

DISKURSUSEANALÜÜSI MUUTUV SUHE KVANTIFITSEERIMISEGA INGLIS- JA PRANTSUSKEELSES TRADITSIOONIS

MARGE KÄSPER, RAILI MARLING

Märksõna *diskursuseanalüüs* tekitab alati teatavat segadust: kas juttu on lingvistikast või ühiskonnateadustest, kas lähenemine on kriitiline või kirjeldav, on see meetod või laiem uurimissuund? Diskursusest kirjutatakse paljudes distsipliinides ning see toob endaga kaasa mõiste tähendusvarjundite arvu suurenemise. Siinne artikkel ei plaani kehtestada ühest tähendust, kuid soovib avada diskursuseanalüüsi (järgnevalt DA) laiemat arengukonteksti, näidates erinevaid arusaamu sellest, kuidas suhestada omavahel uuritav keelematerjal ja seda ümbritsev kontekst ning mida ja kuidas diskursuses uurida. Kuna inglisis- ja prantsuskeelses maailmas on DA arenguloogika olnud põnevalt eripalgeline, on artikli eesmärk vaadelda neid arenguid suhtumises keelematerjali analüüsiviisidesse kahes akadeemilises keeleruumis.

Kaasates keeleanalüüsi sotsiaalteaduslikku või ka filosoofilist ideestikku, on DA tänapäeval nii neis kahes keeleruumis kui ka Eestis osa laiast ja mitmekesisest uurimisväljast, kus põimuvad keeleteadus ja ühiskonnateadused. Diskursuses väljenduvaid või varju jäävaid tõekspidamisi ja ühiskonda iseloomustavaid suhtumisi uurivad sotsiaalteadlased (DA rakendamisest sotsiaalteadustes vt e k Kalmus 2015; akadeemilisi rakendusi eesti ühiskonna analüüsis vt nt Kiisel 2013). Teisalt, kuna diskursus materialiseerub keeles ja tekstides, on DA lähedane tekstianalüüsiga. Tartu Ülikooli ja Tallinna Ülikooli keeleteadlaste tekstirühmade tegevust vedanud Reet Kasiku (2002: 7) sõnutsi vaatlleb DA pigem sotsiaalseid struktuure ja väärtusi, tekstianalüüs aga materiaalsemat ja konkreetsemat, kindlat vormi omavat teksti. Selle eristuse järgi on need teadlased ise rõhutatult tegelenud pigem teksti- kui diskursuseanalüüsiga („Artikleid tekstianalüüsist” (2002), „Lingvistiline tekstianalüüs” (2004) jt Kasiku toimetatud kogumikud); analüüsides fookuses on olnud tekstide funktsioonid ja tüpoloogilised tunnused (Kasik 2002). Samas on vaatluse all olnud ka erinevate tekstitüüpide ühiskondlik mõju ning läbi tekstiprisma väljenduvad väärtushinnangud (nt Aava 2004; Sarapuu 2004, 2005). Nii seab rõhuasetus ühiskondlikule kontekstile ka keeleteadusliku tekstiuurimise DA raamidesse, teisalt aga eristub DA tavapärastest ühiskonnateadustest just keelematerjalipõhisuse rõhutamisega.

Huvi keelekasutuses materialiseeruvate peidetud tähenduste vastu pärineb DA teoreetilises raamistikus kõige enam Lääne-Euroopa XX sajandi marksislike filosoofide keelekäsitustest (Louis Althusser, Antonio Gramsci). Laiemas plaanis huvituvad keelekasutusest kui reaalsest keelematerjalist (vastu-kaaluks abstraktsele keelesüsteemile) aga ka sotsiolingvistika, retoorika,

diskursusepsühholoogia, vestlusanalüüsi jm keeleteaduse harusuunad. Kõigil neil taustadel ja eristustel me käesolevas artiklis mõistagi peatuda ei jõua. Selge on aga see, et arusaamad diskursuse mõistest, nagu ka selle rakendused eri distsipliinides, võivad olla üsnagi erinevad (vt Mills 2004: 1). DA puhul pole tegemist ei ühtse meetodi ega üheselt piiritletud uurimisvaldkonnaga. DA-d on pigem määratletud kui paljusid valdkondi hõlmavat metodoloogilist distsipliini (Maingueneau 2012 jm). Et terminoloogiliste vaidluste lahkamine liiga pikale ei veniks, sedastagem sissejuhatuseks lihtsalt, et DA uurimisobjektiks on mis tahes vormi või sisuga ühiskonnas ringlevad autentsete tekstid, mis pole pelgalt uurimiseks loodud, ning nende tõlgendus kontekst. Rõhutades ühelt poolt reaalse keelekasutuse põhisust, huvitab DA-d teisalt, kuidas avaldub selles keelematerjalis ühiskondlik kontekst. Tekstidele rakendatavad tõlgendusviisid selle uurimiseks võivad aga vägagi erineda.

Ilmselt on just tõlgendamisega seotud väljakutse, sellega kaasnevad „riskid” ja vajadus analüüsiviise selgitada üks põhjusi, miks sel empiirikaga tegeleva metodoloogilisel väljal on niivõrd palju n-ö teooriat toodetud. Praegustes keeleteaduse debattides suurte andmemahutude ning keele- ja tekstikorpuse automaatselt analüüsivate andmetöötlusprogrammide üle väärub aga tähelepanu asjaolu, et diskursuseanalüüs, alati pigem keeleteaduse piirialadele asetuv, kuid samuti süvitsi keelega tegelev uurimisvaldkond, on analüüsi automatiseerimise ja võimalikult suure hulga andmete hõlmamise võimalusi vaagunud juba aastakümneid. Need püüdlused on esile toonud vägagi erinevaid arusaamu sellest, mida ja kuidas diskursuses uurida, ja see omakorda selgitab, miks võib diskursuse mõiste tänapäeval esineda niivõrd erinevates kasutustes. Siinse artikli eesmärgiks ei ole kindlasti anda ülevaadet kõigist sel uurimisväljal aset leidnud debattidest. Jälgides eri koolkondade ja analüüsimetodite suhet kvantitatiivsuse ja kvalitatiivsusega, võtame pigem eesmärgiks kaardistada neid eriarvamusi ja arenguid.

Kuivõrd DA eesmärgiks on uurida tekste nende suhtes sotsiaalse kontekstiga ehk „paigutumisega” (Maingueneau 2012 jm), on omal kohal ka küsimus sellest, kuhu on paigutunud DA eri uurimissuundi kujundanud tekstid. Et DA on tõlgenduslik, on see valdkond ka teistest keeleteaduse harudest kultuuriliselt varieeruvam. Artiklis püüame niisiis dialoogi panna DA arenguid inglise- ja prantsuskeelses akadeemilises ruumis, kirjeldades, milliseid võimalusi on see metodoloogiline väli neis pakkunud empiiriliste materjalide tõlgendamiseks ning millised on olnud arutelud alates selle tekkest formalismi laineharjal, arvutiteaduse algajastul, kuni tänaste mõttevahetusteni suurte andmehulkade töötlusvõimaluste taustal. Kuna ingliskeelsed allikad on meie teaduses kättesaadavamad ja tuttavamad olnud, avame enam just prantsuse keeleruumis toimunut.

Järgnevalt vaatleme, kuidas DA ajaloolist arengut on mõjutanud laiemad humanitaarteaduslikud intellektuaalsed otsingud. Mõneti võib sellegi uurimisvälja juured lingvistikas tagasi viia Ferdinand de Saussure'i klassikaliseks kujunenud keeleteooriani ning sellest välja kasvanud strukturalistlike ja poststrukturealistlike lähenemisteni keelele ja kultuurile (vt e k Saussure 2017 [1916], aga ka vähemkanoonilist tekstikogu pr k Saussure 2002).¹ Teisalt

¹ Ligi sajand pärast seda, kui 1916. aastal ilmus Ferdinand de Saussure'i õpilaste koostatud „Üldkeeleteaduse kursus” (e k 2017), on prantsuse keeleruumis suurt tähelepanu pälvinud esialgselt süstematiseerivast käsitlestest kõrvale jäänud käsikirjade põhjal

on DA sünd seotud just kriitikaga saussure'liku keeleteaduse kohta – artiklis käsitleme Michel Pêcheux' radikaalselt erinevat arusaama keeleteaduse ülesannetest ja uurimisviisidest. Otsapidi põhineb DA sealjuures formalistlikul keeleteadusel (Zellig Harris) ja strukturalismil, teisalt kujundavad seda pragmaatika ja prantsuse kontekstis lausungilingvistika (Émile Benveniste). Arvutiteaduse ja loomuliku keele töötlemise algaastail kasutasid uurijad Harrise ideid esimeste keeleanalüüsi programmide väljatöötamiseks. 1970. aastate „keeleline pööre” tõi aga diskursuseanalüüsi kvalitatiivsete meetodite esirinda, tingides sotsiaal- ja humanitaarteadustes üsna üleüldise „diskursiivse pöörde” (Hall 1997: 6). Selle arengukaare jooksul on DA pidevalt maadelnud küsimusega, kuidas lahendada paratamatult tekkivat pinget kvantifitseerimise ja formaliseerimise ning keelematerjali kvalitatiivse tõlgendamise vahel.

