

Hyvä sanakirja

Hyvä sanakirja

Tieteellisiä, käytännöllisiä ja eettisiä näkökulmia leksikografiaan

Toimittaneet LAURA TYYSSTERI ja KAISA LEHTOSALO

Turun yliopiston suomen kielen
ja suomalais-ugrilaisen kielentutkimuksen oppiaineen julkaisuja **3**
Publications of the Department of Finnish and Finno-Ugric Languages
of the University of Turku

Turku 2013

Kannen kuva Laura Tyysteri
Taitto Eeva Herrala

ISBN 978-951-29-6010-1

ISSN 2323-5322

Suomen kieli ja suomalais-ugrilainen kielentutkimus, Turun yliopisto
Julkaistu painettuna 2013 ja sähköisenä 2015.

Sisältö

Alkusanat	7
Kaisa Lehtosalo & Laura Tyysteri Sanakirjatyön valinnat ja leksikografian ammattietiikka	9
Igor Kudashev Sanakirjan suunnittelu ja siihen vaikuttavat tekijät	21
Eija-Riitta Grönros Sanakirjojen normatiivisuus ja deskriptiivisyys	35
Kaisa Lehtosalo & Pirkko Kuutti Sanakirja historiassa, historia sanakirjassa	53
Kimmo Granqvist Romanikielen sanakirjatyö ja sen eettiset kysymykset	77
Klaas Ruppel Maahanmuuttajille laadittavat sanakirjat ja etiikka	94
Sirkka Saarinen Suomalais-ugrilaisten vähemmistökielten sanakirjojen laatimisesta	113
Laura Tyysteri Arvosidonnaiset sanat yleiskielen selittävässä sanakirjoissa	131
Nina Pilke & Jaana Puskala Kummitussanat ja kummalliset esimerkit käyttäjänkin kontolla? Sanakirjan käyttäjän etiikka osana kieliasantuntuutta	154
Raimo Jussila Sanakirjat ja tekijänoikeus	174
Lähteet	188
Teoksen kirjoittajat	212

Alkusanat

Tämä kirja on syntynyt halusta koota ja levittää sanakirja-alalle kertynyttä hiljaista tietoa. Kokoelma on ensimmäinen sanakirjatyön eli leksikografian perustietoa kattavasti yksiin kansiin kokoava teos Suomessa. Kokoelman kymmenessä artikkelissa käsitellään sekä sanakirjatyön periaatteita että käytännön työssä vastaan tulevia valintatilanteita. Suurin osa teoksen kirjoittajista työskentelee tai on työskennellyt sanakirjantekijöinä Kotimaisten kielten keskuksessa ja eri yliopistoissa, ja kirjassa esitelläänkin monia erilaisten sanakirjojen toimitustyössä esiin nousseita kysymyksiä. Sanakirjantekijöiden lisäksi mukana on leksikografian tutkijoita, jotka tarkastelevat sanakirjoja tutkijan ja käyttäjän näkökulmasta. Toivomme teoksen sopivan luettavaksi sekä alan ammattilaisille, kieliainneiden opiskelijoille että muille leksikografiasta kiinnostuneille.

Kokoelman artikkeleissa käsitellään enimmäkseen tieteelliseen tutkimukseen pohjautuvia yksi- tai kaksikielisiä suomalaisia sanakirjoja. Lisäksi esitellään myös muita kuin tieteellisiä sanakirjoja sekä tehdään huomioita eri maiden sanakirjatratioista. Kaikki artikkelit lukuun ottamatta ensimmäistä, toimittajien laatimaa tausta-artikkelia, ovat käyneet läpi kaksinkertaisen anonyymien vertaisarvioinnin. Kaikkien artikkelien yhteinen lähdeluettelo on koottu teoksen loppuun. Teoksen sisällä muihin artikkeleihin on viitattu lyhenteellä tk. = tässä kokoelmassa.

Ajatus kokoelmasta syntyi Vaasan yliopiston kääntämisen, erikoisala- viestinnän ja monikielisyyden tutkijayhdistys VAKKI ry:n symposiumissa vuonna 2011 pitämämme esitelmän jälkimainingeissa. Symposiumin teema oli kieli ja etiikka. Teema sai meidät pohtimaan myös sanakirjatyötä uudesta näkökulmasta, ja pian olimmekin jo kokoamassa kirjoittajia teokseen, jossa leksikografiaa katsellaan samasta perspektiivistä. Etiikka nähdään teoksessa laajasti: Sen piiriin luetaan leksikografian ammat-

tiettiikka ja vastuu esimerkiksi työn laadusta sekä erilaiset sanakirjoihin liittyvät yhteiskunnalliset kysymykset. Eettisiä valintoja ovat esimerkiksi päätökset siitä, millaisten käyttäjien tarpeet sanakirjoissa huomioidaan, mitkä kielet ja kielimuodot valitaan sanakirjassa kuvattavaksi ja miten sanakirjan hakusanat valitaan ja selitetään.

Toimitustyössä meitä auttoi eteenpäin into ja päättäväisyys. Kokoelma on meille toimittajille ensimmäinen, mutta onneksi saimme asiantuntevaa apua kokeneemmalta. Suurkiitos professori Kaisa Häkkiselle kaikista tuiki tarpeellisista neuvoista ja muusta tuesta! Ansiostasi onnistuimme välttämään monet aloittelijan virheet ja tiesimme koko ajan, mitä meidän pitää seuraavaksi tehdä. Työtaakkaamme vähensi olennaisesti myös suomen kielen opiskelija Juuli Lindroos, joka teki kokoelman parissa opintoihinsa kuuluvan harjoittelun. Kiitos Juuli!

Kirjaprojektin mahdollistamisesta tahdomme kiittää Koneen Säätiotä, joka myönsi hankkeelle rahoituksen. Kiitos myös nimettömäksi jääville vertaisarvioijille sekä kaikille muille artikkeleita niiden eri vaiheissa kommentoineille. Toitte kokoelmaan mukaan monia mielenkiintoisia näkökulmia! Kotimaisten kielten keskusta kiitämme tilojen lainasta aloitusseminaarissamme sekä mahdollisuudesta osallistua leksikografien Ilpo-ryhmän kokoontumisiin. Kiitos kaikille ilpolaisille ideoista ja tuesta! Kiitos myös Turun yliopiston suomen kielen ja suomalais-ugrilaisen kielentutkimuksen oppiaineelle, joka on ottanut teoksen julkaisusarjaansa.

Suurin kiitos teoksen valmistumisesta kuuluu tietysti kirjoittajille. On ollut hienoa huomata, kuinka tärkeänä monet ovat tätä hanketta pitäneet ja kuinka paljon he ovat olleet valmiita uhraamaan aikaansa sen eteen muun kiireen keskelläkin. Toivomme, että keskustelu esiin nostetuista tärkeistä ja mielenkiintoisista aiheista jatkuu.

Turussa 5.10.2013

Teoksen toimittajat

Laura Tyysteri ja Kaisa Lehtosalu

KAISA LEHTOSALO & LAURA TYYSSTERI

Sanakirjatyön valinnat ja leksikografian ammattietiikka

Se [sanakirja] on erityinen kirja, johon on kerätty kaikki maailman sanat, ja siinä kerrotaan mitä mikin sana merkitsee.
Eduard Uspenski (1977: 14): Krokotiili Gena ja hänen ystävänsä

Moni kielenkäyttäjä varmasti toivoo, että olisi olemassa kirja tai sähköinen tiedosto, johon olisi koottu kaikki maailman sanat ja niiden merkitykset. Kuitenkin jokainen sanakirjoja käyttänyt ja erityisesti niitä tehnyt tietää, että se on käytännössä mahdotonta. Eri kieliä on maailmassa noin 7 000 (Ethnologue.com), ja kun otetaan vielä huomioon kielten eri variantit, toiveikkainkin utopisti alkaa oivaltaa, ettei kaikkien maailman sanojen kokoaminen yhteen ole millään mahdollista. Koska elävät kielet muuttuvat jatkuvasti ja eri kielenkäyttäjillä saattaa olla erilaisia käsityksiä sanojen merkityksistä ja merkitysvivahteista, edes yhden kielimuodon saati kokonaisen kielen kaikkien sanojen selittäminen yksissä kansissa ei yleensä onnistu.

Eduard Uspenskin lastenkirjasta lainattu sanakirjan määritelmä, jonka mukaan sanakirjan tehtävä on koota yhteen ja selittää kaikki sanat, edustaa yhtä ääripäätä. Toisessa ääripäässä voi sanoa olevan näkökulman, jonka mukaan sanojen merkitykset eivät ole etukäteen määriteltyjä, vaan ne muotoutuvat aina vasta kontekstissa (tarkemmin esim. Östman 2000: 78–80). Johdonmukainen seuraus tästä olisi, ettei sanakirjoja voi tehdä, koska kaikkien kuviteltavissa olevien kontekstien huomioiminen sanakirjassa ei ole mahdollista. Myös tämä näkemys on liian jyrkkä, mistä on osoituksena jo se, että sanakirjoja kuitenkin tehdään jatkuvasti ja niistä on myös paljon hyötyä kielenkäyttäjille.

Eri kielimuotoja kuvaavia ja erilaisille käyttäjäryhmille suunnattuja sanakirjoja on nykyään melkein loputtomasti. On yleiskielen ja puhekielen sanakirjoja, erilaisia jargon-, sivistys- ja slangisanakirjoja, yksi- kaksi- ja monikielisiä sanakirjoja, erilaisille käyttäjäryhmille, kuten kielenoppijoille tai maahanmuuttajille, suunnattuja sanakirjoja ja niin edelleen. Suomalaisille käyttäjille tutuin sanakirjaformaatti tuntuu olevan kaksikielinen sanakirja, jossa toisena kielenä on suomi. Yksikielisten sanakirjojen käyttöön ei ole Suomessa totuttu, mikä näkyy esimerkiksi siinä, että Nykysuomen sanakirjaan (NS, 1951–1961) vedotaan edelleen ikään kuin se olisi paras mahdollinen 2000-luvun kielen kuvaaja. Vaikuttaa siltä, että kaikki eivät oivalla, että Nykysuomen sanakirja on jo yli 50 vuotta vanha. Tai ehkä ihmiset eivät vain yleisemmin tiedosta kielen jatkuvaa muutosta ja nykykieltä kuvaavien sanakirjojen päivitystarvetta. Sanakirjojen käytötaidon ja sanakirjatyön arvostuksen parantamiseksi olisikin hyvä saada lisättyä kielenkäyttäjien tietoa sekä sanakirjojen käyttömahdollisuuksista että tekoprosessista ja taustoista.

Tässä artikkelissa kuvataan sanakirjan tekoprosessia yleisellä tasolla. Valotamme niitä valinta- ja karsintaprosesseja, joita sanakirjan tekoon väistämättä liittyy ja jotka vaikuttavat lopputulokseen monin tavoin. Pohdimme artikkelissa myös sanakirjantekijän ammattietiikkaa ja työssä tehtäviin valintoihin vaikuttavia seikkoja. Erilaisiin sanakirjoihin ja sanakirjatyön vaiheisiin liittyviä erityiskysymyksiä käsitellään tämän teoksen muissa artikkeleissa, joille tämä kirjoitus toimii eräänlaisena johdantona. Teoksen muut artikkelit on esitelty tämän kirjoituksen loppupuolella.

Sanakirjatyön valinnat

Sanakirjatyön ensimmäinen välttämätön ratkaisu on kuvattavan kielen ja kielimuodon valinta. Vaikka nykyään on luonnollista, että sanakirjat kuvaavat monia eri kieliä, aina ei suinkaan ole ollut selvää, että kaikki kielet ovat sanakirjan arvoisia. Sanakirjaan otettavan kielen valinta heijastaakin yhteiskunnan arvoja. Esimerkiksi vähemmistökielen sanakirjan puuttuminen voi kertoa myös kielen ja jopa sitä käyttävien ihmisten arvostuksen puutteesta (esim. Salmi–Laakso 2005: 315–319; Granqvist tk.). Se, ettei jostain kielestä ole sanakirjaa, voi luonnollisesti johtua myös siitä, ettei kielestä ole olemassa muutakaan kirjallista kuvausta. Näin on

laita esimerkiksi monien Uuden-Guinean kielten kanssa. Alueella on maailman suurin kielitihentymä, eikä kaikkien kielten ominaispiirteitä ole vielä luokiteltu. (Anhava 1998: 174, 234.) Myös taloudelliset realiteetit ovat usein esteenä sanakirjan teolle. Esimerkiksi kaikista Suomessa puhutuista maahanmuuttajien kielistä ei ole vielä olemassa yhtään sanakirjaa tai edes lyhyempää sanastoa, jossa suomi olisi toisena kielenä (Edu.fi: Sanakirjoja; Ruppel tk.).

Kielen valinnan ohella myös sanakirjaan otettavan kielimuodon valinta kertoo yhteiskunnan arvostuksista. Kielitieteessä oli pitkään vallitsevana tendenssinä suosia kirjoitettua kieltä, vaikka puhuttu kieli onkin kielen primääri muoto (Saussure 1959 [1916]: 23–24). Esimerkiksi Oxford English Dictionary (OED, 1884–) on alusta asti pohjautunut kirjalliseen materiaaliin (Examining the OED.uk: Outline material). Kielissä, joissa ei ole normitettua kirjakieltä tai se on nuori, tilanne on erilainen. Esimerkiksi 1600-luvun sanakirjoissa, joissa suomi on ollut yhtenä kielenä, on käytetty materiaalina myös murteita (Nummila 2010). Kokonaan puhuttuun kieleen perustuu Suomen murteiden sanakirja, jonka aineistoa alettiin kerätä vuonna 1900 (SMS 1985: vi). Suomen murteiden sanakirja on edelleen yksi Kotimaisten kielten keskuksen (vuosina 1976–2011 Kotimaisten kielten tutkimuskeskus) suurista sanakirjaprojekteista. Eri kielimuotojen arvostuksen kannalta historiallisesti kuvaavaa on kuitenkin esimerkiksi se, että Suomen ensimmäinen slangisanakirja julkaistiin vasta 1960-luvulla (Hämäläinen 1963). Materiaalia puhutun kielen erilaisista varianteista ja slangeista oli kyllä kerätty jo 1900-luvun alkupuolella, mutta hankkeet rikollisten, koululaisten, katupoikien ja viinan salapolttajien sanaston julkaisemiseksi unohdettiin, kun sisällissodan jälkeen kielentutkijoiden ihannoimaksi tutkimuskohteeksi nousi syrjäseutujen talonpoikien kielimuoto (Häkkinen 1993: 28–29).

Kielen ja kielimuodon valinnan lisäksi keskeinen kysymys on aineiston valinta. On melko harvinaista, että on olemassa hallitun kokonaisuuden muodostava aineisto, jonka sopivuutta sanakirjaan ei tarvitse arvioida. Poikkeuksia ovat muun muassa yksittäisten teosten tai kirjoittajien tuotannon pohjalta tehdyt sanakirjat. Esimerkiksi Kalevalan kielestä on tehty sanastoja ja sanakirjoja (viimeisimpänä Jussila 2009). Potentiaalisesti tällainen aineisto on käytettävissä kokonaan. Tavallisempaa kuitenkin on, että aineistoa kerätään laajasta materiaalista tiettyyn tarkoitukseen ja kerätyn aineiston sopivuutta arvioidaan ja sen käyttöä tavalla tai

toisella rajataan. Esimerkiksi yleiskielen sanakirjojen aineistoa kartutetaan jatkuvasti (painetusta) mediasta, mutta kaikkea ei siirretä suoraan sanakirjaan.

Myös aineiston hankinnassa käytetään hyvin erilaisia metodeja. Jotkin aineistot ovat helposti saatavissa ja toiset voivat olla kirjaimellisesti kiven takana, kuten esimerkiksi vankilaslangi. Leo Lipsonen (1990: 6–7, 29) keräsi suuren osan Vankilaslangin sanakirjan aineistosta istuessaan itse vankilassa, jossa osa informanteista halusi pysyä täysin anonymine. Myös osa sanastosta haluttiin pitää vain oman viiteryhmän tiedossa. Tämä tietysti vaikuttaa aineiston saatavuuteen ja käsittelymahdollisuuksiin.

Aina ei ole itsestään selvää, että aineistonkeruussa huomioidaan kaikkien kielenkäyttäjien ja aihealueiden sanasto tasapuolisesti. Esimerkiksi Suomen murteiden sanakirjan pohjana käytetyt aineistot kerättiin pitkissä jaksoissa vuosina 1900–1974 (SMS 1985: vi). Keruutyön alkuvaiheessa ei tullut edes mieleen kerätä yhteiskunnallisesti tulenarkaa materiaalia ja seksuaalisanasto oli täydellinen tabu (Häkkinen 2006: 206, 215). Esimerkki kuvastaa hyvin sitä, miten ajan henki voi vaikuttaa myös sanakirjoihin – tiedostamatta tai tiedostetusti.

Kerätyn aineiston sopivuutta sanakirjaan täytyy arvioida sanakirjan funktion kautta. Yleensä sanakirjaan otettavaa sanastoa täytyy karsia, vaikka sanakirjan tarkoitus olisi kuvata tietty kielimuoto mahdollisimman laajasti. Erityyppisissä sanakirjoissa on tietysti erilaiset karsintaperusteet: mikä on sopiva sana murre-sanakirjaan, ei välttämättä ole sopiva yleiskielen sanakirjaan. (Zgusta 1971: 240–245.) Koska sanastoa valitaan eri toimituksissa hyvin erilaisin kriteerein, on mahdotonta antaa yleispäteviä ohjeita siitä, mikä on sanakirjaan sopiva sana. Tiettyjä periaatteita on kuitenkin mahdollista kuvata. Esimerkiksi yleiskieltä kuvaavissa yksi- tai kaksikielisissä sanakirjoissa huomioidaan sanan yleisyys. Vanhentuneita ja perifeerisiä sanoja niihin ei yleensä oteta. Myöskään aivan uusia sanoja ei kelpuuteta sanakirjaan heti, sillä ne eivät välttämättä vakiinnu kieleen. Esimerkiksi OED:ssä on uudissanoille suuntaa-antavana aikarajana kymmenen vuotta (Simpson 2011; OED: FAQ). Jos sana pysyy käytössä tämän ajan, se päivitetään myös sanakirjaan. Myös Kielitoimiston sanakirjassa (KS, 2004–) huomioidaan sanan elinvoimaisuus ja pitkäikäisyys.

Suomen kaltaisessa kielessä, jossa on paljon yhdyssanoja ja johdoksia, kiinnitetään sanakirjaan otettavien sanojen valinnassa huomiota ilmausten läpinäkyvyyteen. Sanakirjoihin ei ole yleensä tapana ottaa sellaisia

yhdyssanoja ja johdoksia, joiden merkityksen ajatellaan olevan tulkittavissa jo niiden pelkkien rakenneosien perusteella (esim. *kieli + tihentymä* > *kielitihentymä*, *selittää + -minen* > *selittäminen*). Leksikografien täytyy myös miettiä, milloin lainasanoista tulee yleiskielen sanoja ja onko sana sopivampaa sijoittaa vaikkapa sivistyssanakirjaan kuin yleiskielen sanakirjaan. Yksi mahdollisuus on ilmaista käyttöalamerkinnoilla sanojen kuulumista tiettyyn kielimuotoon (mm. slangi, erikoiskieli) tai kertoa jotain sanojen tyyliarvosta (mm. arkikielinen, vanhentunut). Käyttöalamerkintöjen avulla voi kuvata siirtymävaiheessa olevien sanojen senhetkistä tilaa ja myös seurata sanojen historiallista kehitystä. Erityisesti ei-äidinkielisille puhujille tyyliarvomerkinnot ovat tärkeitä, sillä niiden kautta he saavat tietoa siitä, miten sanoja yleensä käytetään ja millaisia konnotaatioita muut kielenpuhujat voivat niihin liittää.

Yksikielisissä sanakirjoissa merkityksen selittäminen ja monikielisissä sanakirjoissa vastineiden valinta ovat monin tavoin sanakirjan teon ydintä. Käyttäjätutkimusten mukaan merkityksien ja vastineiden selvittäminen onkin yksi keskeisimmistä syistä käyttää sanakirjaa (vrt. Welker 2010). Kieliyhteisössä vallitsee yleensä yhteisymmärrys siitä, mitkä ovat sanan perusmerkitykset. Ilman jonkinlaista yhteisymmärrystä ei kommunikointi ihmisten välillä olisi mahdollista. Perusmerkitysten lisäksi sanoilla on kuitenkin myös konteksti- ja käyttäjäkohtaisia merkityksiä. Leksikografian haasteena on määritellä, mitkä merkitykset ovat niin olennaisia, että ne kuvataan sanakirjassa. On myös päätettävä, miten valitut merkitykset selitetään. Edelleen leksikografian on määriteltävä merkitysryhmien välinen järjestys. Tämän prosessin aikana leksikografi ottaa kantaa eri merkitysten väliseen hierarkiaan, perustui se sitten yleisyyteen, keskeisyyteen, etymologiaan tai konkreettisuuteen.

Kielelliset merkitykset eivät ole koskaan vain yhden ihmisen omaisuutta, sillä vain yhden kielenkäyttäjän tuntemilla merkityksillä ei voi välittää informaatiota kielenkäyttäjältä toiselle (vrt. Wittgenstein 1999 [1953]: 151–154). Sanakirjatyössä korostuukin kielen sosiaalinen luonne. Olisi järjetöntä selittää merkityksiä, jotka eivät vastaa kieliyhteisön yhteistä näkemystä tai jotka ovat täysin kontekstisidonnaisia. Vaikka omassa kielenkäytössäni *auton* merkitys olisi 'surmanloukku', ei kyseinen konnotaatio ole kieliyhteisössä tavallinen. Merkitys voisi esiintyä esimerkiksi onnettomuusutisoinnissa, mutta yksittäinen käyttöyhteys ei ole sanakirjan tasolla merkittävä.

Jotta merkityksenselitykset olisivat johdonmukaisesti jäsennettyjä, leksikografit nojaavat oman sanakirjansa toimitusohjeisiin (ks. kuitenkin Jussila tk. s. 181–182). Niistä huolimatta sanakirjantekijä saattaa, usein tiedostamattaan, selittää sanoja omasta näkökulmastaan (esim. sukupuolittuneesti tai politisoituneesti). Ammattimainen toimitusryhmä onkin paras keino tuottaa mahdollisimman objektiivisia merkityksenselityksiä. Laajat yleiskielen sanakirjat ovat aina ryhmätyön tulosta. Monien sanakirjojen tekemiseen ovat osallistuneet jo useamman sukupolven leksikografit, sillä sanakirjatyössä on tavallista käyttää apuna aiempia sanakirjoja. Näin varmistetaan, ettei sanaston kuvauksesta unohdu mitään oleellista. (Ks. esim. Winchester 1998: 94.) Esimerkiksi OED:ä on tehty lähes keskeytyksettä jo 1800-luvulta ja KS pohjautuu NS:aan ja Suomen kielen perussanakirjaan (PS, 1990–1994). Vaikka aiempien sanakirjojen käyttö toimitustyössä on järkevää, tekijänoikeuksien kannalta on tärkeää miettiä, miten suoraan varsinkaan muiden kustantajien sanakirjoja voidaan käyttää, ennen kuin on kyse plagioinnista (ks. esim. Hausmann 1989; Lehtosalo–Tyysteri 2011: 220–221; Jussila tk.).

Merkitystenselitysten lisäksi sanojen merkitystä ja käyttöä on mahdollista valottaa esimerkkien avulla. Sanakirja-artikkeleihin tulisi valita esimerkkejä, jotka kuvaavat sanojen tyypillistä käyttöä. Mitään oikotietä tyypillisyyden määrittämiseen ei kuitenkaan ole. Tarkasti rajatuissa aiheistoissa (esim. yksittäisen kirjailijan tuotanto) yleisyys voidaan mitata kvantitatiivisesti, mutta erityisesti elävä nykykieli pakenee määrittäjänsä koko ajan. Sen sanasto karttuu ja muuttuu jatkuvasti, kuten myös sanojen käyttöyhteydet. Leksikografi voi koettaa selvittää tyypillisyyttä esimerkiksi internetin hakukoneiden avulla, mutta ne eivät ole tarpeeksi luotettavia lopullisten ratkaisujen perustaksi. Räikeimmillään esimerkkien valinnassa näkyy arvosidonnaisia kannanottoja (vrt. netissä vapaasti täydennettävissä oleva Urbaani sanakirja (US) s.v. *mamu* ”mamut tekevät paljon rikoksia”). Toimittajan käsitykset sanojen tyypillisyydestä voivat siis näkyä myös esimerkkien valinnassa ja muokkauksessa. (Ks. myös Siiroinen 1988.)

Edellä kuvatut valinnat ovat konkreettisia sanakirjan tekoon liittyviä käytännön kysymyksiä. Samalla ne ovat kuitenkin myös sekä tieteellisiä että eettisiä kysymyksiä. Koska kielenkuvaus ei voi koskaan olla täysin objektiivista, erilaisia ratkaisuja ei läheskään aina voida pitää ehdottoman oikeina tai väärinä. Esimerkiksi poliittinen sanasto nostattaa tunteita ja vaatii leksikografilta tasapainoilua eri näkemysten välillä. Millä sanoilla

ja selitteillä tulisi kuvata esimerkiksi vuoden 1918 tapahtumia Suomessa? Voidaan jopa kysyä, onko poliittista sanastoa mahdollista edes määritellä täysin puolueettomasti (ks. Vihonen 1972; Tyysteri tk.).

Leksikografian ammattietiikka ja yhteiskunnallinen vastuu

Sanakirjoja tekevät monenlaiset ihmiset monenlaisista kielimuodoista. Tässä teoksessa päähuomio on tutkimuslaitoksissa ja yliopistoissa tehtävässä tieteellisen tutkimuksen tuloksiin pohjaavassa sanakirjatyössä, mutta teoksessa pohditaan myös muiden sanakirjatyötä tekevien vastuuta lopputuloksen laadusta. Suurin tieteellinen, ammatillinen ja yhteiskunnallinen vastuu on tieteelliseen tutkimukseen pohjaavien sanakirjojen tekijöillä, sillä heidän työnsä tuloksena syntyvät sanakirjat perustuvat parhaaseen mahdolliseen saatavissa olevaan tietoon ja niiden auktoriteetti on kielenkäyttäjien keskuudessa yleensä suurin (vrt. Airaksinen 1993: 29–35). Lisäksi niiden tekeminen on yleensä yhteiskunnan rahoittamaa, joten tekijät vastaavat sanakirjojen laadusta paitsi tiedeyhteisölle myös ympäröivälle yhteiskunnalle. Esimerkiksi internetissä tavallisten kielenkäyttäjien täydennettävissä olevilta tai pääasiassa kaupallisia tarkoituksia silmälläpitäen tehdyiltä sanakirjoilta ei voi vaatia samanlaista tieteellistä ja ammattieettistä vastuunottoa. Myös niin sanotut tieteelliset sanakirjat ovat eri lähtöviivalla sen suhteen, millaista tutkimustietoa kielestä on valmiiksi saatavilla. Esimerkiksi monista vähemmistökielistä on tehty vain vähän tutkimusta, mikä voi kasvattaa leksikografian vastuuta kielen normittajana. Yhteiskunta ja muut rahoittajat myös ohjaavat eri hankkeisiin eri tavalla resurssseja.

Tieteellisen sanakirjatyön ihanteet ovat samantyyppiset kuin tieteellisen tutkimuksen ja tietokirjoittamisen. Yleisellä tasolla leksikografian ammattietiikkaa määrittävät samat tutkimuseettiset periaatteet kuin muutakin tieteellistä työtä. Sanakirjatyössä on kuitenkin myös omat alakohtaiset erityispiirteensä, jotka on otettava huomioon toimitustyön eettisten periaatteiden määrittelyssä. Konkreettisimmillaan leksikografia ohjaavat sekä tieteellisesti että eettisesti oikeanlaisiin ratkaisuihin kunkin sanakirjan omat toimitusohjeet, joiden tarkoitus on muun muassa taata sanakirjan laatu ja antaa ohjeita käytännön työhön.

Eettisesti hyväksyttävä sekä luotettava ja laadukas tieteellinen tutkimus edellyttää hyvän tieteellisen käytännön noudattamista. Sen periaatteisiin kuuluvat rehellisyys, huolellisuus ja tarkkuus sekä tutkimustyössä, tulosten tallentamisessa että esittämisessä. Myös muiden tekemien tutkimusten ja tulosten arvioinnissa täytyy noudattaa samoja periaatteita. Lisäksi hyvään tieteelliseen käytäntöön kuuluvat tieteellisen tutkimuksen kriteerit täyttävien ja eettisesti kestävien tiedonhankinta-, tutkimus- ja arviointimenetelmien käyttö sekä avoimuus tutkimuksen tulosten julkaisussa. Eettisesti kestävä tutkimus on lisäksi suunniteltu, toteutettu ja raportoitu yksityiskohtaisesti ja tieteelliselle tiedolle asetettujen vaatimusten mukaisesti. (Ks. tarkemmin TENK: käytäntö.)

Sanakirjoissa perinteisesti käytetty formaatti tiivistää valtavasti tietoa pieneen tilaan. Yhtä sanaa ei voida sanakirjassa käsitellä kovin pitkästi, eikä kaiken valmiin teoksen taustalla olevan tutkimuksen yksityiskohtainen raportointi ole mahdollista. Näin ollen edes toinen leksikografi ei yleensä voi pelkästä sanakirjasta saamansa informaation nojalla tarkistaa teoksessa esitettyjä tietoja. Sen lisäksi, että sanakirjan syntyprosessia on vaikea arvioida valmiin teoksen perusteella, leksikografien valta korostuu tavallisiin tutkimusraportteihin verrattuna myös siksi, että useimpien sanakirjojen pitäisi pystyä välittämään mahdollisimman tarkkaa ja totuudenmukaista tietoa kielestä erityisesti tavallisille kielenkäyttäjille, jotka eivät itse voi saada tietoa suoraan lähteestä ja joilla ei myöskään ole keinoja tarkistaa heille annettujen tietojen oikeellisuutta. Suurelle yleisölle kirjoittava onkin erityisessä vastuussa tiedon totuudenmukaisuudesta, laadusta ja ajantasaisuudesta. (Tietokirjoittajan eettisiä ohjeita on julkaissut esimerkiksi Suomen tietokirjailijat ry, ST 2012.) Tutkijan on tärkeää tiedostaa oma valtansa (vrt. Airaksinen 1993: 29–35). Varsinkin ammattimaisesti tehtyjen sanakirjojen kohdalla hyvä osoitus tutkijan vallasta on monien sanakirjankäyttäjien luja usko siihen, että sanakirjoista löytyy lopullinen totuus sanojen merkityksistä, kielestä ja jopa maailmasta (vrt. Vanhatalo 2005a: 2).

Tavallisten kielenkäyttäjien keskuudessa yleinen luulo tuntuu olevan esimerkiksi se, että sanakirjasta voi tarkistaa, onko jokin sana olemassa kielessä vai ei. Jotta sanakirjankäyttäjä voisi ymmärtää useiden sanakirjojen kahtalaisen roolin deskriptiivisinä ja normatiivisinä teoksina (ks. myös Grönros tk.), käyttäjällä olisi oltava ainakin jonkinlaiset perustiedot sanakirjojen toimitusprosessista ja siinä tehtävistä valinnoista.

Itse sanakirjassa ainoa paikka toimitustyössä tehdyistä valinnoista kertomiseen on teoksen johdanto. Monien muiden käyttöohjeiden tapaan myös sanakirjojen johdannot jäävät kuitenkin käyttäjiltä yleensä lukematta (käyttäjätutkimuksista tarkemmin mm. Welker 2010). Sanakirjojen ”perustoimintoja” osataan varmasti käyttää ilman ohjeistusta, mutta todellisuudessa monet sanakirjankäyttäjät tuntevat formaatin vain pintapuolisesti. Johdantojen lisäksi sanakirjatyön läpinäkyvyyttä ja sanakirjankäyttäjien tietoisuutta toimitusprosessista voidaan lisätä kertomalla työssä tehtävistä valinnoista avoimesti ja yleistajuisesti mediassa sekä lisäämällä sanakirjojen käyttöön liittyvää opetusta esimerkiksi kieltenopettajien koulutuksessa.

Varsinaista leksikografian alan koulutusta on Suomessa tarjolla niukasti (ks. esim. Hiidenmaa 2007: 444), ja ammatti opitaankin yleensä vasta työn kautta. Myös leksikografiaan liittyvää tutkimusta on tehty Suomessa vähän, eikä esimerkiksi leksikografian käsitteitä ja teoreettisia lähtökohтия pohtivan metaleksikografian alalta ole juuri kirjoituksia (poikkeuksia mm. Hormia 1961; Romppanen 2001; Vanhatalo 2005b; Hoyer 2012). Esimerkiksi Matti Kuusi (1985) on kiinnittänyt huomiota siihen, että tieteellinen mielenkiinto sanakirjoihin on ollut Suomessa merkillisen vähäistä. Yksittäisille ihmisille ja eri sanakirjojen toimituksiin on kuitenkin kertynyt paljon sanakirjatyötä koskevaa hiljaista tietoa.

Tutkimuseettisen neuvottelukunnan (TENK: käytäntö) mukaan vastuu hyvän tieteellisen käytännön noudattamisesta ja tutkimuksen eettisyydestä kuuluu koko tiedeyhteisölle. Sitoutumisesta tieteen tekemisen eettisiin periaatteisiin vastaa jokainen tutkija itse, mutta vastuu on myös kollektiivista. Eettisesti vastuullisen tiedeyhteisön toimintatapoihin kuuluu myös eettisistä kysymyksistä keskusteleminen. Tieteellistä tutkimusta harjoittavien organisaatioiden tulee huolehtia siitä, että hyvään tieteelliseen käytäntöön perehdyttäminen ja tutkimusetiikan opettaminen ovat osa niiden antamaa koulutusta. Tutkijankoulutuksessa hyvää tieteellistä käytäntöä ja ammattietiikkaa tulisi tarkastella myös oman alan erityiskysymyksien kannalta. (Ks. myös Ammattietiikka yliopistoissa 1999: 8–21.)

