

SÕNAVARA JA KONSTRUKTSIOONID NOORE KEELEÕPPIJA KIRJALIKES TEKSTIDES

Reili Argus, Tiina Rüütmaa

Ülevaade. Artikli eesmärk on anda ülevaade eesti keelt teise keelena omandavate 9–10-aastaste keeleõppijate kirjalike tööde leksikaalsete vahendite ja grammatiliste konstruktsioonide arengust kuue kuu jook-sul (2023. a sügis – 2024. a kevad). Kokku 10 õppija sõnavara ja gram-matiliselt konstrueeritud tekstide kirjeldatakse sõnavara ulatuse, sõnaliigilise jaotumuse, eri keeleoskustasemete sõnavara ulatuse ja grammatiliste konstruktsioonide põhjal. Kirjalike tööde sõnavara ulatus kuue kuuga oluliselt ei muutunud, sõnavara arengut näitas tegusõnalekseeemide ja sidesõnade ning kõrgemate keeletasemete sõnavara lisandumine. Konstruktsioonide arengut ilmnestas peamiselt keele mitmekesisuse suurenemine, üksikud uued konstruktsioonitüübid olid kompleks-semad ja variatiivsemad. Vaadeldes kirjalike tekstide sõnavara ja konst-ruktioonide seoseid, selgus, et pikem tekst tähendas ka ulatuslikumat sõnavara, rohkem kõrgema keeletaseme sõnu ning rohkem eri tüüpi konstruktsioone. Samas ei olnud seos sõnade ja konstruktsioonide vahel üksühene: komplekssemad konstruktsioonid ei pruukinud sisaldada komplekssemaid sõnu ja keerukamaid sõnu ei kasutatud tingimata keerukates konstruktsioonides.*

Võtmesõnad: eesti keel teise keelena, keeleomandamine, keeletase, Guiraud' indeks, noor keeleõppija, eesti keel

1. Sissejuhatuseks

Seoses eestikeelsele haridusele üleminekuga on teravalt tõstatunud eesti keele kui teise keele (E2) omandamise alaste uuringute nappus. Puudub näiteks ülevaade, milline on E2 omandavate algkooliõpilaste eesti keele sõnavara ja grammatika areng algkooli lõpus ehk siis, kui minnakse klassiõpetajasüsteemilt üle eraldi ainetundidele.

* Artikli aluseks olevad andmed on kogutud ja analüüsitud Haridus- ja Teadusministeeriumi rahastatud projekti "Professionaalne eestikeelne õpetaja mitmekeelses klassis" toel.

Aruteludes sõnavara ja grammatika osatähtsuse üle keeleõppes on eri aegadel olnud eri seisukohtadel: rõhutatud on nii sõnade (Daller jt 2007, Juffs 2009, Rogers 2013) kui ka konstruktsioonide (vt Milton, Donzelli 2013: 442) olulisust, kuid jõutud ka arusaamani, et tähenduse väljendamiseks on vaja nii grammatilist kui ka leksikaalset infot (vt Bates, Goodman 1997). Enamasti on käsitletud inglise keele omandamist, soome-ugri keeltele omane vormimoodustuse rohkus võib tähendada grammatika suuremat rolli. Vaadeldes E2 sõnavara ja konstruktsioonide arengut noorte õppijate suulises keeles (Argus, Rüütmaa 2024) selgus, et sõnavara ja grammatika on omandamisel tihedalt seotud.

Teoreetiline raamistik on sõnavara ja grammatika omandamist käsitlevates uurimustes enamasti erinev, st kasutatakse ka erinevaid uurimismeetodeid. Sõnavara omandamist teises keeles on uuritud kõige sagedamini testide abil (vt nt Nation, Gu 2007), millega saab küll andmeid paljudelt informantidelt, kuid mis ei mõõda sõna kasutusoskust. Uuringuid, kus vaadeldaks seotult nii teise keele leksikaalsete kui ka grammatiliste üksuste arengut, st sama materjali põhjal ning sama või sarnase meetodiga, ei ole kuigi palju ei E2 (vt Allkivi-Metsoja 2021) ega ka teiste keelte kohta. Ometi tunnistavad mitmed sõnavara omandamise uurijad (Luchini, Ferreira 2023), et sõna kui leksikaalset elementi ei saa lahutada morfosüntaktilisest infost ehk sellest, kuidas (mis vormis/konstruktsioonis) sõna kasutada saab. Seotud on ka sõnavara ja konstruktsioonide hulk. E2 suulise materjali kohta on leitud, et suurema sõnavaraga laste tekstides on ka rohkem konstruktsioone, kuid kindla sõnaliigi lekseemide, nt verbide hulk otseselt verbikesksete ühendite hulgaga seotud ei ole. (Argus, Rüütmaa 2024) Seega ei pruugi sõnavara ja konstruktsioonide omandamise seosed olla lihtsad.

2. Sõnavara ja konstruktsioonide arengust õppijakeeles

Teise keele sõnavara omandamise uurimustes on enamasti fookuses inglise keele omandamine teise keelena, vaadeldud on nii sõnavara omandamise protsessi ja selle seotust õppija taustateguritega (nt González-Fernández 2024) kui ka sõnavara õppimise strateegiaid, nt sõna tähenduse äraarvamine konteksti põhjal või sõnaraamatu kasutamine (Ivanovska, Xhaferi 2023, Sun 2015). Sõnavara omandamisega seotud muid keelelisi tegureid, sh keele teiste tasandite, nt morfosüntaksi või ortograafia seotust sõnade õppimisega (Forsberg, Bartning 2010), on uuritud vähem.

Konstruktsioonide tasandit on sõnavara omandamisega seotud peamiselt meetodi kaudu, nt kasutatud sõnade tähenduse äraarvamist konteksti ehk konstruktsiooni abil (Mulder jt 2019). Sõnavara hulga ja konstruktsioonide kompleksuse seoste kohta leidub vaid üksikuid uuringuid, nt on prantsuse kui teise keele omandamise puhul (Lundell, Lindqvist 2012) vaadeldud leksikaalse rikkuse ja kollokatsioonide omandamise seoseid edasijõudnud keeleõppijail. Leiti, et kuigi mõne õppija kõne võib leksikaalselt rikkuselt olla samaväärne prantsuse keelt ema-keelena kõnelejate sõnavaraga, eristab teda viimastest kinnistunud struktuuride väiksem hulk (Lundell, Lindqvist 2012: 89). Eesti keele kohta puuduvad andmed, kui suure osa tekstist kinnistunud struktuurid moodustavad, ent rikka morfoloogia ja vabama sõnajärgiga eesti keeles on kinnistunud ühendeid arvatavasti vähem. Ka

näitavad uuringud, et kinnistunud ühendite kasutus kasvab keeletaseme tõustes, nende suurem osakaal tekstis on omane pigem kõrgematele keeletasemetele, ent jääb enamasti siiski madalamaks kui emakeelekõnelejal (Lundell, Lindquist 2012: 89). Rikka morfoloogiaga keelte ja madalamatel keeletasemetel õppijate konstruktsioone tuleks niisiis uurida pigem grammatikatasandi abstraktsemate üksustena, kuna lineaarsed kinnistunud sõnajärged ei pruugi näidata ei kompleksuse tegelikku kasvu ega grammatika kui süsteemi arengut. Siinses artiklis lähtutigi “Eesti keele süntaksi” (Erelt 2017: 54–56) terminoloogiast ning vaadeldi ühelt poolt lausete ja sõnade vahetervikuid ehk fraase, st üksusi, mis koosnevad tuumast (mistahes sõnaliigi sõna v.a finiiitne verbivorm) ja selle laienditest. Enamasti võib fraasituum toimida ka üksi fraasina, ent kuna uurimuses vaadeldi eelkõige õpilaste keele kompleksust, käsitleti siin fraasina ainult üksusi, mille tuumal on vähemalt üks laiend. Teisalt olid vaatluse all laused kui finiiitse verbivormi ehk lausetuuma ja selle juurde kuuluvate laiendite (fraaside) eri kombinatsioonid ning liitlausele omased seosed. Konstruktsioonide tasemekohasuse määramisel lähtuti nagu sõnavara puhulgi Sõnaveebi õpetaja tööriistade tasemekirjeldustest.

Eesti õppijakeele kohta tehtud uuringud (vt tabel 1) on analüüsinud sõnavara alates A2 kuni C1 tasemeni, ent konstruktsioonide omandamist on uuritud üpris vähe ja peamiselt kõrgematel tasemetel. Vähesed uuringud, kus on vaatluse all nii leksikaalsed kui ka grammatilised tunnused, on suures osas keskendunud õppija keeletaseme määramisele.

Tabel 1. Eesti keele kui esimese (E1) ja teise keele (E2) kohta tehtud uuringuid

Uurimus	E1/E2	Uurimisfookus
Kerge jt 2014	E1	Sõnavara omandamine tasemetööde põhjal
Alp jt 2013	E2	Sõnavara omandamine A2–C1 tasemel
Kitsnik 2018	E2	Leksikaalgrammatilise pädevuse arenemine, tasemeeksami kirjutusülesannete põhjal, B1–B2 tasemel
Kitsnik 2019	E2	Konstruktsioonide omandamine mängulise keeleõppe kaudu, taset ei ole määratud
Vajjala, Lõo 2014	E2	Õppija keeletaseme määramiseks olulised tunnused, A2–C1
Kitsnik 2017	E2	Tingiva kõneviisi verbikonstruktsioonide areng, modaaltegusõnade tarvituse B1–B2-tasemel
Allkivi 2016	E1, E2	Tegusõnade kasutusmallide kõrvutus, tasemeeksamite loovkirjutiste leksikaalse keerukuse, sõnaliikide ja -vormide kasutuse iseloomulikud tunnused, A2–C1 tasemel
Allkivi-Metsoja 2021	E2	Keeleoskustasemeid eristavad leksikaalsed ja morfoloogilised tunnused kirjalikes tekstides, C1
Argus, Rütmaa 2024	E2	Sõnavara ja konstruktsioonid eelA1–A2 taseme suulises pildikirjelduses

Varasemad uuringud pole niisiis kirjeldanud leksika ega grammatika omandamist noore eelA1- kuni A2-taseme eesti keele kui teise keele õppija kirjalikes tekstides.