1. Keeleanalüüsi formaliseerimine kui võti ja eesmärk

Formaliseerimise vaatenurgast saab DA jaoks kõik alguse sellest, et arvutid teevad võimalikuks tekstide ja lausungite n-ö üksteise peale asetamise, lineaarsust lõhkva lugemise. DA mitmenäolisus kujuneb paljuski sellest, mida ja kuidas selle esmase kvantitatiivsuse kaudu uurima asutakse.

1.1. Formaliseerimise esimesed sammud Ameerika Ühendriikides

Mõistet *diskursuseanalüüs* hakkas metoodiliselt kasutama ameerika lingvist Zellig Harris (1952), keda seostatakse üldiselt transformatsioonigrammatika arendamisega. Harrist huvitas keele uurimine lausest kõrgemal tasandil ning keelelise ja mittekeelelise käitumise vaheline suhe. Ta uuris detailsemalt just esimest, püüdes kirjeldada keelevormide jaotumist tekstides ning eriti seda, kuidas need eri tekstitüüpides omavahel kombineeruvad. Tema eesmärgiks oli konstrueerida tekstigrammatika ehk formaliseerida loomulikus keeles esinevate tekstide struktuur. Samas tegi Harris ka suhteliselt varakult järgmise tähelepaneku:

Seostatud diskursus eksisteerib mingis konkreetsetes situatsioonis – olgu siis jutt inimesest, kes räägib, või vestlusest või kellestki, kes iganes pikema perioodi vältel maha istub ja mingit konkreetset kirjanduslikku või teaduslikku traditsiooni järgides konkreetset tüüpi raamatut kirjutab. (Harris 1952: 3)

Teisisõnu, konkreetsetele situatsioonidele on omased teatud tüüpilised keelekasutuse viisid. Harrist kui keeleteadlast huvitasidki sellised iseloomulikud keelelised tunnused. See osa Harrise tööst ei leidnud Ameerika Ühendriikides kohe laia kõlapinda, kuid kerkis esile siis, kui Harrise mõtted järgmisel kümnendil marksismimeelsele Prantsusmaale jõudsid (vt allpool). Esialgu oli aga lingvistikas algamas Harrise õpilase Noam Chomsky nimega seostatud

koostatud tekstikogu „Kirjutusi üldkeeleteadusest”, kus muuhulgas nähtuvad ka Saussure'i otsingud selles suunas, kuidas keel kõnelemistegevuses diskursusena toimib (Saussure 2002: 95, 277).

revolutsioon, mis juhtis paljud keeleteadlased formaliseeriva generatiivgrammatika juurde. Ehk on kõnekas tõik, et täna kriitilise DA esindajana tuntud hollandi keeleteadlane Teun van Dijk alustas oma teadlasekarjääri, proovides luua generatiivpoetikat ja generatiivset tekstigrammatikat (Dijk 1972).

Tekstigrammatikad olidki üks Harrise tööst inspiratsiooni ammutanud arenguid ameerika keeleteaduses. Harrise ideedele toetudes leidis näiteks hiljem tema juures õppinud Ameerika Piibliühingu keeleteadlane James A. Lorient (1957) ühe põlisameerika keele tekstide diskursusereeglid, mis polnud otseselt seotud lausestruktuuriga. Keeleteaduslikesse ajakirjadesse jõudis tema töö alles 1970. aastal (Lorient, Hollenbach 1970). Teine Harrise õpilane, Robert Longacre, avaldas 1983. aastal raamatu diskursuse grammatikast, mis oli pühendatud just Lorient'le kui diskursuseuuringute teerajajale. Longacre püüdis leida „pinnagrammatika” all peituvat tähenduslikku grammatikat ning arendas välja süsteemse meetodi diskursuse uurimiseks (vt ka Longacre 2002).

Samas kannustas Harrise analüüs varast uurimistööd loomuliku keele töötlemise vallas, mis oli põhjapaneva tähendusega arvutilingvistikale. Harrise huvi vormi ja tähenduse vastu andis tõuke arvutipõhisele loomuliku keele uurimisele, sh Naomi Sageri tööle meditsiinitekstidega juba 1960. aastatel: nt arendas Sager algoritmi süntaktilise mitmetähenduslikkuse lahendamiseks (1967) ning kujundas inglise keele arvutigrammatika (ingl *computer grammar*) (1981). Inglisekeelse DA arengus need formaliseerivad lähenemised kesket rolli ei mänginud, küll aga juhatavad need arengusuunad sisse diskursuse mõiste rakendamise tähenduses, mis märgib loomulikku, seotud lausungitest koosnevat keelekasutust.

1.2. Subjektsus ja sõnavarastruktuurid Prantsusmaal

Prantsuse keeles ei puudu ka puht tekstitasandile keskenduvad diskursusekäsitused, mis lähtuvad Harrise lausetevahelise sidususe uurimise meetodist (Charolles, Combettes 1999), kuid „meetod võlus analüüsijaid seda enam, et selle tulemused polnud „determineeritud”, ning võimaldasid seega interpretatsioone ja ka kõrvalekalduvaid kasutusi” (Marandin 1979). Harrise 1952. aastal inglise keeles ilmunud artikkel avaldati prantsuse keeles ajakirjas Langages 1969. aastal, kontekstis, kus alles äsja oli esile kerkinud kõnelemistegevuses subjektsust uuriv lausungiteooria (Benveniste 1966) ning kus samal 1969. aastal ilmus trükist veel teisigi DA sotsiaalsete tõlgenduste seisukohast olulisi, kuid hoopis teiselaadilisi diskursusekäsitusi (neist allpool).

Harrise maaletoojaks oli leksikograaf Jean Dubois, mitmete sõnaraamatute ja strukturalistlike grammatikakäsitluste koostaja, kes ise oli ühtlasi Langages'i, juba tollal väga mõjuka ajakirja peatoimetaja ning ka antud numbriga koostaja (Dubois, Sumpf 1969). Nimetus *analyse du discours* oli pealkirjaks nii Harrise artikli tõlkele kui ka kogu numbrile, mis hõlmas üsna eripalgelisi tekste psühhoanalüüsist (Luce Irigaray „L'énoncé en analyse”) tekstitüpoloogiani (Serge Méleuci „Structure de la maxime”). Kontekstis, kus suurem osa lingvistikast tegeles keele formaalsete tunnuste kirjeldamisega, legitimeeris number sümbolsest DA kui lähenemise, kus uued ideed kokku

said, näidates, et „tekst ei räägi mitte ainult maailmast, vaid ka kõnelevast subjektist” (Dubois 1969: 110). Neist tõlgendusist saab alguse üks prantsuse DA-le iseloomulikumaid jooni (Maingueneau 2016): lausujapositsiooni määratlemine diskursuses leiduvate markerite kaudu.

Dubois'd ennast huvitas eelkõige meetodis sisalduv potentsiaal arendada selle abil eri sotsiaalsete gruppide sõnavarakasutuste uuringuid (Matoré 1953; Wagner 1967–1970; Dubois 1962). Sotsiaalsete gruppide või meelsuste iseloomulike sõnavarastruktuuride automaatse tuvastamise ehk diskursuses nende leksikat mõötva leksikomeetria peamiseks arendajaks sai Maurice Tournier' juhtimisel loodud Saint-Cloud' poliitilise leksikoloogia uurimiskeskus ning hiljem ajakiri MOTS (akronüüm, mis kokku tähendab otsetõlkes 'sõnad', kuid mille moodustanud prantsuskeelsed komponendid olid *Mots*, *Ordinateurs*, *Textes*, *Sociétés* – Sõnad, Arvutid, Tekstid, Ühiskonnad). Sõnakasutuse regulaarsuse väljatoomiseks loodud sõnavarastatistika programmide abil kirjeldati Alžeeria sõjast rääkivat ajakirjandust (Maldidier 1969), ametiühingutegelaste sõnavõtte (Marcellesi 1971) jpm, määratledes kõnelejate lausumispositsioone peamiselt valitud märksõnade kasutuserinevuste kaudu (temaatiliselt eristuv parem- ja vasakpoolne sõnavara, töölisliikumist pooldav või eksploateeriv sõnakasutus jms). Lähtumine Harris'e tekstimeetodist oli tegelikult suhteliselt tinglik: tekstide formaalseid tunnuseid ja lausetasandil sidusat kombineerumist kõrvale jättes võeti kasutusele eelkõige idee, et on olemas teatavad semantilised mustrid, kuidas üht mõtet võib väljendada eri grammatiliste ja leksikaalsete konstruktsioonide kaudu. Kui Harris (1952, 1969) osutas, et nt konstruktsioonid *oktoobrikuu lõpus* ja *sügise keskpaiku* või *läheb külmaks* ja *hakatakse kütma* väljendavad funktsionaalselt sama mõtet (neid võib semantiliselt võrdsustada ja analoogselt teistele tähendusklassidele vastandada), siis poliitilise leksikoloogia kontekstis rakendati seda loogikat tekstides avalduvate lausujapositsioonide uurimisel. Leksikomeetrilises faktoranalüüsis iseloomustab kõneleja meelsust esilduvate märksõnade kombinatsioon, isikuliste asesõnade üle- või alakasutus või ka eelistatav verbikonstruktsioon.

