

EESTI KEELE KONSTRUKTIKONI POOLE: ÜLESEHITUSE PEAJOONED

Geda Paulsen, Heete Sahkai, Ene Vainik,
Jelena Kallas, Kertu Saul, Arvi Tavast, Kristina Koppel

Ülevaade. Artikkel käsitleb eesti keele konstruktikoni kontseptsiooni – grammatika esitamist sõnastikulaadselt – ja esitleb esimest visiooni sellise ressursi ülesehitusest. Artikkel on sisuliseks jätkuks varasematele (Vainik jt 2024a, Vainik jt 2024b), milles kirjeldati konstruktsioonipõhist keelekäsitust ning rahvusvahelist konstruktikonide loomise kogemust ja väljakutseid. Siinses artiklis defineerime kõigepealt eesti konstruktikoni kontseptsiooni ja visiooni mõistmiseks vajalikud põhimõisted ning põhjendame ettevõtmist nii rakenduslikust kui ka teoreetilise vaatepunktist. Käsitelu põhiosas esitame konstruktikoni makro- ja mikrostruktuuri kirjeldused ning juhtumianalüüsina pakume välja visandid kahe hulgafrasi konstruktikograafilisest kirjeldusest.*

Võtmesõnad: konstruktsioon, konstruktikon, konstruktikograafia, leksikograafia, eesti keel

1. Sissejuhatus

Sõnavara ja grammatikat on traditsiooniliselt käsitletud keelelise teadmise erinevate osadena ja kirjeldatud seetõttu eraldi ressurssides, sõnastikes ja grammatikates. Konstruktsioonipõhises keeleteoorias on aga jõutud arusaamisele, et leksikaalsed ja grammatilised üksused on oma olemuselt sarnased ja moodustavad pidevustiku. Selle järelduse loomulikuks jätkuks on arusaam, et leksikaalseid ja grammatilisi üksusi tuleks ka kirjeldada samas ressursis või vähemalt samade põhimõtete kohaselt.

Selles artiklis anname põgusa sissevaate Eesti Keele Instituudis käsil olevasse teadusprojekti, mille raames uuritakse grammatilise info sõnastikus kättesaadavaks tegemise võimalusi ja probleeme. Projekti, millest lähemalt juttu teeme, nimetame tinglikult “Eesti keele konstruktikoniks”.¹ Tegemist on rakenduslikust

* Uurimistööd on finantseerinud Eesti Teadusagentuuri grant “Uue aja sõnastik: grammatika ja keelepädevuse kirjeldamine integreeritud multifunktsionaalses leksikograafilises ressursis” (PRG 1978).

¹ Tinglikuks praeguses seisus, kuna kasutame seda nimetust teiste keelte konstruktikonide eeskujul ja teooriast tulenevalt (vt Vainik jt 2024a, Vainik jt 2024b), samas anname aru, et tegu on kohmaka ja halvasti artikuleeritava nimetusega ning kui konstruktikon peaks saama argielu osaks, võib mingil hetkel tekkida vajadus juurutada mõni suupärasem nimetus.

vajadusest lähtuva ning tänapäevasest keeleteooriast – konstruktsioonigrammatikast – motiveeritud ettevõtmisega. Eesti keele konstruktikoni all peame silmas elektroonilist keeleressurssi, mis kirjeldab eesti keele sõnavara ja grammatikat ühtses sõnaraamatulaadses formaadis. Konstruktikoni funktsioonidena näeme eesti keele dokumenteerimist, grammatilise info kasutajale kättesaadavaks tegemist, keeleõppijate toetamist ja keeletehnoloogiliste rakenduste ja teenuste arendamist.

Ettevõtmise teoreetilist alust, konstruktsioonipõhist keelekirjeldust ning selle põhimõisteid, oleme varem tutvustanud (Vainik jt 2024a, Vainik jt 2024b); sealt leiab hulgaliselt viiteid sellele teooriaraamistikule nii eestikeelses kui ingliskeelses teaduskirjanduses. Praktilise vajaduse pakkuda keeleõppijatele ja -õpetajatele mõeldud grammatilist infot vastavalt CEFR-i keeleoskustasemetele oleme samuti varem lahti kirjutanud (Pool jt 2024). Varasem analüüs kätkeb ka ülevaadet käimasolevatest teiste keelte konstruktikonide projektidest ning raskustest, millega on tüüpiliselt kokku puutunud (Vainik jt 2024a). Ettekujutus sellest, kuidas eesti keele konstruktikoni luues neid komistuskive vältida ja kuidas praktiliselt toimida, on seni siiski olnud üsna ebamäärane ja fragmentaarne (Vainik jt 2024b). Siinse artikli eesmärk ongi esitada juba kavakindlamalt eesti keele konstruktikoni kontseptsioon ja visioon koos konkreetsete näidetega. Vajaliku eelsammuna esitame esmalt selgitused põhimõistete kohta ning toome esile ka põhjendused, miks on konstruktikoni loomine meie arvates vajalik.

Artikli põhistruktuuri võtavad seega kokku küsimused *mis? miks? kellele? kuidas?* ning *kuidas konkreetset?* Peatükis 2 avame põhimõisted, peatükis 3 esitame põhjendused ja peatükis 4 põhimõtted nii konstruktsioonide valikuks kui peamiste kirjelduskihtide kohta. Alapeatükis 4.1 esitame konstruktsioonitüübid, mida eesti keele konstruktikon võiks katta. Osas 4.2 pakume ka konkreetseid näiteid, kuidas konstruktikonis saaks kirjeldada keeleõppijale tülikat hulgafraasi. Artikkel lõppeb kokkuvõtte ja aruteluga.

2. Põhimõisted ehk *mis?*

Sissejuhatuses juba kasutasime mõisteid *konstruktsioon* ja *konstruktikon*. *Konstruktsioon* on olnud kasutuses ka traditsioonilises ja strukturalistlikus grammatikakirjelduses, ent siinses käsitluses on sellel terminil spetsiifiline konstruktsioonipõhisest keelekäsitusest lähtuv tähendus (vt Vainik jt 2024a ja sealsed viited konstruktsioonigrammatikale, nt Goldberg 2003, Booij 2010, Gras, Elvira-García 2021).

Konstruktsiooni all mõistame vormi ja tähenduse paari – keeleüksust, milles teatav vorm seostub teatavat laadi tähendusega. Konstruktsioonigrammatikas käsitletakse konstruktsioonidena kõiki keelise teadmise üksusi. Konstruktsioon võib koosneda üksiksõnast, aga ka mitme sõna ühendist, mida iseloomustab konventsionaalne ja lahutamatu vormi-tähenduse kombinatsioon (nt Goldberg 1995: 4, Fried, Östman 2004: 24). Näiteks väljend *mees metsast* esindab konstruktsiooni, milles seostuvad nimisõna ja kohakäändes nimisõna ning mille tähenduseks on iseloomustada fraasi põhjaks oleva nimisõna referenti tema ilmutumiskoha alusel jagatud teadmistesse (seega mitte n-ö tema püüseloomustusena). Sama konstruktsioon on ka väljendites *inimene tänavalt*, *Karlsson katuselt*, aga miks mitte

ka produktiivselt ehk ainukordsetes ühendites, esitledes kaaslasele kedagi kui *Mari tantsutrennist, Jüri seminarilt* vmt. Konstruksioonile selles eritähenduses on iseloomulik, et ta võib sisaldada n-ö lünkasid, mida keelekasutuses sobiva sisuga täidetakse. Eespool mainitud konstruksioonis on kaks lünka ja mõlemas on nimisõnale teatavad sisulised piirangud – oleks imelik öelda nt *?kivi toast*, kuigi ilmselt mitte võimatu. Paratamatult hakkab konstruksiooni kui terviku tähendus siin peale suruma tõlgendust, et kõnealusel kivil on mingi, kas või põgus ajalugu konkreetse toaga.

