

C1-tasemega eesti keele õppijate ja emakeelekõnelejate kirjaliku keelekasutuse võrdlus verbialguliste tetragrammide näitel

KAIS ALLKIVI

Tallinna Ülikool

Ülevaade. Kõrgeim eesti keele kui võõrkeele oskuse tase, milles korraldatakse eksamit ja mida tööandja saab töötajailt nõuda, on vilunud keelekasutaja tase C1. Artiklis võrreldakse vilunud keeleõppija (K2) ja haritud emakeelekõneleja (K1) kirjalikus keelekasutuses eelistatud morfosüntaktilisi mustreid verbialguliste neljasõnaliste järjendite ehk tetragrammide alusel. Tegu on kvantitatiivse kasutuspõhise uurimusega. Rakendatud on klasteranalüüsi, mis toob tekstis esile morfosüntaktiliselt samalaadsed sõnajärgnevused. Keeleaines on pärit Tallinna Ülikooli eesti vahekeele korpuse C1-tasemel esseedest ja emakeelekõnelejate arvamuskorpusel alamkorpusest. K1 ja K2 erisus avaldub selles, mis struktuuriga tetragramme eelistatakse, kui avar või piiritletud on nende morfoloogiline, süntaktiline ning leksikaalne varieerumine. K2-s avaldub suundumus stereotüüpidele (nt seoses eitava kõne, lihtmineviku, *olema*-verbiga); teisalt tuleb esile suurem morfosüntaktiline mitmekesisus.

Võtmesõnad: kasutuspõhine keelekirjeldus; morfosüntaktiline varieerumine; leksikaalne varieerumine; kontrastiivne vahekeele analüüs; klasteranalüüs; teise keele omandamine; eesti keel

1. Uurimuse eesmärgid ja lähted

Euroopa keeleõppe raamdokument, mille kuueastmelisest keeleoskustasemete süsteemist (A1–A2–B1–B2–C1–C2) lähtutakse ka eesti keele kui teise keele (K2) ja võõrkeele õpetamisel ning õppijate keeleteadmisi hinnates, sõnastab keeletasemete kirjeldused väga üldiselt ja funktsionaalselt. C1-tasemega ehk vilunud keelekasutaja oskab raamdokumendi üldskaala järgi “luua selget, loogilist, üksikasjalikku teksti keerukatel teemadel, kasutades sidusvahendeid ja sidusust loovaid võtteid.” Kirjutamisoskuse juures on osaoskuste all välja toodud suutlikkus kirjutada keerukal teemal selgeid, hea ülesehitusega tekste, rõhutada olulist, selgitada oma seisukohti ning kasutada loovtekstides loomulikku lugejaga arvestavat stiili. Samuti on mainitud, et C1-tasemega õppija valdab rikkalikku sõnavara. (Raamdokument 2007: 25, 64–66) Aga kui varieeruv on rikkalik sõnavara ja mis muudab keelekasutuse loomulikuks?

Vajadust avada eesti keele tasemeoskuste tegelik lingvistiline sisu – tuua välja igale tasemele omased keelestruktuurid ja sõnavara – on teiste seas esile tõstnud Pille Eslon jt (2010) ning Mare Kitsnik (2014)¹. C1-taseme eesti õppijakeelt ei ole seni spetsiaalselt uuritud. Eesti õppijakeele alusel on lähemalt vaadeldud B1- ja B2-taset, keskendudes verbivormidele (Kitsnik 2014). Tasemeid eristamata on kirjeldatud eituse väljendamist (Kitsnik 2007) ning *ma-* ja *da-*infinitiivi kasutust (Šmõreitšik 2009), samuti võrreldud eesti kirjakeele ja õppijakeele tekstikasutusmustreid (nt Eslon 2010, 2014b, 2014c).

Krista Kerge jt (2014) on mõõtnud sõnavara rikkust ja ulatust erineva keeleoskustasemega õppijate ja haritud emakeelekõnelejate esseedes ning leidnud, et emakeelekõnelejatega võrreldes on C1-tasemega õppijatel 1000 sagedaima sõna osakaal palju suurem ja 4000 sagedaima sõna hulka mitte kuuluvate sõnade osakaal väiksem. Üldine lekseemide ja sõnade suhe ei erinenud valimites statistiliselt olulisel määral.

¹ Inglise ja soome keele omandamist eri tasemetele omaste lingvistiliste tunnuste kaudu on uuritud nt CEFLING-projekti “Linguistic Basis of the Common European Framework for L2 English and L2 Finnish (2007–2009)” käigus (Martin jt 2010).

Siinse uurimuse objekt on C1-tasemega eesti keele õppija ja haritud emakeelekõneleja kirjalik keelekasutus. Eesmärk on võrrelda K2-le ja K1-le omaste keelekasutusmuustrite struktuure, nende morfosüntakti- list ja leksikaalset varieeruvust ning tuua esile kahe keelekasutusvariandi ühis- ja erijooni. Alameesmärk on avada klasteranalüüsi potentsiaali keeleoskustasemete lingvistilise sisu määramisel.

Lähtun arusaamast, et verb on lause keskne komponent, mis- tõttu olen analüüsimiseks valinud verbist paremal ja vasakul paikneva konteksti. Artiklis kirjeldan verbiga algavaid neljasõnalisi struktuure ehk tetragramme (nt *valib oma meeskonda töötajaid*), mis koonduvad komponentide sõnaliigi alusel klastritesse (vastavalt verb-pro- noomen-substantiiv-substantiiv, lühendatult VPSS). Struktuuride esi- ja lõpukomponendi sõnaliigi põhjal (siinjuhul *valib* ja *töötajaid*) saab need omakorda jagada klassidesse (vastavalt verb-substantiiv ehk V-S), kus tetragrammi kaks keskmist komponenti (siinjuhul *oma meeskonda* ehk pronoomen-substantiiv ehk PS) varieeruvad, erista- des klastreid (nt *valib välja häid töötajaid* ehk verb-adverb-adjektiiv- substantiiv ehk VDAS). Sel viisil saab välja tuua K1 ja K2 üldised keelekasutusmuustrid ning neid omavahel võrrelda. Lähemalt vaatlen K1 ja K2 kuue põhiklassi sagedamaid sõnaliigijärgendeid ning esitan ten- dentsid, mida eelistatuimate verbialguliste klastrite kõrvutav analüüs aitab välja tuua. Analüüsi keskmis on see, kuidas omavahel seondu- vad sõnaliigijärgendite struktuur, selle leksikaalne ja morfosüntaktiline varieerumine.

Oma uurimuses järgin kasutuspõhisuse printsiipi, mis ei käsitle keelt kindlaksmääratud ega staatilise, vaid kasutuse käigus pidevalt muutuva nähtusena, mida mõjutavad keelekasutaja kogemused. Iga kõneleja lingvistiline pädevus rajaneb tema senisel keelelisel sisendil ja keelekasutuskogemusel, mille seaduspärade põhjal saab teha statistilisi üldistusi. (Bybee 2013: 50; Ellis 2015: 5, 23) Nick C. Ellis (2002) on välja toonud, et keeleüksuste ja nende koosesinemise sagedus sisendis mõju- tab õppija keeletöötlust ja -kasutust kõigil keelestruktuuri tasanditel foneemist fraasini.

Keeleõppija on Ellise (2006a: 1–2) järgi intuiitiivne statistik, kes vaeb erinevate tõlgenduste tõenäosusi ja prognoosib, millised konstruktsioonid on mingis kontekstis oodatavad. Keele omandamine tähendab info kogumist vormi ja funktsiooni vastavuste suhteliste sageduste kohta. Nii siis omandab õppija kiiremini ning kasutab ladusamalt sagedasi vorme ja regulaarseid mustreid. Konstruktsioonid on lihtsamini õpitavad, kui neis sisaldub suurima sagedusega lekseeme, sh üldise tähendusega tuumverbid (nt *saama*, *minema*), kui konstruktsiooni komponendid on semantiliselt sidusad ning vormide ja funktsioonide koosinemise tõenäosus suur (Ellis jt 2015: 169).

Keelekasutuse loomulikkus seostub valikute ratsionaalsuse ja optimaalsusega nii keele mõistmisel kui ka produtseerimisel. Erinevalt K1-st ei ole K2 omandamine samavõrd ratsionaalne: hoolimata ohtrast loomuliku keele sisendist ja keelekasutuskogemusest ei pruugi õppija keelelised valikud kokku langeda emakeelekõneleja valikutega. K2 keelekasutus ei esinda teatud vorme ja nende tekstilisi funktsioone sihtkeelele omasel määral. (Ellis 2006a: 1–2, 9) Mõjutegurite hulgas on esile toodud võimalik ülekanne K1-st, alakasutatud keeleüksuste vähene olulisus või liiasus sõnumi edasi andmise seisukohast ning teatud keelenähtuste ja struktuuride vähene esilduvus, nt seotud morfeemid ja funktsiooni-sõnad kipuvad suure sageduse tõttu fonoloogiliselt ümbrusega sulanduma, kuna need ei ole rõhulise intonatsiooniga. (Ellis 2006a: 19, 2006b: 170–171) Esilduvust mõjutab ka sõnajärg: lauselõpuline positsioon on esildavam kui lausekeskne (Bassano jt 2005: 72).

Parimaks allikaks keelestruktuuride sageduste tuvastamisel peetakse tekstikorpust ja korpusanalüüsi (McEnery & Wilson 1996: 12), milles eristatakse korpuspõhist ja korpusest tulenevat suunda. Siinses uurimuses on need kaks lähenemist integreeritud. Ühelt poolt lähtun hüpoteesist, et ka vilunud keelekasutaja tasemele jõudnud K2 kõneleja keeleline väljendus on võrreldes K1 kõnelejaga morfosüntaktiliselt nihestunud – eelistatakse osaliselt erinevaid struktuure ja vorme – ning sõnavara on teatud juhtudel piiratum. Teisalt ilmneb alles korpusanalüüsi käigus see, millised sõnaliigijärjendid, vormi- ja funktsioonielistused on

C1-tasemega õppijale ja emakeelekõnelejale tegelikult omased ning mille poolest on valikud lähedased, mille poolest erinevad.