Seega õnnestus juba üsna varasel arvutiajastul masinaile selgeks õpetada omavahel seotud märksõnade tuvastamine, et kasutussageduste erinevuste järgi kaardistada kontrastseid vastandusi. Hilisem kriitika tollaste uurimuste aadressil on see, et kuna võrreldavuse huvides valiti võimalikult representatiivsed uurimiskorpused, siis õigupoolest otsiti uurija määratletud märksõnu uurija määratletud vastanduse näitamiseks – seega leiti seda, mida otsiti (vt Guilhaumou 2002). Seepärast on paljud uurijad tänini Prantsusmaal leksikomeetriliste meetodite suhtes umbusklikud ning tähelepanu pööratakse pigem lausumistingimuste kontrastsusele: kuidas muutub lausumispositsioon vastavalt kõneformaadile (avalik kõne vs. *teleshov*), publikule, teemale, sündmuste kulgemisele ajas jne (vt Charaudeau 2009). „Suletud”, mingi määratud positsiooni suhtes eeldatavasti representatiivsetelt korpustelt on seega liigunud üha enam „avatud” ja võimalikult variatiivsete uurimiskorpuste suunas. Leksikomeetrilisi meetodeid pole samal ajal siiski hüljatud, nende täiendamisest ja arengust räägime allpool.

1.3. Tähenduse süntaks ja unistus automaatsest ideoloogia-tuvastusest

Hoopis teistlaadse keelematerjali analüüsi formaliseerimise programmi pakub Prantsusmaal samal 1969. aastal välja peamiselt ideoloogiateoreetikuna tunnuks saanud Michel Pêcheux pealtnäha väga formaliseeriva pealkirjaga teoses „Automaatne diskursuseanalüüs”. Kui Dubois’ kaudu Harrise meetodiga seotud vastandlike sõnavarastruktuuride uuringud võib siiski paigutada üsna traditsioonilise strukturalistliku keeleteaduse alla, siis Pêcheux’ ümber koonduvad uurimisgrupp vastandas end sellele täielikult, heites tollasele lingvistikale ette, et see oli „unustanud” arendada teooriat, mis kirjeldaks tähenduse loomist keele kasutamisel, „sest kõnelemine on hoopis midagi muud kui grammatikanäite produtseerimine” (Haroche jt 1971: 98).

Marksistliku filosoofia ja sotsiaalpsühholoogia taustaga Pêcheux’ jaoks oli tekstide tähendus eelkõige sotsiaalne ja täpsemalt ideoloogiline, kuid keeleteaduselt ootas ta tänapäevases mõttes õigupoolest interdistsiplinaarset dialoogi tekstide teistsuguseks, süvastruktuurides peituvat varjatud tähendust välja toovaks lugemiseks. Tundes hästi tollaseid võimalikke tekstitöötlusmeetodeid – sotsiaalpsühholoogia oli sel hetkel valdkond, kus väga edumeelselt oli arendatud sisuanalüüsipõhist automaatset dokumentitöötlust –, ei rahuldanud teda andmete lahterdamine etteantud teemaklassidesse. Ta soovis leida analüüsimumoodust, kus tekstidest leitavad positsioonid oleksid analüüsi tulemus, mitte küsitav kategooria.

Harrise meetodiga oli Pêcheux mingil määral kursis, selles „oli midagi”, mis võis aidata leida keskteed ühelt poolt intuiitiivse hermeneutika ja teiselt poolt positivistliku leksikomeetria vahel (Léon 2010), kuid teda huvitasid eelkõige tekstide lineaarse pealispinna all peituvad eelkonstrueeritud tähendused, mis nähtusid tema arvates mõningates süntaktilistes konstruksioonides, nagu iseloomustavad kõrvallused, nominalisatsioon vms, mida ta lahkas lähemalt oma järgnevas, samuti pigem (marksistlikus) keelefilosoofilises teoses „Les vérités de la Palice” (ilmselged tõed, 1975, ingl k 1982). Nii oli „Automaatne diskursuseanalüüs” küll inspireeritud lennukast soovist luua „automaatne ideoloogia paljastamise masin”, mis kasutaks andmetöötluse võimalusi ühiskonda reguleerivate ideoloogiliste lausungite nähtavaks tegemiseks, kuid automatismide tuvastamise käsiraamatu asemel kujutas teos endast pigem omamoodi teoreetilist arutelu keelemehhanismide üle. „Masin” ise jäigi lõpuks loomata, sest üha enam hakkas Pêcheux’ d paeluma loomuliku keele keerulisus ja isegi agrammatilisus (Gadet, Pêcheux 1981). Pêcheux’ lahkumise järel tema uurimisgrupp küll lagunes, kuid küsimus tähenduse mitmekihilisusest ja selle väljendumisest keekekasutuses jäi. Nii viis unistus automatiseerimisest arusaamiseni, et keele mitmetähenduslikkus vajab kvalitatiivsemat lähenemist.

2. Sotsiaalne tõlgendusruum kui eesmärk

Formaliseerimispuüdlustele vastandudes on teised suunad rõhutanud pigem tekstide sotsiaalse tõlgendamise keerukust ja kompleksust. Kuna selliste diskursusekäsituste fookuses on tekstide sotsiaalsus, tekib nii paljudes keelt

vaagivates käsitlustes kui ka keeleteaduses eneses arutelu selle üle, kuivõrd ja kuidas ühiskondliku konteksti uurimises keelematerjali kasutada ning mis rolli selles analüüsis mängib lingvistiline aparatuur.

2.1. Foucault ja keeleline pööre?

Ingliskeelses maailmas ja meilgi seonduvad diskursuseanalüüs ja nii keeleline kui ka laiem diskursiivne pööre² paljuski prantsuse filosoofi Michel Foucault' nimega, kelle töödest juhinduvad just radikaalselt kvalitatiivsed ja kriitilised diskursuseanalüüsijad. Foucault'lt ilmub – täiesti juhtumisi – samuti 1969. aastal teos „Teadmiste arheoloogia” (2005 [1969]), kus ta oma diskursuseuuri- mise meetodit selgitab, kuid prantsusekeelse DA välja kirjeldava Dominique Maingueneau (2001) sõnutsi oleks „selge eksitus” pidada Foucault'd DA loo- ja k s, kuna tema ambitsioonid selles osas piirdusid vaid „ironiaga, millest ei maksa end petta lasta”.

Foucault' panust diskursuse uurimisse võikski kõige paremini iseloomus- tada märksõnaga *problematiseerimine*, mille all ta mõistab „diskursiivsete või mittediskursiivsete praktikate hulka, mis asetab millegi tõeluse ja vale vahelisse mänguruumi, tehes sellest mõtlemise objekti [---]” (Foucault 1994 [1984]: 670). Ühiskonnas tavapärastena funktsioneerivate mõistete ja suhete problematiseerimisel liigub Foucault ise oma hilisemates teostes pigem sel- liste ruumiliste jm mittediskursiivsete praktikate analüüsimise poole, mis võimaldavad uurida võimu avaldumist. Keeleuurijate jaoks on olulisemad aga tema varasemad teosed, mis tegelevad rohkem tekstitasandil nähtuvate ilmingutega. „Teadmiste arheoloogia” ning sellele eelneva „Sõnade ja asjade” (2015 [1966]) ilmumise ajal tunneb Foucault suurt huvi grammatika, lingvis- tika ja formalismi ning teisalt keelefilosoofia vastu (Revel 2002: 23; Eribon 1991: 204). „Seniolematu ja tähtis oli öelda, et see, mis oli keele kaudu tehtud – poeesia, kirjandus, filosoofia, üldse diskursus –, allub mingitele sisemistele seaduspärastele ja regulaarsetele: keelekasutuse seaduspärastele ja regulaarsetele. Keele kaudu tehtu lingvistiline olemus oli oluline avastus.” (Foucault 1994 [1973]: 539) Samas kirjeldabki ta keelelist analüüsi „Tead- miste arheoloogias” kui selgelt strukturalistlikku heuristilise valimi vaatlust, mida võimaldab suletud korpus:

Muidugi saab mingit lingvistilist süsteemi (kui seda just kunstlikult ei konst- rueerita) rajada üksnes lausungi korpuse või diskursusefaktide kogumi põh- jal, aga siis tuleb meil selle näitliku väärtusega koosluse alusel määratleda reeglid, mis lubavad moodustada ka teistsuguseid võimalikke lausungeid [---]. (Foucault 2005 [1969]: 30)

Diskursuse tasand on see võimalike ja võimatute variatsioonide düna- mika. Kõige üldisemalt öeldes on diskursus Foucault' jaoks kogum lausun- gitest, mis võivad kuuluda täiesti eri valda, kuid mis samas alluvad ühistele

² „Pöördeliste aegade” kohta humanitaarteaduste metodoloogias laiemalt vt Piirimäe 2011 ja teisi artikleid sellesamas Marek Tamme koostatud kogumikus. Kogumiku aluseks on Keele ja Kirjanduse teemanumber „Humanitaarteadused pöördes” (8–9/2008).

toimimisreeglitele, nii et läbivatest kordustest kujunevad diskursuses käibel olevad mõisted, läbivatest mittemainimistest aga ka tähendusrikkad lüngad.