Useilla ammattialoilla on viime vuosikymmeninä julkaistu alakohtaisia eettisiä ohjeistuksia. Leksikografian alalla omaa eettistä ohjeistusta ei ainakaan toistaiseksi ole, mutta esimerkiksi kielitieteen, tietokirjoittamisen, tulkkien ja kääntäjien sekä historian- ja kulttuurintutkimuksen ja museoalan eettiset ohjeistukset sopivat osin sovellettaviksi myös eri-

laisten sanakirjojen toimitustyössä. Näillä aloilla keskeiset tutkimuskohdetta ja -menetelmiä sekä arkistointia koskevat kysymykset ovat yhteisiä monille humanistisille tieteille leksikografia mukaan lukien. (Ks. esim. VTEO 1982; MES 2005; HLEO 2008; LSA 2009; ST 2012.) Nimenomaan alakohtaisen ohjeiston laatiminen voisi kuitenkin olla paikallaan muun muassa siksi, että tutkijoiden on usein vaikea tunnistaa oman alansa eettisesti ongelmallisia kysymyksiä ilman asianmukaista opetusta ja ohjeistusta (Clarkeburn–Mustajoki 2007: 22–25; Ammattietiikka yliopistoissa 1999: 8–21).

Tieteellisen ja ammatillisen vastuun lisäksi leksikografilla on myös yhteiskunnallinen vastuunsa, sillä sanakirjoja laaditaan aina jotakin tarkoitusta varten. Pelkän yleisen tiedonintressin lisäksi sanakirjoilla pyritään yleensä vastaamaan myös jonkinlaiseen käytännön tarpeeseen joko tiedeyhteisössä tai yhteiskunnassa laajemmin. Sekä kuvattavan kielimuodon valintaa että muita valintoja tehdessä onkin olennaista miettiä, millaisia sanakirjoja käyttäjät ja yhteiskunta tarvitsevat, miksi kyseistä sanakirjaa ylipäänsä tehdään ja kenen siitä on tarkoitus hyötyä ja millä tavalla. Myös leksikografit ovat tiedostaneet käyttäjänäkökulman tarpeellisuuden ja pitävätkin nykyään ensisijaisena ohjeenaan lukijoidensa tuntemista (Atkins–Rundell 2008: 5). Käyttäjien ja yhteiskunnan tarpeita on syytä sanakirjatyössä pohtia myös tulevaisuuden kannalta. Olennaisia kysymyksiä ovat esimerkiksi aineistojen säilytystavat sekä sanakirjojen julkaisuformaatin valinta.

Sitä, mikä on yhteiskunnassa toivottavaa, ei voida määritellä tieteellisesti, vaan yhteiskunnallisten tavoitteiden määrittely on aina filosofinen ja eettinen kysymys (Raatikainen 2006: 102). Suomalaisessa yhteiskunnassa vallitsevia arvoja ovat esimerkiksi pyrkimys yhdenvertaisuuteen ja tasa-arvoon. Sanakirjat lisäävät kielenkäyttäjien tietoa kielestä ja edistävät siten kielellistä tasa-arvoa. Kielen objektiivinen kuvaus ja yhteiskunnassa toivottavana pidettyjen arvojen edistäminen voivat kuitenkin olla ristiriidassa esimerkiksi siksi, että kieli ei suinkaan aina kuvaa maailmaa tasapuolisesti. Voidaankin pohtia, onko sanakirjoissa aina pyrittävä kuvaamaan kieltä mahdollisimman objektiivisesti vai voidaanko niissä kielenhuollollisen normittamisen lisäksi pyrkiä ohjailemaan kielenkäyttöä myös arvosidonnaiselta kannalta. Selvää on, että vaikka yhteiskunnassa periaatteessa vallitsevat tietyt useimmille ihmisille yhteiset perusarvot, käytännön moniarvoisessa yhteiskunnassa kaikkien kielenkäyttäjien

mielipiteen huomioiminen vaikkapa sanojen merkityksistä ei ole mahdollista.

Yksilön omat arvot eivät aina ole samoja kuin ne, jotka yhteiskunnassa yleisimmin tunnustetaan. Leksikografin käytännön työtä ohjaavat paitsi ammattietiikka ja sanakirjatyön yhteiskunnallisesti määritellyt tavoitteet myös työntekijän henkilökohtaiset ominaisuudet. Jokainen ammatinharjoittaja on myös yksilö, jonka henkilökohtaiset arvot ja esimerkiksi kiinnostuksen kohteet sekä ammattitaidon taso vaikuttavat työhön (vrt. Airaksinen 1993: 19–20; Ollila 1997: 87–94; Tuominen–Wihersaari 2006: 256–258) – usein alitajuisesti ja tiedostamatta. Erilaisten arvojen taustalla voivat olla esimerkiksi poliittiset tai uskonnolliset näkemykset, mutta myös henkilökohtaiset kokemukset, pelot, uskomukset ja niin edelleen. Arvojen ja asenteiden lisäksi myös muut työntekijän henkilökohtaiset ominaisuudet voivat vaikuttaa työhön. Esimerkiksi leksikografin vahva kiinnostus johonkin tiettyyn alaan tai harrasteeseen voi vaikuttaa sanakirjoihin otettavien sanojen valintaan ja sitä kautta sanakirjankäyttäjien mahdollisuuksiin saada tietoa jonkin erikoisalan sanastosta.

Kokoelman sisältö

Tämä teos sisältää tämän tekstin lisäksi yhdeksän artikkelia, joissa käsitellään sanakirjatyötä erilaisista näkökulmista. Käsitellyiksi tulevat sekä monet käytännön valintatilanteet ja ongelmat että leksikografian periaatteet. Teoksen kantavia teemoja ovat leksikografin ammattietiikan ja yhteiskunnallisten vastuuden pohtiminen. Ammattietiikkaa käsitellään teoksessa ensisijaisesti kielenkuvauksen tarkkuuden, oikeellisuuden ja laadun näkökulmasta. Monet eettisistä kysymyksistä ovat samalla myös tieteellisiä ja käytännöllisiä kysymyksiä. Keskeisiksi nousevat muun muassa objektiivisuuden ja variaation suhde sekä normatiivisuuden ja deskriptiivisyyden rajankäynti. Leksikografin yhteiskunnallisen vastuun ja kielipolitiikan kannalta tärkeitä kysymyksiä ovat kuvattavan kielen ja kielimuodon valinta sekä yksilön kielellisten oikeuksien sekä yhdenvertaisuuden ja tasa-arvon toteutuminen.

Tätä taustoittavaa kirjoitusta seuraavassa artikkelissa Igor Kudashev käsittelee sanakirjan suunnittelua ja siihen vaikuttavia tekijöitä. Kudashev osoittaa artikkelissaan sanakirjaprojektin huolellisen suunnittelun ja

kompromissien tekemisen taidon tärkeyden. Eija-Riitta Grönros käsittelee artikkelissaan sanakirjojen normatiivisuuteen eli ohjailevuuteen ja deskriptiivisyyteen eli kuvailevuuteen liittyviä valintoja ja rajankäyntiä. Kaisa Lehtosalo ja Pirkko Kuutti käsittelevät artikkelissaan historiallisten sanakirjojen toimittamiseen liittyviä erityiskysymyksiä. He summaavat tekstissään myös suomalaisen sanakirjatyön historiaa.

Seuraavat kolme artikkelia käsittelevät erilaisten vähemmistökielten sanakirjoja. Kimmo Granqvist kirjoittaa Suomessa jo pitkään puhutun mutta ulkopuoliselle varsin vieraaksi jääneen romanikielen sanakirjatyöstä. Klaas Ruppel puolestaan käsittelee Suomessa uudempien vähemmistökielten sanakirjatyötä ja kuvaa myös maahanmuuttajasanakirjoihin yleisesti liittyviä kysymyksiä. Sirkka Saarisen artikkelissa siirrytään Suomen ulkopuolella puhuttavien suomalais-ugrialaisten vähemmistökielten sanakirjatyön pariin. Saarinen kuvaa neljän suomalais-ugrilaisen sanakirjan toimitustyössä vastaan tulleita valintatilanteita ja niihin vaikuttaneita tekijöitä.

Laura Tyysteri kuvaa artikkelissaan yhteiskunnassa vallitsevien arvojen heijastumista sanakirjoihin sekä pohtii, miten arvosidonnaisen sanaston erityinen herkkyys voitaisiin huomioida sanakirjatyössä entistä paremmin. Nina Pilke ja Jaana Puskala käsittelevät artikkelissaan sitä, miten sanakirjan käyttäjä sanakirjan näkee ja millaista vastuuta kielialan ammattilaisen tulisi sanakirjan käyttäjänä tuntea. Kokoelman päättää Raimo Jussilan sanakirjatyöhön liittyviä tekijänoikeuskysymyksiä käsittelevä artikkeli, jossa kirjoittaja pohtii muun muassa siteeraamista, lähteiden käyttöä ja merkitsemistä sekä leksikografin omia oikeuksia.

Teoksen ensisijainen tarkoitus on nostaa esiin uudenlainen näkökulma sanakirjoihin sekä välittää eteenpäin sanakirjantekijöille kertynyttä hiljaista tietoa. Teoksessa pohditaan, miten valintatilanteet on käytännön sanakirjatyössä yleensä huomioitu tai miten ne voitaisiin huomioida. Keskustelu uudenglaisista näkökulmista avaa mahdollisuuksia kehittää suomalaista leksikografiaa. Työskentelytapojen ja työssä tehtävien valintojen pohtiminen on ajankohtaista myös sanakirjojen tulevaisuuden kannalta, sillä nopeasti kehittyvä tekniikka ja uudenglaiset julkaisualustat tuovat mukanaan sekä uudenglaisia vaatimuksia että hienoja mahdollisuuksia.

IGOR KUDASHEV

Sanakirjan suunnittelu ja siihen vaikuttavat tekijät

Dieu seul pourrait dicter un dictionnaire parfait.

(Vain Jumala pystyy laatimaan täydellisen sanakirjan.)

Pierre Claude Victor Boiste, ranskalainen leksikografi (1765–1824)

Sanonta ”hyvin suunniteltu on puoliksi tehty” pätee myös sanakirjatyöhön. Käsittelen tässä artikkelissa seuraavia sanakirjan suunnitteluun liittyviä aiheita:

- Miksi sanakirjat tulee suunnitella?
- Mitkä sanakirjan ominaisuudet tulee suunnitella?
- Mitkä tekijät vaikuttavat sanakirjan suunnitteluun ja sanakirjatyöhön?
- Millaisia eettisiä kysymyksiä leksikografi joutuu pohtimaan sanakirjahankkeen aikana?

Artikkelini perustuu pitkälti väitöskirjaani Проектирование переводческих словарей специальной лексики (= Erikoiskielen sanakirjojen suunnittelu kääntäjien tarpeisiin – Kudashev 2007) sekä useassa sanakirjahankkeessa (ks. Kudashev 2000; Suomalais-venäläinen metsäsanakirja 2008; Puusta katsoen 2009; Kudasheva–Kudashev 2008) saatuun käytännön kokemukseen. Tutkijana olen erikoistunut erityisesti kaksija monikielisiin erikoiskielen sanakirjoihin. Ehdottamani sanakirjojen suunnittelumenetelmä on kuitenkin melko universaali, koska se perustuu sanakirjan olennaispiirteisiin.

Koska väitöskirjani on venäjänkielinen, viittaa usein myös venäläisiin lähteisiin. Toivon tämän tuovan lisäarvoa niille lukijoille, jotka osaa- vat venäjää mutta eivät tunne venäläistä tieteellistä tuotantoa leksiko-

grafian alalta. Leksikografialla on Venäjällä pitkät perinteet, ja erityisesti 1970–1990-luvuilla on ilmestynyt useita korkeatasoisia väitöskirjoja, monografioita ja artikkelikokoelmia. Niistä voi ja kannattaa ottaa opiksi vielä tänäkin päivänä.

Miksi sanakirjoja tulee suunnitella?

Sanakirjojen ja sanakirjahankkeiden suunnittelu on tärkeää useasta syystä. Ensinnäkin jokainen sanakirja on ainutlaatuinen. Suunnittelulla pyritään vastaamaan paremmin sanakirjan kohderyhmän tarpeisiin ja kuvaamaan käsiteltävää leksikaalista aineistoa parhaalla mahdollisella tavalla. Esimerkiksi kaksikielisen erikoiskielen sanakirjan tärkeimpiä kohderyhmiä ovat alan asiantuntijat, opiskelijat, tekniset kirjoittajat ja kääntäjät. Näillä kohderyhmillä on eri tarpeita muun muassa tietomäärän ja aineiston järjestyksen suhteen. Erikoisalan opiskelijat ja kääntäjät tarvitsevat paljon selityksiä ja lisätietoja käsitteistä, kun taas tekniset kirjoittajat ja erityisesti alan asiantuntijat eivät niitä juuri kaipaa. Opiskelijalle suunnatun sanakirjan luontevin järjestys on temaattinen (aiheen mukaan), kun taas kääntäjät suosivat ensisijaisesti aakkosjärjestystä. Alan asiantuntijoiden ja teknisten kirjoittajien mieltymykset saattavat riippua tässä suhteessa esimerkiksi siitä, käyttävätkö he sanakirjaa vieraskielisten tekstien lukemiseen vai kirjoittamiseen. Koska sanakirja ei yleensä pysty palvelemaan kaikkia kohderyhmiä ja käyttötilanteita yhtä hyvin, joudutaan suunnittelun yhteydessä tekemään priorisointia ja kompromisseja.

Toiseksi sanakirja vaatii systemaattisuutta. Käyttäjät odottavat, että samanlaiset ilmiöt kuvataan sanakirjassa aina samalla tavalla. Järjestelmällisyys ja tarkkuus ovat sen vuoksi ensisijaisia tekijöitä sanakirjoja laadittaessa. Niitä on kuitenkin vaikea saavuttaa ilman suunnittelua ja ohjeistusta, varsinkin jos sanakirjahankkeeseen osallistuu monta tekijää.

Kolmanneksi suunnittelu on tärkeää käytännön syistä. Kuuluisa leksikografi Ladislav Zgusta kirjoitti aikoinaan *Manual of Lexicography* -käsikirjassaan, ettei hän tunne yhtään sanakirjaprojektia, joka olisi saatu päätökseen alkuperäisellä budjetilla ja aikataululla (Zgusta 1971: 348). Yksityiskohtainen suunnittelu on tärkeä keino saada sanakirjahankkeen aikataulu ja budjetti kuriin.

Mitä sanakirjan ominaisuuksia tulee suunnitella?

Mistä sanakirja koostuu?

Sanakirjan suunnittelua voi verrata uuden talon suunnitteluun. Ensin suunnitellaan kokonaisuutta ja vasta sitten ruvetaan suunnittelemaan yksittäisiä rakenteita: perustusta, seiniä, kattoa, LVI- ja sähköverkkoa ja niin edelleen. Suunnittelu on yleensä syklistä, koska tiettyä rakennetta koskevat valinnat saattavat vaikuttaa muihin rakenteisiin ja niiden kautta kokonaisuuteen (Grinev 1995: 19, 20, 56; Béjoint 2000: 37).

Niin talo kuin sanakirjakin ovat sen verran monimutkaisia kohteita, että ne on pakko pilkkoa suunnittelun aikana pienempiin osiin. Sanakirjan osia ja rakenteita kuvaavat käsitteet ovat tästä syystä erittäin tärkeässä asemassa sanakirjaa suunniteltaessa.

Sanakirjan rakenteeseen on perinteisesti viitattu termeillä *mikrorakenne* ja *makrorakenne*. *Mikrorakenne* tarkoittaa sana-artikkeleiden rakennetta, kun taas *makrorakenne* viittaa artikkeleiden välisiin suhteisiin (esimerkiksi aakkosjärjestykseen). Sittenmin on huomattu, että sanakirjassa voi olla muitakin rakenteita, joiden kuvaamiseen perinteiset termit eivät riitä. Esimerkiksi rungon lisäksi sanakirjassa on usein sellaisia osia kuin esipuhe, käyttöohjeet, lähdeluettelo ja erilaisia liitteitä (engl. *outside matter*), jotka voivat sijaita rungon edessä (*front matter*), sisällä (*middle matter*) tai perässä (*back matter*). Rungon ja lisäosien järjestyksen kuvaamiseen on esitetty termejä *megarakenne* (*megastructure*, ks. Hartmann 2001: 59) ja *kehysrakenne* (*frame structure*, ks. Bergenholtz–Tarp 1995: 211). Lisäksi on kirjoitettu muun muassa viittausten muodostamasta rakenteesta (*cross-reference structure*, ks. Bergenholtz–Tarp 1995: 215 tai *mediostructure*, ks. Hartmann 2001: 65), hakukeinojen muodostamasta rakenteesta (*access structure*, ks. Bergenholtz–Tarp 1995: 219) sekä leksikografisten kuvausten jakaumasta (*distribution structure*, ks. Bergenholtz–Tarp 1995: 188). Uudet käsitteet ja termit eivät ole vakiintuneet eivätkä siirtyneet leksikografien aktiiviseen sanavarastoon, mutta niiden esittäminen osoittaa perinteisen mallin puutteellisuuden.

On myös todettu, että makro- ja mikrorakenteeseen perustuva sanakirjamalli ei ole universaali vaan se sopii parhaiten yksikielisen selittävän sanakirjan rakenteen kuvaamiseen (Hausmann–Wiegand 1989: 344). Sen sijaan esimerkiksi kaksi- ja monikielisillä sanakirjoilla on omat eri-

koispiirteensä, jotka vaativat toisenlaista käsitteistöä (vrt. Kromann–Rieber–Rosbach 1991: 2717; Kudashev 2007: 51–64).

Tietyt piirteet ovat kuitenkin ominaisia kaikille sanakirjoille. Väitöskirjassani olen termiopissa käytettävän käsiteanalyysin avulla löytänyt joukon sanakirjan olennaispiirteitä, jotka erottavat sanakirjan muista hakuteoksista, kuten tietosanakirjoista ja hakemistoista (Kudashev 2007: 15–65). Koska nämä olennaispiirteet koskevat kaikkia sanakirjoja, ne muodostavat hyvän lähtökohdan minkä tahansa sanakirjan suunnitteluun. Seuraavaksi käsittelen lyhyesti näitä olennaispiirteitä.

Sanakirjan olennaispiirteet

Sanakirja on artefakti eli ihmisen laatima esine. Ihmiset luovat artefakteja palvelemaan jotakin tarkoitusta ja kohderyhmää. Sanakirjan tarkoitus ja kohderyhmä ovatkin se perusta, josta sanakirjan suunnittelu lähtee liikkeelle.

Sanakirjan yläkäsitteenä voidaan pitää *teos*-käsitettä. Kannattaa ottaa huomioon se, että sanakirja ei välttämättä ole itsenäinen teos vaan se voi olla jonkin toisen teoksen (esimerkiksi oppikirjan) osa tai kuulua laajempaan kokonaisuuteen, esimerkiksi sanakirjasarjaan. Tällaisessa tapauksessa sanakirjan suunnittelu on osa emoteoksen tai sarjan suunnittelua, ja sanakirjan suhde emoteokseen tai sarjaan on otettava suunnittelussa huomioon. Rungon lisäksi sanakirja yleensä sisältää myös muita osia: esipuheen, käyttöohjeet, lähdeluettelon, hakemiston, erilaisia liitteitä ynnä muuta. Lisäosien sisältö ja järjestys ovat niin ikään suunnittelun kohteina.

Sanakirjan esittämiseen tarvitaan aina tietovälinettä. Muinoin kyse oli savitauluista ja papyruksista, sittemmin paperista ja nyt yhä enenevässä määrin elektronisista laitteista. Leksikografia ei toistaiseksi ole onnistunut ottamaan kaikkea hyötyä irti elektronisista tietovälineistä, ja suurin osa sähköisistä sanakirjoista onkin korkeintaan ”tietokoneistettuja” (Tarp 2009: 28; vrt. Schryver 2003: 146; Atkins 2002: 11). Tällaisten sanakirjojen selailu saattaa olla nopeampaa hakukoneen ja hyperlinkkien avulla, mutta perimmältään kysymys on samasta vanhasta paperikauden rakenteesta. Yksinkertaisemmassakin tapauksessa tietovälineellä on kuitenkin suuri vaikutus sanakirjan toimintaperiaatteeseen. Tulevaisuudessa elektroniset välineet saattavat mullistaa myös sanakirjan rakenteita niin, että staattiset sana-artikkelit korvautuvat dynaamisilla *näkymillä* ja staattiset

hakusanalistat ”virtuaalisanakirjoilla”, joita käyttäjät pääsevät räätälöimään haun ja asetusten avulla (Atkins 2002; Varantola 2002; ks. myös Kudashev–Carlson–Kudasheva 2010).

Muiden hakuteosten tapaan sanakirjoissa on aina kuvauksen kohteita ja niihin liittyviä kuvauksia, teemoja ja reemoja. Sanakirjan laatijan on suunniteltava, mitä kohteita sanakirjassa kuvataan ja millaisia tietoja näistä kohteista halutaan antaa.

Vaikka sana-artikkeleiden otsikoita kutsutaan leksikografiassa vanhasta muistista hakusanoiksi, on tärkeä ymmärtää, että edes yleiskielen sanakirjoissa sana ei ole ainoa mahdollinen kuvauksen kohde. Kuvauksen kohteina voivat olla myös morfeemit, yhdyssanojen osat, lyhenteet, sanaliitot (esim. monisanaiset termit) ja jopa kokonaiset lauseet (esim. fraseologismit ja sanonnat). Sanakirjassa kuvauksen kohteet ovat kuitenkin tavalla tai toisella leksikaalistuneita eli niiden muoto on pieniä vaihteluita lukuun ottamatta vakiintunut. Vapaamuotoisia lausumia ja kokonaisia tekstejä ei sanakirjaan voi ottaa, vaan niiden paikka on esimerkiksi korpuksissa ja käännösmuistiohjelmissä.

Kuvauksen kohteita ei valita sanakirjaan sattumalta, vaan ne muodostavat aina jonkinlaisen kokonaisuuden. Muuten sanakirjan käyttäjien olisi mahdotonta arvata, mitä löytyy mistäkin sanakirjasta. Sanakirjaa suunniteltaessa täytyy aika varhaisessa vaiheessa miettiä, mitkä kielelliset merkit ovat kuvauksen kohteena ja muodostavatko ne sellaisen järjestelmän, jonka käyttäjä pystyy hahmottamaan.

Sanakirjan tarkoitus on auttaa käyttäjiä kielellisten merkkien tulkinnaissa, käytössä tai korvaamisessa, joko samassa kielessä tai eri kielten välillä. Tämä tavoite on se olennaispiirre, jonka avulla pystyy erottamaan sanakirjat muista hakuteoksista, kuten tietosanakirjoista tai käsikirjoista. Vaikka hakusanat ja kuvaukset olisivat esimerkiksi ruoanlaittotaidon sanastossa ja keittokirjassa osittain samoja, sanaston tavoitteena on kuvata ruoanlaittoon liittyviä termejä eikä auttaa ruokien laittamisessa. Sanakirja voi sisältää myös ensyklopedisia tietoja mutta vain silloin, kun nämä tiedot palvelevat edellä mainittua perustavoitetta.

Jos kuvauksen kohteita on sanakirjassa useita, kuvaukset saattavat poiketa toisistaan. Kuvausten sisältö ja esittämistapa on silloin suunniteltava jokaiselle kuvauksen kohteelle erikseen ja sitten mietittävä, miten ne voidaan järkevästi yhdistää. Usein käy niin, että yksi ja sama esittämistapa ei sovi kaikkien kohteiden tai tietotyyppien esittämiseen. Esimerkiksi yleis-

kielen sanakirjoissa käytettävä sanakeskeinen malli ei oikein sovi monisanaisten yksiköiden – kuten sanaliittojen, idiomien tai sananlaskujen – esittämiseen, minkä takia ne usein jäävät kuvaamatta tai niitä on vaikea paikantaa. Sen sijaan terminologiassa käytettävä käsitekeskeinen malli ei sovi yksittäisten sanojen ominaisuuksien esittämiseen. Käsitelähtöisestä tietueesta ei löydy yksittäisille sanoille omistettuja lokeroita eikä haku-sanan sisälle pysty sijoittamaan kovin paljon tietoja. Sama kuvaus jouduttaisiin myös epäkäytännöllisesti toistamaan niin monta kertaa kuin jokin sana esiintyy sanastossa. Tästä johtuen erikoiskielen sanastoissa on useimmiten varsin vähän kielioppitietoja (Bergenholtz–Tarp 1995: 111). Käytännössä leksikografi saattaa siis joutua tekemään kompromisseja kuvauksen täydellisyyden sekä yhtenäisyyden ja selkeyden välillä.

Vaikka jokainen edes hieman tavallista vaikeampi sana on monografi-an arvoinen (Šerba 1958: 72), edes tieteellinen sanakirja ei saa muistuttaa tutkielmaa (Devkin 2000: 17). Muiden hakuteosten tapaan sanakirjat on tarkoitettu antamaan tiivistettyä tietoa. Sanakirja on myös järjestettävä niin, että tiedonhaku on mahdollisimman nopeaa ja helppoa. Tämä tarkoittaa sitä, että sanakirjan laatijan on suunniteltava tiivistämiskeinoja (esimerkiksi lyhenteitä, erikoismerkkejä, viittauksia ja toistuvia tietoja korvaavia merkkejä) sekä hakua nopeuttavia ja helpottavia keinoja, kuten sisällysluetteloa, otsikointia, numerointia, ylä- ja alatunnisteita, hakemis-toja, muotoilua ja graafisia keinoja.

Seuraavassa tiivistän sanakirjan olennaispiirteet luetteloksi.

- Sanakirja on itsenäinen teos tai osa muuta teosta tai sarjaa.
- Sanakirja sijoitetaan jollekin tietovälineelle.
- Sanakirjan tärkein osa on tietyn järjestelmän muodostavat kielelliset merkit leksikografisine kuvauksineen.
- Sanakirjan tarkoitus on antaa sellaisia tietoja, jotka auttavat käyttäjiä kielellisten merkkien tulkinnessa, käytössä tai korvaamisessa.
- Sanakirja antaa tietoja tiivistetyssä muodossa.
- Sanakirja on järjestetty hakua nopeuttavalla tavalla.

Näiden olennaispiirteiden perusteella voidaan erottaa seuraavat sanakirjan ominaisuudet, jotka leksikografian on suunniteltava:

- kohderyhmä(t)
- sanakirjan tarkoitus
- mahdollinen suhde muihin teoksiin (kuten emoteokseen tai sanakirjasarjaan)

- tietoväline
- leksikografisen kuvauksen kohteet
- leksikografisen kuvauksen sisältö
- sanakirjan osat sekä niiden järjestys ja rakenne
- tiivistämis- ja navigointikeinot.

Itse sanakirjan lisäksi on suunniteltava myös sanakirjahanke: prosessi ja sen eteneminen.

Sanakirjan suunnitteluun vaikuttavat tekijät

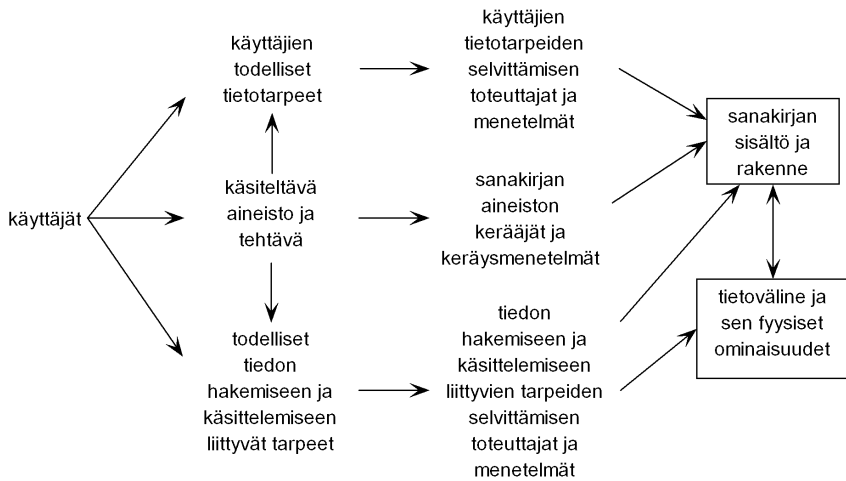
Sanakirjan suunnitteluun vaikuttavat tekijät voidaan jakaa *varsinaisiin leksikografisiin tekijöihin*, jotka perustuvat käyttäjien tarpeisiin, ja *ulkoisiin tekijöihin*, jotka heijastavat ulkomaailman asettamia rajoituksia (Kudashev 2007: 66). Aloitetaan varsinaisista leksikografisista tekijöistä, jotka ovat suurimmaksi osaksi sanakirjan olennaispiirteiden johdannaisia.

Varsinaiset leksikografiset tekijät

Hakuteoksena sanakirja on tarkoitettu käyttäjien tietotarpeiden tyydyttämiseen. Käyttäjien tietotarpeet syntyvät jotakin kieliaineistoa käsiteltäessä, kuten tekstiä käännettäessä. Tietotarpeiden lisäksi käyttäjillä on tiedon hakemiseen ja käsittelyyn liittyviä tarpeita, jotka vaikuttavat tietovälineen valintaan ja sanakirjan fyysisiin ominaisuuksiin (vrt. esim. tieteellinen painettu suursanakirja tai tietokanta vs. turistin taskusanakirja vs. tulkin rakennustyömaalla käyttämä sanaviuhka tai mobiilisovellus).

Ellei sanakirjaa laadita vain omaan käyttöön, leksikografilla ei yleensä ole tarkkoja tietoja käyttäjien todellisista tarpeista. Tästä syystä sanakirjan ominaisuudet riippuvat paljolti myös siitä, kuka kerää tietoja sanakirjan käyttäjistä ja heidän tarpeistaan ja miten tietoa kerätään. Varsinaiset leksikografiset tekijät ja niiden vaikutus sanakirjan ominaisuuksiin on esitetty kuviossa 1.

Varsinaiset leksikografiset tekijät saattavat kilpailla keskenään. Tämä heijastaa niin sanottuja ”leksikografian antinomioita” (ks. Rej-Delesal’ 1983; Dubičinskij 1995; Devkin 2000: 16–23) eli kaikille sanakirjoille ominaisia rajoituksia ja sisäisiä ristiriitoja.



KUVIO 1. Varsinaiset leksikografiset tekijät ja niiden vaikutus sanakirjan ominaisuuksiin.

Tietotarpeet saattavat olla ristiriidassa tiedon käsittelemiseen liittyvien tarpeiden kanssa. Esimerkiksi aakkosjärjestys toisaalta nopeuttaa hakuja mutta toisaalta katkaisee semanttisia suhteita. Toisaalta käyttäjä odottaa sanakirjalta mahdollisimman paljon tietoja, mutta toisaalta taas tietojen runsaus vaikeuttaa tarvittavan tiedon paikallistamista, kasvattaa paperisanakirjan volyymiä ja nostaa sen hintaa. Sanakirjan laatija ei myöskään voi ottaa aivan kaikkia käyttäjiä huomioon, jolloin toiset saattavat valittaa turhista tiedoista ja toiset taas sanakirjan suppeudesta.

Vaikka järjestelmällisyys on sanakirjoille ominaista, ne eivät voi olla järjestelmällisiä joka näkökulmasta. Aineiston lajittelu yhden kriteerin mukaan tuo epäjärjestelmällisyyttä muiden kriteereiden osalta. Sähköisten sanakirjojen räätälöinti voi tuoda helpotusta näihin ongelmiin, mutta käytännössä siinä ei vielä ole päästy kovin pitkälle (esim. Sobkowiak 2002; Granger–Paquot 2009: 90), ja vastaava tutkimus on vielä lapsenkengissä (Trap–Jensen 2010: 1134).

Koska kaikkia kohderyhmiä ei pysty ottamaan huomioon, kannattaa lähteä liikkeelle pääkohderyhmän tarpeista. Ihannetapauksessa on suoritettava oma käyttäjätutkimus, esimerkiksi kysely tai haastattelu. Aikaisempiin käyttäjätutkimuksiin on syytä suhtautua varovasti, koska ne ovat usein maa- ja projektiokohtaisia ja vanhentuvat melko nopeasti. Olen myös

huomannut erikoiskielen sanakirjojen käyttäjäkyselyn tuloksia analysoidessani, että käyttäjät eivät osaa arvioida sellaisia innovaatioita, joista heillä ei vielä ole kokemusta. Tästä syystä kyselyt sopivat pikemminkin vanhojen käytäntöjen tehokkuuden arvioimiseen kuin uusien käytänteiden kehittämiseen. On myös esitetty, että sanakirjan on parhaimmillaan toimittava standardien asettajana ja käyttäjien ”kasvattajana”, jolloin sen suunnittelu ei voikaan perustua pelkästään keskivertokäyttäjien toivomuksiin ja mieltymyksiin (esim. Grinev 1986: 23). Olen tästä asiasta osittain samaa mieltä.

On käytännössä mahdotonta antaa yleispäteviä ohjeita siitä, miten leksikografisia ristiriitoja tulisi ratkaista. Jokainen sanakirjatiimi ratkaisee näitä ongelmia tapauskohtaisesti, joten subjektiivisten tekijöiden merkitys ja vaikutus on leksikografiassa edelleenkin erittäin suuri.