3. Kirjalike tekstide analüüsi põhimõtted

Siinse artikli põhieesmärk on kirjeldada E2 omandava noore keeleõppija kirjalike tööde leksikaalseid vahendeid ja grammatilisi konstruktsioone. Kvalitatiivne uurimus näitab, kuidas õppija sõnavara ja konstruktsioonid kuue kuu jooksul arenevad, st varem omandatud konstruktsioonid muutuvad komplekssemaks ja lisandub uusi. Teeme kindlaks:

- 1) milline on noore õppija kirjalike tööde sõnavara ulatus;
- 2) mis iseloomustab õppija sõnavara sõnaliigilist jaotumust;
- 3) mis taseme sõnavara sisaldavad õppijate tekstid;
- 4) mis tüüpi ja kui palju grammatilisi konstruktsioone õppijate tekstides sisaldub;
- 5) mis keeletasemele õppijate kasutatud konstruktsioonid viitavad;
- 6) millised muutused on kuue kuu jooksul sõnavaras ja eri konstruktsioonide kasutamises toimunud;
- 7) kuidas on seotud konstruktsioonide ja lekseemide komplekssus ehk mis taseme sõnavaraga konstruktsioonitüüpe täidetakse.

Õppijakeele sõnavara uurimiseks on kasutatud nii sekkumisuuringuid (vt nt Hjetland jt 2023), sõnavarateste (vt Puteri 2022) kui ka päevikuvormis enesehindamist (nt Luchini, Ferreira 2023). Vaadeldud on nii passiivset kui ka aktiivset sõnavara (Laufer 1998). Sõnavara hindamisel on rakendatud eri mõõdikuid, kusjuures nii mõnegi põhilise mõõdiku puhul võib leida osaliselt kattuvaid lähenemisi. Näiteks on mõõdetud mõnevõrra erinevalt sõnavara rikkust (ingl *richness*), selle leksikaalset varieeruvust (*lexical variation*) ja mitmekesisust (*lexical diversity*), olgugi et neis uuringuis arvestatakse kaht põhimuutujat, tekstisõnade ehk sõnede ja eri sõnade ehk lekseemide arvu (vt Yu 2009: 238–241).

Siinses uuringus on sõnavara analüüsiks kasutatud:

- 1) Kerge jt (2014) eeskujul väga levinud Guiraud' indeksi G (Guiraud 1960), mis sobib hästi just lühemate tekstide analüüsimiseks ja iseloomustab sõnavara üldist varieeruvust sõnade arvu suhte järgi sõnearvu ruutjuuresse: $G = \frac{\text{sõnade arv}}{\sqrt{\text{sõnede arv}}}$ jagatud ruutjuurega sõnede arvust;
- 2) sõnavara jaotumist eri sõnaliikide vahel (eesti grammatikate traditsiooniline jaotus nimi-, tegu-, omadus-, arv- ja asesõnadeks ning määrsõnad, kaasõnad, partiklid (arvestatud kokku) ja sidesõnad; eraldi kategooria "muu" alla on liigitatud lühendid ja numbrid ning arusaamatud sõnalaadsed üksused. Kui sõna sisaldas õigekirjaviga, kuid tähendus oli arusaadav (nt *vigusid pro vigu*), võeti see lekseemina arvesse;
- 3) sõnavara jaotumist keeleoskustasemeti vaadeldi Sõnaveebi õpetaja tööriistade¹ sõnaloendite põhjal.

Kirjalike tekstide grammatiliste konstruktsioonide puhul kasutati sama meetodit kui suulise kõne analüüsis (Argus, Rüütmaa 2024): töötati läbi laste 2023. ja 2024. a kirjalikud tööd ning leiti neis eri konstruktsioonitüübid. Eraldi vaadeldi fraasi- ja lausetasandit. Fraasitasandil leiti esmalt kirjalikes tekstides esinenud fraasiliigid (edaspidi FL, nt NP *tema ema*, PP *isa all*)². Kuna analüüsi üks eesmärke oli selgitada välja õpilaste keele variatiivsus, jagati fraasiliigid omakorda fraasitüüpideks, st vaadeldi eraldi fraasituuma ja selle laiendite eri

¹ <https://sonaveeb.ee/teacher-tools/#/home> (3.10.2024).

² Lühendid artikli lõpus.

kombinatsioone (edaspidi FT, nt hulgasõnafraasi tüübid Q+N *palju lapsi* või Nu+N *kaks laeva*). Lausetasandil aga leiti lausetuuma (finiitverbi) ja selle laiendite eri kombinatsioonid, (edaspidi VL, nt SVA *Mina olin kodus*). VL kombinatsioonidena käsitleti finiitverbi ja vähemalt selle seotud laiend(eid) (nt SV *Puu kasvab*, SVO *Isa loeb raamatut*). Samuti leiti tekstidest liitlausele omased seosed (rind- ja põimlause) – lausepiirid on tekstides tihti ebamäärased, mistõttu keskendutigi pigem seosetüüpidele (vt täpsemalt 5.2).

Konstruktioonide analüüsiga selgitatakse välja, milliseid konstruktsioone õpilased oma kirjalikes töödes kasutavad; kas on konstruktsioone, mida kasutavad kõik või enamik õpilasi; kas vaadeldud perioodi jooksul lisandub keelerepertuaari uusi konstruktsioone ning kui, siis missuguseid.

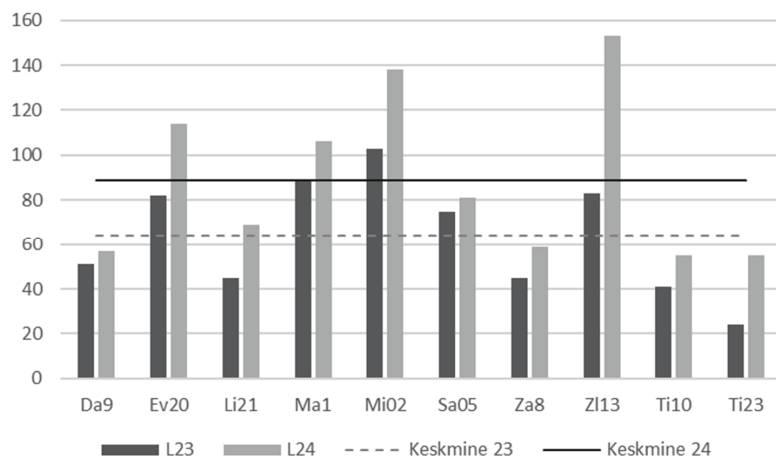
Sõnavara ja konstruktsioonide vaheliste seoste leidmiseks kõrvutati andmeid tekstide üldise pikkuse (sõnedes), lekseemide arvu, A-keelelasemest kõrgemale kuuluvate lekseemide hulga ja eri konstruktsioonitüüpide arvu kohta. Lisaks analüüsiti eraldi konstruktsioonide sõnavara õpilastel, kelle sõnavara ja konstruktsioonide hulk oli keskmisest suurem.

Valimi moodustasid programmis “Professionaalne eestikeelne õpetaja mitmekeelses klassis” alates 2020. aastast E2 õppinud 9–10-aastased õpilased (vt Argus jt 2024). Valimis on nii lapsi, kes projekti alguses eesti keelt veel ei rääkinud, kui ka neid, kes oskasid iseseisvalt lauseid moodustada. Analüüsi ajaks ehk projekti neljandaks õppeaastaks oli õpilaste eesti keele oskuse tase olulisel määral ühtlustunud ja erinev stardipositsioon enam väga palju mõju ei avalda. Kokku kümnekt Tallinna ja Ida-Virumaa koolide õpilaselt koguti 2023. a novembris kaks kirjalikku ülesannet: kiri sõbrale, pereliikmele või jõuluvanale ja (ruumi või pildi) kirjeldus. Vahepealsel perioodil ehk kuue kuu jooksul läbisid õpetajad koolitustsükli kirjaliku väljendusoskuse arendamisest ning said võimaluse omandatut rakendada (kas ja kuidas seda tehti, ei jälgitud, see eeldanuks mahukat sekkumisuuringut). Sama õppeaasta kevadel (aprillis 2024) koguti samadelt õpilastelt uuesti kaks tööd: pildiseeria kirjeldus ja arvamus. Kirjutiste näol oli tegemist õpilaste vaba iseseisva tekstiloomega (klassiruumis kirjutatud tööd), mis näitab mitte ainult üksiku sõna tähenduse tundmist, vaid ka oskust seda tekstis kasutada. Tekstide žanriomased jooned kirjutistes kuigi tugevalt ei esildunud: näiteks arvamused meenutasid temaatikalt üsna olulisel määral kirju (lapsed kirjeldasid, mis neile meeldib ja mis mitte). Kuna uurimuse eesmärk on anda ülevaade laste sõnavara ja konstruktsioonide üldisest arengust ja konkreetse teksti või žanri tasandil järeldusi ei tehta, ühendati mõlemal aastal saadud kirjalik materjal ehk nt 2023. a tööde puhul analüüsiti kirjeldust ja kirja koos ühe tekstina.

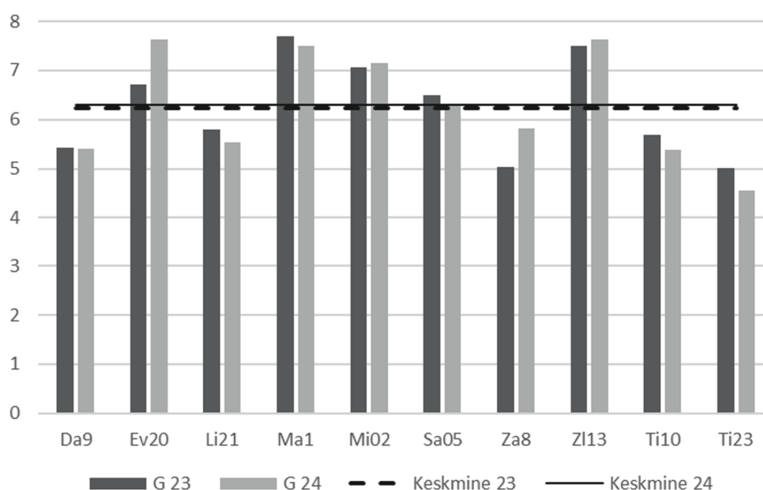
4. Keeleõppijate kirjalike tööde sõnavara

4.1. Sõnavara üldine ulatus

Võrreldes 2023. a sügise tekstidega, mis olid keskmiselt 114 sõnet pikad, olid 2024. a kevade tööd pikemad, keskmiselt 189 sõnet. Ka lekseemide keskmine arv on kevadistes töödes suurem (vt joonis 1), sealjuures on aga sõnavara üldine varieeruvus jäänud peaaegu samaks (vt joonis 2).



