

Kirjanduskeele kauglugemine

Kirjandust väärindava metodoloogia arendamine
tsiteerimisvõtte näitel

MAARJA HELENA MERISTE

Digimeetodite kasutuselevõtt on kirjandusteaduses toomas uurimisviise, mis võivad põhjalikult ümber kujundada senised arusaamad kirjanduskeelest ja kirjandusajaloost laiemalt. Arvutilingvistid on välja arendanud tänuväärseid kvantitatiivse keeleanalüüsi meetodeid ja tööriistu, mis sobivad loomuliku keele uurimiseks ja töötlemiseks. Eesti keeleteaduses on arvutuslikud meetodid juba kanda kinnitanud, ent kirjandusteaduses on neil põhinevaid lähenemisi vaid põgusalt katsetatud. Miks? Keeleteaduses on deskriptiivsed uurimismeetodid (sh arvutustel ja mõõtmistel põhinevad) otseselt kohasemad, sest objekt on inimkeel ise ja eesmärk seda kirjeldada. Kirjanduse keel seevastu on kunstiline keel. Juri Lotmani (1991) traditsioonis mõtestatakse kunstilist keelt teisele modelleeriva süsteemina ehk omaette keelena, mis on tuletatud loomulikust keelest. See teeb kirjandusest märksa keerulisema uurimisobjekti, mis vajab selle eripärale kohandatud uurimisvõttestikku: kirjanduse keel on sihilikult mitmetähenduslik, mänguline, vihjav, varjav, kujundlik, irooniline jne. Tähenduslik ambivalents, tõlgendusvõimaluste paljusus on kunstilise keele essents, mis loob mitmekihilise teksti. Eesti kontekstis on selle hea näide nõukogude aja kirjanduse n-ö ridadevaheline poeetika. Teisisõnu – ilukirjanduses ei pruugi tähendust kandvad elemendid, nendevahelised seosed ja nende struktuur nähtaval olla. See pärast ei saa keeleanalüüsi meetoditega kirjandusteose tähendust selgitada, ent just tähendusprobleemid on kirjandusteaduse kese. Küll aga võib arvutuslikult tuvastada ilukirjandusest keelekasutusmustreid, mida lugemisel ei haara, kuid mis võivad kultuuriloolisse konteksti asetatuna osutada tähendusrikkaks ning täiendada või muuta seniseid tõlgendusi.

Sellest tulenevalt on digitaalse kirjandusuurimuse suurim väljakutse sobiva metoodika arendamine vastavalt uurimisprobleemile, mis on loomult kvalitatiivne ning tähenduslikult kompleksne. Käesoleva artikli eesmärk ongi ühe magistritöö (Meriste 2020) varal tuua esile uurimisvõimalusi, mille avab kirjandusteaduse jaoks digitaalne analüüs. Siinkirjutaja arvates on digimeetodite viljakaim kasutuspotentsiaal kirjandusteaduses vormiuurimus. Seda näitlikustatakse artiklis kirjanduskeele ühe spetsiifilise nähtuse, nn avatud tsitaadi mõistmiseks loodud metoodika ja selle rakendamise kaudu. Terminiga *avatud tsitaat* on prantsuse filosoof François Recanati (2001: 637) tähistanud poetilist võtet, mis seisneb lausungi mingi osa tsitaadiks märkimises nii, et jutumärkidega tõstetakse esile lisa- või moonutatud tähenduskihi (ironilise, eufemistliku vms) olemasolu. Tsiteerimist on koos muude mänguliste võtetega, nagu paroodia, pastišš ja kollaaž, peetud omaseks 1990. aastate eesti kirjandusele (Pruul 2000: 181), ometi pole selle tegelikku kasutust eesti kirjanduses

uuritud. Selle üheks põhjuseks on arvatavasti olnud asjaolu, et niisugused nähtused võivad ühe teose kontekstis olla liiga marginaalsed, et usaldusväärseks uurimuseks vajalikku hulka näiteid kokku saada: vaja on teosteülest vaadet, mida võimaldavadi digimeetodid. Tänu nendele saab kirjandust vaadelda nähtusekeskselt, nii et üksiku teose või autori eritlemise asemel on rõhk vormi uurimisel ehk sellel, kuidas mõni motiiv, stilistiline element vm võte kirjandusväljal ilmneb, kordub ja teiseneb.

Avatud tsitaadid on kirjanduskeele nähtus, mida on hea sel viisil uurida. Erinevalt paljudest muudest kirjanduslikest võtetest on avatud tsitaatidel kindel formaalne tunnus (jutumärgid), mille abil saab need tuvastada. Artiklis avatakse uurimuse tausta, ettevalmistust, analüüsi käiku ja uurimistulemusi.

Artikli põhitees on, et kunstiuurimusele kohase andmeteanduse arendamisel ei tohi kõrvale heita kvalitatiivset lähenemist, vaid see tuleb ühendada kvantitatiivsega, nii et need teineteist täiendaks. Kirjanduse puhul on arvutuslikud meetodid tulemuslikud stilistiliste üldmustrite tuvastamisel, kuid see üksi jääb vähekõnekaks. Kvalitatiivne analüüs aitab kvantitatiivselt saadud teadmist tõlgenduse ja konteksti asetamise kaudu vääridada. Kvalitatiivsete ja kvantitatiivsete meetodite kombineerimine lubab kirjandusteaduslikele probleemidele läheneda nende loomupärast keerukust lihtsustamata ning sellise uurimisprotsessi kujundamine on digitaalse kirjandusuurimuse esmatähtis oskus.

Vormikoolkonnast kauglugemiseni

Paljusid teoseid ühendava stiilivõttestiku uurimise vastu tekkis huvi ammu enne digiajastut. Kirjandusteooria ajaloos on selle lähenemisviisi silmapaistvaimad arendajad XX sajandi alguse vene vormikoolkonna esindajad. Formalistid pidasid oma ülesandeks vabastada kirjandusteadus traditsioonilistest, ent nende silmis väheväärtuslikest uurimistavadest, millest üks põhilisi oli teose seletamine autori isiku ja eluloo kaudu. Selle asemel rõhutasid nad vajadust keskenduda just kirjandusteosele ning mõelda sellest kui käsitööesemest, mille uurimisel pole oluline selle valmistanud meister, vaid selles avalduvad võtted. Seda vaadet näitlikustab hästi formalist Ossip Briki 1923. aasta manifest „Nn „Vormimeetod””, kus ta kirjutab:

Luuletaja sotsiaalset rolli ei saa mõista tema individuaalseid omadusi ja harjumusi analüüsid. **Tarvis on massiliselt tundma õppida poeetilise käsitöö võtteid**, nende erinevusi inimtöö naabervaldkondadest ja nende ajaloolise arengu seadusi. [---]

Luule ajalugu on sõnalise vormimise võtete arengulugu. [---]

Miks on luuletajad teemade töötlemisel kasutanud just neid ja mitte muid võtteid, mis kutsuvad esile uue võtte ilmumise ja kuidas vana välja sureb – need küsimused võetakse teaduslikus poeetikas kõige hoolikama uurimise alla. (Brik 2014)

Briki väljendatud idee on erinevates variantides ilmnenu paljude tema kaas-aegsete ja hilisemate kirjandusuurijate töödes. Tähelepanuväärseimad neist on

Moskva lingvistikaringi liikmed (nt Boriss Jarho), Praha lingvistikaringi teoreetikud (nt Roman Jakobson) ning prantsuse (post)strukturealistid (nt Gérard Genette ja Roland Barthes).

Oleg Sobchuk (2018: 10) sedastab oma väitekirjas, kus ta katsetab digimeetodite kasutamist kultuuri evolutsiooni uurimisel, et just lingvistiline strukturalism pani kirjandusteadlased otsima mustreid. Niisiis on XX sajandi strukturalistliku kirjandusteaduse pärand nüüdisaegse kvantitatiivse kirjandusuurimise filosoofiliseks ja teoreetiliseks aluseks.

