

TARTU ÜLIKOOL
HUMANITAARTEADUSTE JA KUNSTIDE VALDKOND
EESTI JA ÜLDKEELETEADUSE INSTITUUT

Anastasia Šepilenko

Umbmääraste hulgasõnade valikud proportsioonide kirjeldamisel eesti keeles

Magistritöö

Juhendajad Maria Reile (PhD) ja Renate Pajusalu (PhD)

TARTU 2025

Autorsuse kinnitus

Kinnitan, et olen käesoleva lõputöö ise kirjutanud ning toonud korrektselt välja teiste autorite panuse. Töö on kirjutatud, lähtudes Tartu Ülikooli eesti ja üldkeeleteaduse instituudi lõputöö nõuetest, ning on kooskõlas heade akadeemiliste tavadega.

Anastasia Šepilenko

Lühikokkuvõte

Siinses katses uuriti üheksa eestikeelse hulgasõna valikuid erinevate objektide proportsioonide kirjeldamisel. Tööle oli seatud kaks peamist uurimisküsimust. Esiteks, kas ja milliseid eestikeelseid hulgasõnu valitakse objektide hulga erinevate proportsioonide kirjeldamisel visuaalsete stiimulite puhul. Teiseks, kas ja kuivõrd võiksid vastajast tulenevad tegurid mõjutada hulgasõnade valikut küsitud objektide koguse kirjeldamisel. Katsemeetodina kasutati sunnitud valiku katset, millega paluti katses osalejatel valida üks sõna kas siniste või kollaste pallide hulga iseloomustamiseks. Hulgasõnad paiknesid kas formaalloogilisel (*mitte ükski, mõni, mõned, pooled, kõik*) või hinnanguskaalal (*vähe, keskmiselt, palju*). Katsega selgitati välja, et eestikeelsed hulgasõnad vastavad teatud proportsioonide vahemikele, kuid nende vahel ei esine selgeid piire. Vahemikud, kus valiti hulgasõna proportsioonide kirjeldamiseks, kattusid samuti omavahel suurelt. Erandiks jäi vaid hulgasõna *mitte ükski*, mida valiti ainult objekti 0% proportsiooni kirjeldamisel. Hulgasõnade valikul osutusid oluliseks mõjuteguriks ja ennustajateks samuti kirjeldavate objektide proportsioonid. Vastajaspetsiifilised tegurid hulgasõnade valikuid ei mõjutanud ega osutunud ennustamisel olulisteks.

Võtmesõnad: hulgasõnad, proportsioonid, skaalad, sunnitud valiku katse, vastajaspetsiifilised tegurid, semantiline kattumine

Sisukord

Sissejuhatus	7
1. Teoreetiline taust ja ülevaade	9
1.1. Umbmääraste hulgasõnade definitsioon ja iseloomustus	9
1.2. Hulgasõnade võimalikust paiknemisest mentaalsetel skaaladel	10
1.3. Tähelepanekud hulgasõnade süntaksist	11
1.4. Konteksti mõju hulgasõnade tajumisele ja valikule	12
1.5. Vastajaspetsiifiliste tegurite mõju hulgasõnade valikule	14
1.5.1. Erinevate sotsiaaldemograafilised tegurite võimalik mõju	14
1.5.2. Väsimustase ja selle võimalik mõju	16
2. Uuritavate umbmääraste hulgasõnade tähendustest	17
2.1. Skaala <i>mitte ükski, mõni, mõned, pooled, enamik</i> ja <i>kõik</i> sõnade semantikast	17
2.1.1. <i>Ükski</i> ja <i>mitte ükski</i>	18
2.1.2. <i>Mõni</i> ja <i>mõned</i>	18
2.1.3. <i>Pool</i>	19
2.1.4. <i>Enamik</i>	20
2.1.5. <i>Kõik</i>	21
2.2. Skaala <i>vähe, keskmiselt</i> ja <i>palju</i> sõnade semantikast	22
2.2.1. <i>Vähe</i>	22
2.2.2. <i>Keskmiselt</i>	23
2.2.3. <i>Palju</i>	23
3. Teiste uurimuste tulemused hulgasõnade ja proportsioonide vastavuse uurimisel	25
4. Meetod	28
4.1. Katse koostamine ja disain	28
4.1.1. Katse stiimulite koostamine ning hulgasõnade jaotused	28
4.1.2. Uuritavate hulgasõnade paigutus skaaladel	32

4.1.3. Katse pilootimine	32
4.2. Katse ülesehitus.....	33
4.2.1. Katse lühike iseloomustus	33
4.2.1. Katse struktuur	34
4.2.2. Katse protseduur.....	35
4.3. Osalejate valim.....	36
4.3.1. Katses osalejate kriteeriumid	36
4.3.2. Katses osalejate värbamine ja andmete säilitamine	36
4.3.3. Saadud andmestiku puhastamine	36
4.3.4. Ülevaade katses osalejatest	37
4.4. Andmete analüüsimetodid.....	39
5. Katse tulemused ja analüüs.....	41
5.1. Ülevaade hulgasõnade valikutest katses.....	41
5.1.1. Skaala <i>mitte ükski</i> – ... – <i>kõik</i> hulgasõnade valimissagedused	41
5.1.2. Skaala <i>vähe</i> – ... – <i>palju</i> hulgasõnade valimissagedused	46
5.1.3. Ülevaade katses osalejate vastamiskiiruse kohta	49
5.1.4. Vahepealne arutelu	52
5.2. Hulgasõnade valikut mõjutavate tegurite analüüs.....	54
Kokkuvõte	62
Kirjandus	64
On the selection of Estonian vague quantifiers in proportion description. Summary	76
Lisa 1. Näidis katses kasutatud pilt-stiimulitest	78
Lisa 2. Katsesse sissejuhatav tekst.....	79
Lisa 3. Nõusoleku vorm katses osalemiseks	80
Lisa 4. Taustainformatsiooni küsimused.....	82
Lisa 5. Näidis harjutusvooru stiimulist koos harjutusskaalaga.....	84
Lisa 6. Näidis katsevoorude stiimulist koos uuritava skaalaga	85

Lisa 7. Juhised enne harjutusvooruga alustamist	86
Lisa 8. Juhised enne esimese katsevooruga alustamist	87
Lisa 9. Juhised enne teise katsevooruga alustamist	88
Lisa 10. Värbamiskutse	89

Sissejuhatus

Käesoleva magistritöö eesmärk on välja selgitada nn umbmääraste hulgasõnade valikut mõjutavad tegurid eesti keeles. Selleks viisin läbi sunnitud valiku katse, kus palusin katseisikul valida ette antud hulgasõnade seast nende arust kõige sobivama hulgasõna kas siniste või kollaste pallide hulga kirjeldamiseks. Katsega uurisin eesti emakeelega osalejaid vanuses 18 kuni 30 eluaastast.

Eesti keeles on hulgasõnu varasemalt vaadeldud formaalsemantika vaatenurgast (Paldre 1997), uuritud on ka nende kasutust alates 19.sajandist (Lõbu 2011) ja igapäeva kõnes (Pajusalu jt 2004; Pajusalu 2008). Samuti on uuritud hulgasõnade *mõned* ja *kõik* mõistmist lastekeeles (Truus 2015). Hulgasõnade ja arvu vastamist on vaadeldud ka Maria Reile & Mariann Proos (2022). Siiani ei ole aga teada, kuidas tõlgendatakse või kasutatakse erinevaid eestikeelseid hulgasõnu mingi objektide hulga ja ka selle osakaalu suhte ehk proportsiooni kirjeldamisel. Samuti ei ole teada, kuidas tõlgendatakse ja valitakse hulgasõnu juhul, kui hulgasõna esineb üksinda, ilma lauselise ja fraasilise kontekstita. Seetõttu on minu töö peamiseks uurimisküsimuseks, kas ja milliseid eestikeelseid hulgasõnu valitakse objektide hulga erinevate proportsioonide kirjeldamisel visuaalsete stiimulite puhul.

Hulgasõnade valikute varieerumist vastavalt objektide hulga osakaalule on uuritud ka vastavalt konteksti muutustele (nt Newstead, Pollard & Riezebos 1987; Newstead & Coventry 2000), mille mõju võib oletada ka eestikeelsete hulgasõnade valimise puhul. Minu teada on vähe uurimusi (nt Reile & Šepilenko 2024), mis käsitleksid vastajast endast tulenevate erinevuste mõju hulgasõnade tõlgendamisele ja valikule eesti keeles. Seetõttu ei ole nende tegurite olulisusest ja mõjust eestikeelsetele hulgasõnadele väga midagi teada. Mitmed uurimused näitavad (nt Dąbrowska 2012, Wirth jt 2007), et erinevad sotsiaaldemograafilised tegurid (nt omandatud keeled või haridustase) ja ka nt väsimusaste avaldavad erinevat mõju keelelise tajule, siinkohal ka sõnade kasutusele ja valikule. Seega on minu teine uurimisküsimus, kas ja kuivõrd võiksid vastajast tulenevad tegurid mõjutada ka eestikeeleste hulgasõnade valikut küsitud objektide koguse kirjeldamisel visuaalsete stiimulite puhul.

Lisaks kogusin katses osalejate andmeid nende vastamiskiiruse ehk reaktsioonaja (*reaction time*) kohta, mis peegeldavad erinevate keeletöötlusega seotud protsesse (Vorweg 2012: 363; Kaiser 2013: 137–138). Hulgasõnade uurimisel on Stefan Heim jt (2012) täheldanud, et peamise hindava objekti hulga lähenemisel mingile taustal olevale objektide hulgale võiksid

vastamisaja tulemused suureneva töötlemise keerukuse tõttu. Seda võiks oletada ka eestikeelsete hulgasõnade valikute kohta. Katse formaadi tõttu ei saa vastamisaega käsitleda tõetruu parameetrina (vt Vorweg 2012; Kaiser 2013), kuna tegemist ei ole kontrollitud keskkonnaga ning võimalikest segavatest teguritest või muudest kaasnevatest asjaoludest pole midagi teada. Siiski nt märkimisväärne või keskmisest suurem ajaline näitaja võib kaudselt viidata hindamis- ning seekaudu ka sobiva sõna valimisraskustega seotud protsessidele. Silmapaistvad tendentsid või eripärad võivad omakorda olla aineks edasistele uurimustele kontrollitud keskkonnas.

Magistritöö koosneb kolmest peamisest osast: teoreetilisest osast (peatükk 1 kuni 3), meetodi kirjeldusest (peatükk 4) ning analüüsi ja tulemuste esitamises koos aruteluga (peatükk 5). Esimeses peatükis annan ülevaate teoreetilisest taustast, kus tutvustan hulgasõnade definitsiooni, omadusi ja paiknemist skaaladel, tähelepanekuid hulgasõnade süntaktilisest esinemisest ning teguritest, mis võivad mõjutada hulgasõnade tõlgendamist, nagu nt kontekst või vastajast tulenevad tegurid. Teises peatükis tutvustan kahel skaalal paiknevate uuritavate hulgasõnade tähendusi ning võimalikke proportsioonide vahemikke. Kolmandas peatükis tutvustan lühidalt teiste uurimusi, millega on avastatud erinevate indoeuroopakeelsete hulgasõnade vastamist teatud proportsioonide vahemikule. Neljandas peatükis annan ülevaate kasutatud meetodist, selle koostamisest ja ülesehitusest, osalejate valimist ja analüüsimeetoditest.

Viiendas peatükis esitan katsetulemusi ning arutlen nende üle. Peatükis 5.1 annan ülevaate hulgasõnade valimissagedusest ning arutatakse hulgasõnade valikute kohta saadud tulemuste üle. Lisaks vaatan ka vastamiskiiruse tulemusi iga uuritud proportsiooni kohta. Peatükis 5.2 annan ülevaate vastajast tulenevate tegurite mõjust hulgasõnade valikule proportsioonide kirjeldamisel.

1. Teoreetiline taust ja ülevaade

Siinses peatükis annan ülevaadet umbmääraste hulgasõnadega seotud teoreetilisest taustast. Peatükis toon umbmääraste hulgasõnade definitsiooni ja iseloomustuse ning käsitlen ka nende paiknemisest skaalal ja süntaktilisi omadusi. Samuti annan ülevaate konteksti olulisusest ning võimalikest vastajast tulenevatest teguritest, millest mõlemad võivad mõjutada umbmääraste hulgasõnade valikut.

1.1. Umbmääraste hulgasõnade definitsioon ja iseloomustus

Umbmäärased hulgasõnad (ingl *vague quantifier* või *Natural Language Quantifiers*; edasiselt hulgasõnad) on sõnad, mille abil märgitakse entiteetide¹ kogust. Sealjuures nende täpne arvuline hulk jääb teadmata. (Moxey & Sanford 2000: 237; Moxey 2011: 119; Metslang 2017: 463) Näiteks lause *mõni raamat on laual* kaudu ei saa kuulaja teada, kas *raamatuid* oli kokku kaks, kolm või viis. Seevastu suudab kuulaja siiski järeldada, et tegemist on (arvuliselt) pigem väikese kogusega. Kokkuvõttes ei ole midagi teada raamatute arvulise osakaalu kohta, kuid tänu mingile ühisele abstraktsele hulgasõna tähendusele on aimata raamatute vähesust.

Eesti keele süntaksis (2017; edaspidi EKS) ei ole umbmääraseid hulgasõnu eraldi käsitletud, kuid erinevate peatükkide järgi kuuluvad üldiselt hulgasõnade alla nt erinevad asesõnad (nt *mõni* või *kõik*) või hulgamäärsõnad (nt *vähe*, *palju*) (Veismann, Erelt & Metslang 2017: 352–357; Metslang 2017: 463). Hulgasõnadeks on peetud ka (põhi)arvsõnu (nagu nt *viis*) (Erelt 2017: 63; Pajusalu 2017: 383; Veismann, Erelt & Metslang 2017: 348–349), kuid neid ei saa minu arvates arvulise konkreetsuse tõttu pidada umbmäärasteks. Ingliskeelses teadusmaailmas on umbmäärasteks hulgasõnadeks nimetatud ka sagedussõnu (vt Pohl 1981; Newstead & Collis 1987; Schneider & Stone 2016), mida pole eesti keeles ei tavaliste ega umbmääraste hulgasõnade alla eelnevalt liigitatud.

Võib oletada, et umbmääraste hulgasõnade tõlgendamine on subjektiivne, kuna nendega saab inimene edastada lisaks ka oma uskumusi ja ootusi (Moxey & Sanford 2000; Yildirim jt

¹ Tegemist on kõikehõlmava ja abstraktse terminiga, mille all mõeldakse erinevaid kõikvõimalikke maailmas esinevaid objekte (ese, olend) või nähtusi (vt täpsemalt Langacker 2006: 50).

2016). Lisaks on hulgasõnade valik tingitud kontekstist ja selle erinevustest. Näiteks tavalises supikaussis 8 Kalevi kommi võivad magusaarmastaja jaoks olla *vähe*, kuid inimene, kes soovib suhkru tarbimist vältida, hindab antud kogust arvatavasti sõnaga *palju*. Kui supikauss oleks omakorda asendatud 5-liitrise supipotiga, oleks mõlema eelmainitud isiku tõenäoline hinnang *vähe* koos mõne võimaliku intensiivistajaga. Sobiva umbmäärase hulgasõna valimine jääb siinses olukorras kõneleja enda kanda vastavalt tema isiklikele uskumustele ja ootustele. Võib samuti oletada, et hulgasõnade tõlgendamist, kasutust ja valikut võivad mõjutada erinevad vastajast tulenevad tegurid, nagu näiteks väsimusaste või haridustase, mida tutvustan täpsemalt peatükis 1.5.

Lisaks vastab hulgasõna mingile vahemikule ehk *kasutusulatus*ele, millega saab viidata erinevatele objektide proportsioonidele. Hulgasõnade vastamist arvudele või numbrilisele süsteemile on välja pakkunud Bernard M. Bass, Wayne F. Cascio & Edward J. O'Connor (1974) ja Valeri F. Reyna (1981), kuid Linda M. Moxey (1986: 3, 27, 49) toob välja sellise lähenemise probleemkoha: hulgasõnad ei vasta ainult teatud kindlale proportsioonile, vaid esinevad mingis ulatuses (*range of proportions*), mis võivad inimeste vahel erineda. Hulgasõnade vastamist mitmest proportsioonist koosnevale vahemikule on pooldanud või leidnud ka teised (Newstead & Collis 1987; Moxey & Sanford 2000; Pezzelle, Bernardi & Piazza 2018). Kuna hulgasõnad väljendavad entiteedi ligikaudset kogust, on sellest oletatud, et nad moodustavad omaette lisa- n-ö teisejärgulise numbrilise süsteemi (*second/approximate number system*) (Pietroski jt 2009; Coventry jt 2010; Olm jt 2014). Magistritöös oletan, et igal uuritava hulgasõnal esineb oma ligikaudne kasutusulatus, millega saab edastada entiteetide hulga proportsioone.

1.2. Hulgasõnade võimalikust paiknemisest mentaalsetel skaaladel

On pakutud, et hulgasõnad ise võiksid paikneda erinevatel (dünaamilistel) skaaladel inimese mentaalses ruumis (vt nt Holyoak & Glass 1978; Moxey 1986; Newstead & Collis 1987; Routh 1994; Moxey & Sanford 1993; Yildirim jt 2013; Pezzelle, Bernardi & Piazza 2018; Abbondanza jt 2021). Tegemist on vaid oletusega, kuna pole teada, kuidas hulgasõnad on täpsemalt esindatud inimese mentaalses ruumis (Pezzelle, Bernardi & Piazza 2018: 118). Pakutud on hulgasõnade paiknemist ühemõõtmelisel skaalal (*unidimensional scale*) (Holyoak & Glass 1978), ordinaalskaalal (*ordinal scale*) (Newstead & Collis 1987) või järjestatud skaalal

(*ordered scale*), mille hulgasõnade jaotused ehk kasutusulatused ei ole võrdsete vahemikega (Pezzelle, Bernardi & Piazza 2018). Ka mina eeldan, et uuritavad hulgasõnad võiksid moodustada või paikneda teatud mentaalsetel skaaladel, mida tutvustan lähemalt peatükis 2. Kuna eesti keeles ei ole skaadel paiknevaid hulgasõnu uuritud, ei ole ka hulgasõnade kasutusulatustest nendel midagi teada.

1.3. Tähelepanekud hulgasõnade süntaksist

Katsega uurin üheksat eestikeelset hulgasõna, mille olen paigutanud kas formaalloogilisele (*mitte ükski, mõni, mõned, pooled* ja *kõik*) või hinnanguskaalale (*vähe, keskmiselt* ja *palju*). Katse koostamisel olen arvestanud hulgasõnade süntaktiliste omadustega. Järgnevalt tutvustan oma tähelepanekuid lähemalt iga skaala valimi kohta eraldi.

Skaala mitte ükski, mõni, mõned, pooled, enamik, kõik

Suur hulk skaalal toodud hulgasõnu on osa nimisõnafraasist (v.a. *enamik*). Selliseid hulgasõnu peetakse atributiivseteks hulgasõnadeks või määratlejateks (Pajusalu 2017: 382) ning nende funktsioon on laiendada nimisõna. Seetõttu atributiivsed hulgasõnad ka selle nimisõna käändest ja arvust. (Pajusalu 2017: 382–383; Metslang 2017: 463) Erandiks on hulgasõna *enamik*, mis kuulub nimisõnalise hulgafraasi alla (Metslang 2017: 471), kuid võib esineda ka nimisõnafraasi osana (Pilvik jt 2025: 238). Nii võiksid skaala *mitte ükski – ... – kõik* hulgasõnad kuuluda mentaalses ruumis mingisse suuremasse konstruktsiooni, nagu [HULGASÕNA + PALL_{VASTAV KÄÄNE}]. Hulgasõnafraas võiks omakorda paikneda suuremas abstraksemas loogilises üksuses vastavalt katses küsitud pallide värvusele (meetodist peatükis 4). Oletatavalt võiksid katses osalejate mentaalses ruumis järgmise malliga laused:

- mitte ükski: [*mitte ükski* + pall] pole (pildil) kollane/sinine
- mõni: [*mõni* + pall] on (pildil) sinine/kollane
- mõned: [*mõned* + pallid] on (pildil) sinised/kollased
- pooled: [*pooled* + pallid] on (pildil) sinised/kollased
- enamik: [*enamik* + palle] on (pildil) sinised/kollased

- kõik: [*kõik* + pallid] on (pildil) kollased/sinised

Vastavalt eesti keele süntaksi omapäradele (Pajusalu 2017: 382–383; Erelt, Metslang 1998: 659–661), ühilduvad laiendist hulgasõnad (*mitte ükski, mõni, mõned, pooled* ja *kõik*) hüpoteetilise nimisõnaga *pall* nii käändes kui ka arvus. Ainsuses esinev hulgasõna *enamik* nõuab seevastu vaid osastavast laiendit. Eeltoodud süntaktiliste eripärade tõttu olid kõik skaala *mitte ükski* – ... – *kõik* hulgasõnad toodud katses sellises grammatilises vormis, nagu need võiksid minu pakutud abstraktses konstruktsioonis esineda.

Skaala vähe, keskmiselt, palju

Skaalal *vähe* – ... – *palju* toodud hulgasõnad *vähe* ja *palju* on osa määrsõnalisest hulgasõnafraasist, kus tuumaks on hulgamäärus ise (Metslang: 2017: 474–476). Kaasnev nimisõna on tavaliselt osastavas käändes, mille hulk ei ole semantiliselt defineeritud (Metslang 2017: 474). Sõltuvalt fookusest, saavad hulgasõnad *vähe* ja *palju* esineda nii nimisõnaga koos kui ka lahus (Metslang 2017: 476). Seetõttu on võimalik eeldada vaid hulgasõnade kokkukuuluvust fraasiga *kollaseid/siniseid palle*. Hulgasõna *keskmiselt* kirjeldab samuti nt mingi entiteetide hulka, kuid EKSS-i (2009) näidislausete alusel esineb see vaid eraldi peale entiteetide ennast. Võib eeldada, et kolme hulgasõna koosesinemisel võivad katseisikutel moodustuda mentaalses ruumis laused malliga *kollaseid/siniseid palle on vähe/keskmiselt/palju*, kuid välistatud pole ka muud lausemallid.

1.4. Konteksti mõju hulgasõnade tajumisele ja valikule

Klassikalistes semantilis-pragmaatilistes käsitlustes (Grice 1989; Levinson 1981; Levinson 2000) kui ka formaalses semantikas (nt Barwise & Cooper 1981; Westerståhl 1985; Peters & Westerståhl 2006) vaadeldakse erineval tasemel hulgasõnade esinemist terviklause või -lausungi sees. Koolkondade erinevused seisnevad omakorda lähenemisperspektiivis: klassikaline semantilis-pragmaatiline lähenemine vaatleb hulgasõnu ehk skalaarseid implikatuure (*scalar implicature*) implikatuuriteooria vaatenurgast (vt Grice 1989; Horn 1984; Levinson 1981; Levinson 2000). Seevastu formaalne semantika hindab terviklause väärust või tõesust, kus hulgasõnad esinevad (Barwise & Cooper 1981; Westerståhl 1985; Peters &

Westerståhl 2006). Mõlemad lähenemised on olnud aktuaalsed hulgasõnade uurimise valdkonnas (vt nt Papafragou & Musolino 2003; Knowlton jt 2022), kuid mina neid oma analüüsis ei kasuta.

Keeleliste väljendite tõlgendamisel on oluline roll ka keelevälisel kontekstil (*extra-linguistic context*) (Requejo 2007; Dash 2008). Oma töös lähtun arusaamast, et entiteetide hulga proportsiooni tõlgendamisel ja edastamisel mängivad rolli erinevad konteksti keelevälised tegurid, mis mõjutavad ka hulgasõnade valikut. Keelevälise konteksti mõju hulgasõnade tõlgendamisele on tuvastatud ka mitme uurimusega (nt Newstead & Collis 1987; Moxey & Sanford 1993; Coventry jt 2010; Yildirim jt 2016; Stateva jt 2019; Moll, Lücke & Bromme 2021).

Varasemad uurimused on näidanud, et hulgasõna kasutusala võib varieeruda vastavalt entiteedi enda ootuspärasele esinemissagedusele (*baseline frequency*) igapäevaelus (Pepper & Prytulak 1974, Newstead & Collis 1987, Moxey & Sanford 1993) ehk kui tavapäraselt entiteet inimese elus esineb. Näiteks ühes katses palusid Stephen E. Newstead & Janet M. Collis (1987) briti päritoluga katseisikutel anda arvuline hinnang tegevuse sagedussõna sisaldavale propositsioonile, milleks olid kas Ameerika külastamine või kinos käimine. Autorid arvasid, et Ameerika külastamise kirjeldamiseks võib hulgasõnade kasutusulatus olla väiksem, kuna tegevus ise leiab üldjuhul aset harvem. Katsetulemused näitasid, et sagedussõna *always* ('alati') vastas Ameerikas käimise korral vahemikule 9–43 korda aastas, kuid kinos käimise puhul vastas sama sõna vahemikule 20–80 korda aastas, mis kinnitas ka Newsteadi & Collisi (1987) oletust.