Joonis 1. Lekseemide arv (L) lapseti 2023 ja 2024 ning mõlema aasta keskmine



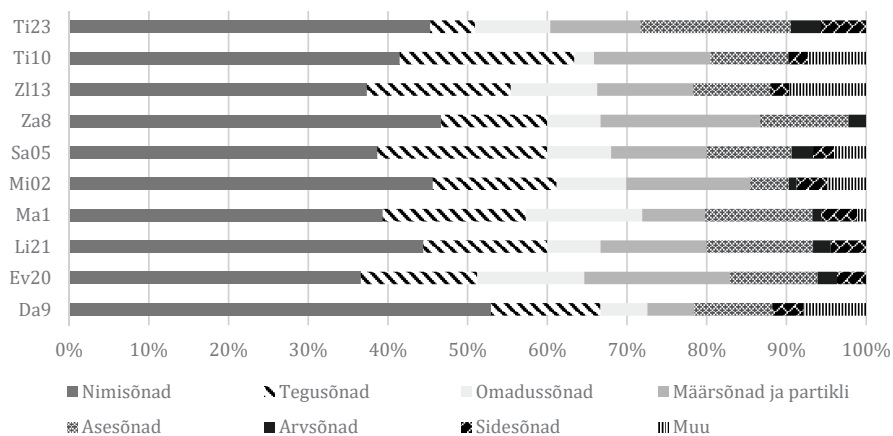
Joonis 2. Sõnavara üldine varieeruvus (G) lapseti 2023 ja 2024 ning mõlema aasta keskmine

Kirjalike tööde lekseemide arv varieerus oluliselt: 2023. a sügisel oli kõige vähem ühe õppija töös 24 eri lekseemi ja kõige rohkem 103 lekseemi, kuus kuud hiljem oli kõige vähem (55 lekseemi) kahe lapse töös ja kõige rohkem, 138, ühe lapse töös.

Sõnavara üldise varieeruvuse Guiraud' indeks jäi sügisel vahemikku 5,01–7,69 ning kevadel vahemikku 4,55–7,64. Seega moodustas kõige vähem varieeruva sõnavaraga kirjutaja teksti sõnavara sügisel 65% ja kevadel 59% kõige rikkalikuma sõnavaraga kirjutaja kasutatud sõnadest ehk vahe eri kirjutajate vahel oli muutunud suuremaks. Samas ei näita Guiraud' indeks kuigi hästi laste sõnavara arengut, sest laste keskmine indeks on jäänud peaaegu samaks ehk laste teksti sõnavara ei ole muutunud varieeruvamaks. Seega, suurem sõnade arv ehk tervikuna pikem tekst justkui tasandab suurema hulga lekseemide kasutuse.

4.2. Sõnavara sõnaliigiti

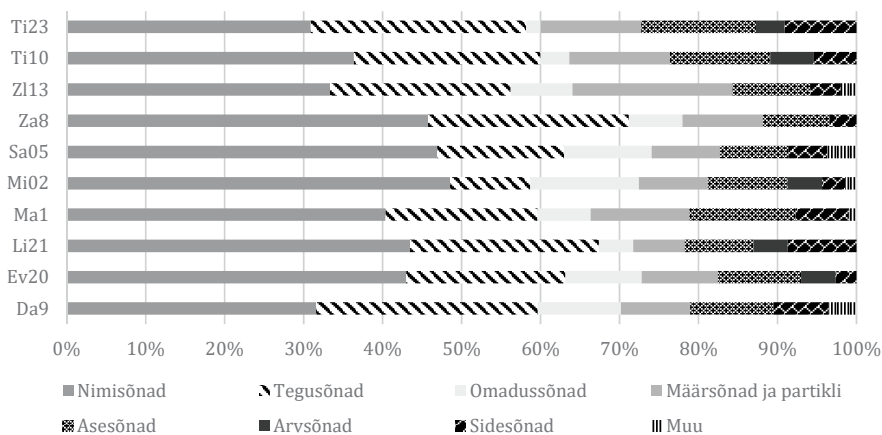
Õpilaste kirjalikes töödes kasutatud lekseemidest moodustasid kõige suurema osa (36–52%) nimisõnad (vt joonis 3).



Joonis 3. Sõnavara (lekseemide) jaotumine sõnaliigiti 2023. a sügise töödes

Keskmiselt sisaldas ühe õpilase tekst 28 nimisõna, 11 tegusõna, 6 omadussõna, 9 määr- või kaassõna või partiklit, 7 asesõna ja 2 sidesõna.

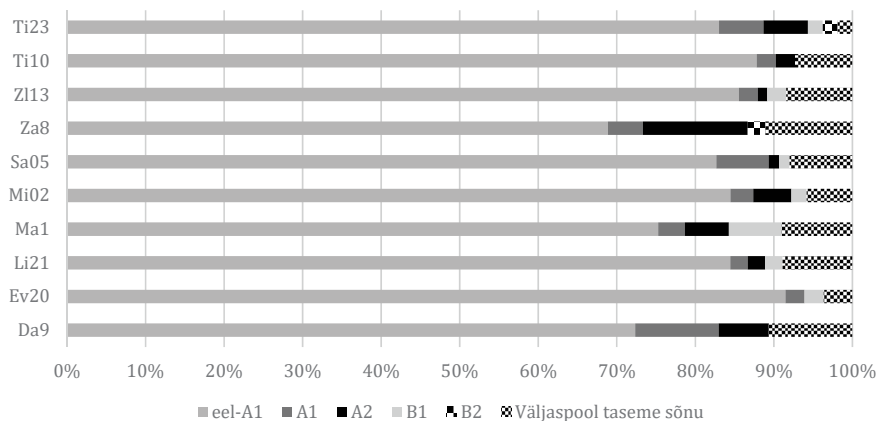
2024. kevade töödes oli peaaegu võrdselt suurenenud kõigi sõnaliikide sõnade arv (vt joonis 4), teiste sõnaliikidega võrreldes on enam kasvanud tegu- ja sidesõnade hulk. Keskmiselt kasutas üks õpilane 34 nimisõna, 17 tegusõna, 7 omadussõna, 10 määr-, kaassõna või partiklit, 9 asesõna ja 4 sidesõna.



Joonis 4. Sõnavara (lekseemide) jaotumine sõnaliigiti 2024. a kevade töödes

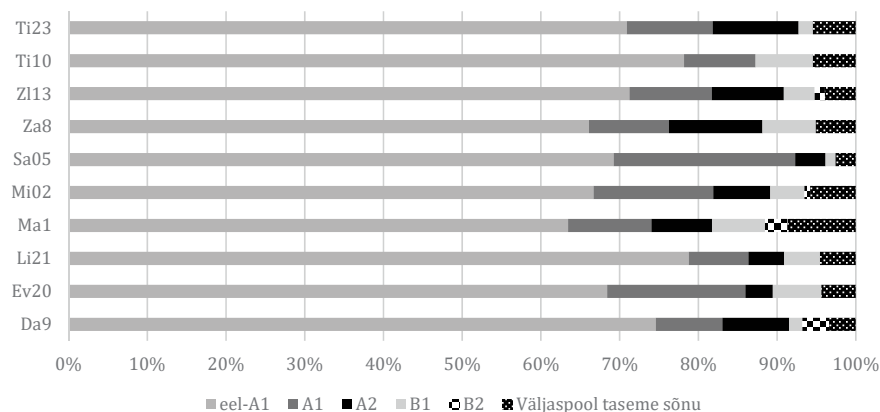
4.3. Sõnavara jaotumine keeleoskustasemeti

2023. a sügise kirjutiste sõnavarast moodustavad peamise osa (vt joonis 5) eelA1-taseme sõnad, mida on kokku keskmiselt 83% kõikidest lekseemidest. Nii A1- kui ka A2-taseme sõnu on keskmiselt 4%, B1-taseme sõnu 2% ning 7% sõnadest jäävad väljapoole tasemekirjeldusi (enamasti pärisnimed, nt *Egiptus*, *Uno*, aga ka liitsõnad, mida ei ole esitatud sõnaloendites, nt *juuksekuum*, *sügisvaheae*g).



Joonis 5. Sõnavara jaotumine keeleoskustasemeti 2023. a sügise kirjalikes töödes

Kaht tekstikogumit võrreldes on näha, et kuue kuuga on lisandunud A1-taseme ja mõnevõrra ka A-2 taseme sõnu. Eel-A1 taseme sõnu oli 2024. a kevade töödes kokku keskmiselt 69%, 13% oli A1-taseme sõnu, 7% sõnadest moodustasid A2-taseme sõnad, B1-taseme sõnu oli 5%, B-2 sõnu 1% ning väljapoole taset jäävaid sõnu 5%. 2024. a moodustasid sõnadest, mis on väljaspool tasemekirjelduste loendeid, suurema osa liitsõnad (*poetöötaja*, *kihelkonnamuuseum*, *tõsielusari*, *kevadtööd*, *lillepõõsad*), vähem leidis tekstides pärisnimesid (nt *Iisaku*).



Joonis 6. Sõnavara jaotumine keeleoskustasemeti 2024. a kevade kirjalikes töödes

Sõnad, mis jäid väljapoole tasemekirjeldusi, olid 2023. a sügisestes töödes enamasti pärisnimed, üldjuhul kas isikunimed (nt *Trofim*, *Vivaldi*), kohanimed (nt *Pariis*) või kaubamärkinimed (nt *aifon*, *Volvo*, *Ford*, *X-box*), kevadistes töödes aga pigem tuletatud omadus- või määrsõnad (*neljakorruseline*, *viiekorruseline*, *normaalselt*, *vihastunud*). Erinevus võis tuleneda mingil määral ka kirjutiste teemast: kui sügisel kirjeldati vaheaja tegevusi, mis sisaldasid rohkelt pärisnimesid, siis kevadiste kirjade teistsuguse teemaringiga tekstides oli pärisnimesid vähem. Liitsõnatuletised võivad aga viidata ka sõnamoodustusprotsesside omandamise algusele.

5. Grammatilised konstruktsioonid

Grammatiliste konstruktsioonide analüüsiks töötati esmalt läbi laste kirjalikud tööd leidmaks neis esinenud fraasid ning verbi ja selle laiendite (VL) eri kombinatsioonid. Seejärel jagati iga lapse konstruktsioonid fraasiliigi (FL), -tüübi (FT) ja VL kombinatsiooni (lihtlause) kaupa tabelisse.

