

Lähteet

- AIKHENVALD, ALEXANDRA 2010: *Imperatives and commands*. Oxford: Oxford University Press.
- BLUM-KULKA, SHOSHANA – HOUSE, JULIANE – KASPER, GABRIELE (toim.) 1989: *Cross-cultural pragmatics. Requests and apologies*. Norwood, NJ: Ablex Publishing Corporation.
- CRAVEN, ALEXANDRA – POTTER, JONATHAN 2010: Directives. Entitlement and contingency in action. – *Discourse Studies* 12 s. 419–442.
- FOX, BARBARA 2007: Principles shaping grammatical practices. An exploration. – *Discourse studies* 9 s. 299–318.
- GOODWIN, MARJORIE 1990: *He-said-she-said. Talk as social organization among black children*. Bloomington, IN: Indiana University Press.
- KÄRKKÄINEN, ELISE – KEISANEN, TIINA 2012: Linguistic and embodied formats for making (concrete) offers. – *Discourse Studies* 14 s. 587–611.
- LAURANTO, YRJÖ 2015: *Direktiivisyyden rajoja. Suomen kielen vaihtokauppasyntaksia*. Helsinki: Helsingin yliopisto.
- ROSSI, GIOVANNI 2012: Bilateral and unilateral requests. The use of imperatives and *Mi X?* interrogatives in Italian. – *Discourse Processes* 49 s. 426–458.
- 2014: When do people not use language to make requests? – Paul Drew & Elizabeth Couper-Kuhlen (toim.), *Requesting in social interaction* s. 303–334. Amsterdam: John Benjamins.
- 2015: *The request system in Italian interaction*. Nijmegen: Radboud University.
- SORJONEN, MARJA-LEENA 2001: Lääkäriin ohjeet. – Marja-Leena Sorjonen, Anssi Peräkylä & Kari Eskola (toim.), *Keskustelu lääkärin vastaanotolla* s. 89–111. Tampere: Vastapaino.
- STEVANOVIC, MELISA – PERÄKYLÄ, ANSSI 2012: Deontic authority in interaction. The right to announce, propose and decide. – *Research on Language and Social Interaction* 45 s. 297–321.
- STEVANOVIC, MELISA – SVENNEVIG, JAN 2015: Introduction. Epistemics and deontics in conversational directives. – *Journal of Pragmatics* 78 s. 1–6.
- VINE, BERNADETTE 2004: *Getting things done at work. The discourse of power in workplace interaction*. Amsterdam: John Benjamins.
- VISK = HAKULINEN, AULI – VILKUNA, MARIA – KORHONEN, RIITTA – KOIVISTO, VESA – HEINONEN, TARJA RIITTA – ALHO, IRJA 2004: *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura. <http://scripta.kotus.fi/visk>.
- WOOTTON, ANTHONY J. 1997: *Interaction and the development of mind*. Cambridge: Cambridge University Press.

Korpuslingvistiikka kohtaa historiantutkimuksen

Anthony McEnery & Helen Baker: *Corpus linguistics and 17th-century prostitution. Computational linguistics and history. Research in Corpus and Discourse*. Lontoo: Bloomsbury 2017. 257 s. ISBN 978-1-4725-0609-2.

Muistiorganisaatiot, kuten kirjastot, arkistot ja museot, ovat viime aikoina niin Suomessa kuin ulkomailla digitoineet kokoelmiaan ja näin avanneet ne uudella tavalla tutkimuksen aineistoiksi. Koska monet digitoituista kokoelmista ovat pääsään-

töisesti tekstiaineistoja, on samalla syntynyt kysyntää korpuslingvistiikan ja laskennallisen kielentutkimuksen menetelmille – myös kielentutkimuksen ulkopuolella. Muistiorganisaatioiden ja historiatieteiden luonnollisen yhteyden vuoksi tämä uusi monitieteinen kenttä on avautunut erityisesti historian tutkijoiden ja lingvistien väliin. Käsillä oleva teos sijoittuu tähän kontekstiin. Teoksen tavoitteena on tutkia prostituution kielellisiä representaatioita 1600-luvun Englannissa hyödyntämällä suurta historiallista korpusaineistoa, niin sanottua EEBO-korpusta (*Early English Books Online*). Teoksen korpuslingvistinen tutkimusote on eksploratiivinen: korpusta tutkitaan ilman suurempia ennakkoletuksia, ja tunnuslukujen, kuten sanojen ja kollokaatioiden frekvenssien, avulla yritetään tunnistaa jokin kiinnostava ilmiö. Teoksessa pyritään erityisesti kollokaatioiden analyysillä luomaan yleiskuvaa prostituutioon viittaavien sanojen käytöstä ja käytön muutoksesta tutkittuna ajanjaksona. Kirjoittajista McEnery on korpuslingvistiikan professori ja Baker historian tutkijatohtori Lancasterin yliopistossa.

