

Sõnastikust konstruktikoniks?

Taustu, eeskujusid ja väljakutseid

ENE VAINIK, GEDA PAULSEN, JELENA KALLAS

1. Kapist välja

Kui aastal 2011 sai konstruksioone käsitleva väitekirja pealkirjaks olla „Teine grammatika” (Sahkai 2011), mis viitas selle lähenemise alternatiivsele iseloomule, siis aastal 2024 ei ole konstruksioonipõhine keelekäsitlus enam midagi uut ega erakordset. See on iseseisvunud mitmesuguseid keelenähtusi käsitlevaks kirjeldusraamistikuks (vt nt Goldberg 2003; Booij 2010; Gras, Elvira-García 2021); sel alal toimub juba 20 aastat eraldi konverentside seeria,¹ välja antakse spetsialiseeritud ajakirju² ning ilmunud on põhjalik käsiraamat (Hoffmann, Trousdale 2013). Konstruksiooni-põhise lähenemise põhiteese on, et keel ei jagune eraldi grammatikaks ja leksikoniks, vaid koosneb terveniisti erineva kompleksuse ja abstraktsusega konstruksioonidest, mis on – kõige lihtsamalt öeldes – tähenduse ja vormi paarid, kus nendevahelisel seosel on kanda roll terviku tähenduses. Näiteks saab lugeja käesoleva artikli pealkirjast „Sõnastikust konstruktikoniks” mingil määral aru isegi siis, kui ta ei tea, mida tähendab sõna *konstruktikon*. Ta mõistab, et tegu on sõnastiku muutumisega millekski muuks, sest nii tõlgendama suunab muutumistähendusega resultatiivkonstruksioon $\{[NF_{\text{elat}}] [NF_{\text{transi}}]\}$.³ Teisi näiteid samast konstruksioonist: *tütarlapsest naiseks*, *ajalehepõisist miljonäriks*, *pätist ettevõtjaks*.

Konstruksioonilise lähenemisviisi eeliseid tutvustades rõhutatakse esiteks selle psühholoogilist realistlikkust, kuivõrd tuginemine konstruksioonidele on osutunud empiirilisel tõendatuks nii esimese (nt Tomasello 2006; Diessel 2013) kui ka teise keele omandamist käsitlevates uurimustes (Ellis 2013), tõendeid on saadud ka psühho- ja neurolingvistika alalt (Bencini 2013; Pulvermüller jt 2013). Teiseks eeliseks on sobitumine kasutuspõhise lingvistika üldisse trendi, mis teoreetilise keele- võime ja uurija konstrueeritud lausungite grammatilise vastuvõetavuse asemel uurib tegelikult aset leidnud keelekasutusjuhtumeid ehk keeletarvitust, mis on dokumenteeritud korpustena (nt Langacker 1987; Bybee 2010, 2013). Keelekasutusmuutrite hindamisel peetakse oluliseks näitajaks nende kasutussagedust (nt Diessel 2019; Perek 2015). Grammatikat tervikuna ei käsitleta mitte abstraktse ja reeglipärase

¹ <https://www.cognitivelinguistics.org/en/event/detail/conferences-on-construction-grammar>

² <https://constructions.journals.hhu.de/>; <https://benjamins.com/catalog/cf>

³ Siin ja edaspidi kasutame konstruksioonide märkimiseks esitusviisi, kus tervikkonstruksioon on piiratud loogeliste sulgudega ning selle osised nurksulgudega (Fillmore jt 2012), seega $\{[NF_{\text{elat}}] [NF_{\text{transi}}]\}$ tuleks lugeda kui „substantiivfraas elatiivis koos substantiivfraasiga translatiivis käsitletuna ühe konstruksioonina”.

süsteemina iseeneses, vaid kasutusmustritest võrsuva dünaamilise ja vajadustega kohaneva süsteemina (Schmid 2020: 2).

Konstruktioonipõhisel lähenemisel on Eestiski arvestatav ajalugu (nt Penjam, Pajusalu 2006; Sahkai 2006) ning seda on korduvalt populariseeritud (Penjam 2009; Sahkai 2012). Käsitlusviis on valitud mitme doktoritöö teoreetiliseks raamistikuks (nt Penjam 2008; Sahkai 2011; Kitsnik 2018; Pilvik 2021) ja juba aastal 2008 on Heete Sahkai öelnud välja algselt Charles Fillmore'i (2006) sõnastatud mõtte, et konstruktioonidel võiks olla koht ka leksikograafias. Selle idee kohaselt võiks ja peaks konstruktioone esitama sõnaraamatus sarnaselt leksikaalsete üksustega, seega täiendades traditsioonilist sõnaraamatut „konstruktikoniga” – mooduliga, milles on võimalik kirjeldada produktiivsete konstruktioonide vormi, tähendust ja kombinatoorseid omadusi (Sahkai 2008). Vajadus sellise lähenemise järele võrsub ühest küljest sellest, et keelt grammatikaks ja leksikoniks jagav traditsiooniline vaade jättis hulga pool-skemaatilisi ja/või erandlikke kasutusmustreid keelekirjeldusest üldse välja. Teisest küljest valmistavad just sellised juhtumid peavalu keeleõppijatele ning nende kirjeldamine tundub hädavajalik (Lyngfelt jt 2018: 52).

Fillmore'i ideest on juba saanud alguse konstruktikonide koostamine mitmete keelte jaoks (vt lähemalt ptk 3), kuid eesti keelekirjeldusest on seni konstruktioone ja leksikograafiat siduv andmekogu puudunud. Nüüdseks on Eesti Keele Instituudis käivitunud teadusprojekt,⁴ mille raames uuritakse sellise andmekogu loomise võimalusi. Seejuures on võetud suund siduda konstruktioonide esitus andmebaasis olemasoleva leksikaalse ja keele omandamise tasemete infoga, et toetada keeleõppijat ja õppematerjalide koostajaid (pedagoogilise suunitlusega konstruktioonigrammatika kohta vt nt Boas 2022).

Siinse artikli eesmärk on kirjeldada raamistikku, kuhu eesti keele konstruktikoni loomise idee paigutub. Esmalt tutvustame konstruktioonigrammatilise keelekäsituse mõningaid põhimõisteid ning seejärel anname ülevaate uudest tegevusalast – konstruktikograafiast – ehk uurime järele, kuhu ollakse teiste keelte konstruktikonide loomisega jõutud, ning teeme kokkuvõtte, millised on olnud seniste konstruktikograafiliste algatuste suurimad väljakutsed. Artikli lõpuosas arutleme põgusalt võimaluste üle luua eesti keele jaoks konstruktikonilaadset ressursi.

2. Kändegrammatikast konstruktikograafiani

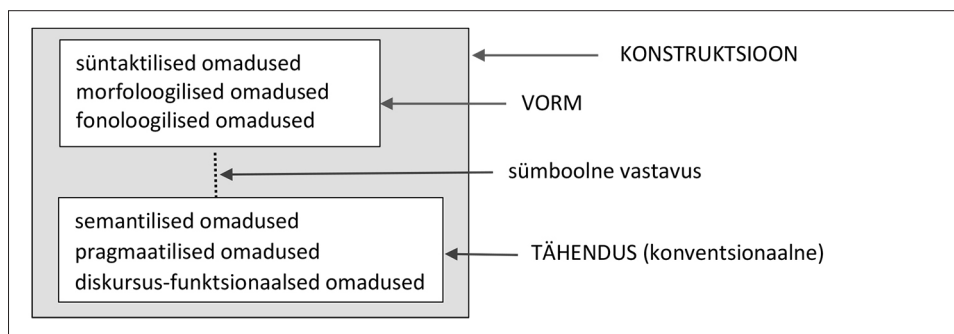
Konstruktioonigrammatika või laiemalt konstruktioonipõhine keelekäsitlus ühendab teoreetilisi suundi, mis lähtuvad eeldusest, et keeleline vorm ja tähendus on omavahel lahutamatult seotud (Langacker 1987; Fillmore jt 1988; Goldberg 1995, 2006; Kay, Fillmore 1999; Croft 2001; Östman, Fried 2004; Penjam, Pajusalu 2006; Penjam 2008; Sahkai 2011). Ühtset vormi ja tähenduse kooslust nimetataksegi *konstruktiooniks*. See mõiste viitab komplekssele, kuid terviklikule keelelise teadmise üksusele,

⁴ „Uue aja sõnastik: grammatika ja keelepädevuse kirjeldamine integreeritud multifunktsionaalses leksikograafilises ressursis” (PRG1978). <https://www.etis.ee/Portal/Projects/Display/66a57ac4-e0c6-40cd-b9df-148048ce9f95>

milles kindel vorm ühendub mingil kindlal viisil kindla tähendusega. Konstruksioone tuleb eristada konstruktidest ehk kasutusjuhtudest: konstruktsioonid on Paul Kay ja Charles Fillmore'i (1999: 2) järgi abstraktsioonid, „grammatika tükid”, konstruktid aga konstruktsioonide realisatsioonid tegelikus diskursuses.

