

EESTI KEELE KUI TEISE KEELE ÕPPIJATE HINNANGUD SÕNASELETUSTELE KEELEÕPPIJA SÕNAVEEBIS

KRISTINA KOPPEL, LIINA LUTSEPP,
MARIA TUULIK

Annotatsioon. Artiklis anname ülevaate küsitlusest, mille korraldasime 2025. aasta lõpus 55 eesti keele kui teise keele õppija seas. Küsitluse eesmärk oli välja selgitada, millised leksikaalsed, morfoloogilised ja süntaktilised nähtused raskendavad keeleõppijal sõnaseletuste mõistmist. Selleks pidid vastajad hindama Keeleõppija Sõnaveebi 20 märksõna kaht erinevalt sõnastatud seletust: üht originaalsõnastuses, teist kohandatud sõnastuses. Leksikaalsete parameetrite analüüs näitas, et õppijad eelistavad kompaktsed sõnastusi, kus on välditud sidesõnu, sisutihje asesõnu ja abstraktseid täpsustusi. Morfoloogiliste parameetrite analüüs näitas, et umbisikulise tegumoe kasutus ei raskenda arusaamist, aga *des*-vormid osutuvad keerukaks. Süntaktiliste parameetrite analüüsist ilmnnes, et arusaamist soodustavad seletused, kus põhisõna on esitatud seletuse alguses ja kus info on piisavalt asjakohane, samas kui täiendite rohkus enne põhisõna võib mõistmist takistada. Küsitluse tulemuste põhjal kohandasime juba olemasolevaid sõnaseletusi Keeleõppija Sõnaveebis ning uuendasime koostamis põhimõtteid, mille alusel saab luua seletusi uutele sõnadele.¹

Võtmesõnad: keeleõpe, õppeleksikograafia, seletuste kohandamine, keeleoskustasemed, eesti keel teise keelena

¹ Artikkel on valminud „Eesti keele õppe ja keeleõppe arendamise“ programmi raames, mida kaasrahastab Haridus- ja Teadusministeerium ning Euroopa Liit. Artikli valmimist on toetanud ka teadusprojektid EKKD-III1 „Suurte keelemudelite rakendamine leksikograafias: uued võimalused ja väljakutsed“ ja PRG1978 „Uue aja sõnastik: grammatika ja keelepädevuse kirjeldamine integreeritud multifunktsionaalses leksikograafilises ressursis“.

1. Sissejuhatus

Eesti Keele Instituudis (EKI) on alates 2019. aastast koostatud EKI ühendsõnastikku (ÜS) – sõnakogu, kuhu on koondatud info mitmest EKI tänapäeva eesti keele sõnastikust (Langemets jt 2021). Need rikkalikud keeleandmed on kasutajatele kättesaadavad keeleportaalis Sõnaveeb (Koppel jt 2019), mille arendamisel lähtuti põhimõttest, et sõnastiku kasutaja ei peaks näiteks sõnade tähendusi otsima ühest, naabersõnade infot teisest ning võõrkeelseid vasteid kolmandast kohast, vaid saaks kogu sõnainfo kätte ühe päringuakna kaudu. Sõnaveeb on mõeldud abistama kõiki eesti keele rääkijaid, keeleuurijaid ja -huvilisi, keele- ja erialaeksperte – üldkeele kõrval kuvatakse infot ka ligikaudu 160 terminibaasist². Et mitte koormata eesti keele kui teise keele õppijaid aina kasvava infomahuga, loodi 2020. aastal Sõnaveebi kõrvale n-ö sõsarleht Keeleõppija Sõnaveeb (Koppel, Kallas 2021). Enne sõsarlehte kuvati keeleõppijale suunatud infot Sõnaveebi avalehel olevate filtrite kaudu, kuid selline keeleinfo esitusviis ei osutunud kasutajasõbralikuks, mistõttu tekkis vajadus eraldiseisva veebilehe järele.

Keeleõppija Sõnaveeb on mõeldud eesti keele kui teise keele õppijale ning seal kuvatav info on suures osas pärit ligikaudu 5000 märksõna sisaldanud „Eesti keele põhisõnavara sõnastikust 2014“ (PSV) (Kallas jt 2014), mida on aastate jooksul täiendatud. Kui PSV oli mõeldud eeskätt A2 ja B1 keeleoskustasemele ehk algajale ja edasijõudnud (täiskasvanud) keeleõppijale, siis Keeleõppija Sõnaveebis on märksõnu juba üle 6000, kusjuures suur osa lisandunud märksõnadest on kõrgema keeleoskustaseme (B2) sõnad. Märksõnade arv on muuhulgas kasvanud keeleõppijatele ja -õpetajatele mõeldud EKI piltsõnastiku (Kallas jt 2024) ja õpetaja tööriistade (Kallas jt 2021) arendamise tulemusena. Mõlemas ressursis on täiskasvanud keeleõppija kõrval kirjeldatud ka noore keeleõppija (sh koolieelikud ja I–III kooliaste) sõnavara.

Keeleõppija Sõnaveebil on Google Analyticsi andmeil ühes kuus umbes 2300 aktiivset kasutajat, nende seas ka keeleõpetajad. Viimastelt oleme koolituste käigus saanud tagasisidet, et sõnade seletused on algajale keeleõppijale liiga keerulised. Sellest tingituna käivitas ESF-i programm „Eesti keele õpe ja keeleõppe arendamine“ ehk Keelesamm

² Siin ja edaspidi kajastatud arvandmed (sh sõnakogude arv, kasutajate arv kuus jmt) pärinevad 2026. aasta algusest.

2025. aastal projekti, mille raames analüüsisime juba olemasolevaid Keeleõppija Sõnaveebi sõnaseletusi. Analüüsitud materjalist kerkisid esile leksikaalsed, morfoloogilised ja süntaktilised nähtused, mis näisid seletusest arusaamist raskendavat ja mille kohta tahtsime küsitluse abil teada õppijate eelistust. Küsitluse tulemuste põhjal uuendasime koostamis- põhimõtteid, millele toetudes oleme juba muutnud olemasolevaid Keeleõppija Sõnaveebi sõnaseletusi, nii et need oleksid kergemini mõistetavad A2–B1-tasemele. Samuti oleme lisanud Keeleõppija Sõnaveebi sada uut B2-taseme sõna, mille seletuste keerukus vastab A2–B1 keeleoskustasemele.

2. Sõnaseletused õppesõnastikes

Sõnatähenduste seletamine on ükskeelsete sõnastike peamine ülesanne. Sõnade seletamisel tuleb arvestada kõiki asjaolusid, eriti seda, kes on sõnastiku sihtgrupp. Oluline on jälgida, missugust infot seletus sisaldab ning milliseid sõnu ja grammatilisi konstruktsioone info edastamiseks kasutatakse. Sõnaseletused peaksid suutma rahuldada kahte tüüpi vajadusi: esimesel juhul pöördub kasutaja sõnastiku poole, kui ta tahab teada saada tundmatu sõna või väljendi tähendust; teisel juhul tahab kasutaja midagi öelda või kirjutada, kuid tal on vaja abi, et enda peas olevat tähendust võimalikult loomulikult ja mõjusalt väljendada. (Atkins, Rundell 2008: 405–407). Geoff Barnbrook (2002: 47) toob välja, et see tekitab õppesõnastike koostajatele keerulise dilemma: arusaamiseks oleks parem, kui seletus prooviks katta suurt hulka erinevaid kasutusi, millega õppija võib kokku puutuda, ent väljendamise abistamiseks tuleks pigem kirjeldada normatiivset, n-ö keskmist kasutust.

Keeleõppija Sõnaveeb ning selle aluseks olnud „Eesti keele põhisõnavara sõnastik 2014“ (PSV) (koostamispõhimõtetest vt Kallas, Tuulik 2011) on mõlemad aktiivset tüüpi sõnastikud: sisaldavad sõna leksikaalse ja süntaktilise käitumise malle ning leksikaalsemantilisi suhteid teiste sõnadega. Mõlemad sõnastikud on mõeldud mitmekülgeks abiliseks eesti keele kui teise keele õppimisel ja õpetamisel. Kasutame siinses artiklis PSV-d ja Keeleõppija Sõnaveebi kohati sünonüümsetena, kuigi PSV on pabersõnastik, mis ilmus 2014. aastal ka omaette veebilehel ja mille sisu püsib staatilisena. PSV-s olev keeleinfo ühendati 2019. aastal ka ÜS-i külge ning on aluseks Keeleõppija Sõnaveebis kuvatud keeleandmetele. Neid

on alates ÜS-iga ühendamisest muudetud ja täiendatud, mis tähendab, et Keeleõppija Sõnaveebis kuvatav sisu on dünaamiline.

PSV märksõnastiku koostamisel lähtuti eelkõige olemasolevate sagedussõnastike andmetest, sõnavara miinimumidest ja sagedusest ühendkorpuses (ühendkorpuste sarja kohta vt Koppel, Kallas 2022). Kuna korpuses pole suulise keele materjali, tuli käsitsi juurde lisada igapäevaelus vajalikke sõnu, mis kirjalikke tekste sisaldavas korpuses sageduse järgi esile ei tulnud (nt *hakkliha*, *pott*, *plaaster* jmt). Seletusi prooviti hoida võimalikult lihtsana ja nii järgiti sõnastiku koostamisel põhimõtet, et sõnade seletamisel kasutatakse vaid neidsamu sõnu ja sõnatähendusi, mis sõnastikus kirjeldatud on. Nii on info kergesti leitav ning kasutaja saab enda jaoks tundmatu sõna ja selle tähenduse sealsamas järele vaadata. Mitmetähenduslikel sõnadest on seletatud ainult sagedamad ja/või need tähendused, mis on keeleõppija kontekstis relevantid. Sageduse väljaselgitamiseks toetatakse üldjuhul keelekorpuse andmetele. Nii on sõnal *tõmbama* ÜS-is (Sõnaveebis) eristatud 30, PSV-s (ja Keeleõppija Sõnaveebis) viis tähendust – näiteks pole keeleõppijale seletatud tähendust 'ligimeelitavalt, ahvatlevalt mõjuma' (*Mind tõmbab teater*).