Monitieteisen tutkimuksen vaikeus

Aiheen valinta kytkee teoksen sosiaalishistorian ja feministisen historian tutkimuksen perinteisiin. Köyhiin, yhteiskunnan marginalisoimiin naisiin fokusoiminen on osa feministisen historian tutkimuksen paradigmat, jossa pyritään tekemään näkyväksi haavoittuvien ja näkymättömien ihmisryhmien historioita. Yhteiskunnan alimpiin kerroksiin kuuluvat henkilöt tulevat historiallisissa lähteissä yleensä esiin vain joutuessaan tekemisiin virallisia dokumentteja tuottavan virkakoneiston kanssa. Köyhien naisten tapauksissa tämä on tarkoittanut tyypillisesti sitä, että heitä on syytetty prostituutiosta. Tästä syystä prostituution historian tutkimus on nojanut raskaasti juuri oikeuden pöytäkirjoihin. Tähän tutkimustraditioon McEnery ja

Baker pyrkivät tuomaan uuden näkökulman tutkimalla prostituoitujen representaatioita laajassa, monenlaisia tekstejä sisältävässä EEBO-korpuksessa.

Teosta voidaan kuitenkin lukea myös osallistumisena keskusteluun monitieteisen tutkimuksen periaatteista, erityisesti sellaisen, jossa yhdistyvät perinteinen humanistinen tutkimus ja määrällinen, tietokoneavusteinen tutkimus. Teoksessa edellistä edustaa laadullinen historian tutkimus ja jälkimmäistä korpuslingvistiikka. Monitieteistä otetta ilmentää se, että teoksen johdannossa tekijät antavat lukuisia varoittavia esimerkkejä siitä, millä tavoin humanistien ei pitäisi hyödyntää suuria aineistoja ja laskennallisia menetelmiä. Tämä huomioon ottaen kokonaisvaikutelma teoksesta onkin jonkinlainen mallitutkimus tai tyyliharjoitus historioitsijan ja korpuslingvistin yhteistyöstä. Teoksen kannalta se on sekä hyvä että huono asia. Yhtäältä teos onnistuu ansiokkaasti havainnollistamaan korpuslingvististen menetelmien mahdollisuuksia aiheesta mitään tietämättömälle humanistille. Tällaiselle teokselle on varmasti kysyntää. Toisaalta teoksessa ei juuri pohdita sitä, kuinka hyvin esitellyt menetelmät istuvat nimenomaan teoksessa toteutetun tutkimuksen kysymyksenasetteluun.

Teoksen alkusanoissa tekijät painottavat sitä, että teoksen ensisijaisena tavoitteena on osoittaa, miten hedelmällistä läheinen yhteistyö kahden tieteenalan välillä voi olla, mikäli se on tasa-arvoista. Tekijöille tämä on tarkoittanut sitä, että kumpikaan tieteenala ei toimi toiselle ikään kuin teknisenä apuna eikä kumpaakaan alisteta toisen alan tiedonintressille. Ajatus on kauris, mutta se tuntuu sivuuttavan sen tosiasian, että tutkimuskysymyksillä on aina sidos tietynlaiseen tiedonintressiin. Tästä johtuukin McEneryn ja Bakerin tutkimuksen perusongelma. Nimittäin välttääkseen tilanteen, jossa jommankumman tutkimusalan tiedonintressi muodostuisi hallitsevaksi, tekijät ovat päätyneet ratkaisuun,

jossa he käytännössä työskentelevät omilla tahoillaan ja tekevät havaintonsa toisistaan riippumatta. Seurauksena on, että osatutkimukset yhdistävää tutkimuskysymystä tai -kysymysten joukkoa ei ole syntynyt. Tätä voi pitää jonkinasteisena epäonnistumisena, mikäli tarkoitus oli osoittaa läheisen yhteistyön hedelmällisyys. Monitieteisen tutkimuksen ongelmat eivät toki missään nimessä ole yksinkertaisia, eikä varmasti kaikkia tällaisen tutkimuksen ideoita ole käytännössä mahdollista saavuttaa.

