

EESTI KEELE VERBIDE SEMANTILISEST LIIGITAMISEST

ILONA TRAGEL, PIIA TAREMAA

Inimeste loomuomane oskus ja vajadus kategoriseerida avaldub väga mitmel tasandil ning muuhulgas ka keeles. Üks keelelisi kategooriaid, mis enamikus keeltest avaldub, on verbikategooria. Keeleti erineb siiski, kui hõlpsasti on verb kui sõnaliik määratletav. Varieerub ka see, kuivõrd on verbikategooria jaotatav alamkategooriateks, kuid üldjuhul on alamkategooriateks jagamine sootuks keerulisem kui verbikategooria enda eristamine. Eesti keeles ei ole üldiselt verbi kui sõnaliigi määratlemine kuigi keeruline tänu morfoloogia rikkusele, kuid selle kategooria sees muust kui puhtalt vormist lähtuvate alamkategooriate moodustamine on Eesti keeleteadlastele olnud raske ülesanne. Ka maailma keeleteadus on näinud ainult üksikuid verbiklassifikatsioone (nt Levin 1993; Kipper Schuler 2005). Sestap pole üllatav, et siiani puudub eesti keele verbide semantiline liigitus ning üldse põhjalikum kõikide eesti keele verbide käsitus. Väga oluline verstepost on kindlasti Huno Rätsepa eesti keele põhiliste verbikesksete lausemustrite¹ analüüs 1978. aastast. Hiljem on verbiteemadega põhjalikumalt tegeldud Haldur Õimu projekti „Lihtlause semantiline analüüs” (2004–2010) raames, üks võimalik verbiliigitus on olemas ka eesti keele Wordneti kujul.

Siinse artikli eesmärk on näidata viise, kuidas saaks eesti keele verbe ja nende lausemustreid analüüsida verbiliigituse jaoks. Rõhuasetus on eri empiiriliste meetodite kombineerimisel ning olemasolevate keeleressursside ja analüüsivahendite optimaalsel ärakasutamisel, et jõuda usaldusväärse verbiliigituseni. Liigitamise problemaatika paremaks avamiseks oleme artiklis näitlikustamiseks valinud tuumverbid ning liikumisverbid, mida oleme ise uurinud. Kasutame järgmisi olulisemaid mõisteid, mille piiritlemisel lähtusime eelkõige verbide liigitamise praktilisest tööprotsessist:

- **verbi põhitähendus** (ingl *core meaning*) on tähendusjoonte kogum, mis iseloomustab verbi kõigis tema lausemustrites, n-ö skemaatiline tähendus;
- **lausemuster** (ingl *sentence/clause pattern, construction, argument structure*) on kooslus, millesse kuuluvad verbiga koos käivad semantiliselt või vormiliselt defineeritud üksused (sh teine verb lauses), aga mitte verb ise²;

¹ Rätsep (1978: 4) kasutab terminit *lausemall*, lähtudes generatiivse grammatika grupi kasutustavast: „Sõna *mall* on TRÜ eesti keele kateedri generatiivse grammatika grupi poolt kasutusele võetud murdesõna (kirde-eesti rannikumurdest), mis on inglise termini *pattern* eesti vaste. Sõna on olemas ka soome keeles.” Sõna *mall* tähendus (ÕS 2018 järgi ‘eeskuju, mudel’) kätkeb endas preskriptiivsust, meie lähenemine on deskriptiivne ning eelistame seetõttu terminit (*lause*)muster (vt ka järgmist allmärkust).

² Kuna ilma keelematerjali uurimata ei ole võimalik ette öelda, kuivõrd on mingi keeleüksuste kogum konstruktsioon, veel vähem on aga alust rääkida sellest, justkui verb

- tunnus (ingl *feature*) on verbiliigilisuse seisukohast määrav verbi, lause või lauses esineva üksuse semantiline või vormiline omadus, mida on võimalik märgendada (ingl *coded feature*).³

Järgnevalt anname ülevaate verbide semantilise liigitamise katsetustest maailma ja Eesti keeleteaduses. Seejärel peatume olulisematel probleemidel, mis verbide liigitamisel esile kerkida võivad, ning vaatleme, milliseid lahendusi tänapäevased meetodid liigitamisprobleemidele pakkuda võiks. Artiklile on oodata järke, milles tutvustame verbide semantilise liigitamise praktilisi lahendusi.

1. Taust

1.1. Maailmas

Maailma keeleteaduse üks olulisemaid verbiklassifikatsioone pärineb Beth Levinilt aastast 1993. Selles klassifikatsioonis on tavalisemad inglise keele verbid tähenduse ning argumentstruktuuride (st argumentide ja nende realiseerumise) järgi jagatud paljudesse üldistesse (nt tajuverbid, eemaldamisverbid ja liikumisverbid) ja spetsiifilisematesse klassidesse (nt nägemisverbid, pühkimisverbid ja liikumisviisiverbid). Levini verbiklassifikatsioon on kindlasti üks mõjukamaid tänapäevalgi ning leiab allikana kasutust paljudes keeleteaduslikes (sh keeletehnoloogilistes) uurimustes ja verbide andmebaasides. Näiteks on selle põhjal koostatud inglise keele verbide VerbNet (vt ka Kipper Schuler 2005). Sama lähtealuse järgi on eri valmidusastmetes ka teiste keelte verbide jaotused (vt lähemalt Verbs Index). Levini käsitluse, inglise ning teiste keelte VerbNet'ide eeskujul võiks olla võimalik luua eesti keele VerbNet (Jentson 2013). Esitame oma seisukoha selle kohta allpool (vt jaotist 2.1).

Verbiklasse ning verbiga lausete tüüpilisi ülesehitusi (lausemustreid) on püütud kindlaks määrata ka freimisemantikas (vt nt Fillmore 1968, 1982). Inglise ja saksa keele kohta on näiteks koostatud FrameNetid, mis koondavad põhilise info nende keelte verbide lausemustrite kohta, liigitades ühtlasi ka verbe endid semantilistesse klassidesse. Ka eesti keelt on freimisemantika vaatenurgast analüüsitud (Õim 1965; Õim, Saluveer 2002; Õim jt 2009, 2010). Samuti on verbe klassifitseeritud eri keelte kohta koostatud tesaurus-tes (WordNet), kus liigitus põhineb tavaliselt tesauruse koostaja intuitsioonil.

Eesti keele verbiliigitusele mõeldes on kindlasti oluline Anneli Pajuneni (2001) soome keele verbide klassifikatsioon. Selles laiaulatuslikus uurimuses liigitatakse soome verbid semantika järgi ning kirjeldatakse seejärel verbide lausemustreid, eeldades, et erinev verbitähendus paistab välja ka selle eripärasest lausemustrist.

„tingiks” ühe või teise argumendi, siis kasutame üldist mõistet *lausemuster*, mitte aga teoreetilisel „raskeid” mõisteid *konstruktsioon* ja *argumentstruktuur*.

³ Statistilises analüüsis on tunnus ühel või teisel moel mõõdetavate või kindlaks tehtavate väärtustega üksus ehk muutuja (ingl *variable*), mis iseloomustab meid huvitava verbiga korpuslauset.