Konstruksiooniteooria algne eesmärk oli tõsta keeleteaduse huviorbiiti (generatiivse grammatika vaatenurgast) perifeersed keelenähtused ning vaadelda keelt laiemalt kui reeglitest ja sõnadest koosnevat süsteemi. See teooria on välja kasvanud Fillmore'i (1968) käändegrammatikast, millest ta arendas edasi freimisemantika (Fillmore 1982). Freimiteooria järgi on keeleüksuse mõistmiseks vaja analüüsida selle taustsituatsiooni (taustfreimi) ning situatsioonist osa võtvaid situatsioonirole. Süntaktilisele seaduspärale – William Crofti ja Alan Cruse'i (2004: 257) järgi ka kõige üldisemale või abstraktsemale konstruktsioonile – vastab mingi semantiline tõlgendus. Seetõttu saab konstruktsioone mõista sümboolsete üksustena, mida konstruktsioonigrammatika kõige laiema käsitluse, radikaalse konstruktsiooniteooria (Croft 2001: 362) põhjal esineb kõigil keele tasanditel,⁵ olgu selleks morfeem, sõna, fraas, lause või diskursus.

Konstruksiooniteoorias ei oletata keelelise teadmise selgesti eristuvat jagunemist sõnavarak ja grammatikareegliteks, vaid keelt nähakse ühe erineva üldisusastmega sümboolsete üksuste võrgustikuna (Sahkai 2008: 181). Konstruktsioonigrammatika uurimisobjektiks on konstruktsioonid kui konventsionaalsed tähenduse ja vormi ühendused ning konstruktsioonigrammatika kognitiivse suuna (Lakoff 1987; Goldberg 1995; Croft, Cruse 2004) järgi sisaldab konstruktsiooni struktuur järgmisi keelelisi tasandeid:



Joonis 1. Konstruktsiooni sümboolne struktuur (Croft, Cruse 2004: 258 joonise 10.1 põhjal).

Nii võib arutleda, et resultatiivkonstruktsiooni {[NF_{elat}] [NF_{transl}]} iseloomustab finiiiverbi puudumine ning konstruktsiooni realisatsioonide kalduvus esineda diskursuses omaette üksusena, näiteks pealkirjana (nagu siinses artiklis) või pildi-alkkirjana. Semantilises plaanis sisaldub konstruktsioonis idee edenemisest ühest

⁵ Crofti (2001: 362) sõnul on konstruktsioon ainus grammatiline primitiiv, mis võib avalduda vormi-tähenduse paarina: „The only type of primitive grammatical units are CONSTRUCTIONS – pairings of form and meaning which may be atomic or complex, schematic or substantive.”

seisundist teise ning pragmaatilises tõlgenduses on translatiiviga märgitud lõppseisund üldjuhul uudsem või hinnatum kui elatiivse käändega tähistatud algseisund. Konstruktsioon on produktiivne, mida näitab selle sõltumatu ilmumine eri temaatikaga pealkirjades, nt aianduskäsiraamat „Naadimetsast paradiisiks” (Palusalu 2023), teadusartikkel „Sõnastikukogust keeleportaalik” (Langemets jt 2021) ja eelinkubatsiooniprogramm „Teadusest äriks!”⁶ kuid sellel põhineb ka kinnistunud väljendeid, nagu *ajalehepoisist miljonäriks*.

Kõnealune konstruktsioon on ilmselgelt tähenduse (ja vormi) poolest sarnane teiste alg- ja lõppseisundit sisaldavate väljendusviisidega, näiteks nendega, milles esinevad muutumisverbid nagu *kasvama*, *saama*, *muutuma* jne. Nende puhul on elatiivne ja translatiivne nimisõnafrasas verbi argumendid ja konstruktsioon ise seega olemuselt pigem lausemall {[NF][V][NF_{elat}][NF_{transl}]}, mille kasutusnäited on: *torm kasvas madalrõhkkonnast orkaaniks; ta muundus konnast printsiks; kuidas allhankijast tootjaks saada?* Kummati osutub, et muutumistähendust saab edasi anda ka siis, kui verb ise välja jätta, kuivõrd alussituatsioon ja selle osalised on olemas keelekasutajate jagatud teadmistes (freimi ehk üldistatud olukorrana, kus keegi/miski [muutub] (kellest/millest) kelleks/milleks // (millisest) milliseks).⁷

Konstruktsioonid võivad olla väga erineva keerukus- ja abstraktsusastmega tähenduse ja vormi kombinatsioonid – nii erandlikud kui ka regulaarsed keele nähtused (Fillmore jt 1988; Goldberg 2013). Esimeste hulka kuuluvad idioomid, mis esindavad fraasistruktuurireeglitest komplekssemate süntaktiliste struktuuride ning semantiliste ja pragmaatiliste tõlgendusprintsipi seoseid (Fillmore jt 1988) – näiteks *verest välja lööma*. Teise rühma alla lähevad produktiivsed lausekonstruktsioonid (Goldberg 1995; Goldberg, Jackendoff 2004), nagu näiteks resultatiivkonstruktsioonid. Grammatika ja leksikoni vahelist kontiinumit ilmestavad järgmised kolm erineva abstraktsusastmega konstruktsiooni, kus fraaselemendid (NF, KF jne) on käsitletavad keelekasutuses konkreetsete lekseemidega täidetavate lünkadena ja leksikaalsed fragmendid kohustuslike kinniselementidena:⁸

1. täisskemaatiline ehk täielikult lünkadest koosnev abstraktne konstruktsioon, näiteks põhjustatud liikumise argumentstruktuuri konstruktsioon {[NF][V][NF][XF_{kuhu}]}⁹, mille kasutamist illustreerib lekseemidega täidetud lause *Uku viib raamatu riiuli peale*;
2. poolskemaatiline ehk osaliselt lünkadega skemaatiline konstruktsioon, mis esindab segatüüpi, näiteks idioomaatiline väljend {[tee viib] [NF_{kuhu}]}, mis sisaldab nii

⁶ <https://ut.ee/et/hep> (28. III 2024).

⁷ Vrd inglise keele freimiga „Undergo_change” (Berkeley FrameNet, <https://framenet.icsi.berkeley.edu/luIndex>).

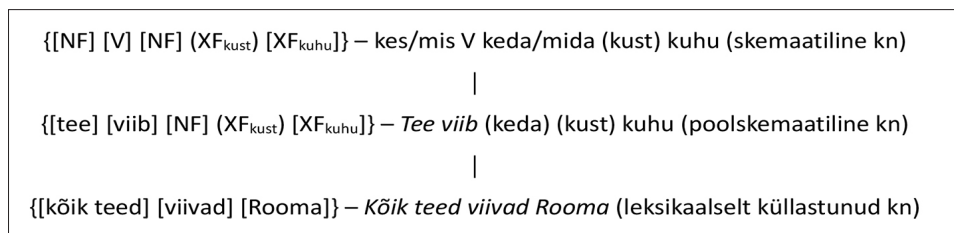
⁸ Lünka (ingl *slot*) täitvate leksikaalsete elementide variatiivsus on piiratud teatud konstruktsiooni seatud piirangutega ning kinniselementideks (*anchor*, *keyword*) on kindel lekseem (vt nt Janda jt 2020).

⁹ Lühendid XF_{kust} ja XF_{kuhu} tähistavad Rätsepa (1978: 44–46) mõistes substitsiooniklasse, mis hõlmavad lähte- ja sihtkohamäärusena funktsioneerivaid moodustajaid (kaasõnafrasaad, määrõnafrasaad, kohakäändes nimisõnafrasaad, proadverbid).

kinniselemente kui ka lünka(sid) ning väljendab fiktiivset liikumist (*Tee viib ränduri märke/metsa*);

- leksikaalselt küllastatud konstruktsioon ehk konkreetne lünkadeta konstruktsioon, näiteks väljend {[*kõik teed viivad Rooma*]}, mis sobitatakse konteksti üldjuhul just sellisel kujul või mis allub piiratud määral modifikatsioonidele.

Erineva abstraktsus- ehk skemaatilisusastmega konstruktsioonid võivad olla omavahel pärilusseostes ning moodustada taksonoomilise võrgustiku (Fillmore jt 1988; Jurafsky 1992; Goldberg 1995, 2006; Croft, Cruse 2004). Pärilusseose all mõistetakse suhet, kus konstruktsiooni alatüüp „pärib” ehk võtab omaks üldisema konstruktsiooni omadused (Lakoff 1987; Goldberg 1995: 67–100) – kui kasutada sugulusmetafoore, on „lapsel” (ingl *child*) spetsiifilisemad omadused kui „vanemal” (*parent*). Ülal esitatud kolm konstruktsiooninäidet esindavad kõik sihtkohta markeerivat põhjustatud liikumise konstruktsiooni, mille vahel valitsev pärilushierarhia on kujutatud joonisel 2.