Teoksen jäsenitys

Teoksen luvut 2–3 sisältävät katsauksen prostituution historiaan ja siitä tehtyyn aiempaan tutkimukseen. Luvussa 2 tarkastellaan prostituutiota sosiaalishistoriallisesta näkökulmasta kulttuurisena, poliittisena ja taloudellisena ilmiönä. Myös kieltä ja kielenkäyttöä sivutaan kirjoittajien esitellessä historian tutkimuksen niin sanotun kielellisen käänteen metodologisia suuntauksia, erityisesti saksalaisen Reinhart Koselleckin käsitehistoriaa ja brittiläisen aatehistorioitsijan Quentin Skinnerin austinilaisesta puheaktiteoriasta ammentavaa traditiota¹. Osuus toimii hyvin lyhyenä kirjallisuuskatsauksena siihen, miten historian tutkimuksen piirissä on keskusteltu kielestä. Luvun pääpaino on kuitenkin siinä, minkälaiset olosuhteet ja muut tekijät saivat naiset ryhtymään prostituoiduiksi ja minkälaisia prostituoitujen elämäntavat olivat.

Luvussa 3 kartoitetaan prostituution näkyvyyttä ja representaatioita 1600-luvun aikalaisteksteissä. Korpuslingvistiselle analyysille annetaan joitain täkyjä, kuten kysymys siitä, missä määrin transaktionaalisen ja transgressiivisen seksin merkitykset painottuvat prostituoituihin viittaavissa sa-

noissa. Transaktionaalisen seksin käsite pitää sisällään seksin käyttämisen vaihdannan ja liiketoiminnan muotona. Transgressiivisessä seksissä taas on kyse yhteisön seksuaalisia normeja rikkovasta käyttäytymisestä. 1600-luvun seksuaalisen moraalin historian kannalta kysymys on relevantti ja liittyy läheisesti siihen, minkälaisia vapauksia naisille on annettu säädellä omaa seksuaalista käyttäytymistään. Samaan tematiikkaan liittyy myös kysymys siitä, voidaanko prostituutio 1600-luvulla nähdä myös voimaantuneiden naisten yrityksenä toimia itsenäisesti erittäin vaikeissa taloudellisissa oloissa vai onko kyse ollut yksisuuntaisesta alistavasta suhteesta, jossa hyvin toimeentulevat miehet ovat marginalisoineet ja alistaneet alemman sosiaaliluokan naisia. Näitä muutamaa luonnosmaista kysymyksenasettelua lukuun ottamatta kielentutkimus ei ole luvuissa 2 ja 3 juuri läsnä, joten kokonaisuuden kannalta luvut toimivat – lukijan mielenkiinnosta riippuen – joko teoksen varsinaisena sisältönä tai taustoituksena myöhempien luvujen korpuslingvistisille havainnoille.

Korpuslingvistinen analyysi on toteutettu luvuissa 4–6. Osatutkimuksen tavoitteena on analysoida prostituoituihin viittaavien sanojen käytön ja merkityksen vaihtelua EEBO-korpuksesta – niiltä osin kuin se on edellä mainittujen frekvenssien ja kollokaattien avulla mahdollista. Aluksi luvussa 4 kuvataan perusteellisesti se prosessi, jolla prostituoituihin viittaavien sanojen joukko on koostettu. Tutkittavan ilmiön kannalta oleellisen sanaston valikoiminen on tyypillinen haaste sekä merkityslähtöisessä historiallisessa korpuslingvistiikassa että korpusavusteisessa historian tutkimuksessa, erityisesti mikäli aihe on kulttuurisesti ja sosiaalisesti laatautunut. Tämä johtuu siitä, että erilaisen parafraasin, eufemismien ja dysfemismien vuoksi relevantteja ilmauksia on paljon ja niiden valikoima on historiallisesti dynaaminen. Tekijät ovat pyrkineet lähestymään ongelmaa systemaattisesti

1. Quentin Skinnerin keskeiset metodologiset interventiot on koottu kokoelmaan Tully toim. 1988. Reinhart Koselleckin käsitehistoriaan toimii johdantona Richter (1995) tai Marjanen (2018).