Leksikograafias ongi tavapärase, et keeleõppijatele suunatud sõnastikes kasutatakse kontrollitud defineerimissõnavara, mis koosneb piiratud hulgast (u 2000–3000) kõrge sagedusega sõnadest. Need on sõnad, mida õppijad peaksid teadma, et sõnade seletustest aru saada. Kuigi defineerimissõnavara võib tuua teatud kitsendusi (nt ristviited seotud sõnade vahel, idiomaatiliste väljendite ümbersõnastamine), on see kõige usaldusväärsem viis selgete ja arusaadavate seletuste loomiseks keeleõppijatele. (Atkins, Rundell 2008: 449; Svensén 2009: 246–248) Selline koostamise viis eeldab, et kasutajal on juba enne sõnastiku kasutamist omandatud mingi hulk põhisõnavara, mida ta juba tunneb, ning mille abil püütakse sõnastikus teiste sõnade seletusi edasi anda (Geeraerts 2006: 91).

Sue Atkins ja Michael Rundell (2008: 450) on öelnud, et sõnaseletused täidavad oma eesmärgi siis, kui nende sisu ja vorm on paigas. See, milised sisu ja vorm olema peavad, sõltub sõnastiku kasutajate vajadustest ja oskustest. Kui seletus ei paku kasutajale piisavalt kergesti mõistetavat infot, on sõnastik oma eesmärgis läbi kukkunud. Näiteks on lihtne lugeda selliseid sõnaseletusi, mis koosnevad reast sünonüümidest, kuid enamasti ei anna selline seletus sõna sisust piisavat ülevaadet. Teisalt ei ole mingit kasu seletusest, mis kirjeldab sõna sisu nii detailselt ja tehniliselt, et see

on arusaamatu neile kasutajatele, kellele sõnastik mõeldud on. Seega on keeleõppijate jaoks sõnaseletuste juures olulisim see, et need pakuksid ainult asjakohast ja hõlpsasti hoomatavat teavet (mitte detailset ja keerukat), olles samas piisavalt täpsed ja informatiivsed, aitamaks sõna tähendusest aru saada ning sõna kasutusele võtta (Atkins, Rundell 2008: 420). Tulemusliku sõnaseletuse jaoks on oluline selles kasutatavate sõnade valik, mis peaks olema mõistetavad lihtsamini kui seletatav sõna ise (Atkins, Rundell 2008: 448).

Niisiis on sõnatähenduste kirjeldamisel oluline silmas pidada seda, kellele seletus on suunatud. Kui sõnastiku sihtgrupp on keeleõppija, piisab sellest, et tähendust kirjeldatakse kergesti mõistetaval viisil, andes selles edasi informatsiooni sõna sisu kohta. (Atkins, Rundell 2008: 413–414) Nii on sõna *kari* seletus 'rühm koos peetavaid koduloomi, eriti lehmi' sobilik just (algajale) eesti keele õppijale (*pro* 'koos peetavate, karjatavate põllumajandusloomade rühm' ÜS-is), kuna selline seletus aitab õppijal mõista sõna tähendust sellistes tekstides, millega ta keelt õppides tõenäoliselt kokku puutub.

Keeleõppijatele loodud seletustes tuleb arvestada seega leksikaalse, aga ka grammatilise keerukusega. Selle kohta, kas seletuse grammatiline keerukus on kohane mingile kindlale keeleoskustasemele, leiab infot nt EKI õpetaja tööriistade lehelt (Sõnaveebis). Leksikaalse keerukusega on keerulisem, kuna tööriistad ei arvesta polüseemiaga. Kirjeldused õpetaja tööriistades põhinevad õpikute ja mitmesuguste korpuste andmestikul ning see on ühtne raamistik eesti keele kui teise keele sõnavara- ja grammatikapädevuse käsitlemiseks. Õpetaja tööriistades on neli moodulit: sõnavaraloendid, grammatikapädevuse kirjeldus, kasutusolukordade kirjeldused ning õppeteksti hindamise tööriist. (Kallas jt 2021) PSV koostamise ajal õpetaja tööriistade lehte veel ei eksisteerinud, mistõttu ei saanud tollal seletuste koostamisel keeleoskustasemete kohta käivale infole toetuda, kuid nüüd on see võimalus olemas.

PSV ja Keeleõppija Sõnaveeb on suunatud eelkõige A2–B1-tasemele ja seega on oluline, et sõnaseletused ka morfoloogilise ja süntaktilise keerukuse poolest sellele tasemele vastaksid. Keelesammu projekti analüüsisime enne küsitluse koostamist läbi ligi 900 Keeleõppija Sõnaveebi sõnaseletust (s.o ligikaudu 10% kõigist seletustest), milles kordusid teatud sõnaseletusmallid, mille leksikaalne, morfoloogiline ja süntaktiline keerukus võib artikli autorite arvates keeleõppijal seletusest

arusaamist raskendada. Morfoloogilise keerukuse aspektist esines sageli nt umbisikulist tegumoodi (kõikides Keeleõppija Sõnaveebi seletustes esines seda kokku 441 korda) ja *des*-vormi (130 korda); süntaktilise keerukuse aspektist nt *kui...siis*-tarindit (97 korda). Õpetaja tööriistades kirjeldatud grammatikamooduli järgi ei kuulu aga A2-taseme pädevuste hulka *des*-vormi ja *kui...siis*-tarindi mõistmine ega kasutus, neid sel tasemel ei kirjeldatagi. B1-tasemel õppija aga juba mõistab tegusõnade *des*-vormi tähendust, oskab seda kasutada mõnedes terviklikuna omandatud struktuurides ning kasutab pea- ja kõrvallause sidumiseks *kui...siis*-tarindit.

Eelnevast ilmneb, et sellise seletuse koostamine, mis erinevatele keeleoskustasemetele jõukohane oleks, pole lihtne ülesanne. Seega otsustasimegi Keeleõppija Sõnaveebi seletuste arusaadavust kasutaja peal testida ning valisime küsitlusse 20 märksõna, mille seletused illustreerisid eelnevalt kirjeldatavaid küsitavaid kohti. Küsitlusse kaasatud probleemsete nähtuste hulk ei olnud lõplik. Järgnev loend tuleneb artikli autorite praktilisest tööst ning toetub nendele nähtustele, mis 900 Keeleõppija Sõnaveebi seletuse analüüsimisel kõige sagedamini esile kerkisid ning töörühmas arutelu tekitasid.

Täpsemalt võtsime vaatluse alla järgmised **leksikaalsed nähtused** (ptk 4.1):

- 1) ase- ja sidesõnade kasutus (nt *ebasobiv*: *selline, mis ei sobi milleski või mis ei vasta mingitele tingimustele*);
- 2) abstraktse iseloomuga täpsustused (nt *meetod*: *kindel viis, kuidas midagi tehakse, et teatud eesmärki saavutada*);
- 3) omasõna vs. võõrsõna (nt *eessõna*: *nimisõna ees asetsev kaassõna vs. prepositsioon*);
- 4) märksõnaga samatüvelised sõnad (nt *aktiivne*: *aktiivselt tegutsev või toimiv*);

morfoloogilised nähtused (ptk 4.2):

- 5) umbisikuline tegumood (nt *miinus*: *matemaatikas: märk (-), mida kasutatakse siis, kui üks arv teisest lahutatakse*);
- 6) *des*-vorm (nt *muusika*: *helid, mida tehakse lauldes või pilli mängides*);
- 7) *v*-kesksõna (nt *elektrooniline*: *elektrit ja väga väikeseid elektriga töötavaid osi kasutav; selliste seadmete vahendusel töötav või toimuv*);

süntaktilised nähtused (ptk 4.3):

- 8) liigendatus, info kompaktsus (komakohad, loetelude esitus, sulud) (nt *koht: millegi (näiteks keha, teksti või ruumi) kitsam piirkond*);
- 9) *kui...siis*-tarind (nt *angiin: kui sul on angiin, siis sul on kurgus äge põletik*);
- 10) ebatäpne kombineerimine vs. topeldamine (nt *grupp: inimesed (või asjad), kes on ühes kohas koos või mingil moel seotud*);
- 11) põhisõna³ asend seletuses (nt *hääli: heli, mida sa rääkides või lauldes teed; esik: korteri välisuksele lähim ruum, kust lähevad ukсед tubadesse; kaebama: kellegi kohta midagi süüdistavat ütleva*);
- 12) piisav info hulk vs. lühidus (nt *kaart: maailma, riigi või linna vähendatud pilt, mille pealt saab vaadata mitmesuguseid objekte, näiteks teid, mägesid või jõgesid vs. skeem, kus on näiteks teed, mäed ja jõed*).