2. Materjal ja metoodika

2.1. Uurimisaines

Uuritav keeleaines pärineb Tallinna Ülikooli eesti vahekeele korpuse (EVKK)² tuumkorpuse esseedest ning Eesti suurimate ajalehtede Postimehe ja Õhtulehe võrguväljaandes 2014. aasta 1.–31. detsembrini ilmunud arvamusalustest, mille hulgas on vähemal määral ka Sirbis, Õpetajate Lehes ja Linnalehes avaldatud kirjutisi. Arvamuslood olen koondanud EVKK alamkorpuse K1 referentskorpus.

Analüüsisin EVKK esseid, mille kolm eesti keele kui K2 kodakondsus- ja riigieksamite parandamise kogemusega eksperti on vastavalt Euroopa Nõukogu keeleoskustasemete süsteemile hinnanud C1-tasemel olevaks. Need esseed pärinevad EVKK tuumkorpusest, kuhu on kogutud valdavalt Ida-Virumaal ja Harjumaal elavate vene emakeele taustaga inimeste kirjutisi, ja eesti keele olümpiaadi tööde alamkorpusest, mis sisaldab eesti keele kui K2 olümpiaadi töid (Eslon 2014a: 438–439).

Kokku analüüsisin 108 teksti, mille maht on 74 342 sõnet, kõrvale on jäetud pealkirjad ning võõrkeelsed fraasid-laused (nt *Я финский студент*). Tekstide keskmine pikkus on 688,4 sõnet. Õppijatest on 80,6% vene emakeelega, 15,7% soome emakeelega, rootsi keel ja jidiš on esimeseks keeleks ühel autoril. 38% informantidest on 17–18-aastased keskkooliõpilased, 51,9% on 19–26-aastased, 6,5% on 27–40-aastased ja üks on 41-aastane või vanem. Naiste osakaal on 71,3% ja meeste osakaal 26,9%. Enim on autorite seas keskharidusega eesti keele õppijaid (58,3%), põhiharidusega informantide osakaal vastab gümnaasistide omale, kahe teksti autor on kõrgharidusega. Kahel juhul on autori esimene keel, vanus, sugu ja haridustase teadmata.

² Vt <http://evkk.tlu.ee> (31.08.2016).

Võrdlusaluseks õppija keelekasutusele arutlevat laadi loomingulistes kirjutistes valisin emakeelekõneleja etaloni ehk kõrgharitud mittefiloogi oma. Kuna keelenorm on kokkuleppeline, ei pruugi see kajastada haritud emakeeleveldajale aktsepteeritavat loomulikku keelekasutust, mida Ute Römer on nimetanud keeleõppija ideaalseks sisendiks. Ka koolis õpetatav keel ei sobi õppija keeleoskustaseme mõõdupuuks, nagu Römer on näidanud inglise keele modaalverbide kasutuse võrdlemisel õpikukeeles ja emakeeleveldajate suulises kõnes, samuti *if*-tingimuslausetel võrdlemisel nende tegelikus sisendis ehk inglise keele õpikus, briti inglise keele koondkorpusest pärit tekstides ja suulises kõnes ning saksa emakeelega õppijate ingliskeelsetes esseedes. Römer on järeldanud, et õpikukeel ei ole vastavuses tegeliku keelekasutusega, ja oletanud, et siin võib olla üks põhjusi, miks on isegi vilunud keelekasutaja tasemega K2 kõnelejal *if*-lausetega probleeme. (Römer 2004a, 2004b, 2007).

Keele loomulikkuse mudeleid välja töötades ja mh kõrgtasemega eesti keele õppijate sõnavara ulatust uurides on K1 loomulikkuse etaloniks seatud kõrgharitud, kuid mitte filoloogiharidusega emakeelekõnelejate spontaanne kõne ja enesekontrolli all kirjutamine (Pajupuu jt 2009; Pajupuu jt 2010; Kerge jt 2014). Ideaalina, mille suunas peaks püüdlema nii K1 kui ka K2 õppija ning mille alusel tuleks keeleoskust mõõta, ei nähta seega standardkeelt ehk normikeelt, vaid haritud emakeelekõneleja suulist ja kirjalikku keelekasutust formaalsemates situatsioonides, nagu ettekanne või essee.

Arvamuslood on küll toimetatud, ent stiililt võimalikult autoritruud. Peamiselt korrigeerivad toimetajad ja keeleteoimetajad ortograafiavigu, mida sinne uurimus ei vaatle. Arvamusavaldused, mida paberlehe tarbeks on kärbitud, ilmuvad veebis sageli täispikana. (Postimehe arvamusportaali toimetaja Liisa Tageli isiklik teade 08.12.2015, Õhtulehe arvamustoimetaja Laur Uudami isiklik teade 14.12.2015)

Kõrgharidusega ja mittefiloloogist emakeelekõnelejate tekste ilmus määratud ajavahemikus Postimehe ja Õhtulehe veebikülje arvamusrubriigis vastavalt 107 ja 78, kokku 185. Nende maht on 107 203 sõnet ning tekstide keskmine pikkus 579,5 sõnet. Kõrvale jätsin pealkirjad

ja alapealkirjad, toimetuse lisatud sissejuhatavad lõigud ja võõrkeelsed väljendid (nt *Arbeit macht frei*). Teksti autori emakeele, haridustaseme ja filoloogihariduse olemasolu selgitasin välja toimetuste abiga. Enamik autoritest (72,6%) on vähemalt 41-aastased, 27–40-aastasi on autorite seas 24,7% ja kuni 26aastasi 2,2%. Ühe autori vanus on teadmata. 72,0% tekstidest on meesterahva ja 27,4% naisterahva kirjutatud, ühel arvamusel on kaks eri soost autorit.

Kuigi võrreldavate tekstikogude maht on erinev, ei ole vormide esinemus korrelatsioonis sõnede arvuga (vt nt Biber jt 1998: 32–34): võrdlus põhineb tetragrammide osakaaludel. Kuna K2 valim on moodustatud C1-taseme alusel, ei oma võrreldavuse seisukohast tähtsust, et osa informantidest on täiskasvanud ja osa õpilased.

2.2. Klasteranalüüs ja statistiline olulisus

Laiemas mõttes olen rakendanud kontrastiivset vahekeele analüüsi (ingl *Contrastive Interlanguage Analysis* ehk CIA), mis kõrvutab K2 omandaja ja sihtkeele kui K1 kõneleja keelekasutust, tuginedes õppijakeele korpusele ja K1 referentskorpusele (vt Granger 1996, 2015). Keelekasutusmuustrite leidmiseks kasutasin klasteranalüüsi Sander Otsa (2012) programmeeritud Klastrileidja abil, mis otsib tekstist sarnaseid morfo- ja süntaksimärgendite lineaarseid järjendeid. Tekstid eelmärgendasin eesti keele süntaksianalüsaatoriga EstCG 1,0 (Müürisep 2000), enne Klastrileidjasse sisestamist ühestasin märgenduse käsitsi.

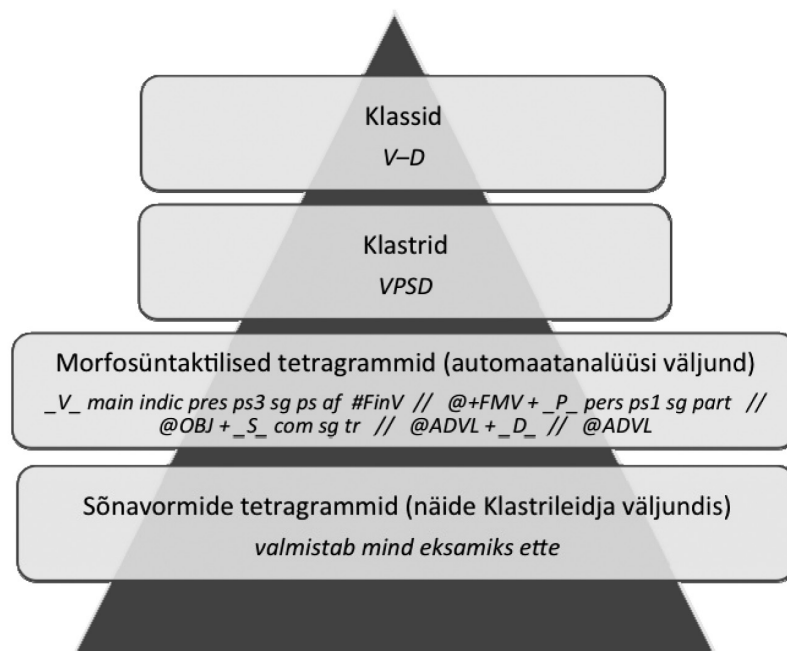
Klastrileidja võimaldab otsida morfoloogilisi ehk vormijärjendeid, süntaktilisi ehk lauseliikmete järjendeid ja morfosüntaktilisi järjendeid, mille osiste vorm ja funktsioon kattuvad (Ots 2012: 8–9). Programm koondab järkjärgulise liitmise teel morfoloogiliselt, süntaktiliselt või morfosüntaktiliselt sarnaseid sõnavormide n-gramme. Arvutilingvistikas tähistab n-gramm n arvu keeleliste elementide (nt sõnavormide, lemmade, foneemide või silpide) jada, mis on leitud etteantud tekstikogust (Aggarwal & Reddy 2014: 382). Klastrileidja tuvastab n-gramme põhimõttel sõna1 sõna2 sõna3 sõna4; sõna2 sõna3 sõna4 sõna5 jne.