yhdistämällä kolme yleisintä menetelmää ongelman ratkaisemiseksi. Sopivia sanoja on etsitty sanakirjoista, aikalaisteksteihin syventyneen historian tutkijan kielitajusta sekä korpuksista. Lopullinen sanojen valikoima on syntynyt näiden kolmen lähteen tarjoamien tulosten yhdistelmänä. Prosessi eteni seuraavasti: Ensin sanakirjoista haettiin joukko sanaehdokkaista, jotka validoitiin historian tutkijan kielellisen intuition ja EEBO-korpuksen frekvenssien avulla. Jotta selviäisi, sisältävätkö aikalaistekstit sellaisia ehdokkaista, jotka ovat jääneet huomaamatta sekä sanakirjojen toimittajilta että historioitsijoilta, korpuksista etsittiin sanoja, jotka esiintyvät samankaltaisissa leksikaalisissa ympäristöissä kuin prostituoituihin tunnetusti viittaavat sanat. Tällä menetelmällä saatiin selville esimerkiksi se, että *miss*-substantiivin monikko isolla alkukirjaimella kirjoitettuna (*Misses*) viittasi korpuksessa tyyppillisesti rakastajattariin, avioliiton ulkopuolisiin suhteisiin ja suoranaiseen seksityöhön.

Luvuissa 5–6 on esitelty tutkimuksessa käytetyt korpuslingvistiset analyysit sekä niistä saadut tulokset. Varsinaiset menetelmät ovat prostituoituihin viittavien sanojen frekvenssien seuraaminen läpi vuosisadan sekä näiden sanojen kollokaatioiden kartoitus ja tulkinta. Kaikki analyysit on tehty käyttäen CQPWeb-korpusohjelmiston työkaluja. Ohjelmisto ei eroa paljoakaan kotoisesta Kielipankin Korpista: se on rakennettu hakukoneen ympärille, ja se näyttää käyttäjälle hakutulosten konkordanssirivit ja tarjoaa hauista joitain valmiiksi koostettuja tietoja, esimerkiksi niiden tilastollisen merkitsevyyden perusteella tunnistetut kollokaatit.

Kollokaatioiden analyysi

Sanaston yleisiä trendejä kartoitetaan teoksessa ryhmittämällä kimpuiksi sellaiset sanat, joiden frekvenssien ajalliset vaihtelut korreloivat voimakkaasti toistensa kanssa. Näin on saatu yleiskuva siitä, mitkä sa-

nat olivat pääasiallisesti käytössä tutkittavan aikakauden eri vaiheissa. Merkitysten muutosten käsittelyssä ensisijaisena tutkimusmetodinä on kollokaatioiden analyysi. Menetelmä on korpuslingvistiikassa hyvin suosittu, sillä kollokaatioiden etsiminen ja käsittely on varsin helppoa (ks. esim. Tognini-Bonelli 2001; Halliday, Teubert, Yallop & Čermáková 2004). Kollokaatiolla tarkoitetaan yleisesti toistuvia sanojen yhteiseiintymiä, mutta tarkempia määrittelyjä on mahdollista tehdä useita esimerkiksi sen mukaan, onko sanojen sijaittava välittömästi toistensa vieressä, montako sanaa kollokaatioon voi kuulua ja mitä toistuvuus tarkoittaa. Esimerkiksi Jantunen (2009: 358) määrittelee kollokaatioksi sellaisen sanaparin tai -ryhmän, joka esiintyy yhdessä tilastollisesti merkitsevästi useammin kuin mitä pelkkien sanojen frekvenssien perusteella voisi odottaa. McEneryn ja Bakerin määritelmässä kollokaation parien on sijaittava toisistaan korkeintaan viiden sanan päässä ja toistuvuus on määritetty *likelihood ratio* -nimisen tunnusluvun avulla.

Kollokaattien analyysissa on ehdottomasti vahvuutensa. Kun kartoitetaan sanan merkitysvariaatiota, on yleisimpien kollokaattien vuosikymmenittäisten listojen läpikäynti huomattavasti vaivattomampaa kuin sanojen esiintymien silmäily suoraan teksteistä. Toisaalta koska kollokointi on aggregoiva menetelmä, joka tuottaa yksittäisistä esiintymistä abstraktin kokonaiskuvan, tulosten tulkinta ei aina välttämättä ole helppoa: kollokoivien sanojen syntaktis-semanttiset suhteet voivat olla monentyyppisiä eikä analyysi erottele näitä keskenään. Tästä syystä yksittäisten kollokaatioiden todistusvoima voi olla hyvinkin summittainen, ja joskus vasta kollokaatiotapausten laadullinen analyysi paljastaa niiden relevanssin.