3. Küsitluse struktuur ja vastajate taust

Küsitlus, mille viisime läbi LimeSurvey keskkonnas, koosnes 25 küsimusest. Viies neist tahtsime saada taustainfot vastaja kohta, sh vanust, eesti keele oskuse taset, emakeelt, aga ka eelistusi keeleinfo otsimisel. Ülejäänud 20 küsimust keskendusid sõnaseletuste hindamisele. Selleks andsime vastajale teatud märksõna (nt *heeringas*) ühele tähendusele kaks seletuse varianti (*teatud piklik kala, kes on hõbeda värvi ja elab meres* ning *piklik kala, kes on hõbeda värvi ja elab meres*) (vt joonis 1). Üks variantidest oli originaalsõnastuses (nagu see oli Keeleõppija Sõnaveebis enne Keelesammu projekti algust) ning teine kohandatuna, kusjuures seletuste esitamise järjekord varieerus märksõnati. Vastaja ülesanne oli hinnata, kumb seletus aitab tal paremini aru saada, mida sõna tähendab. Samuti oli võimalik jätta valik tegemata (valida *ei oska öelda*) ning lisada kommentaar. Piirdusime küsitluses 20 märksõnaga, kuna sellest piisas, et käsitleda analüüsis esile kerkinud probleemseid sõnaseletuste malle.

³ Põhisõna on seletuse põhjaks olev sõna, mida muud osad seletuses täpsustavad.

Seletused

*

sõna: heeringas

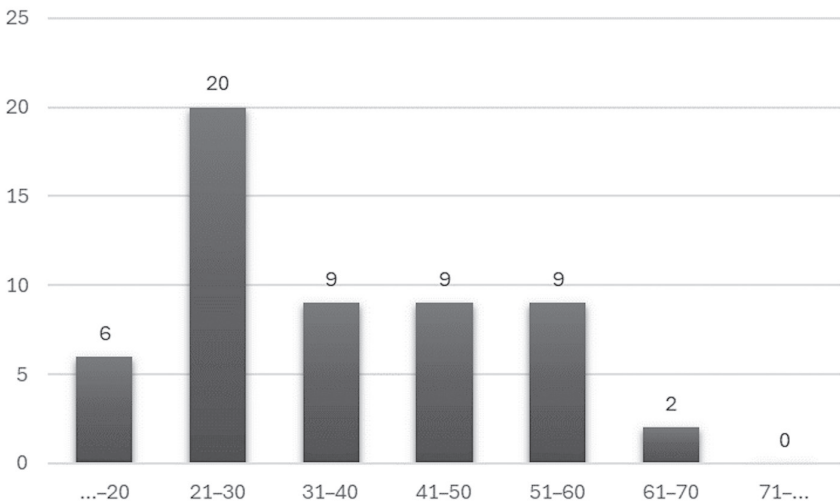
Kumb seletus aitab teil paremini aru saada, mida see sõna tähendab?

- piklik kala, kes on hõbeda värvi ja elab meres
- teatud piklik kala, kes on hõbeda värvi ja elab meres
- ei oska öelda

Palun lisage oma kommentaar siia:

Joonis 1. Näide küsitlusest LimeSurvey keskkonnas

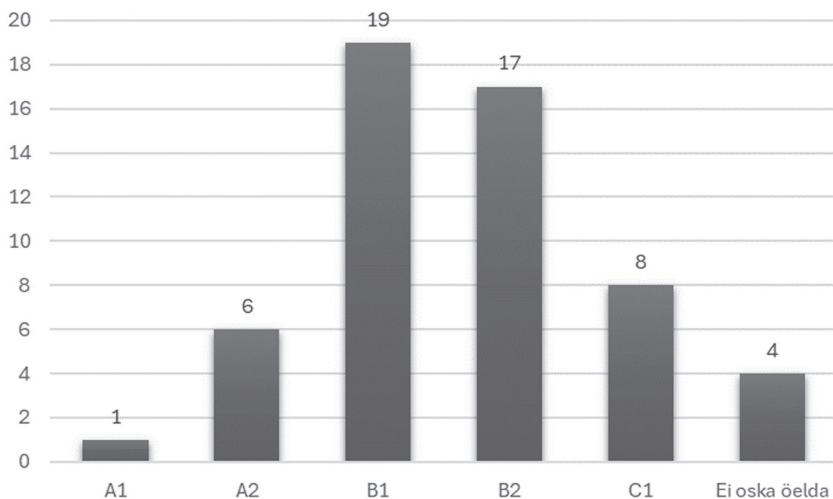
Küsitlust jagasime kohalike ja välisülikoolide, üldhariduskoolide ja täiskasvanutele mõeldud keelekursuste eesti keele kui teise keele õpetajatega ning eesti keele kui teise keele õppijate seas, keeleõpetajatele suunatud EKI õpiampsul ning Facebooki grupis „Räägi minuga“. Viimases on u 25 500 liiget, kellest suure osa moodustavad eesti keelt emakeelena mitte rääkivad inimesed. Küsitlus oli avatud 13 päeva ning selle ajaga kogusime 55 vastust eri vanuseklassidest (vt joonis 2) ning keeleoskustasemelt (vt joonis 3).



Joonis 2. Küsitluses osalenute vanus

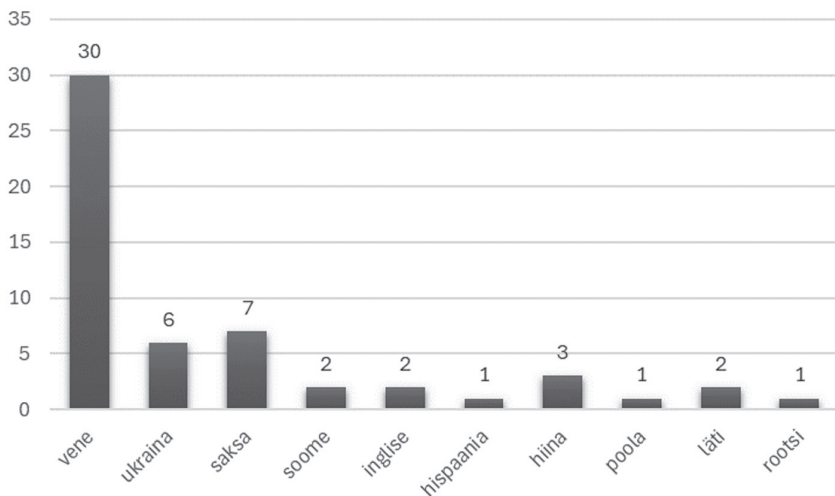
Nagu jooniselt 2 näha, oli kõige rohkem vastajaid vanuseklassist 21–30 – sinna kuulus u 36% vastanuist. Vanuseklassis 31–40, 41–50 ja 51–60 oli vastajaid võrdselt – igaühes ligikaudu 16%. Umbes 11% vastajaist olid kuni 20-aastased ning kõige vähem vastajaid saime vanusevahemikus 61–70 eluaastat (u 4%).

Enamik vastajaist (u 65%) olid iseseisvad keelekasutajad (Raamdokument 2007), kes valdasid eesti keelt B1 või B2 keeleoskustasemel (vt joonis 3). Vilunud (C1-tase) ning algtasemel keeleõppijaid (A1–A2-tase) oli vastajate seas üldjoontes võrdselt (vastavalt 15% ja 13%). Neli vastanut (u 7%) ei osanud oma keeleoskustaset ise määratleda.



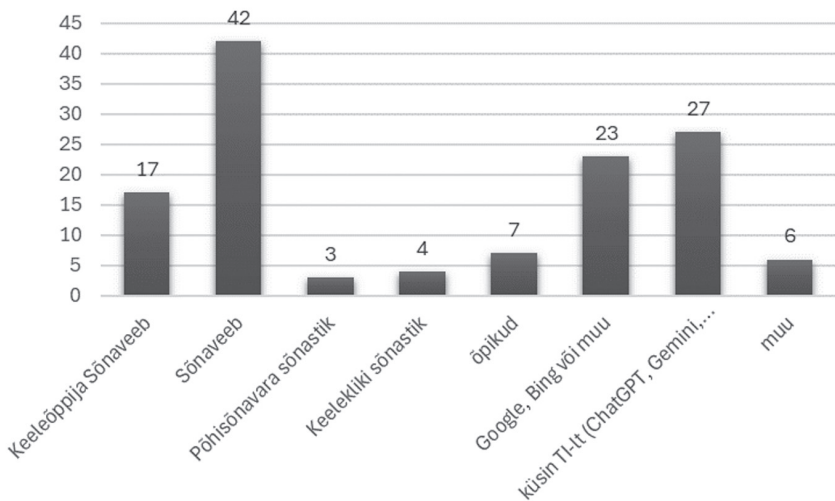
Joonis 3. Küsitluses osalenute keeleoskustase

Küsitluse tulemused näitavad, et valimis olid kõige enam esindatud slaavi keelte taustaga eesti keele kui teise keele õppijad, eriti vene keelt emakeelena rääkijad. Viimaseid oli kokku 30 ehk üle poole vastajaist (u 60%). Seitse inimest märkis oma emakeeleks saksa keele (u 13%), kuus inimest ukraina keele (u 11%) ning kolm inimest hiina keele (u 5%). Kahe vastajaga olid esindatud soome, inglise ja läti keelt emakeelena kõnelejad ning ühe vastajaga hispaania, poola ja rootsi keelt emakeelena kõnelejad (vt joonis 4).



Joonis 4. Küsitluses osalenute emakeel

Soovisime küsitluses teada saada ka seda, kust otsivad keeleõppijad infot, kui nad eestikeelse sõna tähendust ei tea. Andsime selleks ette seitse vastusevarianti, millest valida sai mitu. Vastajail oli võimalik ka ise oma infoallikaid täpsustada ning seda võimalust kasutas kuus inimest (vt joonis 5).