Ulkoiset tekijät

Varsinaisten leksikografisten tekijöiden lisäksi sanakirjan suunnitteluun vaikuttavat myös erilaiset ulkoiset tekijät. Sanakirjatyö vaatii resursseja – henkilöstöresursseja, työinfrastruktuuria ja rahaa. Tahot, jotka tarjoavat näitä resursseja, kuten kustantaja, rahoittaja, työnantaja ynnä muut, saattavat asettaa sanakirjalle ja sanakirjaprojektille omia vaatimuksiaan. Myös laajemmalle ympäristölle – kuten tietylle työryhmälle, koulukunnalle, kieliyhteisölle tai yhteiskunnalle – ominaiset perinteet ja ennakkoodotukset vaikuttavat leksikografian valintoihin.

Ulkomaailman rajoitukset leikkaavat siivet monilta teoreettisesti oikeilta ajatuksilta. Esimerkiksi Lev Šerban ehdotusta siitä, että kaksikielisiä sanakirjoja pitäisi olla neljä jokaista kieliparia kohti, koska erikieliset käyttäjät tarvitsevat eri tietoja äidinkieleen ja vieraaseen kieleen käännettäessä, ei ole koskaan taloudellisista syistä voitu toteuttaa (Berkov 1971: 5).

Jotta sanakirja vastaisi käyttäjien tarpeita parhaalla mahdollisella tavalla, sanakirjan suunnittelu kannattaa kuitenkin aloittaa leksikografisista tekijöistä ja sitten muokata suunnitelmaa ulkoiset tekijät huomioon ottaen. Myös tasapainoilu teorian ja käytännön realiteettien välillä on tärkeää, koska sanakirja on ensisijaisesti käytännön työkalu. Kuten Samuel Johnson on todennut jo 1700-luvulla, leksikografiassa puhdas teoria ei sovi käytännön tehtävien ratkaisemiseen. Sellaisesta koneesta, jota

insinööri ihailee mutta tavallinen työntekijä ei pysty käyttämään, ei ole paljon hyötyä (Johnson 1747: 4–5).

Sanakirjojen suunnittelu on siis pitkälti kompromissien etsimistä kilpailevien tekijöiden välillä: leksikografisten ja ulkoisten, objektiivisten ja subjektiivisten, teoreettisten ja käytännönläheisten. Siinä missä on valintoja, voi kysyä myös valintojen eettisyydestä. Siirryn seuraavaksi käsittelemään leksikografian etiikkaa.

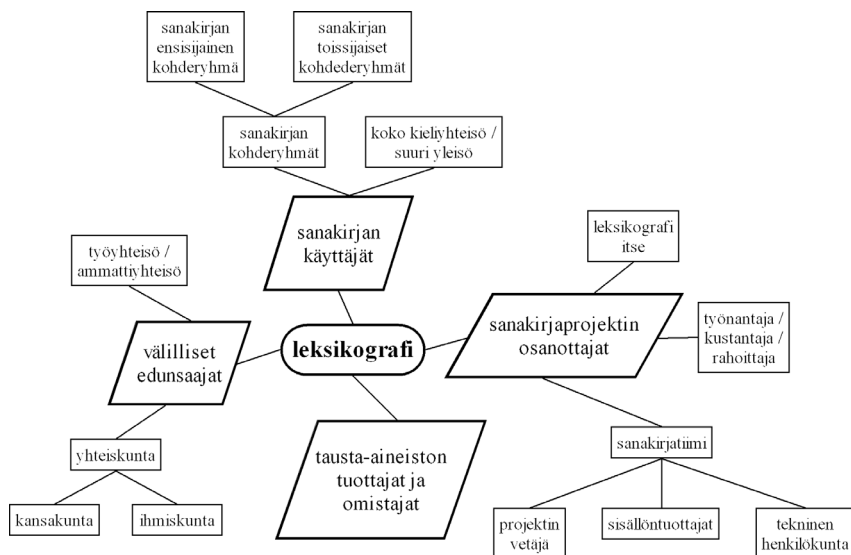
Leksikografian vastuu ja etiikka

Etiikassa on perimmältään kysymys yksilön suhteesta muihin. Kun tämän suhteen laatu on luonteeltaan neutraali tai positiivinen, toisin sanoen kun yksilön toiminta tai toimimattomuus ei vahingoita ketään tai tuo hyötyä, eettiset kysymykset harvoin tulevat esille. Etiikasta ruvetaan puhumaan yleensä silloin, kun jonkun toiminta loukkaa tai voi aikaa myöten loukata muiden intressejä. Eettisyys ei myöskään ole jyrkästi duaali käsite (joko–tai), vaan sitä mitataan suhteessa tuotettuun (tai odotettuun) hyötyyn tai vahinkoon.

Leksikografian eettisten valintojen ymmärtämiseksi täytyy tarkastella häntä ympäröivää ”ekosysteemiä”, joka koostuu sanakirjatyön välittömistä ja välillisistä edunsaajista, muista sidosryhmistä sekä sanakirjahankkeen osanottajista (ks. kuvio 2).

Jos sanakirjahankkeen tavoitteena on sellaisen sanakirjan laatiminen, josta on oikeasti hyötyä käyttäjille, käyttäjien tarpeiden huomioon ottaminen on leksikografian tärkeimpiä prioriteetteja. Aivan kaikkia käyttäjiä sanakirja ei kuitenkaan pysty palvelemaan yhtä hyvin. Erään sanonnan mukaan kaikille tarkoitettua sanakirjaa ei tule käyttämään juuri kukaan (Dubičinskij 1994: 16). Tästä syystä leksikografian tulee jakaa käyttäjät ensisijaisiin ja toissijaisiin käyttäjiin sekä ”suureen yleisöön”. Tämän priorisoinnin ansiosta leksikografi pystyy keskittymään ensisijaisesti pääkohderyhmän tarpeisiin.

Vaikka sanakirjan näkökulma olisikin puhtaasti deskriptiivinen (kuvaileva) eikä normatiivinen (ohjaileva), sanakirjan käyttäjät yleensä pitävät sanakirjaa auktoriteettina tai erotuomarina, jonka avulla he pystyvät ratkaisemaan kielenkäyttöön liittyviä ongelmia ja kiistoja (Gak 1977:12; vrt. Dubois–Dubois 1971: 49–50; Zgusta 1971: 212; Denisov



KUVIO 2. Sanakirjahankkeen ekosysteemi.

1976: 32; Varantola 2001: 219). Jokainen sanakirja on myös eräänlainen ”kielimuseo”, joka kuvaa aikakautensa kielenkäyttöä (Algeo 1989: 29). Tästä syystä leksikografi on vastuussa myös koko kieliyhteisölle, haluaa hän sitä tai ei.

Varsinkaan isoja sanakirjoja ei enää laadita yksin vaan tiimin voimin. Sanakirjatiimin jäsenten välille syntyy usein jännitteitä, jotka liittyvät vastuun, vallan, rahan ja kunnian jakoon. Nämä jännitteet voimistuvat yleensä projektin loppua kohti, ja niiden ehkäisemiseksi on hyvä sopia asioista ennakkoon, mieluummin kirjallisesti. Sopimus ei sinänsä ole eettisyyden tae, mutta se tekee pelisäännöt selviksi. Kaikesta ei tietenkään voi sopia etukäteen, eikä sopimukseen saa korvata omaatuntoa.

Leksikografi on vastuussa sanakirjahankkeen etenemisestä työnantajalle, kustantajalle ja rahoittajalle, jotka eivät välttämättä ole sama taho ja joilla saattaa olla omia tavoitteita ja intressejä. Myös näiden tahojen kanssa on sovittava selkeistä pelisäännöistä ja solmittava kirjalliset sopimukset. Sanakirjan laatimissuunnitelma sekä projektisuunnitelma on syytä hyväksyttävä resurssien tarjoajilla ja myös pitää suunnitelmista kiinni. Sanakirjahankkeet kestävät yleensä useita vuosia, jolloin toisaalta

laatimissuunnitelma ja toisaalta ulkoinen tilanne saattavat muuttua paljonkin. Kummankin osapuolen on oltava joustava, ja muutostarpeista on informoitava yhteistyökumppaneita mahdollisimman pian.

Sanakirjan laatija ei voi olla kokoaikaisesti vain tiettyyn sanakirjaprojektiin sitoutunut leksikografi, vaan hänellä on muitakin ammatillisia ja sosiaalisia rooleja. Leksikografina on osattava olla eettinen myös itseään kohtaan eli hyvässä mielessä egoistinen, mutta toisaalta hänen on tietysti tasapainotettava omat intressinsä muiden intressien kanssa. Suhde muihin on kaksisuuntainen, eli myös muut osapuolet ovat vastuussa leksikografille. Tämä vastuu ei myöskään ole aina mitattavissa vain rahassa.

Eräitä uniikkeja sanakirjoja lukuun ottamatta sanakirja ei yleensä synny tyhjästä vaan se perustuu aikaisempiin sanakirjoihin ja muihin teoksiin. Se, kuinka paljon ja millaisia tietoja muista teoksista saa lainata, on selkeitä tapauksia lukuun ottamatta vaikea ja pitkälti avoin kysymys leksikografiassa (Hausmann 1989; Williams 1992; Gerd 2005; Jussila tk.). Aikaisempien lähteiden mainitseminen tai merkitseminen on kuitenkin tärkeää niin juridisista kuin eettisistäkin syistä.

Sanakirjoista hyötyvät epäsuorasti myös muut kuin sanakirjan käyttäjät. Tieteellisesti ja hallinnollisesti niistä hyötyvät esimerkiksi sanakirjan laatijan ammattiyhteisö ja organisaatio. Sanakirjan käytöstä saadun hyödyn (tai vahingon!) myötä sanakirja vaikuttaa myös koko yhteiskunnan toimintaan. Onko sattumaa, että Japanin ”talousihmettä” on edeltänyt erikoiskielen sanastojen määrän raju kasvu Japanissa (Marčuk 1992: 4)? Ja vaikka joitakin sanakirjoja ei pysty yhdistämään talouskasvuun yhtä suoraviivaisesti, ne ovat joka tapauksessa tärkeä osa kieli- ja kulttuuriperintöä. Uskaltaisin väittää, että sanakirjojen määrä ja laatu ovat kielen, kulttuurin ja koko kieliyhteisön hyvinvoinnin tärkeimpiä indikaattoreita. Yhteiskunnan tulisi ymmärtää ja tunnistaa sanakirjojen tuoma hyöty eikä kuvitella, että sanakirjoja syntyy itsestään (vrt. Tammilahti 1988 : vii-ix).

Kuten edellä olevasta kuviosta 2 näkee, sanakirjan laatija on usein puun ja kuoren välissä. Saattaa helposti käydä niin, että kun yhdelle sidosryhmälle kumartaa, toiselle pyllistää. Käytännössä leksikografi joutuu priorisoimaan, kenen intressejä hän ottaa huomioon ensisijaisesti ja kenen vain mahdollisuuksien mukaan. Myös eettisyys on dynaaminen ja suhteellinen käsite. Sen sijaan että kysyy, onko jokin teko tai menetelmä eettinen vai ei, tulisikin mielestäni mieluummin kysyä, kuinka eettistä tämä on olosuhteisiin nähden – ja ketä kohtaan.

Tiivistelmä

Olen käsitellyt tässä artikkelissa sanakirjan suunnittelua ja siihen vaikuttavia tekijöitä. Suunnittelu on tärkeää, koska sen avulla pystyy räätälöimään sanakirjan vastaamaan paremmin kohderyhmän tarpeita, ohjaamaan sanakirjatyötä ja pitämään sanakirjahankkeen aikataulun ja budjetin kurissa.

Sanakirjan osia, rakenteita ja muita ominaisuuksia kuvaava termistö ei ole vakiintunut, vaan vaihtelee huomattavasti eri käsikirjoissa. Perinteinen mikro- ja makrorakenteeseen perustuva malli on liian karkea eikä sovi kaikkien sanakirjojen suunnitteluun.

Yksi tapa lähestyä sanakirjan suunnittelua on määritellä sanakirja artefaktiksi ja löytää tämän kautta kaikille sanakirjoille ominaiset olennaispiirteet. Näiden piirteiden perusteella voi listata sanakirjan ominaisuuksia, jotka leksikografian on suunniteltava. Suunniteltaviin ominaisuuksiin kuuluvat sanakirjan kohderyhmä(t), tarkoitus, mahdollinen suhde emoteokseen tai sarjaan, tietoväline, leksikografisen kuvauksen kohteet ja sisältö, sanakirjan osien järjestys ja rakenne sekä tiivistämis- ja navigointikeinot. Sanakirjan suunnittelu on yleensä syklistä, koska tiettyyn rakenteeseen liittyvät valinnat saattavat vaikuttaa muihin rakenteisiin ja niiden kautta kokonaisuuteen.

Sanakirjan suunnitteluun vaikuttavat tekijät voidaan jakaa varsinaisiin leksikografisiin tekijöihin, jotka perustuvat käyttäjien tarpeisiin, ja ulkoisiin tekijöihin, jotka heijastavat ulkomaailman asettamia rajoituksia. Käyttäjien tarpeet voidaan jakaa tietotarpeisiin sekä tiedon hakemiseen ja käsittelyyn liittyviin tarpeisiin. Ellei sanakirjaa laadita vain omaan käyttöön, leksikografilla ei yleensä ole tarkkoja tietoja käyttäjien todellisista tarpeista. Tästä syystä sanakirjan ominaisuudet riippuvat paljolti myös siitä, kuka kerää tietoja käyttäjistä ja heidän tarpeistaan ja miten kerätään. Varsinaiset leksikografiset tekijät saattavat kilpailla keskenään. Tämä heijastaa kaikille sanakirjoille ominaisia sisäisiä ristiriitoja ja rajoituksia, niin sanottuja ”leksikografian antinomioita”.

Ulkoisiin tekijöihin kuuluvat esimerkiksi resursseja tarjoavien tahojen vaatimukset sekä laajemmalle ympäristölle – kuten tietylle työryhmälle, koulukunnalle, kieliyhteisölle tai yhteiskunnalle – ominaiset perinteet ja ennako-odotukset. Ulkoiset tekijät saattavat kilpailla niin keskenään kuin varsinaisten leksikografisten tekijöidenkin kanssa. Sanakirjojen

suunnittelu onkin pitkälti kompromissien etsimistä kilpailevien tekijöiden välillä – leksikografisten ja ulkoisten, objektiivisten ja subjektiivisten, teoreettisten ja käytännönläheisten.

Eettisten valintojen ymmärtämiseksi täytyy tarkastella leksikografia ympäröivää ”ekosysteemiä”, joka koostuu sanakirjatyön välittömistä ja välillisistä edunsaajista, muista sidosryhmistä sekä sanakirjahankkeen osallistujista. Käytännössä leksikografi joutuu priorisoimaan, kenen intressejä hän ottaa huomioon ensisijaisesti ja kenen vain mahdollisuuksiensa mukaan. Eettisyys ei ole staattinen eikä absoluuttinen käsite, joten jyrkän vastakkainasettelun asemesta kannattaa kysyä, kuinka eettistä tämä on olosuhteisiin nähden – ja ketä kohtaan.

EIJA-RIITTA GRÖNROS

Sanakirjojen normatiivisuus ja deskriptiivisyys

Normatiivisuus liittyy sanakirjoihin monella tavalla. Joissakin sanakirjoissa se voi olla jo työn lähtökohta, kuten esimerkiksi useimmissa erikoisalan sanakirjoissa ja sanastoissa. Ne pyrkivät usein vakiinnuttamaan ja yhtenäistämään alan terminologiaa ja ohjaavat siksi suositettavien termien käyttöön. Yleiskielen sanakirjoissa suhtautuminen normatiivisuuteen vaihtelee. Monissa esitetään erilaisia sanastoon ja kielenkäyttöön liittyviä ohjeita ja suosituksia. Joissakin pyritään puhtaasti deskriptiivisyyteen, mutta se on vaikeasti toteutettavissa. Deskriptiivisessä kuvauksessakin on aina tehtävä erilaisia valintoja ja rajauksia. Sanakirjan käyttäjä saattaa myös tulkita normatiiviseksi sen, mikä sanakirjan tekijän mielestä on kuvausta. Tarkastelen tässä artikkelissa sitä, miten tietoinen tai tahaton normatiivisuus ja toisaalta deskriptiivisyys toteutuvat eri kielten yksikielisissä yleiskielen sanakirjoissa. Keskeisellä sijalla ovat suomen yleiskielen sanakirjat ja varsinkin Kielitoimiston (KS, 2004–) sanakirja, jonka päätoimittajana työskentelen, mutta käsittelen teemaa myös yleisellä tasolla ja tarkastelen muidenkin kielten yleiskielisten sanakirjojen ratkaisuja.

Mitä tarkoitetaan normatiivisella ja deskriptiivisellä sanakirjalla?

Sanakirjojen luokittelussa deskriptiivisellä tarkoitetaan sanakirjaa, jonka päämääränä on kielen kuvaus, ja normatiivisella (tai preskriptiivisellä) sanakirjaa, joka pyrkii vaikuttamaan kielenkäyttöön. Normatiivinen sanakirja voi olla normeja luova, jos on kyse esimerkiksi kirjakielen luo-

misesta, tai olemassa olevia normeja kuvaava. (Nordisk leksikografisk ordbok 1997, s.v. *normativ ordbok*.)

Käsitteet ”deskriptiivinen sanakirja” ja ”normatiivinen sanakirja” eivät ole suoranaisesti toistensa vastakohtia. ”Deskriptiivinen” kuvastaa leksikografian suhdetta kielelliseen aineistoon ja tarkoittaa, että aineisto ja kuvaus vastaavat toisiaan. ”Normatiivinen” taas kuvastaa leksikografian suhdetta sanakirjan käyttäjään ja tämän tarpeeseen saada ohjeita tekstin tuottamiseen. Käytännössä puhtaasti deskriptiivisiä tai puhtaasti normatiivisia sanakirjoja ei kuitenkaan juuri ole; kyse on enemmänkin aste-eroista. Deskriptiivinen sanakirja myös heijastaa aina jollain tavoin leksikografian kielellisiä arvotuksia, jotka lukija voi tulkita normiksi, ja normatiivinen sanakirja taas lähtee aina deskriptiiviseltä perustalta. (Svensén 2004: 29.)

Nyky-suomen sanakirjan (NS, 1951–61) laatiminen käynnistyi, kun eduskunta kehotti 1927 hallitusta ryhtymään toimenpiteisiin ”suomen kirja- ja sivistyskielen kantasanakirjan aikaansaamiseksi”. Sanakirjan piti sisältää ”kaikki nykyisessä suomen kirja- ja sivistyskielessä käytetyt ja käytettäväksi hyväksyttävät sanat tarpeen mukaan määriteltyinä ja selitettyinä”. (NS: v.) Tätä tehtävänantoa tekijät tulkittivat niin, että sanakirjaan tuli sekä deskriptiivinen että normatiivinen puoli. Deskriptiiviseen puoleen kuului, että sanakirja antoi kuvan aikansa kielenkäytöstä. Siihen kuuluivat sanojen erilaiset käyttötavat, merkitykset, tiedot sanan yleisyydestä tai harvinaisuudesta, tyyliarvosta ja mahdollisesta vanhentuneisuudesta. Tekijät olivat kuitenkin tietoisia siitä, että jo tässä on mukana myös normatiivinen puoli – kuvaushan ei välttämättä vastaa jokaisen sanakirjan käyttäjän käsitystä näistä sanoista. Selvemmin normatiivisuutta edustivat sanakirjan viittaukset suositettavampaan muotoon tai hylättävästä ilmauksesta hyväksyttävään. Niidenkin suhteellisuudesta tekijät olivat tietoisia: ratkaisut voivat olla jonkun mielestä paikoin harhaan osuneita, jonkun mielestä normatiivisuutta voi olla liikaa ja jonkun mielestä liian vähän. Silti katsottiin, että molemmat puolet olisivat hyödyksi käyttäjille. Normatiiviset kannanotot olivat välttämättömiä, mutta myös lavean deskription katsottiin antavan kielestä tietoa sille, joka sitä tarvitsee ja hakee. (Sadeniemi 1952: 5.)

Saako kieltä normittaa?

Sanakirjat ovat pohjimmiltaan kielen kuvauksia. Kuvauksen ja normatiivisuuden suhde on vaihdellut sanakirjoissa eri aikoina ja eri syistä. Normittamiseen on voinut olla erityistä tarvetta esimerkiksi kirjakielen kehityksen tietyissä vaiheissa. Myös sanakirjan luonne, lähtökohta ja kohderyhmä ovat voineet määrittää sanakirjan kuvaustapaa. Toisaalta myös yleinen suhtautuminen normatiivisuuteen ja käsitykset siitä, missä määrin sitä painotetaan, ovat vaihdelleet eri aikoina ja erilaisissa yhteiskunnallisissa tilanteissa. Esimerkiksi ensimmäisten yleiskieltä kuvaavien englannin sanakirjojen päämääränä oli kuvata kieltä deskriptiivisesti, mutta 1700-luvulla leksikografit alkoivat katsoa velvollisuudekseen osoittaa, mikä on huonoa ja mikä hyvää käyttöä (Béjoint 1994: 100). Ranskan leksikografia sai alkunsa 1500-luvulla, mutta vasta 1600-luvun lopulla ilmestyi sanakirjoja, joilla oli merkittävä asema kielen normituksessa. Yksi niistä oli Ranskan Akatemian sanakirja, josta on ilmestynyt myöhemmin useita painoksia. Akatemia julkaisee edelleen suosituksia ja ohjeita, jotka ovat hyvin puristisia, mutta suurimmat ja käytetyimmät nykyranskan sanakirjat eivät ole välttämättä noudattaneet niitä. (Jokinen 1986: 8–9.) Se ei kuitenkaan tarkoita, että nämä sanakirjat eivät olisi normatiivisia. Normatiivisuus voi tulla myös valintojen, rajausten ja kuvausten kautta. Esimerkiksi Le Nouveau Petit Robert (2007) merkitsee monet anglismit erityisellä *anglic.*-lyhenteellä. Se ei viittaa pelkästään alkuperäkieleen, joka on osoitettu erikseen, ja joihinkin vanhoihin, kieleen jo ”hyväksytyihin” englannin sanoihin (esim. *week-end*) tätä merkintää ei ole liitetty. Merkintä viestittää, että sillä osoitettuihin sanoihin ei suhtauduta neutraalisti, kuten lyhenteen selitteestä ilmenee (mts. xxxi).

Varsinkin viime vuosikymmeninä leksikografian piirissä on yleisesti tähdennetty puhtaasti deskriptiivisyyden tärkeyttä. On koettu epämu-kavaksi se yleinen sanakirjan käyttäjien odotus tai oletus, että sanakirja antaa ohjeita siitä, miten kieltä ”pitäisi” käyttää. (Atkins–Rundell 2008: 2.) Osittain pyrkimys deskriptiivisyyteen on kuitenkin teoreettinen: normatiivisuudesta pyritään usein sanoutumaan irti ja deskriptiivisyyttä korostetaan varsinkin sanakirjojen esipuheiden kuvauksissa, vaikka itse sanakirjassa normittavuutta olisikin (Nikula 1992: 49).

Suomessa suurilla yleiskielen sanakirjoilla on ollut aina yhteys viralliseen kielenhuoltoon. Kun Nykysuomen sanakirjaa ryhdyttiin tekemään

valtion rahoituksella, käytännön toteuttajana oli Suomalaisen Kirjallisuuden Seura, jonka yhteydessä oli vuodesta 1928 toiminut myös kielenkäyttöä koskevia suosituksia pohtinut kielivaliokunta. Kielivaliokunnan työtä jatkoi myöhemmin suomen kielen lautakunta ja 1945 perustettu Kielitoimisto, josta tuli 1976 osa Kotimaisten kielten tutkimuskeskusta eli Kotusta. Suomen kielen perussanakirja (PS, 1990–94) ja sen 2000-luvulla ilmestynyt uudempi laitos Kielitoimiston sanakirja (KS, 2004–) on tehty Kotuksessa. Nämä sanakirjat ovat siis aina edustaneet kielenhuollon ja suomen kielen lautakunnan suosituksia, ja niille on annettu normatiivinen tehtävä. Kotimaisten kielten keskukselta annetun lain (2012) perusteluissa todetaan edelleen, että KS kuvaa suomen nykyisen yleiskielen sanastoa, sen käyttöä ja yleiskielen normeja.

Kielessä on erilaisia normeja. Päivi Rintala on luokitellut suomen kirjakielen normit kolmeen ryhmään: otettuihin, tehtyihin ja korrektiivisiin normeihin. Otetut normit on otettu kirjakielen normiksi vakiintuneen käytön perusteella. Niitä on noudatettu jossakin kielimuodossa jo ennen kuin niistä on tullut nykykirjakielen normeja, esimerkiksi varhaisnykysuomen kaudella kirjakielessä tai laajasti murteissa. Tehdyt normit taas koskevat tapauksia, joihin puhekieli ei ole antanut mallia, esimerkiksi monimutkaisia lauserakenteita, joita ei käytetä puhekielessä, sekä oikeinkirjoitusta, lyhenteitä, välimerkkejä ja niin edelleen. Korrektiiviset normit ovat sellaisia tehtyjä normeja, jotka tähtäävät kielen parantamiseen, ja niiden noudattaminen on ainakin osittain makuasia. Näiden normien perustelut ovat voineet olla arvoväritteisiä: jotakin ilmausta on saatettu perustella esimerkiksi niin, että se on ”suomalaisempi” tai ”tarkoituksenmukaisempi”. Kielenkäytöltä on saatettu vaatia johdonmukaisuutta semanttiset ja pragmaattiset näkökohdat unohtaen. (Rintala 1992: 55–62.)

Kielenhuollon periaatteista keskusteltaessa on erotettu kaksi näkemystä: pyritäänkö huoltamaan kieltä sen rakennetta vai käyttöä silmällä pitäen. Kieli voidaan nähdä järjestelmänä, jonka selkeyttä, puhtautta ja alkuperäisyyttä tulee vaalia, tai se voidaan nähdä kommunikaation välineenä, jonka ohjailussa painottuvat vuorovaikutukselliset, sosiaaliset ja kulttuuriset tekijät ja jonka muutokset voidaan hyväksyä, vaikka ne sotisivat rakenteen symmetriaa ja alkuperäisyyttä vastaan. Jälkimmäinen periaate on ollut etusijalla suomen kielenhuollon suosituksista päätettäessä ainakin 1970-luvulta lähtien. (Mantila 2010: 179.)

Tämänsuuntaisia suositusten muutoksia on tehty paljon NS:n jälkeen ilmestyneissä sanakirjoissa. Esimerkiksi svetisismeihin ei ole enää ollut tarvetta suhtautua siten kuin NS:aa tehtäessä (1930- ja 1940-luvulla) ja vielä sen jälkeenkin, jolloin ne nähtiin uhkana. Myös moniin epäortodoksisiin taiputusmuotoihin ja johdoksiin on suhtauduttu ymmärtävämmmin. PS:aa laadittaessa 1980-luvulla ja myöhemmin KS:n aikana 2000-luvulla sanakirjojen toimitukset ovat tehneet suomen kielen lautakunnalle lukuisia esityksiä suositusten muuttamiseksi. Esimerkiksi PS:aan otettiin aikaisemmin esitettyjen taiputusmuotojen rinnalle yleisiä, analogian kautta syntyneitä rinnakkaismuotoja (esim. *viiveen, oria*), samoin sellaisia ”poikkeuksellisesti” muodostettuja sanoja kuin *huilisti* ja *pianismi*, joiden sopivuudesta oli keskusteltu 1970-luvulla. Vastaavasti esimerkiksi *huomioida*-sanan merkitys ’ottaa huomioon’ hyväksyttiin suomen kielen lautakunnan päätöksen mukaan CD-Perussanakirjaan (1997), *ulko-maalainen*-sanan adjektiivinen käyttö KS:n ensimmäiseen versioon (KS 2004), samoin kuin verbit *uutisoida*, *entisöidä* ja *eläköityä*, joiden sijaan oli aiemmin esitetty niiden johtotavan vuoksi muita ilmauksia. Myöhemmin on poistettu suositus esimerkiksi *alipalkattu*-sanan yhteydestä (aiempi suositus *alipalkkainen*) samoin kuin sanoista *muistorikas* ja *tietorikas* (KS 2012). Muutetuissa on ollut yleensä kyse siitä, että suositusta ei ole noudatettu sen olemassaolon aikana juuri lainkaan, tai siitä, että suositusten mukaiset ilmaukset eivät ole olleet luontevia. (PS:n ja KS:n muutoksista ks. Lehtinen 1994, Grönros 2004 ja 2012a, Haapanen 2008.)

Suomessa on käyty kielentutkijoiden keskuudessa keskustelua siitä, saako kielenkäyttöä ylipäänsä normittaa ja jos saa, niin millä tavoin. Keskustelu on koskenut erityisesti niitä näkyviä ohjeita tai suosituksia, joilla kielenhuollossa ja sanakirjoissa on otettu kantaa joihinkin ilmauksiin. Kriittisissä kirjoituksissakin on kuitenkin yleensä todettu, että kirjakieli tarvitsee normitetun rekisterin, mutta kieltä ei pitäisi normittaa kovin pitkälle (esim. Leino 2002 [1996]: 342).

Yleiskielen normeilla – samoin kuin kirjakielen vakiinnuttamisella aikoinaan – on pyritty siihen, että kaikilla olisi käytettävänäan yksi yhteinen kielimuoto erilaisten puhekielen varianttien rinnalla. Kielenhuollon tavoitteena on ollut, että yleiskielestä annetut suositukset ja ohjeet saataan kaikkien tietoon muun muassa sanakirjan avulla. Kielenhuolto kohdistuu pelkästään kirjakieleen, ei puhekieleen. Kun puhekielen käyttöala on yleisesti laajentunut viime vuosikymmeninä, sen on nähty aiheuttavan

painetta yleiskielen normin muuttamiselle. On esitetty, että kirjakielen annettaisiin lähentyä puhekieltä niin, ettei tarvitsisi miettiä, kumpi kahdesta kielellisestä variantista on oikein, vaan puhekielen variantti hyväksyttäisiin yleiskieleen (Paunonen 1996: 553). Ongelma on kuitenkin se, että puhekielen variantteja on usein enemmän kuin yksi. Kaikkien mukaan ottaminen yleiskieleen ja sen kuvaukseen voisi olla hankalaa, mutta jos valitaan vain jokin tai joitakin, se on jälleen uusi rajausta ja kannanotto.

Variaatio kuuluu kieleen, ja myös monia puhekieliseksi katsottuja muotoja ja ilmauksia on siirtynyt ja siirtyy jatkuvasti yleiskieleen. Kielenhuollon ja kirjakielen muutosten on kuitenkin väistämättä oltava jossain määrin hitaita, jotta ajatus yhteisestä kielestä olisi olemassa. Kieliyhteisö kokonaisuutena on myös yleisesti ottaen konservatiivisempaa kuin kielentutkijat: normien muuttaminen ja variaation hyväksyminen nostattaa usein vastustavia ääniä.

Normatiiviset ohjeet sanakirjoissa

NS:n käsittely- ja esitystapa on luonut pohjan, jota myöhemmissä suomen yleiskielen sanakirjoissa on pitkälti noudatettu. NS:ssa on ehdottomasti noudatettaviksi tarkoitettuja ohjeita, joissa on käytetty risti- ja nuoli-symboleja. Risti osoittaa hylättävää käyttöä (esim. † *Määrän perään* 'määrältä', s.v. *perään*), ja nuoli ohjaa hylättävästä suositeltavaan (*itseuhrautuva* → *uhrautuva*, *itsensä uhraava*). Lisäksi on *par.*-lyhenteen avulla merkittyjä paremmin-suosituksia sanojen käytöstä (*itsenäistyttää par. itsenäistää*). Näiden lisäksi voi joskus olla hakusanan jäljessä suorasanaisia kommentteja (esim. hakusanan *oikeutettu* jäljessä "käyttö suositeltavaa vain selvästi verbaalisessa tehtävässä"). PS:ssa ja KS:ssa on käytetty samanlaista kaksitasoista ohjeistusta: ohje "pitää olla" liittyy esimerkiksi oikeinkirjoitussääntöjen vastaisiin ilmauksiin (esim. *kehoittaa* pitää olla: *kehottaa*). Paremmin-sanalla viitataan suositettavampaan muotoon (esim. *geenimuunneltu* paremmin: *muuntogeeninen*). Samanlainen kaksitasoinen ratkaisu (merkinnät po. 'pitää olla' ja par. 'paremmin') on myös Gummeruksen suuressa suomen kielen sanakirjassa (2004).

KS:n paremmin-suosituksissa on erilaisia tapauksia. On esimerkiksi viittauksia merkityksen tai muodostustavan kannalta säännönmukaisempiin johdoksiin tai muihin sanoihin (esim. *annostelijan* sijaan lait-

teesta paremmin *annostelin*, *alekkain* on paremmin *allekkain* ja *asteettain* paremmin *asteittain*). Jotkin suositukset liittyvät sananvalintaan: esimerkiksi *noin-hinta* on paremmin *arvio-* tai *ohjehinta* tai *likimääräinen hinta*, *kanssakilpailija* on paremmin *kilpakumppani*, *lomarakentaminen* paremmin *loma-asuntojen rakentaminen* ja *miestyöpäivä* sukupuolineutraalisti paremmin *henkilötyöpäivä*. Suosituksia on myös vierassanojen kirjoitusasusta (esim. *carambola* paremmin: *karambola*) tai niiden suomenkielisistä vastineista (*domain-nimi* paremmin: *verkkotunnus*). Suositukset voivat liittyä vain johonkin osaan sanan käytöstä, jolloin suositus on esimerkin yhteydessä (*Alin* [paremmin: *pienin*] *numero*; *Korjauksen alla* paremmin: *korjattavana*). Osa suosituksista liittyy erikoisalan termeihin, jolloin ne ovat alan omia suosituksia eivätkä kielenhuollon kannanottoja (esimerkiksi *viherkaihi* paremmin: *glaukooma*, *kierrätyskuitu* paremmin: *uusiokuitu*).