Laste fraasiliikide, -tüüpide ja verbi-laiendi(te) kombinatsioonidest 2023. ja 2024. a kirjalikes töödes annab ülevaate tabel 2; hall taust näitab kasvu, st konstruktsioonitüüpe on 2024. a-l rohkem kui 2023. Näeme, et neljal lapsel on lisandunud kõiki kolme vaadeldud konstruktsioonitüüpi, kolmel kaht ja kahel üht, ainult ühel õpilasel pole kasvu näha. Sõnaveebi³ järgi kasutab eelA1-taseme õppija maksimaalselt 4 fraasiliiki; A1-taseme õppijatel lisandub küll ka määrsõnafraas, ent väga piiratud (ainult laienditega *väga* ja *ni*), ka A2-tasemel piirdub määrsõnafraasi kasutus peamiselt nendesamade ühenditega, seega õppijalt Zl13 fraasiliikide kasvu eeldada ei saagi. Täpsemalt tema arengusse süvenedes näeme siiski, et 2024. a on tal lisandunud määrsõnafraas (ta on ainus valimi õpilane, kes 2024. a määrsõnafraasi kasutas), ent 2024. a töödes pole ta kasutanud ühtki hulgasõnafraasi, mida esines tema 2023. a tekstides. Kuna hulgasõnafraasi kasutusvajadus sõltub suuresti teksti-tüübist ja eeldame, et kord juba omandatud fraasiliik on püsiv, võime arvestada ka Zl13 õpilaste hulka, kel on arengut kõigis konstruktsioonitüüpides.

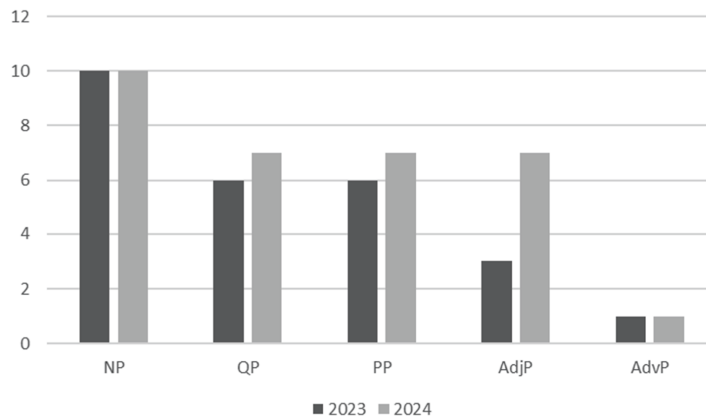
Tabel 2. Konstruktsioonide areng 2023. ja 2024. a kirjalikes töödes lapseti (hall taust näitab kasvu)

Lapse kood	FL23	FL24	FT23	FT24	VL23	VL24
Da9	1	1	3	3	4	6
Ev20	3	4	9	10	3	5
Li21	2	3	3	5	1	3
Ma1	3	4	6	6	6	9
Mi02	3	4	8	8	4	5
Sa05	4	3	7	6	6	5
Za8	1	3	2	4	4	5
Zl13	4	4	7	9	5	8
Ti10	2	3	3	5	3	5
Ti23	3	3	7	6	4	8

³ <https://sonaveeb.ee/teacher-tools/#/grammar?page=1> (3.10.2024).

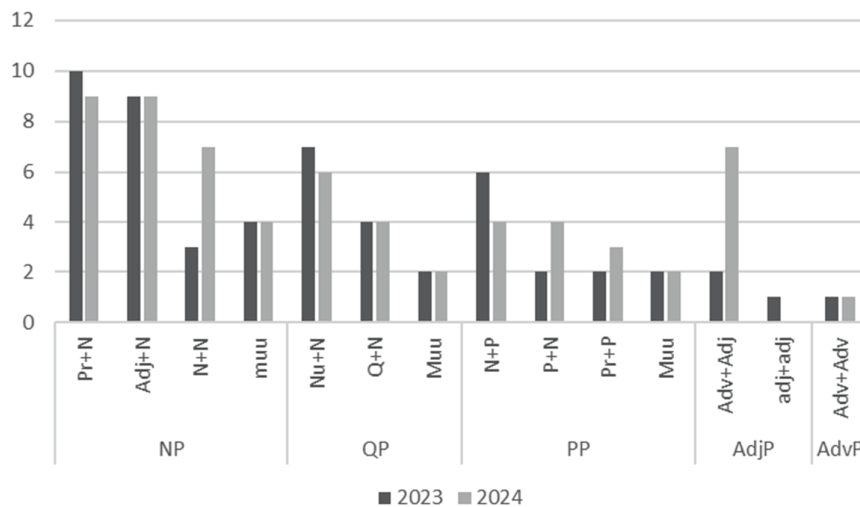
5.1. Fraasitasandi konstruktsioonid

Fraasiliikidest kasutasid õpilased enim nimisõnafraasi (NP): juba 2023. a kirjalikes tekstides esines seda kõigis töödes. 2023. a kasutas hulgasõna- ja kaassõnafraasi kumbagi kuus last, 2024. a oli mõlema kasutajaid ühe võrra rohkem. Kõige suuremat kasvu on näha omadussõnafraasi kasutajate arvus: kui 2023. a oli neid vaid 3, siis 2024. a on selle kasutajaid enam kui kaks korda rohkem, 7 last (vt joonis 7). Osalt võib omadussõnafraasi kasutuse suur kasv olla tingitud teksti žanrist: 2024. a kirjutati arvamussõna, mis motiveerib omadusi hindama, st omadussõnafraase kasutama. 2024. a tekste žanriti võrreldes näeme, et arvamustekstis kasutab omadussõnafraasi kuus, pildiseeria kirjelduses aga ainult kaks õpilast. Määrsõnafraasi kasutus on nii 2023. kui ka 2024. a tekstides pigem juhuslik: seda kasutab kummalgi aastal vaid üks laps.



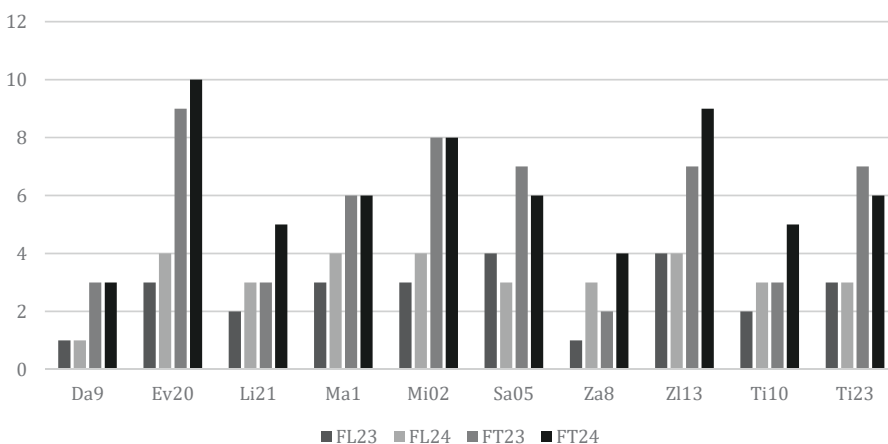
Joonis 7. Eri fraasiliike 2023. ja 2024. a kirjalikes töödes kasutanud laste arv

Fraasitüüpe vaadeldes saab kirjalike tekstide keele mitmekesisusest parema ülevaate kui piiratud arvuga fraasiliikidest lähtudes. Joonis 8 annab ülevaate õpilaste kasutatud fraasiliikide jagunemisest fraasitüüpideks ning sellest, kui palju õpilasi mingit fraasitüüpi kasutas; kategooria “muu” sisaldab enamasti rohkem kui ühe laiendiga fraasitüüpe, mida kasutab vaid 1–2 last ühes tekstis, nt Adj+Adj+N *suur sinine pall*, Q+Adj+N *vähe head toitu*. Ühe laiendiga fraasitüüpe võrreldes 2023. aastaga lisandunud ei ole, neid on isegi ühe võrra vähem (2023 on üks õpilane kasutanud omadussõnafraasi tüüpi Adj+Adj, nt *ilus punane*), küll aga on kasvanud mitmete fraasitüüpide kasutajate arv. Kõige suurem on kasv fraasitüüpides Adv+Adj, nt *väga ilus*, ja N+N, nt *isa tool*. Graafikutel on näha, et fraasitüüpe, millel on vaid 1–2 kasutajat, on 2024. a ainult üks, st valimi keelekasutus on muutunud üldvaates mitmekesisemaks, rohkem lapsi kasutab mitmeid eri fraasitüüpe.



Joonis 8. Fraasitüpe 2023. ja 2024. a kirjalikes töödes kasutanud laste arv

Arengust fraasiliikide ja -tüüpide kasutuses 2023. ja 2024. a annab ülevaate joonis 9: enamikul õpilastest on näha arengut vähemalt ühes fraasitüübis, samale tasemele on jäänud õpilane koodiga Da0, aastast 2023 vähem fraasitasandi konstruktsioone kasutab õpilane Sa05, kelle konstruktsioonide arv on 2023. aastal üks valimi suuremaid (4 FL ja 7 FT), sh on ta üks kahest, kel on 2023. a nn maksimumarv fraasiliike.



Joonis 9. Fraasiliigid (FL) ja -tüübid (FT) lapseti 2023. ja 2024. a töödes

5.2. Lausetasand

Lausetasandil vaadeldi eraldi liht- ja liitlauseid; lihtlausetest eraldi koondlauseid, liitlauseid jagati rind- ja põimlauseteks. Analüüsi raskendas tõsiasi, et enamik õpilasi pole veel omandanud lause vormistamise reegleid, st suures osas tekstidest pole laused lausealguse suure tähe ja lauselõpu punkti järgi eristatavad. Seega ei viita

liitlausete kasutamine tingimata keerukamale tekstile, vaid võib ka tähendada, et nt verbikeskseid üksusi ühendav *ja* on pigem kõnekeelele omane mõttekoht. Kui laused on visuaalselt eristatavad (lause lõpus punkt, enamasti ka suur algustäht alguses), on lause määramisel üldjuhul aluseks võetud nende vormistus. Enamasti aga tuli lähtuda lause kui mõttelise terviku määratlusest, või kombineerida lause visuaalseid ja sisulisi tunnuseid. Niisiis sai lausete liigitamisel leida pigem teatud tüüpi struktuure ja seoseid (koondlause struktuur, rind- või põimlausele omased seosed). Kuna vaadeldi pigem struktuuride omandamist, võeti lausete määramisel arvesse ka liitlause osalause esinenud koondlauseid; kui ühes liitses liitlause oli nii rind- kui põimlausele omaseid seoseid, läksid need eraldi arvesse. Tekstide lauselist koosseisu mõjutas sageli ka õpetaja juhendamine (või selle puudumine), nt mõne klassi õpilased alustavad ja lõpetavad oma tekste samade või sarnaste lausetega, teise klassi omad jälle kirjutavad vaid üks-kaks lauset.

