

EESTIKEELSES TEKSTILOOMES EELISTATUD KONSTRUKTSIOONID JA KÄÄNDEVORMID

Pille Eslon

Ülevaade

Artiklis võrreldakse käändevormide kasutuseelistusi eesti ajakirjanduskeeles ja eesti õppijakeeles (vahekeeles).¹ Kahte korpusainest on analüüsitud statistikal põhinevate programmide ja eesti keele tarkvara abil. Seejärel on võrreldud mõlema korpusetekstiloomes kaks ja enam korda esinenud konstruktsioone, milles on võimalik kasutada ainult kindlaid käändevorme. Saadud tulemused kajastavad mitte üksnes eesti ajakirjandus- ja õppijakeele käändekasutuseelistusi, vaid ka nende kahe keelevariandi diskursuserinevusi. Uurimuse aluseks on korpusainestik, millel rakendatakse korpusest tulenevat analüüsi-suunda.

Võtmesõnad: korpuslingvistika, korpusest tulenev võrdlev analüüs, käändevormide kasutamine, eesti keel

¹ Tööd on toetanud riikliku programmi „Eesti keele keeletehnoloogiline tugi (2006–2010)“ projekt „VAKO: Eesti vahekeele korpuskeeletarkvara ja keeletehnoloogilise ressursi arendamine (2008–2010)“ ja riikliku programmi „Eesti keel ja kultuurimälu“ projekt „REKKi käsikirjaliste materjalide digiteerimine, Eesti vahekeele korpuskeele alamkorpuskeele loomine ja korpuskeele kasutusvõimaluste populariseerimine (2009–2013)“.

1. Lähtepunkt

Tänapäeva keeleteaduses on korpuste kasutamine uurimuste tarbeks tavapärane. Samas on korpust kui allikmaterjali kogu mõistetud erinevalt: uurija enese jaoks kitsal eesmärgil koostatud näiteainesest miljonisõnaliste elektrooniliste korpusteni, millel võib olla oma märgendussüsteem ja kasutajaliides. Erinev on olnud ka arusaam sellest, kuidas korpusi koostada ning uurida, milliseid analüüsimeetodeid ja -vahendeid rakendada, et leida vastuseid tõstatatud küsimustele, tõestada hüpoteese või välja selgitada keelendite funktsionaalne potentsiaal: vt nt inglise keele *if*-konstruktsiooni uurimus Gabrielatos 2007, soome keele muutkondade produktiivsuse indeksi määramine Nikolajev 2007, eesti keele *tulema*-verbide sageduse muutumine kindla kõneviisi ainsuse kolmanda pöörde vormis Kilgi 2006. Niisugust uurimissuunda on korpuslingvistikas enamasti nimetatud korpuspõhiseks keeleanalüüsiks (ingl k *corpus-based language analysis*), mille sünonüümina on Paul Rayson kasutanud mõistet *hypothesis-driven 'hüpoteesist tulenev'* (vt Rayson 2002: 1) ning mida Geoffrey Leech on pidanud orienteerituseks teooriale või lingvistilisele koolkonnale (*theory-oriented paradigm characteristic of some other schools of linguistics*), vt Leech 2004: 61. Töö alguses tõstatatakse uurimisküsimus, luuakse vastavalt märgendatud korpus, mida analüüsitakse (pool)automaatselt ja / või käsitsi. Täiendavate uurimismeetoditena rakendatakse korpusainese kvantitatiivset ja kvalitatiivset analüüsi, mille tulemusi tõlgendatakse lingvistiliselt (vt Rayson 2002: 13; Granger 2003). Korpuspõhine keeleuurimine pole välistanud seda, et korpusi kasutatakse jätkuvalt allikmaterjalina oma teoreetiliste seisukohade illustreerimiseks (*corpus-illustrated language research*).

Teine tänapäevane võimalus keelt uurida on korpusest või korpusainestikust tulenev uurimissuund (*corpus-driven research*

või *data-driven research*), mille puhul ei käsitleta korpust pelgalt oma teoreetilistele seisukohtadele kinnituse leidmise allikana. Leechi järgi on see keeleanalüüsi teine paradigma, mida ta on nimetanud andmetest tulenevaks (*data-driven paradigm characteristic of corpus linguistics*), vt Leech 2004: 61. Kuna korpuse keeleaines kajastab keelekasutusele iseloomulikke nähtusi, siis on oluline leida moodus, kuidas need nähtused esile tuleksid. Võimaluse selleks peaks andma korpusainese uurimine mittelingvistiliste vahenditega, rakendades erinevaid statistikal põhinevaid analüüsimeetodeid (nt keeletarkvara, statistiline andmetöötlus), mis välistavad uurija subjektiivsed valikud ning toovad esile uusi aspekte keeles toimivate seoste ja protsesside kohta, nt varjatult kulgevad keelemuutused ja ootamatud arengutendentsid. Tähelepanu keskmesse tõusevad seosed lingvistiliste struktuuride ja nende kasutussageduse vahel. Sel moel tuleb esile nii lingvistiline kui ka mittelingvistiline teave, mida korpusstatistika uurijale pakub. Just need andmed peaksid andma impulsi teoreetilist laadi mõttekäikudele. Uuritakse, kuidas on omavahel põimunud keelekasutus (ingl k *use of language*), sünkroonne varieerumine (*synchronic variation*) ja diakroonilised keelemuutused (*diachronic change*), vt Leech 2004: 77–78. Selle suuna eeliseks peetakse võimaluses tõstatada hüpoteese ning luua teooriaid uurimisobjektist tulenevalt, mitte etteantud teoreetilistest seisukohtadest lähtudes, mis omakorda lubab viia keelekirjelduse vastavusse kirjeldatava objekti olemusega ja vältida lingvisti subjektiivsete eelistuste ning arusaamade võimalikku eksitavat mõju uurimistulemustele (vt Tognini Bonelli 2002). Korpusest tuleneva uurimise aluseks võib olla kas märgendatud korpus või tekstiarhiiv, mida analüüsitakse kvantitatiivselt. Saadud tulemused näitavad, mida antud korpuse põhjal tasuks üldse uurida, mis on keelekasutuses sage, mis mitte (vt ka Rayson 2002: 14), missugused lingvistilised mustrid on

keelekasutuses olulised, kuidas neis on põimunud leksika ja grammatika. Analüüsis ei lähtuta nt pelgalt sõnavarast või grammatilistest vormidest, sõnade süntaktilistest funktsioonidest, rollidest lauses jne. Oluline on keeleline väljund, mis kajastab semantika, grammatika ja pragmaatika koosmõju lausungis, laiemalt tekstis, diskursuses².