Joonis 2. Pärilushierarhia põhjustatud liikumise konstruktsiooni näitel.

Kui klassikalise konstruktsioonigrammatika (Lakoff 1987: 492) järgi võtab laps¹⁰ täielikult vanema omadused üle, siis kognitiivse konstruktsioonigrammatika Adele Goldbergi juhitud harus on oletus, et laps pärib vaikumisi vanema omadused, kuid võib vanemast mingis suhtes ka eristuda (Goldberg 1995: 73). Goldberg (1995: 75, 81) teoorias võivad konstruktsioonid olla võrgustikus üksteisega seotud nelja tüüpi pärilusseoste kaudu: polüseemiaseos, allosa seos (kus üks konstruktsioon on teise osa), metafoorse laienduse seos ja pretsedendiseos (kus üks konstruktsioon on teise erijuhutum; vt ka Penjam 2008: 21). Et allosa seost esindaval konstruktsioonil võib olla rohkem kui üks seos eri vanema või isegi vanavanemaga, võib rääkida hulgpärilusseosest (ingl *multiple inheritance*, vt Croft, Cruse 2004: 321; Lyngfelt 2018: 7–8).

Konstruktsioonide võrgustik ehk ühe keele konstruktsioonide kogum moodustab konstruktikoni (ingl *ConstructiCon*, nt Goldberg 2003). Benjamin Lyngfelt (2018: 1) rõhutab, et konstruktikoni mõistel on kaks tõlgendust: see on ühelt poolt teoreetiline ettekujutus keelest kui konstruktsioonide võrgustikust (vt nt Diessel 2019), teisalt aga konstruktsioonide kirjelduste praktiline kogum, 'konstruktsioonide sõnaraamat'. Sedalaadi praktiliste konstruktsioonikirjelduste loomise, süstematiseerimise ja talletamisega tegeleb valdkond, mida nimetakse *konstruktikograafiaks* (ingl *constructico-*

¹⁰ Lapseks nimetatakse hierarhias alamal paiknevat konstruktsiooni; vanemaks ülemal paigutatavat.

graphy) (Lyngfelt 2018: 11) ja mis kujutab endast leksikograafia ning konstruktsiooni-grammatika sünteesi. Edasises räägimegi konstruktikonist tähenduses 'konstruktsioonide kogumik, konstruktsioonide sõnaraamat' ning järgmine peatükk annab ülevaate nende koostamise põhimõtetest, ulatusest ja esile kerkinud probleemidest.

3. Konstruktikograafiat seinast sein

Esimene katse konstruktsioonide jaoks kirjeldusstandardit luua tehti juba üle kümne aasta tagasi (Fillmore jt 2012) ja mõned põhiprintsiibid said paika juba siis. Johtuvalt sellest, et definitsiooni järgi kujutavad konstruktsioonid endast vormi ja tähenduse paare, on ka konstruktikoni üksuse ehk konstruktsiooni kirje loomulikuks osaks ühelt poolt vormi ja teisalt tähenduse ja/või funktsiooni esitamine. Vormi poole kirjeldusse kuulub tavaliselt konstruktsiooni koostisosade ehk elementide esiletõmine, kuivõrd tüüpiliselt on tegu komplekssete keelekasutusmuustritega.¹¹ Konstruktioonelementide omavahelist seost nimetatakse sisesüntaksiks ning konstruktsiooni kui terviku süntaktilist funktsiooni selle välissüntaksiks (vt Ohara 2018: 152). Neile vajalikele kirjelduselementidele lisandub konstruktsioonile antud sõnaline nimetus ning valemilaadne esitus, mille abil toimub konstruktide (ehk konstruktsioonide kasutusjuhtumite) märgendamine ning näitelausete kuvamine. Näitelausetes peab olema näha, millised tekstiüksused milliseid konstruktsioonelemente esindavad.

Vorm	<p>Nimi: Muutus</p> <p>Valem: $\{Muutus[AlgseisundNF_{elat}][LõppseisundNF_{trans}]\}$</p> <p>Väline süntaks: T (=tervik): süntaktiliselt seostamata omaette üksus (nt pealkiri, pildiallkiri)</p> <p>Sisesüntaks: O (=osis)1: nimisõna(fraas) elatiivis, mis tähistab Algseisundit O (=osis)2: nimisõna(fraas) translatiivis, mis tähistab Lõppseisundit NB! Verbita mall</p>
Tähendus/ funktsioon	<p>Tõlgendus: Semantiline: idee edenemisest ühest seisundist teise; Pragmaatiline: Lõppseisund üldjuhul uudsem, parem või hinnatum kui Algseisund.</p> <p>Näide: $\{Muutus[Algseisund]inetust\ pardipojast}[Lõppseisund]tõeliseks kaunitariks\}$</p> <p>Näite tähendus: Kirjeldab esteetilist edenemist, mida ei osatud oodata</p>
Suhted	<p>Pärib: Tähenduse ja osiste vormi pärib <i>Muutumise</i> konstruktsioonist $\{[NF][V][NF_{elat}][NF_{trans}]\}$, milles esineb muutumisverb (nt <i>kasvama</i>, <i>muutama</i>, <i>teisenema</i>, <i>kujunema</i>, <i>saama</i> jne)</p> <p>On päritud: Väljendis: <i>ajalehepoisist miljonäriks</i></p>

Joonis 3. Konstruktsiooni *Muutus* esituse visand.

¹¹ Fillmore'i algset kirjeldusstandardist lähtuvas traditsioonis nimetatakse tervikkonstruktsiooni emaks (ingl *mother*) ja osiseid tütardeks (*daughter*), kuid eestikeelses terminoloogias võiks pigem kasutada termineid Tervik (T) ja Osised (O), et vältida liigset sugulussuhete terminite kasutamist kirjelduses; niikuinii on juba kasutusel *pärilusseosed*, kummati ei ole *mother* ja *daughter* nendega seotud – konstruktsiooni osised ei päri midagi tervikkonstruktsioonilt.

Joonisel 3 on visandatud neist põhimõtetest lähtuvalt kirjeldus artikli alguses esitatud verbita resultatiivkonstruktsiooni jaoks, pannes sellele tinglikult nimeks *Muutus* ning lisades väljad, kuhu saab märkida (pärilus)seoseid teiste konstruktsioonidega.

Väljend *ajalehepoisist miljonäriks* tuleks aga konstruktikonis kirjeldada omaette üksusena – leksikaalselt küllastatud konstruktsioonina –, mille kompaktne esitus on kujutatud joonisel 4.

Nimi:	ajalehepoisist miljonäriks
Vorm:	{Muutus[Algseisundajalehepoisist] [Lõppeisund miljonäriks]}
Tähendus:	Võrdkuju enese ülestöötamise, suure edu kohta
Pärib:	Muutus

Joonis 4. Konstruktsiooni *Ajalehepoisist miljonärist* kompaktne esitus.

Idee keele kirjeldamisest konstruktsioonidena on leidnud väljundi mitmetes projektides, milles on kirjeldatud juba arvestatavat hulka eri keelte konstruktsioone ning mõistagi on põrkunud raskustele, sest keele kirjeldamine uudsel viisil nõuab otsuste langetamist nii materjali piiritlemise kui ka kirjelduse detailsuse osas.

Artikli kirjutamise ajaks ehk 2024. aasta kevadeks oli teada kümnekond projekti seitsme keele jaoks: saksa (FrameNet-Konstruktikon des Deutschen), rootsi (Constructicon for Swedish), vene (Russian Constructicon), Brasiilia portugali (FrameNet Brasil), jaapani (Japanese FrameNet) ja ungari (Hungarian Constructicon), kusjuures inglise keele kohta on vähemalt kaks (Berkeley FrameNet ja Birmingham English Constructicon). Suurem osa neist jälgendavad põhimõtetest esimest, nn Berkeley konstruktikoni (Fillmore jt 2012) ja on end seetõttu orgaaniliselt sidunud FrameNeti kui teise leksikograafilis-semantilise andmekoguga.¹² See on muidugi võimalik keelte puhul, mille jaoks on FrameNeti-taoline semantiliste seoste andmebaas olemas. Oluline eelis seisneb selles, et konstruktsiooni semantika kirjeldamisel saab näidata otsest seost alussituatsiooni semantikaga (vt joonisel 3 lahtrit tähendus/funktsioon – sinna peaks kuuluma link vastava freimi juurde). FrameNet ja konstruktikon on tihedalt seotud, kuna on võrsunud samast keele kirjeldamise koolkonnast ning Fillmore on olnud mõlema algataja (Lyngfelt 2018: 10).