See, millistest korduvustest ja sarnasustest need reeglid moodustuvad, on Foucault' jaoks üsna intuiitiivne. Tema analüüsid põhinevad tema enda erudeeritud lugemusel lausungite „arhiivist”, mis pole „mingi tsivilisatsiooni säilitatud tekstide koguhulk ega ka kogum selle hukust päästetud jälgedest, vaid reeglite kooslus, mis määrab ühes kultuuris lausungite ilmumise ja kadumise, nende püsimise ja kustumise, nende paradoksaalse eksisteerimise *sündmuste ja objektidena*” (Foucault 1994 [1968]: 708). Sellega seab Foucault õigupoolest kahtluse alla keeleteadusliku strukturalistliku uuritava valimi ehk *uurimiskorpuse* mõiste, asendades selle dünaamilise *arhiivi* mõistega. Kuna selle arhiivi diskursuste dünaamika moodustus Foucault' järgi mitte ainult keeleliste seaduspärade, vaid ka keeleväliste korduvuste ja seaduspärade kooslusest, siis sellise valimi automaatne moodustamine olekski esiteks olnud väga keeruline, teiseks ei olnud see ka problematiseerimisele orienteeritud Foucault' eesmärk.

2.2. Kvalitatiivne dominant ning kriitilised diskursuseuuringud

Inglisekeelses keeleteaduses oli diskursuseuuringute jaoks oluline M. A. K. Halliday töö 1960.–1980. aastatel, mis vaidlustas mõned Chomsky generatiivlingvistika põhipostulaadid. Nagu Pêcheux' (vt 1.3), nii ka Halliday meelest ei saa keelt samastada ainult grammatiliste lausetega. Halliday leiab samuti, et ainuüksi formaallogikast lähtumine ei ole keele uurimisel piisav (vt nt Halliday 1978). Halliday meelest peaks fookuses olema konkreetne keelekasutus ja keel kui tervik. J. R. Firthist mõjutatuna vaatleb Halliday (2002) keelt kui toimivat süsteemi. Idee polnud iseenesest uus, sest selle oli juba varem välja pakkunud Saussure. Sellest vaatenurgast on keel omavahel seotud tähenduse loomise võimaluste ehk süsteemide võrgustik. Halliday lähenemine keskendub valikutele, mida keelekasutajad keelesüsteemide sees teevad. Suhtluse aluseks on funktsioon. Halliday keeleteaduslikku lähenemist tuntaksegi süsteem-funktsionaalse lingvistika nime all (Halliday 1994). Huvi keele funktsiooni vastu sunnib ta kaasama analüüsi ka ühiskondlikku konteksti:

Kui me ütleme, et lingvistiline struktuur „peegeldab” sotsiaalset struktuuri, siis omistame keelele liiga passiivse rolli [---]. Pigem peaksime ütleva, et lingvistiline struktuur on üks sotsiaalse struktuuri realiseeringuid, mis seda vastastikuse loovuse protsessis aktiivselt sümboliseerib. Kuna keel on üks ühiskonna metafoore, siis ta mitte ainult ei anna ühiskondlikku korda edasi, vaid samuti kindlustab ja potentsiaalselt mugandab seda. (Vaieldamatult just see aitab seletada neid vägivaldseid suhtumisi, mida teatud ühiskondlikes oludes üks grupp rakendab teiste kõne suhtes.) (Halliday 2007: 255)

Samas on ühiskonna rolli rõhutatud ka ühiskonna analüüsis mitte tegelevas keeleteaduslikus traditsioonis. Näiteks Gillian Browne ja George Yule'i (1983: 1) definitsiooni järgi on „[d]iskursuse analüüs [---] paratamatult kasutuses oleva keele analüüs. Sellisena ei saa see piirduda keele-

vormide kirjeldusega, kaasamata eesmäärke ja funktsioone, milleks need vormid on inimtegevuses mõeldud” [meie sõrendus – *M. K, R. M*].

Erinevus on selles, mis roll ühiskondlikule kontekstile antakse ning kui tugevasti see seotakse ideoloogiatega. Halliday analüüs pole teravdatult poliitiline, tema eesmärgiks on keeleteooria arendamine. Kuid et tema keeleteooria on eelkõige keele sotsiaalse toimimise teooria (Jones 2010) ja lähtub ajastule omaselt marksistlikust ideeraamist, langevad tema mõtted paljuski kokku Prantsusmaal keeekasutuse ideoloogilisust postuleerinud ja nii klassikalise keeleteaduse vastu välja astunud Pêcheux’ ideedega (vt 1.3). Ka Pêcheux’ „marksistlik semantika” (Maingueneau 1996: 30) kirjeldas lausungeid kui võimalikke valikuid, mis sõltuvad lausujate sotsiaalsetest jm positsioonidest, ja nägi keeekasutuses ühiskondlikke struktuure taastootvat mehhanismi. Kuid otsides keeekasutusest eelkõige sõlmpunkte, kus nähtuksid diskursuses ideoloogiliselt eelkonstrueeritud tähendused, ei jõudnud Pêcheux väga süsteemse käsituseni keelest kui terviklikust masinavärgist. Keeleteadusest lähtuva Halliday kriitilisest keeekäsitusest arenes ingliskeelses traditsioonis aga terve kriitilise lingvistika koolkond (nt Fowler 1991; Hodge, Kress 1993).

Halliday tööd on olnud olulised ka Reet Kasiku ja Krista Kerge alustatud Eesti tekstianalüüsi koolkonna väljakujunemises (Kasik 2003). Ingliskeelses traditsioonis on Halliday mõjutanud kriitilise lingvistika, kriitilise diskursuseanalüüsi või kriitiliste diskursuseuuringute nime all esinevat, 1980. aastatel alguse saanud lähenemist, kus diskursust vaadeldakse mitte ainult keelena tegevuses, vaid tegevusena, mis on aktiivselt kaasatud ühiskonna kujundamisse. Seda väljendab kujukalt briti kriitilise DA klassiku Norman Fairclough’ idee, et „[d]iskursus on [---] enam kui lihtsalt kasutuses olev keel: see on keeekasutus, olgu siis kõnes või kirjas, mida käsitletakse sotsiaalse praktikana” (Fairclough 1992: 28). Teine kriitilise keeueuurimise hälli juures olnud briti lingvist Roger Fowler (1991: 94) rõhutab, et „keel annab kategooriatele nimed ja nii loob nende piirid ja suhted ning diskursus võimaldab neid nimesid sageli lausuda ja kirjutada, tehes nii need kategooriad reaalseiks ja tuttavaiks”. Fairclough’ jaoks on diskursuse ja ideoloogia suhe otsene: „Diskursus loob ühiskondlikku. [---] Diskursust vormivad võimusuhted ja sellesse on kätketud ideoloogiad.” (Fairclough 1992: 8)

Kriitilises diskursuseanalüüsis vaadeldakse, nagu viimane tsitaat näitab, mitte ainult keelt, vaid ka keele taga olevaid võimustruktuure ning ideoloogiaid selle sõna kõige üldisemas tähenduses. Fairclough (1995: 73) on keelt otsesõnu nimetanud ideoloogia materiaalseks vormiks. Kuna Pêcheux’ ideoloogiat uurivad keeekäsitlused jäid ingliskeelsele maailmale pikalt tundmatuks, siis suhestub kriitiline diskursuseanalüüs keeleteaduse kõrval laiemast kriitilisest teooriast pigem Foucault’ diskursusekäsitlustega, mis ingliskeelses maailmas kiiremini levisid. Kriitilised lingvistid on püüdnud leida nende ideede rakendamiseks keeleteaduslikumat metodoloogiat, et ideoloogia loomulikustatud toimimist nähtavaks teha. Distsipliini arenguloogika seisukohast on aga märkimisväärne, et kui inglise keeles tekkis laiemal formalistlikul keeleteaduslikul väljal vajadus vastandava sildi „kriitiline” järele, polnud prantsuskeelses traditsioonis põhjust DA-d „kriitiliseks” nimetada, kuna see tõukus nagunii juba eos sotsiaalsest kontekstist.

Et diskursuste varjatud ideoloogilise palge paljastamisel tuleb alati väga mitmeti ning täpselt analüüsida nii tekstuaalset kui ka ühiskondlikku konteksti, siis on kriitiline DA konkreetset keelematerjali analüüsidest rõhutatult kvalitatiivne. Sellest ka skepsis katsete suhtes vaadelda tekstimaterjali kvantitatiivselt: suurtes korpustes läheb kaduma ühiskondlik raamistus, milles tekst ilmneb. Nt on Fairclough (1989: 20) irooniliselt märkinud, et korpuslingvistika pole analüüs, vaid tööriist, mida saab analüüsis kasutada. Nagu eespool nähtud, Pêcheux' marksistlik semantika sellise tööriista kasutamist iseenesest ei välistanud, seades vastupidi isegi eesmärgiks selle konstrueerimise, et edukamalt tekstimassidest ideoloogiat püüda; loomuliku keele keerukus oli aga see, mis ta tagasi kvalitatiivsete küsimuste juurde viis.