Kolmas võimalus on lähtuda mõlemast kirjeldatud uurimisuunast, varieerides neid vastavalt uurimiseesmärgile ja analüüsitava materjali iseloomule. Leitakse, et ei korpuspõhine ega ka korpusest tulenev analüüs pole universaalne uurimismeetod, millel on võrdsed võimalused ühelt poolt mistahes keelendite uurimisel ja teisalt diskursusanalüüsis, eriti kui diskursust mõistetakse verbaalsest tekstist avaramalt (vt Baker 2006: 17–21). Samuti arvatakse, et uurimustes, mille aluseks on tasakaalustatud selgetel printsiipidel rajanevad korpused, kus tekstid on rangelt valitud, esindavad kindlat tüüpi diskursuseid jne, on erinevaid analüüsimeetodeid mõttekas rakendada integreeritult (vt Behrens 2008: XXVII; Orpin 2005: 38–39). Samuti soovitatakse kombineerida erinevates teadusharudes rakendatavaid meetodeid – statistilisi, lingvistilisi, psühholingvistilisi, sotsiolingvistilisi jt (vt Taylor 2008: 183). Hindamiseks seda, missuguseid lingvistilisi mustreid leidub erinevates tekstiliikides, millise diskursusega nende mustrite kasutamine on seotud, peetakse korpuslingvistikas vajalikuks võrrel-

² See korpuslingvistiline seisukoht haakub välja teooria ning funktsionaalgrammatilise suunaga keeleteaduses. Alates Praha lingvistilisest koolkonnast ja lõpetades näiteks Peterburi-Moskva keeletüpoloogide (V. S. Hrakovski, V. A. Plungjan jt) ning grammatikateooria uurijatega (A. V. Bondarko, J. A. Pupõnin jt) on olnud läbivaks ideeks keelendite ja kategooriate koosmõju lausungis, kus avaldub nende funktsionaalne potentsiaal, mis on omane eelkõige leksikaalgrammatilisele perifeeriale.

da erinevaid terviktekste ja testida saadud tulemusi suuremahuliste kirjakeele korpuste peal (vt nt Stubbs 1996: 126–154). Niisugust suunda on korpuslingvistikas nimetatud korpuste abil teostatavaks analüüsiks (ingl k *corpus-assisted analysis*). Seda võimalust on kasutatud õpikutekstide ja muude keeleõppematerjalide sõna- ning vormikasutuse autentsuse hindamisel (vt nt Römer 2007), registrierinevuste esiletoomisel (nt Biber 2004) jne. Lisaks klassikalisele ükskeelsele korpusainesele on peetud vajalikuks sama nähtuse uurimisse kaasata ka teisi allikaid (veebimaterjalid, spetsiaalselt antud uurimiseesmärgil produtseeritud tekstid jne), erinevat tüüpi korpusid (nt paralleel- ja tõlkekorpused).

Olles valinud mistahes uurimissuuna (korpuspõhine, korpusest tulenev, integreeritud lähenemine) ning analüüsimeetodid, tuleb neid rakendades olla süstemaatiline, püüda haarata keeleainest kogu selle rikkuses, olgugi et kitsaskohti esineb eelkõige suulise kõne korpustega. Selle metodoloogilise lähtekoha olulisust on rõhutanud ka Jarmo Harri Jantunen (2009), kui ta kirjutab vajadusest uurida keelt mitte ainult süsteemselt, vaid süstemaatiliselt korpusainest kasutades ning erinevaid uurimissuundi, samuti korpuste kvantitatiivse ja kvalitatiivse analüüsi võimalusi integreerides. Niisugune põhimõte on klassikaline: midagi ei saa välistada, kuid olles valiku langetanud, tuleb seda ka süstemaatiliselt rakendada.

Käesolevas artiklis esitatavas eesti keele käändekasutustendentside uurimuses on erinevate korpusainestike võrdlemisel lähtutud korpusest tulenevast analüüsisuunast, rakendatud erinevaid statistikal põhinevaid ning keelest sõltumatuid analüüsiprogramme ja eesti keele tarkvara. Aluseks on Douglas Biberi multidimensionaalse analüüsi idee (vt Biber 2004: 15, 17–18), mis haakub hästi artikli autori üldmetodoloogilise tõekspidamisega uurimistöö ja taksonoomiate loomise mitme-

mõõtmelisusest (vt Eslon 2006). Biberi mõte seisneb selles, et uurimistöös tuleb kasutada erinevat statistikat, eriti faktor- ja klasteranalüüsi, kuna need annavad hea võimaluse registreerida keelevariatiivsusega seotud nähtusi (nt registri varieerumist inglise keele suulises ja kirjalikus esituslaadis) ning võrrelda erinevate keelekasutusvariantide lingvistilisi mustreid. Kui multidimensionaalse analüüsi tulemused viitavad samadele nähtustele, siis ei ole põhjust uurimistulemuste objektiivsuses kahelda.

Uurimisobjekt on praegusaja eesti keelt kõige ehedamalt kajastav ajakirjanduskirjakeel³ ja õppijakeel⁴ – eesti keele kaks aktiivselt kasutatavat varianti. Korpusstatistika alusel on neis välja toodud neli sagedasemat nimisõna *inimene*, *elu*, *aeg* ja *sõna*⁵, mille käändevormide kasutusmustreid kirjalikus tekstiloomes on uuritud statistikal põhineva WordSmith Tools 5,0 programmidega ning eesti keele süntaksianalüsaatori abil; saadud väljundit on töödeldud selleks spetsiaalselt programmeeritud makrode abil⁶. Neid meetodeid omavahel kombineerides ja erineva korpusainese suhtes rakendades õnnestub välja tuua konstruktsioonid, mida eesti ajakirjanduskeele ja õppijakeele tekstiloomes on esinenud kaks ja enam korda ning milles omakorda on võimalik kasutada vaid kindlaid käände-

³ Eesti Keele Instituudi keelekorpus = EKI korpus, vt <http://www.eki.ee/corpus/> (03.09.2008).

⁴ Eesti vahekeele korpus = EVKK, vt <http://evkk.tlu.ee> (03.09.2008).

⁵ Vt Eslon & Matsak 2009; sagedasemate nimisõnade väljatoomise protseduuri on kirjeldatud artiklis Eslon 2008: 33–35.