KS:n eri versioihin on tehty monia normimuutoksia, mutta tarkistusta kaipaavia suosituksiakin on edelleen mukana. Esimerkiksi joidenkin erikoisalojen termien suositukset näyttävät muuttuneen alan omissa julkaisuissa sen jälkeen, kun ne on tarkistettu KS:aa varten. Myös varsinaisia kielenhuollon suosituksia pitää arvioida aina uudelleen. Monet suositukset ovat säilyneet sellaisina kuin ne ovat olleet PS:ssa tai jopa NS:ssa, vaikka tilanne on nykyisin toinen. Esimerkiksi se, että *kaunoluistelun* sanotaan olevan paremmin *taitoluistelu*, on ollut NS:sta PS:aan siirryttäessä aiheellinen ohje, mutta nyt olisi enemmän paikallaan luonnehtia *kaunoluistelua* vanhentuneeksi nimitykseksi. Joitakin muita ilmauksia, joista nyt on suositus, voisi taas merkitä nykyisin ehkä arkikieliseksi ellei sitten jo neutraaliksi yleiskieleksi (esim. *Ongelma on jo sitä luokkaa* [paremmin: sellainen, niin vaikea], *että – –*, *Hankkeen toteutuminen on kiinni* [paremmin: riippuu, kiikastaa] *rahasta*).

Normatiivisista ohjeista voidaan olla eri mieltä, mutta niitä voidaan myös tulkita toisin kuin on tarkoitettu. Sanakirja on hankalammassa tilanteessa kuin kielenhuolto yleensä: kun kielenhuollossa – esimerkiksi Kielitoimiston puhelinneuvonnassa – voidaan keskustella ohjeista ja niiden soveltamisesta erilaisiin tilanteisiin ja käyttöyhteyksiin, sanakirja joutuu esittämään kannan hyvin kategorisesti ja mustavalkoisesti. Se voi johtaa tulkitsemaan asioita toisin ja tiukemmin kuin on tarkoitettu. *Paremm*-sana on myös hiukan ongelmallinen tässä tehtävässä, jonka traditio on synnyttänyt, koska se on merkitykseltään tehtävänsä nähden

liian arvottava. Se on kuitenkin eräänlainen koodi, lyhyt tapa ilmaista asia. Toisaalta pitemmät sanalliset kuvaukset, joita kielenhuollon teksteissä joskus käytetään, ovat usein ongelmallisia nekin: jos käytetään esimerkiksi sellaisia ilmauksia kuin *asiatyylissä*, *asiatekstissä*, *neutraalissa tyyliässä*, ne herättävät yhtä lailla kysymyksiä ja niitä voidaan pitää arvottavina.

Sanakirja voi ottaa kantaa myös muuten kuin näkyvillä ohjeilla. Pelkkä hakusanan muodon valintakin voi olla ohjeellinen. KS:ssa on esimerkiksi vain hakusana *elämäkerta*, ei yleisesti esiintyvää mutta tähän asti vääränä pidettyä muotoa *elämänkerta*. Toisaalta siitä, että jokin sana puuttuu sanakirjasta, ei voi automaattisesti päätellä, että sitä ei pidetä hyvänä tai että sitä ei ole haluttu ottaa sanakirjaan. Sana voi puuttua unohduksen tai muun syyn vuoksi, ja varsinkin uudet sanat voivat puuttua vain siksi, etteivät ne ole vielä ehtineet mukaan.

Ei ole myöskään helppo sanoa, milloin sanan uusi käyttö, uusi rinnakkaisasu tai uusi merkitys on otettava sanakirjaan. *Presidentinvaalit* ja *uutistenlukija* ovat vakiintuneita sanoja, mutta jossain määrin esiintyy myös asuja *presidenttivaalit* ja *uutislukija*. Pitäisikö ne mainita sanakirjassa rinnakkaismuotoina? *Supliikki*-sanaa näkee joskus käytettävän myös adjektiivina; pitäisikö tämä käyttö ottaa mukaan? Kaikkea harvinaista vaihtelua ei pystytä kuvaamaan rajallisessa sanakirjassa, eikä se olisi tarkoituksenmukaistakaan, mutta kaiken yleisen pitäisi olla mukana. On vaikea määrittää sitä, miten yleinen sanan tai käyttötavan pitää olla, kun se otetaan mukaan, ja miten yleisyys mitataan. Mukaanotto ja poisjätöt ovat kuitenkin molemmat kannanottoja.

Sanakirjoissa on erilaisia tapoja esittää normatiivista tietoa. Svensk ordbokissa (SO, 2009) ohjeita on esitetty artikkelien lopussa olevissa erillisissä ruuduissa. Ruuduissa voi olla myös muita tietoja, esimerkiksi sanan etymologiasta, mutta useimmiten ne sisältävät sanan käyttöön liittyviä ohjeita. Ruudussa voi olla pohdiskelua, jossa todetaan jonkin aiemman ohjeen tai käsityksen muuttuneen (esim. s.v. *blommogram*), mutta siinä voi olla myös tiukkoja huomautuksia väärästä käytöstä ja virheistä (esim. s.v. *dess*, *desto*, *fråga*). Samantapaista käytäntöä noudattaa New Oxford Dictionary of English (NODE, 2001), joka antaa ohjeita ja tietoa sanojen oikeasta käytöstä erillisissä *usage*-tietoruuduissa. Ruuduissa voidaan kommentoida sellaisia vanhoja käsityksiä virheellisestä käytöstä, joita ei enää pidetä aiheellisina, mutta annetaan myös voimassa olevia normatiivisia ohjeita sanojen käytöstä (esim. s.v. *if*, *kind*, *majority*).

Den Danske Ordbogin (2003–05) tarkoitus kuvattiin työtä aloitettaessa siten, että se on yleisesti ottaen deskriptiivinen, mutta antaa normatiivisia kannanottoja ortografiasta ja ohjailevaa tietoa muun muassa taivutuksesta, sananmuodostuksesta ja kiinteistä ilmauksista. Mielipiteitä jakavia sanojen uusia merkityksiä sanakirjassa kommentoidaan sanomalla, että merkitys ei monien (tai joskus: joidenkin) mielestä ole korrekki (esim. *stipendiat*-sanalle syntynyt uudempi merkitys 'stipendi'). Vaikka kommentoinnin muotoilu näyttää objektiiviselta, se on silti normatiivinen. (Trap–Jensen 2002: 63–76.)

Virossa deskriptiivisen ja normatiivisen sanakirjan tehtävät on jaettu: Eesti keele seletav sõnaraamat (2009) on deskriptiivinen ja Eesti õigekeelsussõnaraamat ÖS (2006) esittää oikeakielisyyssnormit. Samasta sanasta on siis toisessa sanakirjassa selitetty vain merkitys jättäen kommentit pois ja toisessa on huomautettu siitä, että sanan sijasta suositetaan käytettäväksi toista sanaa.

Vaikka sanakirjassa ei olisi selviä suosituksia, se voi silti sisältää normatiivisia kannanottoja ja rajauksia. Esimerkiksi Cobuild-sanakirjan esipuheessa sanotaan, että sanakirja suhtautuu varovaisesti kielenkäytön uusiin piirteisiin, koska ne voivat osoittautua väliaikaisiksi. Kirja pyrkii esipuheen mukaan kuvaamaan kieltä, joka on tuttua tekijöiden ajattellemalle keskivertokäyttäjälle. (Cobuild 1987: xx–xxi.) Rajauksia on siis tehty valinnoissa ja aineiston mukaanotossa, joten kuvaus ei ole pelkästään deskriptiivistä.

Ovatko tyylilajimerkinnät normittavia?

Yleiskielen sanakirjoissa sanojen kuvaukseen kuuluu lähes aina erilaisia käyttöalamerkintöjä, joilla osoitetaan jotakin lisäpiirrettä verrattuna standardikieleen. Näin kuvataan esimerkiksi johonkin erikoisalaan kuuluvat termit (esim. lääketiede, maatalous) sekä sanat, jotka ovat alueellisesti tai ajallisesti rajoittuneita (murre sanat, vanhentuneet), sosiokulttuurisia (slangi, lastenkieli), asennetta ilmaisevia (halventava, ironinen), frekvenssiltään poikkeavia (harvinainen) tai muodollisuuden asteeltaan eroavia (arkikielinen, ylätyylinen). Merkintäsystemit vaihtelevat eri sanakirjoissa, ja myös yksittäisten merkintöjen käyttöalueet vaihtelevat, joten eri sanakirjojen käyttöalamerkintöjä on usein vaikea verrata kes-

kenään (Svensén 2004: 379). Merkinnt ovat kuitenkin yleisiä eri kielten sanakirjoissa, ja niitä kuvataan yleensä sanakirjan alkusivuilla (esim. NODE: xiv–xvi; SO: xvi).

Kun käyttöalamerkinnällä ilmoitetaan, että ilmaus on standardista poikkeavaa kieltä, huomautus voidaan nähdä arvottavana ja normittavana. Varsinkin rekisteriä kuvaavat tyyllilajimerkinnt, kuten englannin *informal* tai suomen *arkikielinen* (*ark.*), jolla kuvataan paitsi puhekielessä myös asiatyyliä rennommassa kirjoitetussa kielessä käytettäviä sanoja (KS 2012: v), ymmärretään usein ohjeeksi ja normiksi (esim. Leino 2002 [1996]: 341). Leksikografien näkökulmasta tyyllilajimerkintöjä ei ole kuitenkaan tarkoitettu ohjeeksi, vaan osaksi kielen kuvausta. Tätä osoittaa sekin, että yleensä niissäkin sanakirjoissa, jotka ilmoittavat olevansa puhtaasti deskriptiivisiä, on kuitenkin merkitty sanojen tyyllilaji (esimerkiksi Eesti keele seletav sõnaraamatussa on suomen arkikielisyttä vastaava merkintä *kõnekeel*). Merkinntöillä ilmaistaan, että sanaa käytetään epämuodollisissa yhteyksissä, mutta se ei ole kaikissa tilanteissa neutraalia yleiskieltä. Merkinntän pois jättäminen antaisi vajavaisen kuvan sanasta ja sen käytöstä.

Ongelmallisiksi merkinnt tekee se, ettei niiden määrittäminen ole helppoa. Ei ole mitään mittaria tai objektiivista tapaa todeta, että jokin ilmaus on arkikielinen. Leksikografi tekee päätelmät sen perusteella, millaisissa yhteyksissä sanaa yleisimmin käytetään. Myös käsitykset yksittäisten sanojen arkikielisydestä vaihtelevat, ja esimerkiksi eri-ikäisillä voi olla eri käsitys siitä, mikä on arkikieltä ja mikä yleiskieltä. Sanojen tyylliväriin syntyymiseen vaikuttaa usein se, millaisessa tilanteessa ja tekstiympäristössä sanoja on alun perin käytetty. Ensimmäiset käyttöyhteydet vaikuttavat kielenkäyttäjien konnotaatioihin ja määrittelevät sanan tyylliväriin (Mäkinen 1989: 200). Tämän vuoksi esimerkiksi eri sukupolvilla voi olla eri käsitys joidenkin sanojen tyylliarvosta: vanhemmat ovat kokeneet sanan tulon kieleen, nuoremmille se on ollut aina olemassa. Toisaalta koko kieliyhteisönkin käsitykset ilmausten arkisuudesta muuttuvat usein ajan mittaan. Siksi merkintöjä on tarkistettava sanakirjojen uusiin versioihin. (Ks. myös Grönros 2012b.)

Tyyllilajimerkintöjä kritisovien mielestä sanakirjassa ei pitäisi sanoa, onko sana yleiskieltä vai ei, vaan asian ratkaiseminen pitäisi jättää kielenkäytön luonnollisen valinnan huomaa (Leino 2002 [1996]: 342). Samaa tapaan perustelivat myös Webster's Third New International Dictionaryn (W3, 1961) tekijät ratkaisujaan. Täyteen deskriptiivisyyteen

pyrkinyt sanakirja jätti pois sekä normatiiviset kannanotot että myös tyyllilajimerkinnot. Ratkaisu herätti aikanaan paljon kritiikkiä. Sanakirjan toimittajien mukaan se, että tyyllilajileimaa ei merkitty, vaikka lukuisat esiintymät olisivat siihen viitanneet, ei tarkoittanut, etteikö lukija voisi itse liittää leimaa mielessään. Lukijalla ei kuitenkaan ole esiintymätietoja käytettävänä. Leksikografilla on, ja siksi hänen pitäisi tehdä päätelmä. Arvostelu katsoi, että sanakirja oli mieluummin auttamatta lukijaa, kuin otti riskin olla väärässä. (Landau 2001: 255–256.)

Vaikka yksittäisten sanojen arkikielisydestä tai yleiskielisyydestä voidaan olla eri mieltä, yleinen käsitys arkikielisyyden olemassaolosta ja kielen eri rekistereistä on yleensä olemassa. Tyyllilajin merkitsemisellä sanakirja pyrkii palvelemaan sankirjankäyttäjiä kertomalla, että sana ei ehkä ole kaikkien mielestä ja kaikissa yhteyksissä neutraalia yleiskieltä, vaikka se joidenkin mielestä olisikin. Merkinnällä ei ole tarkoitus ohjata kielenkäyttöä eikä kieltää sanan käyttöä, vaan kertoa, miten sana saatetaan kokea. Tämä voi olla tärkeää tietoa nimenomaan niistä sanoista, joiden suhteen käsitykset vaihtelevat. Samoin se on tärkeää tietoa niille, joille kieli ei ole äidinkieli. Kirjoittajan täytyy voida tehdä tietoisia ratkaisuja siinä, käyttääkö tekstissään monien arkikielisenä pitämää sanaa vai ei. Esimerkiksi genreen sopimaton rekisteri tai tyyli voi aiheuttaa ongelmia siinä, miten tekstiin suhtaudutaan ja miten asianmukaisena sitä pidetään (Heikkinen 2012: 11).

Sanakirjantekijät ovat kuitenkin hyvin tietoisia merkintöjen ongelmallisuudesta ja siitä, että ne helposti tulkitaan normittaviksi. Sen vuoksi on pohdittu myös vaihtoehtoja esitystavalle. On esimerkiksi esitetty, että tyyllilajimerkintöjen sijasta kuvattaisiin tekstilajeja, joissa sanat esiintyvät: esimerkiksi englannin *formal* ja *informal* korvattaisiin korpustietojen pohjalta sellaisilla esiintymäyhteyttä kuvaavilla ilmauksilla kuin esimerkiksi *tabloid news*, *broadsheet* tai sellaisilla erikoisalalyhenteitä lähenevillä kuvauksilla kuin *news*, *business*, *financial*, *entertainment*, *academic*, *popular fiction* ja niin edelleen (Landau 2001: 269–272). Tämänäsuuntainen kuvaustapa olisi periaatteessa hyvä, koska ongelma ei ole nykyisin useinkaan se, onko sana arkikielinen vai ei, vaan se, millaisiin yhteyksiin arkikieliset sanat sopivat. Kirjoitetun yleiskielen sijaan yhä useammat tekstit sisältävät kirjoitettua arkikieltä (Lauerma 2012: 54). Tällainen kuvaustapa toisi kuitenkin yhtä paljon ongelmia kuin vanhakin. Esimerkiksi tekstilajit sopisivat huonosti lähtökohdaksi, koska kyse on enemmän rekiste-

ristä, ja samaa rekisteriä käytetään monessa eri tekstilajissa (Voutilainen 2012: 76). Tarkan käyttöyhteyden kuvaaminen olisi sanakirjan tekijän kannalta ongelmallista: mihin sanoihin merkintä tehtäisiin ja miten kuvattaisiin ne arkikieliset sanat, joita ei voisi ryhmitellä tällä tavoin. Lukija taas voisi pohtia, eikö sanoja käytetä tai ”saa käyttää” muissa yhteyksissä.

Tyylilajimerkinnät ovat koodimaisia luonnehdintoja, ja ne ovat aina suhteellisia: suuntaa antavia, aikaan ja käyttöyhteyksiin sidoksissa olevia ja herkästi muuttuvia. Epämuodollisuutta tai arkikielisyyttä ilmaisevat merkintätavat käsittävät myös aina laajan alueen niin, että ne voivat olla joko hyvin lähellä yleiskieltä tai lähellä slangia tai siltä väliltä (Svensén 2004: 379). Tällainen karkea yleisleima voi olla vaikeasti tulkittava, mutta vielä vaikeampi olisi yksityiskohtainen luokitus. On ehdotettu esimerkiksi, että muodollisuuden astetta ilmaistaisiin viisiportaisella asteikolla hyvin epämuodollisesta hyvin muodolliseen siten, että keskellä on tunnusmerkitön nolla (Verkuyyl–Janssen–Jansen 2003: 303). Olisi kuitenkin vaikea sanoa, miten asteet määritettäisiin ja mikä olisi tunnusmerkitöntä. Mitä hienojakoisempi kuvaus on, sitä vaikeampi sitä on toteuttaa ja tulkita. Merkintöjen karkeus ja tulkinnallisuus liittyy muihinkin käytöalalyhenteisiin kuin muodollisuuden astetta kuvaaviin. Ilmausta voidaan luonnehtia esimerkiksi sekä leikilliseksi että halventavaksi. Yleensä esimerkit valaisevat eri käyttötapoja, mutta tilanteinen tulkinta voi jäädä lukijalle. Silti luonnehdinnoista on ehkä enemmän apua kuin siitä, että niitä ei merkittäisi lainkaan.

Voiko sanakirja olla deskriptiivinen?

Normatiivisuuteen liittyy monia eettisiä ongelmia, ja sen vuoksi sanakirjantekijät nykyisin mielellään korostavat deskriptiivisyyttä esipuheissa. Samaa heijastanee sekin, että esimerkiksi NODE esittelee johdannossaan (s. xv) *usage*-tietoruutujensa käyttöä vain sellaisin esimerkein, joissa kumotaan joitakin vanhoja normikäsityksiä ja esitetään nykyinen vapaa-mielisempi kanta. Sanakirjan tietoruuduissa on tällaisten tietojen lisäksi esitetty kuitenkin myös olemassa olevia normeja. Normatiivisuudesta pyritään usein sanoutumaan irti, se koetaan kiusalliseksi (Nikula 2002: 39).

Vaikka monilla sanakirjoilla on pyrkimys deskriptiivisyyteen, sitä ei ole helppo toteuttaa. Sanakirja voi tuskin koskaan olla puhtaasti des-

kriptiivinen. Sanakirjantekijät ovat itsekin tietoisia siitä, että vaikka he sanoisivat esipuheissa vain kuvailevansa kielenkäyttöä, he joutuvat aina tekemään tuhansittain päätöksiä, joihin sisältyy subjektiivisia arvioita (Cobuild 1987: xx–xxi). Kuvattaessa joudutaan aina myös rajaamaan: vaikka kuvattaisiin kuinka runsaasti variaatiota, jossakin on aina raja ja jokin uusin tai harvinaisin jää aina pois. Sekin on valinta tai voidaan tulkita valinnaksi. Yleensä sanakirjassa esitetään se, mikä on tavallista eikä sitä, mikä olisi mahdollista. Silloinkin on kuitenkin otettu kantaa. On vaikea määrittää rajoja sille, milloin sana jätetään pois ja milloin se otetaan mukaan samoin kuin sille, milloin mukaan otettu sana merkitään harvinaiseksi. Onkin sanottu, että kun normatiivisuus perustuu laadun normiin, deskriptiivisyys perustuu määrän normiin: ilmaus otetaan mukaan, jos sitä käyttää tietty määrä ihmisiä. Vaikeutena on kuitenkin määrittää se, miten suuri käyttäjä- tai esiintymämäärä tarvitaan – ja kuka sen päättää. (Béjoint 1994: 100.)

Korpuukset tarjoavat nykyleksikografialle aineistoa ja yleisyystietoja enemmän kuin aikaisemmin on ollut saatavissa, mutta tulkinta ja rajanveto jäävät kuitenkin leksikografian tehtäväksi. Toisaalta myöskään se, että sanakirja kuvaa tarkoin kaikkea korpuksessa olevaa, ei ole tae siitä, että kuvataan kattavasti kielenkäyttöä. Korpuskin on aina rajallinen, ja jo sitä koostettaessa on tehty valintoja. Laajimmat ja monipuolisimmatkaan korpuukset eivät pysty antamaan varmaa vahvistusta sille, että jotakin *ei ole* kielessä, vaikka ne kertovat yleisyydestä paljon (Svensén 2004: 73).

Sanakirjantekijä ei voi myöskään aina vaikuttaa siihen, miten hänen tarkoitustaan tulkitaan. Jos sanotaan, että sanakirja ei kerro, miten sanoja pitäisi käyttää, vaan esittää, miten niitä yleensä käytetään, lukija voi kuitenkin tulkita käytön kuvauksen ohjeelliseksi. Kun englannin sanakirjoissa kuvataan, kuuluuko sana brittienglantiin vai amerikanenglantiin, kuvaus voidaan ymmärtää myös ohjeeksi ja tulkita niin, että sanaa ei pidä käyttää toisessa kielimuodossa. Deskriptiiviseltä näyttävä sanakirja, joka ei ilmoita olevansa normatiivinen mutta joka noudattaa olemassa olevia tai yleisiksi katsottuja normeja valitessaan hakusanoja ja esimerkkejä tai esittäessään taivutusmuotoja ja ääntämisohejeita, on piilonormatiivinen (Nordisk leksikografisk ordbok 1997 s.v. *normativ ordbok*).

Sanakirjan esitystapa vaikuttaa myös siihen, miten asioita tulkitaan. Vaikka pyrittäisiin ottamaan mukaan runsaasti rinnakkaisia muotoja ja

kirjoitusasuja, jokin niistä valitaan päähakusanaksi, johon muista viitataan. Päähakusanaksi valitaan usein vaihtoehtoista yleisin tai alkupe-
räisin, jos sellaisia eroja on mahdollista tehdä, mutta toisinaan rinnak-
kaismuodot ovat täysin samanarvoisia eikä esitystapaan sisälly mitään
ensisijaisuus- tai yleisyysnäkökulmaa. Sanakirjan käyttäjä saattaa kui-
tenkin ymmärrettävästi tulkita asian niin, että päähakusana on aina en-
sisijainen tai ainakin muita yleisempi. Sanakirjantekijän ratkaisut eivät
välttämättä avaudu käyttäjälle siten kuin on tarkoitettu, eikä niitä aina
edes pystytä selittämään täysin aukottomasti. Sanakirjan käyttäjät eivät
yleensä lue kirjan johdanto-osia, joten siellä esitetyn tiedon ei voi myös-
kään ajatella tulleen selväksi kaikille käyttäjille. Kun englannin kielessä
on vaihtelua siinä, kirjoitetaanko ilmaukset kahdeksi sanaksi, yhdyssa-
naksi ilman viivaa vai yhdyssanaksi yhdysmerkkiä käyttäen, NODE on
valinnut näistä hakusanaksi aina vain yhden. Sanakirjan johdannossa (s.
xvii) todetaan, että näin on tehty tilansäästön vuoksi, eikä se tarkoita, että
muutkaan muodot olisivat väärin. On kuitenkin todennäköistä, että sa-
nakirjan käyttäjä tulkitsee tarkoituksen toisin ja olettaa ainoaksi oikeaksi
sen vaihtoehdon, joka on hakusanaksi valittu.

Deskriptiivisyyteen pyrkivä leksikografi saattaa tehdä myös tarkoituk-
sellisia rajauksia, jolloin esityksen ei voi sanoa olevan puhdasta kuvausta.
Englannin sanakirjoissa seksuaalisuuteen liittyvä sanasto oli paljolti tabu
aina 1960-luvun loppupuolelle asti. Esimerkiksi sana *fuck* esiintyi ensi
kerran pienessä englannin sanakirjassa 1965, ja vasta sen jälkeen se otet-
tiin mukaan yleisemminkin. Ei edes kuuluisimmassa deskriptiivisessä
sanakirjassa W3:ssa ollut tätä hakusanaa vielä 1960-luvun alussa. Tekijät
olivat kyllä ottaneet sanan mukaan ja varustaneet sen huomautuksella
”obscene and usu. unprintable”, mutta kustantaja oli vaatinut poistamaan
sen. (Landau 2001: 228.) NS:ssa vastaava keskeisin sanasto oli tuossa vai-
heessa jo esitetty, ja sanoja luonnehdittiin joko alatyylisiksi tai arkikieli-
siksi. Normatiivinen sanakirja voi siis olla joskus deskriptiivisempi kuin
deskriptiivinen. Nykyisin sukupuolisuuteen liittyvä sanasto ei ole enää
sanakirjoissa tabu, mutta rasismiin liittyvä sanasto on: sitä kuvastavia sa-
noja jätetään usein kokonaan pois sanakirjoista (Fjeld 1994: 36; Atkins-
Rundell 2008: 232). Rajaukset ja valinnat koskevat paitsi hakusanoja ja
niiden selittämistä myös sanojen käyttöä ja esimerkistöä. Jos korpuksessa
on loukkaavaa aineistoa – esimerkiksi kansallisuuteen, sukupuoleen tai
johonkin vähemmistöön liittyvää –, se jätetään ottamatta mukaan ja esi-

merkit valitaan toisin. Ratkaisu on luonnollinen ja eettinen, mutta kuvaus ei ole puhtaasti deskriptiivinen. (Landau 2001: 424.)

Selitteet ja esimerkit voivat heijastaa myös muunlaisia normeja kuin kieleen liittyviä, kuten esimerkiksi yhteiskunnassa yleisiä käsityksiä käytäytymisnormeista, asenteita ja ennakkoluuloja. Vaikka pyrittäisiin tietoisesti välttämään asenteellisuutta ja arvottamista, selitteisiin voi tulla tiedostamattomia painotuksia tai merkityksiä. Erityisesti tämä tarkoittaa sitä, että valtakulttuuriin ja yhteiskunnassa enemmistöä edustaviin näkökulmiin liittyvät piirteet vallitsevat ja korostuvat (ks. myös Tyysteri tk.). Esimerkiksi kristinusko saattaa olla eräänlainen oletusarvo uskontoon liittyvien sanojen selitteissä, ja kristinusko saattaa myös joskus painottua esimerkeissä suhteettomasti (Landau 2001: 420). Sukupuolirooleihin liittyvät asenteet voivat heijastua niin ikään varsinkin esimerkistössä. Ne heijastavat paljon myös kirjoitusajankohtaansa, ja ongelma asiasta tulee varsinkin silloin, kun sanakirjan uudempiin laitoksiin ei ole huomattu tehdä tarkistuksia tällaisiin kohtiin. Toisaalta on erotettava ”tarpeeton” ja ”tarpeellinen” asenteellisuus: tarpeetonta on sellaisten asenteellisuutta ilmaisevien esimerkkien käyttö, jotka eivät mitenkään liity itse hakusanaan, ja hakusanaa voitaisiin kuvata yhtä hyvin jollakin toisella esimerkillä. Sen sijaan sellaisia sanontoja kuin *hän kesti sen kuin mies, se on miehen puhetta* ja *teki miehen työn* ei voida jättää pois sanakirjasta, vaikka ne monien mielestä kuvaavat asenteellista suhtautumista mieheen ja naiseen. Ne ovat vakiintuneita sanontoja ja edelleen käytössä.

Vallitsevien asenteiden ja enemmistön näkökulman voima on usein vahva, eikä sanakirjantekijä huomaa aina kyseenalaistaa sitä. Käsityksillä ja näkemyksillä on voinut olla pitkät perinteet, ja niistä on tullut yleisiä normeja. Esimerkiksi heteroseksuaalisuus on pitkään ollut ikään kuin itsestään selvä lähtökohta sanojen selittämisessä sanakirjoissa, ja vasta viime vuosina on tultu enemmän tietoiseksi näkökulman suppeudesta (ks. tarkemmin Tyysteri tk.). Esimerkiksi KS:aan on tehty joitakin tähän liittyviä muutoksia, muun muassa sanan *flaksi* selite oli aiemmin ’menestys toisen sukupuolen keskuudessa, viuhka’ ja versiossa KS 2012 se on ’menestys, onni, tuuri, viuhka (vars. eroottista kiinnostusta tavoiteltaessa)’. Kaikkea ei ole huomattu vielääkään korjata, esimerkiksi ilmaus *Hän ei ole minun tyyppiäni* on selitetty: ’(vastakkaisesta sukupuolesta:) minun makuuni’.

Lopuksi

On usein todettu, että sanakirjojen esipuheissa korostetaan deskriptiivisyyttä ja epänormatiivisuutta enemmän kuin itse sanakirjassa on pystytty toteuttamaan. Voidaan jopa sanoa, että normatiivisuus kuuluu jossain määrin leksikografian luonteeseen ja siihen lähtökohtaan, että tiedot on koottu juuri sanakirjan muotoon (Fjeld 1994: 35). Sanakirja ei siis pysty olemaan puhtaasti deskriptiivinen. Sanakirjantekijä joutuu tekemään valintoja, jotka ovat tai joiden voidaan tulkita olevan kannanottoja. Valinnat voivat olla tietoisia tai tiedostamattomia. Ei ole myöskään mahdollista tehdä sanakirjaa, joka olisi täysin objektiivinen ja jossa ei olisi mitään ideologioihin pohjautuvaa normatiivisuutta. (Mts. 36.)

Kielenkäyttöön ja ilmaisuihin liittyvä normatiivisuus voi olla myös tarkoituksellista, mutta sen laatua ja määrää on pohdittava yhä uudelleen. Millaiset normit ovat tarpeellisia, ja mihin ne perustuvat? Voiko lukija ymmärtää niiden perusteet? Sanakirjan ratkaisuilla on merkitystä, koska niillä voi olla laajaa vaikutusta: sanakirja nähdään usein ehdottomana totuutena, johon saatetaan vedota, jopa riitatilanteissa ja oikeuskäsittelyissä. Sanakirjantekijän on otettava huomioon myös se, että lukija tulkitsee usein asian toisin kuin on ollut tarkoitus; lukija saattaa luoda normin itse.

Sanakirjan vaikutusta ei toisaalta pidä yliarvioida: sanakirja ei pysty koskaan vaikuttamaan koko kieliyhteisöön, ei yleiseen kielenkäyttöön eikä itse kieleen. Kun esimerkiksi PS:aa (1990–94) tehtiin, monet NS:ssa (1951–61) esitetyt kielenkäyttötavat, ohjeet ja varsinkin tyyllilajiluonnehdinnat todettiin vanhentuneiksi. Ne eivät kuvastaneet enää yleistä kielenkäyttöä, ja uuteen sanakirjaan ne muutettiin. Nämä muutokset olivat siis tapahtuneet kielenkäytössä sanakirjan ohjeista ja kuvauksista huolimatta. Vastaavaa on todettu myöhemminkin, ja muutoksia on tehty KS:n eri versioihin. Kieli elää ja kielenkäyttö muuttuu sanakirjoista riippumatta. Elämään jäävät ne ohjeet ja kielenkäyttötavat, jotka kieliyhteisö hyväksyy ja joita se haluaa noudattaa. Kuolleeksi kirjaimeksi muuttuneita normeja ei myöskään kannata pitää yllä. Normeja on tarkistettava ja arvioitava uudelleen, ja sanakirjoja uudistettaessa sisällön ajantasaistaminen on vähintään yhtä tärkeää kuin uuden lisääminen. Vaikka sanakirjan ohjeet eivät tavoita kaikkia eivätkä kaikki niitä noudata, sanakirjan tekijän on ajateltava niitä, jotka suosituksia lukevat, ja arvioitava, miten lukijat niitä

ehkä tulkitsevat. Suosituksiin saattavat vedota jotkin avainryhmät, kuten opettajat tai aktiivisesti suosituksista kiinnostuneet toimittajat, ja siksi niillä on merkitystä.

Se, että sähköisestä sanakirjasta ilmestyy uusia versioita suhteellisen tiheästi, on sekä hyvä että huono asia. Hyvää on se, että korjauksia on mahdollista tehdä, mutta huonoa se, että sanakirjan käyttäjä voi olettaa kaiken tiedon olevan tarkistettua ja vastaavan nykyhetken käsityksiä. Vaikka vanhoja artikkeleita tarkistetaan ja päivitetään mahdollisuuksien mukaan jatkuvasti, kattavaan ja järjestelmälliseen tarkistukseen ei ole versioiden välillä kuitenkaan mahdollisuuksia, koska uuden aineksen lisääminen vaatii yleensä suurimman osan tekijöiden ajasta. Painettujen sanakirjojen osalta voi taas olla joskus käyttäjän kannalta epäselvää, onko kyseessä uusi vai vanha kirja. Jos kirja ilmestyy sisällöltään samana mutta uudennimisenä tai jopa samannimisenä mutta eriasuisena, käyttäjä voi olettaa, että on kyse uusitusta teoksesta, ellei asiaa ole selvästi ilmaistu. Monet uskovat esimerkiksi nimen ja erinäköisten painosten perusteella, että NS:aa on vuosien mittaan uudistettu, vaikka sisältö on ollut kaikissa painoksissa muuttamaton.