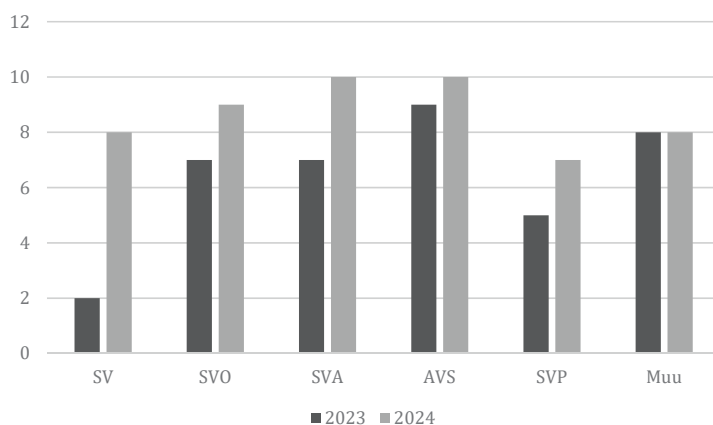
Tekstide lauseline koosseis on väga erinev, vt tabel 3. Lihtlauseid on ühe õpilase tekstides minimaalselt 2 (2024. a) ja maksimaalselt 17 (2023. a). Koondlauseid võivad tekstidest hoopiski puududa, kõige rohkem on neid ühe õpilase tekstides 9 (2024. a). Ka rindlauseid ei pruugi alati tekstis leida (2023. a oli neid vaid kahel lapsel), maksimaalselt on neid lapse kohta 14 (2024. a). Põimlauseid võivad samuti tekstist hoopis puududa (2023. a kolmel õpilasel), maksimaalselt on neid 19 (2024. a). Tabelis 3 on halliga märgitud arengukohad, st õpilase tekstis on vaadeldavat konstruktsiooni 2024. a tekstides enam kui varasemates. Peaaegu kõigil valimi õpilastel on 2024. a-l rohkem nii rind- kui põimlauseid (ühel õpilasel ainult põimlauseid). Kasv pigem liitlausest kui keerulisemates konstruktsioonides viitab tekstide komplitseeritumaks muutumisele. Sellega kaasneb enamasti loogiliselt ka lihtlausete kasutuse vähenemine (lihtlauseid seotakse liitlauseteks), ent on ka neid, kel on rohkem kõigi vaadeldud lauseliikide lauseid (Ev20) või on lisaks liitlausele suurenenud ka koondlausete arv (Mio2, Sa05 ja Ti23).

Tabel 3. Õpilaste kasutatud liht- ja liitlausete arv. Lihtlausest on eraldi näidatud koondlauseid (KL), liitlauseid on jagatud rind- (RL) ja põimlauseteks (PL)

Lapse kood	Lihtlause				Liitlause			
	LL 23	LL 24	KL 23	KL 24	RL 23	RL 24	PL 23	PL 24
Da9	8	5	6	0	0	7	0	7
Ev20	5	8	4	5	3	7	4	6
Li21	6	2	4	2	0	1	0	5
Ma1	6	4	3	2	5	8	4	9
Mio2	14	7	7	9	3	14	4	10
Sa05	17	8	0	6	3	7	1	3
Za8	9	14	3	1	1	1	2	4
ZI13	9	7	3	1	3	12	2	19
Ti10	5	4	2	1	1	3	0	6
Ti23	6	3	2	3	2	3	3	8

Lihtlausete puhul vaadeldi öeldise ja selle laiendite eri kombinatsioonide kasutust (vt joonis 10 ja 11). Eraldi on toodud välja enamlevinud kombinatsioonid: SV, nt *lumi sulas ära*; SVO, nt *Naine looeb oma lemik raamat*; SVA, nt *lapsed tulevad koju*; AVS, nt *aia taga on mets* ja SVP, nt *ta oli suur*. Kuna eesmärk on vaadelda konstruktsiooni omandatust, võeti konstruktsioonide leidmisel arvesse ka lauseid, milles ühe lauseliikme funktsiooni täidab kõrvallause, nt *Laps vaatab, et kedagi ei ole* (SVO). Enim kasutati niisiis ühe või kahe laiendiga konstruktsioone, ent esines ka kolme ja nelja laiendiga konstruktsioone. Kuna nende varieeruvus oli suur – üht kombinatsiooni kasutas maksimaalselt kolm õpilast – on need koondatud kategooriasse “muu” (vt joonis 8) nt *SVAA keegi istus pingil üks*; *SVOA, poiss mängis palli tüdrukuga*. Joonisel 10 kajastub kategooria “muu” konstruktsioone kasutanud laste arv, ent kui üks laps kasutas mitut kategooria kombinatsiooni, siis seda märgitud pole.

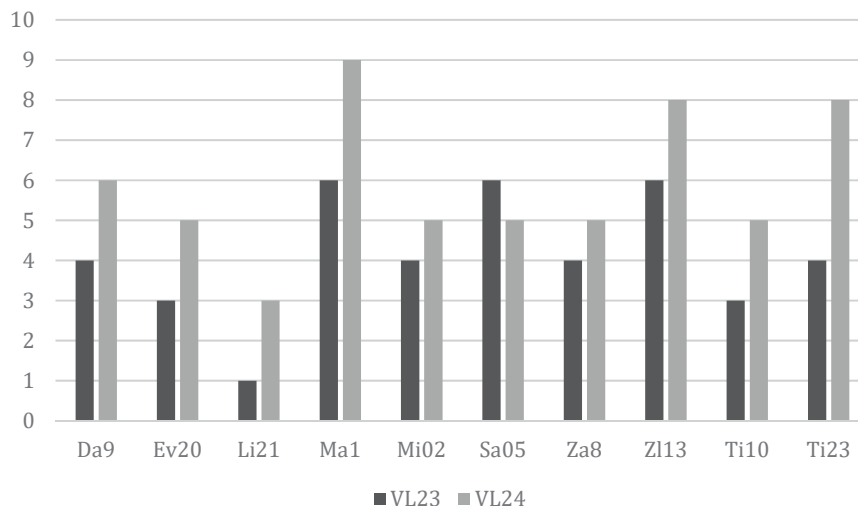
Joonis 10 annab ülevaate enim esinenud verbi ja laiendi (VL) kombinatsioonidest ja neid kasutanud õpilaste arvust. Nii 2023. kui ka 2024. a töödes kasutati enim kombinatsioone AVS (vastavalt 9 ja 10), nt *Pargi juures on palju majju*, veidi vähem kombinatsioone SVO (7 ja 9), nt *Kana sööb putukaid* ja SVA (7 ja 10), nt *Tegevus toimub pargis*. Kõige suurem on kasv kombinatsioonis SV (nt *Puu kasvab*), kahelt õpilaselt kaheksale. See kasv tuleb peamiselt pildiseeriast (pildiseerias kasutab kombinatsiooni 7, arvamuses 2 õpilast), mis ilmselt on ajendanud nimetama piltidel toimuvaid tegevusi. 2024. a tekstides kasutasid kõik valimi õpilased kombinatsioone SVA ja AVS, kombinatsioonil SVO ja SV oli vastavalt üks ja kaks kasutajat vähem. Peamiste VL kombinatsioonide hulgas niisiis muutusi pole, mis on ootuspärane ka Sõnaveebi õpetaja tööriistade grammatikapädevuse hindamisvahendi⁴ järgi: juba A1-tasemel esinevad seal kõik joonisel 8 toodud kombinatsioonid; juurde tuleb siiski üksikuid enam kui kahe laiendiga VL konstruktsioone. Kasvanud on kõigi VL kombinatsioonide kasutajate arv, st valimis on enam õpilasi, kelle tekstis on mitmekesised VL kombinatsioonid. Õpilaste lauseehitus on niisiis üldvaates muutunud mitmekesisemaks. Edasi vaadeldaksegi VL kombinatsioonide kasutust lapseti.



Joonis 10. VL kombinatsioonid 2023. ja 2024. a töödes

⁴ <https://sonaveeb.ee/teacher-tools/#/grammar?page=1> (3.10.2024).

Enamikul õpilastest (üheksal kümnest) on 2024. aastal enam VL kombinatsioone kui 2023. aastal, vt joonis 11. Ainsana on 2024. aasta tekstides vähem VL kombinatsioone õpilasel koodiga Sa05, kes kasutas 2023. aastal kuut, 2024. aastal aga viit kombinatsiooni. Üks VL kombinatsioon vähem võib olla ka juhuslik, ent tähelepanuväärne on, et samal õpilasel on 2024. aastal vähenenud ka nii fraasiliikide kui ka -tüüpide arv. Selle languse põhjusi selle artikli raames analüüsida pole võimalik, see eeldaks teistlaadi andmeid, kui siinse artikli kirjutajate käsutuses on.



Joonis 11. VL kombinatsioonid 2023. ja 2024. a lapseti

6. Sõnavara ja konstruktsioonide seosed

Kõrvutades sõnavara ja konstruktsioonide kolme põhinäitajat, on selgelt näha, et teksti pikkus, sõnavara hulk ja konstruktsioonitüüpide hulk on omavahel seotud (vt tabel 4). Tabeli andmete korrelatsioonianalüüs näitab, et kõige tugevam statistiliselt oluline positiivne korrelatsioon on lekseemide ja sõnade arvu vahel ($r = 0,94$ ja $p = 0,001$), samuti on tugev statistiliselt oluline korrelatsioon ka lekseemide ja sõnade arvu ning liitlausele omaste seoste arvu vahel (vastavalt $r = 0,87$, $p = 0,001$; $r = 0,94$, $p = 0,001$); statistiliselt oluline positiivne seos on fraasitüüpide ja lekseemide arvu ($r = 0,83$, $p = 0,003$) ning fraasitüüpide ja sõnade arvu vahel ($r = 0,77$, $p = 0,009$); A2+ taseme sõnavaral on tugev oluline korrelatsioon nii lekseemide ($r = 0,78$, $p = 0,008$) kui ka sõnade arvuga ($r = 0,67$, $p = 0,033$, aga ka liitlausele omaste seoste arvuga ($r = 0,66$, $P = 0,036$). Niisiis on kõige tugevamas seoses sõnavaranäitajad (lekseemide ja sõnade, aga ka A2+ sõnade arv) omavahel, konstruktsioonidest on sõnavaranäitajatega enim seotud liitlausele omased seosed, veidi vähem aga fraasitüübid (ainult lekseemide ja sõnade arv); VL konstruktsioonidel aga ühegi vaadeldud muutujaga statistiliselt olulist seost pole.