Moodsas kirjandusteaduses tuntakse arvutuslike meetodite abil tehtud vormiuurimusi ka u g l u g e m i s e mõiste all. Kauglugemise looja on USA kirjandusuurija Franco Moretti, kes pälvis tähelepanu julge loosungiga, et uurijad peaksid kirjanduse tegelikuks mõistmiseks lõpetama raamatute lugemise. Sellega osutab Moretti (2013: 48) lähilugemiskeskse kirjandusteaduse kitsaskohale: väheste, tavaliselt kaanonisse kuuluvate teoste põhjal luuakse kirjanduslugu, mis tegelikult pole proportsionaalselt esinduslik.

Moretti arvates tuleb kultuuri läbivate mustrite nägemiseks distantseeruda üksikust tekstist. Selle asemel soovitab ta vaatluse alla võtta paljudest teostest koosneva tekstikorpuse, mis võimaldab uurida üksiku elemendi levimismustreid kirjanduses laiemalt. Suuremas tekstikorpuses ilmnevad teatud nähtuse kasutustendentsid, mis on aluseks üldistusjõulisematele tõlgendustele. (Moretti 2013: 49–50)

Mida suurem on korpus, seda suuremaid kultuuriloolisi, sotsiaalseid vm üldistusi saab materjali põhjal teha. Näiteks ilmnes 5000 ingliskeelse romaani (ilmunud vahemikus 1780–2000) keelekasutust uurides, et raamatud on aja jooksul muutunud aina kergemini loetavaks: laused on aina lühemad, lihtsama struktuuriga, sõnad lühemad (Piper 2018).

Nii Moretti kauglugemise idee kui ka üleüldse digimeetodite kasutamine kirjandusteaduses on viimastel aastatel olnud valdkonna teravamaid aruteluteemasid. Digihumanitaaria skeptikud ei usu, et kauglugemine suudaks anda uut vaadet kirjandusele, heal juhul leiab statistiliselt kinnitust see, mida varem juba teati (vt nt Kirsch 2014). Samuti on palju meetodilisi libastumiskohti seoses korpuse koostamise ning kanoonilise ja mittekanonilise kirjanduse vahekorraga ning universaalselt toimivaid lahendusi napib (vt nt Wilkens 2012). Korpuse koostamise osas on üks värskemaid ja põnevamaid analüüse uurijatelt (Schöch jt 2021), kes koostavad digitaalset Euroopa kirjandustekstide kogu (ELTeC – European Literary Text Collection).

Nimetatud korpust koostatakse Euroopa teaduskoostöövõrgustiku COST projekti „Distant Reading for European Literary History” raames, mille eesmärk ongi luua üleeuroopaline digitaalne kirjanduskorpus, ühtsustada ja arendada kauglugemise meetodeid ning uurida, kuidas suhestuvad kauglugemisega saadud teadmised traditsioonilise kirjanduslooga.

Kuigi kauglugemine vastandub juba terminoloogiliselt lähilugemisele, oleks digihumanitaarse kirjandusuurimuse seisukohalt kõige kasulikum kaht lähene-mist kombineerida. Selle järelaluseni on jõudnud mitmed uurijad (nt van de Ven 2017). Eestis tutvustas sellist lähenemist California ülikooli folkloristika, kirjanduse

ja kultuuriuuringute professor Timothy Tangherlini, kes luges 2019. aastal Tartu Ülikoolis kursust makroskoopilistest meetoditest kultuuri uurimisel. Ta rõhutas, et humanitaarteadustes töötatakse harva sellise andmemahuga, mida võiks pidada suurandmeteks (ingl *big data*), pigem töötatakse keskmiselt väikese andmestikuga (*medium small data*). Klassikalised statistilised ja tõenäosuslikud arvutused ei saa väikese korpuse analüüsimisel eriti tulemuslikud olla, kuid humanitaarteadustes võibki mikro- ja makrovaate kombineerimine olla sobivaim. (Tangherlini 2019)

Selline meetod on eriti oluline teose keerukama tähendusega elemente kaardistades, sest sellised elemendid on tekstis tihtilugu varjatud – näiteks intertekstuaalsus, iroonia, vihjelisus. Ka käesoleva artikli kese, tsiteerimisvõte, on tähendusloome keerukuse poolest nendega sarnane, erinedes selle poolest, et tsitaat on tekstis nähtavalt (jutumärkidega) tähistatud.

Mis on avatud tsitaat ja kuidas seda uurida?

Avatud tsitaat on kirjanduses laialt levinud võte, mida iga lugeja on kohanud, kuid uuritud on seda vähe ning mõiste on ka paljudele keele- ja kirjandusuurijatele võõras. Seepärast on alustuseks otstarbekas tuua näide uuritavast korpusest:

Me oleme ühtmoodi illusioonideta küünikud, kes praktiseerivad „mõningasest metodoloogilisest lahknevusest hoolimata”, nagu Adrian ütleb, sedasama maailmavaadet. (Hiram 1999: 17)

Kui akadeemilises ning igapäevakasutuses mõistetakse tsiteerimist valitud materjali sõnasõnalise ümbertõstmisena uude teksti, siis tähenduse uurimise vaatenurgast on tegemist palju komplekssema nähtusega. Tsitaat ei pruugi alati olla seotud viitega teisele kontekstile, mujal varem öeldud või kirjutatud sõnadele. Antud juhul on tsitaadi miinimumtunnuseks teksti mingi osa teadlik eristamine ülejäänust: jutumärgistades valitud sõnad, kutsutakse tsitaati tõlgendama teisiti kui seda ümbritsevat teksti (Capone 2013: 260, 278).

Sellist tsiteerimisviisi nimetatakse avatuks, et seda vastandada suletule. Suletud tsitaadid on täistähenduslikud ja iseseisvad lausungid – näiteks otsekõne esineb tekstis suletud tsitaatidena (Recanati 2001: 649). Otsekõnelt ei saaks jutumärke eemaldada, sest need on ühe diskursusetüübi konventsionaalne tähistusviis. Avatud tsitaat, nagu näites näha, on aga lausesse süntaktiliselt ja tähenduslikult lõimitud. Sellelt jutumärke eemaldades jääks lause keeleliselt korrektseks, aga kaotaks tuntavalt ühe tähenduskihi. Siit ilmnebki, et avatud tsitaadi eesmärk on esile tõsta lisatähenduse olemasolu – pragmaatilist rikastatust (Recanati 2001: 637).

Millised on need võimalikud lisatähendused? Esiteks evidentsiaalsus ehk see, kelle diskursusesse tsiteeritud sõnad kuuluvad – eeltoodud näites juhivad jutumärgid tähelepanu sellele, et märgitud sõnad ei pärine kõnelejal endalt, vaid teisest allikast. Laiemas mõttes saab niimoodi tähistada võõra allkeele või idiolekti tungimist peadiskursusesse. Peale selle on ilmne distantseerumine: kõneleja sarkastiline

või irooniline hoiak. Samuti võib tsitaadiga rõhutada just sõna tavatähendust, anda sellele ikoonilise jõu (Brendel jt 2011: 5–6).