Joonis 5. Keeleõppijate eelistatud keeleinfo allikad

Nagu jooniselt 5 nähtub, kasutatakse keeleinfo allikana kõige enam Sõnaveebi, millele järgneb tehisintellekt (ChatGPT, Gemini jmt), interneti otsingumootorid (Google, Bing jmt) ning Keeleõppija Sõnaveeb. Vähemal määral kasutatakse õpikuid, Keelekliki sõnastikku ning PSV-d. Kaks vastajat mainis infoallikana tõlkemootorit DeepL ning kakskeelseid sõnastikke, aga ka eesti keele seletavat sõnaraamatut ning Vikipeediat (loodussõnavara, eriti taimede, loomade, seente jms otsimiseks). Üks vastajaist kirjutas, et küsib abi eestlasest vestluspartneri käest, kui temaga rääkides mõnd sõna ei tea ning kui ta ei taha tema juuresolekul telefonist sõnu otsides ebaviisakas olla.

Eraldi küsisime kakskeelsete sõnastike kasutamise kohta, kuna soovisime teada, milliseid neist eestikeelsete sõnade tähenduste vaatamiseks kasutatakse. Vastajad mainisid eesti-vene, inglise-eesti, eesti-saksa, soomeeesti ning eesti-läti sõnaraamatut. Ka selle küsimuse vastustes nimetati mitu korda tõlkeprogrammi DeepL, kus eesti keel on üks 37 keelest, mille vahel tõlkimist võimaldatakse. Veel mainiti mitmekeelset veebisõnastikku Glosbe ja eraldi toodi korduvalt välja ka Google Translate'i. Meie pakutud vastusevariantide hulgas olid küll otsingumootorid Google, Bing vm, aga mitte tõlkeprogrammiosa iseseisva valikuna.

4. Tulemused

Küsitluses olid esindatud kolm kõige sagedasemat sõnaliiki – 13 nimi-sõna, neli tegusõna ja kolm omadussõna. Sõnaseletused, mille küsitlusse valisime, olid seotud teatud parameetritega, mis seletustes sageli kordusid ning mida vastajate peal testida tahtsime (vt loetelu ptk 2). Need parameetrid jaotusid leksika (4 parameetrit, ptk 4.1), morfoloogia (3 parameetrit, ptk 4.2) ja süntaksi (5 nähtust, ptk 4.3) tasandi vahel.

Tabelis 1 on kokkuvõtvalt toodud küsitluses kasutatud märksõnad ning arvandmed selle kohta, kui palju vastajaist eelistas originaalsõnastust, kui palju kohandatud seletust ning kui paljudel vastajatest eelistus puudus. Eelistatud variant on esile toodud paksus kirjas.

Tabel 1. Vastajate eelistused märksõnade seletuse sõnastuse osas. Protsendid on ümardatud täisarvuni

Märksõna + näite number	Originaal-sõnastuses seletus	Kohandatud seletus	Ei oska öelda
aadress (1)	6 (11%)	49 (89%)	0
alla andma (2)	20 (36%)	26 (47%)	9 (16%)
heeringas (3)	12 (22%)	41 (75%)	2 (4%)
hapu (4)	22 (40%)	33 (60%)	0
ajavahemik (5)	22 (40%)	31 (56%)	2 (4%)
ehtima (6)	25 (45%)	27 (49%)	3 (5%)
esineja (7)	11 (20%)	41 (75%)	3 (5%)
arvutimäng (8)	28 (51%)	22 (40%)	5 (9%)
armuke (9)	22 (40%)	24 (44%)	9 (16%)
etendus (10)	39 (71%)	16 (29%)	0
hankima (11)	17 (31%)	32 (58%)	6 (11%)
ettevaatlik (12)	29 (53%)	25 (45%)	1 (2%)
ilm (13)	11 (20%)	44 (80%)	0
elamine (14)	23 (42%)	25 (45%)	7 (13%)
elamus (15)	27 (49%)	25 (45%)	3 (5%)
avama (16)	26 (47%)	27 (49%)	2 (4%)
eeskuju (17)	28 (51%)	20 (36%)	7 (13%)
eestlane (18)	9 (16%)	45 (82%)	1 (2%)
aasta (19)	27 (49%)	27 (49%)	1 (2%)
ametlik (20)	26 (47%)	26 (47%)	3 (5%)

4.1. Leksikaalsed parameetrid

4.1.1. Ase- ja sidesõnade kasutus

Ase- ja sidesõnade mõju hindamiseks valisime välja kahe märksõna (*aadress* ja *alla andma*) seletused, kus ase- ja sidesõnu esines rohkelt, ning sõnastasime originaalseletuse kõrvale uue, kohandatud variandi.

Siinkohal tuleb märkida, et enamasti esineb küsitluses kasutatud seletustes mitu vaadeldavat parameetrit korraga. Näiteks märksõna *aadress* (näide 1) originaalseletuses pidasime probleemseks lisaks sisutihjadele asesõnadele ka selle liigendatust (sidesõnad vs. komad, vt ptk 4.3.1) ja fraaside pikkust. Siin ja edaspidi on näidete all esitatud kaks seletuse varianti, millest esimene on originaalsõnastuses (O) ning teine uues, kohandatud sõnastuses (K). Seletusele järgneb seda varianti eelistanud inimeste arv ning täisarvuni ümardatud protsent.

(1) **aadress**

O: *kellegi kodu või mingi asutuse andmed, näiteks maja number ning tänava ja linna nimi* – 6 (11%)

K: *kodu või asutuse andmed: tänav, maja number, sihtnumber, linn, riik* – 49 (89%)

Kohandatud seletust pidas arusaadavamaks 89% vastanuist. Kolm neist olid lisanud ka kommentaari: koma on selgem kui sidesõnad; kohandatud seletus on selgem; lühikesed märksõnad on paremad. Viimase kommentaari autor pidas *märksõna* all silmas ilmselt nominatiivis sõnavorme loetelus.

Märksõnaga *alla andma* (näide 2) soovisime vaadata, kas ja kui palju segab eesti keele kui teise keele õppijaid *sina*-vormi kasutus seletuses.

(2) **alla andma**

O: *kui sa annad alla, siis sa loobud kellegi vastu võitlemast, lepid mingi olukorraga* – 20 (36%)

K: *kellegi või millegi vastu võitlemast loobuma; mingi olukorraga leppima* – 26 (47%)

Siin hindasid umbes pooled vastajad kohandatud seletuse selgemaks (47%), küllaltki palju eelistati ka originaalsõnastuses seletust (36%). Esile kerkib nende osakaal (16%), kes otsust ühe või teise seletuse kasuks langetada ei osanud. Vastajate jaoks osutus ilmselt keeruliseks konstruktsioon *kellegi vastu võitlemast loobuma*, mis on õpetaja tööriistade järgi kõrgema, B2-taseme kategooria. Ühes kommentaaris tõstetigi esile just *mast*-vormi keerukust ja teises peeti infinitiivi seletuse paremaks kui ainsuse 2. pöoret. *mast*-vormi kasutus Keeleõppija Sõnaveebi seletustes tegelikult sage ei olegi – kokku esines seda kõigis seletustes vaid üheksa korda.

4.1.2. Täpsustused: *teatud, kindel, näiteks*

Keeleõppija Sõnaveebi seletusi analüüsid kerkisid materjalis esile nn täpsustused (nt *teatud, näiteks, kindel*), mida leksikograafid kasutavad ühe konkreetse liigi vms kirjeldamiseks paljude hulga. Kõigis Keeleõppija Sõnaveebi seletustes esines sõna *näiteks* kokku 507, *teatud* 247 ning *kindel* 114 korda. Näiteks on sõna *heeringas* (näide 3) seletuses täpsustatud, et see on *teatud kala*, sest mitte kõik hõbedased ja pikad kalad, kes meres elavad, ei ole heeringad. Saamaks teada, kas õppijad selliseid täpsustusi vajalikuks peavad, palusime hinnata märksõnade *heeringas, hapu* ja *ajavaheimik* originaalsõnastust ja kohandatud seletust (näited 3–5).

(3) **heeringas**

O: *teatud piklik kala, kes on hõbeda värvi ja elab meres* – 12 (22%)

K: *piklik kala, kes on hõbeda värvi ja elab meres* – 41 (75%)

Näites 3 jätsime kohandatud seletuses ainsa muudatusena ära määratluse *teatud* ning suurem osa vastajaist (75%) seda ka eelistas. Üks vastajaist ka kommenteeris, et *teatud* on nn tühi sõna, mis lausele midagi juurde ei lisa.

Näites 4 eelistasid vastajad (60%) samuti kohandatud seletust. Tulemus oli ootuspärane, kuna originaalsõnastuses seletus sisaldab üsna keerukat konstruktsiooni *millelegi iseloomulik olema*, kuigi ka üsna suur hulk vastajaid (40%) pidas originaalsõnastuses seletust arusaadavamaks. Siinkohal ei saa järeldusi teha vastajate keeleoskustaseme kohta, sest originaalsõnastuses seletuse eelistajate seas on peaaegu võrdselt madalama ja kõrgema keeleoskustasemega õppijaid.

(4) **hapu**

O: *teatud tugeva maitsega, mis on iseloomulik näiteks sidrunile või äädikale* – 22 (40%)

K: *sellise maitsega nagu sidrun või äädikas* – 33 (60%)

Näites 5 soovisime seletust lühendada täpsustuste arvelt, st kohandatud seletusest on ära jäetud sõnad *näiteks* ja *kindel*.