Tabel 4. 2024. a tekstides (halliga on märgitud keskmisest suurema väärtusega näitajatega lahtrid)

Lapse kood	FT24	VL24	Liitlausele omaste seoste arv	Lekseemide arv	Sõnede arv	Üle A2-taseme sõnavara
Da9	3	6	14	57	111	5
Ev20	10	5	13	114	223	12
Li21	5	3	6	69	69	6
Ma1	6	9	17	106	200	19
Mi02	8	5	24	138	372	15
Sa05	6	5	10	81	166	3
Za8	4	5	5	59	103	7
Zl13	9	8	31	153	402	14
Ti10	5	5	9	55	104	7
Ti23	6	8	11	55	146	4
Keskmine	6	6	14	89	190	9

Neljal õpilasel, kelle tekstid on keskmisest pikemad ning neis on keskmisest rohkem eri lekseeme ja kõrgema keeletaseme lekseeme, on selgelt ka suurem valik erinevaid konstruktsioone. Õpilastel, kelle tekstides on vähemalt üht tüüpi konstruktsioone keskmisest rohkem (nt laps koodiga Ev20), on üldjuhul rohkem ka lekseeme (ja pikem tekst). Ka liitlausele omaste seoste arv on rohkem neil, kel on vähemalt üht tüüpi konstruktsioone üle keskmise ja pikem tekst.

Eelpool kirjeldatud nelja õpilase tekste analüüsiti Sõnaveebi tekstihindamisvahendiga⁵, mis võimaldas jälgida eri tasemete sõnavara õpilase tekstis, aga ka konstruktsioonides kasutatud grammatiliste vormide keeletaset. Selgus, et vaadeldud õpilaste sõnavarast moodustavad B1-, B2-taseme ja tasememääratluseta lekseemid kokku umbes kümnendiku (ühel lausa 18%), nt *tiik* (B1) *ümbrik* (B1), *beebi* (B1), *streikima* (B2), *topis* (määramata); *rulluisud* (B1), *beebivanker* (määramata), aga ka kõnekeelseid sõnu, nagu *plära* ja *näts* (B1). Tekstides on selges ülekaalus eelA1- ja A1-taseme sõnad, ent see on igapäevateemade käsitlemisel tavaline ka emakeelekõnelejal.

Teksti keerukus võib avalduda ka keerukamate vormide kasutuses, nt esineb vaadeldud õpilaste tekstides lühikese mitmuse osastava, nt *tunde polnud* (B1), *asju* (B1), ja umbisikulise tegumoe vorme, nt *andi* (anti: B1). Tekstides kasutatud grammatilised konstruktsioonid on niisiis keerukamad (vrd nt joonis 6 ja 8) ja/või kasutatakse kõrgemate tasemete sõnavara või vorme. Fraasitasandil kasutatavad vaadeldud lapsed näiteks nii enama kui ühe laiendiga kui ka keerukamaid vorme sisaldavaid fraase, nt *palju huvitavaid asju, väga suured järjekorrad, mõne aja pärast, tema väike beebi, pärast viimast tundi, tema kaallid (:kallid) uued rulluisud; loomade topiseid, enda koerasi* (B1-taseme vormid). Lausetasandil kasutab õpilane enam VL kombinatsioone (vrd joonis 8), kasutatakse keerukamaid vorme ja konstruktsioone (A2+), aga ka keeruka struktuuriga lauseid, nt *Minu meelest võib sööklas noormalset kõhu täis saada. Ma nägin seal palju huvitavaid asju nagu riided, tööriistad ja ametid*; esineb mitmeastmelisi rind- ja põimlausele omasteid seosteid sisaldavaid liitseid liitlauseid, nt *Aga mõne aja pärast ta sai aru, et ta tegi*

⁵ <https://sonaveeb.ee/teacher-tools/#/rating> (3.10.2024).

paha asja, ja tõi tagasi kohta kus ta neid (:need) võttis koos kirjaga; Tüdruk oli väga kurb kui sai aru, et tema rull uisud (:rulluisud) varastasi (:varastati), aga kui järgmisel päeval ta läks kooli nägi ta garderoobis oma rulluisud (:rulluiske)!; Mina ei söö kooli sööklas, sest seal on plära ja kooli toitust (:koolitoidust) mul hakkab valutama kõht ja pea. Kõrgemate tasemete sõnad ja vormid ei esine tingimata alati koos keerulisemate konstruktsioonidega (ja vastupidi), nii näiteks on kaassõnafraasi *pärast viimast tundi* sõnavara eelA1/A1-tasemel, vorm *viimast* aga A2-taseme vorm, ent kahe laiendiga kaassõnafraas *kaassõnaga pärast* iseloomustab B1-taseme keelekasutust; eelpool toodud esimeses liitse liitlause näidetes on kasutatud ainult A-taseme sõnu (sh 87% eelA1) ja vorme (peamiselt eelA1 ja A1). Kokkuvõttes näitavad nii kõrgemate tasemete sõnad või vormid kui ka keerulisemad konstruktsioonid siiski nende õpilaste arenenumat keeleoskust.

Kõnealuse nelja lapse näitel võib niisiis öelda, et nad ületavad selgelt eesti keele kui teise keele õppekava eeldatud taseme A1.2~A2.1 (I kooliastme lõpuks on hea tulemus A1.2, väga hea A2.1, rahuldav A1.1) keeleoskust⁶; aga ka projekti eesmärgi (A2, ilma täpsustusteta)⁷: analüüsitud materjali põhjal võib väita, et neil on vähemalt A2.2-keeleoskus. Kuna õpilased kasutavad lihtsas igapäevatekstis ka mõningaid kõrgematele tasemetele iseloomulikke lekseme, vorme ja konstruktsioone, võib eeldada, et nad kas on juba saavutanud või on väga lähedal B1-tasemele (uurimismaterjal ei võimalda täpsemalt keeletaset määrata, selleks oleks vaja analüüsida nt õpilaste ainetekste).

7. Arutlus ja kokkuvõte

Vaadeldud 9–10-aastaste E2 omandavate õppijate kirjalike tööde sõnavara ulatus (Giraud' indeks) jäi sügisestes töödes 5,01 ja 7,69, kevadistes aga ja 4,55 ja 7,63 vahele. Kui laste keskmine sõnavara ulatuse indeks oli sügisestes töödes 6,24, siis tõus kevadeks 6,30-le olulist erinevust ei märgi: ehkki tekstid olid pikemad ning keskmine lekseemide arv kasvanud 64-lt 89-le, ei ole sõnavara ulatuse kasvu näha. Seega toob tekstide pikemaks muutumine küll kaasa suurema lekseemide arvu, kuid teksti pikkuse kasv muudab indeksi väärtuse samaks. Seetõttu võib oletada, et kasutatud mõõtevahend ehk G indeks võib näidata teksti sõnavara ulatust paremini mingil kindlal ajahetkel, kuid ei sobi ilmselt kuigi hästi kirjeldama sõnavara arengut. Kindlasti oleks vaja oluliselt suuremat valimit, et mõõtevahendi sobivust paremini hinnata.

Kui võrrelda tulemusi E1 kõnelevate 5. klassi õpilaste põhise uurimuse tulemustega (Kerge jt 2014: 163), selgub, et siinses materjalis oli E2 omandavate õppijate sõnavara isegi suurem: Kerge jt andmetel on 5. kl õpilaste loovkirjutiste sõnavara varieeruvuse indeks keskmiselt 5,73 (peaaegu punkti võrra madalam kui siinses uurimuses). Selle erinevuse üks võimalik põhjus on see, et õppijate tekstid on suhteliselt nimisõnarohked, see tõstab G väärtust, emakeelsed kirjutajad võivad kasutada enam eri verbe (aga mitte nii palju rohkem kui on õppijal rohkem nimi-sõnu), mistõttu on õppija sõnavara ulatuse indeks kõrgem. See oletus vajab aga E1 materjali põhjal kontrollimist.

Sõnavara areneb sõnaliigiti erinevalt. Kui 2023. a sügisel oli õpilaste kirjutistes keskmiselt 11 eri tegusõnalekseemi, siis pool aastat hiljem oli nende hulk kasvanud

⁶ <https://oppekava.ee/eesti-keel-teise-keelena-oppeprotsess/> (3.10.2024).

⁷ <https://harno.ee/eestikeelne-opetaja-haridusasutuses> (3.10.2024).

keskmiselt kuue võrra. Teine sõnaliik, milles on näha arengut, on sidesõnad: 2023. a sügise tekstides oli neid keskmiselt kaks, pool aastat hiljem aga neli. Nii suurenenud tegu- kui ka sidesõnade valik näitab keele komplekssemaks muutumist ilmselt oluliselt ilmekamalt kui lekseemide üldine arv või sõnavara varieeruvuse indeks.

Õppijate tekstid koosnevad peaaesjalikult eelA1-taseme sõnadest. Võrreldes 2023. a sügise töödega on nende osakaal aga vähenenud 14%, suurenenud on A1-, A2- ja B1-taseme sõnade osakaal. Mõnevõrra on vähenenud väljapoole tasemekirjeldusi jääv sõnavara, mis tuleneb väiksemast hulgast pärisnimedest 2024. a kevade töödes. Kuna valim on väike, mõjutab iga üksiku õppija töö väga palju tulemusi, seega ühe õppija töös kasutatud kuus kaubamärkinime suurendab kindlasti väljapoole tasemekirjeldusi jäävate sõnade hulka sügisestes töödes. Siiski näitavad väljapoole tasemekirjeldusi jäävad sõnad keelekasutuse muutumist komplekssemaks, eelkõige seetõttu, et nende sõnade puhul oli enamasti tegemist komplekssete sõnadega ehk liitsõnade või tuletistega, mida alati ei loetletagi tasemekirjeldustes eraldi sõnavaraüksusena. Osa erinevusest võib küll tuleneda kahe perioodi teemade erinevusest, samas näitab komplekssemate sõnade kui teemast ja kontekstist vähem sõltuvate lekseemide (võrreldes pärisnimedega) kasutamine siiski väljendusoskuse arengut. Suurema valimi ja eri žanris tekstide põhjal annaks teha kaalukamaid järeldusi, kuid tasuks ka mõelda komplekssete sõnade koha üle tasemekirjeldustes: kas esitada sagedasemad neist sõnaloendites, vähem sagedaste, kuid produktiivsete sõnamoodustusmallide jaoks aga lisada tasemekirjeldusse sõnamoodustuse osa. Igal juhul tekib küsimus, kui sage peaks kompleksne sõna olema, et eeldada selle tervikuna omandamist, või kui produktiivne peaks mall olema, et oletada sõnamoodustust.

