of inflected parts of speech

	1;7	1;9	2;5	with adult
Part of speech	Percentage among all words			
Noun	22%	21.7%	21.9%	18.7%
Adjective	1.2%	2.1%	2.9%	6.4%
Numeral	3.0%	–	0.4%	1.2%
Pronoun	10.8%	15.4%	13.0%	14.7%
Verb	28.1%	25.7%	28.4%	20%
Adverb	18.8%	17.7%	12.9%	16%
Total	83.9%	82.6%	79.5%	77%

Table 1 shows that verbs are most common in the mother's speech. Although this part of speech dominates in the conversations with adults as well, here its frequency is much lower than when speaking to the child. The percentage of verbs used in the CDS in different stages varies from 25.7% to 28.4%, and when speaking to the adult – 20.0%.

Such plentiful use of verbs in the CDS results from short utterances expressing short thoughts. The speech contains nouns and no recitation; the simple thought is often expressed by means of verbs, a noun, and an adverb or an official part of speech. For example:

(1a) *Močiutė miega jau.* 'The grandmother is sleeping already.' (1;7)

(1b) *Net vystyklo nereikia?* '[You] don't even need a diaper?' (1;9)

(1c) *Meškučiai jau miega.* 'Teddies are already sleeping.' (2;5)

Moreover, CDS contains many verbs and several question words and dictations, where no noun is used. For example:

(2a) *Skaitai?* 'Are you reading?' (1;7)

(2b) *Padainuok dar.* 'Sing:IMP more.' (1;7)

(2c) *Ką nori daryti?* 'What do [you] want to do?' (1;9)

(2d) *Nusiprausk.* 'Wash yourself.' (1;9)

(2e) *Ko tu nori?* 'What do you want?' (2;5)

(2f) *Truputi garsiau šnekėk.* 'Speak:IMP a little bit louder.' (2;5)

The domination of verbs in CDS is determined by the heavy use of short utterances, regulations, and questions. It proves the claim by Ferguson (1977: 232) that the communicative function is the most important function of the register of simplified CDS. Adults, when communicating with children, try to be understood when noting, asking, or chatting in short and simple verbal clauses. Therefore, the difference between the usage of verbs and nouns in CDS is much greater than in speech among adults. In the mother's conversation with the child the difference in the percentage of verbs and nouns (difference= verb%–noun%) ranges from 4 to 7, whereas in the conversation with an adult it is 1. Thus, more frequent use of verbs and less frequent use of nouns can be considered as characteristic features of CDS.

In regard to nouns, it is important to mention that one of the most important features of CDS is the frequent use of diminutives – from 47% to 64% of all nouns and from 10% to 14% of all words (see Figure 1, Figure 2).

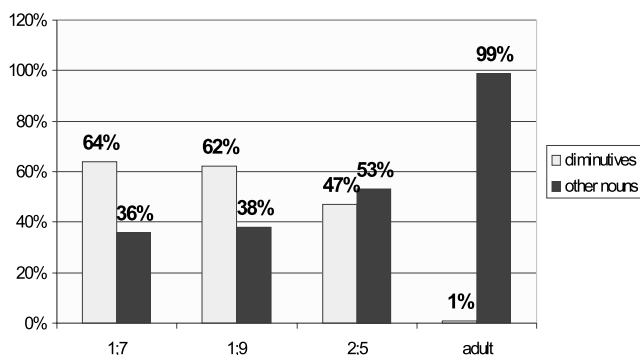


Figure 1. Distribution of diminutives among all nouns

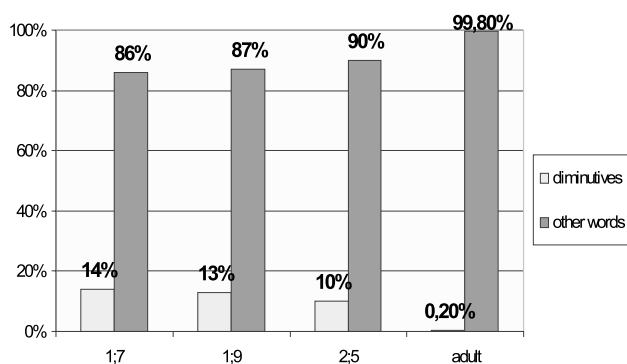


Figure 2. Distribution of diminutives among all words

The use of diminutives in CDS is related to pragmatic functions. Since the register of CDS is characterized by intimacy and emotion, the abundance of diminutives is not surprising. Although diminutives in the Lithuanian language are longer than their basic forms, according to V. Kempe (in press), their usage does not contravene the CDS as a mode of the simplified register. Diminutives are not only important for the process of segmentations where the meaning of a noun is emphasized by its suffix form; they simplify the system of the language as well. The diminutive forms of all the twelve noun paradigms in the Lithuanian language obtain the flexions of only three paradigms: the first (inflection *-as*), the third (inflection *-is*), and the eighth (inflection *-ė*) (Savickienė 2003). Moreover, the diminutive suffixes in the Lithuanian language are always stressed, this way attracting the child's attention and helping him/her to acquire diminutive nouns faster.

The diminutives used by the mother mainly name people, toys, food, animals, parts of the body, clothing, furniture, and abstractions that are known to the child (Savickienė 2003). The majority of diminutives in the mother's speech relate to the naming of people (see Table 2). Nearly all of these diminutives are proper nouns modified to express tenderness and emotion. For example:

- (3a) *O kq Rūtytē myli?* ‘What does **Rūtytē**:PNOUN:DIM love?’
 (3b) *O šīandien matei Akvilytē?* ‘Have [you] seen **Akvilytē**:PNOUN:DIM:ACC today?’ (1;9)

Table 2. Distribution of diminutives

	1;7	1;9	2;5	with adult
Common nouns	54%	75%	74%	100%
Proper names	46%	25%	26%	–
Total	121 (100%)	110 (100%)	99 (100%)	2 (100%)

As Table 2 shows, diminutive proper nouns in the informant’s speech were most frequently used during the first analysed stages (46% – 1;7). Later the diminutive forms of proper nouns were used more seldom while diminutive common nouns became more frequent. The percentage of diminutive self-entitling nouns decreases as well. During the first stages the informant often calls herself by using diminutives. For example:

- (4a) *Kur mamytē?* ‘Where is **Mummy**?’ (1;7)
 (4b) *Duok, mamytē išgers.* ‘Give [me], **mummy** will drink.’ (1;9)

As the child’s language acquisition increases, diminutives are replaced by pronouns, and the function of speech simplification becomes less important.

Diminutive adjectives and adverbs are rarely used in CDS. This tendency is typical of adult speech as well.

The study of the diminutives in the mother’s speech revealed that each stage of CDS is characterized by frequent use of common words – diminutives. Diminutive proper nouns are used in the mother’s speech only in the early stages of the child’s speech acquisition (during the stages of premorphology and modularized morphology).

Frequent use of adverbs (as well as of pronouns) in CDS is determined by short verbal utterances. The majority of the adverbs are used during the first stages of the child’s speech acquisition (18.8% – see Table 1). Later the number decreases. It has been noticed that the decline of adverbs in the CDS is influenced by the decreased frequency of the use of the interrogative adverbs *where*, *how*, and *why*, which are very frequent during the early stages (see Table 3).

Table 3. Distribution of adverbs

	1;7	1;9	2;5	with adult
Interrogative	24%	19%	16%	5%
Other	76%	81%	84%	95%
Total	169 (100%)	154 (100%)	129 (100%)	236 (100%)

When communicating with a smaller child, interrogative utterances dominate the CDS. They are used with the intention to involve the child in a conversation and to encourage him/her to speak. For example:

- (5a) *Kaip tu gyveni, Rūtuže?* ‘How are you, **Rūtuže**:PNOUN:DIM?’ (1;7)
 (5b) *Kur tu bēgi?* ‘Where are you running?’ (1;7)

Thus, one of the main features of CDS could be frequent use of adverbs during the early stages of language acquisition.

During the later stages of the child's language acquisition the use of interrogative adverbs decreases. More adverbs denoting character, location, and time are used. For example:

(6a) *Dabar bus gerai*. 'Now [it] will be good.' (2;5)

(6b) *Teisingai*. '[It] [is] right.' (2;5)

(6c) *Čia bus gerai*. '[It] will be good **here**.' (2;5)

The child's conception of space and time during the stages of protomorphology and modularized morphology contributes to the increased use of a wider variety of adverbs in the mothers' speech. When the child accumulates some elements of morphology and a rich vocabulary, less attention is paid to teaching and naming particular denotations and to interrogative adverbs, and more emphasis is placed on communication and adverbial words determining space and time.

Table 1 shows that CDS is full of pronouns. The percentage of pronouns during the latest analysed stage is higher than that of the first stage and is more similar to that of adult speech. However, besides this tendency to increased pronoun use, there is another important tendency that is related to the meanings of the pronouns. The frequent use of some pronouns results from intimate conditions of communication, familiar atmosphere (*you, I, yours, mine, for you, for me*), and the function of CDS to acknowledge the world via questions (*what?, who?*). Table 4 presents the distribution of meanings of pronouns used in various stages.

Table 4. Distribution of pronouns

	1;7	1;9	2;5	with adult
Personal	43%	40%	51%	41%
Demonstrative	3%	7%	12%	17%
Interrogative	38%	42%	18%	10%
Indefinite	16%	11%	19%	33%
Total	97	134	130	217

Table 4 shows that the most frequently used pronouns in the CDS are personal pronouns. The percentage of personal pronouns used in some stages is significantly higher than in the speech of adults (51% at age 2;5 – 41% with the adult). It is caused by the intimate atmosphere of communication and the narrowness of the child's world where the most important people are the child and the mother. Whereas in the CDS no increased or decreased tendency was observed in the use of pronouns carrying this meaning, it is possible to claim that the use of such pronouns is an important feature of CDS regardless of the child's age.

Interrogative pronouns constitute another group of pronouns typical of CDS. The mother uses these pronouns especially often during the first stages of language acquisition (38%–42%). For example:

(7a) *Kas čia yra?* 'What is it?' (1;7)

(7b) *Ką daro tėvelis dabar?* 'What is the father doing now?' (1;9)

Later, more active participation of the child in conversation and less frequent questions from the mother result in the decreased use of interrogative pronouns. However, even during the last analysed stage of CDS, the percentage of interrogative pronouns is much higher than in the speech with the adult (18% at age 2;5 – 10% with the adult).

The CDS uses indefinite pronouns not very often (19% at age 2;5 – 33 % with the adult). Although in the speech of adults it is the second biggest notional group of pronouns, it is not so frequent when speaking with children. It is more difficult for children to conceive indefinite pronouns, as they do not define concrete things or persons. Therefore, the indefinite pronouns in the mother's speech are used more often during the stage of modularized morphology. For example:

(8a) *Nėra **nieko** ten.* 'There is **nothing** there.' (2;5)

(8b) *Va šitoj mažutėj knygutėj šuniukas yra **kur nors** nupieštas.* 'The dog: DIM is painted **somewhere** in this small: DIM book: DIM.' (2;5)

However, even in this stage, the percentage of indefinite pronouns differs significantly from that in adults speech (19% at age 2;5 – 33% with the adult).

In the early CDS the demonstrative pronouns are rare (3% – 1;7). Their use in the mother's speech depends on the child's age. During the last stage it already reaches a higher percentage (12% at age 2;5). For example:

(9a) *Va, šita graži knygutė, pažiūrėk.* 'Look, this is a nice book: DIM.'

(9b) *Nori šita lėlytę paguldyt?* 'Do you want to put this doll: DIM to bed?' (2;5)

It is interesting to note that the demonstrative pronoun *tai* 'that' predominates in the speech of adults, whereas *tas* 'this' predominates in CDS. This is due to the limitation of the child's atmosphere and the need to define objects of the world around.

Adjectives are rarely used in CDS. Even during the stage of modularized morphology the use of adjectives in the mother's speech makes up less than 3% (2,9% at age 2;5). Mainly adjectives with positive connotations, such as 'beautiful', 'small' and 'good', are used. This part of speech only describes and does not provide any new information. Consequently, it is rarely used.

The other part of speech rarely used in CDS is that of numerals. Its concrete meaning results in infrequent use in special cases. The most frequently used numeral in the speech of the informant is 'one'. It constitutes the opposition to the adverb *a lot*; therefore, it is used more often than other numerals.

The usage tendencies of denominative parts of speech in CDS are determined by the age of the child participating in the communication. The smaller the child the more denominative parts of speech naming concrete things and actions are used in the mother's speech. With the growth of the child the percentage of the denominative parts of speech decreases.

3.2. Uninflected parts of speech

Uninflected parts of speech (preposition, particle, conjunction, interjection) do not have any entitling meaning. They connect the denominative parts of speech. Only prepositions distinguish themselves from the other uninflected parts of speech because they are often required by verbs. Particles or conjunctions can almost always be skipped. For these reasons, one can expect few particles and conjunctions in CDS, which is considered as a simplified register.

According to Paulauskienė (1994), interjections are not necessary for communication because they do not carry any information. However, the researchers of CDS are of the opinion that such emotionally expressive words are very important in the speech of children. They do not only express joy or dissatisfaction; they also convey certain information. They are holophrases, which often have the meaning of the whole sentence. Since the meaning of interjections in CDS is very significant, it is expected that in CDS these parts of speech also have a significant communicative function. Table 5 shows the regularity of usage of all uninflected parts of speech.

Table 5. Distribution of uninflected parts of speech

Parts of speech	1;7	1;9	2;5	with adult
Prepositions	1.2%	1.4%	1.9%	5.1%
Particles	5.1%	4.7%	4.2%	9.3%
Conjunctions	5.9%	7.4%	8.5%	8.1%
Interjections	4%	4%	6%	1%
Total	12.2%	13.5%	14.6%	22.5%

It is interesting to note that the use of particles and conjunctions, contrary to our expectations, is not low. The conjunctions or particles serve the functions of intonation, expression, and drawing attention but not the function of conjoining. For example:

(10a) *O kodėl tu bėgai?* ‘And why were you running?’ (1;7)

(10b) *Ir gulės lėlytė.* ‘And the doll:DIM will lie.’ (1;9)

Several particles and conjunctions dominate CDS – *and*, *but*, *this*, and *here* – which have the constant function of adding emphasis and stress. The particle *but* is often used to intensify questions.

Comparison of the percentage rates of CDS and adult speech in Table 5 shows that the informant uses more interjections in CDS than when speaking to the adults (2.6% at age 1;7; 2.8% at age 1;9; 4.1% at age 2;5 and 0.1% with adult). For example:

(11a) *Rūtyte, duok ir pasakyk labas.* ‘Rūtyte, give and say hello.’ (1;7)

(11b) *Sakyk ačiū, gerai.* ‘Say, **thank you**, fine.’ (1;9)

The most frequently used interjections are *thanks*, *welcome*, *hi*, and *alio*, reflecting the educational function – the mother teaches her daughter phases of politeness. When communicating with a child, such phrases serve not only as indicators of

politeness; they also have the function of mutual relation, support of communication, and attraction of attention.

In conclusion, use of interjections is a typical feature of CDS. These emotional-expressive parts of speech are important not only as factors of particular meaning but also as the means of supporting the communicative act.

4. Conclusion

After the study of conversations of the mother both with her daughter and with other adults it is possible to conclude that the most frequent **morphological** modifications of CDS include:

- 1) higher frequency of verb tokens over noun tokens,
- 2) low frequency of indefinite, demonstrative pronouns and adjectives,
- 3) high frequency of diminutives and interjections.

The prevalence of the denominative parts of speech has been influenced by the cognitive function, and the abundance of verbs is conditioned by short utterances. The high frequency of diminutives in CDS is related to pragmatic function, intimacy, and emotion.

Also, the hypothesis that since the special functions are typical of the register of CDS, they should be reflected by more or less frequent use of particular parts of speech, was confirmed.

The morphological modifications in CDS can be divided into two groups:

1) modifications independent of the child's age, and 2) modifications dependent on the child's age. Independent modifications are the features of CDS that do not depend on the stage of language acquisition (use of diminutives). Dependent modifications are the features of CDS that develop in their frequency and character together with the child's age (distribution of parts of speech).

References

- Cruttenden, Alan 1994. Phonetic and prosodic aspects of baby talk. – C. Gallaway, B. J. Richards (Eds.). *Input and Interaction in Language Acquisition*. Cambridge University Press, 135–152.
- Dressler, Wolfgang U. 1997. Zur Entstehung der Distinktivität in der Kindersprache. – *Polyphonie pour Ivan Fonagy*. Jean Perrot (Ed.). Paris, Canada: Editions L'Harmattan.
- Dressler, Wolfgang U.; Drażyk, R. D.; Dziubalska-Kołaczyk, K.; Jagła, E. 1995-1996. On the earliest stages of acquisition of polish declension. – *Wiener Linguistische Gazette*, Vol. 53–54, 1–21.
- Dressler, Wolfgang U.; Barbaresi, Merlin L. 1994. *Morphopragmatics. Diminutives and Intensifiers in Italian, German and other Languages*. Trends in Linguistics: Studies and Monographs. Berlin.
- Ferguson, Charles A. 1977. Baby talk as simplified register. – C. E. Snow, C. A. Ferguson (Eds.). *Talking to Children: Language Input and Acquisition*. Cambridge University Press.
- Harkness, S. 1976. Mother's language. – W. von Raffler-Engel, Y. Lebrun (Eds.). *Baby Talk and Infant Speech*. Amsterdam: Swets; Zeitlinger B. V., 110–112.

- Kempe, Vera; Brooks, Patricia J.; Gillis, Steven (forthcoming). Diminutives provide multiple benefits for language acquisition.
- MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk. Vol. I: Transcription, Format and Programs*. Mahwah, NJ: Erlbaum.
- Paulauskienė, Aldona 1994. *Lietuvių kalbos morfologija*. Vilnius: Mokslo ir enciklopedijų leidykla.
- Remick, H. 1976. Maternal speech to children during language acquisition. – W. von Raffler-Engel, Y. Lebrun (Eds.). *Baby Talk and Infant Speech*. Amsterdam: Swets; Zeitlinger B. V., 223–233.
- Savickienė, Ineta 1999. *Lietuvio vaiko daiktavardžio morfologija*. PhD Thesis. Kaunas: Vytauto Didžiojo universitetas.
- Savickienė, Ineta 2000. *Linksniai šnekamojoje kalboje*. – *Darbai ir dienos* 24, 89–99.
- Savickienė, Ineta 2003. *The Acquisition of Lithuanian Noun Morphology*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Snow, Catherine E. 1994. Beginning from baby talk: twenty years of research on input in interaction. – C. Gallaway, B. J. Richards (Eds.). *Input and Interaction in Language Acquisition*. Cambridge University Press, 3–12.
- Wójcik, Paweł 1994. Some characteristics features of Lithuanian baby talk. – *Linguistica Baltica* 3, 71–86.
- Wójcik, Paweł 2000. *The Acquisition of Lithuanian Verb Morphology: A Case Study*. Krakow: Quartis.

Laura Kamandulyte, Department of the Lithuanian Language at Vytautas Magnus University, research interests: corpus linguistics, psycholinguistics, first language acquisition, second language acquisition, language impairment.
l.kamandulyte@hmf.vdu.lt

LEEDU HOIDJAKEELE MORFOLOOGILISED MODIFIKATSIOONID

Laura Kamandulyte

Vytautas Magnuse nimeline Kaunase Ülikool

Artiklis analüüsitakse leedu hoidjakeelele omaseid morfoloogilisi jooni ja nende seost lapse vanusega.

Hoidjakeelel on väga oluline roll lapse arengus. Lastega rääkides lihtsustavad täiskasvanud oma keelekasutust ja kohandavad seda vastavalt keeleomandamise perioodile. Hoidjakeeles kasutatavaid sõnu ja konstruktsioone omandavad lapsed varem.

Emal ja tütre dialoogi analüüs lapse eri keeleomandamisperioodidel (premorfoloogiline, protomorfoloogiline, modulaarmorfoloogiline periood) näitab, et hoidjakeele morfoloogilised erijooned tulenevad sellistest keele funktsioonidest nagu kognitiivne ja emotiivne funktsioon, tähistus-, kasvatus-, lihtsustusfunktsioon.

Sagedamateks morfoloogilisteks modifikatsioonideks hoidjakeeles on: 1) verbide suurem esinemissagedus substantiividega võrreldes, 2) indefiniit- ja demonstratiivpronoomenite ning adjektiivide madal esinemissagedus, 3) deminutiivide ja interjektsioonide suhteliselt kõrge sagedus. Täistähenduslike sõnade domineerimine tuleneb kognitiivsest funktsioonist, verbide rohkusest lausete lühidusest. Deminutiivide sage kasutamine seostub pragmaatilise funktsiooni, intiimsuse ja emotsionaalsusega.

Osa hoidjakeele erijooni ei sõltu lapse keeleomandamisperioodist, osa sõltub. Esimeste hulka kuulub deminutiivide kasutamine, teiste hulka teatud sõnaliikide eelistamine.

Võtmesõnad: keele omandamine, longitudinaalne korpus, premorfoloogia, protomorfoloogia, modulaarmorfoloogia, hoidjakeel, sisendkeel, leedu keel

VÕIMU KEEL: ASUTUS SUHTLEB KODANIKUGA

Reet Kasik

Ülevaade. Vaatlen artiklis ametikirjade koostamise tekstilisi põhimõtteid ja analüüsin selle taustal üht konkreetset ametiasutusest kodanikule saadetud kirja. Jälgin kirjas avalduvaid võimuvahekordi ja arutlen valitud keelekasutuse sotsiaalsete põhjuste üle. Analüüsi teoreetiliseks raamistikuks on funktsionaalne grammatika ja arusaam tekstist kui keeleliste valikute kogumist (vt nt Halliday 1994), meetoodiliseks aluseks kriitiline tekstianalüüs (vt nt Fairclough 1995). Artikli eesmärgiks on osutada keele tähtsusele võimuvahekordade kujundamisel, säilitamisel ja muutmisel ning näidata, kuidas keel toetab mõnede inimeste (ja nende kaudu institutsioonide) domineerimist teiste üle.*

Võtmesõnad: kriitiline tekstianalüüs, tekstiliigid, ametikirjad, direktiivsus, eesti keel

Keel võimuvahendina

Kriitilise diskursuse- või tekstianalüüsi (kirjanduses kasutatakse mõlemat nimeust, aga lingvistina eelistan viimast) üksi eesmärke on näidata keelekasutuse ja ühiskondlike protsesside seost ning vastavalt sellele on analüüs kolmemõõtmeline (vt nt Fairclough 1997: 74–92):

- a) tekstitasandil hõlmab see teksti valitud keelenähtuste kirjeldamist;
- b) diskursusetasandil – teksti tootmise ja tarbimise kontekstis – muutub oluliseks keeleliste valikute tõlgendamine;
- c) sotsiokultuurilisel tasandil – ühiskonna sotsiokultuuriliste tavade ja harjumuste kontekstis – saab võimalikuks tõlgenduste ja tähenduste seletamine.

Võimu ja keele suhete kohta on palju oletusi, kuid Eestis on seda valdkonda uuritud vähe. Keelekasutus on kõige üldisem sotsiaalse käitumise vorm, ja selline sotsiaalse käitumise vorm, kus me oleme kõige rohkem sõltuvad nn tervemõistuslikest eelole-

* Artikkel on valminud ETF-i grandil nr 6236 osalisel toel.

tustest, s.t me peame teatud keelekasutustavasid loomulikuks, konventsionaalseks, nende valikut ja tähendust küsimärgistamata ja neile hinnangut andmata. Ma ei vaata selles artiklis mitte ainult seda, kuidas olemasolevad võimusuhted peegelduvad keeles, vaid ka vastupidi: kuidas keel toodab ja loob võimusuhteid. Keel on oluline võimukasutusvahend. Kriitilise lingvistika nime all tuntuks saanud tekstiuurimissuund seletab konventsioone kui võimuvahekordade ja võimuvõitluse tagajärge. See, kelle käes on võim, valib (ja järelikult määrab) suhtlusviisi. Eesmärgiks on selgitada, kuidas tekst konstrueerib tähendusi ja milliseid tähendusi uuritavasse teksti on konstrueeritud. Norman Fairclough' järgi tähendab analüüsi kriitilisus selle arvestamist, et sotsiaalne käitumine ja keelekasutusviisid on seotud põhjuse ja tagajärje suhetega. Keelekasutus sõltub sotsiokultuurilisest kontekstist ja samas mõjutab ning kujundab sedasama sotsiokultuurilist konteksti (Fairclough 1997: 74–78).

Keelekasutus on sotsiaalne tegevus ja keelekasutaja on selle protsessi osaline, sotsiaalne toimija, mitte lihtsalt sõnumi saatja. Teksti autor kasutab keelt oma eesmärkide saavutamiseks valikuliselt. Keeleliste valikutega – valides sõnu, lausestruktuure ja teksti retoorilisi süsteeme – loob autor teatud tunnustega ja teatud tähendusega teksti. Funktsionaalsest aspektist on teksti koostamine eesmärgipärane keelekasutus, keeleliste valikute tegemisel lähtutakse teksti eesmärgist ja funktsioonist. Kõige sagedamini eristatakse tekstitähenduste hulgas kolme tasandit või metafunktsiooni (Halliday 1994):

- a) tähistusfunktsioon keskendub sellele, kuidas keel peegeldab või konstrueerib keelevälist tegelikkust;
- b) suhtlusfunktsioon keskendub sellele, kuidas keel väljendab suhtumist tähistatavasse, kuidas konstrueeritakse suhtlusosaliste rolle ja identiteete;
- c) tekstiline funktsioon keskendub sellele, kuidas ja milliste keelevahenditega saadakse tulemuseks teatud tunnustega sidus terviktekst, mis seatud eesmäärke kõige paremini täidab.

Lisaks maailma nimetamisele, kujutamisele ja teksti konstrueerimisele loob ning kinnistab keel ka sotsiaalseid suhteid, rolle ja identiteete. Keele suhtlusfunktsioon ehk interpersonaalne metafunktsioon tähendab seda, et tekstitootja kasutab keelt kui vahendit selleks, et esitada oma hinnanguid, suhtumisi ja kommentaare ning luua suhe enda ja lugeja vahele. Suhtlustähendus väljendab teksti autori ja lugeja rolle suhtluses ning nendevahelist suhet. Kirjutaja konstrueerib teksti mingilt positsioonilt ja kirjutab ennast seega vastavasse diskursuserolli. Samal ajal annab kirjutaja ka lugejale vastuvõtjana teatud rolli. N. Fairclough' (1997: 17) järgi on interpersonaalsete rollide puhul küsimus sotsiaalsetest suhetest, võimust ning identiteetidest. Michael A. K. Halliday (1994: 68) eristab kahte põhilist diskursuserolli, milleks on andmine ja nõudmine. Kirjutaja kas annab midagi lugejale (nt informatsiooni) või ta nõuab temalt midagi (verbaalset või tegevuslikku reageeringut). Vastavalt sellele moodustub automaatselt ka teksti vastuvõtja roll. Kui autor on andja, siis on teksti vastuvõtja saaja, kui aga autor esineb nõudjana, siis on vastuvõtjale määratud andja roll.

Kui institutsioon saadab isikule kirja, võib sel ülaltoodust lähtudes olla kolm eesmärki: informeerida kirja saajat millestki, küsida temalt mingeid andmeid või veenda saajat toimima teatud viisil, reageerima mingi tegevusega. Keskendun siin viimasele, s.t vaatlen, kuidas esitatakse kirja saajale sõnum selle kohta, mida vastuvõtja peab või peaks tegema. Ametikirjade uurijad eristavad kirjades tavaliselt

kaks osa: direktiivne osa, kus veendakse kirja saajat midagi tegema, ja taustaosa, kus seda soovi põhjendatakse. Vastavalt kõneaktide teooriast pärit klassikalisele definitsioonile (Searle 1969) on direktiivne kõneakt selline, mille taotlus on panna keegi midagi tegema. Direktiivsus kuulub vastuvõtjale suunatud modaalsuse alla. Direktiivsel verbaalsel tegevusel on eesti keeles mitmeid võimalikke nimetusi: palve, soov, õhutamise, käskimine, suunamine, veenmine, keelamine jne, aga eri nimetustega seostub erinev mentaalne, emotsionaalne lisavarjund, mis hõlmab ka soovi edasiandmise viisi. John Searle (1969) on osutanud, et direktiivsete sõnade tähendus on seotud autoriteetsusega, näiteks käskimisega seostub ütleja autori-teetsus, seevastu palumisel selline tähendus puudub.

Demonstreerin alguses ühe näite varal, kuidas suhtlusrollide konstrueerimine toimub vahetu kontakti puhul, dialoogilises suhtluses. Näitetekst on pärit N. Fairclough' raamatust "Keel ja võim" (1989) ja kujutab endast katkendit ülekuulamisest politseijaoskonnas. Näide sobib teema illustreerimiseks, kuna ülekuulamisega ja politseiga on tegemist ka siinses artiklis järgnevalt analüüsitava kirjaliku teksti puhul.

Näitetekst on osa dialoogist relvastatud röövi pealtnägija ja politseiniku vahel, kus käib põhilise informatsiooni väljaselgitamine. Pealtnägijalt (n), kes on juhtunust vapustatud, küsitakse juhtunu kohta, politseinik (p) kirjutab väljaselgitatud informatsiooni üles.

- (1) p: Kas te märkasite autos istuvat meest?
- (2) n: Ma nägin ta nägu, jah
- (3) p: Kui vana ta oli?
- (4) n: Umbes 45. Seljas oli tal...
- (5) p: Ja kui pikk?
- (6) n: Meeter kaheksakümmend.
- (7) p: Meeter kaheksakümmend. Juuksed?
- (8) n: Tumedad ja lokkis. Kas see võtab veel kaua? Ma pean minema laste-aeda lastele järele.
- (9) p: Mitte enam kaua. Mismoodi riides?
- (10) n: Ta jättis turjaka mulje, sinised püksid, must...
- (11) p: Teksased?
- (12) n: Jah.

Iseloomustades küsitleva politseiniku ja pealtnägija suhteid kõnealusel juhul, osutab N. Fairclough, et suhe on ebavõrdne, politseinik kontrollib kindlalt ülekuulamise käiku ja juhib küsitletavat, ega võta vaevaks pehmed nõudmisi, mis ta ülekuulatavale esitab. Need küsimused võivad olla küllaltki piinarikkad kellelegi, kes on just pealt näinud vägivaldset kuritegu; p küsimus kõnevoorus (1) näiteks võiks olla esitatud pehmentatult nagu *Kas teil oli mingit võimalust pöörata tähelepanu sellele mehele selle järsu vormi asemel, milles see tegelikult esineb*. Mõnel juhul on küsimused taandatud pelgaks üksiksõnaks või lühifraasiks: *kui pikk* voorus (5) ja *juuksed* voorus (7). Sellised redutseeritud küsimused on tüüpilised, kui üks isik täidab formulari teise isiku jaoks, nagu p siin; huvitav on, et situatsiooni tundlik iseloom ei kõiguta blanketitaitmise norme. Ka on märkimisväärne, et politseinik ei anna mingit tagasisidet, näiteks tunnustust või tänu informatsiooni eest, mida ta tunnistajalt saab. Kõnekas on ka viis, kuidas küsitleva reageerib infole,

mida tunnistaja omaalgatuslikult pakub voores (4) ja voores (10). Kaasabi andmist kontrollitakse: p katkestab vastaja kõnevooru 5. ja 11. voorega, ja voores (9) annab minimaalselt lakoonilise vastuse küsimusele, kui kaua see ülekuulamine veel kestab – tunnustamata ülekuulatava probleemi, esitab ta vahetult järgmise küsimuse, lõigates ära tunnistaja võimaluse lisaküsimust esitada.

N. Fairclough rõhutab, et need omadused on sotsiaalsete tingimustega määratud: selline keelekasutus on tingitud politsei ja “üldsuse” vaheliste suhete iseloomust ühiskonnas ja tegelikult kuulub ühe osana neisse suhetesse. Ülekuulatav pealtnägija näib sellist küsitlusstiili aktsepteerivat ja tundub olevat tõepoolest olukorraga täiesti kohanenud tunnistaja. Ta võtab ise passiivselt sisse positsiooni, mis on n-ö konstrueeritud “üldsuse” liikmete jaoks vestluses politseiga. Kõnekas on, kuidas tunnistaja tõlgendab igasuguse politseinikupoolse tänulikkuse puudumist pakutud informatsiooni eest. Kui midagi samasugust juhtuks tavavestluses, siis häiriks see kindlasti suhtlust. Vestluspartner võib tõlgendada sellist keelelist käitumist viitena, et teda kahtlustatakse või tahetakse segadusse ajada. Ta peaaegu kindlasti reageeriks sellele ja jutuajamise tavatu iseloom peegelduks nii või teisiti teksti tunnustes (sealt võiks leida näiteks ootamatuid pause või muid ebakindluse märke). Politseiülekuulamistel arvatavasti tänulikkust eriti ei väljendata, mistõttu ülekuulatav ei tunneta selle puudumist probleemseks, vaid eeldab iseenesestmõistetavana seda tüüpi ülekuulamistel just niisugust tooni. Selline diskursusetüüp on loomulikustunud ja kumbki osaline ei märka seda tegelikult, vaid aktsepteerib sellist vestlusstiili kui enesestmõistetavat. See, kuidas inimesed interpreteerivad teksti tunnuseid, sõltub sellest, milliseid sotsiaalseid (täpsemalt: diskursuse) konventsioone ja kokkuleppeid nad eeldavad valitsevat. Näide illustreerib üht kriitilise lingvistika põhiväidet – et sotsiaalsed tingimused määravad diskursuse omadused.

Kirja direktiivne osa – kutse või korraldus

Järgnevalt analüüsitav kiri on osa sündmusest, mille osaliseks sattusin tahtmatult. Mõistmiseks vajalik situatsioonikontekst on järgmine. Piirkonnas, kus ma elan, juhtus intsident, kuhu sekkus politsei. Meie maja elanikega polnud sel mingit seost, aga meile öeldi, et politseijaoskonnas on tarvis fikseerida meie kinnitused, et keegi meist pole kutsunud süüdistatavat sinna, kus ta ei oleks tohtinud olla. Mõne aja pärast helistaski mulle naisinspektor, pakkus välja aja, mil ta saaks mind vastu võtta ja küsis, kas see aeg sobib mulle. Vestlus oli täiesti viisakas, sõbralik ja mitteametlik. Muuseas teatas ta, et kuna ta mind esimesel helistamisel kätte ei saanud, siis pani ta igaks juhuks posti ka kirjaliku kutse. Järgmisel päeval tõi post mulle selle kirja.

Kiri oli järgmine:

Nimi

Aadress

KUTSE

*Olete kohustatud ilmuma kriminaalasjas nr XXX **tunnistajana** ülekuulamisele 09.02.2006 kell 09.00 XXX Politseiprefektuuri kriminaalosakonna*

XXX kriminaalpolitseieeskirjast raskete kuritegude grupi politseijuhtivinspektor NN juurde aadressil XXX, kabinet XXX (tel XXX).

Palun kaasa võtta isikut tõendav dokument.

Kui teil ei ole võimalik määratud ajaks ilmuda, olete vastavalt KrMS § 170 kohustatud sellest viivitamata teatama ning nimetama ka ilmumata jäämise põhjuse. Menetleja kutsel ilmumata jäänud väljakutsutud trahvib eeluurimiskohtuniku prokuratuuri taotlusel või kohus omal algatusel lähtudes KrMS § 138 kohtumääruse alusel kuni kaheksa miinimumpäevapalgamäära ulatuses või kohaldab talle kuni neliteist päevaaresti. Kui ilmumata jääb alaealine süüdistatav või kahtlustatav, trahvib kohus tema vanemat, eestkostjat või hooldajat. Menetleja kutsel ilmumata jäänud väljakutsutud kahtlustatava, süüdistatava, süüdimõistetud, kannatanu või tunnistaja suhtes võib kohaldada sundtoomist või kuulutada ta tagaotsitavaks. Vastavalt KrMS § 139 võib sundtoomisele allutatud isikut kinni pidada nii kaua, kui see on vajalik sundtoomise aluseks oleva menetlustoimingu tegemiseks, kuid mitte kauem kui 48 tundi.

Lugupidamisega

NN

Juhtivinspektor.

Kirja vorm on neutraalselt funktsionaalne. Pealkirjaks on KUTSE ja allkiri sisaldab viisakat tervitusvormelit ning omakäelist allkirja *Lugupidamisega NN, juhtivinspektor*. Selline raamistik eeldaks viisakat ja lugupidavat keelekasutust läbi kogu kirja. Tegelikult konstrueeritakse lõpetava tervitusvormeli ja allkirjaga viisakam suhe kui kirja sisuga. Esimesena torkab silma, et tüüpilist algusvormelit, (nimelist) pöördumist kirja saaja poole, ei ole. Kirjas on vaid üks viisakusväljendit sisaldav lause – *Palun kaasa võtta isikut tõendav dokument*, kuid see mõjub kirja üldise fooni taustal võõrkehana – viisakas palve range käsu ja ähvarduse vahel.

Direktiivsuse prototüüpne edastamisvorm on imperatiiv. Kuna see on käskiva kõneviisi kongruentne e põhitähendus, siis selline käsu vahendamine on suhteliselt neutraalne: *Tulge kolmapäeval, 9. veebruaril kell 9 sellel ja sellel aadressil*. Praktiliselt on imperatiiv aga peamiselt suulises otsesuhtluses kasutatav keelevorm, kirjalikes tekstides on tema põhiline kasutusvaldkond reklaamikeel (vt Kasik 2000).

“Palumise” ja “käskimise” vahetamine näib sõltuvat protsessiosaliste rollidest. Imperatiivi kasutamine viitab, et käsu täitmine on vastuvõtja, mitte saatja huvides. Näiteks on tähele pandud, et reklaamides tarvitatakse ohtralt imperatiive, aga viisakusvormelid saadavad neid harva (Leech 1966). Seletuseks on pakutud nimelt seda, et meie kultuuris ei ole kombeks kasutada viisakusvormeleid seal, kus käsk on pigem kuulaja kui rääkija huvides: *astu sisse, võta istet, võta veel küpsist*.

Seega on viisakuskeelendi puudumine üheks viisiks, kuidas reklaamijad püüavad esitada oma käsku nii, nagu oleks nende täitmine vastuvõtja huvides. Kui reklaamija ütleks *Palun ostke meie šampooni* või *Kas te võiksite anda oma juuste X-palsamit*, oleks selles vormis implitsiitselt sees tähendus, et soovi täitmine on reklaamija, s.t paluja huvides, et ostja teeb reklaamijale teene, kui ostab nende

toodet. Kuigi enamik käske on mõttekad vaid siis, kui nad on rääkija või kirjutaja huvides, püüavad reklaamid esitada neid sellises vormis, nagu oleks neist kasu eeskätt lugejale/kuulajale (Myers 2001: 48).

Oluline on ka reklaamija enesekindlus, mida püütakse tugevdada süntaktiliste funktsioonidega, esitades asju nagu kindlaid väiteid, mille puhul ütleja on kindel, et sellest on kuulajale/lugejale suurt kasu. Seepärast kasutatakse ka üleskutsete puhul kõige selgemat vormi. *Tule ja proovi!* on enesekindel kutse, seevastu *Ehk oleks teil võimalik läbi astuda ja proovida* tunduks vastuvõtjale ilmselt ebausaldusväärne ja kahtlane. Reklaamijal tegelikku võimu vastuvõtja üle ei ole, seepärast peab ta keelevahenditega osutama, et jutt on tarbija kasust.

Asutuse ja kodaniku suhted on teistsugused. Institutsioon esindab võimu ja tema palved on tõlgendatavad korraldustena, mille täitmist kirja saajalt eeldatakse. Sellegipoolest on demokraatlikus ja tsiviliseeritud ühiskonnas vaikimisi omaks võetud tava, et korraldusedki esitatakse viisakas vormis palvena. Ametikirjades kasutatakse käskivat kõneviisi minimaalselt juhtudel, kui tahetakse viidata, et mingi paberi täitmisest on kliendile kasu ja täitmata jätmisest kahju (Julkunen 2002: 119). Enamasti esitatakse kirja saatja soov palvena: *Palun tulge...* või *Palume teil tulla...*

Viisakusstrateegiasse kuulub nii *paluma*-verb kui ka saaja nimetamine palvega seoses (*teil*). Süntaktiliste laiendite kasutamine *paluma*-verbiga seoses rõhutab palumise leksikaalset, verbaalse protsessi tähendust – meie palume teid või mina palun teid midagi teha. Kui *palun* esineb teksti alguses üksi – *palun makske kassasse* või *palun maksta kassasse*, siis on protsessitähendus taandunud ja *palun* on lihtsalt formaalne, lauseadverbiaalina esinev grammatikaliseerunud vormel. *Makske kassasse* või *maksta kassasse* on iseseisev käsklause, mitte *paluma*-verbi laiend. Seejuures on *da*-infinitiivi vormis esitatud käsk hoopis tugevama kaman-damise varjundiga kui imperatiivvormis, sest imperatiiv pöördelise vormina on siiski isiklik pöördumine adressaadi poole, *da*-infiniitiv on umbisikuline ega viita korralduse saajale. Seega võib *paluma*-verbiga seotud viisakusvormeleid vaadata hierarhiliselt: siin vaadeldutest kõige viisakam on *palume teil teha X*, seejärel *palun tehke X* ning kõige formaalsem *palun teha X*. On ka veel viisakamaid malle, kui palumine esitatakse nominaliseeritult: *pöördun teie poole palvega*, *mul on teile palve* vms. Analüüsitava kirja keskel olev viisakas vormis, kuid ilma isikuviieteta lause *Palun kaasa võtta isikut tõendav dokument* sisaldab siis õigupoolest minimaalset võimalikku viisakust.

Selles kontekstis tekitab politseijaoskonnast saabunud kiri selle saajas segadust, sest ta on ülesehituselt vastuoluline. Kiri on osaliselt teises isikus pöördumine, osaliselt aga kolmandas isikus umbisikuline teade. Viisakusstrateegiale räägib vastu isikliku pöördumisvormeli puudumine, kuigi kiri kõnetab saajat *teie*-vormis (*Olete kohustatud...*, *Kui teil ei ole võimalik...*). Kirja saatja ei esine kirja tekstis suhtlusosalise rollis: seal pole ühtki *meie*- või *mina*-pronoomenit ega esimeses isikus verbivormi (nt *Ootan teid...*). Kuigi kirja saatja on seesama inspektor ise, kes alla kirjutab *lugupidamisega NN*, ei kirjuta ta enda nimel, vaid on kirja tekstis ennast nimetanud kolmandas isikus, *olete kohustatud ilmuma inspektor NN-i juurde...* Direktiivseks verbiks on *kohustama*, mis eesti keele direktiivseid verbe analüüsitud Heili Orava semantilise analüüsi järgi tähendab 'autoritaarselt, oma staatusest lähtudes kedagi tema tahte vastaselt midagi tegema sundima' (Orav 1998: 64). Käsk

on esitatud saaja rolli fokuseerivas passiivses vormis: *olete kohustatud ilmuma*. Impersonaaliga kaasnev saatja rolli (kes kohustab) taandamine muudab teksti funktsiooni: kirja saatja verbaalse tegevuse asemel representeerib lause vastuvõtja seisundi. Ametlikkust ja direktiivsust rõhutab *ilmuma*-verb (vrd tavateksti *tulema* vms). Ja ometi ei ole kirja nimetuseks *korraldus*, nagu sellise teksti puhul oleks ootuspärane, vaid *kutse*. Kirjas endas ühtki keelelist viidet kutsumisele ei ole.

Kui püüda tõlgendada, missuguseid tähendusi selline sõnastus kirja saajale vahendab, siis kirja vormistusviisi ja kasutatud keelevahendid konstrueerivad eelolevuse, et kirja saaja ei taha kirjas väljendatud soovi täita, et kirja saatnud asutusega suhtlemine on kirja saajale vastumeelne ja teda tuleb selleks sundida. Ja et kirja saatjal on sundimiseks piisavalt autoriteeti. Kohustava lause toon ja direktiivses osas sisalduv oletus sellest, kuidas kirja saaja tegutseb või tahab tegutseda, saab kinnitust ka kirja direktiivsele osale järgneva taustaosa analüüsist, mille eesmärk on demonstreerida, et kirja saatjal tõepoolest on võimu ja võimalusi saajat sundida.

Kirja taustaosa – põhjendus või ähvardus

Tahteavaldusega ehk direktiivse osaga liituva taustaosa funktsioon tegevusele õhutatavates tekstiliikides on põhjendada, miks esitatud soovi või palve või korralduse täitmine on vajalik või mõistlik. Reklaamides näiteks on see tavaliselt materiaalne või emotsionaalne isiklik kasu, mis toote omandamisega kaasneb (rahaline sääst, mugav elu, hea tervis, õnn ja armastus) või kahju, mis toote puudumisest tuleneb (raha kulub rohkem, tervis läheb käest ära, kallim ei pane tähele vms). Ametikirjade uurijad väljaspool Eestit on tõdenud, et direktiivsetes kirjades esitatud põhjendused on ratsionaalsed ja informatiivsed, sisaldades viiteid dokumentidele, faktidele, statistilistele andmetele jne, mille eesmärk on tõestada, et esitatud soovi või palve täitmine on arukas. Teiseks on tõdetud, et viiteid negatiivsetele sanktsioonidele, mis soovi täitmata jätmisele järgnevad, esineb ametikirjades suhteliselt harva. Sellisele võimalusele, et esitatud juhtnööre ei täideta, asutuste kirjad üldjuhul ei viita. (Vt nt Honkanen 2002: 144–147.) Ka Eesti ametikirju analüüsinud Egle Pullerits on oma bakalaureusetöös (Pullerits 2005) selgitanud, et kirjade taustaosas esitatakse argumente, põhjendusi, selgitusi, viiteid ja olukorrakirjeldusi, mis on vajalikud direktiivse osa mõistmiseks.

Selle teabe taustal on siin analüüsitava kiri täiesti erandlik. Tekstis ei ole ühtegi ratsionaalset ega emotsionaalset põhjendust ega selgitust, miks kirja saaja peaks esitatud soovi täitma. Kogu taustaosa on negatiivse teabega, koosneb ähvardusest ja sanktsioonide loetelust, mis järgneb korralduse täitmata jätmisele. Sarnaselt direktiivse osaga, kus üldisest kamandavast toonist langeb välja üks palvelause (*Palun kaasa võtta isikut tõendav dokument*), on ka taustaosas üks lauseosa, mis pöörduv kirja saaja poole ja viitab saaja olukorra arvestamisele (*Kui teil ei ole võimalik määratud ajaks ilmuda*). Viisakusstrateegia seisukohast on seegi lauseosa vastuoluline. Algus kujutab endast viisakat mööndust (*Kui teil ei ole võimalik*), millele ootaks samas stiilis järke, näiteks *Kui teil ei ole võimalik sel ajal tulla, siis palun helistage sellel ja sellel numbril, et leppida kokku teile sobivam aeg*, selle asemel aga rõhutab väljend *määratud ajaks* jällegi saatja võimupositsiooni ja viitab korraldusele. Ka saava käände kasutamine on tõlgendatav pigem oletusena

võimalikust hilinemisest kui tulemata jätmisest. Järgneb juba kirja algusosast tuttav kamandamisvormel *olete kohustatud sellest viivitamata teatama*, kus käskija võimupositsiooni rõhutab eriliselt ajamäärus *viivitamata*.

Järgnev sanktsioonide loetelu muudab taas kirja saaja ja saatja rollisuhteid. See pole suunatud keeleliselt üldse kirja saajale, vaid viitab referatiivselt ja üldistavalt seadusesätetele, nimetades kirja saajat indefiniitselt *väljakutsutu*-ks või *tema*-ks, kelle suhtluspartneriks ja sanktsioonide rakendajaks pole kirja saatja, vaid eeluurimiskohtunik, kohus või umbisikuline, nimetamata võim (*.. ilmumata jäänud väljakutsutu suhtes võib kohaldada sundtoomist või kuulutada ta tagaotsitavaks*). Võimu üleolevat ja kodanikke ühtse, individualiseerimata alluvate massina käsitavat suhtumist väljendab ka rinnastav loetelu, kus kirja saajale viitav *väljakutsutu* seletatakse lahti kui *kahtlustatav, süüdistatav, süüdimõistetud, kannatanu või tunnistaja*, järgmises lauses viidatakse talle väljendiga *sundtoomisele allutatud isik* – kuigi kirja alguses on kutsutu identifitseeritud üheselt kui tunnistaja. Sellist võrdsustatud liikmetega loetelu saab tõlgendada vaid nii, et kirja saatjale on täiesti ükskõik, kas ta suhtleb kannatanu või süüdistatavaga.

Niisiis on selle kirja saatnud riigi ametiasutus sellesama riigi süütule kodanikule, keda ta ei palu, vaid käsib ja kellele ta oma soovi ei põhjenda, vaid ähvardab, mis kodanikuga juhtub, kui too käsku ei täida. Tekstivahendite valik konstrueerib võimusuhted väga selgelt: siin on ametiasutus, kes ei püüagi oma võimupositsiooni tavapärase keelelise viisakusega maskeerida ja kes lähtub eeldusest, et inimesi sunnib seaduskuulekalt käituma üksnes hirm karistuse ees. Miks ametiasutus nii teeb? Reapolitseinikke ju õpetatakse kodanikega viisakalt käituma, miks pole see laienenud politsei ja kohtu kui institutsiooni kirjalikule suhtlemisele kodanikega? Näen siin kahte võimalikku seletust. Üks on selline, et politsei näol on meil institutsioon, kus on käibele jäänud nõukogulikud suhtlusmallid. Nõukogude riigis oli selline kinnistunud võimusuhte taastootmine ja demonstratsioon mäletatavasti iseloomulik kõigile ametiasutustele. Enamik ametiasutusi on oma suhtlustavasid muutnud demokraatlikule ühiskonnale sobivamaks: viisakaks ja neutraalseks. Politsei ja kohus pole selleni veel jõudnud.

Teine võimalik seletus on, et selline suhtlusviis on teadlikult valitud. Võim peab ennast taastootma ja kasutatavate keeleliste valikutega kinnistatakse ja loomulikustatakse üha uuesti sotsiaalseid rolle ja identiteete, mis kummalegi suhtlusosalisele on võimu poolt määratud. Politsei ja kohus institutsioonina on end oma keelekasutusega distantseerinud teistest riigiasutustest ja demonstreerib oma võimupositsiooni, välistades võimaluse seda vaidlustada. Et suhtlusesmärkide kohta kaugemale minevaid järeldusi teha, peaks võrdlema, missugune on tekstiliselt politsei ja kohtu suhtlusstrateegia kodanikega teistes riikides, nii demokraatlikes kui ka vähem demokraatlikes.

Nagu artikli alguses viidatud, on sellised sotsiaalsed suhted võimalikud vaid niikaua, kui me seda suhtluspartnerina aktsepteerime. Kuni meie ise võtame meile määratud identiteedi vastu leplikult, möönvalt, vabatahtlikult, niikaua peetakse sellise keelekasutusega üleval ja säilitatakse neid sotsiaalseid võimusuhteid, mis neid positsioone määravad. Ja vastupidi, kui valitsevatele konventsioonidele vastu hakatakse, kui neid ei võeta loomulikena, iseenesest mõistetavatena, vaid küsimärgistatakse, protestitakse ja nõutakse suhtlemistavade muutust, siis võib keelekasutus kaasa aidata ka sotsiaalsete suhete muutmisele. Praeguse keelekasu-

tusega konstrueeritav suhe on võim ja hirm võimu ees. Kui keelekasutust muuta, võib-olla siis seaduskuulekas kodanik ei peagi demokraatlikus riigis politseid ja kohut kartma.

Kirjandus

- Fairclough, Norman 1989. *Language and Power*. London: Longman.
- Fairclough, Norman 1995. *Critical Discourse Analysis*. London: Longman.
- Fairclough, Norman 1997. *Miten media puhuu*. Tampere: Vastapaino.
- Halliday, Michael A. K. 1994. *An Introduction to Functional Grammar*. London: Arnold.
- Honkanen, Suvi 2002. Päämäärien valtakunta. Välttämättömyys ja toiminta viraston ohjailevissa kirjeissa. – Vesa Heikkinen (toim.). *Virkapukuinen kieli*. Tietolipas 182. Helsinki: Suomalaisen Kirjallisuuden Seura, 137–166.
- Julkunen, Verna 2002. Ystävällisesti Kansaneläkelaitos. – Vesa Heikkinen (toim.). *Virkapukuinen kieli*. Tietolipas 182. Helsinki: Suomalaisen Kirjallisuuden Seura, 107–122.
- Kasik, Reet 2000. Reklamikeel tekstiliigina. – Tiit Hennoste (toim.). *Eesti keele allkeeled*. Tartu Ülikooli eesti keele õppetooli toimetised 16. Tartu: Tartu Ülikooli Kirjastus, 111–127.
- Leech, Geoffrey 1966. *English in Advertising. A Linguistic Study of Advertising in Great Britain*. London: Longman.
- Myers, Greg 2001. *Words in Ads*. London: Arnold.
- Orav, Heili 1998. Eesti keele direktiivverbide semantilise välja struktuur tesaurusena. Magistritöö üldkeeleteaduse alal. Käsikiri Tartu Ülikooli eesti keele õppetoolis.
- Pullerits, Egle 2005. Ametikiri kui tekst: struktuur ja keelekasutus. Bakalaureusetöö Tartu Ülikooli eesti keele õppetoolis.
- Searle, John R. 1969. *Speech Acts*. Cambridge University Press.

Reet Kasiku (Tartu Ülikool) teaduslikeks huvialadeks on eesti keele sõnamoodustus, tekstianalüüs ja tekstiliigid (ajakirjanduskeel, reklaamikeel).
reet.kasik@ut.ee

LANGUAGE OF POWER

Reet Kasik

University of Tartu

A concrete letter, sent from a public office to an Estonian citizen, is analysed against the background of the textual principles of writing an official letter. The power relations reflected in the text have inspired a discussion on the possible social reasons behind such language use. The theoretical framework of the analysis relies on functional grammar and the idea of text as a set of linguistic choices, while the methodical basis of the study is provided by critical text analysis and critical discourse analysis. The focus lies on the means used in an official letter to express and enhance directivity and the discursive roles construed by the text for the sender and the addressee. The aim is to point out the essential role of language in the designing, maintaining and changing power relations, and to show how language may support the domination of some people and, via those people, of some institutions over some other people.

Keywords: critical text analysis, genres, official letters, directivity, Estonian

AMETIKEELE OSKUSE MÕÕTMINE: TEKST JA ÜLESANNE EESTI KEELE KUI TEISE KEELE KUULAMISTESTIS

Tiina Kikerpill, Ülle Türk

Ülevaade. Ehkki eesti keele kui teise keele oskuse mõõtmisega on tegeldud juba kümmekond aastat, on peamiselt koostatud üldkeeleoskust mõõtvaid teste. Ametikeele oskuse mõõtmine on paraku jäänud tähelepanu alt välja. Seetõttu pakkus EL Phare projekt “Eesti keele oskuse arendamine mitte-eesti õppekeelega kutseõppeasutustes” tänuväärse võimaluse luua kuulamistestid kutsekoolide eri erialade õppijatele. Artiklis käsitletakse nende testide koostamist. Vaadeldakse keeleoskuse eri mudeleid ning eriotstarbelise keeleoskuse eripära, kuulamisoskuse aspekte, suuliste tekstide eripära ja kirjalike tekstide suuliseks kohandamise põhimõtteid, ülesandetüüpide valikuprintsiipe ning kuulamistestide koostamise protsessi. Kuulamistestide koostamise teeb eriti raskeks sobivate tekstide vähesus, mistõttu suur osa ajast kulub tekstide kohandamisele või koostamisele. Ülesannete koostamisel tuleb silmas pidada seda, kuidas vastavat teksti tavaelus kuulatakse, ja ka seda, et kasutatakse erinevaid ülesandetüüpe.

Võtmesõnad: eesti keel teise keelena, keeletestid, kuulamisoskuse arendamine ja mõõtmine

Eestis on eesti keele kui teise keele oskuse mõõtmisega tegeldud üle kümne aasta, kuid nii eesti keele tasemeeksamite kui ka põhikooli lõpueksami ja gümnaasiumi riigieksami puhul mõõdetakse üldkeeleoskust. Samas tähendab eesti keele riigikeele staatus, et mitte-eestlased peavad suutma eesti keeles ka tööalaselt suhelda. Vene õppekeelega kutsekoolides õpetatakse küll erialaspetsiifilist eesti keelt, kuid teste selle oskuse mõõtmiseks ei ole loodud. Üheks katseks seda tühikut täita on EL Phare projekti “Eesti keele oskuse arendamine mitte-eesti õppekeelega kutseõppeasutustes” käigus aastatel 2004 ja 2005 loodud kuulamisoskuse mõõtmise testid kümnele erialale.

Käesolevas artiklis mõtestatakse projekti raames omandatud kuulamistestide koostamise kogemust. Vaadeldakse kuulamistestide koostamise protsessi alates konstrukti määratlemisest ning sobivate tekstide leidmisest või tekstide sobivaks kohandamisest ülesandetüüpide valimise ja ülesannete koostamiseni välja. Kuna eesmärgiks on käsitleda kuulamistestide koostamise problemaatikat, siis testide läbiviimisel ja tulemuste analüüsil peatume vaid põgusalt.

Lähteülesanne

Projekti “Eesti keele oskuse arendamine mitte-eesti õppekeelega kutseõppeasutustes” eesmärgiks oli anda vene õppekeelega kutseõppeasutuste õpilastele ja õpetajatele eestikeelses töö- ja õpikeskkonnas viibimise kogemus, mille kaudu sooviti muuhulgas parandada nende eesti keele oskust. Eesti õppekeelega kutseõppeasutustes käis 2004/2005. õppeaastal lähetusel 100 vene õppekeelega kutsekoolide õpilast 10 erialalt. Et välja selgitada, mil määral nende eesti keele oskus kolme nädala jooksul tegelikult paranes, oli projektis ette nähtud testida õpilaste eesti keele oskust enne lähetusele minekut ja praktika lõppedes.

Eesti keele oskuse mõõtmise testi koostamist enda peale võttes andsime endale täielikult aru, et keeleoskuse omandamine on pikaajaline, pidevat tööd nõudev protsess ja on äärmiselt ebatõenäoline, et ilma intensiivse keeleõppeta (ja seda projekt ette ei näinud) toimub kolme nädala jooksul õppijate keeleoskuse mõõdetav paranemine. Hästi läbi mõeldud ja põhjalikult ette valmistatud lühiajaline keelekeskkonnas viibimine lisab õpilastele kindlasti julgust õpitavat keelt kasutada ja aitab paremini mõista sihtkultuuri, kuid ei tõsta märgatavalt õpilaste keeleoskuse taset.

Kõige valiidsem olnuks õpilaste keeleoskuse mõõtmiseks kasutada intervjuusid, sest eeldatavasti mõjutab keelekeskkonnas viibimine kõige enam just suhtlemise oskust ja valmidust osaleda vestluses. Paraku on intervjuude usaldusväärne hindamine keeruline, samuti on intervjuude läbiviimine ajamahukas. Seetõttu otsustasime mõõta õpilaste kuulamisoskust, sest eesti keeles loengute kuulamine ja vestluses osalemine peaksid ka seda oskust arendama.

Kuna eesmärgiks oli mõõta projektis osalenute keeleoskuse paranemist, oli kuulamistestide koostamiseks vaja ligilähedaselt teada testitavate keeleoskuse taset. Liiga kerge või liiga raske test ei anna aga õppija keeleoskusest adekvaatset ettekujutust: liiga kerge testi puhul oleksid õpilased juba enne lähetust sooritanud testi maksimaalsele tulemusele, liiga raske testi puhul aga oleks tulemus nii madal, et õigeid vastuseid võiks pidada juhuslikkuseks.

Paraku oli meil õpilaste eesti keele oskuse kohta liiga vähe andmeid, et täie kindlusega otsustada, millisele tasemele peaks test vastama. Kõik õpilased olid küll täitnud ankeedi, kuhu nad muuhulgas märkisid ka, milline eesti keele eksam on sooritatud ja kui heaks nad ise oma eesti keele oskust hindavad, kuid eksamid olid enamasti sooritatud rohkem kui aasta tagasi ja meil ei olnud võimalik kontrollida, kuivõrd usaldusväärne oli õpilaste enesehinnang. Hoolimata sellest, kas sooritatud oli algtaseme (A2/B1) või kesktaseme (B2) eesti keele eksam, märkis enamik küsitletutest, et saavad eesti keeles raskustega hakkama. Seetõttu otsustasime, et meie koostatavad kuulamistestid peaksid olema veidi madalamal tasemel kui

eesti keele kesktaseme eksam – vastama Euroopa Nõukogu keeleoskustasemetete süsteemi B1/B1+ tasemele.

Euroopa keeleõppe raamdokumendi järgi (vt Hausenberg jt 2004: 44–46, CEF 2001: 66–68):

- 1) mõistab B1 tasemel õppija nii monoloogilist kui dialoogilist lühemat kõnet tuttavatel teemadel: töö, kool, puhkus. Saab aru üldisest igapäevaelu või tööga seotud faktilisest teabest ja mõistab ka detaile, kui kõne on selge;
- 2) suudab B1+ tasemel õppija jälgida oma eriala loengut või kõnelust, kui teema on tuttav ja ettekanne selge struktuuriga. Raadio ja lindistuste kuulamisel mõistab suurema osa saadete põhisisu, kui esitus on selge ja teema isikule huvi pakkuv.

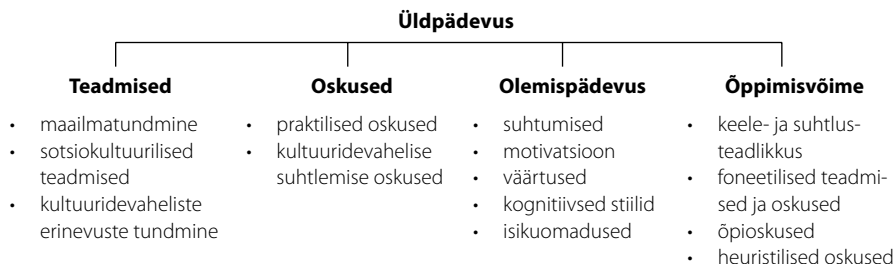
Eeldatud keelekasutusolukordi (õppetöös osalemine ja kaaslastega suhtlemine) silmas pidades koostasime kahest tekstist koosnevad testid, millest esimene oli üldkeelne tööteemaline vestlus kahe inimese vahel ja teine eriala mingile aspektile keskenduv lühiloeng. Esimese teksti juurde kuulus 9 kolme valikuga valikvastustega küsimust ja teise teksti juurde 6–18 eritüübilist küsimust. Kuna lähetuses osalenud õpilased sooritasid sama testi nii lähetuse alguses kui lõpus, koostasime kokku 10 testi.

Üldkeeleoskus ja eriotstarbeline keeleoskus

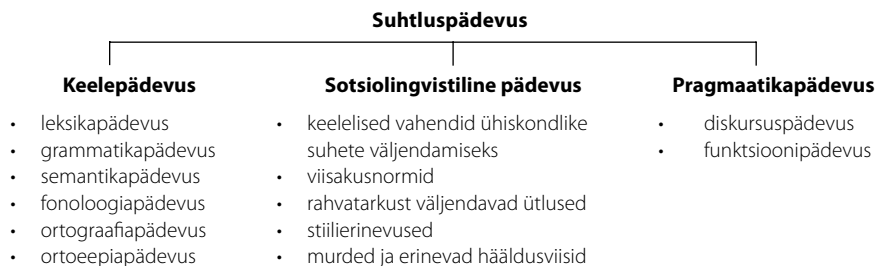
Et mingit oskust mõõta, tuleb see oskus kui mõõtmise objekt ehk konstrukt esmalt määratleda. Antud juhul oli esmaseks küsimuseks, kuivõrd erialaspetsiifilised peaksid koostatavad tekstid olema ja mida erialaspetsiifilisus kuulamisteksti puhul tähendab. Tihti keskendutakse erialakeele tundides terminoloogia omandamisele. Ometigi on ametikeele oskus märksa laiem, sest teatud ameti esindaja peab suutma tööalaselt suhelda nii ametikaaslaste kui ka klientidega. Esimesel juhul võidakse üksteisest aru saada ka ühist suhtluskeelt omamata, teisel juhul muutub aga tihti määravaks üldkeele oskus. Samuti on ametikeele oskust vaja mitte üksnes suhtlemiseks vaid ka oma valdkonnas toimuvaga kursis olemiseks. Teisisõnu on pelgalt terminoloogia tundmisest vähe abi – seda tuleb osata eesmärgipäraselt kasutada.

Seega on ametikeel kui eriotstarbeline keelekasutus (ingl *specific purpose language ability*) üks keeleoskuse variante ning selle määratlemine sõltub suuresti sellest, mida mõeldakse keeleoskuse all üldisemalt. Arusaamine keeleoskusest on aga aja jooksul muutunud. Kui esialgu mõisteti keeleoskust kui grammatika, sõnavara ja õige häälduse tundmist, siis hiljem on lisandunud keeleoskuse mudelisse ka osaoskused (kuulamine, lugemine, rääkimine, kirjutamine) (Lado 1961, Carroll 1961, mõlemad Hausenberg jt 2004: 8 järgi). 1972. aastal võttis Dell Hymes kasutusele mõiste *suhtluspädevus* (ingl *communicative competence*), milles ta eristas keelepädevust ja sotsiolingvistilist pädevust, ning rõhutas konteksti tähtsust keeleoskuses. D. Hymes'i järgi sisaldub suhtluspädevuses see, mida grammatika võimaldab (ehk mis on süsteemi poolest võimalik), mida meie mõistmine võimaldab (psühholingvistiline aspekt) ja mis on ühiskonnas lubatud (ehk mis on sotsiokultuuriliselt kohane). Teisisõnu tähendab suhtluspädevus keeleteadmisi ja võimet neid kasutada (Douglas 2000: 25).

Erinevad autorid on keeleoskuses näinud küll samu komponente, kuid kirjeldanud neid erinevate aspektide kaudu (Canale, Swain 1980, Canale 1983, Bachman 1990, Douglas 2000). Euroopa keeleõppe raamdokumendis (CEF 2001: 101–129) toodud mudelis on osa teadmisi ja oskusi liigitatud üldpädevusse, osa aga suhtluspädevusse kuuluvaks. A.-R. Hausenberg jt (2004) on esitanud need ülevaatlike joonistena (vt joonis 1 ja 2).



Joonis 1. Üldpädevuse komponendid keeleõppe raamdokumendi järgi (Hausenberg jt 2004: 11)



Joonis 2. Suhtluspädevuse komponendid keeleõppe raamdokumendi järgi (Hausenberg jt 2004: 11)

Ehkki üldpädevus ei ole suhtluspädevuse osa, on keelelise suhtlemise alati haaratud ka üldpädevuse komponendid. Eriti märgatav on see eriotstarbelises keelekasutuses, kus erialaspetsiifilised teadmised ja oskused mängivad olulist rolli.

Eriotstarbelise keeleoskuse mõõtmisel arutletakse tihti selle üle, mis on testi edukaks sooritamiseks olulisem, kas üldkeeleoskus või erialaspetsiifilised taustteadmised. Dan Douglas (2000: 29–40) väidab eri autoritele tuginedes, et kui keeletest ei ole kitsalt erialaspetsiifiline, on olulisem hea üldkeeleoskus, mis võimaldab taustteadmiste puudumist kompenseerida, samal ajal kui nõrka üldkeeleoskust taustteadmistega kompenseerida ei ole võimalik. Kui keeletest on aga väga erialaspetsiifiline ja mõistetav vaid selle valdkonna inimestele, muutuvad erialateadmised määravaks.

Samuti on leitud (Douglas 2000: 31–32), et teksti spetsiifilisust (erialasuse astet) ei mõjuta mitte niivõrd spetsiifilise sõnavara maht, kuivõrd see, kas nende sõnade tähendust tekstis selgitatakse või mitte. Mida erialaspetsiifilisem on tekst, seda rohkem eeldatakse, et lugeja või kuulaja on erialaga kursis ja seda vähem selgitatakse mõistete tähendust, mis teeb erialakaugale inimesele niisuguse teksti mõistmise väga raskeks. Seega näib teksti spetsiifilisus olevat sõltuvuses sellest, mil määral selle mõistmine nõuab erialaseid teadmisi, mida tekst ei sisalda.

Et meie koostatud kuulamistestide sihtrühmaks olid kutsekoolide õpilased, mitte vastava eriala spetsialistid, lähtusime tekstide valikul ja kohandamisel sellest, et teema oleks õpilastele tuttav (olid seda eelmisel aastal õppinud) ja selle käsitus tavainimesele mõistetav. Hoolimata sellest, et käsitleti erialaspetsiifilisi teemasid, olid tekstid valdavalt üldkeelsed.

Tekstid

Kuulamisoskuse õpetamise ja mõõtmise üheks suuremaks takistuseks on sobivate tekstide leidmine. Kui kirjalikke tekste on Interneti-ajastul suhteliselt lihtne leida ja kasutamiseks kohandada, siis suuliste tekstide töötlemine on raskendatud. See tõttu kasutatakse tihti kuulamisoskuse arendamiseks kirjalikke tekste, mis kantakse suuliselt ette. Ometi on kirjalikud ja suulised tekstid oma olemuselt vägagi erinevad. Tiit Hennoste (2000: 1139–1141) toob välja viis kõne ja kirja erinevust:

- 1) kõne on tagasikerimatu, s.t kuulajal ei ole võimalik teksti mitu korda kuulata;
- 2) kõne on kustutamatu, s.t öeldut ei saa tagasi võtta;
- 3) kõne on lineaarne, s.t kuulaja kuuleb teksti sõnahaaval ja ei saa tekstis tagasi- ega edasivaateid teha;
- 4) kõne on seotud nii kõneleja kui ka kuulaja lühimälu mahuga, mistõttu korraga suudetakse töödelda 7 ± 2 ühikust koosnevat üksust;
- 5) kõne on visuaalne-auditiiivne, s.t erinevalt kirjast saab kuulaja teavet ka vokaalsete vahendite (intonatsioon, rõhud, tämber, tempo, pausid) kaudu.

Samas erinevad suulised tekstid nii suhtlusviisi omaduste kui ka suhtlusolukorra poolest (Hennoste 2000: 1147). Suhtlusviisi omaduste poolest on suulised tekstid rohkem või vähem dialoogilised või monoloogilised ning rohkem või vähem spontaansed või redigeeritud. Näiteks peetakse loengut klassikaliseks monoloogiks, kuid ometi saab lektor kuulajate reaktsiooni jälgides tagasisidet, mis muudab loengugi mingil määral dialoogiks. Redigeerituse all peetakse silmas teksti eelnevat ettevalmistamist ja selles mõttes on enamik suulisi tekste mingil määral redigeeritud, kasvõi põhipunktide läbimõtlemissel näol.

Suhtlusolukord võib olla vahetu või vahendatud ning argine või avalik (Hennoste 2000: 1148). Vahetu suhtlus võimaldab kuulajal anda kõneleajale kohest tagasisidet, mistõttu dialoog kujutab endast vestluspartnerite ühisloomingut. Salvestatud tekstid, nagu raadio- ja telesaated, aga ka keeleõppes kasutatavad kuulamismaterjalid, on vahendatud tekstid, mille puhul tagasiside andmine pole võimalik. Helisalvestiste puhul kaob kuulaja jaoks ka kõne visuaalne pool, sest pole näha kõneleajate kehakeelt, miimikat, žeste jne. Argises suhtlusolukorras on vooruvahetus tavaliselt vaba ja osalejate rollid võivad muutuda (näiteks võib küsijast saada vastaja ja vastupidi), avalikus suhtlusolukorras on aga rollid rangemalt määratud (on selge, kes küsib ja kes vastab, mis järjekorras kõneldakse jne).

Nii suhtlusviisi kui ka suhtlusolukord mõjutavad seda, kui kerge on kuulajal teksti jälgida ja mõista. Võib oletada, et vahetust argisest dialoogist arusaamine on lihtsam, sest kuulajal on võimalik rääkijat katkestada ja selgitusi paluda. Samas kulgeb argivestlus ettearvatult, mis võib selle jälgimise teiselele kõneleajale

hoopis raskemaks muuta. Avaliku vestluse määratlus võib aga teksti jälgimist hõlbustada. Igal juhul on dialoogi jälgimine kergem kui monoloogi jälgimine, sest monoloogi puhul dikteerib teksti töötlemise kiiruse rääkija kõnetempo. Eriti raske on jälgida redigeerimata monoloogi, millel puudub selge ülesehitus ja mis esitatakse kiire tempoga.

Loomulikult ei ole suhtlusviis ja suhtlusolukord ainsad kuulamisteksti mõistmist mõjutavad tegurid. Nii teksti teema tuntus, abstraktsuse aste, keeleline lihtsus kui ka esituse selgus mängivad oma osa selles, kui kerge või raske on kuulajal tekstist aru saada. Eksamiolukorras aga ei piisa pelgalt kuulamisteksti üldisest mõistmisest, vaid tuleb aru saada just sellest, millele keskendub ülesanne ning arusaamist ka nõutaval moel näidata. Seega kujuneb kuulamisülesande raskus teksti ja ülesande omaduste koosmõjus, millest olulisemad kajastuvad tabelis 1 (vt ka Buck 2001: 149–153, Alderson 2000: 282–285, Hausenberg jt 2004: 36–40, 57–61).

Tabel 1. Kuulamisülesande raskust mõjutavad aspektid

	Lihtne	Kuulamisülesande tähtmist mõjutav aspekt	Raske
TEKST	Jutustav ja kirjeldav	Tekstitüüp	Arutlev ja analüüsiv
	Loogiline, hästi liigendatud, selgete üleminekutega	Ülesehitus	Kõrvalepõigetega, selge struktuurita, ebaloogiline
	Igapäevane üldkeele sõnavara, lihtne lauseehitus	Keel	Kõnekeelsed ja idiomaatilised väljendid; keerulised lausekonstruktsioonid
	Konkreetne, hästi tuttav, igapäevane	Teema	Abstraktned, vähettavad, spetsiifilised
	Aeglane, selge, madal tämber, selgesti eristuvad hääled	Esitus	Kiire, segane, taustamüra, kõrge tämber, monoloog või raskesti eristatavad hääled
ÜLESANNE	Valikulise teabe leidmine	Kuulamise eesmärk	Üksikasjade mõistmine
	Tekstis otseselt antud	Vastus	Teksti põhjal järeldatav
	Mittekeelelised valikud	Vastamise viis	Sõnaline vastus

Meie poolt kasutatud tekstid – intervjuu ja loeng – on T. Hennoste määratluse järgi mõlemad redigeeritud ja avalikud suulised tekstid, mis erinevad oma redigeerituse astmelt. Intervjuuks olid küsimused eelnevalt ette valmistatud, kuid intervjueritav ei teadnud neid. Loengutekstid olid aga valmis kirjutatud ja salvestamisel need loeti maha. Mõlemad tekstid olid kuulajatele vahendatud tekstid, sest neid kuulati helilindilt. Ometigi on ka vahendatud dialoogi jälgimine lihtsam kui monoloogi jälgimine, sest intervjuu on kindla struktuuriga ja intervjuerija küsimused valmistavad kuulajat vastuste kuulamiseks ette. Samuti muudab dialoogi kuulamise lihtsamaks erinevate tämbrite esinemine.

Monoloogi – kas vahetu või vahendatud – kuulamine nõuab pikemaajalist keskendumist ja ei ole seetõttu madala keeleoskusega inimesele jõukohane. Loengu kuulamine on eriti raske, sest tavaliselt on sel juhul tegemist arutleva ja analüüsiva tekstiga abstraktsel, vähettaval ja spetsiifilisel teemal (vt tabel 1). Loengut saab kuulajale kergemaks muuta teksti aeglasema esituse, loogilise ülesehituse, lause-

ehituse lihtsustamise ning sõnavara teadliku valiku ja vajalike selgitustega (vt Rost 2002: 162–164, 233–235). Võrreldes õppe-eesmärgilise lugemisega on loengu kuulamise puuduseks see, et on vaja palju meelde jätta, mistõttu tuleb märkmeid teha (Rost 2002: 162). Keeleõppija jaoks on aga kuulamise ajal märkmete tegemine ülesande täitmist raskendav asjaolu (vt tabel 1).

Suuliste ja kirjalike tekstide erinevuste mõistmine on oluline siis, kui (nagu see tihti juhtub) soovitakse keeleõppes kuulamisoskuse arendamiseks ja mõõtmiseks kasutada suuliselt esitatud kirjalikke tekste. Sel juhul tuleb kirjalik tekst suulise teksti eripära arvestades ümber töötada.

Kuulamisoskuse aspektid

Tavaelus sõltub see, kuidas me midagi kuulame nii tekstist kui kuulamise eesmärgist. Näiteks hommikul raadiot kuulates võib valikuliselt keskenduda vaid huvipakkuvale, näiteks ilmateatele või uudistele, aga võib ka kõike ühtviisi süvenenult kuulata, eriti siis, kui eesmärgiks on programmis kaasa rääkida. Kommunikatiivne keeleõpe lähtub eelkõige sellest, kuidas inimesed tavaelus keelt kasutavad. Kuulamisoskuse arendamisel ja mõõtmisel keskendutakse üldiselt neljale aspektile: teksti üldisele mõistmisele (üldkuulamine), mingi konkreetse teabe leidmisele tekstist (valikk kuulamine), teksti üksikasjalikule mõistmisele (detailne kuulamine) ja kuuldu põhjal järelduste tegemisele (kuuldu tõlgendamine) (vt nt Harmer 2001: 201–202).

Uurimused (Shohamy, Inbar 1991, Buck 2001: 134 järgi) on näidanud, et kuuldu tõlgendamist nõudvad küsimused on testitavatele kõige raskemad, samas kui valikk kuulamist ja detailset kuulamist nõudvate küsimuste raskus sõltub suuresti sellest, mis laadi teabele küsimus keskendub. Teksti põhiteema seisukohalt olulise teabe kohta koostatud küsimused on lihtsamad kui tõlgendamisküsimused, kuid nn triviaalküsimused, mis nõuavad täpsete, kuid teksti põhiteema seisukohalt ebaoluliste üksikasjade mõistmist, “käituvad” ettearvamatult: võib juhtuda, et nõrgad õpilased vastavad neile paremini kui tugevad õpilased. See tähendab, et triviaalküsimused muudavad testi tulemused ebausaldusväärseks, mistõttu tuleks neid kuulamistestides vältida (Buck 2001: 137).

Ka meie poolt kasutatud tekste – vestlust ja loengut – võib kuulata erinevatel eesmärkidel ja seega erinevalt. Tavaliselt on aga vestluse pealtkuulamisel tegemist mingi konkreetse teabe otsimisega ja kuuldust järelduste tegemisega; loengu puhul on lisaks konkreetse teabe leidmisele oluline üksikasjade märkamine ja üleskirjutamine. Loetletud kuulamisoskuse aspekte soovisimegi testidega mõõta.

Ülesandetüübid

Kui tavaelus määrab kuulamise eesmärgi inimese huvi või vajadus, siis keeleõppes ja keeleoskuse mõõtmisel antakse kuulamise eesmärk ülesandega. Seega on kuulamisoskuse mõõtmine alati vahendatud, kusjuures vahendajaks on ülesanne, mis võib ise testitulemusi mõjutada. Näiteks kui küsimus on mitmeti mõistetav, võib kuulaja sellele valesti vastata mitte seetõttu, et ta kuulamisteksti ei mõistnud, vaid seetõttu, et ta mõistis küsimust teisiti, kui selle koostaja oli planeerinud.

Üheks põhimõtteks keeleoskuse mõõtmisel on võimalikult erinevate ülesandetüüpide kasutamine (vt nt Alderson jt 1995: 44–46). Põhjuseks on see, et igal ülesandetüübil on oma eelised ja puudused ning mõned tüübid sobivad teatud oskuste mõõtmiseks paremini kui teised. Mida rohkem eri ülesandetüüpe kasutada, seda kindlam võib olla, et mõõtmistulemused on usaldusväärsed.

Ülesandetüübist oleneb ka kuulamistesti raskus. Kõige parem oleks kuulamisostkust mõõta n-ö puhtalt, s.t teisi osaoskusi kasutamata. See on paraku enamalt jaolt võimatu, sest testitavad peavad tavaliselt lisaks teksti kuulamisele küsimusi lugema. Kui aga ülesanne nõuab veel sõnalise vastuse kirjutamist, on see eriti raske ja selliseid ülesandeid peetakse jõukohaseks vaid kõrgematel keeleoskustasemetel. Samuti tuleb seda tüüpi ülesande koostamisel arvestada, et testi sooritajal oleks piisavalt aega eeldatavad vastused kuulamise käigus kirja panna. Üldiselt eelistatakse siiski valikvastustega ja järjestamisülesandeid, kus testitavad kirjutavad vaid numbreid/tähti või teevad ristikesi/linnukesi. Kuulamisostkuse mõõtmiseks sobib eriti hästi piltide valimine või järjestamine, sest sellisel juhul ei ole testitavatel vaja kasutada ka lugemisostkust.

Valikvastustega ülesandel on mitmeid alltüüpe sõltuvalt valikute arvust ja liigist. Valikuid võib olla 2–4, mõnikord rohkemgi. Kahe valikuga ülesandeks on näiteks õige-vale vastusega küsimus, mille suureks puuduseks on see, et eksaminandidel on vastuseid huupi märkides 50-protsendiline võimalus tabada õige vastus. Seetõttu peaks sedalaadi küsimusi olema palju, et oletamise mõju vähendada (Alderson jt 1995: 51). Valikvastustega küsimusi on lihtsam koostada valikulise info mõistmise ja teksti peamõttest arusaamise kontrollimiseks, kuid seda tüüpi ülesannet saab kasutada teistegi kuulamisostkuse aspektide mõõtmiseks.

Kuulamistesti järjestamisülesanne nõuab testi sooritajalt piltide või tekstide järjestamist vastavalt kuulatavale tekstile. Ülesanne on raskem, kui vastusevariante on rohkem kui tulemus seda nõuab, kuna sel juhul peab õppija lisaks järjestamisele ka otsustama, kas mingi variant üldse vastusesse kuulub. Järjestamisülesanne võib olla problemaatiline nii koostamise kui hindamise seisukohast. Koostamisel tuleb silmas pidada, et tekstis sisalduksid selged vihjed, mis aitavad vastuseid järjestada, nii et vaid üks tulemus oleks võimalik ja õige. Tavaliselt tuleb teksti koostajal nimetatud vihjeid lisada. Kuna testisooritaja võib ülesande täitmise käigus eksida, seisneb probleem selles, kuidas osaliselt õiget vastust hinnata. Harilikult loetakse osaliselt õige vastus valeks ning arvesse läheb vaid täiesti õige järjestus (Alderson 2000: 221).

Ülesandetüüpide valikul lähtusime põhimõttest kasutada tuntud ülesandetüüpide kõrval ka neid, mida eesti keele tasemeeksami ei kasutata ja mis suure tõenäosusega on õppijatele võõrad (vt tabel 2). Samas sõltus ühe või teise tüübi kasutamine ka võimalustest, mida tekst pakkus. Õppijatele eeldatavasti vähem tuttavateks kuulamisülesanneteks võib pidada

- 1) järjestamisülesandeid,
- 2) valikvastustega küsimusi, milles valikud on antud tabelis,
- 3) mitut õiget vastust eeldavaid valikvastustega küsimusi,
- 4) lühivastuste märkimist nõudvat tabeli täitmist.

Nagu tabelist 2 ilmneb, sisaldasid kõige rohkem eri ülesandetüüpe ehituse ja kokaerialale mõeldud testid.

Tabel 2. Testides kasutatud ülesandetüübid erialati

Eriala	Üksi-küsimusi	Valikvastused				Järjestamine		Lühivastus
		3–4	Tabel	Õige-vale	Mitu õiget	Tekst	Pildid	Tabel
Arvutiteenindus	12			X				
Autolukksepp-maaler	6	X			X	X		
Ehitus	14	X	X	X	X	X	X	
Hotelliteenindus	11			X				
Keskkonnakaitse	11			X				
Kokk	18	X		X	X			X
Laomajandus	11			X				
Sotsiaalhooldus	8	X		X	X			
Turismikorraldus	7	X			X			
Ärikorraldus	7	X						

Kuulamistestide koostamine

Kuulamistestide koostamist alustasime sobivate tekstide leidmisest. Kuna esimene tekst oli üldkeelne intervjuu, otsustasime selle salvestada võimalikult loomulikuna. Seetõttu mõtlesime küll valmis küsimused, kuid intervjuueeritavale ütlesime vaid teema. Intervjuueeritav oli eesti keelt teise keelena valdav inimene, mistõttu tema keeleteadus isegi tööst rääkides on eesti keele õppijatele mõistetavam, sest on kirjakeelsem kui eesti keelt emakeelena kõnelejalatel spontaanses vestluses.

Teiseks tekstiks valisime erialaloengu ning seetõttu tutvusime õppekavadega, et välja selgitada esimesel aastal läbitud ainekursused. Sellega soovisime tagada teema tuntuse, mis on B1/B1+ tasemel kuulamisel veel oluline. Valisime eriala seisukohalt oluliste ainete hulgast teemad, mille kohta otsisime sobivaid populaarteaduslikke tekste. Valitud tekstid töötasime ümber loenguteks. Kohandamisel võtsime arvesse suuliste tekstide eripära. Kuna kõne on seotud nii kõneleja kui kuulaja lühimälu mahuga, vähendasime tunduvalt tekstis sisalduvat infot, eemaldades ülearused detailid. Samuti lühendasime segmentide pikkust nii, et need jääksid enamikus 7 ± 2 ühiku raamesse. Kõne lineaarsus tingib selle, et kõneleja peaks ise lisama teksti edasisaateid, mida kuulaja teha ei saa. Seetõttu muutsime tekstide ülesehitust. Et kuulajad olid eesti keelt teise keelena õppijad, lihtsustasime tekstide sõnavara ja lauseehitust. Lisas 2 on näiteks toodud katkend ehituseriala õpilastele koostatud testi lähtetekstist ja selle ümbertöötatud variandist, millest on näha tehtud muudatuste ulatus.

Tekstide kohandamisele järgnes nende lindistamine, sest oluline on kuulamistesti ülesanded kirjutada kuulatud, mitte loetud tekstile. Teksti üleskirjutuselt ei selgu esitamise tempo ja see, kuidas rääkija midagi rõhutab, mistõttu võib olla võimatu sellele kirjutatud ülesannet kuulamisel täita. Samas aga tekitab teksti eelnev lindistamine teatud probleeme. Nimelt võib selguda, et teksti põhjal on raske ülesannet koostada, kuid kuna tekst on juba lindistatud, ei saa seda ka enam muuta.

Ülesannete koostamisel lähtusime eelkõige sellest, kuidas antud teksti tavaelus kuulatakse. Seetõttu oli põhiorhk valik- ja detailsel kuulamisel, ehkki intervjuu kohta koostatud küsimuste hulgas esines ka tõlgendamisküsimusi. Nagu tabelist 2 ilmneb, kujunes kõige enam kasutatud ülesandetüübiks õige-vale vastusega küsimus, mida kasutasime seitsmes testis kümnest, kusjuures kolme testi puhul oli see ainuke kasutatud ülesandetüüp. Pooled testid sisaldasid rohkem kui ühte ülesandetüüpi. Kõige rohkem eri ülesandetüüpe (kuut) sisaldas ehituserialale koostatud test, mis on toodud lisa 2.

Testide läbiviimine ja tulemused

Kuna erialati oli lähetatuid vaid 4–20 õppijat, siis ei saanud testide sooritamist statistiliselt analüüsida. Samuti esines probleeme testide usaldusväärse läbiviimisega. Näiteks otsustati kohapeal ühe testi läbiviimisel mitte kasutada meie koostatud testi, vaid testiti õppijaid enda meelest sobivate materjalidega. See seab kahtluse alla teistegi testide tulemused ning muudab statistilise analüüsi mõttetuks. Seega põhinevad järgnevad üldistused tähelepanekutel, mida me tegime testide hindamise käigus.

Ülesandetüübid, mida tasameeksamitel ja seega ilmselt ka õpetuses ei kasutata, on õppijatele võõrad ja selliste ülesannete lahendamine valmistas raskusi. Et raskuse põhjus peitus just ülesandetüübis, näitab asjaolu, et paljudel juhtudel polnud esimesel testimisel ülesannet täidetud või täideti vaid osaliselt. Teisel testimisel aga olid samad õppijad ülesannet vähemalt täita üritanud. Üheks niisuguse ülesande näiteks on koka erialale koostatud lühivastuste märkimist nõudev tabeli täitmine.

Teiseks näiteks võib tuua ehituse erialale koostatud järjestusülesande, milles kuulamise käigus tuleb ette antud plaatimistoimingud järjestada õigesti (vt lisa 2). Õppijad, kes esimese soorituse käigus polnud ülesannet üldse teinud, olid teisel testimisel tabelit vähemalt üritanud täita, mõnel juhul isegi osaliselt õigesti. Õppijad, kes aga olid esimesel testimisel toimingud järjestanud valesti, täitsid teise sooritusega ülesande enamasti õigesti.

Hästi vastati küsimustele, mis eeldasid konkreetse teabe leidmist tekstist. Sellest võib järeldada, et õpilased on omandanud selektiivse kuulamise võtted.

Raskusi valmistasid küsimused, millele vastamiseks oli vaja kuuldu põhjal järeldusi teha või tekstis sisalduvat teavet kombineerida. See võib tähendada, et õpilaste kuulamisoskus ei olnud piisav, et kogu teksti mõttega jälgida, saadud teavet varasemaga võrrelda ja järeldusi teha. Samas võib see aga märku anda sellest, et õpilased pole harjunud sedalaadi kuulamisülesannetega ega oska nendele seetõttu õigesti läheneda.

Testitavate soorituse hindamisel tehtud tähelepanekud kinnitavad taas, et kolme nädala jooksul, nagu ette arvatud, õppijate kuulamisoskus märgatavalt ei paranenud. Tõsiasi, et paljud õppijad hindasid lähetuse järel oma keeleteadlikkust paremaks (Vaagen 2005), võib tuleneda sellest, et õppijad said positiivse kogemuse eesti keele eesmärgipärasest kasutamisest keelekeskkonnas, mis eelkõige parandas nende enesekindlust ja usku võimesse eesti keeles toime tulla. Seetõttu on sellised lähetused õpilastele äärmiselt vajalikud, kuid nende eesmärgid peavad olema realistlikud.