Normitusta voidaan toteuttaa monella tavalla. Joissakin sanakirjoissa otetaan hakusanoiksi tai esiintymämuodoiksi vain ne, jotka edustavat vanhastaan noudatettua normia. Toisissa otetaan mukaan myös ”väärä” muoto ja esitetään suositus ”oikeasta”. Ensin mainittu tapa luo kuvan deskriptiivisestä kuvauksesta, mutta on yhtä normatiivinen kuin toinenkin. Jälkimmäinen tapa on tavallaan deskriptiivisempikin: se kertoo, että toisenlaistakin käyttöä on ja että normi ei ole kaikkien hyväksymä tai noudattama, mikä voi myös ennakoida normin muuttumista.

Sanakirjan ongelma on sen kaavamainen esitystapa, joka yksinkertaistaa ja pelkistää ja jota siksi voidaan helposti tulkita toisin kuin on tarkoitettu. On vaikea kuvata sanakirjan suppeassa esityksessä sitä, millaisessa yhteydessä ja miksi jokin ilmaus on parempi valinta kuin toinen. Sanakirjojen johdanto-osissa pitäisi kuvata tätä puolta tarkemmin, vaikka se ei ole helppoa ja vaikka sitä harvat sieltä lukisivatkaan. Pitäisi myös etsiä uusia esitystapoja, jotka mahdollistaisivat perustelevamman ja pohdiskelevamman kuvauksen. Sähköisissä sanakirjoissa artikkeleihin linkitetyt taustatekstit ja kommentoinnit olisivat tähän helposti tarjoutuva mahdollisuus. Painetuissa kirjoissa monien sanakirjojen (esim. NODE ja SO) jo nyt käyttämät erilliset tietoruudut antavat myös pohdiskelevamman

mahdollisuuden kommentointiin ja suositusten esiin tuontiin kuin niukat paremmin-ohjeet.

Olipa sanakirjan esitystapa mikä tahansa, aina jää tulkinnanvaraa. Aina jää selitteitä, esimerkkejä, suosituksia ja tyyllilajimerkintöjä, joista voidaan olla eri mieltä, ja aina jää myös tapauksia, jotka lukija tulkitsee toisin kuin sanakirjantekijä on tarkoittanut. Aina tulee myös olemaan sanakirjoja, jotka ovat joltain osin jo vanhentuneet, vaikka eivät olisi vanhojakaan. Se, mitä tarvittaisiin, olisikin eräänlainen sanakirjankäytön opetus. Suomessa ei ole sellaista oman kielen sanakirjan käytön perinnettä ja sanakirjakulttuuria kuin monissa muissa maissa. Sanakirjojen (sekä yksi- että kaksikielisten) käyttöä pitäisi opettaa koulussa. Perehdyttäisiin siihen, miten sanakirjaa pitäisi lukea ja tulkita, miten ja mihin sitä voi käyttää ja miten sen kaavamaista esitystapaa tulisi soveltaa. Opittaisiin myös, että sanakirja ei ole koskaan koko totuus. Samalla voitaisiin oppia paljon kielestä: onko sanakirjan esittämä suositus perusteltu ja paikallaan ja miten se toteutuu kielenkäytössä, noudatetaanko normia ja jos ei, niin miksi ei. Voitaisiin myös pohtia, millaisissa teksteissä ja yhteyksissä sopii käyttää arkikielisiä ilmauksia ja millaisissa ei ja miksi. Sanakirja voisi näin palvella paitsi kielitiedon jakajana myös kielitietoisuuden herättäjänä.

Sanakirja historiassa, historia sanakirjassa

Tämä artikkeli käsittelee historiallisia sanakirjoja. Artikkelin aluksi pohdimme historiallisen sanakirjan käsitettä ja esittelemme joitain historiallisia sanakirjoja. Sen jälkeen luomme yleiskatsauksen sanakirjojen historiaan Suomessa, mihin liittyy läheisesti myös suomen kielen aseman kehittyminen. Lisäksi pohdimme historiallisten sanakirjojen merkitystä sekä erityispiirteitä, jotka erottavat ne muista sanakirjoista. Artikkelin lopuksi lähestymme näitä erityispiirteitä sekä yleisellä tasolla että konkreettisesti omassa työssämme vastaan tulleiden esimerkkien kautta. Pohdimme historiallisen sanakirjan toimitustyötä myös etiikan näkökulmasta. Esimerkit ovat pääasiassa Vanhan kirjasuomen sanakirjasta (VKS, 1985–) ja Turun yliopistossa tekeillä olevasta Agricolan sanakirjasta [tulossa]. Artikkelin kannalta on keskeistä huomata, että molemmat sanakirjat käsittelevät vain kirjoitetun kielen aineistoja, jotka sijoittuvat vanhan kirjasuomen kaudelle. Kausi sijoitetaan yleensä aikavälille 1540–1810. Sitä seuraa varhaisnykysuomen kausi, joka ulottuu 1870-luvun tienoille, jolloin alkaa nykysuomen kausi.

Mikä on historiallinen sanakirja?

Hyvän määritelmän historiallisista sanakirjoista antavat Henning Bergenholtz ja Sven-Göran Malmgren. Heidän mukaansa historiallisia sanakirjoja ovat sanakirjat, jotka kokonaan tai osittain kuvaavat sanavarat ajanjaksolta, joka on selvän ajallisen välimatkan päässä siitä ajankohdasta, jolloin itse sanakirja tehdään. (Bergenholtz–Malmgren 2006: 2.) Mikä sitten on selvä ajallinen välimatka? Kun yleensä ajatellaan, että nykykieltä

kuvaavan sanakirjan tulee tavoittaa kolme elossa olevaa sukupolvea (Svensén 2004: 28), voidaan sanoa, että historiallinen sanakirja kuvaa silloin huomattavasti kaukaisempaa aikaa. Kuitenkin myös nykykieltä kuvaavat diakroniset sanakirjat voivat olla historiallisia sanakirjoja. Tällöin niiden kuvaaman kielen ajallisen alkupisteen on oltava selvästi kauempana kuin parin sukupolven päässä. Hyvä esimerkki tällaisesta sanakirjasta on Oxford English Dictionary (OED, 1884–), jossa on aineistoja jo vanhimmasta säilyneestä kirjallisesta englannista eli noin vuodelta 1100 jaa. (Hartmann 2001: 41–42). Joskus sanotaan historiallisiksi sanakirjoiksi myös vanhoja sanakirjoja, jotka kuvaavat oman aikansa kieltä. Historialliseen sanakirjaan kuitenkin kuuluu ajassa taaksepäin suuntautuva näkökulma.

Bo Svensén (2004: 28) on jaotellut sanakirjat neljään ryhmään sen mukaan, mikä on niiden suhde aikaan. Nämä ryhmät ovat synkroninen aikalais-sanakirja, diakroninen aikalais-sanakirja, synkroninen historiallinen sanakirja ja diakroninen historiallinen sanakirja. Tavallisin sanakirjatyyppi on synkroninen aikalais-sanakirja, joka kuvaa oman aikansa kieltä. Suomessa Kielitoimiston sanakirja (KS, 2004–2012) edustaa tätä tyyppiä. Diakroniset aikalais-sanakirjat puolestaan kuvaavat kielen kehitystä pidemmältä aikajaksolta mukaan lukien oman aikansa nykykielen. Tähän kategoriaan kuuluvia tunnettuja esimerkkejä ovat jo edellä mainittu OED, Grimmin veljesten Deutsches Wörterbuch (DWB, 1854–1961) ja Svenska Akademiens ordbok (SAOB, 1898–). Suomessa ei varsinaisia diakronisia aikalais-sanakirjoja ole tehty, vaikka esimerkiksi Chrisfried Ganander käytti sanakirjansa (käsikirjoitus valmistui 1780-luvun lopulla) aineistona myös vanhemmista suomenkielisistä kirjoista poimittua sanastoa. Tällaista kirjallisuutta oli kuitenkin vähän ja vain satunnaisesti saatavilla. (Hormia 1961: 93, 97–99.) On kuitenkin huomattava, että synkroniset aikalais-sanakirjat heijastavat aina jossain määrin myös historiaa. Ne huomioivat myös varhaisempaa sanastoa, vaikka eivät menekään tarkoituksellisesti yhtä kauas historiassa kuin diakroniset sanakirjat. Synkroniset ja diakroniset historialliset sanakirjat eivät puolestaan kuvaa nykykieltä lainkaan. Synkroniset historialliset sanakirjat kuvaavat tiettyä rajattua ajanjaksoa historiassa. Kotimaisiksi esimerkeiksi sopivat Agricolan sanakirja (tekeillä) sekä Ljungo Tuomaanpojan lainsuomennoksen sanasto (2007). Diakroniset historialliset sanakirjat puolestaan kuvaavat kielen kehitystä pidempänä historiallisena ajanjaksona. (Svensén 2004: 28.) Suomessa VKS kuuluu tähän ryhmään.

Hyvin monissa Euroopan maissa on tehty tai tehdään laajoja diakronisia aikalais- ja historiallisia sanakirjoja. Esimerkiksi DWB, hollantilainen *Woordenboek der Nederlandsche Taal* (1863–2001) sekä SAOB kuvaavat kaikki kohdekielensä kehitystä 1500-luvulta kirjoitusajankohtaan tai nykypäivään asti. OED:n aikajänne on vielä pidempi. Kaikkien neljän sanakirjasarjan tekeminen aloitettiin jo 1800-luvulla. Saksassa ja Hollannissa sanakirjojen toimittaminen kesti yli sata vuotta. SAOB on edelleen työn alla, ja OED:n ylläpito jatkuu myös nykyään. Yhteistä näille kaikille sanakirjasarjoille on siis varsin pitkä toimitusaika.

Suomessa ei ole sitten Henrik Gabriel Porthanin¹ suunniteltu sanakirjaa, joka yhdistäisi kattavasti etymologista ja diakronista tietoa sekä kuvaisi myös nykykieltä, kuten esimerkiksi OED. Sen sijaan päädyttiin tekemään erilliset sanakirjat (ks. esim. Läsimäki 2008: 565). Vuonna 1961 valmistui ensimmäisenä Nykysuomen sanakirja (NS, 1951–1961), joka kuvaa nimensä mukaisesti oman aikansa nykykieltä, vaikka kurkottaakin aineistossaan osittain 1800-luvulle. Suomen kielen etymologinen sanakirja (SKES) valmistui vuonna 1981. VKS ja kansankieltä kuvaava Suomen murteiden sanakirja (SMS, 1985–) ovat edelleen työn alla. (Hakulinen 1999 [1967]: 149–150.)

Suomen laajin historiallinen sanakirja on Kotimaisten kielten keskuksessa (Kotus) valmisteilla oleva VKS. Kovin montaa suoranaista historiallista sanakirjaa ei suomen kielestä ole ylipäänsä tehty. Etymologisia sanakirjoja on toimitettu, mutta ne kuvaavat kielen diakronista kehitystä hieman eri näkökulmasta ja eri painotuksin kuin yksikieliset historialliset sanakirjat, joihin tässä artikkelissa keskitytään. Jo aiemmin mainittujen VKS:n, *Agricolan sanakirjan* ja *Ljungo Tuomaanpojan lainsuomennoksen* sanaston lisäksi *Kalevalan kielestä* tehdyt sanakirjat voidaan laskea historiallista kielimuotoa kuvaaviksi sanakirjoiksi (ks. Turunen 1949; Jussila 2009). Ensi sijassa murre-sanakirjoiksi mielletyt SMS ja Suomen ruotsalaismurteita kuvaava *Ordbok över Finlands svenska folkmål* (FO, 1982–) ovat siinä mielessä historiallisia, että niissä käytetty ja kuvattu aineisto on enimmäkseen kirjoitusajankohtaa huomattavasti varhaisempaa (Vilppula 2006). FO käyttää myös pääasiallisena mallinaan SAOB:a

1 Porthan suunnitteli 1780-luvun alussa etymologista sanakirjaa, joka toisi esiin suomen ja sen sukukielten sanojen väliset suhteet ja erottaisi suomalaiset ainekset vierasperäisistä (Setälä 1892: 309).

(Kotus.fi: Suomen ruotsalaismurteiden sanakirja), mikä kytkee sen osaksi historiallisten sanakirjojen perinnettä. Suurena puutteena Suomen sanakirjaperinteessä voi pitää sitä, ettei varhaisnykysuomen kaudesta ole olemassa eikä myöskään tekeillä perusteellista sanakirjaa. 1800-luvulla suomen kieli kehittyi merkittävästi, kun sanastoa kehitettiin eri alojen tarpeisiin, minkä takia aikakauden kielen tallentaminen sanakirjaan olisi merkittävä kulttuuriteko. (Länsimäki 2008: 565.)

Suomen sanakirjahistoria ja suomen kielen asema historiassa

Kirjakielistyminen tapahtui Suomessa melko myöhään. Myös suomen sanakirjahistoria on lyhyt moniin muihin kielisiin verrattuna. Ensimmäinen tunnettu suomen kielen sanaluettelo on vuodelta 1544, jolloin ilmestyi Sebastian Münsterin *Cosmographia*-teos. Siinä on kahdentoista sanan esimerkkilista suomen kielestä. (Häkkinen 1994: 115.) Hyvin vaatimatoman alun jälkeen sanalistat alkoivat Suomessa laajentua ja muuttua varsinaisiksi sanakirjoiksi 1600-luvulla. Alkuun sanakirjat olivat asiakokonaisuuksien mukaan järjestettyjä tulkkisanakirjoja, joiden ensisijainen tarkoitus oli tukea latinan kielen opintoja (Hakulinen 1999 [1967]: 134).

1600-luvulta on säilynyt kaikkiaan kolme painettua suomea sisältänyttä tulkkisanakirjaa ja niiden uusintapainoksia. Näistä ensimmäinen on vuoden 1637 *Lexicon Latino-Scondicum*, jonka tekijä oli ruotsalainen maisteri Ericus Schroderus. Sanakirjassa on noin 2 400 latinankielistä sanaa, jotka on selitetty ruotsiksi, saksaksi ja suomeksi. Seitsemän vuoden päästä Schroderuksen sanakirjasta ilmestyi *Variarum rerum vocabula Latina, cum Svetica et Finnonica Interpretatione*, jossa nimensä mukaisesti on latinalaisia sanoja tulkittuna ruotsiksi ja suomeksi. Sen tekijä on pysynyt tuntemattomana. Kolmas suomea sisältänyt tulkkisanakirja on Henricus (Henrik) Florinuksen *Nomenclatura Rerum brevissima Latino-Sveco-Finnonica* vuodelta 1678. Sanakirjasta julkaistiin useampia painoksia, joiden sisältö karttui kerta kerran jälkeen, ja selityskieleksi tuli myös saksa. (Hakulinen 1999 [1967]: 134–136.) Nummilan (2010: 190–191, 199) mukaan 1600-luvun tulkkisanakirjat sisältävät pääasiassa yleisesti tunnettua perussanastoa ja niiden murreaines on selvästi läntinen. Neljäs 1600-luvulta säilynyt sanakirja on kielten sukulaisuussuhteiden

ta tutkineen Martin Fogelin käsikirjoitus *Nomenclator Latino Finnicus*, jonka Romppanen (2001: 187–188) mainitsee olevan ensimmäinen suomea sisältänyt kaksikielinen sanakirja. Tämä vuodelta 1669 oleva käsikirjoitus on kuitenkin Nummilan (2012: 166, 168) mukaan melko suora kopio *Variarum rerum vocabulan* vuoden 1668 painoksen suomen- ja latinankielisistä osista eikä siten tarjoa paljon uutta Suomen sanakirjahistoriaan, vaikka siitä onkin pudotettu ruotsi pois selityskielen asemasta.

Hakulinen (1999 [1967]: 136) on sitä mieltä, ettei 1600-luvun sanakirjoja voi vielä nimittää sanakirjoiksi vaan melko alkeellisiksi sanaluetteloksi. Romppanen (2001: 122) kritisoi tätä näkemystä yksinkertaistavaksi ja korostaa, että sanakirjojen lähtökohdat olivat tuolloin kokonaan erilaiset kuin nykyään: niiden tärkein tehtävä oli toimia apuneuvoina latinan opiskelussa. Teemoittain rakennetut sanaluettelot tukivat koulu-laisten keskustelutaidon opiskelua. Romppanen on tutkinut suomalaista sanakirjatradiotiota ja huomauttaa sen seuranneen pitkälle 1800-luvulle Ruotsin mallia, vaikkei Suomi enää vuoden 1809 jälkeen ollutkaan osa Ruotsia (mts. 95). Ruotsissa kaksikielinen leksikografia, jonka pääasiallinen tarkoitus ei ollut latinan opettaminen, alkoi vuoden 1700 tienoilla ja yksikielinen ruotsalainen leksikografia noin vuonna 1850 (Ralph 2012: 40). Suomessa puolestaan Daniel Jusleniuksen sanakirja vuodelta 1745 oli ensimmäinen, jossa suomi oli lähtökielenä. Selvä ero maiden välillä on, että Suomessa yksikielisten sanakirjojen traditio on nuorempi kuin Ruotsissa. Ensimmäinen kiistattomasti yksikielinen sanakirja oli vasta NS 1900-luvun puolivälissä. (Romppanen 2001: 95–96, 125, 133.)

1800-luvulla Suomessa keskityttiin käytännön syistä erityisesti ruot-sin ja suomen välisiin sanakirjoihin (Häkkinen 1994: 118). Vaikka Suomi tuolloin oli osa Venäjää, pysyivät suhteet Ruotsiin vahvoina niin kulttuurisesti, kielellisesti kuin taloudellisestikin. Myös Ruotsin laki pysyi voimassa Suomen autonomisessa suuriruhtinaskunnassa. 1800-luvulla julkaistiin monia kaksikielisiä sanakirjoja, joissa suomi oli lähtö- tai kohdekielenä. Ensimmäinen varsinainen suursanakirja² oli Elias Lönnrotin Suomalais-Ruotsalainen Sanakirja vuosilta 1867–1880. Siinä suomen kielen sanoja selitettiin ruotsiksi, joka tuolloin oli Suomen sivistyneistön kieli. (Romppanen 2001: 90, 147–150, 163.) 1900-luvulla suomalainen

2 Suursanakirjoiksi kutsutaan sanakirjoja, joissa on vähintään 100 000 hakusanaa eli lemmaa, kun Lönnrotilla niitä on noin 200 000 (Romppanen 2001: 163).

sanakirjakenttä monipuolistui. Yleiskieltä kuvaavien yksikielisten sanakirjojen lisäksi toimitettiin muun muassa ensimmäiset slangisanakirjat ja muita eri kielimuotoja ja aineistoja kuvaavia sanakirjoja (esim. Hämäläinen 1963; Jussila–Länsimäki 1994) sekä selkeästi tietylle kohderyhmälle tarkoitettuja sanakirjoja (esim. SKSU, 1999). Myös sanakirjojen kielivalikoima on 1900- ja 2000-luvulla kasvanut ensin muun muassa fennougriastien tutkimuskohteiden ja myöhemmin erityisesti kansainvälistymisen ja maahanmuuton myötä. Esimerkiksi ensimmäinen suomalais-englantilainen sanakirja Siirtolaisen tulkki julkaistiin Yhdysvalloissa 1888, Suomessa näiden kielten välisiä sanakirjoja alettiin toimittaa 1900-luvun alkuvuosina (Koukkunen 2012: 77, 87). 2000-luvun taitteen suuri globaali muutos olivat sähköiset sanakirjat. Ne ovat madaltaneet kynnystä koota myös vapaasti internetissä täydennettäviä sanakirjoja (esim. Wiktionary).

Suomen sanakirjojen historiaa voidaan tarkastella myös siitä näkökulmasta, miten suomen kieltä on arvostettu erityisesti suhteessa ruotsiin ja latinaan. 1500-luvulla reformaatio korosti kansankielen ja lukutaidon merkitystä (Laine–Laine 2010: 258). Toisaalta myös latinan asema pysyi vahvana, koska kirkko piti pitkään hallussaan suomalaista koulutusjärjestelmää. Kirkon keskeinen asema vaikutti siihen, että koko vanhan kirjasuomen kaudella suurin osa suomeksi painetusta kirjallisuudesta oli vielä sisällöltään uskonnollista käännöskirjallisuutta. Suomea kyllä käytettiin asiointikielenä, mutta sillä ei ollut virallista asemaa. 1600-luvun kuluessa suomen kielen asema jopa heikkeni, koska sivistyneistö käytti ruotsia puhekielenä. Turun akatemian perustamisen yhteydessä vuonna 1640 heräsi kiinnostus suomen kieleen, mutta toisaalta yliopistossa käytettiin latinaa ja myöhemmin ruotsia, joten suomalaiset opiskelijat vierottuivat suomen kielestä. (Häkkinen 1994: 12–13.) Lukutaito eteni Suomessa hitaasti, ja vasta 1700-luvulla voi opetuksen tason sanoa parantuneen laajemmin (Laine–Laine 2010: 262). Reformaation vaikutusten lisäksi suomen kielen kehitykseen vaikutti erityisen voimakkaasti 1800-luvun fennomania, joka pyrki parantamaan suomen kielen asemaa. Erityisesti vuosisadan lopulla suomen merkitys alkoikin korostua. (Romppanen 2001: 90.) Tämä muutos heijastui myös sanakirjoihin siten, että niiden katsottiin täyttävän kansallista tehtävää, joten niitä tuli tehdä entistä päämäärätietoisemmin ja monipuolisemmin.

Historiallisten sanakirjojen merkitys

Erilaisilla sanakirjoilla on erilainen suhde yhteiskuntaan. Bo Svensén (2004: 1–2) puhuu dokumentaatioleksikografiasta, joka täyttää esimerkiksi kansallisen, poliittisen tai tieteellisen yhteisön tarpeen säilyttää tietoa tulevaisuuteen. Historialliset sanakirjat täyttävät yhteiskunnassa dokumentoivaa tehtävää. Ne säilyttävät vanhoja sanoja, mutta voivat myös laajentaa ja rikastuttaa nykyistä sanastoa, mistä hyötyvät tutkijoiden lisäksi muutkin kielestä kiinnostuneet.

Kotimaisten kielten keskuksen toteutettavaksi annettujen, suomalaisen yhteiskunnan ja kulttuurin kannalta merkittävien suurten sanakirjahankkeiden ja koko tutkimuskeskuksen tilanne yleisemminkin on viime vuosina ollut keskustelun alla (esim. Vir. 4/2008: 562–587; Rintala 2010: 254–256). Rahoittajan näkökulmasta pitkäkestoiset sanakirjahankkeet voivat olla ongelmallisia. Esimerkiksi OED:n ja SAOB:n historiaa tiedetään, että sanakirjaprojektit helposti venyvät ja työ kestää usein kauemmin kuin on osattu ennustaa. Jopa OED:llä on nykyään painetta nopeuttaa toimitustyötä (Simpson 2011). On kielipoliittinen ja yhteiskunnallinen kysymys, mikä kielimuoto ansaitsee tulla kuvatuksi. Yhteiskunta, jossa kielen ja kielen systemaattisuuden merkitys yhteisön toimivuuden ja yhtenäisyyden kannalta koetaan tärkeäksi, tuottaa sanakirjoja. Jos kieli koetaan tärkeäksi, myös sen historiaa kohtaan tunnetaan mielenkiintoa ja arvostusta.

Lingvistikseen näkökulmasta historiallisten sanakirjojen merkitys on ilmeinen: nykykieli kumpuaa vanhasta kielimuodosta ja nojaa siihen (Vilppula 2006). Sanaston tutkijoille historialliset sanakirjat ovatkin tärkeitä lähdekirjoja. Ne ovat pysyviä muistomerkkejä oman aikansa kielestä ja kuvastavat aikansa kielenkäyttöä ja (ajatus)maailmaa. Tästä juontaa sanakirjojen merkitys kansatieteelle ja historian tutkimukselle. Sanasto ilmentää esimerkiksi yhteiskunnan teknistä ja kulttuurista kehitystä. Näin ollen sanakirjat tallentavat sekä historiaa että kansanperinnettä.

Jo mainittujen ryhmien ohella myös lakikielen tutkijat sekä teologit hyötyvät historiallisista sanakirjoista. Koska varhainen kirjoitettu kieli oli paljolti uskonnollista, teologit ovat kiinnittäneet huomiota historialliseen kieleen ja erityisesti Mikael Agricolan tekemiin Raamatun tekstien käännöksiin (esim. Heininen 1999; Holma 2008). Sanakirjantekijäkin hyötyy teologien kiinnostuksesta, koska nämä tuottavat tietoa, johon lingvistin

koulutus ei yllä. Sama koskee luonnollisesti myös muuta tutkimusta. Historialliset sanakirjat hyötyvät menneisyyteen suuntaavista tieteistä, olivat ne sitten humanistisia tai luonnontieteitä. Esimerkiksi biologit voivat tehdä yhteistyötä lingvistien kanssa ja paljastaa myös sanastontutkimuksen kannalta kiinnostavaa tietoa muun muassa kasvitieteestä (ks. esim. Häkkinen–Lempiäinen 2007, 2011).

Miten historiallisten ja nykykielen sanakirjojen toimittamisen lähtökohdat poikkeavat toisistaan?

Historiallisten sanakirjojen toimitustyössä on joitain eroavaisuuksia verrattuna nykykielen sanakirjojen toimittamiseen. Tässä luvussa pohditaan, mikä erottaa erilaisten sanakirjojen toimitusprosesseja. Kaikki historialliset aineistot ovat jollain tavalla rajallisia, kun taas nykykielen aineistot paisuvat alati. Suomen kielen historiallinen aineisto on moniin vanhempiin kirjakieliin verrattuna niukka niin määrältään kuin tekstilajeiltaan. Suomen mittakaavassa VKS:n aineisto on kuitenkin varsin laaja. Sisältäähän se noin puolimiljoonaisen sanalippuaineiston ja melko kattavan sähköisen tekstikorpuksen (Länsimäki 2008: 568).

Aikalaissanakirjat eivät käytännössä valmistu koskaan, sillä kieli muuttuu koko ajan. Tämän takia niitä päivitetään jatkuvasti. Historialliset sanakirjat sen sijaan on ainakin teoriassa mahdollista saada valmiiksi, ja yleensä ne tehdään vain kerran. (Vilppula 2006.) Historiallinen kieli ei myöskään enää muutu. Joiltain osin tämä tekee historiallisen kielen kuvaamisen helpommaksi, koska se ei pakene kuvaajaansa samalla tavalla kuin nykykieli. Sen sijaan toimittajan työtä vaikeuttaa se, että oma kielitaju ei enää vastaa menneisyyden käytäntöjä. Myös kuvattava maailma siirtyy koko ajan kauemmas, minkä takia on vaikeampaa ymmärtää sen ajan konkreettisia ja abstrakteja tarkoituksia. Ongelmana on myös jo aiemmin mainittu lähteiden niukkuus, joka ei mahdollista aineiston tarkempaa vertailua. Suomessa myös kirjallisen kulttuurin kaipa ja nuoruus vaikuttavat suoraan siihen, miten laajasti vanhaa aineistoa voi tutkia ja ymmärtää (vrt. esim. Lehtonen 1999: 14). Historiallisen sanakirjan tekijällä ei ole myöskään kieliyhteisöä, jolta tarkistaa kieleen liittyviä kysymyksiä. Tästä seuraa leksikografille huomattavia semanttisia ongelmia.

Aineiston käsittelyssä ja sitä kautta esimerkkien valinnassa historiallisen sanakirjan tekijällä on varsin vapaat kädet. Erityistä hienotunteisuutta ei tarvitse miettiä, koska aineiston kuvaamat tapahtumat ja henkilöt ovat jo vanhaa historiaa. Esimerkiksi SMS:n toimitustyössä joudutaan esimerkkien valinnassa huomioimaan se, että pienillä paikkakunnilla ihmiset voivat vielä tunnistaa, kenestä on kyse, vaikka SMS:n aineisto onkin pääosin 1900-luvun alkupuolelta ja siten suurin osa ihmisistä, joihin aineistossa viitataan, on jo kuollut. Nykykielen kuvaajille puolestaan kuuluisien henkilöiden tunnistettavuus voi olla ongelma. Sanakirjantekijän on pohdittava ainakin sitä, millaisessa sävyssä tunnistettavia henkilöitä voi esimerkkeihin nostaa. Historiallisten sanakirjojen kohdalla aikaa on kuitenkin ehtinyt kulua niin paljon, ettei mainittavia ongelmia synny. Esimerkiksi Mikael Agricola ei juuri mainitse itsenäisesti kirjoittamiinsa tekstinpätkissä nimeltä kuin kuninkaita tai muita merkkihenkilöitä omalle ajalleen ja asemalleen sopivaan kunnioittavaan sävyyn. Historiallisen aineiston käsittelyssä ei myöskään tarvitse huomioida tekijänoikeuksia, koska tekijöiden kuolemasta on kulunut yli lain määrittelemät 70 vuotta (Tekijänoikeuslaki 4. luku, 43. §). Tämä mahdollistaa sen, ettei aineistosta tarvitse jättää mitään piiloon.

Yksi keskeinen ja moneen muuhun kysymykseen kytkeytyvä ongelma historiallisen materiaalin kanssa työskentelevälle tutkijalle on, miten tehdä oikeutta kuvattavan ajan ihmisille. Historiantutkijat pitävät tätä laajana eettisenä kysymyksenä, ja myös leksikografien on se syytä huomioida. Tutkijan täytyy varoa, ettei ylenkatso omasta aikaperspektiivistään menneiden aikojen ihmisiä. Historian- ja kirjallisuudentutkijat käyttävät termiä *presentismi* kuvatessaan, miten menneisyyden ihmisille asetetaan vaatimuksia nykypäivän näkökulmasta ja sorrutaan monitahoisiin anakronismeihin (Fischer 1971: 135). Historiallisessa sanakirjassa anakronismi voi olla esimerkiksi jonkin vanhan käytännön kuvaaminen raaaksi tai jopa tietyn aikakauden kuvaaminen siitä lähtökohdasta, että se on ollut erityisen julma ja armoton. Keskiajastahan käytetään usein epiteettejä julma ja pimeä, vaikka historiantutkijat ovat tätä myyttiä pyrkineetkin purkamaan (esim. Klemetilä 2008). Myös nykykielen sanakirjan tekijän tulisi välttää arvottavia adjektiiveja. Esimerkiksi se, että jotain kansaa kuvataan alkukantaiseksi, korostaa melko karkeasti sanakirjan kirjoittajien näkemystä siitä, mikä on kullekin ajalle sopivan modernia (vrt. esim. 1).

- (1) **bušmanni** 'eräs Etelä-Afrikan tasangoilla asuva alkukantainen kääpiökansa' (NS)

Tutkijan pitäisikin aina huomioida tutkimuskohteensa mahdollisimman objektiivisesti. Ongelmana tutkimuskohteen huomioimisessa on se, että tutkija ei voi irtautua omasta yhteiskunnastaan ja ajastaan, jotka määrittävät pitkälti sen, miten historiaa tulkitaan. Myös lukijoilla on historiasta eräänlainen esiyymmärrys, jonka läpi he suodattavat uutta tietoa. Historiantutkijat korostavatkin, että yleisöään ei voi unohtaa, mikäli aikoo tehdä oikeutta tutkimuskohteelle. Tämän takia tutkijan on tiedostettava tapa, jolla ilmaisee tutkimustuloksensa. (Kalela 2000: 51–53, 56.) Myös leksikografit pitävät tärkeänä ohjenuoranaan, että sanakirjan lukija on tunnettava (Atkins–Rundell 2008: 5). Tämä tarkoittaa paitsi sanakirjan järkevää kohdentamista potentiaalisille lukijoille myös lukijan asemaan asettumista. Historiallisen sanakirjan tekijälle tämä merkitsee muun muassa sen miettimistä, mitä lukijan voidaan olettaa tietävän vanhasta kielestä ja sen kuvaamasta maailmasta sekä millaista kieltä tuon maailman kuvaamisessa tulisi käyttää.

Historiantutkijan vastuussa voikin nähdä kaksi eri puolta. Toisaalta on vastuu omia aikalaisia kohtaan ja toisaalta vastuu niitä menneisyyden ihmisiä kohtaan, joihin tutkimus kohdistuu ja jotka eivät enää itse saa ääntään kuuluviin. Näiden kahden puolen erottaminen toisistaan on kuitenkin vaikeaa ellei mahdotonta, koska ne ovat osittain päällekkäisiä. Yhtä kaikki tutkijan on oltava tietoinen välittäjän roolistaan kahden eri kulttuurin välillä. (Gurevich 1994: 73–74, 81–82.) Myös historiallisen sanakirjan tekijää koskee edellä kuvattu kaksinkertainen eettinen peruslähtökohta.

Historiallisten sanakirjojen toimittamisessa pohdittavia kysymyksiä

Tässä luvussa siirrytään historiallisten sanakirjojen toimitustyössä vastaan tuleviin käytännöllisiin ongelmiin ja syvennetään osin jo edellisessä luvussa mainittuja kysymyksiä. Osa ongelmista on selkeämmin eettisiä (myös laadun kautta) ja osa selkeämmin sanakirjan tekoon liittyviä yleisiä ongelmia. Jotkut ongelmista ovat yhteisiä myös muunlaisten sanakirjojen kanssa. Emme kuitenkaan erittele tarkemmin sitä, mitkä ongelmat koskevat yksinomaan historiallisia sanakirjoja.