McEnery ja Baker pyrkivät tarkastelemaan 1600-luvulla yleisimpien prostituoituihin viittaavien sanojen (*whore*, *jilt*, *harlot*, *strumpet* ja *prostitute*) välisiä merkitys-

eroja monipuolisesti. Sanoja tarkastellaan niin konnotaatioiden, tekstien puheaiheiden, paikallisten tekstiyhteyksien kuin puheaktien näkökulmasta. Tulkinat näistä piirteistä muodostetaan vertailemalla yhden sanan kollokaatteja toisen sanan kollokaatteihin. Argumentoinnin vahvistamiseksi kollokaatioita on lisäksi koottu suuremmiksi ryhmiä, jotta päätelmät eivät perustuisi vain yksittäisiin sanayhdistelmiin. Luokitteluperusteita on kaksi: aika ja merkitys. Aikaluokittelu jakaa kollokaatiot neljään luokkaan: pysyviin, syntyviin, kuoleviin ja vaihteleviin. Luokittelu perustuu siihen, miten sanojen esiintymät sijoittuvat vuosikymmentäisiin alakorpuksiin. Pysyvät kollokaatiot säilyttävät asemansa läpi tarkastellun ajanjakson, siinä missä syntyvät kollokaatiot ilmaantuvat aineistoon tarkastelujakson aikana ja kuolevat kollokaatiot vastaavasti katoavat. Vaihtelevat kollokaatiot syntyvät ja kuolevat tarkastelujakson aikana taikka katoavat käytöstä joskin aikaa.

Merkitysluokittelussa kollokaatiot on jaettu niin sanottuihin leksikaalisiin kenttiin. Aihepiirit on valikoitu joko kollokaatiosanojen yhteisinä yläkäsitteinä tai jonkinlaisten assosiatiivisten suhteiden yhteen sitomina aihepiireinä. Esimerkiksi semanttiseen kenttään RIKOLLISUUS kuuluu erilaisia rikollisiin viittaavia substantiiveja (esim. *knave, thief, rogue*) ja rikollisiin tekoihin viittaavia verbejä (esim. *steal*). Termin *leksikaalinen kenttä* käyttö tässä yhteydessä poikkeaa siis melkoisesti yleisemmästä, strukturalistisesta käytöstä, jossa sillä viitataan kielijärjestelmään vakiintuneisiin kategorioihin (ks. esim. Geeraerts 2010).

Tutkimuksen tulokset

Korpuslingvististen havaintojen avulla McEnery ja Baker kykenevät tekemään joitain tarkennuksia historian tutkimuksessa pohdittuihin kysymyksiin. Perinteisen käsityksen mukaan esimerkiksi sa-

naan *whore* sisältyvä transaktionaalisen seksin merkitys on kehittynyt aiemman, transgressiivisuutta ilmaisevan käytettävän pohjalta. Tekijöiden mukaan molemmat merkitykset ovat kuitenkin sisältyneet sanaan tasapuolisesti koko 1600-luvun ajan. Yllättävin havainto on kuitenkin se, että suurin osa *whore*-sanan esiintymistä kuuluu liittomuotoon *whore of Babylon*, joka oli protestanttien keskuudessa yleinen haukkumanimi katoliselle kirkolle. Tämän ilmauksen osana sana ei siis liity prostituutioon kuin metaforisesti. Siihen kysymykseen, missä määrin prostituutio koettiin 1600-luvulla osoitukseksi naisten voimaantuneisuudesta ja agenttiivisuudesta, McEneryn ja Bakerin korpusaineisto ei juuri tarjoa yksiselitteistä vastausta vaan pikemminkin tuloksia sekä puolesta että vastaan.

Korpuslingvististen analyysien toteutus ja saavutetut numeeriset tulokset raportoidaan teoksessa huolellisesti. Itse olisin ehkä kaivannut loppuun kokoavaa menetelmälukua, mutta nytkin teos esittää käytetyt menetelmät ymmärrettävästi myös sellaiselle lukijalle, jolla ei ole aiempaa kokemusta laskennallisesta kielentutkimuksesta.

Kuten jo edellä totesin, teoksessa esitelty tutkimusprosessi on edennyt siten, että historiallinen katsaus ja korpuslingvistiset analyysit on toteutettu toisistaan riippumatta. Luonnosvaiheessa tekijät ovat kuitenkin käyneet korpuslingvististen analyysien tulokset läpi ja pyrkineet tulkitsemaan saatuja tuloksia historian tutkimuksen aikaisempien tulosten valossa. Järjestys on perusteltu, ja kollokaatioiden kontekstualisointi aikalaislähteiden avulla on järkevä lähestymistapa ja potentiaalisesti omiaan ruokkimaan monitieteistä dialogia. Hyvä esimerkki on se, kuinka teoksessa selitetään korpuslingvististä havaintoa siitä, että leksikaaliseen kenttään *PITY* ('säälä, myötätunto') kuuluvat *whore*-sanan kollokaatiot yleistyivät 1660-luvulla ja väistyivät 1680-

luvun jälkeen. Monitieteinen tutkimus kantaa hedelmää, kun havainto kytketään historian tutkimuksen keinoin toisaalta käänteeseen, jossa seksityöläiset alettiin 1600-luvun puolessavälissä nähdä vähäosaisina olosuhteiden uhreina, toisaalta moraalisen ryhtiliikkeen kasvuun vuosisadan loppupuolella (s. 187).