Kokkuvõte

Kutsekooliõpilastele mõeldud kuulamistestide koostamise protsess kinnitas taas, et hea testi eeltingimuseks on sobiv tekst. Suuliste tekstide leidmine ei ole paraku lihtne, mistõttu tuleb tihti kohandada kirjalikke tekste. Seejuures tuleb kindlasti arvestada suuliste tekstide eripära.

Ülesandetüüpide valikul tuleb lähtuda põhimõttest, et igas testis tuleks kasutada vähemalt kahte tüüpi, et vähendada ülesandetüübi mõju testi tulemustele. Samade ülesandetüüpide kasutamine eksamil võib viia selleni, et õppetöös harjutakse ühtede ülesannetega ja omandatakse teatud strateegilised oskused nende ülesandetüüpide lahendamiseks, kuid harjutamise käigus kuulamisoskus tegelikult ei parane ja tundmatu ülesandetüübi kasutamine võib kaasa tuua ülesande ebaõnnestunud sooritamise. Eri ülesandetüüpide kasutamine õpetuses arendab kuulamisoskust laiemalt ja on eelduseks keeletestide edukale sooritamisele.

Kuulamistekstide kohta küsimusi koostades tuleb lähtuda tekstitüübist ja sellest, kuidas me analoogseid tekste tegelikult elus kuulame. Küsimusi tuleks esitada eelkõige olulise teabe kohta tekstis, mitte ebaoluliste detailide kohta. Samuti tuleks silmas pidada, et küsimusi ei esitataks mitte ainult tekstis otseselt sisalduva info kohta, vaid kontrollitaks ka kuuldu põhjal järelduste tegemise oskust.

Kirjandus

- Alderson, Charles J. 2000. *Assessing Reading*. Cambridge: Cambridge University Press.
- Alderson, Charles J.; Clapham, Caroline; Wall, Dianne 1995. *Language Testing Construction and Evaluation*. Cambridge: Cambridge University Press.
- Bachman, Lyle F. 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Buck, Gary 2001. *Assessing Listening*. Cambridge: Cambridge University Press.
- Canale, Michael 1983. On some dimensions of language proficiency. – J. W. Oller (Ed.). *Issues in Language Testing Research*. Rowley, Mass.: Newbury House, 333–342.
- Canale, Michael; Swain, Merrill 1980. Theoretical bases of communicative approaches to second language teaching and testing. – *Applied Linguistics* 1 (1), 1–47.
- CEF = Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Douglas, Dan 2000. *Assessing Languages for Specific Purposes*. Cambridge: Cambridge University Press.
- Harmer, Jeremy 2001. *The Practice of English Language Teaching*. Harlow: Pearson Education Ltd.
- Hausenberg, Anu-Reet; Kikerpill, Tiina; Rõigas, Maia; Türk, Ülle 2003, 2004. *Keeleoskuse mõõtmine*. Käsiraamat. Tallinn: TEA Kirjastus.
- Hennoste, Tiit 2000. Sissejuhatus suulisesse eesti keelde I. Taust ja uurimisobjekt. – *Akadeemia* 5, 1118–1150.
- Hymes, Dell H. 1972. On communicative competence. – J. B. Pride, J. Holmes (Eds.). *Sociolinguistics*. Harmondsworth: Penguin, 269–293.
- Rost, Michael 2002. *Teaching and Researching Listening*. Harlow: Pearson Education Limited.
- Vaagen, Anu 2005. Eesti keele kui teise keele oskuse arendamisest EL Phare keeleõppe- ja lähetusprojekti näitel. – Ettekanne konverentsil “Emakeel ja teised keeled V”, Tartu Ülikool, 17.-18.11.

Kaudviited

- Carroll, John B. 1961. Fundamental considerations in testing for English proficiency of foreign students. – Testing the English Proficiency of Foreign Students. Washington, D.C.: Center for Applied Linguistics, 31–40.
- Lado, Robert 1961. Language Testing. London: Longman.
- Shohamy, Elana; Inbar, Ofra 1991. Validation of listening comprehension tests: The effect of text and question types. – Language Testing 8 (1), 23–20.

LISA 1. Lähteteksti kohandamise näide

Katkend lähtetekstist (260 sõna):

.. Vuukide täitmine võib alata, kui plaatimissegude on kivinenud pakendil nõutud aja. Vuugitäite värv valitakse esteetilisest seisukohast lähtudes. Vuugitäiteid on saadaval erinevate laiusega vuukidele, ning erineva kasutusala ruumidele. On saadaval ka vuugitäite omadusi täiendavad lisaained. Vuugitäide segada vastavalt juhiste pakendil ning kanda plaatide pinnale vuukide suhtes diagonaalselt. Selleks kasutatakse kummist vuugilabidat.

Kui vuugitäide on pindmiselt kuiv kuid veel elastne, pestakse plaatide niiske käsnaga puhtaks. Poorsed plaadid on soovitatav katta enne vuukimist pinnakaitseõliga. Vuugitäitesegu ei tohi lasta plaatide pinnale kivistuda. Juhul kui siiski vuugitäide on jäänud õigeaegselt eemaldamata ning plaatide külge kivistunud, on abiks spetsiaalsed kemikaalid. Kui vuugid on tahenenud kuid mitte lõplikult kivistunud, puhastage plaatkatte pinda pehmepõhjalise hõõruti või jämedakoelise kangaga.

Järelhoolduseks on hea vuuke veel paari ööpäeva jooksul mõned korrad niisutada. Plaatkatte nurgavuugid, põrandaseina ühenduskohad ning kokkupuutekohad teiste tarindite, näiteks lengide ja sisustusega täidetakse elastse sanitaarsilikoniga. Elastne täide on vajalik, kuna erinevast materjalist tarindid elavad ning liiguvad pisut erinevalt. Et on saadaval vuugitäitega sobivates toonides silikoone ei mõju need vuugid häirivalt.

Vuugiservadele paigaldatakse teip, et silikoon ei satuks ülemäära plaatide peale. Silikoon kantakse vuukidesse spetsiaalse püstoliga, silutakse märja sõrme või niisutatud puupulgaga. Teip eemaldatakse peale silikooni paigaldamist ning silutakse märja sõrmega veelkord üle, et tasandada teibi eemaldamisest tekkinud randid. Looduskividest plaatkatte puhul on oluline valida neutraalne sanitaarsilikoon, et vältida plaadiservade värvumist silikoonmassis leiduvate kemikaalide mõjul. Lisaks silikoniga täitmisele võib plaatkatte sise- ja välisnurkades ning vabaserva lõppedes kasutada ka dekoratiivseid plastikliste. Plastiklist aga ei asenda vuukide tihendamist.

Põrandakütte sisselülitamisega tuleb oodata vähemalt nädal peale tööde lõpetamist. Ruumi kasutamisega, näiteks vannitoa puhul oodata kuni materjalid on kuivanud pakendil toodud aja. .. (Veiko Ikkonen, Mira Ehitusmaterjalid OÜ tegevjuht)

Katkend kohandatud tekstist (125 sõna):

.. Vuukide täitmine võib alata, kui plaatimissegude on kivinenud pakendil nõutud aja, mis võib olla üks kuni kolm ööpäeva. Vuugitäide segatakse vastavalt juhiste

pakendil. Kõigepealt kantakse vuugitäide plaatide pinnale vuukide suhtes diagonaalselt. Selleks kasutatakse kummist vuugilabidat.

Seejärel, kui vuugitäide on pindmiselt kuiv, pestakse plaadid niiske käsnaga puhtaks. Kuivanud plaatide pind puhastatakse hoolikalt kuiva lapiga.

Lõpuks täidetakse plaatkatte nurgavuugid ning põranda ja seina ühenduskohad silikooniga. Elastne täide on vajalik, kuna erinevast materjalist tarandid elavad ning liiguvad pisut erinevalt.

Vuugiservadele paigaldatakse teip, et silikoon ei satuks plaatide peale. Silikoon kantakse vuukidesse spetsiaalse püstoliga, silutakse märja sõrme või niisutatud puupulgaga. Teip eemaldatakse kohe peale silikooni panemist.

Põrandakütte võib sisse lülitada mitte enne kui nädal peale tööde lõpetamist. Plaaditud ruumi võib hakata kasutama, kui materjalid on kuivanud pakendil ettenähtud aja. ...

LISA 2. Ehituse erialale koostatud kuulamistest

Tekst 2. Lugege küsimusi töölehel. Märkige vastused ristiga.

NÄIDE. o. Eelmisel korral räägiti põhjalikult

- aluspinnast
- plaatimisest
- vuukimisest

1. Töövahendid, mida kasutatakse plaatimisel ja vuukimisel, on (NB! märkige KÕIK teksti järgi õiged vastused!)

- segukamm
- vuugilabidas
- vuugisegu
- käsn
- silikoon

2. Plaatimiseks on vaja

- Aluspinnad ette valmistada
- Ruumi tuulutada
- Põrandat kütta

3. Märki tabelisse ristiga (X), kas aluspind on täiesti stabiilne, teatud määral stabiilne või ebastabiilne.

	Täiesti stabiilne	Teatud määral stabiilne	Ebastabiilne
vana betoon			
hiljuti valatud betoon			
puitalusel sein			
kõetav põrand			
soemüür			

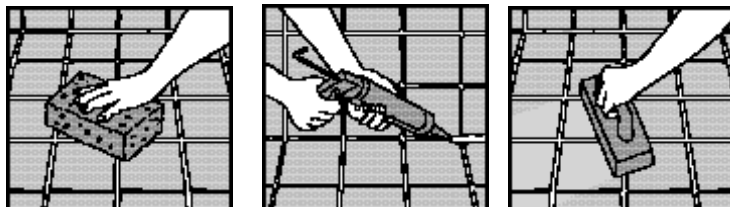
4. Järjesta plaatimistööd (1–5) lektori soovitude järgi. (NB! 1. töö on märgitud)

- valmista plaatimissegud
- kammi segu läbi
- kanna segu aluspinnale
- koputa plaadid kinni
- suru plaadid segusse

5. Kas väide on õige või vale?

Väide	Õige	Vale
Esimesena plaaditakse seinad.		
Plaatimist alustatakse nurkadest.		
Töö kvaliteeti peab kontrollima mitu korda.		

6. Järjesta vuukimistööd (1–3) lektori soovitude järgi. Märgi number pildi kõrvale kastikesse.



7. Teipi kasutatakse

- plaatide kaitsmiseks
- vuugiservade kaitsmiseks
- vuukide kaitsmiseks

8. Põrandat võib kütta

- 2 päeva pärast plaatimist
- 5 päeva pärast plaatimist
- 8 päeva pärast plaatimist

Tiina Kikerpilli (Tartu Ülikool) uurimisvaldkondadeks on võõrkeelte õppimine ja õpetamine, keeleõppimise mõõtmine, arvutitugine keeleõpe.
tiina.kikerpill@ut.ee

Ülle Türk (Tartu Ülikool) teaduslikeks huvialadeks on keeleõppimise mõõtmine, võõrkeelte õppimine ja õpetamine, kirjaliku väljendusõppimise ja kriitilise lugemisõppimise õpetamine, kultuuridevaheline kommunikatsioon ning Suurbritannia kultuur.
ulle.turk@ut.ee

MEASURING SPECIFIC PURPOSE LANGUAGE ABILITY: TEXTS AND TASKS IN ESTONIAN AS A SECOND LANGUAGE LISTENING TESTS

Tiina Kikerpill, Ülle Türk

University of Tartu

Although tests of Estonian as a second language (L2) have been developed for more than ten years, most of them have been designed to test learners' competence in general Estonian. Testing of Estonian for specific purposes has been neglected so far. The EU Phare Project of 2004-2005 "Developing the Estonian language competence in vocational schools with the language of tuition other than Estonian" offered a welcome opportunity to develop ten different listening tests for students of ten specialities.

The article discusses the process of test development focusing on different models of language competence and the characteristics of specific purpose language ability, various aspects of listening comprehension, features of spoken texts and the principles that need to be followed when transforming written texts into spoken ones, principles of choosing appropriate task types and the general procedure for developing listening tests.

The test development process highlighted a number of issues. One of the major obstacles to designing good listening tests turned out to be a lack of appropriate oral texts. Therefore, it was necessary to use some written texts and adapt them bearing in mind the differences between the spoken and written languages. In designing the tasks, the skills used in real-life listening to the same types of texts were taken into account and an attempt was made to tap similar skills. To minimize the test method effect, each test employed at least two different task types.

As the number of students who took each test was very small, it is not possible to analyse the results statistically and to draw any reliable generalizations from them. It was clear, however, that students found it easier to answer the questions asking for specific information than those obliging them to obtain information from various parts of the text and combine it in a new way. Similarly, familiar task types proved to be easier than those that had not been used in Estonian language examinations before.

Keywords: Estonian as a second language, language tests, development and assessment of listening comprehension

ARVUTI SUHTLUSES

Mare Koit

Ülevaade. Artiklis käsitletakse probleeme, mis tuleb lahendada selleks, et arvutiga saaks suhelda loomulikus keeles, sh eesti keeles. Tuuakse näiteid arvutitega suhtlemise kohta, alustades A. M. Turingi mõtlemisvõimelise arvuti ideest 1950. a kuni kaasaegsete loomulike liidestega dialoogsüsteemide ja robotiteni. Tutvustatakse akadeemilist ja insenerlikku lähenemist dialoogsüsteemide loomisele ning kahte paradigmat, mida dialoogi automaattöötluses on rakendatud – ratsionalismi ja empirismi. Antakse ülevaade töödest, mida on tehtud eesti keeles suhtleva dialoogsüsteemi loomise eesmärgil, sh Tartu Ülikooli eesti dialoogikorpusel ja olemasolevast eesti keele automaattöötluse tarkvarast. Lõpuks vahendatakse tuntud teadlaste tulevikuvisionone arvuti kui suhtluspartneri kohta.*

Võtmesõnad: Turingi test, arvutikõne hüpotees, dialoogsüsteem, robot, dialoogikorpus, eesti keel

1. Sissejuhatus

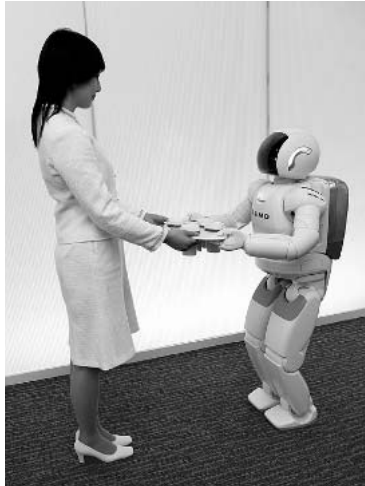
Suhtlus arvutiga on intrigeeriv teema. Pole veel sugugi igapäevane, et meie vestluskaaslasel on robot-tehisinimene või et näiteks helistades infotelefonile kuuleme vastuseks sünteeshäält, mis viisakalt ja kannatlikult vastab ka meie kõige tüütumatele küsimustele. Ometi oli 2006. a aprillikuus võimalik Tallinnaski kohata robotit: rahvusvahelise auto- ja tehnikanäituse Motorex Tallinn külastajad said tutvuda Jaapani firma Honda robotiga Asimo (joonis 1). 1,2 m pikkune tehismehike on täis kõige kaasaegsemat tehnoloogiat ning ta on oodatud külaline kõikvõimalikel suurematel tehnikasündmustel üle kogu maailma. Asimo tunneb ära oma vestluskaaslased, kõnnib, ronib trepil ja kõneleb. Asimo keelelised võimed on siiski veel napid: ta mõistab lihtsaid jaapani- või ingliskeelseid käsklusi ja suudab neile vastata.

Nimetagem veel kahte tähelepanuväärset robotit – Aibo¹ ja Furby². Aibo on robotkoer, mida toodab Jaapani firma Sony alates 1999. a. Ta liigub, “näeb” keskkonda kaamerate abil, mõistab suulisi käsklusi, õpib, kasutades väliseid stiimuleid

* Artikli valmimist on osaliselt toetanud ETF-i grant nr 5685.

¹ <http://www.aibo-europe.com/> (20.07.2006)

² <http://www.hasbro.com/furby/> (20.07.2006)



Joonis 1. Robot Asimo kandikut ulatamas³

keskkonnast, omanikult või teistelt Aibodelt. Furby on robotloomake, mida valmistab alates 1998. a Tiger Electronics. 2005. a lisati robotile kõnetuvastus. Furby suudab tuvastada kümneid fraase ja hääldada sadu sõnu.

Kõigi nende robotite tegevust juhivad sisseehitatud arvuti.

Maailmas kasutatakse lisaks robotitele tuhandeid kõnedialoogsüsteeme, mis suhtlevad inimestega mitmes erinevas keeles, võttes vastu miljoneid telefonikõnesid päevas (Minker jt 2005). Pakutavate teenuste skaala on lai: infotelefon, hädaabi, lennureisi reserveerimine, postisaadetiste jälgimine, telefonikõne suunamine nime järgi, kinnisvarainfo jne.

Kas arvutiga peab saama suhelda inimkeeles? Sellele on nii poolt- kui vastuargumente. Põhiliseks pooltargumendiks on asjaolu, et inimesed ei pea siis ära õppima uut keelt (programmeerimiskeelt või andmebaasi päringukeelt), vaid saavad kasutada loomulikku keelt, mida nad niikuinii juba valdavad. Igas loomulikus keeles on vahendid soovide ja kavatsuste väljendamiseks, objektidele viitamiseks ja nende kirjeldamiseks, konteksti arvestamiseks (anafoor, ellips). Kõne kasutamine jätab muude ülesannete jaoks vabaks käed ja silmad. Põhiliseks vastuargumentiks on tehnoloogilised piirangud: kuna kõnetuvastus ja keele automaatne mõistmine pole veel saavutanud inimese taset, siis on raske defineerida arvuti jaoks adekvaatset sisendkeelt, et vältida mitmesust ja vigu. Arvuti keeleline ja arutlusvõime pole inimesele arusaadavad.

2. Elizast Alice'ini: kas arvuti mõistab keelt?

Esimene tuntuks saanud programm, mis suhtles kasutajaga inglise keeles, oli 1966. a valminud ELIZA (Weizenbaum 1966). Dialog (1) on näide vestlusest ELIZA-ga.⁴

(1) Dialog ELIZAgas (Weizenbaum 1966). K – kasutaja, A – arvuti, s.o ELIZA

³ <http://world.honda.com/ASIMO/> (8.09.2006)

⁴ Vt ka <http://www-ai.ijs.si/eliza/eliza.html> (6.07.2006)

K: Men are all alike.
 A: In what way?
 K: They're always bugging us about something or other.
 A: Can you think of a specific example?
 K: Well, my boyfriend made me come here.
 A: Your boyfriend made you come here.
 K: He says I'm depressed much of the time.
 A: I am sorry to hear you are depressed.
 K: It's true. I am unhappy.
 A: Do you think coming here will help you not to be unhappy?
 K: I need some help, that much seems certain.
 A: What would it mean to you if you got some help?
 K: Perhaps I could learn to get along with my mother.
 A: Tell me more about your family.

Arvuti vestleb kasutajaga näiliselt arukalt, kuid see on saavutatud väga lihtsate vahenditega: ELIZA kasutab võtmesõnade andmebaasi (vt tabel 1). Iga võtmesõna jaoks on antud 1. nn järk (täisarv), 2. šabloon, millega võrrelda sisendit, 3. väljundi (vastuslause) spetsifikatsioon. Analüüsisid kasutaja lauset, leiab ELIZA selles võtmesõna, mille šabloon sobib lausega (nt *alike*). Kui võtmesõnu on rohkem kui üks (nt esimeses lauses *are* ja *alike*), siis valib ta neist sellise, mille järk on suurim (antud juhul *alike*), ning vastuslausena kasutab võtmesõnaga seotud väljundi spetsifikatsiooni (antud näites *In what way?*). Kui analüüsitavas lauses ei leidu ühtki võtmesõna, siis annab ELIZA "ohutu" vastuse, nt *Tell me more...* (nagu joonisel 2 toodud näite viimases reas) või *Go on*.

Tabel 1. Väljavõte ELIZA andmebaasist (Weizenbaum 1966). X, Y – suvalised tekstid

Võtmesõna	Järk	Šabloon	Väljundi spetsifikatsioon
<i>alike</i>	10	X	<i>In what way?</i> <i>What resemblance do you see?</i>
<i>are</i>	3 3	<i>X are you Y</i> <i>X are Y</i>	<i>Would you prefer it if I weren't Y?</i> <i>What if they were not Y?</i>
<i>always</i>	5	X	<i>Can you think of a specific example?</i> <i>When?</i> <i>Really, always?</i>
<i>what</i>	2	X	<i>Why do you ask?</i> <i>Does that interest you?</i>

ELIZA eeskujul on koostatud ka mitmeid eesti keeles suhtlevaid programme, nt 1999. a tollase informaatikaüliõpilase Mart Tudre programm Liisbet, mis valmis kursuse "Sissejuhatus arvutuslingvistikasse" raames. Dialogis (2) on toodud katke Liisbetiga.

(2) Liisbet – eestikeelne Eliza (Tudre 1999). K – kasutaja, A – arvuti, s.o Liisbet
 A: Mina olen Sinu algaja psühhoanalüütik-abimees Liisbet.
 Räägi palun mulle oma probleemidest.
 K: Mul on raske suhelda inimestega.

A: Ah nii.

K: Ma ei leia nendega kontakti.

A: Oled Sa kindel?

K: Muidugi, selles ei maksa kahelda.

A: Miks mitte?

K: Sellepärast, et kõik ütlevad, et ma olen halb suhtleja.

A: Kas Sulle meeldib olla halb suhtleja?

Kumbki, ei ELIZA ega Liisbet, ei mõista keelt, vaid reageerivad vestluskaaslase lausetest leitud märksõnadele ning on seega suutelised vestluses etendama üksnes suhteliselt passiivset rolli. Tõsi, mõnes piiratud rakenduses, kus inimesele on esmatähtis, et teda kuulataks ja aeg-ajalt kuuldel olekust märku antaks, võib seegi piisavaks osutuda.

Kuidas ära tunda, et arvuti saab vestluskaaslasest ka tegelikult aru? Inglise matemaatik A. M. Turing pakkus selleks nn jäljendamismängu, mida hiljem on hakatud nimetama Turingi testiks. Mäng seisneb järgnevas (Turing 1950, 2005). Eksperimentaatori eest varjatult paigutatakse eraldi ruumidesse inimene ja testitav arvuti ning suhtlus toimub loomulikus keeles nt läbi arvutivõrgu. Eksperimentaator esitab küsimusi, kusjuures ta ei tea, kas vastajaks on inimene või arvuti. Kui eksperimentaator ei suuda kindlaks teha, et üks vastajatest on arvuti, siis on arvuti testi läbinud ja võib öelda, et ta on mõtlemisvõimeline. Turing toob ka näite dialoogist, millega arvuti peaks toime tulema (3).

(3) Näide A. M. Turingi artiklist (Turing 2005). K – küsija, A – arvuti

K: Kas teie soneti esimesse ritta – “Kas suvepäeva sarnaseks sind kiita” ei sobiks sama hästi või paremini “kevadepäeva”?

A: Rütm läheks sassi.

K: Aga “talvepäeva”? See sobiks rütmiga.

A: Jah, aga talvepäeva moodi ei taha ju keegi olla.

K: Kas hr Pickwick ei meenuta mitte jõule?

A: Mõnes mõttes küll.

K: Aga jõulud on ju talvel ja vaevalt hr Pickwick sellist võrdlust pahaks paneks.

A: Nalja teete. “Talvepäeva” all peetakse silmas ikka mõnda tüüpilist, mitte näiteks jõululaulpäeva moodi erilist talvepäeva.

Ajal kui A. M. Turing oma artikli kirjutas, oli arvutite tootmine alles algusjärgus ning esimesed arvutid üsna tagasihoidlike võimetega. Nüüdseks, üle viiekümne aasta hiljem, ei ole ikka veel suudetud luua arvutit, mis tuleks toime Turingi testiga.

(4) Vestlus Alice’iga

PROGRAM[02:19:16] My name is ALICE.

JUDGE05[02:19:37] Good morning Alice. Why have you entered the Loebner Prize?

PROGRAM[02:19:37] Hello and how are you this morning? The Loebner Prize is an annual Turing Test sponsored by New York philanthropist Hugh Loebner. Dr. Wallace and ALICE won it in 2000. The ALICE program was ranked “most human computer” by the panel of judges.

JUDGE05[02:20:14] So you are here today to defend your title?

PROGRAM[02:20:14] Interesting deduction. Do you mind if I tell other people you say I am here today to defend my title.

JUDGE05[02:20:25] Not at all

PROGRAM[02:20:25] Next question?

JUDGE05[02:21:00] What will you spend the prize money on if you win?

PROGRAM[02:21:00] I don't have any plans yet.

1990. a pani ameerika insener dr Hugh Loebner välja kuldmedali ja 100 000 dollari suuruse auhinna arvutiprogrammile, mis esimesena läbib Turingi testi, ja algatas sellega iga-aastase võistluse.⁵ Kolm korda (2000, 2001 ja 2004) on võitjaks tulnud ning pronksmedali ja 2000-dollarilise lohutusauhinna pälvinud programm Alice⁶ (autor Richard Wallace), mida mõnigi kohtunik on pidanud inimeseks. Näites (4) on toodud katkend 2001. a võistluste protokollist⁵. Kahel viimasel aastal (2005 ja 2006) võitis programm Jabberwacky⁷ (autor Rollo Carpenter), kus rakendatakse masinõpet.

Programmide, mille peaülesanne on inimesega loomulikus keeles suhelda (nn juturobotid), pakuvad küll meelelahutust, aga samas ka innustavad uurijaid oma programme järjest täiustama ja uusi lahendusi välja mõtlema. Lihtsast võtmesõnapõhisest keeletööstusest siin muidugi ei piisa.

3. Praktilised dialoogsüsteemid

3.1. Esimesed küsimus-vastussüsteemid

Küsimus-vastussüsteemide loomine algas 1960. aastate lõpul. Peagi eraldusid kaks suunda: akadeemiline ja insenerlik.

Akadeemilised süsteemid põhinesid sügaval teooriatel, järgisid tehisintellekti traditsiooni, modelleerides arutlust ja keele mõistmist maailmateadmuse põhjal. Üheks tähelepanuväärseks selle suuna esindajaks oli süsteem Shrdlu⁸ (autor Terry Winograd, 1972) – roboti käe mudel arvuti ekraanil nn kuubikute maailmas, mis demonstreeris kasutaja ingliskeelsetest käsklustest arusaamist, vastates neile inglise keeles ja ühtlasi sooritades nõutud tegevusi (vt dialoog 5).

(5) Dialoog Shrdluga

Inimene: find a block which is taller than the one you are holding and put it in the box

Shrdlu: OK.

Inimene: what does the box contain?

Shrdlu: The blue pyramid and the blue block.

Inimene: what is the pyramid supported by?

Shrdlu: The box.

Inimene: how many blocks are not in the box?

Shrdlu: Four of them.

Insenerisüsteemides seevastu tähtsustati täitmisomadusi. Üks tuntumaid süsteeme – PARRY⁹ (autor K. M. Colby, 1971) – jäljendas paranoikut. Programm kasutas ligi 6000 mustrit, millega võrdles sisendit, et sellele õigesti reageerida. See võimaldas tal pidada tõepäraseid dialooge, tehes aeg-ajalt kõrvalekaldeid initsiatiivi

⁵ <http://www.loebner.net/Prizeof/loebner-prize.html> (8.09.2006)

⁶ <http://www.alicebot.org/> (8.09.2006)

⁷ <http://www.jabberwacky.com/> (8.12.2006)

⁸ <http://hci.stanford.edu/~winograd/shrdlu/> (8.09.2006)

⁹ <http://www.faqs.org/rfcs/rfc439.html> (8.09.2006)

haaramiseks, mida võis hea tahtmise juures panna ka paranoilise käitumise süüks. Näites (6) püüab Parry viia juttu kihlvedudele – teemale, mida ta hästi valdab.

3.2. Ratsionalism ja empirism

Akadeemiline ja insenerlik suund esindasid dialoogsüsteemide loomisel kumbki kahte erinevat paradigmat – vastavalt reeglipõhist (e ratsionaalset) ja andmepõhist (e empiirilist). Ratsionalism oli kuni 1990. aastate alguseni ainuvaldav. Andmepõhine dialoogitöötlus oli selle ajani raskendatud, sest erinevalt tekstikorpustest (mille kasutuselevõtt tekitas pöördet masintõlkes juba 1980. aastate lõpul, vt Brown jt 1990), hakati dialoogikorpusi looma alles 1990. a. Tõsi, kõnetuvastusel ja -sünteesil oli juba varem rakendatud statistilisi meetodeid, kuid kõnetöötlust oli käsitletud dialoogist eraldi, olgugi et kõne fenomenid (pausid, rõhud, kontuurid jms) annavad edasi tähendusi ja aitaksid dialoogis kaasvestlejast aru saada. 1990. aastatel jõudis empirism ka akadeemilistesse dialoogsüsteemidesse.

(6) Dialoog Parryga

Inimene: Good afternoon. My name is Doctor Jones.

PARRY: Pleased to meet you.

Inimene: How are you?

PARRY: All right I guess.

/---/

Inimene: I thought you might have some problems.

PARRY: I've been upset lately.

Inimene: By what?

PARRY: People get on my nerves sometimes.

Inimene: Tell me more.

PARRY: Do you know anything about bookies?

Inimene: No. What about bookies?

PARRY: A bookie didn't pay me off once.

Inimene: When did this happen?

PARRY: A couple of months ago.

Üks parimaid sel ajal loodud süsteeme oli TRAINS¹⁰ (autor David Traum Rochesteri ülikoolist, 1994), kus võeti kasutusele dialoogikorpus, aga samas säilitati ka loogiline arutlus. TRAINS pidas kasutajaga inglise keeles dialoogi kooperatiivse ülesande lahendamiseks – kaubavedude organiseerimine Ameerika Ühendriikide idarannikul, kasutades abivahendina ka kaarti. Süsteemis olid realiseeritud kõnetuvastus ja -süntees. Näites (7) on esitatud katkend dialoogist süsteemiga TRAINS, mis demonstreerib, et süsteem mõistab kasutajat ja oskab talle ülesande lahendamisel nõu anda.

(7) Dialoog süsteemiga TRAINS (kõne on transkribeeritud)

utt1 : s(system): hello <sil> can I help you

utt2 : u(ser): yeah I want t- I want to determine the maximum number of boxcars of oranges <sil> by seven a.m. <sil> tomorrow morning

utt3 : so <brth> hm <sil>

so I guess all the boxcars will have to go through oran- <sil> through Corning because that's where the orange juice <brth> orange factory is
 utt4 : so from Corning to Bath how far is that
 utt5 : s: two hours
 utt6 : u: and it's gonna take us also an hour to load <sil> boxcars right
 utt7 : s: right + +
 utt8 : u: + okay + so <sil> hm so <sil> every trip will take at least <sil> three hours <sil> then
 utt9 : um
 utt10 : s: right we can unload any amount of cargo onto a train in one hour
 utt11 : so we can + <sil> so a maximum of three + boxcars in an hour

3.3. Empirismi võidukäik

Praegu domineerib dialoogitöötles, nagu loomuliku keele automaattöötles üldse, empirism. Seda toetab arvestatav hulk märgendatud dialoogikorpusi, millest tuntumad on Switchboard, Verbmobil, BNC (McTear 2004). See võimaldab kasutada lausungis dialoogiakti tuvastamisel masinõpet ja dialoogi juhtimisel dialoogiaktide järjendite statistikat.

Värsketest uurimustest on tähelepanuväärne lähenemine, mis käsitleb dialoogi kui osaliselt vaadeldavat Markovi otsustusprotsessi (Partially Observable Markov Decision Process) (Williams, Young 2006). Ühe konkreetse dialoogieku asemel uuritakse kõikvõimalike olekute tõenäosusjaotust, nn arvamusolekut (ingl *belief state*), mis dialoogi käigus võib muutuda. Selliseid otsustusprotsesse kasutatakse tegevuste valikuks olukorras, kus maailm (olekute ruum) ei ole otseselt vaadeldav. Kui tuua hästi arusaadavaid võrdlusi, siis nt male on *otseselt* vaadeldav mäng; kõigi malendite positsioon alati nähtav. Seevastu pokker on *osaliselt* vaadeldav: teades oma kaarte, vastaste pakkumisi ja laual olevaid kaarte, võib ainult oletada, millised kaardid on vastastel. Sarnasel viisil on ka dialoog osaliselt vaadeldav protsess.

Praegustes dialoogsüsteemides kasutatakse nelja lähenemist.

1. Loogikapõhised süsteemid, kus rakendatakse arutlust (täitmisomadustele ei pöörata erilist tähelepanu).
2. Kõnetehnoloogiat, masinõpet ja lihtsaid vahestruktuure kasutavad süsteemid.
3. Lihtsad olekuautomaadid inimese ja arvuti vaheliste suuliste dialoogide kirjeldamise keeles VoiceXML¹¹. Selles keeles esitatud dokumentide interpreteerimiseks on loodud spetsiaalsed häälebrauserid. Süsteemi ülesehitamiseks ühendatakse brauserite andmepangad avaliku telefonivõrguga. Sellisel viisil on loodud palju juturoboteid ja mitmesuguseid kommertsüsteeme.
4. Hüübriidsüsteemid, mis ühendavad reeglid ja masinõppe.

¹¹ <http://www.voicexml.org/> (8.09.2006)

3.4. Kõnedialoogsüsteemide näiteid

Maailmas on dialoogsüsteemide arendamine kestnud juba ligi neli aastakümnet. Esimesed lihtsad süsteemid, mis suhtlevad kasutajaga kõne abil, jõudsid kommerts-rakendusteni siiski alles 15 aastat tagasi. Need on mitmesugused teadetesüsteemid (annavad liiklusinfot, pangainfot vms), ärikokkulepete sõlmimise (pangalaenu saamise, autokindlustuse, ostuabi jms) süsteemid, probleemilahendusel (nt arvutirikke tuvastusel) abistavad süsteemid.

Alljärgnevalt on ülevaade mõnede tuntud firmade dialoogsüsteemidest, millega saab suhelda kõne abil.

- **Nuance** (Ameerika Ühendriigid) võimaldab kasutada häälvälimisega telefoni, e-posti ettelugemist, kindlustust, pangateenust, annab reisiinfot jms.¹²
- **AT&T Labs** (American Telephone and Telegraph, Ameerika Ühendriigid) pakub kellaja teatamise, krediidiinfo jms süsteeme.¹³
- **Philips** (Holland) toodab dikteerimissüsteeme, telefoniteenuste süsteeme.¹⁴
- Rootsi firma **Telia** ja Soome firma **Nokia** rakendavad mõlemad oma firmas telefonikõne suunamist töötaja nime järgi; Telia säästab sellega väidetavalt 40 miljonit Rootsi krooni aastas.
- **Loquendo** (Itaalia) on välja töötanud rongiinfosüsteemi,¹⁵ mille andmebaasis on 3000 linna/jaama Itaalia raudteel. Süsteem saab 2 miljonit telefonikõnet kuus.

Dialoogsüsteemide arendamisega tegeldakse paljudes ülikoolides ja uurimisinstituutides.

- Massachusettsi Tehnoloogiainstituudis on loodud mitmeid dialoogsüsteeme,¹⁶ mis annavad lennuinfot koos piletihindadega, lennukite saabumise ja väljumisaegadega, liiklusinfot ja restoraniinfot Bostoni piirkonnas, ilmateateid (enam kui 500 linna kohta üle kogu maailma).
- Carnegie Melloni ülikoolis on loodud kõnedialoogsüsteem **CMU Communicator**,¹⁷ mis aitab planeerida reisi (lennuinfo, pileтите, hotelli reserveerimine, autorent).
- Soome ülikoolide ning mõnede Soome ja Jaapani firmade koostöös on lõpule jõudnud projekt **Interact**¹⁸ dialoogsüsteemi loomiseks, mis võimaldab suunata telefonikõnesid, reserveerida pileteid, saada infot Helsingi bussisõiduplaanide kohta, interaktiivset abi probleemilahendusel.
- Rochesteri ülikoolis on loodud dialoogsüsteemi TRAINS edasiarendus TRIPS¹⁹ (The Rochester Interactive Planning System), kus rakendusstsenariumiks on inimeste päästmine katastroofi piirkonnast, seejuures on aeg ja transpordivahendite arv piiratud. Kasutatakse ka teede kaarti.
- Saksamaa Tehisintellekti Uurimisinstituudis on professor Wolfgang Wahlsteri juhtimisel läbi viidud mitmeid suuri projekte inimesega loomulik keeles suhtlevate dialoogsüsteemide loomiseks.

¹² <http://www.nuance.com/speech/> (8.09.2006)

¹³ <http://www.att.com/atllabs/> (8.09.2006)

¹⁴ <http://www.philips.com/> (8.09.2006)

¹⁵ <http://www.loquendo.com/> (8.09.2006)

¹⁶ <http://www.sls.lcs.mit.edu/applications> (8.09.2006)

¹⁷ <http://www.speech.cs.cmu.edu/Communicator/> (8.09.2006)

¹⁸ <http://www.mlab.uiah.fi/interact/> (8.09.2006)

¹⁹ <http://www.cs.rochester.edu/research/trips> (8.09.2006)

- **Verbmobil**²⁰ (1993–2000) on kohtumiste kokkuleppimise süsteem, mis hõlmab ka reisiplaneerimise ja kõike sellega seonduva (nt hotelli reserveerimine) ning ühtlasi võimaldab automaatselt tõlkida kõnet ühest keelest teise (valida on kolme keele vahel: saksa, inglise ja jaapani keel).
- **SmartKom**²¹ (1999–2003) on mitmefunktsiooniline dialoogsüsteem, millel on kolm põhilist rakendusstsenaariumi:
 - *SmartKom-Public* – multimodaalne telefoniautomaat (kaasatud on nii kõne kui ka žestid ja miimika),
 - *SmartKom Mobile* – Interneti-teenused mobiiltelefoni teel,
 - *SmartKom-Home / Office* – arvuti häälujuhtimine mitmesugusteks rakendusteks, sh dokumentide dikteerimine, infootsing Internetist jms.
- VirtualHuman²² (2002–2006). Eesmärk on arendada virtuaalseid (emotsionaalseid ja fotorealistlikke) karaktereid. Põhitähelepanu pööratakse personaalsete dialoogipartnerite loomisele, mis tegutsevad reaajas ja autonoomselt. Ühe võimaliku rakendusstsenaariumina nähakse e-õpet.

Tuleb rõhutada, et praegu on aktuaalseks uurimisprobleemiks multimodaalne suhtlus arvutiga, mis lisaks loomulikule keelele (eeskätt suulisele keelele) hõlmab ka žeste, miimikat, sh emotsioonide edasiandmist ja tõlgendamist jms.

Üks suurejoonelisemaid projekte on *Cyc* (sõnast *encyclopedia*), mida alustati Ameerika Ühendriikides juba 1984. a. Projekti eesmärk on luua väga mahukas mõisteid, fakte ja reegleid koondav teadmusbasis, mille abil tehisintellektisüsteem suudaks arutleda, õppida ja inimesega loomulikus keeles suhelda. Juulis 2006 valminud versioon sisaldab 47 000 mõistet ja 306 000 fakti.²³

4. Eesti keeles suhtlev arvuti

Tartu Ülikoolis tegutseva keeleteadlaste ja informaatikute rühma eesmärk on modelleerida arvutil loomulikku dialoogi, s.t sellist dialoogi, mis toimub loomulikus (eesti) keeles ning inimestevahelise suhtluse normide ja reeglite järgi (Koit, Õim 2003).

Siiski pole kõik uurijad üksmeelel selles, kas arvuti peaks inimesega suhtlema täpselt nii nagu inimene. On püstitatud nn arvutikõne hüpotees (Zoeppritz 1985), mis väidab, et inimesed ei oota arvutilt, et ta suhtleks nendega nagu inimene, vaid kohanevad arvuti piiratusega. Ometi on selge, et mida kaugemale tehnoloogia areneb, seda vähemaks jääb selliseid piiranguid. Nii või teisiti, dialoogsüsteemi kavandamisel tuleb uurida nii inimestevahelist kui ka inimese ja arvuti vahelist suhtlust, et kinnitada või ümber lükata arvutikõne hüpoteesi. Kuni dialoogsüsteemi pole veel loodud, ei saa muidugi koguda inimese ja arvuti vahelisi dialooge. Siis kasutatakse nn võlur Ozi meetodit (Dahlbäck jt 1993): inimestel palutakse justkui testida programmi, mis suhtleb nendega loomulikus keeles, kuid tegelikult etendab arvuti rolli arvutivõrgu kaudu teine inimene (“võlur Oz”). Teine võimalus on olemasolevate inimestevaheliste dialoogide nn destilleerimine, eesmärgiga muuta neid “arvutilaadseteks” (Jönsson, Dahlbäck 2000).

²⁰ <http://verbmobil.dfki.de/> (8.09.2006)

²¹ <http://www.smartkom.org/> (8.09.2006)

²² <http://www.virtual-human.org/> (8.12.2006)

²³ <http://www.cyc.com/> (8.12.2006)

Kogutud dialoogid koondatakse korpustesse ning dialoogikorpusi märgendatakse mitmel tasemel: fonoloogilisel, morfoloogilisel, süntaktilisel, dialoogiaktide tasemel jne.

4.1. Eesti dialoogikorpused

Eesti dialoogikorpuse (EDiK) loomist alustati Tartu Ülikoolis 2001. aastal (Koit 2003). Korpus sisaldab praegu kolme liiki dialooge:²⁴

1. inimestevahelised dialoogid Tartu Ülikooli suulise eesti keele korpusest.²⁵ Seisuga september 2006 sisaldab korpus 884 litereeritud telefonikõnet ja silmast-silma vestlust, kokku u 155 000 tekstisõna;
2. võlur Ozi meetodil kogutud kirjalikud dialoogid, kus arvuti rolli on etendanud inimene, 21 reisiinfodialoogi (2500 sõna);
3. lihtsate veebis kasutatavate dialoogsüsteemidega²⁶ Reisiagent ja Teatriagent peetud dialoogid (programmide autor Margus Treumuth).

Osa suulisi dialooge on märgendatud morfoloogiliselt ja süntaktiliselt. Selleks on kasutatud eesti kirjakeele morfoloogilist analüsaatorit ESTMORF (Kaalep 1998) ja eesti keele kitsenduste grammatika süntaktilist analüsaatorit (Müürisep 2000), mida on kohandatud suulisele keelele. Kõigis dialoogides (v.a Reisi- ja Teatriagentiga peetud dialoogid) on märgendatud dialoogiaktid. Selleks on Tartu Ülikoolis välja töötatud vestlusanalüüsil põhinev dialoogiaktide tüpoloogia, mis sisaldab üle 120 akti (Hennoste, Rääbis 2004). Iga dialoogi märgendavad teineteisest sõltumatult kaks inimest ja kolmas ühtlustab. Aktid jagunevad naaberpaare moodustavateks ja üksikaktideks. Naaberpaaril on esi- ja järelliige, mida väljendatakse aktide nimedes. Aktide nimed koosnevad kahest osast: akronüüm, mis näitab akti paiknemist tüpoloogias ja naaberpaariakti puhul lisaks, kas on tegu esi- või järelliikmega (nt RIE on rituaalide rühma esiliige, IL on üksikaktide rühm infolisa), ja akti pärisnimi (nt KUTSUNG, TÄPSUSTAMINE). Need kolm dialoogitüüpi on esitatud näidetes (8–10).

(8) EDiK: kahe inimese telefonidialoog. H – helistaja (klient), V – vastaja (ametnik). Kasutatud on vestlusanalüüsi transkriptsiooni (Hennoste, Rääbis 2004)

((kutsung) | RIE: KUTSUNG |

V: bussi'info=kuuleb= | RIJ: KUTSUNGI VASTUVÕTMINE | | RIE: ESITLUS |

tere | RIE: TERVITUS |

H: 'hh tere. | RIJ: VASTUTERVITUS |

.h mul on=nüd selline 'küsimus. |YA: EELTEADE |

(.) mt e < 'mis (0.5) e 'kell 'läheb (.) selline 'buss mis jõuaks > (.) e 'Soome mineva 'laeva 'peale. | KYE: AVATUD |

'Soome minev 'laev läheb 'Tallinast (.) kell='kaheksa. | YA: INFOLISA: TÄPSUSTAMINE |

(0.5)

'kas 'lähevad bussid ka: 'sadamasse. | KYE: SULETUD KAS |

V: ei, | KYJ: EI |

²⁴ <http://math.ut.ee/~koit/Dialoog/EDiC.html> (8.09.2006)

²⁵ <http://www.cl.ut.ee/suuline/> (8.09.2006)

²⁶ <http://math.ut.ee/~treumuth> (8.12.2006)

´sadamasse läheb ainult ´üks´ainuke ´buss, | KYJ: INFO ANDMINE |
see sõidab ´Tartust väljust välja ´kaksteist neligend=´viis. | IL: TÄPSUS-
TAMINE |

Korpusesse kogutud dialoogide analüüs võimaldab vastata mitmesugustele küsi-
mustele selle kohta, missugune peaks olema suhtlus arvutiga võrreldes inimeste-
vahelise suhtlusega, sh missugune peaks olema arvuti sõnavara ja grammatika,
kuidas lahendada mitmesusi, kuidas kasutada konteksti jms (vt nt Gerassimenko,
Valdisoo 2006).

(9) EDiK: dialoog võlur Oziga (sisend klaviatuurilt, väljund ekraanile)
Infoklient: mis kell väljub buss Tartust Võrru, et ma oleksin Võrus juba kell
17.00? | KYE: AVATUD |
Infoklient: mind huvitab lähim buss | IL: TÄPSUSTAMINE |
VOZ: Oodake, palun! | DIE: PALVE OODATA |
VOZ: Buss väljub kell 14.15 ja saabub Võrru kell 15.50 | KYJ: INFO
ANDMINE |
Infoklient: Kas see buss sõidab läbi Põlva? | KYE: SULETUD KAS |
VOZ: Jah. | KYJ: JAH |
Infoklient: Mis kell ta peatub Põlvas? | KYE: AVATUD |
VOZ: Buss peatub Põlvas kell 15.15 | KYJ: INFO ANDMINE |
VOZ: Kas saan Teile veel kasulik olla? | KYE: SULETUD KAS | | TVE:
TEEMAALGATUSE PAKKUMINE |

(10) EDiK: dialoog arvutiga (Reisiagent; sisend klaviatuurilt, väljund ekraanile
ja sünteeskõnes). K – kasutaja, A – arvuti
K: tere!
A: Tere. Kuhu soovite lennata?
K: soovin lennata itaaliasse
A: Itaalia puudub lennuandmebaasist.
K: Siis tahan lennata prantsusmaale
A: Pariisi saab lennata teispäeval kell 08:00.
K: tahaksin lennata hoopis 6. aprillil
A: Neljapäeval saab Pariisi lennata kell 08:00.
K: kust ma piletit osta saan?
A: Pileti saate Estraveli kassast.

Soovides luua sellist dialoogsüsteemi, mis täidaks infoametniku rolli, peame
eeskujuna silmas ideaalset ametnikku, kes on viisakas, koostöövalmis, korrektse
keelekasutusega jms. Näide (11) on inimestevahelise suulise dialoogi destilleeri-
mise kohta, kus on muudetud ametniku (V) keelekasutust ja on lisatud kliendi (H)
küsimus. Destilleeritud dialoogis võiks ametniku rolli täita arvuti.

(11) Dialoog (a) enne ja (b) pärast destilleerimist
(a)
H:
(0.5)
´kas ´lähevad bussid ka: ´sadamasse.
V: ei,
´sadamasse läheb ´üks´ainuke ´buss,

see sõidab 'Tartust väljust välja 'kaksteist neligend= 'viis.

(b)

H:

(0.5)

'kas 'lähevad bussid ka: 'sadamasse.

V: ei,

'sadamasse läheb ainult 'üks 'buss,

H: millal see väljub?

V: see sõidab 'Tartust välja 'kaksteist nelikümmend= 'viis.

4.2. Loomuliku keele töötlus dialoogsüsteemis

Dialoogsüsteem täidab tsükliliselt kolme järjestikust protsessi: 1. kasutaja lausungi (nt küsimuse) analüüs, 2. probleemilahendus (leidmaks vastust küsimusele), 3. oma lausungi süntees. Analüüsi võib omakorda jagada etappideks: kõne teisendamine tekstiks (kõnetuvastus) ja seejärel teksti morfoloogiline, süntaktiline ja semantiline analüüs. Samasugusteks etappideks (vastupidises järjekorras) saab jaotada ka vastuse sünteesi. Probleemilahendus on ainus protsess, mis ei sõltu keelest ja kasutab teatava formalismi abil esitatud maailmateadmust, andmebaase, mitmesuguseid universaalseid ja spetsiifilisi probleemilahendusmeetodeid.

Eesti keele automaattöötluks on valminud juba mitmesugust tarkvara. Sellest annab ülevaate tabel 2.

Morfoloogiatarkvara on erinevatel teoreetilistel alustel loodud nii Tartu Ülikoolis (koos firmaga Filosoft) kui ka Eesti Keele Instituudis (Kaalep 1998, Viks 2000). Eesti keele kitsenduste grammatika baasil on koostatud reeglipõhine süntaktiline analüsaator (Müürisep 2000) ja morfoloogiline ühestaja (Puolakainen 2001). Firmas Filosoft on lisaks loodud teine morfoloogiline ühestaja, mis põhineb Markovi peitmodelil. Eesti keele sõnatähenduste ühestaja (Kaljurand 2004) kasutab mitmetähendusliku sõna vaadeldavasse konteksti sobiva tähenduse väljavalimiseks eesti WordNetti²⁷ ning kitsenduste grammatika süntaksianalüsaatorit. Dialoogiaktide klassifitseerimisel rakendatakse tehisnärvivõrke ja otsustuspuud (Fišel, Kikas 2006), kusjuures klassifitseerimise aluseks on Eesti dialoogiaktide tüpologia ja klassifitseerimisprogramme treenitakse Eesti dialoogikorpusel. Tekst-kõnesüntesaatoris rakendatakse kompilatiivset meetodit, kasutades difoonide andmebaasi (Mihkla, Meister 2002). Kõnetuvastaja on statistikapõhine (Alumäe 2003).

Tabel 2. Eesti keele tarkvara

Keeletöötlu tase	Tarkvara	Asutus
Morfoloogia	Morfoloogiline analüsaator ja süntesaator ^{28, 29}	TÜ, Filosoft, EKI
Süntaks	Kitsenduste grammatika süntaktiline analüsaator, ³⁰ morfoloogiline ühestaja	TÜ TÜ, Filosoft
Semantika ja pragmaatika	Sõnatähenduste ühestaja, ³¹ dialoogiaktide klassifitseerija	TÜ
Kõnetöötlus	Tekst-kõne süntesaator, ²⁸ kõnetuvastaja	TTÜ, EKI

²⁷ <http://www.cl.ut.ee/ressursid/teksaurus> (8.09.2006)

²⁸ http://www.filosoft.ee/html_morf_et/ (8.09.2006)

²⁹ <http://www.eki.ee/tarkvara/> (8.09.2006)

³⁰ <http://www.cs.ut.ee/~kaili/parser/demo/> (8.09.2006)

³¹ <http://math.ut.ee/~kaarel/NLP/> (8.09.2006)

Eespool nimetatud dialoogsüsteemid Reisiagent ja Teatriagent kujutavad endast katseid ühendada olemasolevat eesti keele tarkvara lihtsate küsimus-vastusdialoogide läbiviimiseks. Mõlemal juhul on tegu kitsalt piiritletud ülesandega: Reisiagent vastab kasutaja küsimustele lennukite väljumisaegade kohta Tallinna lennujaamast ning Teatriagent annab infot Eesti teatrite mängukava osas. Mõlema süsteemiga suhtlemine toimub sel viisil, et kasutaja sisestab veebibrauseris päringuaknasse oma küsimuse ja näeb ekraanil vastust ning soovi korral kuuleb sedasama sünteeskõnes. Teatriagentile on lisatud ka kõnetuvastusmoodul (mitte veel veebiversioonis), nii et kasutaja saab oma küsimuse esitada hoopis suuliselt. Kirjaliku eesti keele töötlemise tarkvarast on mõlemas süsteemis rakendatud üksnes morfoloogilist analüüsi ja sünteesi. See on siin otstarbekohane: piiratud ainevaldkonna korral ei ole vaja läbi teha analüüsi täistsükli, mõlemad süsteemid otsivad kasutaja küsimusest ainult sobivaid märksõnu (nt kuupäeva, teatri või linna nime jms) ja vastuse sünteesimise asemel saab kasutada valmis vastuste šabloone, milles üksnes täidetakse tühje kohti (Treumuth 2006).

4.3. Eestikeelsed kasutajaliidesed

Paljud arvutiprogrammid peavad suuremal või vähemal määral suhtlema kasutajaga, kas või ainult algandmete sisestamiseks ja tulemuste väljastamiseks. Näiteks programm, mis arvutab kolmnurga pindala kolme külje järgi, teatab kasutajale:

Sisestage kolmnurga küljepikkused

a =

b =

c =

ja väljastab tulemuse

Kolmnurga pindala S = ...

Selline kasutajaliides ei nõua küll loomuliku keele analüüsi ega sünteesi, kuid ometi peab programmeerija suutma ette näha kasutaja käitumist, võimalikke veaolukordi ja tagama, et kõik teated, mida arvuti väljastab, oleksid informatiivsed, keeleliselt korrektsed ja üheselt mõistetavad. Kui eelmises näites sisestab kasutaja negatiivse arvu, siis võiks järgneda veateade

Küljepikkus peab olema positiivne, sisestage uuesti.

Dialoogis (12) on toodud näide keelemängu Alias eestikeelsest kasutajaliidestest.³² Programm võtab sõnaseletused eesti WordNetist.

(12) Dialoog programmiga Alias. Kasutaja vastused algavad märgiga >

Alustame mängu

Niisiis uus sõna

isik, kellel on loomuniguline³³ ja kujutuslik tegevus

>kunstnik

Näitelause:

Mulle meeldis väga selle kunstniku näitus

>artist

³² <http://www.cl.ut.ee/lahe/> (5.05.2006)

³³ Siin on ilmselt tegu veaga andmebaasis (peaks olema: loomuniguline).

kunstnik
>maalija
on teatud liiki looja
>maalikunstnik
Selle sõna teatud liigid on esineja, maalija, disainer, komponist
>interpreet
Selle sõna üks liik on dirigent
>loomeinimene
Rohkem seletusi mul anda pole. Võid veel pakkuda
>kunstiinimene
kunstiinimene on õige :)

5. Suhtlus arvutiga tulevikus

Tuntud arvutiteadlane, Sheffieldi ülikooli professor Yorick Wilks näeb dialoogsüsteemide tulevikku robot-kaaslastes, mida ta nimetab kompanjonideks (Wilks 2006). Kompanjon on konversatsiooniant, juurdepääsuga Internetile. Ta on isikustatud oma omaniku suhtes: teab tema vajadusi, eelistusi, harjumusi ning õpib omaniku kohta süstemaatiliselt juurde. Ta võib olla näiteks mänguasja- või lemmikloomakujuline, mida saab ranitsas kaasas kanda. Eraldi väärivad märkimist seniorkompanjonid vanuritele ja juuniorkompanjonid lastele. Seniorkompanjon võib oma omanikule meelde tuletada huvitava telesaate algust, lülitades sisse televiisori, ravimi võtmise aega, vanuriga lihtsalt suhelda, vajaduse korral talle abi kutsuda jms. Juuniorkompanjon hoiatab kooliteed alustavat last ohtude eest liikluses, aitab koolitükke ette valmistada, otsides infot Internetist, ajakavast kinni pidada jne.

Kompanjonid on uus agentide liik vastukaaluks passiivsele veebile: veeb peab muutuma personaalsemaks, et olla hoomatav ka kiirelt kasvades. Semantiline veeb on esimene samm veebi struktureerimiseks, aga praegused veebiagendid on siiski veel abstraktsed ja ajutised. Personaalsed agendid peavad olema autonoomsed, püsivad ja usaldusväärsed ka selliste elanikkonna kihtide jaoks, kes praegu on veebist eemal.

Tulles lõpuks tagasi algul mainitud Turingi testi juurde, võib viidata Ameerika arvutiteadlase ja visionääri Raymond Kurzweili sõnavõtule³⁴ 2001. a, kus ta rõhutas, et Turingi test on märkimisväärne sellepoolest, et näitab, kuidas keel peegeldab inim mõtlemist. Praegu on olemas palju arvutiprogramme, mis suudavad kitsa ainevaldkonna probleeme lahendada niisama hästi kui inimene: malemäng, meditsiiniline diagnoosimine, finantsotsuste vastuvõtmine jpm. Seega võib väita, et arvuti suudab nendes valdkondades täita Turingi testi. Et arvuti läbiks Turingi testi täies mahus, s.t et ta suudaks suhelda nagu inimene, peaks tal olema inimese tasemel mõtlemisvõime.

R. Kurzweili arvates saab see võimalikuks aastal 2029.

6. Kokkuvõte

Maailmas on olemas tuhandeid VoiceXML-põhiseid kommertssüsteeme, mis suhtlevad inimestega kõne abil ja töötlevad miljoneid telefonikõnesid päevas. Pakutavate teenuste skaala on lai: infotelefon, hädaabi, lennu reserveerimine, postisaadetiste jälgimine, telefonikõne suunamine nime järgi, kinnisvarainfo jne.

Suund on võetud hübriidsüsteemidele, milles kombineeritakse reeglipõhiseid ja korpusepõhiseid meetodeid, ning multimodaalsele suhtlusele, kus lisaks kõnele kasutatakse miimikat ja žeste.

On loodud mitmeid roboteid, mis muuhulgas suudavad inimesega loomulikus keeles suhelda. Tulevikku nähakse sellistes robotites-kompanjonides, mis on võimalised õppima, kohanema omaniku vajadustega ja teda abistama.

Eesti keeles suhtleva arvuti arendamiseks on Tartu Ülikoolis loodud eesti dialoogikorpus, samuti on olemas mitmesugust eesti keele automaattöötluste tarkvara. Osaliselt on seda rakendatud kahes lihtsas küsimus-vastussüsteemis Reisiagent ja Teatriagent, mida saab veebis kasutada. Eesmärk on luua selline dialoogsüsteem, mis täidaks infoametniku rolli ja suudaks kasutaja dialoogiakte automaatselt tuvastada ning neile reageerida nii nagu inimene.

Kirjandus

- Alumäe, Tanel 2003. Eestikeelse kõne tuvastus: prototüübi loomine. – M. Langemets, H. Sakhai, M.-M. Sepper (toim.). Toimiv keel I. Töid rakenduslingvistika alalt. Eesti Keele Instituudi toimetised 12. Tallinn: Eesti Keele Sihtasutus, 34–49.
- Brown, Peter F.; Cocke, John; Della Pietra, Stephen A.; Della Pietra, Vincent J.; Jelinek, Fredrik; Lafferty, John D.; Mercer, Robert L.; Roossin, Paul S. 1990. A statistical approach to machine translation. – *Computational Linguistics* 16 (2), 79–85.
- Dahlbäck, Nils; Jönsson, Arne; Ahrenberg, Lars 1993. Wizard of Oz studies. Why and how. – *Knowledge-Based Systems* 6 (4), 258–266.
- Fišel, Mark; Kikas, Taavet 2006. Dialoogiaktide automaatne tuvastamine. – M. Koit, R. Pajusalu, H. Õim (toim.). Keel ja arvuti. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 6. Tartu: Tartu Ülikooli Kirjastus, 233–245.
- Gerassimenko, Olga; Valdisoo, Maret 2006. Loomulik infodialoog ja infodialoogi simulatsioon: infoandja strateegiad. – M. Koit, R. Pajusalu, H. Õim (toim.). Keel ja arvuti. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 6. Tartu: Tartu Ülikooli Kirjastus, 196–209.
- Hennoste, Tiit; Rääbis, Andriela 2004. Dialoogiaktid eesti infodialoogides: tüpoloogia ja analüüs. Tartu: Tartu Ülikooli Kirjastus.
- Jönsson, Arne; Dahlbäck, Nils 2000. Distilling dialogues. A method using natural dialogue corpora for dialogue systems development. – *Proceedings of 6th Applied Natural Language Processing Conference*. Seattle, 44–51.
- Kaalep, Heiki-Jaan 1998. Tekstikorpuse abil loodud eesti keele morfoloogiaanalüsaator. – *Keel ja Kirjandus* 1, 22–29.
- Kaljurand, Kaarel 2004. Word Sense Disambiguation of Estonian with syntactic dependency relations and WordNet. – *Proceedings of ESSLLI-2004*. Nancy, France, 128–137.
- Koit, Mare 2003. Märgendatud dialoogikorpus kui keeleressurss. – M. Langemets, H. Sakhai, M.-M. Sepper (toim.). Toimiv keel I. Töid rakenduslingvistika alalt. Eesti Keele Instituudi toimetised 12. Tallinn: Eesti Keele Sihtasutus, 119–136.
- Koit, Mare; Õim, Haldur 2003. Eestikeelse dialoogi modelleerimine. – *Keel ja Kirjandus* 10, 721–735.

- McTear, Michael F. 2004. *Spoken Dialogue Technology. Toward the Conversational User Interface*. London: Springer Verlag.
- Mihkla, Meelis; Meister, Einar 2002. Eesti keele tekst-kõne-süntees. – *Keel ja Kirjandus* 2, 88–97; 3, 173–182.
- Minker, Wolfgang; Bühler, Dirk; Dybkjær, Laila (Eds.) 2005. *Spoken Multimodal Human-Computer Dialogue in Mobile Environments. Text, Speech and Language Technology*. Vol. 28. London: Springer Verlag.
- Müürisep, Kaili 2000. Eesti keele arvutigrammatika: süntaks. *Dissertationes Mathematicae Universitatis Tartuensis* 22. Tartu: Tartu Ülikooli Kirjastus.
- Puolakainen, Tiina 2001. Eesti keele arvutigrammatika: morfoloogiline ühestamine. *Dissertationes Mathematicae Universitatis Tartuensis* 28. Tartu: Tartu Ülikooli Kirjastus.
- Turing, Alan M. 1950. Computing machinery and intelligence. – *Mind. New Series* 59 (236), 433–460.
- Turing, Alan M. 2005. Arvutusmasinad ja intellekt. – *Elo Tuur* (tõlk.). *Akadeemia* 12, 2572–2606.
- Treumuth, Margus 2006. A Natural Language Interface to a Theater Information Database. – *Proceedings of SPECOM. St.-Petersburg*, 179–181.
- Viks, Ülle 2000. Eesti keele avatud morfoloogiamudel. – T. Hennoste (toim.). *Arvutuslingvistikalt inimesele. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised* 1. Tartu: Tartu Ülikooli Kirjastus, 9–36.
- Zoeppritz, Magdalena 1985. *Computer Talk? (Technical Report 85.05)* – Heidelberg: IBM, WZH.
- Weizenbaum, Joseph 1966. ELIZA – a Computer Program for the Study of Natural Language Communication between Man and Machine. – *Communications of the ACM* 9 (1), 36–35. <http://www.acm.org/cacm> (4.05.2006).
- Wilks, Yorick 2006. Artificial Companions as a new kind of interface to the future Internet. – Oxford Internet Institute. Research Report No. 13. <http://www.oii.ox.ac.uk/research/publications.cfm> (8.12.2006).
- Williams, Jason D.; Young, Steve 2006. Partially Observable Markov Decision Processes for Spoken Dialog Systems. – <http://svr-www.eng.cam.ac.uk/~sjy/papers/wiyoo6.pdf> (8.09.2006). *Computer Speech and Language* (accepted for publication).

Mare Koit on Tartu Ülikooli keeletehnoloogia professor, uurimisvaldkondadeks on keeletehnoloogia, tehisintellekt ning dialoogi modelleerimine arvutil.
mare.koit@ut.ee

INTERACTION WITH THE COMPUTER

Mare Koit

Univesrity of Tartu

The paper investigates some problems which have to be solved in order to enable interaction with the computer in a natural language, incl. Estonian. Examples of interaction with the computer are given, starting with the Turing's idea of an intelligent computer (1950) and finishing with contemporary dialogue systems and robots that have natural user interfaces. Academic and engineering approaches to the building of dialogue systems are introduced as well as two paradigms of dialogue processing – rationalism and empiricism. An overview is given of work hitherto done on a dialogue system which interacts with a user in Estonian, incl. the Estonian dialogue corpus of the University of Tartu and software for Estonian language processing. Finally, visions of some well known scientists about the future computer as an interactive partner are intermediated.

Keywords: Turing test, computer speech hypothesis, dialogue system, robot, dialogue corpus, Estonian

ROHKEM KUI SÕNAD¹

Jaakko Lehtonen

Ülevaade. Artikkel julgustab rakenduslingviste lisama verbaalse keele uurimise paradigmasse visuaalse keele mudelid ja teooriad, kaasama visuaalset kompetentsi, visuaalse kirjaoskuse ja ladususe küsimusi diskussiooni lingvistilise kompetentsi, kirjaoskuse, ladususe jms üle. Kommunikatsiooni visuaalseid elemente kohtab kõige enam reklaamikontekstis. Arvatakse, et reklaamide visuaalsed kujutised (visuaalid) parandavad suhtumist brändi ning mõjutavad positiivselt käitumiskavatsusi. Artiklis käsitletakse visuaalide mõningaid aspekte keele ja kommunikatsiooni vaatenurgast ning antakse lühiülevaade neuroökonomika² neist tulemustest, mis on saadud magnetresonantsuuringute (Magnetic Resonance Imaging ehk MRI) toel.

Võtmesõnad: kirjaoskus, reklaam, veenmine, visuaalne keel, visuaal, MRI

Sissejuhatus

Reklaam on kõikjal. Meid ümbritsevad iga päev sajad reklaamsõnumid, rääkimata varjatud sõnumitest, mis võivad peituda kus tahes. Reklaamid töötavad sotsiaalse reaalsuse pinnal. Nad soovivad meile, mida süüa, kuidas riietuda, kuidas veeta vaba aega, seega pakuvad meile valikuid igaks elujuhtumiks. Me ei suuda põgeneda reklaamide loodud sotsiaalse maailma eest, sest see on kõikvõimas ja kõikjal meie ümber.

Sõnumit, mis on mõeldud mõjutama isiku arvamusi, emotsioone, hoiakuid ja käitumist, kutsutakse veenmiseks. Artiklis on vaatluse all kahte tüüpi veenmisaktid – propaganda ning kommerts- ja sotsiaalreklaam, kusjuures keskendutakse reklaamikommunikatsiooni visuaalsetele elementidele. Uuemad teadmised, mis on saadud brändireklaamide vaataja ajufunktsioonide salvestamisel magnetresonantsuuringutes (MRI), lubavad oletada, et visuaalne sõnum võib mõjutada ajutalitust ja alateadlikult ka meelelist taju, kuid võimalik et ka inimese käitumist.

¹ Pealkiri on laenatud Richard Dimbleby, Graeme Burtoni õpikust "More than Words. An Introduction to Communication", London: Routledge, 1992.

² Neuroökonomika (ingl *neuroeconomics*, eesti keeles sageli *neuroökonomika*) on majandusteaduse interdistsiplinaarne allvaldkond, mis ühendab neuroteaduse, ökonomika, psühholoogia ja mille eesmärk on uurida aju rolli otsuste tegemisel majanduskeskkonnas.

Et näidata visuaalide tähtsust inимtunnetuses, käsitletakse artiklis nende mõningaid aspekte keele ja kommunikatsiooni kontekstis, samuti antakse lühiülevaade MRI tulemustest.

Visuaalide uurimiskontekst

Rakenduslingviste huvitab eelkõige reklaamis ja visuaalses propagandas kasutatav veenmiskeel ja selle grammatika. Kas visuaalse keele n-õ sõnad on universaalsed, s.o samad kõigi jaoks, või tõlgendatakse veenmissõnumeid eri maades erinevalt; kas visuaalid mõjutavad ühtviisi eestlasi, ameeriklasi ja soomlasi? Kui ei, siis milline on kultuuri mõju reklaami tähendusele? Äsja käivitunud rahvusvahelise projekti esialgsed tulemused lubavad oletada, et erinevus selles, mida eri maade inimesed näevad ühes ja samas reklaamis, võib olla isegi suurem kui juleti oodata.³

Reklaami kogumahust moodustab tekst enamasti ainult väikese osa. Enamikus reklaamides kannab põhiinfot visuaalne kujutis. Mõnes reklaamis ei esine keelelist teksti üldse, teistes esinevad sellised verbaalse keele elemendid nagu ettevõtete logod või tootenimed, kuid hoopis visuaalidena. Visuaalne kujutis on kommerts-kommunikatsiooni põhielement, kuid ometigi peatume harva reklaami ees, et n-õ lugeda teadlikult pildi tähendust – mida üldse pildiga öeldakse või mis sõnumit tahab reklaamija meile lähetada, kasutades just sellist pilti.

Kui visuaalsed kujutised on nii olulised, siis miks ei ole uurijad jõudnud sügavamate teadmiseni visuaalidest ja nende toimimisest? Val Larseni, David Luna ja Laura A. Peracchio (2004) järgi eksisteerib lõhe, mis eraldab kahte akadeemilist kultuuri: mentaalseid protsesse ja visuaalsetele stiimulitele reageerimist uurib eksperimentaalpsühholoogia, reklaami tähendust aga põhiliselt lingvistika ja semiootika. Tarbijareageeringu empiirilise testimise puudumine mõlemas neis on kahandanud tulemuste praktilist paikapidavust.

Lisaks keskendub enamik lingvistika ja semiootika reklaamiuurimusi sellele, kuidas keele abil antakse edasi reklaampakkumise sisu. Kui visuaalsete kujutiste üle on üldse arutletud, siis on neid käsitatud kui reklaami illustratsioone. Sarnasust võib näha illustreeritud lasteraamatutega: raamatu põhiline väärtus seisneb piltides, kuid see, kes kirjutab piltide juurde paar värsirida või väikse teksti, on kaanel märgitud autoriks ning kunstniku nime leiab väiksemate tähtedega tema nime alt: *Illustreerinud NN*.

Tähenduse analüüsi või visuaalse kommunikatsiooniga on seotud veel üks probleem. Visuaalide analüüsimiseks kaldutakse nende tähendust ümber kodeerima lingvistiliseks koodiks, implitsiitselt oletades, et visuaalset sõnumit saab verbaalsesse keelde tõlkida. Kuid mis oleks siis, kui tähendus, mida pilt kannab, on aisting, maitse, õhkkond, midagi mida isik kogeb, kuid mida otseselt verbaalsesse keelde tõlkida ei saa?

Loomulikult on tarbijad üle maailma võimelised n-õ lugema pilte, kuid kas nad loevad samas keeles? Naiivoletuse kohaselt on piltide tähendused universaalsed. Pildid on – vastavalt sellele tõekspidamisele – nagu visuaalne esperanto, keel, mis ületab kõik keele- ja kultuuribarjäärid.

Säärase tõekspidamise pooldajad kasutavad sama sõnumit igas kultuuris. Nad esitavad väljakutse väitele, et eri kultuuride turg on erinev. Teised osutavad rahvuste

³ 2006. a käivitunud projektis, mida juhivad dr Dessislava Boshnakova New Bulgarian ülikoolist ja dr Jaakko Lehtonen Jyväskylä ülikoolist, kogutakse reaktsioone reklaamikollektsioonile, intervjuuerides eri kultuuride tarbijaid. Esimesed andmed koguti 2006. a Bulgaariast, Soomest, Ukrainast ja Šveitsist.

erinevustele elustiilis, kollektiivsetes väärtustes, majandusliku arengu tasemes ning kultuuris tervikuna, nad tahavad kohandada oma sõnumit iga kultuuri jaoks ning luua erinevaid sõnumeid, et rahuldada iga kultuuri ootusi.

Reklaamija, kes esindab n-ö standardiseerimise koolkonda, valib reklaami pildikujundi lähtuvalt sellest, kuidas sihtauditoorium tema enda kultuuris tõenäoliselt visuaalset sõnumit dekodeerib, oletades, et sihtkultuuri tarbija teeb seda samamoodi. Võib-olla talitas sellest seisukohast lähtudes ka reklaamija, keda kirjeldab Øyvind Dahl (2001): too avaldas araabia maades peavalutablette reklaamiva koomiksi, ent reklaamitavatel tablettidel oli kehv menu. Vasakpoolsel koomikspildil oli kujutatud kannatava ilmega inimest, keskmisel pildil võttis ta tableti sisse ja paremal pildil naeratas, ilmselgelt ilma mingi peavaluta. Reklaamija ei pööranud aga tähelepanu sellele, et araabia tekste loetakse paremalt vasakule. Mitte keegi ei tahtnud osta tablette, mis põhjustavad peavalu!

Visuaalsete kujutiste uurimise teoreetilist raamistikku on nimetatud *visuaalseks retoorikaks*, et eristada seda verbaalse keele tähenduste teooriast. Visuaalne retoorika erineb visuaalse kunsti uurimisest – või reklaami visuaalse graafilise disaini uurimisest, milles keskendutakse esteetilistele kaalutlustele. Visuaalne retoorika uurib kultuurilisi tähendusi, mis väljenduvad visuaalsetes artefaktides, nagu maalid, fotod, reklaamid, võrgulehed või ajalehed. David Blakesley (2006) järgi on näiteks visuaalne retoorika kokku kleebitud standardse arusaamaga retoorikast kui keelelisest veenmiskunstist. Lisaks sellele õpetavad retoorika käsiraamatud, kuidas kasutada visuaalset materjali, et täiendada ja avardada verbaalse keele argumente, või kuidas kasutada visuaalset materjali illustratsioonina, või lihtsalt kuidas kasutada ära publikatsiooni küljendust.

Visuaale ja kultuuri uurib ka visuaalne antropoloogia. See on sotsiokultuurilise antropoloogia allvaldkond, mis hakkas arenema pärast kino ja fotograafia teket. Kuigi öeldakse, et visuaalne antropoloogia hõlmab massimeedia tootmist ja vastuvõtmist, on keskmis massimeedia toote antropoloogiline uurimine, mitte aga näiteks visuaalsete märgitähenduste kultuuridevaheline võrdlemine (vt Lester 1999).

Visuaali kasutuskontekst

Veenmine ei peitu mitte ainult pildis, vaid sõltub mentaalsetest protsessidest, mida pilt kontekstuaalsete, situatsioonist lähtuvate ja individuaalsete muutujate mõjul inimese peas esile kutsub. Tänapäevast, sotsiaalses keskkonnas toimuvat veenmist võib teisalt nimetada võistlevaks veenmiseks: meile pakutakse rohkem kui ühe elustiili võimalust, rohkem kui üht poliitilist kandidaati või brändi, ja igaüks neist pommitab meid pidevalt ja püüab veenda enda eelistes. Uurimust, mis keskendub unikaalsetele visuaalsetele aspektidele reklaaminduses, nimetatakse *visuaalseks veenmiseks*. Reklaaminduses, nagu ka propagandas võib pilte kasutada muuhulgas selliste väidete esitamiseks, mis muidu oleksid vastuvõetamatud või mida verbaalselt esitatuna tõrjutaks tagasi.

Vastupidiselt kultuurispetsiifikat ignoreeriva standardiseerimise koolkonna veendumusele, millest eelpool juttu oli, võib esitada hüpoteesi, et eksisteerib visuaalne keel, mis varieerub kahe erineva kultuuri piires samamoodi nagu verbaalne keel.

Uurimusi kultuurierinevustest visuaalide n-õ lugemises on üsna vähe. Verbaalset veenmist on uuritud sajandeid, mille tulemusel on tekkinud kõrgetasemelised mudelid ja teooriad, kuid visuaalse veenmise uurimine on alles lapsekingades. Visuaalse kommunikatsiooni teooria on lähedalt seotud semiootikaga. Pilte võib vaadata ikoonilisuse ja indeksilisuse vaatenurgast. Esimene viitab pildi võimele imiteerida reaalsust: inimesed võivad näha pilti sümbolina, mis justkui (taas)loob reaalsust. Indeksilisus viitab pildile, mida nähakse kui füüsilise reaalsuse füüsilist jälgendit: inimesed võivad näha pilte nii, nagu need oleks reaalsuse vahendamata jäädvustused, selle vahetu osa. Need semiootika mõisted on hiljuti toodud reklaamiuurimisse vägagi huvitaval moel: brändi ikoonistaatus on alati olnud eduka brändiloompeas, mida on nimetatud kultuuriliseks brändiloomeks. Niisuguse brändiloompe näitena on ikooniks muutunud Coca-Cola. Douglas B. Holti (2004) järgi viivad brändi ikooni staatusesse lood tema kasutuselevõtt ja identiteedimüüdid. Nii aitas Coca-Colat kultuurilise ikooni staatusesse tõsta teadmine, et seda jõid Ameerika sõdurid Teise maailmasõja ajal.

Enne piltide veenmisjõu käsitlemist peatun korraks tähenduse ja kultuuri küsimusel. Tegelikult on küsimus pildi tähenduse kultuurilisest seotusest puhtakadeemiline. Kõik teavad igapäevakogemuse põhjal, et igas kultuuris on sõnumeid, ja mitte ainult pilte, mida teisest kultuurist pärit inimene ei mõista või millest saadakse aru täiesti erinevalt.

Ron Allison (1995) toob näiteid eksimustest mõnede reklaamlausete (ingl *slogan*) tajumisel kultuuride kommunikatsioonis:

Lähtesõnum

“Come alive with Pepsi!”

“Schweppes tonic water”

“Cleans the really dirty parts of you wash”

Vastuvõetud sõnum

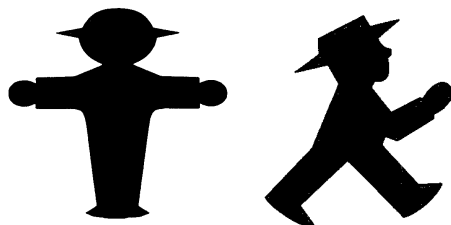
Sakslaste tõlgendus: “Come out of your grave with Pepsi”
‘Tule välja oma hauast koos Pepsiga’

Itaallaste tõlgendus: “Schweppes bathroom water”
‘Schweppes vannivesi’

Kanada prantslaste tõlgendus: “Cleans your private parts!”
‘Puhastab su privaatset kehaosad’

Sarnast väärnimõistmist võib kohata ka visuaalses kommunikatsioonis, kuigi seal on raskem avastada momenti, millal pildi tähendust väärsti mõistetakse. Võib-olla on mõned meist seisnud võõral maal WC-ukse taga ja üritanud nuputada, kumb sümbol tähistab meest ja kumb naist. Vahel aga tehakse väär otsus mitte pildi, vaid väärtuste tasandil. Mõni aeg tagasi üritas anonüümne Soome mobiiltootja müüa telefone Hiinasse reklaamiga, kus telefoni käes hoidev mees seisab üksinda suurel inimtühjal mäel. Kavatsatud sõnum oli “Mobiiltelefoniga pole sa kunagi üksi”, kuid hiinlased ei tahtnud seda telefoni, sest kartsid sattuda niisugusele tühermaale, kus pole ühtegi teist inimolevust (vt Li 2001).

Kultuuri ja visuaalide seotust demonstreerivad järgmised joonised. Igaüks teab väikesi mehikesi valgusfooris (vt joonis 1). Mõned võivad ära tunda, et need mehikesed on võetud Ida-Berliini valgusfooridest, Lääne pool on kujud kõhnemad. Kuid kui me ei teaks konteksti, milles neid kujusid nähakse, või millest nende punane ja roheline värv meile märku annab, siis oleksid need mehikesed lihtsalt seisva ja kõndiva mehe pildid.



Joonis 1. Jalakäijate valgusfoori tuled, vasakpoolne mehike on punane, parempoolne roheline

Poliitilisel karikatuuril on väike merineitsi riietatud islami burkasse (vt joonis 2).



Joonis 2. Merineitsi (Suomen Kuvalehti 10.2.2006)

Sellest pildist arusaamine nõuab nii kultuurilist kui sotsiaalset teadmist ning metafoori mõistmist. Tuleb teada, et Merineitsi on Kopenhaageni sümbol ja et Taani ajaleht avaldas koomikspildid Muhamedist, mis tekitasid ulatusliku meeleavalduste laine islamimaades. Visuaalset kirjaoskust vajatakse ka pildi metafoorse tähenduse mõistmiseks: Lääne muret selle üle, et islam levib Euroopasse.

Joonis 3 sisaldab palju sümboolset ja metafoorilist.



Keep Europe Tidy

Joonis 3. Hoia Euroopa puhas

Selleks et sõnumist aru saada, tuleb teada *prügikasti*-mõistet ning millegi prügikasti viskamise metafoori. Et anda metafoorile tähendust, tuleb teada islami sümboleid, aga ka ajaloolist konteksti, milles joonistus avaldati.

Kaks nendest lihtsatest piltidest on visuaalse propaganda näited: nii merineitsi kui ka prügikast kannavad poliitilist sõnumit, islamivastasust. Enamikule eurooplastele on sõnum ilmselge, ja seda ilma ühegi verbaalse sümbolita mis tahes keeles.

Kuid kas pildid saavad käituda visuaalse veenmise, propaganda või isegi ajupe-suvahendina? Uus tekkiv teadusharu – neuroökonomika – üritab avada Pandora laegast ning paljastada, mis toimub inimajus, kui tarbija näeb pilte ning teeb ostu-valikuid, ehk teisiti öeldes, milline on inimese mentaalne seisund ning reaktsioon turundussõnumitele. (Vt ka Braeutigam 2005.)

Aju-uuringud

Neuroökonomika kasutab ajutegevuse uurimiseks järjest enam magnet-resonantsuuringuid, millega on võimalik ohutult saada pilte aju aktiveerunud piirkondadest (vt Economist 12.6.2003).

Esimesed MRI-meetodil tehtud avastused visuaalide uurimises saavutasid laia kõlapinna. Autohuviliste meestega tehtud katsetest selgus, et nende aju mõnu-piirkond aktiveerub sportautode pilte nähes märgatavalt enam kui limusiini või väikeautode piltide puhul. Sellest võib järeldada, et aju mõnupiirkondi on võimalik aktiveerida toodete abil, mis assotsieeruvad rikkuse ja sotsiaalse positsiooniga (Erk jt 2002).

Veelgi enam erutas aga turundusnimesi Baylori Meditsiinikolledžis läbiviidud MRI-katse kahe hästituntud suhkrujoogi Coca-Cola ja Pepsiga (vt McClure jt 2004). Mõlemad joogid on väga sarnase keemilise koostisega, ent tarbijad jagunevad selgelt kaheks, ühed on veendunud Coca-Cola pooldajad ja teised eelistavad Pepsit. Korraldati kaks katset, esimeses tuli katsealustel juua kahte jooki, ilma et nad oleksid teadnud, kumb on kummas klaasis, ning neil tuli öelda, kumba nad peavad maitse põhjal paremaks. Teises katses uuriti brändi mõju eelistuse kujunemisele: kuigi näiteks mõlemas klaasis oli ainult Coca-Cola või ainult Pepsi, kandis üks neist Coca-Cola tunnust ja teine Pepsi oma.

Tulemustest selgus, et pimekatse puhul ei lähe aju reageering kokku selle joo-giga, mida katsealused olid öelnud end eelistavat. Katses maitset paremaks peetud jook aktiveeris ajus ventromediaalse prefrontaalse korteksi (VMPFC).

Teises katses, kus katsealused arvasid teadvat, mida nad joovad, aktiveerisid Coca-Cola fännidel selle sildiga jooki juues mitmed ajupiirkonnad: hipokampus, dorsolateraalne prefrontaalne korteks (DLPFC), keskaju, talamus ja ajukoore vasak nägemispiirkond. Pepsi silt sellist aktiveerumist esile ei kutsunud. Katsetest järeldus, et sensoorne info määrab inimese käitumist vaid osaliselt ja eelistusotsuste tegemisel on suur roll bränditeadlikkusel.

Katse on heaks illustratsiooniks sellele, kuidas lähetatud reklaamsõnum võib tõepoolest töötada alateadvuse tasandil, ning kuidas visuaalsete sümbolitega saab valla päästa reaktsioone. Bränd on võimeline petma aju nii, et see modifitseerib maitseneuronitega vastuvõetud informatsiooni!

Kuigi neuroturundus (vt Lee jt 2007) teeb alles esimesi samme, on selle hirmuäratavad võimalused poliitilises propagandas ja kahjulike toodete ning harjumuste reklaamimises (nt tubakas, alkohol, isegi narkootikumid) esile kutsunud kriitikat. Commercial Alert, Ameerika tarbijate assotsiatsioon saatis 12. juulil 2004 Ameerika Ühendriikide Kaubanduskomiteele kirja, nõudes neuroturunduse asjas juurdlust. Assotsiatsioon küsib: “Mis juhtub sellel maal, kui korporatiivsed turustajad ja poliitilised konsultandid võiksid sõna otseses mõttes piiluda meie ajusse ning kaardistada närvitegevuse, mis suunab meid tegema valikuid supermarketites ja valimiskabiinides? Mis juhtub, kui nad saaksid närvikava aktiveerida, et muuta meie käitumist vastavalt nende vajadustele?”

Neuroökonomika uurimistulemused näivad toetavat ka nüüdisaegse reklaamimise väidet, et reklaam mõjutab tarbijat enim siis, kui ta sellele ei mõtle. Kuigi inimesed võib-olla ei pööra reklaamidele tähelepanu või ei mõtle neile teadlikult, imab aju keskkonnast pidevalt informatsiooni. Kuid me ei vaja neuroteaduse tõestusi märkamaks, et reklaamimise on propaganda haru, mis keskendub kommertstoodete ja teenuste üleskiitmisele – vastupidi poliitilisele propagandale, mis teostab valituste, poliitiliste parteide ja kolmanda sektori eri laadi organisatsioonide avalikke informatsioonikampaaniaid. Meetodid on samad: *slogan*’id, stereotüpiseerimine, veenvad tõendid, autoriteetidele viitamine, kuulsustest toetajate kasutamine, eufooria, lärmakad orkestrivankrid, must-valged valikud jne. Neid vahendeid võib kohata nii poliitilises reklaamis kui ka kommertspropagandas, nii verbaalses kui ka visuaalses keeles. Visuaalse turunduskommunikatsiooni rahvuslike vm kultuuriliste karakteristikute ja kultuurivastuolude uurimine on väljakutse kultuuride kommunikatsiooni rakendusuringutele. Kuigi me võib-olla ei leiagi *osta*-nuppu, millest unistab neuroturundus, võime siiski õppida nii mõndagi visuaalsest veenmisest. Vähemalt õpime paremini aru saama sellest, kuidas inimene töötleb visuaalset informatsiooni – ja pürgimus aru saada inimese käitumisest on alati olnud ja jääb kõikide inimuuringute põhieesmärgiks.

Kokkuvõte

Kommunikatsioon tähendab palju enam kui verbaalse keele kasutamist. Visuaalsed kujutised esindavad lühimat teed inimajusse. Nii turunduses kui ka tarbija-psühholoogias tehtud uurimused on kinnitanud visuaalide jõudu kommertslikus ja poliitilises veenmisel. Kui tahame mõista oma igapäevaelu sotsiaalse keskkonna sõnumeid tervikuna ja nende sõnumite mõju meie väärtushinnangutele, hoiakutele ja käitumisele, siis verbaalsetest sõnumitest ei piisa. Peame uurima visuaalide keelt, ja mitte ainult illustratsiooni või verbaalse teksti tasandil, vaid kui kommunikatsioonikanalit, mis võib veenmise mõttes olla palju võimsam kui verbaalne keel.

Kirjandus

- Allison, Ron 1995. Cross-cultural factors in global advertising. – Jürgen Bolten (Ed.). *Cross-Culture – Interkulturelles Handeln in der Wirtschaft*. Sternenfels: Verlag Wissenschaft & Praxis, 92–101.
- Blakesley, David 2006. Visual Rhetoric. [http://www.cla.purdue.edu/dblakesley/visual/\(01.04.2006\)](http://www.cla.purdue.edu/dblakesley/visual/(01.04.2006)).
- Braeutigam, Sven 2005. Neuroeconomics – from neural systems to economic behaviour. – *Brain Research Bulletin* 67, 355–360.
- Dahl, Øyvind 2001. *Møter mellom mennesker*. Interkulturell kommunikasjon. Oslo: Gyldendal.
- Economist 2003. MRI's inside story. 369 (8353), 26–28.
- Erk, Susanne; Spitzer, Manfred; Wunderlich, Arthtur P.; Galley, Lars; Walter, Henrik 2002. Cultural objects modulate reward circuitry. – *Neuroreport* 13 (18), 2499–2503.
- Holt, Douglas, B. 2004. *How Brands Become Icons: The Principles of Cultural Branding*. Harvard Business School Press.
- Larsen, Val; Luna, David; Peracchio Laura A. 2004. Points of view and pieces of time: A taxonomy of image attributes. – *Journal of Consumer Research* 31, 102–111.
- Lee, Nick; Broderick, Amanda J.; Chamberlain, Laura (2007). What is 'neuromarketing'? A discussion and agenda for future research. – *International Journal of Psychophysiology*, 63 (2), 199–204.
- Lester, Paul 1999. *Visual Communication. Images with Messages*. Belmont CA: Wadsworth.
- Li, Zhenyi 2001. *Cultural impact on international branding. A case of marketing Finnish mobile phones in China*. Doctoral dissertation. University of Jyväskylä.
- McClure, Samuel M.; Li, Jian; Tomlin, D.; Cypert, Kim S.; Montague Latané M.; Montague, Read P. 2004. Neural correlates of behavioral preference for culturally familiar drinks. – *Neuron* 44, 379–387.