(5) **ajavaheimik**

O: *aeg, mis kestab näiteks mingist kindlast päevast või aastast mingi teise kindla päevani või aastani* – 22 (40%)

K: *aeg, mis kestab mingist päevast või aastast mingi teise päevani või aastani* – 31 (56%)

Ka näite 5 puhul eelistas rohkem vastajaid (56%) kohandatud seletust.

4.1.3. Omasõna vs. võõrsõna

Leksikograafias üldjuhul taunitakse sünonüümide kaudu tähenduste seletamist, kuna enamik sünonüüme ei ole semantiliselt identsed ning poliuseemia ja tähendusnihked võivad kasutajat eksitada, mistõttu sünonüümid võivad küll seletust toetada, kuid mitte olla seletuse ainsaks vahendiks (Atkins, Rundell 2008: 450–422). Uurimaks, kas õppijad eelistavad võõrsõnalist ja lühemat või omasõnalist, aga täpsemat ja keerukamat seletust, palusime hinnata märksõna *ehtima* (näide 6) seletusi (täheanduses 'kujundades, värve, kaunistusi vm lisades kauni(ma)ks, ilusa(ma)ks tegema'). Siin valisime uue kohandatud seletuse sõnastuseks sünonüümi *dekoreerima*, mis on ÜS-is selles tähenduses *ehtima* täissünonüüm ehk jagab sama tähendust, kuigi *dekoreerima* ei kuulu keeleõppija sõnavarasse ega ole ei PSV-s ega Keeleõppija Sõnaveebis omaette märksõnana esitatud. Koostamispõhimõtete järgi sellist sõnaseletust seega esitada ei saa, ent tahtsime siiski uurida, kas kasutajad eelistavad seletuses lühemat võõrsõnalist sünonüümi või pikemat seletust omasõnadega. Eeldasime, et madalama keeleoskustasemega vastaja eelistab pigem esimest varianti, kuna see viitab sarnasele ingliskeelsele sõnakujule ja tähendusele. Selguse huvides tuleb siin mainida, et *ehtima* teised tähendused (sh 'millegi kaunistuseks, ehteks olema' ja 'end millekski valmis seadma') ei kuulu A2–B1-taseme sõnavarasse.

(6) *ehtima*

O: *kedagi või midagi ilusamaks tegema* – 25 (45%)

K: *dekoreerima* – 27 (49%)

Näite 6 puhul tulebki esile erinevus A2–B1 ja B2–C1 keeleoskustaseme vahel. 62% A2–B1-tasemega vastajaist eelistas võõrsõnalist seletust, B2–C1-taseme hulgast hindas kohandatud seletust paremaks vaid 36% – suurem osa neist eelistas omasõnalist ja täpsemat seletust. Neli vastajat kasutas võimalust kommenteerida: üks ühendaks need seletused; teine arvas, et mõlemad sobivad hästi; kolmas leidis, et mõlemad seletused võiksid olemas olla; neljas nentis, et sõna *dekoreerima* on lihtsam mõista, kuna see on ingliskeelse sõnaga sarnane. Kommentaaride põhjal võib arvata, et mõni vastaja võis tajuda pakutud sõnaseletuste tähenduserinevust, kuigi meie eesmärk oli esitada sama tähenduse kaks seletuse varianti.

4.1.4. Märksõnaga samatüvelised sõnad seletustes

Üks PSV koostamispõhimõtteid oli see, et seletustes välditakse üldiselt tähenduse edasiandmiseks märksõnaga samatüveliste sõnade kasutamist⁴. Alati seda põhimõtet siiski ei järgitud, nagu nt märksõnades *esineja* (näide 7), *arvutimäng* (näide 8) ja *avama* (näide 16).

(7) *esineja*

O: *keegi, kes mingil üritusel esineb* – 11 (20%)

K: *keegi, kes esineb, näiteks laulab, mängib pilli või peab kõnet* – 49 (89%)

Näite 7 puhul valisid vastajad ülekaalukalt (89%) arusaadavamaks kohandatud seletuse, kuigi ka selles esinev põhiverb on sama tüvega kui märksõna. Kohandatud seletuse paremale mõistmisele aitavad kaasa näited. Märksõna *esineja* juures kasutas kaheksa vastajat võimalust kommenteerida. Neljas kommentaaris pöörati tähelepanu just sellele probleemile, et märksõna *esineja* tähendus ei selgu *esinema* kaudu. Samuti leiti näiteks, et sõna *üritus* raskendab arusaamist; aga ka, et mõlemad variandid sobivad.

Ka märksõna *arvutimäng* (näide 8) seletuse puhul toodi esile, et sama tüve kasutamine ei tule seletuse mõistmisele kasuks.

(8) *arvutimäng*

O: *mäng, mida mängitakse arvutiga* – 28 (51%)

K: *mäng, mida saab mängida arvutiga* – 22 (40%)

Üks vastaja kommenteeris, et liitsõna (*arvutimäng*) ei tohiks seletada selle komponentide (*arvuti* ja *mäng*) kaudu, sest sõnastiku kasutaja ei pruugi teada, mida need komponendid tähendavad. PSV-d koostades kasutati kohati sellist seletuse malli, kuna kõik seletustes ja näidetes kasutatavad sõnad ning liitsõna komponendid on omaette märksõnad ja veebisõnastikus klikitavad – seega saab kasutaja tundmatu sõna tähendust kohe sealtsamast mugavalt järele vaadata.

⁴ PSV koostamisjuhend on kättesaadav EKI OneDrive'i pilvesalvestusruumis.

4.2. Morfoloogilised parameetrid

4.2.1. Umbisikuline tegumood

Eesti keele kui teise keele õpikute tekstide analüüsist on selgunud, et umbisikulist tegumoodi kasutatakse vähesel määral küll A1- ja A2-taseme tekstides, mõnevõrra tõuseb kasutus B1-taseme tekstides, kuid hüppeliselt on seda näha B2-taseme tekstides (Koppel 2019: 111). Ka õpetaja tööriistade järgi mõistab keeleõppija A2-tasemel umbisikulise tegumoe oleviku jaatavate vormide tähendust ning moodustama hakatakse neid alles B1-tasemel. PSV sõnaseletustes prooviti umbisikulist tegumoodi üldjuhul vältida. Juhul kui muu sõnastus (nt *saab teha* asemel *tehakse*) mõjus kohmakalt ja ebaloomulikult, kasutati seletamisel siiski umbisikut. Küsitluses tahtsime teada, kas ja kui võrd raskendab umbisikulise tegumoe kasutus seletusest arusaamist. Vastandasime selleks umbisikulises tegumoes seletuse isikulises tegumoes seletusega kolme sõna juures: *armuke* (näide 9), *etendus* (näide 10) ja *arvutimäng* (näide 8, vt ptk 4.1.4).

(9) *armuke*

O: *kindel partner seksis väljaspool abielu* – 22 (40%)

K: *kindel inimene, kellega väljaspool abielu seksitakse* – 24 (44%)

(10) *etendus*

O: *üritus, kus publikule esitatakse näiteks näidendit, ooperit või balletti* – 39 (71%)

K: *näiteks näidendi, ooperi või balleti esitus publikule* – 16 (29%)

44% vastanuist hindas sõna *armuke* seletustest arusaadavamaks kohandatud seletuse, kus on kasutatud umbisikulist tegumoodi. 40% vastanuist eelistas originaalsõnastuses seletust, kus umbisikulist tegumoodi ei kasutatud. 16% vastajatest ei osanud öelda, kumb seletus on selgem.

Sõna *etendus* puhul eelistas originaalsõnastust 71% ja sõna *arvutimäng* puhul 51% vastajaist. Kõigi kolme näite puhul ei mõjuta tõenäoliselt seletusest arusaamist mitte umbisikuline tegumood, vaid muud parameetrid, nt põhisõna asend (näide 10) või see, et seletuses kasutatakse märksõnas sisalduvaid sõnattüvesid (näide 8).

4.2.2. *des*-vorm

Õpetaja tööriistade järgi kuulub *des*-vorm keerukuse poolest B2-taseme oskuste hulka. Ka eesti keele kui teise keele õpikuid analüüsid on selgunud, et *des*-vorm ilmub tekstidesse alles B1-tasemel, kuid rohkem hakatakse kasutama seda B2-taseme lausetes (Koppel 2019: 108). Kuna Keeleõppija Sõnaveebi sihtgrupp on eeskätt A2–B1-tasemel õppijad, kohandasime seletused ümber nii, et need *des*-vormi ei sisaldaks. Ühtlasi soovisime teada, kas ja kui võrd segab õppijaid *des*-vormi olemasolu seletuses. Järgnevalt on esitatud märksõna *hankima* (näide 11).

(11) **hankima**

O: *otsides või asju ajades midagi saama* – 17 (31%)

K: *otsima või asju ajama ja siis midagi saama* – 32 (58%)

Kohandatud, ilma *des*-vormita seletust eelistasid vastajad mõnevõrra rohkem (58%). Siin tuleb esile ka A2–B1-taseme ja B2–C1-taseme eelistuste erinevus: kõrgema taseme vastanuist eelistas kohandatud seletust 68%, madalama taseme vastanuist eelistas kohandatud seletust 54%. Üsna suur oli nende osakaal (11%), kes ei osanud otsustada, kumb seletus arusaadavam on.

