

LINGVISTILISEST KOLONIALISMIST SÕNARAAMATUS

MARGIT LANGEMETS

Sissejuhatus

Ajaloost on teada, et sõnaraamatuid või muid sarnaseid teatmeteoseid (ametnike, preestrite käsiraamatuid, äritegemise abivahendeid jm) on maailmas koostatud juba umbes 3000 aastat tagasi ilmse rakendusliku eesmärgiga korrastada ja seletada ümbritsevat maailma. Kui kõige vanemad, idamaade sõnaraamatud (Pärsias, Hiinas, Egiptuses jm) on olnud ükskeelsed, et aidata mõista oma keele vanu ja pühaks peetavaid tekste, siis Euroopa omad on valdavalt olnud kakskeelsed ja vahendanud muukeelset (praktilist) tarkust, algul peamiselt kreeka- ja ladinakeelset. Kuigi ka sõnaraamatud, eriti ükskeelsed, tegelesid sõna tähendusega, sisaldades raskete sõnade lühiseletusi (glosse) ja/või sünonüüme, on eelistatud tähendusõpetuse alguseks lugeda pigem hilisemaid filosoofilisi arutlusi sõna tähenduse ja vormi üle, nagu on sedastanud Renate Pajusalu oma raamatus „Sõna ja tähendus” (2009: 34). Sõnaraamatusse puutumata on kokkuvõtte keeleteadusliku semantika peamistest suundadest teinud Silvi Tenjes (õpperaamatus „Semantika”, 2010). Semantika enda „tulemist” moodsasse keeleteadusse, süntaksi kõrvale, on eesti keeles peaaesjalikult selgitanud Haldur Õim, sh ka populaarsetes vormis (raamatus „Semantika”, 1974).

Sõnaraamatutöö on mingis mõttes nagu kahe kalda vahel voolav jõgi, kus ühel kaldal on lingvistika ning teisel seisab tegelik kasutaja. Sõnaraamat on neist mõlemast püüdnud ühekorraga kinni hoida, kord ühele, kord teisele lõivu makstes. Leksikograafia vahekorra keeleteadusega olen käsitlenud juba varem, peamiselt pilguga lingvistika poole (Langemets 2002, 2010). Meenutan seda alustuseks väga lühidalt. Siinse artikli vaatenurk tõukub samuti lingvistikast, aga pilk on suunatud mujale: esmalt jõe ehk leksikograafia endale, sellele kui tugevalt rakenduslikule alale, aga samavõrra ka kasutajale – liialdades võiks suisa öelda, et lingvistika kaldast hoopis ära. Sellist suunamuutust on mõned leksikograafia teoreetikud sõnastanud lausa kui pööret leksikograafias, kui sõnaraamatutöö olemuslikku teisenemist. Leksikograafia all pean ennekõike silmas kõike seda, mis seondub elektroonilise sõnaraamatuga.

Tühi koht lingvistika ja leksikograafia vahel

Leksikograafias on sõna tähendusliku külje avamist läbi aegade peetud (seletavate) sõnaraamatute põhiliseks teoreetiliseks probleemiks (vt viiteid Langemets 2010: 18–19, laiemalt Rundell 2012). Umbes 1960–1970-ndaist on leksikograafias hakatud mõtlema sellele, kuidas praktilist tööd toetada teo-



reetilise keeleteadusega, ja igatsetud, et lingvistika ütleks, kuidas asjad keeles on. Just ses osas on sõnaraamat pälvinud ka kriitikat: üks esimesi tuntud keeleteadlasi, kes kritiseeris sõnaraamatu semantilist esitust teoreetilisest vaatepunktist, oli Uriel Weinreich (1964). Ka uueaegne „Oxford Dictionary of English” (1998, 3. trükk 2005), mis oma eessõnas on väitnud end järgivat kognitiivse semantika teooriat, on pälvinud kriitikat (Meer 2000) just teoreetilise järjekindlusetuse ja meetodi ebatäpsuse osas. Ikka ja jälle on kahjatsetud, et „teoreetiline lingvistika on sõnaraamatut puudutanud parimal juhul vaid riivamisi” (Béjoint [1994] 2000: 177).

Samal ajal on (peamiselt alates 1970-ndatest) jätkuvalt soovitud ja proovitud sõnaraamatut keeleteadusega ühendada, on püütud r a k e n d a d a leksikaalsemantikas tehtud teoreetilisi järeldusi, jätkuvalt arutletud lingvistilise teooria ja leksikograafia vahekorra üle (Rundell 2012: 64 jj; Langemets 2010: 13–14). Võimalus selleks tekib jaokaupa, sõltudes ühe või teise keelenähtuse uuritusest, tänapäeval ka keeletehnoloogiast: korpuse töötlemise vahenditest, sellest, kuidas on korraldatud sõnaraamatu andmebaas (kas see on rohkem sõnaraamat või rohkem andmebaas) jpm.

Näiteks Thierry Fontenelle (1997) ehitas kakskeelsest sõnaraamatust leksikaalsemantilise andmebaasi, otsides Igor Melčuki leksikaalsete funktsioonide teooria toel lahendusi kahe keele kollokaadipaaride (süntagmaatiliseks ja paradigmaatiliseks) ühendamiseks. Baasi iseloomustab järgnev näide, kus (ühekordset) toidukogust väljendavad hulgasõnad on seotud funktsiooniga ‘Sing(food)’: Sing(food) = gulp, mouthful, morsel, portion, ration ...

Vaatasin kõrvutavalt, kuidas meie kakskeelsete sõnaraamatute põhjal sellist funktsiooni kokku saada: lähtusin otsisõnadest *portsjon* (näide 1) ja *suutäis* (näide 2) (näidetes alla joonitud), nii eesti-vene (a) kui ka vene-eesti (b) sõnaraamatus.¹ Funktsiooniga Sing(toit) hõlmatavad (eesti ja vene) hulgasõnad on esitatud näites 3. Vene-eesti poolel (b) on vaadatud sõnu, mis eesti-vene poolel (a) esinevad märksõna, seletuse osa või vastena. Näited annavad aimu, kui keeruline oleks olemasolevate, inimese koostatud sõnaraamatute põhjal saavutada teoreetiliselt korrastatud leksikaalsemantilist andmebaasi (siin nähtavaid ebaühtlusi on loetletud järgmises osas).

- (1a) **ports** .. **2. kõnek** (*[toidu]portsjon*) порция
portsjon .. (*kindlamõõduline hulk, kogus*) порция, порцион
- (1b) порция .. **portsjon**, norm, annus
 порцион .. **portsjon**, norm; *sõj. aj. toiduraha*
- (2a) **amps** .. **2. kõnek** (*hammustatud suutäis*) кусок
lõmps .. (*lusika- v suutäis*) глоток, хлебок
pala .. **1. (suutäis)** кусок ..; (*toit*) кушанье
suutäis .. **1. (korraga suhu mahtuv v väikene söögikogus)** кусок, кусочек, кус; (*joodava kohta*) глоток
- (2b) кусок .. **tükk**; pala; osa; кусок хлеба leivakäär, -tükk
 глоток .. lonks, rüüp, sõõm (*ka ülel.*)
 кусочек vt. кусок

¹ Sõnaraamatute veebiversioonid Eesti Keele Instituudi kodulehel <http://www.eki.ee/dict/ves> ja <http://www.eki.ee/dict/evs>. Kasutasin veebilehel käsku „Sama päring eesti-vene / vene-eesti sõnaraamatus” (25. V 2012).