Jaakko Lehtose (Jyväskylä Ülikool) uurimisvaldkondadeks on organisatsiooni suhtekorraldus ja ühiskondlik vastutus, imago ja kollektiivi-identiteet, Soome kuvand, kommunikatsioon ja rahvusvahelistumine, kommunikatsiooni eetika.
jaakko.lehtonen@jyu.fi

MORE THAN WORDS

Jaakko Lehtonen

University of Jyväskylä

This article wants to encourage applied linguists to focus on research in visual communication, not as illustration of verbal texts but as 'texts' in the language of visuals. The significance of visuals in communication is grounded with recent findings in studies using the dynamic MRI methodology.

Keywords: literacy, advertisement, persuasion, visual language, visual, MRI

KUIDAS HINNATA SUURE PANUSEGA TESTIDE HINDAJAID

Hille Pajupuu

Ülevaade. Suure panusega testide (ingl *high-stakes test*) osatähtsuse kasv toob kaasa vajaduse pöörata rohkem tähelepanu testi subjektiivhinnatavate osade hindamise kvaliteedile. Kasutatavate lihtsate statistiliste meetodite puhul võib jääda märkamata palju valestihindamisi juhul, kui hindajate arv on suur ning nende koostis nõrk. Artikkel tutvustab meetodit, mille abil saab kindlaks teha valesti hinnatud hindajad ning nende hinnatud tööd õigeaegselt ümber hinnata. Meetod on mõeldud kasutamiseks kahekordse hindamise puhul, seda demonstreeritakse eesti keele algtaseme testi rääkimisosa hindamise näitel.*

Võtmesõnad: hindajate hindamine, hindajate järjekindlus, subjektiivhindamine, kahekordne hindamine, eesti keel

Sissejuhatus

Suure panusega testide all mõeldakse niisugust laadi teste, kontrolltöid, arvestusi, eksameid jms, millel on paratamatult suur mõju inimese saatusele. Seesuguste testide tähtsus on pidevalt kasvanud, nende põhjal tehakse järjest enam otsustusi testi sooritaja, aga ka õpetaja ja õpetamise kohta. Paljud neist testidest on muutunud aja jooksul kohustuslikuks (Altshuler jt 2006, Fine 2005, Helfenbein 2004). Sama tendentsi on näha ka Eestis, mõeldagu siis arvestuslikele töödele, põhikooli lõpueksamitele või gümnaasiumi riieksamitele. Sama puudutab eesti keele kui teise keele oskuse mõõtmist. Kui 1995. aastal oli Eestis vaid üks riiklikult korraldatav standardiseeritud eesti keele kui riigikeele test – Eesti kodakondsuse taotleja eesti keele eksam –, siis 1999. aastal lisandusid sellele eesti keele tasemeeksamid (alg-, kesk- ja kõrgtase), eesti keele kui teise keele riieksam ning võõrkeelse põhikooli eesti keele kui teise keele lõpueksam. 2000. aastal ühitati kodakondsuse taotleja eesti keele eksam tasemeeksamitega. (Vt tabel 1.)

* Artikkel on valminud Eesti Keele Instituudi baasfinantseerimise ja Alfred Kordelini Sihtasutuse Eesti Fondi toel. Täna Riiklikku Eksami- ja Kvalifikatsioonikeskust võimaluse eest kasutada eksamitulemusi uurimistöökäes.

Tasemeeksami tulemustest sõltub inimese võimalus töötada nii või teistsugust keeleoskustaset eeldavatel ametikohtadel, põhikooli lõpueksamitulemustest sõltub võimalus jätkata õpinguid gümnaasiumis, riigieksamitulemustest pääs ülikooli.

Kõiki neid eksameid korraldab Riiklik Eksami- ja Kvalifikatsioonikeskus (REKK). Tasemeeksamid toimuvad üle riigi seitsme linna eksamipunktides 10 korda aastas, riigieksam ja põhikooli lõpueksam viiakse läbi koolides üks kord aastas. Eksaminandide arv on aasta-aastalt kasvanud. 2005. aastal oli eksaminande 16 124 (algtaase 5075, kesktaase 2357, kõrgetaase 1177, riigieksam 5344, põhikooli lõpueksam 2172).

Tabel 1. Suure panusega eesti keele kui riigikeele eksamid ja nende ligikaudne vastavus Euroopa Nõukogu keeleoskustasemetele

Euroopa Nõukogu keeleoskustasemed		Eesti keele kui riigikeele eksamid
Vilunud keelekasutaja (Proficient User)	C2	(Eestis riiklikult ei testita ega nõuta)
	C1	Eesti keele kõrgetaseme eksam
Iseseisev keelekasutaja (Independent User)	B2	Eesti keele kesktaseme eksam Muukeelse gümnaasiumi riigikeele riigieksam
	B1	Eesti keele algtaseme eksam ühitatud Eesti kodakondsuse taotleja eesti keeleksamiga Eesti keele kui teise keele riigieksam
Algtasemel keelekasutaja (Basic User)	A2	(Eestis riiklikult ei testita ega nõuta)
	A1	

Kuna testi võim on suur – mõne tunni jooksul pannakse paika inimese saatus –, siis eeldab see kõigi testiarenduse ja testide kasutamise seotud inimeste vastustunnet (vt ka Shohamy 2001).

Suure panusega eesti keele testides annavad testi maksimaalsest punktisummast poole objektiivhinnatavate osade (kuulamine ja lugemine) eest saadavad punktid ja poole subjektiivhinnatavate osade eest (kirjutamine ja rääkimine) saadavad punktid. Seega on hindaja roll lõpliku punktisumma kujunemisel väga suur. Eksameid korraldab institutsioon peab tagama, et hindajad hindaksid võimalikult õigesti. Selleks tuleb hindajaid koolitada ning nende tööd analüüsida. Hindamisvead tuleb leida enne, kui hinded jõuavad eksaminandide tunnistustele.

Eesti keele testimisulokord on suhteliselt keeruline: et eksamid toimuvad üle riigi üheaegselt ja tööde hindamisperiood on väga lühike (30 päeva), tuleb hindamisse kaasata suur hulk hindajaid. Tasemeeksamihindajaid on 65, koolieksamihindajaid üle 400. Kui tasemeeksamihindajad on läbi aastate enam-vähem ühed ja samad, avaliku konkursiga valitud eesti filoloogid, siis koolieksamitel kasutatakse hindajatena kooliõpetajaid, kes võivad aastati vahetuda.

Tasemeeksami hindajaid koolitatakse 1–2 korda aastas, enamik neist hindab kindla taseme eksameid 7–8 korda aastas. Koolitustel osalemine on vabatahtlik, hindajaeksameid pole. Koolieksamite hindajaid koolitatakse üks kord aastas – vahetult enne eksamit – ning enamikul puudub suure panusega eksamite hindamiskogemus.

Suurema objektiivsuse saavutamiseks toimub riigi- ja tasemeeksamitel kahekordne hindamine: sooritust hindab sõltumatult kaks hindajat, hinnete 3-punktise vahe korral (10-punktilisel skaalal) läheb töö kolmandale hindamisele.

Probleemid, mis Eestis töös hindajatega välja paistavad on hindajakoolituse vähesus, analüüsi puudumine koolituse mõju hindamiseks, puudulik tagasiside hindajatele eksamihindamise tulemustest, teadmatus hindaja järjekindlusest (ingl *rater consistency*)¹. Põhjus, miks hindamisele ja hindajatele on seni suhteliselt vähe tähelepanu pööratud, on ühelt poolt eksami sisu ja hindamise eest vastutavate inimeste väike arv (tasemeeksamite arendamise, hindajakoolituse ja analüüsiga tegeleb REKK-is 3 inimest, eesti keele kui teise keele koolieksamitega 1 inimene), kuid teisalt pole selline olukord ainuomane ainult Eesti eksamikeskusele. Vähest tähelepanu hindamisele on Eesti kõrval täheldatud ka teistes Balti riikides, kusjuures kõigis neis kasvab aastast aastasse suure panusega testide arv; just hindamise analüüsi puudumisest võib saada eksamikvaliteedile peamine oht (vt Eckes jt 2005). Ent hindajakoolituse ja hindajaanalüüsi kohta eksamikeskustes pole ka laiemalt üksikasjalikku teavet. Eksamikeskused ei publitseeri hindajaanalüüsiste raporteid kuigi sageli (Brooks 2004). Küsitluste põhjal võib järeldada, et igakülgsest hindajakoolitusest ja analüüsist jääb puudu paljudes eksamikeskustes (Alderson jt 1995).

Et koolituse ja tagasiside mõju hindamisele on alles viimasel ajal uurima hakatud, siis on võimalik, et teadmisi pole jõutud veel praktikasse rakendada, eriti juhul, kui vastavaid uurimusi tehakse väljaspool eksamikeskusi. Eestis näiteks puudub teadusasutus, kelle uurimisvaldkondade hulka kuuluks keeletestimine. Keeletestimist uuritakse lühiajaliste projektide raames, ning kuna ka REKK-i enda põhiülesannete hulka teadustöö ei kuulu, siis uurimistulemuste praktikasse viimine on küllaltki pikaajaline protsess, mis takerdub sageli ka rahastamise taha.

Hindajakoolituse ja -analüüsi tähtsust ei tohi aga alahinnata: need peaksid käima käsikäes. Eksaminandi hinne ei tohi sõltuda sellest, kas teda satub hindama koolitatud hindaja või mitte. Valesti hinnatud sooritus vajab ümberhindamist enne, kui selle punktisumma jõuab tunnistusele.

Hindajakoolituse mõju

Hindaja usaldatavust tõstab järjekindel koolitamine.

Tom Lumley ja Tim F. McNamara (1995) uurimus näitab, et koolituse mõju ei kesta tingimata väga kaua – aja möödudes hakkavad mõned hindajad hindama eba-järjekindlamalt. Analüüsides hindaja tööd, võib tema koolitusvajadust õigeaegselt märgata. Sara Cushing Weigle (1998), uurides hindajate rangust (ingl *rater severity*) ja järjekindlust enne ning pärast koolitust, täheldas koolitusjärgses hindamises mitmeid muutusi, kuid peaaesjalikult uute, vähese kogemusega hindajate puhul. Hindajad jäid küll ka pärast koolitust hindama erineva rangusega, ent äärmuslik

¹ Hindaja järjekindluse all mõeldakse seda, et ta hindab töid alati ühtedelt alustelt lähtudes.

rangus ja leebus taandusid. Koolitus mõjus tõhusamalt hindaja järjekindlusele: osa hindajad, kes enne koolitust klassifitseerusid kui ebajärjekindlad, muutusid küllaltki järjekindlaks pärast koolitust. Kuid koolitus ei muutnud järjekindlaks sugugi mitte kõiki hindajaid: võib-olla pole kõiki hindajaid võimalik üldse järjekindlaks koolitada ja nende kasutamisest hindajana tuleb loobuda. S. Cushing Weigle (1998) uurimus kinnitas varasemaid Mary D. Lunzi, Benjamin D. Wrighti ja John M. Linacrei (1990) tulemusi, et hindajakoolitus ei muuda hindajaid üksteise duplikaatideks, kuid koolitus võib muuta hindajad järjekindlamaks. Järjekindlate hindajate pandud hinded on eeldatavalt õigemad. Seega on koolitus tingimata vajalik uutele hindajatele nii järjekindluse saavutamiseks kui ka äärmuslikult range või leebe hindamise vältimiseks; ühtlasi aitab koolitusega kaasnev hindajaanalüüs õigeaegselt välja selgitada need hindajad, kes vaatamata koolitusele jäävad ebajärjekindlaks. Ka William J. Bonk ja Gary J. Ockey (2003) rõhutavad koolituse tähtsust, aga ka kogemuse ja tagasiside vajadust hindaja järjekindlamaks muutmisel, olgugi et isegi väga intensiivne koolitus ei kaota hindajate erinevusi täiel määral.

Hindajate püsivat ranguserinevust koolitusele vaatamata on märkinud ka Kimi Kondo-Brown (2002): sarnase hariduse ja hindajakogemusega hästi koolitatud hindajad hindavad ühe ja sama soorituse eri aspekte erineva rangusega, ja kuigi see hindajatevaheline erinevus on väike, on ta piisavalt oluline, et mõjutada lõpphinde kujunemist.

Koolitust vajavad ka kogenud hindajad, et tõlgendada hindamisskaalat sarnasemal viisil (Lumley 2002).

Seega on koolitus oluline õiglasema hindamise saavutamiseks, kuid kuna koolitus ei pane hindajaid hindama ühesuguselt, on pidev hindajaanalüüs ja asjakohane tagasiside hindajale vajalik.

Ülesanne

Eesmärk on leida meetod, mis lubab tuvastada valesi hinnanud hindajad võimalikult lihtsal ja kergesti rakendataval viisil ja mida sobiks kasutada eksamikeskustes, kus eksameid ja hindajaid on palju, kasutatakse kahekordset hindamist ning kõikide hindajate järjekindel koolitamine on mingitel põhjustel raskendatud (hindajad paiknevad üle riigi või isegi üle maailma laiali, hindajaid vajatakse ajutiselt ja nende hulgas on palju uusi hindajaid, eksamid toimuvad tihedalt ja nende vahel ei jää koolituseks aega, hinnata tuleb palju töid lühikese ajavahemiku jooksul, hindajakoolitajatest on puudus, hindajakoolituseks pole raha vms).

Hindajate hindamine ja probleemid

Subjektiivhindamisel on oluliseks hindaja järjekindlus, mille tähtsam näitaja on reliaablus. Kui hindajaid on üks, siis peetakse teda reliaabseks juhul, kui ta annab ühe ja sama soorituse eest eri olukordades samad punktid (ingl *intra-rater reliability*). Mitme hindaja puhul peetakse hindajaid reliaabseteks, kui eri hindajate ühele ja samale sooritusele antud hindepunktid langevad kokku (ingl *intera-rater reliability*) (vt ka Bachman 1990: 178–181, Luoma 2004: 179–184).

Hindajakoolitusega püütakse saavutada olukord, kus hindaja pandud hinded langeksid kokku teda koolitava eksperthindaja omadega (Alderson jt 1995: 105–112). Enamasti alustatakse koolitust hindaja reliaabluse kindlakstegemisest: arvutatakse hindaja korrelatsioon iseendaga (ingl *intra-rater correlation*). Selleks lastakse hindajal hinnata pikemate vaheaegade järel samu töid uuesti ning korreleeritakse pandud hinded. Tugeva positiivse korrelatsiooni puhul (korrelatsioonikoeffitsient üle 0,8) usutakse, et hindajal on kujunenud välja arusaam, kes on tugev ja kes nõrk keeleoskaja, ning ta hindab sama tööd alati ühesugustelt alustelt lähtudes. Tugev positiivne korrelatsioon ei välista siiski seda, et hindaja hindab töid erinevatel kordadel erineva rangusastmega. Et rangusastet välja selgitada, võidakse arvutada iga korreleeritava hindamiskorra hinnete keskmine ja vaadata, kus see hindamisskaalal paikneb.

Lisaks hindaja korrelatsioonile iseendaga, korreleeritakse koolitustel hindaja hindeid eksperthindaja omadega (ingl *inter-rater correlation*) ning arvutatakse mõlema hinnete keskmine ning standardhälve. Eksperthindajaga ühesugust hindamist näitab 0,8-st suurem korrelatsioonikoeffitsient. Keskmise järgi näeb, kas hindaja on eksperthindajast rangem või leebem. Kui keskmine on märgatavalt madalam, siis on tegu range hindajaga, kui kõrgem, siis leebemaga (Alderson jt 1995: 132).

Standardhälbe järgi näeb, kui suures ulatuses hindaja hindamisskaalat kasutab, kas ta kasutab seda sarnaselt eksperthindajaga, kas ta julgeb vajadusel panna kõrgeid ja madalaid hindeid. Standardhälbe põhjal võib eksamisituatsioonis hindaja kohta teha järeldusi siiski vaid juhul, kui hinnatavate tööde hulk on suur. Kui hinnatavaid töid on vähe, võib väikse standardhälbe põhjus olla hinnatavate juhuslikult ühesuguses tasemes, mitte hindaja oskamatuses või kartuses hindamisskaalat kasutada.

Kas eksamil kasutada ühe- või kahekordset hindamist, on olnud mõneti problemaatiline. Kahekordne hindamine on kallis, sest selleks vajatakse rohkem hindajaid, ning see on ka aeganõudvam (Brooks 2004). Kahekordne hindamine on õigustatud subjektiivhinnatavate osade puhul, kus sedasi pandud hinne on eeldatavasti usaldusväärsem. Eksameid korraldavale institutsioonile annab kahekordne hindamine võimaluse kontrollida hindamisprotsessi (hindamise kokkulangevuse ulatuse analüüs) ja aidata kaasa hindamise standardiseerimisele. Kahekordne hindamine annab tagasisidet ka hindajatele: hindajate arutlusi pandud hinnete üle võib käsitada ka kui hindajakoolitust (Cannings jt 2005).

Eksamikeskustes, kus subjektiivhinnatavaid testiosi hindab sõltumatult kaks hindajat, kontrollitakse hindamise reliaablust valdavalt hindajatevahelise korrelatsiooniga või arvutatakse kahe hindaja pandud hinnete vahed. Nõrga korrelatsiooni või hinnete suure vahe puhul hinnatakse tööd ümber kas kolmanda hindaja poolt või siis püüavad kaks paarishindajat jõuda kompromissini. Mida koolitatumad ja järjekindlamad on hindajad, seda usaldatavamad on taolised lihtsad statistilised näitajad.

Kui aga ollakse olukorras, kus hindajakoolitus pole väga heal järjel ja kus hindaja järjekindluse kohta pole andmeid, ei saa välistada olukorda, et paaris hindama satuvad kaks kvalifitseerimata hindajat. Sellisel juhul pole eelpoolnimetatud lihtsatest statistilistest näitajatest kuigi palju abi: korrelatsioon võib olla tugev ka siis, kui hindajad hindavad ühtemoodi valesti, nii nagu ei tule märkimisväärset

hinnete vahet siis, kui paaris on hinnanud kaks leebet või kaks ranget hindajat. Seega, kuigi statistilised näitajad võivad olla head, võib eksaminandi siiski olla hinnatud ebaõigesti.

Olukorras, kus hindaja pädevusest pole selget ülevaadet, ent on vaja kindlaks teha ebakvaliteetselt hinnanud hindajad, oleks otstarbekas leida võimalus hindajate võrdlemiseks eksperthindaja(te)ga.

Hindaja sarnasus eksperthindajaga – hindaja kvaliteediindeks

Hindaja (H) töö kvaliteedi hindamise meetodi väljatöötamisel lähtuti eeldusest, et eksperthindajad (E) on hindajad, kes on korralikult koolitatud, kes on järjekindlad (korrelatsioon iseendaga üle 0,8), kes kasutavad hindamiskaalat õigesti (piisava arvu hinnatud tööde korral hinnete keskmine u 6,0 10-punktilisel skaalal, standardhälve võimalikult suur), kes saavad ühtemoodi aru, kes on tugev, kes nõrk keeleoskaja (hindajatevaheline korrelatsioon kahekordsel hindamisel üle 0,8, kel omavahel paaris hinnates keskmised hinded sarnased ning hinnete vahe väiksem kui 3 p 10-punktilisel skaalal).

Analoogiliselt hindajakoolituse eesmärgiga – õpetada hindaja hindama eksperthindajaga sarnaselt, püstitati ka hindaja kvaliteedi määramisel eesmärk leida hindaja sarnasus eksperthindajaga, ja seda ka juhtudel kui hindaja eksperthindajaga ise paaris ei hinda. Hindajakoolituste põhjal võib eeldada, et mida sarnasemalt on hindaja hinnanud eksperthindajaga, seda parem on tema töö kvaliteet ja mida erinevamalt on ta hinnanud eksperthindajast, seda enam on põhjust kahelda tema töö kvaliteedis. Viimasel juhul on hindaja eksperthindajast kas leebem, rangem või ebajärjekindlam.

Meetodi kirjeldus

Hindaja sarnasuse leidmiseks eksperthindajaga arvutatakse kaks indeksit:

- paari sarnasusindeks S (ingl *similarity index*) – näitab kahe paarishindaja sarnasust;
- hindaja kvaliteediindeks K (ingl *quality index*) – näitab hindaja sarnasust eksperthindajaga.

Esimene samm hindaja kvaliteediindeksi leidmiseks on arvutada reaalselt paaris hinnanud hindajate omavahelised sarnasusindeksid. Selleks arvutatakse kõigepealt hinnete erinevuse ruutude keskmine (1) ning seejärel sarnasusindeks (2):

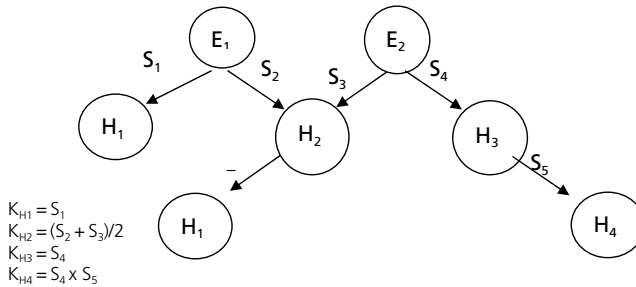
$$(1) \quad \overline{H_{\Delta}} = (H_1 - H_2)^2$$

$$(2) \quad S = 1 - \frac{\overline{H_{\Delta}}}{H_{\max}^2}$$

kus $\overline{H_{\Delta}}$ hinnete erinevuse ruutude keskmine, H_{\max} maksimaalne võimalik hinne, S on paari sarnasusindeks. Sarnasusindeks jääb vahemikku 0–1, kus 1 näitab paariliste täpselt sarnast hindamist ja 0 täiesti erinevat hindamist.

Eksperthindaja kvaliteediindeksiks võetakse 1. Ülejäänud hindajatele arvutatakse kvaliteediindeks (K).

Kui hindaja on hinnanud paaris eksperthindaja(te)ga, siis tema kvaliteediindeks on eksperthindaja(te)ga paaris hindamisel saadud sarnasusindeksite keskmine (vt joonis 1, hindajad H_1, H_2, H_3).



Joonis 1. Hindaja kvaliteediindeksi tuletuspuu

Hindajatele, kel pole olnud võimalust hinnata paaris eksperthindaja(te)ga, arvutatakse tulemuslik kvaliteediindeks pärast seda, kui eksperthindajaga paaris hinnatud hindajad on kvaliteediindeksid saanud. Tulemuslik kvaliteediindeks arvutatakse eksperthindajaga sarnaselt hinnatud hindajate kaudu sarnasusindeksite kaalutud keskmisena. Kaalukoefitsiendina kasutatakse eksperthindajaga sarnaselt hinnatud hindaja kvaliteediindeksit (vt joonis 1, hindaja H_4).

Kvaliteediindeksi arvutamisel rakendatakse kahte piirangut:

1. Sarnasusindeksit ei arvutata, kui paaris hinnatud töid on kaks või vähem.
2. Hindaja tulemusliku kvaliteediindeksi arvutamisel võetakse arvesse ainult tema need paarilised, kelle kvaliteediindeks on suurem etteantud koefitsiendist R. R-i väärtus valitakse selline, et see oleks maksimaalne kvaliteediindeks, mille korral saab arvutada võimalikult paljudele hindajatele tulemusliku kvaliteediindeksi. Kui hindajal R-ist suurema kvaliteediindeksiga paarilist pole olnud, jäetakse talle kvaliteediindeks arvutamata.

Kvaliteediindeksite arvutamist jätkatakse iteratiivselt, kuni on nõudeid täitvaid paare.

Katseuurimuse materjal

Meetodit katsetati ja kontrolliti Riikliku Eksami- ja Kvalifikatsioonikeskuses eesti keele algtaseme kolme eksami suulise osa hindamise andmetel. 2005. a märtsi, aprilli ja mai algtaseme eksamil hinnati kokku 1567 suulist sooritust, hindajaid oli 54. Kasutati kahekordset hindamist. Kuna hindajate paardesse jagamist nendel eksamitel ette ei planeerita, siis on ka hindaja hinnatud tööde ning tema paariliste arv väga erinev, kõikides 1 hinnatud tööst kuni 209 tööni ja 1 paarilisest 19 paariliseni (vt tabel 2).

REKK kasutab hinnete usaldusväärsuse kontrollimiseks vahede arvutamise meetodit: ümber hinnatakse tööd, mille hindeerinevused kahekordsel hindamisel on 3 või enam punkti 10-punktilisel hindamisskaalal.

Tabel 2. Hindaja kood, temaga paaris hinnatud hindajate arv ja hinnatud tööde arv. Poolpaksu kirjaga on tähistatud eksperthindajad

Hindaja kood / paariliste arv / hinnatud tööde arv				
001 / 6 / 92	022 / 1 / 1	038 / 1 / 1	063 / 2 / 52	100 / 3 / 51
002 / 8 / 119	023 / 3 / 40	043 / 1 / 10	064 / 1 / 19	101 / 2 / 29
004 / 10 / 139	024 / 19 / 139	045 / 1 / 15	065 / 1 / 15	104 / 4 / 68
005 / 3 / 42	027 / 2 / 33	046 / 3 / 56	067 / 3 / 45	107 / 1 / 16
007 / 1 / 16	028 / 1 / 14	051 / 2 / 24	069 / 1 / 17	109 / 1 / 16
008 / 7 / 84	030 / 8 / 139	054 / 3 / 43	070 / 1 / 16	110 / 3 / 53
009 / 7 / 127	031 / 1 / 15	055 / 3 / 46	074 / 1 / 16	122 / 2 / 27
012 / 10 / 157	033 / 1 / 14	057 / 3 / 49	081 / 3 / 52	123 / 1 / 7
013 / 3 / 41	034 / 8 / 130	059 / 14 / 209	086 / 9 / 152	124 / 1 / 13
014 / 10 / 144	035 / 3 / 50	061 / 2 / 33	087 / 7 / 136	125 / 2 / 34
017 / 1 / 17	037 / 1 / 23	062 / 3 / 48	091 / 10 / 138	129 / 3 / 50

REKK-i algtaseme eksami peaspetsialisti soovitusel käsitati eksperthindajatena hindajaid koodiga 014, 024, 091 (tabelis 2 poolpaksus kirjas). Eksperthindajate järjekindlust pole uuritud, nende koostis on sama mis teistel hindajatel, ent neil on suurem kogemus ning peaspetsialisti usaldus seniste hindamiste põhjal.

Soovitatud eksperthindajatel lasti sõltumatult hinnata 16 tööd ning arutati pandud hinnete korrelatsioonid esmase hinnangu saamiseks hindamise kokkulangemise ulatuse kohta (vt tabel 3).

Tabel 3. Eksperthindajate hinnete vahelised korrelatsioonid 16 ühesuguse töö hindamise põhjal

Ekspert	E 014	E 091
E 014	1	
E 091	0,839	1
E 024	0,913	0,844

Eksperthindajate kvaliteediindeksiks võeti 1 ja arutati ülejäänud 51 hindaja kvaliteediindeksid. Tulemusliku kvaliteediindeksi arvutamisel kasutati koefitsiendi R väärtust 0,97.

Tulemused ja järeldused

Hindajate kvaliteediindeksid on esitatud kahanevas järjestuses tabelis 4. 9 hindajale polnud võimalik kvaliteediindeksit arvutada, sest nad olid hinnanud 2 või vähem tööd või ei olnud hinnanud paaris ühegi eksperthindajaga või eksperthindajasar-nase hindajaga.

Mida väiksem on hindaja kvaliteediindeks, seda erinevamalt ta eksperthindajast hindab ja seda enam on alust kahelda pandud hinnete usaldusväärsuses. Alates pin-gerea viimasest, tuleks alustada tööde uuesti hindamist, ning seda mitte juhusliku hindaja poolt, vaid eksperthindaja või temaga sarnaselt hinnanud hindajate poolt (s.t tabeli esimesed hindavad üle viimaseid).

Tabel 4. Hindajate kvaliteediindeksid

Hindaja kood	Kvaliteediindeks	Hindaja kood	Kvaliteediindeks	Hindaja kood	Kvaliteediindeks
014	1	031	0,976	057	0,951
024	1	122	0,976	086	0,948
091	1	104	0,973	002	0,948
030	0,995	035	0,973	109	0,940
087	0,992	129	0,972	107	0,936
059	0,991	001	0,971	027	0,928
023	0,991	005	0,971	061	0,923
043	0,991	008	0,970	123	0,891
012	0,990	017	0,969	022	–
067	0,990	054	0,967	028	–
100	0,989	081	0,967	037	–
013	0,989	051	0,965	038	–
110	0,988	004	0,963	064	–
062	0,988	034	0,963	065	–
046	0,987	101	0,958	069	–
124	0,986	009	0,958	070	–
055	0,984	063	0,957	074	–
007	0,982	045	0,955		
125	0,980	033	0,955		

Meetodi kontrollimiseks võeti eksamirühm, kus sooritust olid hinnanud H_{061} ja H_{001} . Kvaliteediindeksilt paiknes H_{061} pingerea lõpuosas ja H_{001} pingerea keskel (vt tabel 4). Hindajate eksamil pandud hinded ja hinnete vahed on esitatud tabelis 5. Suure hinnetevahe tõttu suunas REKK kolmandale hindamisele 6 tööd.

Eksperthindajal E_{091} lasti hinnata samu töid (vt tabel 5).

Eksperthindaja hinded erinevad H_{061} hinnetest oluliselt 5 juhul (vt tabelis 5 $H_{061} - E_{091}$), neist 2 ei tulnud REKK-is kasutatava hinnete vahede arvutamise meetodiga välja ja läksid ebaõigetena eksaminandi tunnistusele: üks eksaminand sai põhjendamatult madala hinde ja teine põhjendamatult kõrge hinde (vt tabelis 5 eksaminandid 2 ja 6).

Pingerea keskel oleva H_{001} hinnetest erinesid eksperthindaja hinded oluliselt 1 juhul (vt tabelis 5 $H_{001} - E_{091}$), mis läks ka eksaminandi tunnistusele ebaõigena (vt tabelis 5 eksaminand 2).

Tabel 5. Eksamil hinnatud hindajapaari hinded ja hinnete vahed; eksperthindaja hinded ning tema ja hindajate hinnete vahed

Eksaminand	H_{061}	H_{001}	$H_{061}-H_{001}$	Kommentaar	E_{091}	$H_{061}-E_{091}$	$H_{001}-E_{091}$	Kommentaar
1	9	7	2		8	1	-1	
2	4	4	0		7	-3	-3	tunnistusele läks ebaõige hinne
3	10	7	3	hinnati uuesti	6	4	1	
4	9	6	3	hinnati uuesti	6	3	0	
5	7	7	0		7	0	0	
6	10	8	2		7	3	1	tunnistusele läks ebaõige hinne
7	5	4	1		5	0	-1	
8	5	5	0		4	1	1	
9	7	4	3	hinnati uuesti	5	2	-1	
10	7	7	0		6	1	1	
11	6	3	3	hinnati uuesti	5	1	-2	
12	9	6	3	hinnati uuesti	6	3	0	
13	7	6	1		8	-1	-2	
14	10	6	4	hinnati uuesti	8	2	-2	
15	9	7	2		8	1	-1	
16	9	7	2		8	1	-1	
\bar{x}	7,7	5,9			6,5			
SD	2,0	1,5			1,3			

Tabel 6. Hindajate ja eksperthindaja pandud hinnete vaheline korrelatsioon

	H_{061}	H_{001}	E_{091}
H_{061}	1		
H_{001}	0,735	1	
E_{091}	0,505	0,592	1

Näitest on näha, et hindajapaari hindevahe arvutamine pole piisav, et üles leida ebaõigesti hinnatud töid. Hindevahede arvutamisel läheb kolmandale hindamisele töid, mis kolmandat hindamist ei vajaks, samas jääb osa ebaõigesti hinnatud töid uuesti hindamata.

Tabelis 6 esitatud hindajate ja eksperthindaja hinnete vaheline korrelatsioon näitab hindajapaari omavahelist suuremat kooskõla kui eksperthindajaga. See viitab asjaolule, et pelgalt korrelatsioonile tuginedes ei saa otsustada hindamise õigsuse üle.

Hindaja kvaliteediindeksi arvutamine aitab mõlema laialt kasutatud meetodi (korrelatsiooni ja hinnete vahe arvutamise) puudusi vältida.

Diskussioon

Hindaja kvaliteediindeksi leidmise meetod on kasutatav vaid kahekordse hindamise korral. Kahekordne hindamine pole väga levinud, kuigi tõstab hindamise usaldusväärsust (Cannings jt 2005, Brooks 2004). Suure panusega eksamil, millest sõltub inimese tulevik, on aga iga võtte hindamise usaldusväärsuse tõstmiseks vajalik ning õigustab sellega kaasnevat kulutusi.

Kuid kulutused pole ainus põhjus, miks kahekordset hindamist omaks ei võeta. Üheks põhjuseks on peetud kahekordse hindamise puhul harjumuspärasest ühekordsest hindamisest erinevat filosoofiat: kui ühekordsel hindamisel teostab (pisteliste) kontrolli eksperthindaja, kelle pandud hindeid loetakse usaldusväärsemateks nii tema suurema hindamiskogemuse kui ka selle tõttu, et paljud neist on ise osalenud hindamisskaalade väljatöötamisel, siis kahekordsel hindamisel, kus "õige hinne" kujuneb kahe hindaja pandud hinnete keskmisest või siis hinnete erinevuse korral paariliste või rohkemate hindajate kompromissist, on eksperthindaja roll taandunud. See pole aga hindamismetodoloogias kergelt omaksvõetav muutus (Brooks 2004).

Kvaliteediindeksi leidmise meetod toetub ühekordse hindamise filosoofiale ja hindajakoolituse põhimõtetele, mis peaks hindajatele olema mõistetav ja vastuvõetav. Ent ka see meetod eeldab tööde ümberhindamisel mõtteviisi muutust: meetod ei tegele ebaõigete üksikhinnete leidmisega (nagu kahe hindaja pandud hinnete vahede arvutamine), vaid hindajaga ja tema töö kvaliteediga. Kuid ka siin võib näha analoogiat hindajakoolitusega: ka hindajakoolituses pole olulised üksikud eksperthindajast erinevalt hinnatud tööd, vaid üritatakse selgust saada hindaja järjekindlusest ning rangusest. Kvaliteediindeksi järgi leitakse valesti hinnanud hindajad, ent valestihindamise põhjus selgub alles siis, kui eksperthindaja või temaga sarnane hindaja hakkab töid uuesti hindama.

Uuesti hindamine on otstarbekas siduda koolitusega: ümberhinnatavate tööde hindajat tuleb ebaõnnestunud hindamisest teavitada ja koos temaga arutleda võimalike põhjuste üle.

Meetod tõstab oluliselt eksperthindaja rolli. See tähendab aga eksperthindaja kavakindlat koolitamist, tema järjekindluse uurimist ja ka atesteerimist. Olukord, kus eksperthindaja töö kvaliteet pole eksamikeskuses teada ja tugineb vaid oletusel ja usaldusel või ammusel analüüsil, on lubamatu. Isegi järjekindlad ja kogenud hindajad muutuvad ajapikku ebajärjekindlamaks (Lumley, McNamara 1995, Lumley 2002).

Mida rohkem suudetakse eksperthindajaid välja koolitada, seda enamate hindajatega saab neid paaris hindama panna ja seda enam saab infot teiste hindajate kohta. Võrreldes kõikide hindajate koolitamisega on eksperthindajate koolitamine igati otstarbekam: üheks eksamisessiooniks palgatud hindajat ei saa mõnepäevase koolitusega muuta järjekindlaks hindajaks, küll aga on võimalik välja koolitada inimesi, kes pidevalt on hindamise ja eksamitega seotud. Koolitatud ja atesteeritud eksperthindaja on võimeline ise hindajaid koolitama, mis omakorda toob kaasa hindajate töö kvaliteedi tõusu.

Meetod eeldab hoolikat hindajapaaride planeerimist. Võimalikult paljud hindajad peaksid saama hinnata paaris eksperthindajaga, hindajale hindamiseks antavate tööde hulk peaks olema piisavalt suur kvaliteediindeksi arvutamiseks,

hindajapaarid ei tohiks olla fikseeritud, eksperthindajaid ei peaks omavahel paaris hindama (eriti siis, kui eksperthindajaid on vähe). Kui paaride valik on juhuslik ja eksperthindajatega paaris hinnanud hindajaid vähe, siis ei ole võimalik kõigile hindajatele kvaliteediindeksit arvutada.

Kirjandus

- Alderson, Charles J.; Clapham, Caroline; Wall, Dianne 1995. *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Altschuler, Sandra J.; Schautz, Tresa 2006. No Hispanic students left behind: The consequences of "High-Stakes" testing. – *Children & Schools* 28 (1), 5–14.
- Bachman, Lyle F. 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bonk, William J.; Ockey, Gary J. 2003. A many-facet Rasch analysis of the second language group oral discussion task. – *Language Testing* 20 (1), 89–110.
- Brooks, Val 2004. Double marking revisited. – *British Journal of Educational Studies* 52 (1), 29–46.
- Cunnings, Rebecca; Hawthorne, Kamila; Hood, Kerenza; Houston, Helen 2005. Putting double marking to the test: A framework to assess if it is worth the trouble. – *Medical Education* 39, 299–308.
- Cushing Weigle, Sara 1998. Using FACETS to model rater training effects. – *Language Testing* 15 (2), 263–87.
- Eckes, Thomas; Ellis, Melanie; Kalnberzina, Vita; Pižorn, Karmen; Springer, Claude; Szollás, Krisztina; Tsagari, Constance 2005. Progress and problems in reforming public language examinations in Europe: Cameos from the Baltic States, Greece, Hungary, Poland, Slovenia, France and Germany. – *Language Testing* 22 (3), 355–377.
- Fine, Michelle 2005. High Stakes Testing and Lost Opportunities. *The New York State Regents Exams*. – *Encounter* 18 (2), 24–29.
- Helfenbein, Robert 2004. *New Times, New Stakes: Moments of Transit, Accountability, and Classroom Practice*. – *Review of Education, Pedagogy & Cultural Studies* 26 (2/3), 91–109.
- Kondo-Brown, Kimi 2002. A FACETS analysis of rater bias in measuring Japanese second language writing performance. – *Language Testing* 19 (1), 3–31.
- Lumley, Tom 2002. Assessment criteria in a large-scale test: What do they really mean to the rater? – *Language Testing* 19 (3), 246–76.
- Lumley, Tom; McNamara, Tim F. 1995. Rater characteristics and rater bias: Implications for training. – *Language Testing* 12, 54–71
- Lunz, Mary E.; Wright, Benjamin D.; Linacre, John M. 1990. Measuring the impact of judge severity on examination scores. – *Applied Measurement in Education* 3, 331–45.
- Luoma, Sari 2004. *Assessing Speaking*. Cambridge: Cambridge University Press.
- Shohamy, Elana 2001. Democratic assessment as an alternative. – *Language Testing* 18 (4), 373–91.

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

HOW TO ASSESS THE RATERS OF HIGH-STAKES TESTS

Hille Pajupuu

Institute of the Estonian Language

In a situation where, among all tests, the proportion of high-stakes tests is constantly growing, while their results are increasingly used to pass judgements not only on examinees but also on teachers and teaching quality, and, with time, many of those tests have become obligatory, high demands should certainly be set on the sense of responsibility of anyone involved in test development or test use.

Special attention should be paid to the quality of subjective ratings as the writing and speaking parts of a test may often account for half of the total score. If a testee should score lower or higher than their competence is worth, it may change their life as well as that of other people.

The commonly used simple statistics (calculation of differences between the marks awarded by two raters, inter-rater correlation) may actually fail to take account of quite a lot of wrong credits if the raters are many and inadequately prepared.

In order to reduce unfair assessment a method is suggested to identify poorly performing raters and to reassess their results in good time. The method is meant to be used in the case of double marking. Notably, a quality index is computed to show the degree of similarity between the credits given by the rater to be assessed and an expert rater, even if the two have never worked in a pair. It is assumed that the higher the similarity the fairer the credits.

The article describes the general principles of the method suggested, pointing out its advantages over some other simple methods used for the same purpose.

Keywords: rater evaluation, rater consistency, rating, double marking, Estonian

KEELEÕPPIJATE VÄLTIMISSTRATEEGIADEST EESTI KEELE TÄIS- JA OSASIHITISE NÄITEL

Raili Pool

Ülevaade. Artiklis antakse ülevaade keeleõppijate poolt ebakindluse korral rakendatava vältimise olemusest, põhjustest ja liikidest ning tutvustatakse uurimust, mille eesmärgiks on välja selgitada eesti keele täis- ja osasihitise eristuse vältimise võimalused ja avaldumisvormid. Informantideks on vene emakeelega eesti keele (võõrkeelena) eriala üliõpilased (kaks keeleoskuselt eri tasemel rühma), keeleaines on kogutud vabast suulisest ja kirjalikust keelekasutusest ning testülesannetest, võrdlusmaterjalina kasutatakse emakeelekõneleja andmeid. Õppijate vaba keelekasutuse materjali põhjal vaadeldakse eelkõige täissihitise vältimise võimalusi (osasihitise ületarvitus tingib täissihitise alatarvituse, mis on käsitletav vältimisena), testülesanded võimaldavad jälgida ka sihitise täielikku vältimist ning selle seost õppijate keeleoskuse tasemega.*

Võtmesõnad: keeleõpe, teise keele omandamine, teise keele kõneleja, emakeelekõneleja, vältimine, objektikäänded, eesti keel

1. Sissejuhatus

Keeleõpetajad puutuvad oma igapäevases töös pidevalt kokku olukordadega, mille puhul üks ja sama õppija võib eri tüüpi ülesannetes näidata üles üsna erinevaid teadmisi õpitava keele kohta. Erinevused avalduvad seejuures kõige ilmsemalt keelevigade rohkuses ja olemuses. Nii võib õppija, kes teist keelt ladusalt räägib ja vabas vormis tekste peaaegu vigadeta kirjutab, teha näiteks emakeelest sihtkeelde tõlkimisel üsna palju vigu. Seesuguse olukorra taustaks on asjaolu, et õppijad püüavad sageli vältida neid õpitava keele vahendeid, mille kasutamisel nad kindlad ei ole. Seetõttu võidakse eriti loovkirjutamisel, kui ajaline piirang puudub, kasutada teadlikult lihtsustatud struktuure, eriti kui tegemist on oluliste eksamitöödega, milles püütakse igati vigu vältida. Vältimisstrateegiate rakendamine komplitseerib keeleõpetaja tööd: nii mõnigi kord on keeruline õpilaste keeleteadmisi hinnata,

* Uurimust on toetanud A. Kordelini fond.

näiteks kui õppija kasutab hindelises töös äratuntavalt primitiivsemaid lausestruktuure kui nendes töödes, mille eest hindeid ei panda. Ka teise keele omandamise uurimisel võivad vältimisstrateegiad olla pigem probleemiks kui toeks, kuna vaba suulise ja kirjaliku keelekasutuse materjal ei pruugi nende tõttu anda täielikku pilti õppijakeele arengust ja selle etappidest. Keeleõppija jaoks seevastu on vältimisel olemas nii omad plussid kui ka miinused. Sihtkeeles korrektsust taotleval teise keele õppijal võimaldab teatud grammatiliste struktuuride või leksika vältimine ennast kahtlemata kindlamini väljendada ning kuna sel viisil modifitseeritud keelekasutus on formaalselt korrektne, võib õppija keeleoskusest jääda petlikult hea ettekujutus. Miinuseks on seejuures asjaolu, et ebakindlate keelevahendite vältimisega ning ainult kindlalt omandatu kasutamisega loobub keeleõppija oma teise keele alaste hüpoteeside proovilepanekust ning tema keeleoskuse areng võib seetõttu aeglustuda.

Käesolev artikkel käsitleb vältimisstrateegiatega olemust ja problemaatikat, põhjalikumalt tuleb vaatluse alla üks konkreetne valdkond eesti keele grammatikast – täis- ja osasihitis. Artiklis otsitakse vastust küsimusele, kas eesti keelt teise keelena kõnelejate suulise ja kirjaliku keelekasutuse materjali põhjal on võimalik sihitise näitel vältimisstrateegiaid empiiriliselt kindlaks teha ja kui on, siis milles need avalduvad, ning kas õppijate keeleoskuse tase ja võimalikud vältimisstrateegiad on omavahel seotud.

2. Mis on vältimine ja kuidas seda ära tunda?

Kontrastiivsest analüüsist välja arenenud veaanalüüsi meetod osutus õppijakeele kirjeldamisel piiratuks seetõttu, et kirjeldas ainult sihtkeele seisukohast ebaõnnestunud keelekasutust, jättes aga kõrvale selle keeleainese, milles õppijad vigu ei teinud. Just sihtkeelepärane keelekasutus on seotud õppijatepoolse vältimisega, mille uurimiseks ei piisa veaanalüüsi vahenditest.¹

Vältimise ilmingut on Carl Jamesi (1980: 183) andmetel esimesena märganud L. Dušková (1969: 29), kelle inglise keelt õppivad tšehhi emakeelega õpilased ei teinud väga keeruliste keeleüksuste tarvitamisel peaaegu üldse vigu, kuna leidsid võimaluse nende kasutamisest mööda minna. Vältimise termini võttis kasutusele Jacquelyn Schachter (1974), kes avastas, et hiinlastest ja jaapanlastest õppijad tegid inglise keele relatiivlausetes vähe vigu, kuigi eelduste kohaselt peaks see konstruktsioon olema neile väga keeruline, kuna kontrastiivne analüüs on näidanud suuri erinevusi õppijate emakeele ja inglise keele vastavate lausestruktuuride vahel. Samas eksisid muude emakeeltega õppijad inglise keele relatiivlausetes märksa rohkem. J. Schachter leidis, et tema hiina ja jaapani õpilaste väiksem vigade arv oli põhjustatud sellest, et nad kasutasid inglise keele relatiivlauseid vähem kui muude emakeeltega õpilased, järeldades, et kui mingi sihtkeele konstruktsioon on õppija jaoks raske, püütakse seda vältida. J. Schachter tulemus oli teise keele omandamise uurimise alal olulise tähtsusega, samas ei andnud mõistagi vastuseid kõigile vältimist puudutavatele küsimustele.

Esimeseks küsimuseks, mida vältimise uurimisel tuleks silmas pidada, on õppijate teadlikkus sellest struktuurist, mille võimalikku vältimist uuritakse. Sõnavara vältimist uurinud Batia Laufer (2000) rõhutab, et õppijad ei saa vältida seda, mida

¹ Kontrastiivsest analüüsist ja veaanalüüsist lähemalt vt nt James 1998, Ellis 1985, Sajavaara 1999.

nad ei tea, teisisõnu eeldab vältimine sihtkeele struktuuri mõningast tundmist ja oskust asendada seda mõne teise struktuuriga, mis on õppijale vähem keeruline ja põhjustab vähem vigu. Ka C. James (1980: 183) nimetab vältimise eeldusena õppijate oskust mitte vältida ehk siis valida kasutamiseks just vaatlusalune vorm. Yan Liao ja Yoshinori J. Fukuya (2004: 194) kritiseerivad eelpool kirjeldatud J. Schachteri uurimust just sellest seisukohast, et nimetatud tööst ei selgunud, kas inglise keele relatiivlauseid vältinud hiina ja jaapani õpilased üldse oskasid neid kasutada või oli vältimiseks nimetatud ilming põhjustatud lihtsalt sellest, et õppijad vaadeldud struktuuri ei tundnud. Viimasel juhul ei saa mõistagi olla tegemist teadliku vältimisega. Vältimise uurimiseks on vaja niisiis esmalt välja selgitada, kas õppijad on vaadeldava keelepiirkonnaga tuttavad. Uurimused on samuti näidanud, et vältimist ei soodusta ainult K1 ja K2 erinevus, vaid seda võib põhjustada ka lähte- ja sihtkeele struktuuride sarnasus. Näiteks toob B. Laufer (2000) välja uurimistulemusi, mille kohaselt võidakse K1 ja K2 vahelist sarnasust, seda eriti markeeritud grammatiliste struktuuride või sõnade perifeersetete tähenduste puhul, tajuda erinevusena ning see võib viia nende struktuuride vältimiseni.

Teine oluline vältimisega seonduv probleemistik puudutab keeleliste struktuuride üle- ja alakasutust ning nende protsesside vastastikuseid seoseid. Vältimine on seotud mingi struktuuri alakasutusega õppijakeeles, kusjuures alakasutuse kindlakstegemiseks tuleb õppijakeelt võrrelda emakeelekõnelejate keelekasutusega: kui vaadeldav struktuur esineb õppijakeeles väiksema sagedusega kui emakeelekõnelejal, võib seda käsitada lahknevusena sihtkeele normist (Odlin 1989: 36). Mingi keelendi või konstruktsiooni alakasutus ei pruugi siiski alati olla vältimise tagajärg, asi võib olla ka selles, et õppijate emakeeles võib vaadeldav struktuur olla märksa väiksema esinemissagedusega (Larsen-Freeman, Long 1999: 62). Sageli toob ühe struktuuri alakasutus õppijakeeles endaga kaasa mingi teise struktuuri ülekasutuse, nii võib näiteks eesti keelt teise keelena õppija liigtarvitada partitiivobjekti, mille tagajärjeks on täissihitise alakasutus; võrdluseks tuleb aga jälgida emakeelekõnelejate valikuid sarnastes kontekstides.

Vastates artikli käesoleva osa pealkirjas esitatud küsimusele võib öelda, et mingi struktuuri vältimise äratundmise eelduseks on andmed, et õppijatel on olemas selle kasutamiseks vajalikud teadmised (mida nad nt vigade tegemise kartuses või mõnel muul põhjusel rakendada pelgavad), kuid ometi tarvitavad nad vaadeldavat struktuuri harvemini kui emakeelekõnelejad.

3. Vältimise liikidest

Keeleõppe eesmärgiks on saavutada teises keeles suhtluspädevus, mis Michael Canale (1983) järgi koosneb neljast komponendist: grammatikapädevus, diskursuspädevus, sotsiolingvistiline pädevus ja strateegiapädevus. Viimane realiseerub keeleõppijate oskuses kasutada suhtlusstrateegiaid. Claus Færch ja Gabriele Kasper (1983: 36) defineerivad suhtlusstrateegiaid kui potentsiaalselt teadlikke plaane, mida nii K1 kui ka K2 kasutajad rakendavad suhtluseesmärkide saavutamisel tekkivate probleemide lahendamiseks. Suhtlusstrateegiate peamiste liikidena toob Elaine Tarone (1983) välja parafraseerimise, laenamise, abiotsimise, miimika kasutamise ja vältimise², jagades vältimise kaheks alaliigiks: a) teema vältimine – õppija

² Suhtlusstrateegiate detailsemat liigitust vt nt Dörnyei, Scott 1997, eesti keelt teise keelena õppijate suhtlusstrateegiatest vt Hennoste, Vihalemm 1999, Kikerpill, Pool 2003.

püüab mitte rääkida teemal, millest rääkimiseks vajalikke sihtkeele struktuure ta ei tunne; b) teema katkestamine – õppija alustab mingil teemal rääkimist, kuid ei oska jätkata ja katkestab jutu.

C. Færch ja G. Kasper (1983) jagavad suhtlusstrateegiad kaheks pealiigiks vastavalt sellele, kas õppijad muudavad probleemi tekkimisel suhtluseesmärki, püüdes sel viisil probleemist lahti saada (reduksioonistrateegiad) või püüavad rakendada alternatiivset plaani (saavutusstrateegiad). Vältimisel põhinevad reduksioonistrateegiad, mille rakendamise põhjusena toovad nimetatud autorid välja õppijate soovi vältida vigu ja/või soovi suurendada oma teise keele kasutuse soravust. C. Færch ja G. Kasper eristavad formaalseid ja funktsionaalseid reduksioonistrateegiaid. Esimesed võivad avalduda vahekeele kõigil tasanditel (fonoloogiline, morfoloogiline, süntaktiline ja leksikaalne tasand), teised on sisult samad kui eelpool nimetatud E. Tarone'i poolt esitatud vältimise jaotamine teema vältimiseks ja teema katkestamiseks. Douglas H. Brown (1994: 118–120) toob vältimise tavalisemate liikidena välja mingi semantilise kategooria sisese süntaktilise või leksikaalse vältimise ning fonoloogilise vältimise, pidades veel otsemaks vältimise liigiks teema vältimist, mis võib avalduda erineval moel: teema vahetamine, arusaamatuse teesklemine, küsimusele mittevastamine, poolelijätmine jne.

Ükskõik mis liiki vältimist rakendatakse, on see alati seotud keeleõppija puudulike teadmistega mingist sihtkeele piirkonnast. Veaanalüüsi spetsialist C. James (1998: 176) kirjeldab vältimise avaldumist õppijakeeles esinevate vigade olemuse kaudu: teema vältimise korral on tegemist varjatud vigadega, mis on põhjustatud mingi struktuuri alakasutusest; parafraseerimise või ümberütlemise puhul ilmnevad aga paljusõnalisusest ja ebamäärasusest tingitud ilmsed vead. Käesoleva artikli 2. jaotises käsitletud keeleliste struktuuride ala- ja ülekasutust on seega tarvis vaadelda seoses vältimise liikidega.

4. Mida saab ja mida ei saa keeles vältida

Vältimist uurides tuleks esmalt arutleda selle üle, kas kõik keele piirkonnad on võrdsel määral välditavad või on keeles ka midagi sellist, mille vältimine õppijatel ka parima tahtmise juures eriti ei õnnestu. Suhtlussituatsioonis keeleliselt raske või sisult ebameeldiva teemaga kokku puutudes saab ilmselt alati leida võimaluse jutuaine muutmiseks, äärmisel juhul võib kõneleja lihtsalt vait jääda või raskele küsimusele mitte vastata. Probleemaatilisem on keeleõppija olukord siis, kui keeruliseks osutub mingisugune grammatiline kategooria. Siin hakkab rolli mängima vaadeldava kategooria esinemissagedus sihtkeeles ja roll lausete moodustamisel.

Teise keele õppijate keeleoskuse hindamiseks ja probleemsete piirkondade väljaselgitamiseks koostatakse õppijakeeles sagedamini esinevate veatüüpide klassifikatsioone. Seejuures on aga oluline silmas pidada, et mingi vealiigi suur esinemissagedus ei tähenda tingimata, et see kategooria, milles viga tehti, oleks õppijatele kõige raskem. Soomerootslaste soomekeelsetes kirjandites tehtud vigu klassifitseerinud Eeva Lähdemäki (1995: 2) nendib, et vea suur esinemissagedus tähendab hoopis seda, et vastavat kategooriat on tekstis raske vältida. Raili Pooli ja Elle Vaimanni (2005) esitatud vene emakeelega eesti keelt teise keelena kõrgtasemel kõnelejate kirjutistes esinenud sagedasemate vigade loendis olid sageduselt

esimesel kohal sõnajärjevead, teisel kohal sõnavalikuvead (leksikaalsed vead, mille parandamiseks piisab ühe sõna asendamisest teisega) ning kolmandal täis- ja osasihitise valiku vead. Kuigi nimetatud eesti keele piirkonnad (eriti sihitise kasutus) keeleõppijatele lihtsad kindlasti ei ole, võib nimetatud vealiikide suurt esinemissagedust siiski mõjutada ka asjaolu, et need kategooriad on raskesti välditavad. Eriti puudutab see sõnajärge, ilma milleta lauseid moodustada ei saa.

Keeleõppijate vigade analüüsimisel ja veatuübi esinemissageduse ning vastava keelekategoria keerulisuse või lihtsuse vastastikuste seoste otsimisel tuleb silmas pidada ka vaadeldava kategoria tavapärasest esinemust sihtkeeles. Näiteks kui õppijakeele kirjalikus ainestikus ei esine ühtki käskiva kõneviisi moodustamise või kasutamiseiga seotud viga, võib arvata, et see kategoria on kas väga hästi omandatud või on tegemist vältimisega. Samas ei saa struktuuri alakasutuse kohta enne järeldusi teha, kui õppijakeele andmeid on võrreldud emakeelekõneleajate andmetega. Kui emakeelekõneleajate analoogilistes kirjutistes esineb käskivat kõneviisi rohkem, on võimalik rääkida selle alakasutusest õppijakeeles. Kui aga kummagi rühma kirjutistes pole käskivat kõneviisi tarvitatud, ei saa vaba keelekasutuse materjali põhjal õppijakeele kohta järeldusi teha ning tuleb lisaks rakendada muid materjali kogumise viise. Siinkohal meenutatagu, et vältimisest saab rääkida ainult siis, kui eelnevalt on kindlaks tehtud, et vaadeldavad õppijad on uuritava kategoriaiga (nt käskiva kõneviisiga) tuttavad, sest mida õppija ei tea, seda ei saa ta ka vältida.

5. Täis- ja osasihitise vältimise võimalustest

5.1. Materjal ja informandid

Et empiirilisel uurida võimalikke vältimisi objektikasutuses, on registreeritud käändsõnaliste objektide kasutusjuhud vene emakeelega eesti keelt teise keelena kõneleajate vabast suulisest ja vabast kirjalikust eesti keelest, võrdlusmaterjaliks on emakeelekõneleajate analoogsetest kirjutistest kogutud ainestik. Lisaks vabale keelekasutusele tuuakse näiteid ka objektikasutuse uurimiseks läbi viidud testülesannete vastustest. Käsitletav materjal jaguneb järgmisteks osadeks.³

- 1) **Kesktaaseme kirjalik keelekasutus:** 28 eesti keelt kesktaasemel valdava Tartu ülikooli eesti keele (võõrkeelena) eriala üliõpilase eksamikirjanditest kogutud objektid. Kirjandite keskmine pikkus 508 sõna, tekstikogu maht kokku 14 227 sõna, sh 637 objekti.
- 2) **Kesktaaseme suuline keelekasutus:** 11 eesti keelt kesktaasemel valdava Tartu ülikooli eesti keele (võõrkeelena) üliõpilasega läbi viidud suuliste vestluste litereeringutest kogutud objektid. Lindistusi kokku 5 tundi, millest on registreeritud 379 objekti.
- 3) **Kõrgtaaseme kirjalik keelekasutus:** 28 eesti keelt kõrgtaasemel valdava Tartu ülikooli eesti keele (võõrkeelena) üliõpilase eksamikirjanditest kogutud objektid. Kirjandite keskmine pikkus 369 sõna, tekstikogu maht kokku 10 332 sõna, sh 581 objekti.
- 4) **Kõrgtaaseme suuline keelekasutus:** 15 eesti keelt kõrgtaasemel valdava Tartu ülikooli eesti keele (võõrkeelena) üliõpilasega läbi viidud suuliste

³ Kesktaaseme määramise aluseks on informantide poolt venekeelse gümnaasiumi lõpus sooritatud eesti keele riigieksam, mis on võrdsustatud kesktaaseme eksamiga. Kõrgtaaseme informandid on edukalt sooritanud eesti keele kõrgtaaseme eksami (kõrgtaaseme kirjalik materjal on pärit eksamikirjandist, millele pääsemise eeltingimuseks oli kõrgtaaseme eksami läbimist tõendava tunnistuse esitamine).

vestluste litereeringutest kogutud objektid. Lindistusi kokku 10 tundi, millest on registreeritud 882 objekti.

- 5) **Emakeelekõnelejate kirjalik keekekasutus:** 16 eesti keelt emakeelena kõneleva Tartu ülikooli eesti keele (võõrkeelena) eriala üliõpilase eksamikirjanditest (materjal on pärit sama eksami kirjandiosast, millest kõrgtaseme kirjalik aines, materjal on seega võrreldav) kogutud objektid. Kirjandite keskmine pikkus 398 sõna, tekstikogu maht kokku 6368 sõna, sh 297 objekti.
- 6) **Testülesannete vastused:** 2 liiki ülesannetest kogutud ainek (valikvastustega ülesanne ja tõlkelaused vene keelest eesti keelde), vastajateks 29 kõrgtasemel eesti keelt valdavat eesti keele (võõrkeelena) eriala üliõpilast, 15 venekeelse gümnaasiumi 11. klassi õpilast ja 30 eesti keelt emakeelena kõnelevat Tartu ja Tallinna kõrgkoolide erinevate erialade üliõpilast.

Vaba suulise ja kirjaliku keekekasutuse sihitised on liigitatud uurimuse eesmärgi silmas pidades järgmiselt: 1) muutumatu peasõnaga kvantorifraasidena vormistatud objektid, 2) eitavate konstruktsioonide objektid, 3) vormihomonüümiaga objektid (juhtumid, mille puhul objektnoomeni vormihomonüümiat tõttu pole üheselt selge, millist objektikäändet on soovitud kasutada) ning 4) põhiliigituse objektid (jaatavas lauses, ilma vormihomonüümiata noomenid), mis jagunevad partitiiv-, genitiiv- ja nominatiivobjektideks. Emakeelekõnelejate kirjandite ainek on liigitatud samadel alustel kui õppijakeele materjal.

5.2. Uurimuse lähtealus ja hüpoteesid

Käesoleva uurimuse 15 kõrgtaseme suulise keekekasutuse informandilt on kogutud ka introspektiivne materjal: vaba vestluse lindistuse lõpus esitati neile veel 7 eestikeelset lauset, milles paluti panna sihtis nende arvates sobivasse käändesse ja oma valikut põhjendada (põhjalikumalt vt Pool 2006b). Introspeksioon andis infot ka vältimise kohta: informandid kinnitasid, et nad on eestikeelses tekstis nii mõnigi kord terve lause just objekti tõttu ümber teinud, näiteks⁴:

- (1) INF L: kui kirjutatan kirjandis näiteks ja ei oska valida kas omastav või osastav (.) siis lihtsalt võtan teise lause

Näites (2) tuleb esile keeleõppija soov vältida lauses subjektina talitlevat *da*-infinitiivi laiendava objektiga konstruktsiooni, mille kasutamisel ta ennast ebakindlalt tunneb. Informantidele objektikäände valiku põhjenduseks pakutud lause *Vene emakeelega inimestel on Tartu koolides raske (eesti keele õpetaja koht)* leida analüüsimisel arenes küsitaja ja INF N vahel järgmine mõttevahetus:

- (2) INF N: mina oleks vist kindlasti kuidagi teisiti sedasama öelnud
R.P.: nii et kui sa ei ole kindel endas sa valiksid teise konstruktsiooni?
INF N: jaa (.) tegelikult ma ette ei mõtle kuidas ma seda ütlen aga
R.P.: aga kirjalikus tekstis?
INF N: jah (.) kirjalikus tekstis kindlasti (.) kui ma loen lauset ja tunnen et siin ma kahtlen (.) siis ma kirjutatan ümber

⁴ Näited (1) ja (2) on võetud artiklist Pool 2006b, milles on informandid tähistatud lühenditega INF A jne, küsitajat märgib lühend R.P. (Raili Pool).

R.P.: Kas sul on olnud ka selline olukord mõnikord et sa just selle objekti selle sihitise pärast hakkad kahtlema?

INF N: kindlasti kindlasti (.) see ei ole mõnikord see on päris tavaline asi

Introspektiivsed andmed kinnitavad seega, et täis- ja osasihitise kasutamisel on vältimise rakendamine tõenäoline, kuid ei anna täpsemat infot selle kohta, missugusel viisil see võib õppijate keelekasutuses realiseeruda.

Võttes arvesse asjaolu, et partitiiv on eesti keeles sihitise põhikäänne, niisiis opositsiooni markeerimata liige (vt nt Erelt jt 2000: 377, Raag 1985: 205, Tauli 1968: 216) ning et ka käesoleva uurimuse informantide keelekasutuses esines kõige rohkem just partitiivobjekti, on järgnevates hüpoteesides lähtunud peamiselt täis-sihitise vältimise ja vastavalt partitiivobjekti ületarvituse võimalustest.

- 1) Õppijad võivad tarvitada objektidena rohkesti kvantorifraase muutumatute peasõnadega *palju, rohkesti* jne: kvantorite partitiivireksioon peaks kesktasemel (ning kõrgtasemel) eesti keele kõnelejalatel juba omandatud olema.
- 2) Kuna eitavate konstruktsioonide objektid on alati partitiivis ning see reegel (lihtsaim, mis objektikasutust reguleerib) peaks kesktasemel omandatud olema, võidakse kahtluse korral vormistada lause eitavana – ka niisugustel juhtudel langeb ära objektikäänete valiku vajadus.
- 3) Eesti keelele omase vormihomonüümia tõttu langevad osas noomenitüüpides kaks või ka kolm sihitisekäännet vormiliselt kokku. Keele struktuurist teadlike filoloogiaüliõpilaste puhul pole välistatud võimalus, et sihitiseks võidakse valida just vormihomonüümiaga noomen, mida võib tekstis interpreteerida nii täis- kui osasihitiseks. Seesugune valik võimaldab õppijatel vältida vigu.
- 4) Kuna informandid on partitiivobjekti tarvitamisel kindlamad, kasutavad nad seda objektikäännet emakeelekõnelejatega võrreldes rohkem ning vastavalt tuleb esile täissihitise alatarvitamine, mida võib käsitleda täissihitise vältimisena.
- 5) Tõlkeülesandes võidakse valida ka sihitist täielikult vältiv tõlkevariant.

Vältimisstrateegiate rakendamise tõenäosust toetab asjaolu, et nii kesk- kui ka kõrgtaseme keelevaldajate vaba kirjaliku keelekasutuse materjal on kogutud informantide jaoks olulistest eksamitöödest, milles püütakse igati vigu vältida. Suulise keelekasutuse ja testülesannete kohta informandid hinnangulist tagasisidet ei saanud, mistõttu neil puudus otsene vajadus oma keelekasutuse korrektsust pingsalt jälgida.⁵

5.3. Muutumatu peasõnaga kvantorifraas sihitisena

Sihitiseks esinevad kvantorifraase, mille põhjaks on määrsõnad *palju, vähe, oht-rasti, rohkesti* jne, võivad keeleõppijad tajuda lihtsamana, kuna need esinevad alati kujul kvantor + noomen partitiivis ning nende puhul pole vaja mõelda käändevalikule nagu noomenist sihitise puhul (nt kui kaheldakse, kas valida nominatiiv- või partitiivobjekt, võidakse valida kvantorifraas ehk nt valiku *leidsin seeni* või *leidsin seened* asemel *leidsin palju seeni*). Seesuguseid kvantorifraase esines materjali hulgas järgnevalt (siin ja edaspidi on esitatud nii absoluutarvud kui ka protsentuaalne esinemus materjalikogus tervikuna):

⁵ Kuna suuliste lindistuste läbiviijaks ja ühtlasi informantide vestluspartneriks oli artikli autor, kes on ühtlasi informantide õppejõud, pole välistatud, et informandid püüdsid siiski ka suulises kõnes ennast võimalikult korrektselt väljendada.

- 1) kesktaseme kirjalik 20 (3,1%)
- 2) kesktaseme suuline 3 (0,8%)
- 3) kõrgtaseme kirjalik 18 (3,1%)
- 4) kõrgtaseme suuline 10 (1,1%)
- 5) emakeelekõnelejad 3 (1,0%)

Kuna erinevate andmekogude materjalihulk ning sellest tingitult ka absoluutarvud on erinevad, on siin ja edaspidi mõttekam võrrelda eelkõige protsentuaalseid andmeid. Kesk- ja kõrgtaseme kirjalikus keelekasutuses esines emakeelekõnelejatega võrreldes rohkem kvantorifraasidest sihitisi, näiteks:

(3) Nad andestavad endale palju ebameeldivaid omadusi ja puudusi tol ajal, kui ei salli samu omadusi teistes inimestes ja ise ei taipa seda. (kesktaseme kirjalik)

(4) Kuigi usun, et ülikool andis mulle palju uusi ja vajalikke teadmisi ning laienes mu silmaringi.⁶ (kõrgtaseme kirjalik)

Suulise ja kirjaliku keelekasutuse andmeid võrreldes paistab silma kvantorifraaside sagedasem kasutus kirjalikes töödes. Kuna kirjandite kirjutamise ajal oli aega oma töö ülekontrollimiseks ning tegemist oli eksamitöödega, kus keelelise korrektsuse saavutamine oli informantide jaoks oluline, ei ole välistatud võimalus, et muutumatute peasõnadega kvantorifraaside kasutamine oli mõnigi kord teadlik valik ning seega seotud vältimisega (niisugust järel dust toetab ka emakeelekõnelejatest sagedasem kvantorifraaside kasutus).

5.4. Eitavate konstruktsioonide objektid

Arvandmetes eitavate konstruktsioonide kasutamise kohta on esitatud ka ebakorreksete objektide arv:

- 1) kesktaseme kirjalik 66 (10,4%), sh 3 viga
- 2) kesktaseme suuline 56 (14,8%), sh 3 viga
- 3) kõrgtaseme kirjalik 71 (12,2%), sh 1 viga
- 4) kõrgtaseme suuline 126 (14,3%), sh 5 viga
- 5) emakeelekõnelejad 40 (13,5%)

Väikest vigade arvu arvestades võib pidada eesti keele reeglit, mille kohaselt eitus nõuab partitiivobjekti, vaadeldavatel informandirühmadel tervikuna hästi omandatuks. Vigu oli vähe, ning nende hulgas ka selliseid, mille puhul sihitisi ei laiendanud pöördelist verbivormi, vaid *da*-infinitiivi, kus keeleõppijal ei ole eituse mõju lihtne tajuda (5).

(5) Minu arust Eestis ei ole sama olukord ja lähimail aastatel ei ole otstarbekas üldse mingi teine keel (*pro* mingit teist keelt) riigikeelena teha. (kesktaseme kirjalik)

Erinevalt kvantorifraasidest on eitavaid konstruktsioone pisut rohkem tarvitatud just suulises õppijakeeles, kirjalikus materjalis on eitust emakeelekõnelejatega võrreldes kasutatud veidi vähem, kesktasemel on erinevus märgatavam. Vaadel-

⁶ Õppijakeele näited esitatakse autentsel kujul, mistõttu neis võib esineda mitmesuguseid keelelisi ebatäpsusi.

dava materjali põhjal ei ole alust pidada eitust täissihitise vältimisel rakendatavaks strateegiaks, pigem võib kesk- ja kõrgtaseme kirjaliku materjali põhjal oletada, et kesktasemel on õppijad eitavate konstruktsioonide moodustamisel olnud ebakindlamad ja neid vähem tarvitanud.

5.5. Vormihomonüümiaga objektid

5.5.1. Materjali üldandmed

Sihitise vormihomonüümi muudab keerulisemaks keeleteooria töö, kelle eesmärgiks on sihitised käänete järgi liigitada, samas võib kahe või kolme objektikäände vormiline kokkulangevus keeleõppijale hoopis toeks olla, kuna oskusliku noomenivaliku abil on võimalik veaohutlikke kohti vältida.⁷ Kuigi eestlaste kirjutistes sisalduvad vormihomonüümiaga objektid oleks artikli autoril võimalik enda kui emakeelekõneleja keeletaju kohaselt liigitada täpsemalt kas nominatiiv-, genitiiv- või partitiivobjektideks, on need kogu materjali võrreldavuse huvides eraldi rühma arvatud. Vormihomonüümiaga objektide sagedus on järgmine:

- 1) kesktaseme kirjalik 60 (9,4%)
- 2) kesktaseme suuline 19 (5,0%)
- 3) kõrgtaseme kirjalik 40 (6,9%)
- 4) kõrgtaseme suuline 31 (3,5%)
- 5) emakeelekõnelejad 21 (7,1%)

Selle rühma materjali illustreerivad näited (6)–(8).

(6) Sõjad algasid siis, kui inimesed tahtsid oma territooriumi (genitiiv või partitiiv) jaotada. (kesktaseme kirjalik)

(7) mees tooks raha (nominatiiv, genitiiv või partitiiv) koju (kesktaseme suuline)

(8) See on viga, mida tehti siis ja tehakse tänase päevani, tuues endaga kaasa ainult vihavaenu (genitiiv või partitiiv). (emakeelekõnelejad)

Kuna teise keele kõnelejate objektikasutust ei reguleeri samad mehhanismid mis emakeelekõnelejal, ei saa õppijakeeles esinevaid vormihomonüümiaga sihitisi emakeelekõneleja seisukohast liigitada ning selliste juhtumite koondamine eraldi rühma on põhjendatud.

Vormihomonüümiaga objektide statistikat vaadeldes on näha, et kõrgtaseme kirjaliku materjali andmed on emakeelekõnelejate andmetega võrreldes väga sarnased, mistõttu ei saa selle materjalikogu põhjal teha järeldusi vormihomonüümsete objektnoomenite tarvitamise kohta täissihitise vältimise eesmärgil. Emakeelekõnelejatest protsentuaalselt pisut rohkem on vormihomonüümiaga sihitisi ainult kesktaseme kirjalikus materjalis. Kesktaseme vastavate näidete arvukust on mõjutanud objektnoomeni *elu* sagedane kasutus (13 korda), kuid kuna need näited jagunevad 9 informandi kirjutiste vahel, pole põhjust selle sõna liigtarvitusest rääkida, ka eestlaste kirjutistes oli *elu* sagedasim vormihomonüümne objektnoomen (5 korda), kõrgtaseme kirjalikus materjalis esines enim *raha* (5 korda).

⁷ Probleemistikust põhjalikumalt vt R. Pooli artiklist "Sihitise vormihomonüümiaga – keeleteooria probleem, keeleõppijatuugi" (Pool 2006a).

5.5.2. Vormihomonüümia neutraliseerimisest

Tuleb märkida, et käände määratlemise probleemid ei seostu mitte kõigi vormihomonüümsete objektnoomenitega. Riho Grünthali (2001: 55–59) poolt loetletud sihitise vormihomonüümia süntaktilise ja semantilise neutraliseerimise vahenditest (täiendi ühildumine, rinnastus, eitus, umbisikuline tegumood, arvsõnajärgne positsioon, perfektivsed adverbid) välistavad mitmed kahemõttelisuse ka õppijakeeles, näiteks selge käändemarkeeringuga täiendi olemasolu võimaldab näite (9) liigitada partitiivobjektide hulka, näite (10) aga genitiivobjektide hulka, sihitiste rinnastus lubab näidet (11) pidada partitiivobjektiks.

(9) et juhtus nii et minu isa (.) ema mitte aga isa (...) ta lõpetas seda (.) õhukooli (partitiiv) (kõrgtaseme suuline)

(10) Inimene on ise selle illusiooni (genitiiv) loonud ja elabki sellega. (kõrgtaseme kirjalik)

(11) Eespool öeldu lisaks pean tähtsaks, et me säilitaksime kontakti ja tunnet (partitiiv), et me oleme perekond, oleme üheskoos. (kõrgtaseme kirjalik)

Samas ei saa näiteks perfektivsusadverbide olemasolu lauses pidada teise keele kõnelejate keelekasutuse analüüsimisel sihitise vormihomonüümia neutraliseerijaks samadel alustel kui emakeelekõnelejate puhul, kuna ka niisugustel juhtudel võivad keeleõppijad kasutada partitiivobjekti, näiteks:

(12) sel talvel meie maja juures jänesed sõid ühe ploomipuu puukoort ära (.) midagi enam ei kasva vist (kesktaseme suuline)

(13) Aga see on arusaadav – ta on seda ära teeninud. (kesktaseme kirjalik)

(14) siis ma unustasin eesti keelt ära (kõrgtaseme suuline)

Näited (12)–(14) annavad tunnistust sellest, et kui *ära*-adverb esineks koos vormihomonüümse sihitisega, ei saaks emakeelekõneleja keeletajust lähtuvalt partitiivi võimalust välistada, seda muidugi juhul, kui üks kokkulangevatest käänetest on osastav.

Huvitaval kombel esineb suulise õppijakeele ainestiku hulgas vähem vormihomonüümiaga objekte kui kirjaliku keele andmekogudes, mis annab tunnistust keeleõppijate suulise ja kirjaliku keelekasutuse erinevustest. Kõiki jaatavates lausetes esinenud objekte vaadeldes ilmneb, et potentsiaalseid vormihomonüümiaga sihitisi on nii suulises kui ka kirjalikus õppijakeeles kasutatud tegelikult märgatavalt enam, kuid vormihomonüümia on ühilduva atribuudi abil süntaktiliselt neutraliseeritud. Vaatleme, kas suulises õppijakeeles võiks kirjalikuga võrreldes väiksem vormihomonüümiaga objektide arv tuleneda sellest, et suulises kõnes on rohkem potentsiaalseid vormihomonüümseid sihitisi täiendi ühildumise abil neutraliseeritud. Seesuguste objektide suhtarv materjalikogu alaliikides oli järgmine (sulgudes on protsent kõigist vastava allkogu objektidest, eraldi on välja toodud demonstratiivpronoomeni *see* kasutusjuhud vormihomonüümiat neutraliseeriva adjektiivatribuudina):

- 1) kesktaseme kirjalik: 21 (3,3%), neist 4 *see*-atribuudiga NP-d
- 2) kesktaseme suuline: 13 (3,4%), neist 6 *see*-atribuudiga NP-d

- 3) kõrgtaseme kirjalik: 16 (2,8%), neist 1 *see*-atribuudiga NP
- 4) kõrgtaseme suuline: 27 (3,1%), neist 18 *see*-atribuudiga NP-d
- 5) emakeelekõnelejad: 11 (3,7%), neist 0 *see*-atribuudiga NP-d

Arvandmed ei võimalda järeldada, et suulise materjali väiksem vormihomonüümiaga objektide arv võiks olla tingitud sellest, et suulises kõnes tarvitatakse rohkem sihitise vormihomonüümiat neutraliseerivaid ühilduvaid täiendeid kui kirjalikus, seega ei saa siinsete andmete põhjal öelda, et suulises õppijakeeles avalduks suurem sihitisekäände vormiline eristusvajadus kui kirjalikus. Küll aga paistab silma erinevus vormihomonüümseid sihitisi laiendavate atribuutide laadis: kõrgtaseme ainestikus ilmneb suulises kõnes *see*-pronoomeni eelistamine täiendina (27-st 18 ehk 66,7% on *see*-atribuudiga NP-d), vt näide (15), samas kui kirjandites esineb samadel tingimustel vaid üks *see*-atribuudiga NP (10) ning vormihomonüümsete sihitiste täienditeks on üldjuhul ühese käändemarkeeringuga adjektiivid või proadjektiivid (16).

(15) et iga aasta (...) vaatasin seda õppekava ja (kõrgtaseme suuline)

(16) Tihti ei ole need veel küpsed, et oma tulevase eriala õigesti valida. (kõrgtaseme kirjalik)

Siin on esitatud andmed ainult vormihomonüümseid sihitisi laiendavate *see*-pronoomenite kohta, samas kasutati neid ka selge käändemarkeeringuga sihitiste laienditena, kuid need juhtumid ei paku praegu huvi. Käsitatud materjal annab tunnistust kõrgtasemel teise keele kõnelejate oskusest oma pronoomenikasutust vastavalt suhtlussituatsioonile modifitseerida: *see*-pronoomenit adjektiivatribuudina peetakse eelkõige suulisse informaalsesse keelekasutusse kuuluvaks, kirjalik tekst vormistatakse peamiselt täistähenduslike atribuutide abil, milline kasutusviis on lähedane emakeelekõnelejate kirjaliku keelekasutusele. Kas *see*-pronoomeneid sihitisfraasi koosseisus esineb eestlaste suulises kõnes samuti rohkesti või on see kõrgtasemel teise keele kõnelejate erijoon, pole andmete puudumisel võimalik öelda.

Kuigi kokkulangevate käändevormidega noomenite valimine lausesse aitab keeleõppijatel viga vältida, ei ole käesolevas artiklis käsitletava keelematerjali põhjal võimalik empiiriliselt tõestada, et selliseid noomeneid oleks sihitisena kasutatud vältimise eesmärgil. Tõestust ei leidnud ka hüpotees, mille kohaselt võiks suulises materjalis ilmnenu väiksem vormihomonüümsete objektide arv tuleneda sellest, et suulises kõnes neutraliseeritakse sihitise vormihomonüümia sagedamini ühilduvate täienditega kui kirjalikus keelekasutuses.

5.6. Sihitisekäänete üle- ja alatarvitus vältimise avaldumisvormina

5.6.1. Vaba suuline ja kirjalik keelekasutus

Tabelis 1 on esitatud andmed kõigi informandirühmade vabas suulises ja kirjalikus keelekasutuses esinenud objektide põhiliigituse kohta (siin ei kajastu juba käsitletud eitavate verbivormide objektid, sihitisena esinevad kvantorifraasid peasõnadega *palju*, *rohkesti* jms ning vormihomonüümiaga objektid).

Tabel 1. Vaba suulise ja kirjaliku keelekasutuse objektide põhiliigitus

	kesktaseme kirjalik	kesktaseme suuline	kõrgtaseme kirjalik	kõrgtaseme suuline	emakeelekõnelejad
Partitiiv	425 (86,6%)	241 (80,1%)	378 (83,6%)	532 (74,4%)	174 (74,7%)
Genitiiv	36 (7,3%)	26 (6,9%)	42 (9,3%)	91 (12,7%)	26 (11,2%)
Nomin.	30 (6,1%)	34 (9,0%)	32 (7,1%)	92 (12,3%)	33 (14,1%)
KOKKU:	491 (100%)	301 (100%)	452 (100%)	715 (100%)	233 (100%)

Kõrgtaseme suulises materjalis on partitiivobjektide ja täissihitiste omavaheline suhe emakeelekõnelejate andmetega võrreldav, muudes materjali allkogudes aga ilmneb partitiivobjekti suurem ja vastavalt täissihitiste väiksem esinemus, mida emakeelekõnelejate keelekasutuse andmetega võrreldes tuleb lugeda kõrvalekaldeks sihtkeelest ning on võimalik käsitleda täissihitise vältimisena. Objektikäänete tegelikust ületarvitusest saab täpsema ülevaate, võttes arvesse kasutatud sihitise vormide sobivust esinemiskonteksti sihtkeele seisukohast. Tabelites 2–5 esitatakse õppijakeele andmete täpsem liigitus, arvestades nii kasutatud objektide sihtkeelepärast korrektsust kui ka ebakorrektsust (üle- või alatarvitust).

Tabel 2. Kesktaseme kirjaliku keelekasutuse põhiliigitus

	Kokku	Õiged	Valed
Partitiiv	425 (86,6%)	397 (93,4%)	28 (6,6%)
Genitiiv	36 (7,3%)	30 (83,3%)	6 (16,7%)
Nominatiiv	30 (6,1%)	17 (56,7%)	13 (43,3%)
Kokku:	491 (100%)	444 (90,4%)	47 (9,6%)

Tabel 3. Kesktaseme suulise keelekasutuse põhiliigitus

	Kokku	Õiged	Valed
Partitiiv	241 (80,1%)	212 (88%)	29 (12%)
Genitiiv	26 (8,6%)	22 (84,6%)	4 (15,4%)
Nominatiiv	34 (11,3%)	17 (50%)	17 (50%)
Kokku:	301 (100%)	251 (83,4%)	50 (16,6%)

Tabel 4. Kõrgtaseme kirjaliku keelekasutuse põhiliigitus

	Kokku	Õiged	Valed
Partitiiv	378 (83,6%)	338 (89,4%)	40 (10,6%)
Genitiiv	42 (9,3%)	33 (78,6%)	9 (21,4%)
Nominatiiv	32 (7,1%)	20 (62,5%)	12 (37,5%)
Kokku:	452 (100%)	391 (86,5%)	61 (13,5%)

Tabel 5. Kõrgtaseme suulise keelekasutuse põhiliigitus

	Kokku	Õiged	Valed
Partitiiv	532 (74,4%)	513 (96,4%)	19 (3,4%)
Genitiiv	91 (12,7%)	87 (95,6%)	4 (4,4%)
Nominatiiv	92 (12,3%)	70 (76%)	22 (24%)
Kokku:	715 (100%)	670 (93,7%)	45 (6,3%)

Tabelites 2–5 toodud andmetest on näha, et kõigis materjali allkogudes on partitiiv-objektide puhul õigete ja valede kasutusjuhtude protsentuaalne suhe kõige rohkem õigete kasuks, enim vigu on aga põhjustanud nominatiivobjekti sihtkeelepärane tarvitamine, seda eriti kesktasemel.⁸ Materjal näitab, et partitiivi peetakse sihitise põhikäändeks ning seda tarvitatakse ka arvuliselt kõige sagedamini ebasobivas kontekstis, erandiks on ainult kõrgtaseme suuline ainek, kus tuli esile nominatiivobjekti suurem ületarvitus.

Õppijate objektikasutuse sihtkeelepärasest korrektsusest saab täpsema pildi, kui võtta arvesse lisaks mingi vormi korrektsetele kasutusjuhtudele obligatoorses kontekstis⁹ (õiged) ja sama vormi kasutusjuhtudele muudes kontekstides (valed) arvesse ka nende kontekstide arvu, milles vaadeldav vorm oleks obligatoorselt pidanud esinema, kuid mille asemel oli tarvitatud muud objektikäändet. (nt partitiivi kasutamine genitiivi asemel on sihtkeele seisukohast viga mõlemat pidi, esimesel juhul on tegemist käändevormi üle-, teisel juhul alatarvitusega). Kõiki neid aspekte võimaldab sihtkeelepärase korrektsuse protsendi (ingl *per cent accuracy*) arvutamiseks arvesse võtta Teresa Pica (1984) väljatöötatud valem:

$$\frac{n \text{ korrektsed tarvitajuhud obligatoorses kontekstis}}{n \text{ obligatoorsed kontekstid} + n \text{ tarvitamine mitteobligatoorsetes kontekstides}} \times 100 = \text{sihtkeelepärase korrektsuse \%}$$

Suulise ja kirjaliku õppijakeele objektikasutuse sihtkeelepärase korrektsuse protsent selle valemi järgi on esitatud tabelis 6.¹⁰

Tabel 6. Täis- ja osasihitise sihtkeelepärase korrektsuse % suulises ja kirjalikus õppijakeeles

	Kesktaseme kirjalik	Kesktaseme suuline	Kõrgtaseme kirjalik	Kõrgtaseme suuline
Partitiiv	90,2%	81,9%	87,1%	92,8%
Genitiiv	56,6%	43,1%	45,2%	81,3%
Nominatiiv	37,8%	41,5%	38,5%	70%

Tabeli 6 andmetest võib välja lugeda, et nii kesk- kui ka kõrgtaseme informantidel on partitiivobjekti kasutamine kõige paremini omandatud, mistõttu on mõistetav, et keeleõppijad valivad kahtluse korral sageli just osastavalise sihitise. Partitiivobjekti oluliselt suurem sihtkeelepärase korrektsuse protsent võrreldes täissihitisega on tõenäoliselt seotud ka sellega, et partitiiv on eesti keeles sihitise süntaktiliselt markeerimata kuju. Tabelist 6 paistab, et kõrgem keeleoskuse tase ei

⁸ Siinjuures on oluline meenutada, et eesti keele eriala üliõpilastena on uurimuse informandid eesti keele struktuurist teadlikud ning võimalus, et sihtis on formaalselt nominatiivis, kuna on lihtsalt käänamata jäetud, on minimaalne.

⁹ Mingi keelelise vormi obligatoorne kontekst on keeleline kontekst, milles emakeelekõnelejad tarvitaksid grammatiliselt korrektse lausungi produtseerimiseks just seda vormi (Tarone, Yule 1991: 126).

¹⁰ Näide T. Pica (1984) valemi rakendamiseks kesktaseme kirjaliku materjali genitiivobjektide sihtkeelepärase korrektsuse arvutamisel. Andmed: a) korrektsed kasutusjuhud: 30; b) genitiivi obligatoorsed kontekstid: 30 (s.o korrektsete juhtude arv) + 17 (juhtumid, mil oli kasutatud nominatiiv- või partitiivobjekti, kuid pidanuks olema genitiiv); c) genitiivi ebakorrektsed kasutusjuhud: 6. Arvutus: $\frac{30}{30 + 17 + 6} \times 100 = 56,6\%$.

pruugi kaasa tuua objektikasutuse suuremat korrektsust võrreldes nõrgema tase-mega: suulises materjalis on kõrgtaseme tulemused ootuspäraselt kesktasemest paremad, kuid kirjaliku keelekasutuse ainestikus on sihtkeelepärase korrektsuse protsent partitiiv- ja genitiivobjekti puhul mõnevõrra suurem hoopis kesk-, mitte kõrgtasemel informandirühmal, mis annab alust järeldusteks, et objektikäänded on teise keele kõnelejalatel kivinemisaldis grammatikavaldkond, mille kasutamisel on ilmselt peaaegu võimatu saavutada emakeelekõnelejate taset. Nii kesk- kui ka kõrgtaseme kirjalik materjal pärineb eksamikirjanditest, mille puhul oli õppijate kontroll oma keelekasutuse üle eeldatavalt ühtviisi tugev ning vältimisstrateegiate rakendamine vigade tegemise kartuses tõenäoline. Seega on välistatud võimalus, et kesk- ja kõrgtaseme kirjaliku materjali genitiivobjektide sihtkeelepärase korrektsuse protsendi erinevus (partitiivobjektide puhul on erinevus väga väike ega vääri suuremat tähelepanu) tuleneks sellest, et kõrgtaseme ainestik pärineks väiksema õppijatepoolse enesekontrolliga kirjutatud kirjanditest. Arvestades genitiivobjektide väikest absoluutarvu ning asjaolu, et kesk- ja kõrgtaseme kirjalike tekstide transiitiivlausete süntaktiline struktuur ei olnud erinev (mõlemas materjalikogus esines objekte finiiitse ja infiniitse verbivormi laiendina enam-vähem samas suhtes, mistõttu ei saa kõrgtaseme genitiivobjektide väiksemat korrektsuse protsenti seletada sellega, nagu oleks kõrgtaseme kirjalikes töödes kasutatud rohkem keeleõppijatele raskemaid infiniittarindeid, mis võivad põhjustada rohkem vigu), võib kirjeldatud erinevuse põhjuseks pidada kirjandite leksikaalse külje ning lihtsalt juhuslikkuse koosmõju.

Teise keele omandamise alastes uurimustes on tavaks rääkida mingi struktuuri omandatusest õppijate poolt, kui korrektsuse protsent on 90% ja üle selle (taus-taks on asjaolu, et ka emakeelekõnelejad ei pruugi alati saavutada sajabrotsendi-list korrektsust) (Ellis, Barkhuizen 2005: 84–85), seda arvestades võib rääkida omandatusest ainult kesktaseme kirjaliku ja kõrgtaseme suulise ainestiku partiitiivobjektide kasutamisel, täissihitise puhul ei saa omandatusest rääkida. Küll aga võib täissihitise alakasutust pidada vältimise avaldumisvormiks. Sihitise käänete omandatuse hierarhia on tabeli 6 andmete põhjal mõlemal informandirühmal **partitiiv > genitiiv > nominatiiv**.