1.2. Eestis

Eesti keele verbide põhjalikum semantiline klassifikatsioon seni puudub. Verbid on klassifitseeritud pigem vormilistest tunnustest lähtudes ning morfoloogilise ülevaate andmiseks ja seetõttu eeskätt muuttüüpideks (Tauli 1973: 109–112; EKG I: 344–349; Kaalep 2015; ülevaadet muuttüüpidest sõnaraamatutes vt Viks 2001). Lisaks morfoloogilistele liikidele on verbe jagatud ka tule-tusmalli järgi (Tauli 1973: 145–169; Kasik 2015: 108–182). Tavaline on verbide jagamine kaheks suureks klassiks, transitiiivseteks ja intransitiiivseteks verbi-deks (Tauli 1973: 145; EKG II: 46–47; Metslang 2017: 261–262), mis seostub grammatikates siiski rohkem süntaksi kui verbisemantikaga. Eesti keele õpi-kutes on verbid siiani esitatud enamasti vaid struktuurikeskselt (finiitsed ja infiniitsed verbivormid; verbi häälikuline struktuur, sh tüvemuutused, tüüp-sõnad jne). Verbivormistike esituse kõrval on verbide seosed semantikaga jää-nud õpikutes peaaegu tähelepanuta.

Siiski on näiteks eesti keele grammatikates tavaline, et üht või teist süntaksi teemat mitte üksnes ei näitlikustata konkreetsete verbide abil, vaid sageli seletatakse süntaksi erijooni just verbiliikide kaudu. Teatud teemade puhul ongi lausemuster või mingi iseloomulik joon lausemustris see, mille alusel verbe (semantilistesse klassidesse) liigitatakse. Näiteks on verbe jagatud eri klassidesse võimaliku sihitise vormi järgi, mis on ühtlasi ka üks põhjaliku-malt uuritud valdkondi verbiliigituse mõttes (vt nt Tauli 1983: 45–49; EKG II: 49–51; Klaas 1999; Tamm 2007, 2012; Metslang 2017: 264–271).

Verbiliikide osatähtsust süntaksikäsitlustes ilmestab seegi, kui palju on erinevaid verbiliike, mida süntaksi kirjeldamiseks on nimetatud. Näiteks eesti keele grammatikas (EKG I ja II) ja uusimas eesti keele süntaksi käsitluses (Erelt, Metslang 2017) on ühe või teise teema käsitlemisel (sealhulgas sõna-moodustuse peatükkides) nimetatud umbes 30–40 üldisemat või spetsiifili-semat verbirühma (nt aspektverbe, liikumisverbe, tajuverbe, olenemisverbe, teatamisverbe). Ka Reet Kasik (2015: 108–144) loetleb eesti keele sõnamoo-dustuse ülevaates tähenduse ja sõnamoodustusmalli järgi paljusid verbirühmi (nt põhjustamisverbe, tõlkimisverbe, kõnetusverbe). Verbiliikide nimetamisel on eesmärgiks olnud kirjeldada ja/või seletada eesti keele süntaksi, morfo-logia või sõnamoodustuse eri tahke, mis osutab sellele, et verbisemantika ja lausestruktuurid on tihedasti seotud (vt ka Õim 1965; Taremaa 2017). Veelgi enam, keele grammatika mõistmiseks, seletamiseks ja õpetamiseks on mööda-pääsmatu teada, millised verbiliigid keeles on.

Nagu artikli algul mainitud, on eesti keele verbide ja lausemustrite võima-liku liigituse olulisemaid tähiseid Huno Rätsepa verbikesksete lausemustrite analüüs 1978. aastast. Rätsepa käsitlus on oluline mitmel põhjusel. Esiteks on see seni põhjalikum ülevaade eesti keele verbide võimalikest lausemustri-test. Teiseks on Rätsepa ülevaade käsitletav korpusajastu omalaadse esimese pääsukesena – kuigi lausemustrid põhinevad uurija enda intuitsioonil, on autor siiski kasutanud n-õ minikorpust ehk käsitsi kogutud kirjalikku keele-materjali, sest „Tekstides kohtab sellise ehitusega lauseid, selliseid ütle-mis-võimalusi, millele lingvistil iseseisvalt oleks raske tulla” (Rätsep 1978: 7). See on mõte, mis iseloomustab tänapäeva korpuslingvistikat tervikuna (Biber jt 1998; McEnery, Wilson 2001).

Sama olulised on ka Haldur Õimu (1965, 1970; Õim, Saluveer 2002) uurimused eesti keele verbide tähendusest ning lausete semantilisest struktuurist, mis on olulised nii semantika kirjeldamise teoreetilisest kui ka keelematerjali empiirilise analüüsi olulisuse rõhutamise vaatenurgast. Freimisemantikast tõukus näiteks Õimu juhitud projekt „Eesti keele lihtlause analüüs”, mille raames ning automaatse semantikaanalüüsi eesmärgil uuriti eesti keele verbe ning lausemustrite semantilist struktuuri (vt nt Õim jt 2009, 2010).

Verbe ja nende võimalikke liike puudutavatest uurimustest tasuks märkida mitmeid teisigi viimaste aastakümnete uurimusi: teatud liiki verbidest Trigel 2003 (tuumverbid), Pajusalu, Trigel 2007 (muutumisverbid), Taremaa 2017 (liikumisverbid), Muischnek 2006 (tugiverbid); verbikonstruktsioonidest Penjam 2008, Muischnek, Kaalep 2010; murdekeele ja regilaulu verbidest ja verbikonstruktsioonidest Uiboaed 2013, Labi 2006a, 2006b; sagedase verbi polüseemiast Pajusalu 2001, Habicht jt 2010, Habicht, Trigel 2014, Trigel, Habicht 2017. Siiski pole eesti keele kohta esitatud ühtegi verbiliigitust, mis hõlmaks kõiki põhilisi eesti keele verbe. Samuti on oluline märkida, et kuigi näiteks eesti keele grammatikad ja ülevaated käsitlevad paljusid verbiliike, on verbipõhistele empiirilistele uurimustele saadud tugineda ainult piiratud mahu, kuna eesti keele verbide semantika ning lausemustrid on vähesel määral uuritud.

Rääkides verbide semantilisest liigitamisest, kerkivad uurija ette kaks põhilist probleemi. Üks on seotud sellega, kuidas üldse verbi tähenduseni jõuda, et verbe semantiliselt alustelt liigitada. Teine probleem on seotud sellega, millistelt semantiliselt alustelt verbe liigitada, kui erinevad semantilised jooned verbitähenduses kombineeruvad või on verb sootuks polüseemne. Vaatleme neid kahte teemat järgemööda.

2. Kuidas jõuda verbitähenduseni?

Eesti keeles puuduvad üldtunnustatud verbivormilised parameetrid, mille alusel verbe tähenduspõhistesse klassidesse jagada.⁴ Seetõttu peab eesti keele verbide liigitus paratamatult lähtuma verbisemantikast. Laias laastus on verbisemantiline analüüs võimalik kahte moodi: verbide liigitamine tähenduse või lausemuistri järgi.