Hulgasõnade valikut võivad suurel määral mõjutada ka entiteedi enda omadused. Näiteks entiteetide hulga suurus (*set size*), olgu see füüsiline (nt raamat, kohviod) või sageduslik (nt esinemised, seminarid) (Newstead, Pollard & Riezebos 1987, Newstead & Coventry 2000). Stephen E. Newsteadi & Kenny R. Coventry (2000) katses, kus katseisikud pidid hindama pallide kogust vastavalt pallide ja kausi suurusele ja asetusele, eelistati suuremate pallide korral ka suurema suurusjärgu (*magnitude*) sõnu, nagu *many* ('palju') ja *lots* ('palju'). Autorid rõhutavad, et tegemist on suhtelise suuruse (*relative size*) hindamisega: palle hinnati vastavalt kausi suurusele, milles need parajasti paiknesid. Samuti leidsid autorid, et oluline on ka entiteetide või nendega seotud objektide funktsionaalsus. Nii tõlgendasid nende katses osalejad kaussi anumana, mis hoidis või hoiustas palle. Kokkuvõttes, võib hulgasõnade tõlgendamist ja valikut mõjutada ka mittekeeleline kontekst. Kuigi siinses töös omaette konteksti mõju eesti

hulgasõna kasutusulatustele ja valikule ei uurita, olen eeltoodud teguritega arvestanud magistritöö katse koostamisel (meetodist peatükis 4).

1.5. Vastajaspetsiifiliste tegurite mõju hulgasõnade valikule

Järgnevalt tutvustan erinevaid vastajatest tulenevaid tegureid, mis võiksid mõjutada hulgasõnade valikut, nagu sotsiaaldemograafilised tegurid ning katseisiku väsimustase.

1.5.1. Erinevate sotsiaaldemograafilised tegurite võimalik mõju

Peale keelevälise konteksti saab rääkida ka keelevälistest teguritest, mis samuti võivad mõjutada keeleüksuste tajumist (*extra-linguistic factors*) (vt Ginzburg 1974), nt inimese enda teadmised maailma kohta (Pinker 1995), tema uskumused (Yildirim jt 2016; Dubravskaja & Soroka 2024: 120) või erinevad sotsiaaldemograafilised tegurid (Ulanska jt 2020; Garimella, Bane & Mihalcea 2023). Järgnevalt tutvustan lähemalt erinevaid sotsiaaldemograafilisi tegureid, mis võiksid avaldada mõju hulgasõnade valikule entiteetide hulga kirjeldamisel.

Elukoht

Keel on tugevalt sotsiaalne fenomen, mis võib erineda inimrühmade vahel vastavalt nende arealsele paiknemisele. Geograafilise paiknemise vaatenurgast võib see olla erinev nii riigi, kui ka maakondade või linnade tasandil (Johnstone 2004: 65). Erinevusi võib täheldada erinevatel tasanditel, nagu nt sõnavaras, morfoloogias või häälikute hääldamisel. (Reed 2019: 1–2; Wardhaugh & Fuller 2021: 2, 6–7) Ka eesti keele sõnavara on suure varieeruvusega sõltuvalt murdealast (Oja 2014; Vaba 2014). Sellest lähtuvalt võin oletada, et eesti keele siseselt võivad hulgasõnade tähendused ja tõlgendamine erineda vastavalt inimese elukohale ning mõjutada katseisikute vastuseid objektide proportsioonide kirjeldamisel. Katse raames kogun andmeid maakonna täpsusega, mis peaks olema minu arvates piisav keeleliste erinevuste ja nende mõju tuvastamiseks hulgasõnade valikul objekti proportsiooni kirjeldamisel.

(Omandatud) haridustase

Haridustaseme ja keelevõimete korrelatsiooni on üldiselt tuvastatud mitmes uurimuses (Mackenzie 2000; Dąbrowska 2012), sh sõnavara uurimisega seotud uurimustes (Van der Elst jt 2005; Mulder & Hulstijn 2011; Treffers-Daller 2013). Kõrgeima haridustasemega katseisikud näitasid parimaid tulemusi uute sõnade õppimisel (Van der Elst jt 2005) ning sõnavara/leksikaalsete teadmise ja mäluga seotud ülesandes (Mulder & Hulstijn 2011). Dąbrowska (2012: 219) on oletanud, et kõrgema haridustasemega vastajate tulemused on kõrgemad tänu mitmekülgsele keelelisele kogemusele, mis võimaldab järeldada suuremate keeleliste malle omandamist. Seega võivad esineda ka erinevused hulgasõnade valikus entiteetide proportsioonide kirjeldamisel erinevate haridustasemetega inimeste seas.

Võõrkeelte valdamine

Mitme muu keele (L2) valdamine võib avaldada mõju inimese emakeelele (L1), selle oskustele ja tunnetusele ning ka erinevatele kognitiivsetele protsessidele (Bialystok jt 2009: 89, 92). Välja on pakutud erinevaid mudeleid sellest, kuidas võivad osatud keeled inimese ajus koos esineda (de Bot 2004: 17–18; Harley 2014: 155; Traxley 2023: 422–425), kuid kindel vastus keelte kooseksisteerimisest inimese mentaalses ruumis puudub. On tõenäoline, et omandatud keeled paiknevad ühes suuremas mentaalses süsteemis (Kroll jt 2015: 1). Seega muude keelte valdamine (ka ühe konkreetse keele oskus) võiks avaldada mõju ka inimese hulgasõnade tõlgendamisele ja seega nende valikule emakeeles hulga kirjeldamisel. Täheldada tuleb, et igas keeles on hulgasõnadel oma spetsiifilised tähendusvarjud, mis ei pruugi teiste keelte omadega sarnaneda või kokku langeda, nagu on seda näidanud Stateva jt (2019) uurimus.

Sugu

Naiste ja meeste erinevusi kognitiivsete võimete juures on täheldatud mitme uurimusega, nagu nt naiste suuremat edukust keeleliste ja mäluga seotud võimete (Hirnshein, Coloma Andrews & Hausmann 2014) ning meeste visuaal-ruumiliste võimete testimisel (Weiss jt 2003; Downing jt 2008). Erinevusi on tuvastatud ka semantilise töötluse tasandil, mis näitavad naiste edukust. Nii võtab naistel semantiline töötlemine vähem kognitiivseid ressursse (Wegesin 1998; Jaušovec & Jaušovec 2009) ning töötlemine ise on meeste omast kiirem (Wirth jt 2007; Daltrozzo, Wioland & Kotchoubey 2007). Sellest võin oletada, et hulgasõnade valikutes

objekti hulga kirjeldamisel võiksid esineda erinevused naiste ja meeste vahel. Siinse uurimuse puhul on tegemist laiemal oletusega, kuna sugu vaadatakse ilma vastamisaja tulemusteta, nagu on seda eeltoodud uurimustes ühel või teisel määral tehtud. Minu eesmärk on leida üldisemat erinemist, mingi soo paremuse tuvastamine hulgasõnade valikul ei ole siinse töö fookuses.

1.5.2. Väsimustase ja selle võimalik mõju

Inimese väsimustase (*tiredness* või *mental/cognitive fatigue*²) katse sooritamisel võib osutuda üheks suurimaks mõjuteguriks sobiva hulgasõna valimise etteantud proportsiooni hindamisel. Siiani on mitmed uurimused näidanud korrelatsiooni inimese kognitiivsete võimete vähenemise ning väsimustaseme vahel (Healy jt 2004; Kurzban jt 2013; Cutsem jt 2022). Näiteks suurema väsimuse korral on tuvastatud eksimuste suurenemist (Boksem jt 2006; Lorient, Boksem & Ridderinkhof 2013), raskusi tähelepanu saavutamises (Van der Linden & Elling 2006) ja säilitamisel (Boksem jt 2005). Seega võib suurem väsimusaste mõjutada hulgasõnade tajumist negatiivselt ning näiteks soodustada vigade tekkimist. Katsega ei ole aga võimalik vaadelda tähelepanuga seotud faktoreid hulgasõnade valimisel katse disaini tõttu (meetodit peatükis 4), kuid selle esinemine ja mõju ei ole üldiselt katse ajal välistatud.

² Ingliskeelses teadusruumis tehakse vahet terminite *tiredness* ja *fatigue* vahel. *Tiredness* on lühiajaline ning iseenesest kaduv olek, mis on põhjustatud mingi kindla tegevusest (Moncrieff & Fletcher 2007: 334). *Fatigue* on üldiselt n-ö püsiv kurnatus, mis ei lähe iseseisvalt ära ning mõjutab oluliselt igapäeva või kognitiivset tegevust (Lorient, Boksem & Ridderinkhof 2005: 199; Fritsch & Quinn 2010: 33; Boksem & Tops 2008: 126). Siinses katses nende mõistete vahel vahet ei tehtud.

2. Uuritavate umbmääraste hulgasõnade tähendustest

Järgnevalt esitan ja arutan üheksa eestikeelse hulgasõna tähenduse ja kasutusulatuse üle. Esimesena tutvustan peatükis 3.1 hulgasõnu *mitte ükski – mõni – mõned – pooled – enamik – kõik*, mis moodustab omaette nn formalloogilise skaala. Peatükis 3.2 kirjeldan lähemalt hinnanguskaala *vähe – keskmiselt – palju* hulgasõnade semantikat. Teen samuti oletusi selle kohta, kuidas võiksid hulgasõnad jaotuda skaalal vastajate seas.

Peatükis käsitlen vaid hulgasõnade uurimuse seisukohalt relevantseid tähendusi. Juhul, kui sõna on polüseemne, käsitletakse ainult ühte uurimuse vaatekohast relevantset vastet. Kõik peatükis toodud näitelaused on minu enda välja mõeldud.

2.1. Skaala *mitte ükski, mõni, mõned, pooled, enamik ja kõik* sõnade semantikast

Üheks uuritavaks skaalaks on kohandatud formaalloogiline skaala *mitte ükski, mõni ja mõned, pooled, enamik ja kõik* (edasiselt: *mitte ükski – ... – kõik*), mis tuleneb vastandite ruudust või aristoteelse ruudust (eesti keeles ka loogiline ruut).

Vastandite ruut põhineb Aristoteelse (4 saj. eKr) süллоogilisel õpetusel positiivsest ja negatiivsest vanstandumisest või suhetest (Smessaert & Demey 2023: 1–2), mida on omakorda arendanud ja kujundanud Boethius 6. sajandi alguses (Westerståhl 2012: 3). Inglise keeles koosneb vastandite ruut kahest polaarsest paarist: universaalne positiivne *all* ja negatiivne *no/none* poolus ning suhteline positiivne *some* ja negatiivne *not all* poolus. (Parsons 2021) Universaalsete hulgasõnade paariks võiksid eestikeelses tõlkes olla *kõik* (positiivne) ja *(mitte) ükski* (negatiivne), suhteliseks hulgasõnade paariks *mõni* (positiivne) ja *mitte kõik* (negatiivne). Lisaks on hulgasõnad seotud omavahel erinevate loogiliste suhtega (Westerståhl 2012: 3–4; Smessaert & Demey 2023: 1–2; Parsons 2021), mida siin töös aga ei käsitleta. Oma uurimuses asendasin omakorda *mitte kõik* propositsiooni hulgasõnaga *enamik*. Vastavalt piloottestide tulemustele oli samuti lisatud hulgasõna *pooled* (vt peatüki 4.1.3). Edasiselt vaatlen täpsemalt iga skaala *mitte ükski – ... – kõik* hulgasõna tähendust eraldi koos võimalike proportsioonide tähistamisega.

Järgnevalt annan ülevaate hulgasõnade skaala *mitte ükski* - ... - *kõik* tähendustest ja nende tähenduse nüanssidest.

2.1.1. *Ükski ja mitte ükski*

Sõna *ükski* on sõnaliigi poolest asesõna, mida kasutatakse peamiselt koos eitusega (Erelt 2017: 185). Umbmäärase asesõnaga saab eitada nt mingi entiteedi liiki või kuuluvusrühma. Viimase alla võivad kuuluda nii füüsilised (nt inimene, maja) kui ka abstraktsed (nt sõna, reegel) entiteetid. Süntaktiliselt on *ükski* sel juhul täiendi rollis. (EKSS 2009) Näide sellisest kasutusest on esitatud lauses 1.

(1) *Mari ei allu ühelegi reeglile.*

Lauses 1 kasutab kõneleja *ükski*-hulgasõna ainult kõikide maailmas võimalike reeglite eitamiseks. Seevastu ei tähenda see, et Mari ei allu omakorda mingile muule asjaolule, nagu näiteks kohustused.

Tuleks samuti täheldada, et eesti keeles võib hulgasõna *ükski* esineda samuti koos eituspartikliga *mitte*, mis toonitab/intensiivistab hulgasõna negatiivset tähendust. *Mitte* kasutus on kohustuslik juhul, kui *ükski* esineb üksi ilma verbita. (Erelt 2017: 185–186) *Mitte ükski* võiks samuti olla üks viisidest väljendada entiteedi puudumist, nagu on seda oletatud ingliskeelse vaste *none* puhul (Tsouhlaris 2024). Siinses töös on esitatud *ükski* hulgasõna koos eituspartikliga, mida käsitletakse ühe tervikuna. Lähtuvalt EKSS-i (2009) ning Mati Ereli (2017: 185–186) iseloomustusest oletan, et skaalal *mitte ükski* – ... – *kõik* vastab *mitte ükski* kasutusulatus ainult 0% proportsioonile.

2.1.2. *Mõni ja mõned*

Sõna *mõni* on määramatu hulgaline asesõna (Pajusalu 2017: 383), mida kasutatakse kas vähesele loendatavale hulgale või umbmääraste entiteetide hulga osutamiseks (Pajusalu jt 2004: 84). Esimese toodud tähenduse korral saab hulgasõnaga viidata nii spontaanselt kokkuloetavale kui ka arvuliselt (äärmiselt) ebamäärasele hulgale. Näited mõlemast kasutusest

on esitatud lausetes 2a ja 2b. Mõlemal korral täiendab hulgasõna teatud nimisõna, olles osa nimisõnafraasist (NP). (EKSS 2009)

(2a) *Mõni päev on töö esitamiseni jäänud.*

(2b) *Mõned saapad on kapis.*

Mõlemas näitlauses on hulgasõna *mõni* osa NP-st (*mõni päev* ja *mõni kogemus*). Lauses 2a esineva *mõni* all mõeldakse mingit väiksemat kogust päevi (nt kaks-kolm), mida on võimalik kalendri abil kokku lugeda. Lause 2b korral viitab kõneleja mingile *saabaste* kogusele, kuid kuulajal ei ole võimalik nende täpset kogust aimata ning nende täpne arv jääb seega teadmata.

Võrreldes nt ingliskeelse vastega *some* (Stateva jt 2019) või itaaliakeelse *alcuni* (Pezzelle, Bernardi & Piazza 2018), on eestikeelsel hulgasõnal *mõni* nii ainsuse kui ka mitmuse vorm. Vormide tähenduslike erinevusi on eesti keeleteadusest käsitletud vaid EKK-is (2020), kus autorite meelest ei esine vormide vahel tähenduses erinevusi. Keele ikoonilisuse kvantiteedi põhimõtte (*iconicity of quantity*) järgi tingib keerulisem/laiem tähendus ka keerukamat vormi (Hapselmath 2008: 2, 4). Sellest suurem kogus või entiteetide hulk peaks seetõttu väljenduma mitmusena (Jakobson 1971: 352; viidatud Hapselmath 2008: 4). Seega markeeritud mitmuse vorm *mõned* peaks viitama hulga poolest suuremale kogumusele, olles *mõni*-vormist tähenduse vaatenurgast hierarhiliselt kõrgemal positsioonil. Vorme *mõni* ja *mõned* käsitletakse edasiselt kahe eraldi hulgasõnana.

Skaalal *mitte ükski* – ... – *kõik* võiksid hulgasõnade *mõni* ja *mõned* kasutusulatus vahemikus 10% – 40%, kuna 0% objekti proportsiooni leksikaalseks vasteks on oodatud hulgasõna *mitte ükski* (vt eelmist peatükki) ning 50% objekti proportsiooni vasteks *pooled* (vt järgmist peatükki). Samuti võiks mitmuse vorm olla eelistatum suurema ehk abstraktsema koguse kirjeldamisel vastavalt ikoonilisuse kvantiteedi printsiibile.

2.1.3. *Pool*

Sõna *pool* kuulub arvsõnade (täpsemalt põhiarvsõnade) klassi, millega edastatakse entiteetide ligikaudset arvulist hulka (Erelt 2017: 59). Arvsõna *pool* saab kasutada nii mingi koguse kahendiku kui ka nt mediaani või keskpunkti markeerimiseks (EKSS 2009). Näited mõlemast kasutusest on esitatud lausetes 3a ja 3b.

(3a) *Pool tuhat.*

(3b) *Tahan ürituselt poole pealt lahkuda.*

Lauses 3a tähistab hulgasõna *pool* täpse arvulist kahendikku ehk 50% tuhandest. Lauses 3b viitab kõneleja hulgasõna *pool* abil ürituse ajalisele keskpunktile. Sellise esinemise korral, nagu lauses 3b, peab hulgasõna olema osa mingist suuremast määruslikust konstruktsioonist.

Veel üks hulgasõna *pool* kasutus sarnaneb lauses 3a esitatud kasutusele, kuid seda kasutatakse mingi terviku koguse kahe võrdse osa kirjeldamisel (EKSS 2009). Näide sellisest kasutusest on toodud lauses 4.

(4) *Pooled kohad on tühjad.*

Lauses 4 kasutatud hulgasõnaga tähistab kõneleja, et kahendik ruumis olevatest kohtadest on täitmata või vabad. Sellise kasutuse antonüümiks on *kõik*, kuid mõnede kasutuste korral võivad olla nt *terve* või *kogu* (EKSS 2009).

Lähtudes selle ülaltoodud kirjeldusest eeldan, et skaalal *mitte ükski* – ... – *kõik* valitakse hulgasõna *pooled* ainult objekti 50% proportsiooni kirjeldamisel.

2.1.4. Enamik

Sõna *enamik* on nimisõna, mida kasutatakse *osa-terviku* suhte kirjeldamisel (Metslang 2017: 470). EKSS (2009) annab hulgasõna iseloomustuseks “*suuremat osa*”. Inglisekeelse vaste *most* kohta on järeldatud, et see hõlmab vaikimisi enda all vähemalt poolt ehk 50% entiteetide hulgast (Ariel 2004; Hackl 2009³), mida võiks oletada ka uuritava eestikeelse *enamik* puhul. Näide hulgasõna kasutusest on esitatud lauses 5.

(5) *Enamik minu kursusekaaslastest kaitseb oma magistritööd sel suvel.*

Lauses 5 tähistab *enamik*, et üle poole, kuid mitte kõik kõneleja kursusekaaslastest kaitsevad suvel magistritööd. Kursusekaaslaste täpne hulk jääb kuulaja jaoks teadmata, kuid kõneleja meelest võiks tegemist olla suurema osaga.

Lähtudes hulgasõna eelkirjeldatud tähendusest oletan, et skaalal *mitte ükski* – ... – *kõik* võiks hulgasõna *enamik* kasutusulatus entiteetide proportsionaalse hulga kirjeldamisel olla

³ Hackl (2009) uurimus põhineb formaalsemantikal.

vahemikus 60% kuni 90% ning piirneb üleval hulgasõnaga *kõik*. Siinsest tööst jääb välja hulgasõna mitmuse vorm, kuna hulgasõna *enamik* tähendus tähistab vaikumisi suuremat hulka (EKSS 2009) ning selle esinemissagedus Eesti keele Ühendkorpuse (Koppel jt 2023) järgi on harv⁴. Seetõttu katset koostades oletasin, et *enamikud* lisamine ei oleks toonud ka erilisi tulemusi.

Tuleb täheldada, et hulgasõnal *enamik* esineb ka sünonüümne sõna *enamus* (Pilvik jt 2025). Kuna *enamus* võib tähistada ka entiteetide hulgalist ülekaalusust (EKSS 2009), jäädi siinse uurimise koostamisel hulgasõna *enamik* juurde, millel on EKSS-i (2009) vaid üks kindel tähendus.

2.1.5. *Kõik*

Sõna *kõik* on hulgaline asesõna (Pajusalu 2017: 383), mida on peetud universaalseks hulgasõnaks (Erelt 1997: 11; Pajusalu 2008; Erelt 2017: 204). Hulgasõnaga saab viidata nii juba teada olevale/eelnevalt mainitud entiteetide hulga kui ka mingile kogusele (Pajusalu jt 2004: 62, 71–72; Pajusalu 2008; EKSS 2009). Järgnevalt keskendun vaid selle omadusele viidata hulga.

EKSS (2009) iseloomustab hulgasõna *kõik* esimest tähendust kui “*igaiüks, viimane kui üks*”. Niisiis viidetakse sellega entiteedi tervikhulgale, mis omakorda koosneb igast entiteedi eraldi juhtumist (vrd EKSS 2009). Seetõttu tähistab ka hulgasõna *kõik* vaikumisi mitmust (Erelt 2017: 204; EKSS 2009). Näited hulgasõna erinevast kasutusest eelkirjeldatud tähenduses on toodud lausetes 6a ja 6b. Lauses 6a käitub hulgasõna nimisõnataoliselt, lauses 6b omadussõnataoliselt.

(6a) *Kõik puudusid eile loengust.*

(6b) *Kõik keeleteaduse üliõpilased lõpetavad see aasta ülikooli.*

Mõlema lause puhul mõtleb kõneleja *kõik*-hulgasõna all entiteetide hulga igat eraldi esinevat juhtumit, kuid hulgasõna roll nendes on erinev: nimisõnaline *kõik* osutab iga loengu osalejale ning omadussõnaline *kõik* edastab informatsiooni keeleteaduse üliõpilaste arvu kohta. Mõlemal juhul markerib *kõik* samuti mitmust, mis ühildub arvuliselt tegusõna või lisaks ka esineva

⁴ Sõnana esineb hulgasõna *enamik* 0.0077% ning lemmana 0.013% kogu korpusest ning *enamikud* sõnana 0.000017% ning lemmana null korda kogu korpusest. (Koppel jt 2023)

fraasi nimisõnaga. Lähtudes eeltoodud hulgasõna kirjeldusest, oletan, et skaalal *mitte ükski* – ... – *kõik* saab hulgasõna *kõik* valida ainult objekti 100% proportsiooni kirjeldamiseks.

2.2. Skaala *vähe*, *keskmiselt* ja *palju* sõnade semantikast

Järgnevalt tutvustatakse hulgasõna skaala *vähe* – *keskmiselt* – *palju* (edasi: *vähe* – ... – *palju*) sõnade tähendusi ja nende tähendusnüansse.

2.2.1. *Vähe*

Sõna *vähe* on süntaktiliselt hulgamäärsõna (Metslang 2017: 462–463). Hulgasõna abil on võimalik kirjeldada mingit väikest kogust, kuid selle täpne arvuline osakaal jääb kuulaja jaoks teadmata (EKSS 2009). Eeltoodud hulgasõna kasutus on esitatud näites 7.

(7) *Ma olen täna vähe kohvi joonud.*

Nii viitab kõneleja näites 7 hulgasõnaga *vähe*, et tarbitud kohvi kogus päeva jooksul oli kõneleja võimaliku koguse suhtes väike. Täpne kohvi kogus (siinkohal kohvitasside arv) jääb kuulajale seevastu teadmata. Sellise kasutuse korral on võimalik *vähe* asendada nt sõnadega *veidi* või *natuke* (EKSS 2009). Seetõttu saab *vähe* pidada nn väikese/väiksema suurusjärguga hulgasõnaks (*low magnitude quantifier*), nagu on seda tehtud ingliskeelse hulgasõna *a few/few* (Newstead & Coventry 2000; Coventry jt 2010) ning itaaliakeelse *pochi* kohta (Pezzelle, Bernardi & Piazza 2018). *Vähe*-hulgasõna antonüümiks on hulgasõna *palju* (EKSS 2009).

Sarnaselt hulgasõna ingliskeelsetele vastetele *few* ja *little* (Katriz & Ariel 2023: 2, 25), ei saa ka eestikeelne *vähe*-hulgasõna sisuliselt tähistada/võrduda nulliga või puuduva hulgaga. Lauses 7 ei saa kuulaja järeldada, et kõneleja pole tilkagi kohvi joonud. Vastasel juhul oleks esinenud loogikaviga: joomine tegevusena implitseerib koguse kohvi (ehk kohvitasside) alla neelamist ning tegevuse mittesooritamisel ei saa kõneleja väita, et ta on seda tegevust sooritanud. Nicole Katriz & Mira Ariel (2023: 2, 25) eeskujul võib väita, et hulgasõna *vähe* kasutus eeldab vähemalt ühe entiteedi eksemplari olemasolu.

Lähtudes ülaltoodud EKSS-i (2009) hulgasõna iseloomustuse kirjeldusest ning (Katriz & Ariel 2023: 2, 25) tähelepanekust oletan, et eestikeelne hulgasõna *vähe* kasutusulatus entiteedi hulga proportsiooni kirjeldamisel võiks võrduda vahemikuga 10% kuni 30%.

2.2.2. Keskmiselt

Sõna *keskmiselt* on määrsõna, mille abil saab edastada kas entiteedi hulka täpset arvulist mediaani või umbmäärast keskpunkti. Viimase puhul on tegemist hinnangulise ja subjektiivse keskmisega, millega entiteedi koguse arvu ei tähistata. Sellisel juhul on hulgasõna sisuliselt asendatav sõnaga *rahuldavalt*. (EKSS 2009) Näide hulgasõna *keskmise* kui umbmäärasest keskpunktist on esitatud lauses 8.