Sõnavara-grammatika pidevustik ehk kontiinuum kätkeb olulist teoreetilist arusaama, et klassikalises mõttes jääka piiri leksikoni ja grammatika vahele ei saa tõmmata (Croft 2001, Goldberg 2003, Hoffmann, Trousdale 2013). Pidevustiku üht otsa saab kirjeldada kui leksikaalselt küllastatud konstruksioone (näiteks idiomaaatilised väljendid, nagu *võta näpust* või *kuidas külvad, nõnda lõikad*, mis koosnevad konkreetsetest sõnadest). Pidevustiku teise otsa paigutuvad abstraktsed skemaatilised n-ö lünkadega konstruksioonid, milles saab esineda teatud valik sõnu (näiteks argumentstruktuurikonstruksioonid, nagu intransitiivne subjekti-predikaadi konstruksioon (*Mia loeb, Peeter jalutab*) või transitiivne konstruksioon, kus objektiks on da-infinitiivifraas (*Mia armastab lugeda, Peeter tahab jalutada*). Leksikaalselt küllastatud ja täisskemaatiliste konstruksioonide vahele jääb aga hulk poolskemaatilisi ehk osaliselt lünkadega või varieeruvate liikmetega konstruksioone. Näiteks *Mia pole koristajat* esindab ühe leksikaalselt küllastatud liikmega (olemaNEG) konstruksiooni, kus kõneleja hindab negatiivselt elatiivse nimisõnafrasiga tähistatud osaleja suutlikkust täita ülesannet, mida tähistab tüüpiliselt *ja*-tegijanimega moodustatud partitiivis nimisõnafraas. Teine näide võiks olla *minema* 3. isikus koos tegevust väljendava nimisõnaga ainsuse translatiivis: *läheb löömaks, läks aruteluks, võib minna kakluseks* jne. Sellel puhul on üks liige leksikaalselt küllastatud (ja ka vormipiirangutega), kuid teist liiget saab täita erinevate sisult sobivate nimisõnadega – ka sellistega, mida *ad hoc* alles tuletatakse tegusõnast näiteks *mine*-liitega.

Konstruktikon on termin, millega siinses artiklis viitame eeskätt teatavat tüüpi keelekirjeldusele – konstruksioonide kogumikule² (nii nagu tavamõistes on leksikon sõnade kogumik või grammatika reeglite kogumik).

Nii konstruksioonide kogumikke kui ka keelesüsteemi tervikuna iseloomustab **võrgustikulaadne ülesehitus** (nt Diessel 2019, 2023). Eri taseme konstruksioonide vahel on võrgustikus eri tüüpi seosed. Näiteks **hierarhiline** nn **pärlusseos** on üldjuhul abstraktse konstruksiooni ja tema konkreetsema väljenduse vahel: “laps” on “vanema” spetsiifilisem variant (Hudson 2007, Traugott, Trousdale 2013). **Horisontaalset** seosetüüpi võib kirjeldada “naabri” kujundi kaudu (Van de Velde 2014: 147): seda eristatakse sama tasandi konstruksioonide, aga ka konstruksiooni lünga ja selles esinevate sõnade vahel (Goldberg 1995: 78). Naaberkonstruksioonideks võib pidada hulgafrase *kott õunu* ja *kott suhkrut*, kus hulgasõna laiendi vorm oleneb selle referendi tähendusest (vt konstruksiooni-võrgustiku näidet joonisel 4 ptk 4.2). Seega on konstruktikonil võrgustikuloomuse tõttu – kui jätkata analoogiaga – teatud sarnasus ka tesaurustega. Lisaks püüdlavad konstruktikonid keeltevaheliste vastavuste loomise suunas (Lorenzi jt 2024), sarnanedes selles osas WordNetiga (Geller jt 2020). Konstruksioonidevaheliste

² Mõõndusega eeskätt sellepärast, et erialakirjanduses on *konstruktikon* polüseemne: see tähendab nii keelt ennast kui konstruksioonide kogumit kui ka keelekirjeldust konstruksioonide kogumikuna (Lyngfelt 2018).

seoste esitamine võimaldab teha üldistusi ning on osutunud keeleõpet tõhustavaks faktoriks, nagu on näidanud González-García (2019) edasijõudnud L2 õppijate uuring.

Konstruktikograafia on valdkond, kus tegeldakse keeleüksuste konstruktioonidena kirjeldamisega konstruktikonis (lihtne analoog on siin leksikograafia kui tegevusvaldkond). Konstruktikograafia ongi välja kasvanud sellest, et konstruktioonigrammatika teoorias käsitletakse kõiki keelelise teadmise üksusi samasuguste vormi ja tähenduse paaridena nagu on sõnad. Sellest tuleneb loogiliselt, et kogu keelelist teadmist on otstarbekas kirjeldada ühes formaadis ja just sellises, nagu seni on kirjeldatud sõnu (vt ülevaadet Vainik jt 2024a).

Võrdlismõiste (ingl *comparative concept*) on termin, millega viidatakse keeletüpoloogias üldistele katusmõistetele (Croft 2016, 2022, Haspelmath 2010), ning rahvusvahelises konstruktikograafias on need võetud kasutusele selleks, et oleks võimalik fikseerida vastavusi eri keelte konstruktioonide vahel (Borin, Lyngfelt 2025, Lyngfelt jt 2022).

Erinevalt teistest konstruktikonidest kasutame siinses kirjelduses ka leksikograafiaditsioonist pärit mõisteid nagu makrostruktuur ja mikrostruktuur (Svensén 2019: 344–377). **Makrostruktuur** hõlmab otsustamist selle üle, millised üksused kuuluvad kirjeldamisele. Traditsioonilises leksikograafias esindab seda näiteks märksõnastik, aga ka üldisemad põhimõtted selles osas, millist laadi üksusi kaasata. **Mikrostruktuur** kätkeb endas selle määratlemist, millist infot keeleüksuste kohta esitatakse – traditsioonilises leksikograafias näiteks hääldus, tähenduse selgitus, vormimoodustus jpm.

Millise sisu saab lõpuks termin **eesti keele konstruktikon**, selgub kirjeldatava projekti tulemustest, kuid praeguse seisuga mõistame selle all eesti keele morfosüntaktiliste kasutusmustrite esitamist sõnastikulaadselt – vormi ja tähenduse paaridena. Meie plaan on luua konstruktikon sõnastiku- ja terminibaasisüsteemi Ekilex (Tavast jt 2018) EKI ühendsõnastiku (Langemets jt 2021, Langemets jt 2023) laiendusena.

3. Põhjendused ehk miks ja kellele?

Nagu sissejuhatuses öeldud, on eesti keele konstruktikoni loomisel nii teoreetilised kui rakenduslikud põhjendused. Teoreetiline põhjendus on konstruktioonipõhisest keeleteooriast tulenev arusaam, et sõnavara ja grammatika ei ole kaks eraldi moodulit, vaid moodustavad pidevustiku. Seetõttu on otstarbekas ka nende kirjeldamine ühtses formaadis ja ressursis. Grammatika kirjeldamine sõnaraamatulaadses vormis ei ole iseenesest uus idee, vaid selleks on nähtud vajadust ka varem. Näited grammatika kirjeldamisest sõnaraamatulaadses vormis on valentsi- ja reksiooni-sõnaraamatud ja süntaktiliste mallide sõnaraamatud (*pattern dictionary*) (Colman, Tiberius 2018, Hanks 2004, 2008, Herbst, Uhrig 2019); morfoloogiasõnaraamatud (Viks 1992, 2008); sõnastikulaadsed ressursid, kus keele grammatika on esitatud Euroopa keeleõppe raamdokumendi tasemete kaupa (nt English Grammar Profile (Harrison 2015, O’Keeffe, Mark 2022)); sõnamoodustussõnaraamatud (nt Vare 2002); tuletusliidete esitamine traditsioonilise sõnaraamatu märksõnadena (nt TLFi).

Konstruktikoni loomise rakenduslikke põhjendusi on mitu. Esiteks on sõnade grammatilise käitumise kirjeldamine juba praegu üks traditsioonilise leksikograafia eesmärk tähenduse kirjeldamise kõrval. Näiteks EKI ühendsõnastikus (Langemets jt 2021) on sel eesmärgil esitatud sõna täielik vormiparadigma, sõnaliik, naabersõnad, rektisioon, grammatilised kommentaarid ja näitelauseid. Samuti sisaldab sõnaraamat grammatilisi sõnu, nagu *et*, *kui* jne. Praegu sõnaraamatutes kasutusel olevate sõnade süntaktilise käitumise kirjeldamise vahendite võimalused on aga piiratud. Nii ei ole võimalik kirjeldada sõnade esinemist konstruktsioonides, mida ei saa moodustada kõigi mingisse sõnaliiki kuuluvate sõnadega. Teiselt poolt ei saa selliste konstruktsioonide komponente täielikult kirjeldada ka traditsioonilise grammatika formaadis, sest see eeldab loetelu sõnadest, mis saavad selles konstruktsioonis esineda. Grammatikates aga tavaliselt ei esitata sellisel puhul täielikku loetelu, vaid üksnes mõned näited ja tähenduse iseloomustus. Näiteks võib tuua kvantorina toimivate nimisõnade kirjelduse EKI ühendsõnastikus ja nimisõnalise hulgafraasi kirjelduse “Eesti grammatikas” (Metslang jt 2023: 756–759). Grammatikas on kirjeldatud nimisõnalise hulgafraasi süntaksit ning loetletud kvantorina esinevate nimisõnade tähendusrühmad ja toodud igast rühmast mõned näited, nagu *liiter (piima)*, *klaas (vett)*, *komplekt (mööblit)*, *osa (mu sõpru)*, kuid grammatika formaat ei võimalda esitada selliste sõnade täielikku loetelu. Seevastu ühendsõnastikus on kvantorina esinevate nimisõnade märksõnaartiklis kirjeldatud eraldi tähendusplakis seda, et nad võivad väljendada kogust, kuid sõnaraamatus ei ole mingit viidet sellele, et kogusetähenduses käitub nimisõna süntaktiliselt hoopis teisiti kui muudes tähendustes. Sõnaraamatus ei ole võimalik nimisõnaliste kvantorite süntaktilist käitumist kirjeldada, sest seda ei saa teha sõnaliigimärgendiga, mis on praegu ainus viis ühendsõnastikus seda tüüpi süntaktilise info esitamiseks. Seda võib pidada puuduseks, sest nagu öeldud, sõnaraamatu koostamisel on sõnade kohta sellise info esitamine eesmärgiks seatud. Sõnaraamatus ja grammatikas oleva info koondamine ühte ressursi võimaldaks seega anda samal ajal täieliku ülevaate kvantorina kasutatavatest nimisõnadest ja nende süntaktilisest käitumisest.