2.1. Verbi tähendus

Üks – ning sageli kasutatud – võimalus on liigitada verbe nende enda tähenduse järgi, määrates põhitähenduse või põhitähenduse moodustavad olulisemad tähendusjooned nii, et keelelised kasutuskontekstid seda võimalikult vähe mõjutaksid. Piltlikult kirjeldades on see olukord, kus uurija võtab ette verbiloetelu ning jagab verbid mingitesse (enda arvates) võimalikult loomulikesse klassidesse. Ideaaljuhul võiks selline klassidesse jaotamine lähtuda

⁴ Paljudes slaavi keeltes jagatakse verbid klassidesse morfoloogiliste tunnuste alusel. Näiteks sloveeni keele verbid jagunevad aspekti alusel perfektiivseteks ja imperfektiivseteks ning kannavad vastavaid morfoloogilisi markereid (nt prefiksid, tüvemuutused) (Lečič 2012).

verbi põhitähendusest. Põhitähendus oleks skemaatiline ning võimalikult vaba keelekasutuses verbiga kaasnevate lausemustrite tähendusest.

Verbi põhitähenduse olemasolu või selle määramise võimalikkus on pälvitud keeleteaduses palju tähelepanu, eriti konstruktsioonilistes lähenemistes, kus postuleeritakse, et verbil on oma tähendus ning konstruktsioonil oma tähendus (Goldberg 1995, 2006; Fried, Östman 2004).

Verbi skemaatilise tähenduse põhjal tehtud liigitus võib olla intuiitiivne, st uurija intuitsioonil põhinev tähendus(t)e määratlemine. Sõna tähenduse peale on aga üsna keeruline mõelda nii, et ei aktiveeruks sõna võimalikud situatsioonilised ja keelelised kasutuskontekstid. Tihtipeale kasutatakse lisaks korpustest pärit keelematerjali, mis toetab intuiitiivse analüüsi varal verbi tähendusklassidesse jagamist. Eriti puudutab see verbe, mis on tähenduselt kompleksed ja/või polüseemsed ning mille põhitähendus ei ole seetõttu nii ilmselge (nt *saama*) kui mõne teise verbi tähendus (nt *riisuma*).

Intuiitiivne verbiliigitus võib olla küllaltki informatiivne ning samas ajasäästlik töömeetod, mis võimaldab kiiresti liigitada suure hulga verbe (vt nt erinevate tesauruste verbitaksonoomiaid⁵). Meetodi miinuseks on tulemuste põhinemine väga väiksel valimil. See tähendab, et valimisse kuuluvad ainult verbe liigitavad uurijad ise, mitte aga suurem hulk keelekasutajaid. See omakorda mõjutab tulemuste usaldusväärsust: üksikute uurijate intuitsioon ei pruugi olla piisav, et esindada laiema kasutajaskonna ega isegi mitte mõne teise keeleuurija keelestruktuuri.

Et liigitus oleks usaldusväärsem, on standardmeetodina kasutatav intuitsioonide summa meetod – katse ehk eksperiment. See tähendab, et kasutades katselisi meetodeid verbitähenduse määramiseks, on liigitus empiiriline ja koondab paljude inimeste intuitsioone. Eesti keele kohta võib tuua näiteks abstraktse tähendusega verbi (nt *mõtleva*, *elama*) katse, kus oli kombineeritud joonistuskatse ja häälega mõtlemise katse (Tragel, Klavan: valmimas): katseisikutel paluti verbi tähendust kujutada skemaatiliselt (st suunda joonistades) ning samal ajal kirjeldada, mida nad mõtlevad.

Sageli on katseline lähenemine möödapääsmatu, kuna verbitähenduse määramiseks on vaja pääseda ligi implitsiitsele, teadvustamata teadmisele. Näiteks eesti keele liikumisverbide ühe olulise tähendusjoone – kirjeldatava liikumise kiiruse – hindamiseks on tehtud katse (Taremaa 2017), kus suur hulk katseisikuid andsid hinnangu ilma kontekstita esitatud verbi kiiruse kohta. Katses tuli liugskaalal märkida, kas verb märgib pigem aeglast, kiiret või mingi vahepealse kiirusega liikumist. Katse tulemuste järgi saab verbid liigitada klassidesse selle põhjal, kui kiiret liikumist verb esindab. Näiteks *lonkima* on aeglase liikumise verb, *kihutama* aga väga kiire liikumise verb.

Mõistagi võib intuiitiivset ja katselist lähenemist kombineerida, hinnates intuiitiivselt määratud verbiklasse katseliste meetoditega. Sama kehtib ka juba olemasolevate, sh näiteks teiste keelte klassifikatsioonide kohta. Tahaksime siiski rõhutada, et ühe keele (nt inglise keele) põhjal tehtud liigitus ei saa olla teise keele verbi liigituse ainsaks aluseks, kuna pole põhjust arvata, nagu

⁵ Eesti keele Wordnetis on verbi sünohulki ehk mõisteid 6712 ning eri verbe 9836 (13. II 2019 seisuga; andmed Heili Oravalt), mis on semantiliste suhete (peamiselt hüponüümia) alusel üksteisega suhestatud ehk klassifitseeritud (Kerner jt 2010; Orav jt 2011, 2015).

oleksid verbide alamkategoriad universaalsed. Vastupidi – isegi sõnaliigina eristatav verbikategooria ise ei ole universaalne (vt nt Haspelmath 2012).

2.2. Verbi lausemuster

Teine võimalus on läheneda verbisemantikale kaudseid teid pidi, määrates tähenduse verbi tegelike keeleliste kasutuskontekstide järgi. Tegelik kasutus võib hõlmata erinevaid aspekte, millest lihtsamini määratav on lausemuster. Lausemusteri kaudu lähenemist toetab korpuslingvistikas ning kognitiivses lingvistikas palju kinnitust leidnud tõsiasi, et tähenduslik erinevus peegeldub vormilises erinevuses, st verbi tähendus avaldub ühel või teisel moel verbi lausemusteris (vt nt Harris 1954; Bolinger 1968; Öim 1965, 1970; Hanks 1996; Tragel 2003; Kuznetsova 2015; Taremaa 2017).

Freimisemantika ja konstruktsioonigrammatika lähenemised (Fillmore 1968, 1982; Goldberg 1995, 2010) ning eriti nende praktilised rakendused (nt FrameNet) on samuti selgesti näidanud, et konstruktsioonide semantiline ülesehitus erineb verbiti ja mis veelgi olulisem, verbiliigiti, väga olulisel määral. Seetõttu on alust verbide liigitamisel kasutada distributsionaalset lähenemist. See tähendab, et verbiga samas lauses (osalauses) esinevate teiste oluliste elementide vormi, tähenduse ja omavaheliste kombinatsioonide, aga ka muude kasutust ilmestavate tunnuste põhjal on võimalik vähemalt mingil määral verbi tähendusi määrata ning verbe klassifitseerida. Nii võimaldaksid piisav materjalihulk ja sobivad statistilised analüüsimeetodid anda verbiliigituse, milles infokadu oleks võimalikult väike (vt täpsemalt 5. ptk).

Korpuslingvistilised ja statistilised meetodid võimaldavad hinnata ka polüseemsete verbide selliseid kasutusi, mis on seotud mingi kindla verbivormi või lausemusteriga. Selline lausemusterisse puutuv, konstruktsiooniline info võib olla oluline lähtekoht nii eri verbide klassidesse jagamisel kui ka selle üle otsustamisel, kas peaks sama verbi esitama üheaegselt mitmes verbiklassis.