Nõnda võiks avatud tsitaati mõista kui teksti vastuvõtmise ühtlase, harjumuspärase tempo teadlikku häirimist: lugemiskulg pidurdatakse, lugeja suunatakse osutama tähelepanu esiletõstetud tekstikillule. Mõnesõnaline tsitaat võib keelilise eneseleosutuse kaudu toimida otsekui päästik mõne suurema objekti või nähtuse kujutlusse äratamisel. Tsiteeritud sõnad muutuvad ikooniks mingist laiemast diskursusest, mida nad esindavad (Recanati 2001: 645). Järgnevas näites kasutatakse avatud tsitaati selleks, et osutada ühiskondlikule standardsuhtumisele:

Tol ajal käis küll koolis palgaline koristaja, aga tema käis kord nädalas ja pesi ainult treppe ja koridoripõrandaid, õpilased pidid oma klassiruumi ikkagi ise koristama ja tüdrukud tegema *naistetöid* nagu tolmupühkimine ja põrandapesu, sest „klass oli ju ometi nende teine kodu”. (Viiding 2008: 14)

Niiviisi viitab avatud tsitaat millelegi endast suuremale: suhtumisele, maailma-vaatele, kogukonnale, tekstile, kultuurile, teosele vms. Sellega haakub Anneli Mihkelevi seletus allusiivsusest kirjanduses:

Allusioon saab alguse mikrotasandilt kui vihjav signaal ehk marker, mis asub teksti pindmises kihis. Sellel markeril on esilekutsuv jõud ning käivitab allusiooni makrotasandil, st laia tähenduste ning assotsiatsioonide võrgustiku ehk allusiooni laiemas mõttes. (Mihkelev 2005: 144)

Eesti kirjandusuurimustest puudutavad tsiteerimisvõtteid Mihkelevi eelviidatud väitekirja ja Katrin Puiki doktoritöö irooniast Heiti Talviku ja Betti Alveri luules (2009). Mõlemad eritlevad kvalitatiivselt kirjandusnähtust, mis on olemuselt ambivalentne ja kindlate vormiliste tunnusteta. Käesolevas uurimuses lähenetakse teistpidi: nähtus piiritletakse esmalt formaalse tunnuse järgi ja seejärel vaadatakse, milliseid sisulisi omadusi see kätkeb.

Välismaistest kirjandusuurimustest on tsiteerimisvõttega tihedalt seotud Mihhail Bahtini kontseptsioon võõrast kõnест (Bakhtin 1982) ja Gérard Genette'i (1992: 8) lähenemine intertekstuaalsusele, mis tema järgi esineb tekstis tsitaadi, allusiooni või plagiaadina. Samuti suhestub tsiteerimise teemaga Bahtini ideede edasiarendaja Julia Kristeva (1980: 65–66), kes peab teksti teiste tekstidega suhtlemise dünaamiliseks väljaks, kusjuures teksti all ei pea ta silmas üksnes verbaalset, vaid üleüldse kultuuriteksti.

Seega saab öelda, et avatud tsitaat, mis tekstis esineb pelgalt mõne sõna tüpograafilise esiletõstuna ja mis näiliselt on seotud vaid konkreetse narratiivse hetkega, võib mõõtmatult rikastada teksti tähendust. Seda näiteks intertekstuaalsuse, vihjelisuse, tonaalsuse vm kaudu.

Avatud tsitaatide uurimise eesmärk on kirjeldada teatud ilukirjanduslikku võtet, selle levikut ja mõju ilukirjanduskeecele. Täpsemad uurimisküsimused on järgmised: milline on avatud tsitaatide levik korpuses? Millised on võtte tähenduslikud funk-

sioonid tekstis? Kas ja millised kasutustendentsid ilmnevad ning kas on märgata-
vaid erisusi eri autorite/teoste puhul? Mida võib selle põhjal öelda analüüsitud teoste
kohta?

Uurimuse ettevalmistus

Tsiteerimisvõtte efektiivse uurimise meetodi väljatöötamisel olid abiks kolm peamist
küsimust. Neist tõukunud mõttearendused pole pelgalt tsitaadispetsiifilised – neist
võiks olla kasu ka teiste sarnaste uurimuste ülesehitamisel.

Millistest teostest võiks korpus ideaalis koosneda, et
uuritavast nähtusest hea ülevaade saada? Siinkohal peab uurija
tegema kompromissi kahe vastandpõhimõtte vahel. Ühelt poolt tuleks korpus loo-
mulikult koostada teostest, kus otsitav nähtus esineb. Teiselt poolt on kauglugemise
üks põhimõtteid rakendada materjali valikul võimalikult vähe filtreid, et „kõneleda”
saaksid ka need teosed, mis tavaliselt uurijate huviorbiidist välja jäävad.

Esimese nõude puhul aitab uurijal fookust seada lugemus ja kirjandusloolised
teadmised. Näiteks tsiteerimine kuulub tekstivõtete hulka, mis on iseloomulikud
taasiseseisvumisjärgsele eesti kirjandusele. Poliitiline ja sotsiaalne murranguaeg
peegeldus kirjanduses mängulisuse, stiliseerituse, intertekstuaalsuse, allkeelte-
rohkuse, ironia ja erinevate jäljendusvõtete kasvuna (Velsker 2001: 620–622). Seega
võis tsiteerimisvõtte esinemist selle perioodi kirjanduses pidada ootuspäraseks, kuigi
statistilist kinnitust sellele pole.

Teise põhimõtte puhul tuleb arvestada sellega, et mingi valiku peab uurija niikui-
nii tegema. Eriti möödapääsmatu on see juhul, kui analüüs hõlmab ka kvalitatiivset
osa ja materjali maht peab seda arvestades olema hoomatav. Olgu valitud piirangud
ajalised, žanrilised vms, oluline on tulemustest järelduste ja üldistuste tegemisel neid
teadvustada. Näiteks ei tohiks teatud perioodi proosakirjanduse ilminguid kerge-
käeliselt üldistada kogu toonasele ilukirjandusele, mis hõlmab ka luulet.

Kust saada uurimismaterjali? Digitaalse kirjandusuurimuse läbi-
viimiseks vajab uurija teoste tekstifaile. Eesti kirjanduse puhul piiravad materjali
peamiselt kaks asjaolu: autoriõigused ja skaneerimise teel digiteeritud raamatud.
Eesti kirjandust digiteeritakse Kultuuriministeeriumi projekti „Kultuuripärandi
digiteerimine 2018–2023” (Kultuuriministeerium 2018) raames hoogsalt, Eesti
Rahvusraamatukogu digitaalarhiiv Digar täieneb pidevalt. Kahjuks aga saavad raa-
matukogud autoriõigusega kaitstud teoseid digitaalselt kättesaadavaks teha vaid
kohapealseks lugemiseks (AutÕS § 20). Autoriõigused kaitsevad teost autori eluajal
ja 70 aastat pärast surma (AutÕS § 38), mis tähendab, et praegu on uurijale kätte-
saadavad ainult enne 1951. aastat lahkunud autorite teosed. Märkimisväärne osa
digitaalselt kättesaadavast vanemast kirjandusest on aga digiteeritud skaneerituna,
mis pole tekstitöötlusvalmis formaat. Käesolevaks uurimuseks oli niisiis tarvis teo-
seid, mis pole vabalt kättesaadavad, mistõttu sõltus korpuse koosseis suuresti sellest,
kas kirjanikud ja kirjastajad olid nõus uurimuse jaoks jagama teoste elektroonilisi
tekstifaile.

Kui suur t korpust on vaja? Korpuse koostamisel on üks nõue selle esinduslikkus: et materjal tõepoolest aitaks vastata uurimisküsimustele. Kauglugemise seisukohalt tuleb esinduslikkuse idee suhtes tähelepanelik olla. Kui teatud perioodi esindusliku kirjanduskorpuse koosseisu arvata ainult teosed, mis kuuluvad kirjanduslooga kaanonisse, on juba mööda mindud kauglugemise ühest alusprintsibist. Sellest lähtuvalt võiks korpusesse hõlmata nii palju teoseid, kui on vähegi kättesaadav. Samal ajal tuleb arvestada ka kättesaadavuse mõjuga korpuse sisule ja sedakaudu uurimistulemustele. Siit järeldub, et praegustes oludes on ideaalset digikorpust koostada raske, kui mitte võimatu. Korpuse koostamise põhimõtted mõjutavad otseselt selle põhjal tehtud uurimusi.

Oluline on teadvustada korpuse koostamisel tehtud valikuid ja piiranguid, isegi kui need on uurijast sõltumatud ja paratamatud, ja olla ka töö tulemusi tõlgendades tähelepanelik väidete tugevuse osas. Tulemuste üldistamine kogu kirjanduspildile pole õigustatud, kuid mööndusi arvesse võttes võib teatavaid suundumusi siiski välja tuua.

Uurimiskorpus sisaldas 35 eesti autori 50 ilukirjanduslikku proosateost (romaanid ja lühiproosa, välja jäeti laste- ja noortekirjandus), mis on ilmunud aastatel 1990–2010 (vt lisa). Korpuse kogumaht oli 2 609 062 sõna.