4.2.3. *v*-kesksõna

Õpetajate tööriistade grammatika mooduli järgi ei kuulu *v*-kesksõna mõistmine A2-taseme pädevuste hulka, B1-tasemel õppija mõistab selle tähendust ja oskab seda moodustada mõnest õpitud tegusõnast, B2-taseme õppija oskab seda moodustada tuttavatest tegusõnadest ja kasutada täiendina ning õpitud juhtudel ka öeldistäitena. Pidades silmas seda, et seletused oleksid kohased just A2–B1-tasemele, uurisime, kuidas hindavad õppijad *v*-kesksõna sisaldava seletuse arusaadavust, kasutades märksõna *ettevaatlik*.

(12) **ettevaatlik**

O: *tähelepanelik, vigadest ja õnnetustest hoiduda püüdev* – 29 (53%)

K: *selline, kes püüab vigadest ja õnnetustest hoiduda, kes on tähelepanelik* – 25 (45%)

Kuigi kohandatud seletuses vältisime *v*-kesksõna, eelistas *v*-kesksõna sisaldanud originaalsõnastuses seletust pisut üle poole vastajaist (53%).

Tulemustest võib järeldada, et v-kesksõna seletusest arusaamist ei mõjuta, küll aga võis seda teha B2-taseme konstruktsioon *hoiduma* [MILLEST]. Võimalik, et originaalsõnastuses seletust eelistatakse rohkem, kuna see algab sünonüümset tähendust kandva põhisõnaga. Kohandatud seletust alustab asesõna, mis tähendust edasi ei anna.

4.3. Süntaktilised parameetrid

4.3.1. Liigendatus ja info kompaktsus

Mõne märksõna (nt *ilm*, näide 13) tähendus tingib seletuse, mis koosneb mitmest osalausest ja sisaldab loetelu ning millest A2–B1-taseme õppijal võib olla keeruline aru saada. Eesmärk oli sõnastada kohandatud seletus selliselt, et info oleks võimalikult kompaktselt esitatud. Palusime õppijail hinnata märksõnade *ilm* ja *aadress* (näide 1, vt ptk 4.1.1) seletusi.

(13) *ilm*

O: *teatud ajal väljas valitsevad tingimused, näiteks temperatuur, tuul ja see, kas sajab vihma või paistab päike* – 11 (20%)

K: *tingimused väljas: temperatuur, tuul, sajab vihma või paistab päike* – 44 (80%)

Mõlema märksõna puhul eelistasid vastajad ülekaalukalt kohandatud seletust: märksõna *ilm* puhul 80% ja märksõna *aadress* puhul 89%.

Soovisime ka teada saada, kas sulud seletustes mõjutavad eesti keele kui teise keele õppijaid. Selleks palusime hinnata märksõnade *elamine* (näide 14) ja *elamus* (näide 15) seletusi.

(14) *elamine*

O: *oma elu veetmine (mingil viisil)* – 23 (42%)

K: *oma elu veetmine, mingil viisil olemine* – 25 (45%)

(15) *elamus*

O: *tugev (hea) kogemus* – 27 (49%)

K: *tugev ja enamasti hea kogemus* – 25 (45%)

Mõlema märksõna puhul jagunevad originaalsõnastuses ja kohandatud seletuse eelistused peaaegu võrdselt. Märksõna *elamine* puhul torkab silma ka nende inimeste osakaal (13%), kes ei osanud valida, kumb seletus selgem on. Sulud seletustes ei paista vastajaid segavat. Segavaks osutub muu: kolmes kommentaaris on välja toodud, et omadussõna *hea* seletustes ei ole kohane, kuna *elamus* võib olla ka negatiivne.

Komade arvu lauses on uuritud ka õppesõnastike ja eesti keele kui teise keele õpikute näitelausestes (Koppel 2020). Uurimuse tulemusena selgus, et õppijale suunatud näited on üldjuhul lihtlauseid ning seetõttu eelistatakse õppesõnastike näitelauseste korpuspõhisel automaattuvastusel lauseid, milles esineb kuni üks koma (vältimaks osalauseste esinemist).

4.3.2. *kui...siis*-sõnaseletusmall

PSV-d koostades oli *kui...siis*-sõnaseletusmalli kasutamise põhjuseks soov näidata ühend- ja väljendverbide komponentide järjekorda lause tasandil ning vältida komplekssemaid konstruktsioone ja keerukamat sõnavara. Õpetaja tööriistade grammatika mooduli põhjal ei kuulu aga A2-taseme pädevuste hulka pealause ja kõrvallause sidumine *kui...siis*-tarindi abil, küll peaks seda oskama kasutada B1-tasemel õppija. Uurisime, kas ja kuidas segab vastajaid pikem seletus, kus liitlause osalused *kui...siis* abil ühendatud on. Selleks palusime hinnata märksõnade *avama* (näide 16) ja *alla andma* (näide 2, vt ptk 4.1.1) seletusi.

(16) *avama*

O: *kui mingi pood, restoran vms on avatud, siis töötajad on kohal ja inimesed saavad seal käia* – 26 (47%)

K: *poodi, restorani vms lahti tegema* – 27 (49%)

Üldse mitte ootuspäraselt eelistasid märksõna *avama* seletustest u pooled vastanuist (47%) originaalsõnastuses pikemat varianti (kohandatud varianti eelistas 49%). Märksõna *alla andma* puhul eelistati kohandatud varianti (47%) veidi rohkem kui originaalsõnastuses (36%), kuid esile kerkib ka nende osakaal (16%), kes valida ei osanud. Kui vaadata A1–B1- ja B2–C1-taseme tulemusi eraldi, torkab silma, et kõrgema keeleoskustasemega õppijad eelistavad mõlema näite puhul kohandatud seletust veidi rohkem: *avama* puhul B2–C1 56% vs. A1–B1 42%, *alla andma* puhul B2–C1 64% vs. A1–B1 38%.

4.3.3. Ebatäpne kombineerimine vs. topeldamine

Mõnel juhul on PSV sõnaseletuses soovitud võimalikult palju infot esitada võimalikult kompaktselt, nagu nt märksõna *eeskuju* (näide 17) puhul. Originaalsõnastuses seletus kirjeldab seda kui *inimest või käitumise viisi*, millele järgnevas osaluses viidatakse nii inimesele kui ka käitumise

viisile tagasi asesõnaga *mis*, kuid see pole grammatiliselt täpne (inimesele pidanuks viitama asesõnaga *kes*). Kohandatud seletuse sõnastasime selliselt, et andsime varasemalt kokku surutud info kahe järjestikku esitatud tähendusena.

(17) **eeskuju**

O: *inimene või käitumise viis, mis on mudeliks, mida teised inimesed järgivad* – 29 (53%)

K: *inimene, keda teised järgivad; käitumise viis, mida teised järgivad* – 25 (45%)

Vastajaist veidi rohkem kui pooled (53%) eelistavad originaalsõnastuses seletust ja esile kerkib ka see, et B2–C1-tasemel õppijaist 60% peab originaalsõnastuses seletust arusaadavamaks. Kaks neist ka põhjendavad kommentaaris oma valikut: ühe jaoks lihtsustab arusaamist laensõna *mudel*, teise jaoks saab originaalsõnastuses kombineeritud seletuse abil tähendus selgeks lühema ajakuluga.

4.3.4. Põhisõna asend

Keeleõppija Sõnaveebi seletusi analüüsid märkasime, et põhisõna võib asuda seletuses nii alguses (*fotograaf* – inimene, kelle töö on pildistada), keskel (*friikartul* – pikkadeks pulkadeks lõigatud kartul, mis on õlis pehmeks keedetud) kui ka lõpus (*esindus* – riiki või ettevõtet esindav asutus). Uurisime ka seda, kuidas mõjutab põhisõna asend seletuses selle mõistmist. Selleks palusime hinnata märksõnade *eestlane* (näide 18), *etendus* (näide 10, ptk 4.2.1) ja *aasta* (näide 19) seletusi.

(18) **eestlane**

O: *Eesti päritolu inimene* – 9 (16%)

K: *inimene, kes on pärit Eestist* – 45 (82%)

Märksõna *eestlane* puhul eelistasid vastajad ülekaalukalt (82%) kohandatud seletust, kus põhisõna seletust alustab. Märksõna *etendus* puhul eelistasid vastajad (71%) originaalsõnastuses seletust, kus põhisõna samuti alustab seletust. *Etenduse* kohandatud seletuses asub see aga peaaegu lõpus.

A1-taseme märksõna *aasta* kohandatud seletust lühendasime põhisõna *periood* täiendite arvelt, nii et põhisõna *periood* alustaks seletust.

(19) **aasta**

O: 365 või 366 päeva pikkune periood, mis algab 1. jaanuaril ja lõpeb 31. detsembril – 27 (49%)

K: periood, mis algab 1. jaanuaril ja lõpeb 31. detsembril – 27 (49%)

Siin aga jagunevad eelistused võrdselt. Numbrid 365 ja 366 on sõltumata emakeelest märksõnaga *aasta* seostatavad, nii et hulk täiendeid põhisõna ees ei raskenda arusaamist. Ühes kommentaaris pakuti välja, et seletus peaks algama nii: „*periood, mis kestab 365 või 366 päeva ja...*“.

4.3.5. Piisav info hulk vs. lühidus

Kui keeleõppijal on vaja endale selgeks teha uusi sõnatähendusi, siis ei soovi ta kulutada selleks liiga palju aega. Sellest lähtudes võib oletada, et lühemat seletust eelistatakse rohkem, kuna nii on võimalik info kiiremini kätte saada. Et uurida, kuidas seletuse pikkus mõjutab vastajaid, palusime hinnata märksõnade *esineja* (näide 7, ptk 4.1.4), *avama* (16, ptk 4.3.2) ja *ametlik* (20) seletusi.