Teisipidi saab lausemusteri kaudu hinnata olemasolevaid, verbi enda tähenduse põhjal pakutud klassifikatsioone, nagu käsitletud eelmises peatükis. Näiteks liikumisverbide uurimuses (Taremaa 2017) on vastandatud kontekstivabalt, nii intuitiivselt kui ka eksperimentaalselt määratud liikumisverbide klassid ning korpusmaterjali põhjal leitud lausemusterid. Need musterid aitavad kinnitada liikumisverbide klassifitseerimisotsuste kehtivust ning samas näitavad ka verbitähenduste kompleksust.

3. Mida teha tähendusliku kompleksusega?

Mõlemal, nii ainult verbi kui ka lausemusteri kaudu lähenemisel on oma puudused. Verbide kontekstivaba ja intuitsioonile toetuva hinnangu puuduseks on vähene usaldusväärsus. Katseline lähenemine on küll tegelikule keelekasutusele lähemal, kui seda võimaldab üksnes uurija intuitsiooni usaldamine, aga nii on võimalik verbe uurida ainult piiratud raamides, mistõttu on see aja- ja ressursikulukas meetod. Lähenedes kasutuskontekstide kaudu, põrkume keelekasutuse suure variatiivsusega. Esiteks ei pruugi olla selge,

millised lause muud elemendid on uuritava verbi seisukohalt olulised. Nii võivad jääda olulised tunnused analüüsi võtmata, samal ajal kui väheolulised tunnused pälvivad liigselt uurija tähelepanu. Teiseks on loomuliku keele lausetes verbiliigituse mõttes palju müra, st nende järgi klassifitseerimine ei pruugi anda puhtaid klasse. See omakorda võib olla tingitud sellest, et verbi-semantic on tavaliselt kompleksne ning ka verbiga seotud võimalike lausemustrite hulk võib olla suur, kusjuures lausemustreid mõjutavaid tegureid on palju. Kõiki tegureid ei pruugi olla võimalik kindlaks teha, veel vähem auto- maatsesse klassifitseerimisse kaasata.

Peale selle ei ole selge taksonoomia tavaliselt võimalik seetõttu, et verbide ja verbitähenduste liigitamisel võib lähtuda väga erinevatest semantilistest tunnustest. Kokkuvõttes toob see kaasa selle, et näiteks ühte ja sama verbi võib iseloomustada ning tihti iseloomustabki mitmesse alamklassi kuulumine (vt ka Levin 1993; Pajunen 2001; Labi 2006a, 2006b; Taremaa 2017). Samuti teevad verbide liigitamise keeruliseks fraasiverbid, näiteks võib olla keeruline otsustada, kas lihtverb ja sama verb ühendverbi koosseisus peaksid viitama kuulumisele eri verbiklassidesse või mitte.

Isegi võrdlemisi lihtsad verbid võivad olla klassifitseerimise mõttes keerulised, kuna nende põhitähendus võib koosneda mitmest selgelt eristatavast tähendusjoonest. Näiteks võib tuua verbi *ukerdama*: liikumisverbile omaselt (vt ka Taremaa 2017) on selle verbi tähendus iseloomustatav mitme tunnuse järgi, kusjuures iga tunnus võib olla liigitusaluseks. Nii on *ukerdama* ühtaegu viisiverb, väljendades seda, kuidas liikumine toimub; vahendita liikumise verb; aeglast liikumist väljendav verb; raskusastme järgi rasket/raskendatud liikumist väljendav verb; tüüpilise liikuja järgi elusolendi liikumist väljendav verb; liikumise asukoha järgi pigem pinnal liikumist väljendav verb; ning agenttiivsuse järgi agenttiivse liikumise verb. Seega võiks verb *ukerdama* sõltuvalt liigitusalusest kuuluda samaaegselt vähemalt seitsmesse liikumisverbide alamklassi.

Niisiis võib klassifitseerimine olla keerukas juba ainuüksi siis, kui verbil on suhteliselt piiritletud tähendus. Nagu eespool kirjeldatud, on paljud tuumverbid polüseemsed (ühel verbil rohkem kui üks tähendus): näiteks verbil *saama* on EKSS-is 12 tähendust. Põhitähenduselt on see omandamist väljendav (nt *Sain isalt sünnipäevaks nukumaja*). Põhitähendus omakorda võimaldab selle verbi kasutamist järgmistes tähendustes: õnnestumine (nt *Ta sai töötaja nõusse*), kogemine (nt *Saime näha suurepärase etendust*), muutumine (nt *Kindad said märjaks*) ja tulemuslikkus ehk resultatiivsus (nt *Artikkel sai valmis*) (Habicht, Tragel 2014; Tragel, Habicht 2017). Seega verbitähendusest liigitamist alustades võib *saama* olla omandamisverb, õnnestumisverb, kogemisverb, muutumisverb ja resultatiivverb. Lisaks on verbil *saama* ka väga sage modaaltähenduses kasutus, nt *Me ei saa homme teie juurde tulla*.

On koguni väidetud, et eesti keele verbidest on umbes kolmandik polüseemsed, samal ajal kui üldiselt polüseemiarikkaks peetavas inglise keeles on polüseemseid verbe vaid umbes viiendik (Langemets 2003: 103, 113). Niisiis on polüseemia kindlasti üks olulisemaid valdkondi, millega verbiliigituse juures tegelema peab. Sagedasemad verbid on lisaks laialt kasutatavad eri abstraktsetes valdkondades, mis suurendab võimalikku eri semantilistesse klassi-

desse kuulumist veelgi. Omaette küsimus on see, kuivõrd peaks polüseemne või ülekantud tähenduses kasutatud verb olema liigitatud eri klassidesse.

Kokkuvõttes on verbide semantiline liigitamine seotud komplekssete probleemidega. Probleemid omakorda on tingitud keelest endast – keel on paratamatult ääretult kompleksne ning mitmeteguriline (ingl *multifactorial*) nähtus. Komplekssuse haldamiseks, või isegi hindamiseks, on vaja rakendada analüüsimetodeid, mis on kas loodudki komplekssete nähtuste uurimiseks või mis annavad selliste nähtuste uurimisse väikse, aga olulise osa. Peatume verbiliigituse metodoloogilistel aspektidel järgnevas peatükis.

4. Meetod

Tänapäeva keeleteaduses on osutatud (Evans, Levinson 2009), et tegelikke keeleuniversaale (Greenberg 2005 [1966]) on äärmiselt vähe. Järelikult pole kuigivõrd usutav (ja kindlasti mitte tõestatud), et eri keelte verbiliigid võiksid olla täpselt samad. Nagu juba eespool mainitud, ei tohiks niisiis mis tahes keele kohta koostatud klassifikatsiooni otse teisele keelele üle kanda. Ent teiste keelte kohta koostatud verbiliigitused võivad olla väärtuslikuks allikaks verbide taksonoomia väljatöötamisel.