Berkeley konstruktikoni arhetüübist lähtuvad jaapani, Brasiilia portugali ning saksa keele konstruktsioonide kogud. Viimane neist on algusest peale seadnud eesmärgiks keeltevaheliste vastavuste leidmise konstruktsioonide tasandil ning seetõttu alustati inglise konstruktsioonidele saksakeelsete vastete otsimisest (Ziem jt 2019). Rootsi keele konstruktikon ja FrameNet on siiski kaks eraldi andmebaasi, mis on omavahel seotud mitmekülgse taristus, kuhu kuulub keskne sõnastik, korpuse-tööriistad jm (vt Lyngfelt jt 2018: 49). Ka inglise keele teine, Birminghami konstruktikon on linkidega seotud Berkeley FrameNetiga, kuid on andmekoguna siiski

¹² Berkeley FrameNet sisaldab 1224 alussituatsiooni kirjeldust omavahelistes pärilusseostes olevaina ning osutab, millised leksikaalsed üksused inglise keeles neid situatsioone esile kutsuvad (vt <https://framenet.icsi.berkeley.edu/>).

iseseisev ja loodud korpusest abstraheritud grammatikamuustritele tuginedes (Perek, Patten 2019). Vene keele konstruktikoni ühe alusressursina on mainitud vene keele FrameBanki (vt Janda jt 2018); teisest küljest on Laura Janda väitnud, et vaid 8% vene konstruktikoni üksustest omab mingit seost freimidega ja seega pole FrameNeti-taolise keelekirjelduse olemasolu konstruktikoni loomise eeltingimus (Janda jt 2023). Freimidest ja FrameNeti kirjeldustest sõltumatu algatus ongi teada: selleks on ungari keele konstruktikoni projekt, mis tõlgendab ümber ungari üldsõnastiku ainekku ja võimaldab esitada konstruktsioonidena nii mitmesõnalisi üksusi kui ka üksikuid sõnu ja neid moodustavaid morfeeme (Sass 2023).

Valdav osa seniseid projekte on seadnud esmaseks eesmärgiks kajastada selliseid poolskemaatilisi keeleüksusi, mis oma kompleksse ja/või varieeruvate komponentidega koostise tõttu ei ole leidnud omaette üksusena kajastamist leksikograafilistes andmekogudes ega grammatikakirjelduses (nt Berkeley, rootsi, vene, saksa, jaapani, Brasiilia portugali keele konstruktikonid). Erandina on ungari keele konstruktikoni projekt alustanud leksikaalselt küllastatud ehk lünkadeta ühendite esitamist sama digitaalse sõnastikubaasi sees (Sass 2023). Birminghami konstruktikon seevastu sisaldab ka (ja isegi eeskätt) täisskemaatilisi, nt argumentstruktuuri konstruktsiooni.¹³ Kõige ambitsioonikamad projektid (rootsi, vene, saksa) on deklareerinud, et soovivad edaspidi katta kogu keele grammatikat, st mitte ainult süntaktilisi, vaid ka morfoloogilisi ja diskursuse tasandi üksusi (Lyngfelt jt 2018; Janda jt 2018; Ziem jt 2019). Need kolm projekti on loonud endale selleks ka kõige paremad eeldused eraldi andmebaasidena, mis on disainitud konstruktsioonide põhjalikuks kirjeldamiseks, st vastavad teooriast tulenevatele nõudmistele konstruktsiooniartiklite ja nende omavaheliste (pärilus)seoste kirjeldamiseks.

Senised konstruktikonid erinevad ka oma mahu ja konstruktsioonide käsitlemise põhjalikkuse poolest. Skaala ühes otsas on Berkeley FrameNeti konstruktikon väga põhjalikult käsitletud 73 konstruktsiooniga (Fillmore jt 2012) ja skaala teises otsas ungari projekt automaatselt genereeritud 13 000 kirjega, mis fikseerivad eeskätt ühendeid ning lingivad neid osa–terviku seoseid pidi teiste üksustega leksikaalses andmebaasis (Sass 2023). Vahepeelsele tasemele jäävad rootsi, saksa, Brasiilia portugali ning vene keele konstruktikonid, mis samas ei tee allahindlust konstruktsioonide struktuuri ja tähenduse kirjeldamises, märgendatud näitelausetega varustatuses ega omavaheliste pärilusseoste osutamises. Nende andmekogude loomisse on kaasatud arvestatav hulk infotehnoloogide ja keeleteadlaste tööd; tegemist on kollektiivide vähemalt kümne aasta pikkuste pingutuste tulemustega, mis on sellegipoolest alles arendamise ja täiendamise faasis.

Vaadeldud konstruktikonid on võrsunud mõnevõrra erinevatest eesmärkidest: osa on pidanud silmas eeskätt võimalikult detailset ja keeleekspertidele suunatud kirjeldust (selline on nt Berkeley FrameNeti konstruktikon, aga ka saksa, jaapani ja Brasiilia portugali keele kogud); teised on jällegi püüdnud ühitada keeleõppija huviseid nii ekspertide kui ka võimalike edasiste rakenduste ja keeletehnoloogia vajadustega (st on silmas peetud ka masinloetavust). Sellised mitmeid funktsioone taotlevad

¹³ Samalaadne lähenemine on ka Saksamaal Erlangenis arendataval verbi laiendusmallide andmekogul, mis end samuti konstruktikoniks nimetab (Herbst, Uhrig 2019).

projektid on nt rootsi (vt Lyngfelt jt 2018) ja vene keele algatused. Vene keele konstruktikon muu hulgas eristab eri keeleoskustasemetelt eeldatavaid konstruktsioone ja ongi juba seotud keeleõppijaile ja õpetajatele mõeldud rakendusega, mille abil saab konstruktsioone iseseisvalt harjutada.¹⁴ Osale (nt Birminghami inglise konstruktkonile ning saksa konstruktikonile) on õppijaile suunatud orientatsioon lisandunud projekti edenedes.

4. Peamised väljakutsed

Konstruktikonide loomine on üldjuhul vaeva ja aega nõudev protsess, mille käigus tuleb lahendada hulk probleeme. Esiteks, mis üldse kvalifitseerub konstruktsiooniks? See on ennekõike, andmekogu moodustamist puudutav praktilise kaaluga küsimus – teoreetiliselt on kõik tähenduse ja vormi paarid keeles konstruktsioonid. Tegelikuses on ikkagi raske kogu keele kirjeldamist lähteülesandeks võtta ja on kaheldav, kas sellel näiteks sõnatasandil on mõtetki, kuivõrd sõnade esitamiseks on leksikograafias traditsioon välja kujunenud (Lyngfelt 2018: 9).

Teine küsimus on see, kuidas toimida samade konstruktsioonide esitamisega erinevatel abstraktsuse ja skemaatilise astmetel. Näiteks, kas esitada mingi malli levinud ja kinnistunud realisatsioonid omaette konstruktsioonidena (st eraldi kirjetena andmebaasis)? Kas siinses artiklis illustreeritud verbita resultatiivkonstruktsiooni tuntumaid realisatsioone *ajalehepõisist miljonäriks* peaks olema omaette täieõiguslik kirje või pigem esitatud skemaatilise konstruktsiooni {^{Muutus}[^{Algeisund}NF_{elat}][^{Lõppseisund}NF_{transl}]} näidete rubriigis? Vene konstruktikoni loomise käigus on leitud, et leksikaalselt küllastatud juhtumite esitamine on vajalik, kuna konstruktikonist otsitakse just levinud väljendeid (vt nt Janda jt 2023), seevastu rootsi keele konstruktikon on seadnud kriteeriumiks, et vähemalt üks liige peab olema varieeruv, st abstraktselmal astmel esindatud lüngaga (Lyngfelt jt 2018). Kokkuvõtvalt võib öelda, et raskusi tekitab see, et kirjeldusformaad peab sobima leksikoni–grammatika kontiinumil nii leksikapoolses otsas olevate üksuste (nt idioomid, fraseologismid), vahepealsete ning ka täiesti skemaatiliste puhtgrammatiliste konstruktsioonide jaoks (Ziem jt 2023).