(8) *Kuulajaid oli saalist keskmiselt.*

Lauses 8 kasutab kõneleja hulgasõna *keskmiselt* kuulajate mõõduka hulga väljendamiseks.

Kuna hulgasõna edastab umbmäärast keskpunkti tähenduses *rahuldavalt* (EKSS 2009), ei viita selline kirjeldus otseselt range kahendiku väljendamisele, vaid pigem laiema hulgale. Seega oletan, et *keskmiselt* võiks entiteedi hulga iseloomustamisel jääda vahemikku 40% kuni 60% hulga proportsioonide kirjeldamisel.

2.2.3. Palju

Sõna *palju* on hulgamäärsõna (Metslang 2017: 462–463). Hulgasõna *palju* saab kasutada entiteedi suurema umbmäärase koguse edastamiseks (EKSS 2009; Pajusalu jt 2004: 87). Renate Pajusalu jt pe(2004: 87) järgi saab hulgasõnaga edastada kõneleja subjektiivset hinnangut kõnealustele entiteetidele⁵. Selline kasutus on esitatud näites 9.

(9) *Mul on palju raha.*

Lauses 9 viitab kõneleja *palju*-hulgasõna abil, et temal hetkel olev raha kogus on võimaliku koguse suhtes suur. Kuulajale omakorda jääb kogu raha koguse summa teadmata. Loogiliselt aga peaks hulk ületama vähemalt 50% kõneleja jaoks võimalikku rahalist hulka. Seetõttu võib

⁵ Sellest oletan, et ka hulgasõnad *vähe* ja *keskmiselt* võiksid olla hinnangulised.

palju pidada nn suure/suurema suurusjärguga hulgasõnaks (*high magnitude quantifier*), nagu on seda tehtud ingliskeelse hulgasõnade *many* ja *lots* (Newstead & Coventry 2000; Coventry jt 2010) ning itaaliakeelse *molti* kohta (Pezzelle, Bernardi & Piazza 2018).

Sarnaselt oma antonüümile *vähe* (EKSS 2009; *vähe*-hulgasõnast vt täpsemalt 3.2.1.), ei saa hulgasõna *palju* tähistada hulga absoluutset maksimumi ehk 100%. Vastaval juhul tekib nt lause 9 puhul loogika viga: kõneleja ei oma ega saa hüpoteetiliselt omada kogu maailmas olevat raha, mistõttu ei saa ka kuulaja hulgasõna võrdsustada 100% esinemisega.

Lähtudes eeltoodud ülevaatest hulgasõna tähendusest oletan, et *palju* kasutusulatus võiks jääda 70% kuni 90% piiridesse, kuna tegemist on suurema hulgaga, kuid 60% puhul kehtib hulgasõna *keskmiselt* kasutusulatus.

3. Teiste uurimuste tulemused hulgasõnade ja proportsioonide vastavuse uurimisel

Siinses uurimisvaldkonnas on tehtud mitu erinevat uurimust, millega on proovitud kirjeldada erinevate indoeuroopa keelte hulgasõnade tõlgendamist ja valikut, nagu nt itaalia (Pezzelle, Bernardi & Piazza 2018), inglise (Moxey 1986; Stateva jt 2019) ning saksa, prantsuse ja sloveenia (Stateva jt 2019). Sellegi poolest on nende metodoloogia olnud siinsest uurimusest erinev: näiteks Moxey (1986: 26–102)⁶ katse on kombinatsioon produktsiooni- ja tajumiskatsest; Sandro Pezzelle, Raffaella Bernardi & Manuella Piazza (2018) uurisid korraga mitme skaala sõnu, kuid sõnad esinesid samuti üksinda; Penka Stateva jt (2019) tulemused põhinevad hulgasõnu sisaldavate lausete sobivuse hindamisel viiese Likerti skaala alusel vastavalt kahe lausega kirjeldatud olukorrale. Stateva jt (2019) kaks konteksti loovat testlauset sisaldasid ka numbrilist väärtust. Järgnevalt teen kokkuvõtliku ülevaate iga võimaliku uuritava eestikeelse hulgasõna vaste kohta, tuginedes peamiselt eelkirjeldatud uurimustele.

3.1. Skaala *mitte ükski* – ... – *kõik* hulgasõnade muukeelsed vasted

Järgnevalt tutvustan sõna kaupa skaala *mitte ükski* – ... – *kõik* hulgasõnade võimalike vastete kohta käivaid tulemusi.

Mitte ükski

Hulgasõna *mitte ükski* pole teistelt uurijatelt palju tähelepanu saanud. Näiteks Stateva jt (2019: 5) jätsid oma katsest meelega välja ingliskeelse sõna *no* ja selle saksa, prantsuse ja sloveenia vasteid. Autorite meelest oli tegemist väga kitsatähendusega sõnaga, mis saab vastata vaid objekti 0% proportsioonile. Ka Pezzelle, Bernardi & Piazza (2018: 121) tulemuste järgi tähistas itaaliakeelne *nessuno* üldtulemuste järgi objekti 0% proportsiooni.

⁶ Kasutan vaid ühe katsele tulemusi kaheksas (vt täpsemalt Moxey 1986).

Mõni ja mõned

Moxey (1986: 52) tulemuste järgi osutus ingliskeelse *some* vaste kasutusulatus olema vahemikus 25% kuni 75%. Itaaliakeelne *alcuni* oli seevastu kasutusulatusena 10% kuni 75% Pezzelle, Bernardi & Piazza (2018: 121). Stateva jt (2019: 8) tulemused aga näitasid, et ingliskeelset *some* hinnati sobivaks väga laias proportsionaalses vahemikus ehk see oli suure aktsepteeritavusulatusena (umbes 3% kuni 79%), kui selle saksakeelne *einige*, prantsuskeelne *quelque* ja sloveeniakeelne *nekaj*, mille ulatus ei läinud üle 50% piiri ning nende aktsepteeritavustase oli kõrgem skaala alumises otsas (umbes kuni 20%).

Pooled

Sarnaselt hulgasõnale *ükski*, ei ole ka hulgasõna *pool* uurijatelt erilist tähelepanu saanud. Stateva jt (2019: 9) tulemused näitasid, et prantsuse *la moitié*, saksa *halbe*, inglise *half* ja sloveenia *polovica* jäid väikese varieeruvusega üldiselt ligikaudsesse vahemikku 40% kuni 60%. Hulgasõnad olid kõrgeima aktsepteeritavuse tasemega umbes 50% lähenduses (Stateva 2019: 9).

Enamik

Ingliskeelse *most(ly)* puhul leiti, et selle kasutusulatus jäi vahemikku 25% kuni 90% (Moxey 1986: 51). Itaaliakeelne *la maggior parte* tähistas üldtulemuste järgi vahemikku 35% kuni 90% (Pezzelle, Bernardi & Piazza 2018). Stateva jt (2019: 10) tulemuste järgi algas saksakeelse *die meisten*, prantsuskeelse *la plupart*, sloveeniakeelse *vecina* ning ingliskeelse *most* kasutusulatus ligikaudselt 51% läheduses ning aktsepteeritavustase oli kõrge umbes alates 60%. Erandiks osutus vaid saksakeelne vaste *die meisten*, mis oli kõrgema aktsepteeritavusega juba 53% juures (Stateva jt 2019: 10). Nende vastete aktsepteeritavustase langes saksa ja sloveenia keeles 97% juures ning inglise ja prantsuse keeles 99% lähedal (Stateva jt 2019: 10).

Kõik

Sarnaselt hulgasõnadele *ükski* ja *pool*, ei ole hulgasõna *kõik* kasutusulatust objekti kirjeldamisel teistes keeltes eriti uuritud, kuna see saab oletatavalt tähistada vaid 100% (Moxey 1986: 60; Stateva jt 2019: 5). Pezzelle, Bernardi & Piazza (2018: 121) üldtulemused näitasid,

et itaaliakeelne *tutti* katab samuti ainult 100% proportsiooni, kuid objekti kokkuloetavuse võimaluse korral kasutati hulgasõna teatud määral ka 80% ja 90% kirjeldamiseks.

3.1. Skaala vähe – ... – palju hulgasõnade muukeelsed vasted

Järgnevalt tutvustan sõna kaupa skaala vähe – ... – *palju* hulgasõnade võimalike vastete kohta käivaid tulemusi.

Vähe

Moxey (1986: 51) leidis, et ingliskeelsed *a few/few* kasutusulatus vastab vahemikule 10% kuni 25% objekti proportsiooni kirjeldamisel. Itaaliakeelne *pochi* jäi Pezzelle, Bernardi & Piazza (2018: 121) üldtulemuste järgi vahemikku umbes 10% kuni 50%, kuid subitiseerimisefektita⁷ ligikaudselt 25% kuni 75%. Stateva jt (2019: 7) tulemuste järgi tajuti hulgasõna *vähe* vasteid erinevalt: ingliskeelne *few* kasutusulatus lõppes umbes 41% juures, prantsuskeelse *un peu* umbes 35% juures ning saksakeelse *wenig* 19% juures.

Keskmiselt

Hulgasõna *keskmiselt* on objekti kirjeldamise uurimustest välja jäetud (vt nt Newstead & Coventry 2000; Rett 2017; Pezzelle, Bernardi & Piazza 2018), mistõttu ei ole teada ei tema indoeuroopa vastete võimaliku proportsioonide vahemiku ega aksepteeritavustaseme kohta midagi teada.

Palju

Moxey (1986: 51) tulemuste järgi kattus ingliskeelse *many* kasutusulatus ilma täiendi(te)ta vahemikus 40% kuni 90%. Itaaliakeelne *pochi* jäi üldtulemuste järgi vahemikku umbes 35% kuni 95% (Pezzelle, Bernardi & Piazza 2018: 121).

⁷ Subitiseerimise all mõeldakse inimese võimet aru saada entiteetide arvust seda kokku lugemata. Efekt esineb kuni nelja entiteedi olemasolul. (Balakrishnan & Ashby 1992: 80)

4. Meetod

Järnevalt tutvustan lähemalt katses kasutatud meetodit ning sellega kaasuvaid detaile. Samuti esitan ülevaate andmetest ja analüüsimeetoditest. Peatükis 4.1. annan ülevaate katse koostamisprotsessist, hulgasõnade oodatud kasutusulatuses, hulgasõnade paiknemisest katse ajal ning pilootkatse tulemustest. Peatükis 4.2. tutvustan katse ülesehitust ja protseduuri. Peatükis 4.3 kajastan osalejate valimiga seotud külgi. Peatükis 4.4 tutvustan analüüsimeetodeid.

4.1. Katse koostamine ja disain

Järgnevalt tutvustan erinevaid katse koostamisega seotud etappe. Kirjeldan täpsemalt stiimulite (ka harjutusstiimulite) koostamise protsessi ning tutvustan ka uuritavate hulgasõnade oodatud jaotusi kahel skaalal. Samuti kirjeldan ja põhjendan hulgasõnade paiknemist ja järjekorda katsevoorude ajal. Lisaks toon pilootkatsega saadud olulisi muudatusi.

4.1.1. Katse stiimulite koostamine ning hulgasõnade jaotused

Inspiratsiooni katseks saadi osaliselt Heim jt (2012) uurimusest, kus katses osalejatel paluti hinnata lause tõesust vastavalt kuulnud propositsioonile ja ekraanil kuvatud kollaste ja siniste pallikeste hulga. Eestikeelsete hulgasõnade valikute uurimiseks hulga edastamisel, loodi Heim jt (2012) uurimuses kasutatud stiimulite põhjal, kuid on läbinud väikseid muudatusi. Piltstiimulite joonistamisel kasutati Tartu Ülikooli psühholoogia instituudi eksperimentaaluuringute spetsialisti Kristiina Averini tehnilist abi. Sisuliselt kohandas piltstiimuleid magistritöö autor ise. Piltstiimulite loomiseks kasutati tarkvara MATLAB.

Vastavalt Heim jt (2012) koostati unikaalseid piltstiimuleid. Igal pildil oli kujutatud hallil taustal 50 juhusliku paiknevuse ja suurusega palli. Samuti valiti sarnaselt Heimi jt (2012) uurimusele pallide värviks kollane ja sinine. Pallide hulk oli samamoodi jagatud kaheks:

1. Testobjektid. Testobjektide all mõeldakse samavärvilisi palle (nt kollaseid palle), mis on seatud uurimisfookuse alla. Testobjektide hulga minimaalne väärtus on

0 (0%) ning maksimaalne 50 (100%). Testpallide hulga vahepealsed väärtused ehk proportsioonide olid 5 (10%), 10 (20%), 15 (30%), 20 (40%), 25 (50%), 30 (60%), 35 (70%), 40 (80%) ja 45 (90%). Testpallid paiknevad pilt-stiimulil alati ühe klastrina koos;

2. Taustobjektid. Taustobjektid on muu värvusega pallid, mille suhtes hindavad katses osalejad testobjektide hulka. Taustobjektide kogus varieerus vastavalt testpallide hulgale, vastates valemile $testobjektid + taustobjektid = 50$.

Täpsem testobjektide ja taustobjektide jaotus stiimulitel on esitatud tabelis 1. Näidis1 stiimulist on esitatud Lisas 1.

Tabel 1. Testobjektide ja taustobjektide jaotus pildil.

Testobjektide proportsioon pildil taustobjektide suhtes (100%-st)	Testobjektide arv	Taustobjektide arv
0%	0	50
10%	5	45
20%	10	40
30%	15	35
40%	20	30
50%	25	25
60%	30	20
70%	35	15
80%	40	10
90%	45	5
100%	50	0

Heim jt (2012) poolt valitud pallide hulk on piisav, et vältida subitiseerimise efekti tekkimist (efektist täpsemalt peatükis 3.1).

Tarkvara abil määrati nii test- kui ka taustobjektide pallide suurused juhuslikult. Samuti oli juhuslik kõikide nende (ka testobjekti klatri) paiknevus. Sellega prooviti välistada assimilatsiooni ja suksessiivkontrasti⁸ tekkimist. Pallide vahelised kaugused olid siinses uurimuses erinevalt Heim jt (2012) seevastu väiksemad, mis oli tingitud MATLAB-i tarkvara eripäradest.

Katse jaoks loodi algselt kokku 22 pilt-stiimulit: ühest testobjektide proportsioonist oli algselt tehtud kaks varianti ehk 11 kollaste testobjektidega ja 11 sinistega testobjektidega pilt-stiimulit. Katsesse võeti kasutusele 20 pilt-stiimulit, kuna hulgasõna *vähe* ja *palju* abil ei saa tähistada vastavalt objekti null- ja maksimumhulka (nende tähendustest peatükkides 2.2.1 ja 2.2.3).

Pilt-stiimulite koostamisel olen samuti koostanud tabelleid oodatud tulemustega sellest, millist hulgasõna millise proportsiooni korral valitakse. Oletusi tehes põhinesin EKSS-i (2009) hulgasõnade iseloomustustele, võttes arvesse just nende eestikeelset tähendusspetsiifikat. Hulgasõnade oodatud jaotusi skaaladel on esitatud tabelites 2 ja 3.

Tabel 2. Umbmääraste hulgasõnade oodatud kasutusulatus vastavalt testobjektide proportsioonidele katsevooru skaala *mitte ükski* - ... - *kõik* puhul.

Oodatud hulgasõna	Arvuline väärtus (proportsioon)
Mitte ükski	0 (0%)
Mõni	5 (10%)
	10 (20%)
Mõned	15 (30%)
	20 (40%)

⁸ Assimilatsiooni ja suksessiivkontrasti tuleb vaadelda koos (McKenna 1984). Assimilatsiooni efekti all mõeldakse olukorda, kus katseisik tajub stiimulit n-ö tuttavana lähtuvalt eelmisest stiimulist või suuremast kontekstist üldiselt (Arielli 2012: 59). Suksessiivkontrast tekib siis, kui katseisik tajub esitatud stiimulit n-ö kaugemana või drastiliselt teistsugusena lähtuvalt eelmisest stiimulist (Arielli 2012: 59).

Pooled	25 (50%)
Enamik	30 (60%)
	35 (70%)
	40 (80%)
	45 (90%)
Kõik	50 (100%)

Tabel 3. Umbmääraste hulgasõnade oodatud kasutusulatus vastavalt testobjektide proportsioonidele katsevooru skaala *vähe* – ... – *palju* puhul.

Oodatud hulgasõna	Arvuline väärtus (proportsioon)
Vähe	5 (10%)
	10 (20%)
	15 (30%)
Keskmiselt	20 (40%)
	25 (50%)
	30 (60%)
Palju	35 (70%)
	40 (80%)
	45 (90%)

Hulgasõnade oodatud kasutusulatuse koostamisel lähtusin peamiselt EKSS-is (2009) toodud kirjeldustest.

Heim jt (2012) eeskujul koostati samuti harjutusvooru jaoks kolm eraldi pilt-stiimulit, kuid testobjektid võtsin osakaaluga 6% (3 palli), 64% (32 palli) ning 92% (46 palli). Nendest kahe

harjutusstiimuli testobjektide värvus oli sinine (64% ja 92%) ning ühe kollane (6%). Harjutusvooru stiimulitega koos uuritava skaalaga ja pilt-stiimuliga võib tutvuda Lisas 2.

4.1.2. Uuritavate hulgasõnade paigutus skaaladel

Hulgasõnade järjestamisel skaalasiseselt lähtusin Newsteadi & Collisi (1987: 1458) soovituselt skaalade uurimise kohta. Olen toonud piisavalt hulgasõnu (leksikaalseid väärtusi) lähtuvalt eesti keele eripäradest, mis kataksid kõiki võimalikke proportsioonide väärtusi ning välistaks valikutega manipuleerimist (Newstead & Collis 1987: 1458). Juhul, kui mõni hulgasõna osutub vastaja jaoks üleliigseks, jääb see tal ka valimata (Newstead & Collis 1987: 1458). Uuritavaid hulgasõnu skaalal paigutasin vastavalt nende loogilisele suurusjärgule (*magnitude*) väiksemast väärtusest suuremani, mis teeb katseisiku jaoks vastamise mugavamaks (Newstead & Collis 1987: 1458, vt ka Newstead & Coventy 2000; Coventry jt 2010). Tuleks täheldada, et mõned hulgasõnade uurimused on lähtunud just vastupidisest printsiibist (Pezzelle, Bernardi & Piazza 2018; Stateva jt 2019). Pezzelle, Bernardi & Piazza (2018) meelest võib selline lähenemine viia n-ö tutvumiseefektideni⁹ (*familiarization effects*), mis võivad teatud määral kallutada uurimuse tulemusi (Baumann, McFadyen & Humphreys 2020). Kuna tutvumiseefektide mõjust ja selle esinemisest on vähe uuritud (vt Baumann, McFadyen & Humphreys 2020), lähtusin oma katse koostamisel Newsteadi & Collisi (1987) soovitustest.

4.1.3. Katse pilootimine

Enne avalikustamist läbis katse kaks pilootimisvooru. Esimeses piloottestis osales kaks naissoost ja kaks meessoost testisikut vanuses 21 kuni 28 (keskmine: ~ 24 aastat). Kaks mõlemast soost testisikut läbisid katseversiooni koos harjutusvooriga, kaks ilma. Peale katse läbimist vastasid testisikud katse kohta käivate küsimustele, mille vastused protokolliti.

Esimese piloodikatse tulemusena:

⁹ Tutvumiseefekt tuleneb mingi katseosaga (nt stiimulid, katse vastused) liigse tutvumise tagajärjel. Katseisikul tekivad teatud assotsiatsioonid katseosade vahel, mis kokkuvõttes mõjutavad nende valikut. (Baumann, McFadyen & Humphreys 2020)

- likvideeriti katses esinenud väiksemad keelelised ebatäpsused ning tehnilised puudujäägid,
- pikendati skaala *mitte ükski* – ... – *kõik* ajalist piirangut, kuna mõnel pilootisikul ei jagunud aega stiimuli ja hulgasõnade skaalaga tutvumiseks ning seega otsuse tegemiseks,
- lisati *mitte ükski* - ... - *kõik* skaalale sõna *pooled* (eelnevalt ei olnud skaalal esitatud), kuna pilootisikute meelest ei sobinud *mõni* ja *mõned* hästi objekti 50% proportsiooni kirjeldamiseks ning skaalale jäi sisse leksikaalne tühik; *pooled* oli pakutud ka mitmuse vormina,
- muudeti uuritavate skaalade paigutust. Eelnevalt ülevalt-alla esitatud skaala, esitati nüüd vasakult-paremale, mis võimaldas katseisikul paremat ja mugavamat navigeerimist ning minimaliseeris võimalikke keskkonna visuaalse küljega seotud (nt ekraanisuurus) ebamugavusi,
- jäeti sisse harjutuskorrad katsega tutvumise ja harjumise eesmärgil.

Võttes arvesse esimese piloottesti tulemusi ning olles viinud sisse vastavad parandused, viidi läbi teine piloottest. Viimasest võttis seekord osa vaid kaks naissoost isikut vanuses 20 ja 22. Teise piloottestiga veenduti ja kinnitati tehtud muudatuste toimimine ja nende mugavus.

4.2. Katse ülesehitus

Siinses peatükis tutvustan täpsemalt katse ülesehitust. Annan ülevaate katse täpsemast struktuurist ning kirjeldan katse protseduuri.

4.2.1. Katse lühike iseloomustus

Katsega sooviti uurida üheksat erinevat eestikeelset hulgasõna, millest kuus paiknesid formaalloogilisel skaalal *mitte ükski* – *mõni* – *mõned* – *pooled* – *enamik* – *kõik* (lühidalt: *mitte ükski* – ... – *kõik*) ning kolm hinnanguskaalal *vähe* – *keskmiselt* – *palju* (lühidalt: *vähe* – ... –

palju). Hulgasonade valikuid objekti proportsiooni kirjeldamisel kasutati meetodina sunnitud valiku katset¹⁰ osalejasisese katseplaaniga¹¹.

4.2.1. Katse struktuur

Katseisik sai katset läbida LimeSurvey keskkonnas endale sobival hetkel.

Katse koosnes kahest katsevoorst ning neile eelnenud harjutusvoorst. Enne katsega alustamist kuvati katses osalejatele sissejuhatav tekst, millega tutvustati lühidalt katse sisu, eesmärki ja ülesannet (vt Lisa 2). Samuti enne alustamist paluti katseisikutel anda oma nõusolek katses osalemiseks (vt Lisa 3) ning ära täita taustainformatsiooni küsimustik (vt Lisa 4), milles küsiti erinevaid isikuandmeid, nagu vanus, sugu, domineeriv keel, muu keeleoskus, haridustase, emakeel ja väsimustase. Lisaks teavitati enne katse algust katseisikuid katse ja uurimuse eesmärgist üldiselt, vältides katseisiku mõjutamist ning soovitud vastuse poole suunamist. Katse läbiviimiseks eetilise jälgimine ja vajaliku andmekaitse tagamine oli kinnitanud Tartu Ülikooli inimuuringute eetika komitee poolt (taotluse number: 394/T-28).

Peale taustainformatsiooni küsimustiku täitmist algas harjutusvoor, mille järgnes kaks katsevooru. Harjutusvoorus näidati katses juhuslikus järjekorras selle kolmselle jaoks ettevalmistatud stiimulit, mis koosnes küsimusest ja pilt-stiimulist. Testobjektide osakaalu ning värvust valiti juhumeetodiga keskkonna RANDOM.ORG (Haahr 2025) abil. Vastamiseks oli all esitatud harjutusskaala *napilt – mõõdukalt – ohtralt*, mille vastuseid vaatasin vaid katsevooru vastuste valiidsuse hindamisel. Skaala võimaldas kõrvaldada erinevaid skaalade harjumisega seotud efekte (nt priming või ülehindamine), tagades samuti spontaanseid vastuseid, mis ei põhine eelmisel hinnangul. Ühe harjutusvooru stiimuliga on võimalik tutvuda Lisas 5.