3. Vastanduselt komplementaarsusele

Tekstide sotsiaalse paigutumise uurimisel on keelekasutuse analüüsi automatiseerimise katsetajad seega püüdnud lähtuda hoomatavatest keelekategooriast (sõnavaraühikud, määratletud grammatilised konstruktsioonid); analüüsi kompleksuse rõhutajad on soovinud hõlmata laiemat pilti, otsides diskursusemustrites vm iseloomulike joontena kirjeldatavates kasutustes tegelikult samuti korduvusi, kuid eelistades enam süüvida materjali iseloomu, mõistmaks, mida oluliselt pidada. Kvalitatiivse lähenemise kasutajad on seejuures eri analüüsisisendite kombineerimise vajadust rõhutanud, et nii tõlgendusi testida ja süsteemsust leida (vt nt Claudel jt 2013). Teisalt osutavad juba esimeste leksikomeetriliste analüüside esitajad, et nende „kirjeldatud meetod pole mõistagi tõendamise, vaid avastamise” vahend (Marcellesi 1971: 33), sest ka statistiliste andmete hulgast tuleb ära tunda olulisemad ja kõige kõnekamalt tekstivalimit kirjeldavad faktorid. Nende määratlemist seletades rõhutab André Salem (1982: 155), et väär praktika oleks korpuse kohandamine kõnekamate statistiliste tulemuste saamiseks, see ei vii leksikomeetrist uurimust sugugi edasi. Tänapäeval rõhutab enamik meetodikäsitusi, et kvalitatiivne ja kvantitatiivne uurimisviis annavad küll erisugust informatsiooni, kuid toimivad kõige paremini teineteist täiendades, seda nii uurimistöö ettevalmistavas kui ka tõlgendavas osas (vt Õunapuu 2014: 68). DA jaoks on küsimus endiselt selles, kuidas ühest küljest kõige paremini kasutada keelset tekstimaterjali selle struktuuride väljatoomisel, teisalt aga näidata selle materjali kontekstipõhisust tõlgendustes. Kas ja kuidas on aja jooksul arenenud andmetöötlusprogrammid ja võimalikud töödeldavad materjalimahud muutnud DA-d ja selle esitatavaid küsimusi?

Suurte andmehulkade kerge kättesaadavus ja otsingute tegemise lihtsus tänapäeval on mõistagi tõstatanud küsimuse, kas ja kuidas vahepeal radikaalselt kvalitatiivset rada käinud diskursuseuuringud kvantitatiivse uurimisviisi juurde (tagasi) tuua. Viimastel aastakümnetel on digiteeritud väga suured andmekogumid. Nt Google'i, Internet Archive'i ja terve rea ülikoolide koostöös loodud HathiTrust sisaldas juba 2015. aastal 13,7 miljonit raamatut, mis esmakordselt on tekitanud võimaluse uurida mitte ainult üksiktekste, vaid näiteks kogu sajandi jooksul mingil teemal loodud teoseid. Sarnaseid keskondi, kus tehakse kättesaadavaks digiteeritud meediatekstdid, poliitilised

kõned ja arhiivimaterjalid, on eri keeltes palju ning nendesse on sisse ehitatud vähemalt mingid otsingutööriistad. Sellised andmehulgad võimaldavad uurida erinevaid sõnavalikuid, sõnade kasutustähendust diskursustes, diskursiivseid mustreid, tekstistruktuure. Nende arengute valguses on DA kasutajatel kiususatus, kui mitte suisa vajadus liikuda üksikute juhtumianalüüside juurest laiahaardelisemaid üldistusi lubavate tööde juurde.

Tähendusstruktuuride uurimiseks on selles kvantitatiivsuskeskses intellektuaalses kontekstis taas üles võetud korpuslingvistika alla kuuluv sõnade ja fraaside sageduse ning kollokatsioonide uurimine. Korpusepõhine diskursuseuuring (ingl *Corpus-Assisted Discourse Studies* ehk CADS) on üks uurimisviise, mis proovib kvalitatiivsust ja kvantitatiivsust lepitada. Selles kasutatakse korpuslingvistika vahendeid erinevate diskursusetüüpide uurimiseks ja võrdlemiseks, et tuvastada neist tähendusi, mis pole pinnapealsel vaatlusel ilmsed (Partington 2008: 96–97). Kuna fookuses on konkreetne diskursusetüüp, siis luuakse *ad-hoc*-korpusi, kasutades üldkorpusi võrdlusalusena. Selliste uuringute hulk on 1990. aastate teisest poolest üha suurenenud, sest andmepõhine kvantitatiivne lähenemine suurematele tekstikorpustele lisab varem tihti subjektiivsuses ja andmete valikulises tõlgendamises süüdistatud DA-le usaldusväarsust (Baker, Levon 2015: 222). Seepärast on ka kriitilistes diskursuseuuringutes kvantitatiivsusele toetuv uurimisviis üha populaarsem (vt nt Taylor, Marchi 2018).

Jätkuvalt valmistab aga kvantitatiivse uurimisviisi puhul raskusi DA huvi tekstide kui laiemate analüüsiühikute vastu. Nende kvantifitseerimisel tekib ridamisi küsimusi. Michael Jones ja Melody Dye (2018) märgivad, et on kerge tuvastada mingi sõna või foneemi sagedust, aga kui räägime mitmelauseelistest jadadest, siis tõuseb erinevaid keerukaid küsimusi. Nii nagu uurija valitud kontrastidele rajatud algsete uurimuste tekstivalimid võimaldasid välja tuua ennekõike eeldatavaid kontraste (vt leksikomeetrilise lähene-mise kriitikat alaosas 1.2), nii rõhutab digihumanitaar Ted Underwood (2014: 66) ka tänapäeva ülisuurte korpuste puhul, et ilma refleksioonita sõnaotsing võimaldab meil uurida seda, mida me juba teame oodata. See on eriti probleemne, kui uurime ajaloolisi tekste, milles sisalduvad teadmised võivad olla organiseeritud teistmoodi, kui me eeldame. Gerlinde Mautner võrdleb ühes intervjuus andmeanalüüsi tööriistu uute mänguasjadega, mille kasutamisel ei tohiks unustada, et nendega saab teha ainult seda, milleks nad on loodud, ning seepärast soovib ta kvantitatiivsete ja kvalitatiivsete meetodite kombineerimist (Fruttaldo, Cambria 2017: 293).

Paul Baker ja Erez Levon (2015) on teinud huvitava kvantitatiivsete ja kvalitatiivsete meetodite sobivuse katse, analüüsides sama korpust erinevalt: üks uuris 41,5 miljoni sõnalist meediaartiklite korpust, teine selle baasil loodud 51-artiklilist alakorpust. Suure korpuse analüüs võimaldas – lisaks ilmsele, ehk siis võimalusele uurida suuremat tekstimassiivi – välja tuua korduvad leksi-kaalsed mustrid ning representatsioonide esinemissagedused. Selle analüüsi- viisi nõrkus oli aga materjali kirjeldamine ilma tõlgenduseta ning samuti võimalus, et suures materjalihulgas ei näe analüüsija kasutuse irdumist üldisest mustrist, mille jaoks on vajalik kontekst (Baker, Levon 2015: 231). Probleemitu pole ka kvalitatiivne uurimus, mis võib taastoota uurijate – või meid ümbritseva ühiskonna – eeldusi ning kus tekstide poeetilisse struktuuri süvenemine ei

pruugi uurijal võimaldada näha laiemaid mustreid. Samas aga tõi just kvalitaativne lähenemine esile tekstides vaikumisi esinevaid jooni ning võimaldas analüüsida tekstide struktuurset poolt ning selle mõju diskursustele – ehk siis nii keelelise kui ka ühiskondliku poole nüansse (Baker, Levon 2015: 233). Just sel põhjusel jõudsid Baker ja Levon järeldusele, et suurte andmehulkade olemasolu korral võiks kvalitatiivset ja kvantitatiivset uurimisviisi kasutada teineteist täiendavalt, st järelduste valiidsuse ja andmetest parema arusaamise huvides erinevaid andmehulki, uurimissisendeid või seletusviise kombineerides.