Kieli kuvauksen kohteena – ja välineenä

Sanakirjan teossa kieli on aina sekä kuvauksen kohteena että välineenä. Historiallisten sanakirjojen kohteena on vanha kielimuoto, mutta kuvauksen välineenä nykykieli. Tämä tekee sanakirjatyöstä kerroksellista ja korostaa leksikografian kaksinkertaista vastuuta vanhan ja uuden kulttuurin keskisenä välittäjänä. Sanakirjan kerroksellisuus korostuu entisestään, kun siirrytään ajassa eteenpäin. Yleensä historianhistoriallisia sanakirjoja ei päivitetä samalla tavoin kuin yksinomaan nykykieleen keskittyviä sanakirjoja, joten tulevaisuudessa edes kuvauksen välineenä oleva selityskieli ei ole nykykieltä, vaan sanakirjan kirjoitusajankohdan kieltä, joka vanhenee koko ajan. Sanakirjan käyttö voi olla myöhemmin vaikeaa, jos teoksessa ei selitetä sellaisia merkityksiä, jotka ovat samanlaisia tekohten kielessä ja kuvattavassa kielimuodossa. Esimerkiksi VKS:ssa (1994) ei ole selitetty sanoille *keskiviikko*, *keskiviikkosaarna* ja *kirkkotie*. Toisaalta ajatonta selittämistapaa ei ole olemassakaan, joten leksikografien täytyy vain luottaa siihen, että lukijat tulevaisuudessakin osaavat käyttää sanakirjoja – myös tulkittaessaan muita sanakirjoja.

Myös pitkät toimitusaika lisää sanakirjan kerroksellisuutta. Esimerkiksi SAOB:aa on toimitettu yli sata vuotta, joten sen kirjoitusajankohdan ”nykypäivä” on varsin pitkä. Sen toimituskäytäntöihin ovat ehtineet antaa leimansa kielitieteessä yleisesti vaikuttavat teoriat, ja kasvava kerroksellisuus on näkyvässä erityisesti artikkelien rakenteessa (Nilsson 2012). SAOB:ssa näkyvät paikoitellen myös ruotsin kielen kirjoitusreformit, esimerkiksi yritys merkitä lyhyt *e*-äänne *ä*:llä tai *v*-äänteen kirjoitustavan muutos *hv*:stä *v*:ksi (ks. tark. Larsson 2009: 36–37). Myös kielen käytäntöjen ja maailmankuvan muutokset näkyvät pitkäkestoisissa sanakirjahankkeissa. Esimerkiksi VKS:ssa (1994) on artikkelissa *kansa* alihaikusana *musta kansa*, jonka selite on ’afrikkalaisista, neekereistä’. *Neekeri*-sanana käyttö on VKS:ssa yksittäistapaus, mutta nykyään sitä vältettäisiin systemaattisesti.

Kun pohditaan uusimman tutkimuksen ja kielitieteen uusien tuulien soveltamista sanakirjatyössä, vastakkain ovat lukijan oikeus saada ajantasaista tietoa ja sanakirjan toimittamisen realiteetit, eli työn hidastuminen. Myös sanakirjan yhdenmukaisuus saattaa kärsiä, jos kesken toimitusprosessin muutetaan käytäntöjä. Joissain asioissa muutokset tuntuvat ilmeisiltä ja väistämättömiltä, kuten rasistisiksi koetuista sanoista luopuminen merkityksenselityksissä, mutta joissain tapauksissa on mietittävä, mikä

on välttämätöntä. Onko esimerkiksi tarpeen huomioida uudet tutkimustulokset vaikkapa sanakirja-aineistona olevan tekstin syntyajankohdasta tai kirjoittajan henkilöydestä? Esimerkiksi VKS:ssa ns. herra Martin maanlainsuomennoksen esimerkkien lähdekoodeissa ei ole vuosilukua, mutta sitaatin sijainti sanakirja-artikkelissa tukeutuu tulkintaan myöhäisemmästä valmistumisajankohdasta. Tutkijat ovat esittäneet erilaisia näkemyksiä niin tekstin ikäämisestä kuin kirjoittajan henkilöydestäkin. Käsikirjoituksen syntyajankohdaksi on arveltu vuotta 1548 erään käsikirjoitusniteen kannessa olevan vuosiluvun mukaan, mutta eräiden muiden todisteiden perusteella käännöksen voi arvioida valmistuneen 1580-luvun alussa. Käännöksen tekijäksi on arveltu etenkin virsirunoilijana tunnettua, huomattavan kirjallisen uran tehnyttä Jacobus Finnoa, mutta todennäköisesti suomentaja on herra Martti eli Martinus Olai. (Blomstedt 1987: 7–12; Kuutti 2010: 111.) Tällaiset uudet tiedot saattaisivat vaikuttaa sanakirja-artikkelin sisäiseen kronologiseen järjestykseen, mikä aiheuttaa ristiriidan lukijan tarpeessa saada uusinta tietoa ja sanakirjan yhdenmukaisuudessa. Yksi vaihtoehto on päivittää aiemmat tiedot, mikä onnistuu sähköisissä sanakirjoissa. Sen sijaan vain painettuna ilmestyvään sanakirjaan tällaisia muutoksia ei voi tehdä kuin korkeintaan uusintapainoksessa, mikäli sellainen otetaan. Toinen vaihtoehto on tehdä päivityksiä sanakirjan myöhempimpiin osiin ja tiedottaa niistä. Ongelmana kaiken kaikkiaan on, ettei leksikografi voi uppoutua loputtomasti uuden tiedon metsästämiseen, koska muuten itse sanakirja-artikkelien työstämiseen ei jää aikaa.

Lukijan tarpeita tulee pohtia myös muista näkökulmista. Ongelmallista voi olla esimerkiksi se, ettei lukijalla ole kovin tarkkaa näkemystä historiallisesta maailmasta ja kielestä. Esimerkiksi tieto siitä, mitä ylipäätään on saanut julkaista vaikkapa 1600-luvun Ruotsissa, voi lukijalta puuttua, ja siksi hän voi nähdä kyseistä aikaa kuvaavan sanakirjan ylikorostavan uskonnollista sanastoa. Tulisi myös pohtia sitä, miten lukijan kielitaito tulisi huomioida. Esimerkiksi VKS pohjaa pitkälti siihen, että lukija tuntee melko hyvin historiallisen kielimuodon säännönmukaisuudet morfofonologiassa ja ortografiassa, mikä saattaa vaikeuttaa sanakirjan tulkintaa. Toisaalta historiallisia sanakirjoja lukevat todennäköisimmin tutkijat, joilta voi odottaa kielihistorian tuntemusta. Hankalalta vaikuttava sanakirja voi kuitenkin karkottaa potentiaalisia lukijoita, minkä takia jonkinlainen kompromissi on tässä(kin) asiassa tehtävä.

Toimittaja polttopisteessä

Leksikografien pätevyysvaatimukset vaihtelevat sanakirjakohtaisesti. Historiallisen sanakirjan tekijän pätevyysvaatimuksiin kuuluu kyky lukea vanhoja tekstejä sekä tieto vanhan kielen syntaktisista, morfologisista ja äänteellisistä erityispiirteistä. Lisäksi historiallisen aikakauden tuntemus on eduksi. Koska suuri osa teksteistä on Suomessa 1800-luvulle asti ollut käännoiksi erityisesti latinasta, ruotsista tai saksasta, on sanakirjantekijän hyvä osata näitä kieliä.

Varhaisiin kirjoittajiin on vaikuttanut suomen kielen vakiintumattomuus sekä suomenkielisten kirjojen ja tietysti myös sanakirjojen puute. Myöskään suomen kielen kieliopeja ei esimerkiksi Agricolalla ollut käytössä, mikä on luonnollisesti vaikeuttanut hänen työtään. Siitä säteilee ongelmia myös nykypäivän leksikografille, joka pyrkii määrittämään sanojen sanaluokkia ja merkityksiä. Leksikografien haasteena onkin tuntee varhainen kielimuoto kokonaisuutena. Sanasto on tietysti kaiken lähtökohta, mutta myös syntaktiset ja morfologiset rakenteet pitäisi hallita. Esimerkiksi tulisi havaita vanhassa kirjasuomessa yleiset passiivin infinitiivirakenteet, kuten *ruoskittaa*, jonka nykykielisenä vastineena voisi olla kontekstista riippuen *ruoskia*, *ruoskittavaksi* tai *ruoskittavan* (ks. tark. esim. Merimaa 2007: 110–111). Passiivimuoto pitää tunnistaa muun muassa siksi, että pystyy määrittämään verbin transitiivisuuden tai intransitiivisuuden, mikä puolestaan voi vaikuttaa sanakirja-artikkelin rakenteeseen.

Leksikografien toimenkuvaan voi sanoa kuuluvan ripauksen turhautumista ja riittämättömyyden tunnetta. Sanakirjantekijän on oikeastaan mahdollonta hallita kaikkea sitä tietoa, minkä avulla voisi saada aikaan täydellisen sanakirjan. Historiallisen sanakirjan tekijän yksi keino sietää turhautuneisuuden tuntemuksia on uskoa siihen, että ihminen on pohjimmiltaan melko lailla samanlainen, oli aikakausi mikä tahansa. Sanaston kohdalla tämä näkyy esimerkiksi tarpeena nimetä ihmisille tärkeitä asioita. Ajatus siitä, että ihmisen perusominaisuudet ovat pysyneet samanlaisina, kaventaa kuilua menneisyyden ja nykyisyyden välillä. Myöskään suomen kielen diakronisessa kehityksessä ei ole suuria hyppäyksiä, vaikka kirjakielen luominen onkin tietysti ollut merkittävä muutos (ks. esim. Häkkinen 2002: 20).

Kielitieteellisten ja historiallisten vaatimusten lisäksi leksikografien olisi hyvä tiedostaa työssään myös omat näkemyksensä ja asenteensa. Historiallisten sanakirjojen kanssa työskenteleville on keskeistä miettiä esimer-

kiksi sitä, miten oma suhtautuminen uskonnollisuuteen ja uskontoon näkyy esimerkkivalinnoissa. Valitseeko toimittaja esimerkkejä, jotka ovat uskonnon kannalta jotenkin yleviä, vai esimerkkejä, jotka näyttävät uskonnon ja menneen ajan maailman brutaalimman puolen tai kuvastavat kirjoittajan maallista näkökulmaa? Valintatilanteet näkyvät konkreettisesti Agricolan teoksista löytyvissä esimerkeissä 2 ja 3, joissa sanakirjan toimittaja joutuu tekemään valinnan maallisen ja uskonnollisen kontekstin sekä uskonnollisen toimituksen hyvässä tai huonossa valossa esittävän esimerkin välillä.

- (2) **paikka** 'maaston) kohta, alue, seutu, tienoo, paikkakunta'³
A-III-463-[Jes-5:2]⁴ Minun Ysteuellen on yxi Winatarha, ydhes Lihavas paicas
[Minun ystävällän on yksi viinatarha, yhdēs lihavas paikās]
A-III-679-[Mik-5:0] hen osotta Bethlehem, henen Syndymisens paican
[hän osottaa Betlehem, hänen syntymisens paikan]
- (3) **kastaa** 'kristillisen (tai juutalaisen) seurakunnan yhteyteen ottamisesta; kastotoimituksesta; myös *kuv.*'⁵
A-0-013-[Mk-16:16] Joca vscopi, ia castetan, hen tule autuaxi
[Joka uskopi ja kastetaan, hän tulee autuaaksi]
A-III-005 ombi sängen coctolinen ia oikein, ettei ioopunut Tomppeli olisi se Pappi, eike mw Tyly sallitaisi ia sodhaisi Castamahan, Jollei ylimeinen Häte ia pakko watisi
[ompi sängen kohtuulinen ja oikein, ettei juopunut tomppeli olisi se pappi, eikä muu tyly sallitaisi ja suodaisi kastamahan, jollei ylimäinen hätä ja pakko vaatisi]

Mikäli sanakirjan toimitusperiaatteisiin kuuluu laajan esimerkkiaineiston nostaminen artikkeleihin, ongelmaa ei välttämättä synny. Kuitenkin yleensä sanakirjoissa on vain rajallisesti tilaa, joten kaikkia mahdollisia esimerkkejä ei voi sanakirjaan ottaa. Näin ollen leksikografian

3 Merkityksenselitys Agricolan sanakirjasta. Esimerkit aineistopalvelu Kainosta Vanhan kirjasuomen korpuksesta (Kaino.fi).

4 Lyhenne A-III-463 merkitsee, että esimerkki on Agricolan teosten näköispainoksen (1987) kolmannesta osasta sivulta 463.

5 Merkityksenselitys sama VKS:ssa ja Agricolan sanakirjassa. Esimerkit aineistopalvelu Kainosta Vanhan kirjasuomen korpuksesta (Kaino.fi).

on tiedostettava, onko hänellä toistuva tapa valita esimerkiksi uskontoa ylentäviä tai alentavia esimerkkejä. Toisinaan sanakirjoissa voi näkyä myös häveliäisyys. Esimerkiksi VKS:sta puuttuu Agricolan sana *kuutautti* 'kuukautiset'. Syynä voi tietysti olla pelkkä vahinko mutta yhtä hyvin sensuuri. Selkeätä sensuuria on ollut ainakin Turusen (1949) Kalevalan sanakirjassa, jossa seksuaalisanastoa on jätetty kokonaan pois tai selitetty epämääräisesti (Lehtosalo–Vaittinen 2009: 208).

Omat arvot voivat vaikuttaa myös siihen, viitataan sanakirjan selitteissä esimerkiksi kristinuskon perusoppeihin tosiasioina vai näkemyksellisinä seikkoina. Esimerkiksi *pakana*-sanan käyttö selitteessä on herättänyt keskustelua: onko sen käyttö korrektaa vai ylenkatsetta muita kuin kristittyjä kohtaan? Toisaalta jos koko kuvattava kielimuoto liittyy kristilliseen, puhdasoppiseen luterilaiseen kulttuuriin, voi *pakana*-selite olla vain taloudellinen tapa ilmaista jonkin sanan merkitys. Ylipäänsä sen, että asioita pyritään selittämään näkemyksellisinä eikä tosiasioina, voi perustellusti nähdä kiistanalaisena, jos koko aikakausi tulkitaan kristinuskon perusoppeja totena pitäneeksi. Vaarana on syyllistyä liialliseen modernismiin tai oman näkemyksen korostamiseen.

Toimituksellisten ratkaisujen avoimuus ja niistä kertominen on asia, jota leksikografien olisi myös mietittävä. Sanakirjaa tehdessä on käytännössä mahdotonta pitää tarkkaa lukua siitä, miten ja millaisiin päätöksiin kunkin sanan kohdalla on päädytty lähtien hakusanan asusta, selitteen muodosta, esimerkistä jne. Koska yksittäisten sanojen kohdalla ei voi selittää kaikkia päätöksiä, on tärkeää, että toimitusperiaatteet ovat mahdollisimman avoimet. Toimitusperiaatteet helpottavat leksikografien työtä suuresti, samalla kun ne auttavat myös lukijaa oivaltamaan, millaisten prosessien kautta sanakirja-artikkelit ovat muotonsa saaneet. Yksi avoimuutta lisäävä tekijä on kertoa muiden sanakirjojen käytöstä toimitustyössä. Esimerkiksi Agricolan sanakirjan toimittamisessa käytetään hyvin paljon apuna VKS:aa, mikä myös kerrotaan jokaisen artikkelin kohdalla erikseen. Kaikkien sanakirjojen rakenne ei mahdollista lähteiden esiin tuomista yhtä tarkasti. Esimerkiksi VKS hyödyntää jonkin verran NS:aa, mutta sitä ei kovin eksplisiittisesti tuoda esiin. Ongelmaa ei kenties ole Kotuksessa tehtyjen sanakirjojen välillä (esimerkiksi PS:n ja KS:n suhde), mutta muuten sanakirjojen toimituksissa tulisi miettiä, miten lähteiden käyttö ilmoitetaan. Ylipäätään leksikografien olisi hyvä miettiä, pitäisikö muiden sanakirjojen käyttöä problematisoida. Perinteisesti sanakirjat

hyödyntävät toisia sanakirjoja melko suruttomasti. Ainakin johdannois-
sa toimintatapa tulisi mainita (ks. myös Jussila tk.).

Käännöskielen vaikutus

Suuri osa vanhan kirjasuomen aikaisista suomenkielisistä teksteistä on käännöksiä. Erityisesti uskonnollisen tekstin kääntäminen oli pitkään jopa kuolemanvakavaa toimintaa: Euroopassa on poltettu roviolla kääntäjiä, joiden on katsottu poikenneen liiaksi pyhinä pidetyistä alkuperäis-
teksteistä. Sanasta sanaan -kääntämisperiaatteen ohella myös kansankie-
len kehittymättömyys ja apuneuvojen puute näkyvät kääntäjien työssä, minkä takia varhaisissa teksteissä on paljon vieraskielisiä vaikutteita. (Saksa 2004: 54, 57, 61, 64.)

Vanhan käännöskielen rakenteelliset ja sanastolliset erikoisuudet heijastuvat myös historiallista kieltä käsitteleviin sanakirjoihin. Erityisesti tämä näkyy sanastossa, jossa on runsaasti vierasperäisiä lainasanoja. Lisäksi vieraat kielenpiirteet näkyvät esimerkeissä sekä hakusanojen ja alihakusanojen asussa. Tällaisten helpommin havaittavien rakenteellisten piirteiden ohella käännöskielisyys on huomioitava myös merkityksen-
selityksissä. Toisinaan pelkästä kontekstista on vaikea päätellä, mikä on tietyn ilmauksen merkitys. Erityisesti uskonnollisen kielen merkitysten määrittämisessä alkukieliset tekstit sekä muut käännökset samasta tekstistä ovatkin tärkeitä tiedonlähteitä. Esimerkiksi Agricolan *kätkeä*-verbin merkityksiä on vaikea määritellä yksinomaan kontekstin perusteella. Tällainen esiintymä on esimerkiksi 2. Pietarin kirjeessä merkityksessä 'pitää, säilyttää':

- (4) A-II-614-[2Pt-2:9] Herra taita ne Jumalalliset kiusauxista pelasta,
Mutta ne wäret Domiopeiuen asti ketke pijnatta⁶
[Herra taitaa ne jumalalliset kiusauksista pelastaa, mutta ne väärät
duomiopäivään asti kätkee piinattaa]

Kaikista esiintymistä ei myöskään ole tiedossa lähtökielisiä vastineita. Mihin leksikografin silloin tulisi nojata? Vaarana on, että merkityksen voi tulkita aivan väärin. Leksikografi voi myös langeta vaivattomalta tuntu-

6 Esimerkki aineistopalvelu Kainosta Vanhan kirjasuomen korpuksesta (Kaino.fi).

vaan ratkaisuun ja tulkita kaikkien outojen esiintymien olevan kuvallista tai abstraktia käyttöä, jolloin merkitystä ei ole välttämätöntä selittää tarkemmin. Toisaalta erityisesti uskonnollisissa tekstissä kuvallisen merkityksen tulkinta on usein syytäkin jättää lukijan vastuulle ja kunkin oman näkemyksen varaan.

Se, että leksikografi käyttää työssään paljon lähtökielisiä vastineita sekä muita käännöksiä apuna, herättää välillä kysymyksen, siirtyykö hän itse asiassa teologiien ja kääntäjien tontille. Raamatun eriaikaisissa käännöksissä on paljon eroja käännöstekniikan ja myös käännöksen taustalla olevan uskonnollisen näkemyksen suhteen. Nykyinen dynaaminen käännöstekniikka, jossa on keskeistä merkityksen ja tyylin siirtäminen, on varsin kaukana esimerkiksi Agricolan sanatarkasta kääntämistavasta eikä näin ollen aina tarjoa tukea jonkin yksittäisen sanan merkityksen ymmärtämiseen. Leksikografi ei myöskään voi ruveta halkomaan hiusta teologisissa kysymyksissä. Vaarana on kadottaa lingvistinen luotettavuus sekä käyttää liikaa aikaa asiaan, joka ei varsinaisesti leksikografille kuulu. Painopisteen tulisi kuitenkin olla kuvattavan kielen systeemissä. Myöskään lähtö- tai vertailukielisten tekstien tutkiminen ei ole oikopolku täydelliseen tekstin ymmärtämiseen. Esimerkiksi sanakirjan aineistoon kuuluvassa tekstissä voi olla käännösvirhe tai muusta syystä aiheutunutta ristiriitaa vertailutekstien suhteen. Kumpaa merkitystä leksikografilla silloin tulisi selittää? Kielentutkijalla ei voi myöskään olettaa olevan kattavaa teologista tuntemusta alkukielisten tekstien merkityksestä ja kyseessä olevan ajan historiallisista ja uskonnollisista painotuksista. Sama pätee tietysti myös muihin tekstilajeihin. Esimerkiksi lakitekstit ovat varhainen ilmentymä erikoiskielestä, johon ovat pitkän ajan kuluessa vaikuttaneet vanha suullinen perinne ja muut kielet (Nordman 2010: 73, 75). Lakikieli käännettiin suomeen yksinomaan ruotsista, ja sen kääntämistä ja ymmärrettävyyttä häihtasivat alkuvaiheessa ruotsin kieleen vakiintuneet lakikielen termit, joille ei suomessa ollut vastinetta, myöhemmin muun muassa latinan- ja ranskankieliset lainasanat (Kuutti 2010: 107, 126). Klassinen esimerkki on Martin maanlainsuomennoksen *paatin siirtäminen*, joka perustuu ruotsin kielen pesänkavallusta tarkoittavan *bodrägt*-sanan väärinymmärrykseen. Kääntäjä ei ole tunnistanut lakitermiä eikä yhdyssanan rajakohtaa oikein, joten *bo+drägt* hahmottui muodoksi *båt+drag*. (Blomstedt 1987: 13.) Myös ajalle tyypillinen vakiintumaton ortografia on voinut vaikuttaa väärinymmärrykseen.

Aineiston rajaaminen ja käsittely

Kaikissa sanakirjaprojekteissa pitää miettiä aineiston rajaamista. Hankalaksi rajaaminen muuttuu pohdittaessa kriteerejä, joilla materiaalia rajataan sopivan kokoiseksi. Agricolan sanakirjassa rajaaminen on poikkeuksellisen helppoa, koska tarkoitus on kuvata Agricolan kieltä mahdollisimman kattavasti. Sen takia aineistoon on otettu kaikki Agricolan painetut teokset ja lisäksi Agricolan käsin kirjoittama sana *katti*, jonka oletetaan olevan hänen ensimmäinen suomeksi kirjoittamansa sana (ks. esim. Heininen 1976: 52). Painetuista teoksista jätetään pois ainoastaan latinaksi kirjoitetut katkelmat. Toisinaan on tietysti hieman vaikea määrittellä, milloin jokin vieraskielinen sana on paremminkin sitaattilaina, mutta kriteerinä on pidetty sanojen esiintymistä muuten suomenkielissä ympäristössä.

VKS:n aineiston rajaaminen on ollut huomattavasti vaikeampaa, koska aineistoon on haluttu ottaa mukaan mahdollisimman paljon suomenkielisiä tekstejä 1500-luvulta 1800-luvun alkuun. Tekstilajien arvottamista on käytetty aineiston rajaamisperusteena. Esimerkiksi arkkiveisuja⁷ on aineistossa vain muutamia, koska niitä ei ole pidetty sivistyneeseen kirjalliseen kulttuuriin kuuluvina. Myös kirjallisuusinstituutiot ovat mieltäneet arkkiveisut – suosiostaan huolimatta – roskakirjallisuudeksi (Varpio 1999: 103). Myös SAOB:ssa on tehty jaottelua parempiin ja huonompiin kirjailijoihin, minkä on sanottu näkyvän esimerkiksi siinä, että runoutta on siteerattu liian paljon suhteessa proosaan tai että työväestölle tarkoitettut lehdet eivät juurikaan näy sitaateissa (Larsson 2009: 43). Tällainen aineiston valinta ja käsittely herättävät kysymyksen sanakirjan normatiivisuudesta. Yleensä historianhistorialliset sanakirjat mieltävät itsensä deskriptiivisiksi, jolloin aineiston käsittelyn tulisi olla mahdollisimman objektiivista. Asia ei kuitenkaan ole niin yksioikoinen. Esimerkiksi SAOB:n käsitys itsestään deskriptiivisenä sanakirjana on melko uusi; alkujaan sen oli tarkoituskin olla normatiivinen (Larsson 2012: 385–386). Voidaan kysyä, onko selkeästi deskriptiivisiä sanakirjoja olemassakaan. Voidaan

7 Arkkiveisut ovat runoja, jotka käsittelivät historiallisia tapahtumia (erityisesti sotia), luonnonkatastrofeja, ihmeitä ja rakkautta. Niitä painettiin runsaasti 1600–1800-luvuilla erityisesti suhteutettuna Suomen muuten niukkaan painotuotantoon. Varsinkin suuri yleisö luki niitä ahkerasti. (Varpio 1999: 92, 103.)

puhua myös sanakirjojen piilonormatiivisuudesta, erityisesti silloin kun lukijan ja tekijän välillä on tulkintaero sanakirjan luonteesta (Lehtosalo–Tyysteri 2011: 213–214; ks. myös Grönros tk.). Piilonormatiivisuus korostuu historiallisissa sanakirjoissa ehkä tavallista enemmän, koska lukija saattaa pitää sanakirjan tulkintoja lopullisina totuuksina, vaikka sanakirjoissa joudutaan tekemään valintoja, jolloin esimerkiksi sanamuodoilla voi olla useampia mahdollisia tulkintoja.

Yksi yleisesti tiedostettu aineiston arvottamisperuste historiallisissa (ja jossain määrin nykykieltäkin kuvaavissa) sanakirjoissa on sukupuoli. Monissa leksikografisissa tutkimuksissa on havaittu, että sanakirjojen arvolutautuneisuus näkyy naisten huomiotta jättämisenä. Naislähteiden aliedustus on todettu niin OED:ssa (Baigent–Brewer–Larminie 2005), SAOB:ssa (Mattisson 2006) kuin ranskalaisissakin historiallisissa sanakirjoissa (Baider 2007). Myöskään VKS:n aineistossa ei ole naislähteitä. Tämä tietenkin kuvastaa vanhan kirjasuomen aikaista yhteiskuntaa: ennen 1800-lukua Suomessa ei ylipäänsä ollut paljon naisia, jotka olisivat tuottaneet tekstejä. SAOB:sta löytyy esimerkkejä sukupuoleen kohdistuvista asenteista myös selitetasolla (Larsson 2009: 43–45).

Sanakirjojen sukupuolittuneet lähtökohdat ovat selkeästi yhteydessä aikakauteen, jolloin niitä on alettu tehdä. Toimittajakunta oli pitkään jos ei täysin niin erittäin miehinen. Sukupuolittuneisuusongelma kertaantuu helposti myös nykypäivään, koska sanakirjat käyttävät lähteinä vanhempia sanakirjoja. Tällöin käy helposti niin, että viittaukset alkuperäisiin lähteisiin katoavat, mikä jättää erityisesti naiskirjailijoita unohduksiin. Ongelmana on, että auktorisoidut lähteet auktorisoivat itse itsensä ja vahvistavat siten ennestäänkin omaa näkyvyyttään, samalla kun muut jäävät huomiotta. (Baider 2007.) Nykyään ainakin Suomessa sanakirjojen toimittajat ovat usein naisia, minkä voi odottaa näkyvän jo nyt tai tulevaisuudessa muun muassa sanakirjoihin valikoitavissa esimerkeissä.

Toisinaan leksikografian puolueettomuutta aineiston käsittelyssä voi haitata ylimielisyys lähteitä kohtaan. Tietyistä aineistolähteistä voi esimerkiksi syntyä mielikuva, jonka mukaan ne ovat aina samanlaisia, yksiuolotteisia ja siten sanakirjan kannalta vähemmän kiinnostavia. Tarkemman tutkinnan avulla saattaa kuitenkin huomata aineistossa moniuolotteisuutta, jota ei osannut odottaa. Esimerkki 4 on sitaatti Anders Björkqvistin Postillasta (1801), jossa Jumalan sanan tai sen julistajan vertauskuvana yleisesti käytettyä *lyhty*-sanaa käytetään mielenkiinto-

sesti varioivalla tavalla: *vimmaavaiset lyhdyt* viittaa uskovaisiin, jotka liialla uskonnon tuputtamisella aiheuttavat asialleen enemmän harmia kuin hyötyä.

- (5) UskHarj1801b-194 Nämät vimmaavaiset lyhdyt juoxevat Jumalan eteen ja hajottavat monta piskuista armon työn alkua itzensä ja muitten tykönä heidän sokian kiivaudensa ohesa⁸
[Nämät vimmaavaiset lyhdyt juoksevat Jumalan eteen ja hajottavat monta piskuista armon työn alkua itsensä ja muitten tykönä heidän sokian kiivaudensa ohessa]

Aineiston käsittelyyn vaikuttaa myös se, mikä on sanakirjan linja esimerkistön suhteen. Aiemmin oli jo puhetta siitä, miten leksikografian omat arvosidonnaisuudet voivat vaikuttaa esimerkkien valintaan. Merkitystä on myös sillä, otetaanko sanakirjaan esimerkkejä paljon vai vähän. Jos sitaatteja on runsaasti, voidaan esimerkistön valinnalla nostaa esiin muun muassa merkitysten tulkinnanvaraisuutta. Samoin esimerkistön avulla voi todistaa jonkin yllättävän merkityksen puolesta silloin kun nykykielen mukainen kielitaju ei tulkinnassa auta. Esimerkkeinä käytettyjä sitaatteja saatetaan muokata ja lyhentää paljonkin lähinnä havainnollisuuden ja tilanpuutteen takia. Esimerkiksi asetusteksteissä on todella pitkiä virkkeitä, joiden merkitystä saa lyhentämällä selkeytettyä huomattavasti. Lyhentämisen haittapuolena on, että vanhan kielen syntaktiset rakenteet ja runotekstin rytmi saattavat hävitä tai kärsiä lyhentämisestä. Lukijalta katoaa myös näköala siihen, millainen teksti ja tekstilaji todella on kyseessä. Toisaalta sanakirjan ensisijainen tarkoitus ei ole perehdyttää lukijaa erityisesti tekstilajiin vaan sanastoon.

Merkitystenselitysten tarkkuudesta

Merkitystenselityksissä on keskeistä, että nykylukija ymmärtää ne. Ymmärrettävyyden takaamiseksi tulisi käyttää yleiskielen sanastoa ja välttää sivistyssanoja. Esimerkiksi VKS:ssa (1994) käytetään *kunnia*-artikkelissa selitettä 'valtasuuruus', joka voi olla nykylukijalle jo melko vieras ilmaus. NS:sta sen vielä löytää merkityksessä 'Jumalan majesteettisesta suuruudesta, vallasta', mutta PS:ssa ja KS:ssa sitä ei enää ole. Nykylukijan kannal-

8 Esimerkki aineistopalvelu Kainosta Vanhan kirjasuomen korpuksesta (Kaino.fi).

ta sen käyttäminen selitteenä ei siis ole perusteltua, vaikka VKS:n kannalta se on ollut kätevä ja tilaa säästävä selite.

Ymmärrettävyyden lisäksi merkitystenselitysten pitäisi olla mahdollisimman kuvaavia. Sen takia esimerkiksi kehäselitteet eivät ole suositeltavia. Ljungo Tuomaanpojan lainsuomennoksen sanastossa on melko tyyppillistä, että hakusanan selitteenä on sama sana kuin itse hakusana, esimerkiksi *korva* 'korva' ja *rakastaa* 'rakastaa'. Ilmeisesti selitteillä pyritään ilmaisemaan, että merkitys on sama kuin nykykielessä, mutta ongelmana ovat monimerkityksiset sanat. Esimerkiksi *korva*- ja *rakastaa*-sanoilla on KS:ssa useampi kuin yksi merkitysryhmä, joten pelkän nykykielen perusteella ei voi päätellä, mitä merkitystä erityisesti tarkoitetaan. Toinen ongelma kehäselitteissä on, että merkitykset myös ajan kuluessa muuttuvat. Tulevia lukijoita ajatellen tarvitaan siis tarkempi selite. Aina Ljungo Tuomaanpojan lainsuomennoksen sanastossa ei ole selitettä ollenkaan, mikä myös on kyseenalaista lukijan näkökulmasta. Esimerkiksi *ampujaislapa*-artikkelissa ei ole selitettä. Näytteenä olevasta sitaatista selviää, että sanalla tarkoitetaan ammutun riistan lapaa, mutta yhdyssana ei ole niin ilmeinen, että lukija ymmärtäisi merkityksen itsestään selvästi. Sanastossa annetaan sitaattilainoille yleensä ruotsinkieliset vastineet, mikä tietysti tarkentaa merkitystä, mutta ei ole ruotsia taitamattomalle kovinkaan valaisevaa.

Historiallisessa sanakirjassa ongelmaksi voi muodostua myös liian nykyaikaisten termien käyttäminen, koska ne saattavat luoda väärää mielikuvaa menneestä ja olla liian täsmällisiä. Esimerkiksi kasvinnimien selitteissä on varottava liian täsmällisiä lajinmäärittäyksiä, jotka saattavat joskus olla harhaanjohtavampia kuin ylimalkaisemmat, esimerkiksi kasvin ulkonäköön ja käyttötarkoitukseen viittaavat, luonnehdinnat. Merkityksiä pitääkin selittää auki tavallista enemmän, mitä taas voi pitää luisumisena ensyklopedian suuntaan. Perinteisesti on ajateltu, että sanakirjat kuvaavat kieltä ja ensyklopediat kielenulkoista maailmaa. Monien leksikografien (toisin kuin joidenkin muiden kielentutkijoiden) mielestä rajanveto on kuitenkin vaikeaa ja ero parhaimmillaankin hiuksenhieno. (Svensén 2004: 354.) Erityisesti historiallisissa sanakirjoissa ensyklopediset selitykset ovat väistämättömiä, koska lukijoiden ei voi olettaa arkitiedon perusteella tietävän yhtä paljon historiallisesta maailmasta kuin nyky-yhteiskunnasta. Ylivoimaisesti eniten ensyklopediseksi luokiteltavaa tietoa annetaan substantiiveista, jotka viittaavat konkreet-

tisiin entiteetteihin. Toisinaan on kuitenkin lähes mahdotonta selvittää tarkasti kaikkien entiteettien olemusta. Myös työmäärä täsmällisen tiedon saavuttamiseksi on sanakirjan teon reunaehdot eli aikataulu- ja kustannuskysymykset huomioiden usein kohtuuton. Näin ollen sanakirjantekijä joutuu miettimään, tyytyäkö luomaan mielikuva esimerkiksi tarkemmin määrittelemättömästä häkkyrästä 1600-luvun ruotsalaisen sotilaan asussa vai etsiäkö yksityiskohtaisempaa tietoa. Joka tapauksessa on selvää, että lähteiden merkitys kasvaa sitä mukaa kuin omakohtainen tieto aikakaudesta ohenee. Samalla tietysti myös lähdekritiikin merkitys kasvaa.