Teiseks aitab konstruktkon parandada keeleinfo kättesaadavust. Kui eesti keele sõnavara kohta on info ühe klikiga leitav kaasaegsest veebiressursist (sonaveeb.ee; Koppel jt 2019), siis grammatika on kirjeldatud raamatu vormis (nt Metslang jt 2023, Kerge 2001), kust info ei pruugi olla kergelt leitav.

Kolmandaks võimaldaks konstruktkon koondada ja süstematiseerida juba ilmunud ja edaspidi ilmuvaid konstruktsioonipõhiseid käsitlusi eesti keele üksik-konstruktsioonidest (ülevaadet ilmunud käsitluste kohta vt Vainik jt 2024a). Konstruktsioonigrammatika teooria ennustab, et eesti keele grammatika sisaldab suurel hulgal erineva produktiivsuse- ja skemaatilisusastmega konstruktsioone, mida veel avastatud ega kirjeldatud ei ole. Konstruktkoni koostamine hõlmaks selliste konstruktsioonide süstemaatilist tuvastamist, kirjeldamist ja ühtsesse ressursi talletamist ning seega eesti keele kirjelduse täiendamist. Selline eesti keele süstemaatiline konstruktsioonipõhine kirjeldamine tähendaks osalt ka juba kirjeldatud keelenähtuste konstruktsioonipõhist ümbermõtestamist. Näiteks sõnade kombineerumist seotud laienditega on seni sõnaraamatutes näidatud sõnapõhiste rektisioonimallidena. Konstruktsioonipõhine lähenemine tähendaks rektisioonimallide käsitlemist skemaatiliste argumentstruktuurikonstruktsioonidena (Goldberg 1995), millel on nii vorm kui tähendus ja mis kombineeruvad erinevate sõnadega. Näiteks

põhjustatud liikumise konstruktsioonis [kes V mida (kust) kuhu] esinevad põhjustatud liikumist väljendavad verbid nagu *viima*, *viskama*, *saatma* jne. Et konstruktsioon ise kannab tähendust, ilmneb sellest, et põhjustatud liikumise tähendus säilib ka siis, kui konstruktsioonis esineb verb, mis ise seda tähendust ei kannab, nt *Vilistasin koera õuest tuppa*, *Ehmatasin kassi puu otsa*. Teatavaid eesti keele argumentstruktuurikonstruktsioone ongi traditsiooniliselt kirjeldatud sõnaütleste lausetüüpidega, mida on iseloomustatud tähenduse kaudu (eksistentsiaallause, kogeja-omaja lause jne, vt nt Metslang jt 2023: 498–500). Konstruktsioonipõhine kirjeldus tähendaks seega sellise lähenemise laiendamist ka neile argumentstruktuurikonstruktsioonidele, mida seni on kirjeldatud sõnastikes sõnapõhiste rektsoonimallidena – ehk senise keelekirjelduse ühtlustamist. Lisaks võiks see olla otstarbekas keeleõppe seisukohalt, sest rektsooni õppimine konstruktsioonipõhiselt ehk sarnase rektsooni ja tähendusega sõnarühmade kaupa võiks olla efektiivsem kui õppimine üksiksõnade kaupa (vt Patten, Perek 2022: 180–181).

Neljandaks saab konstruktikon toetada keeleõppijaid. Keeleõppe seisukohalt võiks konstruktikon täita praegu puuduva pedagoogilise õppijagrammatika funktsiooni (vt Pool jt 2024). Vastavalt pedagoogilise konstruktsioonigrammatika põhimõtetele tähendab keeleõppimine mitte üksikute sõnade ja grammatikareeglite, vaid konstruktsioonide õppimist (Herbst 2016: 18), mida tulebki esitada leksikaalse infoga seotuna vormi ja tähenduse paaridena.

Viiendaks saab konstruktikon teenindada erinevaid keeletehnoloogilisi vajadusi, mille fookuses on loomuliku keele analüüs (sh konstruktsioonide märgendus, vt nt Weissweiler jt 2024) ja loomuliku keele mõistmine (nt masintõlkesüsteemid, Borin jt 2018). Konstruktikoni andmeid saab kasutada ka generatiivsete AI mudelite treening- ja testandmetena (Torrent jt 2023).

Lõpetuseks plaanime linkida eesti konstruktikoni rahvusvahelise projekti Model of Comparative Concepts for Aligning Constructicons (MoCCA) andmebaasiga³ (Lorenzi jt 2024), mis võimaldab asetada eesti keele keeletüpoloogilisse konteksti ja seostada teiste keelte konstruktikonidega; näeme selles kasu nii keeleteooria edasiviimise jaoks kui ka rakenduslikult, teiskeelsete kasutajate huve silmas pidades.

4. Eesti konstruktikoni koostamise põhimõtted ehk kuidas?

Kuivõrd konstruktikoni põhiidee on kirjeldada grammatikat sõnastikulaadselt ja teha seda sõnastiku sees, on vaja määratleda samad dimensioonid nagu leksikograafias ikka – panna paika makrostruktuur ehk kirjeldatavate üksuste koosseis (4.1) ning mõelda läbi mikrostruktuur ehk millist infot ja mis järjekorras esitada (4.2). Konstruktikoni võrgustikusüsteemi loomisega seotud küsimused tuleb seejuures lahendada nii makro- kui mikrostruktuuri tasandil.

4.1. Makrostruktuur ehk konstruktikoni koosseis

Konstruktikoni makrostruktuuri koostamine algab küsimusest, millist tüüpi üksusi konstruktikon peaks sisaldama. Paratamatult tuleb konstruktikoni kuuluvatest konstruktsioonidest teha valik ning kõikmõeldavaid konstruktsioonitüüpe on esialgu keeruline määratleda. Konstruktsioonide täielikku nimekirja ei ole võimalik projekti alguses koostada juba seetõttu, et plaanime arendada sõltuvussüntaktiliselt märgendatud korpusest automaatselt uute konstruktsioonide tuvastamise meetodeid (Saul 2024), mis võib nimekirja lisada teoreetiliselt kirjeldamata konstruktsioone. Teine viis leida konstruktsioonide kandidaate on läheneda ülalt alla ehk koguda grammatikakirjeldustest ja käsitlustest kokku juhtumid ning registreerida need andmebaasis. Neid meetodeid saab kombineerida, mis on näiteks Rootsi konstruktikoni koostajate põhimõte (Lyngfelt jt 2018: 52).

Konstruktsioonide valikul lähtume peatükis 2 mainitud asjaolust, et konstruktsioone leidub erineval abstraktsuse ehk skemaatilisuse tasemel. Kuna nii lihtsaid kui kompleksseid leksikaalselt küllastunud üksusi (sõnu, mitmesõnalisi üksusi ja väljendeid) on eesti keele sõnaraamatutes, nt EKI ühendsõnastikus, juba kirjeldatud, keskendume konstruktikoniprojektis täielikult või osaliselt skemaatilistele konstruktsioonidele, mis võivad olla nii lihtsad (nt sõnamoodustuskonstruktsioonid) kui kompleksed (nt fraasikonstruktsioonid)⁴. Tabelis 1 anname ülevaate konstruktsioonitüüpidest, mida soovime pikemas perspektiivis eesti keele konstruktikonis esitada.