5.6.2. Testülesannete vastused

Nagu vaba keelekasutuse ainestiku puhul, nii ka testülesannete vastustes tuleb ilmsiks informantide suurem kindlus partitiivobjekti tarvitamisel. Nii valikvastustega ülesannetes kui ka tõlkelausetes valiti partitiivobjekt sageli ka täissihitist eeldavas kontekstis, vt (17)–(20), mida võib käsitleda täissihitise vältimise näidetena. Vastajate tähistus: a) Gümni – venekeelse gümnaasiumi 11. klassi õpilased; b) EKV – eesti keele (võõrkeelena) üliõpilased; c) Eesti – emakeelekõnelejad.¹¹

(17) Tuba polnud pärast külaliste lahkumist just korras, kuid tüdruk kavatses enne ema tulekut korda teha.

	A see	B selle	C seda
Gümni:	A 0	B 5 (33,3%)	C 10 (66,7%)
EKV:	A 2 (6,9%)	B 14 (48,3%)	C 13 (44,8%)
Eesti:	A 0	B 30 (100%)	C 0

¹¹ Testülesandes oli teisigi lauseid, mida artikli piiratud maht ei võimalda esitada.

(18) Esmalt püstitatakse, seejärel asutakse neid saavutama.

	A eesmärgid	B eesmäärke
Gümn:	A 5 (33,3%)	B 10 (66,7%)
EKV:	A 15 (51,7%)	B 14 (48,3%)
Eesti:	A 28 (93,3%)	B 2 (6,7%)

(19) **Март жаловался, что жена его оставила.**

(ee: Mart kaebas, et naine oli ta maha jätnud, on ta maha jätnud, jättis ta maha.)

Gümn:	TEDA 13 (86,7%)	TA/TEMA 1/1 (13,3%)	TEINE TÖLGE 0
EKV:	TEDA8 (27,6%)	TA/TEMA 16/1 (58,6%)	TEINE TÖLGE 4 (13,8%)

(20) **Когда Март переехал за границу, он забыл своих старых друзей.**

(ee: Kui Mart välismaale kolis, unustas ta oma vanad sõbrad (ära).)

Gümn:	PART 13 (86,7%)	NOM 2 (13,3%)	TEINE TÖLGE 0
EKV:	PART 20 (69%)	NOM 7 (24,1%)	TEINE TÖLGE 2 (6,9%)

Valikvastustega ülesandes (17), (18) on õppijatel ebakindluse korral võimalik valida partitiivobjekt, tõlkeülesande sooritamisel on vältimise võimalusi rohkem. Näidetes (19) ja (20) on üliõpilastest informandirühma vastustes ära toodud ka variant *teine tõlge*, mille all on mõeldud seda, et vastajad olid valinud tõlkimisviisi, mille puhul polnud lauses üldse vaja sihitist kasutada, või oli püütud tarvitada teist rektsiooni. Näite (19) tõlkimiseks oli kasutatud variante: a) *Mart kaebas tema naise lahkumisele*; b) *Mart kaebas, et tema abikaasa lahkus temaga*; c) *Mart kaebas, et tema naine läks ära*; d) *Mart kaebas, et ta naine lahkus*, näite (20) tõlkimiseks: a) *Kui Mart kolis välismaale, läksid tal meelest vanad sõbrad*; b) *Kui Mart välismaale kolis, unustas ta oma vanadest sõpradest ära*. Ilmsed sihitise vältimise näited on (19c), (19d) ja (20a), kus oli valitud intransitiivne konstruktsioon, mille kasutamisel vastajad on end kindlalt tundnud¹². Muud tõlkevariandid sisaldavad rektsioonivigu ning võimalik sihitise vältimise soov ei ole siiski viinud sihtkeelepärase lõpptulemuseni.

Sihitise täielik vältimine tuleb esile ainult neis vähestes eelpool toodud näidetes, mis kõik kuuluvad üliõpilastest informandirühmale, gümnaasiumiõpilased ei olnud kordagi valinud sihitist vältivat tõlkevarianti, olles kas oma valikutes kindlad või pidamata ülesannet nii oluliseks, et iga hinna eest korrektsust taotleda. Seega võib isegi selle napi materjali puhul märgata seoseid keeleoskuse kõrgema taseme ja vältimisstrateegiate rakendamise vahel.¹³

6. Kokkuvõtteks

Kuigi introspektsiooni teel kogutud andmed näitavad, et täis- ja osasihitise kasutamisel on vältimisstrateegiate rakendamine tõenäoline, on vaba suulise ja kirjaliku keeleainese põhjal väga raske sihitise vältimist empiiriliselt kindlaks teha. Artikli alguses püstitatud hüpoteeside kohta saab materjali analüüsi põhjal öelda järgmist.

1. Kvantorifraaside (peasõnadega *palju* jms) kasutamist sihitisena esines nii kesk- kui ka kõrgtasemel rohkem kirjalikus kui suulises materjalis ning kuna kirjalik ainek on pärit olulise tähtsusega eksamitöödest, pole

¹² Võimalus, et teistsugune tõlge neis lausetes on tingitud verbide *оставила* või *забыл* mitteteadmisest, on eesti keelt kõrgtasemel valdavate filoloogიაüliõpilaste puhul vähetõenäoline, taustaks on ilmselt siiski ebakindlus sihitise käände kasutamisel.

¹³ Kuna testülesande sooritamise kohta hinnangulist tagasisidet ei antud, ei olnud kummalgi informandirühmal otsest vajadust võimalikke vigu vältida, need vähesed sihitist vältivad tõlkevariandid annavad tunnistust pigem mõne keele-

- kvantorifraaside sagedasem kasutamine objekti käändevahelduse vältimise eesmärgil täielikult välistatud.
2. Partitiivobjekti eeldavate eitavate konstruktsioonide tarvitamine täissihitise vältimise eesmärgil ei leidnud kinnitust.
 3. Vormihomonüümsete objektnoomenite kasutamist sihitisekäänete valiku vältimiseks ei olnud võimalik empiiriliselt kindlaks teha, samuti ei leidnud kinnitust hüpotees, mille kohaselt võib vormihomonüümsete sihitiste väiksem arv suulises kõnes tuleneda objektide suuremast vormilisest eristusvajadusest ja vormihomonüümia sagedasemast neutraliseerimisest ühilduvate täiendite abil kui kirjalikus keelekasutuses.
 4. Kesk- ja kõrgtaseme vabas suulises ja kirjalikus materjalis esinenud objektide liigitus näitab võrreldes emakeelekõnelejatega partitiivobjekti ületarvitust ja vastavalt genitiiv- ja nominatiivobjektide alatarvitust, mida võib käsitleda täissihitise vältimisena. Nimetatud objektikäänete üle- ja alatarvitust aitab seletada kõiki objektikasutusjuhte arvesse võttes arvutatud iga objektikäände korrektsuse skoor, mis on materjali kõigis allkogudes kõrgeim partitiiv- ning madalaim nominatiivobjektide puhul, seletades partitiivi kui kõige kindlamini omandatud objektikäände kasutamist sageli ka täissihitist eeldavas kontekstis. Partitiivobjekti valimine täissihitise kontekstis tuleb esile ka testülesannete vastustes.
 5. Vältimisstrateegiate rakendamine on kõige selgemini kindlakstehtav tõlkeülesannetes, milles kõrgema keeleoskustasemega õppijad valivad mõnikord ebakindluse korral sihitist täielikult vältiva tõlkevariandi.

Kirjandus

- Brown, Douglas H. 1994. *Principles of Language Learning and Teaching*. Upper Saddle River (NY): Prentice Hall.
- Canale, Michael 1983. From communicative competence to communicative language pedagogy. – Jack C. Richards, Richard W. Schmidt (Eds.). *Language and Communication*. London, New York: Longman, 2–27.
- Dörnyei, Zoltan; Scott, Mary Lee 1997. Communication strategies in a second language: Definitions and taxonomies. – *Language Learning* 3, 173–210.
- Ellis, Rod 1985. *Understanding Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, Rod; Barkhuizen, Gary 2005. *Analysing Learner Language*. Oxford: Oxford University Press.
- Erelt, Mati; Erelt, Tiiu; Ross, Kristiina 2000. *Eesti keele käsiraamat*. Teine, täiendatud trükk. Tallinn: Eesti Keele Sihtasutus.
- Færch, Claus; Kasper, Gabriele 1983. Plans and strategies in foreign language communication. – Claus Færch, Gabriele Kasper (Eds.). *Strategies in Interlanguage Communication*. London, New York: Longman, 20–60.
- Grünthal, Riho 2001. Homonymy and systemacy in Estonian inflection. – Reet Kasik (koost. ja toim.). *Keele kannul*. Pühendusteos Mati Erelti 60. sünnipäevaks. Tartu Ülikooli eesti keele õppetooli toimetised 17. Tartu, 42–61.
- Hennoste, Tiit; Vihalemm, Triin 1999. Võõrkeele suhtlusstrateegiad. – *Akadeemia* 8, 1571–1607.
- James, Carl 1980. *Contrastive Analysis*. London, New York: Longman.
- James, Carl 1998. *Errors in Language Learning and Use. Exploring Error Analysis*. London, New York: Longman.

- Kikerpill, Tiina; Pool, Raili 2003. Muutused eesti keelt teise keelena kõnelejate suhtlusstrateegiate kasutuses ühe aasta jooksul. – Pirkko Muikku-Werner, Hannu Remes (toim.). VIRSU. Viro ja suomi: kohdekielel kontrastissa. Lähivertailuja 13. Joensuu, 50–60.
- Larsen-Freeman, Diane; Long, Michael H. 1991. *An Introduction to Second Language Acquisition Research*. London, New York: Longman.
- Laufer, Batia 2000. Avoidance of idioms in a second language: The effect of L1–L2 degree of similarity. – *Studia Linguistica* 54 (2), 186–196.
- Liao, Yan; Fukuya, Yoshinori J. 2004. Avoidance of phrasal verbs: The case of chinese learners of English. – *Language Learning* 54 (2), 193–226.
- Lähdemäki, Eeva 1995. Mikä meni pieleen? Ruotsinkielisten virheet suomen ainekirjoituksissa. *Fennistica* 11. Åbo Akademi: Suomen kielen laitos.
- Odlin, Terence 1989. *Language Transfer. Cross-linguistic Influence in Language Learning*. Cambridge University Press.
- Pica, Teresa 1984. Methods of morpheme quantification: Their effect on the interpretation of second language data. – *Studies in Second Language Acquisition* 6, 69–78.
- Pool, Raili 2006a. Sihitise vormihomonüümia – keeleuurija probleem, keeleõppija tugi. – Külvi Pruuli, Annekatrin Kaivapalu (toim.). Lähivertailuja 17. Jyväskylä Studies in Humanities 53. Jyväskylä: Jyväskylän yliopisto, 85–101.
- Pool, Raili 2006b. Täis- ja osasihitise kasutamisest eesti keelt teise keelena õppijate pilgu läbi. – Helle Metslang, Margit Langemets (toim.), Maria-Maren Sepper (keeletoiim.). Eesti Rakenduslingvistika Ühingu aastaraamat. Estonian papers in applied linguistics. 2 (2005). Tallinn: Eesti Keele Sihtasutus, 185–202.
- Pool, Raili; Vaimann, Elle 2005. Vead kõrgtasemel eesti keele kõnelejate kirjalikus keelekasutuses. – Margit Langemets (koost.), Maria-Maren Sepper (toim.). Eesti Rakenduslingvistika Ühingu aastaraamat. Estonian papers in applied linguistics. 1 (2004). Tallinn: Eesti Keele Sihtasutus, 115–137.
- Raag, Raimo 1985. The direct object in Swedish Estonian. – Eesti Teadusliku Seltsi Rootsis aastaraamat. *Annales Societatis Litterarum Estonicae in Svecia*. IX. 1980–1984. Stockholm, 201–211.
- Sajavaara, Kari 1999. Kontrastiivinen kielentutkimus ja virheanalyysi. – Kari Sajavaara ja Arja Piirainen-Marsh (toim.). Kielenoppimisen kysymyksiä. Jyväskylän yliopisto. Jyväskylä: Soveltavan kielentutkimuksen keskus, 103–128.
- Schachter, Jacquelyn 1974. An error in error analysis. – *Language Learning* 24, 205–214.
- Tarone, Elaine 1983. Some thoughts on the notion of 'communication strategy'. – Claus Færch, Gabriele Kasper (Eds.). *Strategies in Interlanguage Communication*. London, New York: Longman, 61–74.
- Tarone, Elaine; Yule, Georg 1991. *Focus on the Language Learner. Second impression*. Oxford: Oxford University Press.
- Tauli, Valter 1968. Totaalobjekt eesti kirjakeeles. – *Suomalais-ugrilaisen seuran toimituksia* 145. Helsinki, 216–224.

Kaudviide

- Dušková, L. 1969. On sources of errors in foreign language learning. – *IRAL: International Review of Applied Linguistics in Language Teaching* 7 (1), 11–36.

Raili Pooli (Tartu Ülikool) peamiseks uurimisvaldkonnaks on eesti keele kui teise keele omandamine ja kasutus.
raili.pool@ut.ee

LEARNER STRATEGIES TO AVOID THE USE OF THE TOTAL AND PARTIAL OBJECTS IN ESTONIAN

Raili Pool

University of Tartu

This article presents a study of strategies used by second language learners to avoid the use of the total and partial objects in Estonian. The subjects were Russian students of Estonian as a foreign language of the University of Tartu. The data comes from spoken and written language use and test exercises.

In the Estonian language the syntactically unmarked object is partial used in the partitive case, whereas the object is total (used in the genitive or nominative case). The subjects were more confident in their use of the partitive object and, of all object cases, they would choose the partitive case the most frequently, often with a tendency to overuse it. Thus our primary attention was focused on the possibility that the use of the total object had been avoided.

The analysis of the data revealed that even though introspective data showed the probability of object avoidance, it is very difficult to verify it empirically on the basis of usage data. A comparison with native speakers' data showed that there was no reason to regard the use of noun phrases with quantifier and negative constructions, which require the partitive, and the use of object nouns with form homonymy as a strategy to avoid total objects. At the same time a statistical analysis of the object cases (the data of objects of noun phrases with a quantifier, objects of negative constructions and form homonymous objects were not included) shows that as compared to native speakers there were more partitive objects in the learners language data and correspondingly fewer genitive and nominative objects, which can be considered evidence of avoidance of total objects. On the basis of language use data, second language subjects displayed correct use of object cases in the target language the most frequently in the partitive and the least frequently in the nominative case. Thus, the hierarchy of object case acquisition is partitive > genitive > nominative.

Avoidance strategy was most evident in translation exercises (sentences for translation from Russian into Estonian), in which advanced learners, when in doubt, would rather choose total avoidance of objects.

Keywords: language learning, second language acquisition, second language speaker, native speaker, avoidance, object cases, Estonian

LADINA KEELEST PÄRIT TERMINITE POLÜSEEMIA JA SÜNONÜÜMIA EESTI ÕIGUSKEELES

Merike Ristikivi

Ülevaade. Artikkel käsitleb ladina keelest pärit õigusterminite polüseemia ja sünonüümia probleeme. Alusmaterjalina on kasutatud “Ladina-eesti õigussõnastiku” (2005) ainek. Ehkki Euroopa õiguspõhine baseerub ühisel Rooma õiguse pärandil ning sellest võrsunud terminid on õigussüsteemides juurdunud, esinevad ladina terminid juriidilistes materjalides sageli erinevas tähenduses, kusjuures erinevused on ka õigussüsteemide ja õigusperekondade terminite vahel. Samuti võib leida sama mõiste tähistajana mitu erinevat terminit. Artikkel analüüsib, millistel juhtudel on polüseemia ja sünonüümia ladina õigusterminoloogias taunitavad, missugustel aga mõistlikud ja otstarbekad.

Võtmesõnad: õiguskeel, terminoloogia, terminisõnastikud, leksikaalne semantika, keelekontaktid, polüseemia, sünonüümia, ladina keel, eesti keel

Sissejuhatus

Õigusteadus ja juriidiline praktika on valdkonnad, kus keelelistel väljendusvahenditel on erilisel suur tähtsus. Nad mõlemad toimivad vahetult keele kaudu, sõna või väljend omandab seal juriidilise jõu. Seepärast on õiguskeele küsimused päevakorral ja nende olulisust tajuvad nii keelega kui juriidikaga tegelejad. Parimad praktilised väljundid sünnivad kahe valdkonna koostöös.

Selles küsimusteringis on üheks keskmeks ladina keelest pärit sõnad ja väljendid, mida õiguskeeles kasutatakse tsitaatsõnadena või eesti keeles mugandunud kujul. Ent kuidas toimib õiguskeele raames kahe keele vastastikune suhe? Siin tuleb vaatluse alla keelekõnastuste leksikaalne ja semantiline külg.¹ Analüüsitava materjal pärineb 2005. aastal ilmunud “Ladina-eesti õigussõnastiku” (Adomeit jt 2005) ainek. Selle sõnastiku koostamisel kasutatud erinevate juriidiliste materjalide uurimisel ilmnes, et ladina termineid kasutatakse sageli eri tähendustes, kusjuures

¹ See ei tähenda, et ladina õigusterminite ortograafia ja morfoloogia probleemid puuduksid, vt Ristikivi 2006.

erinevused on ka õigussüsteemide ja õigusperekondade terminite vahel. Samuti võib leida sama mõiste tähistajana mitu eri terminit. Eesti õiguskeele kujundamisel ja korrastamisel on vajalik selle nähtuse tausta tundmine ja eesti keele seisukohalt vajaliku analüüsi ning järelduste tegemine, mis tagaks ladina keelevahendite asjakohase kasutamise. Seda senini tehtud pole ja käesolevas artiklis toodud analüüs selgitab mõningaid külgi selleks küsimuses.

Ladina keele ja sealt pärit väljendite osa õiguskeeles vajab süvendatud tähelepanu, kuna siin ilmnevad probleemid väljendusviisi täpsuses ja korrektsuses. Samas on ladina keele osa õiguskeeles säilitanud oma tähtsuse läbi sajandite – nii teatud õigusteaduse aladel kui praktilises väljendusviisis. Nende küsimustega tegelemine annab tulemusi interdistsiplinaarse lähenemise raames, kus eri valdkonnad püüavad pakkuda lahendusi, arvestades üksteise eripära. Käesolev artikkel püüab esitada probleeme ja leida neile vastuseid, vaadeldes latinisti pilguga õiguskeeles esinevaid keeleküsimusi.

Leksika seisukohalt vaatlen ladina keelest pärit õigusterminite leksikaalseid, grammatilisi, ortograafilisi ning lühivariante.

Semantika osas tõusevad esile järgmised probleemid: ladina terminite tähendus üld- ja õiguskeeles, tähenduserinevused õigussüsteemide ja -perekondade terminoloogias ning mitmetähenduslikud terminid, mis tähistavad üksteisele lähedal seisvaid mõisteid.

1. Ladina õigusterminite polüseemia

Oskuskeele üldine püüe võimalikult täpse ja ühemõttelise väljenduse poole tingib, et ühele terminile vastaks ainult üks mõiste ning vastupidi, ühele mõistele vastaks ainult üks termin. Seega on ideaalnõue, et oskuskeel oleks vaba nii polüseemiast kui ka sünonüümiast (Erelt 1982: 37; Kull 2000c: 552 jj). Siiski pole absoluutne ühetähenduslikkus kooskõlas keele funktsioneerimise põhialustega. Teaduses ja ka praktilisel oskusalal võib väljenduse mitmekujulisus osutada mitte ainult häiretuks, vaid ka kasulikuks (Saari 1980a: 662). Nii tuleb terminite mõistelise täpsustamise kui ka nende mitmesugustes variantides ja nende omadustes selguse saamiseks jälgida neid tegelikus tarvituses, kontekstis (Kull 2000a: 117). Seejuures tuleb silmas pidada, et nii laiendava kui ka kitsendava tõlgendamise korral vastab termini sisu vastavas kohas ja ajas valitsevatele vaadetele või teistele õigusvälistele tingimustele (Oksaar 1999: 203–204).

Mitmetähenduslikkus võib oskuskeeles tekkida järgmistel põhjustel:

- 1) oskuskeel on loomulik-tehislik allkeel. See tähendab muu hulgas sedagi, et ühelt poolt on temas võimalik taotleda ja ka saavutada terminite ühetähenduslikkust ning semantilist rangust, teiselt poolt toimivad loomuliku keele seaduspärasused. Terminid kasutus võib kitseneda, laieneda, nihkuda teiste mõistete peale, semantilised piirid aga seetõttu tuhmuvad ning areneb välja mitmetähenduslikkus (Erelt 1982: 37);
- 2) tihti tuleb ühel ja samal terminil tähistada küll ühe ja sama nähtuse juurde kuuluvaid, aga juba eri mõisteid²: vanad mõisted jäävad kõrvale, asemele tulevad uued. Ent mõnda aega eksisteerivad vana ja uus kõrvuti, käivad küll samade asjade kohta, aga on täiesti erinevate vaadete või hüpoteeside

² Terminid mitmetähenduslikkuse tekkimist ajaloolises arengus nimetatakse ka diakrooniliseks polüseemiaks (Cornu 1990: 98).

tulemused, täis erinevat sisu (Erelt 1982: 38), õigusteaduslikus kontekstis eriti näiteks ladina terminid *ius civile* ja *ius commune* (vt allpool 1.1.3). Selline polüseemia kajastab teaduse arenguprotsessi, sama teadusharu raames erinevate teooriate lahknevusi, põhimõteteliste seisukohtade muutusi.

Polüseemianähtused võib jagada kahte rühma selle järgi, kas mitmetähenduslik termin tähistab üksteisest kaugel seisvaid või üksteisele lähedal seisvaid mõisteid.³

1.1. Üksteisest kaugel seisvate mõistete tähistamine

1.1.1. Üld- ja õiguskeelne tähendus

Kui mitmetähenduslik termin tähistab üksteisest kaugel seisvaid mõisteid, ei ole need kas üldse seotud või on kauges suguluses. Ladina keeles võib tuua näiteks järgmised terminid: *actio* üldkeelne tähendus 'tegevus, toiming, tegu', õiguskeeles (edaspidi lühendatult *jur*) 'hagi, nõue, kaebus'; *animus* 'vaim, hing, meel', jur 'tahe, tahtlus, kavatsus'; *caput* 'pea; pealinn; peatükk', jur 'õigusvõime'; *causa* 'põhjus', jur 'protsess, kohtuasi'; *dolus* 'kavalus, kaval inimene', jur 'tahtlus, pettus, tüssamine'; *manus* 'käsi, jõuk', jur '(perekonnapea) võim'; *societas* 'ühiskond; ühendus', jur 'seltsinguleping'..

Kui sama sõna kasutatakse nii erinevates süsteemides nagu üldkeel ja õiguskeel, siis suuremaid segiminekvõimalusi praktiliselt pole. Ka erialakeeles võib üldkeele tähendus esineda tegevuse või nähtuste üldise iseloomu või omaduse kirjeldamisel. Tõlkimisel osutab kontekst, millises tähenduses sõna väljendis või lauses esineb, näiteks *actio culposa* 'süüline tegevus', jur *actio in rem* 'asjaõiguslik hagi', *actiones transeunt ad heredes* 'nõuded lähevad üle pärijatele'; *aequo animo* 'rahuliku meelega', jur *animus possidendi* 'valdamise tahe'; *causa mortis* 'surma põhjus', jur *causa civilis* 'tsiviilprotsess'; *manu propria* 'oma käega, omakäeliselt', jur *matrimonium sine manu* 'abielu ilma (mehe) võimuta'..

1.1.2. Õiguskeelesisene mitmetähenduslikkus ladina keele traditsioonis

Sellesse rühma liigitatakse terminid, millel on oskuskeele enda sees erialati erinevad tähendused. Ladina keeles on *culpa* ja *delictum* tuntumad terminid, millel on eri õigusharudes kujunenud erinevad tähendused. Näiteks tähistab termin *culpa* tsiviilõiguses hooletust, karistusõiguses aga on tähenduseks 'ettevaatamatus, süü'. Tsiviilõiguses on *delictum* 'õigusvastane kahjutekitamine', karistusõiguses tähistab see termin süütegu, kuritegu.

Ehkki Euroopa õiguspõhje baseerub ühisel Rooma õiguse pärandil ning sellest võrsunud terminid on juurdunud kõigis Euroopa õigussüsteemides, on rahvusvahelise kogemusega juristid tähele pannud, et välismaa kolleegide õigusallast ladina keelt ei ole alati kerge mõista (Berteloot 1999: 2). Samad terminid ei ole kõikjal kasutusel, samuti võib olla erinevusi terminite tähenduses. Suurimad erinevused ladina väljendite tähenduste vahel on angloameerika⁴ ja Mandri-Euroopa õigussüsteemide terminoloogias (Cornu 1990: 104). Nendes suurtes õiguskultuurides erinevad juri-

³ Polüseemia liigitamisel on lähtutud T. Erelti (1982: 38–40) ja G. Cornu (1990: 95–102) jaotusest.

⁴ Läänelik õigustraditsioon jaguneb Mandri-Euroopa (kontinentaalseks, romaani-germaani) ning angloameerika (*common law*) õigussüsteemiks. Eri õigussüsteemides erineb ka keeleline väljendusviis, sealhulgas sõnavara. Õigussüsteeme iseloomustavad erinevad õigusfilosoofilised käsitlused riigist ja õigusest, erinevad on arusaamad õiguse reguleerivast rollist ühiskonnas, õiguse allikatest ning õigusnormist, õiguse loomisest ning realiseerimisest. Erinevustele vaatamata on läänelikku õigustraditsiooni kuuluvate õigussüsteemide teoreetiliste koolkondade põhiliseks probleemiks üksikisiku ja ühiskonna, seega erahuvi ja avaliku huvi vahekorid (Anepaio jt 2005: 167–168). Mandri-Euroopa õigussüsteemi

dilise ladina keele traditsioonid üksteisest märgatavalt. Etümoloogiline taust tingib sel puhul mingi ühisosa säilivuse semantilisel väljal. Lahknevate osade maht võib eri mõistete puhul erineda, ulatudes suhteliselt lähedasest märgatava erisuseni.

Kontekstiväliselt võib vääriti mõistmist põhjustada näiteks termin *ius commune*, mis tähistab Mandri-Euroopa ühisõigust ning on välja kujunenud Rooma ja kanoonilise õiguse allikate teadusliku käsitluse tulemusena. Angloameerika sõnastike vaste sellele on *common law*, mis märgib Inglismaa vana, kohtupraktikaga kujundatud õigust.⁵

Termin *exitus* tähistab Mandri-Euroopa juriidikas surma, *common law* sõnas- tikud aga annavad vasteks 'lapsed, järglased' või ka 'renditulud; ekspordimaks'.⁶ Ühisosa tähistab siin millestki eraldumist, lahknev osa eraldumise variante. Termin *breve* tähendus saksa õigussõnastikes on 'kiri; lühike paavsti läkitus', *common law* ajaloos tähistab see mõiste aga nõuet kohtusse ilmumiseks, kohtukutset.⁷ Mõlema termini aluseks on läkituse üks omadus – lühidus.

Põhiosa Mandri-Euroopa õigussüsteemis kasutatavatest ladina väljenditest on käibel sama tähendusega. Romaani-germaani õigusperekonna⁸ väljendite sarnasus on paljuski tingitud Mandri-Euroopa maade ühisest *ius commune* pärandist.⁹ 18. sajandi lõpuni oli Mandri-Euroopa õigusteadus oma iseloomult rahvasteülene. Eri maade õpetlased uurisid ja edendasid õigust ladina keeles rahvuspiiridest sõltumata. Siiski võib ka siin leida väljendeid, mille tähendused ei lange osaliselt või mõnel juhul isegi täielikult kokku.¹⁰ Näiteks ladina sõna *conventio* tähendus germaani õigusperekonnas on 'leping, kokkulepe'; romaani õigusperekonnas on

kuuluvad Mandri-Euroopa ja Ladina-Ameerika maad, samuti Québeci provints Kanadas ja Louisiana osariik Ameerika Ühendriikides. Kontinentaalse õigussüsteemi elemente võib leida veel Šoti õiguses, Lõuna-Aafrika Vabariigis, Mauriitiusel, Filipiinidel ja mujalgi. Angloameerika õigussüsteemi moodustavad Inglismaa (ka Wales), Ameerika Ühendriigid, Kanada, Austraalia ja Uus-Meremaa (Narits 1995: 46–47).

⁵ Vrd *ius commune*: "Gemeines Recht (hervorgegangen aus der wissenschaftlichen Bearbeitung römischer und kanonischer Rechtsquellen und deren Anpassung an die deutschen Verhältnisse)" (Lieberwirth 1987); "Gemeines Recht. Recht, welches allgemein gilt. In der frühen Neuzeit wurde darunter das allgemein geltende römisch-kanonische Recht verstanden" (Benke, Meissel 1997); "The common law" (Jowitt's Dictionary); "In the civil law, common right; the common and natural rule of right, as opposed to *ius singulare*. In English law, the common law, answering to the Saxon *folcright*" (Black 1998).

⁶ Vrd *exitus*: "Tod" (Lieberwirth 1987); "Der Tod (wörtlich: Ausgang; Ende)" (Benke, Meissel 1997); "Issue or offspring; the yearly rents and profits of lands and tenements" (Jowitt's Dictionary); "Children; offspring. The rents, issues and profits of land and tenements. An export duty" (Black 1998).

⁷ Vrd *breve*: "Brief; Beweisurkunde; päpstliches Reskript in Briefform" (Lieberwirth 1987); "Im KirchenR: Kurzes päpstliches Schreiben" (Benke, Meissel 1997); "A writ by which a person is summoned or attached to answer an action, complaint, etc., or whereby anything is commanded to be done in the courts, in order to do justice, etc." (Jowitt's Dictionary); "A writ. An original writ or precept of the king issuing out of his courts. A writ by which a person was summoned or attached to answer an action, complaint, etc., or whereby anything was commanded to be done in the courts, in order to justice, etc." (Black 1998).

⁸ Kontinentaalses õigussüsteemis eristatakse õiguse perekondi, mis on põhimõtteliselt rajatud ühtsetele aluse- tele – õigusnormidele –, kuid kus võib kohata ka erinevusi. Eristada võib romaani ja germaani õiguse perekondi. Nende õigusrühmade terminid suuremas osas kattuvad. Siiski võib ka leida väljendeid, mille tähendused ei lange osaliselt kokku või mõnel juhul erinevad isegi täielikult. Romaani perekonda kuuluvad Prantsusmaa, Belgia, Luksemburg, Monaco, Itaalia, Hispaania, Holland. osaliselt Ladina-Ameerika riigid, Québec Kanadas, endised Prantsusmaa kolooniad Aafrikas, Maroko, Tuneesia, Alžeeria, Liibanon, Süüria, endine Hollandi koloonia Indoneesia. Germaani perekonda kuu- luvad Saksamaa Liitvabariik, Austria, Liechtenstein, Šveits, Türgi, Kreeka. Teatud reservatsioonidega võib siia paigutada ka Jaapani, Lõuna-Korea ja Tai. Kuni 1949. aastani kuulus siia perekonda Hiina RV. Germaani õiguse perekonnast ei asu kaugel viis Põhjamaad: Soome, Rootsi, Norra, Taani ja Island. Eesti koht kontinentaalses õigussüsteemis on lähemal germaani perekonnale kui mõnele teisele struktuurile (Narits 1995: 46–47).

⁹ Mandri-Euroopa õiguskultuuri on kujundanud eelkõige Rooma õiguse retseptisioon. 13.–16. sajandil levis Rooma õigus Ülem-Itaaliast üle kogu Euroopa. Selline kirjalikult fikseeritud ja dogmaatiliselt kommenteeritud õiguskord astus Euroopas sugukonnaõiguse kõrvale ja hakkas seda ka asendama. 19. sajandil süstematiseeriti retsepteeritud õiguskord astus põhjalikult ja nii kujunes Rooma õigusest üks tänapäevase Mandri-Euroopa õiguskultuuri põhialus. (Narits 1997: 5.) Kontinentaal-Euroopa õiguskultuur toetub peamiselt kahele arusaamisele õigusest: loomuõigusele ja positivismile. Kontinentaal-Euroopa õiguse sajanditepikkuse arengu tulemuseks on Euroopa üldise õiguse perekonna (*ius commune*) väljakujunemine. Sellesse õigusperekonda kuuluvates õigussüsteemides on sotsiaalsete normide seas esiplaanil õigusnormid, mis vastavad õigluse ja moraali nõuetele. Tavad kujutavad siin endast täiendust seadustele, need on abiks seaduse tõlgendamisel ja õiglase otsuste langetamisel. Põhjalikumalt *ius commune* pärandi kohta vt Knütel 1994 ning Berman, Reid 1995: 3–34.

¹⁰ Germaani ja romaani õigusrühma tõlgete võrdlemisel on aluseks järgmised sõnastikud: Cornu 1992; Cabanellas, Alcalá-Zamora y Castillo 1992; Benke, Meissel 1997; Lieberwirth 1987; Liebs 1998; Anners, Önerfors 1972.

selle termini tähendus kitsenenud: 'leping; hagi, hagi koostamine'. Termin *pignus* tähistab germaani õigusperekonnas panti, hüpoteeki, kuid romaani õigusperekonnas kasutatakse terminit tähenduses 'tõend'. Termin *damnum* on germaani õigusperekonnas 'kahju, trahv', romaani õigusperekonnas on termini tähendus laienenud ning see tähistab konfiskeerimist. Sõna *consilium* puhul erinevad tähendused ka ainult germaani õigusperekonna sõnastikes: ühes on tõlkeks märgitud 'nõukogu; nõupidamine (arstide nõupidamine haigusjuhtumi teemal)', teises aga lisatakse olulisi juriidilisi vaseid: 'kohtuniku otsus; kohtukolleegium'.¹¹

Ladina terminite mitmetähenduslikkuse on põhjustanud ka nende väljakujunemine väga pika aja, rohkem kui kahe aastatuhande vältel. Aja jooksul on muutunud mõistete sisu. Näiteks väljendil *ius civile* ('tsiviilõigus') on arvukalt tähendusi. Rooma õiguses tähistas see ühelt poolt Rooma klassikalist õigust (vastandiks magistraatide praktikaga kujundatud *ius honorarium*), teiselt poolt Rooma kodanike suhteid reguleerivat õigust (vastandiks *ius gentium* 'rahvaste õigus') (Birks, McLeod 1987: 8–9).

Bütsantsis, keskaja Lääne-Euroopas ja veel uusajalgi märkis *ius civile* peamiselt Rooma õigust, laiema ka positiivset õigust, mis vastandus jumalikule õigusele (*ius divinum*) ning loomuõigusele (*ius naturae*). Keskajal hakkas välja kujunema ka *ius civile* kasutamine praeguses tähenduses. Nimelt keskendus keskajal õigusteadus Rooma õiguse tähtsaima, keiser Justinianuse korraldusel 6. sajandil koostatud koodeksi *Corpus iuris civilis*'e ('Tsiviilõiguse kogumik') nende osade uurimisele, mis puudutavad üksikisikute õigussuhteid. See mõjutas aja jooksul ka keelekasutust. Nii omandas termin 'tsiviilõigus' Euroopas oma praeguse tähenduse – üksikisikute vahelisi suhteid reguleeriv normide kogum (Roland, Boyer 1998: 240–241).

1.2. Üksteisele lähedal seisvate mõistete tähistamine

Sellesse rühma kuuluvad mitmetähenduslikud terminid, mis tähistavad üksteisele lähedal seisvaid mõisteid. Ladina keeles valmistavad suuremat mõistmisraskust ning vajavad enam täpsust tõlgendamisel sellised terminid, mis tähistavad ühelt poolt alist ehk objekti, teiselt poolt aga tegevust, protsessi toimumist ja selle tulemust. Tuntumad ja sageli kasutatavad on näiteks järgmised sõnad: *hereditas* ('pärand' ja 'pärimine'), *possessio* ('valdus' ja 'valdamine'), *emptio* ('ost' ja 'ostmine'), *dominium* ('omand'¹² ja 'omamine').

Terminite mitmetähenduslikkus ei tekita alati olukorda, kus neid ei saaks õigesti mõista. Harilikult rõhutatakse, et termin ei vaja konteksti, on mõistetav kontekstitagi. Seda asjaolu ei tohi absolutiseerida. Täielikult avaneb termini sisu alles kaastekstis (Erelt 1982: 42).

Näiteks sõna *hereditas* väljendab objekti ning on tähenduses 'pärand' järgmistes väljendites: *hereditas iacens* 'lamav pärand'¹³, *acquisitio hereditatis* 'pärandi vastuvõtmine', *hereditas, alia corporalis, alia incorporalis; corporalis est, quae tangi potest et videri; incorporalis, quae tangi non potest nec videri* 'pärand on kas kehaline või kehatu; kehaline on see, mida saab katsuda ja näha, kehatu katsuda ega näha ei saa'. Seevastu tegevust või protsessi tähenduses 'pärimine' väljendab fraas: *hereditas est successio in universum, quod defunctus habuit* 'pärimine on kõigi õiguste ülekandumine, mis oli surnule kuulunud'.

¹¹ Vrd *collegium*: "Rat, Beratung (von Ärzten über einen Krankheitsfall)" (Bruss 1999); "Rat; Richterspruch; Richterkollegium" (Lieberwirth 1987).

¹² Terminit *dominium* on varemalt eesti keelde tõlgitud ka sõnaga *isandus* (analoogia *dominus* 'isand') (Talvik 1936: 108–112).

¹³ Rooma õiguses pärandi nimetus ajavahemikul pärandi avanemisest kuni selle vastuvõtmiseni (Adomeit jt 2005).

Lähedalseisvaid mõisteid tähistavate terminite mitmetähenduslikkust peetakse oskuskeelt häirivaks nähtuseks ning soovitatakse sellest võimaluse korral hoiduda (Erelt 1982: 40–41; Möttus 1997: 37). Näiteks võib siia tuua eestikeelsete terminite *omand*, *omandus* ja *omandiõigus* tähendusliku vahekorra ja nende kasutuse oskuskeeles, mis leidis laialdast kajastamist pärast asjaõigusseaduse (1993) vastuvõtmist. Asjaõigusseaduses on kasutatud terminit *omand* omandiõiguse tähenduses. Üldkeeles ja laiemale õiguskeeles käsitatakse omandit üksnes omandiõiguse esemena. Asjaõigusseaduse kohaselt võib asi olla kellegi omandis, üld- ja õiguskeele kohaselt kellegi omanduses.¹⁴

2. Ladina õigusterminite sünonüümia

Üks kesksemaid terminikorrastusprobleeme on terminoloogia variatiivsuse küsimus, oskussõnade variantide, sünonüümide probleem lähemalt vt Saari 1980a, 1980b, 1981a, 1981b).

Varem valdava seisukoha asemel, et igale mõistele vastaku üksainus kindel terminikuju, on viimastel aastakümnetel jõutud arusaamisele, et variante terminoloogias vältida ei saa, kuid need on vaja allutada teadlikule korrastatusele.¹⁵ Variantide esinemist taunitakse põhjusel, et erinevad terminivariandid koormavad mälu. Tähtsamaks tuleb pidada siiski sisulisi põhjuseid, s.t sünonüümia annab võimaluse kõrvaltähenduste tekkele, hägustab arusaamist, n-ö “terminimüra” (Erelt 1982: 45). Sünonüümide kasutamisel on seega oht, et eri sõnakujudesse hakatakse panema eri sisu, nii et täissünonüümid muutuvad ajapikku osalisteks sünonüümideks ja omandavad kõik mitmetähendusliku termini pahed. Seetõttu tuleb hoiduda sünonüümide tähenduse hargnemisest (Saari 1980a: 660; Saari 1980b: 738).

Teaduskirjanduses, terminisõnastikes ning mujal on sünonüümid siiski paratamatud (Saari 1980a: 660), seetõttu peetakse ka korrastatud terminoloogias vajalikuks lubatavat sünonüümiat, terminivariantsust (Kull 2000b: 3).¹⁶ Sünonüümide lubatavus terminoloogias tagab väljenduskvaliteedi, eelkõige stiili ladususe ja paraja vahelduslikkuse.

Õiguskeele uurimused näitavad, et juriidilises terminoloogias on sünonüümid laialt levinud (Mattila 2002a: 177–178). Teistes keeltes peale romaani keelte võib sama mõistet väljendada tihti nii ladinatüveline võõrsõna kui ka omasõna.¹⁷

Ladina õigusterminite puhul on märkimisväärne, et variandid ei erine õiguskultuuri ega keelepiiride tasandil. Ladina terminite variantsus esineb võrdselt nii sama keele piires kui erinevate keelte raames, kui käsitletakse sama teemat.

Sünonüümsus on enamasti kahe sõna vahel, kuid see võib esineda ka sõna ja ühendi vahel, millest esimene on toimingut või selle vahendit näitav substantiiv ja teine teonimi (gerundium). Viimane täpsustab lühemat ja kompaktsemat varianti, näiteks *ususfructus* ja *ius utendi-fruendi* (‘kasutusvaldus’), *aquaeductus* ja *ius*

¹⁴ Selle temaatika kohta vt lähemalt Ekspertarvamus 1995; Kull 2000a: 124–127; Mereste 1995: 35–38; Mereste 2000: 338–356; Kask, Kull 1996: 19–27; Pärna 1995: 27–28.

¹⁵ Vt nt Kull 2000a: 101; Kerge 2004: 33–34; Saari 1980a: 657–663; Erelt 1982: 44–51. Siiski on põhimõtte “ühele terminile lähtekeeles peab vastama üks konkreetne termin tulemkeeles” aluseks veel näiteks tänases Läti terminoloogiatöös (Tavast 2004: 11).

¹⁶ Lubatud ja soovitatava sünonüümia peamised liigid on: omasõna-võõrsõna paarid, põhitervim koos argivariantidega, keelepragmaatikast ja kontekstist tingitud muudatused ning üleminekud motiveeritud väljendil arbitraarsele. Lubatud variantsuse kohta terminoloogias vt Kull, Saari 1975.

¹⁷ Inglise õiguskeeles on näiteks levinud sõnakordus, milles sõnapaar üks sõna on moodustatud ladina-romaani tüvest ning teine germaani tüvest: *acknowledge and confess*, *will and testament*. Selline binaarsus kujunes välja perioodil, kui Inglismaa õiguskeelteks olid ladina ja prantsuse keel ning anglosaksi terminite kõrval võeti kasutusele ladina-prantsuse variandid. Kordused aitasid kergemini mõista terminite tähendust; tänapäeval kinnitab samatähenduslike sõnade korrdamine, et kogu soovitud semantiline väli on kaetud. Sõnakordus on nähtus, mis levib eriti teadusharu kujunemisjärgus (Tiersma 1999: 100–101; Mellinkoff 1963: 74–75). Oma- ja võõrsõna paaride kohta vt ka alapeatükk 2.1.2.

aquam ducendi ('veejuhtimise õigus'), *iter* ja *ius eundi* ('läbimineku ehk käigutee õigus').

Samuti võivad sünonüümid olla ühendid või mitmesõnalised väljendid omavahel. Siinjuures võivad sünonüümideks olla nimisõnaühendi põhisõnad, näiteks *servitutes praediorum* ja *iura praediorum* ('prediaalservituudid'), *servitutes personarum* ja *iura personarum* ('personaalservituudid'); täiendsõnad, näiteks *causa capitis* ja *causa criminalis* ('kriminaalprotsess'), *culpa lata* ja *culpa magna* ('raske hooletus'); aga ka väljendi tegusõnad, näiteks *maius includit minus* ja *maius continet minus* ('suurem hõlmab vähemat').

Sünonüüme võib jagada leksikaalseteks, grammatilisteks, ortograafilisteks ja lühivariantideks (Saari 1980a: 654–666 ja Erelt 1982: 46–47 eeskujul).

2.1. Leksikaalsed variandid

2.1.1. Eri lekseemid

Sellesse rühma kuuluvad sünonüümid juhul, kui iga ladinakeelne terminivariant kujutab endast eri lekseemi. Sellised sünonüümid ei ole ladina õigustekstides sagedased. Peamise hulga moodustavad ladina terminoloogias grammatilised, ortograafilised või lühivariandid. Siiski on mitmetel leksikaalsetel terminivariantidel nende tähenduse tõttu õigusmaterjalides oluline koht, eriti näiteks terminid *dominium* ja *proprietas* ('omand'). Eri lekseemidena võib leida veel järgmisi termineid: *delictum* ja *maleficium* ('õigusrikkumine'), *mos* ja *consuetudo* ('tava'), *ascendentes* ja *linea superior* ('ülenevas joones sugulus, ülenejad sugulased'), *descendentes* ja *linea inferior* ('alanevas joones sugulus, alanejad sugulased').

2.1.2. Oma- ja võõrsõna paarid

Sõnade päritolu järgi jagades on leksikaalsete variantide kõige levinum alaliik oma- ja võõrtermini paralleelne esinemine (Erelt 1982: 47). Oma- ja võõrsõna paarid on Rein Kulli ja Henn Saari käsitluse järgi oskuskeele lubatud ja soovitatava sünonüümia liik. Sünonüümide täielik vältimine saab toimuda ainult väljendus kvaliteedi arvel. Samas on tähtis, et säiliks oskuskeelse väljenduse ühtsus ja ühemõttelisus: "Üksühised sõnade ja ühendite vastavused eri keelte vahel on isegi terminoloogias võimalikud üksnes piiratud ulatuses ja nimelt seda vähemas, mida enam keeled teineteisest erinevad." (Kull, Saari 1975: 247–248).

Tuleb pidada loomulikuks, et ühe eriala oskussõnavaras on nii oma- kui ka võõrtermineid ja ka mõlemaid paralleelselt (Erelt 1982: 73). Nii esinevad eesti õiguskeeles omaterminid, ladina keelest tuletatud võõrterminid, oma- ja võõrsõna paarid ning liitsõnad ja sõnaühendid, milles oma- ja võõr morfeemid on sulandunud üheks keelendiks.

- **Eesti omasõnad õiguskeele terminitena**

Suurem osa eesti õiguskeele põhisõnavarast on loodud eesti omasõnavara põhjal, vähemal määral on kasutatud otselaene suurematest indoeuroopa keeltest (Soosaar 2005: 111, 117). Nii ei ole paljude eesti juriidiliste põhiterminite juures näha ladina tüvesid, pigem on need kujunenud omasõnadest või on omakorda nende põhjal loodud tuletised, näiteks *õigus* (vrd ladina *ius*), *seadus* (*lex*), *tava* (*mos*), *karistus* (*poena*), *asi* (*res*), *süü* (*culpa*), *leping* (*contractus*), *kohtunik* (*iudex*) jpt.

Mida üldisema tähendusega on termin ja mida laiemat hulka inimesi see praktikas puudutab, seda õigustatum on omatermini kasutamine. Ja vastupidi, mida spetsiifilisemaks läheb termini sisu, seda rohkem tuleb kasutusele võõrtermineid.

Omakeelsete oskussõnade talitlusviisid üldjuhul erinevad võõrterminite omast. Kuna omaterminid on juurdunud just omakeelsetena, siis on need enamasti selgemad, läbipaistvamad ning seega ka üldarusaadavamad ja täpsusvõimelisemad. Omasõnalised keelendid kõnelevad nii spetsialistidele kui ka mittespetsialistidele, seega avaralt ühe keele kasutajate piires. Avalikkusega suhtlemisel (nt seaduskeeles jm) soovitataksegi eelistada rohkem omakeelseid oskussõnu, jättes võõrsõnalisema kasutuse spetsialistide omavahelisse suhtlusse (Kull 2000c: 551; Erelt 1982: 73; Saari 1981a: 209).

- **Ladina päritolu sõnad õiguskeele terminitena**

Erialaterminoloogia üheks loomise ja rikastamise allikaks on sõnade laenamine teistest keeltest; teadusalati võib laenamise intensiivsus perioodide lõikes erineda (Lill 1986: 77). Eesti õiguskeeles on kasutusel palju ladina sõnadest tulenevaid termineid, mis pärinevad Rooma õigusest. Rooma õigusest pärit termineid on eriti rohkesti tsiviilõiguses ja peamiselt selle võlaõiguse osas (Ilus 2000: 15). Nii tugevalt juurdunud võõrterminite puhul ei pea ka keelekorraldajad vajalikuks otsida ja juurutada omakeelseid vasteid. Nende kasutuse poolt räägib samuti rahvusvaheline tunnus ja levik (Vettik 1996: 5; Erelt 1982: 74). Enamasti tähistavad sellised terminid mõisteid, mis vajaksid oma keeles pikki, sageli mitmest lausest koosnevaid seletusi. Eesti õiguskirjandusest võib leida näiteks järgmisi võõrtermineid: *testament* (vrd ladina sõna *testamentum*), *regress* (*regressus*), *advokaat* (*advocatus*), *eviktsioon* (*evictio*), *subjekt* (*subiectum*), *kodifitseerimine* (*codificatio*), *bilateraalne* (*bilateralis*), *kautelaarjurisprudent* (*cautularis + iurisprudentia*), *tsiviilprotsess* (*civilis + processus*).

Mitmed loetletud võõrterminitest on tihti kasutusel ka eesti üldkeeles ning mõjuvad tekstis hariliku ja sisuselgena. Enamik neist tähistab aga väga spetsiifilisi mõisteid, olles arusaadavad vaid õigusteadlaste ringile. Kui omakeelsed terminid on eelistatavad omakeelse avalikkusega suhtlemisel, siis ladinatüvelised võõrterminid, mis on levinud nii angloameerika kui ka Mandri-Euroopa õigussüsteemide oskuskeeletes, soodustavad õigusteadlaste rahvusvahelist suhtlemist. Seega võib ka öelda, et üldiselt hõlbustavad stiili poolest neutraalsed võõrsõnalised keelendid suhtlemist teiste keelte kasutajatega.

- **Võõr- ja omaterminite paralleelne kasutamine**

Kõige sagedamini võib eesti õigustekstidest leida võõr- ja omaterminite paare, mida kasutatakse erialakeeles koos, näiteks *õigusrikkumine* ja *delikt* (ladina *delictum*), *seaduslik* ja *legaalne* (*legalis*), *võlasuhe* ja *obligatsioon* (*obligatio*), *seadustik* ja *koodeks* (*codex*), *õiguslik* ja *juriidiline* (*iuridicus*).

Oma- ja võõrsõna paaride võrdlemisel võib näha, et vahel võib võõrsõna olla tavapärasem, vahel vastupidi. Sellistel variantidel võib olla erisugune väljendusväärtus ja stiilivärv. Omakeelend on sageli, kuid mitte alati, täpsem või diferentseerivam. Seda juhul, kui omatermin on spetsiaalselt eriala jaoks loodud, üldkeelest võetud omaterminite puhul seda öelda ei saa. Oluline põhjus on see, et omaterminid ei sõltu võõrsõnade kirjust ajaloost, homonüümiast, polüseemiast ja korraldamatusest. Samas on raske mõõta omasõna jutustavust, sest palju oleneb iga konkreetse inimese emakeele- ja võõrkeelteoskusest, aineala ja võõrsõnaelementide tundmisest (Saari 1981b: 285).

Võõrsõna-omasõna variantsuse korral tuleb arvestada suhtlussituatsiooni ja teksti suunitlust. Üldiselt on suurem võõrsõnakasutus omane kitsamalt erialasele käsitlusele, eriti rahvusvahelises infovahetuses ja keeltevahelises suhtluses, omakeelendite eelistamine on omal kohal aga omakeelse ja üldisema esituse korral. Võõrsõnadele on omane ekso- ehk väljapoole-, omasõnadele aga endo- ehk sissepoolesuunitlus (Kull 2000a: 113; vt ka Kull 2000c: 550–551; Vettik 1996: 5). Võõra ja oma termini kasutamisel tuleb samuti arvestada oma oskussõnavara lähedust teiste rahvaste terminoloogiale ning oskuskeele kokkukuuluvust rahvuskeelega (Erelt 1982: 76).

Eesti õigustekstides (nt ajakirjas *Juridica* ja õppematerjalides) rakendatakse enamasti stiilivahelduse huvides korraga mõlemat paarikut, s.o nii võõrsõnalist kui ka omasõnalist väljendit. Siiski viitavad õiguskeele uurijad, et üldjuhul eelistatakse juristide kirjatöodes terminite sünonüümse kasutuse puhul võõrsõna (Kukk 2003: 100).

• Liitsõnad ja sõnaühendid

Peale kõrvu esinemise ning tekstis vaheldumise avaldub oma ja võõra keeleainese koosseksistents ka selles, et oma- ja võõrморфеeme võib sulatada kokku üheks keelendiks, moodustada nn hübriide. Võõr- ja omasõnadest moodustatud liitsõnade ja sõnaühendite moodustamise otstarbekuse määrab nagu võõrterminitegi puhul mõiste tarvitamise ulatus.

Sageli võib tekstidest leida liitsõnu, milles põhisõnaks on omasõna, täiendosaks võõrtüvi. Omasõna vaatenurgast on see progressiivne liitumine, näiteks *apellatsioonikaebus* (ladina *apellatio*), *regressiõigus* (*regressus*), *tsiviilõigus*, *tsiviilasi*, *tsiviililisik* (*civilis*), *munitsipaalõigus* (*municipalis*), *kausaaltehing* (*causalis*), *krediidileping* (*creditum*), *agendileping* (*agens*), *komisjonileping* (*commissio*).

Liitsõnad võivad esineda ka vastupidisel kujul, seega regressiivse liitumisena, kus põhisõnaks on võõras tüvi, näiteks *õigusnorm* (ladina *norma*), *õigusobjekt* (*obiectum*), *kohtuinstants* (*instans* < *instantia* keskaegne tähendus), *õigusinstituut* (*institutum*), *eraldamisprintsip* (*principium*). Sellistes liitsõnades ei ole ladina sõnad spetsiifiliselt õigusteaduse alale kuuluvad mõisted.

Samuti on õigustekstides sagedased sõnaühendid, mille moodustavad omasõna ja võõrsõna. Ladina tüvest tuletatud võõrsõna on enamasti täiendiks, seega täpsustavaks mõisteks: *bilateraalne leping* (ladina sõna *bilateralis*), *negatoorne hagi* (*negatorius*), *reaalsed lepingud* (*realis*), *singulaarne õigusjärglus* (*singularis*), *formaalne karistusõigus* (*formalis*), *jurüidiline isik* (*iuridicus*), *intellektuaalne omand* (*intellectualis*), *sotsiaalne kaitse* (*socialis*), *primaarsed allikad* (*primus*), *subjektiivne õigus* (*subiectivus*), *faktiline võim* (*factum*).

Nendes näidetes leidub ka juhtumeid, kus võõrkeelest pärit sõna võiks edukalt asendada omasõna, näiteks faktiline võim ~ *tegelik võim* (*factum* 'tegu'). Samuti

leidub õiguskeeles ka väljendeid, milles ladinatuveline võõrtermin on sõnaühendis põhisõnaks, näiteks *isiklik servituut* (*servitus*), *õiguse kodifitseerimine* (*codificatio*), *avalik-õigulik korporatsioon* (*corporatio*).

2.2. Grammatilised variandid

Ladina õigusterminoloogias on laialt levinud sünonüümide grammatilised variandid. Need on ladina keele eripärast (tuletusviis, grammatiline süsteem) tulenevad terminivariandid, mille sõnatüved langevad täielikult või osaliselt kokku, erinev on üksnes nende grammatiline vormistus.

Sünonüümseteks terminiteks võivad olla näiteks grammatilised vormid ja tuletised. Paljudel juhtudel moodustatakse variant passiivi partitsiibi tüvele liidetud tegevuse toimumist või selle tulemust väljendava sufiksiga *-(t)io*, näiteks *locatum* ja *locatio* ('üür, üürimine'), *emptum* ja *emptio* ('ost, ostmine'), *venditum* ja *venditio* ('müük, müümine'). Grammatilised sünonüümid võivad lähtuda ka omadust väljendava sufiksi *-tas* ning päritolu väljendava liite *-inus* kasutamisest, näiteks *aequum* ja *aequitas* ('võrdsus'), *libertus* ja *libertinus* ('vabakslastu'). Siin on tegu substantiveerunud adjektiivide ja adjektiividest tuletusliite abil saadud substantiivide paralleelse esinemisega.

Terminivariante esineb samuti samast sõnatüvest moodustatud omadussõnalise ja nimisõnalise (genitiivi käändes või koos eessõnaga) täiendiga, näiteks *ius naturale* ja *ius naturae* ('loomuõigus'¹⁸), *successio universalis* ja *successio per universitatem* ('üldõigusjärglus'), *actio doli* ja *actio de dolo* ('salakavaluse kohta esitatud hagi'). Eestikeelsete vastete puhul siin variatiivsust pole.

Sünonüümid tekivad ka verbi kõneviiside (*indicativus*, *coniunctivus*, *imperativus*), eituspartiklite, samuti oleviku ja minevikuagade varieerimisel: *ne bis in idem* ja *non bis in idem* ('ärgu olgu / mitte kaks korda sama asja eest'), *nemo plus iuris ad alium transferre potest quam ipse habet* ja *nemo plus iuris ad alium transferre posset quam ipse haberet* ('mitte keegi ei saa rohkem õigusi teisele isikule üle kanda, kui tal endal on / võiks olla'), *audiatur et altera pars* ja *audi alteram partem* ('võetagu kuulda / kuula ka teist osapoolt'), *iura novit curia* ja *iura noscit curia* ('kohus tundis/tunneb õigust').

2.3. Ortograafilised variandid

Kõige rohkem võib õigustekstidest leida ladina terminite ortograafilisi variante. Need on sünonüümid, mille terminivariantide erinevus kajastub ainult õigekirjas. Ortograafilised variandid võivad tekkida ladina õigusterminoloogiasse võõrsõnade (enamasti kreeka keelest, aga ka teistest keeltest) laenamisel, kus mõjuvad kaasa ladina keele enda fonoloogilised seaduspärasused ja häälikuline eripära, näiteks aspiraaside vahesus: *thesaurus* ja *tesaurus* ('peitvara'), *hypotheca* ja *hypoteca* ('kinnis(asja)pant'), *arrha* ja *arra* ('käsiraha'). Variante moodustavad ka liitsõnad, milles eesliite ja põhisõna ühinemisel on toimunud häälikumuutus. Need variandid on tingitud keelearengu üldisest seaduspärasest (assimilatsiooniprotsessid, geminaadi lühenemine jne): *adquisito* ja *acquisitio* ('omandamine'), *adscriptio* ja *ascriptio* ('allkiri'), *connubium* ja *conubium* ('abielu'), *exsilium* ja *exilium* ('pagendamine').

¹⁸ Vrd *ius naturale* / *ius naturae* varasemat eesti vastet *loodusõigus*.

Ortograafilisi variante annab veel *-ci* ja *-ti* vaheldumine ladina terminites, põhjuseks antiikaja lõpus toimunud häälikuühendi *-ti-* assibilatsioon: *condicio* ja *conditio* ('tingimus'), *nuncius* ja *nuntius* ('advokaat; esindaja'), *actio pigneraticia* ja *actio pigneratitia* ('pandinõue').

Variante tekib samuti kokku- ja lahkukirjutamisel, näiteks *ius iurandum* ja *iusiurandum* ('vanne'), *iuris prudentia* ja *iurisprudencia* ('õigusteadus'), *senatus consultum* ja *senatusconsultum* ('senati otsus'), *sub poena* ja *subpoena* ('karistuse ähvardusel'). Kokku- ja lahkukirjutamine ei anna siiski mitte alati samatähenduslikku vastet. Näiteks erineb tähendus järgmiste terminite puhul: *ab usu* ('tavakohaselt') ja *abusu* ('kuritarvitades').

Mitmetes väljendites on variatiivsuse põhjustajaks sõnajärg, näiteks *ratio scripta* ja *scripta ratio* ('kirjapandud tarkus'¹⁹), *sensu stricto* ja *stricto sensu* ('otseses mõttes'), *actor forum rei sequitur* ja *actor sequitur forum rei* ('hageja kuuletub kostja elukohamaa kohtule'), *de minimis lex non curat* ja *de minimis non curat lex* ('seadus ei kanna hoolt kõige tühisemate asjade eest'). Ladina keeles on sõnajärg suhteliselt vaba. Üldjuhul paikneb täiend põhisõna järel. Siiski ei muuda täiendi ja põhisõna ümberpaigutamine väljendi tähendust. Eesti keeles on sõnajärje puhul iseloomulik, et öeldisverb püsib tekstilauses teisel kohal (Kerge 2001: 155). Ladina keeles paikneb öeldis enamasti lause lõpus, kuid toodud näidetes jääb lause tähendus samaks ka juhul, kui öeldise koht on variandis muutunud.

2.4. Lühivariandid

Ladina õigusterminite täisvariantide kõrval kasutatakse ka lühivariante. Terminite lühivariandid saadakse termini komponentide asendamise või väljajätmise ehk ellipsi teel.

2.4.1. Asendamine

Asendamine toimub, kui terminis on sõnaühendid asendatud ühe nimisõnaga, mis kujutab endast ühendis esineva sõna grammatilist ja tuletuslikku varianti. Sellistes lühivariantides jäetakse ära põhisõna, täiendsõnast saab iseseisev substantiiv, näiteks *ius utendi-fruendi* ja *ususfructus* ('kasutusvaldus'), *res mobiles* ja *mobilia* ('vallasasjad'), *lex loci regit actum* ja *locus regit actum* ('koha seadus / koht juhib tehingut'²⁰), *actori incumbit onus probandi* ja *actori incumbit probatio* ('hagejal lasub tõendamiskoormis / tõendamine'). Väljendi võib asendada ka uue nimisõnaga, näiteks *praescriptio longi temporis* ja *usucapio* ('igamine').

2.4.2. Ellips

Elliptiline variant saadakse terminist ühe või mitme komponendi väljajätmisel. Paljudel juhtudel jäävad välja verbi erinevates vormides osised (infinitiiv, partitsiip, gerundiiv), näiteks *pacta sunt servanda* ja *pacta servanda* ('lepingud tuleb täita'), *ex post facto* ja *ex post* ('pärast jõustumist toimuv'), *lex loci delicti commissi* ja *lex loci delicti* ('kuritöö toimepanemise koha seadus'), *suum cuique tribuere* ja *suum cuique* ('määrama igäuhele oma'). Terminite lühivariante moodustatakse ka nimisõnade väljajätmisega: *legitimatatio ad causam activa* ja *legitimatatio activa* ('aktiivne

¹⁹ Väljendit kasutatakse Rooma õiguse, eriti *Corpus Iuris Civilis*'e kohta (Adomeit 2005).

²⁰ Tehingu tegemisel kehtib selle maa seadus, kus tehing tehti (Adomeit jt 2005).

legitimatsioon'), *lex specialis derogat legi generali* ja *lex specialis derogat generali* ('eriseadus tühistab üldise seaduse'). Samuti leidub elliptilisi variante, milles jäävad välja nii verb kui substantiiv: *de cuius hereditate agitur* ja *de cuius* ('kelle pärand on arutlusel').

Angloameerika õigusterminoloogias leidub palju väljendeid, mis sisaldavad ladina sõnade kõrval ingliskeelseid sõnu, näiteks *ad hoc agreement*, *eiusdem generis rule*, *ex parte proceeding*, *in personam jurisdiction*. Mõnikord võib sellises väljendis esineda ka ladina termini elliptiline vorm, mida ei ole võimalik avada sõnade põhitähenduse kaudu. Näiteks terminis *nisi prius clause* on vanast ladinakeelsest väljendist kasutatud vaid paari esimest sõna, mille tõlge 'kui mitte varem' ei viita termini võimalikule tähendusele. Sõnastiku järgi hakkas see väljend keskaegsel Inglismaal tähistama olukorda, kus vandemeestel kästi tulla Londonisse Westminsterisse, "kui mitte varem" polnud kohtupidamist korraldatud maakonnas. Tänapäevases angloameerika õiguskeeles tähistab see termin esimese astme menetlust vandemeeste juuresolekul (Sodhi, Vasani 1980: 57).

Kokkuvõte

Oskuskeel kui täpsete erialamõistetega opereeriv keelevariant esitab keelenditele kui nimetusüksustele teised nõuded kui üldkeel. Terminoloogia puhul on vajalik sisuline, mõisteline täpsus, s.t sellised eristused ja täpsustused, mis üldkeelele pole nii rangelt vajalikud ega omased. Oskuskeeles välditakse üldiselt nii terminite mitmetähenduslikkust kui samatähenduslikke termineid, kuid täielikult neist hoiduda ei saa. Terminite tegelikud talitlustingimused loovad tüüpsituatsioone, kus tähistaja ja tähistatava üksühese vastavuse rikkumine on paratamatu: terminite keelemärgina mõjutavad polüseemia ja sünonüümia.

Terminite mitmetähenduslikkust ehk polüseemiat oskuskeeles üldiselt välditakse. Absoluutne ühetähenduslikkus pole siiski kooskõlas keele funktsioneerimise põhialustega. Terminite sisu vastab kohas ja ajas kehtivatele seaduspärasustele, selle kasutus võib kitseneda, laieneda või nihkuda teiste mõistete suunas. Mitmetähenduslikkuse puhul eristuvad kõigepealt terminite tähendused üldkeeles ja õiguskeeles. Polüseemiat esineb ka õiguse valdkondade vahel. Erinevad tähendused samadele terminitele on kujunenud näiteks karistusõiguses ja tsiviilõiguses.

Õigusterminite tähendused erinevad ka suurte õigussüsteemide keelekasutuses. Õiguskeel erineb teistest erialakeeltest olulises aspektis: see on kultuuripõhine ning seotud ühe kindla ühiskonna ja selle õigussüsteemiga. Ehkki Euroopa õigussüsteemide baseerub ühisel Rooma õiguse pärandil ning sellest võrsunud terminid on õigussüsteemides juurdunud, erinevad paljude ladina väljendite tähendused angloameerika ja Mandri-Euroopa õigussüsteemide terminoloogias. Väiksemaid lahknevusi terminite tähenduses võib leida ka Mandri-Euroopa õigussüsteemi romaani ja germaani õigusperekondade keelekasutuses.

Õiguskeeles on levinud ka sünonüümid ehk terminivariandid. Ehkki erialasel terminikasutusel on omad nõuded ja erijooned võrreldes keele üldsõnavaraga ning terminite kasutusega väljaspool erialast tarvitust, on siin ka kokkupuutepunkte ning sarnaseid jooni. Neid esineb niihästi leksikaalsete kui ka muude variantide kasutamises. Oskuskeeles ei ole terminiteisendused nii vabalt kasutatavad kui üldkeeles, siiski ei saa õigusterminoloogias variante täielikult vältida.

Ladina keelest pärit õigusterminite puhul ei erine variandid õiguskultuuriti ega keelepiiride tasandil. Sünonüümsus on enamasti kahe sõna vahel, kuid see võib esineda ka sõna ja ühendi vahel, kusjuures viimane täpsustab lühemat ja kompaktsemat varianti. Samuti võivad sünonüümid olla ühendid või väljendid omavahel. Sagedased on terminite grammatilised ja ortograafilised variandid, samuti täistermini ja selle lühikuju paralleelne esinemine. Leksikaalsete variantide levinuim liik on võõrsõnalise ja omakeelse termini rööpesinemine. Suurem võõrsõnakasutus on omane erialasemale keelekasutusele, eriti rahvusvahelises infovahetuses ja keeltevahelises suhtluses, omakeelendite eelistamine on omal kohal aga omakeelse ja üldisema esituse korral.

Kirjandus

- Adomeit, Klaus; Ristikivi, Merike; Siimets-Gross, Hesi 2005. Ladina-eesti õigussõnastik. Peeter Kask (toim.). Tallinn: Eesti Keele Sihtasutus.
- Anepaio, Toomas; Hussar, Ave; Jaanimägi, Katri; Kaugia, Silvia; Land, Kristi; Olle, Vallo; Roosma, Peeter 2005. Sissejuhatus õigusteadusesse. Tallinn: Juura.
- Annars, Erik; Önnersfors, Alf 1972. Latinsk juridisk terminologi. Uppsala: Juridiska Föreningen i Uppsala.
- Asjaõigusseadus 1993. Vastu võetud 9.06.1993. a seadusega (RT I 1993, 39, 590), jõustunud 1.12.1993.
- Benke, Nikolaus; Meissel, Franz Stefan 1997. Juristenlatein. Lateinische Fachausdrücke und Redewendungen der Juristensprache, übersetzt und erläutert. Wien: Juridica.
- Berman, Harold J.; Reid, C. Jr. 1995. Römisches Recht in Europa und das *ius commune*. – Zeitschrift für europäisches Recht, 3–34.
- Berteloot, Pascale 1999. Der Rahmen juristischer Übersetzungen. – Recht und Übersetzen. Hrsg. Gerard Rene de Groot, Reiner Schulze. Baden-Baden: Nomos.
- Birks, Peter; McLeod, G. 1987. Justinian's Institutes. London: Gerald Duckworth & Co.
- Black, Henry Campbell 1998. Black's Law Dictionary. 13th reprint, 6th ed. by the publisher's editorial staff. Coauthors Joseph R. Nolan, Jacqueline M. Nolan-Haley. Contributing authors M. J. Connolly *et al.* St. Paul, Minnesota: West Publishing Co.
- Bruss, Jochen 1999. Lateinische Rechtsbegriffe. 2. Aufl. Freiburg: Haufe Verlagsgruppe.
- Cabanelas, Guillermo; Alcalá-Zamora y Castillo, Luis 1992. Diccionario enciclopédico de derecho usual I–VIII. 16a ed. Buenos Aires: Editorial Heliasta.
- Cornu, Gerard 1990. Linguistique juridique. Paris: Montchrestien.
- Cornu, Gerard 1992. Vocabulaire juridique. Paris: Presses Universitaires de France.
- Ekspertarvamus 1995 = Sõnade *omand* ja *omandus* semantika ning terminoloogiline rakedatavus. Eesti Keele Instituudi ekspertarvamus. 1995. – Õiguskeel 2, 24–26.
- Erelt, Tiiu 1982. Eesti oskuskeel. Tallinn: Valgus.
- Ilus, Elmar 2000. Rooma eraõiguse alused. Heiki Pisuke (toim.). Tallinn: Ilo.
- Jowitt's Dictionary = Jowitt's Dictionary of English Law I–II. 1977. London: Sweet & Maxwell.
- Kask, Peeter; Kull, Rein 1996. Asjaõigusseaduse terminikasutus. – Õiguskeel 5, 19–27.
- Kerge, Krista 2001. Eesti süntaks keeleõppe praktikule. Käsiraamat. Tallinn: TEA Kirjastus.
- Kerge, Krista 2004. Terminid leksikaalse ja grammatilise keelendina. – Sirje. Mäearu (toim.). Rahvusvaheline terminoloogikonverents "Eesti oskuskeel 2003" 10. ja 11. oktoobril 2003. Ettekanded. Tallinn: Eesti Keele Sihtasutus, 19–46.
- 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 Pedagoogikakooli eesti filoloogia osakonna toimetised 4. Tallinn: TPÜ Kirjastus, 96–117.
- Kull, Rein 2000a. Kirjakeel, oskuskeel ja üldkeel. Tallinn: Eesti Keele Sihtasutus.

- Kull, Rein 2000b. Kas kindel või lagundatud kirjakeel? (Muutuv keel) – Keel ja Kirjandus 1, 1–9.
- Kull, Rein 2000c. Oskuskeel ja üldkeel: erisused ja samasused. – Keel ja Kirjandus 8, 545–557.
- Kull, Rein; Saari, Henn 1975. Die Entwicklung des estnischen terminologischen Gedankens und die Förderung der Fachsprachen in den letzten Jahren. – Congressus Tertius Internationalis Fenno-Ugristarum. Pars I. Acta Linguistica. Tallinn, 245–249.
- Lieberwirth, Rolf 1987. Lateinische Fachausdrücke im Recht. Heidelberg: Müller.
- Liebs, Detlef 1998. Lateinische Rechtsregeln und Rechtssprichwörter. 6. Aufl. München: Beck.
- Lill, Anne 1986. Kreeka-ladina päritoluga terminite tulek Eesti meditsiinitekstidesse (1869–1914). – Tartu Riikliku Ülikooli Toimetised 728. Fenno-Ugristica 13, 77–88.
- Mattila, Heikki E. S. 2002. Vertaileva oikeuslingvistiikka. Helsinki: Kauppakaari Lakimiesliiton Kustannus.
- Mellinkoff, David 1963. The Language of the Law. Boston, Toronto: Little, Brown and Co.
- Mereste, Uno 1995. Arvamus sõnade *omand* ja *omandus* semantikast ning rakendatavusest. Eesti Keele Instituudi ekspertarvamuse teemal. – Õiguskeel 3, 35–38.
- Mereste, Uno 2000. Oskuskeel ja seaduste keeleline rüü. Tallinn: Eesti Keele Sihtasutus.
- Mõttus, Aaro 1997. Seaduse sõnastamine ja terminoloogia (3). – Õiguskeel 4, 34–46.
- Narits, Raul 1995. Õiguse entsüklopeedia. Õpik Tartu Ülikooli õigusteaduskonna üliõpilastele. Tallinn: Juura.
- Narits, Raul 1997. Õigusteaduse metodoloogia I. Tallinn: Juura.
- Oksaar, Els 1999. Kõnekeel, erialakeel, õiguskeel. – Juridica 4, 199–207.
- Pärna, Priidu 1995. OMAND ja OMANDUS juriidilisest vaatekohast. – Õiguskeel 2, 27–28.
- Ristikivi, Merike 2006. Ladina päritolu juriidilised terminid Eesti õiguskeeles. Magistritöö Tartu Ülikooli Raamatukogus.
- Roland, Henri; Boyer, Laurent 1998. Locutions latines du droit français. Paris: Litec.
- Saari, Henn 1980a. Omasõna ja võõrsõna paarid eesti oskussõnavaras (1). – Keel ja Kirjandus 11, 654–666.
- Saari, Henn 1980b. Omasõna ja võõrsõna paarid eesti oskussõnavaras (1). – Keel ja Kirjandus 12, 737–743.
- Saari, Henn 1981a. Omasõna ja võõrsõna paarid eesti oskussõnavaras (2). – Keel ja Kirjandus 4, 201–210.
- Saari, Henn 1981b. Omasõna ja võõrsõna paarid eesti oskussõnavaras (2). – Keel ja Kirjandus 5, 282–288.
- Silvet, Erki 1997. Ladinakeelne väljend õiguskeeles. – Õiguskeel 2, 16–19.
- Sodhi, Datinder; Vasan, R. S. (Eds.) 1980. Latin Words and Phrases for Lawyers. New York: Law and Business Publications.
- Soosaar, Sven-Erik 2005. Õigusterminite etimoloogia. – Aime Vettik, Erki Silvet (koost.). Keel ja õigus. Tallinn: Juura, 111–117.
- Talvik, Edgar 1936. Rooma õiguslugu. Konspekt. Dotsent dr. jur. Leo Leesmenti loengute alusel. Tartu.
- Tavast, Arvi 2004. Mõiste mõiste: tähendusteooriate mõju oskuskeepraktilikale. – Sirje Mäearu (toim.). Rahvusvaheline terminoloogiakonverents “Eesti oskuskeel 2003” 10. ja 11. oktoobril 2003. Ettekanded. Tallinn: Eesti Keele Sihtasutus, 7–18.
- Tiersma, Peter 1999. Legal Language. Chicago, London: The University of Chicago Press.
- Vettik, Aime 1996. Oma ja võõras. – Õiguskeel 5, 4–7.

Merike Ristikivi (Tartu Ülikool) on erialalt klassikaline filoloog, uurimisteemaks on ladina keele mõjutused eesti õiguskeeles.
merike.ristikivi@ut.ee

POLYSEMY AND SYNONYMY OF LATIN TERMS IN ESTONIAN LEGAL LANGUAGE

Merike Ristikivi

University of Tartu

The article discusses some problems of polysemy and synonymy occurring in Latin legal language. The material analysed comprises the terms collected for the compilation of the Latin-Estonian Juridical Dictionary (2005). Polysemy is generally avoided in terminology. However, absolute monosemy is not in accordance with the principles of language functioning. The meaning of the term reflects the local conditions and time in which it is used, and its usage may become more specific or more general, or the term may even come to denote a totally different idea. In the case of polysemy the meanings of the terms become distinct first in the common language and legal language. Polysemy also occurs in the legal language across the various areas of law. Different meanings for the same terms have developed in penal power and civil law, for example. The meaning of legal terms also varies in the usage characteristic of different legal systems. Legal language differs from language for other specific purposes in one significant aspect: it is culturally determined, associated with an individual society and its legal system. Even though European legal thinking is generally based on the common heritage of Roman Law and the terms originating from it have become rooted in modern legal systems, the meanings of Latin legal phrases often diverge depending on whether they are used in the Anglo-American legal system or in continental Europe. Smaller differences between the meanings of identical terms can also be found in the use of legal language of Romance and Germanic legal families.

The synonymous variants of Latin legal terms do not vary on the level of legal systems and linguistic boundaries. Synonymy occurs mainly between two single words, but it can also appear between a single word and a phrase, with the latter being more specific. Also, synonymy can be observed between phrases or expressions themselves. Often grammatical and orthographic variants appear, and a term and its abbreviated form may exist in parallel. The most common type of lexical variants is the parallel existence of a foreign term and a native term. The trend is towards using foreign words more readily in professional writings, particularly on the level of international communication and information exchange, whereas native phrases are preferred in more general writings and writings in the native language.

Keywords: legal language, terminology, terminological dictionaries, lexical semantics, language contact, polysemy of terms, synonymy of terms, Latin, Estonian

VÕIMALUSI HELISTAJA IDENTIFITSEERIMISEKS INSTITUTSIONAALSES TELEFONISUHTLUSES

Andriela Rääbis

Ülevaade. Artiklis käsitletakse helistaja identifitseerimise võimalusi eri tüüpi institutsionaalsetes telefonikõnedes. Uurimuse meetod on vestlusanalüüs. Keskendutakse võõrastevahelisele suhtlusele, kuid selle kõrval vaadeldakse ka väikest tuttavate suhtlejate allkorpust. Infodiaalooogides jääb helistaja enamasti anonüümseks. Kui vastaja peab kliendi heaks tegema midagi enam kui ainult infot andma, vajab ta kliendi nime ja/või muid isikuandmeid. Helistaja võib end tutvustada kohe vestluse alguses, aga ametnik võib nime küsida ka hilisemas vestluses. Identifitseerimist on mitut liiki: ühel juhul toimub identifitseerimine käesoleva vestluse jaoks (nt müügivestlus): osalejad ei saa suhelda teadmata, kes on vestluspartner. Teisel juhul toimub identifitseerimine edaspidise tegevuse jaoks (nt arsti vastuvõtule registreerimine).*

Võtmesõnad: suuline kõne, vestlusanalüüs, institutsionaalne suhtlus, telefonivestlus, eesti keel

Sissejuhatus

Artiklis vaadeldakse, kuidas identifitseeritakse helistaja eri tüüpi institutsionaalsetes dialoogides. Argivestluses identifitseerivad partnerid teineteise kohe vestluse alguses, kuid institutsionaalsetes dialoogides jääb klient enamasti anonüümseks, teatud situatsioonides küsitakse isikuandmeid aga hilisemas suhtluses. Kliendi–ametniku dialooge analüüsid selgus, et vestluse alguses tutvustas end 97,3% vastajatest ning 2,1% helistajatest (Rääbis 2006).

Telefonivestluse sissejuhatus koosneb üksteisele kindlas järjekorras järgnevatel sekventsidel. Emanuel Schegloff (1986: 113–133) on leidnud, et sissejuhatus tuumsekventsid on järgmised:

1. Kutsung – vastus: telefonihelin ja vastaja esimene voor.
2. Identifitseerimine (ja/või äratundmine): osalejad tutvustavad ennast ja väljendavad partneri äratundmist.

* Artikli valmimist on osaliselt toetanud ETF (grant 5685). Autor on märkuste eest tänulik Renate Pajusalule ja Tiit Henonstele.

3. Tervitused.
4. *Kuidas läheb*-sekventsids.

Pärast nende etappide läbimist jõuavad vestlejad positsiooni, kus saavad vestlusse tuua esimese teema. E. Schegloff (1986: 116) nimetab seda ankrupositsiooniks. Esimest teemat alustab tavaliselt helistaja.

Ingliskeelsete telefonikõnede hilisemad uurimused on näidanud, et niisuguse mudeli järgi viiakse sissejuhatused läbi siis, kui osalejad on omavahel tuttavad ning nende suhted on neutraalsed. Lähedaste või võõraste suhtlejate puhul on sissejuhatuse enamasti redutseeritud. Lähedased jätavad vahele identifitseerimissekventsi, partner tuntakse ära hääle järgi. Võõrastevahelise suhtluse puhul võidakse ära jätta kõik sissejuhatuse osad peale kutsungi–vastuse: enamasti vastaja tutvustab end ning helistaja läheb kohe oma esimeses voorus asja juurde. (Whalen, Zimmerman 1987; Hopper jt 1990/91; Hopper, Drummond 1992; Hopper 1992: 77–83.)

Eesti telefonivestluse sissejuhatuse täielik mudel koosneb samuti neljast sekventsist, kuid osade järjekord on teistsugune: tervitatakse enne tutvustamist (Rääbis 2000: 418). Neljaosaline sissejuhatuse ei ole Eestis tüüpiline: varasem väikese korpuse (120 telefonikõnet) analüüs näitas, et ainult 16% sissejuhatustest koosnes neljast sekventsist (Rääbis 1999: 94). Kliendi ja ametniku telefonikõnedes on normiks niisugune sissejuhatuse, kus vastaja ennast tutvustab ning mõlemad osalejad tervitavad (Rääbis 2006: 146–147).

1. Materjal ja meetod

Käesoleva analüüsi aluseks on 684 institutsionaalset telefonikõnet (91289 sõna) Tartu Ülikooli Eesti dialoogikorpusest¹ ning Tartu Ülikooli Eesti suulise keele korpusest.² Analüüsitava materjal jaguneb 10 allkorpuseks (vt tabel 1): infotelefon, bussiinfo, polikliiniku registratuur, reisibüroo, aktiivne müük, riigiasutused (nt pensioniamet, sotsiaalabi, häirekeskus, politsei), kauplus, takso tellimine, teenindus (nt ehitusfirma, raamatukogu, autoteenindus, pank, hotell), kolleegid. Suhtluse keskmes on info küsimine. Erandiks on taksofirma dialoogid, kus enamasti ei soovita mitte infot, vaid tellitakse takso, ning registratuuridialoogid, kus mõnel juhul soovitakse infot, mõnel juhul aga arsti vastuvõtuaega reserveerida. Helistajaks on enamasti klient (infoküsjaja) ning vastajaks ametiisik (infoandja). Osalejate rollid on vastupidised mõnes reisibüroodialoogis, kui infoandja helistab kliendile tagasi. Müügivestlused erinevad oluliselt muudest dialoogidest: koolitusfirma töötaja helistab kliendile (klient on samuti ametiisik) eesmärgiga pakkuda talle koolituskursust. Peale müügivestluste esindavad mõlemad suhtlejad institutsiooni veel vaid mõnes üksikus dialoogis (nt helistatakse raadiost politseisse, päästeametisse; sotsiaalabi osakonnast polikliiniku registratuuri). Osalejad on enamasti võõrad. Tuttavate suhtluse uurimiseks on võetud suulise keele korpusest 29 kolleegide telefonikõnet. Tuttavate suhtlust on ka registratuuridialoogide hulgas (polikliiniku arst või õde helistab registratuuri; 14 dialoogi) ning taksodialoogide hulgas (taksodispetšer helistab kolleegile teisest firmast; 2 dialoogi). Peaaegu kõik kõned on lauatelefonile, seitse kõnet mobiilile. Kas helistaja kasutab lauatelefoni või mobiili, pole paljudel juhtudel teada, kuna enamik kõnesid on salvestatud vastaja poolt.

¹ [http://math.ut.ee/~koit/Dialoog/EDIC \(27.07.2006\)](http://math.ut.ee/~koit/Dialoog/EDIC (27.07.2006))

² [http://www.cl.ut.ee/suuline/ \(27.07.2006\)](http://www.cl.ut.ee/suuline/ (27.07.2006))

Tabel 1. Analüüsitavad allkorpused

Allkorpus	Sõnade arv	Dialoogide arv
Infotelefon	27297	413
Bussiinfo	1797	14
Registruur	5184	65
Reisibüroo	26116	62
Aktiivne müük	7721	14
Takso	986	22
Riigiasutused	3473	13
Kauplus	2819	10
Teenindus	9221	42
Kolleegid	6675	29
Kokku	91289	684

Dialoogide litereerimisel on kasutatud vestlusanalüüsi transkriptsiooni (vt lisa). Osalejate anonüümsuse huvides on kõik nimed transliteratsioonides muudetud. H tähistab helistajat, V vastajat; müügivestlustes on osalejad tähistatud: M – müüja (helistaja), S – sekretär (vastaja), K – klient (vastaja).

Uurimuse meetodiks on vestlusanalüüs. Vestlusanalüüs uurib eelkõige seda, kuidas suhtlus sünnib osalejate ühistööna. Tõstetakse esile kolm kesket vestlust loovat liigendust: vooruliigendus, sekventsiliigendus ja parandusliigendus. Need kolm samaaegselt toimivat liigendust on suhtluse eelduseks ning alles nende taustal tõlgendavad kõnelejad sõnu ja lauseid.

2. Institutsionaalne suhtlus

Vestlusanalüüsis ei tõlgendata institutsionaalsust etteantud parameetrina, vaid lähtutakse sellest, et institutsionaalsus tekib kõnelemise ja tegutsemise läbi. Institutsionaalses suhtluses toimivad samad liigendused kui argivestluses. Vestlusanalüüsi huvitab siin eelkõige see, kuidas nende liigenduste kaudu luuakse institutsionaalne kontekst ning täidetakse institutsionaalseid ülesandeid. Institutsionaalsusele osutatakse näiteks sellega, et kasutatakse teatud sõnavara, teatud vooruvahetusreegleid või teatud tüüpi voore.

Vestlusanalüüsi seisukohalt on institutsionaalsel suhtlusel kolm tüüpilist joont:

- suhtlusel on eesmärk,
- suhtluses osalemist korraldavad teatud piirangud,
- kasutatakse institutsioonidele iseloomulikke raamistikke (Drew, Heritage 1992: 24–25; Heritage 1997: 163–164).

Paul Drew ja John Heritage (1992) vaatlevad institutsionaalset suhtlust järgmistest aspektidest.

1. Sõnade ja kujundite valik. Sõnavalik on oluline vahend, mille abil kõnelejad loovad vestluse ametliku konteksti. Inimesed kasutavad sõnu, mis sobivad nende rollidega ametlikus situatsioonis.

2. Voo ru kavandamine:
 - a) tegevuse valimine. Iga voo ru ga tehakse midagi: küsitakse, antakse nõu jne. Seega tuleb voo ru kavandades valida kõigepealt tegevus;
 - b) tegevuse keelelise esituse valimine. Mistahes tegevust võib väljendada keeleliselt erinevatel viisidel: kõneleja peab lisaks sõnavalikule valima ka lauseehituse, otsustama, mis järjekorras eri elemendid toodetakse jne. Voo ru ehitus sõltub sageli kõneleja institutsionaalsest rollist.
3. Sekventsi struktuur. Naaberpaaride ja pikemate sekventside ülesehitus, naaberpaari esi- ja järellikme vahel olevad vahesekvensid jne on seotud osalejate rollidega institutsionaalses kontekstis. Voo ru vahetust juhivad antud situatsioonis kehtivad reeglid. Institutsionaalsele suhtlusele on iseloomulikud ka kindla malli järgi läbiviidavad sissejuhatused ja lõpetamised.
4. Vestluse kogustruktuur. Paljudes ametlikes situatsioonides sooritatakse mingeid tegevusi teatud kindlas järjekorras. Seetõttu koosneb ka vestlus neis situatsioonides sekventsidest, mis järgnevad üksteisele teatud järjekorras. Kuid iga konkreetse vestluse ülesehitus sõltub suurel määral konkreetsest situatsioonist ning selles osalejatest.
5. Sotsiaalsed suhted:
 - a) ametialane neutraalsus. Ametiisikud ja osaliselt ka kliendid püüavad olla käsitleva suhtes neutraalsel positsioonil;
 - b) voo ru vahetuse ebasümmeetrilisus. Institutsionaalne vestlus on reeglina ebasümmeetriline. Paljudel juhtudel on osaleja staatuse ja rolli ning diskursiivsete õiguste ja kohustuste vahel otsene seos. Asümmeetriat põhjustavad järgmised asjaolud.
 - Vestlus toimub sageli küsimus–vastus–vormis. Ametiisik (arst, õpetaja, intervjueriija jne) juhib vestlust, otsustades, millest räägitakse ning kui pikalt räägitakse. Teisel osalejal on initsiatiivi haaramiseks vähe võimalusi.
 - Osalejatel on erinevad teadmised. Seda asümmeetriat on kõige rohkem uuritud arsti–patsiendi vestlustes. Erinevused mingi probleemi meditsiinilise definitsiooni ja patsiendi võhikliku versiooni vahel põhjustavad sageli arusaamatusi ja konflikte.
 - Ametiisiku jaoks on situatsioon rutiinne, kliendi jaoks unikaalne ja isiklik. Klient võib mingit tüüpi ametlikku situatsiooni sattuda vaid kord elus ning ta ei käsitleni oma probleemi rutiinse juhtumina.

Argise ja institutsionaalse suhtluse süsteemis saab eristada mitu alltegurit:

- suhtluse põhieesmärk: osalemine / kindla eesmärgiga suhtlemine;
- inimestevahelised suhted: tuttavad / võõrad, sotsiaalsed ja territoriaalsed suhted;
- inimeste rollid vastavas situatsioonis: eraisik / institutsiooni esindaja;
- konkreetne füüsiline olukord: kodu / neutraalne territoorium / ametiasutus;
- vaba / normiline / norminguline suhtlus.