кус .. suutäis; tükk, pala (*ka ülel.*)
 подачка .. (loomale visatav) pala, toiduraas, suutäis

- (3) Sing(toit) = ports, amps, lõmps, pala, suutäis ...
 Sing(еда) = кус, подачка, глоток ...

Omaette küsimus on, kas – ja kui jah, siis milleks – seda on vaja? Fontenelle (1997) kujutleb sellist andmebaasi ühelt poolt keeleteooriate katseväljana (s.t lingvistika jaoks), teisalt kollokatsioonide õpetamise abivahendina arvutipõhistes keeleõppesüsteemides (s.t rakenduslingvistika jaoks). Kui pabersõnaraamatu puhul võib mõelda, et kasutaja kohtub iga kord vaid üheainsa sõnaga, siis moodsates arvutirakendustes kuvatakse mingil viisil sarnased sõnad/seletused mõne sekundiga igäühe silme ette. Arvuti paljastab halastamatult pabersõnaraamatu, s.t inimese (ja sõnapõhise koostamise) juurde kuuluva loomuliku ebaühtluse, mida on võimalik vältida, kui kasutada mõistepõhiselt ülesehitatud andmebaasi.

Sõnapõhisus vs. mõistepõhisus

Eestis on mõistepõhiselt koostatud (onomasioloogilise) sõnaraamatu edemust sõnapõhise (semasioloogilise) ees propageerinud jõuliselt Arvi Tavast (2008), kes on väitnud end lähtuvat radikaalsest funktsionalismist, mille puhul semantilist tähendust pole olemas – viimane on üksnes „uurija enda abstraksioon kohatud kõnelejatähendustest” (2008: 130). Tavasti meelest on mõistlik vaade keelele (ja sõnaraamatule) instrumentalistlik (ehk rakenduslik), millele vastandub keelekeskne lähenemine (Fuertes-Oliveira, Bergenholtz 2011 on viimast nimetanud tekstilingvistiliseks lähenemiseks).

Samal ajal on Tavast rõhutanud, et mõistepõhisus käib üksnes koostamise, s.t andmebaasi ülesehituse, mitte sõnaraamatu kasutamise kohta. Kasutaja kohtub endiselt harjumuspärase tähestikulise sõnakaupa esitusega, mis on igati „loomulik ja vältimatu” (Tavast 2008: 144). Leksikograafidelt eeldab sel viisil koostamine teatavat söakust, et mugavusraamidest väljuda, samuti teatavat annust „mõtteviisilist mõistepõhisust” (Tavast 2008: 146–147). Kui mõistepõhisus on teada-tuntud põhimõtte oskussõnastike puhul, siis Tavasti meelest sobiks see hästi ka üldkeele sõnaraamatu jaoks, sest aitab lahendada tavalisi sõnaraamatu hädasid: vältida eikuhugi viivaid tühje viiteid ja sünonüümivastuolusid, reguleerida vastete mõistlikku hulka, samuti vastete ja lühiseletuste vahetada, vähendada üldist ebaühtlust jm, s.t parandada sõnaraamatu kvaliteeti. Sünonüümivastuolusid peegeldab näide 2, kus nt *lõmps* on *глоток*, aga *глоток* on *lonks*, *rüüp*, *sõõm* ning *lõmpsi* vaste *хлебок* puudub märksõnaloendist hoopis; *toiduraha*, mis vene-eestis (näide 1b) on sõna *порцион* (märgendiga markeeritud) vaste, esineb eesti-vene poolel märksõna *toit+* all kasutusnäitena koos tõlkega *деньги на пропитание ~ на прокорм* jne.

Samalaadsete sünonüümivastuoludega oleme kokku puutunud ka üheköitelist eesti keele (seletavat) sõnaraamatut tehes (Langemets jt 2010). Suures seletavas sõnaraamatus (EKSS 2009) on seletustes sünonüüme kasutatud üsna ulatuslikult (näide 4a) ning, juhtudes alustama sõnast *plõksutama*, leiab kasutaja seletuse pärast mitme lähisünonüümi läbivaatamist, siinsel



juhul alles seitsmendal katsel, sõna *klõpsuma* juurest (näites alla joonitud). Ühekõitelise sõnaraamatu üheks põhimõtteks oleme võtnud tähendust seletada, mitte loetleda sünonüüme. Tulemus, milleni hetkel oleme jõudnud (näide 4b), kergitab endiselt sellesama küsimuse: kas on tegemist sama mõistega või ei ole? Kui on, siis võiks seletus seda toetada, s.t olla üks ja seesama kõigi sama mõistet esindavate sõnade puhul. Nüüd peab kasutaja ise järeldama, kas „lühike kõlav heli” (*plõksutama, plõksima*) on sama mis „kerge lühike heli” (*klõpsutama, klõpsima*); kas pildistades üksnes plõksutatakse, aga ei klõpsutata; kas semikooloniga eraldatud seletuse osa (*plõksutama*) on mingil viisil teistsuguse tähtsusega kui sõnaga „või” eraldatud osa (*klõpsutama*)? Tavast on eriliselt rõhutanud, et kui mõiste ei ole ühe sõnaraamatu piires üks ja sama, siis „on hädavajalik kasutajale öelda, mis neil siis vahet on” (2008: 147). Kahjuks tuleb nentida, et tähenduste selline ühtlustamine sõnapõhises baasis on ilmselt võimatu ning et mõistepõhiseski baasis võib see olla üle mõistuse ajakulukas.

(4a) **plõksutama** plõksu v. plõkse tekitama; midagi plõksu(de)ga tegema; plõksima; klõpsutama
plõks annab edasi plõksatust; klõps
plõksima plõksutama
plõksatus plõks, klõpsatus
klõpsutama klõpsuma panema; klõpsumist tekitades midagi tegema
klõpsatus klõpsatamine
klõpsatama järsku, korraks klõpsuma; klõpsuga uude seisu minema
klõpsuma järsku lühikest (teravat) heli v. seesuguseid helisid andma, neid kuuldavale tooma

(4b) **plõksutama** nupule vm vajutades (korduvalt) lühikest kõlavat heli tekitama; sellist heli tekitades korduvalt nupule vm vajutama või vajutades midagi tegema (nt pildistama)
plõksima lühikest kõlavat heli või selliseid helisid tekitama; sellist heli tekitades korduvalt nupule vm vajutama või vajutades midagi tegema (nt pildistama)
klõpsutama kergelt lühikest heli tekitama või sellist heli tekitades korduvalt midagi vajutama
klõpsima kergelt lühikest heli kuuldavale tuues korduvalt vajutama

Üht Tavasti nõuannet oleme siiski püüdnud ühekõitelises eesti keele sõnaraamatus (Langemets jt 2010) järgida: kõige kasutajasõbralikum on sõnapõhises (termini)baasis korrata kogu infot kõigi sünonüümide juures koos vastastikuste viidetega (Tavast 2008: 157). Hõlpsamini hõlmatavamatel juhtudel kui eelmainitud kõlasõnad (näide 4) oleme seda põhimõtet rakendanud (näide 5, alla on joonitud lingid sama sisuga sünonüümidele).