⁶ Programmeerimistöö on teinud Erika Matsak projekti „VAKO: Eesti vahekeele korpusse keeletarkvara ja keeletehnoloogilise ressursi arendamine (2008–2010)” raames. Valimi representatiivsuse ning andmete võrreldavuse põhjendust vt Eslon & Matsak 2009.

vorme⁷. Ühelt poolt tulevad nii ilmsiks loomuliku keelekasutuse ning eesti õppijakeele diskursuse sarnasused ja erinevused. Teisalt annab korpusest tulenev statistikal põhinev analüüs teavet konstruktsioonitüüpide ja grammatiliste vormide valikust, kuvades pildi sellest, mida kujutab endast loomulik keelekasutus, mida õppijakeel. Ilmnevatel erinevustel on nii teoreetiline kui ka praktiline väärtus: ühelt poolt näitab uurimus olulisi tendentse eesti keele käändegrammatika arengus, teisalt aga võimaldab välja tuua loomuliku eesti keele ja eesti õppijakeele mõningad diskursuserinevused, tekstiloomes aktiivselt kasutatavad konstruktsioonitüübid ja käändevormid. Eristatakse 1) mõlema keelevariandi stereotüüpseid ehk kõrgsagedasi konstruktsioone, nt ajakirjanduskeeles *on (käes) teo-inimeste aeg; sõna otseses mõttes (tähenduses)* ja õppijakeeles *vabadus (vajadus) planeerida aega* jne; 2) ühe keelevariandi keskseid ehk ainult sellele keelevariandile tüüpilisi konstruktsioone, nt õppijakeelele iseloomulikud konstruktsioonid *igal inimesel on (peab olema); sõnal võib olla* jne; 3) erinevusi samalaadsete konstruktsioonide sageduses ja leksikaalgrammatilises varieeruvuses, nt kvantorfraasi kasutamine ajakirjanduskeeles *ligi (veel, umbes) 200 (kolm, 800) inimest; hukkus kolm (kaheksa, kuus, viis, üksteist) inimest* jt ning õppijakeeles *ainult (umbes) kuus (sada) inimest; ning kaheksa (üksteist) inimest, kaheksa (kuus) inimest 32-st; üksteist inimest kirjutasid* jne. Neid konstruktsioone nimetatakse tavapäraseks.

Analüüsiüksus eraldatakse formaalselt ning piiratakse kolmest sõnast koosneva keeleüksusega, mida korpuse statistikas esineb antud kujul kaks ja enam korda. Need keeleüksused ei järgi konstruktsioonipõhist süntaksimudelit, mis tugineb verbi argumendistruktuurile või noomenisüntaksile (vt Goldberg

⁷ Statistilise analüüsi on teinud Erika Matsak, vt Eslon & Matsak 2009.

1995). Ka pole need sõnakombinatsioonid otseselt seotud loogiliste struktuuridega, vaid on välja toodud statistiliselt. Käesolevas uurimuses on konstruktsioonide sageduse määraks võetud viis ja enam korda, millest piisab ülevaate saamiseks olulisematest tendentsidest eesti ajakirjandus- ja õppijakeele tekstiloomes eelistatud konstruktsioonidest ning käändevormidest. Seejuures ei jälgita, kas tegu on mingi kindlat tüüpi fraasi, moodustaja süntaktilise funktsiooni, lausekonstruktsiooni, vaba sõnaühendi, püsiühendi või idioomiga. Kolmest sõnast koosnevad keeleüksused ehk konstruktsioonid on statistika põhjal välja toodud sõnavormide kombinatsioonid. Samuti ei saa neid pidada mitmesõnalisteks väljenditeks (ingl k *multi-word expressions*), kuna sellel mõistel on kindel sisu. Näiteks Francesca Masini vaatleb mitmesõnaliste väljendite all ka partikkelverbe (Masini 2005: 145–146), Kadri Muischnek idioome, poolidiomaatilisi ja kollokatiivseid ühendeid (Muischnek 2006: 12–22, 26–27). Nadja Nesselhauf räägib sageduse alusel välja toodud süntaktilistest ja semantilistest terviküksustest (Nesselhauf 2005: 21).

Käesolevas uurimuses on statistilise analüüsi tulemusel leitud korpusainesest hulk samalaadseid korduvaid konstruktsioone, millest osa moodustavad suvalised tähenduseta, sisutühjad ja absurdsed sõnavormide, arvude, sidendite, isikunimedede, morfeemide, tähemärkide jne kooslused (nt *25 0 inimest, ka vägi elus, kolmsada kolmesaja inimene, inimene inimest küläs, d aeg lustama, 120 sõna töö* jne). Need statistiliselt välja toodud üksused eraldati tähendust omavatest konstruktsioonidest (nt *sõna sekka öelda, ehk teiste sõnadega, tol ajal oli, sellel ajal oli, sel ajal tekkis* jne) ning jäeti analüüsist välja. Seetõttu kasutataksegi antud uurimuses mõistet *konstruktsioon*, mille all peetakse silmas erinevaid struktuure, mis on välja toodud statistiliselt, mis koosnevad kolmest keeleüksusest, millel on lauses oma kindel tähendus ja süntaktiline funktsioon, nt lausekonstruk-

sioon *et inimesed on* alustab kõrvallauset – *Ants ütles mulle, et inimesed on talle palju liiga teinud; Et inimesed on enamasti unustanud, mis toimus viiskümmend aastat tagasi, siis on meie kohus ...* jne. Igal konstruktsioonil võib olla analüüsitava keelevariandile iseloomulik leksikaalgrammatiline varieeruvus (nt ajakirjanduskeele stereotüüpses konstruktsioonis *sõna otseses mõttes (tähenduses) ~ selle sõna otseses (kõige <otsesemas mõttes, tähenduses>)*), mida nimetatakse antud konstruktsiooni kasutusmustriks. Viimase invariandiks on tekstiloomes kõige sagedamini kasutatud konstruktsioon – antud juhul *sõna otseses mõttes*. Samuti võivad konstruktsioonid olla omavahel seotud kui süntaktilised sünonüümid, nt *mõni aeg hiljem ~ mõni aeg pärast <mida> ~ mõne aja pärast* jne.

Niisiis peetakse käesolevas uurimuses oluliseks, missuguseid kolmest keeleüksusest koosnevaid statistilisi konstruktsioone eelistatakse tekstiloomes kasutada. See võimaldab konstruktsioone tõlgendada kasutuspõhiselt.