Analüüsi käik

Järjest enam on saadaval erinevaid pakette tekstianalüüsi teostamiseks, nagu veebi-põhine Voyant Tools, Sketch Engine, RStudio programmeerimiskeskonnas toimiv Stylo (vt Eder jt 2016), mida on kerge kasutada ja mis ei eelda kultuuriuurijalt programmeerimisoskust. Käesolevas uurimuses kasutati Voyant Toolsi korpuse üldiste arvuliste andmete tuvastamiseks (korpuse üldmaht, teoste mahud, proportsioonid jms). Tsitaatide eristamiseks siiski ükski valmistarkvara ei sobinud, selleks tuli sobiv kood ise RStudios kirjutada. Järgnevalt kirjeldan kokkuvõtlikult antud uurimuse tehnilise osa etappe.

1. Esialgne korpus sisaldas teoseid eri failivormingutes. Kõik failid viidi TXT-kujule ning teostelt eemaldati metaandmed (ilmumisaeg, kirjastus jms).
2. Kõik tekstid eraldati üksikuteks lauseteks (seda on kõige lihtsam teha lauselõpumärkide järgi). Sobivas mõõdus lähikontekstiks, milles tsitaati vaadelda, määrati üks lause.
3. Kõikide muude lausete seast eraldati avatud tsitaati sisaldavad laused, kasutades indikaatorina lõpujutumärkide esinemist lause sees (otsekõnes eelneb lõpujutumärkidele kas lauselõpumärk või koma, avatud tsitaadis mitte).
4. Väljafiltreeritud laused kanti edasiseks töötluks Exceli tabelisse. Seal tuli laused käsitsi puhastada samadele otsingukriteeriumidele vastanud lausetest, mis ei sisaldanud tsitaati, vaid hoopis jutumärgistatud nime, pealkirja vms.
5. Allesjäänud laused ehk kõik korpuses esinenud avatud tsitaatide kasutusjuhud jaotati tähendusjärgsesse kategooriasse. Selles etapis lisandus uuri-

musele kvalitatiivne lähenemisviis. Kui ühelauselisest kontekstist jäi tsitaadi mõistmiseks väheks, otsiti teose tekstifailis üles algkontekst.

6. Seejärel vaadeldi erinevate kategooriate esinemis- ja kokkulangemissagedust, tähendusloomet ning iseloomustati sedakaudu tsiteerimisvõtte funktsioone kirjanduskeeles.

Just viimase kahe etapi käigus kasvas kvantitatiivsest tekstianalüüsist kirjandus-teaduslik uurimus.

Tsitaatide rühmitamise eesmärk oli see, et kategooriad peegeldaksid tsiteerimisvõtte peamisi funktsioone tekstis ja et neist moodustuks mudel, mille abil saaks omakorda muudes, korpusesse mittekuulunud tekstides tsitaadikasutust analüüsida. Avatud tsitaat, mille ülesanne on sõnade algse tähenduse moonutamine või neile lisatähenduse andmine, allub, nagu paljud teisedki kirjanduskeeles nähtused, küllaltki raskelt jäikadele jaotustele ja ühestele tõlgendustele. Tähendusliku ambivalentsi säilitamine ongi kirjanduskeeles põhiomadusi. Seda silmas pidades peaksid ka selliste nähtuste alusel loodavad mudelid võimaldama tõlgendusruumi. Avatud tsitaatide rühmitamisel väljendus see selles, et üks tsitaat võis kuuluda üheaegselt mitmesse kategooriasse. Nii ei kahandata võtte loomupärast mitmetähenduslikkust.

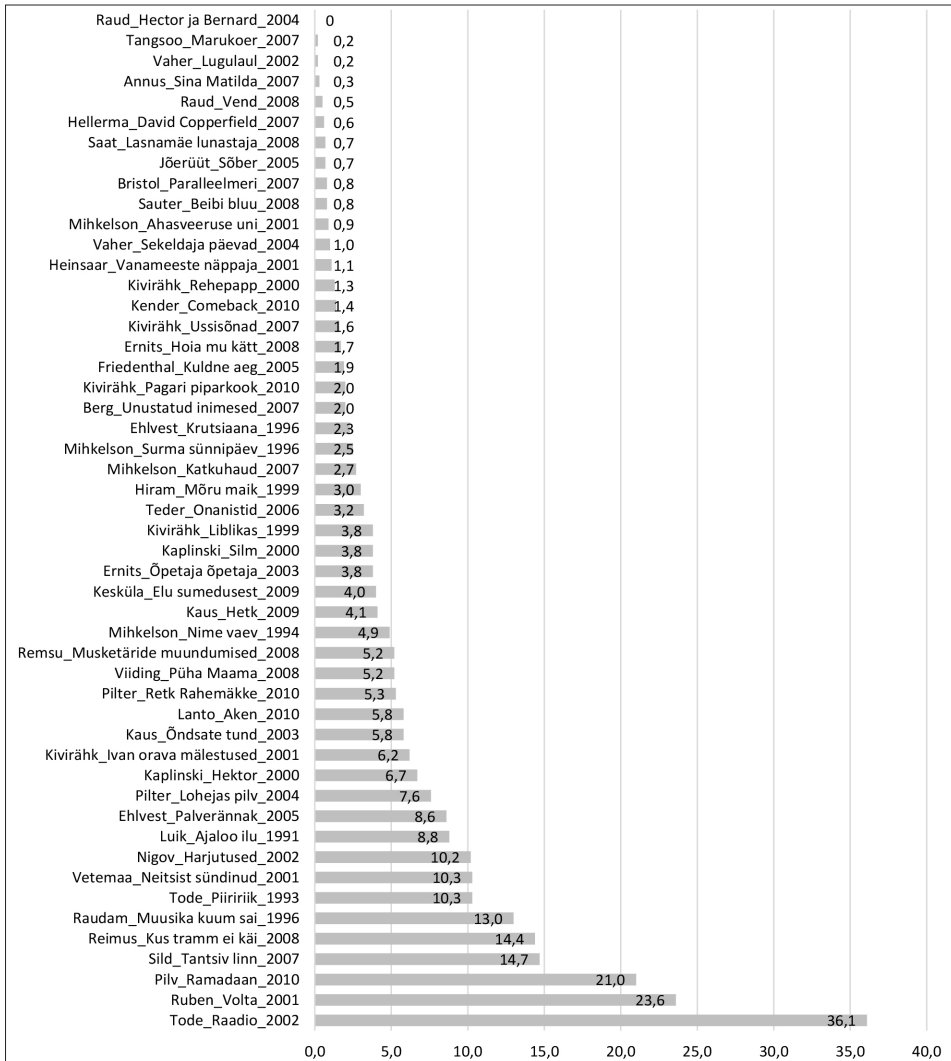
Uurimistulemused

Avatud tsitaate sisaldavaid lauseid oli korpuses kokku 1856. Teoste pikkusest tingituna arvutati tsitaatide esinemissageduse võrdlemiseks nende suhteline sagedus 10 000 sõna kohta (joonis 1). Kümme tsitaaditihedamat teost olid järgmised:

1. Emil Tode, „Raadio” (36,1 avatud tsitaati 10 000 sõna kohta)
2. Aarne Ruben, „Volta annab kaeblikku vilet” (23,6)
3. Aare Pily, „Ramadaan” (21,0)
4. Ivar Sild, „Tantsiv linn” (14,7)
5. Katrin Reimus, „Kus tramm ei käi” (14,4)
6. Toomas Raudam, „Muusika kuum sai” (13,0)
7. Emil Tode, „Piiririik” (10,3)
8. Enn Vetemaa, „Neitsist sündinud” (10,3)
9. Anton Nigov, „Harjutused” (10,2)
10. Viivi Luik, „Ajaloole ilu” (8,8)

On tähelepanuväärne, et arvulised tulemused hakkasid kvalitatiivset analüüsi suunama autoripõhistes tsiteerimismustrites. Esialgu polnud plaanis tsitaadikasutust uurida teoseti ega autoriti. Statistiliselt ilmnas aga, et eri autorid on võtet kasutanud ootamatult erineval määral.