(5) **läinud**² <adj> käesolevale, praegusele vahetult eelnenud, eelmine, möödunud *Läinud nädalal haigestus kollatõppe veel kümmeaastane inimene.*
eelmine <adj> käesolevale, praegusele vahetult eelnenud, läinud², möödunud *Käisin Roomas eelmisel aastal.*
möödunud <adj> käesolevale, praegusele vahetult eelnenud, eelmine, läinud² *Möödunud aastat võib pidada kordaläinuks.*





Üldkeele sõnaraamatu puhul katsetatakse Eestis esimest korda² mõistepõhise koostamise meetodit keskmise mahuga (40 000 märksõna) eesti-läti/läti-eesti sõnaraamatu jaoks, mille koostamisel tekivad korraga mõlemad keelesuunad (üks pool, eesti oma, on andmebaasis automaatselt juba olemas).³ Sõnaraamatutegu algab loodetavasti 2012. aasta lõpus Läti Keeleagentuuri ja Eesti Keele Instituudi ühistööna. Kava järgi peaks sõnaraamat(ud) valmima väga ruttu, kolme aasta pärast. Mõistepõhisest sõnaraamatutööst tuleb kindlasti juttu ka edaspidi.

Lingvistiline kolonialism

Tunnistan, et olen seni enam-vähem arvanud nagu Atkins jt (2008: 130), et kuigi töö, mida leksikograafid teevad, on loomu poolest rakenduslingvistika, on väga hea, et sõnaraamatu külge „r a k e n d a t a k s e lingvistikat” – kui väljenduda Karlfried Knappi sõnadega siinses ajakirjanumbris –, nagu hobust rakendatakse vankri või tõlla ette. Olen mõelnud, et teoreetiline keeleteadus (eriti leksikaalsemantika) võiks mingil moel olla info esitamise alus, niivõrd, kui võrd võib aimata, et lingvistidel on õnnestunud „leida üksiktähenduste tagant semantiline süsteem” (Pajusalu 2009: 34). Seepärast uuringi järgnevas lähemalt teooriat, mis end ja sõnaraamatut sellele otsustavalt vastandab.

Lingvistilise kolonialismi mõiste on kasutusele võtnud Århusi Ülikooli leksikograafiakeskuse⁴ ümber koondunud teadlased Henning Bergenholtz, Rufus Goyws, Sven Tarp ja mitmed teised (vt Fuertes-Oliveira, Bergenholtz 2011), pidades selle all silmas teoreetilist leksikograafiat, mis määratleb ennast lingvistika osana ja on leksikograafias valitsenud vähemalt 1970-ndatest peale. Selle tuntumad esindajad on Ladislav Zgusta (1971), Franz Josef Hausmann (Hausmann jt 1989–1991), Herbert Ernst Wiegand (1977, 1984) jt. Selle teooria järgi on sõnaraamatu mõte tegelda ennekõike sõnaraamatu sisuga, kirjeldada keelt võimalikult täielikult ja võimalikult lingvistiliselt, s.t teaduslikult. Sõnaraamatut nähakse kui teksti iseeneses. Lingvistiline kolonialism valitseb Århusi koolkonna meelest sõnaraamatus kõike: märksõnade valikut ja esitust, sõnartikli struktuuri, keeleüksuste kirjeldust, sõnatähenduste esitust jne. Sõnaraamatu lingvistilist struktuuri ja põhimõtteid kirjeldatakse üksikasjalikus leksikograafilises profiilis (Atkins jt 2008), meil vanemas pruugis sõnaraamatu prospektis (EKSS 1961, EMS 1968). Selle suuna kurbloomis on, et vaatamata leksikograafide lingvistilistele pingutustele pole lingvistid enamasti sõnaraamatut keeleteadusena tunnistanud ning leksikograafia ja lingvistika vahel laiutab ikka tühi koht.

Århusi koolkond kõneleb uue, sõnaraamatutöö põhimõtteid põhjalikult muutva funktsionaalse leksikograafia teooria poolt ning nad sõnastavad seda kui pööret leksikograafias (ingl *paradigm shift*, Fuertes-Oliveira, Bergenholtz 2011: 27, 55, 121). Ühtlasi tulevad nad meelde, et esimesed sõnaraamatud olid samuti väga praktilised käsiraamatud (Eestis näiteks saksa pastorite abi-

² Mõistepõhiselt on üles ehitatud ka eesti keele fraseoloogia elektrooniline andmebaas (ligi 21 000 leksikaalset üksust, vt Öim, A., Öim, K. 2011).

³ Sõnaraamatu lähtebaasiks on eesti-X sõnastikupõhi <http://exsa.eki.ee/> (25. V 2012), mida eesti-läti/läti-eesti projekti jaoks on oluliselt muudetud.

⁴ Aarhus University, Centre for Lexicography <http://bcom.au.dk/centlex/> (25. V 2012).





mehed misjonitööl). Lingvistika ülemvõim on suhteliselt hiline nähtus ning „vabastusliikumise” (*linguistic liberation*, Fuertes-Oliveira, Bergenholtz 2011: 25) avangard on elektroonilised ehk veebisõnaraamatud.

Funktsionaalse leksikograafia teooria

Funktsionaalse leksikograafia teooria⁵ (*function theory of lexicography*, Fuertes-Oliveira, Bergenholtz 2011) on oma olemuselt üdini praktiline (rakenduslik, instrumentalistlik) ning ennekõike suunatud uutele kanalitele ja kasutajate uutele harjumustele: elektroonilisele keelevarale, väga heale oskusele navigeerida internetis, nutitefonis vm. Teooria tugineb potentsiaalsele kasutajale ja kasutamissituatsioonile ning tegeleb tõelise e-leksikograafiaga, mitte pabersõnaraamatute elektrooniliste versioonidega. Kasutaja vajadustest, andmete liiasusest, „tõelistest” ehk virtuaalsetest e-sõnaraamatutest jpm on kirjutanud juba 15 aastat varem ka Atkins [1996] (2002), siiski on Arhusi teadlaste meelest hilisem Atkins jt (2008) ikkagi lingvistilise kolonialismi meelevallas.

Toon järgnevas välja mõned põhipunktid, mis funktsionaalset lähenemist iseloomustavad. Kuna need korduvad üsna üheselt eri autorite peatükkides, viitan üldiselt kogu raamatule (Fuertes-Oliveira, Bergenholtz 2011), mis on komponeeritud tervikteosena (nt on peatükid läbivalt nummerdatud). Vaid mõnel juhul toon mõne autori eraldi välja. Püüan vahendatavat kõrvutada ka meie oludega.

Leksikograafia ei ole lingvistika, leksikograafid ei ole lingvistid. Sõnaraamat on praktiline abivahend, ja seda ei tohi hetkekski meelest lasta. Mis tahes sõnaraamatuga on tegemist, kas etümoloogilise, seletava, õigekeelsuse vm sõnaraamatuga, tuleb seda koostada kindlaksmääratud praktilise sihiga (kellele?, mis otstarbeks?, mis situatsioonis?), mitte lingvistilise teooria esitamise mõttega. Teooria nimetuses esinev fraas „leksikograafiline funktsioon” tähendabki seda, et iga sõnaraamatu puhul on praktiline siht hoolikalt läbi mõeldud. Just internetiajastu on see, mis „vabastab” sõnaraamatu olemast lingvistika allvaldkond, samal ajal kohustab täitma hoopis laiemat, infoteadusega piirnevat rolli, nagu kohustab ka mõtlema, kes leksikograaf tänapäeval/tulevikus ikkagi on: üldjuhul meeskonnaliige, mitte üksiktegitaja, kusjuures meeskonda kuuluvad leksikograafid (mitte lingvistid!), keeletehnoloogid ja erialaspetsialistid, kui tegemist on oskussõnastikega.