Kuidas tagada, et eri kasutajagruppide vajadused oleksid rahuldatud? Keeleõppijad vajaksid suurt hulka võimalikult lihtsas metakeeles kirjeldatud ja keeleoskustasemete põhjal eristatud konstruktsioone lihtsate ja tüüpiliste kasutusnäidete. Keeleeksperdid seevastu sooviksid nii põhjalikku infot kui võimalik, seda nii prototüüpsete kui ka eripärasemate kasutuste kohta, ning samuti võimalust näha autentseid, kuid konstruktsioonide koostiselementide järgi märgendatud korpuse-lauseid. Masinloetavuse tagamiseks peab kasutama vastava funktsiooni või rakenduse jaoks mõeldud märgendust. Metakeele küsimus kätkeb endas ka seda, mille järgi konstruktsioone üldse kasutaja jaoks nimetada: kas lähtuda vormist (nt konstruktsiooni keske(te)st kinnistunud sõna(de)st, fraasitüübist) või hoopiski sisust ja/või funktsioonist? Kuidas sel puhul teha konstruktsioonide info kättesaadavaks teis-

¹⁴ Vt <https://constructicon.github.io/construxercise-rus/>

keelsetele kasutajatele? Kui leksikaalsetele üksustele saab anda lihtsalt vaste, siis kuidas peaks teise keele õppija leidma konstruktsioonide kogust enda soovitud tähenduse väljendamiseks loomulikema viisi sihtkeeles?

Konstruktsioonide eristamise ja nimetamisega on seotud samuti küsimus, kuidas konstruktikoni üldiselt organiseerida ja struktureerida. Lihtne alfabeetiline järjestus ei näi olevat mõttekas, kuna nn kinniselemendid (millel on olemas fonoloogiline vorm ja algustäht) võivad olla erinevates rollides, millel tüüpiliselt konstruktsiooni tähendusega otseseost ei ole. Ka ei lähe selliselt koostatud nimekirja kokku konstruktsioonide võrgustikulaadse ja dünaamilise rühmitumise oletusega. Kas jälgendada kategooriate esitamist grammatikates? Millest lähtuvalt luua katuskategooriaid, taksonoomiat vm tüübistikku? Kuivõrd konstruktikoni kontsept on alles arenemisjärgus, siis pole välja kujunenud ühtset traditsiooni ja vähemasti rootsi keele konstruktikoni puhul on tõdetud, et struktuur peab olema paindlik ja kohandatav¹⁵ sedamööda, kuidas kogusse lisandub uusi (ja võimalik et ka erinevat tüüpi) konstruktsioone (Borin, Lyngfelt ilmumas: 13). Vene keele konstruktikoni jaoks on välja arendatud semantiline klassifikatsioon (Endresen, Janda 2020). Ungari keele konstruktikoni jaoks on loodud otsimisirakendus, mis analüüsib sisestatud vaba tekstijada morfeemideks ja pakub loetelu konstruktsioonidest, kus need komponendidena esinevad (Sass 2023).

Konstruktikoni süstematiseerimisel tekib ka küsimus, kuidas ja mil määral esitada konstruktsioonide omavahelisi pärilus- ja muid seoseid. Kuivõrd konstruktsioonigrammatika postuleerib keele konstruktsioonide võrgustikuna, on ootuspärane, et need seosed on keelekirjelduses eksplitsiitselt esile toodud. Eri keelte konstruktikonid on siin läinud mõnevõrra eri teid, kuna seoste esitamine oleneb andmekogu üldisest eesmärgist: mis on otstarbekas pedagoogilise suunitlusega konstruktikoni jaoks, ei pruugi seda olla keeletehnoloogiale orienteeritud lahenduses. Teine seoste laad, mida konstruktikonid erineval määral esitavad, on nn lünga ja täite (ingl *slot-filler*) seosed, millest tekib vajadus esitada andmekogus teavet selle kohta, millised leksikaalsed üksused saavad poolskemaatilise konstruktsiooni realisatsioonides esile tulla. Seda on kergem teostada, kui konstruktikon on seotud suurema leksikograafilise andmebaasiga. Saksa, rootsi, vene ja Birminghami inglise keele konstruktikonid on selliseid seoseid võimaldavate lahenduste näideteks.

Mõtlemisainet pakub ka selle otsustamine, milline peaks nägema välja konstruktikoniüksuse kirje. Lars Borin ja Benjamin Lyngfelt (ilmumas: 10) rõhutavad, et peale teooriast tulenevate struktuursete osade, nagu vormi ja tähenduse esitamine, on oluline ka see, kas konstruktikon on seotud mõne muu leksikograafilise andmekoguga (nt FrameNet) ja mis on andmekogu üldine eesmärk. Kirjeldusformaad on alati mingit laadi kompromiss ja lihtsus – olgu siis teooria või keeletehnoloogia jaoks vajaliku formaliseerituse seisukohast. Et konstruktikon suudaks katta arvestatava hulga üksusi, ei saa iga üksiku kirje koostamise peale liiga palju aega kulutada. Seega peab konstruktikoni kirje olema asjakohane, lühike, kasutajasõbralik ning samas formaliseeritud (Lyngfelt jt 2018: 48). Näiteks rootsi keele konstruktikon

¹⁵ Nn orgaanilise kasvu strateegia (Lyngfelt jt 2018: 95).

sisaldab konstruktsioonide kohta nii inimloetavaid ja vabalt sõnastatud selgitusi kui ka üliformaalseid valemilaadseid esitusi, mis teenivad eeskätt keetehnoloogia ja edasiste rakenduste loomise huve (Borin, Lyngfelt ilmmas: 11).

Eri keelte konstruktsioonide koostajad, eriti need, kes lähtuvad Berkeley Frame-Neti struktuurist, on asunud looma ka põhimõtteid, kuidas eri keeltes leiduvaid konstruktsioone omavahel vastavusse seada (vt Lyngfelt jt 2022). Katseid luua vastavusi FrameNeti põhjal on peetud ebapiisavaks. Alternatiivse ja täiendava ideena on pakutud keeletüpoloogias tuntud nn võrdlusmõisteid (ingl *comparative concepts*), mis oleksid sõltumatud mis tahes keelest ja väljendaksid üldisi kategooriaid, nt eesti keelde tõlgituna: *osalause, võrdluskonstruktsioon, intransitiivne konstruktsioon, agent, aste, elutu, kujundi ja tausta vastandus, modifitseerimine*. Samas rõhutatakse, et plaanis pole luua ühtset andmebaasi, vaid pigem kirjelduskategooriate süsteem, mille alusel saaks eri keelte konstruktsioone vastavusse seada. (Borin, Lyngfelt ilmmas: 18)

Kõigi nende küsimuste taustal on loomulik mõelda, kas ja kuidas tulla nende väljakutsetega toime eesti keele jaoks konstruktsioonide andmekogu ehk konstruktsiooni luues. Üksikasjalikumalt kirjutame oma kavatsustest ja katsetustest järgmistes artiklites. Siin esitame põgusalt plaanitava eesti konstruktsiooni peajooned, mida suuresti liigendab meie nägemus sellest, kuidas ületada eespool kirjeldatud tüüpilised raskused.

Eesti keele konstruktsioon on plaanitud EKI ühendsõnastiku laiendusena (st osaresursina alusbaasis Ekilex, vt Tavast jt 2018), kust vajadusel kuvatakse sõnavaraüksuste juurde süntaktilist kasutusinfot. Esimeses etapis esitame süstemaatilisel konstruktsioonidena baasis juba leiduva süntaktilise informatsiooni lekseemide kohta (nt fraasistruktuurimallid kollokatsiooniplokist, rektsioonimallid grammatikaplokist). Teises etapis loome kirjeldusstandardi, millesse sobituvad leksikoni–grammatika kontiinumi eri osades paiknevad keeleüksused: nii idiomatilised väljendid, poolskemaatilised lünkadega konstruktsioonid kui ka täiesti abstraktsed mallid nt lihtlausete moodustamiseks (Rätsep 1978; Saul 2024). Argumentstruktuuri konstruktsioonid on kavas ekstraheerida morfoloogiliselt ja sõltuvussüntaktiliselt märgendatud korpusest ning andmebaasis esitamise mõttes liikuda realisatsioonilt abstraktsiooni ehk täisskemaatilise esituse suunas.

Eesti konstruktsiooni sihtgrupp on eesti keele teise keelena õppijad, kuid loodud kirjeldused peavad sobima ka keeleuurijatele, laiale keelehuviliste ringile ning keetehnoloogiliste rakenduste sisendiks. Seetõttu tuleb luua paralleelsed nimetused ja konstruktsioonide kirjeldused vastavalt nii intuiitiivselt lihtsas, keeleteaduslikult täpses kui ka masinloetavas metakeeles. Konstruktsioonid varustatakse mitmete atribuutidega nende semantika, kategooriaalse kuuluvuse jm alusel, mis annab konstruktsioonile sisemise struktuuri ja kasutajale mitmekülgsed otsimisvõimalused. Kavastatakse ka konstruktsioonidevahelisi hierarhilisi pärilus-, osa–terviku- jm seoseid, samuti – ja võib-olla esmajoones – lünga ja täite seoseid sõnade ja konstruktsioonide vahel. Kõige esmasem juurdepääs konstruktsioonidele on neis esinevate sõnade kaudu.