Kiinnostava on myös UNDESIRABLES-kenttään ('marginaaliset, hyljeksityt') sijoittuvien kollokaattien analyysi. Korpuslingvistisen analyysin mukaan kollokaatit uudistuivat 1680-luvun tienoilla siten, että niiden merkityspiirteissä alkoi korostua 'julkijuopottelu' ja 'kiroilu' aiemmin yleisten 'lian' ja 'juonikkuuden' asemesta. Historiantutkimuksen selitys tälle on, että 1680-luvulla yleistyneiden moraalisten ryhtiliikkeiden myötä prostituutio alkoi hahmottua osaksi laajempaa, huonoksi käytökseksi tunnistettua sosiaalista ongelmaa. (S. 172–173.) Vaikka tällaisen ylimalkaisen tulkinnan olisi lingvistikin voinut kollokaatteja katsomalla tehdä, tarvitaan historioitsija yhdistämään ilmiö historiallisiin toimijoihin, tässä tapauksessa erityisesti Society for the Reformation of Manners -liikkeeseen.

Valitettavasti suurimmassa osassa tapauksista yhteistyö ei kuitenkaan ole yhtä hedelmällistä. Paikoin tämä johtuu siitä, että korpuslingvistisesti saadut havainnot ovat luonteeltaan monitulkintaisia siten, että yksi havainto tukee historian tutkimuksen vallitsevaa käsitystä, samalla kun toinen horjuttaa sitä. Toisinaan taas havainnot tukevat historian tutkimuksen vakiintunutta käsitystä tarjoamatta ilmiöön uutta näkökulmaa. Tällöin historian tutkijan panos jää usein vähäiseksi, lähinnä nyökkäykseksi. Tällaisissa tapauksissa lingvistilukija olisi mielellään lukenut enemmän siitä, miten historioitsija kontekstualisoi lingvistisesti havaitut muutokset osaksi historiallisia prosesseja. Vaikka havainnot historian tutkimuksen näkökulmasta olisivatkin hieman itsestäänselviä, monitieteisessä teoksessa niille on paikkansa.

Korpuslingvistiikan kannalta voi huomauttaa vielä siitä, että teoksessa ei onnistuta valituilla korpusmenetelmillä tuomaan selvyttä siihen, mikä on prostituutioihin viittaavien lähisyronymien *whore* ja *harlot* välinen ero (ks. s. 189–193). Erilaisia aihioita ja täkyjä nostetaan esiin useampia, mutta kummastakaan sanasta ei kollokaatioiden tarkastelun avulla löydy piirrettä, jota toisen kohdalla ei tavattaisi. Tämä havainnollistaa hyvin sitä, miten korpusevidenssi voi usein olla ristiriitaista ja moniselitteistä. Kielenkäytöstä harvoin löytyy ilmiöitä, joita valitut analyttiset kategoriat kuvaisivat täydellisesti. Hienosyiset jatkumot ja reunoiltaan sumeat kategoriat ovat ennemminkin sääntö kuin poikkeus. Tämä näkyy erityisesti silloin, kun kategorioiden välisiä suhteita lähestytään laskennallisesti.

Osasyynä McEneryn ja Bakerin tutkimuksen korpuslingvistisiin heikkouksiin pidän sitä, että valitut analyysit rajoittuivat melko tarkasti siihen, mitä tekijöiden valitsemalla CQPWeb-ohjelmistolla on suoraan pystynyt tekemään. Tekijät ovat siis lähteneet analysoimaan lähisyronymien merkityseroja ikään kuin menetelmä eikä tutkimuskysymys edellä. Sama menetelmällinen rajoittuneisuus näkyy myös silloin kun kirjoittajat pohjivat näytelmätekstien vaikutusta siihen, että prostituoitujen nimityksistä on kehittynyt solvauksia. Teatteri oli Englannin vallankumouksen aikana kielletty eikä näytelmäkirjallisuutta julkaistu. Näytelmätekstien vaikutus on hyvä selitys tutkittujen kollokaatioiden vaihtelulle, mutta analyysi jää spekulatiiviselle tasolle, koska korpusta ei ole annotoitu tekstilajeittain. Ohjelmisto ei ylipäänsä tarjoa välineitä tekstilajittaisen vaihtelun analyysiin. Tämänkaltaiset valinnat horjuttavatkin hieman työn uskottavuutta. On vaikea hyväksyä, että yksittäisen tietokoneohjelman rajoitusten tulisi sanalla tieteellisen tutkimuksen mahdollisuuksia.