Tulevik üheskoos minevikuga. Kuigi selle koolkonna esindajad vastandavad end lingvistilisele kolonialismile, ei viska nad koos pesuveega last välja: traditsiooniline sõnaraamat sisaldab palju väärtuslikku, mida tuleks lihtsalt palju paremini ära kasutada. Traditsioonil on leksikograafias väga suur roll, mistõttu nüüdisaja e-sõnaraamat peaks ennekõike keskenduma uutele kanalitele ning inimeste uutele harjumustele. Vana kuld tuleb pakkida uut moodi. Näiteks CD-sõnaraamatut ei peeta kuigi vääriliseks esindama tänapäeva leksikograafiat, sest see on üldjuhul pelk pabersõnaraamatu e-versioon.

⁵ Ingliskeelses kirjanduses on seda nimetatud ka leksikograafiliste funktsioonide teooriaks (*theory of lexicographical functions*), samas on tugevasti kaheldud, kas see üldse teooria mõõtu välja kannab (Rundell 2012: 57).



Pabersõnaraamatud ei ole kuhugi kadunud ega kao arvatavasti veel niipea. E-sõnastike kaudu võivad nendesegi jõuda mõned uued ideed, kusjuures põhimõtte peaks olema sama: sõnaraamatu ülem eesmärk on aidata kasutajat, tabada tema soove, harjumusi ja kasutamissituatsioone. Inimese pea sisse muidugi ei näe, aga veebisõnaraamatute puhul on mõnel määral abi loogifailidesse talletunud infost. Põhiline on, et leksikograafias nähtaks selgelt kaht eristuvat poolt: paber- ja veebisõnaraamatut.

E-sõnaraamat ei ole sõnaraamat, vaid tööriist. E-sõnaraamat on info juurde pääsemise abivahend ning leksikograafia on tänapäeval infotehnoloogia. Tööriistana peaks e-sõnaraamat võimaldama ligipääsu ka (sellest) sõnaraamatust väljapoole jäävale infole, muu maailma infole. Tavapärase sõnaraamatu asemel tuleks kõnelda uut tüüpi, palju laiemast infotehnoloogilisest teatmeteosest. Kogu tulevane areng on puhtalt infotehnoloogiline, hõlmates kolme külge, millest ükski ei tohiks domineerida: kasutajat, ligipääsu infole ja andmeid.

Sven Tarp (2011: 58–62) on e-leksikograafia jaganud kolme kategooriasse, lähtudes sõnaraamatu kui tööriista olemuslikust sisust.⁶

1. Masinloetavad sõnaraamatud ehk kopeerijad (ingl *copycats*) – need on sõnaraamatud, mis on paberilt arvutisse skaneeritud või ka lihtviisil sisse tipitud ning sellisena veebis kättesaadavaks tehtud (sageli pdf-i kujul). Sel viisil ligipäätavad sõnaraamatud võivad originaalis olla käsikirjalised (s.t vanad) või muul moel haruldased või esindada keeli/maid, milles elektroonilisus pole (veel) nii valdav (nt mari sõnastik⁷, hispaania–aimara sõnastik⁸). Ent seda moodust on kasutatud Eestis ka tänapäevase pabersõnastiku rööpkujuna (nt väike paronüümisõnastik⁹). Kopeerimine on laialt kasutusel lepingute, juhendite, manuaalide jt tehniliste tekstide puhul, kus on oluline dokumenti täht-tähelt paljundada. Paberteksti kopeeriv sõnaraamat on aga kasutaja jaoks suhteliselt ebamugav, eeskätt oma aegluse ja kohmakuse tõttu.

2. Paindlikumad e-sõnaraamatud ehk kiiremad hobused (*faster horses*) – sellele kategooriale on nime andnud autotööstur Henry Fordiga seostuv lugu, mis räägib sellest, et kui ta oleks inimestelt küsinud, mida nad tahavad, siis oleksid nad vastanud, et kiiremat hobust (s.t mitte autot). Sama on ka sellesse rühma kuuluvate sõnaraamatutega: üldjuhul on nad pabersõnastiku veebiversioonid, kus on küll kasutatud uusi tehnoloogilisi võimalusi, samas üsna piiratud moel, näiteks võimaldades linke või metamärkidega otsijärjendeid. See tagab mõistagi andmete kiirema otsimise, aga ei enam. Tüüpiliseks „kiiremaks hobuseks” peab Tarp (2011) Longmani ükskeelset sõnaraamatut.¹⁰ Katseks vaadatud sõna *salt* ('sool') ilmub seal päringu peale

⁶ Tegelikult on jaotisi neli – mina olen viimases kategoorias ühendanud kolmanda ja neljanda rühma.

⁷ Galkin, I. S. 1990 (toim). Словарь марийского языка I–VI. = Marij muter I–VI. Joškar-Ola. <http://www.marlamuter.ru/downloads/01-tom.pdf> (25. V 2012).

⁸ Yatiqirinaka Aru Pirwa. Peru: Lima, 2005. http://portal.perueduca.edu.pe/Docentes/xtras/pdf/dicc_aymara.pdf (25. V 2012).

⁹ Väike paronüümisõnastik. Koostanud Helen Plado ja Katrin Mandra. Tartu, 2008. http://engine.koduleht.net/templates/keelehooldekeskus/files/mdl_files.php/Mandra_Plado_V_aike_paronuumisonastik.indd.pdf (25. V 2012).

¹⁰ The Longman Dictionary of Contemporary English Online <http://www.ldoceonline.com/> (25. V 2012).

esiteks üksnes koos valdkonnamärgenditega (näites 6 alla joonitud), mille klõpsates saab liikuda – sugugi mitte tarviliku sõnatähenduse juurde, nagu ma ootasin –, vaid värvilisena esitatud kokanduse tähendusvälja juurde, mille hulgas leidub ka sõna *salt*. Kolmanda klõpsu alt avaneb sõnaseletustega artikkel koos soolatospi pildiga ning märksõna ja näidete kuulduva hääldusega.

- (6) **salt1** noun Cooking, Chemistry
salt2 verb Food, Cooking
salt3 adjective Food

Meil esindavad seda tüüpi sõnaraamatut kõik Eesti Keele Instituudi kodulehel olevad sõnaraamatud, sh suur eesti keele seletav (EKSS 2009) ja õigekeelsussõnaraamat ÕS 2006.¹¹ Paindlikkus seisneb otsimisvõimaluses sõnaartikli osadesse paigutatud infoüksuste kaupa (märksõna, tähendus/seletus, kasutusnäited/lausenäited, ühendmärksõna osa, tüübinumber, sõnaliik jm). Näidetes 6 ja 16 kirjeldatud jaokaupa lähenemist meil (veel) ei ole.