¹⁰ Sunnitud valiku katse (*forced-choice preference task*) on katsemeetodites, mis võimaldab koguda tajumise ja mõistmisega seotuid andmeid mingi keelelise nähtuse kohta. Katseformaad näeb ette seda, et ülesande täitmiseks antakse katseisikule vaid piiratud arv vastuseid, mis olid eelnevalt uurija(te) poolt ette määratud. Katseisik omakorda peab valima ainult pakutud vastuste seast. Seetõttu on n-ö muude vastuste produtseerimine ja andmine välistatud. Rohkem objektiivsete tulemuste ja katseisiku valikuvabaduse tagamiseks peab katsekoostaja esitama võimalikult palju erinevaid vastuseid. (Gillioz & Zufferey 2020: 92–94; Klavan, Veismann & Jürine 2013: 23–24)

¹¹ Osalejasisene katseplaan (*within-subject design*, ka *repeated-measures design*) on üks andmete kogumisviisidest. Katseplaan põhineb n-ö “kõik läbivad kõike” põhimõttel ehk kõik osalejad võtavad kõikidest katseosadest (võrdselt) osa, sõltumata tingimuste (modaalsuste) muutumisest. Katseplaan sobib juhul, kui tahetakse uurida sõltuvat tunnust ennast ja selle esinemist inimeste seas. (Gillioz & Zufferey 2020: 19–21, 44)

Järgnevalt liiguti katsevoorude juurde, mille järjestus oli juhuslik. Igas katsevoorus uuriti korraga vaid ühte skaalat – kas formaalloogilist või hinnangulist. Katsevoorude stiimulid olid sarnase kujundusega, mis harjutusvooru omad (vt Lisa 6). Igale katseisikule kuvati katse jooksul ainult üht versiooni kahest sama proportsiooniga stiimulist. Korraga nähti ainult ühte pilti. Kui üks variant kindlast proportsioonist esines katse esimeses voorus, kuvati katses osalejale selle teine versioon katse teises voorus. Seetõttu kasutati skaala *mitte ükski* – ... – *kõik* korral kokku lähtuvalt selle dünaamikast ja sõnade enda tähendustest (vt peatükk 2.1) 6 kollaste testobjektidega ja 7 sinistega testobjektidega pilti-stiimulit. Skaala *vähe* – ... – *palju* korral kasutati ainult testobjektide vahepealsete väärtustega 5 (10%) ... 45 (90%) 5 kollaste testobjektidega ja 4 sinistega testobjektidega lähtuvalt hulgasõnade enda semantikast (vt peatükk 2.2). Kokkuvõttes, kuvati skaala *mitte ükski* – ... – *kõik* katsevooru jooksul osalejale 11 pilt-stiimulit ning skaala *vähe* – ... – *palju* katsevooru jooksul 9 pilt-stiimulit. Pilt-stiimulid jaotati kahe katseosa vahel juhuslikkuse alusel, kasutades keskkonda RANDOM.ORG (Haahr 2025), tingimusel, et sama värv ei korduks samas proportsioonis.

Katse läbimiseks pidi kuluma kuni 15 minutit.

4.2.2. Katse protseduur

Osaleja ülesandeks katses oli hinnata pilt-stiimulil esitatud testobjektide proportsioone nii harjutusvooru kui ka katsevoorude puhul. Katseisikutel paluti valida iga stiimuli kohta üks sõna alltoodud valimist, mis oleks lähtuvalt tema hinnangust testobjektide hulka kõige paremini iseloomustanud. Küsimuse lugemiseks ja sobiva hulgasõna valimiseks anti katses osalejale aega kas 10 (skaala *vähe* – ... – *palju*) või 12 sekundit (skaala *mitte ükski* – ... – *kõik*). Aeg oli piiratud, et maandada ülemõtlemise ja pallide ülelugemise riski. Mõlemad ajalised piirangud olid seatud vastavalt uuritava skaala enda pikkusele (koos küsimuse lugemise ja pallide koguse hindamisega) ning kinnitatud piloottestidega (vt täpsemalt peatükis 4.1.4). Katseisikul ei olnud võimalik katse ajal eelneva küsimuse juurde tagasi liikuda. Edasi liikumiseks pidi katseisik vajutama nupu „Järgmine“, mis paiknes ekraani all paremas nurgas (vt Lisa 5 või Lisa 6).

Enne harjutusvooru ning esimest ja teist katsevooru kuivati katseisikutele täpsemaid ülesande instruktsioone, millega saab vastavalt tutvuda Lisas 7, Lisas 8 ja Lisas 9.

4.3. Osalejate valim

Siinses peatükis tutvustan lähemalt katses osalejatega seotud aspekte. Annan ülevaate osalemiseks seatud kriteeriumitest, osalejate värbamisest ning andmete säilitamisest. Lisaks tutvustan lõppandmestikku ning selle puhastamise protsessi.

4.3.1. Katses osalejate kriteeriumid

Katsesse oodati eesti emakeelega inimesi vanuses 18 kuni 30 eluaastat. Tegemist oli mugavusvalimiga, millega sooviti hoida vanusest tekkida võivad keelelise eelistuse erisused kontrolli all. Katsesse ei sobinud isikud, kellel oli eelnevalt diagnoositud värvipimedus.

Katsesse eeldati saada võrdne hulk nais- ja meessoost katseisikuid. Esialgselt sooviti saada kuni 40 inimese vastuseid.

4.3.2. Katses osalejate värbamine ja andmete säilitamine

Katseisikud värvati isiklikult või Tartu Ülikooli meililistide kaudu (värbamiskutset vt Lisa 10). Andmete kogumine kestis ligikaudselt poolteist kuud LimeSurvey veebikeskkonnas.

Andmeid koguti isikustamata kujul. Katse käigus ei kogutud IP-aadresse ega isikut otse identifitseerida võimaldavaid andmeid. Isiku võimaliku kaudse tuvastamise vältimiseks kasutati andmete agregeerimist ja vajadusel ümberkodeerimist (nt grupeerime vastajad vanusrühmadesse).

4.3.3. Saadud andmestiku puhastamine

Katsest võttis kokku osa 78 vastajat. Enne analüüsi juurde asumist, vaadati käsitsi üle iga katses osaleja vastused ning nende vastavused etteseadud kriteeriumitele. Andmestiku puhastamise käigus eemaldati vastuseid, mis:

1. vastaja ei vastanud peatükis 4.3.1. toodud kriteeriumitele (nt vanusevahest välja jäämine; muu domineeriv keel, muu emakeel või muu sugu);
2. olid jäänud poolikuks kas enne või esimese katsekorra ajal;
3. olid ebakorrektse stiiliga ehk n-ö naljavastused.

Kokkuvõttes jäid alles 56 vastaja andmed.

4.3.4. Ülevaade katses osalejatest

Kokku kaasati analüüsi 56 katseisiku vastused. Keskmine vastajate vanus oli 25 eluaastat. Katses osales kokku 47 naist ning 9 meest. Keskmiselt osati kahte võõrkeelt, millest kolm sagedasemat olid inglise (~98%), vene (~51%) ja saksa (~26%). Katseisikute jaotus vastavalt haridustasemele on esitatud tabelis 4.

Tabel 4. Ülevaade haridustasemete jaotusest katses osalejate seas.

Katses osalejate haridustase	Vastajate arv
Keskharidus	23
Keskeri- või kutseharidus	1
Rakendus- või kutsekõrgharidus	2
Kõrgharidusepõhi- ehk bakalaureusekraad	20
Kõrghariduse kraadiõpe (magistri- või/ja doktorikraad)	10

Tabeli 4 järgi osales kõige rohkem keskharidusega katseisikuid (~41%) ning ligikaudselt sama palju kõrghariduse põhi- ehk bakalaureusekraadiga (~36%). ~18% vastajatest oli kõrghariduse kraadiõppega. Ülejäänud vastajad olid muu haridustasemega.

Tabel 5. Ülevaade katses osalejate jaotusest vastavalt elukohale.

Katses osalejale püsiv elukoht	Vastajate arv
Harju maakond	9
Tartu maakond	40
Muu	6 + 1(täpsustamata)

Tabeli 5 järgi märkisid ~71% katses osalejat püsivaks elukohaks Tartu maakonna. ~16% märkisid püsivaks elukohaks Harju maakonna ning 12,5% mõne muu piirkonna. Viimase hulgas olid nt Võrumaa, Viljandimaa ja Uppsala (Rootsi Kuningriik).

Tabel 6. Ülevaade katses osalejate väsimustasemetest vastavalt katse läbimishetkele.

Katses osalejate väsimustase	Vastajate arv
Väga väsinud	2
Väsinud	7
Pigem väsinud	11
Keskmine (ei väsinud ega ergas)	18
Pigem ergas	11
Ergas	6
Väga ergas	1

Tabeli 6 järgi hindas ~32% vastanutest oma väsimusastet enne katse läbimist keskmisena, ~20% pigem erksana ning ka ~20% pigem väsinuna. Ülejäänud katses osalejad (~28,5%%) valisid oma hetkelise väsimustase hindamiseks mõne muu vastusevariandi.

Katse tegemisele kulus katseisikutel keskmiselt 4 minutit ja 49 sekundit.

4.4. Andmete analüüsimeetodid

Andmete analüüsimiseks kasutasin vabalist statistikatarkvara R (R Core Team 2020).

Kahe skaala uuritavate hulgasõnade valikute ja kasutusalade uurimiseks kasutan kirjeldava statistika meetodeid (vt Winter 2020: 53–68). Esitan ja kirjeldan sagedustabeleid hulgasõnade valikutest testobjektide proportsioonide hindamisel ning võrdlen ülaltoodud varasemate uurimuste tulemustega. Lisaks toon ka tulemusi vastamiskiirusest. Vastamiskiiruse keskmise arvutamiseks testobjektide proportsioonide kohta kasutasin funktsiooni *mean*. Samuti arvutan ka standardhälve funktsiooniga *sd* vastuste kiiruste varieeruvuse vaatamiseks (Winter 2020: 55–57). Erinevate joonistuste visualiseerimiseks kasutan paketti *ggplot2* (Wickham 2016). Meetodiga loodan saada vastuse nii oma esimesele uurimisküsimusele kui ka pilku heita kaudsetele keeletöötlemisega seotud protsessidele, nagu nt sobiva hulgasõna valimisraskuste tuvastamine.

Selleks, et tuvastada erinevate tegurite mõju kahe skaala hulgasõnade valikule, kasutan kahte statistilist mudelit:

- tingimuslikke otsustuspuid (*conditional inference trees*). Tingimusliku otsustuspuu jagunemise aluseks on p -väärtus (alates < 0.05) (vt täpsemalt Levshina 2020: 611–643); ning
- tingimuslikke juhumetsi (*conditional random forest*). Tingimusliku juhumetsa jagunemise aluseks on samuti p -väärtus (alates < 0.05) (vt täpsemalt Levshina 2020: 611–643).

Nii tingimuslik otsustuspuu kui ka tingimuslik juhumets sobivad mitteparameetrilise jaotusega andmestikule, millega saab uurida ka keerulisemaid vastasmõjusid paljude seletavate tunnuste vahel. Tingimusliku otsustuspuu eelis on andmete intuitiivne tõlgendamine ning selle sobivus ka väiksema arvu vastuste korral. Tingimuslik juhumets on tingimuslikust otsustuspuid täpsem, kuna uuritava tunnuse ennustamiseks tehakse mitmeid erinevaid juhuslikke puid, mille abil saab paremini ennustada seletavate tegurite mõju uuritavale tunnusele. (Levshina 2020: 612–613) Minu uurimuses on seletavateks tunnusteks vastajast tulenevad tegurid ning uuritavaks tunnuseks hulgasõnade valikud vastavalt proportsioonidele. Mudelite abil soovin saada vastuse nii esimesele kui ka teisele uurimisküsimusele.

Tingimuslike otsustuspuude tegemiseks kasutan ühe skaala jaoks paketti *partykit* (Hothorn & Zeileis 2015) funktsiooni *ctree* (Hothorn, Hornik & Zeileis 2015) ning teise skaala jaoks paketti *party* (Hothorn jt 2025) funktsiooni *ctree* (Hothorn, Hornik & Zeileis 2015). Tingimuslike juhumetsade tegemiseks kasutan paketti *partykit* funktsiooni *cforest* (Strobl, Hothorn & Zeileis 2009). Mõlemate mudelite headuse hindamiseks toon samuti makro-keskmise AUC-i (*macro-average AUC*), kasutades paketti *pROC* funktsiooni *roc* (Robin jt 2021). Mudelite visualiseerimiseks kasutan samuti paketti *ggplot2* (Wickham 2016). Mudelite klassifitseerumistest tegin ja tõlgendasin vastavalt Maarja-Liisa Pilviku (ilmumas) käsitlusele. Sama käsitlust kasutasin ka makro-keskmise AUC-i tulemuste tõlgendamisel.

5. Katse tulemused ja analüüs

Siinses peatükis esitan katsega saadud tulemusi ning arutan nende üle. Peatükis 5.1 toon tulemused mõlema uuritava skaala hulgasõnade kasutusulatuste kohta ning võrdlen erinevate olemasolevate uurimuste tulemustega. Samuti vaatlen põgusalt vastamiskiirust. Peatükis 5.2 vaatlen erinevate katsega kogutud vastajatest tulenevate tegurite mõju uuritavate hulgasõnade valikule tingimuslike otsustuspuude ja tingimuslike juhumetsade abil.

5.1. Ülevaade hulgasõnade valikutest katses

Järgnevalt annan ülevaate kahe skaala hulgasõnade valikutest. Esimesena vaatan skaala *mitte ükski* – ... – *kõik* hulgasõnade valikut testobjektide proportsioonide kirjeldamisel. Selle järgneksid skaala *vähe* – ... – *palju* hulgasõnade valikute kirjeldamine. Siin esitaksin ülevaate vastamiskiiruse tulemuste kohta. Lõpus teen kokkuvõtte saadud tulemustest ning arutan nende üle.

5.1.1. Skaala *mitte ükski* – ... – *kõik* hulgasõnade valimissagedused

Koondvaade skaala *mitte ükski* – ... – *kõik* hulgasõnade valikutest vastavalt testobjektide (järgnevalt: objektide) proportsioonidele on esitatud tabelis 7. Tabelis esitatud tulemused põhinevad 56 katses osaleja vastusele. Analüüsist on välja jäetud puuduvad vastused ehk vastused, mille puhul jäi katseisikul ebamäärastel põhjustel vastus andmata. Lisaks arvulisele väärtusele on toodud samuti ka vastuste protsendiline osakaal vastavalt vastuste hulgale.

Lähtuvalt tabelis 7 esitatud tulemustest, võib teha teatud järeldusi skaala *mitte ükski* – ... – *kõik* hulgasõnade valikute kohta ning ka võrrelda neid tulemustega muukeelsete hulgasõnade kohta. Alustan analüüsiga skaala kõige väiksemast väärtusest ning liigun järjest suurema väärtuse juurde.

Tabel 7. Ülevaade skaala *mitte ükski* – ... – *kõik* hulgasõnade valikust testobjektide arvulise proportsiooni kirjeldamisel.

	(Test)objektide proportsioonid hulga tervikust										
Hulgasõnad	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
<i>Mitte ükski</i>	51 (91,1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
<i>Mõni</i>	0 (0%)	38 (69,1%)	13 (23,2%)	1 (1,8%)	1 (1,85%)	0 (0%)	0 (0%)	0 (0%)	1 (1,8%)	0 (0%)	0 (0%)
<i>Mõned</i>	0 (0%)	12 (21,8%)	43 (76,8%)	52 (94,55%)	40 (74,1%)	5 (8,9%)	2 (3,5%)	1 (1,8%)	0 (0%)	0 (0%)	0 (0%)
<i>Pooled</i>	0 (0%)	0 (0%)	0 (0%)	2 (3,65%)	12 (22,2%)	51 (91,1%)	19 (34%)	1 (1,8%)	0 (0%)	0 (0%)	0 (0%)
<i>Enamik</i>	0 (0%)	5 (9,1%)	0 (0%)	0 (0%)	1 (1,85%)	0 (0%)	35 (62,5%)	54 (96,4%)	54 (98,2%)	55 (98,2%)	1 (1,8%)
<i>Kõik</i>	5 (8,9%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1,8%)	55 (98,2%)
Vastuseid kokku	56	55	56	55	54	56	56	55	55	56	56

Mitte ükski

Hulgasõna *mitte ükski* kasutusulatus jäi tulemuste järgi 0% proportsiooni juurde objekti kirjeldamisel. Selline tulemus on kooskõlas Zaina Tsouhlarisi (2024) oletusega ingliskeelse vaste *none* kui puuduva hulga edastaja kohta. Selline tulemus seab omakorda küsimuse alla selle, kas hulgasõna *mitte ükski* saab pidada umbmääraseks, nagu on EKSS (2009) iseloomustanud hulgasõna *ükski*. Samuti sarnaneb eesti keele *mitte ükski* itaaliakeelse *nessuno*'ga, mis Pezzelle, Bernardi ja Piazza (2018: 121) järgi tähistab samamoodi 0% proportsiooni.

Lähtuvalt EKSS-i (2009) definitsioonist olen eelnevalt oletanud (vt ka peatükki 2.1 ja 4.1.1), et hulgasõna *mitte ükski* kasutusalal siinsel skaala võiks jääda 0% proportsiooni juurde. Tulemuste järgi vastas minu oletus hulgasõna jaotusele.

Mõni ja mõned

Hulgasõna *mõni* kasutusulatus jäi vahemikku 10% kuni 40% ning *mõned* vahemikku 10% kuni 50%. *Mõni* valiti sageli objekti 10% proportsiooni korral, vähesel määral ka 20% proportsiooni kirjeldamisel. Objekti 30% ja 40% proportsiooni markeerimisel valiti hulgasõna vaid üks kord. Seevastu *mõned* valiti rohkem vahemikus 20% kuni 40%, kuid hulgasõna valiti ka vähesel määral objekti 10% proportsiooni kirjeldamisel. Sellest järeldub, et hulgasõna *mõni* on väiksema kasutusulatusega kui *mõned*. Samuti valitakse *mõni* rohkem objekti 10% proportsiooni kirjeldamisel, hulgasõna *mõned* vahemiku 20% kuni 40% kirjeldamiseks. *Mõned*-hulgasõna eelistamine suuremate hulkade kirjeldamisel läheb kokku teoorias mainitud ikoonilisuse kvantiteedi põhimõttele (Jackobson 1971: 352; viidetud Haspelmath 2008: 2, 4), mille järgi viidetakse rohkematest üksustest koosneva sõnaga ka suuremale kogusele. Seega EKK-i (2020) väide *mõni* ja *mõned* hulgasõnade tähenduse samasusest ei leia katsetulemustest kinnitust vaatamata sellele, et mõlema hulgasõna kasutusulatuste vahe tundub olevat väike.

Tuleks täheldada, et võrreldes eeltoodud uurimustega, ei arvestanud mina hulgasõnade *mõni* ja *mõned* valikuid alatest 60% nende kasutusulatuste alla. Arvatavasti oli antud juhul tegemist *pöördvastustega*, mille korral katseisikud andsid vastuse mitte testobjekti, vaid taustobjekti kohta. Pöördvastuste esinemise üle arutatakse edasi peatükis 5.1.4.

Võrreldes teiste uurimustega indoeuroopakeelsete vastete kohta, tundub ka *mõni* ja *mõned* ühine kasutusulatus olevat rohkem spetsiifilisem. Hulgasõnade jaotus eesti keeles on väiksem

nii ingliskeelsest kui itaaliakeelsest vastest: ingliskeelse *some* kasutusulatus jääb Moxey (1986: 56) tulemustes vahemikku 25% kuni 75% ning itaaliakeelne *alcuni* aga Pezzelle, Bernardi & Piazza (2018:121) tulemuste järgi 10% kuni 75%. Kumbki neist jääb hulgasõnade kasutusulatuse 10% kuni 50% vahemikust suurem. Ka Stateva jt (2019: 8) näitavad samuti, et ingliskeelse *some* kasutusulatus on eestikeelsetest *mõni* ja *mõned* ühisvahemikust suurem, kuid saksakeelne *einige* ning prantsuskeelne *quelque* kasutusulatus lõpeb sarnaselt umbes 50% juures. Sloveeniakeelse *nekaj* kasutusulatus lõpeb seevastu juba umbes 39% juures (Stateva 2019: 8), mis sarnaneb *mõni*-hulgasõna vahemikuga.

Lähtuvalt EKSS-i (2009) iseloomustusest ja ikoonilisuse kvantiteedi printsiibist (Jackobson 1971: 352; viidatud Hapselmath 2008: 4) olen oletanud (vt ka peatükki 2.1 ja 4.1.1), et hulgasõnade *mõni* ja *mõned* kasutusulatused võiksid jääda skaala vastavalt 10% kuni 20% proportsioonide ning 30% kuni 40% vahemikutele. Tulemused aga näitasid, et *mõni* ja *mõned* kasutusulatus on suurem ning mõlemaid saab valida kuni 40% proportsiooni kirjeldamiseks. Seetõttu ei vasta nende oodatud jaotused minu tulemustega.

Pooled

Hulgasõna *pooled* kasutusulatus jääb vahemikku 30% kuni 70%. Kõige rohkem valiti hulgasõna objekti 50% proportsiooni kirjeldamisel, kuid suuremal määral ka 40% ja 60% proportsioonide puhul. Hulgasõna *pooled* valiti harva 30% ning 70% proportsiooni markeerimisel.

Stateva jt (2019: 9) toodud saksa-, itaalia-, prantsuse ja sloveeniakeelse vastete tulemuste järgi tundub, et siinses katses on eestikeelne *pooled* suurema kasutusulatusega, kuna kõik neli indoeuroopakeelset vastet on sarnase ulatusega umbes 40% kuni 60%. Seevastu eestikeelne jääb vahemikku 30% kuni 70%.

Lähtuvalt EKSS-i (2009) definitsioonile olen oletanud (vt ka peatükki 2.1 ja 4.1.1), et hulgasõna *pooled* kasutus võib jääda ainult 50% proportsiooni juurde. Tulemused seevastu näitasid, et hulgasõna *pooled* ei vasta ainult 50% proportsioonile, vaid on suurema kasutusulatusega. Seega hulgasõna oodatud jaotus skaalal ei leidnud kinnitust.

Tuleks täheldada, et minu töös on esitatud ainult hulgasõna mitmuse vorm *pooled*, mistõttu pole teada, kas hulgasõna ainsuse vormi toomine oleks uurimuse tulemusi muutnud.

Enamik

Hulgasõna *enamik* jäi kasutusulatusena 60% kuni 100%. Hulgasõna valiti palju objekti 70% kuni 90%, aga ka suurel määral 60% proportsiooni edastamisel. Üksikkorral valiti hulgasõna ka 100% proportsiooni markeerimisel. Tuleks täheldada, et hulgasõna valikuid 10% ja 40% proportsioonide korral pidasin samuti pöördvastusteks ning need tulemused ei kajastatud hulgasõna kasutusulatuse all.

Eestikeelse hulgasõna *enamik* kasutusala on osutunud väiksemaks võrreldes Moxey (1986: 51) ja Pezzelle, Bernardi & Piazza (2018: 121) tulemustega: ingliskeelne *most(ly)* jäi vahemiku 25% kuni 90% proportsiooni kirjeldamisel ning itaaliakeelse *la maggior parte* üldine kasutusulatus vahemikku 35% kuni 90%. Ka Stateva jt (2019: 10) uurimus on inglise-, saksa-, itaalia- ja sloveeniakeelsete vastete kasutusulatused suuremad, kuna algavad ligikaudselt umbes 49% või 51% juures. Seevastu eestikeelse *enamik* hakkab siinse katse raames alles 60% proportsiooni juures. Sellest võrdlusest lähenedes, on eestikeelne *enamik* kasutusulatuse poolest indoeuroopa vastete ulatustest väiksem ning seega erinev.

Mira Ariel (2004) on järeldanud, et kontekstis saab ingliskeelne *most* viidata kogu hulga, hõlmates igat entiteedi tervikkoguses ning olles sünonüümiks hulgasõnale *all*. Selline tähendusvarjund eestikeelsel hulgasõnal *enamik* pole ei EKSS-is (2009) ega teiste eesti keeleteadlaste poolt eelnevalt kajastatud. Siinse katse puhul ei pruugi siiski tegemist olla mingi suurema keelelise tendentsiga, kuna katses kasutatud valim on selleks järelduseks liiga väike. Seetõttu võib olla tegemist ka isikliku tajumisega seotud eripäraga, mitte suurema nähtusega. Kindlama järelduse jaoks on vajalik uus suurema valimiga uuring.

Lähtuvalt EKSS-i (2009) definitsioonist oletasin (vt ka peatükki 2.1 ja 4.1.1), et hulgasõna *enamik* kasutusulatus vastab hulga vahemikule 60% kuni 90%. Katsetulemused seevastu näitasid, et võimalik on ka hulgasõna valik 100% juures. Seetõttu hulgasõna oodatud jaotus kinnitust ei leidnud.

Kõik

Hulgasõna *kõik* kasutusulatus jäi vahemikku 90% kuni 100%. Hulgasõna valik domineeris 100% proportsiooni juures, kuid üksikkasutusena ka 90% juures. Pezzelle, Bernardi & Piazza (2018: 121) üldiste tulemuste järgi jäi ka itaalia keelne *tutti* samuti 100% juurde, kuid juhul, kui testobjekte oli võimalik kokku lugeda, valiti hulgasõna ka 80% ja 90% iseloomustamiseks.

Seega võib eeldada, et ka eestikeelse hulgasõna puhul võiks esineda sarnane tendents, kuid siinne valim on liiga väike sellise järelduse tegemiseks. Tuleks täheldada, et hulgasõna valikut 0% juures loeti siinses katses pöördvastuseks ega liigitanud neid hulgasõnade kasutusulatuse juurde.

Lähtuvalt EKSS-i (2009) iseloomustusest olen oletanud (vt ka peatükki 2.1 ja 4.1.1), et hulgasõna *kõik* kasutusulatus jääb 100% proportsiooni juurde. Katsetulemused näitasid, et hulgasõna võidakse valida ka objekti 90% proportsiooni kirjeldamisel. Seetõttu hulgasõna oodatud jaotus kinnitust ei leidnud.

Kokkuvõttes võib minu katsetulemuste põhjal järeldada, et skaala *mitte ükski* – ... *kõik* hulgasõnu valitakse teatud vahemikus, millega saab iseloomustada objekti hulka, kuid hulgasõnade kasutusulatuste piirid jäävad siis hägusateks. Järgnevalt toon tulemusi skaala *vähe* – ... – *palju* hulgasõnade ja nende kasutusulatuste kohta.