Tabel 1. Konstruktikoni kaasatavaid võimalikke konstruktsioonitüüpe

Konstruktsiooni tüüp	Näiteid
fraasikonstruktsioonid	nimisõnalise hulga fraasi konstruktsioon: <i>kott õunu, tass kohvi</i> (vt ptk 4.2)
argumentstruktuuri konstruktsioonid	soovimis-kavatsemiskonstruktsioon: <i>Ta tahab ujuda</i> (Penjam 2008: 75)
määruskonstruktsioonid	adessiivne viisi- ja põhjusemääruskonstruktsioon: <i>Raamat ilmus autori kirjastusel</i> (Sahkai 2006)
liitpredikaadikonstruktsioonid, mis ei sisaldu paradigmas	inhoatiivne passiivikonstruktsioon <i>minema/tulema</i> + <i>V_{misele}</i> : <i>Maja läheb lammutamisele</i> (Sahkai 2005)
sõnamoodustuskonstruktsioonid	teonimekonstruktsioon <i>V+mine</i> (Pilvik 2021)
lause põhitüübid	küsilause
kõrvallused	et-lause

Tegeliku koostamistöö käigus võib ilmned vajadus kaasata tabelis 1 nimetatata konstruktsioonitüüpe ning seetõttu peab kirjeldusformaat olema piisavalt paindlik ja kohandatav uut tüüpi konstruktsioonide kirjeldamiseks. Rootsi konstruktikoni eeskujul (Lyngfelt jt 2018: 58) võib valida tee, kus ei eeldata keerulisi konstruktsioonialamtüüpe, nagu näiteks mööndkõrvallause (Metslang jt 2023: 686), vaid omistatakse sellele konstruktsioonile kaks erinevat tüüpi: ühelt poolt möönduse ja teiselt poolt kõrvallause tüüp.

Makrostruktuuri puudutavad ka küsimused, kuidas konstruktikoni üldiselt organiseerida ja kuidas teostada konstruktikoni võrgustikuideed (vrd ptk 2).

⁴ Pool- ja täisskemaatiliste konstruktsioonide ehk vähemalt ühe lüngaga konstruktsioonide põhjal on teinud oma konstruktsioonivaliku ka näiteks rootsi keele konstruktikoni loojad (Lyngfelt jt 2018). Ungari keele konstruktikoniprojekt soovis alustada just leksikaalselt küllastatud konstruktsioonidest (Sass 2023).

Kavatseme esitada konstruktsioonide hierarhilisi pärilusseoseid (vt Vainik jt 2024a: 630), aga ka horisontaalseid (näiteks konstruktsiooni komponentide vahelisi) seoseid. Selleks peab olema selgelt määratletud koht konstruktsiooni mikrostruktuuris, mida kirjeldame järgmises peatükis. Konstruktsioonidevahelisi seoseid illustreerime hulgaformaasi näitel joonisel 4.

4.2. Mikrostruktuur ehk kirjete sisu

Konstruktsioonidel on konstruktikonis artikkel (ja andmebaasis kirje), nii nagu märksõnal on sõnastikus sõnaartikkel (ja andmebaasis kirje). Joonisel 1 võtame kokku konstruktikonikirje infoväljad, mis meie arvates on konstruktsiooni kirjeldamiseks vajalikud:

nimi
ekspertdile mõeldud nimi + prototüüpne näide
tavakasutajale mõeldud nimi + prototüüpne näide
definiitsioon
tüüp
näited
vorm
fraasitüüp
moodustus- ja sõltuvusstruktuur
süntaktiline funktsioon
komponendid
komponentide kirjeldus, lünga-täite suhted
tähendus
osalejate semantilised rollid
pragmaatilised omadused
suhtlusfunktsioon
register
võrdlusmõiste
konstruktsioonidevahelised seosed
ülemkateegooria
alamkateegooria
võrreldav kateegooria
vastanduv kateegooria
õppekommentaar
keeleoskustase
sagedus
produktiivsus
korpusepäring
otsisõnad
kirjandusviited

Joonis 1. Konstruktikoni kirje visand

Mõtestame konstruktsioonikirje väljad lahti kahe nimisõnalise hulgakonstruktsiooni näitel; nende konstruktsioonide kirjed on esitatud joonistel 2–3. Kuna oleme konstruktikoni põhilisteks sihtrühmadeks määratlenud ekspertkasutajad, keeleõppijad ja keeletehnoloogid, kajastuvad nende vaatenurgast olulised infokihid kirjes nii vormi kui tähenduse tasanditel; erinevatele kasutajarühmadele saab kuvada kirje erinevaid osi.

nimi:	eksperdile: loendamatu laiendiga nimisõnaline hulgafras: <i>tass kohvi</i> tavakasutajale: nimisõna väljendab aine hulka: <i>tass kohvi</i>
definiitsioon:	hulganimisõna määratleb ainenimisõnaga väljendatud aine hulka
tüüp:	fraasikonstruktsioon
näited:	<i>lusikatäis mett, kübe irooni, torn jäätist</i> [lingid andmebaasis olevatele konstruktsiooni esindavatele näidetele]
vorm:	fraasitüüp: hulgafras süntaktiline funktsioon: subjekt, objekt, määrus, öeldistäide, täiend väline moodustusstruktuur: HF sisemine moodustusstruktuur: [HF [[N [NF]]]] sõltuvusstruktuur: N → NF
komponendid:	N: väljendab kogust, mahutit, osa, vormi, liiki [lingid andmebaasis olevatele hulgasõnana esinevatele nimisõnadele] NF: väljendab ainet
tähendus:	osalejate semantilised rollid: [mõõt] _{Mõõt} mõõdab [ainet] _{Tema} mõõtev ASI + mõõdetavat väärtust väljendav AINE
pragmaatilised omadused:	suhtlusfunktsioon: infovahetus register: neutraalne
võrdlusmõiste:	cxn:partitive-construction; cxn:mensural-modification-construction, str:pseudopartitive
konstruktsioonidevahelised seosed:	ülemkategorია: hulgafras alamkategorია: – võrreldav kategorია: – vastanduv kategorია: asjahulgafras
õppekommentaari:	Pane tähele, kuidas muutub mõlema sõna vorm ainehulgafrasis, näiteks <i>kott suhkrut</i> . Kui hulgasõna (<i>kott</i>) on nimetavas käändes, on ainesõna (<i>suhkur</i>) ainsuse osastavas käändes (<i>kott suhkrut</i>). Kui hulgasõna (<i>kott</i>) on omastavas, on ainesõna (<i>suhkur</i>) kas osastavas (<i>Ostsin kotti suhkrut</i>) või omastavas käändes (<i>Maksin kotti suhkrut eest ühe euro</i>). Kui hulgasõna (<i>kott</i>) on osastavas, sisse-, sees- ja seestütlevas, alale-, alal- ja alaltütlevas või saavas käändes, on ainesõna (<i>suhkur</i>) samas käändes (<i>Ma ei ostmud kotti suhkrut, Mulle piisab kottist suhkrust</i>). Rajava, oleva, ilmatütleva ja kaasaütleva käände puhul on vastavas käändes ainult ainesõna (<i>suhkur</i>), kuid hulgasõna (<i>kott</i>) on alati omastavas käändes (<i>kotti suhkruta, kotti suhkruga</i>). Ainesõna (<i>suhkur</i>) on alati ainsuses (vastupidiselt asjasõnahulgafrasisile -> link asjahulgafrasi-konstruktsioonile).
keeleoskustase:	A2–B2
sagedus:	*
produktiivsus:	*
korpuspäring:	
märksõnad:	kogus
kirjandusviited:	EKG II (144–146), EG (752–762), Kerge 2001: 39–41)

Joonis 2. Ainehulgafrasi kirje visand⁵

Konstruktsiooni **nimi** on ühtlasi konstruktikoni märksõna. Et konstruktsioonid erinevad suuresti juba skemaatilise astme osas, on nende nimetamine üks konstruktikonide loomisega seotud suurtest väljakutsetest (Vainik jt 2024a, Vainik jt 2024b). Rootsi konstruktikoni (Lyngfelt jt 2018: 82) eeskujuks võttes piiritleme nime nii konstruktsiooni keeleteadusliku nimetuse kui ühe prototüüpse näite abil. Lisaks esitame tavakasutajale mõeldud lihtsustatud metakeeles nimetuse, kasutades võimalikult läbipaistvaid ja eesti keelde juurdunud termineid. Eksperdile mõeldud nime illustreerivad joonistel 2 ja 3 “loendamatu laiendiga nimisõnaline hulgafras: *tass kohvi*” ja “loendatava laiendiga nimisõnaline hulgafras: *kott õunu*”, milles on kasutatud eesti keeleteaduses traditsiooniliselt käibivaid termineid. Tavakasutajale mõeldud nimetustena pakume välja vastavalt “nimisõna väljendab aine hulka: *tass kohvi*” ja “nimisõna väljendab asjade hulka: *kott õunu*”.