2. EKI tekstikorpuses ja EVKK tekstides eelistatud konstruktsioonid ning käändevormid

Konstruktsioonid eristuvad vastavalt sagedusele, tüübile ja leksikaalgrammatilisele varieeruvusele. Omaette jaotustena kirjeldatakse eesti ajakirjandus- ja õppijakeele stereotüüpsed, tavapäraseid ning keelevariandi keskseid ehk tüüpilisi konstruktsioone ja nende leksikaalgrammatilisi variante.

2.1. Stereotüüpsed konstruktsioonid ja nende leksikaalgrammatilised variandid

Stereotüüpsete konstruktsioonide all mõistetakse ühes või teises keelekasutusvariandis regulaarselt ilmnevaid konstruktsioone.

sioone ja nende leksikaalgrammatilisi variante, mida kasutatakse tavaliselt hinnangut väljendava markeri või tekstisidususvahendina. Ülisuure sageduse tõttu peetakse neid tekstiloomes enamasti parasiitväljenditeks, mida keelekorraldajad soovivad vältida. Samas kasutatakse eesti ajakirjanduskeeles näiteks sõnade *aeg* ja *sõna* käändevorme kindlates stereotüüpsetes konstruktsioonides, mida on vaja läinud hinnangu edastamiseks või siis tekstisidususvahendina: *on (käes) teoinimeste aeg; sõna otseses mõttes (täheenduses); selle sõna otseses (kõige <otsesemas mõttes, täheenduses>); tema (oma) sõnade kohaselt (järgi); tehnika viimane sõna ~ oma viimane sõna ~ oma viimase sõna; oma sõna öelda (ütleva).*

Need kirjakeele stereotüübid ei ole omased õppijakeelele – seda iseloomustavad teistlaadsed leksikaalgrammatilised konstruktsioonid, mille ülisage korduvus korpusaineses viitab ühelt poolt keeleõppe teemakesksusele ja teisalt õppijakeele korpuse tekstide annoteerimisega seonduvatele asjaoludele.

Teemakesksus väljendub nt konstruktsioonis *vabadus (vajadus) planeerida aega*, mida on kasutatud stereotüüpsena seoses oma positiivse suhtumise väljendamisega aja planeerimise vajalikkusse jms; konstruktsiooni *noorkeraamika (kammkeraamika) inimesed sarnanesid* kasutamine stereotüüpsena on seletatav kultuurilooeteemaliste esseede ja kontrolltööde rohkusega EVKKs. Neis tekstides on pikalt kirjutatud Eesti riigi ajaloost. Samal põhjusel kuuluvad stereotüüpsete hulka ka konstruktsioonid sõna *aeg* kindlate käändevormidega: 1) lausekonstruktsioonid *et (kuid, sest) sel ajal ~ sest tol ajal ~ ja samal ajal*; 2) verbikesksed lausekonstruktsioonid *tol (sellel) ajal tekkis (kasutati) ~ samal ajal läksid ~ praegusel (sel, sellel) ajal on ~ vabariigi (vene, *tsari) ajal oli ~ sõja ajal polnud ~ vabal (vabariigi) ajal teha*; 3) noomenikonstruktsioonid *samal ajal ajaloolised <sündmused> ~ tol ajal inimesed ~ sel ajal laulud*; 4) modaalne konstruktsioon *külmal ajal peab*.

Õppijakeele korpuse tekstide annoteerimisega seonduvatest asjaoludest johtuvalt on stereotüüpide hulka sattunud ka tööjuhised, mis EVKKs on tekstist eraldamata. See küsimus lahendatakse korpuse edasise arendamise käigus, kui luuakse uusi alamkorpusi, täiendatakse ja ühtlustatakse metainfot, vaadatakse veel kord läbi korpuses sisalduvad tekstid, parandatakse sisestamisapsud, kontrollitakse üle sisseskanneeritud tekstid, eraldatakse õppija kirjutatud tekstiosa pealkirjast ning tööjuhistest jms. Praegu on tööjuhised õppijakeele stereotüüpsete konstruktsioonide hulgas, milles on *sõna* käändevorme kasutatud: 1) imperatiivses lausekonstruktsioonis *pane sõna(d) õigesse <vormi>; moodustage etteantud sõnadega* ja 2) kvantorfraasis *umbes 80 (120, 160) sõna*. Iseenesest pole selles midagi taunitavat, et tööjuhised kuuluvad õppijakeele stereotüüpide alla, sest need konstruktsioonid on keeleõppeprotsessi lahutamatu osa.

2.2. Tavapärased konstruktsioonid ja nende leksikaalgrammatilised variandid

Võrreldavate keelevariantide tavapäraseid konstruktsioone iseloomustab samalaadsus, erinevusi ilmneb vaid konstruktsioonide sageduses ning leksikaalgrammatilises varieeruvuses. EKI tekstikorpuse ja EVKK tekstiloomes eelistatud tavapärastes konstruktsioonides on eesti keele neljateistkümnest käändest kasutatud vaid grammatilisi – nominatiivi, genitiivi ja partitiivi, kusjuures valdav käändevorm on nominatiiv. Järgnevalt kirjeldame neid konstruktsioone käänete kaupa.

2.2.1. Ainsuse ja mitmuse nominatiiv

Nii loomulikus keelekasutuses kui ka eesti õppijakeeles on nominatiivi sisaldavad konstruktsioonid olnud kõige sagedasemad ja kõige rikkalikuma leksikaalgrammatilise varieeru-

vusega. Lahknevused esinevad nii konstruktsiooni tüüpides kui ka nende leksikaalgrammatilistes variantides, mis annab alust rääkida eesti ajakirjandus- ja õppijakeele diskursuseri-
nevustest sõnade *inimene, elu, aeg* ja *sõna* kasutamisel.