Märksõna *esineja* seletustest eelistasid vastajad ülekaalukalt (89%) kohandatud seletust, vaatamata sellele, et see on võrreldes originaalsõnastuses seletusega oluliselt pikem. Sellest nähtub, et õppijate jaoks on tähtis piisav info hulk, mitte tingimata seletuse lühidus.

Ka märksõnade *avama* ja *ametlik* põhjal saab teha sarnase järelduse, sest automaatselt ei eelistata kohandatud ja lühemat seletust. Mõlema sõna puhul eelistavad (ligikaudu) pooled vastajaist võrdselt originaalsõnastuses ja pikemat ning pooled vastajaist kohandatud ja lühemat seletust.

(20) **ametlik**

O: valitsuse või muude võimu esindavate organisatsioonide või isikutega seotud; nende kehtestatud korrale vastav – 26 (47%)

K: valitsuse või riigiga seotud; kehtestatud korrale vastav – 26 (47%)

5. Põhjäreldused ja arutelu

Küsitlusele vastasid peamiselt B1–B2 keeleoskustasemel õppijad, samas kui A1- ja A2-taseme (kokku 7 vastajat) ning C1-taseme (8 vastajat) esindatus oli ligikaudu võrdne. Kuigi selline jaotus võimaldab teataval määral võrrelda eri keeleoskustasemega õppijate eelistusi, raskendab vastajate ebavõrdne jaotus keeleoskustasemete lõikes põhjapanevate järelduste

tegemist. Edaspidi oleks otstarbekas kujundada küsitluse valim selliselt, et keeleoskustasemetes esindatus oleks rohkem tasakaalus.

Seda, et küsitluses kasutatud seletustes esines enamasti mitu vaadeldavat parameetrit korraga, võib vaadelda nii plussi kui ka miinusena. Ühelt poolt oleks järelduste tegemine lihtsam, kui ei peaks seletuste puhul arvestama erinevate parameetrite mõju, kuid samas loomulikus keeles ja seletuste loomises ei pruugigi uuritavaid elemente n-ö puhtalt esineda. Leksikograafid kasutavad ikka erinevaid väljendusvahendeid koos – näiteks on ÜS-is seletatud *aktsiisimaksu* kui *kaudset riigimaksu*, *millega maksustatakse teatud valmistooteid* (*nt tubakatooteid, alkohoolseid jooke, vedelkütust*), milles esineb meie uuritavatest parameetritest umbisikulist tegumoodi, täpsustusi (*teatud*), näidete esitamist ja liigendatust (sulud ja osalause arv) – ning seetõttu peab ka analüüsis paratamatult arvestama erinevaid võimalikke tegureid, mis võivad seletuse mõistmist mõjutada. Järelduste täpsuse huvides võiks tulevikus küsitluse kaasata vaid selliseid seletusi, milles esineb ainult üks parameeter korraga.

Küsitluse oli kaasatud kolm sõnaliiki – nimisõnad, tegusõnad ja omadussõnad. Need on sisusõnad ehk täistähenduslikud sõnad, mis kannavad lauses kindlat tähendussisu ning mille puhul võiks eeldada, et kuna nende seletused on konkreetsemad, on ka seletustest arusaamine paremini katsetatav. Need kolm sõnaliiki on Keeleõppija Sõnaveebis kõige sagedamad, moodustades kokku 84% kõigist märksõnadest. Ka ÜS-is, mis sisaldab ligikaudu 180 000 eestikeelset märksõna, on need sõnaliigid kõige sagedamad: nimisõnu on ligikaudu 75% (sh pärisnimed 6%), omadussõnu 11% ja tegusõnu 8%. Eesti keele ühendkorpuses 2023 on sagedamad sõnaliigid nimisõnad, tegusõnad, määrsõnad, sidesõnad, asesõnad ja omadussõnad; kuid ühendkorpusest sõnaliikide sagedusi vaadates tuleb arvestada seda, et sõnaliigid on sõnadele määratud automaatselt ning ei pruugi alati olla õiged. Et teha põhjanevamaid järeldusi, saaks tulevikus keeleõppijail lasta hinnata ka teiste sõnaliikide seletusi. Samuti võiks kaaluda rohkemate tegusõnade ja määrsõnade kaasamist, kuna need on abstraktsemad ning neid on raskem seletada.

Jätšime küsitluse ka võimaluse oma eelistust mitte märkida, mõeldes olukordadele, kus mõlemad seletused võivad tunduda vastajale vastuvõetavad. Me ei tahtnud kunstlikult tekitada eelistust sinna, kus seda ei pruugi olla. Samas tekitas see vastusevariant ka ebamäärasust järelduste tegemisel, nt kui 47% vastajatest eelistasid kohandatut seletust, 36%

originaalsõnastuses seletust, kuid samas 16% polnud eelistust, ei saa väita, et pakutud kohandatud seletus oleks tingimata parem.

Küsitluse tulemused osutavad siiski mitmele tendentsile, mis mõjutavad eesti keele kui teise keele õppijate arusaamist sõnaseletustest.

Leksikaalsete parameetrite analüüsis ilmnes, et ase- ja sidesõnade kasutus seletustes raskendab tähenduse mõistmist. Õppijad eelistasid kompaktsemaid sõnastusi, kus info on esitatud loetelude või lühemate fraasidena ning kus välditakse sidesõnu ja sisutühje asesõnu. Samuti selgus, et õppijad eelistasid seletusi, millest on eemaldatud abstraktse tähendusega täpsustused, nagu *teatud*, *näiteks* ja *kindel*. Ilmnes ka keeleoskustasemest sõltuv erinevus võõrsõnade eelistamisel. A2–B1-tasemel õppijad kaldusid eelistama tuttavat ja rahvusvaheliselt läbipaistvat võõrsõna, samas kui kõrgema keeleoskustasemega õppijad eelistasid pigem omasõnalist ja semantilist täpsemat seletust. Rahvusvaheliselt läbipaistev tüvi pakub mõistmisel ilmselt tuge eriti madalama taseme keeleõppijatele, aga siinkohal on oluline rõhutada, et eesti keele kui teise keele õppijad on erineva emakeelega ning mitte kõigi jaoks ei pruugi võõrsõna olla arusaadavam. Nagu eespool ka mainitud, sünonüümide kaudu tähenduste seletamist leksikograafias üldiselt välditakse, pigem püütakse mõiste sisu avada täpsema kirjeldusega. Küsimus jääb, kas selline lähenemine algtaseme õppijaid silmas pidades on kõige otstarbekam, ent andes ainult ühesõnalise võõrsõnaga seletuse, läheneks õppesõnastik juba kakskeelsele sõnastikule.

Morfoloogiliste parameetrite analüüs näitas üllatuslikult, et umbisikulise tegumoe kasutus ei raskenda seletusest arusaamist. Siiski eelistasid umbisikulise tegumoe seletust suuremal määral kõrgema tasemega keeleõppijad. Arusaamist mõjutavad pigem muud tegurid, nagu seletuse ülesehitus või märksõnaga samatüvelised sõnad seletuses. Seevastu *des*-vormi sisaldav seletus, mis on sündmuse kompaktne esitamine teise sündmuse kontekstis, osutus õppijatele pigem keerukaks. *v*-kesksõna puhul oleks tarvilik kaaluda kohandatud sõnastuse muutmist ja küsitluse kordamist. Kui vaadata sõnaseletustes *v*-kesksõna sisaldavate fraaside keerukust, võib eeldada, et nende mõistmine valmistab eriti madalamatel keeleoskustasemetel raskusi. Siinsetest tulemustest ei selgu aga ei originaalsõnastuses seletuse ega kohandatud seletuse eelistus: valikuid on peaaegu pooleks. Ükski vastaja ei kasutanud ka võimalust selgitada oma eelistust. Seega võis vastaja langetada otsuse ühe või teise seletuse kasuks juhuslikult, pikemalt süvenemata, ja on võimalik, et vastaja tegi

küll valiku, aga selgusetuks jääb, kas mõlemad seletused olid ühtviisi arusaamatud või ühtviisi mõistetavad.

Süntaktiliste parameetrite analüüsi põhjal võib öelda, et täiendite rohkus põhisoona ees võib arusaamist takistada. Õppijad eelistasid seletusi, kus põhisoona paikneb seletuse alguses ja loob järgnevale infole kohe semantilise raamistiku. Samas ei osutunud määravaks mitte seletuse lühidus, vaid piisav ja asjakohane infohulk. Mitmel juhul eelistati pikemat seletust, kui see andis tähendusest selgema ja konkreetsema ülevaate, näiteks näidete abil. *kui...siis*-sõnaseletusmalli puhul tuleks samuti kaaluda kohandatud sõnastuse muutmist ja küsitluse kordamist. Kui vaadata *kui...siis*-sõnaseletusmalliga seletuste pikkust ja seda, et oluline info paikneb neis lõpus, võib eeldada, et nende mõistmine on keeruline just madalamatel keeleoskustasemetel. Kuna eelistused jagunesid pooleks ja kommentaare ei lisatud, võis ka siin olla vastaja valik juhuslik.

Küsitluse tulemusi oleme juba rakendanud olemasolevate seletuste kohandamisel Keeleõppija Sõnaveebis ning saja uue, B2-taseme märksõna seletuste koostamisel. Tulevikus on plaanis Keeleõppija Sõnaveebi märksõnastikku veelgi laiendada ning selleks appi võtta suured keelemudelid. Suuri keelemudeleid saaks promptida uute põhimõtete ja eeskujulike seletusnäidete abil ning sel viisil automaatselt luua leksikograafie seletuste mustandid, mida järeldoimetada. Keelemudelite pakutud mustandite kasutamine võimaldaks koostamist oluliselt kiirendada ja seeläbi suurendada lühema ajaga Keeleõppija Sõnaveebi märksõnastikku. Keelesammu projekti raames loodud andmestikku (uusi kohandatud sõnaseletusi) saabki edaspidi kasutada suurte keelemudelite sisendina automaatsete seletuste loomisel (vt ka Ide jt 2026).