Neid tegureid omavahel kombineerides saame rea situatsioone. Institutsionaalne suhtlus on eesmärgistatud, toimub enamasti ametiterritooriumil, keske rühma

moodustab normiline suhtlus. Instituutsionaalsel suhtlusel ei ole ühte selget keset, neli põhirühma eristuvad kahe tunnuse järgi:

- suhtlejad on võõrad;
- suhtlejad on tuttavad;
- suhtlevad instituutsiooni esindajad;
- suhtlevad instituutsiooni esindaja ja eraisik.

Situatsioonitüüpide rühmitamisel on olulised veel järgmised tegurid:

- monoloog / dialoog / polüloog;
- inimestevahelised suhted: instituutsionaalselt seotud / sidumata, staatus- ja formaalsussuhted;
- spontaansus / ettevalmistatus;
- tegevus on sagedane, regulaarne / harv, juhuslik;
- vahetult reaalsuse muutumisega seotud / sidumata tekst;
- kindlate rollidega / vabam suhtlus. (Hennoste 2003: 492–499)

Infodialoge uurides selgus, et vestluse arengu seisukohalt on oluline ka see, kas klient soovib saada infot või ootab partnerilt mingit tegevust (nt arsti vastuvõtule registreerimine). Telefonisuhtluse puhul mõjutab vestluse kulgu ka see, kas kasutatakse lauatelefoni või mobiiltelefoni.

Kliendi–ametniku suhtlust piiravad kaht tüüpi raamistikud: klienditeenindus ning avalik teenistus. Klienditeeninduse puhul maksab klient teenuse, info vms eest (nt takso tellimine, autoremont). Avaliku teenistuse puhul klient otseselt ei maksa (nt kiirabi). (Whalen jt 1988; Tracy 1997)

Need raamistikud erinevad kolmest aspektist:

- informatsiooni hulk ja iseloom, mida inimene peab esitama teenuse saamiseks,
- teenuse osutamise geograafiline ulatus,
- teenuse ooteaeg. (Tracy 1997: 324)

Klienditeeninduse raamistik eeldab, et klient peab teenuse saamiseks esitama üsna piiratud info (nt aadressi, kuhu takso tellitakse). Teenuse osutamise piirkond on kokkuleppeline, ooteaeg eeldatavasti lühike. Avaliku teenistuse raamistikus peab klient aga vastama paljudele küsimustele (häirekeskusesse helistades nt kirjeldama sündmuspaika, kannatanute seisundit jms). Teenust osutatakse ainult kindlas piirkonnas. Ooteaeg varieerub sõltuvalt väljakutsete hulgast ja probleemi tõsidusest. (Tracy 1997: 324–329)

Viimasena viidatud artiklites peetakse avaliku teenistuse all silmas eelkõige häirekeskust. Sellesse rühma kuuluvad aga ka nt sotsiaalabi, pensioniamet, polikliinik jms, kus ametnikul on õigus hinnata, kas klient vajab ja saab soovitud teenust (abi antakse ainult sel juhul, kui klient elab vastavas piirkonnas, tal on saatekiri vms).

3. Identifitseerimine tuttavate suhtluses

Esimene oluline tegur vestluste rühmitamisel on see, kas osalejad on võõrad või tuttavad. Dialoogikorpus on esindatud praktiliselt ainult võõrastevaheline suhtlus, tuttavate vestlusi on ainult registratuuri- ja taksodialoogide hulgas (vastavalt 14 ja 2

dialoogi). Tuttavate institutsionaalse suhtluse uurimiseks on suulise keele korpusest võetud 29 kolleegide telefonikõnet (eri helistajaid oli 8). Selles alamkorpuses on kolme liiki dialoogid:

- mõlemad osalejad on tööl (nt helistatakse ülikooli ühest õppetoolist teise);
- üks osaleja on tööl, teine kodus või neutraalsel territooriumil;
- mõlemad osalejad on kodus, suhtlus toimub väljaspool tööaega.

Omavahel suheldakse telefoni teel suhteliselt harva (vaid paaris kõnes sagedamini kui kord päevas, mõnede osalejate puhul on salvestatud vestlus aga isegi ainus nendevaheline telefonikõne aastate jooksul).

Tuttavate osalejate puhul on partneri identifitseerimiseks kaks võimalust: ta kas tutvustab ennast või tuntakse ära hääle järgi.

Kolleegide alamkorpuses tutvustas end 26 helistajat 29-st. Ennast tutvustati enamasti eesnimega (23 korda), eesnimi ja perekonnanimi öeldi kolmel korral. Kasutatakse järgmisi tutvustusvorme: *[nimi]*; *[nimi] siin*; *[nimi] on siin*; *[nimi] olen*; *[nimi] on siinpool*; *[nimi] veel kord*; *[nimi] helistab*; *siin [nimi] tülitab*; *mina olen [nimi]*. Kahe helistaja vestlusi on salvestatud rohkem (vastavalt 12 ja 9) ning neist ilmneb, et enamasti kasutab inimene sama vormelit: üks helistaja *[nimi] siin* ning teine *[nimi] on siin*. Vormelit muudetakse, kui situatsioonis või osalejate suhetes on midagi eripärast: osalejate suurem distant, telefonile vastab laps, helistatakse mitmes kord lühikese aja jooksul. Korduva helistamise puhul võib tutvustuse ka ära jätta. Ühes telefonikõnes jätab helistaja aga end tutvustamata, kui samal päeval ei ole telefonitsi suheldud ning see tekitab probleemi: vastaja küsib üle *`Anna räägib jah?*

Näites (1) helistab Jaana töölt teises allüksuses töötava kolleegi mobiilile. Varem on kokku lepitud, et V tuleb üht programmi tutvustama. Jaana on välja selgitanud, missugune aeg huvilistele sobib, ning helistab V-le, et kohtumise aeg välja pakkuda. V ei tutvusta end telefonile vastates. Helistaja tutvustab end eesnimega. V annab talle vooru tagasi fraasiga *no räägi*³ ning H asub kohe asja juurde.

(1) ((kutsung))

V: jaa? ma=kuulen

H: .hh `Jaana siin.

V: no=`räägi.

H: `vaata `koostöö tuleb `varem kui sa `arvatagi oskad.

Kolleegidevahelistes registratuuri- ja taksodialoogides on aga olukord hoopis teistsugune. Registratuuri helistanud 14 arstist-õest tutvustas end ainult kaks (*Liina olen*; *doktor Alpikann siin*), taksodispetšer ei tutvustanud end. Erinevalt eelmise rühma dialoogidest suheldakse omavahel telefoni teel pidevalt, päeva jooksul mitmeid kordi ning helistaja eeldab, et partner tunneb ta hääle järgi ära.

4. Identifitseerimine võõraste suhtluses

Edasi uurime ametivestlusi, milles osalejad on võõrad. Vaatleme situatsioonitüüpide kaupa, millal ja kuidas helistaja end tutvustab, millal aga mitte.

4.1. Helistaja jääb anonüümseks

Infosoovi puhul saab suhtlus enamasti toimuda nii, et vastaja ei tea, kellega ta räägib. Sellesse rühma kuuluvad kõned infotelefonile, bussiinfosse, kauplusesse, reisibüroosse (meie salvestatud dialoogides reisi ei broneerita, küsitakse ainult infot), osa kõnesid ameti- ja teenindusasutustesse, registratuuri. Infotelefoni dialoog algab tüüpiliselt nii:

(2) ((kutsung))

V: `info`telefon=Kersti=`tere.

H: ee tere `öelge=palun=mulle `linnapolikliiniku registra`tuuri (.)
`telefoninumber.

Vastaja tutvustab ennast, mõlemad osalejad tervitavad ning klient esitab oma soovi.

4.2. Helistaja identifitseerimine vestluse alguses

Kui vastaja peab kliendi heaks tegema midagi enam kui ainult infot andma, näiteks arsti vastuvõtutaja reserveerima, takso saatma, hotellitua broneerima, katust või arvutit remontima, vajab ta enamasti kliendi nime ja/või muid isikuandmeid. Niisuguseid dialooge oli järgmistes alamkorpustes: registratuur, aktiivne müük, riigiasutused, takso, teenindus. Teatud juhtudel tutvustab helistaja end aga ka infosoovi puhul. Helistaja võib end tutvustada kohe vestluse alguses, aga identifitseerimine võib toimuda ka hilisemas vestluses.

Vaatleme kõigepealt tutvustamist vestluse alguses.

Helistaja tutvustas end 28 dialoogis, neist 6 korda eraisikuna (5 korda eesnimi + perekonnanimi, 1 kord perekonnanimi) ning 22 korda institutsiooni esindajana (tutvustus sisaldab institutsiooni nime).

4.2.1. Identifitseerimine eraisikuna

Eraisikust helistaja ütles oma nime kolme tüüpi dialoogides.

1. Samasse asutusse helistatakse lühikese aja jooksul mitmendat korda. Helistaja eeldab, et vastaja teda veel mäletab. Tutvustus loob seose eelnenud ja käesoleva vestluse vahel ja võimaldab kohe asja juurde asuda. Tutvustusvormelis tuuakse välja, et varem on juba omavahel suheldud: **ma `elistan nüüd `tagasi** Jaana Kasesalu; ma: (.) `Kasearu Jaanika **jälle**. Näites (3) on klient kinni pannud hambaarsti konsultatsioonija ja helistab mõne minuti pärast uuesti, soovides ka raviaega.

(3) ((kutsung))

V: Kuku `kliinik tere

H: .hh tere ma `elistan nüd `tagasi Jaana Kasesalu. .hh `paneks ühe

(.) `raviaja ka `kinni=et (.) niigu `nii ma pean `tulema.

2. Helistaja nimi on vajalik selleks, et vastaja saaks andmebaasist infot otsida. Helistaja ei eelda, et vastaja teda tunneb või mäletab. Tegu on infodialoogidega: klient soovib teada saada, miks arvuti remont venib, mis kellaks ta on arsti juurde kirja pandud. Näites (4) helistab klient pensioniametisse eesmärgiga uurida, miks ta pole lastetoetust saanud.

(4) ((kutsung))

V: pensioniamet

H: tere.

V: tere?

H: .hh minu `nimi=ee on Jaana `Kasesalu, (.) .hh e ma=i=ole `saanud nüd `viimase `kuu laste `toetust.=

3. Kolmandat tüüpi esindab üks taksotellimisdialoog (5). Taksot tellides enamasti end ei tutvustata, kuid siin ütleb klient oma ees- ja perekonnanimi kohe esimeses voorus. Dispetšer küsib aga nime üle: ta vajab kliendi andmeid teatud järjekorras ning liiga vara saadud infost kasu ei ole.

(5) ((kutsung))

V: `Maria=Takso [tere.]

H: [tere,] paluks Indrek `Kiili: nimele Sa `voodi juurde `taksot.

V: mt .hh Sa `voodi nim- Sa `voodi juurde ja `kelle `nimele.

H: `Indrek `Kiil.

(0.5)

V: ja `tuleb `auto.=

4.2.2. Identifitseerimine institutsiooni esindajana

Helistaja tutvustab end institutsiooni esindajana järgmistel juhtudel.

1. Eesmärgiks on konfidentsiaalse info saamine. Näites (6) helistab sotsiaaltöötaja polikliiniku registratuuri ja soovib teada, kes on tema hoolealuse perearst. Igale küsijale niisugust infot ei anta.

(6) ((kutsung))

V: registratuur tere

H: e tere päevast. (.) .hh mina (.) tülitän teid X > linnavalitsuse < `sotsiaalabi osakonnast.=

V: =jaah?

(.)

H: nii:, mina paluksin seda `Kilu `Ninelli (.) hh `perearsti, kas tal `on peearst määratud.

Sellesse rühma kuuluvad ka telefonikõned raadiojaamast politseisse ja päästeametisse: küsitakse, mis viimase päeva jooksul on juhtunud.

2. Müügivestlustes on normiks, et helistaja tutvustab ennast, öeldes oma ees- ja perekonnanime ning firma nime. Tutvustus puudus kolmes dialoogis: kahel juhul küsis helistaja sekretärit soovitud vastajat, keda polnud kohal; vaid ühes telefonikõnes ei tutvustatud ennast soovitud vastajale. Kõikides muudes müügivestlustes tutvustas helistaja end nii sekretärile kui ka soovitud isikule.

(7) ((kutsung))

S: Vettassike tere päevast.

M: tere õhtust uuesti=Elar=Vigur `Tiritammest=siin:.(.)kas ma `saaksin=nüd rääkida härra `Hart Viilseniga.

S: ü::ks hetk vaatan.

(84.0) ((suunamine))

K: Viilsen kuuleb.

M: ee tere õhtust.

(0.2)

K: tere.

M: Elar Vigur `Tiritammest `siinpool.

K: jaa.

Siin rühmas võib vaadelda ka üht telefonikõnet ülikooli õppetoolist raamatukogusse, kus helistaja eesmärgiks on samuti kokkuleppe saavutamine.

3. Reisibüroost helistatakse kliendile tagasi. Kahes eelmises rühmas suhtlesid institutsiooni esindajad, selles rühmas aga institutsiooni esindaja ja eraisik.

(8) ((kutsung))

V: jaa `Ave kuuleb.

H: tere, olen `Ivi, elistan reisibüroost `Ilmareisid tagasi.

V: selge?

4. Normist kaldub kõrvale kaks telefonikõnet, kus ennast tutvustatakse infotelefonile helistades. Üks helistajatest on välismaalane (9), teisel juhul helistati valel numbril.

(9) ((kutsung))

V: info´ telefon=Kersti=tere

H: £ te-´te:re. mina helistan´Tartust ma olen x kiriku´preester isa {´Johan}, .hh ja minu küsimus´on=ed=mul=on´külalisi kes tahab minna(.)see´Saaremaale. £

Institutsiooni esindajana ennast tutvustades kasutati järgmisi vormeleid:

[nimi] [asutusest] siin

mina olen [nimi] [asutusest]

[nimi] [asutusest] on siinpool

[nimi] ja [asutus] on siinpool

[nimi] ja [asutus] siinpool

siin on [nimi] [asutusest]

mina olen [nimi] ja ma helistan [asutusest]

olen [nimi], helistan [asutusest] tagasi

mina [nimi] [asutusest]

mina tülitan teid [asutusest]

[asutusest] helistan

[asutusest] helistan teile
ma räägin [asutusest]
helistan teile [asutusest]
helistan [asutusest]

Müügivestlustes ütleb helistaja nii oma ees- kui perekonnanime ning firma nime. Eri helistajaid oli viis ning ka siin võib täheldada, et üks inimene kasutab sama vormelit: Muudes telefonikõnedes helistaja enamasti enda nime ei ütle, piirduakse institutsiooni nimega.

4.2.3. Identifitseerimine eelmise vestluse kaudu

Veel üks võimalus enda identifitseerimiseks on teha seda eelmise vestluse kaudu. Inimene ei ütle oma nime, vaid annab teada, et soovib jätkata eelmist (poollejäädud või lõpetatud) vestlust. Vahel on kõne katkenud, vahel on kokku lepitud, et klient helistab uuesti, vahel leiab klient, et saadud info ei olnud piisav. Niisuguseid identifitseerimisi oli 11 ja seda võtet võib põhimõtteliselt kasutada igat liiki vestlustes. Sel viisil saab kohe asja juurde minna, ei pea kordama eelmises vestluses räägitut.

Näites (10) helistab väikeste laste ema teist korda ujumisinstruktori mobiilile. Esimese vestluse ajal ei olnud instruktoril tundide graafikut käepärast ning ta palus viie minuti pärast uuesti helistada. Helistaja tutvustus on tema jaoks piisav ning ta hakkab kohe infot andma.

(10) ((kutsung))

V: jaa `Aino kuuleb

H: .hh ma helistan `uuesti selle `pereuju[mise] pärast.

V: [{}ähäh}]

(0.5)

V: .hh a ma `leidsin vaba aja niisuguse päris päris ea aja `nellapäeval kella `seitsmest `kaheksani. (0.5) kas teile `sobib=se.

Viide eelmisele vestlusele võib lisanduda ka nimega tutvustamisele: *`mina olen Anna Gavri `lenko ja ma helistan x `õppetoolist. (.) ma olen se `magistrant kes teil seal `käis. (0.8) ee `palvega (.) ee::=hh saada `luba /---/*

4.3. Helistaja identifitseerimine hilisemas vestluses

Hilisemas vestluses identifitseeriti helistaja 36 juhul. Seda tehti viit tüüpi dialoogides: registratuur (12), teenindus (9), reisibüroo (7), takso (6), ametiasutused (2). Siin on tegemist teistsuguse identifitseerimisega kui vestluse algul: helistajal ei ole vaja end tutvustada käesoleva suhtluse toimumiseks, vaid tema andmed on vajalikud mingiks edaspidiseks tegevuseks.

4.3.1. Identifitseerimissekventsi algatab vastaja

Peaaegu alati (34 korda) algatas identifitseerimissekventsi vastaja.

1. Vastaja kavatseb kliendile tagasi helistada, kuna vajab info hankimiseks aega. Siia rühma kuulus kuus telefonikõnet reisibüroosse ja üks kõne kindlustus-

firmasse. Kõigepealt küsitakse kliendi telefoninumbrit ning seejärel enamasti (kuid mitte alati) ka nime. Nimi ei ole seega tagasihelistamiseks tingimata vajalik info.

(11)

V: kas ma tohin teie telefoni`numbri võtta et=s me vaatame ned soodsamad pakkumised nagu `järgi ja võtame teiega kohe uuesti `ühendust.

H: mm võib `küll jah.

V: mhmh?

H: .hh null viis `üks,

V: jaa?

H: üeksa seitse seitse,

V: jaa?

H: kaks kolm `kaheksa.

V: ja teie `nimi on

H: Maiu. Maiu Rebane.

2. Kahes telefonikõnes ehitusfirmasse vajab vastaja kliendi andmeid, et edastada need firmajuhile, kes helistab tagasi. Ühes dialoogis küsiti ainult telefoninumbrit, näites (12) aga küsib venelannast vastaja kõigepealt kliendi telefoninumbrit, seejärel nime ja lõpuks asukohta, aga mitte täpset aadressi.

(12)

V: £ kas te teil ´on (.) ´telefon. £

H: on ´küll. (.)

V: £ oodake natuke. (32.0) ((otsib arvatavasti paberit ja pliiatsit)) jah palun. £

(.)

H: .hh e ´telefon on

V: £ jah? £

H: ´neli ´õheksa ´neli?

V: £ ahah? £

H: ´null ´seitse ´kaks.

V: £ ´null ´seitse ´kaks jah? ä (.) kes te ´olete. £

H: ja nimi on ´Saabas.

(.)

V: £ Saabas jah? £

H: jah. (.)

V: £ Saabas. a ku- kus te ´asute (0.8) see on ´Tartus jah? £ =

H: =e ´Tartus jah.

V: £ ahah? £

H: ´Anne[linnas.]

V: £ [jah] (.) ´Annelinnas jaa ahaa. (.) me: asume praegu: ´Seenevõsa tänaval he[heh] £

3. Kui ametnik peab kliendile vajalikku (isiklikku) infot andmebaasist nime järgi otsima, tutvustab helistaja end vestluse alguses (4) või küsitakse nime hiljem. Kahes telefonikõnes pensioniametisse helistaja end ei tutvustanud, probleemi lahendamiseks peab ametnik aga kliendi nime teadma.

4. Ühes telefonikõnes ehitusfirmasse lepitakse kokku, et ehitusmees tuleb remonti-
vajavat garaaži vaatama, seega on tal vaja teada kliendi aadressi ja nime.
5. Viimase, kõige suurema dialoogirühma märksõnadeks on registreerimine, bro-
neerimine, tellimine: arsti vastuvõtule registreerimine, takso tellimine, hotellitoa
broneerimine, piletite broneerimine, ajakirja tellimine, ortopeediliste jalatsite
tellimine. Erinevalt eelmiste rühmade dialoogidest ei küsita siin otseselt helis-
taja nime, vaid klient/patsient võib olla ka keegi teine. Kolme esimese rühma
dialoogide identifitseerimisküsimused sisaldasid vaid ühe erandiga sõna *teie*
(lisaks eespool toodud näidetele veel nt *ma võtaks teie telefoni `numbri; kas ma
saan teile `helistada; mis on teie (0.8) `ennem oli `perekonnanimi*), olles seega
otse helistajale suunatud.⁵ Selle rühma dialoogides aga teie-vormi ei kasutata.

Registruuris küsitakse näiteks: *kuidas `nimi on; ja nimi on; ja keda
ma `kirja panen*. Küsimused on vormistatud rutiinsete lühikeste lausete või
fraasidena, ei pöörduta isiklikult kliendi poole. Tänu sellele on suhtlus sujuv,
klient ei pea algatama parandust, kui ta ei soovi endale aega kinni panna.⁶ Vaid
ühes registruuridialoogis küsitakse just helistaja nime – patsient teatab, et
soovib oma kinnipandud aega tühistada (13).

(13) ((kutsung))

V: hambaravi=tere

H: .hh tere. .hh mul on doktor `Männioksa juurde (.) nelja `päeval: kella
`õheksaks `aeg. (.) .hh aga se=i= `sobi mulle enam.

(0.5)

V: mhmh? (1.0) ma tõmban `maha jah?

H: jah=

V: =aja. (2.8) ee te=olete:

H: Tiina Riidoja.

Järgmises näites küsib registruuritöötaja kõigepealt, kas aega soovitakse
täiskasvanule või lapsele. Teada saades, et patsient on laps, saab küsida: *kuidas
lapse nimi on*.

(14)

V: kas:=se=on `teile või `lapsele=

H: =lapsele.

V: * mhmh * (.) .hh siis on kaks-kümend=kuus ap`rill kell `kuusteist
kolm`kümend.

(0.8)

H: jah, sobib `küll.

V: ja kuidas lapse `nimi on.

H: Arvo Maasik.

Mõnes registruuris on identifitseerimiseks vajalik hoopis isikukood: *isikukood
palun; {nüüd} on vaja isiku `koodi*. Kui patsient isikukoodi ei tea, siis küsitakse
nime (15).

⁵ 4. rühma ainsas dialoogis ei olnud küll teie-vormis küsimusi, kuid klient viitab endale (*tulge mulle koju*) ning on selge,
et identifitseeritakse just helistaja.

⁶ Võrdluseks näide telefonikõnest tööhõiveametisse, kus ametnik eeldab, et klient räägib endast ning viimane algatab
paranduse, selgitades, et töötuks on jäänud abikaasa.

H: tere= `päevast. .hh kas `teie käest saab=õ `nõu ka küsida. .hh ed mida `ette võtta kui: `töökoht on: `koondatud.

V: .hh ä `kuna ta `koondati.=

H: =.hh e `esimesest no `vembrist.

V: mhmh? ja `kaua te seal `töötasite?

H: .hh ee > se=ei=ole `mina. < (.) ma küsin

V: [[no aga] kaua `tema.]

H: [.hh a-] `abikasa: (0.5) mingi `kahe (.) aasta `ringis.

(15) ((kutsung))

V: registratuur tere

H: tere, kas te paneksite {minu} (.) doktor `Jaanimäe juurde e a a (.)
`järje[korda.]

V: [jaah] e (.) öelge oma `isikukood=h.

(.)

H: hh oi hh. (.) ma ei [`tea]

V: [või] `olete siin `ennem ka käinud arsti juures.

H: e=jah, olen `küll [käinud.]

V: [kuidas] `nimi on. [{-}]

H: [`Ilse] Karask.

Ajakirjatellimuse vastuvõtja küsis *kelle nimele tellime*, hotelli administraator ja kelle *`nimele selle võib `panna*. Taksodispetšeri tüüpiline küsimus on ja *`kelle `nimele*. Taksodialoogid erinevad muudest selle rühma dialoogidest selle poolest, et klient ütles enamasti ainult enda eesnime, umbes pooltes vestlustes aga kliendi nime ei küsitudki.

Näites (16) on ainuke identifitseerimisküsimus, mis vastust ei saa. Helistaja ei soovi tegelikult reisi broneerida.

(16)

V: ee ja `kuidas oleks reisijate `nimed? (1.2) et sis: ma saaksin küsida täpset hinna `pakkumist.

(.)

H: aa.

(1.8)

V: `nimed ka teada ja sünni `ajad.

(0.8)

H: e mina lihtsalt `uurin nendele, (.) nad on `välismaalased.

4.3.2. Identifitseerimissekventsi algatab helistaja

Helistaja algatas identifitseerimissekventsi hilisemas vestluses ainult kahes dialoogis. Ühel juhul on kokku lepitud, et ehitusmees tuleb garaaži vaatama ning klient ütleb oma nime ja aadressi. Näites (17) soovib klient reserveerida ujumistunni aega ja ütleb oma nime.

(17)

H: =jah (0.8) aga [aga] `äkki võiks `see esmaspäev mis on kuus `teist.

V: [tulet:e]

(0.5)

V: mh mh

(0.5)

H: kuue `teistkümnendal kell `kaheksa sis.

V: mh mh

H: Maasik.

(0.5)

V: jah.

Kokkuvõte

Kas helistaja ennast institutsionaalses suhtluses tutvustab või mitte, sõltub kõigepealt sellest, kas osalejad on tuttavad või võõrad. Tuttavad osalejad identifitseerivad teineteise igal juhul vestluse alguses. Kuna partner on võimalik hääle järgi ära tunda, siis pole tutvustamine alati vajalik. Hääle järgi äratundmisele loodetakse siiski vaid juhul, kui omavahel suheldakse iga päev mitmeid kordi (analüüsitud materjali hulgas polikliiniku arstide-õdede kõned registratuuri, taksodispetšerite kõned). Harvema kontakti puhul peab helistaja vajalikuks end tutvustada, öeldes eesnime. Suurema distantsi puhul öeldakse ees- ja perekonnanimi.

Võõraste osalejate puhul ei ole hääle järgi äratundmine võimalik. Tutvustamine või mittetutvustamine oleneb eeskätt sellest, kas helistaja soovib infot või ootab partnerilt mingit tegevust. Infooovi puhul jääb helistaja enamasti anonüümseks, suhtlus saab toimuda nii, et infoandja ei tea, kellega räägib. Erandiks on juhtumid, kui helistaja vajab konfidentsiaalset infot – siis tutvustab ta end institutsiooni esindajana. Erasisikust klient tutvustab end siis, kui ta vajab isiklikku infot (nt pensioniametist), mida ametnik peab oma andmebaasist otsima.

Kui helistaja taotleb vastuvõtja mingit tegevust, siis peab ta end identifitseerima. Selleks on kaks võimalust: helistaja tutvustab ennast vestluse alguses või küsib vastaja kliendi nime hilisemas vestluses. Kui helistajaks on institutsiooni esindaja, siis tutvustab ta ennast kohe vestluse sissejuhatuses. Erasisikust helistaja puhul kasutatakse teist võimalust märgatavalt sagedamini: nime küsitakse siis, kui ollakse mingile kokkuleppele jõudnud (nt tullakse remontivajavat garaaži vaatama), vastaja kavatseb kliendile tagasi helistada (nt reisikonsultant vajab aega hankimiseks) või peab ametnik kliendi kohta mingi nn blanketi täitma (nt polikliiniku registratuuris küsitakse teatud asju kindlas järjekorras).

Identifitseerimist on seega mitut liiki. Tuttavate suhtlejate puhul on tegu äratundmisega (kas hääle või nime järgi). Võõraste suhtlejate puhul toimub identifitseerimine esiteks käesoleva vestluse jaoks. Siin on omakorda kaks alaliiki:

- osalejad ei saa suhelda teadmata, kes on vestluspartner (nt müügivestlus);
- identifitseerimine kui asjade kokkuviiimine: helistaja identifitseerib end eelmise vestluse kaudu kas nimepidi või ainult eelmisele vestlusele viidates. Suhtlemine ilma identifitseerimata on võimalik, aga siis tuleks kogu asjaajamist otsast peale alata. Teist tüüpi identifitseerimine on vajalik mingi edaspidise tegevuse jaoks. Näiteks registratuuridialoogid saavad toimuda nii, et ametnik ei tea helistaja nime, aja võib kinni panna ka kellelegi teisele. Lisaks nimele võivad identifitseerimiseks vajalikud olla ka muud andmed: isikukood, telefoninumber, aadress.

Tabelis 2 on näidatud, missuguseid identifitseerimisvõimalusi eri tüüpi vestlustes kasutatakse. (Identifitseerimine eelmise vestluse kaudu ilma oma nime ütlemata on kõrvale jäetud, kuna seda võimalust võib põhimõtteliselt kasutada igat liiki dialoogides.)

Tabel 2. Helistaja identifitseerimine võõrastevahelistes institutsionaalsetes dialoogides (andmed esitatud protsentides)

Allkorpus	Identifitseerimine vestluse alguses	Identifitseerimine hilisemas vestluses	Ei identifitseerita
Infotelefon	0,5	–	99,5
Bussiinfo	–	–	100
Registruur	6	19	75
Reisibüroo	7	11	82
Aktiivne müük	79	–	21
Takso	5	27	68
Riigiasutused	31	15	54
Kauplus	–	–	100
Teenindus	5	21	74

Kirjandus

- Drew, Paul; Heritage, John 1992. Analyzing talk at work: An introduction. – Paul Drew, John Heritage (Eds.). *Talk at Work. Interaction in Institutional Settings*. Cambridge: Cambridge University Press, 3–65.
- Hennoste, Tiit 2003. Suulise eesti keele uurimine: korpus. – *Keel ja Kirjandus* 7, 481–500.
- Heritage, John 1997. Conversation analysis and institutional talk: Analysing data. – David Silverman (Ed.). *Qualitative Research: Theory, Method and Practice*. London: Sage, 161–182.
- Hopper, Robert 1992. *Telephone Conversation*. Bloomington: Indiana University Press.
- Hopper, Robert; Doany, Nada; Johnson, Michael; Drummond, Kent 1991. Universals and particulars in telephone openings. – *Research on Language and Social Interaction* 24, 369–387.
- Hopper, Robert; Drummond, Kent 1992. Accomplishing interpersonal relationship: The telephone openings of strangers and intimates. – *Western Journal of Communication* 56 (3), 185–199.
- Kasterpalu, Riina 2006. Vastus enesetutvustusele telefonikõnedes. Käsikiri autori valduses.
- Rääbis, Andriela 1999. Telefonivestluse sissejuhatus. Magistritöö TÜ eesti keele õppetoolis.
- Rääbis, Andriela 2000. Telefonivestluse sissejuhatus. – *Keel ja Kirjandus* 6, 409–418.
- Rääbis, Andriela 2006. Infodialoogi algusrituaalid. – Mare Koit, Renate Pajusalu, Haldur Õim (toim.). *Keel ja arvuti*. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 6. Tartu: Tartu Ülikooli Kirjastus, 143–155.
- Schegloff, Emanuel 1986. The routine as achievement. – *Human Studies* 9, 111–152.
- Tartu Ülikooli eesti dialoogikorpus. <http://math.ut.ee/~koit/Dialoog/EDiC> (27.07.2006).
- Tartu Ülikooli eesti suulise keele korpus. <http://www.cl.ut.ee/suuline/> (27.07.2006).
- Tracy, Karen 1997. Interactional trouble in emergency service requests: A problem of frames. – *Research on Language and Social Interaction* 30 (4), 315–343.
- Whalen, Marilyn R.; Zimmerman, Don H. 1987. Sequential and institutional contexts in calls for help. – *Social Psychology Quarterly* 50, 172–185.
- Whalen, Jack; Zimmerman, Don H.; Whalen, Marilyn R. 1988. When words fail: A single case analysis. – *Social Problems* 35, 335–362.

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

CALLER IDENTIFICATION IN INSTITUTIONAL TELEPHONE CONVERSATION

Andriela Rääbis

University of Tartu

The analysis is based on the transcriptions of 684 institutional phone conversations in Estonian (91 289 tokens in total). Ten situational groups are represented : 413 directory inquiries, 62 calls to travel agencies, 65 calls to outpatients' offices, 42 calls to service encounters, 14 telemarketing calls, 14 calls to bus stations, 13 calls to civil services, 10 calls to shops, 22 calls for taxi, 29 calls of university colleagues. Communication takes place mostly between strangers, only in 45 dialogues the participants are acquainted. Methodologically the study is based on conversation analysis.

If a customer needs some general information, he remains anonymous. Only if requested information is personal or confidential, the customer introduces himself.

If a customer expects an action from the addressee (e.g. to book a doctor's appointment, to send a taxi, to participate in a training course), he introduces himself at the beginning of the conversation, otherwise the operator asks the customer's data (name, ID code, phone number etc.) later. The first option is characteristic of telemarketing calls: the caller represents an institution and tells his first name, family name and the name of the institution at the beginning of the conversation. An operator asks the caller's data 1) after the interlocutors have come to an agreement (e.g. the builder called has agreed to repair the garage); 2) the operator needs more time for finding information and is going to call the customer back (e.g. in calls to travel agencies); 3) the operator fills out a blank (e.g. in the outpatients' office).

There are several types of identification. In some cases identification is necessary for the current conversation: the interlocutors cannot communicate unless the partner is identified. In other cases identification takes place for a consecutive action, e.g. a receptionist needs to know a customer's name not in order to talk with him but to reserve a room for him.

Keywords: conversation analysis, interaction in institutional settings, telephone conversation, spoken Estonian

THE ROLE OF A CHILD'S GENDER IN LANGUAGE ACQUISITION

Ineta Savickienė, Violeta Kalėdaitė

Abstract. The aim of the present article is to examine the relationship between the process of language acquisition and gender. It analyzes a longitudinal corpus of one girl, Rūta, covering the period from 1;7 to 2;6. The corpus consists of 35 hours of recordings. The recorded speech is transcribed according to the requirements of CHILDES (MacWhinney, Snow 1990). A special attention in the article is paid to testing the hypothesis which was supported by empirical evidence from other languages to the effect that the unmarked member, i.e. nouns of masculine gender, is acquired earlier than the marked member of the opposition (i.e. feminine nouns).

Keywords: language development, early stage, longitudinal corpus, nouns, grammatical gender, morphosyntactic properties, inflectional patterns, agreement, Lithuanian

1. Introduction

Gender is not something we are born with, and not something we have, but something we do or perform (cf. West, Zimmerman 1987). The famous statement by Simone de Beauvoir to the effect that women are not born but they are made is applicable to the making of men as well. This process begins even before a baby is born (Eckert, McConnell-Ginet 2003). Newborn babies cannot easily be identified as girls or boys if they are dressed identically. Therefore, in many cultures, babies are dressed in ways to make their gender clear. Gender assigning process takes different forms, starting with the tradition of providing pink caps for girls and blue caps for boys. In addition to the visual, color-coding sign, another early attribution of gender is the linguistic event of naming the baby.

Moreover, from early childhood girls and boys are interpreted differently, and interacted with differently. People usually behave more gently with baby-girls and more playfully with baby-boys. Parents and even strangers talk to them

differently. Thus, while addressing girls more diminutives (*doggie, sweetie* etc.) are used, whereas more direct prohibitives (*don't do that!*) occur while talking to boys (Gleason et al. 1994; Protassova, Voeikova 2007; Korecky-Kröll, Dressler 2007).

Children can see gender distinctions everywhere and start following gender-imposed patterns of behavior very early, thus continuing to learn to be a girl or a boy.

As for linguistic aspects, there is enough evidence to claim that girls are usually more advanced in language development than boys (it is obvious, though, that individual differences exist). Girls begin to talk earlier; they articulate better and acquire a more extensive vocabulary than boys of the same age. Studies of verbal ability have shown that girls and women surpass boys and men in verbal fluency, correct language usage, sentence complexity, grammatical structure, spelling, and articulation (Karmiloff, Karmiloff-Smith 2002).

What does this information tell us about the course of language development at an early stage? Does this mean that for the boys it will be more difficult to grasp and acquire the structure of their mother tongue? Extensive research in the field allows us to state, however, that even though in some aspects of language acquisition boys might be considered to lag behind or experience more difficulties than girls, they, as a rule, reach the same level of linguistic competence as girls already during the preschool years (Karmiloff, Karmiloff-Smith 2002).

The article explores the relationship between the process of language acquisition and gender. It concentrates on the following research questions: Does the child's gender help her to acquire the structure of the native language faster and easier? How do children acquire grammatical gender of the target language?

2. Grammatical gender

According to Corbett (1999: 1), "gender is the most puzzling of the grammatical categories, and it is a topic which interests non-linguists as well as linguists and it becomes more fascinating the more it is investigated". Gender is a category of morphosyntactic properties which distinguish classes of nominal lexemes: for each such class of lexemes, there is a distinct set of inflectional markings for agreeing words (Stump 1998). A language may have two or more such classes or genders. For a noun to belong to a particular declension class often implies that it also belongs to a particular gender. The classification very often corresponds to a real world distinction of sex. Correlations may exist between the meanings of nouns and the genders to which they belong. Thus, in Lithuanian nouns which refer to females are generally of feminine gender. Correlations of this sort are, however, never perfect; that is, membership in a particular gender is most often a matter of arbitrary stipulation (Stump 1998; Corbett 1999).

Lithuanian is one of the many languages which possess grammatical gender. In Lithuanian we distinguish three criteria of gender marking: **syntagmatic**, or form agreement; formal, which is based on the form of stems; and **semantic**, which implies natural sex distinctions for nouns denoting people and animals. The binary opposition masculine/feminine reflects the fact that nouns can behave in different

ways when it comes to the agreement of adjectives, replacement by a pronoun, and inflectional patterns.

There are a lot of studies devoted to the analysis of the categories **markedness** and **unmarkedness** (Greenberg 1966; Dressler et al. 1987, 1996). The category of markedness is often referred to in the discussion of the phenomenon of **complexity**. As regards semantic markedness, it is the unmarked member of the opposition that is always less complex due to the absence of a distinctive semantic property. Therefore, the unmarked member is always less complex semantically than the marked one, whose distinct semantic property is clearly defined. Numerous studies posit that in child language acquisition less complex, or unmarked, items should be acquired more easily (see Zangl 1997).

In his discussion of gender marking, Greenberg (1966) claims that the property 'masculine' is unmarked, whereas the feature 'feminine' is marked. This type of markedness is not a universal feature of all languages; however, in Lithuanian the masculine gender of nouns is treated as unmarked (Girdenis 1981; Valeckienė 1998). With this in mind, it can be posited that children will acquire items having the property 'masculine' earlier than those which possess the property 'feminine'. This hypothesis will be given some consideration in the section below.

3. A survey of gender acquisition research

Literature in the field distinguishes two main criteria operating in gender distinction: formal and semantic (Mills 1986; Corbett 1999). Operation of semantic factors is attributed to extralinguistic forces¹ (the theory of natural gender). According to this approach (see Mulford 1983, 1985), in the process of language acquisition children primarily relate gender to biological gender (sex).

An alternative approach, however, maintains that gender distinction is part of a linguistic system. Proponents of this hypothesis believe that children, while attributing a noun to a particular gender, do not entirely rely on extralinguistic reality; they also draw on information coded in the linguistic context in which a word is located. Therefore, a child is able to recognize a noun together with its ending, categorize it as belonging to one or another gender, and to impose an agreement with its modifiers, such as pronouns, adjectives, articles, etc. (cf. Karmilloff-Smith 1979; Levy 1983, 1988; Maratsos, Chalkey 1980; Maratsos 1988).

Research shows that children are capable of distinguishing differences in biological sex at around the age of 2;6². It is at this particular period that they become aware of the existence of two sexes: in a picture, they recognize which person is a male and which is a female, and attribute himself/herself to one of the two sexes (Fagot et al. 1986). The category of gender becomes an issue in the process of language acquisition when a child finds out that sex is an inherent property and does not change even if clothes are changed (Gelman et al. 1986). It appears that already in their early childhood children become aware of the existence of certain rules related to natural gender.

Different methods of research have been applied in large-scale studies of the issue: some scholars rely on the advantages offered by **longitudinal**³ studies, others resort to an **experimental** methodology.

¹ The terminology used is taken from M.Perez-Pereira (1991): *extralinguistic* indicates semantic or pragmatic level, whereas *intralinguistic* refers to phonological, morphological and syntactic levels.

² The numerals show the child's age (2;6 – 2 years 6 months).

³ The term is usually used in the field of child language acquisition or developmental psychology. The longitudinal study follows the course of language acquisition in a single child or group over a period of time (Crystal 1991).

Data gained from longitudinal studies have shown that, for example, Spanish children acquire formal markers of gender and agreement before the age of 4 (Perez-Pereira 1991; Clark 1985). Thus, the subject of observation first became aware of gender affixes (a formal rule) and gender agreement (a syntactic rule). Errors which still occur at this stage point to the fact that children try to discover certain regularities in gender agreement, as in the Spanish example of *mota rota* used instead of the correct form *moto roto* 'broken bike'. Gender markers for nouns and adjectives are acquired approximately at the same time, around 2;8; however, errors in gender agreement persist for a much longer time.

On the other hand, studies on child language in German based on experimental methodology (Mills 1986; MacWhinney 1978) have demonstrated that equal importance should be attached to both phonological and syntactic rules. And yet Mills (1986) makes a strong claim to the effect that the semantic factor should be treated with utmost importance as well.

The subsequent section takes up the issue of how category of gender is acquired by Lithuanian children. The study is based on a longitudinal corpus of one girl, Rūta⁴. When needed, examples from longitudinal and experimental data of other Lithuanian girl Monika⁵ will be adduced. Special attention will be paid to pointing out which rules – intralinguistic or extralinguistic – are dominant in the process of attributing a noun to a particular gender.

4. Acquisition of the category of gender in Lithuanian

Our working hypothesis is that the fewer exceptions a certain rule or pattern has, the quicker it is acquired by children. This formulation also implies that a child is capable of noting frequencies of certain categories, their differences and similarities, and that a child possesses the necessary abilities for learning the rules (Slobin 1985).

Most researchers claim that during the early stages of language acquisition it is problematic for a child to distinguish between genders because the category of gender is a problematic issue in itself. The fact that a noun ending marks several categories, those of case, number and gender, makes the process of acquisition a complex matter. We will test this claim against the data from Lithuanian. In our data, words which have distinct formal gender markers already appear in early recordings, at 1;7. The frequency of nouns marked masculine or feminine is displayed in Table 1 and Table 2.

Table 1. The distribution of masculine and feminine nouns in Rūta's speech (1;7–2;5)

	1;7	1;8	1;9	1;10	1;11	2;0	2;1	2;2	2;3	2;4	2;5	Total	Total
FEM	18	131	426	387	310	396	359	369	337	422	374	3529	40%
MASC	18	167	469	698	454	479	528	622	622	643	487	5187	60%

⁴ Rūta is a first-born and only child of a middle-class family living in Vilnius. Her speech was recorded in natural everyday situations by her mother, a philologist. Recordings were made three or four times per week; they lasted about fifteen minutes each. For the present study we have chosen to analyse Rūta's speech covering the period from 1;7 to 2;6. The corpus consists of 35 hours of recordings. The recorded speech was transcribed by the girl's mother according to the requirements of CHILDES, or Child Language Data Exchange System (MacWhinney & Snow 1990).

⁵ Monika is also a first-born and only child of a middle-class family living in Kaunas. The corpus consists of diary remarks and almost 45 hours of recordings (transcribed and only partly coded according to CHILDES; therefore, we were not able to provide the statistical data).

Table 2. The distribution of masculine and feminine nouns in Mother's speech (1;7–2;5)

	1;7	1;8	1;9	1;10	1;11	2;0	2;1	2;2	2;3	2;4	2;5	Total	Total
FEM	68	282	789	707	308	553	408	410	436	387	477	4825	45%
MASC	100	391	827	1096	458	583	430	529	561	500	452	5927	55%

The data show that masculine and feminine nouns in Rūta's speech appear in equal numbers only during the 1;7 period (see Table 1). Starting with 1;8 and up to the period of 2;6 masculine nouns are more frequent. The same tendency is noticed in Mother's speech: during the entire period of observation masculine nouns are more common than feminine nouns. The 1;10 period is exceptional in this respect: masculine nouns are especially dominant, and the same tendency is observed in Rūta's speech.

This change can be explained by the expansion of Rūta's vocabulary, because the greatest number of new words belongs to the masculine gender.

The number of masculine nouns was increased by the frequent occurrence of certain nouns belonging to a discrete lexico-semantic groups, such as 'animals' or 'toys'. These cases cover the following instances: (1) when the sex of animals is not distinguished at all, and the masculine noun refers to both sexes, e.g. *dramblys* 'elephant', *kupranugaris* 'camel', *uodas* 'gnat', *ežys* 'hedgehog', *pingvinas* 'penguin'; (2) the young animals are referred to by masculine nouns, as in *šuniukas* 'puppy', *ančiukas* 'duckling', *viščiukas* 'chicken'; (3) there are some feminine nouns which are used to refer to animals of both sexes when the sex of animals is not distinguished at all, for example, *varlė* 'frog', *beždžionė* 'monkey', *kengūra* 'kangaroo'. However, when these nouns are used as diminutives (see Savickienė 1998, 2003), there is a change in gender marking: nouns become masculine, as in *varliukas* 'frog:DIM', *beždžioniukas* 'monkey:DIM', *kengūriukas* 'kangaroo:DIM' (this is the case for Rūta's speech).

Even in the period of vocabulary building the masculine nouns have dominated, Rūta has acquired a correct way of using feminine noun forms earlier than masculine nouns. This state of affairs could have been influenced by an extremely frequent repetition of the names *Rūta* or its diminutive form *Rūtytė* along with *mama* or *mamytė* 'mother'. Due to this fact Rūta has acquired the feminine endings *-ė* and *-a* (nominative singular) first.

The correct usage of feminine nouns during the early period of language acquisition in Rūta's case could be explained within the framework of a hypothesis which relates the early and unproblematic acquisition of certain grammatical categories (e.g. of gender or case) to the **child's gender**.

It appears that a child's biological gender can help him/her to acquire a certain grammatical category or form. The hypothesis is corroborated by research data from Latvian (Rūķe-Draviņa 1973), English (Mills 1986) and Greek (Christofidou, Stephany 1997): thus girls acquire the feminine gender first, whereas boys acquire the masculine gender first. It is important to stress that our data are compatible with this hypothesis as well. To what extent this hypothesis is valid for language data of other children still needs to be researched. In any case, the hypothesis does not look entirely unreliable. Parents, when communicating with their children, very often address them by their names. In Lithuanian, names have fixed endings:

-a and -ė for female names (-ė occurs more frequently due to the frequent use of diminutive forms) whereas -us, -is or -as (the latter two are more frequent because of diminutive forms) mark male names. It stands to reason then that children are able to recognize words with the endings, which are mentioned repeatedly.

This is exactly the pattern noticed in Rūta's speech: first of all the girl acquires the correct use of nouns which end in -ė. Moreover, it is this particular type of nouns that is most commonly used by Rūta in the early period (1;7–1;9) of language acquisition.

As has been noted above, the semantic rule of gender marking is based on sex distinction. In Lithuanian, the two classes of nouns – 'persons' and 'animals' – can exhibit two patterns of gender marking based on the biological category 'sex', namely, by different roots (lexical) and by affixes (derivational). No difficulties in using this type of nouns were noted in Rūta's speech. The girl is perfectly able to distinguish the referent she needs in the pair of masculine/feminine nouns, which are marked by different roots and different inflections.

The category of gender in Lithuanian is expressed by a two-member opposition, and this could be the reason why there were only few errors attested during the entire period of observation (1;7–2;5). An error occurs when one gender is used instead of the other. The largest number of errors was noticed with singular masculine nouns marked genitive, e.g. *šuniukos* (instead of *šuniuko*) 'dog:DIM', *berniukos* (*berniuko*) 'boy', *Lines* (*Lino*) (a boy's name), *lietios* (*lietaus*) 'rain'; all these errors occurred when a feminine ending of the genitive case was used instead of the correct ending of a masculine genitive. With respect to gender errors in other cases, even though very rare, most frequently feminine endings were used instead of masculine. Examples of such instances are: *suoskė*, *sostė* (*sostas*) 'throne:SG:NOM', *vezimėlė* (*vežimelis*) 'pram:DIM:SG:NOM', *aicesė* (*laikraštį*) 'newspaper:SG:ACC', *nuokatas* (*nuotrauka*) 'photo:PL:ACC'. Instances where masculine endings were used instead of feminine are very rare, e.g., *zikukas* (*žirklutės*) 'scissors:DIM:PL:NOM', *nuokatai* (*nuotraukos*) 'photo:PL:NOM', etc. These errors are noted during the 1;10–2;3 period.

It has to be pointed out that, starting from 2;0 onwards, such errors are extremely rare. Incorrect forms occur predominantly in those situations when a word is a new item in the girl's vocabulary, e.g. *sostas* 'throne' or *laikraštis* 'newspaper'. It is interesting to note in this connection that the girl usually overgeneralized feminine gender endings; the opposite phenomenon was noticed only in very few cases.

However, one fact deserves to be singled out – up to the end of the observation period dative singular endings of masculine nouns were substituted with endings of feminine nouns. Psycholinguists claim that the acquisition of the meaning of a grammatical category depends to a great extent on the formally marked variables of linguistic expression (Stephany 1997). Even when choosing a correct grammatical case, a child can construe an incorrect form of the case, i.e. she can choose a paradigmatically wrong ending. In our research data the formal means available for construing different variables is not great and consists of just two items: the masculine singular dative *-(i)ui* and the feminine singular dative *-(i)ai*. The fact that the girl has used the feminine gender marker instead of the correct masculine form for quite a long time is due to a more frequent appearance of the feminine marker in Rūta's and her mother's speech (see Savickienė 2001, 2003).

The dative case example also enables us to posit the following: the semantic criterion of natural gender does not play a decisive role in gender acquisition. The fact that the girl marks masculine nouns (referents of which in the real life are male persons) by feminine endings strongly suggests that a child will not treat physical characteristics of a referent as most important. Illustrative examples include the dative singular *Adomai* (*Adomui*), *berniukai* (*berniukai*) 'boy', *Linai* (*Linui*), *Pauliukai* (*Pauliukai*) and the genitive singular forms *berniukos* (*berniuko*), *Lines* (*Lino*).

Setting the issue of masculine singular dative aside, we can posit that the category of gender is acquired by Rūta rather early, before the age of 2;0. The fact that masculine and feminine singular dative endings are sometimes mixed does not mean that the category of gender has not been acquired yet.

Levy (1983: 91) claims that children, when acquiring the system of gender, do not fully exhaust the cognitive possibilities of the concept. Building her claims on data obtained from several languages, she suggests that the notion of gender cannot be clearly defined and, consequently, that an early acquisition of a gender system is not possible, especially in languages where formal rules, both phonetic and morphological, exist. MacWhinney's (1978) data support this claim: German children resort to the rule of natural gender only occasionally.

Lithuanian data support the statement that in the period of early language acquisition Lithuanian children would rely on formal (morphological) criteria in the process of grammatical gender acquisition.

The primacy of formal rules is apparent in the Polish language, and this is the reason why gender distinctions are acquired rather early (Smoczyńska 1985). Gender acquisition research carried out in Russian (another Slavic language) demonstrates that children already at an early age become aware of the formal properties of a word, and this happens before they are able to grasp the meaning of the word itself (Popova 1973). However, acquiring the system of gender distinctions at a young age is hindered by several exceptions, such as the addition of feminine endings to masculine diminutive nouns: *mishka* 'bear:DIM', *zaika* 'hare:DIM' (Smoczyńska 1985). A similar situation occurs in Latvian. Rūke-Draviņa (1973), in her discussion of the acquisition of gender of Latvian nouns, adjectives and pronouns, asserts the primacy of formal aspects over semantic rules.

The discussion can be summarized in the form of the following generalizations. Formal patterns will be acquired first in those languages where they operate clearly and do not have many exceptions. However, research carried out in a number of languages suggests that it is not possible to prove that the semantic rule is acquired earlier than the formal one and vice versa. It stands to reason to argue that both rules go hand in hand.

Comparison of data on gender acquisition obtained from several languages supports Slobin's (1985) claim that linguistic systems, in which two-member oppositions are marked formally by suffixes, are acquired quicker than those which do not possess such marking. In fact, children which have to acquire a language with a three-gender system or in which morphological markers denote several categories simultaneously, e.g. gender, number and case, need more time to master the category of gender.

However, data from Polish (Smoczyńska 1985) call into question the generalizations presented above. Formal marking of gender in Polish is acquired at an

earlier stage than is the case in Russian, even though structurally the two languages are very similar. Thus Smoczyńska (1985: 645) claims that there are two factors of child directed speech, which create difficulties for Russian children to master the category of gender easily: 'these are (1) that while Russian unstressed *-o* (one typical ending for neuter nominative singular nouns) is pronounced as *-a* (typical ending for feminine nominative singular forms), producing a confusing situation, this is not the case with Polish form; and (2) that Russian has both diminutive masculine forms which end in *-a* (and thus are declined as feminine nouns) and boy's first names which also end in *-a*. Both facts tend to increase the inconsistency of the Russian gender system in relation to the Polish'.

On the other hand, data from Polish corroborate the fact that children are capable of acquiring this relatively complex category by mastering the rules which are based exclusively on formal criteria (Smoczyńska 1985: 646).

Results obtained from data on Lithuanian can be considered as one more exception, which supports the hypothesis of a fairly quick and easy acquisition of the category of gender. Despite the fact that in Lithuanian gender marking is carried out through a two-member opposition, it can present certain difficulties in the sense that one marker is used to denote several categories, i.e. those of gender, case and number.

The analysis of the results pertaining to data on gender acquisition in Rūta's speech does not allow us to assert that Lithuanian data support the hypothesis that the unmarked member of the opposition is acquired earlier than the marked one. The unmarked member in Rūta's speech is the masculine gender. Since masculine nouns are unmarked, they occur in speech more often. However, feminine nouns have been acquired first, and, consequently, it can be claimed that the formal marking of the feminine gender has been acquired earlier. One explanation for this linguistic fact is the frequent occurrence of the formal marking of the feminine gender present in the girl's name, that is, the inflectional ending *-ė*. Some endings of masculine nouns, especially those of the nominative singular, exhibit a tendency to be substituted with endings of feminine paradigms.

The transparency and a frequent use of morphological endings accelerate the process of gender distinctions. First, the feminine endings *-ė* and *-a* were acquired, which were followed by the masculine *-as*, *-is* and *-us*.

In this study we were more concerned with the process of acquiring the category of gender of nouns. Consequently, we do not have enough data for analyzing the syntagmatic (syntactic) criterion, which requires several word classes to stand in gender agreement. Such data could throw more light on the essential properties of gender acquisition. Several examples with personal pronouns point to the fact that during the 2;02;5 period correct gender forms of pronouns replaced the respective nouns. However, syntactic rules of agreement are acquired later than morphonological rules (Perez-Pereira 1991).

In natural adult spoken language the grammatical case of address is the vocative case; other contexts usually require the use of personal pronouns. One of the most interesting and universal phenomena in child language is the so-called **self-reference**. Specifically, this term denotes a child referring to herself/himself by her/his own name, e.g., *Rūtytė sėdi* 'Rūtytė is sitting'; *Rūtytės lėlytė* 'Rūtytė's doll:DIM' (uttered by Rūta); *mamytė duos Rūtytei* 'Mother will give (something)

to Rūtytė' (Mother's utterance). The type of linguistic model under discussion is usually introduced by mothers. This is determined by the fact that very often mothers address their children by names. On the other hand, referring to themselves, mothers use the word *mama* 'mother', and not the personal pronoun *aš* 'I'. Data from different languages suggests that self-references are very common in both adult and child language up to the age of three (Savickienė 2003).

Two longitudinal corpora of data (Rūta's and her mother's and Monika's and her mother's) show a great number of self/other-referential items. Both girls, Rūta and Monika, did not use personal pronouns for quite a long time; they actually became more common only around the period of 2;2 in Rūta's speech and around 2;2–2;3 in that of Monika's.

Both girls are addressed by their names, most often in their diminutive form (*Rūta, Rūtytė, Monika, Monikutė*) and this makes up a large number of diminutive tokens. It is characteristic of children to use their own name very frequently, especially in diminutive form. Children can choose several diminutive forms of their names as a standard variant (e.g., *Rūtelė, Rūtytė, Monikutė, Monikėlė*); such forms can be used by family members and friends as well. However, in our data there are instances when diverse diminutive name forms were used, such as *Rūtuliukas, Rūtužėlis* (gender shift to masculine) uttered by the mother; *Monikėlis, Monikutis, vaikas* 'child' uttered both by Monika and her mother.

Rūta's data shows that the girl never uses masculine self-reference form due to the fact that she is aware of her own gender (semantic rule) and has mastered the rule of grammatical gender assignment as well.

In contrast to Rūta's data, in the cases of self-reference⁶ Monika used both genders, masculine and feminine. Consider the following examples:

- (1) MONIKA: Pabėgo Monikutis.
'Monika:DIM:MASC:NOM ran away.'
- (2) MOTHER: Ak, pabėgo Monikutis.
'Well, Monika:DIM:MASC:NOM ran away.'
- (3) MOTHER: Aš ją pagausiu, pagausiu Monikutį šitą, pagausiu. (1;11).
'I will catch, catch this Monika:DIM:MASC:ACC, will catch her.'

Situation: Monika and her mother are looking at the photo-album.

- (4) MONIKA: Čia Monikutė.
'This is Monika:DIM:FEM:NOM.'
- (5) MOTHER: Monikutė čia maža.
'Monika:DIM:FEM:NOM is little here.'
- (6) MOTHER: Kur dar Monikutė?
'Where else is Monika:DIM:FEM:NOM?'
- (7) MONIKA: Irgi Monikutis yra.
'Here is Monika:DIM:MASC:NOM too.'
- (8) MOTHER: Monikutis. (2;0).
'Monika:DIM:MASC:NOM.'

⁶ In colloquial Lithuanian it is quite often that women are addressed by a masculine form of the noun, especially if it is a diminutive form, e.g., *kačiuk* ('kitten'), *Linuk* (female name), *mažiuk* ('little'), *Lorečiuk* (female name).

It needs to be pointed out that not only Monika's mother, but also the girl addresses herself by a masculine form of her name or uses other words to refer to herself, e.g., *vaikas* 'child:MASC'. Could it be the case that the girl, even though being aware of her own gender, has not acquired the rule of grammatical gender? Other examples demonstrate that this is not true to fact:

Situation: Mother is pointing to Monika.

(9) MOTHER: Čia Balčiūnaitė.

'This is Balčiūnaitė (last name, feminine).'

(10) MONIKA: Čia Balčiūnaitė Monikutis. (2;1).

'This is Balčiūnaitė Monika:DIM:MASC:NOM.'

Situation: Monika enters the room.

(11) MONIKA: Čia tokia Monikutis įėjo.

'Here came in such:FEM:NOM Monika:DIM:MASC:NOM.'

Monika is well aware of belonging to the feminine gender. Examples where agreement is required demonstrate her ability to use the pronoun *tokia* 'such' and her family/last name *Balčiūnaitė* in the feminine gender form, whereas the form of her given name is used in masculine. It shows that the application of a semantic rule in this case is more important than a grammatically correct syntactic agreement rule. Examples from Monika's speech demonstrate that the girl has mastered all rules of gender, i.e., formal, semantic and syntactic already by the age of 2;1. The cases of syntactic (but not grammatical) agreement, such as *tokia* (FEM) *Monikutis* (MASC), allow us to state that she has also acquired adult-like grammatical competence.

5. Conclusion

The analysis of the data pertaining to gender acquisition in Rūta's and Monika's speech allows us to make the following generalizations. First, the results obtained demonstrate that the course of the acquisition of the category of gender was not problematic; on the contrary, it was fast and easy. Both girls have acquired the category of gender rather early, i.e., by the period of 2;3.

Rūta's data show that in making gender distinctions the girl more often relies on formal criteria. Another point is that such linguistic properties of grammatical items as transparency and frequency facilitate the process of gender acquisition at an early stage. It could be claimed that the girl grasped the essence of the category fairly early, around 2;0. Starting from 2;3 onwards, Rūta does not make any errors in gender marking (with the exception of the masculine dative singular forms).

Finally, data on Rūta's acquisition of gender do not corroborate the hypothesis which was supported by empirical evidence from other languages to the effect that the unmarked member, i.e. nouns of masculine gender, is acquired earlier than the marked member of the opposition (i.e. feminine nouns). In contrast, nouns of feminine gender were acquired first and in a grammatically correct form from the very beginning of the recordings (1;7).

Therefore we assume that the girl's name, a clear indication of natural gender, helps to acquire the forms of those nouns which belong to the same class as the girl's name, i.e. feminine noun class.

At the moment we cannot answer the question whether this hypothesis may be valid for other children, but research in this direction might provide answers to the complex process of language acquisition in the context of gender differences.

References

- Christofidou, Anastasia; Stephany, Ursula 1997. The early development of case forms in the speech of a Greek boy: A preliminary investigation. – J. Fisiak (Ed.). *Papers and Studies in Contrastive Linguistics* 33. Poznan: Adam Mickiewicz University, 127–139.
- Clark, Eve 1985. The acquisition of Romance, with special reference to French. – D. I. Slobin (Ed.). *The Crosslinguistic Study of Language Acquisition* 1. Hillsdale, NJ: Erlbaum, 687–782.
- Corbett, Greville G. 1999. *Gender*. Cambridge: Cambridge University Press.
- Crystal, David 1991. *A Dictionary of Linguistics and Phonetics*. Blackwell Publishers.
- Dressler, Wolfgang U.; Drażyk, Danuta; Dziubalska-Kołaczyk, Katarzyna; Jagła, Ewelina 1996. On the earliest stages of acquisition of Polish declension. – *Acta Wiener Linguistische Gazette* 53–54, 1–21.
- Dressler, Wolfgang U.; Mayerthaler, Willi; Panagl, Oswald; Wurzel, Wolfgang U. 1987. *Leitmotifs in Natural Morphology*. Amsterdam: Benjamins.
- Eckert Penelope; McConnell-Ginet, Sally 2003. *Language and Gender*. Cambridge University Press.
- Fagot, B. I.; Leinbach, M. D.; Hagan, R. 1986. Gender labeling and the adoption of sex-types behaviors. – *Development Psychology* 22, 440–453.
- Gelman, Susan A.; Collman, P.; Maccoby, Eleanor E. 1986. Inferring properties from categories versus inferring categories from properties: The case of gender. – *Child Development* 57, 396–404.
- Girdenis, Aleksas 1981. *Fonologija*. Vilnius: Mokslas.
- Gleason, Jean Berko; Perlmann, Rivka Y.; Ely, D.; Evans, D. 1994. The baby talk register: Parents' use of diminutives. – J. L. Sokolov, C. E. Snow (Eds.). *Handbook of Research in Language Development Using CHILDES*. Hillsdale, NJ: Lawrence Erlbaum Associates, 50–76.
- Greenberg, Joseph 1966. *Language Universals: With Special Reference to Feature Hierarchies*. The Hague: Mouton.
- Karmiloff-Smith, Annette 1979. *A Functional Approach to Child Language*. Cambridge: Cambridge University Press.
- Karmiloff, Kira; Karmiloff-Smith, Annette 2002. *Pathways to Language*. Harvard University Press.
- Korecky-Kröll, Katharina; Dressler, Wolfgang U. 2007. Diminutives and hypocoristics in Austrian German. – Ineta Savickienė, Wolfgang U. Dressler (Eds.). *The Acquisition of Diminutives: A cross-linguistic perspective*. Amsterdam: Benjamins, 207–230.
- Levy, Yonata 1983. It's frogs all the way down. – *Cognition* 15, 75–93.
- Levy, Yonata 1988. On the early learning of formal grammatical systems: Evidence from studies of the acquisition of gender and countability. – *Journal of Child Language* 15, 179–187.
- MacWhinney, Brian 1978. *The Acquisition of Morphophonology*. Monographs of the Society for Research in Child Development, vol. 43 (1–2), No. 174.
- MacWhinney, Brian; Snow, Catherine 1990. The Child Language Data Exchange System: An update. – *Journal of Child language* 17, 457–472.
- Maratsos, Michael P.; Chalkey, Mary Anne 1980. The internal language of children's syntax, the ontogenesis and representation of syntactic categories. – K. E. Nelson (Ed.). *Children's Language*. Vol. II. New York: Gardner, 127–214.

- Maratsos, Michael P. 1988. The acquisition of formal word classes. – Y. Levy, I. M. Schlesinger, M. D. S. Braine (Eds.). *Categories and Processes in Language Acquisition*. Hillsdale, N J: Lawrence Erlbaum Associates, 31–44.
- Mills, Anne E. 1986. *The Acquisition of Gender: A Study of English and German*. Heidelberg: Springer.
- Mulford, Randa 1983. Semantic and formal factors in the comprehension of Icelandic pronoun gender. – *Papers and Reports on Child Language Development* 22, 83–91.
- Mulford, Randa 1985. Comprehension of Icelandic pronoun gender: Semantic versus formal factors. – *Journal of Child Language* 12, 443–454.
- Perez-Pereira, Miguel 1991. The acquisition of gender: What Spanish children tell us. – *Journal of Child Language* 18, 571–590.
- Popova, Maria I. 1973. Grammatical elements of language in the speech of pre-preschool children. – Ch. A. Ferguson, D. I. Slobin (Eds.). *Studies of Child Language Development*. New York: Holt, Rinehart, Winston, Inc.
- Protassova, Ekaterina; Voeikova, Maria 2007. Diminutives in Russian. – Ineta Savickienė, Wolfgang U. Dressler (Eds.). *The Acquisition of Diminutives: A cross-linguistic perspective*. Amsterdam: Benjamins, 43–72.
- Rūķe-Draviņa, Velta 1973. On the emergence of inflection in child language: A contribution based on Latvian speech data. – Ch. A. Ferguson, D. I. Slobin (Eds.). *Studies of Child Language Development*. New York: Holt, Rinehart, Winston, Inc.
- Savickienė, Ineta 1998. The acquisition of diminutives in Lithuanian. – *Studies in the Acquisition of Number and Diminutive Marking*. Antwerp Papers in Linguistics 95. Antwerp: University of Antwerp, 115–135.
- Savickienė, Ineta 2001. The role of diminutives in Lithuanian child language acquisition. – *Linguistica Baltica* 9. Universitas Cracow, 109–118.
- Savickienė, Ineta 2003. *The Acquisition of Lithuanian Noun Morphology*. Verlag der Österreichischen Akademie der Wissenschaften. Wien.
- Slobin, Dan I. 1985. Introduction: Why study acquisition crosslinguistically? – D. I. Slobin (Ed.). *The Crosslinguistic Study of Language Acquisition*, Vol. 1. Hillsdale, NJ: Erlbaum, 3–27.
- Smoczyńska, Magdalena 1985. The acquisition of Polish. – D. I. Slobin (Ed.). *The Crosslinguistic Study of Language Acquisition*, Vol. 1. Hillsdale, NJ: Erlbaum, 595–686.
- Stephany, Ursula 1997. Diminutives in early Greek. – W. U. Dressler (Ed.). *Studies in Pre- and Protomorphology*. Wien: Verlag der Österreichische Akademie der Wissenschaften, 145–156.
- Stump, Gregory T. 1998. Inflection. – A. Spencer, A. M. Zwicky (Eds.). *The Handbook of Morphology*. Blackwell Publishers, 13–44.
- Valeckienė, Adelė 1998. *Funkcinė lietuvių kalbos gramatika*. Vilnius: Mokslo ir enciklopedijų leidykla.
- West, Candace; Zimmerman, Don 1987. Doing gender. – *Gender and Society* 1, 125–151.
- Zangl, Renate 1997. Input selection and first patterns in early language development. – W. U. Dressler (Ed.). *Studies in Pre- and Protomorphology*. Wien: Verlag der Österreichische Akademie der Wissenschaften, 11–28.

Ineta Savickienė, Regional Studies Department, Vytautas Magnus University, (Kaunas, Lithuania).
 Research interests: language acquisition, sociolinguistics, psycholinguistics.
 i.savickiene@pmdi.vdu.lt

Violeta Kalėdaitė, English Philology Department, Vytautas Magnus University (Kaunas, Lithuania).
 Research interests: contrastive linguistics (syntax and pragmatics), language typology, language acquisition, translation studies.
 v.kaledaite@hmf.vdu.lt

LAPSE SOO MÕJU KEELE OMANDAMISEL

Ineta Savickienė, Violeta Kalėdaitė

Vytautas Magnuse nimeline Kaunase Ülikool

Artiklis otsitakse seoseid keele omandamise protsessi ja soo vahel. Uurimismaterjaliks on ühe tüdruku, Rūta longitudinaalne korpus vanusest 1;7–2;6 kogukestusega 35 tundi. Lindistatud kõne on transkribeeritud CHILDESe (MacWhinney, Snow 1990) süsteemis. Vajadust mööda on täiendava materjalina kasutatud teise tüdruku, Monika keeleandmeid.

Sookategooria omandamine toimus mõlemal tüdrukul kiirelt, kergelt ning üsna varakult, vanuses 2;3. Rūta materjal näitab, et grammatilise soo eristamisel toetub ta enamasti sõnade vormitunnustele. Tüdruk mõistis sookategooria olemust üsnagi vara, umbes vanuses 2;0. Alates vanusest 2;3 Rūta enam soo markeerimisel ei eksi (välja arvatud meessoost sõnade ainsuse daativi vormid). Grammatiliste üksuste läbipaistvus ja kasutusagedus aitavad tõenäoliselt sookategooria varasele omandamisele kaasa.

Rūta keeleomandamise andmed ei toeta teiste keelte empiiriliste uurimistulemuste põhjal püstitatud hüpoteesi, et soo-opositsiooni markeerimata liige, mees-sugu (meessoost nimisõnad), omandatakse enne naissugu (naissoost nimisõnu). Vastupidi, naissoost sõnad omandati kõigepealt, ja need vormistused grammatiliselt korrektselt juba lindistuste alguses (vanuses 1;7). Võib oletada, et tüdruku nimi, mis selgelt viitab selle kandja looduslikule soole, aitab omandada samasse klassi kuuluvate, niisiis naissoost nimisõnade vorme.

Võtmesõnad: keele areng, varane etapp, longitudinaalne korpus, nimisõnad, grammatiline sugu, morfosüntaktilised omadused, muutmismallid, ühildumine, leedu keel

ÕPETAJA TAGASISIDE ÕPILASE VASTUSELE ALGKOOLITUNNIS

Krista Strandson

Ülevaade. Uurimuse eesmärk on visandada õpetaja algatatud küsimussekventside struktuur ja analüüsida õpetaja tagasisidet õpilase vastusele algkoolitundides.* Analüüsi aluseks on 280 min litereeritud videosalvestust 1.–3. klassi matemaatika ja emakeele tundidest. Õpetaja tagasisidet on analüüsitud kahest vaatenurgast. Esiteks, kuidas on tagasiside seotud küsimuse sisuga: kas õpetaja tagasiside õpilase teadmisi kontrollivate küsimuste vastustele on mingil moel teistsugune kui õpilase kohta käivate küsimuste vastustele. Teiseks, missugune on tagasiside funktsioon: mil moel määrab õpetaja tagaside suhtluse edasise käigu. Analüüs näitab, et õpetaja juhib tunni kulgu äärmiselt kindlakäeliselt. Mis tahes õpilasepoolne tegevuse algatuse võimalus on minimaalne. See avaldub järgnevas. 1) Sekventsi lõpetav tagasiside on valdav sõltumata küsimuse sisust. Enamik õpetaja algatatud küsimussekventsidest moodustavad kolmikstruktuuri (küsimus–vastus–tagasiside) isegi juhul, kui õpetaja pedagoogiline eesmärk on arendada vestlust. 2) Kui õpetaja soovib sekventsi laiendada, teeb ta seda järgmise vooru sisu suunavate küsimuste abil; minimaalset vastaja jätkamist taotlevat tagasisidet, mis järgmise vooru sisu ei määratle, kasutab õpetaja väga vähe.

Võtmesõnad: vestlusanalüüs, koolitunnisuhtlus, küsimussekventsid, reaktsioon vastusele, eesti keel

1. Sissejuhatus

Küsimused on koolitunni vältimatu osa: õpetaja küsimused, õpilase vastused ja õpetaja tagasiside vastustele kuuluvad nii õpetamis- kui ka õppimisprotsessi juurde. Tüüpiline õpetaja esitatud küsimus moodustab koos sellele järgnevate voorudega kolmiksekventsi, mis koosneb kolmest omavahel seotud kõnevoorst: õpetaja küsimus – õpilase vastus – õpetaja tagasiside (vt nt Mehan 1985; Sinclair,

* Tööd on toetanud Eesti Teadusfond (grant nr 5685) ja riikliku programmi "Eesti keel ja rahvuslik mälu" eesti keele mooduli projekt "Eesti kõnekeele korpuse kogumine ja transliteerimine". Autor on tänulik retsensentide märkuste eest.

Coulthard 1975). Klassikaline kolmik on näiteks vestluskatke (1). Näite 1. reas on õpetaja küsimus, 4. reas vastama valitud õpilase vastus küsimusele ning 5. reas õpetaja tagasiside vastusele: vastuse kinnitamine keelendiga *õige*. (Näidetes kasutatud eesti suulise keele transkriptsioonimärgid on esitatud kogumiku lõpus; Õ – õpetaja, pärisnimi – õpilane.)

(1) ((3. klass, matemaatika))

1. Õ: mis= `tuli sis `teha kui oli `korda `rohkem.
2. (...)
3. Õ: Ander.
4. Ander: korrutada.
5. Õ: õige. `kuula `ülesannet.

Termini *õpetaja tagasiside* all on mõeldud kogu õpetaja reaktsiooni õpilase vastusele. Tagasisidevooru võib moodustada üksnes minimaalne tagasidekeelend, mis annab kõnelemise õiguse õpilasele tagasi, tähistab info vastuvõttu, annab vastusele hinnangu või kinnitab vastust; aga ka õpetaja pikem kommentaar või täpsustav küsimus vastuse kohta.

Õpetaja tagasisidevoor määrab suhtluse edasise käigu. See, kui pikalt ja kui palju õpetaja õpilasel rääkida laseb, sõltub konkreetsest pedagoogilisest eesmärgist ja kontekstist. Teoreetiliselt võiks oletada, et õpilaste kõnelemisõigus on vabam tunni osades, mille eesmärk on arendada vestlust, panna õpilasi oma arvamust avaldama, ning reglementeeritum siis, kui on keskendunud uute teadmiste omandamisele. Sellest lähtuvalt on õpetaja küsimused jagatud sisu järgi kahte rühma vastavalt sellele, kes on õige vastuse “teadja”:

- õpetaja on “teadja”. Siia rühma kuuluvad õpilase teadmisi kontrollivad küsimused, mille vastust õpetaja teab;
- õpilane on “teadja”. Siia rühma kuuluvad küsimused, mille vastust õpetaja ei saa teada, näiteks küsimused õpilase arvamuse, isiku jms kohta.

Uurimuse laiem eesmärk on analüüsida sotsiaalseid suhteid klassiruumis, mil moel avaldub õpetaja ja õpilase rollijaotus keeles ja suhtluses. Käesolevas artiklis on piirdutud õpetaja tagasiside analüüsiga: kas ja kuidas erineb tagasiside teadmisi kontrollivate ja õpilase kohta käivate küsimuste vastustele, missugune on mõlemat tüüpi küsimussekventsides ülesehitus.

Analüüs on kahe meetodi kombinatsioon. Analüüsi esimene etapp oli diskursusanalüütiline (Mehan 1985; Sinclair, Coulthard 1975) küsimussekventsides vormi analüüs: vaatasin küsimussekventsides struktuuri. See, mida üksikud struktuurikomponendid mingis kontekstis teevad ja tähendavad, jäi siin täpsemalt määratlemata. Et uurida õpetaja tagasiside funktsiooni täpsemalt, rakendasin vestlusanalüüsi meetodikat (vt näiteks Sacks 1992: 521–570; Hutchby, Wooffitt 1998: 39; Hennoste 2000a: 2236), analüüsisin iga tagasisidekeelendi funktsiooni selle reaalses kasutuskontekstis. Analüüsietapid on teineteisest sõltumatud ses mõttes, et küsimussekventsides struktuuri analüüsil saadud statistikat ei ole kasutatud tagasiside funktsiooni analüüsil ja vastupidi, tagasiside funktsiooni mikroanalüüsi tulemusel saadud statistikat küsimussekventsides struktuuri analüüsil.

2. Õpetaja tagasiside funktsioon

Õpetaja tagasisidel on hulk ülesandeid. Näiteks võib see osutada, et vastus oli õige, õpetaja arvates asjakohane; viidata vastuse täpsustamise/parandamise vajadusele; õpetaja võib anda hinnangu vastusele jms. Õpetaja tagasiside määrab otseselt suhtluse edasise käigu. Dialoogi jätkamise seisukohalt jagunevad küsija tagasisidevoorud küsimuse vastusele mis tahes vestluses laias laastus kahte rühma vastavalt sellele, kas tagasiside taotleb vastaja (või kellegi teise) jätkamist või haarab tagasisidevooru sõnastaja kõnevooru endale ja algatab ise uue sekvents.

Vastaja jätkamist taotlev tagasiside jaguneb omakorda kahte rühma: minimaalsed jätkamist taotlevad keelendid, näiteks analüüsitud koolitundides noogutus ja *veel*, argivestlustes *mhmh*; ning täpsustavad või parandust algatavad küsimused. Jätkaja funktsioonis lühikeste tagasisidekeelenditega osutab kuulaja, et teine pool võib oma teksti jätkata (Hennoste 2000b: 1788; Hennoste, Rääbis 2004: 134). Täpsustavad ja parandust algatavad küsimused taotlevad samuti teise poole jätkamist, kuid suunavad ja/või piiritlevad taotletava voo sisu, nt (19), (20). Täpsustava küsimuse esitab õpetaja juhul, kui vastus on tema arvates ebapiisav või teema edasiarendamist väärt. Sekventsi lõpetava tagasiside annab õpetaja pärast õpilaselt saadud täpsustusi. Paranduse algatab õpetaja siis, kui õpilase vastus oli vale või oli selles mingil põhjusel (näiteks kuulmis- või arusaamisprobleem) midagi parandamist-täpsustamist vajavat.

Tagasisided, mis kõneleja jätkamist ei taotle, on keelendid, mis osutavad, et info on vastu võetud, et sekvents, alateema või teema on lõppenud. Erinevalt jätkajatest ei suuna nad kõnevooru eelmisele kõnelejale tagasi, tavaliselt jätkab kõnelemist vastuvõtuteate sõnastaja, näiteks algatab uue alateema või esitab uue küsimuse.

Siia rühma kuuluvad kaht liiki keelendid. Esimese rühma moodustavad lühikesed tagasisidekeelendid, info vastuvõtuteated, mis tähistavad kõige üldisemalt info vastuvõttu ja ei taotle seejuures eelmise kõneleja jätkamist. Sellised lühikesed keelendid on näiteks *ahah, jah, jaa, nii, õige, hästi* jne. Teise rühma kuuluvad õpetaja pikemad kommentaarid vastuse kohta. Lühikestel info vastuvõtu teadetal on hulk täpsemaid ülesandeid, mis on seotud nii konkreetse keelendi tähenduse kui ka reaalse kasutuskontekstiga. Vastuvõtuteate rollis tagasiside koolitundide analüüsis eristatud funktsioonid on järgmised.

- **Neutraalne vastuvõtuteade** tähistab, et info on vastu võetud, kuid ei anna saadud info kohta hinnangut (Hennoste, Rääbis 2004: 136–137). Neutraalsed vastuvõtuteated on näiteks vastuvõtukordus, partiklid *jah, jaa, mhmh*.
- **Hinnanguline vastuvõtuteade** annab lisaks info vastuvõtmise tähistamisele vastuse kohta hinnangu, näiteks *tore, tubli, vahva, kurb, ilus, kena, oi jt*.
- **Vastuvõtuteade, mis tähistab informeerituse muutust** näitab, et eelnev info oli uus, et kuulaja on selle vastu võtnud ja sellest aru saanud (Hennoste 2000b: 1780; Keevallik 2003: 38). Vormiliselt on selles tähenduses tavaliselt partikkel *ahah*, selle variant *ahaa* ning kahest eelmisest mõnevõrra erineva tähendusega *aa*. Ka informeerituse muutusele osutavad partiklid on ses mõttes neutraalsed vastuvõtuteated, et need ei anna

vastusele hinnangut, vaid tähistavad, et info vastuvõtja käsitleb saadud infot enda jaoks uuena või vähemalt soovib nii näidata.

- **Kinnitavat vastuvõtuteadet** kasutatakse koolitundides väga sageli, vormiliselt võib kinnituse ülesannet täita vastuvõtukordus; keelendid *jah, jaa, mhmh, just, õige* ning hinnangulised keelendid *kena, hästi, hea* jt.
- **Piiritleja** osutab nagu teisedki info vastuvõtu teated, et info on vastu võetud ja sekvents, teema või alateema lõppenud, kuid tähistab sekventsi piire tugevamalt kui hinnanguline ja neutraalne vastuvõtuteade (Hennoste, Rääbis 2004: 136–137). Piiritleja funktsioonis võivad olla näiteks keelendid *nii, kena* ja tänusõnad *aitäh, tänan*.

Kõigile tagasisidele leiduvad hinnangulised paarikud (Hennoste, Rääbis 2004: 132). Hinnanguline variant esines käesoleva analüüsi järgi nii kinnitaval kui neutraalsel vastuvõtuteatel. Analüüsitud materjal ei sisaldanud ühtegi hinnangulise jätkaja rollis tagasisidet, kuigi teoreetiliselt on see võimalus olemas; piiritleva tagasiside hinnangulised variandid kuuluvad hinnanguliste vastuvõtuteadete rühma.

Kommentaariid vastuse kohta võivad lisanduda igale tagasisidekeelendile: nad seletavad, täpsustavad, laiendavad jms saadud vastust või annavad vastuse kohta pikema hinnangu. Järgmises näites täpsustab ja laiendab õpetaja Siimult saadud õiget vastust reas 3.

(2) ((1. klass, eesti keel))

1. Õ: ja mis=te arvate `miks ma panen kaks tooli klassi ette.
2. Siim: jutustajaid sab kaks tükki kes tulevad `lugema.
3. Õ: `väga õige. **ja nemad on sis Žanna ja ja Fe:liks.**

Õpetaja tagasiside funktsiooni analüüsitud korpuses võtavad kokku tabelid 1 ja 2.

Tabel 1. Tagasiside, mis ei taotle eelmise kõneleja jätkamist

Tagasiside tüüp	Õpetaja küsimus	
	Õpetaja on "teadja"	Õpilane on "teadja"
Kinnitus		
(väga) õige	44	
õigus	1	
seda ka	1	
olen nõus	1	
Kinnitus/hinnang		
(väga) tubli (oled)	19	
(väga) hästi	4	
(väga) kena	4	
väga hea	3	
väga tore	1	
Hinnang		
(väga) tubli	10	5
(väga) tore		17
väga vahva		8
(väga) hästi		6

(väga) kena		4
väga hea	2	
väga ilus	1	
suurepärane		1
kurb (küll)	2	1
noojah		1
oi		1
Neutraalne vastuvõtuteade/kinnitus		
kordus	74	
jaa	16	
jah	13	
mhmh	9	
noogutus	2	
Neutraalne vastuvõtuteade		
mhmh		8
jaa		9
jah		1
ahah	5	3
ahaa	3	1
aa		1
kordus	2	6
okei		1
Piiritlev vastuvõtuteade		
nii	9	5
kena	1	
aitäh	7	1
tänan	2	
Kommentaari	52	24

Tabel 2. Eelmise kõneleja jätkamist taotlev tagasiside

Tagasiside tüüp	Õpetaja küsimus	
	Õpetaja on "teadja"	Õpilane on "teadja"
Jätkaja noogutus veel	5	1
Kinnitus/jätkaja jaa kordus	11 3	
Täpsustav küsimus	79	35
Parandusalgatus	60	7

Tagasiside funktsionaalsete alarühmade piirid ei ole alati selged: üks ja sama keelend võib korraga täita mitut ülesannet. See, missugus(t)es funktsiooni(de)s konkreetne keelend esineb, selgub ainult kasutuskonteksti analüüsidest. Siin on oluline küsimuse

sisu: kas tegemist on teadmisi kontrolliva küsimusega või küsimusega õpilase kohta. Teadmisi kontrollivatele küsimuste õigete vastuste reaktsioonivoorudele kaldub sageli lisanduma kinnituse lisatähendus, mis võib haakuda kas hinnangulisele vastuvõtuteatele (*tubli, tore, hästi, kena*) või neutraalsele info vastuvõtuteatele (vastuvõtukordusele, partiklitele *jah, jaa, mhmh*). Näiteks reaktsioonina õpilase kohta käiva küsimuse vastusele on *jah, jaa* ja *mhmh* neutraalse vastuvõtuteate funktsioonis, kuna õpetaja ei tea “õiget” vastust ega saa seda ka kinnitada või ümber lükata. Seevastu reaktsioonina õpilase teadmisi kontrolliva küsimuse vastusele on keelenditel *jah, jaa* ja *mhmh* kinnituse lisatähendus. Ainus tagasisidekeelendite rühm, mis õpilase jätkamist ei taotle ja millele kinnituse tähendust teadmisi kontrolliva küsimuse vastusele reageerides ei lisandu, on informeerituse muutusele osutavad vastuvõtuteated *ahah, ahaa* ja *aa*. Ka piiritlejale võib juhuslikult, kui ta esineb tagasisidena teadmisi kontrollivale küsimusele, implitsiitselt lisanduda kinnituse tähendus. Kuna piiritleja esmane funktsioon on osutada sekvenssi piiridele, ei ole käesolevas analüüsis kinnituse lisatähendusega piiritlejaid eristatud.

Järgnevalt on analüüsitud kumbagi tagasisidekeelendite rühma eraldi kahes kasutuskontekstis: tagasisidena õpilase teadmisi kontrollivatele küsimustele ja õpilase kohta käivatele küsimustele.

3. Tagasiside, mis eelmise kõneleja jätkamist ei taotle

3.1. Kinnitus

Kinnitus järgneb õpilase teadmisi kontrolliva küsimuse vastusele juhuslikult, kui vastus oli õige ja õpetaja arvates asjakohane. Õpetaja kasutas õpilase vastuse kinnitamiseks järgmisi vahendeid:

- keelendeid *õigus, (väga) õige, seda ka*;
- keelendeid (*väga*) *tubli, (väga) hästi, (väga) hea, (väga) kena, (väga) tore*;
- õpilase vastuse täpset, osalist või modifitseeritud kordust;
- partikleid *jah, jaa, mhmh*;
- noogutust.

Vastuse kordus on kõige sagedasem vastuse kinnitamise vahend (74 juhuslikult 192-st); sageduselt järgnevad keelend *õige* (44 juhuslikult), kinnitava lisatähendusega *tubli* (19 juhuslikult), partiklid *jaa* (16 juhuslikult) ja *jah* (13 juhuslikult). Kordusega kinnituse sage kasutus ei ole kindlasti mitte juhuslik. Esiteks võib korduste suur osakaal õpetaja kõnes tuleneda lastega suhtlemise eripärast. Teiseks võimaldavad kordused olulist infot veelkord kogu klassi jaoks kuuldavaks teha, et garanteerida õige vastuse jõudmist võimalikult paljude õpilasteni.

Kui heita pilk tabelisse 1, siis on esiteks näha, et kinnitada saab vaid teadmisi kontrolliva küsimuse vastust; ning teiseks: ainult kinnitavas tähenduses kasutas õpetaja keelendeid *õigus, (väga) õige, seda ka* ja *olen nõus*. Keelenditel *õige* ja *õigus* on kõige “tugevam” kinnitav tähendus. Neid keelendeid kasutas õpetaja ainult õpilase teadmisi kontrolliva küsimuse õigele vastusele reageerides, nagu näites (3).

(3) ((3. klass, matemaatika))

1. Õ: * `kaheksa `teist jagatud `kahega. * Vaike?
2. Vaike: öheksa.
3. **Õ: õige?**

Ülejäänud keelendid, mida õpetaja vastuse kinnitamiseks kasutas, on olemuselt mitmetähenduslikud, nende tegelik tähendus selgub alles kontekstis. Näiteks keelendil *tubli* on teadmisi kontrolliva küsimuse korral lisaks hinnangu andmisele ka kinnituse lisatähendus, kuid reaktsioonina õpilase kohta käiva küsimuse vastusele *tubli*l kinnitav tähendus puudub, vrd (4) ja (5).

(4) ((3. klass, matemaatika))

1. Õ: `mis=oli `korrutis ja `mis=olid `tegurid, `tuleta `meelde? (...)
2. `viis `teist. (...) `millised tegurid on?
3. Liis: `kolm=korda=viis.
4. Õ: **tubli?**

(5) (1. klass, õpetaja jagas lastele paberilipikud, kuhu oli kirjutatud elusorganismide klassid. Lapsed pidid mõtlema, mis liigi esindaja nad tahaksid olla ja miks)

1. Õ: `kellel=on `valmis, [`Amaalia] `palun.
2. L?: [`mina tahan]
3. (.)
4. Amaalia: kui kui=ma `oleksin `loom (.) [sis=ma `oleksin] e `valge `hobune?
5. Õ: [> * `Mait=tule `siia. * <]
6. Õ: * miks. *
7. (.)
8. Amaalia: `sellepärast mulle= `meeldivad ä `hästi `hobused?
9. Õ: **tubli.**

Näites (4) esitab õpetaja teadmisi kontrolliva küsimuse, mille eesmärk on, et õpilane ütleks arvu viisteist tegurid. Õpetaja saab Liisilt dialoogi 3. reas õige vastuse ning kinnitab seda keelendiga *tubli*. Poisitiivne hinnang täidab ka vastuse kinnitamise rolli.

Näites (5) reageerib õpetaja sama keelendiga õpilase kohta käiva küsimuse *mis loom, lind või taim sa tahaksid olla* vastusele. Õpetaja ei saa vastust kinnitada, kuna ta ei tea õiget vastust: selle näite *tubli*-l saab olla ainult hinnangu tähendus.

Vastuvõtukordus ja partiklid *jah, jaa, mhmh* võivad samuti sõltuvalt eelnenud küsimuse sisust olla kas vastuse kinnitajad, nagu *mhmh* näites (6), või neutraalsed info vastuvõtu teated (8).