EKI tekstikorpus

Ainsuse nominatiivi vormi sisaldavad konstruktsioonid:

- 1) loogiline implikatsioon *kui inimene on ~ et inimene on*
- 2) eitav verbikeskne lausekonstruktsioon *inimene ei ole (ei saa)*
- 3) lausekonstruktsioon *on ilus (paras, viimane, õige) aeg; aeg on kallis ~ aeg on möödas (läbi, käes); oli aeg mil*
- 4) ajatähenduslik kvantorfraas *mõni aeg hiljem ~ mõni aeg pärast <mida>; on (oli, olnud, ajanud) kogu aeg*

Mitmuse nominatiivi vormi sisaldavad konstruktsioonid:

- 1) lausekonstruktsioon *need inimesed kes (kellele)*
- 2) jaatav (eitav) verbikeskne lausekonstruktsioon *et inimesed on (ei ole) ~ et (ka) need inimesed; need inimesed on; tema (ta) elu on*

EVKK

Ainsuse nominatiivi vormi sisaldavad konstruktsioonid:

- 1) loogiline implikatsioon *kui inimene on* (variandid: *kui inimene tahab, kogeb, elab, suhtleb*) *~ et inimene peab ~ ja kui (et kui, sest kui) inimene*
- 2) verbikeskne, enamasti modaalne lausekonstruktsioon *inimene peab olema (maksma, teadma) ~ inimene võib töötada; haritud (iga) inimene peab (saab, võib, tahab); inimese elu on (võib, sõltub, erineb); elu sõltub elukohast, <meie> elu erineb Ukraina <elust>; elu on väga*
- 3) mittemodaalne lausekonstruktsioon *see aeg on ~ mõni aeg oli; muutub (on) kogu aeg; on aeg kus; esimene sõna mis; lendab aeg väga*
- 4) verbikeskne eitav lausekonstruktsioon *inimene ei saa (ei ole)*

5) noomenikonstruktsioon *sõna eesti kultuur, väga levinud sõna ~ nõ esimene sõna*

Mitmuse nominatiivi vormi sisaldavad konstruktsioonid:

1) jaatav (eitav) verbikeskne lausekonstruktsioon *kõik (need) inimesed on (hoidsid, tahavad); inimesed kes on (ei ole)*

2.2.2. Ainsuse ja mitmuse genitiiv

EKI tekstikorpuses esineb suhteliselt harva genitiivi sisaldavaid konstruktsioone. Ajakirjanduskeeles tuli esile vaid ainsuse genitiivi sisaldav kvantorfraas. Ka EVKKs esines seda käännet teiste grammatiliste käännete vormidega võrreldes vähem. Õppijakeeles kasutati ainsuse genitiivi kahte liiki lausekonstruktsioonides, ajatähenduslikus kvantorfraasis ja noomenikonstruktsioonis. Mitmuse vormi kasutati küll sagedamini, kuid peamiselt ülivõrret sisaldavas noomenikonstruktsioonis ja lausekonstruktsioonis *on läbi aegade ~ läbi aegade on; on kõigi aegade*.

EKI tekstikorpus

Ainsuse genitiivi sisaldavad konstruktsioonid:

1) kvantorfraas *üle 200 (5000, 300, tuhande, saja) inimese*

EVKK

Ainsuse genitiivi sisaldavad konstruktsioonid:

1) jaatav (eitav) verbikeskne lausekonstruktsioon *mõjutab (ei mõjuta) inimese elu ~ inimese elu mõjutab*

2) lausekonstruktsioon *ja vanemate elu ~ et noorte elu ~ et inimese elu*

3) ajatähenduslik kvantorfraas *mõne aja pärast <tuli (oli)>*

4) noomenikonstruktsioon *inimese elu kõige ~ ?inimese elu haridus ~ *inimese elu tema (?teema) ~ *inimese elu ma (?maal)*

Mitmuse genitiivi sisaldavad konstruktsioonid:

- 1) verbikeskne lausekonstruktsioon *kahjustab inimeste tervist*
- 2) noomenikonstruktsioon *kõigi aegade suurim (parim, esimene, teine); ESTO-d läbi aegade; Eesti kõigi aegade; hümmni sõnade kirjutaja (autor); Eesti hümmni sõnade*
- 3) lausekonstruktsioon *on läbi aegade ~ läbi aegade on; on kõigi aegade; sõnade autor oli; et hümmni sõnade*

2.2.3. Ainsuse ja mitmuse partitiiv

EKI tekstikorpuse valimis on aktiivselt kasutatavatest konstruktsioonidest ainsuse partitiiviga eelistatud rikkaliku leksikaalse varieeruvusega kvantorfraasi. Mitmuse vormi esines sageli lausekonstruktsioonis *inimesi kellel on ~ inimesi kes ei; neid inimesi kes; on inimesi kes*. EVKKs on ainsuse genitiivi vormiga kasutatud erinevat tüüpi konstruktsioone rohkem kui EKI tekstikorpuses, ent valdavaks osutus samuti kvantorfraas. Ka mitmuse vorm esines samas konstruktsioonitüübis.

EKI tekstikorpus

Ainsuse partitiivi sisaldavad konstruktsioonid:

- 1) kvantorfraas *ligi (veel, umbes) 200 (kolm, 800) inimest; hukkus kolm (kaheksa, kuus, viis, üksteist) inimest; kuu aega tagasi (hiljem, enne); umbes (ligi, ja) kuu aega; juba pikka aega ~ väga pikka aega; oli (on) pikka aega ~ pikka aega on; tükk aega tagasi (pärast) ~ nädal aega tagasi; on veel aega ~ vajan veidi aega*

Mitmuse partitiivi sisaldavad konstruktsioonid:

- 1) lausekonstruktsioon *inimesi kellel on ~ inimesi kes ei; neid inimesi kes; on inimesi kes*

EVKK

Ainsuse partitiivi sisaldavad konstruktsioonid:

- 1) kvantorfraas *ainult (umbes) kuus (sada) inimest; ning kaheksa (üksteist) inimest, kaheksa (kuus) inimest 32-st; üksteist inimest kirjutasid; väga vähe aega ~ nii palju aega ~ liiga palju aega*
- 2) modaalne kvantorfraas **pidis (*pidab, pidada, pidama) mõnda aega*
- 3) verbikeskne lausekonstruktsioon *inimese elu mõjutab ~ mõjutavad inimese elu; kuulen sõna kultuur, ma kuulen sõna, kui kuulete sõna, on mõeldud sõna ~ <on> mõeldud sõna neljakesi*
- 4) jaatav (eitav) verbikeskne lausekonstruktsioon *ei ole (ei olnud) aega*

Mitmuse partitiivi sisaldavad konstruktsioonid:

- 1) kvantorfraas *väga (nii) palju inimesi*

2.3. Eesti ajakirjandus- ja õppijakeelekesksed konstruktsioonid ning nende leksikaalgrammatilised variandid

Keelevariandi kesksed ehk tüüpilised konstruktsioonid on isoleeritud vaid ühele võrreldavatest keelevariantidest. Kuigi nad kuuluvad tekstiloomes eelistatud konstruktsioonide alla, ei saa neid samastada stereotüüpsete konstruktsioonidega, kuna nad pole ülisagedased ega teatud määral tuhmunud tähendusega hinnangulisust edastavad või tekstisidususe eesmärgil kasutatavad markerid. Õppijakeeles kuuluvad siia need konstruktsioonid, milles on sageli kasutatud semantiliste käänete vorme: ainsuse adessiivi, inessiivi ja mitmuse komitatiivi.