Mitte just sama korpuse, aga sama nähtuse erisuunaliste analüüsides kõrvuti asetamise näitena võib tuua ajakirja Semen ühe numbri (2015, nr 38, „Pragmatique de la répétition”), kus statistiliste meetodite kaasabil vaadeldakse prantsuse keeles stilistiliselt häiriva sõnakorduse esinemist kirjandus-tekstides. Kuna leksikomeetriselised meetodid võimaldavad välja tuua tekstides igasuguseid korduvaid segmente, on see suurte korpuste kaasamiseks üks võimalikke rakenduskohti. Ühe autori, Jean Giraudeau’ stiili eripärast lähtudes uurib Étienne Brunet (2015) selle nähtuse esinemisloogikat kõigepealt Girardeau’ ühes romaanis, siis kogu autori loomingus, laiendades lõpuks vaatenurka Google Booksi kaudu kättesaadavale prantsuse kirjanduse koguhulgale, küsides, kuidas asetub üks stiilinähtus teadlike valikute ja tekstižanrite jm tekstiloomme mehhanismide suhtes. Ajavahemikus 1800–2000 oluliselt vähenenud kordust sisaldavate konstruktsioonide analüüsi kaudu näitab ta ühest küljest „mikrotekstiliste” lõigete abil leksikomeetria potentsiaali keelekasutuse uurimisevahendina, kuid jõuab teisalt tagasi korpuse moodustumise küsimuste juurde, meenutades, et erinevalt arvukast ja ka tüpoloogiliselt varieeruvast tänapäevasest kirjavarast jõuavad kasutatavasse korpusesse vaid need varasemad tekstid, mis on leidnud arhiveerimist. Vastassuunas liigub Agathe Mezadri (2015), kes toob leksikomeetria abil välja korduste olulisuse ühe ajaloolise autori, Féneloni teostes kõigepealt kogu Prantsusmaa rahvusraamatukogu andmebaasis Gallica leiduva materjali taustal ja analüüsib siis korduste toimimist mälumehhanismidena lõigu- ja tekstitasandil, pürgides nii stilistilise joone vaatlemise kaudu „mikroteksti” kvalitatiivse laiendamise poole.

„Korpustevaimustuses”, mis selle sajandi algusest saadik on vallanud kõige erinevatel viisidel keelekasutuse uurijaid, on eristatud laias laastus kaht korpuste kasutamise viisi: korpus kui teooriate testimisruum ja korpus kui dünaamiline vaadeldav ruum, millest järeldusi tehakse *a posteriori* (Mayaffre 2005). Kutsudes üles lõpetama külma sõda nende vahel, kes keelekasutusnäidetega vaid illustreerivad teooriat, ja teisalt nende vahel, kes kirjeldusest igasuguse teooria kõrvale jätavad, seab Damon Mayaffre’i kaasabil ilmuv, 2002. aastal asutatud ajakiri Corpus eesmärgiks uurija ja keelefaktide teadliku vahendamise (Mellet 2002). Kuidas selles „pingelõdvendusprotsessis” kvantitatiivseid ja kvalitatiivseid uurimisviise teineteisele lähendada, on olnud loogiliselt järgnevate arvukate arutelude teema. Ühest küljest on järjest arendatud leksikomeetrisi tööriistu ning ajaloolasest diskursuseanalüüsija Jacques Guilhaumou (2002) arvates näiteks hakkab suurtele andmemahtudele avatud uurimustes leksikomeetria kaasabil esilduvate märksõnade ja konstruktsioonide kaudu tööle Foucault’ intuiitselt praktiseeritud „arheoloogiline” meetod: uurijale, kes ei pea end piirama sihipäraselt valitud allikatega, võib leksikomeetria kätte näidata hoopis uusi uurimissihthe. Teisest küljest

osutab Marie-Anne Paveau (2014), et paljude diskursuseanalüüsijate jaoks on esmane ja tavaline, „markeerimata” lähenemine jätkuvalt kvalitatiivne uurimisviis, mida tihtipeale ei määratletagi, samal ajal kui kvantitatiivne lähenemine end ikka ja jälle õigustama peab. Viimase kümnendi jooksul ridamisi korraldatud konverentsidel ja ilmunud ajakirjanumbrites, kus kvantitatiivsete ja kvalitatiivsete meetodite vahel konsensust või tasakaalu on otsitud, on läbivaks märksõnaks siiski olnud „komplementaarsus”.

Nii toonitab ka leksikomeetriast suisa mõttemale mõõtva logomeetria arendanud Mayaffre, kes ise on uurinud Prantsusmaa presidentide kõnesid läbi kogu Teise maailmasõja järgse perioodi, oma meetodit selgitades, et esimese tõlgendava valiku teeb ka kvantitatiivsete meetoditega töötav diskursuseuurija nii või teisiti oma uurimiskorpuse jaoks valimit koostades (Mayaffre 2014). Kuid suuremahulistes korpustes saavad leksikomeetria vm automatiseeritud meetodid esiteks osutada proportsionaalselt esilduvatele või ka proportsionaalselt alakasutust leidvatele sõnadele, korduvatele segmentidele või sõnatüüpidele, teiseks hõlbustada ja näitlikustada nende kasutuste tõlgendamist, mis jääb ikkagi uurija kvalitatiivse analüüsi osaks. Näiteks 2007–2012 Prantsusmaa presidendiks olnud Nicolas Sarkozy sõnavõtude korpusest esilduvat „ülekasutatavat” umbmäärast määrsõna *ça* ja selle variatsioone süntaktiliselt täpsemalt analüüsides näitab Mayaffre (2014), et see pole lihtsalt märk moodsast kõnekeelsest väljenduslaadist võrdluses varasemate presidentidega, vaid hägustava ja konsensust eeldava väljenduse laiem diskursiivne strateegia.

Need on vaid mõned näited praegustest väga mitmepalgelistest uuringutest DA üldraamis. Tervikuna on aga huvitav tõdeda, et tekstihulkade keelelise uurimise kontekstis on diskursuseuurijad jõudnud omamoodi ringiga tagasi Harrise 1950-ndatel väljajoonistatud kaheharulise traditsiooni juurde, mis küsib ühelt poolt teksti ja teiselt poolt sotsiaalsete kontekstide struktuuri järele. Varieerides rõhuasetusi kasutuses oleva keele ja tekstide sotsiaalse paigutumise uurimisel, jagavad nii ingliskui ka prantsuskeelne traditsioon jätkuvalt põhimõtet, et DA jaoks on oluline uurida loomulikult esinevat, mitte katsete abil toodetud tekstimaterjali. DA uuritavaks võib olla igasugune keelematerjal väärtkirjandusest surmakuulutuste ja juhuslike märkmeteni, pole n-õ privilegeeritud tekste ega valmis kogumeid. Püütakse vaagida võimalikult laia tekstide ringi, et tuvastada erinevaid keelekasutuse viise, žanre, aga ka mõttemustreid. Selle materjali kvantifitseerimist ega analüüsi automatiseerimist ei peljata, ka tänapäeva suured korpused on DA jaoks väärtuslik ja põnev uurimisala. Samas seisneb DA kui fenomeni eripära Dominique Maingueneau (2002: 45) sõnutsi ka selles, et DA uurimisobjektiks võivad olla kõik ühiskonnas ringlevad lausungid, k.a juhuslikud, kõrvalised jms lausungid, mis suurtesse valmiskorpustesse ei pruugi jõuda. Nii on diskursuseuurijad, erinevalt näiteks korpuslingvistidest, kogunud ja koostanud oma uurimiskorpuse rohkem ise, lähtudes mõnikord võib-olla intuiitvsetest valikutest, kuid kui need valikud on olulised sotsiaalse tõlgendusruumi problematiseerimise seisukohast, on need samuti osa – ja isegi oluline osa – DA-st. Näiteks kutsub tänapäeva digimaailma konteksti vaadeldes üks selle valdkonna innukaid avastajaid Marie-Anne Paveau diskursuseuurijaid empiirikat kirjeldama pigem kui uutmoodi „ökosüsteemi”, milles loendamisest olulisem oleks „osutamine” keeleliselt kirjeldavatele nähtustele (Paveau 2014).

4. Kokkuvõte

Nagu nähtub eelnevast põgusast ülevaatest, on tekstipõhist keeleanalüüsi ja sotsiaalset tõlgendust siduva DA suhe formaliseerimise ja kvantifitseerimisega ajas tugevasti muutunud. Kvantitatiivsus ise pole aga debattidesse tulnud mitte viimastel aegadel, vaid üle 60 aasta tagasi, mil arenev arvutiteadus pakkus uurijaile ahvatleva väljakutse tekstide lugemist ja analüüsi keeleliselt formaliseerida. N-ö tekste üksteise peale asetav lugemisviis lõi võimaluse otsida tekstidest seni kättesaamatuid mustreid, pääsedes samas üksikjuhtumite analüüsiga paratamatult kaasnevast subjektiivsuseohust. Keele toimimist ühiskonnakriitiliselt tõlgendavate lugemiste tuultes kujunes vahepeal keeleteaduse väljal enam iseseisvunud ja samas paljude külgnevate distsipliinidega haakuv DA kvalitatiivsete lähenemiste lipulaevaks, seda ka väljaspool lingvistikat. Tänapäeval rõhutatakse, muuhulgas tänu vahepeal arendatud kriitilistele ja teoretiseerivatele käsitlustele, siiski pigem kahe uurimisviisi teineteist täiendavat komplementaarsust.