3. Tõelised e-sõnaraamatud ehk nutisõnaraamatud¹² – siia rühma kuuluvad dünaamiliste andmete ja dünaamiliste artiklitega e-sõnaraamatud, mis püüavad arvestada kasutaja eri vajadustega eri kasutamissituatsioonides ehk leksikograafiliste funktsioonidega. Au olla siia arvatud on Århusi enda sõnaraamatukeskuses ettevalmistatud e-sõnaraamatutel.¹³ Kasutaja saab – kui oskab – määratleda oma tegevust (kasutamissituatsiooni/leksikograafilist funktsiooni), vastavalt sellele kuvatakse infot, ning valida lühemaid või pikemaid infoüksusi.

Aste kaugemale lähevad nutisõnaraamatud, mis võimaldavat personaalset (individualiseeritud) lähenemist ning kombineerivad sõnaraamatuotsingut (leksikograafi poolt läbi mõeldud) veebilehitsemisega koos andmete töötluse ja taaskasutusega. Tarp (2011) ei tea nimetada veel ühtki leksikograafilist tööriista, mis vastaks sellistele nõudmistele. Tõsi, vastuväiteid esitatakse sssamas raamatus: Serge Verlinde (2011) meelest on seda tüüpi tööriistad juba olemas, näiteks Hollandis väljatöötatud Interactive Language Toolbox.¹⁴

Nutisõnaraamatute klassi võiks ilmselt kuuluda automaatselt koostatud kollokatsioonide sõnaraamat,¹⁵ mis sõnaartikli kokkupanuks kasutab korpuse töötlemise programmi Sketch Engine (ja funktsiooni GDEX 'Good Dictionary Example') (Kilgarriff jt 2004, 2008), mis Eestiski tarvitusel on (Jürviste jt 2011). Eesti sõnaartikkel *sool* saaks selle tööriista abil järgmise kuju (valisin ainult kõige kõrgema esilduvusindeksiga kollokatsioonid, vt tabelit 1). Ühtlasi tulevad mustvalgel ilmsiks tööriista, täpsemalt eesti keele korpuse

¹¹ Eesti Keele Instituudi keelekogud <http://portaal.eki.ee/keelekogud.html> (25. V 2012).

¹² Nutisõnaraamat on minu ühendav termin [M. L.]. Tarp (2011) püsib autotööstuse metafooride juures: kolmas rühm on tal Fordi T-mudelid (tootmises 1908–1927), neljas Rolls-Royce'id.

¹³ Århusi leksikograafiline looming kätkeb endas peamiselt oskussõnastikke, nii kakskeelseid (taani-inglise juurasõnastik ja taani-hispaania börsisõnastik) kui ka ükskeelseid (muusika- ja raamatupidamissõnastik), aga esindatud on ka üldkeele sõnaraamatud: suur (paberil 6-köiteline) ükskeelne veebisõnaraamat („Den Danske Ordbog”, 110 000 märksõna) ja taani püsiväljendite sõnastik (22 000 väljendit). Vt <http://bcom.au.dk/centlex/> (25. V 2012).

¹⁴ Interactive Language Toolbox esitlus, vt http://videlectures.net/elex2011_verlinde_age/ (25. V 2012).

¹⁵ The GDEX Demo Dictionary <http://forbetterenglish.com/> (25. V 2012).



senise märgendamise probleemid (tabelis alla joonitud): morfoloogiline märgendus ei erista pärisnimesid (Surnumeri, *saaga* – „Vee ja soola saaga”) ega homonüüme (sool : soola / sool : soole); kuna puudub süntaktiline märgendus, siis esineb *raputamine* üksiku kollokaadi staatuses (*pro* püsiühend *lahtistele haavadele soola raputama*), kvantorid *teelusikatäis*, *puud*, *tiba* ja *näpuotsatäis* on analüüsitud erinevalt (seeasemel võiks olla omaette grammatiline seos [part_modifies]), fraasist *soola ja pipraga* on *sool* analüüsitud objektiks jne.

Tabel 1.

Sõna *sool* artikkel automaatses kollokatsioonide sõnaraamatus (grammatiline seos, kollokaat, näide)

[a_modifier]	jäme :	Kõigile taimedele peaks viie liitri vee kohta lisama 2 spl jämedat soola .
	puistatav :	Soolarante ei tekita mitte ainult tänavatele puistatav sool .
	teelusikatäis :	Eestlane sööb päevas sisse <u>kaks-kolm teelusikatäit soola</u> .
[subject_of]	lahustuma :	Kui sool ära lahustub , kerkib kast pinnale.
[object_of]	maitsestama :	Vajadusel maitsesta soola ja pipraga.
	puistama :	Kas kõnniteedele võib puistata soola ?
[gen_modifier]	puud :	Me oleme koos söönud ära <u>seitse puuda soola</u> ja oleme sõbrad juba vähemalt 15 aastat.
	surnumeri :	Meresoola müüakse apteegis, eriti sobiv on Surnumere sool .
	glutamiinhape :	Inimeste ülitundlikkust glutamiinhappe soolade suhtes kutsutaksegi hiina restorani sündroomiks.
	nõudepesumasin :	Nõudepesumasina sool on hooldusvahend, mis soodustab nõudepesu, sest pehmedab pesuvett ja takistab katlakivi teket.
	tiba :	Sinna ei ole pandud isegi mitte tiba soola .
	strontsium :	Igale lahusele lisatakse 10 ml strontsiumi soola lahust(3.4) ja a täiendatakse kuni 100 ml-ni veega.
	näpuotsatäis :	Ka võib piimale lisada <u>näpuotsatäie soola</u> – nii püsib see kauem värskel.
[gen_modifies]	mikrofloora :	Raskete haiguste ja antibiootikumide ravi järel aitab see taastada soolte mikrofloorat ja tuua tagasi söögiisu.
	saaga :	1999. aastal sai Joonase ja Kanguri lugu " Vee ja soola saag " teise koha.
	ladestumine :	Migreeni põhjuseks võib olla soolade ladestumine organismi.
	raputamine :	Kogu see protsess on meie <u>lahtistele haavadele soola raputamine</u> .
	seedetegevus :	<u>Mao ja soole seedetegevus</u> paraneb, maksa töökoormus väheneb.
[ja/või]	pipar :	2 spl mett, näpuotsatäis musta pipart ja soola .

E-sõnaraamat ei ole sõnaraamat, vaid info panipaik. Kogu info on (ühelikordset, ilma dubleerimiseta) paigutatud (relatsioonilisse) andmebaasi, millest erinevate linkide ja seoste abil luuakse (virtuaalne) sõnaraamat. Näiteks pole taani raamatupidamissõnastikus traditsioonilist märksõnaloendit – andmebaasis „töötab” seeasemel spetsiaalne andmeesisstruktuur. E-sõnaraamat defineeritakse elementidega, mida selles konkreetses sõnaraamatus vajatakse. On oluline, et ühest ja samast andmekogust saab tuletada uusi (ja eri tüüpi) sõnastikke.





Iga andmeelement olgu ainukordne. Kõige tähtsam andmebaasi juures on vältida liiasust. Info peaks andmebaasis paigutuma nii, et üksused esinevad – kui vähegi võimalik – üks kord ühes kohas. Vaid nii saab vältida ebaühtlusi, mis tekivad samade andmete paiknedes mitmes kohas. Või vältida ajakulu, mis kulub kõigi paranduste sisseviimiseks. Meie EELEXi sõnastikusüsteemis on selle tarvis loodud spetsiaalne hulgi-paranduste funktsioon (aga ka see ei taga vigadeta tulemust, nagu teame). Seda poleks vaja, kui andmebaas oleks teisiti üles ehitatud.