Ensyklopedinen tieto on sanakirjoissa pääosin merkityksenselityksissä. Kuitenkin myös esimerkeissä on mahdollista tuoda esiin kielenulkoista tietoa. Esimerkiksi Agricolan sanakirjaan valitaan usein esimerkkejä, joissa Agricola itse (lähinnä reunahuomautuksissa) selittää jonkin sanan merkitystä. Alla olevassa esimerkissä 6 Agricola kuvailee käyttämänsä *turilas*-sanana merkitystä. Mihinkään tieteelliseen yksiselitteisyyteen kuvailun perusteella ei ole mahdollista päästä, vaan Agricola on käyttänyt sanaa jonkinlaisena pienten tuhoeläinten yleisnimityksenä (Häkkinen 2007: 15). Vaikkei reunahuomautuksen avulla pystykään tekemään lajinmäärittystä, on esimerkin etuna kuitenkin sen antama historiallinen näkökulma sekä näyte Agricolan omasta kynänjäljestä. Lukija kenties oivaltaa raamatunsäe-esimerkkiä paremmin, millainen kuva koulutetulla ihmisellä on 1500-luvulla Suomessa eläimistöä ollut.

- (6) A-III-325-reun2 Turilat) Ouat politain Mateleuaiset ia Maaliset piskuiset Eleimet, Ninquin, Hijret – – Politain Lendeueiset, quin ouat Parmat – – Politain Molemista Osalset, quin Hemhekit Luteet, Teit, Coit, ia mw dh wahingoliset Turilat.⁹
[Turilaat ovat puolittain mateleuaiset ja maaliset piskuiset eläimet, niinkuin, hiiret – – puolittain lentäväiset, kuin ovat parmat – – puolittain molemista osalset, kuin hämhäkit luteet, täit, koit, ja muut vahingoliset turilaat.]

9 Lyhenne A-III-325-reun2 merkitsee, että esimerkki on Mikael Agricolan teosten näköispainoksen (1987) kolmannen osan toinen reunahuomautus sivulta 325. Ajatusviivat ovat esimerkissä poiston merkinä.

Agricolan sanakirjassa on mahdollista välittää ensyklopedista tietoa myös vapaamuotoisessa etymologiaosuudessa. Esimerkiksi sanasta *aami* voi merkityksenselityksen ’vetomitta, 157 litraa’ lisäksi kertoa sen vaihtoehdoisesta merkityksestä ja tarkemmasta historiasta (esim. 7).

- (7) Aami on lainattu muinaisruotsin sanasta *aam*. Alkuaan se juontuu kreikan vesisankoa merkitsevästä sanasta *āmē, hāmē*. Sana tunnetaan laajalti suomen länsimurteissa, ja se on voinut tarkoittaa myös heinämittaa (60 leiviskää). Sana on alkanut jäädä pois aktiivisesta käytöstä 1800-luvun lopulla, kun on siirrytty mittayksiköiden kymmenjärjestelmään.

Kaikkea tietoa on turha ladata merkityksenselitykseen, mutta taustoitus valottaa sanaa tarkemmin niille, joita asia kiinnostaa. Perinteisessä sanakirjaformaattissa, jossa artikkelin rakenne on tiukemmin määrätty, tällaista mahdollisuutta ei ole. Sähköisten sanakirjojen kiistämättömänä etuna voidaankin pitää sitä, että niillä on käytössään rajoittamaton määrä tilaa. Niihin on esimerkiksi mahdollista linkittää lisätietoa, joka on sanakirjan kannalta vähemmän relevanttia, mutta asian kannalta kiinnostavaa.

Lopuksi

Olemme pyrkineet tässä artikkelissa valottamaan historiallisten sanakirjojen tekoon liittyviä kysymyksiä sekä selventämään ylipäänsä leksikografin työnkuvaa, joka ei ehkä ole kovin tuttu muille kuin sanakirjojen kanssa työskenteleville. Työn problematiikan kuvaaminen auttaa uskoaksemme myös sanakirjojen käyttäjiä ymmärtämään paremmin sanakirjoja ja niiden erilaisia taustoja. Historiallisen sanakirjan toimittajalle on keskeistä pohtia, miten hänen omat arvonsa ja asenteensa vaikuttavat toimitustyöhön. Erityisesti suhtautuminen uskontoon on kysymys, joka voi heijastua valmiiseen sanakirjaan esimerkkien valinnassa ja merkityksenselityksissä. Omat arvot saattavat näkyä jopa siinä, mitä hakusanoja sanakirjaan valitaan. Historiallisen sanakirjan toimittajan olisi myös tiedostettava oma asemansa menneiden ja nykyisten sukupolvien keskenä välittäjänä. Tähän tehtävään liittyy vastuu siitä, miten objektiivisen kuvan leksikografi menneistä sukupolvista nykyiselle välittää. Olisi ideaalista, mikäli sanakirjatyo olisi mahdollisimman avointa kaikilta osin. Avoi-

muus työprosessien kuvaamisessa on kuitenkin haaste, koska sanakirja tekstimuotona pyrkii pelkistämään, jolloin suuri osa selvitystyöstä, ajattelutyöstä ja yksittäistapausten ratkaisuista jää dokumentoimatta. Joskus on sanottukin, että sanakirjanteon ja tutkimuksen välinen ero on siinä, että tutkimuksessa tuodaan esille se työ, joka sanakirjatyössä myös tehdään, mutta ei kerrota.

KIMMO GRANQVIST

Romanikielen sanakirjatyö ja sen eettiset kysymykset

Romanikieli kuuluu indoeurooppalaisen kielikunnan indoiranilaisen haaran indoarjalaiseen alaryhmään. Se on yksi niistä Intian kielistä, joita Intiasta lähteneet kiertelevät ryhmät puhuvat Intian ulkopuolella. Romanikieltä on keskiajalta lähtien ainoana indoarjalaisista kielistä puhuttu yksinomaan Euroopassa. Nykyisin Euroopasta lähteneet romaniryhmät puhuvat romanikieltä myös Yhdysvalloissa, Latalaisessa Amerikassa, Afrikassa ja Aasiassa. Romanikielen kantamuodon protoromanin oletetaan muotoutuneen Keski-Intiassa eikä aiempien käsitysten mukaisesti Luoteis-Intiassa. Protoromanin muotoutuminen ajoitetaan siirtymäkaudelle muinaisindoarjalaisista kielistä (sanskrit) keski-indoarjaan (prakrit, pali) eli vuoden 500 eKr. jälkeisenä aikana. Kontakti Bysantin ajan kreikan kanssa noin 900–1200 jKr. vaikutti suuresti romanikieleen. Kreikan vaikutus romanikieleen oli sekä sanastollista, äänteellistä että morfosyntaktista. Romanikielen nykyiset päämurteet kehittyivät romanien lähdettyä Etelä-Balkanilta. Murteiden keskinäiset erot kasvoivat edelleen eri romaniryhmien irtaannuttua muuttajien pääjoukosta ja asetuttua asuinalueelleen 1200–1300-luvuilla. Suomen romanikieli kuuluu niin sanottuun pohjoisten murteiden ryhmään, jonka keskuksena ovat Euroopan saksankieliset alueet. Pohjoismurteet jakautuvat edelleen kahtia luoteisiin ja koillisiin murteisiin. Ainoat yhä elävät luoteismurteet ovat sinti, jota puhutaan mm. Saksassa, ja Suomen romanikieli.

Romaneja on asunut Suomessa 1500-luvulta. Nykyään Suomen romaneja arvioidaan olevan Suomessa noin 10 000 ja Ruotsissa 3 000–4 000. Romaneja asuu aivan pohjoisinta Lappia lukuun ottamatta kaikkialla Suomessa, mutta valtaosa heistä asuu nykyisin Etelä-Suomen kaupungeissa. Suomen romanikieli jakautuu kahteen päämurteeseen, itä- ja länsimur-

teeseen. Itämurteen puhujat – suurin osa Suomen romaneista – ovat levittäytyneet jokseenkin kaikkialle maahan; länsimurretta puhutaan lähinnä Etelä-Pohjanmaalla. Alamurteiden välillä on suomen murteita vastaavia fonologisia eroja muun muassa pitkien vokaalien diftongiutumisessa ja švaa-vokaalin esiintymisessä sekä poikkeamia soinnillisen affrikaatan toteutumisessa ja esiintymisessä. Varsinkin soinnillisen affrikaatan toteutumien erot (esim. *džuuli* 'nainen' > Pohjanmaalla *tšuuli*, *tsuuli* tai *suuli*, Itä-Suomessa *djuuli* tai *juuli*) ovat herättäneet romanien keskuudessa affektioita. Suomen romanikieli on nykyisin vakavasti uhanalainen. Hedmanin (2009) selvityksen mukaan sitä hallitsee vähintään tyydyttävästi noin 60 % Suomen romaneista mutta hyvin enää kolmannes.

Tässä artikkelissa tarkastelen Suomen romanikielen sanakirjatyötä ja keskeisimpiä siihen liittyviä eettisiä kysymyksiä. Tarkasteltavat kysymykset liittyvät yhtäältä romanikielen omistajuuteen, yhteisöllisyyteen ja asemaan salakielenä sekä toisaalta sanakirjojen asemaan romanikielen normittajina ja siihen liittyvään vastuuseen ja romanien kielellisiin tabuihin. Osa eettisistä kysymyksistä koskee kaikkea romanikielen tutkimusta ja romanikielistä julkaisemista, osa on spesifimmin sanakirjatyön kysymyksiä. Osa kysymyksistä on erityisen keskeisiä tilanteessa, jossa romanikieltä tutkivat ja kuvaavat yhtäältä romanit, joilta puuttuu teoreettista kielitietoa ja lingvististä koulutusta, ja toisaalta pääväestöön kuuluvat ammattitutkijat, joilla on kielitieteellistä osaamista ja jopa useiden Euroopan romanimurteiden taitoa mutta joiden henkilökohtainen kosketus romanikieleen ja romaniyhteisöön on ohuempi. Muita tutkimuseettisiä kysymyksiä liittyy ennen kaikkea lähteiden valintaan, arvottamiseen ja keskinäiseen asemaan. Julkaistuissa Suomen romanikielen sanakirjoissa ja sanastoissa ei ole havaittavissa suoranaista toisten töiden plagiointia. Oskari Jalkion julkaisemattoman käsikirjoituksen lopun taivutuskaavat on mitä ilmeisimmin kopioitu Arthur Thesleffin (1901) sanakirjan lopusta, mutta 1990- ja 2000-lukujen sanakirjat perustuvat aiempien sanastolähteiden lisäksi vahvasti laatijansa omaan romanikielen taitoon.

Suomen romanikielen leksikografian historiaa

Romanikielen ja sen sanaston tutkimuksella on Suomessa pitkät perinteet. Varhaisin tunnettu lähde on Kristfrid Gananderin kilpakirjoitus *Undersökning om De så kallade TATTARE eller Zigeuner, Cingari, Bohemiens, Deras härkomst, Lefnadssätt, språk m.m. Samt om, när och hwarest några satt sig ner i Sverige?* (1780) Ruotsin kuninkaallisen kaunokirjallisuusakatemian arkistossa. Tutkielmallaan Ganander osallistui Ruotsin Kuninkaallisen Kaunokirjallisuusakatemian kirjoituskilpailuun, jonka aiheeksi määrättiin vuoden 1779 joulukuussa romanit sekä heidän alkuperänsä ja kielensä. Gananderin tutkielma jakautuu 34 pykälään, joista valtaosa käsittelee romanien historiaa ja kansatiedettä. Pykälät 22–32 sisältävät ensimmäiset muistiinpanot Suomen romanikielestä vain noin 200 vuotta sen jälkeen, kun ensimmäisten romanien tiedetään saapuneen Suomeen. Kilpakirjoituksessa on noin 150 romanikielen sanaa, jotka esiintyvät myös esimerkkilauseiden sisällä. Kilpakirjoitusta ei ole vielä kokonaisuudessaan julkaistu. Jussila (1997: 77) mainitsee, että perimätiedon mukaan Ganander olisi laatinut myös romanikielen sanakirjan käsikirjoituksen, joka on hävinnyt.

1800-luvulla romanikielestä kiinnostui useita tutkijoita ja harrastelijoita. Adolf Ivar Arwidsson (1791–1858) oli poliitikko, kirjailija, runoilija, historiantutkija ja sanomalehtimies. Vuonna 1817 hän kirjoitti Padasjoella muistiin romanikielen sanoja, jotka Sophus Bugge (1858) julkaisi osana artikkeliaan *Vermischtes aus der Sprache der Zigeuner*. Artikkelisi sisältyi A. Kuhnin ja A. Schleicherin toimittamaan teokseen *Beiträge zur vergleichenden Sprachforschung*. Arwidssonin muistiinpanot ovat suppeammat kuin Gananderin kilpakirjoituksen kielelliset havainnot, mutta ne sisältävät melko runsaasti taivutusesimerkkejä.

Claes Johan (Klaus Juhana) Kemell (1805–1832) oli Alavieskan vt. kappalainen, kääntäjä ja runoilija. Hän teki niin ikään muistiinpanoja romanikielestä ja laati sanakirjan käsikirjoituksen. Kemellin kokoama romanikielen sanaluettelo poltettiin hänen kuoltuaan jumalattomana työnä, mutta Arthur Thesleff rakensi myöhemmin oman sanakirjansa (1901) säilyneen konseptin pohjalta. Pertti Valtosen (1968: 33) mukaan tämä konsepti joutui huutokaupassa professori Otto Donnerille, mutta on sittemmin kadonnut. Jürgensen ja Schmidt välittivät 1860-luvulla tietoja Franz Miklosichille, joka julkaisi ne; samat glossat sisältyvät myös

Thesleffin (1901) sanakirjaan. Lydia von Essen ja Lydia Bergroth laativat 1860-luvulla suppean Suomen romanikielen sanaluettelon, jonka käsi-kirjoitus on Suomalaisen Kirjallisuuden Seuran hallussa.

Henrik August Reinholm (1819–1883) oli ensimmäinen varsinainen kansantietouden suurkerääjä. Kansanrunouden ohella hän tallensi valtavasti tietoja hyvin erilaisilta alueilta. Erityisesti on mainittava hänen perustavanlaatuisen toimintansa kansanleikkien ja -tanssien tallentajana. Reinholm valmistui maisteriksi 1844 ja sai filosofian tohtorin arvon 1857. Hänet nimitettiin 1856 Viaporin vankilasaarnaajaksi ja sotilaspapiksi, ja tässä virassa hän toimi kuolemaansa asti. Hän teki Viaporin vankilasaarnaajana toimiessaan ja Turun kehruhuoneessa laajat muistiinpanot romaneista ja heidän kielestään. Reinholmin muistiinpanot ovat tällä hetkellä Kansallismuseossa kahdessa kansiossa (numero 87), joiden otsikkona on ”Finlands zigenare”. Kansiossa 1 on vanhoja julkaisuja, lehti-leikkeitä ja yksi molemmin puolin käsin kirjoitettu arkki. Kansiossa 2 on 892 numeroitua käsin kirjoitettua sivua. (Sirku Dölle, henkilökohtainen tiedonanto 19.1.2004.) Reinholmin laaja-alainen ja kirjava kielenaines käsitti noin 2 000 sanan glossaarion, jotka Thesleff (1901) osittain (n. 1 000 sanaa) käytti sanakirjassaan, ja lisäksi kieliopillisia havaintoja mm. nominien ja verbien taivutuksesta. Osan Reinholmin kulttuurihistoriallisista muistiinpanoista julkaisi Johan Reinhold Aspelin Uudessa Suomettaressa (1894–1895).

Arthur Thesleff (1871–1920) oli kasvitieteilijä ja romanitutkija. Thesleff tutki sieniä ja puita, kirjoitti Tukholman rikollisten slangisanakirjan sekä perusti siirtokunnan Argentiinaan. Kaikkein pysyvimmin Thesleff muistetaan romanitutkijana: hän kirjoitti useita artikkeleita romanikulttuurista ja hänen tutkimusretkensä ulottuivat kaikkialle maailmaan. Thesleff toimi myös muutaman vuoden kansainvälisen romanitutkimusyhdistyksen Gypsy Lore Societyn presidenttinä. Hänen sanakirjansa *Wörterbuch des Dialekts der finnländische Zigeuner* (1901) oli ensimmäinen painettu Suomen romanikielen sanakirja. Lähinnä Länsi-Suomen romanien kieltä kuvaavassa sanakirjassa on yhteensä 7 574 hakusanaa. Romanikieltä varsinkin 1960-luvulla tutkinut Pertti Valtonen (1968: 48) laskee varsinaisia kantasanoja olevan Thesleffillä noin 2 100. Thesleff (1901) itse mainitsi Suomen romanikielen kantasanojen määräksi vain noin 2 200, joten hänen kantasanoistaan voidaan kuitenkin pitää varsin kattavana. Sanaston lisäksi kirjassa on myös sanojen taivutusmuodot sisältävä loppuluku.

1900-luvulta on säilynyt käsikirjoituksina kolme romanikielen sanasto, Oskar Jalkion (Johanssonin) (1882–1952), Axel Kronqvistin (1871–1956) sekä Yrjö Temon (k. 1985) aineistot. Suomen Mustalaislähetys ry:n (nykyään Romano Missio ry:n) perustajan Oskari Jalkion (Johanssonin) (1882–1952) sanottiin olleen ainoa ”valkolainen”, joka hallitsi sujuvasti romanikieltä (Valtonen 1968: 51). Jalkio kirjoitti jo 1900-luvun alussa lehtiartikkeleita romanikielestä ja romanikielellä, käänsi romaniksi lauluja ja virsiä sekä laati kirjasia romaneista. Jalkion keräämä sanastomateriaali vastasi laajuudeltaan suunnilleen Thesleffin (1901) sanakirjaa. Sanaston julkaisemattoman käsikirjoitteen lopussa on Thesleffin sanakirjan tavoin taivutuskaavoja. Axel Kronqvist (1871–1956) laati 1950-luvun alussa länsimurteen pohjalle romanikielen sanakirjan, jota on ollut olemassa kaksi kappaletta: toinen yksityishenkilön hallussa Helsingissä ja toinen Romano Mission hallussa. (Valtonen 1968: 54.) Yrjö Temon (k. 1985) noin 5 000 hakusanan laajuinen suomi–romani-sanaluettelon käsikirjoitus luovutettiin vuonna 1984 Kotimaisten kielten tutkimuskeskukselle. Kotimaisten kielten tutkimuskeskuksen oli tarkoitus laatia käsikirjoituksen pohjalta kaksisuuntainen romani–suomi–romani-sanasto, mutta työ jäi kesken.

Vuonna 1971 julkaistiin Opetusministeriön asettaman romanikielen ortografiatyöryhmän mietintönä Mustalaiskielen normatiivi sanasto (MNS, 1971), jossa esitetään lähes 3 000 romanikielen sanaa työryhmän ehdottaman oikeinkirjoitusmallin mukaisesti. Pertti Valtonen julkaisi vuonna 1972 Suomen romanikielen etymologisen sanakirjan. 1980-luvulla Opetusministeriö asetti romanityöryhmän, joka ehdotti mietinnössään kieliopin, sanakirjan ja lukukirjan laatimista (Jussila 1997: 78). Viljo Koivisto ryhtyi 1980-luvun puolivälissä Kotimaisten kielten tutkimuskeskuksen tuella laatimaan romani–suomi–englanti-sanakirjaa. Hanketta rahoitti opetusministeriö. Vuonna 1994 (2. painos 2005) ilmestynyt romani–suomi–englanti-sanakirja sisältää n. 5 500 hakusanaa. Sana-artikkelit sisältävät tietoja sanojen taivutuksesta, substantiivien suvusta sekä runsaasti esimerkkilauseita, jotka on käännetty suomeksi. Vuonna 2001 (2. painos 2011) ilmestynyt suomi–romani-sanakirja on laajin tähän saakka ilmestynyt Suomen romanikielen sanakirja. Sanakirjassa on noin 23 000 hakusanaa. Mukana on sekä arkikielen sanoja että eri alojen erikoistermejä mm. opettajien ja kääntäjien avuksi. Koiviston romani–suomi-sanakirja ilmestyi vuonna 2011 Opetushallituksen kustantamana.

Teos perustui melko suoraan tekijän kahden aiemmin laatiman sanakirjan aineksiin tietokoneen avulla uudelleen järjestettyinä.

Vuonna 1997 valmistui Kimmo Granqvistin Kotimaisten kielten tutkimuskeskuksen romanikielen korpuksista koostama suomi–romani-sanaluettelo, ja vuonna 1999 julkaistiin Kimmo Granqvistin koostama ja toimittama käänteissanasto. Henry Hedmanilla on tekeillä romanikielen uudissanasto, johon tulee noin 3 000 eri alojen hakusanaa.

Valtaosa Suomen romanikieltä ja sen sanastoa tutkineista tai kuvaneista on ei-romaneja. Romanitaustaisia ovat vain Yrjö Temo, Viljo Koivisto ja Henry Hedman. Yrjö Temo (k. 1985) oli maallikko, joka sanastotyön lisäksi käänsi romanikielelle raamatuntekstejä. Viljo Koivisto (s. 1931) on toiminut pitkään hengellisessä työssä, päätoimittajana, romanikielisten kirjoitusten laatijana ja romanikulttuuriasioiden hoitajana, hengellisten levyjen tuottajana, laulajana, opettajana ja romanikielen oppi- ja sanakirjojen laatijana. Henry Hedman (s. 1952) on saarnaaja, romaniaktivisti, romanikielen opettaja, oppimateriaalien ja sanakirjojen laatija sekä tutkija. Hän on julkaissut muun muassa romanikielen opettajien kielioppioppaan sekä selvityksen romanikielen säilymisestä ja käytäloista. Hän on kääntänyt romanikielelle useita teoksia.

Kielen omistajuus

Kysymys kielen omistajuudesta koskettaa niin sanakirjatyötä kuin kaikkea muutakin romanikielisen ja romanikieltä tarkastelevan tieteellisen ja populaarin materiaalin tuottamista. Toimijoissa on ollut kokonaan romanitaustaisia, puoliromaneja ja ei-romaneja. Romanikieltä pidetään tärkeänä kulttuuri-identiteetin ja ryhmän kiinteyden vahvistajana sekä ”aidon mustalaisuuden” symbolina (Åkerlund 2002: 126; Hedman 2004: 43). Se on liitetty romanien historiaan, traditioihin ja tapakulttuuriin (Åkerlund 2002: 126). Useimmista muista Euroopan romaniryhmistä poiketen Suomen romanit, Ruotsin tatarit ja Saksan sintit ovat haluttomia paljastamaan ulkopuolisille kieltään ja kulttuuriaan. Ehkä kaikkein ankarimmin asiaan suhtautuvat Saksan sintit, jotka eivät saa edes puhua romanikielen murretaan ulkopuolisten kuullen; tähän ovat olleet vähintään osasyynä kansallissosialismin aikaiset vainot. Suomen romanit ovat joskus luonnehtineet romanikieltä ainoaksi omaisuudekseen, jota ei ole

haluttu antaa ulkopuolisten käsiin (Åkerlund 2002: 126). Hedman (2004: 42) toteaa, että romanikielen merkitys ilmenee paitsi identifioitumisena ryhmään ja sitoutumisena romanien tapoihin ja elämään, myös rajana valtaväestöön nähden.

Näkemykset romanikielestä pelkästään romaneille kuuluvana heijastuvat myös verkkokeskusteluissa. Seuraavat kaksi esimerkkiä (1 ja 2) ovat Suomi24.fi-sivuston Romano-keskustelupalstalta keskustelusta ”Kaajeet ja romanikieli”.

- (1) o ihan se rakli vastaa
Sinä Mirkku- markku olet täysi rakli, tai vanha kaaji. Oikia kaajeen vanha piika. Aloittaja- nappaa tukasta kiini sitä kaajia ,te les riso mui. Ei kaajeet saa puhua meidän kieltä. Onneksi tuo mirkku kaaji ei sitä osaa oikein, mitä lie sian saksaa. (Suomi24.fi, Romano-keskustelupalsta/Manu Manne 20.4.2010 20:44.)
- (2) Ei ehdottomasti ei saa kaajeet alkaa puhumaan kaaleen kieltä. Ne ketä on pitäny joskus akkanaa kaajetta niin ne akat on oppineet sitte siinä.Pitääkö meiltä kaikki viijä? Eikö saa tummaheimo elää ja olla? On hieno asija että on oma kieli. On oma kulttuuri tavat ja perinteet. Entis aikoja ei haikailla eikä ne hyvijä muistoja kaikille olekkaa...vaa on niin että kautta historian on siitä ollu vaa haittaa jos on kaajeet tietoisia kaaleen asijoista. Ja tää on totuus.Kaajeittwen maassa kumminki asutaan ja eletään eli tullaan toimeen. Ei mitään henkilökohtaista ja vastaan mutta minun mielipide on Kaaleen kieli kuuluu vaa KAALEIL-LE. (Suomi24.fi, Romano-keskustelupalsta/R***** 23.4.2010 12:49.)

Romanien keskuudessa katsotaan vahvasti myös, että romanikieltä saavat toisille romaneille opettaa ja sen oppimateriaaleja tuottaa vain romanitaustaiset. Käytännössä tämä on heijastunut myös siten, ettei edes puoliromanien ole aiemmin sallittu yksin laatia oppimateriaaleja, vaan on edellytetty, että tekijöiden joukossa on myös täysin romanitaustaisia. Pääväestöön kuuluvien osallistumista oppimateriaalien ja sanakirjojen tuottamiseen on pyritty estämään. Ei-romanien laatimia sanasto- ja oppimateriaaleja on nimitetty romaniväestön keskuudessa muun muassa ”pilatuiksi” tai ”pilallisiksi”.

Romaniyhteisön ajattelutapaan vaikuttaa myös sen gerontokraattinen ja hierarkkinen järjestäytyminen. Yhteisössä arvostetuimpia ovat vanhat miehet ja alimmalla portaalla nuoret naiset. (Viljanen 1974, 1994; Vil-

janen–Saira 1979; Granqvist & Viljanen 2002; Steiner 2004.) Yhteisössä ajatellaan monesti, että asioiden tekemiseen tarvitaan jonkun yhteisössä arvostetun jäsenen lupa. Esimerkiksi Viljo Koivisto pyysi aikoinaan oppimateriaalien tuottamiseen julkisesti luvan romaninyhteisöltä. Myös romanikielen tutkimuksen ja yliopisto-opetuksen sanotaan romanien piirissä usein tapahtuvan romanikielen lautakunnan aktiivisimpien romanijäsenten luvalla, heidän toimiessaan ”puskureina” muun romaninyhteisön ja pääväestöön kuuluvien tutkijoiden välillä. Toisaalta nuoremmat eivät tee yhteisön vanhempien jäsenten jo kieltämiä asioita.

Romanikielen kirjallisen materiaalin julkisuus

Romanikielen omistajuuteen kytkeytyy osaltaan myös kielen julkisuuden ja salakielitehtävän välinen jännitteinen ristiriita. Romanikielen statuksesta salakielenä Suomessa ja muualla ovat maininneet monet (Ganander 1780; Valtonen 1968: 241–245; Holzinger 1995; Grönfors–Virolainen–Åkerlund–Lounela 1997: 175; Bakker–Kyuchukov 2000: 30; Pirttisaari 2002: 17–18). Hedman (2004: 45) liittää salakielistatuksen romaneihin eri aikoina kohdistettuihin julmuuksiin sekä kirkon harjoittamaan politiikkaan ja tutkijoiden toimintaan, josta on koitunut romaninyhteisölle ajoittain vahinkoa. Salakielenä romanikieli on ollut suoja ja turva, ja se on tarjonnut mahdollisuuden puhua perheen sisäisistä asioista vieraissa paikoissa aikana, jolloin romanit vielä kiersivät. Romanikielestä on ollut salakielenä myös suoranaista hyötyä kurinpidon, kehotusten, varoitusten ja neuvojen välittäjänä esimerkiksi kaupanteossa ja asioitaessa viranomaisten kanssa (Tolkki 1951: 264–266; Valtonen 1968: 214; Hedman 2004: 43–45).

Sekä sanakirjatyö ja muu tutkimus että koko kirjallisen perinteen käynnistyminen ovat kohdanneet vaikeuksia juuri kielen yhteisöllisen ja salakieliluonteen vuoksi, kun kirjallisen muodon on pelätty auttavan ulkopuolisia oppimaan kieltä. Hedmanin (1996) opettajan opasta saattoivat tilata Opetushallituksesta romanikielen opettajat ja muut romaniväestöön kuuluvat, muttei pääväestö. Romanilehdissä romanikielisiä kirjoituksia saattoi pitkään julkaista, kunhan suomennosta ei liitetty mukaan. Nykyään kuitenkin suomennoksia on alettu toivoa rinnalle. Hedman (2009) on selvittänyt romanien suhtautumista romanikielen julkisuuteen

ja romanikielisiin painettuihin julkaisuihin. Hedmanin (2009: 57) selvityksestä käy ilmi, että romanien perinteiset asenteet heijastuivat vieläkin asenteissa romanikielisiin painettuihin julkaisuihin, vaikka romanikielen julkiseen käyttöön mediassa ja painotuotteissa suhtaudutaan nykyään myönteisemmin. Kirjallisista materiaaleista romanit toivoivat eniten kielen oppikirjoja, koska romanien tarve ja halu oppia lisää kieltään ovat ilmeisiä ja koska oppikirjat on kohdistettu ainoastaan romanien käyttöön. Muiden romanikielisten kirjojen kuin oppikirjojen julkaisemiseen kielteisesti suhtautuvia tai asiasta epävarmoja oli yhteensä 32 %. Romanit suhtautuivat kielteisesti esimerkiksi romanilehtien romanikielisiin kirjoituksiin, jotka on käännetty myös suomeksi.

Sanakirjat romanikielen normittajina

Suomessa puhuttavaa romanikielen murretta eli Suomen romanikieltä on huollettu liki neljänkymmenen vuoden ajan. Virallisen kielenhuollon alku sijoittuu 1960- ja 1970-lukujen taitteeseen, jolloin opetusministeriön asettama ortografiakomitea sai tehtäväkseen laatia romanikielelle helppolukuisen ortografian ja sen pohjalta normatiivin sanaston (MNS, 1971). Kotimaisten kielten tutkimuskeskuksessa romanikielen huolto sai muodollisesti alkunsa vuonna 1997, jolloin romanikielen lautakunta aloitti ensimmäisen kolmivuotisen toimikautensa.

Romanikielen lautakunnan tehtävänä on päättää romanikielen käyttöä koskevista periaatteellisista ja yleisluonteisista suosituksista. Suositukset koskevat sekä kielen rakennetta että sen sanastoa. Suurehko osa lautakunnan työstä on termityötä, jonka tavoitteena on kehittää kieltä toimivammaksi viestintävälineeksi nyky-yhteiskunnassa.

Romanikielen normitustyön ollessa yhä kesken romanien keskuudessa käydään runsaasti keskustelua siitä, millainen ja kenen romanin romani-kieli on hyväksyttävintä. Karjalan romanien mielestä Karjalan romanimurre on kunnollista romanikieltä, Pohjanmaan romanien mielestä taas heidän oma murteensa. Yhteisesti hyväksytyyn yleiskielen vielä puuttuessa joudutaan romanikielen sanakirjatyössä ottamaan huomioon ainakin kahden päämurteen puhujat. Näkemykset ”oikeimmasta” romanikielestä vaihtelevat ja henkilöityvät pienessä käyttäjäyhteisössä, mitä seuraavat Suomi24.fi-sivuston keskustelusta poimitut esimerkit hyvin kuvastavat

(3–6):

- (3) kaaleen kielestä
onhan se hyvä että kaaleen kieltä pidetään yllä mutta ne jotka sitä voisi opettaa niin ei ainakaan ne jotka on muuttanut aiton kaaleenkielen ihan muuksi en hyväksy mitään kansan välistä kaaleen kieltä siis pitäs ottaa kantaa vanhempien ihmisten tähän asiaa ei opiia vuolasrannoilta eikä hedmanneilta koska ne ei varmaan edes ymmärrä oikeeta kaaleen kieltä pidetään se kaaleenkieli joka on ollut aina eikä sitä tule muuttaa ja se kirjoitetaan niin kun se lausutaan t mustalainen. (Suomi24.fi, Romano-keskustelupalsta/ mustalaismies 29.1.2011 09:29.)
- (4) – – se vielä kun opetetaan väärää mustalais kieltä minun lapsille jos joku kävis opettamaan tuota kieltä niin ei hyvä seuraisi olen itse oppinut mustalaiskielen kotona sekä myös lapseni ei pitäisi vääristellä meidän omaa kieltä ja mustalais kieli kirjoitetaan samalla tavalla kun se lausutaan ja vielä miten opettaja kaaleet opettaa mustalaisuutta kaa-jeen vaatteet päällä sen vaan sanon että jos minun kuullen joku puhuu kansan välistä mustalais kieltä niin voin sanoa että lääkärille tulee sillä asiaa ja jos joku haluaa tietään kuka olen ja yhteystietoja niin kysy vaan rohkeesti kerron kyllä ps hyvä tenho. (Suomi24.fi, Romano-keskustelupalsta/ vanhan ajan kaalo 19.10.2012 06:28.)
- (5) haluaisin sanoa kaaleen kielestä sen verran niille jotka sitä muok-kaavat ja opettavat valtaväistölle että älkää muuttako mustalais kieltä miksikään kansanväliseksi kielelle ja siksi toisekseen aitot mustalaiset kirjoittavat mustalaiskielen sanan samalla tavalla kun se lausutaan eikä tämä kieli kuulu valtaväistön oppiin älkää käykö muuttamaan mustalaiskieltä eikä mustalaiskulttuuria. jos joku tykkää tästä huonoa niin kysy vaan kuka tämän kirjoitti niin kerron kyllä nimeni ja numeroni. (Suomi24.fi, Romano-keskustelupalsta/ mustalainen maalta 3.10.2012 06:44.)
- (6) se mistä sinä nyt on puhe, on sitä kaksosien kieltä, jokainen jolla on ollut kaksoilapsia tietää miten helposti nämä tekevät oman keskinäisen kielen. Se kieli mistä puhuit, hedaman ja mirannan opettama, sikäli kun ymmärän se kieli on saatu 1900 luvun vaihteen jälkeen syntyneiden romanien kielestä, ja on kielenä aidompi ja kästeeltään likkaapi, koska yleinen kehitys ei sitä vielä ollut turmellut, kuten nyt on valtaväen kieli. (Suomi24.fi, Romano-keskustelupalsta/ tuomo 29.1.2011 12:48.)