Definiitsioon on mõeldud eelkõige tavakasutajale ja võtab lühidalt kokku konstruktsiooni vormi, tähenduse ja pragmaatilise funktsiooni. Näiteks loendamatu

⁵ Lühend HF tähistab hulgafrasi, N nimisõna, NF nimisõnafrasi ja V tegusõna.

nimi	ekspertdile: loendatava laiendiga nimisõnaline hulgafras: <i>kott õunu</i> tavakasutajale: nimisõna väljendab asjade hulka: <i>kott õunu</i>
definiitsioon:	hulganimisõna määratleb asjanimisõnaga väljendatud asjade hulka
tüüp:	fraasikonstruktsioon
näited:	<i>rühm inimesi, valik filme, hummik pabereid</i> [lingid andmebaasis olevatele konstruktsiooni esindavatele näidetele]
vorm:	fraasitüüp: kvantorifraas süntaktiline funktsioon: subjekt, objekt, määrus, öeldistäide, täiend väline moodustusstruktuur: HF sisemine moodustusstruktuur: [HF [[N [NF]]]] sõltuvusstruktuur: N → NF
komponendid ja nende kirjeldus:	N: väljendab kogust, mahutit, osa, vormi, liiki [lingid andmebaasis olevatele hulgasõnana esinevatele nimisõnadele] NF: väljendab asja.
tähendus:	osalejate semantilised rollid: [mõõt] _{Mõõt} mõõdab [asju] _{Teema} mõõtev ASI + mõõdetavat väärtust väljendav ASI
pragmaatilised omadused	suhtlusfunktsioon: infovahetus register: neutraalne
võrdlusmõiste:	cxn: partitive-construction; cxn: mensural-modification-construction, str: pseudo-partitive
konstruktsioonidevahelised seosed:	ülemkategorია: hulgafras alamkategorია: – võrreldav kategorია: – vastanduv kategorია: ainehulgafras
õppekommentaar:	Pane tähele, kuidas muutub mõlema sõna vorm asjahulgafrasis, näiteks <i>kott õunu</i> . Kui hulgasõna (<i>kott</i>) on nimetavas käändes, on asjasõna (<i>õunad</i>) mitmuse osastavas käändes (<i>kott õunu</i>). Kui hulgasõna (<i>kott</i>) on omastavas, on asjasõna (<i>õunad</i>) mitmuse osastavas (<i>Ostsin kotti õunu</i>) või mitmuse omastavas käändes (<i>Maksin kotti õunte eest ühe euro</i>). Kui hulgasõna on osastavas, sisse-, sees- ja seestütlevas, alale-, alal- ja alaltütlevas või saavas käändes, on asjasõna (<i>õunad</i>) samas käändes (<i>Ma ei ostnud kotti õunu, Mulle piisab kotist õuntest</i>). Rajava, oleva, ilmutitava ja kaasatitava käände puhul on vastavas käändes ainult ainesõna (<i>õunad</i>), kuid hulgasõna on omastavas käändes (<i>koti õunteta, koti õuntega</i>). Asjasõna on alati mitmuses (vastupidiselt ainehulgasõnafrasaile -> link ainehulgakonstruktsioonile).
keeleoskustase:	A2–B2
sagedus:	*
produktiivsus:	*
corpusepäring:	
märksõnad:	kogus
kirjandusviited:	EKG II (144–146), EG (752–762), Kerge 2001: 39–41)

Joonis 3. Asjahulgafrasi kirje visand

laiendiga nimisõnalise hulgakonstruktsiooni definiitsioon (joonis 2) on “hulganimisõna tähistab ainenimisõnaga väljendatud aine hulka”.

Konstruktsiooni tüüp valitakse tabelis 1 esitatud tüüpide hulgast. Näiteks hulgakonstruktsioon kuulub fraasikonstruktsioonide hulka. Konstruktsiooni olemust illustreerivad kasutus**näited**, mida võiks olla vähemalt kolm (joonisel 2 illustreeritud hulgafrasi puhul *tass kohvi, tonn jäätist, kübe irooniati*).

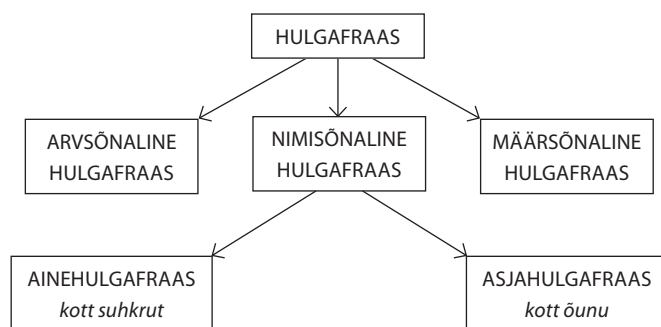
Konstruktsiooni **vormi** kirjeldatakse nii tervikkonstruktsiooni (fraasitüüp, süntaktiline funktsioon, moodustus- ja sõltuvusstruktuur) kui komponentide morfosüntaktilise struktuuri põhjal, seega jaguneb süntaktiline kirjeldus mõtteliselt kahte ossa: välissüntaks ehk kogu konstruktsiooni vormiklass (sõnaliik või fraasitüüp) ning funktsioon ja sisesüntaks ehk konstruktsiooni liikmed ja nende omavahelised suhted (Ohara 2018: 152). Hulgafrasi väline moodustusstruktuur vastab hulgafrasaile (HF) ning väliselt süntaktiliselt funktsioonilt võib ta olla subjekt, objekt, määrus, öeldistäide, täiend. Hulgafrasi sisemine moodustusstruktuur on [[N][NF]] ja sisemine sõltuvusstruktuur N → NF (nimisõnafrasa laiendab nimisõna).

Eraldi infoväli on mõeldud konstruktsiooni **komponentide** kirjelduse jaoks. Siin esitatakse ka poolproduktiivsete konstruktsioonide varieeruvate komponentide ehk lünkadega seotud piirangud ning lingid konstruktikoniüksustele (sõnadele), mis lüngas esineda saavad. Näiteks joonisel 2 kirjeldatud hulgafraasi põhjaks olev nimisõna väljendab kogust, mahutit, osa, vormi või liiki ja seda laiendav nimisõnafraas tähistab ainet. Kavas on esitada ka konstruktsioonis esinevate leksikaalsete üksuste (sõnade) ja konstruktsiooni enda nn kollostruktsiooni seosetugevus (vt Stefanowitsch, Gries 2003) ning keeleoskustase, millel konkreetse sõna ja konstruktsiooni kombinatsioon omandatakse.

Tähenduse kirjeldus vajab veel tõsist läbimõtlemist, kuid kindlasti hõlmab see kogu konstruktsiooni semantilist kirjeldust (liikumine, suund, seisund, põhjustamine jne) ning konstruktsiooni osavõtjarollide (agent, teema, kogeja, instrument jne) määramist. Näiteks hulgafraasi kirje (joonis 2) poolformaliseeritud vabatekstiline kirjeldus võiks olla [mõõt]_{Mõõt} mõõdab [ainet]_{Teema}. Semantilist kirjeldust täiendavad konstruktsiooniga võimalikult seotud **pragmaatilised** funktsioonid, nagu viisakus, hinnang, infovahetus jt. Lisaks on osa konstruktsiooni kasutusest ka register, mida saab tähistada näiteks stiilmärgendite abil konstruktikoni kirjes. Näide stiilmärgendit KÕNEKEELNE nõudvast konstruktsioonist on produktiivne liitpredikaadikonstruktsioon, mis koosneb tegusõna *minema* kolmanda isiku vormist ja translatiivis nimisõnast ning mis väljendab inhoatiivset aspekti (nt *läheb kakhuseks*).

Järgneb **võrdlusmõiste** ID keeltevahelisi vastavusi esitavas andmebaasis MoCCA. Hulgafraasi võrdlusmõisted joonistel 2 ja 3 seostavad hulgafraasi näiteks partitiivkonstruktsiooniga ja pseudopartitiivi strateegiat kasutava hulgafräendi konstruktsiooniga (vt täpsemalt Croft 2022, ptk 5.5.2).

Kirjes kajastuvad konstruktsiooni võimalikud vertikaalsed ja horisontaalsed seosed. Konstruktsiooni vertikaalset seost konstruktsioonide võrgustikus kajastavad konstruktsiooni **ülem- ja alamkategoriad**. Näiteks nii loendamatu kui loendatava laiendiga nimisõnalise hulgafraasi ülemkategoria on nimisõnaline hulgafraas, mille ülemkategoria omakorda on hulgafraas (vt joonis 4). Ainesõna-hulgafraas (joonis 2) sarnaneb (**võrreldav kategoria**) ja samas ka kontrasteerub (**vastanduv kategoria**) sama taseme asjahulgafrasiga (joonis 3), mis on horisontaalsed seosed.