Näiteks võiks üks ja sama kasutusnäide teenida mitut märksõna. Kui vaadata eespool jutuks olnud sõna *suutäis* esinemist kasutusnäitena, näeme, kuidas see esineb eesti-vene (näide 8) ja vene-eesti (näide 9), nagu ka eesti-vene aluseks olnud suures seletavas (näide 7) mitu korda üsna lähedases, mõnikord lausa samas vormis (näited 7a, b, c, 8a, b, 9).

(7a) Anna üks soolane suutäis. [*sub* soolane]

(7b) Marineeritud seemed on mõnus suutäis. [*sub* suutäis]

(7c) Parem suutäis soolast kui maotäis magedat. *Vs.* [*sub* mage, maotäis, parem, soolane]

(7d) Mõnus soolane suutäis. [*sub* mõnus]

(8a) anna üks soolane suutäis [*sub* soolane]

(8b) marineeritud seemed on mõnus suutäis [*sub* suutäis]

(9) rasvane suutäis [kollokatsiooni tõlkena *sub* жирный ja кусок ja лакомый]

Sõnaraamat olgu monofunktsionaalne, mitte polüfunktsionaalne. Pole mõtet kuhjata sõnaraamatusse kõike, kujutledes, et see teeb ta tõeliselt polüfunktsionaalseks, palju olulisem on osata olla monofunktsionaalne, tuua välja eri ülesannete lahendamiseks tarvilikke momente. Mahukad väga mitmekülgsed (polüfunktsionaalsed) sõnaraamatud võivad oma personaalse infosoo-viga tulnud kasutaja päris ära kurnata. Iga internetikasutaja on arvatavasti kokku puutunud infostressiga, võib-olla ka infosurmaga, s.t ta on uppunud infolaviini, enne kui on vajaliku infoga kaldale jõudnud. Suur ei pruugi olla hea (põhjusel, et ta on suur), samas võib hea olla väike sõnaraamat, kui kasutaja ja kasutamissituatsioon on õnnestunult määratletud.

Ärhusi koolkond on alustuseks välja pakkunud aktiivse/passiivse sõnaraamatu skaalat pidi kulgevad valikud (leksikograafilised funktsioonid): 1) kommunikatiivne funktsioon – hõlmab omakorda teksti moodustamist, teksti mõistmist, tõlkimist jm; 2) kognitiivne funktsioon – pakub kasutajale mitmesugust lisainfot, näiteks entsüklopeedilist, on „tahan kõike teada”-funktsioon; 3) keele dokumenteerimise funktsioon jm. Kasutajale peaks pakkuma võimalust oma tegevust ja vajadusi määratleda (mis minu arvates pole see ülesanne sugugi kergete killast).

Kui veebisõnaraamat on info panipaik, nagu ühe põhimõttena eespool öeldud, siis peaks leksikograaf suutma otsustada või kasutajauuringute põhjal järeldama, mida milleks vajatakse, et seejärel infot valikuliselt välja anda. Taani püsiväljendite sõnaraamatu (22 000 väljendit) põhjal on esitada ka ühe aasta põhjal saadud tulemused (Fuertes-Oliveira, Bergenholtz 2011: 52). Kui



kasutajatele oli loodud võimalus kitsendada (sõnapõhist) otsingut oma ligema huvi järgi, s.t eritleda leksikograafilisi funktsioone, siis umbes pooled (55,5 %) kasutajatest valisid „tahan kõike teada“-funktsiooni ja rohkem kui kolmandik (37,6 %) kasutas passiivset, teksti mõistmise funktsiooni. Abi teksti moodustamisel vajasid püsiühendite kasutajatest vaid vähesed: 3,5 % kasutas infot väljendi kasutamise kohta ning 3,4 % otsis spetsiifilist tähendust.

Numbrid, mis kirjeldavad taani ükskeelse (teksti moodustamiseks mõeldud) sõnaraamatu kasutamist, on kõnekad (Fuertes-Oliveira, Bergenholtz 2011: 192). Sõnaraamatus on 110 000 märksõna – sellest umbes pooli sõnu (ligi 60 000) polnud nelja aasta jooksul mitte keegi (isegi ebaõnnestunult) otsinud! Kahjuks ei koorunud Århusi teadlastel selle kohta välja mingit süsteemset seletust: kummaski rühmas, nii otsitute kui ka unustatute hulgas, olid üsna ühtmoodi sõnad (igapäevasõnad, harvaesinevad, neologismid, halvustavad jm).

Kasutajad on eri tasemega – vastavalt tuleks esitada ka info. Enamik on nn keskmised kasutajad, tavainimesed ilma suuremate nõudmisteta, mõned on pooleksperdid, aga üks tuhandest võib olla ekspert, erihuvidega kasutaja. Seepärast peaks kasutaja saama oma eelistust määrata, süsteem valib seejärel info esitamiseks sobiva taseme. Samuti võiks tööriist suuta kasutajat suunata, talle üht-teist kasulikku lisaks soovitada. Loomulikult ei hakata määratlema igatühe taset eraldi, mõistlik on opereerida paari-kolme tasemega: nt tavainimene (ja üliõpilane), poolekspert ja ekspert (aga võib olla ka muid, nt õpilane).

Kui mõelda meie praeguste võimaluste peale, saaks tähenduse seletusi esitada kahel eri raskusastmel: algtasemel (nagu põhisõnavara sõnastikus, näide 10) või kesktasemel (nagu ühekõitelises eesti keele sõnaraamatus, näide 11). Sünonüümid on lingid vastavasse artiklisse (näites alla joonitud). Nagu juba eespool öeldud, on ühekõitelises tähendusi seletatud, kohati lihtsustatud, mitte piiratud krüptiliseks jääda võivate sünonüümidega, nagu suures seletavas (näide 12b, vt ka näidet 4a).

- (10a) **sool** valge aine, mida saab näiteks merest ja mida inimesed lisavad toidule, et anda tugevamat maitset
- (10b) **jaotama 1** midagi osadeks tegema, *sama mis jagama* (1. täh)
2 kellelegi midagi osade kaupa andma, *sama mis jagama* (2. täh)
- (11a) **sool¹** valge kristalne mereveele omase maitsega aine, mida kasutatakse toidu maitsestamiseks ja konserveerimiseks, keedusool, söögisool
- (põhi- või lisainena mitmesugustes segudes)
 - miski, mis lisab loole, naljale vm värskust või pikantsust
- (11b) **jaotama 1** tükkideks või (mõttelisteks) osadeks tegema, jagama (1. täh)
- liikideks või rühmadeks jagama, teatud klassi paigutama, liigitama
- 2** mitmele või paljudele midagi (osade kaupa) andma, jagama (2. täh)
- (12b) **jaotama 1** jagama (1. täh)
2 jagama (3. täh)

Või teine näide. Häälde märkide ja muutmisinfo puhul võiks üldkeele sõnaraamatus (kirjapildis) eristada vähemalt kaht taset: 1) ainult häälde või



2) täielik morfoloogiline transkriptsioon. Veebisõnaraamatus täiendab kirjapilti ka kuuldav hääldus (esialgu on see olemas üksnes eesti keele põhisõnaraava mahu ulatuses, vt Jürviste jt 2011).