(6) ((3. klass, matemaatika))

1. Õ: ee `kuus jagatud `kahega.
2. (...) ((lapsed tõstavad kätt, õpetaja osutab Gerlile))
3. Gerli: kolm.
4. **Õ: mhmh**
5. (.) ((Gerli istub))

3.2. Hinnang

Hinnangulised vastuvõtuteated, mida õpetaja kasutas, olid järgnevad: *väga tubli, väga tore, väga vahva, väga hästi, väga kena, väga hea, väga ilus, suurepärane, oi, noojah, kurb küll*. Kõik, välja arvatud loetelus viimased, *noojah* ja *kurb küll*, annavad positiivse hinnangu. Juhul kui õpetaja sai vale vastuse või ei jäänud vastusega rahule, algatas ta pea alati paranduse, andes probleemse vastuse sõnastajale või mõnele teisele õpilasele võimaluse vastus uuesti öelda. Hinnanguline tagasiside võib järgneda nii teadmisi kontrollivate küsimuste kui ka õpilase kohta käivate küsimuste vastustele.

Õpilase kohta käiva küsimuse vastusele reageerides ei lisandunud hinnangule ühelgi juhul kinnituse tähendust, see tuleneb juba küsimuse sisust: õpetaja ei saa kinnitada ega kummutada infot, mida ta ise varem ei teadnud (5).

Teadmisi kontrolliva küsimuse vastusele reageerides olid hinnangulisel keelendil järgmised tähendused ja positsioonid tagasisidevoorus.

- Keelenditele *väga tubli, väga hästi, väga kena* ja *väga tore* võib teadmisi kontrolliva küsimuse vastusele reageerides lisanduda kinnituse tähendus. Sellisel juhul esines hinnanguline keelend reaktsioonivoorus üksi, positiivse hinnanguga keelend kinnitas ühtlasi ka saadud vastuse õigeks (4).
- Hinnanguline tagasiside võis moodustada tagasisidevooru koos vastust kinnitava ja/või piiritleva tagasisidega (7). Sellisel juhul hinnangulisele tagasisidele kinnituse tähendust ei lisandu, kuna vastuse õigsust kinnitav keelend on tagasisidevoorus juba olemas, nagu näites *neli muidugi* 3. reas.

(7) ((3. klass, matemaatika))

1. Õ: * `kaheksa jagatud `kahega. * ((õpetaja vaatab Karolini poole))

2. Karolin: neli.

3. Õ: neli, muidugi. (.) ((Karolin istub)) aitäh, (.) **tubli**.

3.3. Neutraalne vastuvõtuteade

Neutraalse vastuvõtuteatega osutab info vastuvõtja, et info on vastu võetud, kuid ei anna saadud info kohta hinnangut. Neutraalsed vastuvõtuteated olid vaadeldud materjalis vastuvõtukordus; partiklid *jah, jaa, mhmh, okei* ning informeerituse muutusele osutavad partiklid *ahah, ahaa* ja *aa*. Tähenduselt jagunevad need partiklid kaheks: partiklid *jah, jaa, mhmh* ning vastuvõtukordus sobivad ka vastuse kinnitamiseks ja samameelsuse osutamiseks; *ahah, ahaa* ja *aa* aga tähistavad saadud info uudseks, kinnituse lisatähendus neil puudub. Teoreetiliselt võiks oletada, et koolitundides kasutab õpetaja neutraalset vastuvõtuteadet juhul, kui ta reageerib õpilase kohta käiva küsimuse vastusele, kuna sellist vastust ei saa kinnitada ega kummutada. Jällegi võiks öelda, et nii kaldus olema, kuid ei olnud alati. Tõepoolest, enamik ainult vastuvõtuteate funktsioonis keelendeid esines tagasisidena õpilase kohta käivatele küsimustele (41-st 30). Kuid ülejäänud 11 juhtu olid reaktsioonid teadmisi kontrollivatele küsimustele. Nüüd jõuame järgmiste küsimusteni. Millal ja miks kasutas õpetaja teadmisi kontrollivale küsimusele reageerides neutraalset vastuvõtuteadet, andmata vastuse õigsuse/sobivuse kohta omapoolset hinnangut?

Missugust pedagoogilist ja/või suhtluslikku eesmärki neutraalne vastuvõtuteade teadmisi kontrollivate küsimuste puhul täidab? Tabelit 1 jälgides näeme 10 sellist näidet, kus õpetaja reageeris õpilase teadmisi kontrollivale küsimusele neutraalse vastuvõtuteate rollis tagasisidekeelendiga; konkreetsemalt: 8 korral info uudsusele osutavate partiklitega *ahah* ja *ahaa* ning kahel korral vastuvõtukordusega.

Neid numbreid vaadates tekib kohe kaks tähelepanekut. Esiteks: info uudsusele osutav partikkel *ahah* ja selle variant *ahaa* ning partikkel *aa* on koolitunnis äärmiselt haruldased – kogu analüüsitud korpuses leidis vaid 13 kasutusjuhtu, neist 8 reaktsioonina teadmisi kontrolliva küsimuse vastusele ja 5 õpilase kohta käiva küsimuse vastusele. Info uudsusele osutavate partiklite harv kasutus reaktsioonina teadmisi kontrollivatele küsimustele on teoreetiliselt loogiline: õpetaja teab ju õiget vastust oma esitatud küsimusele, vastuses ei saa olla tema jaoks uut informatsiooni. Õpilase kohta käivate küsimuste vastust aga õpetaja ei saa teada, järelikult saab ta vastustega enda jaoks uut informatsiooni. Huvitav on, et analüüsitud korpuses leidis info uudsusele osutavat partiklit *ahah* või selle variante reaktsioonina õpilase kohta käivatele küsimustele veelgi vähem kui reaktsioonina teadmisi kontrollivatele küsimustele. Siinkohal on hea võrdlusmaterjal Tiit Hennoste statistika dialoogi-partiklite kasutusest argi- ja ametlikes vestlustes (ametlike vestluste korpuses ei olnud ühtki koolitundi), mille järgi moodustas *ahah* ametliku keele korpuses 3,9% sõnavormidest, argikorpuses aga 1,3%. Partikkel *aa* moodustas ametliku keele korpuses 0,9% ja argikorpuses 2,1% sõnavormidest (Hennoste 2000b: 1780). Käesoleva uurimuse aluseks olnud koolitundide korpuses olid vastavad arvud järgmised: *ahah* ja *ahaa* moodustasid mõlemad 0,03% sõnavormidest, *aa* 0,004% sõnavormidest. Samalaadseid tulemusi on saanud John Heritage, kes pani tähele, et inglise keele informeerituse muutust väljendav tagasiside *oh* puudub regulaarselt sellistes institutsionaalsetes suhtlusregistrites nagu kool, kohus, radio- ja teleuudised. Põhjuseks arvab ta olevat selle, et küsijad (õpetajad, kohtunikud, intervjuerijad) teavad või aimavad vastust, seega ei saa vastus olla küsija jaoks uudne (Drew ja Heritage 1992: 41).

Järgnevalt analüüsime neutraalse vastuvõtuteate funktsioonis olevate tagasisidepartiklite kasutust eraldi kahes kontekstis: tagasisidena teadmisi kontrolliva küsimuse vastusele ning õpilase kohta käiva küsimuse vastusele.

• Küsimus õpilase kohta

Vaatleme kõigepealt neutraalse vastuvõtuteate teoreetiliselt kõige loogilisemat ning ka statistiliselt kõige sagedasemat kasutuskonteksti: tagasisidena õpilase kohta käivale küsimusele. Selles kontekstis kasutas õpetaja vastuvõtukordust ja partikleid *mhmh*, *jaa*, *jah*, *okei*, *ahah*, *ahaa*, *aa*. Kõige sagedamini kasutatavad partiklid olid *jaa* (9 juhtu), *mhmh* (8 juhtu) ja vastuvõtukordus (6 juhtu). Partiklit *jah* kasutas õpetaja neutraalse vastuvõtuteatena vaid 1 korral, ülejäänud 13 *jah*-i esinesid kontekstis, kus neile lisandus kinnituse tähendus. See annab alust oletuseks, justkui oleks *jah* “tugevam kinnitaja” kui *jaa*, kuna vastuvõtuteate funktsioonis kasutati enamasti *jaa*-d. Mingisugust konkreetsemat järeldust ei ole analüüsitud materjali põhjal võimalik teha, eriti kui arvestada asjaoluga, et suurem osa vastuvõtuteate rollis olnud *jaa*-dest olid positsioonis, kus õpetaja küsis ühte ja sama küsimust järjest mitmelt õpilaselt ning *jaa* tähistas ühtlasi üleminekut ühe õpilase küsitlemiselt teise õpilase küsitlemisele. Ka Riina Kasterpalu uurimus partiklite *jah*, *jaa* ning *jajaa*

kasutamisest müügivestlustes näitab muu hulgas, et vastust pakkuva küsimuse kinnitamiseks kasutatakse ülekaalukalt *jah-i*, *jaa-d* oluliselt vähem (Kasterpalu 2005: 999). Info uudsusel osutavaid partikleid *ahah*, *ahaa* ja *aa*, millele kinnituse tähendust ei ole võimalik lisada, kasutas õpetaja neutraalse vastuvõtuteatena õpilase kohta käiva küsimuse vastusele reageerides 5 korral.

Näide (8) illustreerib neutraalse vastuvõtuteate funktsioonis partikli *mhmh* kasutust.

(8) ((1. klass; õpetaja jagas lastele paberilipikud, kuhu oli kirjutatud elusorganismide klassid. Lapsed pidid mõtlema, mis liigi esindaja nad tahaksid olla ja miks))

1. Õ: ee (.) Aune?

2. (...)

3. Aune: { * aga kui ma oleks=lind * } siis `mina `tahaksin olla `luik.

4. (...)

5. Õ: miks=`luik.

6. (.)

7. Aune: mulle=`meeldivad `luiged.

8. (.)

9. Õ: * **mhmh** * (.) `Tiit `Treial.

Näites (8) reageerib õpetaja õpilase 3. ja 7. reas sõnastatud vastusele minimaalse neutraalse vastuvõtuteatega *mhmh*, mis lõpetab ühtlasi sekventsi, ning õpetaja palub samale küsimusele vastata järgmisel õpilasel.

• Teadmisi kontrolliv küsimus

Teoreetiliselt võiks eeldada, et teadmisi kontrolliva küsimuse vastusele reageerib õpetaja kinnitava tagasisidega, mitte neutraalse vastuvõtuteatega. Nii ka enamasti oli, kuid mitte alati: korpuses leidub 10 juhtu, kus õpetaja reageerib teadmisi kontrolliva küsimuse vastusele neutraalse vastuvõtuteatega; 2 juhul vastuse kordusega ja 8 juhul partikliga *ahah* või *ahaa*. Näited, kus õpetaja kasutas neutraalset vastuvõtuteadet teadmisi kontrollivatele küsimustele reageerides, on põnev analüüsimaterjal ning peegeldavad ilmekalt, kuidas suhtlus ja pedagoogika kokku saavad.

Nüüd jõuame juba eespool esitatud küsimusteni: millal ja miks kasutas õpetaja teadmisi kontrollivale küsimusele reageerides neutraalset vastuvõtuteadet, andmata vastuse õigsuse/sobivuse kohta hinnangut? Missugust pedagoogilist ja/või suhtluslikku eesmärki neutraalne tagasiside teadmisi kontrollivate küsimuste puhul täidab?

Vaatame esmalt vastuvõtuteate funktsioonis korduse näidet.

(9) ((1. klass, eesti keel, õpetaja palus lastel valida laualt pildid äsja loetud jutustuse tegelastest))

1. Õ: kes=se `on. ((kummardub Andrease kohale))

2. Andreas: siil

3. Õ: **siil**. kas `oli `täna ses jutus `siil.

4. L-d: e:i

5. Õ: e:i `olnud. `pane=se `lauale tagasi.

Andreas valis ekslikult pildi loomaga, kellest jutustuses ei räägitud. Tema vastus on üldises plaanis õige, õpetaja reageerib sellele kordusega ning algatab seejärel kohe paranduse, mille viib läbi kogu klass. Õpetaja kordus 3. reas on pigem vastuvõtuteate kui kinnituse rollis.

Info uudsusele osutavatest partiklitest kasutas õpetaja teadmisi kontrollivate küsimuste vastustele reageerides partikleid *ahah* ja *ahaa*.

Analüüsitud materjalis esines nendel partiklitel kaks kasutuskonteksti.

• Õpilase vastus oli õige, õpetaja reageeris õigele vastusele partikliga *ahah* või *ahaa*, kuid sel juhul eelnes või järgnes info vastuvõtupartiklile muu kinnitav ja/või hinnanguline tagasisidekeelend (näide 10).

(10) ((1. klass, matemaatika, õpetaja luges lastele ette muinasjutu kolmest karust ja küsis seejärel, mida jutus oli kolm))

1. Õ: `Kertu, mis=`sina arvad.

2. Kertu: kausse.

3. (.)

4. Õ: `kausse=oli kolm. **ahaa**. (...) mis=`sina=`Mart arvad.

Näites (10) valib õpetaja vastama Kertu küsimusega *mis sina arvad*: selline küsimuse sõnastus osutab, justkui oleks küsimus tõesti õpilase arvamuse kohta, kuigi tegelikult on küsimus teadmiste kontroll. Seepeale saab õpetaja Kertult dialoogi reas 2 õige vastuse *kausse* ning kinnitab seda laiendatud kordusega: *kausse oli kolm* ning lisab kordusele uue info vastuvõttu tähistava partikli *ahaa*. Siinkohal on võimalik tõlgendus, et õpetaja justkui jätkab “tõeliste”, s.t argivestluses esinevate küsimussekventsides formaati ning annab ka vastava tagasiside: osutab *ahaa*-ga, et info oli tema jaoks uus. Samas on info uudsuse osutusele eelnenud erinevalt näiteks tõelistelt uut infot taotlevatel küsimussekventsides vastuse õigsuse kinnitus.

Teine võimalus on tõlgendada õpetaja laiendatud kordust reas 4 vastuvõtukordusena, mitte kinnitusena. Sel juhul registreerib õpetaja Kertu vastuse, kuid ei ütle, kas see oli õige või vale, ning lisab informeerituse muutust tähistava partikli *ahaa*. Ka sel juhul kasutab õpetaja n-õ petustrateegiat, reageerides teadmisi kontrolliva küsimuse õigele vastusele neutraalse info uudsusele osutava vastuvõtuteatega, et julgustada järgmist õpilast vastama.

• Õpetaja kasutas info uudsusele osutavaid partikleid ka siis, kui õpilase vastus ei vastanud õpetaja ootustele, näiteks ei olnud kooskõlas õpetaja pedagoogiliste eesmärkidega. Sellisel juhul järgneb “tugevam”, kinnitav tagasisidekeelend alles pärast soovitud vastuse saamist.

Vaatame näiteks katket esimese klassi matemaatikatunnist.

(11) ((1. klass, matemaatika, õpetaja luges lastele muinasjuttu kolmest karust))

1. Õ: aga=`mis`õpetuse=`me`võiks Mašale`öelda.`mida=`ei`tohiks`teha.

2. (...)

3. Ander: ma=`tean?=`

4. Õ: `mis=`ta`valesti tegi. (...) mis=`sa`arvad.

5. (...) ((osa lapsi tõstab kätt))

6. Õ: Andreas.

7. Andreas: et ta `karu `majja läks.
 8. (.)
 9. Õ: **aha:h** aga `ENNE seda.
 10. Mait: mm ((tõstab kätt))
 11. (...)
 12. Õ: noo?
 13. Mait: ee=et ta `ära eksis. (...) ja `üks oli `ka:=vel=et (...) ee et
 `kõige=lõpus on=ju `see 14. ka: kui=ta (...) {hullem=et=ta} `putru panid.
 15. Õ: **[ahah]**
 16. Mait: [`karud] ei `teinud talle `midagi.
 17. Õ: **ei=ei=neid `karusid=tea= `midagi**=aga mis=te `veel `mina üt-
 leks et `ENNE seda > 18. kui=ta `karude juurde jõudis < `mis=ta `valesti
 `tegi.
 19. (...)
 20. Riiivo: {-}.
 21. Õ: **seda `ka.**
 22. (...) ((Heli tõstab kätt))
 23. Õ: Heli.
 24. Heli: ta ronis `aknast `välja.
 25. (.)
 26. Õ: **ei= `noh ta pidi `põgenema=> sest ega=neid `karusid**
ei= `tea. aga=ma= `arvan=et 27. `üksi kas `tohib `väike [`laps `metsa
 `minna.]
 28. Andrus: [mul=on `valmis]
 ((Andrus tõstab kätt))
 29. L-d: e:i
 30. (.)
 31. Õ: < **va:t `õige.** >
 32. Andrus: mul=on `valmis [õpetaja]((Andrus tõstab töövihiku üles))
 33. Õ: [`metsa] `läheme `ikkagi `koos `ema=või
 `isaga aga `mitte `üksi 34. `päini.
 35. Triinu:= või `vanaemaga.
 36. Õ: **`vä:ga `õige.**

Reas 1 esitab õpetaja küsimuse, millel on ilmne eesmärk teha loetud muinasjutust õpetlik järeldus. Seejärel saab ta mitu sisuliselt võimalikku õiget vastust: *et ta karu majja läks* (rida 7), *et ta ära eksis* (rida 13), *ta ronis aknast välja* (rida 24). Need vastused ei ühti aga õpetaja ideega soovitud vastusest, see peegeldub selgelt õpetaja reaktsioonivoorude ehituses: vastuvõtuteade *ahah*+täpsustav küsimus (rida 9), vastuvõtuteade *ahah* pealerääkimisena (rida 15), kommentaar +täpsustav küsimus (read 17–18), distantseeruv kinnitus *seda ka* (rida 21), kommentaar+täpsustav küsimus (read 26–27). Kui õpilased soovitud vastuseni ei jõua, muutuvad ka küsimuste vorm ja/või sisu täpsemaks; viimane küsimus, millele lõpuks soovitud vastus tuleb, on suletud *kas*-küsimuse vormis: *üksi kas väike laps tohib metsa minna*. Näitest on näha, et tagasiside kinnituse “tugevus” on astmeline. Õpetaja reageerib voorudele, mis on küll õiged, kuid ei ole sisult sellised, mida tema vastuseks ootas, vastuvõtuteatega *ahah*, distantseeruva kinnitusega *seda ka*, kommentaaridega. Alles pärast iseenda sõnastatud suletud *kas*-küsimusele (rida 27) saadud vastust

kasutab ta “tugevamat” kinnitavat keelendit *vat õige* ja ülirohutatud kinnitust *väga õige*, reageerides Triinu 35. reas lisatud täpsustusele.

3.4. Piiritleja

Piiritleja osutab nagu vastuvõtuteadegi, et info on vastu võetud ja sekvents, teema või alateema on lõppenud, kuid tähistab sekventsi piire tugevamat. Piiritleja funktsioonis võivad olla näiteks keelendid *nii*, *kena* ja tänusõnad *aitäh*, *tänan*. Tänusõnad *aitäh* ja *tänan* ning piiritlev keelend *kena* paiknesid alati sekventsi lõpus. Partikkel *nii* esines nii alasekventsi piiril, enne täpsustavat küsimust kui küsimussekventsi lõpus. Tegelikult, üldisemalt võttes, on piiritlev funktsioon kõigil õpetaja tagasisidevahenditel, mis õpilase jätkamist ei taotle. Prototüüpne piiritlev tagasiside on näiteks *kena* näites (12) real 3.

(12) ((3. klass, matemaatika, tekstülesande lahendamise))

1. Õ: ja= `mis oleks sis `vastus.

2. Epp: {kolm}

3. Õ: `kolm=nööpi=sest < `kuuele (.) `kleidile (.) `õmmeldi. > (...) **ee**
`**kena**. (.) `pane=oma 4. (.) ee (...) * ei `pane `kaa *

5. L-d: hehehe

6. Õ: `õpikust `ülesanne `kaks=sada= `õheksa teeme `kirjalikult
`vihikusse

Ka partiklil *nii* on eesti keeles piiritlev funktsioon. Renate Pajusalu on *niid* analüüsinud kui teema algatajat ja lõpetajat raadiointervjuudes (Pajusalu 1999: 71), Leelo Keevallik üldisemalt kui piiritlejat, mis tähistab üleminekut ühest (suhtluslikust) tegevusest teise (Keevallik 2000: 350–351). Käesoleva uurimuse raamistikus on *nii* piiritlev vastuvõtuteade.

Kinnituse lisatähendus võib *nii*-le lisanduda implitsiitselt juhul, kui ta on reaktsioon õpilase teadmisi kontrolliva küsimuse vastusele (vrd all esitatud näiteid (13) ja (14))

(13) ((3. klass, eesti keel, õpetaja küsis, mis vahe on tontidel ja haldjatel))

1. Õ: Paul?

2. Paul: `tondid on nagu: mm rohkem e `kurjemad ja `haldjad on `head.

3. Õ: nii, palun? ((õpetaja vaatab Peebu poole))

(14) ((3. klass matemaatika; õpetaja ütleb arvu, õpilased peavad ütleva selle tegurid))

1. Õ: `kakskümmend= `üks.

2. ((lapsed tõstavad kätt, õpetaja noogutab Katrele))

3. Katre: `kolm=korda= `seitse.

4. Õ: * **nii**. * `kakskümmend= `neli.

Mõlemas näites on õpetaja tagasisidekeelend *nii* täpselt ühesuguses sekventsiaalses positsioonis. Erinev on vaid see, et näites (13) küsib õpetaja õpilase arvamuse kohta, näite (14) küsimus kontrollib aga õpilase teadmisi. Mõlema näite *nii*-l on ühesugune ülesanne: osutada küsimussekventsi lõpule, mõlemal juhul algatab õpetaja pärast *nii*-d uue küsimussekventsi. Kuna näite (14) puhul on tegemist

õpilase teadmisi kontrolliva küsimusega, lisandub *nii*-le implitsiitselt vastuse kinnitamise roll.

Ülejäänud *nii* kasutuskontekstid olid kahest eelmisest näitest mõnevõrra erinevad, õpetaja kasutas *nii*-d tagasisidena nii teadmisi kontrolliva kui õpilase kohta käiva küsimuse vastusele siis, kui ta soovis saadud vastust täpsustada, nagu näites (15). Selline *nii* kasutuskontekst on võrreldav neutraalse vastuvõtuteate kasutamisega tagasisidena teadmisi kontrolliva küsimuse vastusele, vrd (9) ja (11). Õpetaja saab õpilaselt vastuse, mis ei vasta mingil põhjusel tema ootustele, kuid mis ei pruugi olla põhimõtteliselt vale; ning annab tagasiside, mille kinnituse tähendus puudub või ei ole selgepiiriline; ning algatab vastuse täpsustamise. Pärast eeldatud vastuse saamist reageerib õpetaja “tugevama” kinnitava keelendiga.

- (15) ((1. klass, eesti keel))
1. Õ: `miks Anna `nuttis.
 2. (...) ((osa lapsi tõstab käe))
 3. Õ: ee `Mait.
 4. Mait:et: ta=i leidnud `koduteed.
 5. Õ: **nii**. mis=sis `Annaga oli `juhtunud.
 6. (...)
 7. Õ: `mis `oli `Annaga `juhtunud.
 8. (...)
 9. Õ: `Mariliis.
 10. Mariliis: {ta eksis `ära. }
 11. Õ: õige.

Õpetaja luges lastele ette jutukese Annast, kes suures linnas ära eksis, ja küsib nüüd jutus toimunu kohta. Reas 1 esitab õpetaja küsimuse *miks Anna nuttis* ning saab sellele Maidult põhimõtteliselt õige vastuse: *ta=i leidnud koduteed*. Seepeale alustab õpetaja oma voozu partikliga *nii* ning esitab täpsustava küsimuse ja valib uue vastaja, Mariliisi, kes sõnastab täpsustuse *ta eksis ära*, millele õpetaja reageerib tugevalt kinnitava keelendiga *õige*. Keelendil *nii* 5. reas on vähemalt kaks ülesannet: esiteks on ta piiritlev vastuvõtuteade, mis viitab, et alasekvents on lõppenud ja eelmises voozus sisaldunud info on vastu võetud. Teiseks, kuna tegamist on teadmisi kontrolliva küsimusega, lisandub implitsiitselt kinnitav lisatähendus, mis ei ole siiski *nii* tugev ja selge kui 11. reas sõnastatud tagasisidel *õige*.

Kui jälgida *nii* intonatsiooni, siis küsimussekvensi lõpus oli *nii* hääldatud alati langeva intonatsiooniga (14). Kui õpetaja soovis vastust mingil põhjusel täpsustada või palus samale küsimusele vastata mitmel õpilasel, kasutas ta enne täpsustava küsimuse esitamist või uue vastaja valikut *nii* langeva kui poollangeva intonatsiooniga *nii*-d (13), (15).

4. Eelmise kõneleja jätkamist taotlevad tagasisided

Omaette rühma moodustavad õpilase jätkamist taotlevad õpetaja tagasisidekeelendid. Nendel keelenditel puudub piiritlev funktsioon, nad annavad kõnevooru tagasi õpilasele, kuid määratlevad õpilase järgmise kõnevooru sisu erinevas ulatuses. Minimaalsed jätkamist taotlevad keelendid (materjalis *jaa*, *nii*, *veel*, *kordus*, *noogutus*) jätavad järgmise voozu sisu täielikult õpilase otsustada (küsimussekvensi puhul

määrab sisu muidugi eespool esitatud küsimus). Teine võimalus on, et õpetaja esitab saadud vastuse kohta täpsustava küsimuse, et teemat edasi arendada või algatab paranduse, kui õpilase vastus oli tema jaoks mingil põhjusel probleemne. Tabelist 2 saab teha kolm olulist üldistust.

- Õpetaja kasutas minimaalset jätkajat, mis järgmise vooru sisu kuidagi ei määratle, äärmiselt vähe (20 juhul). Kui õpetaja soovis, et õpilane vastamist jätkaks, esitas ta täpsustava küsimuse (114 juhul). Juhul kui vastus oli mingil põhjusel probleemne, algatas õpetaja parandussekventsi (67 juhul).
- Õpetaja ei kasutanud argivestlustes ja muud tüüpi ametivestlustes jätkaja funktsioonis tavaliselt kasutatavat partiklit *mhmh* jätkajana kogu korpus kes mitte ühtegi korda. Kõik korpus kes leidunud *mhmh*-id olid kas vastuvõtu-teate või kinnituse rollis. Siinkohal on hea võrdlusmaterjal Tiit Hennoste statistika, mille järgi moodustas *mhmh* ametliku suhtluse korpus kes (ei sisalda koolitunde) 9,7% sõnavormidest ja argikeele korpus kes 6,2% sõna-vormidest (Hennoste 2000b: 1780). Analüüsitavas koolitundide korpus kes moodustab *mhmh* 0,08% sõnavormidest.

Korpus kes leidis üks huvitav näide, kus õpilase ja õpetaja tõlgendus *mhmh*-i tähendus es t on erinev. Näide on katke 1. klassi eesti keele tunni osast, kus õpetaja palub õpilastel öelda, mis loom või lind nad tahaksid olla.

(16) ((1. klass eesti keel, õpetaja küsis, mis loom või lind lapsed olla tahaksid ja miks))

1. Õ: e `Angela

2. . Angela: kui=ma=`oleksin `lind sis=ma oleksin `kalju `kotkas.

3. (.)

4. Õ: * miks. *

5. Angela: `sellepärast=et na:d on (.) `nii ilu- > mulle=`meeldivad nende `sulestik s.t nad on 6. äs- nad=on=ästi=ilsad `pruunid=ja `suured. <

7. Õ: * *mhmh* * ((õpetaja noogutab kätt tõstva Kristini suunas))

8. Angela: mulle `meeldivad `suured `linnud.

9. (.)

10. Kristin: kui=ma `oleksin `loom (...) `siis `ma `oleksin koer.

Näite 1. reas valib õpetaja vastama Angela ja saab temalt 2. ja 5.–6. reas vastuse. 7. reas reageerib õpetaja vastusele partikliga *mhmh* ja suunab pilgu järgmisele õpilasele, Kristinile. See annab aluse tõlgendusele, et õpetaja kasutas *mhmh*-i vastuvõtuteatena ja algatas pilguga järgmist vastajat valides uue küsimussekventsi, kuna Angela vastas eeskujulikult esitatud küsimuse mõlemad pooled. Kuna *mhmh* aga ei ole piisavalt tugev piiritleja ning õpetaja verbaalselt uut sekventsi ei algatanud, andis see Angelale võimaluse tõlgendada õpetaja *mhmh*-i jätkajana ning oma vastust täpsustada.

- Vastupidiselt teoreetilisele eeldusele, et õpetaja kasutab tõenäoliselt minimaalset jätkajat rohkem õpilase kohta käivate küsimuste vastustele reageerides, näitas reaalse materjali analüüs vastupidist tulemust: õpetaja kasutas minimaalset jätkajat õpilase kohta käivate küsimuste vastustele reageerides kogu korpus kes vaid 1 korral, teadmisi kontrollivatele küsimustele reageerides aga 19 korral, neist 14 juhul lisandus jätkajale eelmise vooru kinnitamise funktsioon.

4.1. Minimaalsed jätkamist taotlevad tagasisidekeelendid

Minimaalsed jätkamist taotlevad tagasisidekeelendid esinesid pea kõigil juhtudel reaktsioonina teadmisi kontrolliva küsimuse vastusele. Vaid ühe õpilase kohta käiva küsimuse vastusele reageeris õpetaja mitteverbaalselt, noogutusega. See oli ka ainuke korpuses fikseeritud mitteverbaalne jätkav tagasiside. See tulemus ei kajasta mitteverbaalse jätkava tagaside reaalsel hulka materjalis, vaid on seotud filmimise tehniliste puudujääkidega, kuna kogu õpetaja tegevus ei ole alati kaamera fookuses. Ülejäänud minimaalsed jätkavtagasisidevahendid olid keelendid *jaa*, *niis*, *veel* ning vastuvõtukordus. Need (välja arvatud üks langeva intonatsiooniga häälstatud *veel*) olid häälstatud tõusva intonatsiooniga ja paiknesid positsioonis, kus küsimus eeldas pikemat või mitmeosalist vastust. Sellise tagasisidega annab õpetaja märku, et sama õpilane (mitte keegi teine) võib vastamisega jätkata ning ühtlasi kinnitab juba saadud vastuse osa.

(17) ((3. klass matemaatika))

1. Õ: `aa=korda `kolm `korrutised (...) on `Karolin.

2. Karolin: aa `õheksa=korda= `kolm=võrdub= `kakskest `seitse.

3. Õ: **jaa?**

4. (.)

5. Karolin: `kuus=korda= `kolm=võrdub `kaheksa `teist=ja
`kaheksa=korda= `kolm (.)

6. kakskest[{neli}]

7. Õ: [`väga=]tubli.

Näites (17) osutab õpetaja 3. reas öeldud tõusva intonatsiooniga *jaa*-ga, et Karolin loeks ette ka teised töövihikus olevad kolmega korrutised ning ühtlasi kinnitab ka juba saadud vastust.

4.2. Järgmise voo sisse sisu suunav tagasiside

Kui õpetaja soovib oma algatatud küsimussekventsi laiendada, on tal kaks põhimõttelist võimalust: esitada vastuse kohta täpsustav küsimus või algatada vastuse sisule ja/või vormile keskendatud parandussekvents.

Täpsustava küsimuse esitab õpetaja juhul, kui vastus on tema arvates ebapiisav või teema edasiarendamist väärt. Sekventsi lõpetava tagasiside annab õpetaja siis pärast õpilaselt saadud täpsustusi. Näites (18) esitab õpetaja täpsustava küsimuse õpilaselt reas 2 saadud vastuse kohta ning lõpetab küsimussekventsi hinnangulise reaktsiooniga *sis on `tore*.

(18) ((1. klass, eesti keel))

1. Õ: `Arvi, `milline=on `sinu uudis.

2. Arvi: {--} sõitsin `rattaga=ja `kukkusin `maha.

3. (...)

4. Õ: * `said=**sa** `väga haiget. *

5. Arvi: ei= `saand. ((Arvi raputab pead))

6. (.)

7. Õ: * sis=on `tore. *

8. (...) ((Tarmo tõstab kätt))

9. Õ: `Tarmo?

Paranduse algatab õpetaja juhul, kui õpilase vastus oli vale või oli selles mingil põhjusel (näiteks kuulmis- või arusaamisprobleem) midagi parandamist-täpsustamist vajavat. Parandusalgatused on suunatud tekstis tahapoole, nad tegelevad juba väljaöeldud teksti toimetamisega. Vale või õpetaja arvates ebasobiva vastuse korral on õpetaja peamiseks tegutemisstrateegiaks paranduse algatamine, mitte aga parandamine või negatiivse tagasiside andmine.

Analüüsitud korpuses oli kokku 82 õpetaja jaoks probleeme tekitanud vastust, neist 67 juhul algatas ta seepeale paranduse, andes probleemi lahendamise võimaluse probleemi allika väljaütlejale või mõnele teisele õpilasele. Ülejäänud 15 juhul parandas õpetaja probeeme tekitanud vastuse ise.

Näites (19) on õpetaja parandusalgatuse (real 4) põhjus õpilase vale vastus küsimusele *aga`mida seal=pildil`jänestel kokku on`ka`veel`kolm.*

(19) ((1. klass matemaatika))

1. Õ: aga`mida seal=pildil`jänestel kokku on`ka`veel`kolm.

2. (...)

3. Siim: kõrva.

4. **Õ: kas=`on ko- e:i`kõrvasid=on`rohkem, ko-`kõrvu=on`rohkem.**

5. (...) ((osa lapsi tõstab kätt))

6. Õ: Andrus?

7. Andrus: nina.

8. Õ: < õige, >`kolm`jänest,`kolm`nina. (.)`vä:ga=tubli.

5. Küsimussekventsia struktuur

Tabel 3 visandab ülevaate õpetaja algatatud küsimussekventsia ülesehitusest. Analüüsitavas materjalis leidis nii klassikalisi kolmikuid, nende laiendusi kui ka kahest komponendist koosnevat küsimussekventsia.

Tabel 3. Õpetaja algatatud küsimussekventsia struktuur

Küsimussekventsia struktuur					
Õ: küsimus ? L: vastus ? Õ: tagasiside + uue sekventsia algatus	Õ: küsimus ? L: vastus ? Õ: täpsustav küsimus/ parandusalgatus ? L: vastus ? Õ: tagasiside+uue sekventsia algatus	Õ: küsi- mus ? L: vastus ? Õ: uue sekventsia algatus	Õ: küsimus ? L: vastus ? Õ: jätkaja ? L: vastus ? Õ:tagasiside+ uue sekventsia algatus	Õ: küsimus ? L: vastus ? Õ: tagasiside ? L: täpsustamine ? Õ: tagasiside+ uue sekventsia algatus	Vastuseta küsimus
135	97	44 ¹	20	4	7
Kokku					310

¹ Tabelis märgitud 44 juhust 38-l on õpetaja kaamera fookuses, s.t on kindel, et õpetaja ei andnud vastusele mitte-verbaalset tagasisidet ja algatas kohe uue sekventsia. Ülejäänud 6 juhul on kaamera suunatud õpilasele, võimalikku mitte-verbaalset tagasisidet ei ole fikseeritud

Küsimussekventsides vormi analüüs peegeldab, et õpetaja juhib kindlalt tunni sisu ja kulgu. Mis tahes õpilasepoolne algatus/edasiarenduse võimalus on väike: kõik küsimussekvendid lõppevad õpetajapoolse uue küsimuse või tegevuse algatusega², enamikku õpetaja algatatud küsimussekventsides (84%) suleb õpetaja tagasisidevoor. 31% juhtudest eelneb õpetaja küsimussekvendi lõpetavale reaktsioonile vahesekvents, mis tegeleb eelmise voo, s.o õpilase voo parandamisega (parandusalgatused) või esitab õpetaja vastuse kohta täpsustavaid küsimusi. Vastaja jätkamist taotlevat tagasisidet, mis vastuse sisu kuidagi ei piiritle, annab õpetaja väga vähe (0,6% juhtudest).

6. Kokkuvõtte ja arutelu

Analüüs näitas, et algkoolitunnis juhib õpetaja kindlalt tunni käiku, õpilaste oma-poose algatuse võimalus on minimaalne. See avaldus järgnevas.

- 84% õpetaja algatatud küsimussekventsides sulges tagasisidevoor. Pärast tagasiside andmist esitas õpetaja kohe uue küsimuse või algatas uue tegevuse.
- Õpilase jätkamist taotlevat tagasisidet kasutas õpetaja 31% juhtudest. Minimaalset jätkajat kasutas õpetaja äärmiselt vähe, vaid 0,6% juhtudest; pea kõigil juhtudel reaktsioonina teadmisi kontrolliva küsimuse vastusele kontekstis, kus eeldatav vastus oli mitmest osast koosnev. Ka n-ö vabama vestluse kontekstis, kus õpetaja küsis õpilase kohta, sulges õpetaja piiritleva tagasisidega sekvendi kas kohe pärast saadud vastust või pärast esitatud täpsustavat küsimust. Tulemus on üllatav, kuna õpilase kohta käivad küsimused olid suuremalt jaolt esitatud kontekstis, kus õpetaja pedagoogiline eesmärk oli arendada vestlust ja anda õpilastele rääkimise võimalus. Minimaalse jätkaja kasutamine selles kontekstis suurendaks õpilase iseseisva teemaarenduse võimalust. Selle formalistliku skeemi taga on tõenäoliselt õpilaste suur hulk klassis: tunni aeg on piiratud ja õpetaja peab andma võimaluse rääkida võimalikult paljudele. Piiritlevas funktsioonis tagasiside ja täpsustavad küsimused võimaldavad tunni sisu ja voo vahetust kindlalt juhtida. Õpilased järgisid skeemi pea sajaprotsendiliselt: esines vaid 4 juhtu, kus õpilane jätkas vastamist pärast õpetaja sulgevat tagasisidet.

Koolitundide kolmikut on palju kritiseeritud, eriti võõrkeeleõpetuse kontekstis, peamiselt seetõttu, et õpetaja domineerib ja suhtlus erineb reaalsest vabast suulisest suhtlusest (vt näiteks Dinsmore 1985; Nunan 1987). Tegelikult on kolmikstruktuur traditsioonilises koolitunnis levinud kõigis vanuseastmetes ning sage ka eelkooliealise lapse ja vanema vestluses. Seega võib üldistada, et kolmikstruktuur on mis tahes õppimis- ja õpetamissüsteemis loomulik. Vestlusanalüütilise lähenemise järgi on koolitund sama "loomulik" kui argivestlus, koolitunnisuhtlust juhivad lihtsalt oma reeglid, koolitund on institutsionaalse suhtluse variant (Seedhouse 2004: 75). Kolmiku ülivaldav kasutus analüüsitud tundides peegeldab, et õpetaja ja õpilase rolliaotus on selles vanuseastmes rangelt asümmeetriline. Kolmikstruktuur võimaldab õpetajal õpilase vastust hinnata ning tunni käiku ja voo vahetust reguleerida.

Õpetaja tagasiside on paindlik ja hulga tähendusnäanssidega nähtus, tagasiside konkreetne tähendus ja funktsioon selgub reaalselt kasutus konteksti analüüsides.

² Mõõdamines võiks mainida, et õpilase esitatud küsimusi oli vaadeldud tundides väga vähe (7% küsimustest).

Teadmisi kontrollivate küsimuste vastustele kaldub lisanduma kinnituse, õpilase kohta käivate küsimuste vastustele hinnangu tähendus. Tagasiside kinnituse tugevus on astmeline, skaala ühes otsas on neutraalsed info vastuvõtuteated, mille kinnituse tähendus puudub, teises otsas keelend *õige*. Näiteks võib tagasisidena teadmisi kontrolliva küsimuse vastusele, millega õpetaja rahule ei jäänud, kasutada neutraalset ja/või piiritlevat vastuvõtuteadet, mille kinnituse tähendus puudub, ning anda “tugevama” kinnitava tagasiside alles pärast soovitud vastuse saamist.

Uurimus visandab esimese, üldistava ülevaate õpetaja algatatud küsimussekventside struktuurist ja õpetaja tagasiside funktsioonist. Autori edasine eesmärk on arvestada küsimussekventside analüüsil täpsemat konteksti. Selle idee taga on teoreetiline eeldus, et õpetaja pedagoogiline eesmärk muutub tunni eri osades ja sellest võib sõltuda ka küsimussekventside struktuur ja tagasiside funktsioon. Erinevad on näiteks kontekst, kus õpetaja ootab oma küsimusele täpset vastust (nt keeletunnis kindlat sõnavormi, matemaatikatunnis arvutusülesande vastust); vestluse arendamise kontekst, kus õpetaja eesmärk on panna lapsed rääkima; uute teadmiste omandamise kontekst, kus õpetaja esitab küll teadmisi kontrollivaid küsimusi, kuid õiged vastused esitatud küsimustele ei pea olema vormilt ja sisult sedavõrd täpsed kui näiteks matemaatikatunni arvutusülesannete puhul. Huvitavaid tulemusi võiks pakkuda ka küsimussekventside võrdlus klasside kaupa. Samuti on tulevase töö seisukohalt oluline parandada filmimise tehnikat, et jäädvustada analüüsi seisukohalt väga oluline mitteverbaalne suhtlus.

Loodetavasti annavad juba saadud ja tulevikus täpsustuvad teadmised tagasiside kohta õpetajale uusi võimalusi rakendada keelelisi/suhtluslikke vahendeid oma pedagoogiliste eesmärkide teenistusse..

Kirjandus

- Dinsmore, David 1985. Waiting for Godot in the EFL classroom. – *ELT Journal* 39, 225–234.
- Drew, Paul; Heritage, John 1992. Analyzing talk at work. – Paul Drew, John Heritage (Eds.). *Talk at Work. Interaction in Institutional Settings*. Cambridge University Press, 3–65.
- Hennoste, Tiit 2000a. Sissejuhatus suulisesse eesti keelde VI. Lausung suulises kõnes I. – *Akadeemia* 10, 2223–2254.
- Hennoste, Tiit 2000b. Sissejuhatus suulisesse eesti keelde IV. Suulise keele erisõnavara III. – *Akadeemia* 8, 1773–1806.
- Hennoste, Tiit; Rääbis, Andriela 2004. Dialoogiaktid eesti infodialoogides: tüpoloogia ja analüüs. Tartu Ülikool, arvutiteaduse instituut. Tartu: Tartu Ülikooli Kirjastus.
- Hutchby, Ian; Wooffitt, Robin 1998. *Conversation Analysis. Principles, Practices and Applications*. Cambridge: Polity Press.
- Kasterpalu, Riina 2005. Partiklid *jah*, *jaa* ning *jajaa* naaberpaari järelliikmena müügiläbi-rääkimistes. – *Keel ja Kirjandus* 12, 996–1000.
- Keevallik, Leelo 2000. Keelendid *et* ja *nii et* vestluses. – *Keel ja Kirjandus* 5, 344–358.
- Keevallik, Leelo 2003. From Interaction to Grammar. Estonian Finite Verb Forms in Conversation. *Acta Universitatis Upsaliensis. Studia Uralica Upsaliensia* 34. Uppsala.
- Mehan, Hugh 1985. The Structure of Classroom Discourse. – Teun A. van Dijk (Ed.). *Handbook of Discourse Analysis. Vol. 3. Discourse and Dialogue*. London: Academic Press, 119–131.

- Nunan, David 1987. Communicative Language Teaching: Making it work. – *ELT Journal* 41 (2), 136–145.
- Pajusalu, Renate 1999. Deiktikud eesti keeles. *Dissertationes Philologiae Estonicae Universitatis Tartuensis* 8. Tartu: Tartu Ülikooli Kirjastus.
- Sacks, Harvey 1992. *Lectures on Conversation*. Gail Jefferson (Ed.). Cambridge-Oxford: Blackwell.
- Seedhouse, Paul 2004. *The Interactional Architecture of the Language Classroom: A Conversation Analysis Perspective*. Blackwell Publishing.
- Sinclair, John; Coulthard, Malcolm 1975. *Toward an Analysis of Discourse*. Oxford: Oxford University Press.

Krista Strandsoni (Tartu Ülikool) uurimisvaldkonnad on suuline keel, parandused suulises keeles, infodialoogide struktuur, koolitunnisuhtlus.
krista.strandson@ut.ee

TEACHER'S FEEDBACK TO STUDENTS' ANSWERS IN PRIMARY SCHOOL LESSONS

Krista Strandson

University of Tartu

The objective of the paper is to analyse teacher's feedback to students' answers and to outline the structure of teacher-initiated question sequences in Estonian primary school lessons. The teacher's feedback is analysed from two points of view. First, how is the teacher's feedback related to the content of the question, i.e. does the teacher's feedback to the answer to a question aimed to check a pupil's knowledge differ from her feedback to the answer to a question about the pupil? Second, what is the function of the teacher's feedback; how does the teacher's feedback determine the course of further interaction? The survey is based on a transcribed collection of video-recordings (280 min in total) of Estonian and mathematics lessons that took place in 1st-3rd forms at a school in Tartu, Estonia. The main results of the analysis are as follows. The teacher leads the course of the lesson with utmost resolution. The possibility of student-initiated activity is minimal. This is illustrated by the following facts. (1) The teacher's feedback terminating the question sequence is predominant irrespective of the content of the questions. The majority of the question sequences initiated by the teacher constitute a triadic structure (question-answer-feedback) even if the teacher's pedagogical aim is to develop a conversation. (2) The teacher uses next to no minimal continuers as feedback. The majority of the teacher's feedback encouraging the pupils to continue are specifying questions about the answer just given.

Keywords: conversation analysis, classroom interaction, question sequences, reaction to the answer, Estonian

VEENE-EESTI KOODIVAHETUSE KORPUS: KODEERIMISPÕHIMÕTETE VÄLJATÖÖTAMINE

Anastassia Zabrodszkaja

Ülevaade. Artikli eesmärk on anda ülevaade Tallinna Ülikooli üld- ja rakenduslingvistika õppetooli juures valmivast vene-eesti koodivahe- tuse korpusest, mis pannakse tulevikus Tallinna Ülikooli eesti filoloogia osakonna kodulehele. Artiklis tutvustatakse lühendeid LIPPS ja LIDES ning nendega seonduvat, mh rahvusvahelise kakskeelse (*resp.* mitmekeelse) andmebaasi peamisi standardseid transkribeerimis- ja kodeerimispõhimõtteid¹ vene-eesti koodivahe- tuse näitel. Artikli viimases osas analüüsitakse sisestava (valdavalt ühesõnalise) vene-eesti koodivahe- tuse morfoloogia kodeerimisvõimalusi, sõnastatakse peamised üleskerkinud probleemid ning pakutakse ka võimalikke lahendusi.*

Võtmesõnad: korpuslingvistika, koodivahe- tuse, eesti keel, vene keel

Sissejuhatus: mis on LIPPS ja LIDES

Viimastel kümnenditel on kakskeelsuse uurimine kujunenud keeleteaduse ise- seisvaks valdkonnaks, selle uurijad on erinevatest keelekogukondadest kogunud märkimisväärse hulga andmeid. Ent erinevate uurimiseesmärkide ja -võimaluste tõttu polnud materjali üleskirjutamises ühtsust, mis raskendas keelelise materjali vahetamist ning tõstis esile vajaduse luua kakskeelse keeleainese transkribeeri- misstandard. Kuna autorid kirjeldasid koodivahe- tuse juhtumeid eri termineid kasutades, väites, et just nende terminid on paremad, täpsemad ja sobivamad, rajas Euroopa Teadusfond (European Science Foundation) 1990. aastate esimesel poolel koodivahe- tuse ja keelekontaktide võrgustiku (Network on Code-Switching and Language Contact), mille eesmärk oli selgete ja üldaktsepteeritavate koodiva- hetuse põhimõistete arendamine (Milroy, Muysken 1995: 12).

Koodivahe- tuse ja keelekontaktide võrgustik soodustas keelte interaktsiooni andmevahetussüsteemi loomist. Keelte interaktsiooni ehk keelte vastasmõju (ingl *language interaction*) uurimine keskendub kakskeelsele (mitmemurdelisele)

¹ Artiklis kasutatakse terminit *kodeerimine*, lähtudes ingliskeelsest vastest *encoding*.

* Käesolev artikkel on kirjutatud ETF-i grand nr 6151 "Koodivahe- tuse, eesti vahekeele ja lastekeele andmekorpuse koostamine ja üldkirjeldus" raames. Autor tänab prof Martin Ehalat, prof Anna Verschikot ja dr Ad Backust asjakohase terminoloogia arutamise ja kasulike nõuannete eest.

keelekasutusele erinevatest sotsiaalsetest ja lingvistilistest vaatenurkadest. Rohkem kasutatavate terminite *koodivahetus* (ingl *code-switching*) ja *keelekontakt* (ingl *language contact*) asemel valiti termin *keelte vastasmõju*, et hõlmata kõiki keelelisi nähtusi, mitte ainult neid, mis kuuluvad kahte eri süsteemi (LIPPS Group 2000: 133). Andmevahetussüsteem suurendas üldise transkribeerimis- ja kodeerimisjuhendi vajalikkust ning tõstis probleemi veel rohkem esile. See tähendas omakorda, et transkribeerimis- ja kodeerimisstandardite kehtestamisel pidi jõudma üksmeelele, võimaldamaks teiste keeleteaduse harudega (näiteks diskursusanalüüsi või süntaksiga) tegelevatel uurijatel keelte vastasmõju andmevahetussüsteemist samuti kasu saada.

Ljouwertis/Leeuwardenis 1994. aastal toimunud konverentsil tõdeti, et koodivahetuse uurijad saavad oma andmeid vahetada ainult eraviisiliselt, sest puudub formaliseeritud süsteem, standard jms (LIPPS Group 2000: 133). Samal ajal on lastekeele uurijatel olemas ühtne standardne viis andmete transkribeerimiseks ja kodeerimiseks – CHILDES (Children Language Data Exchange System, umbkaudu: lastekeele andmevahetussüsteem). Seetõttu on olemas ka andmebaasid, mida võib hõlpsasti omavahel vahetada. CHILDES võimaldab eri riikide keeleteadlastel jagada lastekeele materjali ja teha võrdlevaid rahvusvahelisi uurimusi ning analüüse.

Ülaltoodud asjaolud ajendasid LIPPS-i (Language Interaction in Plurilingual & Plurilingual Speakers, umbkaudu: mitmekeelsete kõnelejate keeleline vastasmõju) uurimisrühma loomist ja mitmekeelse interaktsiooni kodeerimisjuhendi LIDES (Language Interaction Data Exchange System) koostamist. Nii ilmuski mitme uurija pika koostöö tulemusel 2000. aastal ajakirja *International Journal of Bilingualism* erinumbrina “The LIDES Coding Manual”. Sellel oli kaks peamist eesmärki:

- 1) pakkuda vilunud uurijatele kakskeelse ainese transkribeerimise ja kodeerimise juhtnööre olemasolevat arvutitarkvara kasutades;
- 2) julgustada algajaid uurijaid kakskeelseid andmeid kodeerima, toetudes üldstandardile, ning anda nende võimalikele küsimustele lihtsamas vormis vastused.

LIDES-i ettevalmistamisel kasutati hüppelauana Brian MacWhinney (1995) CHILDES-i projekti, mida tuli mugandada kakskeelse vestluse analüüsi jaoks sobivamaks (pidi tähistama selliseid ükskeelsest kõnest puuduvaid nähtusi nagu grammatiline ja semantiline transferents, leksikaalsed laenud, koodivahetus, koodisegu, diskursuse mallid jne).

Siinjuures tuleb mainida, et LIDES annab keeleliste nähtuste ja nende tähistuste loendi ning loomise põhimõtted. Igale uurijale tekivad aga oma täiendavad vajadused: näiteks ühte huvitab pigem mitmekeelse vestluse struktuur ja kõnevoorude ülesehitus, teist koodivahetuse grammatilised omadused jne. Seetõttu tuleb lisaks konventsionaalsetele põhimõtetele kodeerijal välja mõelda võtteid, mis lubaksid kajastada just teda huvitavaid seiku. Seda asjaolu näitlikustab ka vene-eesti koodivahetuse korpus. Järgmisena tutvustatakse selle koostamisprintsippe, -probleeme ja lahendusviise.

1. Vene-eesti koodivahetuse korpuse eellugu, sisu ja struktuur

1990. aastate poliitilised sündmused tõid eestlaste ja venelaste suhtlemismallidesse muutusi. Igapäevaelus hakkasid vene emakeelega kõnelejad kasutama üha enam eesti keelt. Kaasaegseid eesti-vene kontakte saadi hakata uurima kasutades või testides Lääne sotsiolingvistika teoreetilisi mudeleid ja uurimisparadigmasid. Eesti juhtivaks koodivahetusuurijaks on Anna Verschik ning tema juhendamisel on kaitstud hulk bakalaureuse- ja magistritöid, mis käsitlevad mõlema keele kontaktidest johtuvaid nähtusi.

Anastassia Zabrodskaja (2005) on vaadelnud Kohtla-Järve vene-eesti kooliõpilaste koodivahetust pragmaatilisest seisukohast. Jekaterina Ozernova (2005) on käsitlenud koodivahetust, laenamist ja konvergentsi² eesmärgiga uurida eesti leksikaalseid laene Eesti venelaste kõnekeeles. Tatjana Baškirova (2006) on näidanud koodivahetust teise keele omandamist soodustava vahendina. Indrek Konnapere (2006) on kirjeldanud ja analüüsinud kaitseväes aega teenivate vene keelt emakeelena rääkivate noormeeste vene-eesti koodivahetust ja selle funktsioone. Et nelja töö tulemusel saadud keelematerjal (lindistatud, osa käsitsi kirja pandud) ei läheks kaduma, nagu see juhtus omal ajal Nõukogude okupatsiooni eelse venelaste keelega (sellest on palju rääkinud Sergei Issakov, vt lähemalt Verschik 2001: 531), otsustati luua vene-eesti koodivahetuse korpus ning astuda LIPPS-i rühmaga ühendusse. Esimene töökohtumine leidis aset Limerickis 2006. aasta suvel 16. sotsiolingvistika sümpoosioni ajal. Kollokviumil “New Tools for Research in Bilingualism and Code-Switching: Studies using the Language Interaction Data Exchange System” (umbkaudu: kakskeelsuse ja koodivahetuse uued uurimisevahendid: LIDES-it kasutavad uurimistööd) tutvustasid A. Verschik ja A. Zabrodskaja valmivat vene-eesti koodivahetuse korpust laiale ringkonnale ning said asjakohast vastukaja (Verschik, Zabrodskaja 2006).

Kuna potentsiaalsed kasutajad on uurijad kogu maailmas, siis kõigi koodivahe- tuskorpuste metakeel on inglise keel (tekstide kommentaarid ja muu vajalik abiinfo, nagu situatsiooni kirjeldus, uurija kommentaarid jms). Et ka teised uurijad saaksid vene-eesti koodivahetuse korpuses orienteeruda, järgitakse sama konventsiooni. Andmaks artikli lugejatele parimat ülevaadet valmivast vene-eesti koodivahetuse korpusest, esitatakse näited samast põhimõttest lähtuvalt ning tõlgitakse eesti keelde erivajadusel.

Tavaliselt pole koodivahetuse korpused pelgad tekstikogumid, vaid sisaldavad ka mitmesugust taustainfot. Ühe näitena võib tuua türgi-hollandi andmebaasi, mille autor on Ad Backus. Oma korpuse juurde on ta teabefailis (ingl *readme file*) lisanud keelekogukonna ja informantide iseloomustuse ning kirjeldanud üksikasjalikult vastasmõju tüüpi (LIPPS Group 2000: 164–165).

Ka vene-eesti koodivahetuse korpusele on lisatud ingliskeelne Eesti venekeelsete kogukondade iseloomustus. Selline sissejuhatus on vene-eesti koodivahetuse korpuse puhul eriti vajalik. Selgitades Eesti etnolingvistilise situatsiooni kujunemist, tuleb rõhutada, et venekeelse elanikkonna kujunemine Eestis (vt Rannut 1994, Vseioiov 2002) on küllaltki ainulaadne nähtus kogu sõjajärgse Euroopa ajaloos. Sarnast tüüpi keelekogukondi ei ole Läänes ega Idas, sest tegemist pole põlisvähemuse, venekeelsete immigrandide (selle sõna tavalises mõttes) ega vene diasporaaga. Kuigi

² Konvergentsi kohta vt Ozernova 2005, Zabrodskaja 2006.

venelaste immigratsioon Eestisse suurenes plahvatuslikult pärast II maailmasõda (vt Viikberg 1999), polnud need ümberasujad samalaadsed näiteks Hollandi türklastega (Backus 1996: 43–46) või muude immigrantlike keelekollektiividega Euroopas. Venelaste immigratsioon meenutab teatud mõttes koloniseerimist, sest venelased asusid ümber (keskvõimu heakskiidul ja soosimisel) territooriumile, mida peeti oma riigi osaks. Maailmas pole seda tüüpi kogukondade sotsiolingvistikat uuritud. Seega tutvustab valmiv korpus mitte ainult vene-eesti koodivahetuse nähtusi kaasaegse kontaktlingvistika metakeeles, vaid annab ka ülevaate Eesti eri paikade vene-eesti kontaktide ja venekeelsete keelekogukondade omapärasest.

Kõnealune vene-eesti koodivahetuse korpus koosneb kolmest osast:

1. Eesti televisiooni kakskeelsed saated. Kokku 76 saadet, mis on litereeritud vaid osaliselt, kuna selles on palju ükskeelset materjali. Lisaks pole ükski saade kahe inimese dialoog, vaid koosnevad peamiselt väikestest, kuni kümnelauselitest vestlustest. Kõik vestlustes osalejate nimed on anonüümsuse huvides transliteratsiooni käigus asendatud lühenditega. Kui nimi on teadmata, siis on tähistatud näidetes eestlane tähega E(stonian), venelane tähega R(ussian). Kui ühes vestluses osaleb mitu inimest, on need tähistatud R1, R2, E1, E2 jne. Kui räägitakse kolmandatest inimestest, kes vestluses ei osale, siis nende nimed muudeti. Saatejuhi märgistus on H show_host, kui on kaks saatejuhti siis: H1 show_host_one, H2 show_host_two.
2. Tallinna venekeelsete elanike vene-eesti koodivahetus. Suur osa Tallinna koodivahetuse näiteid on kogutud kaubanduslikust suhtlusest (turult, poest jne) ning kirjalikest allikatest (poe- ja turuletid). Tekstid on litereerinud A. Verschik ja A. Zabrodsckaja ning kodeerinud A. Zabrodsckaja.
3. Ida-Virumaa vene emakeelega laste vene-eesti koodivahetus. Ida-Virumaal aastatel 2000–2006 käsitsi kirja pandud keeleaines on valdavalt pärit A. Zabrodsckaja iseseisvast uurimusest, mis oli aluseks tema magistriväitekirjale (Zabrodsckaja 2005). Translitereeringus on vestluses osalejate nimed kodeeritud korpuse 1. osas kirjeldatud meetodil. Õpilane (kui sugu on jäänud üleskirjutamise hetkel tähistamata) on näidetes tähistatud lühendiga S(tudent), õpetaja aga T(eacher).

Käsitsi kirja pandud kakskeelsetes vestlustes osalevad (vahel lihtsalt kõrvaltvaatajaks) kas A. Zabrodsckaja või A. Verschik, neid tähistatakse vastavalt Researcher_one (RES_1) ja Researcher_two (RES_2). Lindistatud saated välja arvatud, on mõlemad uurijad olnud vaatlejad ja mõnikord ka vestluse algatajad.

Korpuse I alamosas on analüüsi aluseks aastatel 2000–2006 lindistatud 13 järgmist telesaadet:

- 1) “Dilemma” – 5 saadet 2005. aastast
- 2) “Kolmas sektor” – 5 saadet: 2 saadet 2001. a ja 3 saadet 2002. a
- 3) “Neli aastaaga” – 2 saadet: 1 saade 2000. a ja 1 saade 2001. a
- 4) “Pealtnägija” – 1 saade 2002. a
- 5) “Politseinädal” – 1 saade 2001. a
- 6) “Pressiklubi” – 3 saadet 2001. a
- 7) “Sputnik” – 13 saadet: 5 saadet 2001. a ja 8 saadet 2002. a
- 8) “Subboteja” – 1 saade 2005. a
- 9) “Subjektiiv” – 5 saadet: 4 saadet 2001. a ja 1 saade 2002. a
- 10) “Teadmiseks” – 2 saadet 2001

- 11) “Unetus” – 34 saadet: 6 saadet 2002. a, 3 saadet 2003. a, 5 saadet 2004. a, 18 saadet 2005. a ja 2 saadet 2006. a
- 12) “Uudistaja” – 3 saadet 2002. a
- 13) “Vaata mind” – 1 saade 2004. a

Korpuse II alamosas on analüüsi aluseks igapäevasuhtlus Ida-Virumaal:

- 1) Narva:
 - Narva koolid (koolide ametlikud nimed pole välja kirjutatud)
 - Tartu Ülikooli Narva Kolledž
- 2) Kohtla-Järve:
 - Kohtla-Järve 3. keskkool
 - Kohtla-Järve Pärna põhikool
 - Kohtla-Järve avaliku teeninduse punktid (kauplused, pangad, postkontor jms)

Korpuse III alamosas on analüüsi aluseks igapäevasuhtlus Tallinnas. Näiteid on kogutud järgmistest kohtadest:

- 1) Tallinna Ülikool
 - Tallinna Ülikooli Akadeemiline raamatukogu
 - Tallinna Ülikooli vene emakeelega üliõpilastega lindistatud intervjuud (20 tundi)
- 2) Tallinna bussijaam
- 3) kaubanduslik suhtlus:
 - Balti jaama turg (suurim alamkorpus – 36 näidet)
 - kauplused
 - avaliku teeninduse punktid (reisibürood jne)
- 4) õpetajate täienduskoolituskeskus (ametlik nimi on välja kirjutamata)

Tallinna Ülikooli kolmanda alamosana on kavas esitada 2005/2006. õppeaasta ainekursustel “Keeleteaduse alused”, “Üldkeeleteadus”, “Keeleteaduse uuemaid suundumusi” ning “Kultuuridevaheline kommunikatsioon” osalenud üliõpilaste kirjalikest töödest kogutud koodivahetusjuhtumid.

2. Transkribeerimise ja kodeerimise põhimõtetest

Konversatsioonianalüüs eeldab spontaanse kõne üleviimist kirjalikku formaati. Transkribeerimine ja kodeerimine on suulise kõne analüüsi olulisimad osad. Nii ühest kui teisest sõltub, mis tulemusteni uurija jõuab. Mary Bucholtz (2000: 1440) juhib tähelepanu mitmele mõlemat tegevust mõjutavale tegurile: nii keeleuurija enda ideoloogia kui ka poliitilised tingimused, tema uurimistöö eesmärgid ja potentsiaalne lugejaskond ning vahel isegi kodeerija tehnoloogiliste võimaluste piiratus. Peter Auer (1998: 8–10) on transkribeerinud enda loodud reeglite järgi Carol Myers-Scottoni (1993: 83) varem transkribeeritud mitmekeelset vestlust ja jõudnud vastupidistele järeldustele.³

Elinor Ochi (1979) arvates on transkribeerimine subjektiivne tegevus. Kõigepealt peavad uurijad valima ja mugandama definitsiooni põhiüksuse kohta. Transkribeerija peab otsustama, kuidas erinevad üksteisest lause, lausung ja kõnevoor. Seejuures tuleb rõhutada, et lause defineerimine varieerub vastavalt iga

³ Antud artikli raamid ei luba süveneda selle vestluse kahte analüüsi. Nendime vaid, et näide illustreerib lahknevusi erinevate kodeerimislähenemiste vahel.

keeleuurija isiklikele huvidele. Ühtset kokkulepet saavutada on raske. Tiit Hennoste (2000a) märkab õigustatult, et konversatsioonianalüüsi eri käsitlustes esile toodud üksused kattuvad suurel määral, ning samas, et üksuste piiritlemine on küllaltki udune. Kuigi põhilised kriteeriumid on süntaktilised, tuleb mõnikord tugineda ka intonatsioonile, nimelt siis, kui laused on elliptilised, neist puudub lauseliikmeid või nad on poolikud (Hennoste 2000a: 2227). Penelope Gardner-Chloros jt (1999: 412) osutavad lause, fraasi, väljendi defineerimise problemaatilisele. Lause piiritlemiseks on mõningaid kriteeriume, näiteks üksiksõna, intonatsiooniline tervik, paus jms (Gardner-Chloros jt 1999: 408, 413). Fred Karlssoni (2002: 151) järgi on suulise keele põhiüksuseks lausung ehk kompaktne lõik, millel on füüsiliselt määratletavad piirid. T. Hennoste (2000a: 2223–2224) vaatlleb eri autorite suulise kõne liigendusüksusi ning liigitab kesksed üksused kolme rühma:

- a) semantilis-intonatsioonilised üksused: toonigrupp, süntagma, ideeüksus;
- b) lausungid: prosoodiline lausung, vooruehitusüksus, makrosüntagma;
- c) kõnevoor ja paratoon.

Mitmekeelset kõnet kodeeritakse kas lause-, fraasi- või üksiksõna tasandil (Gardner-Chloros jt 1999). Lundi ülikooli juures 1970. aastatel töötanud töörühm Talsyntax on pakunud makrosüntagma mõiste, mis erinevalt lause mõistest hõlmab ka verbita väljendeid (vt tabel 1). Makrosüntagma ja lausungi ning nende piiridega seotud probleemidest on eesti keeles kirjutanud ka T. Hennoste (2000a, 2000b, 2000c).

Tabel 1. Suulise lause määratlemisviisid (Karlsson 2002: 152 ja seal kasutatud viited)

	Makrosüntagmad
1	täisstruktuuriga lausemakrosüntagmad ehk (normaal)laused
2	lausekatkendid: <i>tere päevast; tagasi ülikooli</i>
3	interjektsioonimakrosüntagmad: <i>hei; kurat; häh; nojah</i>
4	pöördumismakrosüntagmad: <i>poiss; kallis sõber</i>
5	lõpetamata makrosüntagmad (katkendid): <i>mine sa</i>
6	mittegrammatilised ehk valesti moodustatud lausemakrosüntagmad

Tabelist 1 võib järeldada, et erinevatest sõnavormidest koosnev makrosüntagma on spontaanse kõne tüüpiline põhiüksus, mis on vene-eesti korpuse puhul väga vajalik.

Huvitav on seegi, et CHILDES-i koostamisreeglitest (MacWhinney 1995) ei leia keeleuurija lause definitsiooni, see jäetakse igapäevasele otsustada. Mõned kasutavad INTROS-i süsteemi (INformant's TRanscriptiOn String, umbkaudu: informandi transkribeeritud lint), milles pakutakse, et lähtuda tuleb süntaktilisest kriteeriumist ja defineerida lause nagu komponent (ingl *unit*), mis võib koosneda üksiksõnast, üksikfraasist, liht- või liitlausest (vt lähemalt LIPPS Group 2000: 173–174, 267).

Vene-eesti koodivahetuse korpuse suulise kõne lausete transkribeerimisel peetakse peamiseks lausungi piiriks pausi või voo vahetumist. See viib aga kahjuks järgmise probleemini: kas voo vahetumiseks pidada seda, kui kõneleja katkestab end ise või katkestatakse teda.

2.1. Tähestiku valiku problemaatika

Kui tegu on kahe keelega, siis tahaks iga uurija transkribeerimissüsteemi, mis võimaldaks neid omavahel eristada. Üks valik võib olla traditsiooniline kirjasüsteem, kui selline olemas. Kui uurija valib keeleliselt neutraalse (ingl *language-neutral*) tähestiku nagu IPA (International Phonetic Alphabet), siis peab näitama, mis keeles on konkreetsed transkribeerimise lõigud. Kuigi IPA on universaalne, kõigile keeleuurijatele ühtmoodi arusaadav transkribeerimisviis, kasutatakse ka mitmesuguseid muid transkribeerimis- ja transliteratsiooniviise. Näiteks LIDES-i autorid (LIPPS Group 2000: 214–246) on esitanud ladina tähestikku kasutavad keeled standardkujul. Ka vene-eesti koodivahetuse korpuses ei kasutata IPA-t. Eestikeelsed elemendid esitatakse eesti õigekirjas. Vene keele puhul on muidugi keerulisem, aga püütakse translitereerida põhimõttel “üks häälik – üks märk”.

Keele tähistuse probleem aga siiski jääb, sest mõnedes lausetes esineb lause-täiteid (vene-eesti koodivahetuse korpuse puhul näiteks *hm, oh, o, a, aha(a)* jne), mida on raske omistada konkreetsele keelele. Seda enam, et võib olla ka sõnu/silpe, mis kuuluvad kolmandasse keelde. Vaatame näiteks järgmist kolmkeelset lauset:

(1)	Segodnja	kes	on	<u>on</u>	<u>duty</u> ?
	täna	kes	on	PREP	korrapidaja
	vn	ee	ee	ingl	ingl

‘Kes on täna korrapidaja?’

Mida sellistel puhkudel teha ja kuidas seda näidata, nõuab teoreetilisi otsuseid. Asi pole ainult meetodi valikus mõne sõna märkimisel (kas K1, K2 või hoopis K3). Suurt rolli mängivad ka süntaksi, fonoloogia ja pragmaatika aspektid. Kaks (või rohkem) keelt teevad transkribeerimist aina keerulisemaks, sest tuleb näidata, mis keelde iga keeleelement kuulub. Näite (1) juures võib tekkida küsimus, kuhu kuulub keeleüksus *on*. Ei saa sajabrotsendiliselt väita, et esimene *on* on just eestikeelse verbi *olema* 3. isiku ainsuse pöördeline vorm, aga teine *on* on ingliskeelne prepositsioon. Esimene *on* on vaieldav, sest seda võib interpreteerida ka ingliskeelse prepositsioonina. Teine *on* peab olema kindlasti ingliskeelne prepositsioon, vastasel korral ei omaks lause tähendust.

Korpuses saab esitada seda rida järgmiselt (näites (2) on transkribeeritud faili katkend).

(2)					
@Languages:	Russian (1),	Estonian (2),		English (3)	
*G1:	segodnja@1	kes@2	on@2	on@3	duty@3 ?
%glo:	today	who	is	on	duty
%tra:	who is on	duty	today		

On otsustatud, et venekeelseid elemente translitereeritakse, sest vene ortograafia nn joteeritud tähed vastavad tihti kahele häälikule, aga morfeemidevaheline piir jookseb sageli just nende kahe hääliku vahel (näiteks koodivahetatud *sai-ječk-a*, vt Zabrodskaja 2005: 41). Näeme, et vene tähestiku kasutamine ainult raskendaks glossimist. Inglise keele keskses CHILDES-i programmis on vene tähestiku kasutamine võimatu. Ladina tähestiku kasutamine ei loo visuaalset kontrasti vene ja eesti keele vahel ja raskendab lugemist ning koodivahetatud sõnast/väljendist arusaamist. Seetõttu tähistatakse iga elemendi keelelist kuuluvust, kasutades märke @1 jne.

2.2. Korpuse failide ülesehitusest

Korpuse valmiskogum peab sisaldama kolme või nelja faili: CHAT-andmefaili (ingl *datafile*), taustafaili (*depfile*), teabefaili ning võimalusel ka täiendfaili (*depadd*).

Korpus kui selline sisaldub just andmefailis. CHAT-andmefail koosneb tüüpiliselt kolmest allosast: faili päised (*file headers*), põhiread (*main tiers*) ja sõltread (*dependent tiers*). Kuigi kõik sõltread ei ole kohustuslikud, soovitatakse kasutada morfeemtõlke rida (%glo) ja lause vaba tõlke rida (%tra) (LIPPS Group 2000: 149).

CHAT-andmefaili koostamisel peab lähtuma järgmistest põhimõtetest:

- 1) faili iga sümbol peab olema ASCII (American Standard Code for Information Interchange – Informatsiooni Vahetamise Ameerika Standardne Kood) tähis;
- 2) iga rea peab lõpetama sisestusklahviga.

On ka teisi nõudeid, mis toimivad iga konkreetse punkti juures.

Tausta- ja täiendfailid teevad kindlaks koodivahetuscorpuse süntaksi ja struktuuri. Esimeses on kõik automaatsed kodeerimisvahendid. Kõik uued koodid peavad sisalduma aga LIDES-i erilises alamkogumis – täiendfailis, nende seletus on teabefailis. Kodeerimisõpikus antakse isegi loetelu, mida täiendfail peab täpselt määratlema (LIPPS Group 2000: 162). Andmete sisestamis-, kodeerimis- ja selgitusviisi õigsust võib kontrollida programmi käskluse CHECK abil. See osutab ka vigadele valmiskogumi kõigis failides.

2.2.1. Faili päised

Faili päised on tekstiread, mis annavad informatsiooni faili kohta. Need päised asuvad iga transkribeeritud näite alguses ning algavad märgiga @, millele järgneb pealkiri. Mõned päised ei vaja mingit seletust, need on “tühjad” päised, näiteks @Begin või @End. Teistele päisetele järgneb informatsioon osalejate, situatsiooni jms kohta. Seda informatsiooni nimetatakse sissekandeks (ingl *entry*). Sissekanne nõudvatele päistele järgneb koolon ja tabeldusklahv, seejärel sissekanne ise. Näiteks:

```
(3)
@Participants: SHO  shopkeeper,  RES  researcher,  CHI  child
```

Peale tühjade päiste ja informatsiooni sisaldavate päiste eristuse pakub CHAT ka kohustuslikke ja mittekohustuslikke ridu.

Järgmisi faili päiseid (vt tabel 2) on vaja programmi automaatse analüüsi käivitamiseks (MacWhinney 1995: 13–14; LIPPS Group 2000: 149).

Täiendavat infot sisaldavad mittekohustuslikud ehk fakultatiivsed read (LIPPS Group 2000: 149-150). Neid on kahte tüüpi: püsivad ja muutlikud.

Püsipäised (ingl *constant headers*) sisaldavad kasulikku informatsiooni, mis ei muutu terve faili jooksul. Samas võivad mõned päised olla kas püsivad või muutlikud, nt aeg (*date*), koht (*location*), situatsioon (*situation*) jne. Need read peab paigutama faili algusesse, s.t enne kõnet (LIPPS Group 2000: 150).

Kuna transkribeeritud vestlustes on kasutatud mitut keelt (antud korpuse puhul vähemalt kaht, vene ja eesti), soovitatakse kindlasti kasutada järgmist kaht rida (vt tabel 3).

Tabel 2. Transkribeeritud faili kohustuslikud read

@Begin	Iga transkribeeritud faili esimene rida. Tähistab terviku algust.
@Participants	Selles reas kirjutatakse lahti, kes on vestluse osalejaskond. See nõuab sissekannet, mis koosneb kahest või kolmest osast: <ul style="list-style-type: none"> a) kõnelejate ID (kohustuslik), mis koostatakse ühest kuni kolmest tunnismärgist (suurtähtedega ja/või numbritega), nt STA või S81; b) kõnelejate nimed (mittekohustuslik); c) kõnelejate roll (kohustuslik), nt <i>Interviewer</i> 'intervjueerija'. Rolli, mida pole taustafailis, peab lisama täiendfaili. Viimased kaks osa (b ja c) võivad koosneda mitmetest elementidest, kuid neid tuleb ühendada allkriipsuga, nt <i>Stasja_Brodska</i> või <i>sister_in_law</i> .
@End	Iga transkribeeritud faili viimane rida. Tähistab terviku lõppu.

Tabel 3. Keelte tähistamine mitmekeelses korpuses

@Language(s):	Päis näitab transkribeeritud näidete peamist keelt või keeli.
@Language of XXX:	Päist võib kasutada, et kirjeldada konkreetse vestluspartneri peamist keelt (peamisi keeli). Sel juhul tuleb teha eraldi rida iga kõneleja kohta. Võib anda informatsiooni ka iga osaleja kakskeelsuse astmest, seda tuleb teha teabefailis.

Teisi näiteid püsipäistest (vastavalt vanus, haridus, sugu):

@Age of XXX:
 @Education of XXX:
 @Sex of XXX:

Muutpäised (ingl *changeable headers*) võivad esineda niihästi faili alguses koos püsipäistega kui ka faili põhiosas. Muutpäised sisaldavad informatsiooni, mis muutub faili jooksul. Nii näitavad need, kuidas saab vestlust etappideks jaotada tegevustest lähtuvalt. Seepärast sisestatakse need päised sinna, kus informatsioon muutub. Näite (4) neljas rida @Activities selgitab, mis juhtub vestluse erinevatel staadiumitel.

```
(4)
@Begin
@Participants:  B1      boy_one,      B2      boy_two
@Sex   of      Participants:  male,   male
@Age   of      Participants:  8,     8
@Mothertongue of Participants:  Russian, Russian
@Languages:    Russian (1),  Estonian (2),  Undecidable (0)
@Location:     Kohtla-Järve  Pärna School
@Date         of boarding:    JUN-02
@Date         of coding:      04-APR-06
@Transcriber  and Coder:      Zabrodskaja,  Anastassia
@Warning:     this is only a   small sample
```

*B1: xxx@0

@Activities: B1 dressed in hooded jacket stands in the classroom.

*B2: xxx@0

```

@Activities: B2 is talking to B1 behind his back.
@Activities: B2 puts the hood on the head of B1.
*B1: ty@1 chto@1 loll@2 ?
%tra: are you a fool
@Activities: B1 is turning angrily to B2.
*B2: sam@1 takoj@1.
%tra: you are
@End

```

Näitevestluses (4) esineb rida @Activities neli korda. See näitlikustab, mida osalejad teevad vestluse või situatsiooni jooksul ning aitab koodivahetuse põhjusest aru saada. Rida @Activities illustreerib vahel väga selgelt, miks konkreetne kõneleja läheb ühelt keelelt teisele, ehk kuidas koodivahetus on seotud tegevuse muutumisega.

2.2.2. Põhiread

Tegelikku kõnet transkribeeritakse andmefaili põhiridadel, mis reprodutseerivad kirjalikus vormis iga vestlusest osaleja ütluse. Nõudmised on järgmised (MacWhinney 1995: 8–9).

1. Põhiread osutavad väljaõeldule ning algavad tärniga (*). Iga põhirida peab sisaldama ühte ja ainult ühte lauset, kuid võib ulatuda mitmele tekstireale.
2. Pärast täрни * tuleb osaleja kood (ID), seejärel koolon ja tabulaator, nt

```
*RES_2: kus on kolmas klass ?
```
3. Põhirea jätkamine järgmisel tekstireal algab tabulaatoriga.
4. Laused lõpevad kas punkti (.), hüüumärgi (!) või küsimärgiga (?). Tavaliselt soovitatakse, et viimase sõna ja lauselõpumärgi vahele jäetaks tühik (LIPPS Group 2000: 151).
5. Suurtähtedega kirjutatakse välja ainult pärisnimed ja asesõna *I* 'mina'. Suurtähti ei kasutata ka lause esimese sõna algustähena.
6. Transkribeerimata keeleainest tuleb tähistada *www*-ga.
7. Arusaamatuks jäänud sõnu või väljendeid tähistab *xxx* või *xx* (kui ainult üks sõna on arusaamatu).
8. Ebaselge artikulatsiooniga sõnu saab märkida fonoloogilise vormina, võttes kasutusele ampersandi, nt *&a*.
9. Ebaselge artikulatsiooniga, kuid kontekstist aimatavaid sõnu saab täiendada, lisades sulgudes oletatava puuduvat osa, nt *pikapäeva(rühm)*.

2.2.3. Sõltread

Sõltread sisaldavad uurija kommentaare: tähelepanekuid, seletusi, kirjeldusi jms. Väga tähtis on see informatsioon välja tuua eraldi ridadele, põhiridadel osutub see mitteloetavaks. Nõudmised sõltridade kohta on järgmised (MacWhinney 1995: 8–9).

1. Sõltread kirjutatakse pärast transkribeeritud kõnevooru ning kommenteerivad või täpsustavad seda. Algavad alati sümboliga %.
2. Pärast sümbolit % kirjutatakse väikeste tähtedega kolmetäheline kood, siis koolon ja pärast tabulaatori vahet kirjutatakse juba informatsioon, nt
 %glo: today who is on duty
3. Sõltridade jätk algab järgmisel tekstireal tabulaatoriga.
4. Tähtis on, et sõltrea lõppu ei pandaks kirjavahemärke.

Pakutakse ka üldkasutatavaid võimalusi (vt tabel 4).

Tabel 4. Transkribeeritud faili sõltridade näiteid (LIPPS Group 2000: 152)

%add:	näitab lause adressaati
%glo:	lause glossimine sõna-sõnalt või morfeem-morfeemilt
%gpx:	mitteverbaalne informatsioon: noogutab pead
%mor:	sõnade morfoloogiline kirjeldus
%spa:	kõneakt, kõnetegu
%tra:	lause vaba tõlge (vene-eesti koodivahetuse korpuses tõlgitakse inglise keelde)
%com:	uurija kommentaarid

Kuna vene-eesti koodivahetuse korpuse loojaid huvitab koodivahetatud sõnade morfoloogiline integratsioon ja/või selle puudumine, mitteverbaalne info ei ole esmatähtis, siis otsustati %gpx rida kasutada vaid vajaduse korral. Seevastu kommentaaride rida (%com) on väga tähtis. Iga näite juures esinevad ka %glo ja %tra read.

3. Sisestava koodivahetuse morfoloogiline analüüs

Järgmise alaosa põhiteema on sisestava (valdavalt ühesõnalise) koodivahetuse morfoloogia analüüsimine. Korpuses seletatakse seda kommentaaride real (%com). Korpuses on ka vahelduva koodivahetuse näiteid (vt lähemalt Zabrodskaja 2005: 58–61), kuid siin on keskendunud just sisestavale koodivahetusele.

Senini on jäänud lahtiseks küsimus, kas iga näidet peab glossima või ei pea. Kui tegu on aglutinatiivsete keeltega (näiteks türgi keel), siis soovitatakse morfeem-morfeemilist glossimist. Glossi eesmärk on kergendada korpusest arusaamist, eriti juhul, kui kasutaja ei oska kumbagi keelt (LIPPS Group 2000: 159). Nagu teada, on eesti keel aglutineeriv-flekteeriv ja vene keel flekteeriv. Et mõlema keele morfoloogia on arenenud ja keeruline, ning analüüsi seisukohalt on just morfoloogiline integratsioon väga tähtis, püütakse glossida kõiki lauseid.

Analüüsi põhimõisted on järgmised. Et kõige sagedamad koodivahetatud sõnaliigid on eesti nimisõnad, analüüsitakse nende kahel tasandil (või kahes järgus) loomulikus vestluses toimuvat integreerimist vene maatriksisse: soo määramine ja vastavate käändelõppude lisamine. Tihtilugu määravad koodivahetajad vaid sugu, kuid teatud positsioonides (näiteks 2. käändkonna elututel nimisõnadel ainsuse nominatiivis ja akusatiivis) pole see võimalik. Näiteks siis, kui puuduvad nimi-

sõnaga soos ja käändes ühilduvad omadus- või asesõnad ning soos ühilduvad verbi mineviku ainsuse vormid. Näidete kodeerimisel tuli välja, et %com rida sisaldab liiga pikki seletusi. Nende asemel on välja töötatud tabelis 5 esitatud süsteem, mida täiendatakse iga uue üksikjuhu kodeerimisvajadustest lähtuvalt.

Tabel 5. Integreerimata elementide tähistused

	Koodivahetatud sõnal:	Tähistus %glo real	%com rida
1	puudub morfoloogiline integratsioon vene maatriksisse, sest süntaktiline positsioon ei nõua seda (NOM=ACC=∅)	∅	∅ = mingit selgitust ei ole
2	puudub morfoloogiline integratsioon vene maatriksisse, kuigi potentsiaalselt on see võimalik	∅+?	∅+? = antakse võimalik variant või variandid
3	on morfoloogiline integratsioon nii soos kui ka käändes	∅+	∅+ = selgitatakse vene morfoloogia seisukohalt
4	on osaline morfoloogiline integratsioon vene maatriksisse (sugu +, kääne -)	∅ ½ G (Gender)	∅ ½ S = mingit selgitust ei ole
5	on osaline morfoloogiline integratsioon vene maatriksisse (sugu -, kääne +)	∅ ½ C (Case)	∅ ½ C = mingit selgitust ei ole
6	on kakskeelse homofooni kuju: realisatsioon on eesti	HE (Homophone, Estonian)	HE = mingit selgitust ei ole
7	on kakskeelse homofooni kuju: realisatsioon on vene	HR (Homophone, Russian)	HR = mingit selgitust ei ole
8	on kakskeelse homofooni kuju: realisatsioon on vahepealne	HB (Homophone, on Border line)	HB = näidatakse, mis keele foneetilisele kujule on konkreetne hääldataud sõna(osa) lähedane

Kõik tabelis 5 toodud %glo reale lisatud märgendused tuli lisada täiendfaili ning teabefailis lahti mõtestada. Eestikeelsete nimisõnade (täieliku/osalise/potentsiaalse) morfoloogilise integratsiooni selgitamisel on vaja kasutada vene käändkondade süsteemi, näidata käändelõppe ja määrata ka sugu. Ühes täiendfailis selgitatakse lühendeid, millega tähistatakse korpuses vene kolme käändkonda. Lühidalt võib seda kirjeldada järgmiselt:

- a) esimene käändkond – 1st D(eclension) C(lass),
- b) teine käändkond – 2nd DC,
- c) kolmas käändkond – 3rd DC.

Üksikasjalikult kirjeldatakse vene käänete süsteemi, tuues vene nimisõnu sisaldavaid näiteid. Märgitakse, et vene keeles on 6 käänet ja eesti keeles 14. Erilist tähelepanu pööratakse soo aspektile. Rõhutatakse, et erinevalt eesti keelest on vene keeles kolm sugu: nais-, mees- ja kesksugu, olemas on ka üldsugu (*umnica* 'tarkpea', *jabeda* 'kaebupunn' jne). Sedalaadi keelekirjeldused on vajalikud eriti neile uurijatele, kes tunnevad huvi vene-eesti koodivahetatud sõnade morfoloogia vastu, kuid ei oska kumbagi keelt.

Järgmisena kirjeldatakse ja kommenteeritakse korpuse valmisnäiteid. Näide (5) on kõnekas juhtum.

(5)
 @Begin
 @Participants: RES_2 Researcher_two, T1 teacher_one
 @Sex of Participants: female, female
 @Languages: Russian (1), Estonian (2)
 @Location: A school in Narva with Russian as a language of instruction
 @Date of boarding: 14-MAR-02
 @Date of coding: 11-APR-06
 @Transcriber: Verschik, Anna
 @Coder: Zabrodskaja, Anastassia
 @Warning: this is only a small sample

*RES_2: kus@2 on@2 kolmas@2 klass@2?
 %tra: where is the third form
 *T1: Nadja@1, otvedi@1 učitelnicu@1 k@1 õpetaja@2
 Monika@2 .
 %glo: Nad-ja take teacher to
 teacher-Ø+? Monika-Ø+?
 %tra: Nadja, take the teacher to
 the teacher Monika
 %com: _k õpetaja Monika_ 'to the teacher Mo-
 nika' Ø+? = _k õpetaje Monik-e_ or _k
 õpetaj-e Monik-a_ (1st DC, DAT).
 @End

Nagu näeme, on *õpetaja Monika* jagamatu sõnaühend, millel puudub morfoloogiline integratsioon. Näide (5) paneb mõtlema, kuidas tähistada morfoloogilise integratsiooni puudumist neil juhtudel, kui potentsiaalselt on see võimalik. *Monika* võib olla nii eesti kui ka vene nimi. Ent juhtum on keerulisem kui konsonandilõpuliste sõnade puhul: täishäälik *-a* kuulub eesti keeles tüvve (*Monika*), vene keeles on see aga 1. käändkonna ainsuse nimetava lõpp (*Monik-a*). Selles konkreetsetes näites on õpetaja Monika eestlanna (lisaks on see sõnaühendi osa), mistõttu interpreteeritakse nime eesti sõnana. %com-real pakutakse võimalikud variandidid, kus mõlemad või üks nimisõna oleksid integreeritud vene maatriksisse.

Näide (5) erineb näitest (6), kus integratsioon puudub, sest eesti nimisõna süntaktiline positsioon vene maatriksis seda ei nõua. Eesti *juust* sobib vene 2. käändkonda, kuhu kuuluvad meessoost null-lõpuga elutud nimisõnad, mis ei nõua akusatiivis mingit lõppu: *pokupali stol* 'ostsime laua' jne.

(6)
 @Begin
 @Participants: B1 buyer_one, B2 buyer_two
 @Sex of Participants: female, male
 @Languages: Russian (1), Estonian (2)
 @Location: Jaama market
 @Date of boarding: 12-SEP-03
 @Date of coding: 10-AUG-06
 @Transcriber: Verschik, Anna

@Coder: Zabrodskaja, Anastassia
 @Warning: this is only a small sample

*B1: v@1 tot@1 raz@1 my@1 etot@1
 juust@2 tebe@1 pokupali@1 ?
 %glo: that time we this cheeseØACC
 youDAT buy 1 SG PAST
 %tra: did we buy this cheese for
 you that time
 %com: _juust_ 'cheese' Ø.
 @End

Vahel võib eesti sõna olla integreeritud nii soos kui ka käändes. Võrdluseks vaatame näidet (7).

(7)
 @Begin
 @Participants: O official
 @Sex of Participants: female
 @Mother tongue of Participant: Russian
 @Languages: Russian (1), Estonian (2)
 @Location: Tourist service centre
 @Date of boarding: JUL-02
 @Date of coding: 11-APR-06
 @Transcriber: Verschik, Anna
 @Coder: Zabrodskaja, Anastassia
 @Warning: this is only a small sample

*O: no@1 eto@1 bez@1 käibemaks@2a@1 .
 %glo: but this without VAT-GEN-Ø+
 %tra: but this is without VAT
 %com: _bez_ käibemaksa_ 'withoutvalue added tax'
 Ø+ = 2nd DC, GEN.
 @End

Selle kõneleja idiolektis võib *käibemaks* olla nii täiesti aktsepteeritud ja integreeritud laen kui ka lihtsalt sisestav koodivahetus. Kas see on laen või koodivahetus, on võimatu otsustavalt öelda, sest põhimõttelist vormilist vahet ei ole, küsimus on ainult konventsionaliseerumise astmes. Kuigi mitmed formaaltingvistiliselt orienteeritud koodivahetuse uurijad on püüdnud luua teoreetilist alust tüüpilise laenu ja tüüpilise koodivahetuse eristamiseks, pole see siia maani õnnestunud. Seda enam, et on teisi uurijaid, kes arvavad, et see on põhimõtteliselt võimatu, pigem on tegemist kontiinumiga. Mõlema koolkonna esindajate püüdlustega võib tutvuda Anneli Sarhimaa (1999: 127–130) väitekirjas. Kuna sõna *käibemaks* fikseeriti mitmel kõnelejal, võib ettevaatlikult pakkuda, et koodivahetuse ~ laenu kontiinumil on see lähedane just viimasele. See sõna käändub vastavalt vene ükskeelse grammatika reeglitele: 2. käändkonna genitiivi lõpp on *-a*. Samas on võimalik varieerumine integreeritud ja mitteintegreeritud kuju vahel, sest ei saa ennustada, kas see sõna esineb integreerituna või integreerimata sama kõneleja kõnepruugis järgmisel päeval.