Ainsuse adessiivi sisaldavad konstruktsioonid:

- 1) loogiline implikatsioon *kui inimesel on (ei ole)*
- 2) modaalne verbikeskne lausekonstruktsioon *igal inimesel on (peab olema); sõnal võib olla*

Ainsuse inessiivi sisaldavad konstruktsioonid:

1) verbikeskne lausekonstruktsioon *minu elus on (oli) ~ meie elus on; elus on palju (väga)*

Mitmuse komitatiivi sisaldavad konstruktsioonid:

1) verbikeskne lausekonstruktsioon <minu> *jaoks seostub sõnadega*

Ajakirjanduskeele tekstiloomele polnud semantiliste käänete vorme sisaldavad konstruktsioonid sedavõrd omased, et oleksid mahtunud kasutuseelistuste tippu. Ajakirjanduskeelekeskses osutus EKI tekstikorpuse valimi alusel hoopis nominaatiivne substantiivkonstruktsioon (*inimene ja seadus ~ inimene ja loodus; haridus ja elu ~ kultuur ja elu*), mida kasutatakse ajalehele iseloomulikus funktsioonis – enamasti pealkirja või eraldi fraasina lugeja tähelepanu köitmiseks.

Õppija- ja ajakirjanduskeelekesksed konstruktsioonid ei ole omavahel üksüheses vastavuses, kuid seda ei saa hinnata ei õppijakeele puudusena ega loomuliku keelekasutuse rikkusena. Tegu on mõlemale eesti keele kasutusvariandile funktsionaalselt omaste konstruktsioonitüüpidega.

3. Kokkuvõte

EKI tekstikorpuse ja EVKK analüüsi alusel välja toodud tekstiloomes aktiivselt kasutatavad konstruktsioonid jagunesid kolme rühma: stereotüüpsed, tavapärased ja ajakirjandus- või õppijakeelekesksed ehk tüüpilised. Esimesse ja viimasesse rühma kuuluvad konstruktsioonid kajastavad mõlema keelevariandi eripära. Tavapäraste konstruktsioonide võrdlemine annab aga ettekujutuse loomuliku keelekasutuse ja õppijakeele universaalidest ning võimaldab välja tuua erinevusi konstruktsioonide leksikaalgrammatilises varieeruvuses ja kasutusmustrites.

3.1. Universaalsed nähtused

Selgus, et kõige universaalsem on loogilist implikatsiooni sisaldav konstruktsioon sõna *inimene* ainsuse nominatiivi vormiga: 1) ajakirjanduskeel – *kui inimene on ~ et inimene on*; 2) õppijakeel – *kui inimene on ~ et inimene peab*. Erinevus loomuliku keelekasutuse ja õppijakeele vahel seisnes vaid konstruktsiooni leksikaalses varieerumises olemis- ja kogemisverbidega, mis oli iseloomulik õppijakeele tekstiloomele: *kui inimene on / tahab / kogeb / elab / suhtleb*.

Teine universaalne konstruktsioon sõna *inimene* ainsuse nominatiivi vormiga on seotud eitusega: 1) ajakirjanduskeel – *inimene ei ole / ei saa*; 2) õppijakeel – *inimene ei saa / ei ole*. Erinevus seisneb vaid selles, et õppijakeeles on eelistatum modaalne predikaat, loomulikus keelekasutuses aga eksistentsiaalne.

Kolmas universaalne konstruktsioon on jaatava / eitava kõneliigi verbikeskne lausekonstruktsioon, milles on kasutatud sõna *inimene* mitmuse nominatiivi vormis: 1) ajakirjanduskeel – *need inimesed on*; 2) õppijakeel – *kõik / need inimesed on / hoidsid / tahavad*. Õppijakeeles on selles konstruktsioonis eelistatud täiendit *kõik*; samuti tuleb ette verbi leksikaalgrammatilist varieerumist (*on / hoidsid / tahavad*).

Neljas universaalne konstruktsioon on kvantorfraas, milles sõna *inimene* kasutatakse ainsuse partitiivi vormis: 1) ajakirjanduskeel – *ligi / veel / umbes 200 / kolm / 800 inimest*; 2) õppijakeel – *ainult / umbes kuus / sada inimest*. Nii loomulikus keelekasutuses kui ka õppijakeeles esineb kvantori leksikaalset varieerumist.

3.2. Erinevused

Erinevused loomuliku keelekasutuse ja õppijakeele vahel on seotud ajatähendusliku kvantorfraasiga, mis võib kuuluda

kaassõnafraasi. Selles konstruktsioonitüübis on EKI tekstikorpuse valimis eelistatud sõna *aeg* ainsuse nominatiivi, EVKKs aga ainsuse genitiivi: 1) ajakirjanduskeel – *mõni aeg hiljem ~ mõni aeg pärast <mida>*; 2) õppijakeel – *mõne aja pärast <tuli / oli>*. Tegu on ajatähendusliku kvantorfraasi kahe leksikaalgrammatilise variandiga, millest mõlemad on aktsepteeritavad ja vastavad eesti kirjakeele normile. Samas on loomulikus keelekasutuses eelistatum nominatiivi sisaldav konstruktsioon.

Teine erinevus ajakirjandus- ja õppijakeele vahel on seotud samuti ajatähendusliku kvantorfraasiga, milles sõna *aeg* kasutatakse ainsuse partitiivi vormis, kuid kvantorid varieeruvad: 1) ajakirjanduskeel – *väga pikka aega ~ juba tükk aega / kuu aega / nädal aega*; 2) õppijakeel – *väga vähe aega ~ nii / liiga palju aega*. Loomulikus keelekasutuses eelistatakse rõhusõnaga esile tuua pikemat ajaperioodi (*pikka aega, tükk aega, kuu / nädal aega*), samas kui õppijakeeles konstateeritakse kas ajanappust või seda, et aega on küllalt (*vähe aega ja palju aega*). Kvantorite varieerumine näitab, et ajatähendusliku kvantorfraasi kasutamine on õppijakeeles piiratum kui loomulikus keelekasutuses. Sedalaadi erinevuste väljatoomine on keeleõppe tarvis oluline, kuna need viitavad mõningatele vajakajäämistele õpikutes ja õppematerjalides. Viimaseid oleks vaja mitmekesistada, viies sisse uusi (ala)teemasid ja lisades õppeülesandeid, mida täites saab aktiveerida ka pikemat ajaperioodi väljendava ajatähendusliku kvantorfraasi kasutamist.