Sama meetodi abil on seni vaadeldud kolmest komponendist koosnevaid üksusi ehk trigramme. Kirjeldatud on eesti ilukirjanduskeele üldist leksikaalgrammatilist varieerumist (Trainis & Allkivi 2014) ning samatähenduslike analüütiliste ja sünteetiliste verbide kasutusmustreid (Paeoja 2015; Eslon & Paeoja 2015), samuti võrreldud verbist vasakule jäävat konteksti ilukirjandus- ja õppijakeeles, tuues välja kasutusmustrite sarnasust ja erinevusi (Eslon 2014b) ning keskendudes adverbi sisaldavate struktuuride tekstifunktsioonidele (Eslon 2014c). Soome keelt uurides on klasteranalüüsi rakendanud Ilmari Ivaska (2015), otsides morfomärgendite alusel mono-, bi- ja trigramme, mille kasutussagedus eristab K1 ja K2 kõnelejaid ning erineva emakeelega soome keele õppijaid. Scott Jarvis ja Magali Paquot (2012) on võrrelnud eri pikkusega sõnavormijärjendite (nt *going to*, *I think that*) kasutust erineva emakeele taustaga õppijate inglise keeles.

Käesoleva uurimuse tarvis analüüsisin nelja komponendiga sõnavormijärjendeid ehk tetragramme (nt *valmistab mind eksamiseks ette*) ning rühmitasin need morfoloogilise ja süntaktilise samalaadsuse alusel (vastavalt põhiverb indikatiivis ainsuse 3. pöörde preesensis predikaadina + ainsuse 1. isiku pronoomen ainsuse partitiivis objektina + üldnimeline substantiiv ainsuse translatiivis adverbialina + adverb adverbialina, mille automaatse morfosüntaktilise analüüsi väljund on `_V_ main indic pres ps3 sg ps af #FinV // @+FMV + _P_ pers ps1 sg part // @OBJ + _S_ com sg tr // @ADV_L + _D_ // @ADV_L`).

Tetragrammist üldisem aste klassifikatsioonis on klaster, mis koondatab sõnaliigijärjendilt samalaadseid tetragramme. Klastreid nimetatakse morfoloogiliste sõnaliigimärgendite alusel (nt VPSD ehk verb-pronoomen-substantiiv-adverb). Olen need leidnud Klastreidija väljundi töötlemisel Microsoft Excelis, kus sorteerisin tetragramme komponentide sõnaliigilise kuuluvuse järgi. Klastrid on algus- ja lõpukomponendi alusel jagatud üldisematesse klassidesse, kus varieerub keskmiste komponentide sõnaliik (nt klassi V–D kuuluvad teiste seas klastrid VPSD ja VSSD).

Lingvistilise klasteranalüüsi hierarhiline liigitus võib uurimusest sõltuvalt varieeruda, siin kasutatud hierarhia on välja toodud joonisel 1.



JOONIS 1. *Klasteranalüüsi hierarhia*

Leitud tetragrammidest jätsin kõrvale ainukordsed klastrid, mis ei sarnane osiste sõnaliigi poolest ühegi teise tetragrammiga. Kokkuvõttes kätkeb K1 valim 11 437 ning K2 valim 8364 verbialgulist tetragrammi.

Klasside kirjeldamisel olen arvestanud optimaalse objektide ehk tetragrammide hulgaga klastris, mis on leitud valemi $k \approx n2$ järgi (Mardia jt 1979). N tähistab valemis kõigi objektide arvu vastavas valimis. Valemi põhjal saab K1 valimis statistiliselt oluliseks pidada klastreid, milles tetragrammide hulk on vähemalt 76, ning K2 valimis klastreid, kuhu kuulub vähemalt 65 tetragrammi. Võrrelnud olen niisuguste klastrite kasutust, mis on vähemalt ühes valimis optimaalsel määral esindatud, kuna osa klastreid, mis mahuvad ühes valimis optimaalse esinemissagedusega klastrite hulka, jäävad teises valimis allapoole määratud künnist. Tetragrammide tasandil analüüsisin klasside sagedaimat ehk invariantset klastrit kummaski valimis. Klasside osakaalud arvutasin kõigi verbialgulist klasside sageduse alusel, klastrite osakaalud neid hõlmava klassi sageduse põhjal. Oluline on aga teha vahet, millal erinevad osakaalud olulisel määral ja millal mitte.

Klasside ja klastrite statistilise sarnasuse või erinevuse esiletoomiseks kasutasin olulisustestina χ^2 testi 2 x 2 tabelis, mida nt Eslon ja Matsak (2009: 87–88) on rakendanud käändevormide esindatuse võrdlemisel. Siinses uurimuses näitab iga klassi kohta koostatud tabel kahes valimis selle klassi sageduse suhet kõigi ülejäänud klassidesse kuuluvate tetragrammide arvuga ning iga vaadeldava klatri kohta koostatud tabel näitab kummaski valimis selle sageduse suhet sama klassi ülejäänud klastritesse kuuluvate tetragrammide hulga (vt tabel 1).

TABEL 1. 2 x 2 tabeli näide klasside ja klastrite puhul

	Tetragrammide arv klassis / klastris	Tetragrammide arv ülejäänud klassides / sama klassi ülejäänud klastrites
K1 arvamuslood	a	b
K2 esseed	c	d

χ^2 arvutamiseks kasutasin valemit, kus $n = a + b + c + d$:

$$\chi^2 = \frac{(n - 1)(ad - bc)^2}{(a + b)(a + c)(b + d)(c + d)},$$

χ^2 -jaotuse kriitiline väärtus, millest alates võib erinevust lugeda statistiliselt oluliseks, sõltub vabadusastmete arvust (df) ja olulisusnivoost (α) ehk maksimaalsest eksimise tõenäosusest, milleks on siinjuhul valitud 5%. Kui vaatluse all on objekti (antud juhul klassi või klatri) kaks tunnust, millel on kummalgi kaks väärtust, siis vabadusastmete arv $df = 1$. Tulemus on saadud valemi $(m - 1)(k - 1)$ järgi, kus m on ühe tunnuse erinevate väärtuste hulk ja k teise tunnuse erinevate väärtuste hulk. χ^2 kriitiline väärtus $df = 1$ ja $\alpha = 0,05$ korral on 3,841. (Kaart 2004) Sellest olen lähtunud ka käesolevas uurimuses klasside ja klastrite esindatuse võrdlemisel K1 ja K2 valimis. Kui χ^2 väärtus ületab kriitilist väärtust, siis on olulisustõenäosus (p) väiksem kui olulisusnivoo ehk väiksem kui 0,05 ning seega on erinevus tõestatud.

Klasside invariantsetesse klastritesse kuuluvate tetragrammide kõrvutus arvestab morfosüntaktilist ja leksikaalset varieerumist ning selle

piire. Olen vaadelnud õppijate ja emakeelekõnelejate eelistusi vormide ja funktsioonide kombineerimisel. Kui mõni nii vormilt kui ka funktsioonilt samalaadsete tetragrammide kogum teiste seas esile tõuseb, siis toon välja ka invariantse morfosüntaktilise struktuuri.

Klastrite leksikaalset rikkust saab hinnata ainukordsete lekseemide osakaalu põhjal kõigi sõnede hulgas. Mida suurem on ainukordsete lekseemide osatähtsus, seda avaram on sõnavara, sest seda vähem leidub korduvaid lekseeme (vt nt Jockers 2014: 59). Oma uurimuses vaatlesin leksikaalset varieeruvust klastrite iga komponendi puhul eraldi, nt arvasin ainukordsete verbide ja kõigi verbikasutuste suhte.

3. Üldised sõnaliigimustrid: verbiaalgulised klassid ja klastrid

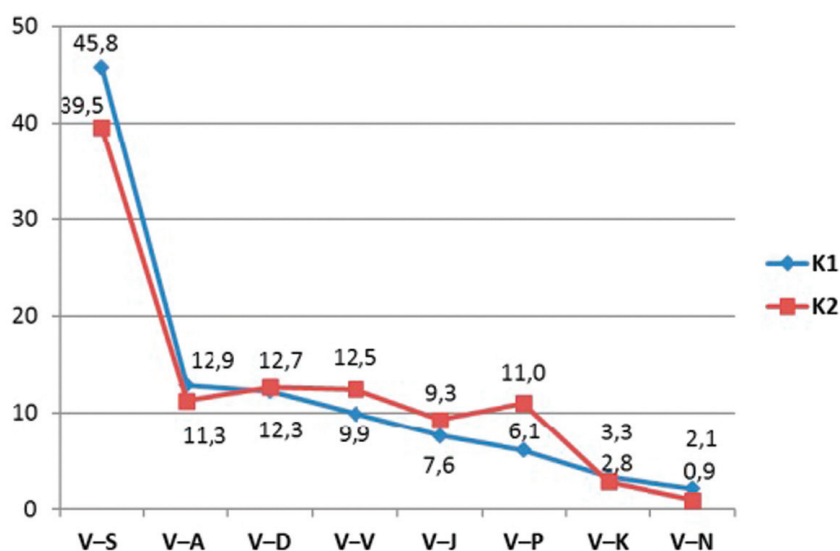
Tetragrammide esi- ja lõpukomponendi sõnaliigilisel samalaadsusel põhinevate klasside võrdlus rajaneb nende osakaalul K1 ja K2 valimites (vt joonis 2) ning χ^2 testi andmetel (vt tabel 2, kus on varjutatud klastrid, mille puhul on χ^2 suurem kui kriitiline väärtus 3,841).

Nii emakeelekõnelejate arvamused kui ka C1-tasemega eesti keele õppijate esseedes on verbiga algavate tetragrammide seas ülekaalukalt sagedaimad substantiivilõpulised tetragrammid ehk klass V–S, mille osakaal on emakeelekõnelejal 45,8% (nt *avastas Marsi pinnalt kraatri*) ja õppijatel 39,5% (nt *on igal inimesel õigus*).