Näiteks selgus, et Tõnu Õnnepalu kolmes teoses („Piiririik”, „Raadio”, „Harjutused”), mis moodustasid 9,6% korpuse kogumahust, esines 616 avatud tsitaati, mis on 33% kõigist leitud tsitaadidest. Seevastu skaala teises otsas paiknev Rein Raua „Hector ja Bernard” ei sisaldanud ühtegi avatud tsitaati.



Joonis 1. Avatud tsitaatide suhteline sagedus teoseti 10 000 sõna kohta.

Sellest tuleneb kaks olulist järeldust. Esiteks, antud korpusese toob pelgalt ühe autori loomingusse suure kõikumise, mis annab teavet konkreetse autori stiilist kui ajastu üldpildist. Teiseks ärgitabki tsitaadirohkus tõlgendama leitud autori loomingut ja stiili kontekstis. Olles seda asjaolu varem ise uurinud, on siinkirjutaja jõudnud järeldusele, et Önnepalu teostes on tegevustikku ja tegelaskonda minimaalselt, ent neid asendab intensiivne dialoog minajatustaja ja teiste diskursuste vahel. Kõrgendatud keeleline tundlikkus avaldub ohtras tsiteerimises ja eri stiilide (ironilises) matkimises ning selle kaudu väljendatakse sageli oma hoiakut tsitaadiga esindatud maailmavaate või teema suhtes (vt lähemalt Meriste 2016).

Avatud tsitaate kvalitatiivselt eritledes tuli esile kaheksa peamist omadust, mille põhjal jaotati tsitaadid kategooriatesse ja kirjeldati neid. Sedakaudu sai täidetud üks

uurimiseesmärkidest: avatud tsitaadi kui kirjanduskeelise võtte funktsioonide kirjeldamine. Siinses artiklis on kategooriate kirjeldused lühendatud kujul (pikemalt Meriste 2020: 38–50). Kategooriate kirjeldused tuginevad peale kvalitatiivse analüüsi statistilistele tähelepanekutele kategooriate kokkulangevusest ja esinemissagedusest, mis on siin esitatud pärast kategooriate kirjeldusi.

Järgnevalt kategooriatest lähemalt.

1. **Avaliku diskursuse** kategoorias kasutatakse tsiteerimisvõtet ametlik-poliitilise žargooni, meediakeele või rahvalike käibetõdede rõhutamiseks ja nende suhtes võõristustunde märkimiseks. Sageli kaasneb sellega iroonia või vastumeelsus, tsitaadi abil vastandatakse ja eristatakse kõneleva indiviidi ning kollektiivset, ametlikku, poliitilist mõtlemisviisi. Selle kategooria avatud tsitaadid paljastavad keele ideoloogilise laetuse, mis uuritud kirjanduses seostus enim totalitaarsele režiimile iseloomuliku keelekasutusega.

Olin valmis otsima vähetasuva, kuid „moraalselt puhta” töö ning elama ainult tema nimel. (Sild 2007: 86)

Kui tuhahunnikute ja söestunud korstnate keskelt saadad Saksa võimudele palvekirja, et keegi on olnud „kommunist puht teenistuse mõttes”, kuid hea pottsepana on teda vaja, siis pisiasjade üle enam ei tingita. (Mihkelson 2007: 179)

2. **Rolli** kategooria avatud tsitaatide all mõeldakse inimestele omistatud identiteete tähistavaid sõnu, näiteks hüüdnimed, tiitlid, identifitseerivad omadussõnad, sotsiaalsed määratlused jms. Ka rolli kategooriasse kuuluvad tsitaadid märkisid eelkõige vastumeelsust antud rolli või sellega kaasnevate ootuste suhtes.

Ja siis pidingi mina olema see, kes „ainsa mehena” Mõmmi korrale kutsub, läheb talle metsaserva järele ja viib kurva karu koju. (Kivirähk 2007: 146)

3. **Deiksise** kategooriasse kuuluvad iseseisva (kontekstist sõltumatu) tähendusega keelendid, näiteks asesõnad. Deiktikud märgivad tekstisisest või tekstist välja-poolse ulatuvat viitesuhet. Seda arvestades oli deiktikute kasutamine tsitaadina või selle osana üllatav leid: milleks rõhutada tsiteerimisvõttega midagi, mille tõlgendamiseks on igal juhul vaja tuge kontekstist? Selgus, et üheks oluliseks funktsiooniks on justnimelt tahe rõhutada, et välditakse otseütlemist ehk et luuakse teadlikult läbi-paistmatust.

Alevi poolt ei julgenud igatahes ükski hing nõmmele lilli korjama ega lambaid karjatama tulla, sest kumu veidrast king-olendist oli seal veel kümme korda jubedamaks paisutatud, kui asi tegelikult oli, nii et isegi Paap Mõrumeelest hoiti parema meelega kaugemale, sest tema ju „tollega” seal mässas. (Heinsaar 2001: 79)

Naiseiga koos istugu Valeri autosse ning neid viiakse otsemaid „sinna kuhu tarvis”. (Vetemaa 2001: 129)

4. *Pseudotsitaadiks* on siinkohal kategoriseeritud sellised avatud tsitaadid, mis on otsekui ütluse vormis kujutlused kellegi või millegi hoiakutest, olemusest, kuigi neid ütlusi pole tegelikult kõlanud. Sellest tulenevalt on pseudotsitaadid illustreerivad ning tihtilugu tugevalt stereotüüpsed. Pseudotsitaadiga on võimalik ütlusi omistada isegi kõnevõimetutele objektidele, nagu näoilme või ellusuhtumine.

Nad lobisesid ning naeratasid üksteisele ühtmoodi võluvalt selle tülgastava „jah, meil on raskeid aegu olnud, aga koos saame me kõigest üle”-ilmega. (Hiram 1999: 332)

Alles palju hiljem sain aru, et ründasin proua Bintat just selle end väikeseks inimeseks kuulutamise pärast, „keskmised-koju-tulevad”-eluvaate pärast. (Mihkelson 2007: 103–104)

Pseudotsitaadi abil saab üks tegelane esile manada kujutluse teise arvatavast reaktsioonist, mis näitab kujutatud tegelast selgelt stereotüüpses, isegi karikeerivas valguses. Teisisõnu peegelduvad selles pigem tsiteerija kujutlus ja rõhuasetus kui kujutlusobjekti tegelik olemus.

Aimasin juba ette tema vastuväiteid, mis järgnevad kohe, kui ma ütlen, et olen väsinud ja tahan magama minna – „aga ta on ju su õde”, „tema nii hoolitses sinu eest, kui sa väike olid”, „me oleme üks pere, me peame üksteist aitama”, „me ei tohi jääda võõraks”. (Kivirähk 2007: 147)

5. *Intertekstiks* liigitati avatud tsitaadid, mille lähikontekstis oli selge viide tsitaadi verbaalsele allikale või mis on äratuntavad kultuuriviited. Intertekstuaalse avatud tsitaadi üks dominantseid funktsioone on tsiteeritud diskursuse kohta arvamuse avaldamine ja/või kõneleja distantseerumine sellest. Tavalise otsekõnena see nõnda efektselt ei ilmneks, sest avatud tsitaat on justkui ootamatu tähendusreljeefi tekitav kiil teise tegelase diskursuse sees.

Ühed teadsid, et see on pikk blond, „ilus puhas eesti mees”, nagu üks naine väitis, väga šikilt riides ja ülimalt rafineeritud. (Berg 2007: 158)

6. *Alternatiivtähenduse* all peeti silmas avatud tsitaate, mis modifitseerivad sõnade neutraaltähendust. Peamiselt tähistas see eufemisme, ironiat või muul viisil kujundlikkust. Need on head näited kirjanduskeele ambivalentisusest ja mängulisusest, mis jäikade määratlustega analüüsile hästi ei allu. Tsitaadimärkidega rõhutatakse moondunud või lisatähenduse olemasolu, samal ajal ei ole see otsesõnu välja öeldud ning jääb tahes-tahtmata sõltuma lugeja tõlgendusest.