6. Kokkuvõte

Andsime artiklis ülevaate küsitlusest, mille korraldasime 2025. aasta lõpus 55 eesti keele kui teise keele õppija seas. Küsitlusega tahtsime välja selgitada, millised leksikaalsed, morfoloogilised ja süntaktilised nähtused raskendavad keeleõppijal sõnaseletuste mõistmist. Selleks hindasid vastajad 20 märksõna kaht erinevalt sõnastatud seletust, mis pärinesid eesti keele kui teise keele õppijale suunatud Keeleõppija Sõnaveebist.

Küsitluse tulemused andsid esialgse ülevaate uuritavatest nähtustest. Analüüs näitas, et ase- ja sidesõnade liigne kasutus võib raskendada

seletuse mõistmist ja õppijad eelistavad pigem täpsustusteta seletusi, kus pole kasutatud nt sõnu *teatud, näiteks* ja *kindel*. Samuti ilmnes, et A2–B1-tasemel eelistatakse seletustes pigem tuttavaid võõrsõnu, samas kui kõrgema keeleoskustasemega õppijatel sellist eelistust ei täheldatud. Üllatuslikult ei raskendanud tulemuste põhjal umbisikulise tegumoe kasutamine seletustest arusaamist, küll aga *des*-vormide kasutus. Lisaks selgus, et täiendite rohkus põhisõna ees mõjutab seletusest arusaamist ning eelistatud on seletused, milles põhisõna paikneb esikohal. Oluliseks teguriks osutus ka info hulk: piisava info hulk on arusaadavuse seisukohalt määravam kui seletuse lühidus.

Küsitluse tulemuste põhjal uuendasime koostamis põhimõtteid, mille alusel saame kohandada juba olemasolevaid seletusi Keeleõppija Sõnaveebis ning luua seletusi uutele lisatavatele B2-taseme märksõnadele.

Kirjandus

- Atkins, Sue; Michael Rundell 2008.** The Oxford Guide to Practical Lexicography. Oxford: Oxford University Press.
- Barnbrook, Geoff 2002.** Defining Language: A Local Grammar of Definition Sentences. (= Studies in Corpus Linguistics 11.) John Benjamins Publishing Company.
- DeepL.** Tõlketarkvara. <https://www.deepl.com/en/translator>.
- Geeraerts, Dirk 2006.** Meaning and definition. – A Practical Guide to Lexicography. John Benjamins Publishing Company, 83–93.
- Glosbe.** Mitmekeelne veebisõnastik. <https://et.glosbe.com>.
- Google Translate.** Tõlketeenus. <https://translate.google.com>.
- Ide, Yusuke; Adam Nohejl; Joshua Tanner; Hitomi Yanaka; Christopher Lindsay; Taro Watanabe 2026.** Towards Automated Lexicography: Generating and Evaluating Definitions for Learner’s Dictionaries. arXiv preprint arXiv:2601.01842.
- Kallas, Jelena; Kristina Koppel; Raili Pool; Katrin Tsepelina; Tiiu Üksik; Pilvi Alp; Anu Epner 2021.** Eesti keele kui teise keele õpetaja tööriistad Eesti Keele Instituudi keeleportaalis Sõnaveeb. – Eesti Rakenduslingvistika Ühingu aastaraamat, 17, 61–80. <https://doi.org/10.5128/ERYa17.04>.
- Kallas, Jelena; Kristina Koppel; Katrin Tsepelina 2024.** The EKI Picture Dictionary. – XXI EURALEX International Congress: Lexicography and Semantics. Book of Abstracts: Cavtat, Croatia, 8–12 October 2024. Institut za hrvatski jezik, 108–112.
- Kallas, Jelena; Mai Tiits; Maria Tuulik; Kristina Koppel; Madis Jürviste 2014.** Eesti keele põhisõnavara sõnastik. Tallinn: Eesti Keele Sihtasutus.

- Kallas, Jelena; Maria Tuulik 2011.** Eesti keele põhisõnavara sõnastik: ajalooline kontekst ja koostamispõhimõtted. – Eesti Rakenduslingvistika Ühingu aastaraamat 7, 59–75. <https://doi.org/10.5128/ERYa7.04>.
- Keelekliki sõnastik.** www.keeleklikk.ee/en/dictionary.
- Keelesamm.** Eesti Keele Instituut. <https://eki.ee/keelesamm>.
- Keeleõppija Sõnaveeb.** Eesti Keele Instituut. <https://sonaveeb.ee/lite>.
- Koppel, Kristina 2019.** Eesti keele kui teise keele õpikute lausete analüüs ja selle rakendamine eri keeleoskustasemetel sõnastike näitelauseste automaatsel valikul. – Eesti Rakenduslingvistika Ühingu aastaraamat 15, 99–119. <https://doi.org/10.5128/ERYa15.06>.
- Koppel, Kristina 2020.** Näitelauseste korpuspõhine automaattuvastus eesti keele õppesõnastikele. (= *Dissertationes linguisticae Universitatis Tartuensis* 38.) Tartu: Tartu Ülikooli Kirjastus.
- Koppel, Kristina; Jelena Kallas 2021.** Keeleõppija Sõnaveeb eesti keele kui teise keele õppija abilisenä. – *Oma Keel* 43, 23–30.
- Koppel, Kristina; Jelena Kallas 2022.** Eesti keele ühendkorpusete sari 2013–2021: mahukaim eestikeelsete digitekstide kogum. – Eesti Rakenduslingvistika Ühingu aastaraamat 18, 207–228. <https://doi.org/10.5128/ERYa18.12>.
- Koppel, Kristina; Arvi Tavast; Margit Langemets; Jelena Kallas 2019.** Aggregating dictionaries into the language portal Sõnaveeb: Issues with and without a solution. – *Proceedings of the eLex 2019 conference*. 1–3 October, Sintra, Portugal. Brno: Lexical Computing CZ, s.r.o, 434–452.
- Langemets, Margit; Kristina Koppel; Jelena Kallas; Arvi Tavast 2021.** Sõnastikukogust keeleportaaliks. – *Keel ja Kirjandus* 8–9, 755–770. <https://doi.org/10.54013/kk764a6>.
- LimeSurvey.** Küsitlustarkvara. <https://www.limesurvey.org>.
- Piltsõnastik.** Eesti Keele Instituut. <https://sonaveeb.ee/wordgame>.
- PSV = Jelena Kallas; Mai Tiits; Maria Tuulik; Kristina Koppel; Madis Jürviste 2014.** Eesti keele põhisõnavara sõnastik. Tallinn: Eesti Keele Sihtasutus. <https://arhiiv.eki.ee/dict/psv>.
- Raamdokument 2007 = Euroopa keeleõppe raamdokument: õppimine, õpetamine, hindamine.** Tartu: Haridus- ja Teadusministeerium, 2007.
- Räägi minuga.** Facebook. <https://www.facebook.com/groups/raagiminuga>.
- Svensén, Bo 2009.** *A Handbook of Lexicography*. Cambridge University Press.
- Sõnaveeb.** Eesti Keele Instituut. <https://sonaveeb.ee>.
- Õpetaja tööriistad.** Eesti Keele Instituut. <https://sonaveeb.ee/teacher-tools/#/home>.
- ÜS = EKI ühendsõnastik.**

Learner evaluation of dictionary explanations in Learner's Sõnaveeb for Estonian as a second language

KRISTINA KOPPEL, LIINA LUTSEPP,
MARIA TUULIK

This article reports on a survey of 55 learners of Estonian as a second language that examined which lexical, morphological, and syntactic features hinder the comprehension of dictionary definitions. Participants evaluated two alternative definitions for 20 headwords from Learner's Sõnaveeb, one original and one adapted according to learner-oriented principles.

The results indicate a preference for compact definitions that avoid excessive use of conjunctions, semantically empty pronouns, and vague or abstract modifiers. Words such as *teatud* 'certain', *näiteks* 'for example', and *kindel* 'specific' were found to reduce clarity. Learners at A2–B1 levels preferred familiar loanwords, whereas no such preference was observed at higher proficiency levels. A morphological analysis showed that the impersonal voice does not significantly impede comprehension, while *-des* constructions are problematic. From a syntactic perspective, comprehension was facilitated when the headword appeared at the beginning of the definition, whereas dense premodification hindered understanding. Overall, providing sufficiently relevant information proved more important than minimizing definition length.

Based on these findings, the principles for compiling definitions in Learner's Sõnaveeb were revised and applied to both existing entries and new B2-level vocabulary.

Keywords: language learning, pedagogical lexicography, adaptation of definitions, proficiency levels, Estonian as a second language

Kristina Koppel
Eesti Keele Instituut
Roosikrantsi 6
10119 Tallinn
kristina.koppel@eki.ee

Liina Lutsepp
Tallinna Ülikool
humanitaarteaduste instituut
Narva mnt 25
10120 Tallinn
Eesti Keele Instituut
Roosikrantsi 6
10119 Tallinn
llutsepp@tlu.ee

Maria Tuulik
Eesti Keele Instituut
Roosikrantsi 6
10119 Tallinn
maria.tuulik@eki.ee