5.1.2. Skaala *vähe* – ... – *palju* hulgasõnade valimissagedused

Koondvaade skaala *vähe* – ... – *palju* hulgasõnade valikutest vastavalt testobjektide (järgnevalt: objektide) proportsioonidele on esitatud tabelis 8. Tabelis esitatud tulemused põhinevad 56 katses osaleja vastusele. Lisaks arvulisele väärtusele on toodud samuti ka vastuste protsendiline osakaal vastavalt vastuste hulgale.

Lähtuvalt tabelis 8 esitatud tulemustest, võib teha teatud järeldusi skaala *vähe* – ... – *palju* hulgasõnade valikute kohta ning ka võrrelda neid tulemustega muukeelsete hulgasõnade kohta. Alustan analüüsiga skaala kõige väiksemast väärtusest ning liigun järjest suurema väärtuse juurde.

Tabel 8. Ülevaade skaala *vähe* – ... – *palju* hulgasõnade valikutest testobjektide arvulise proportsiooni kirjeldamisel.

	(Test)objektide proportsioonid hulga tervikust								
Hulgasõnad	10%	20%	30%	40%	50%	60%	70%	80%	90%
<i>Vähe</i>	56 (100%)	56 (100%)	47 (83,9%)	18 (32,1%)	2 (3,6%)	0 (0%)	2 (3,6%)	0 (0%)	1 (1,8%)
<i>Keskmiselt</i>	0 (0%)	0 (0%)	8 (14,3%)	37 (66,1%)	50 (89,3%)	50 (89,3%)	12 (21,4%)	0 (0%)	0 (0%)
<i>Palju</i>	0 (0%)	0 (0%)	1 (1,8%)	1 (1,8%)	4 (7,1%)	6 (10,7 %)	42 (75%)	56 (100%)	55 (98,2 %)
Vastuseid kokku	56	56	56	56	56	56	56	56	56

Vähe

Hulgasõna *vähe* kasutusulatus siinse katsega jäi vahemikku 10% kuni 50%. Hulgasõna osutus absoluutseks valikuks testobjekti 10% ning 20% proportsiooni kirjeldamisel, kõrgelt eelistatuks 30% testobjektide korral. Samuti valiti *vähe*-hulgasõna vähesel määral 40% ja 50% testobjektide proportsiooni kirjeldamisel. Tuleks täheldada, et hulgasõna valikut 70% ja 90% proportsiooni edastamisel loeti hulgasõna esinemist pöördvastuseks ega liigitanud neid hulgasõnade kasutusulatuse juurde.

Eestikeelne hulgasõna *vähe* indoeuroopa vastetega võrreldes, on *vähe* kasutusulatus suurem. Juba eelmainitud uurimustest kattub *vähe* kasutusulatus vaid itaaliakeelse *pochi* üldise kasutusulatusega, mis samuti vastas vahemikule 10% kuni 50% (Pezzelle, Bernardi & Piazza 2018: 121). Seevastu Moxey (1986: 51) tuvastatud ingliskeelse *a few/few* ulatus esines vahemikus 10% kuni 25%, mis on eestikeelsest *vähe* väiksem. Ka Stateva jt (2019: 7) toodud tulemustest prantsuse-, saksa- ja ingliskeelse vastete kohta on eestikeelne *vähe* suurema kasutusulatusega, kuna kõige suurema kasutusulatusega *few* lõppeb umbes 41% juures. Nii sarnaneb *vähe* kasutusulatus vaid itaaliakeelse vastega; saksa-, prantsuse ja ingliskeelse vastetest on see aga suurem.

Lähtuvalt EKSS-i (2009) iseloomustusest olen oletanud (vt ka peatükki 2.2 ja 4.1.1), et hulgasõna *vähe* vastab 10% kuni 30% kasutusulatusele. Katsetulemused seevastu näitasid, et hulgasõna võidakse valida ka 40% ja 50% proportsiooni kirjeldamise korral. Seetõttu hulgasõna oodatud jaotus kinnitust ei leidnud.

Keskmiselt

Hulgasõna *keskmiselt* kasutusulatus jäi vahemikku 30% kuni 70% proportsioonide kirjeldamisel. Hulgasõna valiti tihti testobjekti 50% ja 60% proportsiooni kirjeldamisel, veidi vähem valiti 40% proportsiooni puhul. Vähesel määral valiti hulgasõna 30% ja 70% testobjektide proportsioonide kirjeldamisel.

Eelnevalt oletasin (vt ka peatükki 2.2 ja 4.1.1), et hulgasõna *keskmiselt* kasutusulatus vastab vahemikule hulga 40% kuni 60%. Katsetulemused omakorda näitasid, et hulgasõna kasutusulatus on testobjektide proportsioonide kirjeldamisel tegelikkuses suurem – 30% kuni 70%. Seetõttu hulgasõna oodatud jaotus ei leidnud kinnitust.

Palju

Hulgasõna *palju* valiti testobjekti 50% kuni 90% proportsiooni kirjeldamisel. Hulgasõna osutus absoluutseks valikuks 80% proportsiooni ning peaaegu ainuvalikuks 90% proportsiooni juures. Hulgasõna *palju* oli suure valimissagedusega ka 70% proportsiooni korral. Vähesel määral valiti hulgasõna 50% ja 60% proportsiooni kirjeldamisel. Hulgasõna valikut 30% ja 40% proportsiooni kirjeldamisel loeti siinses katses pöördvastuseks ega liigitanud neid hulgasõnade kasutusulatuse juurde.

Hulgasõna *palju* kasutusulatus ei vasta täpselt selle itaalia- ning ingliskeelse vastete tulemustele: ingliskeelse *many* kasutusulatus jääb vahemikku 40% kuni 90% (Moxey 1986: 51) ning itaaliakeelse *pochi* oma üldiste tulemuste järgi jääb vahemikku 35% kuni 95% (Pezzelle, Bernardi & Piazza 2018: 121). Seega eestikeelse *palju* kasutusulatus 50% kuni 90% on väiksem, siit ka spetsiifilisem, kui itaaliakeelse ja ingliskeelse vaste oma. Kokkuvõttes, on eestikeelne *palju* itaalia- ja ingliskeelsest vastest väiksema kasutusulatusega.

Eelnevalt oletasin (vt ka peatükki 2.2 ja 4.1.1), et hulgasõna *palju* jääb valikute poolest vahemikku 70% kuni 90% testobjekti proportsioonide kirjeldamisel. Katsetulemused omakorda näitasid, et hulgasõna kasutusulatus algab juba 50% proportsioonist. Seetõttu hulgasõna oodatud jaotus ei leidnud kinnitust.

Kokkuvõttes võib minu katsetulemuste põhjal järeldada, et skaala *vähe* – ... – *palju* hulgasõnad vastavad teatud kasutusulatusele, millega saab iseloomustada objektide hulka. Seevastu kasutusulatuste piirid on ebamäärased. Täpsem arutelu saadud tulemuste üle jätkub peatükis

5.1.4.

5.1.3. Ülevaade katses osalejate vastamiskiiruse kohta

Siinses peatükis teen ülevaate katses osalejate vastamiskiirustest kahe uuritava skaala hulgasõnade valimisel. Enne analüüsi juurde asumist, vaatasin manuaalselt üle iga katseisiku vastamiskiirust mõlema skaala proportsioonide kohta ning veendusin, et andmestikus ei esine liiga suuri aegu, mis viitaksid katseisikute hõivatusle muude tegevustega ning tekitaksid analüüsimisel teatud “müra”.

Vastamisaja puhul vaatasin peamiselt kahte parameetrit – katses osalejate vastamisaja keskmine (*mean*) ning standardhälve (*standard deviation*) iga proportsiooni kohta. Standardhälve näitab katseisikute vastustamiskiirse tulemuste hajuvust võrreldes keskmisega (Winter 2020: 55–57). Seega viitab suurem standardhälve sellele, et katses osalejate vastamiskiirused on olnud ebaühtlased. Standardhälve võib nii osutada teatud valimiskeskustele nagu sobiva hulgasõna valimisele või muudele katseisiku keskkonnast tulevatele teguritele, mida pole võimalik siinse katseformaadi tõttu tuvastada. Skaala *mitte ükski* – ... – *kõik* vastamiskiiruse tulemused on esitatud tabelis 9, skaala *vähe* – ... – *palju* tulemused tabelis 10.

Tabel 9. Ülevaade katses osalejate vastamiskiirustest skaala *mitte ükski* – ... – *kõik* hulgasõnade valikul 11 erineva objekti proportsiooni kirjeldamisel.

Aeg	Osalejaid	Keskmine	St.hälve
Testobjekt_0%	56	5.5	2.0
Testobjekt_10%	56	6.8	2.9
Testobjekt_20%	56	7.2	2.4
Testobjekt_30%	56	7.0	2.5
Testobjekt_40%	56	7.6	2.6
Testobjekt_50%	56	6.2	2.3
Testobjekt_60%	56	7.4	2.8
Testobjekt_70%	56	6.4	2.6
Testobjekt_80%	56	5.9	2.4
Testobjekt_90%	56	5.3	1.8
Testobjekt_100%	56	5.0	2.1

Tabeli 9 põhjal on näha, et katse osalejate seas esines ajaline varieeruvus sobiva hulgasõna valimisel erinevate proportsioonide kirjeldamisel. Kõige suuremad vastamisaegade keskmised esinesid 40% ja 60% proportsiooni kirjeldamise juures. Nende suured standardhälved viitavad omakorda samuti vastamiskiiruste suurele varieeruvusele katseisikute seas. Seega saaks oletada, et sobiva hulgasõna valik oli katses osalejate jaoks üldiselt ajakulukas, kuid mõnede osalejate jaoks osutus hulgasõna valik ka keskmisest raskemaks. Kõrge standardhälve esines ka 10% proportsiooni hindamisel, mille keskmine ei ole seevastu kõige kõrge. See võiks omakorda viidata teatud valimiskeskustele proportsiooni kirjeldamisel mõne katseisiku puhul. Kõige väiksema vastamisaja keskmisega oli proportsioonide 90% ja 100% kirjeldamisel koos madalate standardhälvetega, millest sobiva sõna valik ei osutunud katseisikute jaoks

keeruliseks. Ka 0% proportsiooni kirjeldamine sujus keskmise alusel kiiresti. Niisiis oli tõenäoliselt katses osalejate jaoks kõige raskem leida sobivat hulgasõna 40% ja 60% ning lihtsam 90% ja 100% proportsioonide kirjeldamiseks.

Järgnevalt vaatan ka skaala *vähe* – ... – *palju* vastamiskiiruse tulemusi erinevate proportsioonide kohta.

Tabel 10. Ülevaade katses osalejate vastamiskiirustest skaala *vähe* – ... – *palju* hulgasõnade valikul 9 erineva objekti proportsiooni kirjeldamisel.

Aeg	Osalejaid	Keskmine	St.hälve
Testobjekt_10%	56	4.9	1.5
Testobjekt_20%	56	5.3	1.7
Testobjekt_30%	56	6.1	2.1
Testobjekt_40%	56	6.4	2.1
Testobjekt_50%	56	5.1	1.7
Testobjekt_60%	56	5.2	1.7
Testobjekt_70%	56	5.7	2.0
Testobjekt_80%	56	5.2	1.7
Testobjekt_90%	56	5.0	1.8

Tabeli 10 järgi võib järeldada, et katses osalejate seas esines ajaline varieeruvus sobiva hulgasõna valimisel erinevate proportsioonide kirjeldamisel. Kõige kõrgemad ajalised keskmised esinesid testobjekti 30% ja 40% proportsioonide kirjeldamisel, mille standardhälbed on samuti kõrged ning viitavad vastamiskiiruste varieeruvusele katses osalejate seas. Sellest võiks oletada, et sobiva hulgasõna valik oli katses osalejate jaoks üldiselt ajakulukas, kuid mõnede osalejate jaoks osutus hulgasõnade valik raskemaks. Kõrge standardhälve esines ka 70% proportsiooni korral, mis osutab samuti vastuste varieeruvusele. Koos keskmisega võib aga järeldada, et vaid mõne üksiku katses osaleja jaoks osutus hulgasõna valik keeruliseks. Kõige väiksemad ajalised keskmised esinesid 10% ja 90% proportsioonide kirjeldamisel, mille standardhälve ei viita suurele vastuste varieeruvusele, mis väljendab omakorda kindlust vastamisel vastajate seas. Kokkuvõttes, oli tõenäoliselt katses osalejate jaoks kõige raskem leida sobivat hulga 30% ja 40% ning lihtsam 10% ja 90% proportsiooni kirjeldamiseks.

Heim jt (2012) on oma uurimuses tuvastanud, et testobjekti hulga lähenemisel taustobjekti omale võiks vastamisaja tulemused suureneada töötlemiskeerukuse tõttu. Minu katse põhjal ei saa otseselt midagi järeldada töötlemise kohta, kuid saab mõlema uuritava skaala kohta

kaudselt välja tuua nende hulgasõnade valimisega seotud raskusi vastamisaja keskmise ja standardhälve põhjal. Minu tulemuste põhjal osutus hulgasõnade valik skaala *mitte ükski* – ... – *kõik* keerukaks 40% ja 60% (ka üksikute jaoks 10%) proportsioonide ning skaala *vähe* – ... – *palju* 30% ja 40% (üksikute jaoks 70%) proportsioonide kirjeldamine. Seega Heim jt (2012) järeldusele vastavad mõlemad skaalad vaid osaliselt. Siiski selle kinnitamine vajaks katse läbiviimist kontrollitud keskkonnas, kus vastamisaega saaks käsitleda objektiivse näitajana.

Katsetulemused näitavad, et skaala *mitte ükski* – ... – *kõik* puhul esines suurem ajakulu ja vastamiskiiruste hajutavus 40% ja 60% proportsiooni juures ning skaala *vähe* – ... – *palju* puhul ilmnes sama tendents 30% ja 40% proportsiooni kirjeldamisel. Vähem varieeruvust vastamisaegade seas esines mõlema skaala äärmuste puhul, mis olid ka väikse vastamiskiiruse keskmisega.

5.1.4. Vahepealne arutelu

Skaalade *mitte ükski* – ... – *kõik* ja *vähe* – ... – *palju* hulgasõnade valimissageduse põhjal võib järeldada, et eestikeelsed hulgasõnad vastavad teatud kasutusulatustele, mis on sisuliselt laiad kindlate piirideta. Peatükis 5.1.1 ja 5.1.2 esitasin ja arutasin katse uuritud kahe skaala hulgasõnade kasutusulatuste ja valikute kohta ning võrdlesin neid olemasolevate uurimuse tulemustega võimalike vastete kohta. Siinse katse tulemuste põhjal jaotuvad hulgasõnad skaaladel järgnevalt:

1. *Mitte ükski* (0%) < *mõni* (10%–40%) ≤ *mõned* (10%–50%) ≤ *pool* (40%–70%) ≤ *enamik* (60%–100%) ≤ *kõik* (90%–100%);

2. *vähe* (10%–50%) ≤ *keskmiselt* (30%–70%) ≤ *palju* (50%–90%).

Sellise kokkuvõtte järgi võib näha, et suurem osa hulgasõnade kasutusulatusi kattuvad omavahel objekti proportsioonide kirjeldamisel, mida on eelnevalt ühel või teisel määral tuvastatud ka nt inglise- ja itaaliakeelsete vastete kohta (vt nt Moxey 1986; Newstead & Coventry 2000; Pezzelle, Bernardi & Piazza 2018). Seega saab järeldada, et eestikeelsetele hulgasõnadele on omane teatud kasutusulatus, kuid selle piirid on ebamäärased. Erandiks on vaid hulgasõna *mitte ükski*, mida valiti ainult 0% proportsiooniga hulga kirjeldamisel. See omakorda tõstatab küsimuse, kas *mitte ükski* saaks üldse käsitleda umbmäärase hulgasõnana või tegemist oleks selgelt tühihulka väljendava lekseemiga, nagu on seda eelnevalt täheldanud

Tsouhlaris (2024) ingliskeelse *none* kohta. Seevastu selle vastand *kõik* valiti vaid üks kord objekti 90% proportsiooni kirjeldamisel. Samuti valiti ka hulgasõna *enamik* objekti 100% proportsiooni edastamisel. *Enamik* valiku korral 100% juures võis esineda sama efekt, mis esines ka Newstead & Coventry (2000) uurimuses, millega leiti, et katses osalejad võisid tajuda pallidega kaussi omaette anumana. Seega on võimalik, et ka siinse katse korral hindas katses osaleja testobjekti 100% proportsiooni korral kogu stiimuli ruutu ühe pallide hoiustava anumana, kuid ei olnud pallidega piisavalt täidetud, jättes “mittetäidetud” mahuti tunde, ning sellest valis hulgasõna *enamik*. Siiski on tegemist vaid oletusega, kuid nähtus vajaks edasist uurimist. Kokkuvõttes, on skaalal *mitte ükski* – ... – *kõik* hulgasõnad ebaühtlase jaotusega, neist kõige suurema kasutusulatusena sõnad *mõned* ja *enamik*. Skaalal *vähe* – ... – *palju* jaotusid hulgasõnad aga võrdselt.

Suurema ebaühtlase jaotusega hulgasõnad esinevad samuti skaalal *mitte ükski* – ... – *kõik* puhul. Pezzelle, Bernardi & Piazza (2018) on üheksa uuritava itaaliakeelse hulgasõna kohta järeldanud, et nendest koosneval skaalal toimub suurem kattumine just väiksema väärtustega hulgasõnade vahel, kuid suuremajärgulised jaotuvad ühtlasemalt. Sarnast tahaks kinnitada ka *mitte ükski* – ... – *kõik* skaala hulgasõnade kohta, kuid sinne katse ei soosi sellist järeldust. Peamiseks põhjuseks võiks olla skaalale saanud hulgasõnade valim, mille seas esines rohkem väiksemaid väärtusi kattuvaid leksikaalseid üksusi (*mõni* ja *mõned*) ning skaala ülemise otsa katmiseks toodi uurimusse vaid hulgasõna *enamik* (selle põhjusest peatükis 3.1.4). Seega ei ole teada, kas hulgasõna *enamik* mitmuse vormi esinemine oleks antud skaala jaotusi muutnud, kuid sellisel juhul oleks saanud järeldada rohkem skaala hulgasõnade ebaühtlaste jaotuste kohta. Rohkem objektiivsete tulemuste saavutamiseks tuleks minu meelest viia läbi korduskatse, kus on sisse arvestatud ka *enamik* mitmuse vorm *enamikud*, mille esinemine keeles on üks eesti keele eripäradest. Selle puudumist võib antud juhul pidada siinse uurimuse üheks puudujäägiks. Samuti ei ole teada, kas hulgasõnade jaotusi oleks samuti mõjutanud ka hulgasõnade *pooled* ainsuse vormi *pool* ning *enamik* sünonüümi *enamus*’e esinemine skaalal. Skaalal *vähe* – ... – *palju* hulgasõnade jaotused olid ühtlased, nende kasutusulatuste kattumisele vaatamata.

Tulemustest selgus, et hulgasõnade valik objekti proportsioonide kirjeldamisel vajaks edasist uurimist. Katse üheks suuremaks puudujäägiks on puuduvate vastuste esinemine. Nende esinemine tekkis keskkonna tehniliste tõrgete tõttu, kuna sunnitud valiku katse ei näe ette vastuse mitte andmise võimalust. Siiski võivad nad kaudselt viidata sobiva hulgasõnade valimisega seotud keerukustustele, eriti nt keskmisest suurem ajakulu või ajalise piirangu ületamine. Taas, puuduvate vastuste esinemist võib taandada kontrollitud keskkonna abil.

Samuti ei ole teada midagi nt kuidas oleksid skaala *vähe* – ... – *palju* hulgasõnade jaotusestest juhul, kui hulgasõnad *vähe* ja *palju* oleksid esitatud koos intensiivistajatega (nendest Metslang 2017: 476; vt ka Aasa 2024). Võib samuti oletada, et eestikeelsed hulgasõnad vastavad ligikaudsele arvulisele süsteemile, mida on oletatud nt ingliskeelsete hulgasõnade uurimisel (Pietroski jt 2009; Coventry jt 2010, Olm jt 2014), kuid vajaks edasist uurimist.

Siinse katse tulemused näitavad, et testobjekti hulga proportsiooni kirjeldamisel valisid katseisikud erinevaid hulgasõnu. Sellest võib järeldada, et hulgasõnade kasutuspiirid on ebamääraseks. Siiski eksisteerib eesti emakeele kõnelejate seas mingi üldisem ja abstraktsem arusaam hulgasõnade võimalikest kasutusest. Vastamiskiiruse tulemused viitavad samuti, et teatud proportsiooni kirjeldamise puhul on sobiva hulgasõna valik keerukam. Kokkuvõttes vastavad hulgasõnad teatud kasutusulatustele, kuid nende piirid on hägusad ning sõltuvad arvatavasti inimese enda eeldustest hulkade proportsioonide kohta. Leian, et olen analüüsi ja aruteluga saanud vastuse oma esimesele uurimisküsimusele.

5.2. Hulgasõnade valikut mõjutavate tegurite analüüs

Järgnevalt vaatan lisaks proportsioonidele ka vastajaspetsiilifiste tegurite mõju hulgasõnade valikul ja ennustamisel. Erinevate tegurite mõju hulgasõnade valikule testobjekti proportsioonide kirjeldamisel väljaselgitamiseks kasutan kahte statistilist mudelit (vt Levshina 2020: 611–643): tingimuslikke otsustuspuid (*conditional trees*) ning tingimuslikke juhumetsi (*conditional random forest*). Iga mudeli juures toon ka selle headuse näitajad ja klassifitseerimistäpsuse. Kuna minu valim võis osutada statistiliste mudelite jaoks väikseks, olen mõlema tingimusliku otsustuspuid korral suurendanud *minbucket*’iga oma vaatluste arvu 50ni.

Järgnevalt esitan analüüsi kummagi skaala kohta eraldi.

Skaala mitte ükski – ... – kõik

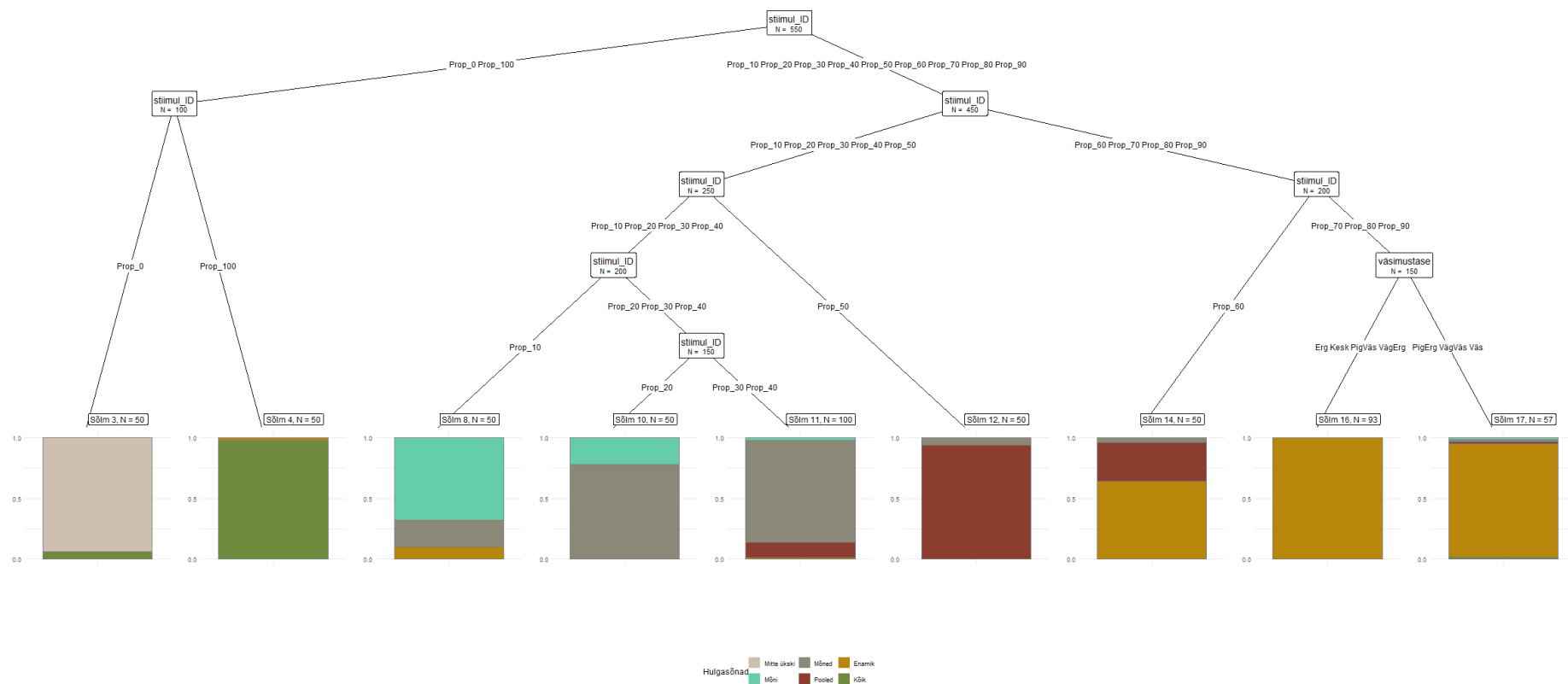
Selleks, et saaksin rakendada soovitud mudeleid, pidin algselt viima oma andmestiku laiaast kujust pikale, kasutades *tidyr* paketi funktsiooni *pivot_longer* (Mailund 2022). Sellega sain andmestiku, kus iga uuritav hulgasõna oli eraldi real. Mõlema mudeli tegemiseks kasutasin

vaid 52 vastaja andmeid, kelle vastuse seast ei leidunud ühtegi puuduvat vastust. Seega mõlemad järgmised mudelid põhinevad 572 vastusel (52 vastajat, 11 proportsiooni).