Teisalt on nii teadusliku lähenemise kui ka mõistelise organiseerituse seisukohalt paremini põhjendatud selline suund, milles keeleüksusi liigitama asudes ei võeta lähtekohaks mitte teise keele põhjal koostatud liigitust, vaid meetod (ja selle võimalik kriitika), mida liigitamisel on kasutatud. Nüüdisaegsed keeleteaduse meetodid võimaldavad statistiliste vahendite ja nende rakenduste abil uurida keele mis tahes mõistelist organiseeritust (sealhulgas verbiliigitust) ning teha selle jagunemise kohta usaldusväärseid järeldusi. Järgnevalt käsitlemegi seda, kuidas eri meetodite kombineerimisega luua verbide semantiline liigitus nii, et ühe keele mõistelised eripärad kaduma ei läheks, aga samas kasutataks võimalikult palju ära teiste keelte põhjal tehtud uurimuste tulemusi ja järeldusi. Järgmises peatükis vaatame näitena, kuidas verbitähenduseni jõuda lausemustrite(k) kaudu.

Usaldusväärse verbiklassifikatsiooni saamise eesmärgil on oluline rakendada eri meetodeid ning kasutada tegelikku keelematerjali. Seetõttu on meie lähenemises oluline koht induktiivsel töömeetodil, mis aga kombineerub deduktiivse lähenemisega. Teisisõnu, toetame kahesuunalise keeleanalüüsi kombinatsiooni (reaalsest keelematerjalist tõukunud verbide empiiriline liigitamine *vs.* olemasoleva intuitiivse või teiste keelte kohta tehtud verbiliigituse kontrollimine reaalse keelekasutuse põhjal).

Järgnevalt peatume sellel, kuidas meetodite kombineerimine (triangulatsioon) ning eri lähenemissuundade rakendamine saaks võimaldada verbide semantilist liigitamist. Seejuures peame oluliseks arvestada kahte põhilist aspekti:

1. Verbide semantiline liigitus peab kajastama loomulikke keelekasutust ja keelest arusaamist võimalikult hästi ja võimalikult vähese infoaoga, et verbiliigitusel oleks piisavalt alust ja usaldusväärset nii keeleteaduslikus mõttes kui ka näiteks keeleõppija jaoks.

2. Verbide empiiriline liigitamine peab olema reaalselt võimalik, st arvestama realistlikult ajakulu ja empiirilise analüüsi võimalustega (nt korpusete suurusega, analüüsimeetodite miinustega, keelematerjali kogumiseks ja märgendamiseks kuluva ajaga, katsete läbiviimiseks kuluva ajaga, katseteks vajaliku tehnika olemasoluga jne).

Sisuliselt tähendab see, et minimaalsete vahendite ja ajakuluga (mis empiirilise verbiliigituse ülesande puhul on siiski märkimisväärselt suur) tuleks jõuda liigituseni, mis – olgugi et teatavate münduste ja raskete otsustega – oleks võimalikult adekvaatne eesti keele verbide taksonoomia esitus.

5. Verbiliigituse tööprotsess ja võimalikud probleemkohad

Verbiliigituse aluseks olevate uurimuste tööprotsessi esitame skemaatiliselt tabelis 1, jagades tööprotsessi väiksemateks etappideks. Toome välja vastava etapi eesmärgi, vajalikud taustateadmised ja oskused, tegevused ning tulemused, samuti võimalikud probleemkohad.

Nagu iga teaduslik uurimus, peab ka verbiliigituseni viiva uurimuse iga tahk algama uurimisküsimuse püstimisega (1). Seejärel tuleb valida ja koguda materjal (2). Mõnel juhul võib korpusetest materjali kogumine olla keerukas või võimatu – siis tuleb kaaluda materjali kogumise muude empiiriliste viiside, näiteks katsete, kasutamist. Järgmine on kogutud materjali puhastamise ja korrastamise etapp (3), mille tulemusena saame märgendamiseks vajalikul kujul oleva sobiva materjalihulga. Järgneb kõige aja- ja töömahukam etapp: materjali märgendamine uurimisküsimusest lähtuvate relevantsete tunnustega (4). Selle etapi tulemus on edasiseks statistiliseks analüüsiks sobivad diskreeted, mündetavad üksused. Tunnused võivad olla semantilised ja vormilised, kusjuures üks või mitu märgendatavatest semantilistest tunnustest näitab verbi (põhi)tähendust. (Tunnuste märgendamist ja sellega seotut tutvustame lähemalt artikli peagi ilmavas järjes.) Järgneb statistiline andmeanalüüs (5). Verbide lausemustreid saab kirjeldada, kasutades näiteks kollokatsioonianalüüsi või tunnustepõhise (tuntud ka kui käitumisprofiili) analüüsi tehnikaid, sh klassifitseerimis- ja mudeldamistehnikaid (vt nt Geeraerts jt 1994; Baayen 2008; Gries, Divjak 2009; Tagliamonte, Baayen 2012; Levshina 2015). Pärast seda saame teha esimesed järeldused ning otsustada, kas andmeanalüüsiga on võimalik tuvastada lausemustreid. Järgmises etapis (6) saab asuda tulemusi tõlgendama ning verbirühmi määrama. Selles etapis võib selguda, et läbiviidud statistiline analüüs ei anna oodatud tulemusi ning tuleb kasutada teisi meetodeid (7). Sellise tööprotsessi tulemusena jõuame esmase verbiliigituseni.

Tabelis 1 on välja toodud ka iga etapi võimalikud probleemkohad. Laias laastus võib need jagada kaheks: esiteks keele mitmetegurilisus, mis muudab keerukaks statistilise analüüsi ning põhjustab infokadu mittediskreetsete üksuste diskreetseks muutmiseks; ning teiseks (osalt esimesest tingitud) tulemuste ületõlgendamise oht.

Esimest liiki probleemid ilmnevad eelkõige märgendamise etapis, mis on, nagu eespool mainitud, protsessi kõige töö- ja ajamahukam etapp. Mis tahes

Tabel 1.