Artiklis püüdsime sama dünaamikat kaardistada nii inglis- kui ka prantsuskeelses DA traditsioonis, näidates nende sarnaseid lähtekohti ja samas erinevaid arenguid. Ühiskondliku ja akadeemilise konteksti erinevuse tõttu on nii kvalitatiivsed kui ka kvantitatiivsed püüdlused olnud kahes keeleruumis üsna eripalgelised (ja tihti vastastikku tundmatud). Osutasime, et arvutilingvistika ja tekstigrammatika arendamisele aluse pannud Harrise ideed jõudsid prantsuse keelde nii ajalise kui ka rakendusliku nihkega, leides Prantsusmaal kohe sotsiaalse sisu, samas kui ingliskeelses maailmas jäi diskursuse mõiste esialgu seotuks pigem tekstide ja loomuliku keele struktuuridega. Teisalt, kui ühiskonnakriitilise prantsuse filosoofi Foucault' ideed mõjutasid kiirelt ka ingliskeelset DA-d, siis näiteks teise prantslase Pêcheux' keelefilosoofilisem automaatse diskursuseanalüüsi raamat ilmus ingliskeelses tõlkes alles aastal 1995, palju hiljem kui tema 1982. aastal tõlgitud raamat keelest, semantikast ja ideoloogiast, ning teda on vaadeldud pigem filosoofi kui keeleteadurina. Peamiselt ingliskeelses traditsioonis tekkinud kriitilist diskursuseanalüüsi vaadeldes tuli esile, et prantsuskeelses traditsioonis polnud põhjust DA-d „kriitiliseks” nimetada, sest tekkinud uurimissuund oli juba eos ühiskonnakriitiline, samas kui inglise keeles oli laiemal formalistlikul keeleteaduslikul väljal vajadus vastandava kriitilise sildi järele. Selliseid erinevaid haakumisi, k.a diskursuse mõiste erinevate tõlgenduste vahel, võiks kindlasti vaadelda veel pikemaltki, kuid küsimused tänases üleilmastuvas maailmas on ühised.

Diskursuse kui kasutuses oleva, loomuliku keele uurimise seisukohast on oluline DA-d saatnud refleksioon uurimiskorpuste mahu, iseloomu ja paigutumiste üle. Keelekasutuse mustrite kirjeldamiseks on DA jaoks väärtuslikud suured, automaatselt töödeldavad andmekorpused. Tekstihulkade toimimise uurimisel analüüse automatiseerides lähenetakse nii paljuski üha enam korpuslingvistikale. Ent oluliseks dominandiks jääb diskursuseuurijate jaoks tekstide eripära, sotsiaalsete paigutumiste ja tõlgendusväljundite tähtsustamine. Kvalitatiivsus ja kvantitatiivsus vastandamise juurest on diskursuseuurijad praeguseks üldiselt jõudnud järelduseni, et kõige viljakam on kahe lähenemise tugevusi kasutada komplementaarselt. Kui andmetöötluse alg-etapis oli analüüsi formaliseerimine ka mingil määral eesmärk omaette – see

oli põnev ning lisas intuitsioonidele kaalu –, siis praktika arenedes on järjest teadlikumalt hakatud kasutama analüüsi automatiseerimist kui vahendit, et tõlgendusi mitte ainult kaalukamaks, vaid ka sisuliselt täpsemaks arendada. Sotsiaalse tõlgendusruumi seisukohalt on DA jaoks olulisim küsida problematiseerivaid küsimusi. Selles ülesandes aitab kvantitatiivsus uuritavaid nähtusi teravapilgulisemalt märgata, problematiseeriv ja selgitav analüüsivaim neid avada.

Artikli avaldamist on toetanud Euroopa Liit Euroopa Regionaalarengu Fondi kaudu (Eesti-uuringute Tippkeskus), see on seotud Eesti Haridus- ja Teadusministeeriumi uurimisprojektiga PUT-1481 „Imaginaarsete narratiivsete stsenaariumide roll kultuuri-dünaamikas”.

Kirjandus

- Aava, Katrin 2004. Narratiiv müüdi loome teenistuses: arhetüübid meediatekstides. – Tekstid ja taustad III. Lingvistiline tekstianalüüs. Toim Reet Kasik. Tartu: Tartu Ülikooli Kirjastus, lk 10–32.
- Baker, Paul, Levon, Erez 2015. Picking the right cherries? A comparison of corpus-based and qualitative analyses of news articles about masculinity. – Discourse & Communication, kd 9, nr 2, lk 221–236.
- Benveniste, Émile 1966. Problèmes de linguistique générale I. Paris: Gallimard.
- Brown, Gillian, Yule, George 1983. Discourse Analysis. Cambridge: Cambridge University Press.
- Brunet, Étienne 2015. La répétition dans la phrase. Etude statistique. – Semen, nr 38. <https://journals.openedition.org/semen/10287> (24. VII 2018).
- Charaudeau, Patrick 2009. Dis-moi quel est ton corpus, je te dirai quelle est ta problématique. – Corpus, nr 8, lk 37–66. <http://journals.openedition.org/corpus/1674> (24.VII 2018).
- Charolles, Michel, Combettes, Bernard 1999. Contribution pour une histoire récente de l'analyse du discours. – Langue française, nr 121, lk 76–116.
- Claudel, Chantal, Münchow, Patricia von, Pordeus Ribeiro, Michele, Pugnère-Saavedra, Frédéric, Tréguer-Felten, Geneviève (toim) 2013. Cultures, discours, langues. Nouveaux abordages. Limoges: Lambert Lucas.
- Dijk, Teun van 1972. Some Aspects of Text Grammars. The Hague: Mouton.
- Dubois, Jean 1962. Le Vocabulaire social et politique en France de 1869 à 1872. Paris: Larousse.
- Dubois, Jean 1969. Énoncé et énonciation. – Langages, nr 13, lk 100–110.
- Dubois, Jean, Sumpf, Joseph (toim) 1969. Langages, nr 13, Analyse du discours.
- Eribon, Didier 1991. Michel Foucault. Paris: Flammarion.
- Fairclough, Norman 1989. Language and Power. London: Longman.
- Fairclough, Norman 1992. Discourse and Social Change. Cambridge: Polity Press.
- Fairclough, Norman 1995. Critical Discourse Analysis. The Critical Study of Language. Harlow: Longman.
- Foucault, Michel 1994 [1968]. Sur l'archéologie des sciences. Réponse au Cercle d'épistémologie. – Dits et écrits 1954–1988, kd I, tekst nr 59. Toim Daniel Defert, François Ewald. Paris: Gallimard, lk 696–731.

- Foucault, Michel 1994 [1973]. A verdade e as formas Jurídicas. La vérité et les formes juridiques. – Dits et écrits 1954–1988, kd II, tekst nr 139. Toim Daniel Defert, François Ewald. Paris: Gallimard, lk 538–646.
- Foucault, Michel 1994 [1984]. Le souci de la vérité (entretien avec F. Ewald). – Dits et écrits 1954–1988, kd IV, tekst nr 350. Toim Daniel Defert, François Ewald. Paris: Gallimard, lk 668–678.
- Foucault, Michel 2005 [1969]. Teadmiste arheoloogia. Tlk Kaia Sisask. Tartu: Tartu Ülikooli Kirjastus.
- Foucault, Michel 2015 [1966]. Sõnad ja asjad. Tlk Mirjam Lepikult. Tallinn: Vararak.
- Fowler, Roger 1991. Language in the News. Discourse and Ideology in the Press. London–New York: Routledge.
- Fruttaldo, Antonio, Cambria, Mariavita 2017. Combining qualitative and quantitative approaches to discourse analysis. In conversation with Gerlinde Mautner and Alan Partington. – Im@go. A Journal of the Social Imaginary, nr 9, lk 284–304.
- Gadet, Françoise, Pêcheux, Michel 1981. La langue introuvable. Paris: Maspero.
- Guilhaumou, Jacques 2002. Le corpus en analyse de discours: perspective historique. – Corpus, nr 1. <http://journals.openedition.org/corpus/8> (24. VII 2018).
- Hall, Stuart 1997. Introduction. – Representation. Cultural Representations and Signifying Practices. Toim S. Hall. London: Sage, lk 1–12.
- Halliday, M. A. K. 1978. Ideas about Language. – Arts: The Journal of the Sydney University Arts Association, nr 11, lk 20–38.
- Halliday, M. A. K. 1994. Introduction to Functional Grammar. 2. tr. London: Edward Arnold.
- Halliday, M. A. K. 2002. On Grammar. London: Continuum.
- Halliday, M. A. K. 2007. Language and Society. London: Continuum.
- Haroche, Claudine, Henry, Paul, Pêcheux, Michel 1971. La sémantique et la coupure saussurienne: langue, langage, discours. – Langages, nr 24, lk 93–106.
- Harris, Zellig Sabbetai 1952. Discourse analysis. – Language, kd 28, nr 1, lk 1–30.
- Harris, Zellig Sabbetai 1969. Analyse du discours. Tlk Françoise Dubois-Charlier. – Langages, nr 13, lk 8–45.
- Hodge, Robert, Kress, Gunther 1993. Language and Ideology. 2. tr. London: Routledge.
- Jones, Alan 2010. Michael Halliday: An appreciation by Alan Jones. – IH. Journal of Education and Development, kd 28. <http://ihjournal.com/michael-halliday-an-appreciation> (24. VII 2018).
- Jones, Michael N., Dye, Melody W. 2018. Big data approaches to study discourse processes. – The Routledge Handbook of Discourse Processes. 2. tr. Toim Michael F. Schober, David N. Rapp, M. Anne Britt. New York: Routledge, lk 117–124.
- Kalmus, Veronika 2015. Diskursusanalüüs. – Sotsiaalse analüüsi meetodite ja metodoloogia õpibaas. Tartu: Tartu Ülikool. samm.ut.ee/diskursusanalüüs (24. VII 2018).
- Kasik, Reet 2002. Tekstil on mitu palet. – Tekstid ja taustad. Artikleid tekstianalüüsist. Toim R. Kasik. Tartu: Tartu Ülikooli Kirjastus, lk 7–11.
- Kasik, Reet 2003. Funktsionaalse tekstianalüüsi eri suunad. – Tekstid ja taustad II. Artikleid tekstianalüüsist. Toim R. Kasik. Tartu: Tartu Ülikooli Kirjastus, lk 7–9.