Sellepärast ma kadestangi vanu ja mõnikord, kui tulevik eriti rõhub, nagu nendel segastel päevadel, kui maailm tundub jälle vaikselt, aga nüri vältimatusega veerevat „õigusemõistmise”, ühesõnaga pigem sõja kui rahu poole, kadestan ma ka surnuid. (Nigov 2002: 23)

7. Metatähendusega avatud tsitaatide eesmärk on juhtida otseselt tähelepanu keelekasutusele. Selle abil rõhutatakse teadlikkust sõnade kaastähendustest, nende konnotatsioonidest, sõnade tavatähenduse küsimärgistamist. Enamasti kaasneb seda tüüpi tsitaadiga keelekasutuse otsesõnaline kommenteerimine lähi-kontekstis. Kuna üksik tsitaat esindab suuremat diskursust (nt poliitilist allkeelt), saab üksikute sõnade kaudu kujutada hoiakut kogu diskursuse suhtes.

Mida rohkem räägiti sõltumatusest ja demokraatiast, maailmakorrast ja maailmarahust, seda kummalisemaid tähendusvarjundeid omandas sõna „demokraatia”. (Mihkelson 2007: 319)

„Ema”, „kummipüksid”, „kintsud” ja „külmlilm” kõlavad tema omastarust-juba-vilunud-kõrvale nii süütult, nagu võiksid nad tähendada eriti tumedaid ja kahtlasi asju. (Luik 1991: 92)

8. Epiteedina kategoriseeriti avatud tsitaadina esitatud täiendeid ja öeldis-täiteid. Tsiteerimisvõtte abil tõstetakse rõhulisse positsiooni epiteedid, mis lause tavapärasest semantilisest struktuuris on teisejärgulised (st mitte põhisõnad). Sellisena tekib omamoodi autonomaasia efekt, mis on levinud lüroepikas ja ka eesti rahvaluules (Laugaste 1975: 192) – omadus muutub pärisnimega võrdväärseks identifitseerijaks ja võib pärisnime isegi asendada. Uurimiskorpuses esines see valdavalt negatiivse sildistamisena.

Kui Nõukogude armee asjus „kõlbmatuks”, „mittekuuluvaks” ja „väljalülitatuks” tunnistamine kõrvale jätta. (Teder 2006: 169)

Kas ma tõesti näen välja nii universaalselt „kohalik”, et mind on peetud Moskvast venelaseks, Prahast tšehhiks, Hollandist hollandlaseks, Londonist inglaseks? (Pilt 2010: 216)

Enim esines meta- ja alternatiivtähenduse kategooria tunnustega tsitaate (vastavalt 30,9% ja 27,7%). Kõige vähem oli deiksise ja pseudotsitaadi tunnustega näiteid (vastavalt 7% ja 6,8%). Rolli kategooria tuvastati 22,8%, intertekst 21,6%, avalik diskursus 16,9% ja epiteet 11,5% juhtudest.

Kokkulangevust kajastab tabel 1.

Tabel 1. Kahe kategooria koosinemised absoluut- ja suhtarvudena.

	Avalik	Roll	Deiksis	Pseudo-tsitaat	Intertekst	Alternatiiv	Meta	Epiteet
Avalik	314	69	11	22	32	87	68	40
Roll	69	423	25	5	52	101	79	28
Deiksis	11	25	130	3	20	5	41	9
Pseudotsitaat	22	5	3	127	24	11	17	6
Intertekst	32	52	20	24	401	21	105	33
Alternatiiv	87	101	5	11	21	514	42	52
Meta	68	79	41	17	105	42	573	35
Epiteet	40	28	9	6	33	52	35	214

	Avalik	Roll	Deiksis	Pseudo- tsitaat	Inter- tekst	Alter- natiiv	Meta	Epiteet
Avalik	100%	22%	4%	7%	10%	28%	22%	13%
Roll	16%	100%	6%	1%	12%	24%	19%	7%
Deiksis	8%	19%	100%	2%	15%	4%	32%	7%
Pseudotsitaat	17%	4%	2%	100%	19%	9%	13%	5%
Intertekst	8%	13%	5%	6%	100%	5%	26%	8%
Alternatiiv	17%	20%	1%	2%	4%	100%	8%	10%
Meta	12%	14%	7%	3%	18%	7%	100%	6%
Epiteet	19%	13%	4%	3%	15%	24%	16%	100%

Kokkuvõte

Artiklis uuriti avatud tsitaati, mis on teksti tähendusstruktuuri tugevalt mõjutav stilistiline võte. Seda tehti kauglugemise kontseptuaalses raamistikus, mis ei vaatle kirjandust eraldiseisvate üksikteoste kogumina, vaid tekstuaalsete eneseväljenduspraktikate väljana, kus ilmnevad teatavad mustrid. Sellist vaatepunkti võimaldavad rakendada digimeetodid. Kauglugemise kasuks otsustati uurimisel eelkõige seetõttu, et üksikuid teoseid lähilugedes oleks võimatu koguda uuritava nähtuse kohta nii palju kasutusnäiteid, et nende põhjal saaks teha üldistusjõulise analüüsi. Et kesken- duda detailile, tuleb see kontekstist välja filtreerida ja asetada kõrvuti suure hulga sama nähtuse esinemisjuhtudega mujal.

Selgus, et avatud tsitaat on vahend, mis loob ilukirjanduslikku lausesse justkui tähendusreljeefi: tsitaadina esile tõstetud tekstiosa on väliselt vaid väike detail, kuid selle abil tuuakse teksti lai seostevõrgustik, tühistatakse ühemõttelise tõlgendamise võimalus ja rõhutatakse kõrgendatud keelelist teadlikkust. Tsiteerimine võib sõna tähenduse muuta teisejärguliseks või isegi valeks – teisisõnu muuta selle tõeväärtust. Tsiteerimisvõttega saab midagi väita ja samal ajal end sellest distantseerida, näiteks selleks, et märkida ironiat. Üheks huvitavaimaks avatud tsitaadi tähenduskatego- oriaks osutus pseudotsitaat, mida siinkirjutaja teadmiste kohaselt ei ole varem teoree- tiliselt käsitletud. Pseudotsitaat, otseku verbaalne illustratsioon, asendab tavapäras- t seletust või kirjeldust. See lähtub tsiteerija kujutlusest, mida keegi või miski võiks öelda, iseloomustades sedakaudu nii illustratsiooni objekti kui ka suhtumist sellesse.

Jutumärkide kui diskursuste selgete eristajate ülesanne võib olla tsitaadiga esi- datud inimese, maailmavaate, valdkonna vms võõristamine ja varjatud või ilmse seisukoha väljendamine. Näidetest ilmnes, et selle üks oluline ülesanne taasiseseis- vumisjärgses eesti kirjanduses on ühtse väärtussüsteemi eitamine ja tõe relatiivseks muutmine, milles võib näha reaktsiooni eelnenud 50 aastale, nõukogude „ühe tõe” režiimile. Keel annab edasi kategooriaid, mis struktureerivad ka arusaamu maa- ilmast, ning avatud tsitaat võimaldab protestida vaikimisi kinnistunud standardite vastu: olgu nendeks sõnaga omistatud identiteet ja sellega kaasnevad ootused või kollektiivsed käibetõed. Sellisena on avatud tsitaat selgelt seotud subjektikeskse vaate esiletõusuga.

Kvalitatiivset ja kvantitatiivset lähenemist kombineeriva analüüsi tulemusena sai kaardistada ühe stilistilise võtte olemust. Ühtpidi saab sel viisil täiendada seniseid arusaamu uurimismaterjali moodustanud kirjandusteostest. Kuid lisaks saaks neid teadmisi kasutada uue uurimismaterjali juures, mis ei pea olema tingimata ilukirjanduslik. Näiteks ajakirjanduslikus kontekstis kasutatakse avatud tsitaati selliste suhtumiste varjamiseks, mida otsesõnu välja ei soovita öelda, või väite infostruktuuri hägustamiseks.