Verbide liigitamise praktiline tööprotsess

etapid	1	2	3	4	5	6	7
eesmärgid	uurimisküsimuse püstistamine ja sõnastamine	materjali kogumine korpustest; materjali kogumine muul viisil (nt katsega)	materjali puhastamine ja korrastamine	materjali märgendamine	statistiline analüüs	tulemuste tõlgendamine, verbiühmade määramine	esmade analüüsitulemuste valideerimine teiste meetoditega
vajalikud taustateadmised ja oskused	lingvistilis-tüpoloogilised teadmised; teadmised meetodi võimaluste kohta	oskus korpus-päringut koostada; eksperimenditaalsete meetodite tundmine	oskus vastavalt uurimis-küsimusele materjali puhastada	oskus vastavalt muuseumile uuritava keele struktuuri tundmine	statistiliste vahendite tundmine ja kasutamisoskus	lingvistilis-tüpoloogilised teadmised; teadmised meetodi võimaluste ja puuduste kohta; tõlgendusoskus	eri meetodite (sh korpus- ja katseliste meetodite) tundmine
tegevused	kirjanduse analüüs uurimisküsimuse püstistamiseks ja lahendusvõimaluste leidmiseks; keelematerjali intuiitvne analüüs	korpuspäringu koostamine; eksperimentide kavandamine ja läbiviimine	eesmärgi piiritlemine ja fokuseerimine	automaatne (mõned vormilised parameetrid) ja käsitsi märgendamine	eri statistiliste vahendite rakendamine	statistilise analüüsi tulemuste tõlgendamine, vajadusel uute tunnuste kaasamine jne	sama materjali analüüs teise meetodiga
tulemused	mõistliku ajahulga ja olemasolevate meetoditega teostatav selgelt sõnastatud uurimisküsimus	sobiv hulk korpuslauseid (üksusi); katsega kogutud keelematerjal	märgendamiseks sobiv materjal	diskreetset üksused edasiseks (statistiliseks) analüüsiks	statistiliselt analüüsitud andmestik	lausemustrid, kombinatsioonid, korrelatsioonid	verbiühituseks vajalikud lähteandmed
võimalikud probleemid	liiga lai ja tehniliselt teostatamatult uurimust eeldav uurimisküsimus	corpuse eba-piisavus (korpus pole representatiivne); katselise meetodi sisuline piiratus	oht liigse puhastamise ja mugavusega midagi olulist materjalist välja jätta	infokadu, mis tuleneb mitte-diskreetsete kategooriate diskreetseteks muutmise; oluliste tunnuste (nt leksi-kaalse aspekti) mittemärgendamise nähtuse keerukuse tõttu, märgendamise võimatus või liigne ajamahukus	meetodite puudujäägid keelematerjali analüüsiks, liiga vähe materjali (lahendus: juurde märgendada)	tulemuste võimalik ületõlgendamine	ületõlgendamine, korpuste mitte-representatiivsus, meetodite puudujääkidega mitte-arvestamine

statistilise analüüsi eelduseks on diskreetsete üksuste olemasolu. Kui diskreetset üksust peab looma nähtustest, mis on olemuslikult mittediskreetset, toob see kaasa paratamatu infokao. Infokadu tuleb kindlasti üritada minimeerida, kuid keele kui olemuslikult mitmetahulise nähtuse puhul on mõningane infokadu loomulik ja paratamatu. Eksperimendid võimaldavad uurida keele nähtust piiratud tingimustes ning kontrollitud situatsiooni saavutamiseks tuleb enamasti infokadu juba katsekavandisse sisse arvestada.

Teise ohukoha – andmete ületõlgendamise – juures tasub rõhutada, et keele olemuslikult intuiitiivne loomus lubab keele uurimisel kasutada intuitsiooni vabamalt kui reaalteadustes. Verbide (ja mis tahes teiste keeleüksuste) uurimisel on intuitsiooni ja tüpoloogiliste teadmiste kombineerimine oluline juba tööprotsessi esimeses etapis uurimisküsimust formuleerides, rääkimata keelematerjali valimisest, analüüsimisest ja statistilise analüüsi jaoks ette valmistamisest. Tõdegem, et kõigele vaatamata on enamasti võimalik (deduktiivselt) määrata verbi üldine liik (näiteks verbide intuiitiivne jagamine asendi-verbideks, nägemisverbideks, hääleverbideks jne), toetudes seejuures intuitsioonile, sõnaraamatutele või varasematele teaduslikele käsitlustele. Oluline on tõlgendamise käigus uuesti läbi mõelda ka kasutatud andmestiku representatiivsus: kas nende andmete põhjal saab tegelikult teha selliseid järeldusi, nagu statistilise andmeanalüüsi tulemused võimaldavad?

6. Kokkuvõtteks

Artiklis on astutud esimesi samme teel, mille lõppsiht on verbide semantilise liigituse teaduslik metodika, mis lähtub tänapäeva keeleteaduse kasutada olevatest vahenditest ning võtab arvesse varasemaid liigitusi ja käsitlusi. Artikli esimeses pooles on lühidalt tutvustatud nii maailma kui ka Eesti vastavateemalisi allikaid. Artikli põhiosas keskendusime verbiliigituse koostamise protsessile ja võimalikele probleemkohtadele.

Verbitähenduseni on võimalik jõuda nii verbi enda tähendust kui ka lausemustrite tähendust ja ülesehitust uurides. Mõlemal lähenemisel on oma eelised ja puudused ning omaette küsimus on seegi, kui võrd on üldse võimalik verbitähendust tõlgendada või kindlaks teha, võtmata arvesse lausemustreid, kus see verb esineb. Verbi tähendust määrates peame aktsepteerima, et intuitsiooni kasutamine on õigustatud. Lausemustreid tõlgendades tuleb aga ettevaatlik olla põhjendamatu ületõlgenduste suhtes. Võimalikke probleemkohti on teisigi. Oluline on kasutada verbiliigituse loomise eesmärgil kõiki praegusi teadmisi verbidest ning keeleuurimise meetoditest kogu oma laias amplituudis ning kõikvõimalikes kombinatsioonides, mis kokkuvõttes tähendab meetodite pluralismi ja triangulatsiooni. Teisisõnu tähendab see mitmete meetoditega sama asja uurimist ning saadud uurimistulemuste valideerimist uute meetoditega (Creswell, Plano Clark 2007; Arppe jt 2010). Kuidas selline analüüs praktilises keeleuurimises välja näeb, tutvustamegi artikli järeles.

Artikli valmimist on toetanud Euroopa Liidu Regionaalarengu Fond (Eesti-uuringute Tippkeskus).

Kirjandus

- Arppe, Antti, Gilquin, Gaëtanelle, Glyn n, Dylan, Hilpert, Martin, Zeschel, Arne 2010. Cognitive corpus linguistics: Five points of debate on current theory and methodology. – *Corpora*, kd 5, nr 1, lk 1–27.
- Baayen, Harald R. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.
- Biber, Douglas, Conrad, Susan, Reppen, Randi 1998. *Corpus Linguistics: Investigating Language Structure and Use*. (Cambridge Approaches to Linguistics.) Cambridge: Cambridge University Press.
- Bolinger, Dwight Le Merton 1968. *Aspects of Language*. New York–Chicago–San Francisco–Atlanta: Harcourt, Brace & World.
- Creswell, John W., Plano Clark, Vicki L. 2007. *Designing and Conducting Mixed Methods Research*. Thousand Oaks–London–New Delhi: SAGE Publishing.
- EKG I = Mati Erelt, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael, Silvi Vare. *Eesti keele grammatika I. Morfoloogia*. Sõnamoodustus. Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut, 1995.
- EKG II = Mati Erelt, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael, Silvi Vare. *Eesti keele grammatika II. Süntaks*. Lisa: kiri. Tallinn: Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut, 1993.
- Erelt, Mati, Metslang, Helle (toim) 2017. *Eesti keele süntaks*. (Eesti keele varamu III.) Tartu: Tartu Ülikooli Kirjastus.
- Evans, Nicholas, Levinson, Stephen C. 2009. The myth of language universals: Language diversity and its importance for cognitive science. – *Behavioral and Brain Sciences*, kd 32, nr 5, lk 429–448.
- Fillmore, Charles J. 1968. The case for case. – *Universals in Linguistic Theory*. Toim Emmon Bach, Robert T. Harms. New York: Holt, Rinehart & Winston.
- Fillmore, Charles 1982. *Frame semantics*. – *Linguistics in the Morning Calm: Selected papers from SICOL-1981*. Linguistic Society of Korea. Seoul: Hanshin, lk 111–137.
- Fried, Mirjam, Östman, Jan-Ola 2004. *Construction grammar: A thumbnail sketch*. – *Construction Grammar in a Cross-Language Perspective*. (Constructional Approaches to Language 2.) Toim M. Fried, J-O. Östman. Amsterdam–Philadelphia: John Benjamins Publishing Company, lk 11–86.
- Geeraerts, Dirk, Grondelaers, Stefan, Bakema, Peter 1994. *The Structure of Lexical Variation: Meaning, Naming, and Context*. (Cognitive Linguistics Research 5.) Berlin–New York: Mouton de Gruyter.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure*. Chicago–London: The University of Chicago Press.
- Goldberg, Adele 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford–New York: Oxford University Press.
- Goldberg, Adele E. 2010. *Verbs, constructions and semantic frames*. – *Lexical Semantics, Syntax, and Event Structure*. Toim Malka Rappaport Hovav, Edit Doron, Ivy Sichel. Oxford: Oxford University Press, lk 39–58.
- Greenberg, Joseph H. 2005 [1966]. *Language Universals*. With a preface by Martin Haspelmath. Berlin: Walter de Gruyter.