Erinevate vastajast tulenevate tegurite mõju uurimiseks hulgasõnade valikule kasutasin tingimuslikku otsustuspuud, mis on esitatud joonisel 4. Mudel on klassifitseerimistäpsusega 86.61%, mis näitab, et mudel suudab ennustada klasse hästi. Ka klasse eristab see hästi (*macro-average AUC* = 0.973). Alltoodud tingimusliku otsustuspuu mudelis esineb uuritava tunnusega hulgasõnade valik ning seletavateks tunnusteks olid määratud katses osalejate sugu, väsimustase, haridustase, omandatud keeled ja elukoht ning samuti ka stiimulitel esitatud objektide hulga proportsioonid (stiimul_ID).

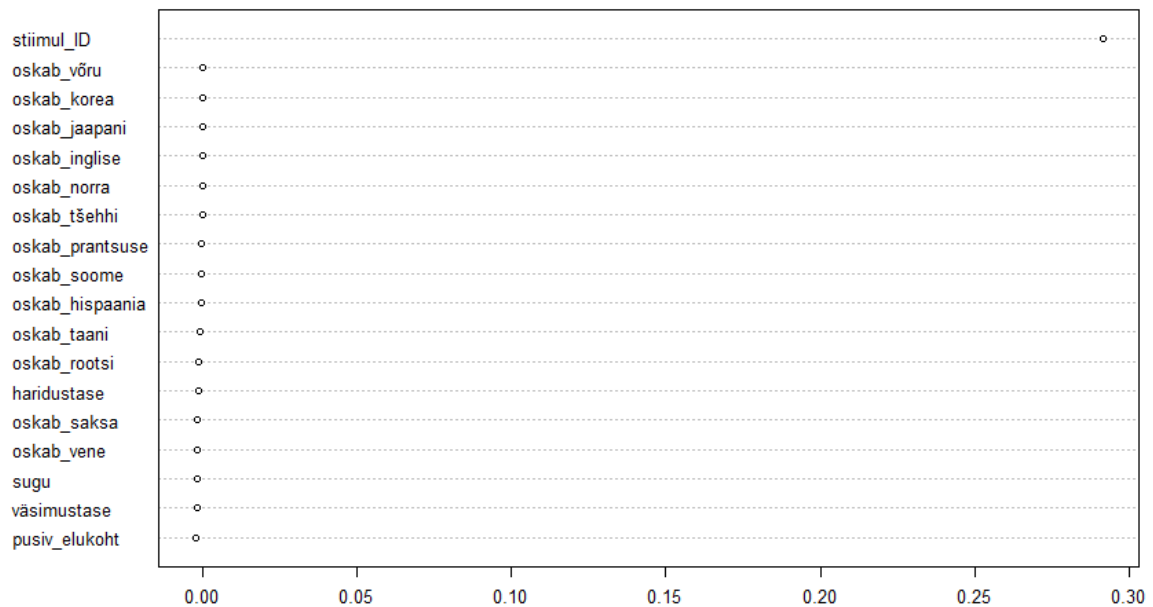
Joonisel 1 kujutatud mudeli järgi on näha, et peamiseks hulgasõnade valikut seletavaks tunnuseks on tunnus Stiimul_ID ehk stiimulitel toodud objektide proportsioonide, mis on ka mõjutanud hulgasõnade valikut. Mudeli järgi on näha, et hulgasõnade *mitte ükski* ja *kõik* valik on mõjutatud ka vastavalt hulga äärmuslikke proportsioonide (joonisel Prop_...) esinemisest, nagu 0% ja 100%. Hulgasõnu *mõni* ja *mõned* valimiseks on oluline objektide hulga vastamine proportsioonide vahemikule 10% kuni 40%, kuid *mõni* saab rohkem vaid 10% korral. *Pooled-*hulgasõna valik oli mõjutatud objektide hulga esinemisest 50% proportsioonis. Objektide 60% proportsioon võib tingida suurema hulgasõnade varieeruvuse valimisel. 70% kuni 90% proportsioonide juures osutus oluliseks mõjuteguriks väsimustase. On näha, et *väga ergas* (VägErg), *ergas* (Erg), *keskmine* (Kesk) ja *pigem väsinud* (PigVäs) väsimustasemetega katses osalejad valivad ühtlaselt objektide 70% kuni 90% proportsioonide kirjeldamiseks hulgasõna *enamik*. Seevastu väsimustasemetega *pigem ergas* (PigErg), *väsinud* (Väs) ja *väga väsinud* (VägVas) katseisikud valivad ka teisi skaalal esitatud hulgasõnu proportsiooni edastamiseks. Teiste oletatud vastajaspetsiifiliste tegurite mõju hulgasõnade valikule objektide proportsioonide kirjeldamisel ei leitud.

Niisiis langevad siinse tingimusliku otsustuspuu tulemused kokku ka hulgasõna valikute sageduse ja osalejate vastamiskiiruste tulemustega (vt peatükke 5.1.1 ja 5.1.3): suurem hulgasõnade valikute varieerumine esineb 10%, 30%, 40% ja 60% proportsioonide korral, mille keskmine aeg koos standardhällbega oli samuti kõrgem, osutades kaudsetele valimiskäskudele. Sellega leian taas vastuse oma esimesele uurimisküsimusele skaala *mitte ükski* – ... – *kõik* kohta.



Joonis 1. Tingimusliku otsustuspuu mudel ($minbucket = 50$, $alpha = 0.01$) erinevate tegurite olulisusest skaala *mitte ükski* – ... – *kõik* hulgasõnade valikule.

Järgnevalt toon tingimusliku juhumetsa mudeli erinevate tegurite olulisuse ennustamiseks hulgasõnade valikule, millega saab vaadelda täpsemalt ka väsimustaseme mõju. Mudel on esitatud joonisel 2. Selle klassifitseerimistäpsus on 80% ehk suudab hästi ennustada klaase. Mudeli makro-AUC on samuti ligikaudselt 0.72, mis näitab, et mudel eristab tunnuste põhjal klasse hästi.



Joonis 2. Tingimusliku juhumetsa mudel erinevatest skaala *mitte ükski* – ... – *kõik* hulgasõnade valikut ennustavatest teguritest.

Joonisel 2 toodud mudeli järgi võib näha, et skaala *mitte ükski* – ... – *kõik* hulgasõnade valikut ennustab vaid objektide enda proportsioonid. Vastajast tulenevad tegurid, nagu sugu, väsimustase, omandatud keeled (siin iga keel eraldi tunnusealgusega oskab_...), elukoht ja haridustase ei ennusta seevastu hulgasõnade valikut. Sellest võib järeldada, et väsimustase mõju 70% kuni 90% proportsiooni juures oli arvatavasti juhuslik. Selline vastus võis olla omakorda oletatavalt tingitud, mille korral osutus üksikkasutus statistiliselt oluliseks väheste andmete ja seega vastuste varieeruvuse tõttu. Sellega leian vastuse oma teise uurimisküsimusele skaala *mitte ükski* – ... – *kõik* kohta.

Kokkuvõttes osutusid skaala *mitte ükski* – ... – *kõik* hulgasõnade valikul olulisteks teguriteks ja ennustajateks vaid objektide enda proportsioonid. Vastajast tulenevate teguritest osutus tingimusliku otsustuspuu alusel oluliseks vaid väsimustase objektide 70% kuni 90%

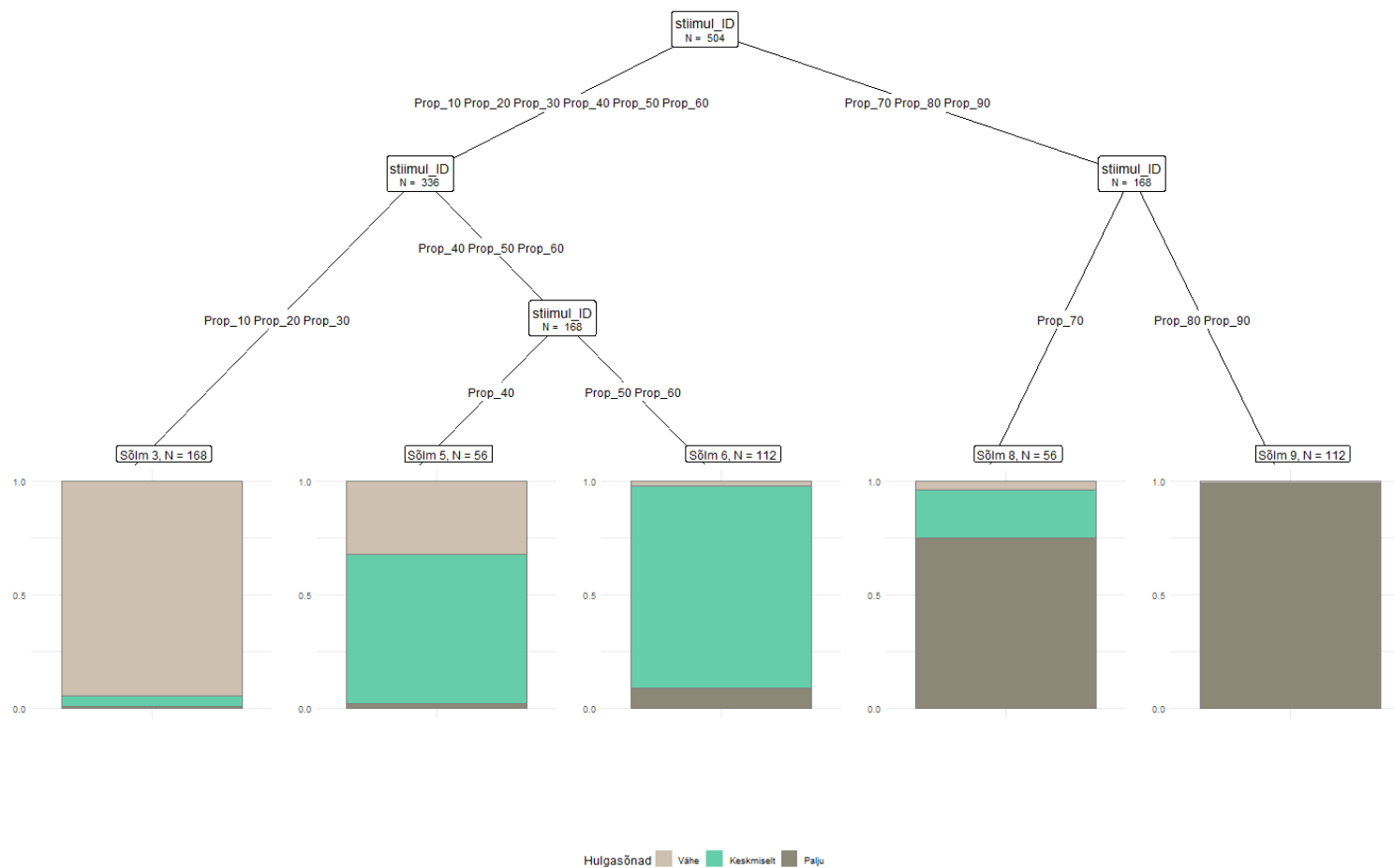
proportsioonide kirjeldamisel, mida pidasin juhuslikuks. Ükski teine vastajaspetsiifiline tegur ei osutunud oluliseks ega ennustanud hulgasõnade valikut.

Skaala vähe – ... – palju

Selleks, et saaksin rakendada soovitud mudeleid, viisin uuesti andmestiku laiast kujust pikale, kasutades *tidyr* paketi funktsiooni *pivot_longer* (Mailund 2022). Sellega sain andmestiku, kus iga uuritav hulgasõna esines eraldi real. Seega 56 vastaja andmete põhjal sain 9 stiimuli (= 9 proportsiooni) puhul kokku 504 vastust.

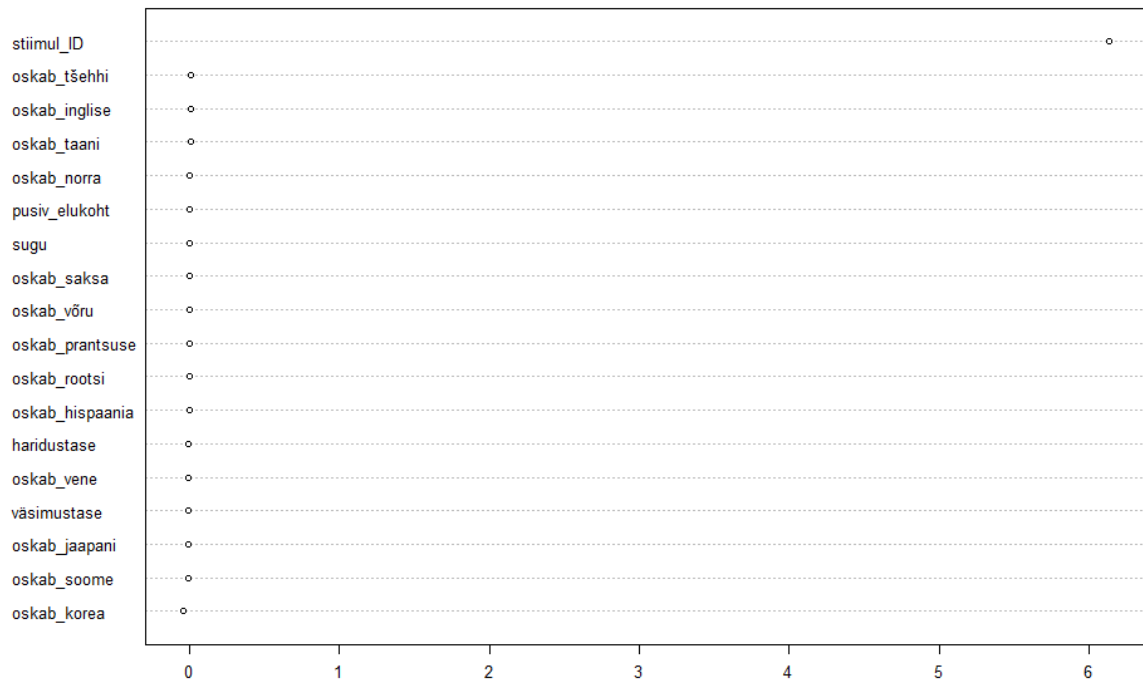
Erinevate vastajast tulenevate tegurite mõju uurimiseks hulgasõnade valikule kasutasin tingimuslikku otsustuspuud, mis on esitatud joonisel 3. Mudel on klassifitseerimistäpsusega 89.09%, mistõttu ennustab mudel hästi erinevaid klasse. Samuti eristab hästi ka klasse (*macro-average AUC* = 0.96). Tingimusliku otsustuspuu mudel näitab, et ükski vastajast tulenev tegur ei osutunud oluliseks. Sarnaselt eelmise skaala mudelile, on ka siinse mudeli uuritavaks tunnuseks hulgasõnade valik ning seletavateks tunnusteks katses osalejate sugu, väsimustase, haridustase, omandatud keeled ja elukoht ning samuti ka stiimulitel esitatud objektide hulga proportsioonid (*stiimul_ID*).

Joonisel 3 toodud mudeli alusel võib näha, et, sarnaselt eelmisele skaalale, on peamiseks hulgasõnade valikut mõjutavaks teguriks objektide enda proportsioonid. Hulgasõna *palju* olulisteks mõjutajateks on proportsioonid (tabeli tunnusealgusega *Prop_...*) 70% kuni 90%, kuid 70% proportsiooni korral esineb rohkem varieerumist hulgasõnade valikutes. Proportsioonid 40% kuni 60% kutsuvad esile rohkem hulgasõna *keskmiselt* valikut, kuid 40% proportsiooni juures on üldiselt vastuste varieerumine suurem ning esineda saab ka hulgasõna *vähe*. Hulgasõna *vähe* valikut tingivad proportsioonid 10% kuni 30%, mille puhul on 30% proportsioon suurema valikute varieeruvusega. Seega langevad tingimusliku otsustuspuu tulemused skaala hulgasõnade valikute sageduse ning osalejate vastamiskiiruste tulemustega (vt peatükke 5.1.1 ja 5.1.3): suurem hulgasõnade valikute varieerumine esineb 30%, 40% ja 70% proportsioonide puhul, mille keskmine aeg koos standardhälbega oli samuti kõrgem, mis omakorda viitab kaudselt valimisraskustele. Sellest saan taas vastuse oma esimesele uurimisküsimusele skaala *vähe – ... – palju* kohta.



Joonis 3. Tingimusliku otsustuspuu mudel ($minbucket = 50$, $alpha = 0.01$) erinevate tegurite olulisusest skaala vähe – ... – palju hulgasõnade valikule.

Järgnevalt esitan tingimusliku juhumetsa mudeli erinevate tegurite olulisuse ennustamiseks hulgasõnade valikule. Mudel on kujutatud joonisel 4, mis on klassifitseerimistäpsusega on umbes 90%, mis näitab, et mudel on hea klasside ennustamisvõimega. Mudeli makro-AUC on ligikaudselt 0.985, mis näitab, et mudel eristab tunnuste põhjal klasse hästi.



Joonis 4. Tingimusliku juhumetsa mudel erinevatest skaala *vähe* – ... – *palju* hulgasõnade valikut ennustavatest teguritest.

Sarnaselt eelmise skaala tingimusliku juhumetsa mudelile, näitab ka antud mudel vaid objektide enda proportsioonide olulisust hulgasõnade valiku ennustamisel. Seevastu ükski vastajast tulenev tegur ei ennusta mudeli järgi hulgasõnade valikut. Sellest saan ka vastuse oma teisele uurimisküsimusele skaala *vähe* – ... – *palju* kohta.

Kokkuvõttes võib järeldada, et skaala *vähe* – ... – *palju* hulgasõnade valikul on oluline vaid objektide enda proportsioonid, mis suudavad samuti ennustada hulgasõnade valikut. Ükski teine vastajaspetsiifiline tegur ei osutunud oluliseks ega ennustanud hulgasõnade valikut, mis sarnaneb ka ülaltoodud skaala *mitte ükski* – ... – *kõik* tulemustele.

Kokkuvõte ja arutelu

Eelnevalt esitasin statistilise analüüsi kahe hulgasõnade skaala kohta. Analüüsiga proovisin tuvastada erinevate tegurite mõju hulgasõnade valikule objektide hulga kirjeldamisel. Tulemuste saamiseks kasutasin tingimuslike otsustuspuude ja tingimuslike juhumetsade mudeleid. Skaala *mitte ükski* – ... – *kõik* mõjutegurite analüüsimiseks kasutati 52 vastaja tulemusi (kokku 572 vastust), skaala *vähe* – ... – *palju* korral 56 vastaja tulemusi (kokku 504 vastust). Analüüsiga leidsin vastuseid nii esimesele kui ka teisele uurimisküsimusele.

Mõlema skaala analüüsimisel kahe mudeli abil leidsin, et kahe skaala hulgasõnade valikut ennustavad ja mõjutavad ainult objektide enda proportsioonid. Ükski vastajast tulenev tegur, nagu sugu, haridustase, omandatud keeled, väsimustase ja elukoht, oluliseks ei osutunud. Võin oletada, et vastajaspetsiifiliste tegurite mõju puudumine võiks olla osaliselt tingitud vastuste vähesusest, mille tõttu ei taganud see ka suuremat varieerumist. Seetõttu esineb tõenäosus, et suurema andmestiku korral võib mõni tunnus osutada oluliseks mõjutajaks ja ennustajaks.

Kokkuvõte

Umbmääraseid hulgasõnu kasutatakse erinevatele kogustele viitamisel, kuid nende kasutusest ei ole võimalik oletada nende täpset arvulist kogust. Magistritöös uurisin sunnitud valiku katse abil üheksa erineva hulgasõna valikut objektide proportsiooni kirjeldamisel: *mitte ükski, mõni, mõned, pooled, enamik* ja *kõik* ning *vähe, keskmiselt* ja *palju*. Hulgasõnad olid paigutatud kas formaalloogilisele- või hinnanguskaalale. Samuti vaatasin vastajatest tulenevate tegurite mõju hulgasõnade valikule. Katses osales 56 eesti emakeelega katseisikut vanuses 18 kuni 30 eluaastat. Tulemuste analüüsimiseks kasutasin erinevaid statistilisi meetodeid, nagu kirjeldav analüüs, tingimuslikke otsustuspuid ja tingimuslikke juhumetsi. Katse viidi läbi veebikeskkonnas LimeSurvey.

Esimese analüüsiga vastati uurimisküsimusele, kas ja milliseid eestikeelseid hulgasõnu valitakse objektide hulga erinevate proportsioonide kirjeldamisel visuaalsete stiimulite puhul. Kirjeldava statistika meetoditega tuvastasin üheksa uuritava hulgasõna kasutusulatust ehk jaotust kahel erineval skaalal objektide proportsioonide kirjeldamisel. Katsetulemuste põhjal selgus, et paljude eestikeelsete hulgasõnade kasutusulatused kattuvad omavahel ning nende vahel puuduvad kindlad semantilised piirid. Hulgasõnade jaotus skaala *mitte ükski* – – *kõik* osutus samuti jaotuste poolest ebahürtlaseks. Kõige suurema kasutusulatusega osutused hulgasõnad *mõned* ja *enamik*. Hulgasõna *mitte ükski* valiti ainult objekti 0% proportsiooni kirjeldamise, kuid *kõik* valiti üksikkorraga ka 90% (muidu aga 100%) proportsiooni märkimisel. Skaala *vähe* – ... – *palju* hulgasõnade jaotus oli seevastu võrdne. Vastamisaegade analüüs näitas, et skaala *mitte ükski* – ... – *kõik* puhul esines suurem ajakulu ja vastamiskiiruste hajutavus objekti 40% ja 60% proportsiooni juures ning skaala *vähe* – ... – *palju* puhul ilmnes sama objekti tendents 30% ja 40% proportsiooni kirjeldamisel. Samuti leidsin, et eestikeelsed hulgasõnad võiksid kattuda ka ligikaudse numbrilise süsteemiga, kuid minu tulemuste põhjal on seda järeldust keeruline teha. Lisaks tõin välja katse puudujääke, nagu näiteks puuduvate vastuste esinemine, mis esinesid keskkonnast tulenevate tehniliste eripärade tõttu, kuid ei olnud katseformaadi poolt tegelikkuses ette nähtud. Kokkuvõttes järeldasin, et eestikeelsed hulgasõnad vastavad teatud proportsioonide vahemikele, kuid nende piirid on umbmäärased, välja arvatud hulgasõna *mitte ükski*. Küsimus eestikeelsete hulgasõnade valikust ja jaotumisest skaaladel vajaks edasist uurimist.

Teise analüüsiga otsisin vastust uurimisküsimusele, kas ja kuidas võiksid vastajast tulenevad tegurid mõjutada hulgasõnade valikut küsitud objekti hulga kirjeldamisel.

Tingimuslikud otsustuspuud ja tingimuslikud juhumetsad näitasid, et vastajatest tulenevad tegurid (väsimustase, sugu, haridustase, omandatud keeled ja elukoht) ei avalda mõju hulgasõnade valikule ega ennustanud seda. Oluliseks teguriks kui ka ennustajaks, osutusid objektide proportsioonid ise. Selle tulemusega sain täiendava vastuse oma esimesele uurimisküsimusele.

Siinne magistritöö teema vajaks minu arvates edasi uurimist. Tööga näitasin, et eestikeelsed hulgasõnad vastavad teatud proportsioonide vahemikele, kuid saadud tulemused käivad vaid kindlate proportsioonide kohta. Samuti tuli esile hulgasõnade *mõni* ja *mõned* vahe kasutusulatustes, mida ei olnud võimalik siinse uurimusega täpselt välja selgitada. Lisamist uurimustesse vajaksid ka hulgasõnad *pool* ja *enamus* ning ka *enamikud*. Samuti vajaks lisauurimist vastajaspetsiifiliste tegurite mõju hulgasõnade valikule, kuna siinse magistritöö andmestik jäi selleks liiga väikseks.

Kirjandus

Aasa, T. (2024). *Omadussõnade intensiivistamine eesti noortekeeles kahe korpuse võrdluses* (Magistritöö; Tartu Ülikool; Eesti ja üldkeeleteaduse instituut).

Abbondanza jt 2021 = **Abbondanza, M., Rinaldi, L., Foppolo, F., & Marelli, M. (2021).** The mental representation of nonnumerical quantifiers: The spatial-linguistic association of response codes (SLARC) effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(12), 2021–2028. <https://doi.org/10.1037/xlm0001037>.

Ariel, M. (2004). Most. *Language*, 80(4), 658–706.
<https://doi.org/10.1353/lan.2004.0162>.

Arielli, E. (2012). Contrast and assimilation in aesthetic judgments of visual artworks. *Empirical Studies of the Arts*, 30(1), 59–74. <https://doi.org/10.2190/EM.30.1.F>.