Kvantitatiivsete ja kvalitatiivsete meetodite kombineerimine annab digihumaniaria uurimuste alusoskuse: kvalitatiivne hüpotees tuleb kohandada kvantitatiivse lahenduskäiguga ja seejärel jõuda omakorda kvalitatiivsete järeldesteni.

Sellega seoses loodan, et arvutuslikud meetodid saavad peagi humanitaarteaduste õppeprogrammi loomulikuks osaks, kuid veel enam näen vajadust valdkonnaülese koostöö järele. Käesolev uurimus on üksik katseprojekt, mis ei esinda efektiivset ja jätkusuutlikku viisi selles valdkonnas edasi liikumiseks. Vaja on luua igapäevaselt koos töötavaid, mõtlemaid, loovaid uurimisgrupe, kuhu kuuluvad keele-, arvuti- ja kirjandusteadlased ning teised asjakohased eksperdid. See võimaldaks keskenduda mitme formaalse näitaja paralleelsele uurimisele ja näha muutujatevahelist suhet – näiteks antud uurimuse põhjal, kas rohke tsitaadikasutus on korrelatsioonis autori sõnavara rikkalikkuse, lauseehituse keerukuse või muu taolise näitajaga. Usun, et tähendust ei mõjuta niivõrd ühe elemendi esinemisviis, kuivõrd koosinevate elementide omavaheline suhe ning nende suhete muutumine ajateljel. Digitaalne areng on andnud selliste dünaamiliste tähendusvõrgustike uurimiseks vajalikud tehnilised vahendid, nende kasutusvõimaluste avastamine on uurijate teha.

VEEBIVARAD

Distant Reading for European Literary History. <https://www.distant-reading.net>

Sketch Engine. <http://www.sketchengine.eu>

Voyant Tools. <http://voyant-tools.org>

KIRJANDUS

AutÕS = Autoriõiguse seadus. Vastu võetud 11. XI 1992, jõustumine 12. XII 1992.

Bakhtin, Mihhail 1982. Discourse in the novel. – *The Dialogic Imagination: Four Essays* by M. M. Bakhtin. Toim Michael Holquist. Austin: University of Texas Press, lk 244–245.

Berg, Maimu 2007. Unustatud inimesed. Tallinn: Tuum.

Brendel, Elke; Meibauer, Jörg; Steinbach, Markus 2011. Exploring the meaning of quotation. – *Understanding Quotation*. (Mouton Series in Pragmatics 7.) Toim E. Brendel, J. Meibauer, M. Steinbach. Berlin: De Gruyter, lk 1–33. <https://doi.org/10.1515/9783110240085.1>

Brik, Ossip 2014. Nn „Vormimeetod”. – Kirjandus kui selline. Valik vene vormikoolkonna tekste. (Gigantum Humeris.) Koost ja toim Märt Väljataga. Tallinn: Tallinna Ülikooli Kirjastus, lk 46–47.

Capone, Alessandro 2013. The pragmatics of quotation, explicatures and modularity of mind. – *Pragmatics and Society*, kd 3, lk 259–284. <https://doi.org/10.1075/ps.4.3.01cap>

- Eder, Maciej; Rybicki, Jan; Kestemont, Mike 2016.** Stylometry with R: a package for computational text analysis. – *The R Journal*, kd 8, nr 1, lk 107–121. <https://journal.r-project.org/archive/2016/RJ-2016-007/index.html>
- Genette, Gérard 1992.** *Palimpsestes: la littérature au second degré*. Paris: Seuil.
- Heinsaar, Mehis 2001.** *Vanameeste näppaja*. Tallinn: Tuum.
- Hiram 1999.** *Mõru maik*. Tallinn: Tuum.
- Kirsch, Adam 2014.** Technology is taking over English departments: The false promise of the digital humanities. – *The New Republic* 2. V.
- Kivirähk, Andrus 2007.** *Mees, kes teadis ussisõnu*. Tallinn: EKSA.
- Kristeva, Julia 1980.** Word, dialogue, and novel. – J. Kristeva, *Desire in Language: A Semiotic Approach to Literature and Art*. Toim Leon S. Roudiez. New York: Columbia University Press, lk 64–91.
- Kultuuriministeerium 2018** = „Kultuuripärandi digiteerimine 2018–2023” tegevuskava. Tallinn: Kultuuriministeerium. <https://www.kul.ee/kultuurivaartused-ja-digitaalne-kultuuriparand/digitaalne-kultuuriparand/kultuuriparandi> (15. II 2021).
- Laugaste, Eduard 1975.** *Eesti rahvaluule*. Tallinn: Valgus.
- Lotman, Juri 1991.** Kunst modelleerivate süsteemide reas. – J. Lotman, *Kultuurisemiootika. Tekst – kirjandus – kultuur*. Tlk Pärt Lias, Inta Soms, Rein Veidemann. Tallinn: Olion, lk 8–31.
- Luik, Viivi 1991.** *Ajaloo ilu*. Tallinn: Eesti Raamat.
- Meriste, Maarja Helena 2016.** *Võõras kõne Tõnu Õnnepalu „Hinnas”*. Bakalaureusetöö. Tartu Ülikool, kirjanduse ja teatriteaduse osakond.
- Meriste, Maarja Helena 2020.** *Eesti 1990.–2010. aastate kirjanduse kauglugemise katse avatud tsitaatide kasutusmustrite näitel*. Magistritöö. Tallinna Ülikool, humanitaarteaduste instituut.
- Mihkelev, Anneli 2005.** *Vihjamise poeetika*. (Collegium litterarum 19.) Tallinn: Underi ja Tuglase Kirjanduskeskus.
- Mihkelson, Ene 2007.** *Katkuhaud*. Tallinn: Varrak.
- Moretti, Franco 2013.** *Distant Reading*. London–New York: Verso.
- Nigov, Anton [= Tõnu Õnnepalu] 2002.** *Harjutused*. Tallinn: EKSA.
- Pilv, Aare 2010.** *Ramadaan. Vahtsõliinah, Lillekülās, in Matiano, a Roma*. Tallinn: Tuum.
- Piper, Andrew 2018.** Are novels getting easier to read? – *TXTLAB* 10. I. <https://txtlab.org/2018/01/are-novels-getting-easier-to-read> (28. VI 2021).
- Pruul, Kajar 2000.** *Vana sajandi uued luuletajad*. – *Varjatud ilus haigus. Valik sajandilõpu eesti luuletajaid*. Koost K. Pruul. Tartu: Eesti Kostabi Selts, lk 175–207.
- Puik, Katrin 2009.** *Iroonia Heiti Talviku ja Betti Alveri luules*. (Dissertationes litterarum et contemplationis comparativae Universitatis Tartuensis 8.) Tartu: Tartu Ülikooli Kirjastus.
- Recanati, François 2001.** Open quotation. – *Mind*, kd 110, nr 439, lk 637–687. <https://doi.org/10.1093/mind/110.439.637>
- Schöch, Christof; Erjavec, Tomaž; Patras, Roxana; Santos, Diana 2021.** *Creating the European Literary Text Collection (ELTeC): Challenges and perspectives*. – *Modern Languages Open* [ilmumas].
- Sild, Ivar 2007.** *Tantsiv linn*. Tallinn: Tuum.