Joonis 4. Konstruktsioonidevahelised seosed hulgafraside näitel

Õppekommentaari väljal saab keeleõppijale selgitada konstruktsiooniga seotud vaeohtlikke kohti. Kirjes on ka koht konstruktsiooni omandamise madalaimale **keeleoskustasemele**, mille määramisel lähtume õpiku- ja õppijakorpuste analüüsist ning EKI-s koostatud grammatikapädevuse kirjeldusest (Kallas jt 2021).

Konstruktsiooni kasutust iseloomustavad tema **esinemissagedus** ning **produktiivsusanäitaja**. Produktiivsuse mõõtmise aluseks võtame realiseerunud produktiivsuse ehk tüübisageduse ning potentsiaalse produktiivsuse ehk *hapax*'ite ja sõnede suhte (nt Baayen, 2009, Muischnek, Sahkai 2010, Pilvik 2021). Nii sagedust kui produktiivsust väljendame kolmeastmelisel skaalal, lähtudes sama tüüpi konstruktsioonide sagedusest ja produktiivsusest üksteisega võrreldes (harv/ebaproductiivne, keskmise sageduse/productiivsusega, sage/productiivne).

Igale konstruktsioonile lisatakse **korpusepäring**, mis võimaldab konstruktikoniga liidestatud korpusest konstruktsiooni esindavaid näiteid pärida ning on vajalik ka keeletehnoloogiliste rakenduste jaoks. Kuna nimisõnalised kvantorid märgendatakse näiteks Universal Dependencies märgendussüsteemis⁶ (Muischnek, Müürisep 2016) samamoodi nagu nimisõnade nimisõnalised täiendid, kujuneb otsingupäring hulga fraasi puhul üsna keerukaks; see peab hõlmama loetelu kvantorina esineda saavatest nimisõnadest, mis on nimisõnaga nimisõnalise täiendi seoses (nmod); lisaks tuleb täpsustada otsitavate sõnade morfoloogilist vormi, et välistada päringutulemusest otsitavate hulganimisõnade muud kasutused nimisõnalise täiendina.

Otsisõnad võimaldavad kasutajal konstruktsioone erinevate tunnuste alusel konstruktikonist leida. Näiteks nimisõnalise hulga fraasi konstruktsioonid varustatakse otsisõnaga "kogus", mis aitab kasutajal leida nii need konstruktsioonid kui ka näiteks määrsõnalise ja arvsõnalise hulga fraasi konstruktsiooni (*palju õunu, kaks õuna*), elatiivse laiendiga hulga konstruktsiooni (*katkend ooperist*) jt.

Kirje lõpuvälja saab koondada viited konstruktsiooniga seotud uurimistöödele.

5. Kokkuvõte ja edasised sammud

Tutvustasime artiklis oma nägemust sellest, kuidas esitada grammatilist infot eesti keele kohta sõnastikulaadses ressursis ehk konstruktikonis. Põhjendasime vajadust sellise ressursi järele nii teoreetilisest kui rakenduslikust vaatenurgast.

Teoreetiliselt lähtub konstruktikoni vajadus konstruktsioonipõhisest keeleteooriast, mis käsitleb sõnavara ja grammatikat ühe pidevustikuna, mistõttu on otstarbekas neid kirjeldada ühtses formaadis. See lähenemine võimaldab süstemaatiliselt kirjeldada ja koondada ka konstruktsioone, mida pole võimalik klassikalises grammatikas täpselt esitada.

Praktilistest vajadustest lähtudes aitaks konstruktikon ühtlustada ja täiendada olemasolevaid leksikaalseid ja grammatilisi kirjeldusi, pakkudes ühtset veebipõhist keeleressurssi, mis oleks hõlpsasti kättesaadav nii õppijatele, teadlastele kui ka keeletehnoloogia arendajatele. See parandaks ligipääsu keeleinfole ja oleks rakendatav keeleõppes, aidates omandada sõnavara ja grammatikat terviklikumalt. Lisaks toetaks konstruktikon keeletehnoloogilisi vajadusi, sealhulgas loomuliku keele analüüsi, masintõlget ja dialoogisüsteemide arendamist.

Artikli üks eesmärke oli piiritleda konstruktikoni ülesehituse põhimõtted ehk selle makro- ja mikrostruktuur. Määratlesime laias laastus, millist tüüpi konstruktisioone tahame kirjeldada, kuidas esitada nendevahelisi suhteid ning millised infokihid peavad kuuluma konstruktisiooni kirjesse. Tõdesime, et konstruktikoni kasutajaskonna heterogeensus avaldab mõju nii kavandatava ressursi makro- kui mikrostruktuurile. Kavas on haarata laia konstruktisioonitüüpide skaalat ning esitada infot eraldi metakeeltes: üldarusaadavas, lingvistiliselt täpses ning masinloetavas, formaliseeritud keeles. Illustreerimaks potentsiaalseid konstruktikoniüksusi, esitasime visandid “mitmepalgelise” hulga fraasi kirjeldamiseks.

Käsitlusest jäid välja konstruktikoni loomisega seotud tehnoloogilised küsimused ning nendega seotud võimalused ja piirangud. Et meie esialgne plaan on siduda konstruktikon mingil moel EKI ühendsõnastikuga, toob see kindlasti kaasa uusi kaalutlusi ja otsustusi, mille läbiarutamine on järgmine samm. Kuna ühendsõnastiku üks pakilisi grammatika esitamise ülesandeid on kollokatsioonide ja rektisioonide esituse ümberkorraldamine, tundub mõttekas lahendada see koos teistegi konstruktisioonide süstemaatilise esitusega konstruktikonina. Olemasolevate üksuste uuel viisil esitamise teed on läinud nt ungari keele konstruktikoni projekt (vt Sass 2023). Seega tundub loomulik mõelda, et EKI ühendsõnastiku andmebaas ning sõnastiku- ja terminibaasisüsteem Ekilex (Tavast jt 2018) saaks kodus ka eesti konstruktikonile. Kuivõrd me ei tunne veel kogu konstruktisioonide mitmekesisuse spektrit, on raske ennustada, kas kõigi nende kirjeldamiseks sobib ideaalselt just Ekilexi andmemudel, mis on disainitud leksikograafilise info talletamiseks ning toimetamiseks. Loodetavasti on võimalik sinna konstruktisioonide jaoks lisamoodulite abil omaette ja paindlikumad kirjeldusvõimalused luua. Praegu ei ole veel selge, kas ja kuidas mahutada konstruktisioonide kirjeid ühendsõnastikku (vt esialgseid ideid Vainik jt 2024a).

Teine valdkond, mis vajab edaspidi läbiuurimist, tööprotsesside loomist ja valideerimist, on konstruktisioonide tuvastamine märgendatud korpustest, märgendatud näidete importimine ning kuvamine; konstruktisioonide kohta sagedusstatistika ning seosetugevusnäidike arvutamine. Üks suuri väljakutseid saab olema ka tähenduse ja seoste esitamise formaadi väljatöötamine.

Lühendid

HF	hulga fraas
N	nimisõna
NF	nimisõna fraas
V	tegusõna

Viidatud kirjandus

- Baayen, Harald R. 2009. Corpus linguistics in morphology: Morphological productivity. – Anke Lüdeling, Merja Kytö (Eds.), *Corpus Linguistics: An International Handbook. Handbooks of Linguistics and Communication Science*, 29.2. Berlin–New York: Mouton de Gruyter, 900–919. <https://doi.org/10.1515/9783110213881.2.899>
- Booij, Geert E. 2010 *Construction Morphology*. Oxford Linguistics. Oxford: Oxford University Press.
- Borin, Lars; Dannélls, Dana; Grūzītis, Normunds 2018. Linguistics vs. language technology in constructicon building and use. – Benjamin Lyngfelt, Lars Borin, Kyoko Ohara,