5. Lõpetuseks

Ajal, mil sõnastikud ega grammatikad ei pea enam mahtuma paberkandjale ja ruumipuudus pole enam teemaks, on mitmel pool maailmas leitud, et leksikaalset ja grammatilist infot võiks ja peaks pakkuma üheskoos, ühendatud andmekoguna. Seda enam, et moodsate keeleteooriate järgi ei olegi grammatika ja leksikaalsete üksuste vahel selget ja üheselt mõistetavat piiri, vaid tegu on sujuva üleminekuga ning keelekasutust suunavad märkamatult lisaks sõnadele ja selgelt formuleeritavaatele reeglitele ka (ja võimalik et esmajoones) mitmesugused keelekasutuse pealt üldistatud pool- või täisskemaatilised konstruktsioonid.

Eelnevas võtsime vaatluse alla konstruktsioonipõhise keelekäsitluse, tegime sissevaate teiste keelte jaoks loodud või loomisel olevatesse konstruktsioonidesse – konstruktsioonide kogumikesse – ning käsitlesime väljakutseid, mis välismaa kolleegidel on keele konstruktsioonidena kirjeldamise käigus esile kerkinud. Olemasolevaid konstruktsioonidena kirjeldades võib järeldada, et nende loomiseks ja esitamiseks ei leidu ühtset tava – peale printsiibi, et konstruktsioonid peavad olema esile toodud vormi ja tähenduse „paarisrakendina”. Konstruktsioonid seob siiski võimalus kas või osaliselt liidestada juba mõne olemasoleva, enamasti FrameNeti tüüpi andmekoguga.

Konstruktsioonidel on traditsioonilise sõnaraamatuga nii ühis- kui ka erijooni. Nagu sõnaraamat, pakuvad vaadeldud konstruktsioonid keeleüksuste loetelusid. Erinevalt sõnaraamatust, kus leksikograafiline selgitus on enamasti võimalikult napp ja üldistav, võib konstruktsiooniraafiline info kohati olla detailirohke, kirjeldades nii tüüpilist kui ka ebatüüpilist kasutust. Sõnaraamat seab esitusvormi ideaali, millega tuleb sobitada konstruktsioonide kirjeldamiseks vajalik grammatiline info. Viimase tehniline loomus on üks peamisi proovikive teel täiusliku konstruktsiooni poole. Veelgi suurema väljakutse esitab mitmekeelsuse, st eri keelte konstruktsioonide omavahel vastavusse seadmise taotlus. (Lyngfelt 2018: 12)

Kogu üritust võib piltlikult kirjeldada rätsepatööna, kuivõrd konstruktsiooni metakeel, üldine disain ja ülesehitus, selle üksikasjalikkus, läbitöötatus ja kasutusning kuvamisvõimalused sõltuvad konstruktsiooni loomise eesmärgist – kellele ja milleks see on esmajoones mõeldud. Teise piltliku printsiibina võib öelda, et töö on tellija materjalist: niivõrd ulatuslikku keelendite uudsel viisil kuvamist saab ja on mõttekas alustada tuginemisest mingile olemasolevale leksikaalsele andmekogule. Kui selleks on FrameNet, on ka konstruktsioonikirjete esitamise viis sellest mõjutatud (nt Torrent jt 2014); kui selleks on aga keele üldsõnastik (nt Sass 2023) või süntaktiliste muustrite andmebaas (Perek, Patten 2019), tundub mõttekana alustada kirjeldust nendele ressurssidele tuginedes.

Tuleb nentida, pääsu pole ka kolmandast piltlikust printsiibist: enne tegutsema asumist tuleb kõvasti mõõtu võtta, st õppida seni parimatest praktikatest. Eelneva ülevaate otstarve oligi seniseid lahendusi kaardistada, et teha parimad valikud eesti keele konstruktsiooni kavandama asudes. Käimasoleva EKI uurimisprojekti üks sihte on töötada välja võimalused, kuidas konstruktsioonilist infot esitada EKI ühend-sõnastiku (vt Langemets jt 2021) aluseks olevas Ekilexi andmebaasis (vt Tavast jt

2018). Siinses artiklis on visandatud vaid esialgne visioon eesti konstruktikonist grammatilise info koondajana ning ühenduslülina korpusmaterjali juurde, kus võib näha konstruktsioone toimimas elavas keelekasutuses.

Kuna me alles töötame välja konstruktikoni kontseptsiooni eesti keele jaoks, on mõistetav, et meil on pakkuda esialgu rohkem küsimusi kui vastuseid. Samuti anname endale aru, et tee selle visiooni teostamise suunas ei saa kindlasti olema kerge ei keeleteaduslikus ega infotehnoloogilises mõttes ega ka ühiskondlikku retseptiooni ette kujutada. Näiteks on ette näha, et kõikmõeldavate, sh huvitavate ja isikupäraste konstruktsioonide kirjeldamine ühtses leksikaalses andmekogus saab hõlpsasti normatiivse tõlgenduse (vt Lyngfelt jt 2018: 59 ja sealsed viited). Seepärast tuleb algusest peale eraldi pingutada selle nimel, et nii puhtleksikaalsete üksuste kui ka konstruktsioonide kirjelduste juures saaksid kindlaks tehtud ja ära märgitud võimalikud registreeritud piirangud.

Artikkel on valminud Eesti Teadusagentuuri projekti „Uue aja sõnastik: grammatika ja keelepädevuse kirjeldamine integreeritud multifunktsionaalses leksikograafilises ressursis“ (PRG 1978) raames.

LÜHENDID

AF = omadussõnafaas; **elat** = elatiiv; **KF** = kaassõnafaas; **kn** = konstruktsioon; **NF** = nimi-sõnafaas; **O** = osis; **T** = tervik; **transl** = translatiiv; **V** = tegusõnafaas; **XF_{kust}** = lähtekohta tähistav fraas; **XF_{kuhu}** = sihtkohta tähistav fraas.

VEEBIVARAD

Berkeley FrameNet. <https://framenet.icsi.berkeley.edu>

Birmingham English Constructicon. <https://englishconstructicon.bham.ac.uk>

Constructicon for Swedish. <https://www.gu.se/en/research/a-constructicon-for-swedish>

EKI ühendsõnastik. Eesti Keele Instituut, Sõnaveeb. <https://sonaveeb.ee>

FrameBank. <http://framebank.ru>

FrameNet Brasil. <https://webtool.framenetbr.ufjf.br/index.php/webtool/report/cxn/main>

FrameNet-Konstruktikon des Deutschen. <https://gsw.phil.hhu.de>

Hungarian Constructicon [kasutaja: eLex2023; salasõna: letssee]. <https://ccn.nytud.hu>

Japanese FrameNet. <https://jfn.st.hc.keio.ac.jp>

Russian Constructicon. <https://constructicon.github.io/russian>

KIRJANDUS

Bencini, Giulia M. L. 2013. Psycholinguistics. – The Oxford Handbook of Construction Grammar. (Oxford Handbooks in Linguistics.) Toim Thomas Hoffmann, Graeme Trousdale. New York–Oxford: Oxford University Press, lk 379–396.
<https://doi.org/10.1093/oxfordhb/9780195396683.013.0021>

- Boas, Hans C. (toim) 2022.** Directions for Pedagogical Construction Grammar: Learning and Teaching (with) Constructions. (Applications of Cognitive Linguistics 49.) Berlin–Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110746723>
- Booij, Geert 2010.** Construction Morphology. (Oxford Linguistics.) Oxford: Oxford University Press.
- Borin, Lars; Lyngfelt, Benjamin (ilmumas).** Framenets and constructiCons. – The Cambridge Handbook of Construction Grammar. (Cambridge Handbooks in Language and Linguistics.) Cambridge: Cambridge University Press.
- Bybee, Joan 2010.** Language, Usage, and Cognition. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511750526>
- Bybee, Joan 2013.** Usage-based theory and exemplar representations of constructions. – The Oxford Handbook of Construction Grammar. (Oxford Handbooks in Linguistics.) Toim Thomas Hoffmann, Graeme Trousdale. New York–Oxford: Oxford University Press, lk 49–69. <https://doi.org/10.1093/oxfordhb/9780195396683.013.0004>
- Croft, William 2001.** Radical Construction Grammar: Syntactic Theory in Typological Perspective. (Oxford Linguistics.) Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198299554.001.0001>
- Croft, William; Cruse, D. Alan 2004.** Cognitive Linguistics. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511803864>
- Diessel, Holger 2013.** Construction grammar and first language acquisition. – The Oxford Handbook of Construction Grammar. (Oxford Handbooks in Linguistics.) Toim Thomas Hoffmann, Graeme Trousdale. New York–Oxford: Oxford University Press, lk 347–364. <https://doi.org/10.1093/oxfordhb/9780195396683.013.0019>
- Diessel, Holger 2019.** The Grammar Network: How Linguistic Structure Is Shaped by Language Use. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108671040>
- Ellis, Nick C. 2013.** Construction grammar and second language acquisition. – The Oxford Handbook of Construction Grammar. (Oxford Handbooks in Linguistics.) Toim Thomas Hoffmann, Graeme Trousdale. New York–Oxford: Oxford University Press, lk 365–378. <https://doi.org/10.1093/oxfordhb/9780195396683.013.0020>
- Endresen, Anna; Janda, Laura A. 2020.** Taking construction grammar one step further: Families, clusters, and networks of evaluative constructions in Russian. – *Frontiers in Psychology*, nr 11, lk 1–22. <https://doi.org/10.3389/fpsyg.2020.574353>
- Fillmore, Charles J. 1968.** The case for case. – *Universals in Linguistic Theory*. Toim Emmon Bach, Robert T. Harms. New York: Holt, Rinehart, and Winston, lk 1–88.
- Fillmore, Charles J. 1982.** Frame semantics. – *Linguistics in the Morning Calm: Selected Papers from SICOL-1981*. Seoul: Hanshin Pub. Co.
- Fillmore, Charles J. 2006.** The articulation of lexicon and constructicon. – *Plenaarettekanne 4. rahvusvahelisel konstruktsioonigrammatika konverentsil Tokyo ülikoolis 3. IX 2006*.
- Fillmore, Charles J.; Kay, Paul; O'Connor, Mary Catherine 1988.** Regularity and idiomaticity in grammatical constructions: The case of *let alone*. – *Language*, kd 64, nr 3, lk 501–538. <https://doi.org/10.2307/414531>