- Sobchuk, Oleg 2018.** Charting Artistic Evolution: an Essay in Theory. (Dissertationes litterarum et contemplationis comparativae Universitatis Tartuensis 18.) Tartu: Tartu Ülikooli Kirjastus.
- Tangherlini, Timothy 2019.** Macroscopic Methods for Culture Analytics. Loengukursus. Tartu Ülikool, digihumanitaaria ja infoühiskonna keskus, 3.–13. IX.
- Teder, Tarmo 2006.** Onanistid. Tallinn: Tuum.
- van de Ven, Inge 2017.** Creative reading in the information age: Paradoxes of close and distant reading. – *The Journal of Creative Behavior*, kd 53, nr 1. <https://doi.org/10.1002/jocb.186>
- Velsker, Mart 2001.** Kirjandus ja ühiskond. – Epp Annus, Luule Epner, Ants Järv, Sirje Olesk, Ele Süvalep, M. Velsker, Eesti kirjanduslugu. Tallinn: Koolibri, lk 608–623.
- Vetemaa, Enn 2001.** Neitsist sündinud. Tallinn: Tänapäev.
- Viiding, Elo 2008.** Püha Maama. Tallinn: Tuum.
- Wilkens, Matthew 2012.** Canons, close reading, and the evolution of method. – *Debates in the Digital Humanities*. Toim Matthew K. Gold. Minneapolis–London: University of Minnesota Press, lk 249–258. <https://doi.org/10.5749/minnesota/9780816677948.003.0026>

LISA. TEOSED UURIMISKORPUSES

- Annus, Epp 2007. Sina, Matilda. Tallinn: Tuum.
- Berg, Maimu 2007. Unustatud inimesed. Tallinn: Tuum.
- Bristol, Piret 2007. Paralleelmeri. Tallinn: Tuum.
- Ehlvest, Jüri 1996. Krutsiaania. Tallinn: Tuum.
- Ehlvest, Jüri 2005. Palverännak. Tallinn: Tuum.
- Ernits, Marje 2003. Õpetaja, õpetaja. Tallinn: Tuum.
- Ernits, Marje 2008. Hoia mu kätt. Tallinn: Tuum.
- Friedenthal, Meelis 2005. Kuldne aeg. Tallinn: Tuum.
- Heinsaar, Mehis 2001. Vanameeste näppaja. Tallinn: Tuum.
- Hellerma, Kärt 2007. Ma armastasin David Copperfieldi. Tallinn: Tuum.
- Hiram 1999. Mõru maik. Tallinn: Tuum.
- Jõerüüt, Jaak 2005. Sõber. Tallinn: Tuum.
- Kaplinski, Jaan 2000. Silm. Tallinn: Tänapäev.
- Kaplinski, Jaan 2000. Hektor. Tallinn: Tänapäev.
- Kaus, Jan 2003. Õndsate tund. Tallinn: Tuum.
- Kaus, Jan 2009. Hetk. Tallinn: Tuum.
- Kender, Kaur 2010. Comeback. Tallinn: EKSA.
- Kesküla, Kalev 2009. Elu sumedusest. Tallinn: Tuum.
- Kivirähk, Andrus 1999. Liblikas. Tallinn: Tuum.
- Kivirähk, Andrus 2000. Rehepapp, ehk, November. Tallinn: Varrak.
- Kivirähk, Andrus 2001. Ivan Orava mälestused, ehk, Minevik kui helesinised mäed. Tallinn: Varrak.
- Kivirähk, Andrus 2007. Mees, kes teadis ussisõnu. Tallinn: EKSA.
- Kivirähk, Andrus 2010. Pagari piparkook: Kivirähk XX sajandil. Tallinn: EKSA.
- Lanto, Ari 2010. Aken. Tallinn: Eesti Raamat.

- Luik, Viivi 1991. Ajaloo ilu. Tallinn: Eesti Raamat.
- Mihkelson, Ene 1994. Nime vaev. Tartu: Ilmamaa.
- Mihkelson, Ene 1996. Surma sünnipäev. Tallinn: Tuum.
- Mihkelson, Ene 2001. Ahasveeruse uni. Tallinn: Tuum.
- Mihkelson, Ene 2007. Katkuhaud. Tallinn: Varrak.
- Nigov, Anton 2002. Harjutused. Tallinn: EKSA.
- Pilter, Lauri 2004. Lohejas pilv. Tallinn: Tuum.
- Pilter, Lauri 2010. Retk Rahemäkke. Tallinn: Tuum.
- Pilv, Aare 2010. Ramadaan. Vahtsõlinah, Lillekülas, in Matiano, a Roma. Tallinn: Tuum.
- Raud, Rein 2004. Hector ja Bernard. Tallinn: Tuum.
- Raud, Rein 2008. Vend. Tallinn: Tuum.
- Raudam, Toomas 1996. Muusika kuum sai. Tallinn: Tuum.
- Reimus, Katrin 2008. Kus tramm ei käi. Tallinn: Tuum.
- Remsu, Olev 2008. Musketäride muundumised. Tallinn: EKSA.
- Ruben, Aarne 2001. Volta annab kaeblikku vilet. Tallinn: Tänapäev.
- Saat, Mari 2008. Lasnamäe lunastaja. Tallinn: Tuum.
- Sauter, Peeter 2008. Beibi bluu. Tallinn: Tuum.
- Sild, Ivar 2007. Tantsiv linn. Tallinn: Tuum.
- Tangsoo, Jaan 2007. Marukoer. Tallinn: Tuum.
- Teder, Tarmo 2006. Onanistid. Tallinn: Tuum.
- Tode, Emil 1993. Piiririik. Tallinn: Tuum.
- Tode, Emil 2002. Raadio. Tallinn: EKSA.
- Vaher, Berk 2002. Lugulaul. Tallinn: Tuum.
- Vaher, Berk 2004. Sekeldaja päevad. Tallinn: Tuum.
- Vetemaa, Enn 2001. Neitsist sündinud. Tallinn: Tänapäev.
- Viiding, Elo 2008. Püha Maama. Tallinn: Tuum.

Maarja Helena Meriste (sünd 1992), MA, ajakirja Värsk Rõhk toimetaja (Vanemuise 19, 51003 Tartu), maarjahelena.meriste@va.ee

Distant reading of literary language: Developing a methodology using quotations as an example

Keywords: digital humanities, distant reading, quotations, literary style, formalism

This paper aims to take advantage of modern digital methods to address research topics that, until now, have mostly been subjects of assumption and interpretation based on rather few examples.

More specifically, the purpose of this article is to explore some usage patterns of open quotation as a literary device and its impact on meaning. Open quotation, as defined by the French philosopher François Recanati, is a type of quotation where

a string of words in an utterance is marked with quotation marks to imply the presence of multiple or deformed meanings.

Alongside with other stylistic devices like parody, irony and pastiche, open quotation is regarded as a distinctive feature of postmodern literary language, yet it has not been empirically studied. One of the reasons may be the fact that it is difficult to obtain a systematic view of such textual features by means of traditional, i.e., close reading. However, the emergence of digital methods in humanities provide better access to such research problems. In literary studies, application of computational methods is associated with Franco Moretti's approach of distant reading. However, distant reading does not offer a specific method to solve complex research problems in literary studies. Therefore, the second purpose of this study is to create a suitable methodology to best answer quotation-related research questions.

In studies of Estonian culture, the emergence of postmodern aesthetics has been associated with the collapse of the Soviet system at the end of the 1980s. Therefore, the text corpus of this study consists of 50 Estonian novels published in 1990–2010. All cases of open quotation were extracted and divided into 8 categories based on qualitative features, and then explored in detail.

Implying metalinguistic awareness was identified as the most important function of open quotation. Most commonly, it was used to reveal Soviet era related totalitarian ideologies, judgments and values inherent in language. Another interesting type of quotation previously undescribed in theoretical literature was distinguished: the category of pseudo-quotations aiming to illustrate both animate and inanimate objects by attributing to them an imaginary verbal act. In general, the use of open quotations reflects the decay of a monolithic value system by undermining the possibility of there being only one correct interpretation.

Regarding methodology, computational methods proved to be effective in establishing high-level stylistic patterns. However, the patterns alone remained inconclusive, requiring qualitative analysis to reach meaningful conclusions. This suggests a fruitful way to approach complex research questions in humanities: first, a qualitative hypothesis is transformed into a specific quantitative method; secondly, the obtained results are interpreted qualitatively.

Maarja Helena Meriste (b. 1992), MA, editor of literary magazine *Värske Rõhk* (Vanemuise 19, 51003 Tartu), maarihelena.meriste@va.ee