Balakrishnan, J. D., & Ashby, F. G. (1992). Subitizing: Magical numbers or mere superstition? *Psychological Research*, 54(2), 80–90. <https://doi.org/10.1007/BF00937136>.

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. In: J. Kulas, J. H. Fetzer, & T. L. Rankin (Eds.), *Philosophy, language, and artificial intelligence* (Studies in Cognitive Systems, Vol. 2, pp. 241–301). Springer. https://doi.org/10.1007/978-94-009-2727-8_10.

Bass, B. M., Cascio, W. F., & O'Connor, E. J. (1974). Magnitude estimations of expressions of frequency and amount. *Journal of Applied Psychology*, 59(3), 313–320. <https://doi.org/10.1037/h0036653>.

Baumann, O., McFadyen, J., & Humphreys, M. S. (2020). Behavioral and neural effects of familiarization on object-background associations. *Frontiers in Psychology*, 11, Article 591231. <https://doi.org/10.3389/fpsyg.2020.591231>.

Bialystok jt 2009 = **Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009).** Bilingual minds. *Psychological Science in the Public Interest*, 10(3), 89–129. <https://doi.org/10.1177/1529100610387084>.

Boksem, M. A., & Tops, M. (2008). Mental fatigue: costs and benefits. *Brain research reviews*, 59(1), 125–139. <https://doi.org/10.1016/j.brainresrev.2008.07.001>.

Boksem, M. A., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: An ERP study. *Cognitive Brain Research*, 25(1), 107–116.

<https://doi.org/10.1016/j.cogbrainres.2005.04.011>.

Boksem, M. A., Meijman, T. F., & Lorist, M. M. (2006). Mental fatigue, motivation and action monitoring. *Biological psychology*, 72(2), 123–132.

<https://doi.org/10.1016/j.biopsycho.2005.08.007>

Breiman jt 2022 = **Breiman, L., Cutler, A., Liaw, A., Wiener, M., & Liaw, M. A. (2022).** *randomForest: Breiman and Cutler's random forests for classification and regression* [R package]. University of California, Berkeley. <https://cran.r-project.org/package=randomForest>.

Coventry jt 2010 = **Coventry, K. R., Cangelosi, A., Newstead, S. E., & Bugmann, D. (2010).** Talking about quantities in space: Vague quantifiers, context and similarity. *Language and Cognition*, 2(2), 221–241. <https://doi.org/10.1515/langcog.2010.009>.

Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, 2(3), 219–253.

Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, 2(3), 219–253.

Dąbrowska, E. (2019). Individual differences in grammatical knowledge. In E. Dąbrowska & D. Divjak (Eds.), *Cognitive linguistics – key topics* (pp. 231–250). De Gruyter Mouton. <https://doi.org/10.1515/9783110626438-012>.

Daltrozzo, J., Wioland, N., & Kotchoubey, B. (2007). Sex differences in two event-related potentials components related to semantic priming. *Archives of sexual behavior* 36(4), 555–568. <https://doi.org/10.1007/s10508-006-9161-0>.

Dash, N. S. (2008). Context and contextual word meaning. *SKASE Journal of Theoretical Linguistics*, 5(2), 21–31.

de Bot, K. (2004). The Multilingual Lexicon: Modelling Selection and Control. *International Journal of Multilingualism* 1(1), 17–32. <https://doi.org/10.1080/14790710408668176>.

Downing jt 2008 = **Downing, K., Chan, S., Downing, W., Kwong, T., & Lam, T. (2008).** Measuring gender differences in cognitive functioning. *Multicultural Education & Technology Journal*, 2(1), 4–18. <https://doi.org/10.1108/17504970810867124>.

Dubravska, D., & Soroka, H. (2024). Linguistic and extralinguistic factors of creating neologisms in modern English. *Світ наукових досліджень. Випуск 27: матеріали Міжнародної мультидисциплінарної наукової інтернет-конференції (м. Тернопіль, Україна, м. Ополе, Польща, 22-23 лютого 2024 р.)* (ред. О. Патряк та ін.), 116–265. Тернопіль: ФО-П Шпак ВБ.

EKK 2020 = **Erelt, M., Erelt, T., & Ross, K. (2020).** *Eesti keele käsiraamat*. Tallinn: EKSA.

EKS 2017 = **Erelt, M., & Metslang, H. (Eds.). (2017).** *Eesti keele süntaks* (Eesti keele varamu III). Tartu Ülikooli Kirjastus.

EKSS 2009 = **Laanemets, M., Tiits, M., Valdre, T., Veskis, L., Viks, Ü., & Voll, P. (Eds.). (2009).** *Eesti keele seletav sõnaraamat* (2. trükk). Eesti Keele SA. <https://www.eki.ee/dict/ekss/>. (vaadatud 25.05.2025).

Erelt, M. (2017). Sissejuhatus süntaksisse. In M. Erelt & H. Metslang (Eds.), *Eesti keele süntaks* (Eesti keele varamu III) (lk 53–92). Tartu: Tartu Ülikooli Kirjastus.

Erelt, M., & Metslang, H. (1998). Oma või võõras? *Keel ja Kirjandus*, (10), 657–668.

Fritschi, C., & Quinn, L. (2010). Fatigue in patients with diabetes: A review. *Journal of Psychosomatic Research*, 69(1), 33–41.

Fritschi, C., & Quinn, L. (2010). Fatigue in patients with diabetes: A review. *Journal of Psychosomatic Research*, 69(1), 33–41.

Fritschi, C., & Quinn, L. (2010). Fatigue in patients with diabetes: a review. *Journal of psychosomatic research*, 69(1), 33-41.

Garimella, A., Banea, C., & Mihalcea, R. (2023). Reflection of demographic background on word usage. *Computational Linguistics*, 49(2), 373–394. https://doi.org/10.1162/coli_a_00475.

Gillioz, C., & Zufferey, S. (2020). *Introduction to experimental linguistics*. John Wiley & Sons.

Ginzburg, R. (1974). Meaning and the relation between intra-linguistic and extra-linguistic factors in lexical units. *Journal of Literary Semantics*, 3(Jahresband), 91–100. <https://doi.org/10.1515/jlse.1974.3.1.91>.

Grice, H. P. (1989). *Studies in the way of words*. Harvard University Press.

Haahr, M. (2025). *RANDOM.ORG: tõeliste juhuarvude teenus*. Saadaval <https://www.random.org> (vaadatud 27.05.2025).

Hackl, M. (2009). On the grammar and processing of proportional quantifiers: *most* versus *more than half*. *Natural Language Semantics*, 17(1), 63–98. <https://doi.org/10.1007/s11050-008-9039-x>.

Harley, T. A. (2014). *The psychology of language: From data to theory*. Routledge.

Haspelmath, Martin. (2008). Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics*, 19(1), 1–33. <https://doi.org/10.1515/COG.2008.001>.

Hawthorn jt 2015 = **Hothorn, T., Hornik, K., Strobl, C., Zeileis, A., & Hothorn, M. T. (2015).** *party: A laboratory for recursive partytioning* [R package]. <https://cran.r-project.org/package=party>.

Healy jt 2004 = **Healy, A. F., Kole, J. A., Buck-Gengle, C. J., & Bourne, L. E., Jr (2004).** Effects of prolonged work on data entry speed and accuracy. *Journal of experimental psychology. Applied*, 10(3), 188–199. <https://doi.org/10.1037/1076-898X.10.3.188>.

Heim jt 2012 = **Heim, S., Amunts, K., Drai, D., Eickhoff, S. B., Hautvast, S., & Grodzinsky, Y. (2012).** The language–number interface in the brain: A complex parametric study of quantifiers and quantities. *Frontiers in Evolutionary Neuroscience*, 4. <https://doi.org/10.3389/fnevo.2012.00004>.

Hirnstein, M., Coloma Andrews, L., & Hausmann, M. (2014). Gender-stereotyping and cognitive sex differences in mixed- and same-sex groups. *Archives of sexual behavior* 43(8), 1663–1673. <https://doi.org/10.1007/s10508-014-0311-5>.

Holyoak, K. J., & Glass, A. L. (1978). Recognition confusions among quantifiers. *Journal of Verbal Learning and Verbal Behavior*, 17(3), 249–264. [https://doi.org/10.1016/S0022-5371\(78\)90165-2](https://doi.org/10.1016/S0022-5371(78)90165-2).

Hothorn, T., & Zeileis, A. (2015). partykit: A modular toolkit for recursive partytioning in R. *Journal of Machine Learning Research*, 16(1), 3905–3909.

<https://jmlr.org/papers/v16/hothorn15a.html>.

Hothorn, T., Hornik, K., & Zeileis, A. (2015). ctree: Conditional inference trees [R package documentation]. The Comprehensive R Archive Network. <https://cran.r-project.org/web/packages/partykit/index.html>.

Jakobson, R. (1971). Quest for the essence of language. In *Selected writings II: Word and language* (pp. 345–359). The Hague & Paris: Mouton.

Jausovec, N., & Jausovec, K. (2009). Gender related differences in visual and auditory processing of verbal and figural tasks. *Brain research* 1300, 135–145.

<https://doi.org/10.1016/j.brainres.2009.08.093>.

Johnstone, B. (2004). Place, globalization, and linguistic variation. In: C. Fought (Ed.), *Sociolinguistic variation: Critical reflections* (pp. 65–83). Oxford University Press.

<https://doi.org/10.1093/oso/9780195170399.003.0005>.

Kaiser, E. (2013). Experimental paradigms in psycholinguistics. In: R. J. Podesva & D. Sharma (Eds.), *Research methods in linguistics* (pp. 135–168). Cambridge University Press.

Katzir, N., & Ariel, M. (2024). A few or several? Construal, quantity, and argumentativity. *Language and Cognition*, 16(1), 148–175.

<https://doi.org/10.1017/langcog.2023.25>.

Klavan, J., Veismann, A., & Jürine, A. (2013). Katselised meetodid tähenduse uurimisel. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 4(1), 17–34. <https://doi.org/10.12697/jeful.2013.4.1.02>.

Knowlton jt 2021 = **Knowlton, T., Pietroski, P., Halberda, J., & Lidz, J. (2022).** The mental representation of universal quantifiers. *Linguistics and Philosophy*, 45, 911–941.

<https://doi.org/10.1007/s10988-021-09337-8>.

Koppel jt 2023 = **Koppel, K., Kallas, J., Jürviste, M., & Kaljumäe, H. (2023).** *Eesti keele ühendkorpus 2023*. Eesti Keele Instituut.

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind, and brain. *Annual Review of Linguistics*, 1, 377–394. <https://doi.org/10.1146/annurev-linguist-030514-124937>.

Kurzban jt 2013 = **Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013).** An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, 36(6), 661–679. <https://doi.org/10.1017/S0140525X12003196>.

Langacker, R. W. (2006). Cognitive grammar: Introduction to concept, image, and symbol. In: D. Geeraerts, R. Dirven, & J. R. Taylor (Eds.), *Cognitive linguistics: Basic readings* (pp. 29–67). Mouton de Gruyter.

Levinson, S. C. (1981). Some pre-observations on the modelling of dialogue. *Discourse Processes*, 4(2), 93–116.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. MIT Press.

Levshina, N. (2020). Conditional inference trees and random forests. In: M. Paquot & S. T. Gries (Eds.), *A practical handbook of corpus linguistics* (pp. 535–551). Springer, Cham. https://doi.org/10.1007/978-3-030-46216-1_25

Lorist, M. M., Boksem, M. A., & Ridderinkhof, K. R. (2005). Impaired cognitive control and reduced cingulate activity during mental fatigue. *Cognitive Research. Cognitive Brain Research*, 24(2), 199–205. <https://doi.org/10.1016/j.cogbrainres.2005.01.018>.

Lõbu, T. (2011). *Hulgasõnade kasutamisest eesti kirjakeeles* (Magistritöö, Tartu Ülikool, Tartu, Eesti keele õppetool).

Mackenzie C. (2000). Adult spoken discourse: the influences of age and education. *International journal of language & communication disorders*, 35(2), 269–285. <https://doi.org/10.1080/136828200247188>.

Mailund, T. (2022). Pipelines: *magrittr*. In *R 4 Data Science Quick Reference: A Pocket Guide to APIs, Libraries, and Packages* (pp. 77–87). Berkeley, CA: Apress.

MATLAB = **The MathWorks Inc. (2022).** MATLAB version: 9.13.0 (R2022b), Natick, Massachusetts: The MathWorks Inc. <https://www.mathworks.com>.

McKenna, F. P. (1984). Assimilation and contrast in perceptual judgments. *Quarterly Journal of Experimental Psychology, Section A*, 36 (3), 531–548.
doi:10.1080/14640748408402176

Metslang, H. (2017). Kvantorifraas. In M. Erelt & H. Metslang (Eds.), *Eesti keele süntaks (Eesti keele varamu III)* (lk 463–480). Tartu: Tartu Ülikooli Kirjastus.

Moll, R., Lücke, A. J., & Bromme, R. (2021). How much are “many people” on Facebook? Interpretations of vague quantifiers in online and offline contexts. *SAGE Open*, 11(3). <https://doi.org/10.1177/21582440211032205>.

Moncrieff, G., & Fletcher, J. (2007). Tiredness. *BMJ (Clinical research ed.)*, 334(7605), 1221. <https://doi.org/10.1136/bmj.39182.615405.94>.

Moxey, L. M. (1986). *Attentional focusing with quantifiers in production and comprehension* (Unpublished doctoral dissertation). University of Glasgow.

Moxey, L. M. (2011). Mechanisms underlying linguistic framing effects. In: G. Keren (Ed.), *Perspectives on framing* (pp. 119–134). Psychology Press.

Moxey, L. M., & Sanford, A. J. (1993). *Communicating quantities*. Lawrence Erlbaum Associates.

Moxey, L. M., & Sanford, A. J. (2000). Communicating quantities: A review of psycholinguistic evidence of how expressions determine perspectives. *Applied Cognitive Psychology*, 14(3), 237–255. [https://doi.org/10.1002/\(SICI\)1099-0720\(200005/06\)14:3<237::AID-ACP641>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1099-0720(200005/06)14:3<237::AID-ACP641>3.0.CO;2-R).

Mulder, K., & Hulstijn, J. H. (2011). Linguistic skills of adult native speakers, as a function of age and level of education. *Applied Linguistics*, 32(5), 475–494. <https://doi.org/10.1093/applin/amr016>.

Newstead, S. E., & Collis, J. M. (1987). Context and the interpretation of quantifiers of frequency. *Ergonomics*, 30(10), 1447–1462. <https://doi.org/10.1080/00140138708966038>.

Newstead, S. E., & Coventry, K. R. (2000). The role of context and functionality in the interpretation of quantifiers. *European Journal of Cognitive Psychology*, 12(2), 243–259. <https://doi.org/10.1080/095414400382145>.

Newstead, S. E., Pollard, P., & Riezebos, D. (1987). The effect of set size on the interpretation of quantifiers used in rating scales. *Applied Ergonomics*, 18(3), 178–182. [https://doi.org/10.1016/0003-6870\(87\)90001-9](https://doi.org/10.1016/0003-6870(87)90001-9).

Oja, V. (2014). Tähendussuhetest eesti murretes. *Keel ja Kirjandus*, 57(8–9), 684–699.

Olm jt 2014 = **Olm, C. A., McMillan, C. T., Spotorno, N., Clark, R., & Grossman, M. (2014).** The relative contributions of frontal and parietal cortex for generalized quantifier comprehension. *Frontiers in Human Neuroscience*, 8.

<https://doi.org/10.3389/fnhum.2014.00610>.

Pajusalu, R. (2008). Pragmatics of quantifiers: The case of Estonian kõik 'all'. *Journal of Pragmatics*, 40(11), 1950–1965. <https://doi.org/10.1016/j.pragma.2008.04.006>.

Pajusalu, R. (2017). Nimisõnafraas. In M. Ereht & H. Metslang (Eds.), *Eesti keele süntaks (Eesti keele varamu III)* (lk 463–480). Tartu: Tartu Ülikooli Kirjastus.

Pajusalu, R., Tragel, I., Veismann, A., & Vija, M. (2004). *Tuumsõnade semantikat ja pragmaatikat*. Tartu: Tartu Ülikooli Kirjastus.

Paldre, L. (1997). Semantics of Estonian universal quantifiers. *Estonian: Typological Studies II. Tartu Ülikooli eesti keele õppetooli toimetised*, 8, 178–198.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics–pragmatics interface. *Cognition*, 86(3), 253–282. [https://doi.org/10.1016/S0010-0277\(02\)00179-8](https://doi.org/10.1016/S0010-0277(02)00179-8).

Parsons, T. (2021). *The traditional square of opposition*. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition). Stanford University. <https://plato.stanford.edu/archives/fall2021/entries/square/>. (Vaadatud 25.05.2025).

Pepper, S., & Prytulak, L. S. (1974). Sometimes frequently means seldom: Context effects in the interpretation of quantitative expressions. *Journal of Research in Personality*, 8(1), 95–101. [https://doi.org/10.1016/0092-6566\(74\)90049-X](https://doi.org/10.1016/0092-6566(74)90049-X).

Pezzelle, S., Bernardi, R., & Piazza, M. (2018). Probing the mental representation of quantifiers. *Cognition*, 181, 117–126. <https://doi.org/10.1016/j.cognition.2018.08.009>.

Peters, S., & Westerståhl, D. (2006). *Taking exception to quantified statements*. Paper presented at the 16th Annual Conference on Semantics and Linguistic Theory (SALT 16), Tokyo University of Foreign Studies, Tokyo, Japan.

Pietroski 2009 = **Pietroski, P., Lidz, J., Hunter, T., & Halberda, J. (2009).** The meaning of 'most': Semantics, numerosity and psychology. *Mind and Language*, 24(5), 554–585. <https://doi.org/10.1111/j.1468-0017.2009.01374.x>.

Pilvik jt 2025 = **Pilvik, M. L., Lindström, L., Plado, H., & Simmul, C. E. (2025).** Nimisõnafraasi ja hulgafrasib piirail: „osa“, „enamik“ ja „enamus“ hulgasõnadena. *Eesti Rakenduslingvistika Ühingu aastaraamat*, 21, 237–262.

Pilvik, M.-L. (ilmumas). Korpusandmete statistiline analüüs. Teoses L. Lindström, A. Veismann, M.-L. Pilvik, K. Muischnek & J. Padrik (toim.), *Korpused lingvistilises uurimistöös. Praktiline käsiraamat*. Tartu: Tartu Ülikooli Kirjastus.

Pinker, S. (1995). *The language instinct: The new science of language and mind*. HarperCollins Publishers.

Pohl, N. F. (1981). Scale considerations in using vague quantifiers. *The Journal of Experimental Education*, 49(4), 235–240.

Proos, M., & Reile, M. (2022). *Vague quantifiers in Estonian: Evidence from a picture choice task*. The 8th Conference of the Scandinavian Association for Language and Cognition, Åbo Akademi University.

R Core Team. (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>. (Vaadatud 26.05.2025).

Reed, P. E. (2020). Place and language: Links between speech, region, and connection to place. *Wiley Interdisciplinary Reviews: Cognitive Science*, 11(3), e1524. <https://doi.org/10.1002/wcs.1524>.

Reile, M. & Šepilenko, A. (2024). *What does harva ‘rarely’ really mean? The effects of context and social expectations for the agent on the selection of Estonian vague frequency words*. The Scandinavian Association for Language and Cognition (SALC 9), Norwegian University of Science and Technology.

Requejo, M. D. P. (2007). The role of context in word meaning construction: A case study. *International Journal of English Studies*, 7(1), 169–173.

Reyna, V. F. (1981). The language of possibility and probability: Effects of negation on meaning. *Memory and Cognition*, 9, 642–650. <https://doi.org/10.3758/BF03202359>.

Robin jt 2021 = **Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., Müller, M., Siegert, S., Doering, M., Billings, Z., & Robin, M. X. (2021).** *pROC: Display and analyze ROC curves* [R package]. <https://cran.r-project.org/package=pROC>.

Routh, D. A. (1994). On the representation of quantifiers. *Journal of Semantics*, 11, 199–214.

Schneider, S., & Stone, A. A. (2016). The meaning of vaguely quantified frequency response options on a quality of life scale depends on respondents' medical status and age. *Quality of Life Research*, 25(10), 2511–2521. <https://doi.org/10.1007/s11136-016-1293-7>.

Smessaert, H., & Demey, L. (2023). Aristotelian diagrams for the proportional quantifier ‘most’. *Axioms*, 12(3), 236. <https://doi.org/10.3390/axioms12030236>.

Stateva, P., Stepanov, A., Déprez, V., Dupuy, L. E., & Reboul, A. C. (2019). Cross-linguistic variation in the meaning of quantifiers: Implications for pragmatic enrichment. *Frontiers in Psychology*, 10, 957. <https://doi.org/10.3389/fpsyg.2019.00957>.

Strobl, C., Hothorn, T., & Zeileis, A. (2009). Conditional variable importance for random forests. *The R Journal*, 1(2), 1–9. <https://digitalcommons.unl.edu/r-journal/328/>. (Vaadatud 26.05.2025).

Traxler, M. J. (2023). *Introduction to psycholinguistics: Understanding language science*. John Wiley & Sons.

Treffers-Daller, J. (2013). Measuring lexical diversity among L2 learners of French: An exploration of the validity of D, MTL-D and HD-D as measures of language ability. In S. Jarvis & M. Daller (Eds.), *Vocabulary knowledge: Human ratings and automated measures* (pp. 79–104). Benjamins. <https://centaur.reading.ac.uk/28712/>.

Truus, M. (2015). *Hulgasõnade kõik ja mõned mõistmine eesti lapsekeeles pildivalikukatse põhjal* (Bakalaureusetöö, Tartu Ülikool, Eesti keele õppetool).

Tsouhlaris, Z. (2024). Quantifier none and its relation to the concept of zero. *Open Journal of Modern Linguistics*, 14(1), 95–115. <https://doi.org/10.4236/ojml.2024.141006>.

Ulanska jt 2020 = **Ulanska, T., Kuzmanovska, D., Todorova, M., & Kirova, S. (2020).** Extra linguistic factors behind lexical change in Modern English, Spanish and German. In *INTED2020 Proceedings* (pp. 2307–2311). IATED.

Vaba, L. (2014). Levik kui etümoloogiline kriteerium (eeskätt eesti sõnavara näitel). *Keel ja Kirjandus*, 57(8–9), 700–707.

Van Cutsem, J., Van Schuerbeek, P., Pattyn, N., Raeymaekers, H., De Mey, J., Meeusen, R., & Roelands, B. (2022). A drop in cognitive performance, whodunit? Subjective mental fatigue, brain deactivation or increased parasympathetic activity? It's complicated!. *Cortex; a journal devoted to the study of the nervous system and behavior*, 155, 30–45. <https://doi.org/10.1016/j.cortex.2022.06.006>.

Van der Elst, W., van Boxtel, M. P., van Breukelen, G. J., & Jolles, J. (2005). Rey's verbal learning test: normative data for 1855 healthy participants aged 24-81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society : JINS*, 11(3), 290–302. <https://doi.org/10.1017/S1355617705050344>.

Van der Linden, D., & Eling, P. (2006). Mental fatigue disturbs local processing more than global processing. *Psychological research*, 70(5), 395–402. <https://doi.org/10.1007/s00426-005-0228-7>.

Wardhaugh, R., & Fuller, J. M. (2021). *An introduction to sociolinguistics*. John Wiley & Sons.

Wegesin, D. J. (1998). Event-related potentials in homosexual and heterosexual men and women: Sex-dimorphic patterns in verbal asymmetries and mental rotation. *Brain and Cognition*, 36(1), 73–92. <https://doi.org/10.1006/brcg.1997.0964>.

Veismann, A., Ereht, E., & Metslang, H. (2017). Määrus. In M. Ereht & H. Metslang (Eds.), *Eesti keele süntaks (Eesti keele varamu III)* (lk 300–378). Tartu: Tartu Ülikooli Kirjastus.

Weiss jt 2003 = Weiss, E. M., Kemmler, G., Deisenhammer, E. A., Fleischhacker, W. W., & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and Individual Differences*, 35(4), 863–875. [https://doi.org/10.1016/S0191-8869\(02\)00288-X](https://doi.org/10.1016/S0191-8869(02)00288-X).

Westerståhl, D. (1985). Logical constants in quantifier languages. *Linguistics and Philosophy*, 8(4), 387–413. <https://doi.org/10.1007/BF00637410>.

Westerståhl, D. (2012). Classical vs. modern squares of opposition, and beyond. In: J.-Y. Béziau & D. Jacquette (Eds.), *The square of opposition: A general framework for cognition* (pp. 195–229). Peter Lang.

Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*, 3(2), 180–185. <https://doi.org/10.1002/wics.147>.

Winter, B. (2020). *Statistics for linguists: An introduction using R*. Routledge.

Wirth jt 2007 = **Wirth, M., Horn, H., Koenig, T., Stein, M., Federspiel, A., Meier, B., Michel, C. M., & Strik, W. (2007).** Sex differences in semantic processing: Event-related brain potentials distinguish between lower and higher order semantic analysis during word reading. *Cerebral Cortex*, 17(9), 1987–1997. <https://doi.org/10.1093/cercor/bhl121>.