Kolmas erinevus on seotud konstruktsioonidega, milles sõna *inimene* on kasutatud mitmuse partitiivi vormis: 1) ajakirjanduskeel – eelistatud on dünaamilist lausekonstruktsiooni ja selle leksikaalgrammatilisi variante *inimesi kellel on ~ inimesi kes ei; neid inimesi kes; on inimesi kes*; 2) õppijakeel – eelistatud on rõhusõna sisaldavat kvantorfraasi *väga / nii palju inimesi*. Erinevad nii tekstiloomes aktiivselt kasutatavad konstruktsioonid.

sioonitüübid kui ka esituslaad: kirjakeelele on omane jutustav-dünaamiline, õppijakeelt iseloomustab noomenisrktuuridele omane kirjeldav esituslaad. Niisugused erinevused näitavad kätte suuna, milles keeleõppija oskusi tuleks arendada.

Neljandat liiki erinevused, mis ilmnevad ajakirjandus- ja õppijakeeles eelistatud konstruktsioonide võrdlusest, viitavad sellele, et teatud konstruktsioonid varieeruvad õppijakeeles leksikaalgrammatiliselt tunduvalt rikkalikumalt kui loomulik keelekasutuses. Ühelt poolt on näiteks eitava kõneliigi verbikeskne ainsuse nominatiivi sisaldav lausekonstruktsioon *inimene ei ole / ei saa* (vt eespool) loomuliku keelekasutuse ja õppijakeele universaal. Teisalt aga on õppijakeeles sama konstruktsiooni jaatava kõneliigi variandis esindatud rikkaliku leksikaalsemantilise varieeruvusega verbikeskne modaalne lausekonstruktsioon *inimene peab olema / maksuma / teadma ~ inimene võib töötada; haritud (iga) inimene peab / saab / võib / tahab*, mis ajakirjanduskeele eelistuste seas pole statistiliselt esile tulnud. Ometi peaks eeldatavalt just hinnanguid sisaldav isiksuseline esituslaad olema omane publitsistlikule arutlusele ning modaalised konstruktsioonid ajakirjanduskeelele. Samas pole ajakirjandustekstide analüüs seda kinnitanud. Tekib küsimus, miks? Kas korpusest tuleneva keeleanalüüsi andmeid saab interpreteerida kui meie päevalehtedele omaseks saanud tendentsi – vähe publitsistikat ja rohkesti isikupäratuid “lugusid”? Kas ajakirjanikud on hakanud hinnangute andmist vältima ning kirjutavad informatiivseid nupukesti? Samas on meeldiv tõdeda, et eesti keele õpetamisel on tehtud rõhuasetus oma arvamuse ja hinnangute väljendamisoskuse arendamisele – on ju õppijakeeles jaatava kõneliigi modaalne lausekonstruktsioon kõige sagedasem.

3.3. Kaks väljundit

EKI tekstikorpuse ja EVKK korpusest tulenev võrdlev uurimine suunab meie tähelepanu küsimustele, mis traditsioonilise korpuspõhise analüüsi puhul ei pruugi üles kerkida. Eelkõige puudutab see loomuliku keele kasutuseelistusi (konstruktsiooni tüüp, selle leksikaalgrammatiliste variantide olemasolu, rangelt valikuline käändevormide kasutus), mille taga võib aimata ja tänu millele saab selgemalt piiritleda mõningaid keelesüsteemi arengutendentse. Käesoleva uurimuse alusel sai kinnitust üldteada seisukoht, et loomuliku eesti keele tekstiloomes eelistatakse grammatilisi käändevorme ja valdavalt nominatiivi sisaldavaid konstruktsioone. Sama tendents ilmnes ka õppijakeele tekstiloomes. Tasub uurida, mis seda põhjustab. Kas keelesisene ökonoomia, suundumus käändeparadigma lühenemisele, mis lõppeks võib viia eesti keele nominatiivistumiseni? Või on tegu hoopis kontaktsituatsioonis tekkinud keelemuutusliku protsessiga, mida suunavad eesti keelt teise / võõrkeelena kõnelevad inimesed või eestlased, kes on omandanud teise ja kolmanda keele kõrgtasemel?

Teine tähelepanek seostub loomuliku eesti keele ja õppijakeele konstruktsioonierinevustega, millest annab teha järeldusi vajakajäämiste kohta keeleõppes. Nende kahe eesti keele variandi korpusest tulenev võrdlev analüüs võimaldab saada rangema piirilise ettekujutuse sellest, mida õppijad on omandanud ning mida nad tegelikult peaksid omandama, et hakata eesti keelt kasutama samalaadselt eestlastega. Loomulikult jääb õhku küsimus, kas ikka kõik, mis ilmneb eestlase emakeelekasutuses, võib olla vaieldamatuks eeskujuks eesti keele kui teise / võõrkeele õppijale (nt jaatava kõneliigi modaalsete konstruktsioonide väljajäämine kasutuseelistuste hulgast). Lihtsaim moodus selles veendumiseks on eesti keele õigekirjaspelleri kasutamine ajakirjandustekstidel. Lisaks erinevuste väljatoomisele

tavapäraste konstruktsioonide kasutamisel loomulikus keelekasutuses ja õppijakeeles on siinkohal olulised loomuliku keelekasutuse stereotüüpsed konstruktsioonid, eriti hinnangu-, rõhumarkerite või sidususvahenditena.

Kirjandus

Baker, Paul 2006. *Using corpora in discourse analysis*. London: Continuum.

Behrens, Heike (Ed.) 2008. *Corpora in language acquisition research: History, methods, perspectives*. Amsterdam / Philadelphia: John Benjamins Publ. Co. Trends in language acquisition research 6.