- Fillmore, Charles J.; Lee-Goldman, Russell R.; Rhodes, Russell 2012.** The FrameNet constructicon. – Sign-Based Construction Grammar. Toim Hans C. Boas, Ivan A. Sag. Stanford: CSLI Publications, lk 309–372.
- Goldberg, Adele E. 1995.** *Constructions: A Construction Grammar Approach to Argument Structure.* (Cognitive Theory of Language and Culture.) Chicago: The University of Chicago Press.
- Goldberg, Adele E. 2003.** *Constructions: A new theoretical approach to language.* – Trends in Cognitive Sciences, kd 7, nr 5, lk 219–224.
[https://doi.org/10.1016/S1364-6613\(03\)00080-9](https://doi.org/10.1016/S1364-6613(03)00080-9)
- Goldberg, Adele E. 2006.** *Constructions at Work: The Nature of Generalization in Language.* (Oxford Linguistics.) Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199268511.001.0001>
- Goldberg, Adele E. 2013.** *Constructionist approaches.* – The Oxford Handbook of Construction Grammar. (Oxford Handbooks in Linguistics.) Toim Thomas Hoffmann, Graeme Trousdale. New York–Oxford: Oxford University Press, lk 15–31.
<https://doi.org/10.1093/oxfordhb/9780195396683.013.0002>
- Goldberg, Adele E.; Jackendoff, Ray 2004.** The English resultative as a family of constructions. – Language, kd 80, nr 3, lk 532–568. <https://doi.org/10.1353/lan.2004.0129>
- Gras, Pedro; Elvira-García, Wendy 2021.** The role of intonation in Construction Grammar: On prosodic constructions. – Journal of Pragmatics, nr 180, lk 232–247.
<https://doi.org/10.1016/j.pragma.2021.05.010>
- Herbst, Thomas; Uhrig, Peter 2019.** Towards a valency and argument structure constructicon of English: Turning the valency patternbank into a constructicon. – Lexicographica, nr 35, lk 87–104. <https://doi.org/10.1515/lex-2019-0006>
- Hoffmann, Thomas; Trousdale, Graeme (toim) 2013.** *The Oxford Handbook of Construction Grammar.* (Oxford Handbooks in Linguistics.) New York–Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195396683.001.0001>
- Janda, Laura A.; Lyashevskaya, Olga; Nessel, Tore; Rakhilina, Ekaterina; Tyers, Francis M. 2018.** A constructicon for Russian: Filling in the gaps. – Constructicography: Constructicon Development Across Languages. (Constructional Approaches to Language 22.) Toim Benjamin Lyngfelt, Lars Borin, Kyoko Ohara, Tiago Timponi Torrent. Amsterdam: John Benjamins, lk 165–181. <https://doi.org/10.1075/cal.22.06jan>
- Janda, Laura A.; Endresen, Anna; Zhukova, Valentina; Mordashova, Daria; Rakhilina, Ekaterina 2020.** How to build a constructicon in five years: The Russian example. – Belgian Journal of Linguistics, kd 34, nr 1, lk 161–173.
<https://doi.org/10.1075/bjl.00043.jan>
- Janda, Laura A.; Zhukova, Valentina; Endresen, Anna; Mordashova, Daria; Rakhilina, Ekaterina; Lyashevskaya, Olga 2023.** What happens if you try to build a constructicon for a whole language? – International FrameNet Workshop 2023: Cognitively Grounded Approaches to Applied Language Description (IFNW2023). Düsseldorf, August 7–11, 2023. <https://www.globalframenet.org/ifnw-2023>
- Jurafsky, Daniel 1992.** *An On-line Computational Model of Human Sentence Interpretation: A Theory of the Representation and Use of Linguistic Knowledge.* Doctoral dissertation. Berkeley: University of California. <https://doi.org/10.21236/ADA604298>

- Kay, Paul; Fillmore, Charles J. 1999. Grammatical constructions and linguistic generalizations: The *What's X doing Y?* construction. – *Language*, kd 75, nr 1, lk 1–33.
<https://doi.org/10.2307/417472>
- Kitsnik, Mare 2018. Iga asi omal ajal: eesti keele B1- ja B2-taseme verbikonstruktsioonid keeleoskuse arengu näitajana. (Tallinna Ülikooli humanitaarteaduste dissertatsioonid 43.) Tallinn: Tallinna Ülikooli Kirjastus.
- Lakoff, George 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
<https://doi.org/10.7208/chicago/9780226471013.001.0001>
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar. Kd 1: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langemets, Margit; Koppel, Kristina; Kallas, Jelena; Tavast, Arvi 2021. Sõnastikukogust keeleportaalsiks. – *Keel ja Kirjandus*, nr 8–9, lk 755–770.
<https://doi.org/10.54013/kk764a6>
- Lyngfelt, Benjamin 2018. Introduction: Constructicons and constructicography. – *Constructicography: Constructicon Development Across Languages*. (Constructional Approaches to Language 22.) Toim B. Lyngfelt, Lars Borin, Kyoko Ohara, Tiago Timponi Torrent. Amsterdam: John Benjamins, lk 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. – *Constructicography: Constructicon Development Across Languages*. (Constructional Approaches to Language 22.) Toim B. Lyngfelt, Lars Borin, Kyoko Ohara, Tiago Timponi Torrent. Amsterdam: John Benjamins, lk 41–106.
<https://doi.org/10.1075/cal.22.03lyn>
- Lyngfelt, Benjamin; Torrent, Tiago Timponi; da Silva Matos, Ely Edison; Bäckström, Linnéa 2022. Comparative concepts as a resource for multilingual constructicography. – *Valency and Constructions: Perspectives on Combining Words*. (Meijerbergs arkiv för svensk ordforskning 46.) Toim Kristian Blensenius. Göteborg: MASO, lk 101–130.
- Ohara, Kyoko 2018. Relations between frames and constructions: A proposal from the Japanese FrameNet constructicon. – *Constructicography: Constructicon Development Across Languages*. (Constructional Approaches to Language 22.) Toim Benjamin Lyngfelt, Lars Borin, K. Ohara, Tiago Timponi Torrent. Amsterdam: John Benjamins, lk 141–163. <https://doi.org/10.1075/cal.22.05oha>
- Palusalu, Maria 2023. Naadimetsast paradiisiks ehk kuidas saada silmapaistvat ja vähe hoolt nõudvat aeda. Tallinn: Varrak.
- Penjam, Pille 2008. Eesti kirjakeele *da*- ja *ma*-infinitiiviga konstruktsioonid. (Dissertationes philologiae estonicae Universitatis Tartuensis 23.) Tartu: Tartu Ülikooli Kirjastus.
- Penjam, Pille 2009. Mis on konstruktsioonigrammatika? – *Oma Keel*, nr 2, lk 5–12.
- Penjam, Pille; Pajusalu, Renate 2006. Konstruktsioonigrammatika. – *Teoreetiline keeleteadus Eestis II*. (Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 7.) Toim Ilona Traegel, Haldur Õim. Tartu: Tartu Ülikool, lk 141–155.
- Perek, Florent 2015. *Argument Structure in Usage-Based Construction Grammar: Experimental and Corpus-Based Perspectives*. (Constructional Approaches to Language 17.) Amsterdam: John Benjamins. <https://doi.org/10.1075/cal.17>