Lopuksi

Kokonaisuutena McEneryn ja Bakerin teoksen voi sanoa pyrkivän kahdella rintamalla kohti sisäisesti ristiriitaisia tavoitteita. Yhtäältä tekijät ovat tavoitelleet kokonaisuutta, jonka metodologia voisi muodostaa havainnollisen mallin tuleville tutkimuksille. Tämä tavoite tuntuu selittävän monet teoksen valinnoista, kuten sen, että analyysit on tehty korpusohjelmiston ehdoilla. Tästä on seurannut, että saavutetut tulokset jäävät jossain määrin laihoiksi. Tieteellisen tutkimuksen luonteeseen kuuluu toki se, ettei hyvästä tutkimusideasta aina seuraa mullistavia uusia tuloksia. McEnery ja Baker esittävät kyllä lupaavia havaintoja esimerkiksi transgressiivisuudesta ja transaktionaalisuudesta, mutta menetelmien eksploraatiivisesta luonteesta johtuen havainnot eivät vie pintaa syvemmälle. Tämä on jossain määrin eksploraatiivisten menetelmien tarkoituskin: ne herättävät enemmän kysymyksiä kuin tarjoavat vastauksia. Samalla voi kuitenkin kysyä, vakuuttuvatko lukijat menetelmien hyödyllisyydestä, jos ne eivät tuota kuin viitteellisiä havaintoja mahdollisten jatkotutkimusten pohjaksi. Voisi sanoa, että menetelmien havainnollistettavuuden ja hyödynnettävyyden välillä on tässä kohtaa ristiriita.

Toinen ristiriita liittyy monitieteiseen yhteistyöhön, jossa eri tieteenaloja edustavat tutkijat tavoittelevat läheistä yhteistyötä ilman kompromisseja. Vakiintunut monitieteisen yhteistyön muoto on artikkelikokoelma, jossa eri alojen tutkijat lähestyvät yhteistä ilmiötä kukin omissa artikkeleissaan. Tällöin jokaisen osatutkimuksen kysymyksenasettelu nousee tutkijan oman tieteenalan tiedonintresseistä. Tällaisesta tutkimuksesta, jossa erityisesti kielen tutkijat ja historioitsijat yhdistävät voimansa, on hyvä esimerkki Laitisen ja Mikkolan (2013) toimittama teos *Kynällä kyntäjät*. Sen, mitä tällainen tutkimus häviää tieteidenvälisessä vuoropuhelussa,

se saa takaisin mahdollisuudessa selkeiden tutkimuskysymysten muotoiluun. Arvioidussa teoksessa kahden alan tutkijoiden yhteistyö oli yhtäältä niin läheistä, etteivät he voineet tavoitella omia, itsenäisiä tutkimuskysymyksiään, toisaalta niin etäistä, ettei kokoavaa yhteistä näkökulmaa tai intressiä syntynyt. Kokonaisuuden hajanaisuuden vuoksi teos ei tarjoa historian-tutkimuksen tavoittelemaa uskottavaa kuvausta historiallisesta prosessista mutta ei myöskään lingvististä kokonaiskuvaa kielen ilmiöstä. Havainnollisena esityksenä korpusmenetelmien soveltamisesta teos on kuitenkin ansiokas ja tarpeellinen lisä olemassa olevan kirjallisuuden joukkoon.

ANTTI KANNER
etunimi.sukunimi@helsinki.fi

Kirjoittaja on suomen kielen väitöskirjatutkija Helsingin yliopiston digitaalisten ihmistieteiden osastossa.