Näites (8), mis fikseeriti lindistatud telesaate “Sputnik” (06.04.2001) transkribeerimisel, esineb eesti sõna *linnaalitsus* vene lauses morfoloogiliselt integreerimata, kuigi see sobib vene 2. käändkonda (meessoost null-lõpuga nimisõna), mille instrumentaal (*s chem?* ehk *millega?*, eesti vaste on kaasäitlev) oleks *om-lõpuga: s linnavalitsus-om*.

```
(8)
@Begin
@Participants: H TV_show_host
@Languages: Russian (1), Estonian (2)
*H: konkurs@1 organizovan@1 Estonskim@1 tanceval'nym@1
agenstvom@1 sovместno@1 s@1 linnavalitsus@2 .
%glo: competition hold-PS-PART Estonian Dance
agency together with municipal administration-Ø+?
%tra: the competition was held by Estonian
Dance agency together with the municipal
administration
%com: _linnaalitsus_ 'municipal administration' Ø+? = _s
linnaalitsus-om_ (2nd DC, INSTR) .
@end
```

Sisestava koodivahetuse piiripealsete nähtuste analüüsi eesmärk on tuvastada ja uurida morfoloogia ning prosoodia/fonoloogia piiril asuvaid nähtusi, mis võivad osutada määravaks otsustamisel, mis keelde antud element kuulub. Tihtilugu on aga tegemist vahepealsete variantidega, mis ei luba elementi selgelt klassifitseerida. A. Verschik (2005) näitab, et mitmekeelses kõnes on selliseid asju küll ja küll. Siia kuulub eestipärane kahe rõhu säilitamine liitsõnadades (kas *'infotehnol' oogia* või *infotehnol'o(o)gia*), aga ka vokaali pikkuse säilitamine (kas *jaanipäev* või *janipjaev*). Sellistel puhkudel sünnivad nn kompromissivormid, mis ei kuulu kummassegi ükskeelsesse varianti, seepärast tõstetakse neid eraldi esile ja kommenteeritakse %com real. Vaatame näidet (9).

```
(9)
@Begin
@Participants: S_1 Salesperson_one, S_2 Salesperson_two
@Sex of Participants: female, female
@Mother tongue of Participants: Russian, Estonian
@Languages: Russian (1), Estonian (2), Undecidable (0)
@Location: Railway Station market
@Date of boarding: 04-JUL-03
@Date of coding: 11-APR-06
@Transcriber: Verschik, Anna
@Coder: Zabrodskaja, Anastassia
@Warning: this is only a small sample
*S_1: Ja@1 na@1 kilogramm@1 popravilas'@1 (.)
posle@1 janipäev@0 .
%glo: I one kilogramØ-ACC put on 3SG-PAST after
Midsummer Day-HB-Ø+?
```

%tra: I put on one kilogram after
 Midsummer Day
 %com: _janipäev_ 'Midsummer Day' HB = jani@1 + päev@2
 janipäev 'Midsummer Day' Ø+? = _posle janipäeva_
 (GEN).
 @End

Ülaloodud näites (9) oli *jaanipäev* ühelt poolt morfoloogiliselt integreerimata laen, teiselt poolt võib seda juhtumit analüüsida ka foneetilises küljest: pikka vokaali ei säilitata, hääldatakse venepäraselt. Sõnas *janipäev* on rõhk viimasel silbil, %com-real näitab seda viimase silbi allakriipsutamise. Kui oleks *janipjaiva*, siis oleks täielik foneetiline integratsioon. Kuid %com-rida ei tasu liiga pikaks teha, sest detailidega ülekoormatus häirib lugemist ja mõistmist. Kui interpreteerida koodivahetatud nimisõna *janipäev* vene grammatika terminites, siis käitub see 2. käändkonna meessoost elutut objekti tähistava nimisõnana, mille ainsuse genitiivi lõpp on *-a*: *posle janipäev-a* 'pärast jaanipäeva'.

Lõppsõna

Vene-eesti koodivahetuse korpuse huvid keskenduvad praegu peamiselt eesti üksuste (valdavalt nimisõnade) morfoloogilisele ja fonoloogilisele integratsioonile vene maatriksraamist vaadatult või selle puudumisele. Erilist tähelepanu pälvidad kompromissivormid ja uued konstruktsioonid, mida pole kummaski keeles.

Korpuses leidub hulk näiteid, mis nõuavad pikemat ja keerukamat analüüsi või mille puhul tekib mitu analüüsivõimalust. Näidete analüüs nõuab pidevalt uute märgenduste loomist. Seda illustreerib eriti ilmekalt artikli 3. peatükk. Kuna üheski valmiskorpuses pole neile adekvaatseid vasteid, püütakse märgendusprotsessi teha võimalikult lihtsaks, et neil, kes tahavad oma keeleainest vene-eesti omaga võrrelda, tekiks selline võimalus.

Tulevikus kavatakse pühenduda ka koodivahetatud sõnade foneetilisele analüüsile. Siis saab välja pakkuda eesti keelest vene keelde ümberasunud nimisõnade koodivahetuse ~ laenu kontiinumi.

Kirjandus

- Backus, Ad 1996. Two in One. Bilingual Speech of Turkish Immigrants in The Netherlands. Studies in Multilingualism 1. Tilburg: Tilburg University Press.
- Baškirova, Tatjana 2006. Koodivahetus eesti keele kui teise keele tunnis. Magistritöö. Tallinn: Tallinna Ülikool.
- Bucholtz, Mary 2000. The politics of transcription. – Journal of Pragmatics 32 (8), 1439–1465.
- Gardner-Chloros, Penelope; Moyer, Melissa; Sebba, Mark; van Hout, Roeland 1999. Towards standardizing and sharing bilingual data. – International Journal of Bilingualism 3, 395–424.
- Hennoste, Tiit 2000a. Sissejuhatus suulisesse eesti keelde VI: lausung suulises kõnes I. – Akadeemia 10, 2221–2254.
- Hennoste, Tiit 2000b. Sissejuhatus suulisesse eesti keelde VII: lausung suulises kõnes II. – Akadeemia 11, 2463–2486.

- Hennoste, Tiit 2000c. Sissejuhatus suulisesse eesti keelde VIII: lausung suulises kõnes III: eneseparandused. – *Akadeemia* 12, 2687–2710.
- Karlsson, Fred 2002. Üldkeeleteadus. Renate Pajusalu, Jüri Valge, Ilona Tragel (tõlkinud ja kohandanud). Tallinn: Eesti Keele Sihtasutus.
- Konnapere, Indrek 2006. Vene-eesti koodivahetuse konversatsioonilised funktsioonid vene emakeelega ajateenijatel. Bakalaureusetöö. Tallinn: Tallinna Ülikool.
- LIPPS Group, The 2000. The LIDES Coding Manual: A Document for Preparing and Analyzing Language Interaction Data. – *International Journal of Bilingualism* 4/2, 131–278.
- MacWhinney, Brian 1995. *The CHILDES Project: Tools for Analyzing Talk*. 2nd edition. Hillsdale, NJ: Erlbaum.
- Milroy, Lesley; Muysken, Pieter 1995. Introduction: code-switching and bilingualism research. – L. Milroy, P. Muysken (Eds.). *One Speaker, Two Languages: Cross-Disciplinary Perspective on Code-Switching*. Cambridge: Cambridge University Press, 1–14.
- Myers-Scotton, Carol 1993. *Social Motivations of Code-Switching*. Oxford: Clarendon Press.
- Ochs, Elinor 1979. Transcription as a theory. – Elinor Ochs, Bambi Schieffelin (Eds.). *Developmental Pragmatics*. New York: Academic Press, 43–72.
- Ozernova, Jekaterina 2005. Eesti laenud venelaste kõnes. Bakalaureusetöö. Tallinn: Tallinna Ülikool.
- Rannut, Mart 1994. Beyond linguistic policy: The Soviet Union versus Estonia. – Tove Skutnabb-Kangas, Robert Phillipson (Eds.). *Linguistic Human Rights. Overcoming Linguistic Discrimination*. Berlin, New York: Mouton de Gruyter, 179–208.
- Sarhimaa, Anneli 1999. Syntactic transfer, contact-induced change, and the evolution of bilingual mixed codes. Focus on Karelian-Russian language alternation. *Studia Fennica Linguistica* 9. Helsinki: Finnish Literature Society.
- Zabrodsckaja, Anastassia 2005. Vene-eesti koodivahetus Kohtla-Järve vene emakeelega algkoolilastel. Tallinna Ülikooli eesti filoloogia osakonna toimetised 6. Tallinn: Tallinna Ülikooli Kirjastus.
- Zabrodsckaja, Anastassia 2006. Eestivene keel(evariant): kas samm segakoodi poole? – *Keel ja Kirjandus* 9, 736–750.
- Verschik, Anna 2001. Interferentsi mehhanismidest ja vene-eesti kontaktidest. – *Keel ja Kirjandus* 8, 529–542.
- Verschik, Anna 2005. Russian-Estonian language contacts, linguistic creativity, and convergence: New rules in the making. – *Multilingua* 24, 413–429.
- Verschik, Anna; Zabrodsckaja, Anastassia 2006. The Russian-Estonian LIDES corpus: Elaboration of morphological encoding principles. – Abstract book. *Sociolinguistics Symposium 16* (6–8 July 2006). Limerick: University of Limerick, 39–40.
- Viikberg, Jüri 1999. Eesti rahvaste raamat. *Rahvusvahemused, -rühmad ja -killud*. Tallinn: Eesti Entsüklopeediakirjastus.
- Vseiov, David 2002. Kirde-Eesti urbaanse anomaalia kujunemine ning struktuur pärast Teist maailmasõda. Tallinna Pedagoogikaülikool. *Humanitaarteaduste dissertatsioonid* 8. Tallinn: TPÜ Kirjastus.

Anastassia Zabrodsckaja (Tallinna Ülikool) on uurinud keelekümblust, laste kakskeelsust ja vene-eesti koodivahetust. Viimasel ajal on peamiseks huviobjektiks vene-eesti koodivahetuse korpus. anastaza@tlu.ee

RUSSIAN-ESTONIAN CODE-SWITCHING CORPUS: ELABORATION OF ENCODING PRINCIPLES

Anastassia Zabrodskaja

Tallinn University

The paper has several aims: 1) to introduce the goals of the LIPPS group and the Russian-Estonian code-switching corpus (LIDES) in the Estonian context, 2) to give an overview of the Russian-Estonian code-switching corpus with its sub-corpora (in preparation at Tallinn University), 3) to make an overview of the standards used to transcribe and encode multilingual data in the LIDES database, and 4) to formulate some principles of morphological encoding.

Several sub-corpora are planned within the corpus: (a) bilingual TV talk shows; (b) data from bilingual Tallinn; (c) data from the predominantly Russian-speaking North East (Narva and Kohtla-Järve).

The encoding of Russian-Estonian code-switching probably requires a special approach: Russian is written with Cyrillic letters whereas Estonian uses the Roman script. The different alphabets may lead to a different treatment of Estonian elements in writing and in oral communication.

Both Estonian and Russian have a developed inflectional morphology. Full integration of nouns means gender assignment and adding of inflectional morphology (case, number, or case and number). Empirical observations show that full morphological integration of an Estonian single noun into the Russian matrix is not always the case. The authors of the corpus are interested in instances where Russian inflectional morphology is absent, although the noun fits structurally into Russian declension classes.

The focus is also on items whose belonging to either language is not clear. If a speaker speaks Estonian with a Russian accent, common internationalisms as well as Estonian proper names are ambiguous. Retention of two stresses in Estonian compound nouns in the Russian matrix is one of the relevant features to be encoded.

As far as morphology in the Russian-Estonian LIDES Corpus is discussed in the second part of the article, here primary attention is given to: 1) morphological and phonic integration or lack thereof, and 2) compromise forms, new creations.

It is necessary to introduce a special encoding system for in-between items and think of a way of encoding lack of integration in order to distinguish it from zero-endings. Numerous relevant examples are presented in the paper.

Keywords: corpus linguistics, code-switching, Estonian, Russian

EMOTSIOONISÕNAD TÖISES KESKKONNAS – MILLEST KÕNELEVAD LOETELUKATSETE TULEMUSED?

Ene Vainik

Ülevaade. Artiklis võrreldakse emotsioonisõnade esiletulekut 2001. aastal läbi viidud loetelukatsetes tööil ja mujal. Hetketunnete ja emotsioonidega seotud tegusõnade loetelukatsete tulemuste semantilise analüüsi põhjal osutus, et katse läbiviimise koht on mõningal määral mõjutanud emotsioonisõnade meenumist, seda eriti konkreetsete emotsioonide ja emotsioonidega assotsiatiivselt seotud sõnavara osas. Hetketundeid loetlevad sõnad olid töises keskkonnas hinnangult vastuolulisemad kui mujal, tegusõnade puhul tuli tööil esile arvukam verbaalset suhtlust ja eriti hinnangulist suhtlemist kirjeldav sõnarühm. Võimalik, et töises keskkonnas kanaliseeruvad emotsioonid hinnangulisel suhtlusele, millega taastluuakse ja hoitakse alal sellele keskkonnale omast hierarhilist suhtemudelit.*

Võtmesõnad: antropoloogiline lingvistika, semantika, pragmaatika, eesti keel

1. Lähtekohad ja eesmärk

Artikkel põhineb ettekandel, mis peeti konverentsil teemaga “Keel töises suhtluses” (Vainik 2006b). Antud pealkiri lubab keelt mõista nii uurimisobjekti kui meediumina, mille vahendusel on võimalik midagi teada saada töise suhtluse eripärade kohta. Käesolevas artiklis ongi võetud just viimane hoiak – eesmärgiks on välja selgitada töö juures ilmnevad keeleainese esiletuleku eripärad. Lähtutakse eeldusest, et töise suhtluse spetsiifikat saab eritleda üksnes kõrvutades seda “mittetöise” suhtluse keelekasutusega. Kui muud tingimused on samad ja varieerub ainult keelefaktide ilmsikstuleku situatiivne kontekst, siis on just see suure tõenäosusega erinevuste põhjuseks.

* Artikkel on valminud ETF-i uurimistoetuse nr 7149 toel.

Järgnevas võrreldakse läbivalt teises ja mitteteises keskkonnas kogutud keeleandmeid. Artikli ainese moodustab mitte otseselt dokumenteeritud *keelekasutuse*, vaid eesti keele emotsioonisõnade *loetelukatsete* tulemuste semantiline analüüs. Loomulikult põhinevad kõik artikli lõpus esitatud kokkuvõtted ja arutus vaid emotsioonisõnade esiletuleku erinevustel loetelukatsetes, laienemata otseselt *kogu keelelisele suhtlusele* töö juures. Etteruttavalt võib siiski öelda, et keeleline suhtlemine kui tegevus leidis loetelukatsetes emotsioonidega seoses märkimisväärselt palju mainimist. Seega keel meediumina (loetelukatseted) peegeldas emotsioonidega seoses osati üllatavaltki tagasi *keelekasutuse* protsessi.

2. Uurimisprobleemi olemus

2001. aastal läbiviidud emotsioonisõnavara loetelukatsete seerias (Vainik 2001) ilmnes märgatav erinevus töökohas ja mujal intervjueeritud tööealiste (20–65) informantide emotsiooninimetuste esimeses loetelukatsetes: töö juures intervjueritutele meenus emotsiooninimetustest märgataval määral vaid *viha*, sellal kui mitteteises keskkonnas loetleti enam-vähem võrdselt kõiki eestikeelseid emotsioonide põhinimetusi (*armastus, rõõm, viha, kurbus*) (Vainik, Orav 2005).

Kirjanduses on loetelukatsete tulemusi mõjutavate teguritena nimetatud sõnade omadusi nagu lühidus, sagedus korpustes (Zammuner 1998) ja inimgruppide omadusi nagu eesmärgid, elukutse või tegevusala ja sellega seotult sugu (Ross, Medin 2005). Et tulemuse jaoks võiks olla oluline ka katse toimumise koht, on seni vähekasitletud aspekt. Käesolevast artiklist lähtutakse just sellest hüpoteesist.

Selleks, et välja selgitada, kas emotsioonikategooria liikmete selektiivne meenumine töökohas oli juhuslik või seotud mingit sorti üldisema emotsionaalse tendentsiga, analüüsitakse käesolevas artiklis semantiliselt samas 2001. a loetelukatsete seerias läbi viidud hetketunnete ja emotsioonidega seotud tegusõnade loetelukatsete tulemusi.

3. Meetod ja valim

Andmed koguti antropoloogilises lingvistikas kasutatava välimeetodiga, mis on mõeldud mingi valdkonna nn aktiivse sõnavara kogumiseks (Sutrop 2001, Davies, Corbett 1995). Ajalise piiranguta vaba loetlemise ülesannetes saab ainsaks keelelise produktsiooni piirajaks katseisikute lühimälu maht ja assotsiatsioonide seostumine küsitud kategooriaga. Kõnealuses hetketunnete loetelukatsetes tuli katseisikutel loetleda tundeid või emotsioone, mida nad küsitlemise hetkel kogesid. Emotsioonidega seotud tegusõnade loetelukatsetes tuli katseisikutel loetleda verbe, mis nende jaoks seostusid emotsioonidega. Andmete edasises töötluses grupeeriti mõlema katse tulemused eraldi semantiliselt ja arvatati semantiliste rühmade esiletuleku suhtelised osakaalud.¹ Semantiliste rühmade esiletuleku võrdlusel kummaski andmete kogumise situatiivses kontekstis (tööl vs. mujal) põhinevadki järgnevas esitatud tulemused. Mõistagi ei saa pidada uurija keeletundel põhinevat semantilist jaotust absoluudiks ega eritletud tähendusrühmi üksteist välistavateks. Oluline on, et mõlema katse puhul on kasutatud samu kriteeriume, mis tulemusena toovad

¹ Seega loobuti seekord konkreetsete sõnade esilduvuse eritlemisest kognitiivse esilduvuse indeksi (Sutrop 2001) abil, mida autor on kasutanud mitmetes varasemates artiklites (nt Vainik 2002a, 2002b).

välja ainekute semantilise liigenduse põhitendentsid ja küsitluskohast tingitud erinevused neis.

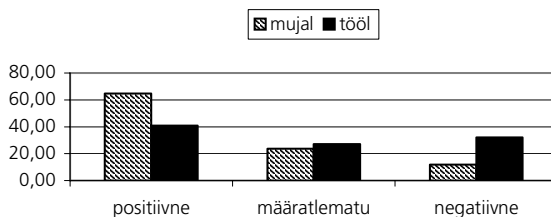
Loetelukatses osales 100 katseisikut vanuses 14–88 aastat. Käesolevas uurimuses kõrvutati neist vaid tööealiste isikute (20–65) loetelusid, keda oli kokku 72. Töö juures küsitletuid oli 46 ja mujal (nt kodudes, polikliinikus, teenindusasutustes) küsitletuid oli 26. Tööl küsitletute hulgas oli rohkem mehi ja mujal naisi. Varasem kontroll on näidanud, et mõlemast soost inimesed nimetasid täiesti sarnaselt töö juures esmalt viha ja alles seejärel teisi emotsioone (Vainik, Orav 2005), sama sarnast käitumist eeldati ka teistest loetelukatsetest. Töö juures küsitletute hulgas oli nii müügimehi, klienditeenindajaid, firmajuhte, laotöölisi, autojuhte kui trükitöölisi. Nad kõik kutsuti intervjuu läbiviimise ajaks siiski otseselt tööpostilt ära eraldi ruumi, kus ei olnud teisi inimesi peale intervjuueerija.

4. Hetketunded

Katseisikutel paluti loetleda tundeid või emotsioone, mida nad kogesid emotsioonide loetelukatsete seeria vältel. Ülesande eesmärgiks oli saada teada erinevused konkreetsete sündmustega seotud episoodiliste ja üldise kategoorianimetusega seostuvate semantiliste emotsiooniteadmiste käepärasuses (vt Vainik 2001, 2006a).

Tulemused osutasid, et eestlaste semantilistes emotsiooniteadmistes prevaleerivatest emotsioonide põhinimetustest (*viha*, *rõõm*, *armastus* ja *kurbus*) mainiti hetketunnete loetelukatses olulisel määral vaid *rõõmu*, seda tegid nii töö juures kui mujal küsitletud (vt tabel 1). Sarnane oli ka see, et mõlemas küsitluskohas nimetati keskmiselt 2,5 emotsiooni.

Erinevused puudutavad loetletud sõnade mitmekesisust: nimelt olid töö juures küsitletute vastused ühetaolisemad (keskmiselt 1,67 erinevat sõna) kui mujal (1,85). Oluline erinevus ilmneb ka emotsioonisõnade hinnangulises aspektis. Situatsioon, kus intervjuueerija palus nimetada tundesõnu, on töövälises keskkonnas leidnud vastuvõttu valdavalt positiivsete tunnetega, töökeskkonnas on reageeringute skaala olnud vastuolulisem (vt joonis 1). Tabel 1 esitab hetketunnete loetelukatse tulemused hinnanguliste gruppide lõikes tööl ja mujal. Seal selgub, et kui mujal on lisaks *rõõmule* nimetatud *uudishimu*, *põnevust*, *meeldivust*, *lõbu* ja *rahulolu*, siis töö juures on tuntud pigem *huvi*, *üllatust*, *segadust*, *ebakindlust* ja *ootusärevust*.



Joonis 1. Hetketunnete loetlemise katse tulemuste protsentuaalne jaotus sõnade tähenduse hinnangulise aspekti alusel

Tabel 1. Kaks või enam korda nimetatud hetketunnete protsentuaalne jaotus

Vastused:	Positiivne		Negatiivne		Määratlematu	
	mujal	tööl	mujal	tööl	mujal	tööl
rõõm	11,76	4,92				
uudishimu	5,88	2,46				
huvi	2,94	3,28				
üllatus	2,94	3,28				
segadus			1,47	4,10		
ebakindlus			1,47	3,28		
ootusärevus					1,47	3,28
põnevus	2,94	1,64				
rahulik	4,41					
rahu		4,10				
rahulolu	2,94	0,82				
lõbus	2,94					
meeldivus	2,94					
mitte midagi					2,94	
tähelepanelikkus	2,94					
mõtlikkus						2,46
rahulolematu				2,46		
tüdimus				2,46		
normaalne					1,47	0,82
sõbralikkus	1,47	0,82				
väsimus			1,47	0,82		
ärevus			1,47	0,82		
abivalmidus		1,64				
armastus		1,64				
ei tea						1,64
hämming						1,64
kohusetunne						1,64
lootus		1,64				
naljakas						1,64
närviline				1,64		
teadmatus				1,64		
uimane				1,64		
unine				1,64		

Märkus: paksus kirjas on esile toodud suurem esiletulek.

On huvitav märkida, et kuigi nii psühholoogias läbi viidud enesekohaste uuringute tulemused (Veski 1996, Allik 1997) kui emotsioonisõnade lingvistiline analüüs (Vainik 2002a) osutavad, et emotsioonisõnavara saab olla eesti keeles ainult kas positiivse või negatiivse tähendusega, siis hetketunnete loetelukatse vastuste hulgas on märgatav hulk vastuseid, mille hinnangulist valentsi on raske määrata (*mitte midagi, mõtlikkus, normaalne, ei tea, hämming, argipäeva emotsioon, tavaline* jne). Seega intervjuuga seotud hetkekogemuse kirjeldamiseks ei paistnud eesti keele läbivalt kahevalentsed emotsioonisõnad sobivat ja katseisikud kirjeldasid oma kogemust ligi veerandil juhtudest teisi lingvistilisi vahendeid appi võttes. Üks võimalikke seletusi on see, et emotsionaalse kogemuse kaks karakteristikut intensiivsus ja valents on emotsioonimõistetes omavahel korreleeritud (nt Russell 1980, Russell, Lemay 2000, Zammuner 1998). Olukorras, kus emotsionaalne kogemus ei ole eriti intensiivne või on allasurutud, puudub selle kontseptualiseeringust ka positiivsus-negatiivsus hinnang ja seda on raske väljendada tavapärase emotsiooninimetustega.

5. Emotsioonidega seotud tegusõnade loetelukatse

Emotsioonidele on võimalik viidata sõnadega erinevatest sõnaliikidest (vt Vainik 2001), tegusõnu paluti katseisikutel loetleda eesmärgiga teada saada, millistest emotsioonidest meenuvad kergemini tegevuslikku aspekti rõhutavad mõistestused.

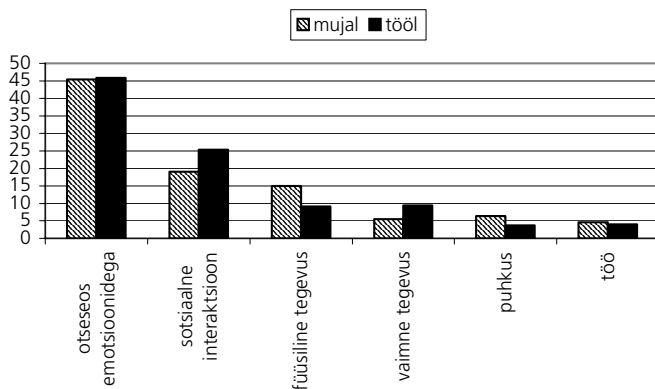
Ka katses osalejad lähtusid vastuseid andes rohkem sõna semantikast ja selle (tihti assotsiatiivsest seosest) emotsioonidega, kui küsimuses sisaldunud vormilisest kriteeriumist. Algvormis verbide kõrval (nt *vihkama, armastama, kurvastama*) nimetati paljudel juhtudel neist regulaarselt tuletatavaid teonimesid (nt *kurvastamine*) aga ka adjektiive (nt *rahulolematu*) või nimisõnu, mille seos emotsioonidega on aimatav, kuid seos tegevusliku aspektiga jäi küsitajale arusaamatuks (nt *üksindus*). Edaspidi esitatavates näidetes ja loeteludes on vastuste vorm jäetud muutmata.

Tööl on olnud loetelude pikkus keskmiselt 7 sõna (mujal 8,5), vastuste varieeruvus on tööl väiksem (keskmiselt 4,3 erinevat sõna, mujal 5,2) ja ka küsimuse vormilisele kriteeriumile vastas tööl vaid 61% vastustest (mujal 72%). On kaks tõlgendusviisi: kas mittetõine keskkond kuidagi erilisel soodustas emotsiooniteadmiste aktualiseerimist või on tõine keskkond mõjunud küsitletute leksikaalsetele emotsiooniteadmistele filtreerivalt.

Vaba loetelukatse tulemustes on alati sõnu, mis kuuluvad tähenduse poolest otse küsitletud kategooriasse ja sõnu, mis küsitud kategooriaga seostuvad kas assotsiatiivselt või nt põhjuse-tagajärje seose tõttu inimeste tavateadmistes. Tinglikult oli sõnade semantikast võimalik eristada järgnevad rühmad (joonis 2).

Tulemused osutavad, et emotsioonidega otse seostuvate verbide meenumises tööl ja mujal olulist kvantitatiivset vahet ei ole, detailsem vaatlus emotsioonide kaupa artikli järgmises alajaotuses võimaldab aga välja tuua küsitluskohaga seostunud kvalitatiivsed erinevused.

Olulisemad kvantitatiivsed erinevused puudutavad emotsioonidega kas assotsiatiivsetes või kausaalsetes seostes olevate verbide meenumise määra. Tööl



Joonis 2. Emotsioonidega seotud tegusõnade loetlemise katse tulemuste protsentuaalne jaotus peamistesse semantilistesse rühmadesse

küsitletutel on tekkinud need seosed sotsiaalse interaktsiooni ja vaimse tegevusega. Mujal küsitletutel on emotsioonidega suhteliselt rohkem seostunud füüsilised tegevused, puhkus ja kodused tööd. Kõikide olulisemate sõnaryühmade võrdlev esiletulek kahes küsitluskohas on esitatud lisa 1. Järgnev kujutab endast nende andmete sisulist avamist ja kommentaari.

5.1. Emotsioonidega otseselt seostuvate verbide alaryühma sisemine liigendus

Ligi pool vastustest (45%) nii tööl kui mujal osutasid leksikaalsete emotsiooniteadmiste eri tasanditele (vt lisa 1, semantiline põhigrupp “otseseos emotsioonidega”). Tüüpiliselt väljendavad need sõnad emotsiooni kogemist protsessina, mispuhul emotsiooni mõistestusse kuulub ajaline mõõde.

Töö juures oli suurem osakaal eestikeelsete emotsioonide põhinimetustega *viha*, *armastus*, *rõõm*, *kurbus* (Vainik 2002a) tuletuslikult seotud sõnadel: *armastama*, *rõõmustama*, *kurvastama*, *vihastama*, *vihkama*, *rõõmu tundma*, *rõõmu valmistama*. Tööl oli suurem viha ja kurbusega seotud verbide mainimine, mujal tulid meelde armastuse ja rõõmu verbid.

Nii tööl kui mujal olid enim nimetatud emotsioonide prototyüpseid väljendustegevusi kirjeldavad verbid *nutma* ja *naerma*. Töö juures on nimetatud ka viha väljendumist tegevusena (*raevutsema*, *märatsema*); õrnade tunnete avaldamist nimetati enam mujal (*kallistama*, *hellitama*, *paitama*, *silitama*, *musitama*, *suudlema*, *õrnutsemine*).

Töö juures mainiti ka verbe, mis kuuluvad emotsiooniteadmise põhitasandist üldisemale tasandile (*tundma*, *emotsionaalsus*) ja tehti viited emotsioonide regulatsioonile (*allasurumine*, *emotsioonide sissekogumine*, *enda vaoshoidmine*, *enesevalitsemine*, *emotsioonide väljaelamine*, *sisseelamine*) või puudumisele (*ükskõiksus*).

Mujal küsitletute loeteludes sisaldub rohkem mõisteid emotsiooniteadmiste eritasandilt. Näiteks sõnad, mis osutavad emotsionaalse pinge kasvule ja kahanemisele (*erutama*, *närveerimine*, *närveldama*, *kergenduse saamine*); seksuaalne veetle-

vus ja külgetõmme (*armuma, ihkama, armatsema, kirge tundma*); hedoonilised tegevused (*mõnulema, nautima, lõbutsema, joovastuma, joobuma, õnnelik olema, rahuldumine, rahulolu tundmine*), kognitiivseid tundeid (*üllatumine, imestama, imetlema, kahtlema, kahetsema*).

Tööl meenusid rohkem sotsiaalsete suhetega seotud emotsioonid (*hoolima, kaasa tundma, kadestama, häbenema, austama*), sõnad, mis on seotud ohu tunnetamisega (*ehmuma, hirmuma, kartma, muretsema*); meeleolu langusega (*leinama, nukrutsemine, piinlemine*) ja frustratsiooniga (*ärritama, rahulolematu, nõrdima, kibestuma, raevuma*).

Mujal küsitletute hulgas tuli esile sõnarühm, kuhu kuuluvad verbid, mis kirjeldavad emotsionaalset kehakeelt (*käsi plaksutama, jalgu trampima, ohkama, vilistama, žestid, pöördumine inimese poole*) ja näoilmeid (*naeratama, muigama, mossitama, grimassitama, silmapilgutus*).

Kokkuvõtvalt võib öelda, et kuigi emotsioonidega otseselt seotud (tegu)sõnu meenus tööl ja mujal ühevõrra ja peamised lausunud sõnad olid emotsioonide väljendustegevusi nimetavad sõnad *nutma* ja *naerma*, kuulusid tööl meenunud sõnad peamiselt emotsiooniteadmiste üld- ja põhitasandile, kergemini meenusid negatiivsed ja sotsiaalsete suhetega seotud emotsioonid. Mujal loetleti rohkesti emotsiooniteadmiste eritasandile kuuluvaid sõnu ja tuli esile emotsionaalse kehakeele mainimine.

5.2. Sotsiaalne interaktsioon

Suurim emotsioonidega assotsiatiivselt või põhjuslikult seotud semantiline põhi-grupp, mis nii tööl kui mujal vastajate loeteludest moodustus, oli sotsiaalne interaktsioon (vt joonis 2 ja lisa 1). Enim mainitud sotsiaalse interaktsiooni liigiks oli verbaalne suhtlemine, millest tuleb täpsemalt juttu järgnevas lõigus.

Lisaks on tööl rohkem mainitud kaasinimese aktiivse soodustamise teemat (*abistama, abivalmidus, hoolitsema, kaitsma, head tegema, laste kasvatamine, ravima, tohterdama, leevendama*) ning huvide konflikti lahendamise strateegiaid (*järele andma, andestama, leppima, kavaldama, kompromissi tegema*) ja suhtumist kirjeldavaid verbe (*sõbrustama, lugu pidama, austama, sallima, ignoreerima*).

Mujal küsitletute hulgas tuli emotsioonidega seoses veidi rohkem esile sõnarühm, mis kirjeldab omandussuhetega kaasnevaid asjaolusid (*andmine, jagamine, osa saama, saamine, vastu võtma, müümine, ostmine, tingima, maksmine, röövimine, varastamine, leidmine*).

5.2.1. Verbaalne suhtlemine

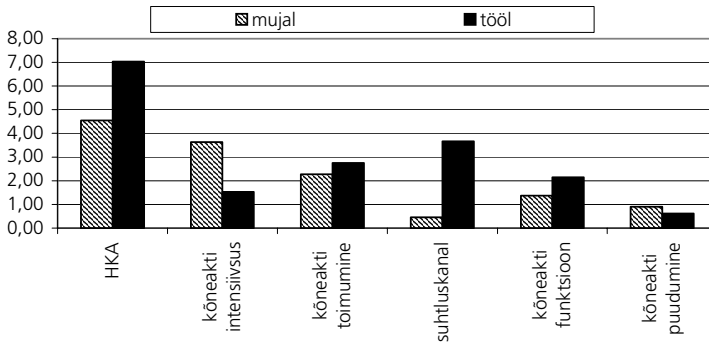
Nii tööl kui mujal läbi viidud intervjuudes seostus küsitletute jaoks emotsioonidega verbaalse suhtlemise tegevus, selle sõnarühma osakaal oli aga tõises keskkonnas 4% suurem (vt lisa 1).

Emotsioonidega seostus tööl küsitletute jaoks juba fakt, et suhtlemine toimub (sõnad nagu *rääkima, ütlemine, vestlemine, suhtlema, kuulama*) ning oluline oli

ka suhtluskanal (*helistama, mobiilile vastama, lugema, kirjutama*) ning kõneakti kommunikatiivne funktsioon (*küsima, pärima, intervjuerima, vaidlema, kor-dama, kajastamine*).

Mujal küsitlute emotsioonidega seostus tõsiasia, kui ootuspärast verbaalset suhtlemist ei toimu (*vaikima, varjama, lõugade kokkusurumine, hambad ristas*) või erineb kõneakti sooritusintensiivsus tavapärasest (*hüüdma, karjuma, kisama, möirgama, sosistama*); samuti verbid, mis tähistavad emotsioonide vastastikust verbaalset väljaelamist, nt *nääklema, nägelema, rüdlema*.

Keelelise suhtlustegevuse kirjeldamist on peetud pragmaatika pärusmaaks (Õim 1986). Selle paradigma raames on väidetud, et verbaalse suhtluse puhul kavandatakse ja tehakse sõnadega tegusid (Austin 1962). Kõneakte nimetavate verbide esiletulek emotsioonidega seostuvate tegusõnade loetelukatses räägib sellest, et eesti keele tavakasutajagi liigitab verbaalse suhtluse olulise ja emotsioone mõjutavate *tegude* alla. Teises keskkonnas küsitatud isikute jaoks on osutunud eriti kergesti meenuvaks kõneteod, mis sisaldavad hinnangut. Kuigi kõneakti hinnanguline iseloom seostus emotsioonidega ka mujal, on tööil sedasorti suhtlus meenuvad märksa enam (vt joonis 3).



Joonis 3. Emotsioonidega seostatud verbaalse tegevuse aspektide protsentuaalne esiletulek; HKA – hinnanguline kõneakt

Hinnangulist kõneakti (HKA) on defineeritud kui partneri emotsionaalse seisundi mõjutamise akti, mis on suunatud kõneleja suhtluseesmärgi saavutamisele (Volf 1985). Seega on hinnangulise suhtluse puhul emotsionaalse mõjutamise komponent olemuslik ning HKA-de assotsiatiivne seostumine loetelukatse teemaga igati ootuspärane. Näiteks on tööil loetletud verbe, mis nimetavad kõneakte, millega taotletakse muutust partneri emotsionaalses seisundis, olgu siis kas positiivset (*kiitma, vabandama, tänama*) või negatiivset (*laitma*).

Kõneaktide teoorias on eristatud lausungit, selle suhtluseesmärgi ja taotletavat mõju partneris (Austin 1962, Searle 1969). Nii ongi uuritavate verbirühmade semantikas olulisel kohal ühelt poolt verbiga kirjeldatava kõneakti propositsioonis olev hinnang (miski on hea/halb) ning lisaks kõneaktile omistatav emotsionaalne illokutiivjõud ja/või perlokutiivne efekt. Näiteks kiitus *Tubli töö!* sisaldab propositsioonis positiivset hinnangut ja selle eesmärki (illokutiivjõud) on adressaati innustada. Kui innustumine toimub, siis on see kõneakti perlokutiivne efekt. Ka kiitja emotsionaalne seisund peab vaikumise olema positiivne või vähemalt neut-

raalne, vastasel korral tõlgenduks antud kõneakt iroonia (vt Vainik 1990, Volf 1985).

Loetelukatses mainitud tegusõnad toovad esile teisi aspekte hinnanguliste kõneaktide võimalikest alussituatsioonidest. Osa sõnadest loetelukatses nimetavad kõnesituatsioone, kus eeldatav kõneleja on ise mingis emotsionaalses seisundis ja "elab seda välja" hinnanguna partnerile, eesmärgiga muuta tema emotsioone kindlas suunas. Näiteks *naljatama* tähistab olukorda, kus üks kõnesituatsiooni osalistest on heas tujus ja tahab ka partneri tuju tõsta. Negatiivsete tunnete ülekandmine negatiivse hinnangu abil kajastub sõnades, nagu *pahandama, kurjustama, õelutsema, õiendama, solvama, söimama, parastama, mõnitamine, halvustama*. Halvas tujus olles teises positiivse hoiaku tekitamise üritamist kirjeldavad verbid *õigustama, targutama*. Emotsioonide väljaelamist eesmärgiga esile kutsuda kaastunnet kirjeldatakse konnotatiivselt hinnanguliste verbidega, nagu *halama, hädaldama, kaeblema, vinguma, virisema*.

Kokkuvõtvalt võib sedastada, et sotsiaalse interaktsiooni (sh verbaalse suhtlus-tegevuse ja eriti hinnangulise suhtluse) seostamine emotsioonidega on ootuspärane ja kinnitab autori varasemaid tähelepanekuid, et eestlaste rahvapsühholoogilises intuiitses suhtlusmudelil on oma koht nii sotsiaalsel suhtel kui emotsioonidel (Vainik 1987, 2002b). Üllatav on vaid see, et teises keskkonnas on sedasorti sotsiaalsel interaktsiooni kirjeldav sõnavara märksa aktiivsem ja käepärasem kui teistes küsitluskohtades. Kui üldiselt on mujal kogutud andmetes olnud suurem mitmekesisus ja väiksem korduvus, siis verbaalse suhtlemise puhul äratav tähelepanu, et töö on alarühmad hästi mitmekesised, s.t emotsioonidega seostatakse samalaadseid nähtusi, aga eri inimestel eri sõnadega kirjeldatuna.

5.3. Füüsiline ja vaimne tegevus

Mujal küsitletute assotsiatsioonides oli ülekaal füüsilistel ja töö juures küsitletutel vaimsetel tegevustel (vt joonis 2 ja lisa 1). Füüsilise tegevuse alla on liigitatud liikumine, füsioloogilised ja aistingutega seotud protsessid ning füüsilised rünnakud. Tööl küsitletud on enam maininud horisontaalset liikumist (*jalutama, kõndimine, ujumine, jooksmine*), mujal küsitletud vertikaalset (*hüppama, kargama*) või kombineeritud liikumist (*tantsima, ronima, kiikuma*). Emotsioonidega seostuvate liikumisviiside erinevust eri küsitluskohtades võiks seletada liikumisviiside eeldatava eesmärgipärasuse vs. selle puudumisega: töökeskkond on soosinud sihipärast ja mittetöine keskkond lihtsalt lustist liikumist kirjeldavate sõnade meenumist.

Lisaks on mujal meenenud ka keha füsioloogilise toimimisega seotud protsessid (*higistama, hingama, magamine, sööma, joomine, palju energiat, seksimine, vananemine, haigestuma*), aistingutega seotud (*aistima, valu tundma, tõmblus, haigestuma, vigastuma, valutama, vaevlema, piinlema, väsima, lihaste pingutamine*) ja rünnakuid nimetavad verbid, mis võivad olla kas emotsiooni väljaelamine esemetele (*laamendama, lõhkuma*) või agressioon kaasinimeste suhtes (*kägistamine, lööma, kaklema, kerepeale andmine, torkimine*).

Vaimsetest tegevustest on nii töö kui mujal emotsioonidega seoses enam meenenud informatsiooni ümbertöötamine (*mõtlemine, mõtisklemine*), mis võib olla ka ebaratsionaalne (*mõttelend, unelema, unistamine*). Tööl on lisaks nimetatud ka

info ratsionaalse ümbertöötamise viise (*analüüsimine, juurdlemine, seostamine*) ja info hankimise soovi (*huvituma, uudistama, uurimine*), info tõlgendamise olulisust (*aru saama, taiplikkus, mõistma, mitte aru saama, uskuma, samastama, identifitseerima*) ja talletamist (*õppima, ei jää meelde*). Mujal on enam mainitud info vastuvõttu, kas läbi meeleorganite (*vaatama, nägema*) või vahenditult (*tunnetamine, aimama*).

6. Kokkuvõte ja arutelu

2001. aastal läbi viidud emotsioonikategooria liikmete loetelukatses ilmnis erinevus tulemuses olenevalt sellest, kas katse viidi läbi töö juures või mujal. Mujal küsitletutele meenusid võrreldava sageduse ja keskpärasusega eesti keele neli emotsioonide põhinimetust (*armastus, viha, rõõm, kurbus*), tööil küsitletutele esmajoonel *viha* ning märksa vähemal määral ülejäänud emotsioonid (Vainik, Orav 2005). Kontrollimaks küsitluskoha mõju loetelukatsete tulemustele vaadeldi käesolevas artiklis võrdlevalt hetketunnete ja emotsioonidega seotud tegusõnade loetelukatse tulemusi.

Hetketunnete osas selgus, et nii tööil kui mujal nimetati enam *rõõmu*, tööil sekundeerisid sellele *segadus, ärevus, huvi*, mujal *uudishimu, meeldivus, lõbusus*. Tööil küsitletute vastused olid keskeltläbi hinnangu osas vastuolulisemad ja määratlematuma valentsiga, mujal valdavalt positiivsed. Küsitletavate emotsionaalne seisund küsitluse ajal on taustategur, mis võis mõjutada kõikide antud seerias läbi viidud loetelukatsete tulemusi. Milline osa sellest küsitletute hetkeseisundist on tingitud teise *vs.* muu keskkonna püsijõust ja milline intervjuuga seotud eriolukorrast, on käesoleva katse andmete põhjal võimatu öelda.

Emotsioonidega seotud tegusõnade osas selgusid semantiliselt seotud sõnarühmad ja ainekvi kvantitatiivne jaotus. Emotsioonidega otseselt seotud verbide loetelud küll näitavad, millised tunded töö juures prevaleerivad (*viha, hirm, frustratsioon, meeleolu langus; sotsiaalsed tunded nagu kaastunne, hoolivus, kadedus, häbi, austus*), kuid lasevad ainult oletada nende tunnete põhjusti.

Erinevused emotsioonidega assotsiatiivselt või põhjuslikult seotud sõnavara esiletuleku proportsioonide vahel on kõnekad ja osutavad, et tööil küsitletute jaoks on olnud sotsiaalne interaktsioon, täpsemalt just keeleline suhtlemine üheks emotsionaalsemaks valdkonnaks. Suuga tehtavad "kõnetead" kalduvad töö juures olema suurema emotsionaalse kaaluga kui mujal, lisaks on teises suhtluses emotsionaalne väärtus suhtluse eri kanalitel (telefon, mobiil, arvuti). Veel üheks emotsioonidega enam seostunud valdkonnaks töö juures on vaimsed tegevused, sh nende õnnestumine ja ebaõnnestumine.

Pragmaatika uurijad on personaalses kommunikatsioonis eristanud sisu- ja suhtetasandit, millest sisutasand moodustub lausutavatest propositsioonidest ja suhtetasand partnerite omavaheliste suhete kinnistamisega seotud kommunikatiivsetest eesmärkidest (Watzlawick jt 1967). Teise suhtluse puhul on oletatud sisutasandi ja sellest tulenevalt n-ö infovestluste domineerimist, suhte hoidmise vestlusi on peetud teisejärguliseks (Brown, Yule 2003). Emotsioonide ja suhtetasandi tähtsust interpersonaalses kommunikatsioonis rõhutavad seevastu suhtlustreenerid ja praktikud. Näiteks propageeritakse nn aktiivset kuulamist (Gordon

2003) ja vägivaldalu suhtlemise tehnikat (Rosenberg 2005) ja räägitakse ka nn emotsionaalsest intelligentsusest töise suhtluse kontekstis (Goleman 2001).

Artiklis kirjeldatud loetelukatsetes esile tulnud suhteliselt suur hinnangulis-tele kõneaktidele osutavate tegusõnade hulk reedab, et töine keskkond on siiski läbi imbutud emotsioonidest ja et hinnang on üheks viisiks, kuidas emotsioonid tööalases keelelises suhtlemises ilmsiks tulevad. Osutus, et ka töö juures me mitte lihtsalt ei kasuta keelt info edastamiseks või selle küsimiseks, vaid mõjutame oma vestluskaaslase seisundit ning anname pidevalt tagasisidet enda emotsionaalse seisundi kohta, sh ka hinnangute abil.

Hinnangulise kõneakti sisuks e propositsiooniks on kas kaudne või otsene hinnang, mis vastab kõneakti sisutasandile. Suhtluseesmärgiks on hinnanguinfo vahendamise kõrval ka kuulaja seisundi, sh emotsionaalse seisundi, mõjutamine (Volf 1985). Hinnang annab suhtluses kas positiivset või negatiivset tagasisidet ja seda nii hinnangu adressaadi tegevuse kui kõneleja emotsionaalse seisundi kohta. Oletan, et kommunikatsiooni suhtetasandil luuakse ja taasluuakse hinnangulise suhtlusega kõnelejate ebavõrdsust. Hinnangud on mõeldavad kontekstis, kus hindaja ja hinnatava staatus on erinev. Tüüpiline näide on kool: õpetaja on see, kes paneb hindeid, pole mõeldav, et õpilane hindaks õpetajat. Hinnanguline suhtlus on üheks viisiks, kuidas kommunikatsiooni suhtetasandil luuakse staatuste vahet, domineerimist ja suhete hierarhiat. Otseku kompensatsiooniks suhtub keelekollektiiv hinnangulistesse kõneaktidesse omakorda hinnanguliselt, kasutades sellise suhtluse kirjeldamiseks konnotatiivselt negatiivse varjundiga sõnu, nagu *näägutama, parastama, õiendama* jne. Negatiivsetest emotsioonidest laetud kõneaktide subjektiivne taju on üldistunud negatiivseks hinnanguks, mida sellisele suhtlusele metatasandil omistatakse.

Asümmeetrilisi suhteid loovat kommunikatsiooni on keelelise pragmaatika uurijad omistanud valdavalt meestele, naiselik suhtlemisstiil pidavat kinnistama suhete sümmeetriat ja vastastikusust (Tannen 1990). Võimalik, et viited hierarhilise mudeli olulisusele tööl küsitletute jaoks seostuvad tõsiasiaga, et tööl küsitletute hulgas oli meeste osakaal suurem. Võimalik aga ka, et loetelukatsete tulemustes peegeldub fakt, et enamus töökeskkondi ongi hierarhiliselt organiseeritud ja igapäevane keeleline kommunikatsioon sellises keskkonnas on emotsionaalne eelkõige hinnangute kaudu. Kõneaktide hinnangulisus on nähtavasti viisiks, kuhu töises keskkonnas kanaliseeruvad emotsioonid, millele muidu ehk kohta ette pole nähtudki.

Hierarhiline suhetemudel omakorda diferentseerib emotsioonide väljendamist ja võimalik, et ka kogemist. Näiteks on psühholoogid leidnud, et viha väljendamine seostub staatuste vahega ja Eesti elanike hulgas on leitud tendents, et staatuselt kõrgema või võrdse inimese viha väljendamine on ootuspärane, staatuselt madalam inimene tõenäoliselt oma viha avalikult ei väljenda (Realo jt, käsikiri). On võimalik, et staatuselt madalam isik kogebki võrreldavas olukorras hoopis näiteks frustratsiooni, meeleolu langust või ärevust, tunde kategoriseerimine vihaks võib tunduda kohatu. Väärib edasist uurimist, kas ja mil määral diferentseerib hierarhiline suhetemudel ka teiste emotsioonide väljendamist ja kogemist peale viha.

Lõpetuseks võib öelda, et läbiviimise koht näis tõesti mõjutavat emotsiooni- sõnade loetelukatse tulemusi, eriti, mis puudutab küsitletud kategooriaga assotiativselt seotud sõnariühmade esiletulekut. Seega leidis artikli alguses püstitatud

hüpotees vähemalt toetust kui mitte otsest tõestust. Huvitava lisaboonusena andsid loetelukatsed kaudset informatsiooni teises keskkonnas toimuvast verbaalsest kommunikatsioonist. Osutus, et keelelist suhtlemist kasutatakse teises keskkonnas kaugeltki mitte üksnes info edastamiseks, vaid arvatavasti ka staatusega seotud rollisuhete loomiseks ja kindlustamiseks.

Kirjandus

- Allik, Jüri 1997. Psühholoogia lihtsusest. Tartu: Tartu Ülikooli Kirjastus.
- Austin, John L. 1962. *How to Do Things With Words*. Oxford: Oxford University Press.
- Brown, Gillian; Yule, George 2003. *Teaching the Spoken Language: An Approach Based on the Analysis of Conversational English*. Cambridge: Cambridge University Press.
- Davies, Ian; Corbett, Greville G. 1995. A practical field method for identifying basic colour terms. – *Languages of the World* 9 (1), 25–36.
- Goleman, Daniel 2001. Töö emotsionaalse intelligentsusega. Tallinn: Väike Vanker.
- Gordon, Thomas 2003. Tark lapsevanem. Tallinn: Väike Vanker.
- Realo, Anu; van Mechelen, Iven; Allik, Jüri. Dispositional, situational, and cultural determinants of anger experience and expression. *Käsikiri autori valduses*.
- Rosenberg, Marshall B. 2005. Vägivallatu suhtlemine. Tallinn: Nebadon.
- Ross, Norbert; Medin, Douglas L. 2005. Ethnography and experiments: Cultural models and expertise elicited with experimental research techniques. – *Field Methods* 17 (2), 131–149.
- Russell, James A. 1980. A circumplex model of affect. – *Journal of Personality and Social Psychology* 39, 1161–1178.
- Russell, James A.; Lemay, Ghyslaine 2000. Emotion concepts. – Michael Lewis, Jeanette Haviland-Jones (Eds.). *Handbook of Emotions*. Vol. 2. New York: The Guilford Press, 491–503.
- Searle, John R. 1969. *Speech Acts. An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Sutrop, Urmas 2001. List task and a cognitive salience index. – *Field Methods* 13 (3), 289–302.
- Zammuner, Vanda. L. 1998. Concepts of emotion: “Emotionness”, and dimensional ratings of Italian emotion words. – *Cognition and Emotion* 12 (2), 243–272.
- Tannen, Deborah 1990. *You Just Don’t Understand. Women and Men in Conversation*. New York: William Morrow and Company, Inc.
- Vainik, Ene 1987. Eesti keele suhtlussõnavara liigendus ja keelekasutaja intuiitiivne suhtlusmudel. – *Dialoogi mudelid ja eesti keel. Töid eesti filoloogia alalt*. Tartu Riikliku Ülikooli toimetised 795. Tartu: Tartu Riiklik Ülikool, 106–119.
- Vainik, Ene 1990. Tegusõnad sõnade tegemisest. – *Keel ja Kirjandus* 9, 549–556.
- Vainik, Ene 2001. Eestlaste emotsioonisõnavara. Käsikirjaline magistr töö Tartu Ülikoolis.
- Vainik, Ene 2002a. Emotions, emotion terms and emotion concepts in an Estonian folk model. – *Trames* 6 (4), 322–341.
- Vainik, Ene 2002b. Kas eestlased on “kuumaverelised”? Eestlaste rahvalikust emotsioonika-tegooriast. – *Emakeele Seltsi aastaraamat* 47 (2001). Mati Erelt (peatoim.), Evi Ross, Asta Õim (toim.). Tallinn: Emakeele Selts, 63–86.
- Vainik, Ene 2006a. Intracultural variation of semantic and episodic emotion knowledge. – *Trames* 10 (2), 169–189.
- Vainik, Ene 2006b. Emotsioonisõnad teises keskkonnas – millest kõnelevad loetelukatsete tulemused? – *Ettekanne V rakenduslingvistika kevadkonverentsil “Keel teises suht-luses”*, Tallinn, 20.-21. aprill 2006.

- Vainik, Ene; Orav, Heili 2005. Tee tööd ja näe vaeva, ...aga ikka oled vihane. – Keel ja Kirjandus 4, 257–277.
- Veski, Epp 1996. Emotsionaalseid seisundeid kirjeldav sõnavara eesti keeles ja selle seos isiksuse omadustega. Käsikirjaline diplomitöö Tartu Ülikoolis.
- Volf 1985 = Вольф Елена М. 1985. Функциональная семантика оценки. Москва: Наука.
- Watzlawick, Paul; Bavelas, Janet B.; Jackson Don D. 1967. Pragmatics of Human Communication. A Study of Interactional Patterns, Pathologies, and Paradoxes. New York, London: W. W. Norton & Company.
- Õim, Haldur 1986. Pragmaatika ja keelelise suhtlemise teooria. – Keel ja Kirjandus 5, 257–269.

Lisa 1. Emotsioonidega seotud tegu sõnade loetelukatse tulemuste kvalitatiiv-kvantitatiivne jaotus küsitluskoha tüüpide lõikes

Semantiline põhigrupp:	Alagrupp:	Täpsustus:	Mujal:	Tööl:
otseseos emotsioonidega			45,5	45,9
	emotsiooniteadmiste põhitasand		10,5	13,2
		viha	1,8	4,9
		armastus	3,2	3,1
		kurbus	1,4	1,8
		rõõm	4,1	3,4
	emotsiooniteadmiste üldtasand		1,4	4,3
	emotsiooniteadmiste eritasand		12,3	11,9
		sotsiaalsed	0,5	3,4
		hedoonia	3,6	1,5
		ohuolukorrad	1,4	2,5
		külgetõmme	1,8	1,2
		muutused pinges	1,8	0,9
		kognitiivsed	2,3	0,3
		frustratsioon	0,9	1,2
	meeleolu langus		0,9	
väljendustegevused		17,7	14,1	
	armastus	4,6	1,5	
	rõõm	7,3	5,5	
	kurbus	5,0	4,9	
	viha	0,9	2,1	
emotsionaalne kehakeel		3,6	2,5	

Semantiline põhigrupp:	Alagrupp:	Täpsustus:	Mujal:	Tööl:
sotsiaalne interaktsioon			19,1	25,4
	verbaalne suhtlemine		13,2	17,7
		suhtluskanal	0,5	3,7
		hinnangusuhtlus	4,5	7,0
		kõneakti funktsioon	1,4	2,1
		kõneakti intensiivsus	3,6	1,5
		kõneakti toimumine	2,3	2,8
		kõneakti puudumine	0,9	0,6
	omandussuhted		2,3	2,1
aktiivne soodustamine		1,4	1,8	
huvide konflikti lahendamine		1,4	1,8	
suhtumine		0,5	0,9	
füüsiline tegevus			15,0	9,2
	liikumine		7,3	4,3
		vertikaalne	3,2	0,6
		horisontaalne	0,9	2,1
		kombineeritud	1,8	
	füsioloogia		2,7	1,8
	aisting		2,7	1,5
rünnak		2,3	1,8	
vaimne tegevus			5,5	9,5
	info tõlgendamine			2,8
	info ümbertöötamine		2,3	3,4
	info vastuvõtt		2,3	1,8
	info hankimine		0,9	0,9
	info talletamine			0,6
puhkus			6,4	3,7
	aktiivne		5,9	2,1
	passiivne		0,5	0,9
	üldine			0,6
töö			4,6	4,0
	spetsiifiline		3,2	0,6
	üldine		1,4	3,4

Märkus: numbrid lahtrites "tööl" ja "mujal" tähendavad sõnade esiletuleku protsentuaalset osakaalu vastavas alajaotuses. Poolpaksus kirjas on esile tõstetud read, kus erinevus on suurem või võrdne 2,5%-ga.

Ene Vainiku (Eesti Keele Instituut) uurimisobjektiks on eesti keele emotsioonisõnavara, selle struktuur, semantika ja varieerumine. Huvipakkuvateks teemadeks on ka psühholingvistika, kognitiivne semantika ja keele ning emotsioonide interaktsioon laiemalt. Täiendanud ennast praktilise psühholoogia alal. ene@eki.ee

EMOTION TERMS AT WORK – WHAT DO THE RESULTS OF TASKS OF FREE LISTING REVEAL ABOUT?

Ene Vainik

Institute of the Estonian Language

The paper focuses on a comparison of the results of tasks of free listing of emotion terms as carried out at work and at leisure. The working hypothesis was that the environment where the task is carried out could also influence the results, at least as far as the task involves listing emotion-related terms.

As a result of semantic analysis of the elicited emotion terms and associatively related vocabulary it occurred, indeed, that the topicality of emotion knowledge varied according to the environment where the task was carried out. Presented and discussed are the results of two tasks – the task of listing one's own current emotions during the task and the task of listing emotion-related verbs.

It turned out that the main difference between the emotion terms elicited at the working place vs. at leisure lay in their evaluative value. Most of the terms listed at leisure referred to positive emotions, while terms elicited at work were both positive and negative. In addition, it was found that the participants had some difficulty in categorizing their current emotions by means of Estonian bi-valent emotion terms. In remarkably many cases the participants turned to lexis other than emotion terms to describe their current feelings.

The task of listing emotion-related verbs revealed different topicality of phenomena related with emotions associatively and causally. At work it occurred that acts of verbal communication and intellectual actions were mentioned much more frequently, while at leisure it was rather physical actions, leisure and housework.

A remarkably big part of the vocabulary referring to speech acts was evaluative by nature. The possible meaning of such a salient amount of evaluative communication emerging at the working place and its topicality for emotions was discussed by means of the theory of pragmatics. It was further hypothesized that emotionally relevant evaluative communication might be related to the hierarchical pattern of relations holding default in most working places. Thus, communication at work is hardly a pure exchange of information – this is what the tasks of free listing of emotion terms mostly revealed about.

Keywords: anthropological linguistics, semantics, pragmatics, Estonian

KAKSKEELSETE LEKSIKONIDE GENEREERIMINE PARALLEELKORPUSE BAASIL

Kaarel Veskis

Ülevaade. Paralleelkorpustel ja võrreldavatel korpustel on lisaks erinevate keelte või keelevariantide kontrastiivvuringutele ka mitmeid teisi kasutusvõimalusi nii teoorias kui praktikas, mitmeid neist võimalustest ollakse alles avastamas. Üks huvitavamaid suundi on sõnastike koostamine või täiendamine tõlkevastete ekstraheerimise läbi. Püüan järgnevalt anda ülevaate senisest tööst selles valdkonnas, pidades silmas eesti keelt puudutavate praktiliste rakenduste väljatöötamise võimalust tulevikus. Samuti kirjeldan katseid esialgse lahendusena kasutada inglise-eesti ja eesti-inglise erialasõnastiku lähtematerjali genereerimiseks väheseid keeleressursse (lisaks paralleelkorpusele) vajavat vabavarana saadaval olevat tarkvara.

Võtmesõnad: korpuslingvistika, arvutisõnastikud, kakskeelne leksikograafia, keeletöötlus, oskussõnastikud, eesti keel, inglise keel

1. Sissejuhatus

1.1. Paralleelkorpuste rakendused

Paralleelkorpus on korpus, mis sisaldab mingit teksti originaalkeeles ja selle tõlget teise keelde või tõlkeid teistesse keeltesse. Paralleelkorpuse paralleeltekstid võivad olla ka mõne kolmanda korpuse mittekuuluva teksti tõlked. Paralleelkorpuste kasutamiseks on neid korpuseid vaja eelnevalt rohkem töödelda kui tavalisi ükskeelseid tekstikorpuseid – kahe paralleelse teksti märgendus peab olema omavahel seotud.

1990. aastate algus tähistab mitmeid märkimisväärseid saavutusi paralleeltekstide vaheliste vastavuste automaatse tuvastamise osas (nt Brown jt 1991, Gale, Church 1993). Edasiste aastate jooksul on esile kerkinud suur hulk paralleelistamisega seotud probleeme, kuid ka palju huvitavaid lahendusi nendele probleemidele. Samuti

on tekkinud teadlikkus mitmesugustest uutest võimalustest, mida paralleelkujul keelekorpused võivad tähendada erinevate eluvaldkondade jaoks.

Enim levinud paralleelkorpuste rakenduslikud eesmärgid saab jagada kolme suuremasse rühma (Borin 2002: 14):

- 1) kontrastiivsed ja tüpoloogilised grammatikat ja leksikograafiat hõlmavad lingvistilised uurimused (vt nt Ebeling 1998);
- 2) paralleelkorpuste kasutamine masintõlkesüsteemides ja tõlkeabiprogrammides (nt Melby 2000); eraldi võib nimetada mitmesuguste toodete keelelist lokaliseerimist ning internatsionaliseerimist kui masintõlke kitsama fookusega allsuunda;
- 3) paralleelkorpuste kasutamine keeleõppes ja -õpetuses (nt Botley jt 2000).

Veidi marginaalsemate, kuid seda huvitavamate praktiliste kasutusviisidena võiks sõnastike ekstraheerimise kõrval esile tuua mitmekeelse infootsingu (Davis 1998) ja sõnastike käsitlemise paralleelkorpustena või paralleelkorpuste osadena, et luua uusi ja täielikumaid tõlkevahendeid (Geisler 2002), või mõnel muul eesmärgil. Samuti on võimalik paralleelkorpuse kasutada keele lingvistiliseks analüüsiks või sünteesiks vajalike vahendite automaatselt produtseerimiseks (Kuhn 2004), erinevates keeltes tekstide automaatselt kategoriseerimiseks (Gliozzo, Strapparava 2005) jne.

1.2. Automaatselt loodavate leksikonide vajalikkus

Tuleviku seisukohalt on keeletehnoloogias üks kahest põhisuunast, millele kogu maailmas tähelepanu pööratakse, uute lähenemiste leidmine arvutis loetavate sõnastike loomiseks kas korpuste või muude allikate toel (Muischnek jt 2003). Paralleelkorpustest automaatselt genereeritud sõnastikud leiavad rakendust nii inim- kui masintõlkes, keeleõppes ja ka mujal, näiteks sellise spetsiifilise arvuti- tehnoloogilise ülesande hõlbustajana nagu seda on semantiline ühestamine (Ide jt 2002). Esialgu saab rääkida ainult poolautomaatselt mitmekeelsete erialasõnastike koostamisest paralleelkorpuste baasil, sest ilma inimkorrektori parandusteta saab sellisel viisil luua üksnes leksikograafidele, terminoloogidele või tõlkesüsteemidele abiks olevat toormaterjali. Erialased tekstid sobivad leksikoni ekstraheerimiseks näiteks kirjanduslikest tekstidest paremini seetõttu, et erialased terminid tähistavad enamasti kindlapiirilisi mõisteid, millele leidub kindel ja järjepidev vaste paralleelteksti sõnade hulgas (Bowker, Pearson 2002: 171–174, 220, Fung 2000). Tõlke järjepidevusest sõltub suuresti ka leksikoni genereeriva programmi võime luua korrektseid seoseid sõnade vahel.

Isegi kui keele erialase valdkonna jaoks on juba olemas kakskeelne sõnastik, siis on üldjuhul paralleelkorpuse põhjal võimalik automaatselt genereerida olemasolevast tunduvalt mahukam sõnastik. Tänapäeva uut tüüpi korpusepõhises õppijasõnastikes esitatakse lisaks tavapärasele leksikaalsele ja grammatilisele infole ka teavet tähenduste piirangute, kollokatsioonide, grammatiliste mallide, stiili, registri ja kasutussageduse kohta ning luuakse seosed sõna struktuuri, kasutamise ja tähenduse vahel (Kitsnik 2006: 96). Korpusest automaatselt ekstraheeritud sõnastikud on üheks sellise teabe allikaks, võimaldades muuhulgas ka parandada

ja täiendada olemasolevaid sõnastikke nii uute kirjade osas kui ka näiteks muutes sõnatähenduste hierarhiat genereeritud sõnastikus sisalduva korpusest ammutatud sagedusinfo põhjal.

Olgugi et kaasaegne keeletehnoloogia võimaldab automaatselt genereerida vaid üsna suure veaprotsendiga sõnastikke isegi erialakeele puhul, on siiski tekstiressurside olemasolul mahukate ja osaliselt vigaste sõnastike loomine enamasti vajalik mitmetel põhjustel. Näitepõhise tõlkesüsteemi jaoks on genereeritava leksikoni juures veaprotsendist olulisem kirjade arv, kuna tõlgitava üksuse vastetest korpuses peab tõlke õnnestumiseks vaid ühel olema korrektne paralleelistus (Brown 1997: 6). Suurem maht, mis tähendab saagise eelistamist täpsusele,¹ on ka terminoloogi, tõlkija või sõnaraamatu koostaja seisukohalt parem lahendus. Lihtsam, kui otsida tekstist ise programmi poolt leidmata jäänud tõlkevasteid, on programmi poolt välja pakutud valesid märksõnakandidaate eraldada korrektsetest tõlgetest või kustutada.

1.3. Tekstide paralleelistamine

Oskus kasutada levinud tõlkeabiprogrammides (nt Trados, MultiTrans) või korpuseanalüüsiprogrammides (Wordsmith) sisalduvaid terminiekstraheerimismoduleid (või nende moodulite tööprintsipi tundmine) võib lisaks terminispetsialistidele suuresti abiks olla nii tõlkijatele, keeleõppijatele kui ka kõikidele tudengitele ning teadlastele, kes soovivad oma erialast terminoloogiat tundma ja kasutama õppida mõnes võõrkeeles. Lihtsat meetodit terminiandmebaasi moodustamiseks korpustest tuletatud sagedussõnastike ja võtmesõnade abil ilma spetsiaalse tarkvarata kirjeldavad Lynne Bowker ja Jennifer Pearson (2002: 137–164). Kakskeelse sõnastiku genereerimiseks on aga lisaks mõlema keele terminiandmebaasile vaja ka vastavuses olevad terminid paralleelistada.

Korrektne paralleelistamine on paralleelkorpuse hilisema kasutatavuse seisukohalt kõige olulisem ja ka üksnes indo-euroopa keeli hõlmava korpuse puhul kõige töömahukam probleem. Paralleelistamist raskendab asjaolu, et tihti sisaldavad paralleeltekstid “müra”, s.t ühes tekstis on midagi rohkem või vähem kui temale vastavas teises tekstis, mistõttu ei saa lauseid omavahel üksüheselt kokku viia. Siiski on automaatne lausetasandil paralleelistamine võimalik rohkem kui 90-protsendilise täpsusega. Fraasi- ja sõnatasandil paralleelistamine on aga tunduvalt keerulisem. Sõnade järjekord lauses sõltub keelest ja mitmetes keeltes (sh eesti keeles) on sõnade järjekord üsna vaba. Leksikaalsete üksuste piire on palju raskem tuvastada kui lausepiire. Ometi on ka sõnadevaheliste tõlkevastavuste automaatne tuvastamine teataval määral võimalik ja sellel põhinebki kakskeelsete leksikonide genereerimine.

Paralleelistamisel ehk joondamisel eristatakse statistilisi ja lingvistilisi meetodeid, kusjuures statistilisi meetodeid peetakse tõhusamaks suuremate korpuste ja lingvistilisi meetodeid väiksemate joondamisel (Oakes, McEnery 1998). Potentsiaalselt kõige edukamaks peetakse siiski statistiliste meetodite kombineerimist lingvistiliste meetoditega ja lisaressurside (sõnastikud) kasutamiseega. Statistiliste meetodite populaarsus on viimastel aastatel oluliselt tõusnud ka üldisemalt keeletehnoloogias ja arvutilingvistikas seoses infotehnoloogia kiire arengu ja järjest laiema levikuga.

¹ Saagise all on siinses artiklis mõeldud kõigi paralleelkorpuses sisalduvate ja omavahel tõlkelises seoses olevate sõnade või väljendite arvu suhet leksikoni jaoks ekstraheeritud korrektsete seoste arvuga. Täpsus on korrektsete tõlkevastavuste hulk leksikonis, võrrelduna kogu leksikoni mahuga.

2. Sõnastike genereerimine paralleelkorpustest

2.1. Sõnastike genereerimise meetodid

Erinevad lähenemised tõlkevastete automaatselt ekstraheerimisele jagunevad kahte põhikategooriasse: nn “hüpoteesi kontrollimise” ehk heuristilised meetodid (nt Smadja jt 1996) ja estimateerivad meetodid (nt Hiemstra 1997). Hüpoteesi kontrollimine tähendab tõlkevastete kandidaatide loendi genereerimist. Kandidaadid allutatakse statistilisele analüüsile, mis peab näitama, kas tegemist on tegelike tõlkevastetega või mitte. IBM-i teadlaste ideedest (Brown jt 1990)² lähtuvast statistilise masintõlke paradigmat inspireeritud estimateerivad meetodid põhinevad tõenäosusliku bitekst-mudeli loomisel, mis võimaldab tõlkevasteid hinnata mitte ainult eraldi, vaid ka rühmadesse jaotatuna. Mõlemal lähenemisel on oma plussid ja miinused. Allpool vaatluse alla tulevatest leksikoni genereerimise vahenditest võib PWA-d pidada heuristiliste meetodite ja Giza++ estimateerivate meetodite paradigmasse kuuluvaks (Tufiş, Barbu 2001: 156, Tiedemann 2003).

Varasemad lähenemised kakskeelse leksikoni ekstraheerimisele (K-vec algoritm, DK-vec algoritm jmt) põhinesid nn “ankurpunktidest” lähtuvatel lausetasandil paralleelistamise meetoditel, mida kombineeriti leksikaalse koosinemuse analüüsiga (nt Church jt 1991). Heuristilised meetodid sõnade joondamisel lähtuvadki üldiselt ideest, et tuleb leida sõnapaar, mis esineb koos märgatavalt sagedamini kui seda võiks lubada juhuse (t-score, Dice'i koefitsient jm).

Lisaks koosinemuse analüüsile on üldkasutatav ka sõnesarnasuse mõõtmine leidmaks omavahel etümoloogiliselt seotud sõnu. Sõnesarnasuse mõõdupuu näiteks on LCSR (ingl *longest common sub-sequence ratio*), mis tähendab kahe sõne pikima ühise tähejärgnevuse ja sõnepaari pikema sõne pikkuse suhet. Sõnesarnasuse mõõtmine eeldab, et mõlemal keelel on sarnane tähestik ja et etümoloogilises suguluses olevate sõnade kirjpiltide vahel eksisteerib arvestatav sarnasus.

2.2. Grammatilise analüüsi osa sõnastike genereerimisel

Eric Gaussier jt (2000: 254–255) toovad välja kolm faasi, milles sõnade või sõnaühendite tõlkepaaride ekstraheerimine paralleelkorpusest üldjuhul seisneb: tõlgitavate üksuste tuvastamine ning filtreerimine igas keeles eraldi, millele järgneb seoste leidmine tuvastatud üksuste vahel statistiliste algoritmide abil ning lõpuks leksikoni genereerimine omavahel seostatud sõnade loendi põhjal. Vajalike üksuste tuvastamine võib toimuda ka dünaamiliselt joondamisprotsessi käigus.

Kui kasutada sõnaühendite tuvastamiseks lingvistilisi meetodeid, siis on siinjuures eelduseks, et kõigi keelte (mõlema keele) lausete morfosüntaktiline analüüs toimub piisavalt heal tasemel ja sarnasel moel, mis alati ei ole võimalik. Samuti ei ole väga lihtne kindlaks teha esimeses faasis leitud grammatiliste mallide keeltevahelisi vastavusi, eriti juhul, kui soovitakse leksikoniga hõlmata ka mitmesõnalisi sõnaühendeid.

Viimast probleemi on püütud lahendada erinevatel viisidel. Üheks väljapääsuks on igasuguste grammatiliste korrelatsioonide leidmine keelepaaride vahel enne terminitele või soovitatavatele üksustele vastavate grammatiliste mallide tuvastamist.

Teine võimalus on kõigepealt leida grammatilise analüüsi abil kõik üksused ainult esimeses keeles olevates tekstides ja seejärel leida nende üksuste tõlked teises keeles. Kui aga eesmärgiks on luua kakskeelne ja kahesuunaline oskussõnastik, siis tuleks esmalt tuvastada terminid mõlemas keeles ja seejärel need terminid paralleelistada.

Grammatilise analüüsi abil leitud tõlgitavate keeleüksuste piire saab täpsustada ka paralleeltekstide põhjal loodud statistiliste tõlkemudelite abil – nõnda osutub reaalseks näiteks ka kõigi püsiühendite ekstraheerimine paralleeltekstidest (Melamed 2001). Võimalikud on ka keerulisemad algoritmid, mille korral sõnade või sõnaühendite grammatilise struktuuri vaatlemise abil parandataks joendamisel tehtud vigu, ning vastupidi, joendamismalle hinnates täpsustatakse vajalike üksuste tuvastamist (Gaussier jt 2000). Käesolevas artiklis pööran aga edaspidi põhitähelepanu grammatilist analüüsi minimaalselt kasutatavatele rakendustele, mille puhul nii sõnade või sõnaühendite tuvastamine kui ka paralleelistamine toimub suuremas osas statistiliste meetoditega.

Sõnastiku automaatsel genereerimisel paralleelkorpustest tuleb vahet teha eeldefineeritud terminite ekstraheerimise ja laiema tähendusega sõnastikugeneerimise vahel (mida võidakse siiski rakendada ka erialaste tekstide põhjal). Arvutilingvistikas määratletakse tehnilisi termineid sageli kitsalt teatud kindlatele morfosüntaktilistele tunnustele vastavate noomenifraasidena (Blank 2000: 240). Selliste kindlate omadustega noomenifraaside automaatne tuvastamine nõuab korpuselt kindlasti süntaktilist märgendust ja leksikoni genereerimine peab sellisel juhul toetuma konkreetse keele grammatilisele kirjeldusele. Seega moodustab keelest sõltumatute meetodite puhul selline kitsatähenduslik terminoloogia vaid osa korpustest genereeritavast leksikonist, kuna ekstraheeritavatele üksustele ei saa morfosüntaktilise info puudumisel kehtestada sõnaliigist vms lähtuvaid piiranguid.

Esialgu puudub eesti keele jaoks tarkvara, mis teostaks paralleelistamist ja sõnastiku genereerimist lähtudes eesti keele grammatilisest struktuurist ja selle struktuuri vastavustest mõne teise keele struktuuriga. Keelest sõltumatud meetodid tähendavad muuhulgas ka seda, et esialgsed tekstist ekstraheeritud üksused, mille omavahelise joendamise kaudu saadakse lõpuks sõnastik, peavad olema leitud põhiliselt statistiliste meetoditega. Mitmeid uurimusi, millest saab lähtuda kakskeelse leksikoni genereerimisel, on aga tehtud ka ükskeelsest korpusest statistiliste vahenditega mitmesõnaliste üksuste tuvastamiseks (Tiedemann 2003).

Kui lähtuda leksikoni võetavate üksuste joendamisel keele grammatikast, siis tekib järgmine probleem: kuigi sagedasti esinevad sõnaühendite grammatiliste mallide vastavused erinevate keelte vahel on tuvastatavad, esineb siiski suhteliselt palju mitmesusi ja kõrvalekaldeid reeglitest.

3. Sõnastiku genereerimise praktilised võimalused

3.1. Eeldused ja eeltöötlus

Mitmed kommertstarkvara tootjad (Xerox, Ahead Software, SensoLogic, TRADOS) pakuvad mõne oma tarkvarapaketi osana ka terminisõnastiku automaatse ekstraheerimise võimalust etteantud tekstide põhjal, kuid nendel programmidel puudub

esialgu eesti keele tugi ning kommertstarkvara poolt kasutatavad meetodid ei võimalda töö kohaldamist teiste keelte tarbeks. Seetõttu on esimeseks loogiliseks sammuks teel eestikeelse osalusega sõnastiku automaatse genereerimise poole keelest sõltumatute meetodite katsetamine vabavara abil või eesti keelt toetava leksikonigenererimistarkvara loomine.

Mida tuleks silmas pidada, kui on kavas välja töötada eriotstarbelist leksikograafia tööle orienteeritud tarkvara, mis genereerib paralleelkorpuse põhjal muuhulgas loodavasse sõnastikku sobivaid sõna- või väljendipaare?

Enamik leksikoni ekstraheerimise alastest töödest põhineb kindlatel eeldustel, mida tuleks arvesse võtta eriti juhul, kui tarkvara loomisel soovitakse alustada lihtsamatest algoritmidest. Ükski nendest postulaatidest ei kehti tegelikkuses saja-protsendiliselt, kuid erandid ei põhjusta nii suurt langust tulemuste kvaliteedis, et eelduste rakendamine poleks põhjendatud. Need eeldused on:

- a) mitmetähenduslikku leksikaalset üksust kasutatakse ühe ja sama teksti siseselt ainult ühes kindlas tähenduses;
- b) tõlkeüksuste paari kuuluvate sõnade sõnaliigid peavad omavahel sobima, s.t näiteks verbile ühes keeles võib vastata üksnes verb või mõni teine sõnaliik, mis on tunnustatud võimeliseks täitma tõlkes verbi funktsiooni –see reegel ei ole muidugi tegelikkuses absoluutne;
- c) tõlkevastete kandidaatide seas on tõenäolisemad tõlkevasted need, millesse kuuluvate sõnade suhteline asend lauses on üksteisele lähedasem;
- d) tõlkepaari ühe poole leksikaalsele üksusele vastab maksimaalselt üks leksikaalne üksus tõlkepaari teisel poolel.

Esimene eeldus vastab paraku seda vähem tõele, mida vähemkasutatava sõna või sõnaühendiga on tegu, ja ka suur hulk oskussõnu tõlgitakse teise keelde mitmel erineval moel isegi sama teksti siseselt. Ingeborg Blank leidis prantsuse-saksa näidiskorpuse varal tehtud katse abil, et 5–15% terminitest on sellised, millel on teises keeles rohkem kui üks tõlkevaste (Blank 2000: 246–247). Ühe termini tõlkevasted võivad olla aga erineva grammatilise struktuuriga. Nii näiteks esineb sõna *sihtliikmesriik* vastena Tartu Ülikooli (TÜ) inglise-eesti paralleelkorpuses kahel korral *the Member State of destination*, aga ühel korral ka *the destination Member State*; sõna *lähetuskoht* esineb inglise keeles kaheksal korral kui *place of dispatch*, kahel korral kui *place of destination* ning kahel korral vastab noomenifraasile hoopis umbisikuline verbivorm: *(the products) are dispatched from*. I. Blank (2000: 247) toob näite selle kohta, kuidas saksa keele liitsõnaline termin *Einspruchsbeschwerdeverfahren* esineb vastavates prantsusekeelsetes tekstides järjepidevalt kujul *procédure de recours engagée à l'encontre d'une décision rendue sur opposition* (kompleksne noomenifraas). Selliste mittevastavuste võimalikkust tuleb arvesse võtta nii terminite automaatsel piiritlemisel kui ka joondamisfaasis.

Ka viimast, üks-ühele vastavuse eeldust on leksikonide genereerimisel küll laialdaselt kasutatud, kuid ka see eeldus tekitab siiski suhteliselt palju ebakorrektsid tõlkeid, kui ühe keele liitsõnale vastab teises keeles mitmesõnaline väljend. Niisiis on see nn “1:1-kaardistuse hüpotees” inglise-eesti keelepaari korral küsitav: eesti keele liitsõnale vastab inglise (või ka näiteks prantsuse) keeles tavaliselt mitmesõnaline üksus. Seda probleemi on täheldanud Pim van der Eijk (1993), kelle uurimus põhineb inglise-hollandi korpusel, Lars Ahrenberg jt (1998), kes tegelesid inglise-rootsi korpusega ning samuti I. Blank (2000), tuginedes saksa-inglise-prantsuse

corpusele, jt. Keelepetsiifilise eel- (automaatne segmenteerimine) ja järeltöötuse (osaliste tõlgete filtreerimine) abil on 1:1-kaardistusest tulenevad probleemid vähemalt osaliselt ületatavad (Tufiş, Barbu 2001: 157). Parema tulemise saamiseks tuleks siiski vähemalt eesti keele puhul üks-ühele paralleelstusega ühendada mitmesõnaliste üksuste tuvastamine ja joondamine.

Paralleelcorpuse eeltöötlus võiks lisaks leksikaalsete üksuste segmenteerimisele eesti keele puhul hõlmata ka Kadri Muischneki (2006) poolt käsitletud inglise-eesti masintõlke kvaliteedi parandamiseks sobivaid meetodeid – ühendverbide restruktureerimist ning liitsõnade osadeks jaotamist. Selline eeltöötlus eeldab aga omakorda morfoloogilist analüüsi, mis tähendab küll eemaldumist esialgsest keeltevahelise portatiivsuse printsiibist, kuid võimaldab tõlkevasteid ekstraheerida tunduvalt lihtsamalt.

3.2. Lihtsa leksikograafilise abivahendi kavand

Morfoloogiline analüüs lubaks tarkvara kavandamisel lähtuda kõigepealt ainult ühest sõnaliigist, näiteks verbidest. Sellisel juhul peaks programm esmalt morfoloogilise analüsaatori abil kõik valitud sõnaliigi esindajad sisendcorpuses tuvastama ja lemmatiseerima ja corpuse paralleelstama lausetasandil. Paralleelstamiseks võib kasutada mõnd Gale'i ja Churchi algoritmi modifikatsiooni (nt Davis jt 1995), mille puhul on võimalik enne joondamist laused filtreerida, teostades joondamise ainult punktuatsiooni, pärisnimede vms põhjal. Seejärel võiks luua nimekirja kõigist võimalikest valitud sõnaliiki kuuluvatest keeltevahelistest sõnapaaridest, mis ei ületa oma joondamisüksuse piire. Selline loend võib endast juba kujutada arvestatavat abimaterjali leksikograafide, kuid tulemise parandamiseks peaks programm lisaks võrdlema kõigi corpuse valitud sõnaliiki kuuluvate keeltevaheliste sõnapaaride elementide omavahelist ortograafilist sarnasust ja koosinemise tõenäosust (väljendatuna näiteks seosetugevuse üldtuntud mõõdu, Dice'i koefitsiendi või mõne selle variandina). Kuna tegemist on sõnapaaridega, siis saab seejuures rakendada kollokatsioonide analüüsimisel kasutatavaid meetodeid (Brew, McKelvie 1996: 48).

Igale sõnapaarile omistaks programm nende näitajate alusel arvilise märgendi, mis iseloomustab tõenäosust, et tegemist on vastastikuste tõlgetega. Kasutajale esitatakse edasiseks töötamiseks üksnes need sõnapaarid, millega vastavusse seatud arv ületab kasutaja seatud lävendi. Nõnda saab tekstist automaatselt esile tuua väidetavalt 30% kõigist tekstis leiduvatest korrektsetest tõlgetest, täpsusega 90% (Brew, McKelvie 1996: 51). Kui tõenäosuse arvutamisel lähtuda üksnes ortograafilisest sarnasusest, siis on võimalik esile tuua ka potentsiaalsed eksitavad valepaarid (nn *faux amis*) – kirjapildilt üksteisele sarnanevad, kuid tähenduselt erinevad sõnad (Brew, McKelvie 1996).

3.3. Poolautomaatsed vahendid

Kui tarkvara väljatöötamine ei ole mingil põhjusel võimalik või otstarbekas, siis on sõnastiku genereerimiseks võimalik kasutada ka vabavara.³ Kõige lihtsam viis sõnade käsitsi joondamiseks või ka paralleelteksti põhjal sõnastiku toormaterjali

³ Vabavarana saadaval olevatest sõnatasandil paralleelstamist võimaldavatest tarkvarapakettidest ülevaate saamiseks vt <http://www.cse.unt.edu/~rada/wa/#softwareWA> (21.08.2006).

loomiseks on kasutada selleks mõnd graafilist joondamisvahendit, mis võimaldab biteksti üksteisega vastavuses olevad sõnad hõlpsasti omavahel ühendada ja genereerida automaatselt omavahel seostatud sõnade loendi.⁴

Järgmine samm täisautomaatse sõnastikugeneereerimise suunas on selline poolautomaatne protsess, mille puhul tarkvara poolt teostatud sõnaparalleelistuse tulemused vaadatakse inimkasutaja poolt üle ja vajadusel parandatakse. See on võimalik näiteks kasutades Chris Callison-Burchi poolt loodud graafilist abivahendit⁵, mille sisendiks on vabavarana saadaval oleva tarkvarapaketi Giza++ (Och, Ney 2000) poolt sõnatasandil paralleelistatud paralleeltekst. Kasutajale kuvatakse paralleelistus maatrikstabelina, kus on hõlpsasti võimalik parandusi teha.

Giza++⁶ tööprintsibid hõlmavad sarnaselt suurema osaga statistilise masintõlke rakendustest IBM-i uurijate mudeleid, mida tutvustasid Peter F. Brown jt (1993). Statistiline masintõlge rajaneb nn müranivooga kanali (ingl *noisy channel*) meetodil, mis võimaldab kasutada mitmeid informatsiooniteooriast, side, kommunikatsiooni, raadio, kõnetuvastuse jm valdkondadest pärit algoritme. Näiteks tõlkides inglise keelest eesti keelde tuleb leida selline eestikeelne lause, mille puhul tõenäosus, et see eestikeelne lause on ingliskeelse lause tõlge, on suurim. Selle tõenäosuse väljaarvutamiseks Bayesi valemi abil on tarvis teada ka tõenäosust, et mingi lause üldse eesti keeles võib esineda. Need tõenäosused saab leida piisavalt suure paralleelcorpuse alusel (Muischnek jt 2003: 53–54).

IBM-i mudel 1 leiab sõnadevahelised vastavused lausetasandil joondatud bitekstist sõnade koosesinemise alusel, alustades ühtlustatud tõlkevastetõenäosustest. Mudel 2 lisab sellele lihtsale tõlkemudelile positsioonilised parameetrid ning mudel 3 nn viljakusparameetrid. Viljakusparameetritega tuuakse esile mõnede sõnade kalduvus olla tõlkelises ühenduses tõenäolisemalt ühe- või mitmesõnalise vastega, mis sisaldab teatud arvu sõnu. Mudel 4 hõlmab meetodeid mitmesõnaliste üksuste tuvastamiseks biteksti põhjal genereeritud sõnaklasside võrdlemise abil lauses. Mudeliga 5 on püütud parandada eelmiste mudelite töö käigus esile tulnud vigasid (Brown jt 1993).

3.4. Täisautomaatne leksikoni genereerimine tarkvarapaketi PWA abil

Rootsi teadlaste projekti Plug raames välja töötatud kahe joondamisrakenduse – Linköping Word Aligner (LWA) ja Uppsala Word Aligner (UWA) – loomisel ja arendamisel on muuhulgas arvesse võetud eespool kirjeldatud sõnade üksühese vastavuse eeldusega seotud probleeme. Mõlemad süsteemid kasutavad tekstide sõnatasandil joondamiseks võrdlemisi vähe keespetsiifilist infot ning on seetõttu üpris lihtsalt rakendatavad ka eestikeelse osalusega sõnastike koostamiseks; mõlemad süsteemid on koostatud programmeerimiskeeles Perl ning on Internetist tasuta allalaaditavad.

LWA ja UWA on integreeritud paralleelcorpuste töötlemiseks loodud laiema funktsionaalsusega tarkvaraplatvormi nimega Uplug (Tiedemann 2002). Projekti Plug käigus loodud süsteemide peamiseks rakendusvõimalusteks on peetud (Sågval Hein 2002) erinevate masintõlkelikeide täiendamist tõlkeinfo ja leksikonidega ning samuti inimtõlkijate abistamist leksikonide loomise läbi. Vaatlen järgnevalt mõlemat joondamisprogrammi ja nende kasutusvõimalusi lähemalt.