Siinse materjali puhul oli ligi 90% sõnu eelA1, A1- või A2-keele taseme sõnavara. Kuigi Alpi jt (2013) analüüsis oli sõnu jagatud sagedussõnastiku baasil sagedasteks ja harvadeks sõnadeks ning võrdlus Sõnaveebi keeleoskustasemete sõnavaraga on mõnevõrra tinglik, võib siiski lugeda A-taseme sõnad sagedasemateks ja sedakaudu väita, et noore õppija tekstides on võrreldes riikliku tasemeeksami tööde tekstidega (kus 80% lekseemidest kuulusid sagedasemate sõnade hulka) ilmselt pisut suurem sagedaste sõnade osakaal.

Võrreldes suulise pildikirjelduskatse tulemustega (vt Argus, Rüütmaa 2024), on kirjalikes tekstides tasemekohase sõnavara areng paremini nähtav, mis võib tuleneda suulise ja kirjaliku teksti eripärast. Lundell ja Lindqvist (2012: 6) on väitnud, et näiteks E1 suuliste tekstide sõnavaras oli suure sagedusega ehk baassõnavarasse kuuluvaid sõnu võrreldes B2-C1-taseme õppijate tekstidega ühepalju. Seega võib just kirjalike tekstide analüüs näidata õppijate liikumist ühelt keeleoskustasemelt teisele paremini kui suulised tekstid.

Konstruksioonitasandil muutus enamiku laste keelekasutus samuti komplekssemaks: uusi üldkasutatavaid fraasiliike, -tüüpe ega verbi ja laiendi kombinatsioone ei lisandunud, küll aga kasvas neist peaaegu kõigi puhul kasutajate hulk. Fraasitasandil on enamikul vaadeldud õpilastest lisandunud kas fraasiliike (seitsmel kümnest), -tüüpe (pooltel) või mõlemaid (neljal); lausetasandil on peaaegu kõigil õpilastel (v.a üks) tulnud juurde uusi verbi ja laiendi kombinatsioone; samuti kasutavad valimi õpilased enam rind- ja põimlausele omaseid seoseid (ühel kasv ainult põimlausestes), mis tähendab enamiku puhul omakorda vähem lihtlauseid (kaheksal õpilasel). Liitlauseste lisandumine on selgelt seotud ka sõnavara

arenguga: kasutatakse ka enam sidesõnu. Arengut on niisiis näha kõigi vaadeldud õpilaste puhul, aga mitte tingimata kõigis vaadeldud konstruktsioonides. Neljast õpilasest (Ev20, Li21, Za8, Ti10), kellel on arengut kõigis vaadeldud fraasitasandi ning verbi ja laiendi konstruktsioonides, on lisandunud ka kõiki vaadeldud lause- liike (liht-, koond-, rind- ja põimlauseid) ainult ühel (Ev20). Temal on ühtlasi nii lekseemide kui ka sõnede arvult ning A2+ sõnavara osakaal keskmisest suurem. Teisalt, ühel õpilasel (Sa05), kel pole arengut üheski vaadeldud fraasitasandi ega verbi ja selle laiendite kombinatsioonis (neid on 2024 isegi vähem kui 2023), on siiski lisandunud koondlauseid ja liitlausele omaseid seoseid. Eelöeldu toetab varasemate grammatiliste struktuuride omandamist käsitlevaid uurimusi, mille järgi areneb keeleoskustaseme tõustes nii lause- kui ka fraasistruktuur, nimisõnafraasi morfoloogia, subjekti-verbi ühildumine jne (vt Forsberg, Bartning 2010). Konstruktsioone ei lisandu õppijakeelde ühtlases tempos (Martin jt 2010, vrd ka Kitsnik 2018), küll aga muutuvad konstruktsioonid ise keerukamaks, nt juba omandatud konstruktsioonis hakatakse kasutama kõrgematele keeletasemetele omaseid vorme (vt Kitsnik 2019) ja sõnavara.

Võrreldes kirjalike tekstide tulemusi suulise pildikirjelduskatse tulemustega (Argus, Rüütmaa 2024) on kirjalike tekstide fraasistruktuur ootuspäraselt keerulisem: kui õpilane kasutas 2024. aasta suulises pildikirjelduses keskmiselt 2,3 fraasiliiki, 4,2 -tüüpi ja 4,8 VL kombinatsiooni, siis kirjalikes tekstides 3,7 fraasiliiki ja 6,2 -tüüpi ja 5,9 VL kombinatsiooni. Niisiis erineb õpilaste suuline ja kirjalik tekst konstruktsioonide variatiivsuse poolest. Suulises kõnes eraldi liitlause arvu ei analüüsitud (enamasti ei tuvastatud selgeid lausepiire), seega pole võimalik kõrvutada liitlause esinemust, ent enamatele liitlausele kirjalikes tekstides viitab keskmine sidesõnalekseemide hulk: suulises tekstis keskmiselt 2 lekseemi lapse kohta, kirjalikes aga 2,9.

Kirjalike tööde tulemusi ja Sõnaveebi noore keeleõppija tasemekirjeldusi⁸ kõrvutades võib öelda, et valimi õpilased on jõudnud projekti sihiks olnud A2 tasemele⁹: kasutatakse kõiki põhilisi konstruktsioonitüüpe. Eesmärgiks olnud või mõne õpilase puhul isegi võimalik, et veidi kõrgemale keeletasemele viitab eelkõige konstruktsioonide täitmine keerukamate vormide ja sõnavaraga, nt kasutatakse kõrgematele keeletasemetele omaseid vorme nagu umbisikuline tegumood või lühike mitmuse osastav, samuti A2 või A2+ sõnavara; aga ka konstruktsioonitüüpide endi keerukamaks muutumine, nt lisandunud on enam kui ühe laiendiga fraaside ja enam kui kahe laiendiga VL kombinatsioonide kasutajaid. Viimati nimetatud konstruktsioonid viitavad juba keele märksa vabamale valdamisele ja kuna võimalike kombinatsioonide arv on avatud, oli igal neist enamasti vaid mõned üksikud kasutajad. Ühesõnaga, peamiste konstruktsioonide valik on valimi seisukohalt jäänud 2023. aasta andmetega võrreldes endiseks, ent samad konstruktsioonid on täidetud keerukama keelematerjaliga (ja igat konstruktsioonitüüpi kasutab 2024. a enam õpilasi kui 2023. a), mis viitab kõrgemale keeletasemele. Kuna kirjalike tööde seired olid suunatud eelkõige A-taseme oskuste tuvastamisele, on raske öelda, kas mõnede õpilaste keeleoskus ka A-taset ületab, ent kirjalike tekstide analüüsi põhjal võib eeldada, et on ka vähemalt B1-tasemele jõudnud õppijaid (B1-taseme määramine eeldaks siiski keerukamate teemade käsitlust).

⁸ <https://sonaveeb.ee/teacher-tools/#/grammar?page=1> (3.10.2024).

⁹ <https://harno.ee/eestikeelne-opetaja-haridusasutuses> (3.10.2024).

Vaadeldes sõnavara ja konstruktsioone seotult, ilmnes, et (sõnedes) pikema teksti kirjutatud lastel oli ka ulatuslikum sõnavara, rohkem kõrgema keeletaseme sõnu ja eri tüüpi konstruktsioone. Sõnavaranäitajate ja fraasitüüpide seost kinnitas ka korrelatsioonianalüüs: nende vahel oli oluline statistiline seos. Seevastu VL konstruktsioonidel polnud olulist statistilist seost ei sõnavaranäitajate ega fraasitüüpidega, st suurem sõnavara ega mitmekesised fraasitüübid ei tähendanud tingimata suuremat hulka lihtlausekonstruktsioone (või vastupidi). Näiteks oli tekste, kus oli küll kasutatud laialdast sõnavara, ent vaid paari konstruktsioonitüüpi. Artikli uurimismaterjali alusel pole võimalik teha kindlaid järeldusi selle erinevuse põhjustest, sellele tuleks tulevastel uurimustel kindlasti tähelepanu pöörata. Seega ei ole seos komplekssemate sõnade ja konstruktsioonide vahel lihtne: komplekssemad konstruktsioonid ei pruugi olla n-ö täidetud komplekssemate sõnadega ja keerukamaid sõnu ei kasutata tingimata keerukates konstruktsioonides. Valimi laste keele komplekskuse kasvule viitab lisaks konstruktsioonide ja sõnavara keerukuse kasvule ka kõrgematele keeletasemetele omaste vormide kasutus. Samas ei pruukinud keerulisemad vormid olla tingimata kasutusel keerulistes konstruktsioonides, laps võis kirjutada ka *sõime õunu*, mis konstruktsioonina on lihtne, aga vormidelt keeruline. Et vormikasutus ei olnud siinse analüüsi objektiks, tuleb see tulemus jätta järgmiste uurimuste üheks võimalikuks uurimisfookuseks.