Vaikimisi (tavainimese jaoks) oleks kirja-pildis antud üksnes hääldus (lihtsustatud foneetilisest transkriptsioonist) ja ainult tsitaatmärksõna jaoks, (eesti keeles ebaregulaarne) rõhk on märgitud silbituuma ees keskpunktiga (·) (näide 13). Sõnamuutmisinfo on esitatud ilma lisamärkideta ortograafilises kirjaviisis. Ekspordil on võimalik morfoloogiaplokis avada ka morfonoloogiline transkriptsioon, mis kuvab nii väldet kui ka rõhku, nii algvormi kui ka sõna muutevormide puhul. Suhteliselt keeruline, aga ekspordi jaoks omandatav, on rõhu esitamine kahel viisil: kui välde on sõnamuutmises oluline (s.t tegemist on I- või II-vältelise silbiga), siis akuudiga (˘) (näide 14); kui välde ei ole oluline, siis III välte märgiga (˘) (näide 15). Muukeelse kasutaja ja eesti keele õppija jaoks on morfonoloogiline transkriptsioon iseenesest vajalikum kui haritud eestlase jaoks, kes oskab niigi (?) sõnu muuta, aga kindlasti peaks see olema kättesaadav kuuldava hääldusena.

(13) *allegro* [alleegro]
croissant [kruass·a(n)]

(14) allergik <all´ergik -u 2 s>

(15) sotsioloogia <+ l`oogia 1 s>

Artiklid kuvatakse jaokskaupa. Kasutaja jaoks üle jõu käivaltpikkade artiklite üle on kurdetud pabersõnaraamatute puhul ammu. E-sõnaraamatus on võimalik materjali organiseerida nii, et infot näidatake jaokskaupa: näiteks suurte sõnade puhul vaid eri üksuste (tähenduste) algusotsa (nt umbes kolm rida, nagu Cobuildi veebisõnastikus,¹⁶ vt näidet 16). Kasutaja saab nõnda sõnadest või tähendustest esmase ülevaate, ja pahatihti polegi tal rohkemat vaja. Et enamik inimesi niikuinii lihtsalt guugeldab ega malda kasutada peenemat tööriista, sellele on osutanud mitmed arvutileksikograafiaga tegelejad (nt Lew 2011). Cobuildi veebisõnastikus on lehe servale, põhiartiklist lahku, paigutatud kollokatsioonid ehk sõnad koos oma kaaslastega (ingl *word partnership*). Sõnartikkel ise on läbinisti lingitav: iga seletuses olev sõna on link vastava artikli juurde (veider on, et (tühjaks) lingiks on tehtud ka sõnad, mis ise märksõnana ei esine, nt *strong-tasting*).



(16) **salt** / sɔlt

/N-UNCOUNT **Salt** is a strong-tasting substance, in the form of white powder or crystals, which is used to improve the flavour of food or to preserve it. Salt occurs naturally in sea water.

More [avaneb 7 üksust, sh 4 tähendust ja 3 püsiühendit]

¹⁶ Cobuild Online Dictionary MyCobuild.com <http://www.mycobuild.com> (25. V 2012).



Robert Lew (2011: 243–244) on inglise veebisõnaraamatuid analüüsisid hinnanud parimaks just artiklit natuke paotavat infojaotamist (nagu Cobuildis). Teisi suvandeid, pelka märksõnaloendit või täit artiklit peab ta kehve-
mat tüüpi lahenduseks. Samamoodi on infot jaotatud Macmillani õppesõnastikus¹⁷ („Näita vähem / Näita rohkem”).

Kokkuvõtteks

Ma ei taha mingil juhul väita, et tuleks loobuda sõnaraamatu lingvistilisest mõõtmest, rääkimata, et pean seda lingvistiliseks kolonialismiks. Vastupidi, olen veendunud, et leksikograafia ilma lingvistikata oleks nonsens: leksikograafid vajavad teoreetilist tuge, et keeleandmetes orienteeruda, tabada reeglipärasust, olla süstemaatiline ja järjekindel, registreerida olulist. Ise küsimus on, kuidas seda esitada. Funktsionaalse leksikograafia teooria polegi nagu päris teooria, kuivõrd ennekõike keskendub ta olemasoleva (lingvistilise) info paremale, läbimõeldumale esitusele, mitte teadmiste uut laadi tõlgendusele, keelenähtuste tõestamisele-seletamisele vmt. Kui jätta kõrvale veidralt vaenulik suhtumine lingvistikasse, isegi korpustesse, siis on Århusi sõnaraamatumudel hea just oma jõuliselt praktilise (rakendusliku, funktsionaalse) meelsuse poolest. Siin on palju, millele mõelda, millest eeskujuga võtta, peamiselt üldisest vaatest tänapäeva sõnaraamatutööle, sõnaraamatule ja selle kasutajale. Nüüdisaja e-sõnaraamat pole niivõrd sõnaraamat, kuivõrd (relatsiooniline) andmebaas, mitte niivõrd valmis, lõplik teos, kuivõrd dünaamiline, kasutaja vajaduste järgi kujutav võttev personaalne (monofunktsionaalne) teos, leksikograafide ja keeletehnoloogide koostöö. Olen ähmaselt unistanud mitmest võimalusest, mida siinses artiklis kirjeldasin.

Eesti Keele Instituudi sõnastikusüsteemis EELEX on koostatud juba ligi 40 sõnastikku. Ent süsteemisest ja sõnastikevahelist funktsionaalsust on siiani üsna vähe. Kasutajale oleme seni pakkunud enam-vähem sedasama, mis on raamatus, ilma rohkemale mõtlema. Sõnastikusüsteemi omaduste tõttu on paljud soovid/hädavajadused seni kättesaamatud: näiteks peaks püsiühendi põhisõnade määramisega hakkama saama lemmatiseerija, lingid peaksid tekkima ja uuenema süsteemi kaudu, infovahetus sõnastikubaaside vahel peaks käima palju hõlpsamini kui käsitsi kopeerida-kleepida jne. Tean küll, kui raske on ehitada uut süsteemi n-ö keskelt, kus põhijõud kulub ära muule: olemasoleva, tihti pelgalt masinloetaval kujul oleva keelevara süsteemikõlblikuks teisendamisele. Palju parem oleks alustada algusest. Aga tasapisi saabub aeg, kus vanad varad on teisendatud ja siis... Igal juhul on mitmed ideed juba õhus, millest mõndagi saaks rakendada juba praegu.

¹⁷ Macmillan English Dictionary Online (25. V 2012).