- Gries, Stefan Th., Divjak, Dagmar 2009. Behavioral profiles: A corpus-based approach to cognitive semantic analysis. – *New Directions in Cognitive Linguistics*. Toim Vyvyan Evans, Stéphanie Pourcel. (Human Cognitive Processing.) Amsterdam–Philadelphia: John Benjamins Publishing Company, lk 57–75.
- Habicht, Külli, Penjam, Pille, Tragel, Ilona 2010. Kas *tahtma* tahab abiverbiks? – *ESUKA/JEFUL*, kd 1, nr 2, lk 115–146.
- Habicht, Külli, Tragel, Ilona 2014. Verbiga *saama* väljendatud leksikaalsed kategooriad ja konstruktsioonid. – *Keel ja Kirjandus*, nr 11, lk 826–844.
- Hanks, Patrick 1996. Contextual dependency and lexical sets. – *International Journal of Corpus Linguistics*, kd 1, nr 1, lk 75–98.
- Harris, Zellig S. 1954. Distributional structure. – *Word*, kd 10, nr 2–3, lk 146–162.
- Haspelmath, Martin 2012. Escaping ethnocentrism in the study of word-class universals. – *Theoretical Linguistics*, kd 38, nr 1–2, lk 91–102.
- Jentson, Indrek 2013. Eesti VerbNet'i loomise võimalikkusest. – *Eesti Raken-
duslingvistika Ühingu aastaraamat*, kd 9, lk 75–83.
- Kaalep, Heiki-Jaan 2015. Eesti verbi vormistik. – *Keel ja Kirjandus*, nr 1, lk 1–15.
- Kasik, Reet 2015. *Sõnamoodustus*. (Eesti keele varamu I.) Tartu: Tartu Ülikooli Kirjastus.
- Kerner, Kadri, Orav, Heili, Parm, Sirli 2010. Growth and revision of Estonian wordnet. – *Principles, Construction and Application of Multilingual Wordnets: Proceedings of the 5th Global Wordnet Conference*. Toim Pushpak Bhattacharyya, Christiane Fellbaum, Piek Vossen. Mumbai: Narosa Publishing House, lk 198–202.
- Kipper Schuler, Karin 2005. *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. PhD dissertation. Pennsylvania: University of Pennsylvania.
- Klaas, Birute 1999. Dependence of the object case on the semantics of the verb in Estonian, Finnish and Lithuanian. – *Estonian: Typological Studies III*. Toim Mati Ereht. (Tartu Ülikooli eesti keele õppetooli toimetised 11.) Tartu: Tartu Ülikooli Kirjastus, lk 47–83.
- Kuznetsova, Julia 2015. *Linguistic Profiles: Going from Form to Meaning via Statistics*. (Cognitive Linguistics Research 53.) Berlin–Boston: De Gruyter Mouton.
- Labi, Kanni 2006a. *Eesti regilaulude verbisemantika*. (Dissertationes philologiae Estonicae Universitatis Tartuensis 18.) Tartu: Tartu Ülikooli Kirjastus.
- Labi, Kanni 2006b. *Kui saan, virve, veeremaie*. Liikumise väljendamine regilaulus. – *Keel ja Kirjandus*, nr 7, lk 538–556.
- Langetse, Margit 2003. *Polüseemia ja leksikograafia*. – *Emakeele Seltsi aastaraamat* 49 (2003). Toim Mati Ereht. Tallinn: Teaduste Akadeemia Kirjastus, lk 97–124.
- Lečič, Rada 2012. *Basic Grammar of the Slovene Language: Language Manual*. Cerklje: Gaya.
- Levin, Beth 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago–London: University of Chicago Press.
- Levshina, Natalia 2015. *How to do Linguistics with R: Data Exploration and Statistical Analysis*. Amsterdam–Philadelphia: John Benjamins Publishing Company.
- McEnery, Anthony M., Wilson, Anita 2001. *Corpus Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.

- Metslang, Helle 2017. Sihitis. – Eesti keele süntaks. Toim Mati Erelt, H. Metslang. (Eesti keele varamu III.) Tartu: Tartu Ülikooli Kirjastus.
- Muischnek, Kadri 2006. Verbi ja noomeni püsiühendid eesti keeles. (Dissertationes philologiae estonicae Universitatis Tartuensis 17.) Tartu: Tartu Ülikooli Kirjastus, lk 258–277.
- Muischnek, Kadri, Kaalep, Heiki-Jaan 2010. The variability of multi-word verbal expressions in Estonian. – Language Resources and Evaluation, kd 44, nr 1–2, lk 115–135.
- Orav, Heili, Kerner, Kadri, Parm, Sirli 2011. Eesti Wordnet'i hetkeseisust. – Keel ja Kirjandus, nr 2, lk 96–106.
- Orav, Heili, Zupping, Sirli, Vare, Kadri 2015. Leksikosemantiliste suhete hägusus Eesti Wordnetis. – Emakeele Seltsi aastaraamat 60 (2014). Toim Mati Erelt. Tallinn: Teaduste Akadeemia Kirjastus, lk 171–193.
- Pajunen, Anneli 2001. Argumenttirakenne. Asian tilojen luokitus ja verbien käyttäytyminen suomen kielessä. (Suomi 187.) Helsinki: Suomalaisen Kirjallisuuden Seura.
- Pajusalu, Renate 2001. Kas *moos* ja *buss* seisavad sarnaselt ehk väike katse verbiga *seisma*. – Keele kannul. Pühendusteos Mati Ereli 60. sünnipäevaks. Toim Reet Kasik. (Tartu Ülikooli eesti keele õppetooli toimetised 17.) Tartu: Tartu Ülikooli Kirjastus, lk 250–274.
- Pajusalu, Renate, Tragem, Ilona 2007. Word and construction as units of categorization: The case of change predicates in Estonian. – Mental States. Kd 2: Language and Cognitive Structure. Toim Andrea C. Schalley, Drew Khlentzos. (Studies in Language Companion Series 93.) Amsterdam–Philadelphia: John Benjamins Publishing Company, lk 289–310.
- Penjam, Pille 2008. Eesti kirjakeele *da-* ja *ma-*infinitiiviga konstruktsioonid. (Dissertationes philologiae estonicae Universitatis Tartuensis 23.) Tartu: Tartu Ülikooli Kirjastus.
- Rätsep, Huno 1978. Eesti keele lihtlausete tüübid. (Eesti NSV TA Emakeele Seltsi toimetised 12.) Tallinn: Valgus.
- Tagliamonte, Sali A., Baayen, R. Harald 2012. Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. – Language Variation and Change, kd 24, nr 2, lk 135–178.
- Tamm, Anne 2007. Perfectivity, telicity and Estonian verbs. – Nordic Journal of Linguistics, kd 30, nr 2, lk 229–255.
- Tamm, Anne 2012. Scalar verb classes: Scalarity, thematic roles, and arguments in the Estonian aspectual lexicon. (Biblioteca di Studi di Filologia Moderna 14.) Firenze: Firenze University Press.
- Taremaa, Piia 2017. Attention Meets Language: A Corpus Study on the Expression of Motion in Estonian. (Dissertationes linguisticae Universitatis Tartuensis 29.) Tartu: University of Tartu Press.
- Tauli, Valter 1973. Standard Estonian grammar. Part I: Phonology, Morphology, Word-formation. (Acta Universitatis Upsaliensis. Studia Uralica et Altaica Upsaliensia 8.) Uppsala: Almqvist & Wiksell.
- Tauli, Valter 1983. Standard Estonian grammar. Part II: Syntax. (Acta Universitatis Upsaliensis. Studia Uralica et Altaica Upsaliensia 14.) Uppsala: Borgströms Tryckeri AB.