K1 tekstiloomes järgnevad sageduselt adjektiivilõpulised tetragrammid (klass V–A), veidi väiksema sagedusega esinevad adverbilõpulised tetragrammid (klass V–D). Harvem tuleb ette verbiaalgulisi tetragramme, mis lõppevad verbiga (klass V–V). Suhteliselt väike on konjunktsiooni- (V–J) ja pronoomenilõpulise klassi (V–P) esinemus. Vähimal määral on esindatud adpositsiooni- (klass V–K) ning numeraalilõpulised tetragrammid (klass V–N).

K2 valimis on klasside osakaalud ühtlasemad. Adverbi- (V–D) ja verbilõpuline klass (V–V) on sagedamad kui adjektiivilõpuline (V–A), mis on K2 kõnelejate keelekasutuses sageduselt alles neljandal kohal. Pronoomenilõpulise klassi (V–P) kasutus on erinevalt K1-st sagedam

kui konjunktsioonilõpulisel klassil (V–J). Analoogselt arvamusaluselise pruugitakse kõige harvem adpositsiooni- ja numeraalilõpulisi klasse V–K ja V–N.



JOONIS 2. Verbialguliste klasside osakaal (%) K1 ja K2 valimis

Suuremad erinevused emakeelekõneleja ja vilunud keeletekasutaja vahel ilmnevad klassis V–S, mille kasutus on tunduvalt sagedam K1 kõnelejal (osakaalude vahe K2-ga 6,3%, χ^2 väärtus 77,975), ja klassis V–P, mille tähtsus on oluliselt suurem K2 kõnelejal (osakaalude vahe K1-ga 4,9%, χ^2 väärtus 151,787). Mõlemas valimis on samal määral esindatud vaid klassid V–D (osakaalude vahe 0,4%, χ^2 väärtus 0,724) ja V–K (osakaalude vahe 0,5%, χ^2 väärtus 3,502). Ülejäänud klasside V–A, V–V, V–J ja V–N puhul on osakaalude erinevus pigem väike (1,2%–2,6%), kuid statistiliselt oluline.

Klasside võrdlusest ilmneb, et K1 ja K2 kõnelejate morfosüntaktilised valikud on mitmeti erinevad. Erinevus eelistustes tuleb veelgi selgemalt esile iga klassi invariantsete klastrite võrdlusest (vt tabel 3).

K1 valimis on suurima osakaaluga tetragrammid, kus verbile järgneb kaks substantiivi. Peaaegu kõigis klassides on sagedaim klaster substantiivikeskne, nt klassis V–D on eelistatuim klaster VSSD (*sõltub inimese tervis väga*). Erandiks on vaid pronoomenilõpuline klass, mille

puhul invariantstes klastris järgnevad algusverbile teine verb ja adverb (nt *on jäänud vaid mõned*).

K2 esseedes seevastu on verbialguliste tetragrammide keskseteks komponentideks kõige sagedamini pronoomen ja substantiiv, mis järgnevad verbile koguni viie klassi invariantstes klastris, nt klassis V–J tõuseb esile klaster VPSJ (nt *satub sellisesse seltskonda ja*). Kahel juhul, adjektiiv- ja pronoomenilõpulisel klassis, on invariantse klastri tuumaks verb ja sellele järgnev adverb, nt (*ei ole tihti kindel, peaks mainima ka seda*). K1 ja K2 valimi invariantssed klastrid kattuvad üksnes pronoomeni- ja numeraalilõpulisel tetragrammide puhul.

Enamasti moodustavad invariantssesse klastrisse kuuluvad tetragrammid kõigist sama klassi tetragrammidest umbes kümnendiku, väiksema klastrivarieerumisega klasside puhul ligi viiendiku või isegi neljandiku. Mida suurem on invariantse klastri osakaal, seda ebahõlpsamad on klastrite osakaalud klassis, ja vastupidi. Nt klasside V–S, V–A ja V–J puhul võib õppijate tekstides täheldada klastrite ühtlasemat jaotumist. Seda, et K2 kõnelejal ei tõuse eelistatud sõnaliigimustrid nõnda selgelt esile ja klasside-klastrite osakaaluvahed on väiksemad kui K1 kõnelejal, on täheldanud ka Eslon (2014b), kes on kõrvutanud õppijakeelt ja ilukirjanduskeelt verbilõpulisel trigrammide alusel ning põhjendanud K2 suuremat variatiivsust sellega, et õppija ei kombineeri keelendeid sama ratsionaalselt ja ökonoomselt kui emakeelekõneleja.

Võttes arvesse kummagi valimi kõiki optimaalse sagedusega klastreid (tervikliku ülevaate saamiseks vt Allkivi 2016), selgub, et emakeelekõnelejad eelistavad üleüldiselt verbi lähemas ümbruses substantiivi- ja adjektiivikeskseid mustreid, s.o täistähenduslikke sisusõnu, õppijad aga kasutavad palju pronoomenikeskseid mustreid, s.o mittetäistähenduslikke sõnu. Nt adverbilõpulisel klassis V–D tarvitab õppija sagedamini klastreid VPDD ja VVPD ning harvemini klastreid VSSD ja VASD, konjunktsioonilõpulisel klassis V–J kasutab õppija sagedamini klastrit VPSJ, kuid harvemini klastreid VSSJ ja VASJ.

TABEL 2. Verbialguliste klasside kasutussagedus K1 ja K2 valimis

Klassid	K1 arvamuslood		K2 esseed			
	a	b	c	d	χ^2	p
V-S	5232	6205	3300	5064	77,975	0,000
V-A	1480	9957	948	7416	11,584	0,001
V-D	1410	10 027	1065	7299	0,724	0,395
V-V	1133	10 304	1042	7322	32,169	0,000
V-J	863	10 484	781	7583	20,375	0,000
V-P	701	10 736	919	7445	151,787	0,000
V-K	376	11 061	236	8128	3,502	0,061
V-N	242	11 195	73	8291	47,688	0,000

TABEL 3. Verbialguliste klasside invariantseid klastrid K1 ja K2 valimites

K1 arvamuslood				K2 esseed			
Klass (tetragramme)	Invariantne klaster	Sagedus	Osa-kaal	Klass (tetragramme)	Invariantne klaster	Sagedus	Osa-kaal
V-S (5232)	VSSS	544	10,4%	V-S (3300)	VPSS	271	8,2%
V-A (1480)	VSSA	157	10,6%	V-D (1065)	VPSD	129	12,1%
V-D (1410)	VSSD	167	11,8%	V-V (1042)	VPSV	97	9,3%
V-V (1130)	VSSV	113	10,2%	V-A (948)	VVDA	84	8,9%
V-J (863)	VSSJ	160	18,5%	V-P (919)	VVDP	79	8,6%
V-P (701)	VVDP	57	8,1%	V-J (781)	VPSJ	137	17,5%
V-K (376)	VSSK	90	23,9%	V-K (236)	VPSK	61	25,9%
V-N (242)	VSSN	27	11,2%	V-N (73)	VSSN	18	24,7%

Avaram pronoomenikasutus on eesti õppijakeele klasteranalüüsi põhjal esile tulnud varemgi (Eslon 2013) ning iseloomustab nähtavasti ka C1-taseme kirjalikku keelekasutust.

4. C1-taseme eesti keele kasutuse morfosüntaktilisi ja leksikaalseid tunnusjooni

Eelistatumate verbialguliste klastrite analüüs tõi morfosüntaksi ja leksika tasandil esile mõningad tendentsid, mis on C1-tasemel omandatud K2 keelekasutusele iseloomulikud. Analüüsi käigus vaatlesin kummaski valimis lähemalt kuue sagedaima klassi invariantset klastrit: klastreid VSSS (K1) ja VPSS (K2), VSSA (K1) ja VVDA (K2), VSSD (K1) ja VPSD (K2), VSSV (K1) ja VPSV (K2), VSSJ (K1) ja VPSJ (K2) ning VVDP (langes K1-s ja K2-s ainsana kokku).

Tehes invariantsete klastrite alusel üldistavaid järeldusi morfoloogiliste, süntaktiliste ja leksikaalsete eelistuste kohta, tuleb silmas pidada, et klasside V–S, V–A, V–D, V–V ja V–J puhul on analüüsitud erinevaid struktuure. Eri struktuurid on aga seotud erinevate vormi-funktsiooni seostega ja leksikaalse varieerumisega. Saamaks aimu, kas K1 ja K2 klasterkasutuses ilmnenu erinevused tulenevad pigem struktuuri või õppijate keeleliste valikute iseärasustest, on tähtis vaadata, kuidas kasutab emakeelekõneleja K2-s eelistatud struktuure ning kuidas kasutab õppija struktuure, mida eelistab K1 kõneleja.

Analüüsi põhjal saab välja tuua järgmised vilunud keelekasutaja tasemega eesti keele õppijale iseloomulikud jooned.

1) Õppija kasutab rohkem eitavat kõnet.

Eitava kõneliigi rohkus tuleb esile adjektiiv- ja pronoomenilõpulisel klassis. Klaster VVDP, mis on nii K1 kui ka K2 valimis invariantne klaster, soosib verbi liitvormide, sh eitusvormide kasutust (nt *ei usu mitte kedagi*), ent kui K1-s algab eitusvormiga pisut vähen kui pool tetragrammidest – 47,4%, siis K2-s tervelt 62,0%. Erinevalt emakeelekõnelejast eelistab õppija ka adjektiivilõpulistest struktuuridest selliseid, mis sisaldavad eitavat kõneliiki. See, et K1-s esineb eitav kõneliik klastris VSSA üksikjuhtudel

(1,9%, nt *pole kohtute statistika avalik*), kuid on K2 klastris VVDA jaatava kõneliigiga peaaegu võrdselt esindatud (46,4%, nt *ei ole enam võõras*), on struktuurist tulenev loogiline erinevus. Kui aga heita pilk VVDA-klastrile K1 valimis, selgub, et emakeelekõneleja kasutab eitust vähemal määral (36,8%) ja selle suurem osatähtsus on siiski iseloomulik K2 kõnelejale.