Biber, Douglas 2004. Conversation text types: A multi-dimensional analysis. – JADT: 7es Journées internationales d'Analyse statistique des Données Textuelles / Ed. by G. Purnelle, C. Fairon, A. Dister. Louvain, 15–34. http://www.cavi.univ-paris3.fr/lexicométrica/jadt/jadt2004/pdf/JADT_000.pdf (9.05.2009).

Gabrielatos, Costas 2007. *If-conditionals as modal colligations: A corpus-based investigation*. – Proceedings of the Corpus Linguistics Conference: Corpus Linguistics 2007 / Ed. by M. Davies, P. Rayson, S. Hunston & Pernilla Danielsson. Birmingham: University of Birmingham. http://www.corpus.bham.ac.uk/corplingproceedings07/paper/256_Paper.pdf (24.03.2009).

Goldberg, Ada 1995. *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Eslon, Pille & Matsak, Erika 2009. Eesti keele kasutusvariandid: korpusest tulenev käändevormide võrdlev analüüs. – Eesti Rakenduslingvistika Ühingu aastaraamat 5 / Toim. H. Metslang, M. Langemets, M.-M. Sepper, R. Argus. Tallinn: Eesti Keele Sihtasutus, 79–110.

Eslon, Pille 2008. Käändevormide kasutussageduse võrdlus eesti õppijakeeles ja kirjakeeles. – Õppijakeele analüüs: võimalused, probleemid, vajadused / Toim. P. Eslon. Eesti filoloogია osakonna toimetised 10. Tallinn: Tallinna Ülikooli kirjastus, 31–66.

- Eslon, Pille 2006. Analoogiast keelte kõrvutamisel. – Keel ja Kirjandus 1, 15–24.
- Granger, Sylviane 2003. Error-tagged learner corpora and CALL: A promising Synergy. – CALICO Journal 20 (3), 465–480.
- Jantunen, Jarmo Harri 2009. Ei pelkästään mielikuvituksen puutteen vuoksi – Kieliaineistojen systemaattinen käyttö kielentutkimuksessa. – Virittäjä 1, 101–113.
- Kilgi, Annika 2006. Mida eestlane teeb? – Oma Keel 12, 20–24.
- Leech, Geoffrey 2004. Recent grammatical change in English: data, description, theory. – Advances in corpus linguistics: Papers from the 23rd international conference on English language research and computerized corpora (ICAME 23) Göteborg 22–26 may 2002 / Ed. by K. Aijmer & B. Altenberg. Amsterdam: Rodopi, 61–81.
- Masini, Francesca 2005. Multi-word expressions between syntax and the Lexicon: the case of Italian verb-particle constructions. – SKY Journal of Linguistics 18, 145–173.
- Muischnek, Kadri 2006. Verbi ja noomeni püsiühendid eesti keeles. Dissertationes philologiae Estonicae Universitatis Tartuensis 17. Tartu: Tartu Ülikooli Kirjastus.
- Nesselhauf, Nadja 2005. Collocations in a Learner Corpus. Amsterdam, Philadelphia: John Benjamins.
- Nikolajev, Alexander 2007. Suomen nominaalisen taivutusjärjestelman kvantitatiivista analyysia. – Kielitieteen päivät. Oulu 24.–25. toukokuuta 2007. Abstraktikirja, 75, <http://www oulu.fi/kielitieteenpaivat2007/Abstraktit.pdf> (25.03.2009).
- Orpin, Debbie 2005. Corpus linguistics and critical discourse analysis: Examining the ideology of sleaze. – International Journal of Corpus Linguistics 10 / 1, 37–61.
- Rayson, Paul 2002. Matrix: A statistical method and software tool for Linguistic analysis through corpus comparasion. Ph.D. thesis in Computer Science. Computing Department Lancaster University. <http://ucrel.lancs.ac.uk/people/paul/publications/phd2003.pdf> (30.04.2009).

Römer, Ute 2007. Learner language and the norms in native corpora and EFL teaching materials: A case study of English conditionals. – Anglistentag 2006. Halle. Proceedings / Ed. by S. Volk-Birke & J. Lipfert. Trier: Wissenschaftlicher Verlag Trier, 355–363.

Stubbs, Michael 1996. Text and corpus analysis: Computer-assisted studies of language and culture. Language in Society 23. Oxford: Blackwell Publ.

Taylor, Charlotte 2008. What is corpus linguistics? What the data says. – ICAME Journal, 32: 179–200. http://icame.uib.no/ij32/ij32_179_200.pdf (24.03.2009).

Tognini Bonelli, Elena 2002. Functionally complete units of meaning across English and Italian: Towards a corpus-driven approach. – Lexis in Contrast. Corpus-based Approaches / Ed. by B. Altenberg & S. Granger. Philadelphia: John Benjamins, 73–95.

Contextual preferences the use of case forms in text production in Estonian

Pille Eslon

Summary

The objective of the study was to compare contextual preferences the use of case forms in two variants of Estonian – standard language (Standard Estonian corpus of the Institute of the Estonian Language) and learner language (Estonian interlanguage corpus of the Tallinn University). The material was most common nouns in Estonian *inimene* ‘person’, *sõna* ‘word’, *elu* ‘life’ ja *aeg* ‘time’. The statistics showed frequency of constructions and grammatical case forms in Text. Statistical representation of lexical units, grammatical forms and constructions has enabled us to reveal such linguistic data

that, despite being characteristic of language use, is not easily accessible by traditional corpus-based analysis. Therefore we believe that corpus-driven studies using statistical and language software have a long-term theoretical and applied value (in language teaching and learning materials).

Analysis of contextual preferences making explicit the hidden tendencies that are working in the language system synchronically. For this purpose, the grammatical constructions and their lexico-grammatical variants (allowing only certain case forms) typical of Standard Estonian and learner Estonian were found out. Mere statistics of case forms would never have provided the data on their use in constructions and on their preference order in text production. Results of this study showed that in both language variants, the clearly preferred case is the nominative. Further research should reveal whether this phenomenon can be considered a process of nominativization, and what its possible motivations could be.

Keywords: corpus linguistics, corpus-driven comparative language analysis, use of case forms in most frequently grammatical constructions, Estonian language

Autor

PhD Pille Eslon, Tallinna Ülikooli eesti keele ja kultuuri instituudi vanemteadur, riikliku programmi „Eesti keele keeletehnoloogia tugi (2006–2010)” projekti „VAKO – Eesti vahekeele korpuse keeletarkvara ja keeletehnoloogilise ressursi arendamine” vastutav täitja, pille.eslon@tlu.ee