- Tragel, Ilona 2003. Eesti keele tuumverbid. (Dissertationes linguisticae Universitatis Tartuensis 3.) Tartu: Tartu Ülikooli Kirjastus.
- Tragel, Ilona, Habicht, Külli 2017. *Saama*-verb grammatilistes konstruktsioonides. – Keel ja Kirjandus, nr 1, lk 22–40.
- Tragel, Ilona, Klavan, Jane (valmimas). Abstract verbs in free form drawing task.
- Uiboaed, Kristel 2013. Verbiühendid eesti murretes. (Dissertationes philologiae estonicae Universitatis Tartuensis 34.) Tartu: Tartu Ülikooli Kirjastus.
- Viks, Ülle 2001. Muuttüübid eesti sõnastikes. – Leksikograafiaseminar: Sõna tänapäeva maailmas. Leksikografinen seminaari: Sanat nykymaailmassa. Toim Margit Langemets. (Eesti Keele Instituudi toimetised 9.) Tallinn: Eesti Keele Sihtasutus, lk 157–185.
- Õim, Haldur 1965. *Tulema, saama ja pidama* tähenduste struktuuriline analüüs. – Keel ja struktuur 1. (Töid struktuuralse ja matemaatilise lingvistika alalt.) Tartu: TRÜ, lk 27–45.
- Õim, Haldur 1970. Eesmärk, taotlema, saavutama, tulemus. Semantiline analüüs. – Keel ja struktuur 4. (Töid struktuuralse ja matemaatilise lingvistika alalt.) Tartu: TRÜ, lk 71–108.
- Õim, Haldur, Orav, Heili, Taremaa, Piia 2009. Lihtlause semantika: teoreetiline kontseptsioon ja arvutianalüüsi võimalused. – Keel ja Kirjandus, nr 7, lk 489–504.
- Õim, Haldur, Orav, Heili, Taremaa, Piia 2010. Lausesemantikast üldkeeleteaduse ja eesti keele kontekstis. – Emakeele Seltsi aastaraamat 55 (2009). Toim Mati Ereht. Tallinn: Teaduste Akadeemia Kirjastus, lk 201–224.
- Õim, Haldur, Saluveer, Madis 2002. Freimid keelekirjelduses. – Akadeemia, nr 12, lk 2663–2678.

Võrguviited

- Eesti keele Wordnet. <http://www.cl.ut.ee/ressursid/teksaurus/?lang=et> (8. II 2018).
- EKSS = Eesti keele seletav sõnaraamat. Veebiversioon. <http://www.eki.ee/dict/ekss/index.cgi> (7. II 2019).
- FrameNet. <https://framenet.icsi.berkeley.edu/fndrupal/> (7. II 2019).
- German FrameNet. <http://www.laits.utexas.edu/gframenet/> (7. II 2019).
- Verbs Index. <https://verbs.colorado.edu> (7. II 2019).
- VerbNet. <http://verbs.colorado.edu/~mpalmer/projects/verbnet.html> (7. II 2019).
- WordNet. <http://globalwordnet.org/wordnets-in-the-world/> (7. II 2019).
- ÕS 2018 = Eesti õigekeelsussõnaraamat ÕS 2018. Veebiversioon. <https://www.eki.ee/dict/qs/> (14. II 2019).

Ilona Tragel (snd 1968), PhD, Tartu Ülikooli eesti ja üldkeeleteaduse instituut, üldkeeleteaduse dotsent (Jakobi 2, 51005 Tartu), ilona.tragel@ut.ee

Piia Taremaa (snd 1982), PhD, Tartu Ülikooli eesti ja üldkeeleteaduse instituut, üldkeeleteaduse teadur (Jakobi 2, 51005 Tartu), piia.taremaa@ut.ee

Semantic classification of Estonian verbs

Keywords: verb, semantics, polysemy, construction, categorisation, linguistic methods

In this paper, the methodology of verb classification is addressed. The main aim is to provide an overview of the possible ways verbs and sentence patterns could be analysed for the purpose of verb classification, and to reveal the possible shortcomings of different methods and problems of verb classification in general. For that purpose, the rare verb classifications available for the languages of the world (mainly English) are introduced focusing on the methodological aspect. In addition, an overview is given of the treatment of verbs as belonging to several semantic classes in Estonian linguistic tradition. Finally, the possible research paths that lead to verb classification are discussed. The two main points of the discussion are as follows. (1) There are two main alternatives to reach a semantic classification of verbs. One is classifying verbs based on their core meaning and disregarding contextual information as much as possible. Such a classification can be based either on the intuition of a (single) researcher or on experiments. The other is classifying verbs based on their sentence patterns, assuming that this is where the semantic differences between verbs are ultimately reflected. Such an approach needs massive (corpus) data and the application of multivariate statistics. (2) The main problems associated with verb categorisation are related to the semantic complexity of the verbs (incl. polysemy), the multifactorial nature of language, disadvantages of particular methods that could be used for verb classification, and difficulties with data coding (e.g., coding non-discrete units as discrete). In summary, we propose that a blend of different approaches to verb meaning – including intuition based approaches, experiments, and corpus analysis – is needed for categorising verbs into semantic classes. Thus, the paper advocates methodological pluralism and triangulation.

Ilona Tragel (b. 1968), PhD, University of Tartu, Institute of Estonian and General Linguistics, Associate Professor of General Linguistics (Jakobi 2, 51005 Tartu), ilona.tragel@ut.ee

Piia Taremaa (b. 1982), PhD, University of Tartu, Institute of Estonian and General Linguistics, Research Fellow in General Linguistics (Jakobi 2, 51005 Tartu), piia.taremaa@ut.ee