- Tiago Timponi Torrent (Eds.), *Constructicography: Constructicon Development Across Languages. Constructional Approaches to Language*, 22. Amsterdam–Philadelphia: John Benjamins, 229–254. <https://doi.org/10.1075/cal.22.08bor>
- Borin, Lars; Lyngfelt, Benjamin 2025. Framenets and ConstructiCons. – Mirjam Fried, Kiki Nikiforidou (Eds.), *The Cambridge Handbook of Construction Grammar*. Cambridge: Cambridge University Press, 71–100. <https://doi.org/10.1017/9781009049139.004>
- Colman, Lut; Tiberius, Carole 2018. A Good Match: A Dutch Collocation, Idiom and Pattern Dictionary Combined. – *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts*. Ljubljana: University of Ljubljana Press, 233–246.
- Croft, William 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198299554.001.0001>
- Croft, William 2016. Comparative concepts and language-specific categories: Theory and practice. – *Linguistic Typology*, 20 (2), 377–393. <https://doi.org/10.1515/lingty-2016-0012>
- Croft, William 2022. *Morphosyntax: Constructions of the World's Languages*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781316145289>
- Diessel, Holger 2019. *The Grammar Network: How Linguistic Structure Is Shaped by Language Use*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108671040>
- Diessel, Holger 2023. *The Constructicon: Taxonomies and Networks. Elements in Construction Grammar*. <https://doi.org/10.1017/9781009327848>
- Fried, Mirjam; Östman, Jan-Ola 2004. *Construction Grammar in a Cross-Language Perspective. Constructional approaches to language*, 2. Amsterdam–Philadelphia: John Benjamins. <https://doi.org/10.1075/cal.2>
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure. Cognitive Theory of Language and Culture*. Chicago: University of Chicago Press.
- Goldberg, Adele E. 2003. Constructions: A New Theoretical Approach to Language. – *Trends in Cognitive Sciences*, 7 (5), 219–224. [https://doi.org/10.1016/S1364-6613\(03\)00080-9](https://doi.org/10.1016/S1364-6613(03)00080-9)
- González-García, Francisco 2019. Exploring the pedagogical potential of vertical and horizontal relations in the constructicon: The case of the family of subjective-transitive constructions with *decir* in Spanish. – *International Review of Applied Linguistics in Language Teaching*, 57 (1), 121–145. <https://doi.org/10.1515/iral-2018-2009>
- Gras, Pedro; Elvira-García, Wendy 2021. The role of intonation in construction grammar: On prosodic constructions. – *Journal of Pragmatics*, 180, 232–247. <https://doi.org/10.1016/j.pragma.2021.05.010>
- Hanks, Patric 2004. *Corpus Pattern Analysis*. – *Proceedings of the 11th EURALEX International Congress, EURALEX 2004, Lorient, France, July 6–10, 2004*, 87–98.
- Hanks, Patric 2008. *Mapping Meaning Onto Use: A Pattern Dictionary of English Verbs*. <https://slideplayer.com/slide/4126510/> (19.3.2025).
- Harrison, Julia 2015. *The English Grammar Profile*. – Julia Harrison, Fiona Barker (Eds.), *English Profile in Practice. English Profile Studies*, 5. Cambridge: Cambridge University Press, 28–48.
- Haspelmath, Martin 2010. Comparative concepts and descriptive categories in crosslinguistic studies. – *Language*, 86 (3), 663–687. <https://doi.org/10.1353/lan.2010.0021>
- Herbst, Thomas 2016. Foreign language learning is construction learning – what else? Moving towards pedagogical construction grammar. – Sabine De Knop Gaëtanelle Gilquin (Eds.), *Applied Construction Grammar. Applications of Cognitive Linguistics*, 32. Berlin–New York: De Gruyter, 21–52. <https://doi.org/10.1515/9783110458268-003>
- Herbst, Thomas; Uhrig, Peter 2019. Towards a valency and argument structure constructicon of English: Turning the valency patternbank into a constructicon. – *Lexicographica*, 35, 87–104. <https://doi.org/10.1515/lex-2019-0006>

- Hoffmann, Thomas; Trousdale, Graeme 2013. *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press.
- Hudson, Richard A. 2007. *Language Networks: The New Word Grammar*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780199267309.001.0001>
- Kallas, Jelena; Koppel, Kristina; Pool, Raili; Tsepelina, Katrin; Üksik, Tiit; Alp, Pilvi; Epner, Anu 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>
- Kerge, Krista 2001. Eesti süntaks keeleõppe praktikule. TEA Kirjastus.
- Koppel, Kristina; Tavast, Arvi; Langemets, Margit; Kallas, Jelena 2019. Aggregating dictionaries into the language portal Sõnaveeb: Issues with and without a solution. – Proceedings of the eLex 2019 Conference. 1–3 October 2019, Sintra, Portugal. Brno: Lexical Computing CZ, s.r.o., 434–452.
- Langemets, Margit; Koppel, Kristina; Kallas, Jelena; Tavast, Arvi 2021. Sõnastikukogust keeleportaalis. – Keel ja Kirjandus, 8–9, 755–770. <https://doi.org/10.54013/kk764a6>
- Langemets, Margit; Hein, Indrek; Jürviste, Madis; Kallas, Jelena; Kiisla, Olga ... Ossipstšuk, Inna 2023. EKI ühend sõnastik 2023. Eesti Keele Instituut.
- Lorenzi, Arthur; Ljunglöf, Peter; Lyngfelt, Ben; Timponi, Torrent Tiago; Croft, William; Ziem, Alexander; Böbel, Nina; Bäckström, Linnéa; Uhrig, Peter; Matos, Ely E. 2024. MoCCA: A Model of Comparative Concepts for Aligning Constructicons. – Proceedings of the 20th Joint ACL – ISO Workshop on Interoperable Semantic Annotation @ LREC-COLING 2024. ELRA, ICCL, 93–98. <https://aclanthology.org/2024.isa-1.12>
- Lyngfelt, Benjamin 2018. Introduction: Constructicons and constructicography. – Benjamin Lyngfelt, Lars Borin, Kyoko Ohara, Tiago Timponi Torrent (Eds.), *Constructicography: Constructicon Development Across Languages. Constructional Approaches to Language*, 22. Amsterdam–Philadelphia: John Benjamins, 1–18. <https://doi.org/10.1075/cal.22.01lyn>
- Lyngfelt, Benjamin; Bäckström, Linnéa; Borin, Lars; Ehrlemark, Anna; Rydstedt, Rudolf 2018. Constructicography at Work. Theory Meets Practice in the Swedish Constructicon. – Benjamin Lyngfelt, Lars Borin, Kyoko Ohara, Tiago Timponi Torrent (Eds.), *Constructicography: Constructicon Development Across Languages. Constructional Approaches to Language*, 22. Amsterdam–Philadelphia: John Benjamins, 41–106. <https://doi.org/10.1075/cal.22.03lyn>
- Lyngfelt, Benjamin; Torrent, Tiago; Eli, Edison da Silva Matos; Bäckström, Linnéa 2022. Comparative Concepts as a Resource for a Multilingual Constructicography. – Kristian Blenselius (Eds.), *Valency and Constructions: Perspectives on Combining Words*. Göteborg: Meijerbergs institut för svensk etymologisk forskning, 101–129.
- Metslang, Helle; Erelt, Mati; Habicht, Külli; Hennoste, Tiit; Kasik, Reet; Teras, Pire; Viht, Annika; Asu-Garcia, Eva Liina; Lindström, Liina; Lippus, Pärtel; Pajusalu, Renate; Plado, Helen; Rääbis, Andriela; Veismann, Ann 2023. Eesti grammatika. Tartu: Tartu Ülikooli Kirjastus. <https://doi.org/10.12697/EG>
- Muischnek, Kadri; Müürisep, Kaili 2016. Eesti keele sõltuvuspuude pank ja selle keeleteoreetilised lähtekohad. – Emakeele Seltsi aastaraamat, 62 1, 122–145. <https://doi.org/10.3176/esa62.04>
- Muischnek, Kadri; Sahkai, Heete 2010. Liitpredikaadid leksikoni-grammatika kontiinumil: Konstruktsioonide produktiivsusest verbiga minema moodustatud liitpredikaatide näitel. – ESUKA/JEFUL, 1 (2), 295–316. <https://doi.org/10.12697/jeful.2010.1.2.16>
- Ohara, Kyoko 2018. Relations Between frames and constructions: A proposal from the Japanese FrameNet constructicon. – Benjamin Lyngfelt, Lars Borin, Kyoko Ohara, Tiago Timponi Torrent (Eds.), *Constructicography: Constructicon Development Across*