Verbi eitusvormide ohtrama kasutuse on õppijakeele eripärana välja toonud ka Eslon (2014b, 2014c), kes täheldas EVKK tuumkorpuse ja ilukirjandustekstide verbilõpulist trigrammide võrdluse põhjal, et eitav kõneliik on teatud klassides ja klastrites K2-le omasem.

C1-tasemega K2 kõneleja lähedus K1 kõnelejale tuleb esile adverbilõpulise klassi invariantsetes klastrites VSSD (K1) ja VPSD (K2), kus verbi eitusvormi finiiitkomponendiga algab vastavalt 7,2% ja 7,8% tetragrammidest, nt *<ei> pidanud riigikohtu juht järelikut, polnud selles poes harilikult*. Eelistatuimates substantiivilõpulistest klastrites VSSS (K1) ja VPSS (K2) ei esine eitavat kõnet ühtviisi peaaegu üldse. Selles peegeldub ka nende klasside eelistatud struktuuride sarnasus verbikasutuse seisukohast.

2) Õppija kasutab vähem mineviku liitajavorme.

Erinevus avaldub eelkõige verbi- ja pronoomenilõpulisel klassis. Kui emakeelekõnelejal sisaldab klatri VVDP tetragrammidest liitajavormi 15,8% (nt *on jäänud vaid mõned*), siis õppijal 8,9% (nt *on öelnud ka seda*). K1 invariantsetes verbilõpulisel klastris VSSV moodustavad algus- ja lõpuverb perfekti- või pluskvamperfektivormi (nt *on meeste haridustase tõusnud*) sagedamini kui K2 invariantsetes klastris VPSV (nt *oli selliste suhetega seotud*), osakaalud on vastavalt 19,8% ja 15,5%. Küsimus ei ole aga selles, et VPSV-struktuur liitaegade kasutust rohkem piiraks, sest K2 kõneleja lihtsalt ei kasuta neid liitvorme oluliselt sagedamini, seda ka klatri VSSV (16,4%). Samas raamivad abiverb *olema* ning *nud-* või *tud-*partitsiip emakeelekõnelejal klatri VPSV tetragramme lausa 34,6%-l juhtudest. K1 valimist lähtudes peaks niisiis hoopis järjend VPSV tingima sagedama liitaegade kasutuse.

Reili Argus (2006: 13) on välja toonud, et eesti keele ajakategooria morfoloogilise markeerituse hierarhiast (Ehala 2009: 45) lähtudes võib

eeldada, et K1 omandamisel õpitakse kõigepealt kasutama preesensit, siis imperfekti ning alles seejärel analüütiliselt moodustatud perfekti ja pluskvamperfekti. Markeeritus, mis on seotud vormide üldise sageduse ja morfosüntaktilise keerukusega, võib mängida rolli ka eesti keele kui K2 omandamisel. Nt inglise ja saksa keele minevikuaegade omandamist uurides on leitud, et imperfekt ilmub õppijakeelde liitaegadest varem (Bardovi-Harlig 2000; Klein 1986).

Teisalt põhjustab vene emakeelega eesti keele õppijaile – enamus siinse uurimuse informantidest – mineviku ajavormi valikul raskusi see, et tänapäeva vene keeles ei ole verbidel analüütilisi minevikuvorme. Eesti keele kolmele minevikuajale vastavad vene keeles imperfektiivse aspekti minevik ja perfektiivse aspekti minevik, aspektipaare moodustatakse enamjaolt liidete abil (nt *рассказывать* ja *рассказать* ‘rääkima, jutustama’), kuid mineviku tunnus *л* koos soo- ja arvutunnustega (*-л, -ла* ja *-ли*) langevad neil ühte (nt *рассказывал(а, и)* ja *рассказал(а, и)*). Perfektiivse aspekti minevikuvormi vasteks võib eesti keeles olla nii imperfekt (nt *он рассказал нам следующую историю* ‘ta jutustas meile järgmise loo’), perfekt (nt *она мне рассказала, что* ‘ta on mulle rääkinud, et’) kui ka pluskvamperfekt (nt *когда они всё рассказали* ‘kui nad olid kõik ära rääkinud’). Pool ja Vaimann (2005: 134) on eesti keelt K2-na kõnelevate vene üliõpilaste vigu analüüsisides täheldanud, et sünteetilist imperfektivormi üldistatakse analüütiliste minevikuaegade vastena. Et imperfekti eelistatakse mineviku liitaegadele, eriti perfektile, tuli ilmsiks ka Maisla (2014: 65–73) analüüsitud materjalist.

3) Õppijal järgneb verbile harvem subjekt.

Emakeelekõneleja invariantset verbialgulised klastrid sisaldavad tunduvalt sagedamini subjekti. Nt K1 klastris VSSS sisaldab noomenit subjekti rollis 28,6% (nt *saabusid platsile dressides mehed*), K2 klastris VPSS 19% tetragrammidest (nt *soovib iga inimene lähedust*). Suurim vahe ilmneb aga adverbilõpuliste klastrite võrdluses. Nii on K1 klastris VSSD üks verbile järgnevaist substantiividest subjekti funktsioonis 42,4%-l juhtudest (nt *kasvatab ema last üksi*), K2 klastris VPSD esineb pronoomen või substantiiv subjektina vaid 19,4%-l juhtudest (nt *jõudsin*

ma kinno juba, aitab see raamat paremini). VPSD-klastri kasutusest K1 kõneleja keelepruugis nähtub, et see struktuur tingibki harvema subjektikasutuse – verbile vahetult järgnev pronoomen on sagedamini eestäiend kui verbile järgnev substantiiv, mis esineb selle võrra sagedamini muudes funktsioonides, sh subjektina. Ometi on subjekti hõlmavate tetragrammide osakaal K1-s märkimisväärselt suurem (28,4%, vahe K2-ga 9%). Õppijal järgneb VSSD-struktuuris verbile subjekt samuti sagedamini (26,9%), ent vahe emakeelekõnelejaga on suur (15,5%).

Oletades, et kehtib ka vastupidine tendents ja K2-s eelneb subjekt verbile sagedamini kui K1-s, vaatlesin subjektikasutust klastreis SSSV ja ASSV, mis on nii K1 kui ka K2 valimis vastavalt eelistatuim substantiivi-alguline ja verbi lõpuline ning eelistatuim adjektiivialguline ja verbi lõpuline klaster. Selgus, et klastri SSSV tetragrammidest sisaldab K1-s subjekti 55% (nt *lapse isa Vello rääkis*) ja K2-s 62% (nt *raamatu lõpus peategelane kirjutas*). Klastris ASSV on subjekti hõlmavate tetragrammide osakaal K1-s 47,7% (nt *kuulsusrikka sõjakäigu saatuse otsustas*) ja K2-s 57,2% (nt *vastavas olukorras noor hakkab*).

Kaivapalu (2010) on välja toonud, et vene emakeelega eesti keele õppija ei järgi nii rangelt V2-reeglit, vaid tajub sõnajärjemalli laiend-verb-subjekt-laiend kõrval loomulikuna ka malli laiend-subjekt-verb-laiend. Õppijate kirjalike tekstide sõnajärg on sageli mõjutatud suulisest kõnest, sest eri registritest pärit sisendi eristamine on õppijaile kuni kõrgeimate keeletasemeteni keerukas (Kaivapalu 2010: 106). Suulises keelekasutuses ei näi V2-sõnajärjetendents aga kirjakeelega võrdsel määral kehtivat (Lindström 2000).

Saab järeldada, et õppija tarvitab predikaadi ja subjekti otsejärge (subjekt-verb) sagedamini ka olukordades, kus emakeelekõneleja eelistab pöördjärge (verb-subjekt), nt *sõprade seltskonnas noored tahavad, <järgmisel> päeval välisminister Meri sõitis* (K2) vs. *<tihti> kasutavad vanemad lapsi ära, <varsti> liikus rahvamass politseibussile järele* (K1).

4) Õppijal järgneb verbile sagedamini objekt.

Kui subjekt esineb K2 verbialgulistes tetragrammides harvem, siis objekti kasutab õppija verbile järgnevas lähikontekstis sagedamini.

Erinevus K1-ga on suurim substantiivi- ja adverbilõpuliste invariantsete klastrite puhul. Kui K1 klastris VSSS kätkeb objekti (peamiselt viimase komponendina) 37,4% tetragrammidest (nt *andnud pättidele karistamatuse tunde*), siis K2 klastris VPSS 47,6% (nt *kuulan teiste inimeste mõtteid*). Klaster VSSD hõlmab K1-s objekti (mis on peamiselt kolmas komponent) 31,7%-l juhtudest (nt *mõjutab maailma riike erinevalt*), samas kui K2 VPSS-struktuuris on objekti sisaldavate tetragrammide osakaal 48,1% (nt *arendas oma keeleoskust edasi*).

Taas tundub tegemist olevat õppijale omase sõnajärje-eelistusega, sest K1-s sisaldab VPSS-klaster objekti tunduvalt vähem (28,4%, vahe K2-ga 19,7%); samuti tarvitab emakeelekõneleja objekti harvem klastris VPSS (38,5%, vahe K2-ga 9,1%).

5) Õppija kasutab rohkem predikatiivikonstruktsioone.

Õppija eelistatud verbialgulistes struktuurides tuleb välja predikatiivi suurem sagedus. Nii esineb K1 klastris VSSS *olema*-verbi rohkest kasutusest hoolimata predikatiivikonstruktsioone vaid 3,9%-l juhtudest (nt *on lääne sepitsuste tulemus*) ja K2 klastris VPSS 9,6%-l juhtudest (nt *on nende normaalsuse piir*).