- Perek, Florent; Patten, Amanda L. 2019.** Towards an English constructicon using patterns and frames. – *International Journal of Corpus Linguistics*, kd 24, nr 3, lk 354–384.
<https://doi.org/10.1075/ijcl.00016.per>
- Pilvik, Maarja-Liisa 2021.** Action Nouns in a Constructional Network: A Corpus-based Investigation of the Productivity and Functions of the Deverbial Suffix *-mine* in Five Different Registers of Estonian. (Dissertationes philologiae estonicae Universitatis Tartuensis 48.) Tartu: Tartu Ülikooli Kirjastus.
- Pulvermüller, Friedemann; Cappelle, Bert; Shtyrov, Yury 2013.** Brain basis of meaning, words, constructions, and grammar. – *The Oxford Handbook of Construction Grammar*. (Oxford Handbooks in Linguistics.) Toim Thomas Hoffmann, Graeme Trousdale. New York–Oxford: Oxford University Press, lk 397–416.
<https://doi.org/10.1093/oxfordhb/9780195396683.013.0022>
- Rätsep, Huno 1978.** Eesti keele lihtlausete tüübid. (Emakeele Seltsi toimetised 12.) Tallinn: Valgus.
- Sahkai, Heete 2006.** Konstruksioonipõhise keelekirjelduse võimalustest adessiivse viisi- ja põhjusemääruse näitel. – *Keel ja Kirjandus*, nr 10, lk 816–831.
- Sahkai, Heete 2008.** Konstruksioonipõhine keelemudel ja sõnaraamatumudel. – *Eesti Rakenduslingvistika Ühingu aastaraamat*, nr 4, lk 177–186.
<http://dx.doi.org/10.5128/ERYa4.11>
- Sahkai, Heete 2011.** Teine grammatika: eesti keele teonimede süntaks konstruksioonipõhises perspektiivis. (Tallinna Ülikooli humanitaarteaduste dissertatsioonid 25.) Tallinn: TLÜ Kirjastus.
- Sahkai, Heete 2012.** Konstruksioonigrammatika. – *Oma Keel*, nr 2, lk 5–10.
- Sass, Balint 2023.** From a dictionary towards the Hungarian constructicon. – *Electronic Lexicography in the 21st Century (eLex 2023): Invisible Lexicography*. Proceedings of the eLex 2023 Conference, Brno, 27–29 June 2023. Brno: Lexical Computing CZ s.r.o., lk 534–544. <https://elex.link/elex2023/wp-content/uploads/105.pdf>
- Saul, Kertu 2024.** Eesti keele lausemallide automaatne tuvastamine liigutamisverbide näitel. Magistritöö. Tartu Ülikool.
- Schmid, Hans-Jörg 2020.** The Dynamics of the Linguistic System: Usage, Conventionalization, and Entrenchment. (Oxford Linguistics.) Oxford: Oxford University Press.
<https://doi.org/10.1093/oso/9780198814771.001.0001>
- Ziem, Alexander; Flick, Johanna; Sandkühler, Phillip 2019.** The German constructicon project: Framework, methodology, resources. – *Lexicographica*, nr 35, lk 15–40.
<https://doi.org/10.1515/lex-2019-0003>
- Ziem, Alexander; Böbel, Nina; Willich, Alexander 2023.** What's in the constructicon? Relating constructional forms and constructional meanings on the full range of the lexicon-grammar continuum. – 16th International Cognitive Linguistics Conference. Book of Abstracts. Düsseldorf: ICLC, lk 115. https://iclc16.github.io/abstracts/ICLC16_BoA.pdf
- 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: Lexicography in Global Contexts, Ljubljana, 17–21 July 2018. Toim Jaka Čibej, Vojko Gorjanc, Iztok Kosem, Simon Krek. Ljubljana: Ljubljana University Press, lk 749–761.

- Tomasello, Michael 2006.** Acquiring linguistic constructions. – Handbook of Child Psychology. Kd II: Cognition, Perception, and Language. Toim William Damon, Deanna Kuhn, Robert S. Siegler. New York: John Wiley & Sons, Inc., lk 255–298.
<https://doi.org/10.1002/9780470147658.chpsy0206>
- Torrent, Timponi Tiago; Meireles Lage, Ludmila; Sampaio, Thais Fernandes; da Silva Tavares, Tatiane; da Silva Matos, Ely Edison 2014.** Revisiting border conflicts between FrameNet and Construction Grammar: Annotation policies for the Brazilian Portuguese constructicon. – *Constructions and Frames*, kd 6, nr 1, lk 34–51.
<https://doi.org/10.1075/cf.6.1.03tor>
- Östman, Jan-Ola; Fried, Mirjam 2004.** Historical and intellectual background of Construction Grammar. – *Construction Grammar in a Cross-Language Perspective*. (Constructional Approaches to Language 2.) Toim M. Fried, J-O. Östman. Amsterdam: John Benjamins, lk 1–10. <https://doi.org/10.1075/cal.2.01ost>

Ene Vainik (sünd 1964), PhD, Eesti Keele Instituudi juhtivteadur (Roosikrantsi 6, 10119 Tallinn), ene.vainik@eki.ee

Geda Paulsen (sünd 1973), PhD, Eesti Keele Instituudi vanemteadur (Roosikrantsi 6, 10119 Tallinn), Uppsala ülikooli eesti keele vanemlektor / kaasprofessor, geda.paulsen@eki.ee, geda.paulsen@moderna.uu.se

Jelena Kallas (sünd 1976), PhD, Eesti Keele Instituudi vanemteadur-projektijuht (Roosikrantsi 6, 10119 Tallinn), jelena.kallas@eki.ee

From dictionary to constructicon: Background, examples, and challenges

Keywords: lexicography, construction-based approach, constructicon, constructicography

A new field has emerged alongside constructional linguistic theories and electronic lexicography, known as constructicography (Lyngfelt 2018: 11). Its goal is to compile a constructicon, a resource that organizes and presents linguistic constructions much like dictionaries present words. This concept stems from the theoretical framework of Construction Grammar (see Fillmore 2006; Goldberg 2003; Hoffmann, Trousdale 2013). The main purpose of this study is to outline the necessary theoretical concepts, map out existing solutions, and identify the main challenges in creating different types of constructicons. This aims to inform decisions when beginning to create a potential Estonian language constructicon.

First, we discuss the main theoretical considerations and concepts essential for constructicon building. The basic assumption of this theory is that there is a continuum-like transition from grammar to lexicon, rather than a sharp boundary

between these two basic linguistic phenomena. According to this view, linguistic units are constructions – pairings of form and meaning with a variable number of lexical components and various levels of schematicity.

We further examine the theoretical and methodological principles behind current constructiographic practices, using examples from eight initiatives across seven languages: English, German, Japanese, Russian, Hungarian, Swedish, and Brazilian Portuguese (Perek, Patten 2019; Ziem et al. 2019; Janda et al. 2020; Lyngfelt et al. 2018; Ohara 2018; Fillmore et al. 2012; Torrent et al. 2014; Sass 2023). Meta-analysis reveals that most constructiographic databases are linked to FrameNet resources created for respective languages. The target audiences for these resources vary from language experts to second language learners and language technology applications. The coverage of constructions ranges from 73 to approximately 13,000 entries. The constructions include semi-schematic constructions, abstract valency patterns, argument structure constructions, and even idiomatic constructions located at the “lexical end” of the grammar–lexicon continuum. As constructions form networks characterized by inheritance relations, describing these relations is a crucial part of constructicon building.

Challenges discussed in the paper include defining a construction – whether narrowly or broadly; deciding if a constructicon should encompass general phrase structure constructions or be limited to idiosyncratic constructions; determining which constructions should be included in the resource and how to arrive at a nomenclature; handling constructions at different levels of abstraction; ensuring that the meta-language of descriptions is accessible to learners, native speakers, linguists, and NLP applications alike; and providing effective search functionalities in the dictionary/database for users.

Based on the analysis results, we highlight questions that need to be addressed if a constructicon is to be created as part of a relational lexicographical database for Estonian, specifically within the database of the EKI Combined Dictionary compiled within DWS Ekilex (Tavast et al. 2018).

Ene Vainik (b. 1964), PhD, Institute of the Estonian Language, Research Professor (Roosikrantsi 6, 10119 Tallinn), ene.vainik@eki.ee

Geda Paulsen (b. 1973), PhD, Institute of the Estonian Language, Senior Researcher (Roosikrantsi 6, 10119 Tallinn), Uppsala university, Senior Lecturer / Associate Professor, geda.paulsen@eki.ee, geda.paulsen@moderna.uu.se

Jelena Kallas (b. 1976), PhD, Institute of the Estonian Language, Senior Researcher-Project Manager (Roosikrantsi 6, 10119 Tallinn), jelena.kallas@eki.ee