Vorweg, C. (2012). Experimental methods in psycholinguistics. In A. Ender, A. Leemann, & B. Wälchli (Eds.), *Methods in contemporary linguistics* (Trends in Linguistics, Studies and Monographs, Vol. 247, pp. 363–386). De Gruyter Mouton.

Yildirim jt 2013 = **Yildirim, I., Degen, J., Tanenhaus, M., & Jaeger, F. (2013).** Linguistic Variability and Adaptation in Quantifier Meanings. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 35, 3835–3840. <https://escholarship.org/uc/item/7h5003md>.

Yildirim jt 2016 = **Yildirim, I., Degen, J., Tanenhaus, M. K., & Jaeger, T. F. (2016).** Talker-specificity and adaptation in quantifier interpretation. *Journal of Memory and Language* 87, 128–143. <https://doi.org/10.1016/j.jml.2015.08.003>.

On the selection of Estonian vague quantifiers in proportion description.

Summary

The aim of the current master's thesis was to investigate selections of different Estonian vague quantifiers in the description of objects' proportions. In this paper, I decided to look into two possible scales of quantifiers: formal-logical *mitte ükski* ('none'), *mõni* (singular 'some'), *mõned* (plural 'some'), *pooled* (plural 'half'), *enamik* ('most'), and *kõik* ('all'), and evaluative *vähe* ('few'), *keskmiselt* ('on average') and *palju* ('many/lots').

To date, little is known about the selection (as well as comprehension) of vague quantifiers in Estonian. Different linguists have focused on various aspects of quantifiers, such as formal semantic approach (Paldre 1997), historical usage and changes (Lõbu 2011), or their use in everyday conversations (Pajusalu jt 2004; 2008). However, not much is known about how Estonian quantifiers are chosen when they are used to describe objects' proportions, as it is known to be so by other Indo-European quantifiers (e.g. Moxey 1986, Pezzelle, Bernardi & Piazza 2018, Stateva 2019 jt). Therefore, my main research question is whether and which quantifiers are chosen when describing quantities of objects by visual stimuli.

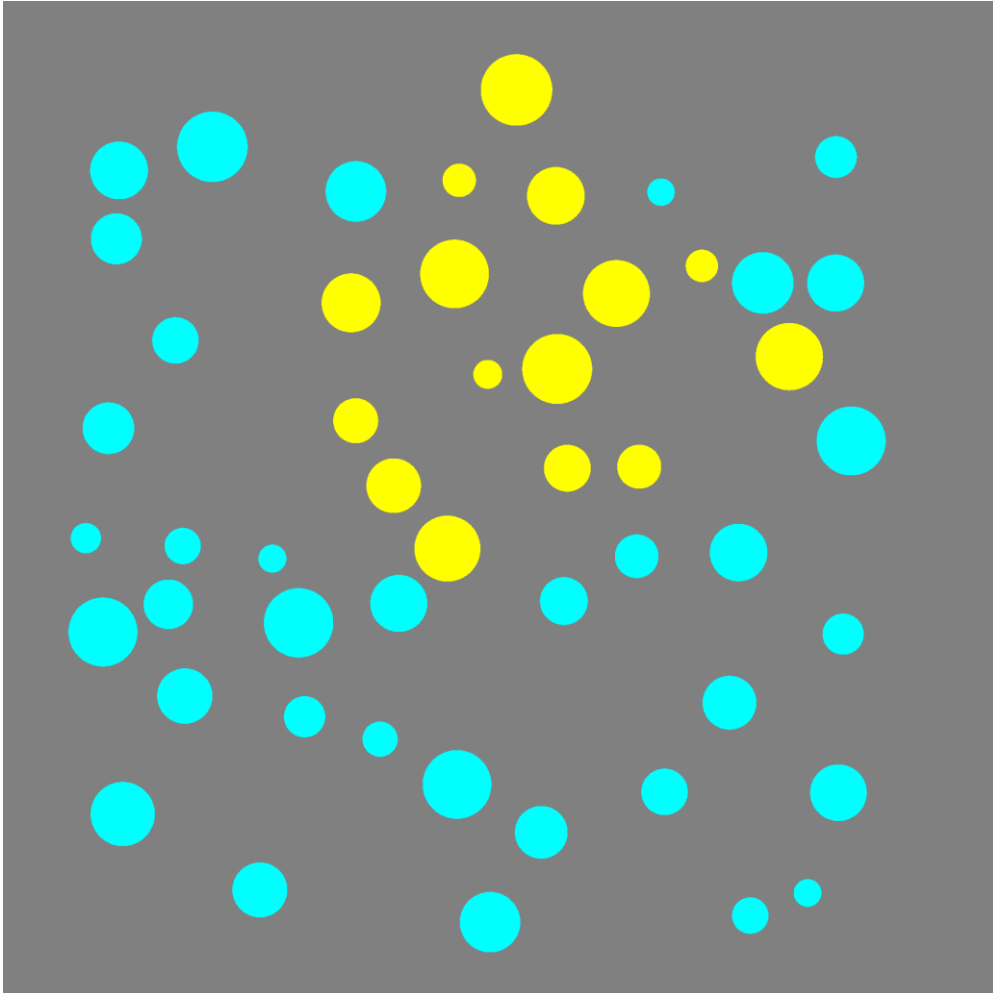
Many studies have also concluded the connection between quantifiers and different types of contexts in Indo-European languages (Newstead, Pollard & Riezebos 1987; Newstead & Coventry 2000) that may also occur by Estonian quantifiers. As far as I know, there is not much research on the influence of individual factors on a quantifier's selection in Estonian (e.g. Reile & Šepilenko 2024). Therefore, my second research question is whether and to what extent can participants' individual factors influence the selection of Estonian quantifiers in visual stimuli.

To meet my goals, I used a forced-choice preference task, in which the participants were asked to select the most suitable quantifier from a given set that best described the quantity of the target object, here yellow and blue balls. The experiment was conducted on a digital platform LimeSurvey. Participants' reaction time was also tracked during the experiment as an indirect measure of semantic processing. All the participants were Estonian native speakers between the age 18 and 30. For data analysis, I used different statistical methods as descriptive statistics, conditional interference trees, and conditional random forests. By conducting my experiment, I hoped to find answers to followed up questions.

The results showed that vague quantifiers in Estonian have some specific ranges of usage in describing objects' proportions, but their boundaries remain vague and unclear. Nearly all

the quantifiers overlapped in their range of usage, except for the quantifier *mitte ükski* that was only selected to mark a 0% proportion of the object. The difference between quantifiers *mõni* and *mõned* was also found. Both quantifiers occurred nearly in the same range from 10% to 40% by *mõni* or 50% by *mõned*, but *mõned* itself had a wider scope and was mainly chosen between 20% to 40%. *Mõni* was more chosen in describing 10%, thus narrowly. The quantifier *pooled* was found to be chosen not only by describing 50% proportion but was chosen in a range from 30% to 70%. Regarding the quantifier *enamik*, it was mainly selected in a range from 60% to 90%, once also by describing objects 100% proportion. The quantifier *kõik* was selected to describe 100%, but once it was chosen by 90%. In terms of the evaluative scale, quantifier *vähe* was chosen in a range 10% to 50%, *keskmiselt* from 30% to 70%, and *palju* from 50% to 90%. As demonstrated, the quantifiers of evaluative scale seem to have a more even distribution and order than those on the formal scale. It might be suggested that Estonian quantifiers can overlap with a so-called approximate number system that was previously suggested by Indo-European quantifiers (see Coventry jt 2010; Pietroski jt 2009; Olm jt 2014). The reaction time indicated the increase in time around the scale's middle point that may indicate some difficulties in quantifiers' selection. Both conditional interference trees and conditional random forest showed no significant influence of participants' varieties on the selection of Estonian quantifiers. Significance was found only in the proportions of the objects described.

Lisa 1. Näidis katses kasutatud pilt-stiimulitest



Pilt 1. Katses kasutatud pilt-stiimul kollaste testobjektidega proportsioonis 30%.

Pildil on esitatud katses kasutatud kollaste testobjektidega stiimul, mille hulga kirjeldamiseks paluti katses osalejatel valida üks sobilik hulgasõna skaalalt *mitte ükski* – ... – *kõik*. Kollased testobjektid esinevad proportsioonis 30%. Sinised pallid on siinkohal taustobjektide rollis, mille suhtes kollasteid testobjekte hinnatakse.

Lisa 2. Katsesse sissejuhatav tekst

Tere!

Oled jõudnud katsesse, millega uuritakse sõnade tõlgendamist eesti emakeelega inimeste seas.

Katse koosneb neljast osast, mille vahel saad soovi korral teha pausi. Sinu peamine ülesanne on esitatud piltidel hinnata pallide hulka ning viia kokku sõnaga, mis sinu meelest vastab küsitud pallide hulgale kõige paremini. Täpsemad juhised tulevad enne iga katseosa algust.

Siinses uuringus ei ole õiget ega valet vastust!

Katse tegemine võtab kuni 15 minutit. Selle läbimiseks on vajalik arvuti. Hiire kasutamine on soovituslik.

Katse viiakse läbi magistritöö raames.

Edasiliikumiseks vajutage nuppu "Järgmine".

Lisa 3. Nõusoleku vorm katses osalemiseks

Hea katsest huvitatud

Katse eesmärk on koguda emakeelse keelekasutaja tõlgendusi eestikeelsete sõnade kohta, mis võimaldab uurida keele ja taju vahelisi seoseid.

Katses palutakse sul erinevate piltide põhjal valida sõna, mis kirjeldab pildil esitatud kindla värvusega pallide hulka kõige paremini. Katse koosneb kahest osast ning kokku on 31 küsimust. Katse tegemiseks kulub kuni 15 minutit. Katse viiakse läbi veebiplatvormil LimeSurvey.

Katsega ei hinnata sinu oskusi ega teadmisi, vaid kogutakse andmeid esitatud sõnade tõlgenduste kohta. Seetõttu ei esine siinkohal "õiget" ega "valet" vastusevarianti.

Katsesse loodame kaasata kuni 40 inimest.

Katses osalemine on **anonüümne**, st et me ei kogu IP aadresse ega muid selliseid andmeid, mille põhjal oleks võimalik sinu isikut tuvastada. Kogutud andmeid kasutatakse ainult eesti keele uurimise eesmärgil. Katse jooksul järgitakse teaduseetika norme ning konfidentsiaalsusprintsipi.

Katseandmed salvestatakse Tartu Ülikooli serverisse, ning säilitatakse tähtajatult edasiseks uurimistööks.

Katses osalemine on **vabatahtlik** – sul ei ole kohustust katsest osa võtta. Samuti on sul õigus soovi korral katse ükskõik mis ajahetkel pooleli jätta.

Katse läbiviimise eest vastutab Tartu Ülikooli eesti ja üldkeeleteaduse instituut. Lisainformatsiooni saamiseks pöördu Anastasia Šepilenko (anastasia.sepilenko@ut.ee) poole. Kui sul tekib küsimusi seoses uuritava õigustega, siis palume sul pöörduda Tartu Ülikooli inimuuringute eetikakomitee poole (etikakomitee@ut.ee). Kaebustega isikuandmete

töötlemise osas pöördu palun Tartu Ülikooli andmekaitespetsialisti poole e-posti aadressil andmekaitse@ut.ee.

- Palun tee märge allpool olevasse nõusolekukasti, et kinnitada oma
- informeeritust ülalmainitud uuringust ning
 - teadlikkust läbiviidava uurimistöö eesmärgist ja metoodikast ning
 - nõusolekut uuringus osaleda.

Katsesse edasi liikumiseks palun kinnita, et sind on uuringust informeeritud ning oled andnud oma nõusoleku uuringus osalemiseks.

Kinnitan, et mind on informeeritud läbiviidavast uuringust ja annan oma nõusoleku uuringus osalemiseks. (märge: raadionupuna)

Lisa 4. Taustainformatsiooni küsimused

1. Mis on sinu vanus?

Vastuse tüüp: numbriline väärtus

2. Mis on sinu sugu?

Vastuse tüüp: ainult üks korraga

Vastuse variandid: Naine, Mees, Muu (kommenteerimisvõimalusega)

3. Mis on sinu emakeel?

Vastuse tüüp: ainult üks korraga

Vastuse variandid: Eesti, Vene, Muu (kommenteerimisvõimalusega)

4. Palun loetle keeli, mida lisaks eesti keelele oskad või oled elu jooksul õppinud.

Vastuse tüüp: vaba loetelu

5. Mis on sinu domineeriv keel? (abistav selgitus: Domineeriv keel on keel, mida praegusel eluperioodil kõige rohkem kasutad ning milles ennast kõige vabamalt väljendada saad.)

Vastuse tüüp: Ainult üks korraga

Vastuse variandid: Eesti, Vene, Inglise, Muu (kommenteerimisvõimalusega)

6. Mis on sinu haridustase? (abistav selgitus: Palun vali oma kõrgeim lõpetatud haridustase.)

Vastuse tüüp: ainult üks korraga

Vastuse variandid: Põhiharidus, Põhiharidus koos kutseharidusega, Keskkharidus, Keskeri- või kutseharidus, Rakendus- või kutsekõrgharidus, Kõrghariduse põhi- ehk bakalaureusekraad, Kõrghariduse kraadiõpe (magistri- või/ja doktorikraad)

7. Mis on sinu püsiv elukoht (maakondliku täpsusega)?

Vastuse tüüp: ainult üks korraga

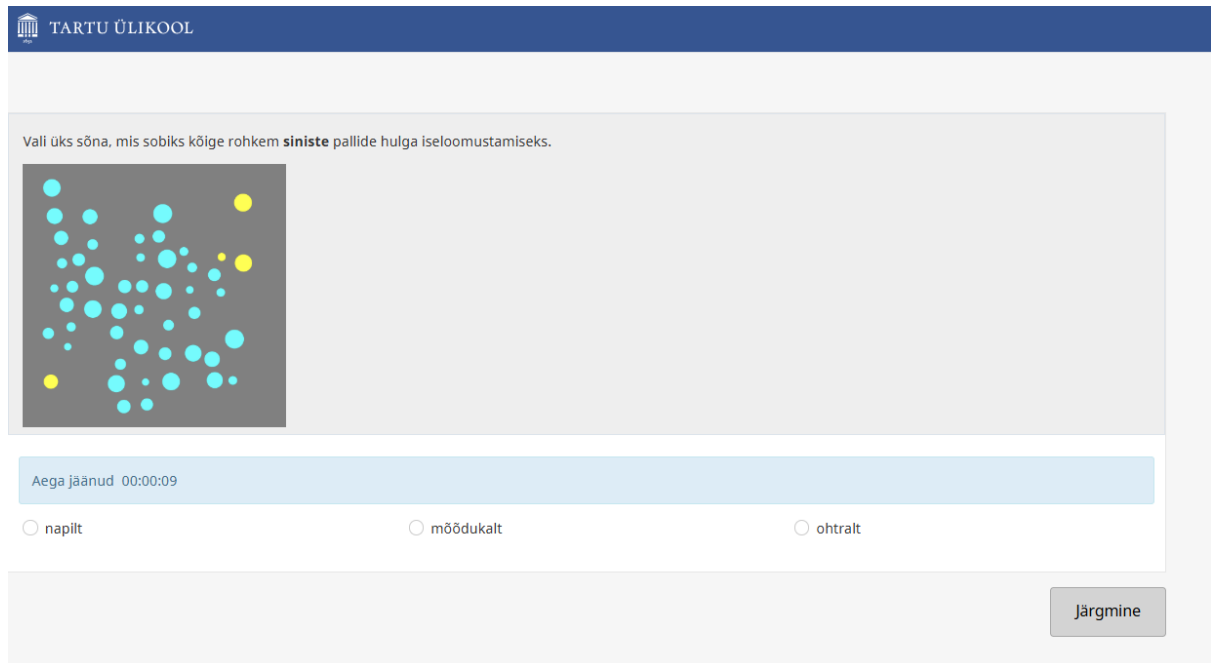
Vastuse variandid: Harju maakond, Tartu maakond, Muu (kommenteerimisvõimalusega)

8. Palun hinda oma väsimustaset praegusel hetkel?

Vastuse tüüp: ainult üks korraga

Vastuse variandid: Väga väsinud, Väsinud, Pigem väsinud, Keskmine (ei väsinud ega ergas),
Pigem ergas, Ergas, Väga ergas

Lisa 5. Näidis harjutusvooru stiimulist koos harjutusskaalaga



TARTU ÜLIKOOL

Vali üks sõna, mis sobiks kõige rohkem **siniste** pallide hulga iseloomustamiseks.

Aega jäänud 00:00:09

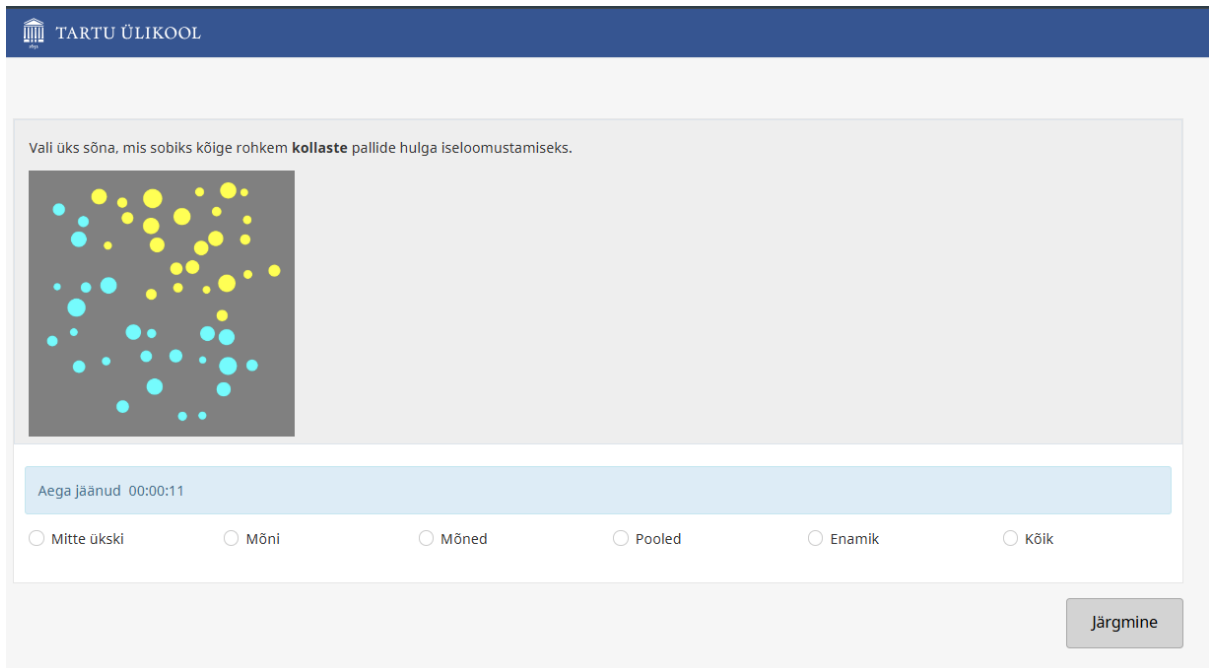
☐ napilt ☐ mõõdukalt ☐ ohtralt

Järgmine

Pilt 2. Kuvatõmmis LimeSurvey keskkonnast harjutusvooru stiimulist siniste testobjektidega proportsioonis 92% (26.05.2025).

Pildil on esitatud kuvatõmmis ühest harjutusvooru stiimulist koos harjutusskaalaga LimeSurvey'i keskkonnast. Siinsel pilt-stiimulil on testobjektiks sinised pallid proportsioonis 92%, mille kohta paluti katseisikutel valida üks sobiv hulgasõna nende koguse iseloomustamiseks. Pilt-stiimuli kohal oli kuvatud ka küsimus pilt-stiimuli kohta ning testobjektide värv oli tehtud mustaks. Pilt-stiimuli oli toodud harjutusskaala *napilt – mõõdukalt – ohtralt*, mille kohal kuvati ka vastamise ajalisi limiite. Nurgas paiknes nupp „Järgmine“, millega sai liikuda edasi järgmise stiimuli juurde.

Lisa 6. Näidis katsevoorude stiimulist koos uuritava skaalaga



TARTU ÜLIKOOL

Vali üks sõna, mis sobiks kõige rohkem kollaste pallide hulga iseloomustamiseks.

Aega jäänud 00:00:11

☐ Mitte ükski ☐ Mõni ☐ Mõned ☐ Pooled ☐ Enamik ☐ Kõik

Järgmine

Pilt 3. Kuvatõmmis LimeSurvey keskkonnast harjutusvooru stiimulist kollaste testobjektidega proportsioonis 50% (26.05.2025).

Sarnaselt harjutusvoorule (vt Lisa 2), on kuvatõmmisel toodud pilt-stiimul koos küsimuse ja uuritava skaalaga LimeSurvey'i keskkonnast. Pilt-stiimulil on testobjektiks kollaste pallide hulk, mis esineb proportsioonis 50%, mille iseloomustamiseks paluti katseisikul valida üks sobiv hulgasõna skaalalt *mitte ükski* – ... – *kõik*. Pilt-stiimuli kohal on esitatud küsimus pilt-stiimuli kohta, kus testobjekti värv on esitatud mustas. Pilt-stiimuli all on esitatud uuritava formaalskaala hulgasõnad, mille kohal on esitatud ka vastamisaja taimer. Paremal nurgas paikneb nupp „Järgmine“, millega sai liikuda edasi järgmise stiimuli juurde.

Lisa 7. Juhised enne harjutusvooruga alustamist

Järgnevalt kuvatakse sulle 3 erinevat pilti, millel on kujutatud kollased ja/või sinised pallid. Iga pildi kohta on esitatud küsimus. Pildi all on valik sõnades.

Sinu **ülesanne** on ekraanil kuvatud pildi põhjal valida esitatud valimist **üks sõna**, mis annab sinu hinnangul **kõige sobilikuma vastuse** pildi kohal esitatud küsimusele.

Enne vastamist loe palun küsimus hoolikalt läbi.

Valikute tegemiseks on antud **10 sekundit**. Ajalimiidi täitmisel liigutakse automaatselt järgmise küsimuse juurde.

Tähelepanu: tegemist on harjutuskorraga.

Katsega alustamiseks vajuta nuppu “Järgmine”.

Lisa 8. Juhised enne esimese katsevooruga alustamist

Järgnevalt kuvatakse sulle 9 või 11 erinevat pilti, millel on kujutatud kollased ja/või sinised pallid. Iga pildi kohta on esitatud küsimus. Pildi all on valik sõnades.

Sinu **ülesanne** on ekraanil kuvatud pildi põhjal valida esitatud valimist **üks sõna**, mis annab sinu hinnangul **kõige sobilikuma vastuse** pildi kohal esitatud küsimusele.

Enne vastamist loe palun küsimus hoolikalt läbi.

Valikute tegemiseks on antud kas **10** või **12 sekundit**. Ajalimiidi täitmisel liigutakse automaatselt järgmise küsimuse juurde.

Katsega alustamiseks vajuta nuppu “Järgmine”.

Lisa 9. Juhised enne teise katsevooruga alustamist

Sarnaselt eelmisele katsevoorule, kuvatakse sulle taas 9 või 11 erinevat pilti, millel on kujutatud kollased ja/või sinised pallid. Iga pildi kohta on esitatud küsimus. Küsimuse all on toodud valik hulgasõnu.

Sinu **ülesanne** on ekraanil kuvatud pildi põhjal valida esitatud valimist **üks sõna**, mis annab sinu hinnangul **kõige sobilikuma vastuse** pildi kohal esitatud küsimusele.

Enne vastamist loe palun küsimus hoolikalt läbi.

Lisa 10. Värbamiskutse

Kutse katsele: Tule ja osale sõnade tõlgendamise kates!

Otsin oma magistritöö jaoks huvilisi keelekatsesse, milles uurin sõnade tõlgendamist eesti keelt emakeelena kõnelejate seas. Katses palun sul omavahel kokku viia pildi ja sõna, mis sellele pildile sinu meelest kõige paremini vastab. Katse tulemusi kasutan magistritöö kirjutamise eesmärgil, nii aitad sa kaasa minu magistritöö valmimisele.

Katset **saad teha** veebiplatvormil **LimeSurvey**. Eelnevalt palun sul samas keskkonnas ära täita taustaküsimustiku. Katse kestus on kuni 15 minutit.

Katses saad osaleda, kui:

- oled vanuses 18–30 aastat,
- räägid eesti keelt emakeelena,
- sul pole diagnoositud värvipimedust.

Katses osalemine on **anonüümne**. Andmete analüüsitulemused avalikustatakse magistritöös.

Katses osalemiseks sisene <https://survey.ut.ee/index.php/653423?lang=et> või kasuta QR koodi. Katse läbimiseks on vajalik arvuti. Hiire kasutamine on soovituslik.



Lisainfo jaoks võta ühendust:

Anastasia Šepilenko, Tartu Ülikool

anastasia.sepilenko@ut.ee

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Anastasia Šepilenko,

1) annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose „Umbmääraste hulgasõnade valikud proportsioonide kirjeldamisel eesti keeles“, mille juhendajad on Maria Reile ja Renate Pajusalu, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada Tartu Ülikooli digitaalarhiivi kuni autoriõiguse kehtivuse lõppemiseni;

2) annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi kaudu Creative Commons'i litsentsiga CC BY NC ND 4.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni;

3) olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile;

4) kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Allkirjastatud digitaalselt

27.05.2025