K2 invariantstes adjektiivilõpulisel klastris VVDA on adjektiiv predikatiivi funktsioonis enam kui pooltel juhtudel (51,2%, nt *teha on alati raske*), samas kui K1 invariantstes klastris VSSA on adjektiiv tüüpiliselt (77,1%) eestäiendi funktsioonis ja predikatiivina esineb harva (8,9%, nt *on sõit Tallinnas tasuta*). Osalt on põhjus struktuuride erinevuses. Ka K1 kõneleja on VVDA-struktuuris adjektiiv rohkem predikatiivina kasutatud (26,4%), ent vahe K2 kõnelejaga on peaaegu kahekordne. Õppija tarvitab ka VSSA-klaster adjektiiv predikatiivina sagedamini (14,3% juhtudest) kui emakeelekõneleja.

Järelikult on predikatiivi suurem esinemus eelkõige K2 iseärasus, lisaks eelistab õppija ka säärast verbialgulist ja adjektiivilõpulist struktuuri, mis seda funktsiooni soosib. Kuna K2 valimis on esindatud peamiselt vene emakeelega inimeste õppijakeel, siis võib selles näha nende K1 mõju. Selle nähtuse põhjusi, nagu teistegi välja toodud erisuste puhul, oleks vaja edaspidi eraldi uurida.

6) Õppija kasutab modaalverbe samal määral või pisut vähem.

Kui varem on välja toodud, et EVKK tekstidele on võrreldes ajakirjanduskeelega tunduvalt omasemad modaalse öeldisega konstruktsioonid (Eslon 2009), siis C1-taseme struktuure minu andmetel modaalverbide rohkus ei iseloomusta. Vilunud keeleteadaja tasemel õppija kasutab verbialgulistes tetragrammides modaalverbe K1-ga võrreldes pigem veidi vähem, ehkki hinnangulisus peaks tekstide arutleva laadi tõttu iseloomustama mõlemat valimit. Modaalverbide kasutust võimaldab võrrelda pronoomenilõpuline klaster VVDP, mis on klassi V–P invariant nii K1-s kui ka K2-s. K1 tetragrammidest algab modaalverbiga 26,3%, moodustades 19,3%-l juhtudest järgneva infiniitverbiga ahelverbi (nt *võiks julgustada ka teisi*). Ülejäänud juhtudel esinevad vaegpöördelised modaalverbid *tulema*, *maksma* ja *tasuma*, mille da-infinitiivne laiend talitleb subjektina (nt *tasub pidada ka selliseid*). K2 kõneleja ei kasuta neid üldse, ahelverbiga algab 15,2% tetragrammidest (nt *peab proovima absoluutselt kõike*).

Klassis V–A eelistab õppija VVDA-struktuuri, mis tingib suurema modaalverbide kasutuse (17,9%, nt *võivad olla ikkagi õnnelikud*) võrreldes emakeelekõneleja eelistatud VSSA-ga, kus modaalverbe esineb väga vähe (1,9%, nt *peaks naiste ajateenistuse kohustuslikuks <muutma>*). Kui aga vaadata VVDA-struktuuri kasutust K1-s, siis on modaalverbide osakaal isegi veidi suurem (21,8%, vahe 3,9%). Klassi V–V invariantsete klastrite VSSV (K1) ja VPSV (K2) puhul on ahelverbide osakaal sarnane, vastavalt 23,3% ja 24,7% (nt *saaks Eesti koolides õpetada, pean seda tööd tegema*). Substantiivi- ja adverbilõpulistest klastrites on modaalverbid esindatud võrdselt vähesel määral.

7) Õppija keeleteaduses ei avaldu verbivormieelistus nii selgelt.

Finiitverbide kasutus on nii K1 kui ka K2 kirjalikus tekstiloomes piiratud 3. isiku vormidega, eelkõige preesensi ainsuse 3. pöördega. Emakeelekõnelejal tõuseb see verbivorm aga sageduse poolest selgemalt esile, samas kui õppijal on eri vormide kasutus ühtlasem. Nii on K1 invariantsetes adverbilõpulistest klastris VSSD preesensi ainsuse 3. pöörde osakaal 38,3% (nt *liigub maksumaksja raha Eestis*), kuid K2 invariantsetes klastris

VPSD algab selle vormiga 26,4% tetragrammidest (nt *väljub sellest võitlusest võitjana*) ning 3. isiku vormide kõrval on suhteliselt sage (10,1%) ka ainsuse 1. isiku preesens (nt *kuulen seda juttu väga*), mida K1 VSSD-klastris peaaegu ei leidu. Samuti on õppija kasutanud rohkem konditsionaalivorme (9,3% vs. 2,4%).

Finiitvormide kasutus varieerub rohkem ka verbilõpulistele klastrite esikomponendi puhul. K1 klatri VSSV tetragrammidest algab 39,7% preesensi ainsuse 3. pöörde vormiga (nt *tekib soov peret luua*), kuid K2 klatri VPSV tetragrammidest 30,9% (nt *loodab iga osaleja võita*). Taas on K2-s erinevalt K1-st suhteliselt sage ka preesensi ainsuse 1. pööre (10,3%). K2 kõneleja vormieelistused ei tule nii selgelt esile samuti konjunktsioonilõpulisel klassis, nii K1-le iseloomulikus klatri VSSJ kui ka K2-le tava-pärasel klatri VPSJ, sest mõlemas on üsna samaväärselt eelistatud kahte verbivormi: indikatiivi ainsuse 3. pöörde preesensit ja *da*-infinitiivi. K1-s on nende osakaal vastavalt 25,6% (nt *käib toast tuppa ega*) ja 23,1% (nt *takistada eestlaste väljasuremist ja*), K2-s vastavalt 21,2% (nt *kaob igasugune vajadus või*) ja 19,7% (nt *täita oma rolli ehk*). K1 kõnelejaga võrreldes kasutab õppija aga pisut sagedamini imperfekti ainsuse ja preesensi mitmuse 3. pööret (nt *õppis sellest kogemusest ja, pööravad talle selja või*).

Kui vaadata nt õppija eelistatud adverbilõpulisel klatri VPSD ja verbilõpulisel klatri VPSV kasutust K1 valimis, siis tõuseb teiste seas esile vähem vorme kui K2-s, nt imperfektivormid on harvemad. Ühtlasem jaotumine tundub niisiis olevat õppija keelekasutuse, mitte struktuuri iseärasus.

8) Õppija kasutab vähem infiniitverbe.

Kui K1 klatri VSSS tetragrammidest algab infiniitvormiga 28,5% (nt *kasutada piiride muutmiseks jõudu*), siis K2 klatri VPSS 22,1% (nt *teadma oma elu eesmäärke*). K1-s on VSSD-klastris verbi infiniitsete vormide osakaal 18,6%, K2 klatri VPSD aga 11,6%. VSSJ-struktuur K1-s algab infiniitverbiga 36,3%-l juhtudest, VPSJ-struktuur K2-s aga 29,2%-l juhtudest.

Et infiniitsete verbivormide osatähtsus tetragrammide algusverbina on õppija eelistatud klatriks palju väiksem, siis näitab see verbi

funktsionaalset mitmekesisust emakeelekõneleja keelekasutuses. Infiiniitsed vormid on predikaadi kõrval ka objekti (nt <tahtis> rääkida oma muredest isale), subjekti (nt <on meile elutähtis> taastada liitlaste abiga maailm), predikatiivi (nt <parim lahendus on> anda talle aega mõtisklusteks), adverbiaali (nt lämisedes sõbra sünnalt saabub) või atribuudi funktsioonis (nt <said võimaluse> jätkata oma esivanemate traditsiooni).

9) Õppija substantiivi-, adjektiivi- ja verbikasutus on leksikaalselt kitsam.

Leksikaalse rikkuse poolest jääb ka C1-taseme saavutanud õppija keelekasutus emakeelekõnelejale alla, iseäranis substantiivide puhul. Kui nt K1 klastris VSSS moodustavad ainukordsed lekseemid eelviimase ja viimase komponendi puhul vastavalt 68,4% ja 70,4% kõigist sõnedest, siis K2 klastris VPSS on ainukordsete lekseemide osatähtsus vastavalt 37,3% ja 54,2%. Suur erinevus substantiivide leksikaalses varieerumises ilmneb ka klastrite VSSD (K1) ja VPSD (K2), VSSV (K1) ja VPSV (K2) ning VSSJ (K1) ja VPSJ (K2) võrdluses – emakeelekõneleja substantiivikasutusest moodustavad ainukordsed substantiivid enam kui 20% võrra suurema osa.

K1 klastris VSSA on kõigi tetragrammide lõpus ainukordsed 70,1% adjektiividest, K2 klastris VVDA on see vaid 51,2%. Tegu on väga erinevate struktuuridega, kuid ka K1 kõnelejal on klastris VVDA ainukordsete adjektiivide osakaal märksa suurem (74,7%), K2 kõnelejal on see osakaal väiksem ka klastris VSSA (63,6%). Märkimist väärib, et kui K1 keelekasutuses moodustavad eestäiendina käituvad partitsiibid (nt *elav*, *õpetatav*, *heitunud*) kolmandiku adjektiividest, siis K2-s neid peaaegu ei esine ning K1-s eelistatud *v*-partitsiibid puuduvad sootuks.