Kirjandus

- Atkins, B. T. Sue [1996] 2002. Bilingual Dictionaries: past, present, future. – Correard, Marie-Helene (toim), *Lexicography and Natural Language Processing. A Festschrift in Honour of B.T.S. Atkins*. Euralex 2002, lk 1–29.
- Atkins, B. T. Sue, Rundell, Michael 2008. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Béjoint, Henri [1994] 2000. *Modern Lexicography: An Introduction*. Oxford: Oxford University Press.
- EKSS 1961 = Eesti kirjakeele seletava sõnaraamatu prospekt. Eesti NSV TA Keele ja Kirjanduse Instituudi sõnaraamatute sektor. Tallinn, 185 lk.
- EKSS 2009 = Eesti keele seletav sõnaraamat I–VI. („Eesti kirjakeele seletussõnaraamatu” 2., täiendatud ja parandatud trükk.) Toim M. Langemets, M. Tiits, T. Valdre, L. Veskis, Ü. Viks, P. Voll. Eesti Keele Instituut. Tallinn: Eesti Keele Sihtasutus.
- EMS 1968 = Eesti murrete sõnaraamatu prospekt. Eesti NSV Teaduste Akadeemia. 207 lk.
- Fontenelle, Thierry 1997. *Turning Bilingual Dictionary into Lexical-Semantic Database*. Tübingen: Max Niemeyer Verlag.
- Fuertes-Olivera, Pedro, Bergholtz, Henning (toim) 2011. *e-Lexicography: The Internet, Digital Initiatives and Lexicography*. Great Britain: Continuum.
- Hausmann, Franz Josef, Reichmann, Oskar, Wiegand, Herbert Ernst, Zgusta, Ladislav 1989, 1990, 1991. *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie (HSK 5)*. Berlin–New York.
- Jürviste, Madis, Kallas, Jelena, Tuulik, Maria, Viks, Ülle 2011. *Extending the functions of the EELex dictionary writing system using the example of the Basic Estonian Dictionary. – Electronic Lexicography in the 21st Century New Applications for New Users: Proceedings of eLex 2011, Bled, 10–12 November 2011*. Ljubljana: Trojina, Institute for Applied Slovenian Studies, lk 106–112.
- Kilgarriff, Adam, Rychlý, Pavel, Smrž, Pavel, Tugwell, David 2004. *The Sketch Engine. – G. Williams, S. Vessier (toim), Proceedings of the XI Euralex International Congress*. Lorient: Université de Bretagne Sud, lk 105–116.
- Kilgarriff, Adam, Husák, Milos, McAdam, Katy, Rundell, Michael, Rychlý, Pavel 2008. *GDEX: Automatically finding good dictionary examples in a corpus. – E. Bernal, J. DeCesaris (toim), Proceedings of the XIII Euralex International Congress*. Barcelona: Universitat Pompeu Fabra, lk 425–432.
- Langemets, Margit 2002. *Sõnaraamat keeleteaduses: generatiivse leksikoni teooria. – Renate Pajusalu, Ilona Tragel, Tiit Hennoste, Haldur Õim (toim). Teoreetiline keeleteadus Eestis. (Tartu Ülikooli üldkeele õppetooli toimetised 4.) Tartu: Tartu Ülikooli Kirjastus, lk 146–163.*
- Langemets, Margit 2010. *Nimisõna süstemaatiline polüseemia eesti keeles ja selle esitus keelevaras*. Eesti Keele Instituut. Tallinn: Eesti Keele Sihtasutus.
- Langemets, Margit, Tiits, Mai, Valdre, Tiia, Voll, Piret 2010. *In spe: ühekõiteline eesti keele sõnaraamat. – Keel ja Kirjandus, nr 11, lk 793–810.*
- Lew, Robert 2011. *Online Dictionaries of English. – P. Fuertes-Olivera, H. Bergholtz (toim), e-Lexicography, lk 230–250.*
- Meer, Geert van der 2000. *Core, Subsense and the New Oxford Dictionary of*

- English (NODE). On How Meanings Hand together, and not Separately. – Proceedings of the Ninth EURALEX International Congress. EURALEX 2000. Vol. I–II. Institut für Machinelle Sprachverarbeitung, Universität Stuttgart, lk 419–431.
- Pajusalu, Renate 2009. Sõna ja tähendus. [Tallinn:] Eesti Keele Sihtasutus.
- Rundell, Michael 2012. 'It works in practice but will it work in theory?' The uneasy relationship between lexicography and matters theoretical. – Proceedings of the 15th EURALEX International Congress. EURALEX Oslo 2012. University of Oslo, Department of Linguistics and Scandinavian Studies, lk 47–92.
- Zgusta, Ladislav 1971. Manual of Lexicography. (Janua Linguarum. Series maior 39.) Prague: Academia / The Hague–Paris: Mouton.
- Tarp, Sven 2011. Lexicographical and Other e-Tools for Consultation Purposes: Towards the Individualization of Needs Satisfaction. – P. Fuertes-Olivera, H. Bergenholtz (toim), e-Lexicography, lk 54–70.
- Tavast, Arvi 2008. The translator is human too: a case for instrumentalism in multilingual specialised communication. (Dissertationes philologiae estonicae Universitatis Tartuensis 21.) Tartu: Tartu Ülikooli Kirjastus.
- Verlinde, Serge 2011. Modelling interactive reading, translating and writing assistants. – P. Fuertes-Olivera, H. Bergenholtz (toim), e-Lexicography, lk 275–286.
- Weinreich, Uriel 1964. Webster's Third: A Critique of its Semantics. – International Journal of American Linguistics, nr 4, kd 30, lk 405–409. [Uustrükk Weinreich, Uriel 1980. – On Semantics. Toim W. Labov, B. S. Weinreich. Philadelphia: University of Pennsylvania Press, lk 361–367.]
- Wiegand, Herbert Ernst 1977. Einige grundlegende semantisch-pragmatische Aspekte von Wörterbucheinträgen. Ein Beitrag zur praktischen Lexikologie. Kopenhagener Beiträge zur germanistischen Linguistik, nr 12, lk 59–149.
- Wiegand, Herbert Ernst 1984. On the Structure and Contents of a General Theory of Lexicography. – R.R.K. Hartmann (toim). LEXeter Proceedings. Papers from the International Conference on Lexicography at Exeter 9.–12. September 1983. Lexicographica. Series Maior 1. Tübingen, lk 13–30.
- Õim, Asta, Õim, Katre 2011. Eesti fraseoloogia leksikograafiline areng. – Keel ja Kirjandus, nr 11, lk 842–862.

On 'Linguistic Colonialism' in the Dictionary

Keywords: function theory of lexicography, word-based vs. concept-based database, dictionary as a tool, Estonian

Dictionary work is, in a sense, like a river with linguistics standing on one bank and the user on the other. The dictionary has been trying to keep hold of both, sometimes at the expense of one, sometimes of the other. My earlier discussions of the relations between lexicography and linguistics (Langemets 2002, 2010) have mainly been focused on linguistics. Although the present article also has its impetus in linguistics, its focus lies elsewhere, notably, on lexicography as a highly applied area, as well as on the user. According to Aarhus School such a turn-



around means quite a revolution in lexicography, bringing an essential change to dictionary work. Their enmity to linguistics and to theoretical lexicography as a strongly linguistically-based discipline is reflected in their neologism 'linguistic colonialism' to be opposed by their own 'function theory of lexicography' (Fuentes-Oliveira, Bergenholtz 2011). Aarhus School's approach to the dictionary is practical (applicational, instrumentalistic) to the bone and it is targeted, first and foremost, to new channels and to the users' new habits. They remind us that the first dictionaries were also very practical handbooks (like, e.g., the language aids that German missionary pastors used in Estonia). The theory relies on the potential user and the using situation and is focused on real e-lexicography, not on electronic versions of paper dictionaries.

The article presents a few key points characteristic of a functional dictionary model and compares them with the current situation at our Institute of the Estonian Language. Several of the options described have, somehow or other, been in the dreams of the author.

Margit Langemets (b. 1961), PhD, Institute of the Estonian Language, researcher, head of the Dictionary Department, margit@eki.ee