Lähteet

- EEBO = Early English Books Online. <https://quod.lib.umich.edu/e/eebogroup> (3.12.2018).
- GEEAERTS, DIRK 2010: *Theories of lexical semantics*. Oxford: Oxford University Press.
- HALLIDAY, MICHAEL A. K. – TEUBERT, WOLFGANG – YALLOP, COLIN – ČERMÁKOVÁ, ANNA 2004: *Lexicology and corpus linguistics. An introduction*. Open Linguistics. Lontoo: Continuum.
- JANTUNEN, JARMO H. 2009: ”Minulla on aivan paljon rahaa” – Fraseologiset yksiköt suomen kielen opetuksessa. – *Virittäjä* 113 s. 356–381.
- LAITINEN, LEA – MIKKOLA, KATI (toim.) 2013: *Kynällä kyntäjät. Kansan kirjallisuuden 1800-luvun Suomessa*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- MARJANEN, JANI 2018: *Begreppshistoria*. – Martin Gustavsson & Yvonne

Svanström (toim.), *Metod. Guide för historiska studier* s. 97–133. Lund: Studentlitteratur.

RICHTER, MELVIN 1995: *The history of political and social concepts*. New York: Oxford University Press.

TOGNINI-BONELLI, ELENA 2001: *Corpus-linguistics at work*. Amsterdam: John Benjamins.

TULLY, JAMES (toim.) 1988: *Meaning and context. Quentin Skinner and his critics*. Princeton: Princeton University Press.

Lukutaito emansipaation välineenä ja yhteisön voimavarana

Tuija Laine: *Aapisen ja katekismuksen tavaamisesta itsenäiseen lukemiseen. Rahvaan lukukulttuurin kehitys varhaismodernina aikana*. Helsinki: Suomalaisen Kirjallisuuden Seura 2017. 210 s. ISBN 978-952-222-862-8.

Tuija Laine piirtää tiiviissä, parinsadan sivun teoksessaan suomalaisten lukutaidon isoja kaaria noin 300 vuoden ajalta, 1500-luvun lopulta 1800-luvulle. Hän tarkastelee rahvaan lukukulttuuria ennen muuta Suomessa mutta sijoittaa maan tarvittaessa Ruotsin yhteyteen – niin kuin noiden vuosisatojen tarkastelussa kuuluukin – ja edelleen osaksi Pohjoismaita ja Eurooppaa.

Kirjassa kuvataan lukutaidon kehitys ulkoluvusta kohti modernia itsenäistä lukemista. Kirja perustuu Laineen työhön kirjahistorian professorina ja yhteistyöhön muiden kirjahistorian tutkijoiden kanssa ennen muuta teologian, kirkkohistorian ja historian viitekehyksessä. Laine edustaa tieteenalan sosiaalishistoriallista suuntausta, jossa lukutaitoa ja kirjoja tarkastellaan yhteisön ja yhteiskunnan näkökulmasta: kenellä oli mahdollisuus lukea, mitä kirjoja oli saatavilla, miten lukutaito vaikutti lukijoiden elämään ja niin edelleen.

Kirjan alaotsikko kohdistaa huomion nimienomaan rahvaan lukukulttuuriin.

”Rahvas” tarkoittaa Laineen kielenkäytössä kansanihmisiä eli talonpoikia ja säätyjen ulkopuolisia. Yläpuolella olevat hallitsijat ja oppineet kannustivat, ohjasivat ja edistivät rahvaan lukutaitoa tai kurittivat lukutaidottomuudesta. Rahvaan lukutaidon tarkasteleminen on vastakkainen – tai täydentävä – suunta sille kirjahistorian painotukselle, jossa kirjojen ja lukutaidon historiaa kirjoitetaan tarkastelemalla kalliita loistopainoksia, hienoja kuvituksia ja käsityötaidon ihmeitä ja jossa lukutaitoa edustavat papisto ja munkit.

Kirjan johdantoluku problematisoi rahvaan näkökulmaa monelta kannalta. Laine lähtee liikkeelle historian tutkija Robert Darntonin ajatuksesta, jonka mukaan kirjahistorian tutkimuksessa tulisi ottaa huomioon taloudellinen, sosiaalinen ja poliittinen ulottuvuus. Lisäksi tulisi tarkastella kaikkien näiden tekijöiden vaikutusta toisiinsa. Laine kohdistaa tarkastelunsa ennen muuta kirkon ja uskonnollisen yhteisön toimintaan, sillä näiden vaikutus lukutaidon varhaisvuosisatoina oli olennainen.

Laine tekee toisenkin olennaisen rajauksen. Hän tarkastelee lukutaitoa sosiaalisena ja yhteiskunnallisena ilmiönä, ei vain yksilön kykynä lukea. Tähän viittaa alaotsikon sana ”lukukulttuuri”. Lukutaitoa määrittävät ja ohjailevat poliittiset ja