Erinevus verbide leksikaalses varieerumises on väiksem. Nt on ainukordsete lekseemide osakaal K1 klatri VSSV algusverbi puhul 22,4% ja lõpuverbi puhul 61,2%, K2 klatri VPSV algus- ja lõpuverbi puhul vastavalt 18,6% ja 55,7%. Samas ilmneb õppijal *olema*-verbi kinnistumine. K2 klatri VVDA kätkeb verbi *olema* põhiverbina kas esi- või teise komponendina 64,3% tetragrammidest. Kui analüüsida sama klatri K1-s, siis

selles on *olema*-verbi kui põhiverbi sisaldavate tetragrammide osakaal tunduvalt väiksem (39,1%, vahe 25,2%).

Muutumatud sõnad, adverbid ja konjunktsioonid, ning suletud käändsõnaliigi pronoomenid varieeruvad leksikaalselt ühtemoodi. Kattuvad ka leksikaalsed eelistused. Nt esinevad adverbid struktuuride VSSD (K1) ja VPSD (K2) lõpus peamiselt rõhu-, kvantumi- ja viisiadverbiaalina, VVDP-s aga kummaski valimis valdavalt rõhuadverbiaalina (sagedaim *ka*), sageli ka eitava adverbina *mitte*. Analüüsi põhjal tuleb seega esile sõnaliigi leksikaalse ja vormirikkuse võrdeline seos sellega, kuivõrd erineb sõnavara ulatus sarnastes K1 ja K2 struktuurides.

6. Kokkuvõte

Erinevused kahe keelekasutusvariandi vahel ilmnevad keelestruktuuri eri tasanditel. Kuigi nt verbialgulisi ja adverbilõpulisi tetragramme koondav klass V–D on K1-s ja K2-s sarnasel määral esindatud ja kuigi verbialgulise ja pronoomenilõpulise klassi V–P sagedaim sõnajärjestruktuur on kahes valimis ühtviisi VVDP, ei tähenda see, et morfosüntaksi ja leksika tasandil erinevusi ei esine. Üldiselt eelistatakse K1-s ja K2-s kasutada erinevaid struktuure. Emakeelekõnelejal tõusevad sagedaimad klastrid ja tüüpilised verbivormid enamasti selgemalt esile, kuid leksikaalne varieerumine on, eriti substantiivide ja adjektiivide puhul, avaram. See-eest on õppijal keelekasutusmustrite ja verbivormide kasutussagedus ühtlasem, sõnavara aga kitsam.

C1-tasemega õppija keelepruugis esineb leksikaalseid ja grammatilisi stereotüüpe ehk teatud lekseemide (nt *olema*-verb) või vormide (verbi eitusvormid, imperfekt) ohtramat tarvitamist kontekstis, kus need on ka K1-s sagedad. Teisalt ei leia peaaegu üldse kasutamist partitsiivid, iseäranis eestäiendi funktsioonis, mistõttu on K2-s verbide funktsionaalne mitmekesisus väiksem. Õppija tarvitab harvem predikaadi ja subjekti pöördjärge (verb-subjekt) ning sagedamini predikatiivikonstruktsioone. Analüüsi tulemused näitasid, et tetragrammi struktuur määrab küll komponentide vormivaliku ja võimalikud funktsioonid, kuid

erinevused teatud vormide ja funktsioonide osakaalus tulevad ilmsiks ka sarnaste struktuuride võrdlemisel.

Õppija keelekasutuse analüüs näitab justkui vastandlikke tendentse: ühelt poolt suundumust stereotüüpide kasutamisele, teisalt vabamat varieerumist ja vormivalikut, milles kindlad mustrid nii selgelt esile ei tule (ilmneb nii klasside, klastrite kui ka morfoloogilise vormi tasandil). Suuremat varieerumist K2-s saab pidada selle väljenduseks, et kindla-piirilised valikud sõnaliikide, vormide ja lekseemide kombineerimisel on veel lõpuni välja kujunemata.

Plaanin edaspidi kõrvutada K1 ja K2 kirjaliku keelekasutuse verbi-algulisi struktuure verbilõpulistega, et saada aimu kasutusmustritest verbi ümbritsevas laiemas kontekstis. Saamaks täpsemat ülevaadet eesti keele õppija C1-taseme ja emakeelekõneleja keelekasutuse samalaadsusest ning erinevuste olemusest ja põhjustest, tuleks invariantsete struktuuride kõrval võrdlevalt vaadelda kõiki statistiliselt olulisi mustreid ja tuua välja, kuidas toimivad K1 ja K2 kõneleja samalaadset sõnajärjendit kasutades. Erineva tasemega õppijate keelekasutusmustrite kõrvutus (nt C1 ja B2, B2 ja B1) annaks võimaluse leida neid eristavad struktuursed ja leksikaalgrammatilised tunnused, mille põhjal saab modelleerida keeleoskustasemete lingvistilise sisu.

Tänuõnad

Uurimistööd toetas Euroopa Liidu Euroopa Sotsiaalfond õpirändeprogrammi DoRa raames, mida viis ellu Sihtasutus Archimedes.

Kirjandus

- Aggarwal, Charu C., Chandan K. Reddy (Eds.) 2014. *Data Classification: Algorithms and Applications*. New York: Chapman and Hall/CRC Press.
- Allkivi, Kais 2016. C1-tasemega eesti keele õppijate kirjalik keelekasutus võrdluses emakeelekõnelejatega: samalaadsusi ja nihkeid verbist paremal paiknevas kontekstis. Magistritöö. Tallinn: Tallinna Ülikool.
- Argus, Reili 2006. Eesti keele verbi ajamorfoloogia ja aspektilisuse omandamisest [‘Acquisition of tense morphology and aspectuality in Estonian’]. – *Emakeele Seltsi aastaraamat* 52, 7–32.

- Bassano, Dominique, Pascale-Elsa Eme, Christian Champaud 2005. A naturalistic study of early lexical development: General processes and inter-individual variations in French children. – *First Language* 25 (1), 67–101.
- Bardovi-Harlig, Kathleen 2000. *Tense and Aspect in Second Language Acquisition: Form, Meaning, and Use*. Oxford: Blackwell.
- Biber, Douglas, Susan Conrad, Randi Reppen 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511804489>
- Bybee, Joan 2013. Usage-based theory and exemplar representations of constructions. – Thomas Hoffmann, Graeme Trousdale (Eds.), *The Oxford Handbook of Construction Grammar*. New York: Oxford University Press, 49–69.
- Ehala, Martin 2009. Linguistic strategies and markedness in Estonian morphology. – *Sprachtypologie und Universalienforschung* 62, 29–48. <http://dx.doi.org/10.1524/stuf.2009.0003>
- Ellis, Nick C. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. – *Studies in Second Language Acquisition* 24 (2), 143–188. <http://dx.doi.org/10.1017/S0272263102002024>
- Ellis, Nick C. 2006a. Language acquisition as rational contingency learning. – *Applied Linguistics* 27 (1), 1–24. <http://dx.doi.org/10.1093/applin/ami038>
- Ellis, Nick C. 2006b. Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. – *Applied Linguistics* 27 (2), 164–194. <http://dx.doi.org/10.1093/applin/aml015>
- Ellis, Nick C. 2015. Implicit and explicit language learning: Their dynamic interface and complexity. – Patrick Rebuschat (Ed.), *Implicit and Explicit Learning of Languages*. Amsterdam: John Benjamins, 3–23. <http://dx.doi.org/10.1075/sibil.48.01ell>
- Ellis, Nick C., Matthew B. O'Donnell, Ute Römer 2015. Usage-based language learning. – Brian MacWhinney, William O'Grady (Eds.), *The Handbook of Language Emergence*. Malden, MA – Oxford: Wiley-Blackwell, 163–180.
- Eslon, Pille 2009. Eestikeelses tekstiloomes eelistatud konstruktsioonid ja käändevormid. – *Korpusuuringute metodoloogia ja märgendamise probleemid*. Tallinna Ülikooli eesti keele ja kultuuri instituudi toimetised 11. Tallinn: Tallinna Ülikooli kirjastus, 30–53.
- Eslon, Pille 2014a. Eesti vahekeele korpus ['Estonian interlanguage corpus']. – *Keel ja Kirjandus* 6, 436–451.

- Eslon, Pille 2014b. Morfosüntaktilise ja leksikaalse varieerumise piiridest: ilukirjandus- ja õppijakeele kasutusmustrite võrdlus ['Constraints on morpho-syntactic and lexical variability']. – Eesti Rakenduslingvistika Ühingu aastaraamat = Estonian Papers in Applied Linguistics 10, 55–71. <http://dx.doi.org/10.5128/ERYa10.04>
- Eslon, Pille 2014c. Adverbi sisaldavate struktuuride tekstifunktsioonidest eesti ilukirjandus- ja õppijakeeles ['On the textual functions of adverbial structures in literary Estonian and in Estonian learner language']. – Lähivõrdlusi. Lähivertailuja 24, 15–46. <http://dx.doi.org/10.5128/LV24.01>
- Eslon, Pille, Helena Metslang 2007. Õppijakeel ja eesti vahekeele korpus ['Learner language and Estonian interlanguage corpus']. – Eesti Rakenduslingvistika Ühingu aastaraamat = Estonian Papers in Applied Linguistics 3, 99–116. <http://dx.doi.org/10.5128/ERYa3.07>
- Eslon, Pille, Katre Õim, Annekatrin Kaivapalu, Reili Argus, Erika Matsak 2010. Kuidas uurida esimese ja teise keele omandamist? ['Perspectives of the research of the first and of the second language acquisition']. – Lähivõrdlusi. Lähivertailuja 20, 11–48. <http://dx.doi.org/10.5128/LV20.01>
- Eslon, Pille, Heleriin Paeoja 2015. Samatähenduslike sünteetiliste ja analüütiliste verbide kasutamine ['Use of the synonymous synthetic and analytic verbs']. – Lähivõrdlusi. Lähivertailuja 25, 63–104. <http://dx.doi.org/10.5128/LV25.04>
- Granger, Sylviane 2015. Contrastive interlanguage analysis: A reappraisal. – International Journal of Learner Corpus Research 1 (1), 7–24. <http://dx.doi.org/10.1075/ijlcr.1.1.01gra>
- Granger, Sylviane 1996. From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. – Karin Aijmer, Bengt Altenberg, Mats Johansson (Eds.), Languages in Contrast: Papers from a Symposium on Text-based Cross-Linguistic Studies. Lund Studies in English 88. Lund: Lund University Press, 37–51.
- Ivaska, Ilmari 2015. Edistyneen oppijansuomen konstruktiopiirteitä korpusvetoisesti: avainrakenneanalyysi. Turun yliopiston julkaisuja = Annales Universitatis Turkuensis. Turu: Turun yliopisto.
- Jarvis, Scott, Magali Paquot 2012. Exploring the role of n-grams in L1 identification. – Scott Jarvis, Scott A. Crossley (Eds.), Approaching Language Transfer through Text Classification. Bristol–Buffalo–Toronto: Multilingual Matters, 71–105.