Lühendid

G	Guiraud' indeks	Nu	arvsõna
E1	eesti keel esimese/emakeelena	P	kaassõna
E2	eesti keel teise keelena	Pr	asesõna
FL	fraasiliik	Q	hulgasõna
FT	fraasitüüp	SV	alus+öeldis
VL	finiitse verbivormi (lausetuuma) ja selle laiendite kombinatsioon	SVO	alus+öeldis+sihitis
AdjP	omadussõnafraas	SVA	alus+öeldis+määrus
AdvP	määrsõnafraas	SVP	alus+öeldis+öeldistäide
NP	nimisõnafraas	AVS	määrus+öeldis+alus
PP	kaassõnafraas	SVAA	alus+öeldis+määrus+määrus
QP	hulgasõnafraas	SVOA	alus+öeldis+sihitis+määrus
Adj	omadussõna	KL	koondlause
Adv	määrsõna	RL	rindlause
N	nimisõna	PL	põimlause

Viidatud kirjandus

- Allkivi, Kais 2016. Verbist paremal ja vasakul paiknev kontekst C1-tasemega eesti keele õppijate ja emakeelekõnelejate kirjutistes. – XII muutuva keele päev. Ettekannete teesid. Tartu: Tartu Ülikool, 18–19. https://www.keel.ut.ee/sites/default/files/www_ut/mkp_2016_teesid.pdf (12.8.2018).
- Allkivi-Metsoja, Kais 2021. Eesti keele A2–C1-taseme kirjalike tekstide võrdlev automaatanalüüs. – Lähivõrdlusi. Lähivertailuja, 31, 13–59. <https://doi.org/10.5128/LV31.01>
- Alp, Pilvi; Kerge, Krista; Pajupuu, Hille 2013. Measuring lexical proficiency in L2 creative writing. – Jozef Colpaert, Mathea Simons, Ann Aerts, Margret Oberhofer (Eds.),

- Language Testing in Europe: Time for a New Framework? Antwerpen: Linguapolis Universiteit Antwerpen, 274–286.
- Argus, Reili; Baird, Piret; Bauer, Annika; Kütt-Leedis, Andra; Lutsepp, Liina; Meristo, Merilyn; Rüütmaa, Tiina; Jõemets, Viivian 2024. Pilootprojekti “Professionaalne eestikeelne õpetaja haridusasutuses” laste keelelise arengu hindamise tulemused. Grammatiliste kategooriate omandamine. Sõnavara ja konstruktsioonide omandamine pildikirjelduskatse põhjal. Kirjaliku väljendusoskuse omandamine.
- Argus, Reili; Rüütmaa, Tiina 2024. Sõnavara ja konstruktsioonide areng eesti keelt teise keelena omandava noore õppija kõnes aasta jooksul. – Eesti Rakenduslingvistika Ühingu aastaraamat, 20, 5–24. <https://doi.org/10.5128/ERYa20.01>
- Bates, Elizabeth; Goodman, Judith C. 1997. On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. – Language and Cognitive Processes, 12 (5/6), 507–584. <https://doi.org/10.1080/016909697386628>
- Daller, Helmut; Milton, James; Treffers-Daller, Jeanine 2007. Series editors' preface. – Helmut Daller, James Milton, Jeanine Treffers-Daller (Eds.), Modelling and Assessing Vocabulary Knowledge. Cambridge Applied Linguistics. Cambridge: Cambridge University Press, xii–xiii. <https://doi.org/10.1017/CBO9780511667268.002>
- Erelt, Mati 2017. Sissejuhatus süntaksisse. – Mati Erelt, Helle Metslang (Toim.), Eesti keele süntaks. Eesti keele varamu, 3. Tartu: Tartu Ülikooli Kirjastus, 33–88.
- Forsberg, Fanny; Bartning, Inge 2010. Can linguistic features discriminate between the communicative CEFR-levels? A pilot study of written L2 French. – Inge Bartning, Maisa Martin, Ineke Vedder (Eds.), Communicative Proficiency and Linguistic Development: Intersections Between SLA and Language Testing Research. EuroSLA Monographs Series, 1. European Second Language Association, 133–158.
- González-Fernández, Beatriz 2024. How is vocabulary learnt? An acquisitional sequence of L2 word knowledge. – TESOL Quarterly, A3342. <https://doi.org/10.1002/tesq.3342>
- Guiraud, Pierre 1960. Problèmes et méthodes de la statistique linguistique. Dordrecht: D. Reidel.
- Hjetland, Hanne Næss; Hofslundsengen, Hilde; Klem, Marianne; Karlsen, Jannicke; Hagen, Åste Mjelve; Engevik, Liv Inger; Geva, Esther; Norbury, Courtenay Frazier; Monsrud, May-Britt; Næss, Kari-Anne B. 2023. Vocabulary interventions for second language (L2) learners up to six years of age. – Cochrane Database of Systematic Reviews, 8, CD014890-CD014890. <https://doi.org/10.1002/14651858.CD014890.pub2>
- Ivanovska, Biljana; Xhaferi, Gzim 2023. Second language vocabulary acquisition and learning strategies. – Folia Linguistica et Litteraria, 44, 359–363. <https://doi.org/10.31902/fll.44.2023.21>
- Juffs, Alan 2009. Second language acquisition of the lexicon, linguistics. – William Ritchie, Tej K. Bhatia (Eds.), The New Handbook of Second Language Acquisition. Leeds: Emerald.
- Kerge, Krista; Uusen, Anne; Põlda, Halliki 2014. Teismee loovkirjutiste sõnavara ja selle hindamine. – Eesti Rakenduslingvistika Ühingu aastaraamat, 10, 157–175. <https://doi.org/10.5128/ERYa10.10>
- Kitsnik, Mare 2017. Modaalverbide *saama* ja *võima* konstruktsioonid B1- ja B2-taseme kirjalikus õppijakeeles kui keeleoskuse arengu näitaja. – Eesti Rakenduslingvistika Ühingu aastaraamat, 13, 19–36. <https://doi.org/10.5128/ERYa13.02>
- Kitsnik, Mare 2018. Iga asi omal ajal: eesti keele B1- ja B2-taseme verbikonstruktsioonid keeleoskuse arengu näitajana. Tallinna Ülikooli humanitaarteaduste dissertatsioonid, 43. Tallinn: Tallinna Ülikool.
- Kitsnik, Mare 2019. Eesti keele kui teise keele õppimine: kas raske töö või kerge lõbu? – Keel ja Kirjandus, 1–2, 39–56. <https://doi.org/10.54013/kk735a4>
- Laufer, Batia 1998. the development of passive and active vocabulary in a second language: Same or different? – Applied Linguistics, 19 (2), 255–271. <https://doi.org/10.1093/applin/19.2.255>

- Luchini, Pedro Luis; Ferreiro, Gabriela Marie 2023. Second language vocabulary teaching and learning: Inspecting alternatives and analyzing results. – *Athens Journal of Philology*, 10 (2), 83–100. <https://doi.org/10.30958/ajp.10-2-1>
- Lundell, Fanny Forsberg; Lindqvist, Christina 2012. Vocabulary aspects of advanced L2 French: Do lexical formulaic sequences and lexical richness develop at the same rate? – *Language, Interaction and Acquisition*, 3 (1), 73–92. <https://doi.org/10.1075/lia.3.1.05for>
- Martin, Maisa; Mustonen, Sanna; Reiman, Niina; Seilonen, Marja 2010. On becoming an independent user. – Inge Bartning, Maisa Martin, Ineke Vedder (Eds.), *Communicative Proficiency and Linguistic Development: Intersections Between SLA and Language Testing Research*. EuroSLA Monographs Series, 1. European Second Language Association, 57–80.
- Milton, James; Donzelli, Giovanna 2013. The lexicon. – Julia Herschensohn, Martha Young-Scholten (Eds.), *The Cambridge Handbook of Second Language Acquisition*. Cambridge Handbooks in Language and Linguistics. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139051729.027>
- Mulder, Evelien; Van de Ven, Marco; Segers, Eliane; Verhoeven, Ludo 2019. Context, word, and student predictors in second language vocabulary learning. – *Applied Psycholinguistics*, 40 (1), 137–166. <https://doi.org/10.1017/S0142716418000504>
- Nation, Paul; Gu, Peter Y. 2007. *Focus on Vocabulary*. Sydney: National Centre for English Language Teaching and Research.
- Puteri, Celiacika Gustisiwi 2022. Second language acquisition: students' vocabulary size and their strategies for retaining it. – *Aksara*, 23, 1, 30–36. <https://doi.org/10.23960/aksara/v23i1.pp30-36>
- Sun, Yanqing 2015. Effective Strategies for Second Language Vocabulary Acquisition. – *Proceedings of the International Conference on Social Science, Education Management and Sports Education (SSEMSE 2015)*. *Advances in Social Science, Education and Humanities Research*, 39. Atlantis Press, 1881–1884. <https://doi.org/10.2991/ssemse-15.2015.476>
- Sõnaveeb: Õpetaja tööriistad. Eesti Keele Instituut. <https://sonaveeb.ee/teacher-tools/#/home>
- Vajjala, Sowmya; Lõo, Kaidi 2014. Automatic CEFR level prediction for Estonian learner text. – *Proceedings of the Third Workshop on NLP for Computer-assisted Language Learning*. NEALT Proceedings Series, 22, 113–127.
- Yu, Guoxing 2009. Lexical diversity in writing and speaking task performances. – *Applied Linguistics*, 31 (2), 236–259. <https://doi.org/10.1093/applin/amp024>

VOCABULARY AND CONSTRUCTIONS IN THE WRITTEN TEXTS OF YOUNG LANGUAGE LEARNERS

Reili Argus, Tiina Rüütma

Tallinn University

The aim of this article is to provide an overview of the development of lexical resources and grammatical constructions in the written works of 9–10-year-old learners acquiring Estonian as a second language over a period of six months (autumn 2023 – spring 2024). The vocabulary and grammatical constructions of a total of 10 learners are described based on variety of vocabulary, distribution of word classes, vocabulary across different language proficiency levels, and grammatical constructions.

The variety of vocabulary in the written texts (measured by the Guiraud's index) did not change significantly over six months. However, vocabulary development was indicated by the addition of verb lemmas, conjunctions, and vocabulary from higher language proficiency levels. The development of constructions was mainly characterized by increased linguistic diversity, with a few new construction types becoming more complex and varied.

Examining the relationship between vocabulary and constructions in the written texts revealed that longer texts contained a broader vocabulary, more words from higher language levels, and a greater variety of constructions. However, the relationship between words and constructions was not one-to-one: more complex constructions did not necessarily include more complex words, and more advanced words were not necessarily used in complex constructions.

Keywords: Estonian as a Second Language, language acquisition, language proficiency level, Guiraud's index, young language learner, Estonian

Reili Argus (Tallinna Ülikool) on keeleomandamise uurija, peamised huvivaldkonnad on morfoloogia ja leksikaalmorfoloogiliste kategooriate varane omandamine.
Narva mnt 25, 10120 Tallinn, Estonia
reili.argus@tlu.ee

Tiina Rüütmaa (Tallinna Ülikool) peamised uurimisvaldkonnad on eesti keele kui teise keele õpetamine ja eesti-ungari kontrastiivlingvistika.
Narva mnt 25, 10120 Tallinn, Estonia
rtiina@tlu.ee