- Kasik, Reet (toim) 2004. Tekstid ja taustad III. Lingvistiline tekstianalüüs. Tartu: Tartu Ülikooli Kirjastus.
- Kiisel, Maie 2013. Problems of Critical Analysis of Communication of Environmental Issues and Risks. (Dissertationes de mediis et communicationibus Universitatis Tartuensis 22.) Tartu: University of Tartu Press. <http://hdl.handle.net/10062/34592>
- Léon, Jacqueline 2010. AAD69: archéologie d'une étrange machine. – Semen, nr 29. <http://journals.openedition.org/semen/8823> (24. VII 2018).
- Longacre, Robert E. 1983. The Grammar of Discourse. Dordrecht: Springer.
- Longacre, Robert E. 2002. Some implications of Zellig Harris' discourse analysis. – The Legacy of Zellig Harris: Language and Information into the 21st Century. Kd 1: Philosophy of Science, Syntax and Semantics. Toim Bruce E. Nevin. Amsterdam: John Benjamins, lk 117–135.
- Loriot, James 1957. Some problems in translating paragraphs idiomatically. – The Bible Translator, kd 8, nr 4, lk 166–169.
- Loriot, James, Hollenbach, Barbara 1970. Shipibo Paragraph Structure. – Foundations of Language, kd 6, nr 1, lk 43–66.
- Maingueneau, Dominique 1996. Les termes clés de l'analyse du discours. Paris: Seuil.
- Maingueneau, Dominique 2001. Archéologie et analyse du discours. – Communication à une table-ronde sur Foucault, 23. VII 1998, Reims. http://www.revue-texto.net/Reperes/Themes/Maingueneau_Archeologie.html (22. VII 2018).
- Maingueneau, Dominique 2002. Discours. – Dictionnaire d'Analyse du Discours. Toim D. Maingueneau, Patrick Charaudeau. Paris: Seuil, lk 185–190.
- Maingueneau, Dominique 2012. Que cherchent les analystes du discours? – Argumentation et Analyse du Discours, nr 9. <http://aad.revues.org/1354> (22. VII 2018).
- Maingueneau, Dominique 2016. Énonciation et analyse du discours. – Corela, nr HS-19. <http://journals.openedition.org/corela/4446> (22. VII 2018).
- Maldidier, Denise 1969. Analyse linguistique du vocabulaire politique de la guerre d'Algérie d'après six quotidiens parisiens. (Thèse de doctorat.) http://classiques.uqac.ca/contemporains/maldidier_denise/analyse_linguistique/analyse_linguistique.html (26. VII 2018).
- Marandin, Jean-Marie 1979. Problèmes d'analyse du discours. Essai de description du discours français sur la Chine. – Langages, nr 55, lk 17–88.
- Marcellesi, Jean-Baptiste 1971. Le Congrès de Tours (décembre 1920). Études sociolinguistiques. Paris: le Pavillon.
- Matoré, Georges 1953. La Méthode en lexicologie. Paris: Didier.
- Mayaffre, Damon 2005. Rôle et place des corpus en linguistique: réflexions introductives. – Texto!, kd 10, nr 4. http://www.revue-texto.net/Corpus/Publications/Mayaffre_Corpus.html (26. VII 2018).
- Mayaffre, Damon 2014. „ça suffit comme ça !”. La fausse opposition quantitatif / qualitatif à l'épreuve du discours sarkozyste. – Corela, nr 15. <https://journals.openedition.org/corela/3543> (26. VII 2018).
- Mellet, Sylvie 2002. Corpus et recherches linguistiques: introduction. – Corpus, nr 1, lk 5–13. <https://journals.openedition.org/corpus/7> (26. VII 2018).

- Mezzadri, Agathe 2015. Répétition lexicale et mémorisation: Une approche lexicométrique de l'infini dans les Œuvres complètes de Fénelon. – Semen, nr 38. <https://journals.openedition.org/semen/10299> (24. VII 2018).
- Mills, Sara 2004. Discourse. 2. tr. London: Routledge.
- Partington, A. S. 2008. The armchair and the machine: Corpus-assisted discourse studies. – Corpora for University Language Teachers. Toim Carol Taylor Torsello, Katherine Ackerley, Erik Castello. Bern: Peter Lang, lk 95–118.
- Paveau, Marie-Anne 2014. L'alternative quantitatif/qualitatif à l'épreuve des univers discursifs numériques. – Corela, nr 15. <https://journals.openedition.org/corela/3598> (26. VII 2018).
- Pêcheux, Michel 1969. Analyse automatique du discours. Paris: Dunod.
- Pêcheux, Michel 1975. Les vérités de la Palice. Paris: Maspero.
- Pêcheux, Michel 1982. Language, Semantics and Ideology. Tlk Harbans Nagpal. New York: St. Martin's Press.
- Pêcheux, Michel 1995. Automatic Discourse Analysis. Tlk David Macey. Amsterdam: Rodopi.
- Piirimäe, Eva 2011. Keeleline pööre. – Humanitaarteaduste metodoloogia. Uusi väljavaateid. Toim Marek Tamm. (Gigantum humeris.) Tallinn: Tallinna Ülikooli Kirjastus, lk 33–58.
- Revel, Judith 2002. Le vocabulaire de Foucault. Paris: Ellipses.
- Sager, Naomi 1967. Syntactic analysis of natural language. – Advances in Computers, nr 8, lk 153–188.
- Sager, Naomi 1981. Natural Language Information Processing: A Computer Grammar of English and Its Applications. Reading, Mass.: Addison-Wesley.
- Salem, André 1982. Analyse factorielle et lexicométrie: synthèse de quelques expériences. – Mots, nr 4. Abus de mots dans le discours. Désabusement dans l'analyse du discours, lk 147–168. (26. VII 2018).
- Sarapu, Kathy 2004. Väärtushinnangute väljendamise juhtkirjades. – Tekstid ja taustad III. Lingvistiline tekstianalüüs. Toim Reet Kasik. Tartu: Tartu Ülikooli Kirjastus, lk 120–137.
- Sarapu, Kathy 2005. Ajalehe juhtkiri õpetajana. – Tekstid ja taustad IV. Teksti liigivaatlusi. Toim Reet Kasik. Tartu: Tartu Ülikooli Kirjastus, lk 100–117.
- Saussure, Ferdinand de 2017 [1916]. Üldkeeleteaduse kursus. Tlk Tiit Kuuskmäe. Tallinn: Varrak.
- Saussure, Ferdinand de 2002. Ecrits de linguistique générale. Paris: Gallimard.
- Taylor, Charlotte, Marchi, Anna 2018. Corpus Approaches to Discourse. London: Routledge.
- Underwood, Ted 2014. Theorizing research practices we forgot to theorize twenty years ago. – Representations, kd 127, nr 1, lk 64–72.
- Wagner, Robert-Léon 1967–1970. Les Vocabulaires français. Kahes köites. Paris: Didier.
- Õunapuu, Lembit 2014. Kvalitatiivne ja kvantitatiivne uurimisviis sotsiaalteadustes. Tartu: Tartu Ülikool. <http://hdl.handle.net/10062/36419> (26. VII 2018).

Marge Käsper (snd 1973), PhD, Tartu Ülikooli prantsuse keele ja lingvistika lektor, marge.kasper@ut.ee

Raili Marling (snd 1973), PhD, Tartu Ülikooli anglistika professor, raili.marling@ut.ee

The shifting relationship between discourse analysis and quantification in the English- and French-language tradition

Keywords: discourse analysis, formalisation, text corpora, lexicometry

The article gives a historical overview of discussions within the Anglophone and French traditions of linguistic discourse analysis about the formalisation and quantification of text data. We begin with the first steps towards the formalisation of text analysis, derived from the work of Zellig Harris (1952). Then we consider the lexical interpretation of Harris' work by the French lexicometric school and in a more syntax-based and critically engaged approach developed by Michel Pêcheux (1969). In contrast to the quantitative methods, we discuss qualitative approaches to discourse analysis within the humanities more generally (Michel Foucault) and specifically within linguistics (M. A. K. Halliday, *Critical Discourse Analysis*). Considering recent trends and discussions on the impact of Big Data on text analysis, we come to the conclusion that discourse analysis has evolved to appreciate the complementary use of quantitative and qualitative methods.

Marge Käsper (b. 1973), PhD, University of Tartu, Lecturer in French Language and Linguistics, marge.kasper@ut.ee

Raili Marling (b. 1973), PhD, University of Tartu, Professor of English, raili.marling@ut.ee