- Jockers, Matthew L. 2014. *Text Analysis with R for Students of Literature*. Cham–Heidelberg – New York – Dordrecht – London: Springer. <http://dx.doi.org/10.1007/978-3-319-03164-4>
- Kaart, Tanel 2004. Biomeetria. Kahemõõtmeline sagedustabel. Hii-ruut test. http://www.eau.ee/~ktanel/kool_ja_too/informaatika_ja_biomeetria/LKI_I_loengud5b.pdf (29.2.2016).
- Kaivapalu, Annekatrin 2010. Mõnede eesti sõnajäremallide psühholingvistilisest reaalsusest [‘On psycholinguistic reality of some word order patterns of Estonian’]. – *Eesti Rakenduslingvistika Ühingu aastaraamat = Estonian Papers in Applied Linguistics* 6, 103–120. <http://dx.doi.org/10.5128/ERYa6.07>
- Kerge, Krista; Hille Pajupuu, Pilvi Alp, Halliki Põlda, Anne Uusen 2014. Towards sophisticated writing. – *Uurimusi keele omandamisest, õppimisest ja korpustest*. Tallinna Ülikooli eesti keele ja kultuuri instituudi toimetised 16. Tallinn: Tallinna Ülikool, 103–115.
- Kitsnik, Mare 2014. Verbivormid B1- ja B2-taseme kirjalikus õppijakeeles [‘Written learner language verb forms at B1 and B2 levels’]. – *Eesti ja soome-ugri keeleteaduse ajakiri / Journal of Estonian and Finno-Ugric Linguistics* 5 (3), 9–35.
- Kitsnik, Mare 2007. *Õppijakeele uurimise ja arendamise võimalusi eesti vahekeele korpuse põhjal (eituse väljendamise näitel)*. Magistritöö. Tallinn: Tallinna Ülikool.
- Klein, Wolfgang 1986. *Second Language Acquisition*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511815058>
- Lindström, Liina 2000. *Narratiiv ja selle sõnajärg [‘Narrative and its word order’]*. – *Keel ja Kirjandus* 3, 190–200.
- Maisla, Diana 2014. *Eesti keele mineviku ajavormid vene emakeelega üliõpilaste kasutuses [‘Past tenses of Estonian language in the usage of Russian-speaking students’]*. *Dissertationes linguisticae Universitatis Tartuensis* 20. Tartu: Tartu Ülikooli kirjastus.
- Mardia, Kantilal Varichand, John T. Kent, John M. Bibby 1979. *Multivariate Analysis*. London–San Diego: Academic Press.
- Martin, Maisa, Sanna Mustonen, Nina Reiman, Marja Seilonen 2010. *On becoming an independent user. – Communicative Proficiency and Linguistic Development: Intersections between SLA and Language Testing Research*. *Eurosla Monographs* 1. European Second Language Association, 57–81.
- McEnery, Tony, Andrew Wilson 1996. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.

- Müürisep, Kaili 2000. Eesti keele arvutigrammatika: süntaks ['Computer grammar of Estonian: syntax']. *Dissertationes Mathematicae Universitatis Tartuensis* 22. Tartu: Tartu Ülikooli Kirjastus.
- Ots, Sander 2012. Statistikapõhise tarkvara loomine morfoloogiliste kollokatsioonide eraldamiseks eesti keele tekstidest. Bakalaureusetöö. Tallinna Ülikool.
- Paeoja, Heleriin 2014. Analüütiliste / sünteetiliste verbipaaride kasutusmustrid 1990. aastate eesti ilukirjanduskeeles. Magistritöö. Tallinna Ülikool.
- Pajupuu, Hille, Krista Kerge, Lya Meister, Eva Liina Asu, Pilvi Alp 2010. Natural speaking and how to assess it. – *Trames: Journal of the Humanities and Social Sciences* 59 (2), 120–140.
- Pajupuu, Hille, Krista Kerge, Pilvi Alp 2009. Sõnavara loomulik rikkus haritud keeleskaja tekstides ['Natural lexical richness in educated language use']. – *Eesti Rakenduslingvistika Ühingu aastaraamat = Estonian Papers in Applied Linguistics* 5, 187–196. <http://dx.doi.org/10.5128/ERYa5.12>
- Pool, Raili, Elle Vaimann 2005. Vead kõrgtasemel eesti keele kõnelejate kirjallikus keelekasutuses ['Errors in written Estonian by advanced level speakers']. – *Eesti Rakenduslingvistika Ühingu aastaraamat = Estonian Papers in Applied Linguistics* 1, 115–137. <http://dx.doi.org/10.5128/ERYa1.06>
- Raamdokument 2007 = Euroopa keeleõppe raamdokument: õppimine, õpetamine, hindamine. Tartu: Haridus- ja Teadusministeerium, 2007. <http://www.ekk.edu.ee/101183> (29.2.2016).
- Römer, Ute 2007. Learner language and the norms in native corpora and EFL Teaching Materials: A Case Study of English Conditionals. – Sabine Volk-Birke, Julia Lippert (Eds.), *Anglistentag 2006 Halle. Proceedings*. Trier: Wissenschaftlicher Verlag Trier, 355–363.
- Römer, Ute 2004a. Comparing real and ideal language learner input: The use of an EFL textbook corpus in corpus linguistics and language teaching. – Guy Aston, Silvia Bernardini, Dominic Stewart (Eds.), *Corpora and Language Learners*. Amsterdam: John Benjamins, 151–168.
- Römer, Ute 2004b. A corpus-driven approach to modal auxiliaries and their didactics. – John McHardy Sinclair (Ed.), *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins, 185–199.
- Šmõreitšik, Anastassia 2009. ma- ja da-infinitiivi kasutuse võrdlus eesti kirjakeeles ja õppijakeeles. Magistritöö. Tallinna Ülikool.

Written language use of C1 learners of Estonian and native speakers in comparison: Analysis of verb-initial fourgrams

KAIS ALLKIVI

Tallinn University

This paper comparatively describes verb-initial morphosyntactic patterns in the written language use of C1 level learners of Estonian, i.e. proficient L2 users, and educated native speakers. For this purpose, verb-initial fourgrams were investigated. A further aim of the study is to help provide research-based linguistic content to define language acquisition levels.

The research material consists of L1 and L2 argumentative writings:

- essays written by L2 learners of Estonian, which have been evaluated to represent the proficient user level C1, were taken from the inter-language corpus of the Tallinn University (75,320 tokens, 8364 fourgrams);
- L1 reference corpus was compiled from opinion articles published in the two leading newspapers in Estonia during December 1–31 in 2014 and written by highly educated non-linguists, considered the benchmark for both L1 and L2 learners (107,590 tokens, 11,437 fourgrams).

Cluster analysis was applied to bring forth identical linear subsequences of morphological and syntactic tags based on frequency. These sequences were grouped into more general patterns of language use. Fourgrams where the word classes of the components coincide (e.g. VPSS, representing verb-pronoun-substantive-substantive) constitute bundles, referred to as clusters. The highest level in the hierarchy has been presented on the basis of the word class of the first and last component of fourgrams. These groups of fourgrams (e.g. V–S, representing verb–substantive) are referred to as morphological classes. Thus, clusters are distinguished by the variation of the middle components of fourgrams.

The analysis revealed mostly statistically relevant differences in the frequencies of classes and clusters of fourgrams. E.g. the adjective-final class V–A is more common in L1 use, while the pronoun-final class V–P is more common

in L2 use. Furthermore, the preferred clusters mostly do not overlap, e.g. the preferred cluster in the class V–A is VSSA (verb-substantive-substantive-adjective) in L1 and VVDA (verb-verb-adverb-adjective) in L2.

The morphological, syntactic and lexical comparison of the preferred clusters of L1 and L2 within the same class (e.g. VSSS favoured by L1 users and VPSS favoured by L2 users in the class V–S) brought out several interesting tendencies:

- 1) negative verb forms are used more frequently in L2 than in L1;
- 2) analytic tense forms (Present and Past Progressive) are less common to L2 language use;
- 3) there is a tendency of the subject to precede the verb more commonly in L2 than in L1, therefore, subjects are rarer in L2 verb-initial structures;
- 4) the object, however, follows the verb more commonly in L2 than in L1;
- 5) in L2 there is a wider use of predicatives that occur together with the existential verb to be;
- 7) lesser use of infinitive verb forms in L2 shows in the compared clusters;
- 8) the use of modal verbs is relatively similar in the two groups;
- 9) a greater lexical diversity shows in L1, especially in substantives, as well as in adjectives and verbs, whereas there is no difference in the lexical diversity of adverbs, conjunctions and pronouns.

Keywords: usage-based language description; second language acquisition; cluster analysis; contrastive interlanguage analysis; morphosyntactic variability, lexical variability; Estonian

Kais Allkivi

Tallinna Ülikool

Narva mnt 25, 10120 Tallinn, Estonia

kais.allkivi@gmail.com