- Languages. Constructional Approaches to Language, 22. Amsterdam–Philadelphia: John Benjamins, 141–163. <https://doi.org/10.1075/cal.22.05oha>
- O’Keeffe, Anne; Mark, Geraldine 2022. The English Grammar Profile of Learner Competence: Methodology and Key Findings. – International Journal of Corpus Linguistics, 22 (4), 457–489. <https://doi.org/10.1075/ijcl.14086.oke>
- Patten, Amanda L.; Perek, Florent 2022. Pedagogic applications of the english constructicon. – Hans C. Boas (Eds.), Directions for Pedagogical Construction Grammar: Learning and Teaching (with) Constructions. Applications of Cognitive Linguistics, 49. Berlin–Boston: De Gruyter, 179–216. <https://doi.org/10.1515/9783110746723-007>
- Penjam, Pille 2008. Eesti kirjakeele *da-* ja *ma-*infinitiiviga konstruktsioonid. Dissertationes philologiae Estonicae Universitatis Tartuensis, 23. Tartu: Tartu Ülikooli kirjastus.
- Pilvik, Maarja-Liisa 2021. Action Nouns in a Constructional Network: A corpus-based Investigation of the Productivity and Functions of the Deverbal Suffix *-mine* in Five Different Registers of Estonian. Dissertationes philologiae Estonicae Universitatis Tartuensis, 48. Tartu: University of Tartu Press.
- Pool, Raili; Kallas, Jelena; Vainik, Ene 2024. Eesti keele kui teise keele tasemepõhine pedagoogiline õppijagrammatika veebikeskkonnas. – Eesti Rakenduslingvistika Ühingu aastaraamat, 20, 199–220. <https://doi.org/10.5128/ERYa20.12>
- Saul, Kertu 2024. Eesti keele lausemallide automaatne tuvastamine liigutamisverbide näitel. MA. Tartu Ülikool. <https://hdl.handle.net/10062/102023>
- Sahkai, Heete 2005. Teonimi perifrastilises verbivormis. – Keel ja Kirjandus, 10, 790–807.
- Sahkai, Heete 2006. Konstruktsioonipõhise keelekirjelduse võimalustest adessiivse viisi- ja põhjusemääruse näitel. – Keel ja Kirjandus, 10, 816–831.
- Sahkai, Heete 2011. Teine grammatika: Eesti keele teonimede süntaks konstruktsioonipõhises perspektiivis. Tallinna Ülikooli humanitaarteaduste dissertatsioonid, 25. Tallinn: Tallinna Ülikool.
- Sass, Bálint 2023. From a dictionary towards the Hungarian constructicon. – Electronic Lexicography in the 21st Century: Invisible Lexicography. Proceedings of the eLex 2023 Conference. Brno, 27–29 June 2023. Brno: Lexical Computing CZ s.r.o., 534–544.
- Stefanowitsch, Anatol; Gries, Stefan Th. 2003. Collocations: Investigating the Interaction of Words and Constructions. – International Journal of Corpus Linguistics, 8 (2), 209–243. <https://doi.org/10.1075/ijcl.8.2.03ste>
- Svensén, Bo 2019. A Handbook of Lexicography: The Theory and Practice of Dictionary-Making. Cambridge: Cambridge University Press.
- Tavast, Arvi; Langemets, Margit; Kallas, Jelena; Koppel, Kristina 2018. Unified data modelling for presenting lexical data: The case of EKILEX. – Proceedings of the XVIII EURALEX International Congress: EURALEX: Lexicography in Global Contexts, Ljubljana, 17–21 July 2018. Ljubljana: Ljubljana University Press, 749–761.
- TLFi = Trésor de la langue Française informatisé. ATILF – CNRS & Université de Lorraine. <http://www.atilf.fr/tlfi>
- Torrent, Tiago Timponi; Hoffmann, Thomas; Almeida, Arthur L.; Turner, Mark 2023. Copilots for Linguists: AI, Constructions, and Frames. Elements in Construction Grammar. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781009439190>
- Traugott, Elisabeth C.; Trousdale, Graeme 2013. Constructionalization and Constructional Changes. Oxford: Oxford University Press.
- Vainik, Ene; Paulsen, Geda; Kallas, Jelena 2024a. Sõnastikust konstruktioniks? Taustu, eeskujusid ja väljakutseid. – Keel ja Kirjandus, 7, 626–644. <https://doi.org/10.54013/kk799a2>
- Vainik, Ene; Paulsen, Geda; Sahkai, Heete; Kallas, Jelena; Tavast, Arvi; Koppel, Kristina 2024b. From a dictionary to a constructicon: Putting the basics on the map. – Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress. Institute for the Croatian Language.

- Van de Velde, Freek 2014. Degeneracy: The maintenance of constructional networks. – Ronny Boogaart, Timothy Coleman, Gijsbert Rutten (Eds.), *Extending the Scope of Construction Grammar*. Cognitive Linguistics Research, 54. Berlin–Boston: De Gruyter Mouton, 141–180. <https://doi.org/10.1515/9783110366273.141>
- Vare, Silvi 2002. Eesti keele sõnapere sõnaraamatu koostamise põhimõtetest. – *Keel ja Kirjandus*, 2, 855–866.
- Viks, Ülle 1992. Väike vormisõnastik. 1: Sissejuhatus ja grammatika; 2: Sõnastik & lisad. Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut. Tallinn.
- Viks, Ülle 2008. Eesti-X-keele sõnastik ja grammatika. – *Eesti Rakenduslingvistika Ühingu aastaraamat*, 4, 247–261. <https://doi.org/10.5128/ERYa4.15>
- Weissweiler, Leonie; Böbel, Nina; Guiller, Kirian; Herrera, Santiago; Scivetti, Wesley; Lorenzi, Arthur; Melnik, Nurit; Bhatia, Archana; Schütze, Hinrich; Levin, Lori; Zeldes, Amir; Nivre, Joakim; Croft, William; Schneider, Nathan 2024. UCxn: Typologically Informed Annotation of Constructions Atop Universal Dependencies. arXiv:2403.17748. <https://doi.org/10.48550/arXiv.2403.17748>

Geda Paulsen (Eesti Keele Instituut, Uppsala Ülikool) on uurinud sõnamoodustuskonstruksioone ning leksikaalset semantikat.

Roosikrantsi 6, 10119 Tallinn, Estonia
geda.paulsen@eki.ee, geda.paulsen@moderna.uu.se

Heete Sahkai (Eesti Keele Instituut) on uurinud eesti keele süntaksit ja prosoodiat.

Roosikrantsi 6, 10119 Tallinn, Estonia
heete.sahkai@eki.ee

Ene Vainik (Eesti Keele Instituut) on uurinud leksikaalset semantikat, hinnangu väljendamist ning kujundlikku keelekasutust; viimasel ajal leksikograafia ja kosnstruktikograafia küsimusi.

Roosikrantsi 6, 10119 Tallinn, Estonia
ene.vainik@eki.ee

Jelena Kallas (Eesti Keele Instituut) juhib Eesti Teadusagentuuri granti “Uue aja sõnastik: grammatika ja keelepädevuse kirjeldamine integreeritud multifunktsionaalses leksikograafilises ressursis” (PRG 1978). Põhilised uurimisvaldkonnad on korpuslingvistika, korpusleksikograafia ja õppeleksikograafia.

Roosikrantsi 6, 10119 Tallinn, Estonia
jelena.kallas@eki.ee

Kertu Saul (Eesti Keele Instituut, Tartu Ülikool) on uurinud arvutilingvistiliste meetoditega eesti keele morfosüntaksit.

Roosikrantsi 6, 10119 Tallinn, Estonia
kertu.saul@eki.ee

Arvi Tavast (Eesti Keele Instituut) on EKI direktor. Uurimisvaldkonnad: leksikaalse info esitus ja andmemudelid.

Roosikrantsi 6, 10119 Tallinn, Estonia
arvi.tavast@eki.ee

Kristina Koppel (Eesti Keele Instituut) on EKI ühendsõnastiku töörühma liige, sh sünonüümide ja inglise keele infokihi töörühma juht. Põhilised uurimisvaldkonnad: korpuslingvistika, korpusleksikograafia.

Roosikrantsi 6, 10119 Tallinn, Estonia
kristina.koppel@eki.ee

TOWARDS THE ESTONIAN CONSTRUCTICON

**Geda Paulsen^{1,2}, Heete Sahkai¹, Ene Vainik¹,
Jelena Kallas¹, Kertu Saul¹, Arvi Tavast¹, Kristina Koppel¹**

Institute of the Estonian Language¹, University of Uppsala²

This article outlines the main approach for presenting grammatical information on Estonian in a dictionary-like resource known as a constructicon. We justified the need for such a resource from both theoretical and practical perspectives. Theoretically, the need for a constructicon arises from construction-based linguistic theory, which treats vocabulary and grammar as a continuum, making it reasonable to describe them in a unified format. This approach would allow for the systematic description and compilation of constructions that cannot be precisely captured in traditional grammar.

From a practical standpoint, the constructicon would help standardize and enhance existing grammatical descriptions, offering a unified, web-based language resource that is easily accessible to learners, researchers, and language technology developers. It would improve access to linguistic information, as grammatical knowledge is often fragmented and available only in book form. The constructicon would have applications in language learning, helping users grasp word and phrase meanings more comprehensively. Additionally, it would support language technology needs, including natural language analysis, machine translation, and dialogue system development.

We thus defined the primary target audience for the constructicon as L2 learners and teachers of Estonian, though we are designing the resource to also serve language experts and language technology applications. The diversity of the target audience influences the macro- and microstructure of the proposed resource. The plan is to cover a wide range of construction types and provide descriptions in multiple metalevels: general, linguistically precise, and machine-readable. We presented sketches for describing complex quantifier phrases as constructicon entries that pose challenges for language learners.

Keywords: construction, constructicon, constructicography, lexicography, Estonian