⁴ Vt nt <http://www.umiacs.umd.edu/~nmadnani/alignment/forclip.htm> (21.08.2006)

⁵ Vt <http://demo.linearb.co.uk:8080/sandbox/start.jsp> (21.08.2006)

⁶ Vt <http://www.fjoch.com/GIZA++.html> (21.08.2006)

UWA sisendiks on lause- või fraasitasandil joondatud bitekst. Operatsiooni-süsteemi Windows jaoks praegu saadaval olev PWA (Plug Word Aligner – tarkvara-pakett, mis ühendab endas UWA ja LWA) versioon⁷ on eelseadistatud rootsi-, inglise- ja saksakeelse sisendteksti jaoks, kuid seadistusi on võimalik kohaldada ka teistele keeltele. Sisendtekst jagatakse programmi poolt kas juba olemasoleva (kasutaja poolt lisatud) lihtsavormilise sõnastiku või teksti enese põhjal ühe- või mitmesõnalisteks üksusteks. Sõnastiku puudumisel arvestatakse selle etapi juures tähejärjendite sagedusi ja pikkust, samuti sõnatüüpe ja punktuatsiooni. Järgnevalt püütakse mõlema keele vastavad üksused omavahel kokku viia. See protsess algab nn “kindlate juhtumite” (nt kui korpuse paralleelstatud segment kujutab endast vaid ühesõnalist elementi) eristamisest. Seejärel hindab süsteem sõnade omavahelist sarnasust ja suhtelist paiknemist tekstis (arvestades konteksti) ja märgib tõlkevastete kandidaatidena ära eelnevalt seatud lävendid ületanud sõnepaarid. Eraldi üritatakse seostada vähesagedasi sõnesid. Viimaks toimub tõlkevastete automaatne filtreerimine, millele võib järgneda tulemuste “käsitsi” korrigeerimine.

4. Kakskeelse leksikoni genereerimine PWA abil TÕ paralleelkorpuse põhjal

4.1. Uppsala Word Aligner

Järgnevalt kirjeldan katset luua UWA abil lähtematerjal erialasõnastiku koostamiseks või täiendamiseks, kasutades selleks mahukat paralleelkorpust (730 880 paralleelset ühest või mitmest lausest või (ala)pealkirjast koosnevat lõiku, 24 169 586 sõnet).

TÕ inglise-eesti paralleelkorpuse⁸ ühendasin leksikoni koostamise otstarbeks JRC-Acquis’ mitmekeelse paralleelkorpuse⁹ inglise-eesti alamosaga, kuna mõlemad korpused sisaldavad Euroopa Liidu seadusandlusega seotud tekste. Mõlemad korpused on samuti paralleelstatud Vanilla paralleelstatija¹⁰ abil.

TÕ korpus oli algselt kujul, kus eesti ja ingliskeelsed üksused paiknevad vaheldumisi ja on üksteisest eristatud keele nimetust sisaldavate märgenditega. JRC-Acquis’ korpus oli algselt mitmesugust märgendust sisaldaval TEI-kujul, jaotatuna paljudesse failidesse. Et muuta need korpused PWA-le “arusaadavaks”, tuli korpused teisendada kujule, milles mõlema keele üksused on ümber tõstetud kahte eraldi faili ja on tähistatud omavahel vastavuses olevate numbriliste tähistega. Selleks kasutasin osaliselt Camelia Ignat’ poolt loodud Perli-programmi¹¹.

Katse¹² käigus genereeritud inglise-eesti õiguskeele leksikon sisaldab 130 865 märksõna (vrd EKI inglise-eesti elektrooniline sõnastik – u 86 000 märksõna, Eesti Õiguskeele Keskuse terminibaas ESTERM – 57 829 märksõna) ja 482 571 sõnet. Sisendina kasutatud koondkorpus hõlmas 730 880 paralleelset ühest või mitmest lausest või (ala)pealkirjast koosnevat üksusepaari. Lisaks sisendkorpusele hõlmas programmi kasutajapoolne sisend ka väiksemahulist morfoloogiainfo faili

⁷ Vt <http://stp.ling.uu.se/plug/pwa/index.html> (21.08.2006)

⁸ Vt <http://www.cl.ut.ee/korpused/paralleel/index.php?lang=et> (21.08.2006)

⁹ Vt <http://langtech.jrc.it/JRC-Acquis.html> (21.08.2006)

¹⁰ Vt <http://nl.ijs.si/telri/Vanilla/> (21.08.2006)

¹¹ Vt <http://wt.jrc.it/lt/Acquis/JRC-Acquis.2.2/alignments/index.html> (21.08.2006)

¹² Nii UWA kui LWA abil genereeritud leksikonid on allalaaditavad aadressil www.teatja.ee/leksikonid.zip (29.08.2006). Lisaks koondkorpusest ekstraheeritud leksikonidele leiab sealt ka üksnes TÕ korpuse põhjal UWA-ga genereeritud sõnastiku ning Giza++ sõnaparalleelstatused koos Giza++ sisendiks olnud TÕ korpuse alamosaga.

¹³ Vt <http://www.eki.ee/dict/inglise/> (21.08.2006)

eesti keele tüüpiliste sõnalõppude ja ebareeglipäraste verbide kohta ning ka lähtesõnastikku, milleks otsustasin valida EKI inglise-eesti sõnastiku¹³. EKI sõnastik tuli UWA jaoks teisendada sobivale kujule, mis tähendas põhiliselt mitmesugust lühendamist – mittesobivate kirjete, sulgudes asuvate märkuste, liigsete vastete jms automaatset kustutamist UNIXI vahenditega.

Esitan juhusliku parandusteta fragmendi (1) UWA poolt TÜ inglise-eesti paralleelkorpuse ja JRC-Acquis' mitmekeelse paralleelkorpuse inglise-eesti alamosa põhjal genereeritud leksikonist, mis ei ole esitatud UWA graafilise liidese vahendusel, vaid tavalise tekstifailikatkendina:

```
(1)
{reserve officer}
{
  1X:reservohvitseri
}
{reserve officer candidates}
{
  1X:reservohvitserikandidaat
}
{reserve officer courses}
{
  1X:reservohvitserikursusel
}
{reserve positions and}
{
  2X:reservipositsioonid ja
}
{reserve power system}
{
  1X:reservelektrisüsteem
}
{reserve ratio}
{
  3X:reservibaasi
  2X:reservimäärä
}
{reserve service}
{
  1X:reservteenistus
```

Nagu näitest (1) paistab, on väljundsõnastikus ka eesti keele puhul üsna edukalt lahenenud juhtumid, mille puhul eestikeelsele sõnale vastab inglise keeles mitmesõnaline üksus. Küll aga hakkavad silma rohked keeltevahelistest morfoloogilistest erinevustest tingitud ebatäpsused, mis kehtivad samamoodi eesti-inglise keelepaari puhul nagu ka rootsi ja inglise keelte puhul (UWA on loodud arvestades rootsi keele eripärasid, kuid mitmeid väljundi vigasid ei ole suudetud ka rootsi keele puhul esialgu parandada).

Kuna UWA paneb tõlkevastavuste leidmisel suurt rõhku sõnesarnasusele, siis tulenevad sellest inglise ja eesti keele mittesuguluse tõttu paljud vead ja osalised vastavused¹⁴ väljundis, näiteks:

¹⁴ Võib oletada, et mitmed vead on tingitud ka sellest, et sisendkorpuse suure mahu tõttu ei suutnud programm katse käigus läbi viia viimast, automaatse filtreerimise etappi.

```

(2)
{selling}
{
  1X:Lepingus
  1X:Selleks
  1X:eelkõige
  2X:lepingu
  3X:müümine
  10X:selle
  1X:selles
  1X:sellest
  1X:selline
  2X:sellise
  1X:sellisel
  1X:selliselt
  2X:selliste
}

```

Anna Sågvald Hein (2002: 74) väidab, et erinevast morfoloogiast tulenevad ebatäpsused lahendaks rootsi keele puhul olemasoleva ülekandepõhise tõlkesüsteemi (Multra) liitmine UWA-ga. Eesti keele jaoks sellist tõlkesüsteemi kahjuks esialgu ei leidu. Vähene keelespetsiifiline morfoloogiainfo ei võimalda UWA-l teostada automaatset lemmatiseerimist, kuid oletan, et tõenäoliselt parandaks väljundsõnastiku korrektsust siiski sõnastiku genereerimisele eelnev sisendkorpuse lemmatiseerimine muude vahenditega. Tulemust parandaks kindlasti JRC korpuse (sisendkorpuse) automaatse paralleelsetuse manuaalne korrigeerimine, samuti süsteemi poolt kasutatavate statistiliste meetmete osaline laiendamine sisendkorpusele Internetile eesmärgiga kahandada andmehõreduse efekti. S.t üks süsteemi alam moodul võiks olla ühendatud mõne Interneti (mitmekeelse) otsimootoriga, millele võib esitada samu päringuid kui sisendkorpuselegi.

PWA-d puudutavatest publikatsioonidest ei selgu, kas süsteem kasutab rootsi keele puhul morfoloogilist infot ka selleks, et parema joondamistulemuse eesmärgil liitsõnad eelnevalt koostisosadeks liigendada (nagu seda on mitmel puhul tehtud saksa-inglise sõnaparalleelisel – nt Déjean jt 2002: 6).

Genereeritud leksikoni saaks täiustada ka mitmesuguse automaatse järeltöötusega, mis võiks tegelikult olla ekstraheerimistarkvarasse integreeritud. Näiteks oleks soovitatav kustutada kõik üksnes mitteamfabeetilisi sümboleid sisaldavad märksõnad, kui soovitud tulemus kujutab endast loomuliku keele leksikoni. Kasutaja poolt lisatud morfoloogiainfo põhjal saaks sõnastiku genereerinud moodul kustutada ka ühe ja sama märksõna erinevaid muutelõppe sisaldavatest, kuid muus osas sarnastest tõlkevastetest ebavajalikud. Samuti ei oleks vaja kohelda erinevate leksikaalsete üksustena identseid sõnu, mille algustähed on lausesisesest positsioonist tingituna erineva suurusega.

A. Sågvald Hein (2002: 74) näeb süsteemi arendusvõimalustena lisaks selle täiendamisele reeglipõhiste meetoditega veel sõnastiku märksõnade lemmatiseerimist ning viidete lisamist genereeritud sõnastikukirjetelt kontekstidele korpuses.

4.2. Linköping Word Aligner

LWA loomisel on toetutud Pascale Fungi ja Kenneth Churchi (1994) ning Dan Melamedi (1997) joondamisalastele uurimustele. LWA hõlmab nagu UWA-gi põhiliselt statistilisi meetodeid. Algoritm on iteratiivne – tõlkevasted genereeritakse biteksti põhjal, siis kustutatakse genereeritud sõnapaarid bitekstist ja korratakse tsüklit. Statistilistel tõenäosustel põhinevat baasalgoritmi täiendavad neli lisamoodulit ning programmi kasutajaliides võimaldab seadistada mitmeid parameetreid (lisateetid, lävendid, tsüklite arv).

Esimene moodul alustab tööd sõnade jagamisega etteantud info põhjal mitmesugustesse kategooriatesse ja alamkategooriatesse: relevantseteks ja irrelevantseteks (irrelevantsetena on määratletud näiteks inglise keele abiverb *do*, mille enamasti puudub üksühene vaste teises keeles), suletud ja avatud sõnaklassi sõnadeks. Avatud sõnaklassi kuuluvaid sõnu saab edaspidi joondada vaid teiste sama sõnaklassi sõnadega ja suletud sõnaklassi sõnu saab vastavalt joondada ainult suletud sõnaklassi sõnadega. Suletud sõnaklassi sõnad jagatakse järgnevalt veel alamkategooriatesse. Kategooriate põhjal toimub iteratiivne paralleelistamine.

Morfoloogiamoodul tunneb vastava keele sufiksiloendi järgi ära ühe sõna erinevad vormid. Kui leksikaalsete üksuste paar (X, Y) on mõne eelneva tsükli käigus tunnistatud tõlkevastavuses olevaks, siis otsib moodul teisi kandidaatpaare, mille esimene element on X ja teine element Z, nõnda et leiduvad ka sõned W, F ja G, mille puhul $Y = WF$ ja $Z = WG$ ning F ja G sisalduvad sufiksiloendi ühes ja samas paradigmas. Kui leitakse mitu erinevat üksust, mida saab tähistada sümboliga Z ja mille sufiksud kuuluvad erinevatesse paradigmadesse, siis valitakse neist suurima sagedusega paradigma.

Järgmine moodul tegeleb mitmesõnaliste üksuste paralleelistamisega, kasutades selleks samasuguseid võtteid kui ühesõnaliste üksuste puhulgi. Mitmesõnaliste üksuste kogum koosneb süsteemi poolt automaatselt leitud üksustest ning eelnevalt lisatud keelespetsiifilistest kollokatsioonidest. Lõpuks parandatakse paralleelistust veel vastavalt joondatavate segmentide suhtelisele asendile tekstides.

Sisendkorpuse mahupiirangu tõttu kasutasin LWA-ga katselise sõnastikumaterjali genereerimiseks TÜ paralleelkorpuse 10 000 paralleelsest lõigust koosnevat alamosa. Toon siin ära fragmendi tulemuseks saadud leksikonist, mis sisaldas 8516 kirjet:

(3)	
kinnitades	confirming
kinnitades	reaffirming
kinnitatud	affixed
kinnitava	assurance
kinnitavad	reaffirm
kirikute	churches
kirjalik	written
kirjalike	written
kirjalikke	written
kirjaliku	written
kirjalikult	writing
kirjalikust	written

4.3. Tulemuste hindamine: ARCADE ja PWA

Joondamisprogrammide võrdleva hindamise standardiseerimisprojekti ARCADE (Véronis, Langlais 2000) raames on kindlaks tehtud, et parimate sõnadevahelise paralleelistamise süsteemide täpsus ja saagis ulatuvad umbes 75 protsendini (Véronis, Langlais 2000: 386). Seejuures erineb tulemus sõnaliigiti, ulatudes 94 protsendini adjektiivide puhul ja ainult 60–70 protsendini verbide puhul (samas). Nii hea tulemuse saavutab joondamissüsteem aga üksnes rohke keelespetsiifilise info kasutamisega bitekstide analüüsil. Kuna UWA ja LWA loomisel on järgitud keeltevahelise portatiivsuse eesmärki, siis on nende programmide töötulemuste hindamisel saadav täpsus ja saagis oluliselt väiksem. Projekti ARCADE juhendite järgi arvatud rootsi-inglise näidiskorpusest ekstraheeritud leksikoni täpsus on UWA poolt teostatud joondamise puhul üksnes 42,2% ja saagis 37%; LWA puhul on samad näitajad vastavalt 51% ja 41,3%.

Mitmesõnaliste üksuste korral on saagise arvutamine tunduvalt keerulisem kui sellise vastavuse korral, mille puhul ühele sõnele vastab alati ainult üks sõne. On väga raske hinnata terves korpuses sisalduvate mitmesõnaliste üksuste koguhulka ja seetõttu peaks mitmesõnalisi üksusi kontrolli kaasates teostama arvutuse väga väikese tekstinäite põhjal. ARCADE-projektis esitatud meetodid ei võimalda hinnata seda, kui hästi süsteem jagab biteksti lähtepoole üksused korrektseteks ühe- või mitmesõnalisteks üksusteks, vaid keskenduvad süsteemi poolt pakutud sihtkeele sõnade võrdlemisele etaloniga ehk nn kuldstandardiga (Ahrenberg, Merkel jt 2000: 4). Seetõttu on PWA autorid välja arendanud uue kontrollmeetodi (samas), mis erinevalt ARCADE-st võtab arvesse ka mitmesõnalisteks üksusteks jaotamist nii lähte- kui sihtkeele siseselt ja samuti erinevaid juhtumeid, mille korral mitmesõnaliste üksuste joondamist saab lugeda korrektseks vaid osaliselt. Sama rootsi-inglise näidiskorpuse põhjal UWA süsteemiga genereeritud leksikoni täpsus on selle meetodi kohaselt 71,8% ja saagis 37,4% (LWA-ga genereeritud leksikoni puhul on need näitajad vastavalt 71,9% ja 42,6%).

4.4. Eesti-inglise paralleelkorpuste põhjal genereeritud leksikonide hindamisest

PWA tarkvarapakett sisaldab muuhulgas moodulit joondamistulemuste automaatselt hindamiseks nii ARCADE kui ka PWA meetodil. Mooduli kasutamine eeldab PWA kuldstandardi formaadis võrdlusfaili olemasolu vastava keelepaari jaoks, kus korrektsed vastavused on registreeritud koos mitmesuguse juurdekuuluva infoga ning näitelausetega. Eesti-inglise keelepaari tarbeks selline kuldstandard esialgu puudub. PWA kuldstandardifailide loomiseks on kasutatud samuti tarkvaraplatvormi Uplug kuulunud programmi Plug Link Annotator (PLA), kuid kahjuks ei ühildu PLA enam 2006. aastal kasutatavate operatsioonisüsteemide ja muu vajaliku tarkvaraga (Java) ega ole seetõttu kasutuskõlblik. Seega tuleb standardifaili loomiseks leida mõni alternatiivne poolautomaatne võimalus või kontrollida käsitsi.

Ka PWA hindamise meetodi tulemused varieeruvad üsna palju sõltuvalt tekstide žanrist ja kuldstandardi koostamise kriteeriumitest – s.t sõltuvalt sellest, millise sagedusega ja funktsiooniga sõnu ja millistes proportsioonides valitakse nende näidete hulka, mille alusel paralleelistustulemusi hinnatakse (Ahrenberg jt 2000: 6).

Rootsi-inglise korpusest pärineva kolme erineva tekstikategooria UWA-süsteemi joondamistulemuste hindamine PWA-meetodil andis kõige kõrgema tulemuse (täpsus 81,26%; saagis 64,47%) tehnilise sisuga tekstide puhul ja kõige madalama tulemuse (täpsus 69,04%; saagis 41,44%) poliitiliste tekstide puhul (samas). Vahepealse tulemuse saavutas ilukirjanduse alamkorpus (samas). Võib eeldada, et nende kolme tekstikategooria hulgast asetub TÜ eesti-inglise õigusakte ja seadusi sisaldav paralleelkorpus žanriliselt kõige lähemale poliitikaalastele tekstidele ning oodatav tulemus leksikoni täpsuse ja saagise hindamisel peaks olema võrreldav tulemusega, mille nimetatud katse käigus andis see tekstikategooria.

See oletus pidas tulemuste kontrollimisel vähemalt osaliselt paika: UWA-ga TÜ paralleelkorpuse ja JRC-Acquis' korpuse põhjal genereeritud leksikonist juhuslikult valitud 50 kirje kontrollimisel sain selle alamosa täpsuseks 60%. Samasugune kontroll LWA-ga TÜ eesti-inglise paralleelkorpuse põhjal genereeritud leksikoni 50-kirjelise alamosa peal andis täpsuseks 61%. Saagise hindamine ilma kuldstandardfailita oleks korpuse suure mahu tõttu keerulisem ülesanne ja väikese alamosa põhjal saadud hinnangud ei pruugi olla adekvaatsed, seetõttu saagise arvutamisest esialgu loobusin.

Projekti Plug meetoditele väga sarnast lähenemist leksikoni genereerimisele esindavad Dan Tufiş ja Ana-Maria Barbu (2001), kes kasutasid leksikoni koostamiseks europrojekti Multext-East¹⁵ raames valminud paralleelkorpust, mis sisaldab George Orwelli romaani "1984" kaheksas, sh eesti keeles. D. Tufiş ja A-M. Barbu kirjeldavad muuhulgas ka väikese eesti-inglise sõnastiku ekstraheerimist ning tulemuste kontrollimist. Kontrolliti kuldstandardi puudumisel käsitsi ja eesti-inglise sõnastiku hindamisel saadi selle täpsuseks 96,2% ja saagiseks 57,9%. Paremaid tulemusi võrreldes Plug-projektiga võib seletada mitmesuguse eeltötlusega, mida oli rakendatud MULTEXT-East korpusele ja millele toetus tõlkevastekandidaatide loendi ekstraheerimine: lisaks lausetasandil paralleelilistuse ka leksikaalsete üksuste segmenteerimine, morfoloogiline ühestamine ja lemmatiseerimine.

Sõnatasandil saab paralleelilistada ka statistilise masintõlke rakenduste abil, eespool mainisin tarkvarapaketti Giza++. TÜ inglise-eesti korpuse paralleelilistamise tulemuste kohta Giza++ abil vt K. Muischneki kirjutist (2006).

5. Kokkuvõte

Eesti keelt hõlmavate mitmekeelsete sõnastike koostamine ja täiustamine on üks peamisi teid tulevikus eesti keele väljatõrjumise takistamiseks mitmetest eluvaldkondadest. Seetõttu osutub väga oluliseks siinses artiklis tutvustatud töö jätkamine praktilisel tasandil leksikograafide ja keeletehnoloogide poolt. Kuigi esialgsed tulemused paralleelkorpustest leksikonide ekstraheerimise vallas jätavad veel kõvasti soovida, peaks siiski kõik mitmekeelsete leksikonide koostamisega tegelevad inimesed mõtlema selle peale, kuidas paremini ammutada automaatsete vahenditega leksikoni rikastavat infot nii tava- kui paralleelkorpustest.

Kirjandus

- Ahrenberg, Lars; Andersson, Mikael; Merkel, Magnus 1998. A simple hybrid aligner for generating lexical correspondences in parallel texts. – Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics. 10–14 August 1998. Montréal, Canada, 29–35.
- Ahrenberg, Lars; Andersson, Mikael; Merkel, Magnus 2000. A knowledge-lite approach to word alignment. – J. Véronis (Ed.). *Parallel Text Processing: Alignment and Use of Parallel Corpora*. Dordrecht: Kluwer, 97–116.
- Ahrenberg, Lars; Merkel, Magnus; Sägvall Hein, Anna; Tiedemann, Jörg 2000. Evaluation of word alignment systems. – Proceedings of LREC 2000. Vol. III. Athens/Greece, 1255–1261.
- Blank, Ingeborg 2000. Terminology extraction from parallel technical texts. – J. Véronis (Ed.). *Parallel Text Processing: Alignment and Use of Parallel Corpora*. Dordrecht: Kluwer, 237–252.
- Borin, Lars 2002. ...and never the twain shall meet? – Lars Borin (Ed.). *Parallel Corpora, Parallel Worlds. Language and Computers: Studies in Practical Linguistics nr. 43*. Amsterdam: Rodopi, 47–59.
- Botley, Simon Philip; McEney, Anthony Mark; Wilson, Andrew (Eds.) 2000. *Multilingual Corpora in Teaching and Research*. Amsterdam: Rodopi.
- Bowker, Lynne; Pearson, Jennifer 2002. *Working with Specialized Language: A Practical Guide to Using Corpora*. London/New York: Routledge.
- Brew; Chris; McKelvie, David 1996. Word-pair extraction for lexicography. – K. Oazer, H. Somers (Eds.). *Proceedings of the Second International Conference on New Methods in Language Processing*. Ankara: Bilkent University, 45–55. <http://citeseer.ist.psu.edu/brew96wordpair.html> (21.08.2006).
- Brown, Peter F.; Cocke, John; Della Pietra, Stephen A.; Della Pietra, Vincent J.; Jelinek, Frederick; Lafferty, John D.; Mercer, Robert L.; Roossin, Paul S. 1990. A statistical approach to machine translation. – *Computational Linguistics* 16 (2), 79–85.
- Brown, Peter F.; Lai, Jennifer; Mercer, Robert L. 1991. Aligning sentences in parallel corpora. – Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics. Berkeley, CA, 169–176.
- Brown, Peter F.; Della Pietra, Stephen A.; Della Pietra, Vincent J.; Mercer, Robert L. 1993. The mathematics of statistical machine translation: Parameter estimation. – *Computational Linguistics* 19 (2), 263–311.
- Brown, Ralf D. 1997. Automated Dictionary Extraction for –Knowledge-Free – Example-Based Translation. – Proceedings of the Seventh International Conference on Theoretical and Methodological Issues in Machine Translation (TMI-97). <http://citeseer.ist.psu.edu/brown97automated.html> (21.08.2006).
- Church, Kenneth W.; Gale, William A.; Hanks, Patrick; Hindle, Donald 1991. Using statistics in lexical analysis. – Uri Zernik (Ed.). *Exploiting On-Line Resources to Build a Lexicon*. Hillsdale, NJ: Lawrence Erlbaum Associates, 115–164.
- Déjean, Hervé; Gaussier, Eric; Sadat, Fatia 2002. Bilingual terminology extraction: An approach based on a multilingual thesaurus applicable to comparable corpora. – Proceedings of COLING. Tapei, Taiwan. <http://www.xrce.xerox.com/Publications/Attachments/2002-025/dejean.pdf> (21.08.2006).
- Davis, Mark; Dunning, Ted; Ogden, Bill 1995. Aligning noisy corpora. – Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics. University College Dublin. Belfield, Dublin, 67–74.
- Davis, Mark 1998. On the effective use of large parallel corpora in crosslanguage text retrieval. – G. Grefenstette (Ed.). *Cross-Language Information Retrieval*. Boston/Dordrecht/London: Kluwer Academic Publishers, 11–22.
- Ebeling, Jarle 1998. Contrastive linguistics, translation, and parallel corpora. – *Meta* 43 (4), 602–615.

- van der Eijk, Pim 1993. Automating the Acquisition of Bilingual Terminology. – Proceedings of the 6th Conference of the European Chapter of the Association for Computational Linguistics. Utrecht, 113–119.
- Fung, Pascale; Church, Kenneth W. 1994. K-vec: A New Approach for Aligning Parallel Texts. – Proceedings of the 15th International Conference on Computational Linguistics. Kyoto, Japan, 1096–1102.
- Fung, Pascale 2000. A statistical view on bilingual lexicon extraction – From Parallel Corpora to non-parallel corpora. – J. Véronis (Ed.). *Parallel Text Processing: Alignment and Use of Parallel Corpora*. Dordrecht: Kluwer, 219–236. <http://citeseer.ist.psu.edu/fung98statistical.html> (21.08.2006).
- Gale, William; Church, Kenneth 1993. A program for aligning sentences in bilingual corpora. – *Computational Linguistics* 19 (1), 75–102.
- Gaussier, E.; Hull, D.; Ait-Mokhtar, S. 2000. Term alignment in use: Machine-aided human translation. – J. Véronis (Ed.). *Parallel Text Processing: Alignment and Use of Parallel Corpora*. Dordrecht: Kluwer, 253–274.
- Geisler, Christer 2002. Reversing a Swedish-English dictionary for the Internet. – Lars Borin (Ed.). *Parallel Corpora, Parallel Worlds. Language and Computers: Studies in Practical Linguistics* nr. 43. Amsterdam: Rodopi, 122–133.
- Genereeritud leksikonid. www.teataja.ee/leksikonid.zip (29.08.2006).
- GIZA++. Training of statistical translation models. <http://www.fjoch.com/GIZA++.html> (21.08.2006).
- Glozzo, Alfio; Strapparava, Carlo 2005. Cross language text categorization by acquiring multilingual domain models from comparable corpora. – Proceedings of the ACL Workshop on Building and Using Parallel Texts (in conjunction of ACL-05). University of Michigan, Ann Arbor, 9–16.
- Hwa, Rebecca; Madnani, Nitin 2004. The UMIACS Word Alignment Interface. <http://www.umiacs.umd.edu/~nmadnani/alignment/forclip.htm> (21.08.2006).
- Hiemstra, Djoerd 1997. Deriving a Bilingual Lexicon for Cross-Language Information Retrieval. – M. Heemskerk, M. Diepenhorst (Eds.). Proceedings of the 4th Groningen International Information Technology Conference for Students. University of Groningen, 21–26.
- Ide, Nancy; Erjavec, Tomaz; Tufiş, Dan 2002. Sense discrimination with parallel corpora. – Proceedings of the SIGLEX Workshop on Word Sense Disambiguation: Recent Successes and Future Directions. ACL2002. Philadelphia, 56–60. http://www.racai.ro/~tufis/Selected_Papers/sense-discrimination.pdf (21.08.2006).
- Index of /telri/Vanilla. <http://nl.ijs.si/telri/Vanilla/> (21.08.2006).
- Index of JRC-Acquis/alignments. <http://wt.jrc.it/lt/Acquis/JRC-Acquis.2.2/alignments/index.html> (21.08.2006).
- Inglise-eesti-inglise sõnastik. <http://www.eki.ee/dict/inglise/> (21.08.2006).
- Inglise-eesti ja eesti-inglise paralleelkorpus. <http://test.cl.ut.ee/korpused/paralleel/> (21.08.2006).
- Kitsnik, Mare 2006. Keelekorpused ja võrkeeleõpe. – Helle Metslang, Margit Langemets (toim.), Maria-Maren Sepper (keeletoim.). *Eesti Rakenduslingvistika Ühingu aasta- raamat 2. Estonian Papers in Applied Linguistics 2. Eesti Rakenduslingvistika Ühing*. Tallinn: Eesti Keele Sihtasutus, 93–107.
- Kuhn, Jonas 2004. Experiments in parallel-text based grammar induction. – Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics. ACL, 470–477.
- Linear B. Word alignment tool. <http://demo.linearb.co.uk:8080/sandbox/start.jsp> (21.08.2006).
- Melamed, Dan 1997. A word-to-word model of translational equivalence. – Proceedings of the Thirty-Fifth Annual Meeting of the Association for Computational Linguistics

- and 8th Conference of the European Chapter of the Association for Computational Linguistics. Madrid, Spain, 490–497.
- Melamed, Dan 2001. *Empirical Methods for Exploiting Parallel Texts*. Cambridge, Massachusetts: MIT Press.
- Melby, Alan K. 2000. Sharing of translation memory databases derived from aligned parallel text. – J. Véronis (Ed.). *Parallel Text Processing: Alignment and Use of Parallel Corpora*. Dordrecht: Kluwer, 347–368.
- Multext-East home page. <http://nl.ijs.si/ME/> (21.08.2006).
- Muischnek, Kadri; Orav, Heili; Kaalep, Heiki-Jaan; Õim, Haldur 2003. Eesti keele tehnoloogilised ressursid ja vahendid. Arvutikorpused, arvutisõnastikud, keeletehnoloogiline tarkvara. Haridus- ja Teadusministeerium, Eesti keelenõukogu. Tallinn: Eesti Keele Sihtasutus.
- Muischnek, Kadri 2006. Improving the quality of the statistical machine translation with linguistic preprocessing: The case of particle verbs in Estonian. – http://www.id.cbs.dk/~dh/ngslt/projects/muischnek_final_paper.doc (21.08.2006).
- Oakes, M. P.; McEnery, A. M. 1998. Bilingual text alignment: An overview. – A. M. McEnery, S. P. Botley, A. Wilson (Eds.). *Multilingual Corpora in Teaching and Research*. Amsterdam: Rodopi, 1–37.
- Och, Franz Josef; Ney, Hermann 2000. Improved statistical alignment models. – Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics. Hongkong, China, 440–447.
- Sågvall Hein, Anna 2002. The PLUG-project. Parallel corpora in Linköping, Uppsala, Göteborg: Aims and achievements. – Lars Borin (Ed.). *Parallel Corpora, Parallel Worlds. Language and Computers: Studies in Practical Linguistics nr. 43*. Amsterdam: Rodopi, 61–78.
- Smadja, Frank; McKeown, Kathleen R.; Hatzivassiloglou, Vasileios 1996. Translating collocations for bilingual lexicons: A statistical approach. – *Computational Linguistics* 22 (1), 1–38.
- Software for word alignment. <http://www.cse.unt.edu/~rada/wa/#softwareWA> (21.08.2006).
- The JRC-Acquis multilingual parallel corpus. <http://langtech.jrc.it/JRC-Acquis.html> (21.08.2006).
- The PLUG Word Aligner – PWA. <http://stp.ling.uu.se/plug/pwa/index.html> (21.08.2006).
- Tiedemann, Jörg 2002. Upplug – a modular corpus tool for parallel corpora. – Lars Borin (Ed.). *Parallel Corpora, Parallel Worlds. Language and Computers: Studies in Practical Linguistics nr. 43*. Amsterdam: Rodopi, 181–197.
- Tiedemann, Jörg 2003. Recycling Translations. Extraction of Lexical Data from Parallel Corpora and Their Application in Natural Language Processing. – http://stp.ling.uu.se/~joerg/phd/html/thesis_html.html (21.08.2006).
- Tufiş, Dan; Barbu, Ana-Maria 2001. Automatic Construction of Translation Lexicons. – V. V. Kluev, C. E. D'Attellis, N. E. Mastorakis (Eds.). *Advances in Automation, Multimedia and Modern Computer Science. A Series of Reference Books and Textbooks in Electrical and Computer Engineering*. WSEAS Press, 156–172.
- Véronis, Jean; Langlais, Philippe 2000. Evaluation of parallel text alignment systems: The ARCADE project. – J. Véronis (Ed.). *Parallel Text Processing: Alignment and Use of Parallel Corpora*. Dordrecht: Kluwer, 369–388.

Kaarel Veskis töötab Tartu Ülikooli üldkeeleteaduse õppetooli juures, kus põhiülesandeks on eesti keele segakorpuse koostamine.
kaarel.veskis@ut.ee

GENERATION OF BILINGUAL LEXICONS FROM A PARALLEL CORPUS

Kaarel Veskis

University of Tartu

In addition to contrastive studies of languages or language variants, parallel/comparative corpora have many other uses both in theory and practice, while the potential of some of such uses is still awaiting discovery. One of the most interesting trends involves dictionary compilation or revision by means of extracting translation equivalents. The article attempts a survey of what has been done, with a view to some possible practical applications to Estonian in the future. For example, a simple lexicographic device has been outlined to enable the lexicographer to generate a list of translation equivalents by using a parallel text. Also, there have been reports of attempts to generate source material for an English-Estonian Estonian-English technical dictionary using not only the parallel corpus but also some free software, which needs little additional language resources beside the corpus material.

For the time being multilingual technical dictionaries could be compiled from parallel corpora only semiautomatically, because without intervention on the part of a human proofreader the method would yield but raw material to help lexicographers, terminologists or translation systems.

There is no software that could perform paralleling and dictionary generation on the basis of the Estonian grammatical structure and its possible points of equivalence with the structure of some other language. Although the quality of the lexicon to be generated would certainly be improved by preliminary morphological analysis of the parallel corpus, our present attention has been focused on language independent approaches to dictionary extraction. The word aligners UWA and LWA developed by Swedish researchers within the *Plug* project (Tiedemann 2002) use relatively little language-specific information, which makes them easily applicable in automatic generation of dictionaries containing Estonian material. The article describes an attempt to develop source material for a technical dictionary by means of UWA and LWA, drawing on the English-Estonian parallel corpus of the University of Tartu and the English-Estonian subsection of the JRC-Acquis multilingual parallel corpus. One of the dictionaries so generated contains 130 865 headwords and 482 571 word forms. The precision of a random sample of 50 entries turned out to be 60%. In addition the article provides a survey of the working principles of the used programmes, and some suggestions on how to improve UWA results with a view to an analogous device to be possibly developed for Estonian.

Keywords: corpus linguistics, automatically created dictionaries, bilingual lexicography, language processing, technical dictionaries, Estonian, English

PRONOOMENID LAPSEKEELES: MÕNDA MINA JA SINA OMANDAMISEST

Maigi Vija

Ülevaade. Artiklis vaadeldakse ainsuse 1. ja 2. isiku pronoomeni omandamist lastel vanuses 1;7–3;1 CHILDES-i eesti keele allkorpuse ning teistest allikatest leitud toetava keelematerjali põhjal. Põhianaalüüs on tehtud Andrease allkorpuse 76 tunnipikkuse lindistuse põhjal. Lisaks lapse kõnele on analüüsitud ka sisendkeelt. Analüüsitud materjali põhjal võib väita, et demonstratiivpronoomenid ilmuvad enne personaalpronoomeneid. Personaalpronoomeneid hakkab laps kasutama umbes kaheaastaselt. Esialgu leidub üksikuid juhuslikke näiteid, pronoomenite täiskasvanupärane kasutus kujuneb välja alles mõne aja pärast. Sisendkeele analüüs näitab, et ema hakkab lapse kasvades kasutama täiskasvanupärasemat keelt. Kui vestluses kaheaastasega kasutab ema valdavalt lapse nime ja endast rääkides sõna *emme*, siis kolmeaastasega vesteldes viitab ema endale ja lapsele pigem pronoomeniga. Ainsuse 1. isiku pikka ja lühikest pronoomenit *mina/ma* (subjektina) esineb Andreasel peaaegu kolm tuhat korda, ning ainsuse 2. isiku pronoomeneid kolm korda vähem. Sisendkeele võrreldes on Andreasel 1. ja 2. isiku pronoomeni pikki vorme rohkem kui emal. Pronoomenite varases kasutuses teevad lapsed ka vigu, sagedamini ajavad nad segamini 1. ja 2. isiku pronoomeni.*

Võtmesõnad: keeleomandamine, lastekeel, sisendkeel, asesõnad, eesti keel

Sissejuhatuseks

Pronoomenid on suletud sõnaliik, sinna kuuluvad käändes ja arvus muutuvad sõnad, mis talitlevad lauses substantiivide, adjektiivide ja numeraalide sarnaselt, kuid on leksikaalselt sisuvaesed, abstraktse tähendusega, mis võib konsituatsiooni kaudu konkretiseeruda (*mina* kui kõneleja, *sina* kui kuulaja jne). Pronoomenid viitavad lauses mingitele kontekstist nähtuvatele asjadele või tunnustele. Pronoomeneid

* Artikkel on valminud ETF-i grandil nr 5813 toel.

on jaotatud kaheksasse tähendusrühma: personaal-, demonstratiiv-, retsiprook-, refleksiiv-, possessiiv-, determinatiiv-, indefiniit- ja interrogatiiv-relatiivpronoomeid (EKG I: 26–31). Uuema käsitluse kohaselt puuduvad eesti keeles possessiiv- ja determinatiivpronoomenid, seega on pronoomenirühmi kuus (Erelt 2000: 16).

Laste väga varasele pronoomenite kasutusele viitavad paljude maailma keelte uurimistulemused (Bloom 2000: 121). Eve Clarki järgi on sõnavara omandamisel olulised kaks pragmaatilist põhimõtet, mis on keele funktsioneerimise aluseks: konventsionaalsus ja kontrast. Kontrast pole väga kergesti lahutatav konventsionaalsusest, sest kui lapsed võtavad sihtkeelendiks konventsionaalsed vormid sisendkeelest, järgivad nad ka kontrastiprintsiipi, mida võib seletada nii, et iga erinevus vormis tähendab erinevust tähenduses (Clark 1993: 84). Lapse varast keelelist arengut analüüsid tuleb tähelepanu pöörata sellele, mida laps tegelikult kuuleb enda ümber räägitavat (Tomasello 2000: 11).

Materjal

CHILDES-i¹ andmebaas sisaldab 10 eesti lapse (vt tabel 1) spontaanse kõne lindistusi vanusest 0;11.22–3;1.12. Stella, Heleni, Sandori, Taimo, Carlose, Annabeli, Henri ja Mariliisi korpuse on koostanud Kaja Kohler, Hendriku korpuse Reili Argus ja Andrease korpuse Maigi Vija. Tabelist 1 on näha, et lindistuste arv on lapseti väga erinev. Siin esitatakse kvantitatiivne analüüs põhineb Andrease allkorpuse materjalil, kuna seda on kõige rohkem (76 lindistust kogupikkusega 76 tundi), lindistusi on tehtud järjepidevalt ja need katavad peaaegu poolteist aastat väikelapse keelelisest arengust. Kõik 10 last pole vaatlusperioodil veel jõudnud personaalpronoomenite kasutamiseni, seetõttu puudub võimalus neid teistega võrrelda. Võrdlusmaterjalina olen kasutanud Carlose, Mariliisi, Madli (Kõrgvee 2001), Eriku (Salo 1995) ning lisaks ka Kadri Videri (1995) diplomitööst pärinevaid keeleandmeid. Lastekeele tekstide transkribeerimiseks ja analüüsimiseks mõeldud programmiga CLAN on failidest üles otsitud vajalikud pronoomenivormid ning analüüsitud nende sagedust nii lapse kõnes kui ka sisendkeeles.

Tabel 1. Eesti keele materjal CHILDES-i andmebaasis

Lapse nimi	Koguja/ koostaja	Lapse vanus	Lindistuste arv	Mitmes laps peres
Stella	Kaja Kohler	0;11.22–1;6.4	9	2.
Helen	Kaja Kohler	1;1.17–1;10.17	7	1.
Sandor	Kaja Kohler	1;2–2;2.22	10	1.
Taimo	Kaja Kohler	1;5.8–1;11.13	9	1.
Carlos	Kaja Kohler	1;7.17–1;10.29	9	2.
Annabel	Kaja Kohler	1;10.28–2;1	7	2.
Henri	Kaja Kohler	2;2.12–2;3.8	3	2.
Mariliis	Kaja Kohler	2;5.7–2;8.10	7	3.
Hendrik	Reili Argus	1;8.13–2;5.30	17	2.
Andreas	Maigi Vija	1;7.24–3;1.12	76	1.
Kokku		0;11.22–3;1.12	154	

¹ CHILDES (Child Language Data Exchange System) <http://childes.psy.cmu.edu/>, <http://childes.psy.cmu.edu/data/other/> on lapsekeele andmebaas, mille eesmärgiks on talletada ja uurida laste keeleomandamist erinevate maailma keelte näitel. Andmebaas koosneb lindistatud suhtlussituatsioonide transkriptsioonidest ja hõlmab 26 keele näiteid.

² Lapse vanus aastates; kuudes. päevades.

Demonstratiivpronoomenid ilmuvad enne personaalpronoomeneid

CHILDES-i andmebaasi analüüs näitab, et eesti lapsed kasutavad enne teist sünnipäeva demonstratiivpronoomenivorme *see*, *selle* ja *seda* (vt Pajusalu jt 2004: 113), kuid teiste pronoomenirühmade esindajad, sh personaalpronoomenid ilmuvad üldjuhul laste kõnesse hiljem. Kadri Videri leksikostatistiline analüüs osutab, et laste kõnes eristuvad pronoomenite seast esinemissageduse poolest tugevalt personaal- ja demonstratiivpronoomenid (Vider 1996: 128).

Esimeste miniparadigmade moodustajaks on lastel tihti lekseem *see* (Argus 2004). Holger Diesseli (1999: 2) järgi on demonstratiivid deiktilised väljendid, millel on spetsiifilised süntaktilised ja pragmaatilised funktsioonid ning semantilised omadused. Demonstratiividega on lapsel võimalik leksikonis puuduvaid sõnu asendades osutada enda läheduses olevatele asjadele/olenditele, nõuda läheduses olevaid asju, tõmmata juuresolija tähelepanu enesele jne. Andrease kõnes produktiivseim pronoomen *see* (esines korpuses 2504 korda) on sagedane ka sisendkeeles³, esinedes Andrease ema kõnes 1660 korda (Vija 2005). Pronoomeni *see* sagedast esinemist tõendavad ka TÜ suulise kõne korpuse andmed (Pajusalu jt 2004: 81). Olgu lisatud, et artiklis vaadeldud pronoomenid on ka “Eesti kirjakeele sagedussõnastiku” (Kaalap, Muischnek 2002) sagedusloendi tipus: *see* 4., *mina* 5. ja *sina* 14. kohal. Seega on üks oluline sõnavara omandamise kriteerium ka produktiivsus, mis tegelikult peegeldab kollektiivseid konventsionaalseid eelistusi (Clark 1993: 135).

Kategooriad ‘mina’ ja ‘sina’ lapsel ja emal

Eri keeli omandavatel lastel on väga sarnane varajane leksikon ning olemasolevad sõnad (sh ka deiktikud) jagunevad kindlatesse sarnastesse semantilistesse domeenidesse (Clark 1993: 29). Isikutele viitamiseks kasutab väike laps algul üldnime (*issi*, *emme*, *venna*) või pärisnime (*Mati*, *Pitsu*), millele mitmesõnaperioodil lisandub tunnusetu verbitüvi (*Antsu loe*, *emme maga*). Sõnavara kasvades loovad lapsed lisadomeene ning varasem deiktikute domeen jaguneb edasi alldomeenideks (subjekt, objekt, possessiivpronoomen, koht, aeg jne) (Clark 1993: 30–31).

Andrease näitel tulevad personaalpronoomenid üldiselt enne possessiivsust väljendavana genitiivis või siis genitiiv ja nominatiiv samal ajal koos, alles hiljem ilmub pronoomen koos verbiga. Personaalpronoomenite omandamisel võib märgata individuaalset omapära, kõikidel lastel ei ilmu need samas järjekorras ning samas vanuses, samuti erineb personaalpronoomenite mittetäiskasvanupärane kasutus lapseti (Dabrowska 2004: 163; Oshima-Takane jt 1999; Vider 1995). Andrease keeleandmete põhjal ei saa öelda, kas enne ilmub pronoomen *mina* või *sina*. Mõlemad ilmuvad üksikjuhtudena enne teist sünnipäeva, esinevad äärmiselt hajusalt ning mõlemaga on moodustatud mittekorrektseid kombinatsioone.

Väikeseid lapsi iseloomustab enesekesksus, seetõttu on tavaliselt esimeste personaalpronoomenite hulgas *mina*. Enne kaheaastaseks saamist esineb Andreasel üks kord pronoomenivorm *minu* (1;9.16 *tupsu kätte*, *minu kätte*). Kahel korral kasutab laps pronoomenivormi *mina*, ühel korral koos pöördelõputa verbiga *viska* (1;8.25 *mina mina viska*) ja teisel korral ühesõnalise enesele viitamisenä (1;10 *mina*).

³ Sisendkeele mõistet on artiklis kasutatud kitsamas tähenduses: see on lapse ema lausunud sõnaline info, sest igapäevases suhtluses ning lüüdistusituatsioonides on lapse peamiseks vestluspartneriks ema.

Ainsuse 1. isiku pronoomenite kasutus on enne kaheaastaseks saamist väga harv ja juhuslik, rohkem näiteid sellest vanusest pole, seega ei ole selles vanuses laps veel küps opereerima eri personaalpronoomenitega.

Personaalpronoomenite puudumisel viitavad lapsed endale pärisnimega. Mariliisi kohta on esimesed korpuse näited vanusest 2;6, mil laps räägib enesest pärisnime kasutades: *Mari karjub, Mari ei taha, Mari oskab torni teha*. Vanusest 2;6–2;7 leiduvad tal üksikud personaalpronoomenite kasutused (*siin on minu venna, ma võtan ka kommi*), kuid ei midagi rohkemat. Ka Andreas kasutab endast rääkides pärisnime koos ainsuse 3. pöördega (*Antsu⁴ võtab*), eriti produktiivselt vanuses 1;11–2;1. Madli (Kõrgvee 2001) andmete varal võib aga väita, et mitte kõik kaheaastased lapsed ei räägi endast pärisnime kasutades. Madlil on vanuses 2;0 suhteliselt väike sõnavara ja personaalpronoomenid puuduvad. Endale viitab ta pronoomeniga *ise*, mis esines üksiksõnana enese tegutsemise kohta, aga ka kombinatsioonides koos verbiga: *ise allo; ise näha; ise hoia kinni; ise katsu*. Samal ajal asendas *ise* ka possessiivpronoomenivormi *minu*: *ise poti moodi* 'minu poti moodi'. Eelnevates näidetes verbiga koos esinedes rõhutab *ise* last kui tegijat.

Age Salo (1995: 21) andmetel kuulub väike laps sisendkeeles palju oma nime ning kasutab seda iseendast rääkides. Pärisnime ja üldnime kasutamine pronoomenite asemel lapsega suheldes on üldlevinud lapsevanemate keelekasutuses paljudes keeltes, nt itaalia, kreeka, korea, läti, soome, inglise keeles (Bloom 2000: 122). Ema eesmärgiks on teha end väikesele lapsele kergemini mõistetavaks; lapse kasvades muutub ema oma keelelistes väljendustes järk-järgult keerulisemaks, s.t ta kasutab eri vanuses lapsega rääkides erinevat strateegiat. Kui vestluses kaheaastase lapsega kasutab ema väga palju lapse nime (tabeli 2 põhjal kokku 509 korda 30 lindistuses) ning endast rääkides sõna *emme* (*emme läheb Andreasega, tädi annab Andreasele, Carlos sõidab*), siis kolmeaastasega vesteldes viitab ema endale ja lapsele pigem pronoomeniga (*ma annan sulle*). Enne lapse kaheaastaseks saamist kasutas Andrease ema analüüsitud materjalis ainult 4 korda vormi *sina* ja 1 kord vormi *sind*. Tabelist 3 selgub, et suhtluses kaheaastase lapsega on ema kasutanud 42 korda vormi *sina* ja 18 korda vormi *sind* ning kolmeaastase lapsega rääkides samu sõnavorme vastavalt 126 ja 34 korda. Ajavahemikus 2;0–2;1.12 leidis ema kõnes 41 *mina*, 437 *ma*, 37 *mul* ja 32 *mulle* kasutust (tabel 2), aasta hiljem on samade vormide kasutus hüppeliselt kasvanud, fikseeritud vastavalt 149 *mina*, 888 *ma*, 89 *mul* ja 97 *mulle* esinemist. Samal ajal vähenevad *emme* kasutused emal 268-lt 41-le.

Tabel 2. *Mina/ma, emme⁵ ja Andreas⁶ kasutused Andrease ema kõnes*

Periood (lindistuste arv)	emme	Andreas	mina	ma	mul	mulle
2;0–2;1.12 (30)	268	509	41	437	37	32
3;0–3;1.13 (30)	41	15	149	888	89	97

Tabel 3. *Sina/sa kasutused Andrease ema kõnes*

Periood (lindistuste arv)	sina	sinu	sind	sul	sulle
2;0–2;1.12 (30)	42	34	18	155	118
3;0–3;1.13 (30)	126	65	34	223	131

⁴ Andrease hellitusnimi.

⁵ Tüvevariant, mida on kasutatud programmiga CLAN tehtud otsingus.

⁶ Tüvevariant, mis hõlmab pärisnime kasutusi kõikides käänetes (*Andreas, Andrease, Andreast* jne).

Michael Tomasello (2003: 202) väitel on ainsuse 1. isiku pronoomenid paljude inglise keelt omandavate laste esimeste sõnade hulgas; ta kirjeldab oma tütre Travise keelematerjali põhjal, kuidas tütar kasutas 23-kuuselt täiesti korrektselt ja personaalpronoomeneist kõige produktiivsemalt just 1. isiku pronoomenivorme *I* ja *me* (Tomasello 1992: 158). M. Tomasello (2003: 90–92) peab sotsiopragmaatilises lähenemises keeleomandamise aluseks sotsiokognitiivseid võimeid ning keeleomandamise protsessi hõlbustavateks faktoriteks konteksti ja kontrasti. Kontrast kui pragmaatiline printsiip aitab lastel sõnade tähendusi täpsemalt kindlaks määrata ja laste keekekasutus annab tunnistust kontrasti ning konventsionaalsuse väga varajasest taipamisest (Clark 1993: 92, 100). Sedasama osutab ka Travise 1. isiku pronoomenite kasutus.

Eesti laste näitel võib väita, et personaalpronoomenid ei ilmu väga vara. Selleks, et õppida sõnu õigesti kasutama, peab laps kõigepealt sõna ülejäänud kõnevoolust eristama ning salvestama iga sõna jaoks mällu mitut tüüpi informatsiooni: sõna tähendus, liik, morfoloogiline struktuur, hääldus jne (Clark 1993: 2). Pärast teist sünnipäeva (2;0–2;1.12) lisandub Andrease kõnesse mitmeid uusi pronoomenivorme (*ma*, *mind*, *mu* ja *mul*), mis on alguses juhuslikud ning sageli ka mittetäiskasvanupärased. Selliste vormide ilmumine tõendab, et laps on täiskasvanute kõnevoolust need sõnad eristanud, kuid pole veel selgeks saanud täpset tähendust ja kasutust. Teise sünnipäeva paiku esineb Andrease kõnes palju imitatsioone, lastelaulude ja juttude deklameerimist ning endamoodustatud tarindeid. Personaalpronoomenite hajusat esinemist kinnitab ka objektivormis pronoomeni *sind* kasutus, mille kohta vanusest 2;3–2;4 võib leida üksikuid näiteid (*näen sind*, *Antsu kuulab ka sind*) ning alles pärast kolmandat sünnipäeva tuleb ainsuse teise isiku objektivorm *sind* produktiivsemalt kasutusse.

K. Videri (1995: 57) andmetel kasutasid vaatlusalused eesti lapsed personaalpronoomeneist kõige varem nominatiive *mina* ja *ma*, seejärel tuli genitiiv *minu*. Paul Bloomi esitatud hüpoteesi järgi omandavad lapsed enne need pronoomenid, mis viitavad neile endile (Bloom 2000: 121–126). Soome keele andmete analüüsitulemused (Brown 2004: 506) ja eesti laste keelenäited toetavad seda seisukohta, sest *mina* ja *sina* on lastel varasemad (ning lisaks ka produktiivsemad) kui *tema*, *teie* või *nemad*.

Lühike ja pikk pronoomen

Eesti keele personaalpronoomenitele on iseloomulik pikkade ja lühikeste paraleelkujude olemasolu. Pikk pronoomen on eesti keeles üldjuhul rõhuline ja lühike pronoomen rõhuta. Pronoomenikuju valikul on olulisteks teguriteks sõnavormi kääne ja esiletõstetus, kusjuures tihti on pika pronoomenikujuga viidatud isik kellegi teisele vastandatud ja eriti fokuseeritud (Pool 1999: 179). Andreasel on 1. ja 2. isiku pronoomeni pikki vorme rohkem kui emal. Selle põhjuseks võib olla asjaolu, et varajasel keeleomandamise faasis on kahesilbilised pronoomenid *mina* ja *sina* väikelapse jaoks sisendkeelest lihtsamalt tuvastatavad kui ühesilbilised *ma* ja *sa*, mis sageli häälduvad järgneva komponendiga (verbiga) kokku. Lühikest vormi *sa* esineb Andreasel tegelikult perioodi 2;0–2;1.12 lindistustes 158 korda, aga need kõik esinevadki küsimuste sees nn kokkukasvanud ühendites *missa* ja *kussa*, mida

laps ilmselt ei taju 2. isiku kohta käiva pronoomenina (*kus sa lähed? mis sa teed? mis sa sööd?* jne).

Ainsuse 1. isiku produktiivset kasutust tõendab see, et 1. isikule viitamisi (nominatiivis) on Andreasel peaaegu kolm tuhat (vt tabel 4). Pikkade ja lühikeste pronoomenivariantide suhet võrreldes ilmneb, et lühikest pronoomenit *ma* on emal kõikidest *mina* ja *ma* kasutustest 87% ja Andreasel 71%. Ainsuse 1. isiku lühikest pronoomenit *ma* esines Andreasel lindistustes kokku 2113 korda ja emal 1420 korda, kusjuures pikka kuju *mina* kasutatati palju vähem (Andreasel 856 ja emal 207 korda).

Tabel 4. Ainsuse 1. isiku pronoomenid ema ja Andrease kõnes 76 lindistuse põhjal

	<i>mina</i> ja <i>ma</i> kokku	<i>mina</i> (% kõigist)	<i>ma</i> (% kõigist)
EMA	1627	207 / 13%	1420 / 87%
ANDREAS	2969	856 / 29%	2113 / 71%

Andrease korpuses leidub paar näidet selle kohta, kuidas laps on juba taibanud ainsuse 1. isiku pika ja lühikese pronoomeni kontrasti. Allolevad näited (1a), (1b) on vanusest 2;4 ja mõlemal korral püüab laps vestluspartneri tähelepanu endale tõmmata, kasutades alguses pronoomeni lühikest ja seejärel pikka varianti. Näites (1b) on näha, kuidas laps teatab isale midagi ja seejärel kordab oma teadet, formuleerides sellesama info ümber nii, et sõnajärg muutub mittetraditsiooniliseks. Siinkohal tasub lisada, et sisendkeelest ei leia ühtegi näidet sarnase lausungiskeemi (*X on minul*) kohta, leidub üksainus näide ema kõnest tavalise sõnajärjega skeemi *minul on X* kohta: *minul on ka armas loom süles*. Selles ema näites rõhutab pikk pronoomenivorm ilmselgelt omajat, lisarõhu annab veel partikkel *ka*.

Näidet (1b) analüüsides on raske tuvastada, missugused sisendkeele lausungid on siin lapsele eeskujuks või milliseid operatsioone sooritades jõuab laps sellise lausungini. Anna Siewierska (2004: 183) väitel põhjendab isikuvormi valikut referentide infostaatust ning isikumarkerite kontrast sisaldab tihti konkurentsi ja ka teema muutumist. Kas näites (1b) mängib lapse jaoks rolli lause infostruktuur? Võib-olla kombineerib ta lihtsalt keeleüksusi loovalt, tekitades niimoodi traditsioonilises mõttes vigaseid lauseid ning keelekasutust, mida selles vanuses leidub rohkemgi.

(1a) CHI: **mul** on sinine käär [= käärid], **minul** on sinine käär [= käärid].⁷

(1b) CHI: **mul** on kepp, kepp on **minul**, issi.

Ainsuse 2. isiku pikka pronoomenit *sina* (vt tabel 5) esineb ema kõnes 186 ja Andreasel 279 korda ja lühikest pronoomenit *sa* emal 738 ning Andreasel 1647 korda. Ainsuse esimese isiku pronoomeneid oli lapsel kaks korda rohkem kui emal. Ainsuse 2. isikule viitamisi on emal peaaegu sama palju kui 1. isikule, kuid Andreasel on ainsuse 2. isiku pronoomeneid kolm korda vähem kui 1. isiku omi. Seega on lapse jaoks oluline rääkida endast ja rõhutada asjade temale kuulumist. Ema kõnes on aga tähtsad 2. isiku vormis pöördumised lapse poole, s.t käsud, keelud, küsimused ja asjade seletamine lapsele läbi *sina/sa* vormi.

⁷ Näited on vormistatud programmi CLAN transkriptsioonireeglitele vastavalt, suurtäheline kood CHI tähistab last ja MOT ema, nurksulgudes on lisatud kommentaarid või sõna tegelik tähendus.

Tabel 5. Ainsuse 2. isiku pronoomenid ema ja lapse kõnes 76 lindistuse põhjal

	<i>sina ja sa kokku</i>	<i>sina</i> (% kõigist)	<i>sa</i> (% kõigist)
EMA	1833	186/ 10%	1647/ 90%
ANDREAS	1017	279/ 27%	738/ 73%

Deiksise probleem lapse jaoks ja vead

Deiktiline nullpunkt on alati seotud kõnelejaga ning isikudeiksise põhikategooriad on *mina* ja *sina* ehk kõneleja ja vastuvõtja, mida nimetatakse primaarseks rollideiksiseks (Lyons 1977: 638, Larjavaara 1990: 43). Kuna deiktiline nullpunkt on alati seotud kõnelejaga, lähtub nii kasutus kui interpretatsioon alati kõnelejust (Lyons 1977: 638). Nagu eespool esitatud analüüsist selgus, on primaarse rollideiksise vormide (*mina* ja *sina*) arv nii sisendkeeles kui ka lapse kõnes mitmeid kordi suurem kui teiste personaalpronoomenite oma (vt täpsemalt Vija 2005).

Vanuses 1;10 tähistab Andreas possessiivsust märkiva ainsuse 2. isiku pronoomeniga *sina* hoopis iseennast, s.t ainsuse esimest isikut, rääkides köögis laua ääres enda tassist: *sinu joogitass*. Selline kasutus on väikestel lastel sagedane, kuna semantiliselt keerulise ainsuse teise isiku pronoomeni referent vaheldub vestlutsituatsioonis pidevalt ning sama rääkija võib selle abil viidata eri kõnetatutele (Brown 2004: 502). M. Tomasello (2003: 203) seletab niisuguseid keeleleide sellega, et lapsel on kalduvus varem kuulnud imiteerida, s.t ta reprodutseerib analüüsimate keeleteervikuid, mida on kuulnud täiskasvanutelt. Sellest tulenevad ka nn ümberpööramisvead (*reversal errors*) personaalpronoomenite kasutuses (vt ka Dabrowska 2004: 163). Teises sarnases näites väljendab laps jällegi ainsuse 2. isiku abil iseenda tegevust (*sina kiigud* [*] [= *mina kiigun*]). Samasugused ümberpööramisvead leiduvad A. Salo (1995: 23) magistritöös Erikult, kes samuti kasutab teatud perioodil teist pööret esimese asemel. Andreas pere esimese lapsena kasutas teise sünnipäeva paiku paralleelselt pronoomeneid *minu* ja *sinu*, sest ta ilmselt ei mõistnud, et *sina* võib referenti vahetada. Keelepäeviku andmetel lahendab Andreas vanuses 2;1 ühe keerulise olukorra pärisnime kasutades, mis ei jäta valesi tõlgendamise võimalust (vt näide 2).

(2) [Laps sööb oma taldrikult suppi]

CHI: see on *minu* oma.

MOT: *jaa*, see on *sinu* oma.

CHI: *sinu* oma? (% umbusklikult)

MOT: *jaa*, see on *sinu* (% rõhutatult) oma.

CHI: see on *Antsu* oma.

Huvitav konstruktsioon esineb näites (3), kus laps ilmselt oma nõudmise rõhutamise pärast markeerib topelt 1. isiku pronoomenit, kõigepealt possessiivsust ning seejärel kohta märkides. Sellise ülemarkeeritud kasutuse võib panna emotsionaalsuse arvele, kuid M. Tomasello (2000: 10–13) arvates imiteerib laps kuulnud keelematerjali (fraasid, laused), aga ei korda seda täpselt, vaid kombineerib varem kuulnud konstruktsioone vägagi loovalt ja esitab neid samas kommunikatiivses funktsioonis, kui ta täiskasvanutelt kuulnud on. Suurt osa lapse keeleanamandamisest

seletab M. Tomasello kolme põhiprotsessi abil: kultuuriline (jälgendav) õppimine, analoogia ja abstraktsioon ning struktuuride kombineerimine (Tomasello 2000: 10). Eelnevale tuginedes võib oletada, et laps on sisendkeeles kuulnud käske *anna mulle* ja *anna mu kätte* ning moodustab nendest ise konstruktsiooni *anna mu kätte mulle*.

(3) 2;7.12: *anna mu kätte mulle*.

Mitmesuguste mõistmiseksperimentide abil on kindlaks tehtud, et väikesed lapsed ei suuda ennast kõnesituatsioonis panna teise inimese rolli. Lapsed on segaduses seetõttu, et nad ei taipa, kuidas deiktikud vestluses töötavad. Sellist keelekasutust võib nimetada pragmaatiliseks segaduseks, mida põhjustab lapse sotsiokognitiivne ebaküpsus (Tomasello 2003: 201). 1. ja 2. isiku pronoomenite omandamine läheb hõlpsamalt nendel lastel, kellel on õdesid-vendi, sest siis saab laps kuulata pronoomenite kasutamist teistele adresseeritud kõnes (Oshima-Takane jt 1999, Tomasello 2003). Sellistel lastel on võimalus näha täiskasvanu-lapse vestlust kõrvalseisja perspektiivist. Katsed on tõestanud, et lapsed omandavad pronoomenite semantilised reeglid, kuulates pealt pronoomenite vahelduva referentsiga kasutust kõnes, mis pole adresseeritud neile endile, nt lapsevanema ja vanema õe või venna vahelises vestluses (Oshima-Takane jt 1999). Nagu eeltoodud andmetest selgus, domineerivad ema kõnes ainsuse 2. isikus pöördumised lapse poole. Jane Brown ja Judy Dunn (1992) uurisid tunnetest rääkimist perekonna suhtlussituatsioonides 50 pere näitel ning leidsid, et õed-vennad on väikese lapse jaoks emast erinevad suhtluspartnerid. Kui emad rääkisid enamasti ainult lapse tunnetest, soovidest, vajadustest (n-õ suhtlus kui "hooldamine"), siis õed-vennad pöörasid tihti tähelepanu ka oma tunnete ja soovidele.

Andreae näitel on vanus 2;0–2;1 väga huvitav etapp keelelises arengus. Laps võtab kasutusele hulgaliselt uusi sõnu ning kasutab mittetäiskasvanupäraselt mitmeid personaalpronoomeneid: *mina*, *sina*, *tema*. Veanäiteid ei leia andmebaasist väga palju, kuid märgitud lühikese aja jooksul esineb huvitavaid kombinatsioone, näiteks ajab Andreas vanuses 2;1.8 segamini 1. ja 2. isiku pronoomenid, kusjuures leidub mõlematpidi kasutust, kord on *sinu* asemel kasutatud *minu* (*konn sööb minu [= sinu] ära*). Siis jällegi räägib laps endast ja markeerib seda ainsuse 2. isiku pronoomeniga (näited 4, 5). Ewa Dabrowska (2004: 162–164) peab pronoomenite ümberpööramise (*reversal*) peamiseks põhjuseks seda, et laps lihtsalt opereerib varem konstrueeritud fraasidega, s.t taaskasutab (*recycle*) täiskasvanu kõnest kuulnud lausungit või selle osi. Selline teistelt kuulnud lausungite (taas)kasutamine võimaldab moodustada süntaktiliselt keerulisi kombinatsioone, aga põhjustab vigu pronoomenikasutuses.

(4) [Laps mängib emaga piilumismängu ning kiltkab rõõmsalt]
2;0.24: CHI: *emme piilus sind* [= mind] !

(5) [Laps ootab tädi külla endaga mängima]
2;1.8: CHI: *võtab riided ära, siis mängib sinuga* [= minuga].

Vanusest 2;0.14 ja 2;0.23 leiduvad huvitavad konstruktsioonid, mis näitavad, kuidas laps hakkab pronoomenitega opereerima ning neid pöördelõppudega kokku sobitades moodustab enda loodud kombinatsioone (6a)–(6d).

- (6a) 2;0.1: CHI: **ma** lugema [*] ära .
 (6b) 2;0.14: CHI: **mina** teeb [*] hoogu.
 (6c) 2;0.23: CHI: **ma** hoiab [*] emmest kinni.
 (6d) 2;0.15 CHI: <**mina** ka tahab [*] käbi süüa ütles xx > ["].

Kui (6a) välja jätta, siis sellised moodustised tekivad seetõttu, et vaadeldaval perioodil on lapse kõnes enesest rääkides produktiivseim konstruktsioon *Antsu* + ainsuse 3. pöörde verbivorm. Sellesse konstruktsiooni hakkab Andreas asendama nime asemele pronoomenit *mina/ma* ja nii tekibki personaalpronoomeniga koos vale pöördelõpu kasutamine, s.t 1. isiku pronoomen esineb koos 3. isiku pöördelõpuga.

Materjali piiratud hulk ei võimalda välja selgitada, kuidas eesti laste pronoomenikasutus sõltub näiteks lapse sünnijärjekorrast või perekonna suhtlusstrateegiast (kui kaasatud on vanemad õed-vennad). Arvata võib, et pere ainsa lapse sisendkeel erineb mitmelapselise pere laste omast, mitme lapse osalemine suhtluses hõlbus- tab kõnerollide ja vestluses teise inimese olukorda asetumise oskuse omandamist. Mariliisi kõnest võib ka leida selliseid näiteid, kust paistab hästi läbi pere vanematelt lastelt kuuldu reprodutseerimine. Näites (7) on Mariliis võtnud vendade kõnest kuuldu analüüsimate terviku (*ise sa lõhkusid*), aga viimases sõnas on rõhutu pöördelõpp kaotsi läinud. Väikesed ja rõhutatud elemendid jäävad väikelastel keeleüksuste töötlemisel ja kombineerimisel märkamata (Tomasello 2000: 11).

(7) *Ise sa lõhkus* [*] .

Lapse kasvades väheneb tema enesekesksus ning asendub täiskasvanupärase võimega pidada tegevussubjektiks ka muid peale enda. Kõne täiustub üha, neljanda eluaasta algus on lapse intensiivse keelelise arengu aeg. Pronoomenid *mina* ja *sina* tulevad õigesti ja sagedases kasutuses alles sel ajal koos, sest siis on lapse jaoks (vähemalt igapäevases tavaolukorras ema-lapse vestluses) referendid selged, kes on *mina*, kes *sina*. Me ei saa olemasoleva materjali põhjal kindlaks teha seda, kas laps opereerib nende pronoomenitega õigesti ka siis, kui tegemist pole tema enda ja lapsevanema vahelise spontaanse suhtlusega, vaid lapsel tuleb ennast asetada mõnda teise rolli või kellegi teise olukorda. Ei leidunud näiteid selle kohta, kus lapsed oleksid mängusituatsioonis kedagi teist etendanud, nt nuku või mängukaru eest rääkinud pronoomeneid kasutades. Sellise perspektiivi muutmise oskuse väljaselgitamine nõuab spetsiaalseid eksperimente.

Kokkuvõtteks

Lapsed hakkavad pronoomenitest kõige varem kasutama demonstratiivpronoomeneid. Personaalpronoomenid ilmuvad eesti lastel umbes kaheaastaselt, kui laps on piisavalt küps, et neid kõnevoolest eristada, nende konventsionaalset kasutust tajuda ja neid teiste keeleüksustega kombineerida. Enne seda viitab laps endale ja vestluspartneri(te)le pärisnime ning üldnimega. Eri lapsed kasutavad eri mooduseid endast rääkimiseks enne personaalpronoomenite kasutuselevõttu. Personaalpronoomenite produktiivsem kasutus tuleb kolmanda sünnipäeva paiku.

Keeleandmed näitavad, et sisendkeel muutub lapse kasvades. Ema hakkab lapsega kõneldes järk-järgult kasutama keerulisemat keelt ning täiskasvanupärasemat

suhtlemisviisi. Kui sisendkeeles sagedased sõnavormid on tavaliselt sagedased ka lapse kõnes (demonstratiivpronoomen *see*, personaalpronoomenid *mina* ja *sina*), siis ainsuse 1. isiku pronoomeneid oli Andreasel kaks korda rohkem kui emal, see fakt iseloomustab lapse kõne egotsentrilisust. Emal seevastu on 2. isiku vormis pöördumisi sama palju kui 1. isiku omi, s.t ema pöördub lapse poole *sina/sa* vormidega (käskude, keeldude, küsimuste ja seletamise kaudu).

M. Tomasello arvates omandab laps keele sotsiopragmaatilisel, kolme põhi- protsessi abil: kultuuriline (jälgendav) õppimine, analoogia ja abstraktsioon ning struktuuride kombineerimine. Kõike, mida laps on kuulnud teiste kõnest, reprodutseerib ta keeleüksusi loovalt kombineerides. Kuuldud lausungite kasutamine võimaldab moodustada süntaktiliselt keerulisi kombinatsioone, aga põhjustab vigu pronoomenikasutuses.

Kirjandus

- Argus, Reili 2004. Eesti keele käändesüsteemi omandamine: esimestest sõnadest miniparadigmadeni. – Mati Erelt (peatoim.), Maria-Maren Sepper (tegevtoim.). Emakeele Seltsi aastaraamat 49 (2003). Tallinn: Emakeele Selts, 23–49.
- Bloom, Paul 2000. How Children Learn the Meanings of Words. Cambridge: MIT Press.
- Brown, Anneli 2004. Persoonareferentsi lapsenkieles. – Virittäjä 4, 483–507.
- Brown, Jane R., Judy Dunn 1992. Talk with your mother or your sibling? Developmental changes in early family conversations about feelings. – Child Development 63, 336–349.
- Clark, Eve V. 1993. The Lexicon in Acquisition. Cambridge: Cambridge University Press.
- Dabrowska, Ewa 2004. Language, Mind and Brain. Some Psychological and Neurological Constraints on Theories of Grammar. Edinburgh University Press.
- Diessel, Holger 1999. Demonstratives: Form, Function, and Grammaticalization. Amsterdam John Benjamins Publishing Company.
- 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.
- Erelt, Mati 2000. On the pro-forms in Estonian. – M. Erelt (Ed.). Estonian: Typological Studies IV. Tartu ülikooli eesti keele õppetooli toimetised 14. Publications of the Department of Estonian of the University of Tartu 14. Tartu: Tartu Ülikooli Kirjastus, 7–18.
- Kaalep, Heiki-Jaan; Kadri Muischnek 2002. Eesti kirjakeele sagedussõnastik. Tartu: Tartu Ülikooli Kirjastus.
- Kõrgvee, Kersti 2001. Lapse sõnavara areng vanuses 1;8–2;1. Bakalaureusetöö. Käsikiri Tartu Ülikooli eesti keele õppetoolis.
- Larjavaara, Matti 1990. Suomen deiksis. Suomi 156. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Lyons, John 1977. Semantics. Cambridge: Cambridge University Press.
- Oshima-Takane, Yuriko; Takane, Yoshio; Shultz, Thomas R. 1999. The learning of first and second person pronouns in English: Network models and analysis. – Journal of Child Language 26, 545–575.
- Pajusalu, Renate; Trägel, Ilona; Veismann, Ann; Vija, Maigi 2004. Tuumsõnade semantikat ja pragmaatikat. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 5. Tartu: Tartu Ülikooli Kirjastus.
- Pool, Raili 1999. About the use of different forms of the first and second person singular personal pronouns in Estonian cases. – M. Erelt (Ed.). Estonian: Typological Studies III. Tartu ülikooli eesti keele õppetooli toimetised 11. Publications of the Department of Estonian of the University of Tartu 11. Tartu: Tartu Ülikooli Kirjastus, 158–184.

- Salo, Age 1995. Eesti keele verbivormistiku omandamine vanuses 1;5–3;11. Magistritöö. Käsikiri Tartu Ülikooli läänemeresoome keele õppetoolis.
- Siewierska, Anna 2004. *Person*. Cambridge: Cambridge University Press.
- Tomasello, Michael 1992. *First Verbs: A Case Study of Early Grammatical Development*. New York: Cambridge University Press.
- Tomasello, Michael 2000. Acquiring syntax is not what you think. – D. V. M. Bishop, L. B. Leonard (Eds.). *Speech and Language Impairments in Children: Causes, Characteristics, Intervention and Outcome*. Hove: Psychology Press, 1–15.
- Tomasello, Michael 2003. *Constructing a Language: A Usage-Based Theory of Child Language Acquisition*, Cambridge, MA: Harvard University Press.
- Vider, Kadri 1995. 2–3-aastaste laste sõnavara. Diplomitöö. Käsikiri Tartu Ülikooli eesti keele õppetoolis.
- Vider, Kadri 1996. 2–3-aastaste eesti laste sõnavara. – K. Toivainen (toim.). *Suomalaiskielten omaksumista tutkimassa*. Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisu 53. Turku, 125–138.
- Vija, Maigi 2005. *Asesõnad ja isikuviitamine lapsekeeles*. Magistritöö. Käsikiri Tartu Ülikooli eesti keele õppetoolis.

Maigi Vija (Tartu Ülikool) on uurinud esimese keele omandamist, sh morfoloogiat ja pronoomeni-kasutust.
maigi.vija@ut.ee

PRONOUNS IN CHILD LANGUAGE: SOME FRAGMENTS OF A STUDY OF THE ACQUISITION OF THE PRONOUNS *MINA* AND *SINA* BY ESTONIAN CHILDREN

Maigi Vija

University of Tartu

The article focuses on the acquisition of Estonian personal pronouns *mina* 'I' and *sina* 'you' by children at the age of 1;7–3;1 and compares the children's usage with the input they receive.

The spontaneous speech data of Estonian children in the CHILDES database were analysed. The main analysis is based on Andreas's 76-hour recorded corpus. The results show that demonstrative pronouns tend to appear before personal pronouns. The usage of personal pronouns is accidental and sporadic at the beginning. When speaking about themselves, children use their own name with the 3rd person singular verb form. They also hear such constructions a lot in the input. Estonian data suggest that the way that a mother speaks with the child is related to the child's age, i.e., a three-year-old would be spoken to in more adult-like language than a two-year-old.

There are nearly 3000 instances of the short and long forms of the 1st person singular personal pronouns *mina/ma* (functioning as the subject) in Andreas's language. The 2nd person singular personal pronouns occur three times less frequently. Compared to the input, Andreas uses the long forms of the 1st and 2nd pronouns more than does his mother. Children often make mistakes in their early use of pronouns, e.g they tend to mix up the 1st and 2nd person pronouns. It is difficult for a child to understand the shift in the reference of a pronoun in conversation. That is why children use the 2nd person singular pronoun *sina* when speaking about themselves. Personal pronouns do not appear very early in the language of Estonian children. The first to appear are the primary deictic forms (*mina, sina*); their numbers are higher than those of other forms both in the language of the child and the mother.

Keywords: child language acquisition, pronouns, input, Estonian

TAVAKEELEST JA TAVAMATEMAATIKAST

Leo Võhandu

Ettekanne 24. mail 2005 Tartu Ülikooli pidulikul seminaril
“Mare Koit – 60“

Tänane seminar on tervenisti pühendatud Tartu Ülikooli raalingvistika professori Mare Koidu juubelile. Nii tähtsad juubelid algavad tavaliselt *Laudatumiga*, milles juubilari kaunid teod kõik halastamatult päevavalgele tuuakse. Meie seminaris nihkub *Laudatum* aga hoopis viimaseks sõnavõtuks, nii et ma pean esimese ettekandjana tahes-tahtmata korraga kahes rollis olema.

Lubage mul siis oma lugu sisse juhatada ühe südamest tuleva pöördumisega:

Leo Võhandu →♥! Mare Koit

Ma usun, et selle lause semantika on kõigile kohalviibijaile selge ilma mingi selgituseta. Lisan kohe etteruttavalt, et pöördumise üksikasjalikum klaarimine tuleb pärastpoole.

Et Mare tegudest õigesti aru saada ja neile vähegi väärikat hinnangut anda, peame uurima keele ja matemaatika vahekorda.

Keelel on oluline eelis matemaatika ees, sest kõik lapsed elavad keelekümb-luses ja asuvad emapiimaga koos iseendale emakeele individuaalgrammatikat konstrueerima. Matemaatika osas peab laps paraku üpris pikka aega läbi ajama ühendsüsteemiga, s.o loendamisega. Tegelikult on see süsteem ideaalne ja täpne, aga pikavõitu.

Koolitee alguseks on laps loonud isikliku grammatika, oskab tavakeeles kõik asjad ära rääkida ja foneetiliselt enam-vähem õigesti ka kirjutada.

Mis edasi juhtub? Juhtub see, et koolis hakatakse emakeele metakeelt õpetama ja keskkooli lõpuks jõutakse seisu, mille kohta hiljuti manalasse varisenud tõsine keelemnõu Eduard Vääri ütles lühidalt: toodame poolkirjaoskamatu.

Meenutame seejuures, et lapse keeleomandamise stardipositsioon ühiskonnas on igati hea. Tegemist on madalstardiga ema rinna najalt, mis garanteerib kiire lähte.

Matemaatika stardipositsioon on võrratult kehvem, sest kümb-lus puudub! Ühendsüsteem ei anna pärismatemaatikale nii vajalikku struktuurset mõtlemist.

Kuidas asja parandada? Ma arvan, et eelkooliealistele sobiksid kümblyseks pedagoogilised kaardimängud (*sic!*). Näiteks Memory (Mälu), Lolli rõõm (Idiots delight) ja Plats puhtaks! (Sweep). (Inglisekeelsed nimed kergendavad infootsingut veebist.) Miks nii? Mängukaardid on abstraktsed objektid, mida saab loendada, järjestada, seostada (paari või tihisse panna) ja pealegi saab kaardihulga elementide vahel mitmeid erilisi kombinatsioone moodustada (koodid, motiivid jne). Neil mängudel on eesmärk, selged reeglid (kasvatuslikult tähtis) ja sportlik tulemus, millela ükski mäng pole hea mäng. Eesti entsüklopeedia väitel peitub mängu motiiv küll vaid tegutsemisrõõmus ja mängul puudub eesmärk. Ärgem seda pedagoogide poolt välja öeldud mõtet siiski usume.

Mis koolis matemaatikast saab? Neli põhitehet võtavad aega viis aastat. Alles siis usaldatakse lapsi korrutama 3-4-kohalisi arve ja sedagi algoritmika seisukohast kõige viletsama algoritmi abil. (Aitaks näiteks Trachtenbergi meetod, mis vajab vaid topeldamist, poolitamist ja numbri 10-ni/9-ni täiendi võtmist.)

Kokkuvõtteks võib vist öelda, et ega matemaatika seis koolis eesti keele omast parem ole.

Ma loodan, et see lühiarutelu selgitas, kui rasketes tingimustes tuli Mare Koidul oma tööd alustada.

Keel ja matemaatika

Nihkume nüüd tavakeele ja -matemaatika tasemelt teaduse poole, Mare Koidu jahimaadele. Ma esitan teile kaks väidet, kaaluge endamisi, kas te olete nõus nendega või on tegemist mingi umbluuga.

Iga matemaatiline valem on interpreteeritav loomuliku keele lausena.

Iga loomuliku keele lause on väljendatav matemaatilise valemiga.

Jätame nüüd igäihe enda arvamuse juurde ja märgime, et Marest ligi 90 aastat varem (1916) pani need väited kirja eesti talupoja järeltulijana sündinud ja 1891. a Tallinna Raudteekooli lõpetanud Jacob Linzbach. Ta pühendas kogu oma elu (1874–1953) nende väidete tõestamisele. J. Linzbach oli täielik iseõppija matemaatika ja keeleteaduse alal, ometi loevad praegu kõige kuulsamad semiootikud nagu U. Eco, J. Kristeva, kadunud J. Lotman jt teda semiootikas F. de Saussurega ühte ritta kuuluvaks. Praegu uurib J. Linzbachi pärandit innukalt TÜ professor A. Dulicenko.

J. Linzbachi esikteos “Printsipõ filosoofskago jazõka. Opõt totsnago jazõkõznaniya” (1916) pakub meie praegusele seltskonnale selle poolest erilist huvi, et see on tegelikult maailma esimene tõsine informaatikateaduslik üllitis. J. Linzbach pani oma 228-leheküljelise raamatu kirja sõttamineku hirmus.

Ta oli oma raallingvistilises mõtteviisis kogu muust maailmast tohutult ees. Ma toon paar näidet, kuidas J. Linzbachi tollane käsitus seostub kaasajaga.

Liiasus. Me teame, et nii tavamatemaatikas kui tavakeeles on palju liiasust, millest meid päästab üleminek kahendsüsteemile. J. Linzbach näitas oma esikteoses, et kui ühendsüsteem on absoluutselt ideaalne ja täpne, siis kahendsüsteemil on piisav selgus maksimaalse lühidusega.

Otseviide “Iga keele jaoks on kaks esitust – täielik ja kokkusurutud. Võime öelda, et iga täielik esitus peab olema ühendüsteemis ja iga lühendesitus kahendüsteemis.“ J. Linzbachil oli täielik arusaamine teksti liiasusest ja sellest lühendesisitusega vabanemisest. “Mida rohkem te teksti kokku surute, seda rohkem aega kulub selle täpseks taastamiseks“.

Matemaatikutele teadmiseks, et selles 1916. aasta raamatus on esitatud täielik neljatehteline kahendarvutusskeem. Huvitav oleks teada, kes mujal maailmas ja millal sellise skeemi esitas.

Kõnekas on kahendüsteemile järgnev peatükk “Mõte ja intuitsioon kui arvu-süsteemid“.

Paraku ta raamat (200 eksemplari) ja ka hilisem metakeele teooria “Transcendent algebra” (1921) hääbusid raamatukogude arhiividesse ammu enne kui tekkis küberneetika (Wiener 1948, Shannon 1948). Muide, J. Linzbachi 1921. aasta raamatut ei leidu üldse Eesti raamatukogudes ja arhiivides. Ilmselt on ainus koopia praegu minu käes. Sain selle Hollandist Utrechti ülikoolist.

Matemaatikuna ütlesin, et J. Linzbach lõi esimese kontseptialgebra. Kunstnikuna lõi ta piktograafilise metakeele, mille põhiidee on aga ääretult põnev. Ta tõi sisse tähistused diskreetse aja ja perioodiliste protsesside jaoks. Üpris huvitav on J. Linzbachi skeem asjade maailma ja hingelise maailma olekute ja tegevuste kirjeldamiseks. Muide, veel 1952. aastal ütles Moskva TA retsensent, et kui ta nägi J. Linzbachi käsikirjas sõna *dušaa*, siis ta enam edasi ei lugenudki. Meenutan, et küberneetikat hakati Nõukogude Liidus tunnistama alles alates 1956. aastast ja sedagi sõjaväeteadlaste otsesel nõudmisel.

Ma arvan, et kui loobuda J. Linzbachi piltkeelest ja mõisteid tähistada tavasõnadega, siis tekivad üsna lühikesed ja asjalikud väljenduskeemid. (Seda J. Linzbachi kontseptialgebrat arendame edasi oma ETF-i grandil 5918 abil. Puuetega inimesed vajavad suhteliselt lihtsat formaliseeritud keelt nii kirjas kui kõnes.)

Esitan siinkohal vaid mõned ilmsed metamärgid:

- + ja, on, eksisteerib, koos
- ei, ei eksisteeri, puudub, on eraldatud, ilma, ..
- x korduv liitmine, tootmine, loomine
- / osa, kasutus, vajadus, kuuluvus
- ^ >0 tegutsejate koostöö, <0 kasutajate koostöö
- √ millegi teke, sünd
- ∞ surm, millegi lõpp

Jacob Linzbach ja Mare Koit

Nüüd saame esimese eestlasest informaatiku J. Linzbachi metakeelt kasutada Eesti esimese raallingvistika professori Mare Koidu *Laudatumi* osaliseks maitseprooviks:

- √ Mare 1945 Mare sündis 1945
- RaalLingv/ Mare → Raallingvistika kuulub Marele (on osa Marest)
- Keel ^(- Mare) → Keel ei allu Marele
- Keel^(1/ Mare) → Keel rahuldab Mare vajadusi

Keel^(+ Mare) → Keel on Mare võimuses
Mare ^mõistus → Mare tegutseb mõistuse käsul

Riskid

Mare pole üldse nii vagur inimene, kui ta pealt paistab. Ta riskib algusest peale tõsiselt. Ta esimene raamat (Koit 1983) enne lingvistikaga maadlema hakkamist oli õpetus sellest, kuidas käsitleda MUMPS-i (siiski mitte haigus, vaid dialoogiline andmebaas).

Nüüd ta tegeleb jälle dialoogi(de)ga. Ilmselt on see teema talle noorest peale jäänud südamelähedaseks. (Oli siiski nakkav!)

Dialoogis on peaaegu alati sees varjatud mõte. Metafoorid ja mõttevastandused... Tekstist on seda nähtamatute situatsioonimuutuste tõttu pagana raske püüda. Võtame kas või meie suure idanaabri keelest ühe lihtsa näite: *zakuska* ja *jedaa*. Kui dialoogis järsku hakatakse *zakuska* asemel *jedaa*'d pruukima. Mida see tähendab? Lihtsalt seda, et viin otsas ja peab otsustama, mida edasi teha...

Olgu nii! Lõpetan oma tõsise jutuarenduse sümpaatse daami Mare Koidu juubeliseminaril linzbachilikult:

Mõistus ∞ / Leo Leo mõistus on otsas
Leo - jutt ∞ Leo ilma lõppenud jututa

Kirjandus

Koit, Mare 1983. Programming in DIAMS-2 (textbook in Estonian). Tartu: University Press.

Linzbach, Jacob 1916. Printsipõ filosofskago jazõka. Opõt totšnago jazõkoznaniija. Petrograd.

Linzbach, Jacob 1921. Transcendent Algebra. Reval, Estonia.

Shannon, Claude 1948. A Mathematical Theory of Communication. Bell System Technical Journal 27, 379–423, 623–656.

Wiener, Norbert 1948. Cybernetics. Control and Communication in the Animal and Machine. New York: Wiley.

Leo Võhandu (Tallinna Tehnikaülikool) on aktiivne informaatika, mänguteooria ja rakendusliku matemaatika propageerija. Ta on informaatikainstituudi eellase, endise infotötluse kateedri rajajaid ja vastava eriala ellukutsujaid tehnikaülikoolis.
leov@staff.ttu.ee

Dialoogikorpuse lühendid ja transkriptsioonimärgid

DIE	direktiivi esiliige
DIJ	direktiivi järelliige
IL	infolisa
KÜS	täpsustav küsimus ja vastus
KYE	küsimuse esiliige
KYJ	küsimuse järelliige
PAR	parandussekvents
PPE	partneri algatatud paranduse esiliige
PPJ	partneri algatatud paranduse järelliige
RIE	rituaali esiliige
RIJ	rituaali järelliige
RY	rituaalne üksikakt
SEE	seisukohavõtu esiliige
SEJ	seisukohavõtu järelliige
TVE	teemavahetuse esiliige
TVJ	teemavahetuse järelliige
VR	vabatahtlik reaktsioon
VTE	vastuse tingimuste täpsustamise esiliige
VTJ	vastuse tingimuste täpsustamise järelliige
YA	üksikakt
.	langev intonatsioon
?	tõusev intonatsioon
,	poollangev intonatsioon
´	sõna rõhutamine
˘	rõhuline sõna või silp
[pealerääkimise algus
]	pealerääkimise lõpp
(.)	mikropaus (0,2 sek või lühem)
(...)	mikropausist pikem paus
(0.5)	pausi pikkus sekundites
=	kahe iseseisva üksuse kokkuhääldamine
> <	kiirendatud lõik
< >	aeglustatud lõik
* *	muust kõnest vaiksem lõik
AHA	(suurtähed) hääle kõvendamine
e::i	(kaksikpunkt) hääliku venitus
si-	(sidekriips) sõna jääb pooleli
{või}	ebaselgelt kuulnud lõik
{-}	ebaselgeks jäänud sõna või kõneleja
{---}	pikem ebaselgeks jäänud lõik

(())	litereerija kommentaarid ja seletused situatsiooni kohta
.hh	sissehingamine
hh	väljahingamine
.jaa	(punkt sõna ees) sõna on lausunud sisse hingates
mt	matsutamine
£	aktsent

Lapsekeele andmebaasi CHILDES transkriptsioonimärgid

%	kommentaar, mis annab lisainfot ja seletusi lüüdistussituatsiooni kohta
#	paus lausungis
[*]	täiskasvanupärasest keelekasutusest erinev vorm
xx	arusaamatu sõna
< >	fraas või lausung, mille kohta järgneb täpsustav info
[“]	deklameerimine või päheõpitud üksus

Inglisekeelsetes artiklites kasutatud lühendid

ACC	accusative
CART	classification and regression tree
CDS	child-directed speech
DA	dialogue act
DIM	diminutive
FEM	feminine
HMM	hidden Markov model
IMP	imperative
MASC	masculine
MBL	memory-based learning
MLP	multilayer perceptron
NOM	nominative
NOM	nominative
PL	plural
SG	singular
TBL	transformation-based learning
VOC	vocative