Sanakirjat ovat oppikirjojen ohella keskeinen käytännön kielen normittaja. Käytännössä niin sanakirjoja kuin oppikirjoja laatii vain kourallinen toimijoita, romaneja ja ei-romaneja. Suurin osa romanitoimijoistakin on akateemisesti koulutettuja, mutta vailla kielitieteellistä ja teoreettista romanilingvististä osaamista. Käsitukset romanikielestä perustuvatkin paljolti heidän romaniyhteisössä kuulemalla oppimaansa kieleen, koska kirjallinen perinne on kovin ohutta ja uutta. Viljo Koiviston romanikielen oppi- ja sanakirjoissa on käytetty pääosin Pohjanmaan romanien murretta, jolle on leimallista Itä-Suomen romanien murretta voimakkaampi ruotsin kontaktivaikutus. Suomen romanikielen kirjakieli onkin lähtenyt muotoutumaan pitkälti länsimurteen pohjalta ennen kaikkea Koiviston normitus- ja sanakirjatyon vuoksi. Käytännössä monet Itä-Suomesta lähtöisin olevat romanit ovat toisaalta suhtautuneet nuivasti sanakirjojen runsaisiin skandinaavisiin lainasanoihin.

Suomen romanikielen dokumentoidusta leksikosta jopa yli 40 % on germaanisperäistä, keskiylä- ja alasaksasta, tanskasta ja ruotsin murteista lainattua. Suomalaisten lainasanojen osuus jää vain 8 %:iin, joko tietoisien välttämisen tuloksena romanikielen salakieliluonteen tai ruotsin prestii- siaseman vuoksi. (Valtonen 1968.) Kuitenkin osa romanikielen sanastoa aiemmin kehittäneistä on pyrkinyt laajentamaan sanastoa nimenomaan mukauttamalla siihen suomalaisperäistä sanastoa, esim. *ankkuris* 'ankkuri', *bananis* 'banaani', *divisioonas* 'divisioona', *eetteris* 'eetteri'.

Nykyään sanakirjojen laatijoiden ja kielenhuoltajien suhtautumista konkreettisten lekseemien lainaamiseen lähikontaktikielistä leimaa purismi, jonka seuraukset ovat usein käänteisiä selkeyden ja ymmärrettävyyden tavoitteelle. Fonologiselta muodoltaan kaikkein läpinäkyvimmin skandinaavisia ja varsinkin suomalaisia lainasanoja pyritään tietoisesti välttämään ja korvaamaan ne kielen alkuperäisinä pidettävien resurssien avulla, mieluiten indoarjalaisiin kieliin eli sanskritiin ja prakritiin palautuvilla sanoilla. Tähän pyrkimykseen liittyy erilaisia ongelmia. Eronteko lainattujen ja alkuperäisten resurssien välillä edellyttää sellaista tietoa romanikielen sanojen historiasta, jota termien kehittäjillä ei ole riittävästi, vaan valinta on usein tunnepohjainen ja intuitioon perustuva. Suomen muotoja muistuttavien romanikielen sananmuotojen välttäminen vaikuttaa myös sanakirjojen taivutuskaavoihin. Romanikielen lainasubstantiivien produktiivisin luokka ovat *-os*-loppuiset maskuliinit, esim. *appelsiin-os* 'appelsiini-NOM.SG', *materiaal-os* 'materiaali-NOM.SG'. Niiden

monikon nominatiivi muodostetaan useimmissa Euroopan romanikielen murteissa kreikasta lainatulla päätteellä *-i*. Koska monikon nominatiivit *appelsiin-i*, *materiaal-i*, ovat monissa tapauksissa homonymisiä suomen yksikön nominatiivin muotojen kanssa, ovat romanikielen sanakirjojen tekijät päätyneet usein *-e*-loppuisiin monikon nominatiiveihin *appelsiin-e*, *materiaal-e*. Monikon nominatiivin *-e* on indoarjalainen päätte ja luonteenomainen esieurooppalaisille *-o*-loppuisille maskuliinisukuisille substantiiveille, esim. *džeen-o* 'mies-NOM.SG', *rakl-o* 'poika-NOM.SG'. Romanikielessä esieurooppalaiset (sanskritista, prakritista, iranilaisista kielistä sekä osin Bysantin ajan kreikasta lainatut) ja Euroopan kielistä lainatut sanat kuuluvat eri taivutustyyppihin. Sanakirjantekijät siten mieluummin päätyvät romanikielen oman, perityn kahtiajaon vastaisiin, virheellisiin muotoihin kuin hyväksyvät suomen sananmuotojen kanssa yhteenlankeavia taivutusmuotoja.

Viime vuosikymmenien sanakirjatyölle leimallista on ollut indoarjalaisen sanastonosan ylikorostaminen. Suomen romanikielen alkuperäinen, indoarjalainen sanasto käsittää kuitenkin vain noin 450 juurta ja on käytännön sanastotyöhön suppea. Euroopan romanimurteissakaan indoarjalaisista sanastoa ei ole kuin yhteensä 650–700 juurta. (Granqvist–Pirtti-saari 2003.) Indoarjalaisia resursseja hyödyntävät sanaliittomuodostukset ovat paitsi monitulkintaisia, tyypillisesti myös rakenteellisesti raskaita, useista peräkkäisistä genetiiveistä koostuvia monimutkaisia nominaalilausekkeita, kuten *bar-o-komu-jen-go-sikjib-os-ko-sentrum* 'iso-M.SG + ihminen-OBL.SG-GEN + opetus-OBL.SG-GEN + keskus = aikuiskoulutuskeskus'.

Sanakirjantekijöiden ja muidenkin kirjallisesti romanikieltä käyttävien uutena tendenssinä leksikon rikastuttamisessa on lainaaminen erityisesti englannista, esim. *addiktos* 'addikti' < engl. *addict*, *international* 'kansainvälinen' < engl. *international*, *objekto* 'kohde' < engl. *object*, *religios* 'uskonto' < engl. *religion*. Englantilaisperäisten ja muiden vieraista kielistä lainattujen sanojen ongelmina ovat ymmärrettävyys ja istuvuus romanikieleen, adaptaation onnistuneisuus (esim. *objekto* on mukautettu alkuperäiseen indoarjalaiseen taivutustyyppiin lainasubstantiivien taivutustyyppien sijaan) ja taivutus (miten *international* taipuu suvussa ja luvussa?).

Koska sanakirjat oppikirjojen ja kielioppiteosten tavoin normittavat romanikieltä, niiden laatijoilla on ratkaisuisaan vastuu kielen sanastosta ja rakenteesta. Yhteisön hyväksynnän lisäksi siten olennaista on teosten

laatijoiden oma romanikielen taito ja taju. Edellä on ollut esimerkkejä kielelle ominaisten jaottelujen purkautumisesta sanakirjoissa, esim. *bolib-os-ko* ~ *bolib-a-ko* 'maailma-OBL.SG-GEN', *merkn-es-ko* ~ *merkn-os-ko* 'merkki-OBL.SG-GEN', *momul-is-ko* ~ *momuj-a-ko* 'kynttilä-OBL.SG-GEN', *fiin-a* ~ *fiin-e* 'viisas-PL', *mortšavitik-a* ~ *mortšavitik-e* 'nahkainen-PL' jne. Niiden lisäksi esimerkiksi verbien taivutuksissa ja johto-opissa esiintyy paljon variaatiota, esimerkiksi johdettuna transitiiviverbinä *lan-sav-a-a* 'hävetä-PRS.1SG-IND', mutta johtamattomana primaariverbinä *lans-eh-a* 'hävetä-PRS.2SG-IND'.

Nykyään oman arvionsa mukaan romanikieltä puhuu hyvin vain kolmasosa Suomen romaneista (Hedman 2009: 13). Nykyään kuitenkin paraskin romanikielen taito on huomattavan rapautunutta verrattuna 1960-luvun taitaviin puhujiin: useimmilla iäkkäilläkin puhujilla esiintyy muun muassa suomen modaaliverbejä, negaatioita ja konjunktioita, substantiivitaivutusten analogista yhdentymistä ja epäkieliopillisia persoonamuotoja tuottavaa verbiparadigmojen yksinkertaistumista, joista ei ollut 1960-luvun parhailla puhujilla vielä merkkejä (Granqvist 2012). Sama rapautuminen heijastuu tietenkin sanakirjoihinkin, ajan kuluessa yhä voimakkaammin.

Lähdekritiikki ja suhtautuminen lähteisiin

Osaltaan purismiin kytkeytyy lähdekritiikki, joka kuitenkin on pikemmin tunnepohjaista ja lähteiden laatijoiden henkilöä koskettavaa kuin objektiivisesti lähteen romanikielen laatua tarkastelevaa. Tämä on nähtävissä muun muassa Viljo Koiviston romanikielen sanakirjoissa (1994, 2001), joiden molempien lähdeviitteissä korostuvat Arthur Thesleffin 1900-luvun alun sanakirja sekä siihen sisältyvät muutamien vanhempien lähteiden hakusanat.

Toisaalta Koiviston romanikielen sanakirjoissa selvimmin aliedustettuna on Yrjö Temon kielenaines. Temon romanikieli on aikakaudelleen poikkeuksellisen konservatiivista, ja siinä ovat tallella muun muassa määräiset artikkelit, lokatiivisija, verbien preesensin persoonapäätteiden pitkä vokaali (esim. *rakkav-eel-a* 'puhua-PRS.3SG-IND') ja prepositiot. Toisaalta Temo on sisällyttänyt sanastoainekseensa huomattavasti muita kieltä dokumentoineita enemmän suomalaisperäisiä lainasanoja, joista

on vaikea sanoa, ovatko ne käytössä kuultuja vai Temon itse kieleen mukauttamia. Suomalaisten lainojen merkittävä osuus lienee yksi syy siihen, että Temon kielenainesta on hyljeksitty ja alettu pitää romanien keskuudessa epäluotettavana. Todellisuudessa kyky lainata ja mukauttaa kieleen lainasanoja on kuitenkin vitaalisuuden merkki.

Romanikielen sanakirjoille leimallinen ajallinen ja alueellinen hajanaisuus on vaikuttanut lähdekritiikkiin ja eri lähteisiin suhtautumiseen. Vanhimmat nykykielen sanakirjojen lähteet ovat 1700- ja 1800-luvun taitteesta. Romanikielen sanakirjojen laatijoiden ja monien romanien luoteutumpana pitämä lähde eli Arthur Thesleffin sanakirja on vuodelta 1901. Lähteiden ajallisen hajanaisuuden seurausta on monien nykykielestä jo kadonneiden muotojen sisällyttäminen sanakirjoihin. Tällaisia muotoja ovat esimerkiksi *sin* 'on' (nyk. *hin*), *daav-* 'antaa' (nyk. *d-*) ja *tša-* 'viipyä' < *ačh-* (nyk. *aaħħ-*). Verbin *s-/h-* 'olla' *s-*alkuiset 3. persoonan muodot *sin*, *si* 'on, ovat' olivat vielä 1700- ja 1800-lukujen taitteessa yleisempiä kuin *h-*alkuiset *hin* ja *hi*, mutta jäivät jo 1800-luvun loppuun mennessä enimmäkseen käytöstä. Preesensin yksikön 1. persoonan *-av-*päätteen lopun labiaalinen heikkeneminen ja kato on romanikielen murteissa yleinen myöhempi muutos, jonka tuloksena voivat olla myös *-aw*, *-ao* (Matras 2002: 144). Reduktio on Suomen romanissa ilmeisesti varsin vanhaa perua, koska jo Gananderilla esiintyy sekä redusoitumattomia että redusoituneita yksikön 1. persoonan muotoja. Redusoitumattomissa muodoissa yksikön 1. persoonan päätteen lopun *-v-* säilyy, esim. *cha-w* 'syödä-PRS.1SG', kun taas redusoituneista muodoista se katoaa (esim. *drabaw-a* 'lukea-PRS.1SG' ja sekä romanikielen että suomen persoonapäätteet sisältävät *drabaw-an* 'lukea-PRS.1SG' ja *akraw-an* 'puhua-PRS.1SG'). Nykykielessä selvimpiä jälkiä preesensin 1. persoonan päätteen *-v-*aineksesta esiintyy verbissä *pi(iv)-* 'juoda', jonka taivutusmuodoista osa noudattaa vokaalivartalaisia verbejä (*pi-l-a* 'juoda-PRS.3SG-IND', *pi-n* 'juoda-PRS.2/3PL') ja osa konsonanttivartalaisia (*piiv-ah-a* 'juoda-PRS.1PL-IND, *piiv-es* 'juoda-PRS.2SG'). Alun perin romanikielen murteissa painollisten verbien preesensin persoonapäätteiden vokaalin (eli teemavokaalin) *-é-* tai *-á-* lyhentyminen on Suomen romanissa melko tuore kehitys, koska vielä Valtonen (1968: 132) saattoi dokumentoida pitkävokaalisia muotoja osalla puhujista (esim. *tenkav-ēh-a* 'ajatella-PRS.2SG-IND'). Nykykielessä teemavokaali on aina lyhyt. Soinniton aspiroitu affrikaatta *čh* on kadonnut kokonaan Suomen romanikielestä. Sen on korvannut jo varhain joko

tš, s(s) tai ħ(ħ) (*čhaj* > *tšai* 'tyttö', *pačh-* > *pass-* 'uskoa', *ačh-* > *aaħħ-* 'olla'). Sanakirjojen *tša-* onkin nykykieleen kuulumaton, historiallinen muoto. Nykyisten sanakirjojen kielenaineksen alueellinen hajanaisuus liittyy niiden laatijoiden omiin murretaustoihin suhteessa eri lähteiden laatijoiden dokumentoimiin murteisiin: romanikielen länsi- vs. itämurteen asuihin ja muotoihin.

Itsesensuuri ja romanien kielelliset tabut

Eri-ikäisten diskurssiin oman leimansa lyövät myös eri sukupuolten sosiaaliseen kanssakäymiseen liittyvät rajoitukset ja suoranaiset tabut. Tabuina pidetyt aiheet ovat sosiaalisesti kiellettyjä. Romanikulttuurissa on suuri joukko puheenaiheita, joihin nuorempien romanien ei ole mahdollista viitata suoraan vanhempien ihmisten kuullen. Tabuja ovat kaikki seksuaalisuuteen ja suvunjatkamiseen liittyvät aiheet, seurustelusuhteet, avioliitto sekä monet ruumiintoiminnot. Tabuoinnissa keskeisenä on ruumiin skeema, jota Anna Maria Viljanen on tarkastellut laajalti. Se jakaa ihmisruumiin rituaalisesti puhtaaseen yläosaan ja likaiseen alaosaan. (Viljanen 1974, 1994; Viljanen–Saira 1979; Granqvist & Viljanen 2002; Steiner 2004.)

Tabuointi ulottuu esineisiin ja vaatekappaleisiin, jotka liittyvät välillisestikin esimerkiksi seksuaalisuuteen tai suvunjatkamiseen tai jotka koskettavat rituaalisesti epäpuhtaina pidettäviä ruumiinosia. Hedelmällisyysikäisten naisten ja miesten alusvaatteet ovat käyttämättöminäkin epäpuhtaita. Ne säilytetään ja pestään erillään rituaalisesti puhtaista vaatteista, eikä niitä pidetä esillä siten, että vanhemmat romanit voisivat ne nähdä. Niistä ei myöskään voida puhua sukupolvienvälisessä diskurssissa mutta kylläkin saman ikäisten romanien kesken. (Granqvist–Viljanen 2002.)

Granqvist ja Viljanen (2002) ovat todenneet, että tabuointi kohdistuu paitsi sanoihin, joiden referenssin kohde on tabu, myös sellaisiin sanoihin, joiden fonologinen muoto muistuttaa varsinaisia tabusanoja. Tällaisia ovat esimerkiksi sanat *beero* 'auto' (vrt. *peelo* 'penis'), *hinnav-* 'ehtiiä' (vrt. *ħin[n]iba* 'ulostaminen'), *rissa* 'kipu' (vrt. *rinša* 'takapuoli'), joiden referenssin kohteet eivät itsessään ole millään tavoin tabuja toisin kuin aitojen tabusanojen merkitysisältö. (Granqvist–Viljanen 2002). Sanaa *piero*

'jalka' pyritään usein välttämään suomen *pieru-* ja *pierrä-*sanoihin johtavien konnotaatioiden vuoksi. Esimerkiksi vuonna 2011 järjestetyssä kieli-seminaarissa kyseinen romanikielen sana herätti häpeänsekaista hilpeyttä osallistuneiden romaninaisten keskuudessa. Muiden romanimurteiden sanaa *nula* "0" ei voida käyttää Suomen romanikielessä, koska romanit yhdistävät yleisesti sen ruotsin verbiin *knulla*. Romanitaustaisten tekijöiden itsetsensuurin vuoksi osa sanoista on vaikeuttanut Suomen romanikielen sanakirjatyötä. Sanakirjojen laatijat ovat pelänneet, että romanilukijat luulevat heidän sisällyttäneen teoksiinsa hävettäviä tabusanoja.

Lopuksi

Tässä artikkelissa olen tarkastellut Suomen romanikielen sanakirjatyön historiaa sekä sanakirjatyöhön liittyviä eettisiä kysymyksiä. Historiallisesta näkökulmasta leksikografialle leimallista on ei-romanitoimijoiden huomattava osuus aina 1980-luvulle saakka, jolloin yhtäältä ensimmäinen romanikielen, -kulttuurin ja -politiikan aktivisti Viljo Koivisto ryhtyi laatimaan sanakirjaa yhteistyössä Kotimaisten kielten tutkimuskeskuksen kanssa ja toisaalta romanitaustainen maallikko Yrjö Temo laati sanakirjansa käsikirjoituksen. Sanakirjojen ja romanikielen parissa työskentelee sekä pääväestöä ja että romaneja. Hyvin keskeinen eettinen kysymys on kielen omistajuus. Tämä korostuu etenkin romanikielessä, jonka osa romaneista haluaa tiukasti pitää omana omaisuutenaan.

Romanikielen omistajuuteen kytkeytyy osaltaan myös kysymys sen julkisuudesta ja salakielitehtävästä. Sekä sanakirjatyo ja muu tutkimus että koko kirjallisen perinteen käynnistyminen ovat kohdanneet vaikeuksia juuri kielen yhteisöllisen ja salakieliluonteen vuoksi. Osa romaneista on edelleenkin sitä mieltä, ettei pääväestö saisi oppia kieltä, osa taas suhtautuu avoimemmin.

Niin romanikielen kuin muidenkin kielten sanakirjat ovat kieltä normittavia. Romanikielen sanakirjojen merkitys kodifoinnissa korostuu, koska Suomen romanikieleltä edelleen *de facto* puuttuu standardi, ja romaniyhteisössä näkemykset oikeasta romanikielestä vaihtelevat suuresti. Hakusanojen valinnassa on pyritty jonkinlaiseen purismiin – lähinnä suomalaisperäisiltä vaikuttavien ja joidenkin skandinaavisten lainasanojen karttamiseen. Sanakirjantekijöiden ja muiden kirjallisesti

romanikieltä käyttävien uutena tendenssinä leksikon rikastuttamisessa on lainaaminen erityisesti englannista, vaikka uudet lainasanat joskus istuvatkin huonosti romanikieleen.

Purismiin kytkeytyy lähdekritiikki, joka kuitenkin on pikemmin tunnepohjaista ja lähteen laatijoiden henkilöä koskettavaa kuin objektiivista lähteessä käytetyn romanikielen tarkastelua. Esimerkiksi Yrjö Temon romanikieli on aikakaudelleen poikkeuksellisen konservatiivista, mutta hänen kielenainestaan on silti muiden romanien keskuudessa aliarvostettu ja pidetty epäluotettavana.

Hakusanojen valintaan vaikuttavat myös romanien kielelliset tabut. Ne vaikuttavat sekä sanakirjojen romanikieliseen että suomenkieliseen ainekseen. Kielelliset tabut ilmenevät romaneilla laajemmin ja monimuotoisemmin kuin pääväestöllä, ja niistä johtuva itesesensuuri on vaikuttanut myös sanakirjatyöhön.

KLAAS RUPPEL

Maahanmuuttajille laadittavat sanakirjat ja etiikka

Artikkelissani tuon esille eettisiä näkökulmia, jotka liittyvät leksikografiseen työhön maahanmuuttajakielten parissa. Aiheen käsittelyssä pääpaino on käytännön leksikografiseen työhön liittyvissä seikoissa. Maahanmuuttajakielillä tarkoitetaan kieliä, joiden puhujia on saapunut Suomeen vasta viime vuosikymmenten aikana. Vähemmän tunnetun kielen lisäksi uudet tulokkaat tuovat mukanaan Suomeen myös oman kulttuurinsa, joka saattaa erota Suomessa vanhastaan olevista kulttuureista enemmän tai vähemmän.

Maahanmuutto on Suomessa suhteellisen nuori ilmiö. Maahanmuutto alkoi kasvaa Suomessa merkittävästi vasta 1990-luvulla Neuvostoliiton hajottua, jolloin inkeriläisten muutto Suomeen alkoi. Somalian sisällissota, entisen Jugoslavian alueella alkanut sota sekä Iranin ja Irakin epävakaa tilanne toivat Suomeen eniten pakolaisia. (Nieminen 2004.)

Maahanmuutto ja lainsäädäntö

Maahanmuuttoon liittyvät poliittiset ratkaisut heijastuvat lainsäädäntöön ja siinä käytettyyn sanastoon. Sana *maahanmuuttaja* esiintyy ensimmäisen kerran jo vuonna 1973 sosiaali- ja terveysministeriön asetuksessa lasten päivähoitosta (16.3.1973/239). Tämän jälkeen sanaa näytetään käytetyn lainsäädännössä vasta vuonna 1992 opetusministeriön päätöksessä ylimääräisen valtionavustuksen myöntämisestä maahanmuuttajien tukiopetukseen peruskoulussa ja lukiossa (1609/1992). Termi *kotouttaminen* esiintyy ensimmäisen kerran 1.5.1999 voimaan tullessa laissa maahanmuuttajien kotouttamisesta ja turvapaikanhakijoiden vastaan-

otosta (493/1999) (*kotouttaa*-sanankuperästä ks. Jussila tk. s. 179). Laki kumosi 31.12.1991 voimaan tulleen lain turvapaikanhakijoiden vastaanottoasemista ja -keskuksista (1465/1991). Vuonna 1999 voimaan tulleen lain puolestaan kumosi 1.9.2011 voimaan tullut laki kotoutumisen edistämisestä (1386/2010).

Voimassa oleva laki 1386/2010 (3. §) määrittelee *kotoutumisen* ja *kotouttamisen* seuraavasti: ”Tässä laissa tarkoitetaan: 1) *kotoutumisella* maahanmuuttajan ja yhteiskunnan vuorovaikutteista kehitystä, jonka tavoitteena on antaa maahanmuuttajalle yhteiskunnassa ja työelämässä tarvittavia tietoja ja taitoja samalla kun tuetaan hänen mahdollisuuksiaan oman kielen ja kulttuurin ylläpitämiseen; 2) *kotouttamisella* kotoutumisen monialaista edistämistä ja tukemista viranomaisten ja muiden tahojen toimenpiteillä ja palveluilla – –.”

Huomiota herättää, että kotoutuminen nähdään maahanmuuttajan ja yhteiskunnan välisenä vuorovaikutuksena. Kotoutuminen ei siis lainsäätäjän näkökulmasta ole pelkästään maahanmuuttajan vastuulla. Ennen kaikkea viranomaisten toimintaa luonnehditaan kotoutumista edistäväksi ja tukevaksi kotouttamiseksi.

Maahanmuutto ja leksikografia

Kulttuurisesti läheltä tulevien maahanmuuttajien kohdalla ei yleensä puhuta kotouttamisesta eli integraatiosta. Pohjoismaalainen tulee hyvin toimeen Suomen yhteiskunnassa, kunhan kieli tulee tutuksi. Kun erot kulttuurissa, uskonnoissa, historiassa, politiikassa ja muilla elämänalueilla suurenevat, kasvaa tarve kotouttamiseen ja sitä tukevaan sanakirjatyöhön. Maahanmuuttajakielten sanakirjoja tehdään siis sellaisille ryhmille, joiden tiedetään tarvitsevan kotoutumistukea ja joiden tarpeisiin ei ole olemassa sopivia sanakirjoja.

Leksikografian kannalta maahanmuutto isojen, lähinnä Euroopassa puhuttujen kielten alueelta ei vaadi välitöntä reaktiota. Suomen kieli on parina monenlaisissa sanakirjoissa englannin, saksan, ranskan, italian, venäjän ja monien muiden kielten kanssa. Ainakin tähän asti näiden kieliparien sanakirjoja on myös pidetty ajan tasalla. Sen sijaan sanakirjojen tarve on suuri niiden kielten osalta, joista ei ole vielä tehty sanakirjoja, joissa suomi on lähde- tai kohdekielenä. Suomalainen maahanmuutta-

jakielten leksikografia keskittyy siis Suomessa uusiin kieliin, jotka ovat vailla sanakirjoja.

Kun sanakirjoja laaditaan ja käytetään saman kulttuurin piirissä, eettiset kysymykset ovat erilaisia kuin kulttuurien välisessä eli kaksi- tai useampikielisessä leksikografiassa. Kulttuurien sisäisiä sanakirjoja ovat esimerkiksi suomen nykykieltä selittävät yksikieliset yleissanakirjat Nykysuomen sanakirja (NS, 1951–1961), Suomen kielen perussanakirja (PS, 1991–1994) ja Kielitoimiston sanakirja (KS, 2004–). Tekijät ja käyttäjät elävät samassa ajassa ja samoissa kulttuurisissa kehyksissä (vrt. esim. historialliset sanakirjat, ks. Lehtosalo & Kuutti tk.). Kaksikieliset sanakirjat ovat välittäjiä kahden kielen välillä. Mitä etäämpänä toisistaan kielten edustamat kulttuurit ovat, sitä tärkeämmiksi nousevat eettiset kysymykset. Tämä koskee erityisesti kielen oppijoille suunnattuja pedagogisia sanakirjoja, sillä niiden käyttäjät opiskelevat kielen lisäksi kulttuuria.

Suomessa on tehty joitain pieniä kielen oppijoille suunnattuja kaksikielisiä sanakirjoja, jotka tosin ovat luonteeltaan enemmänkin sanalistoja, esimerkiksi Suomi–somalı–suomi-taskusanasto (SSS, 2010). Kulttuurien välisen leksikografian tuotteita ovat kuitenkin myös yksikieliset ulkomaalaisille oppijoille tarkoitetut sanakirjat, ja tällaisia sanakirjoja on Suomessa tehtykin. Mainittakoon Timo Nurmen Suomen sanakirja opiskelijoille ja ulkomaalaisille (SSOU, 2009), joka käsittää 17 000 sana-artikkelia, sekä Anne Saarikallen ja Johanna Vilkunan laatima Suomen kielen sanakirja maahanmuuttajille (SSM, 2010), joka on 3 500 sana-artikkelin laajuinen.

Vaikka nämä mainitut yksikieliset sanakirjat ovat molemmat nimenomaan maan rajojen (ja kulttuurin) ulkopuolelta tuleville oppijoille tarkoitettuja pedagogisia teoksia, ne eroavat voimakkaasti toisistaan. SSOU keskittyy laajojen taivutustietojen ja runsaan esimerkistön esittämiseen. Lisäksi sana-artikkeleiksi on otettu runsaasti nimiä (mm. Suomen kaupunkien ja kuntien nimiä). Oman sana-artikkelin saavat myös monet sanonnat. SSM:ssä sen sijaan käytetään mahdollisimman helposti ymmärrettäviä selitteitä täydellisten lauseiden muodossa, ja merkityksiä kuvataan lisäksi lyhyiden esimerkkien sekä synonyymien ja antonyymien kautta. Sanakirja noudattaa sana-artikkeleiden rakenteessa mallia, joka kehitettiin Birminghamin yliopistossa yksikielistä oppijan sanakirjaa varten. Birminghamin yliopiston sanakirja (COBUILD, 1987) perustuu nykykielen korpukseen nimeltä Collins Birmingham University International Language Database, ja tästä sanakirjatyyppistä käytetään lyhennettä

COBUILD. Tämäntapaisissa sanakirjoissa fokus on selvästi semantiikassa (SSM: 5–6). SSOU:ssa (s. 5) paino on enemmän muoto-opissa. Jopa näin tarkkaan määritellyssä sanakirjatyyppissä (yksikieliset pedagogiset sanakirjat) voidaan siis päätyä hyvin erilaisiin ratkaisuihin. Taulukossa 1 olevat hakusanojen selitteet (sana-artikkelien muut osat on jätetty pois) havainnollistavat perinteistä tyyppiä edustavan sanakirjan (SSOU) ja COBUILD-tyyppisen sanakirjan (SSM) välistä eroa.

Maahanmuuttajakielten sanakirjoja tehdään tukemaan kotoutumista. KS määrittelee hakusanan *kotoutua* 'maahanmuuttajista: asettua uuteen yhteiskuntaan omat kulttuuripiirteensä säilyttäen' ja lisää esimerkin *Maahanmuuttajien kotoutumisen edistäminen*. Tässä selitteessä maahanmuuttaja itse siis kotoutuu, ja hänen kotoutumistaan voidaan edistää. Sana-artikkelista ei käy ilmi, kuka on edistäjänä, mutta *kotouttaa*-sanan selitteessä implisiittisesti mukana oleva edistäjä 'auttaa maahanmuuttajaa asettumaan uuteen yhteiskuntaan niin, että hän säilyttää omat kulttuuripiirteensä'. Tähän sana-artikkeliin KS lisää esimerkin *Maahanmuuttajien kotouttaminen Suomeen ja suomalaiseen kulttuuriin*. Kotoutuminen näyttää olevan myönteinen asia, ja sitä voidaan edistää.

SSM tarkastelee termejä nimenomaan maahanmuuttajien näkökulmasta ja nimeää myös kotoutumisen edistäjän: 'Kotouttaminen tarkoittaa sitä, että viranomaiset auttavat maahanmuuttajaa sopeutumaan Suomeen. Esimerkiksi kunta yrittää auttaa maahanmuuttajaa löytämään työtä' (SSM s.v. *kotouttaminen*). Kotouttaminen on siis viranomaistoimintaa, joka auttaa kotoutumaan. Sanat *kotoutua* ja *kotoutuminen* puolestaan selitetään seuraavasti: 'Kun maahanmuuttaja kotoutuu, hän alkaa ajatella, että Suomi on koti' (mt. s.v. *kotoutua*) ja 'Kotoutuminen tarkoittaa sitä, että maahanmuuttaja löytää uuden kodin ja elämän Suomessa' (mt. s.v. *kotoutuminen*).

KS:n määritelmät ovat viileän asiallisia, mutta hakusanojen tarkoittamat käsitteet saavat edistämiseen liittyvien esimerkkien kautta myönteisen sävyn. SSM määrittelee nämä hakusanat subjektiivisesti, maahanmuuttajan näkökulmasta. Selitteet hohkaavat suorastaan tunteen lämpöä. Työ- ja elinkeinoministeriö määrittelee kotouttamisen seuraavasti: "Viranomaisten kotouttamistoimien tavoite on, että Suomeen muuttanut henkilö tuntee yhteiskunnalliset oikeutensa ja velvollisuutensa ja tuntee olevansa suomalaisen yhteiskunnan tervetullut jäsen. Viranomaiset muun muassa tukevat maahanmuuttajaa kieliopinnoissa sekä auttavat työikäisiä löytämään työpaikan." (TEM 24.9.2012.)

TAULUKKO 1. Perinteellisen ja COBUILD-tyyppisen selitteen ero (*kieli, luottaa, politiikka, sanakirja, vieras*).

hakusana	SSOU	SSM
<i>kieli</i>	<p>1 ihmisen ja useimpien selkärankaisten suun alaosaan kiinnittynyt lihaselin, jota käytetään syötäessä ja (ihmisellä) puheen tuottamisessa.</p> <p>2 kuv.</p> <p>3 muista kieltä (1) muistuttavista rakenteista:</p> <p>4 <i>mus.</i> eräiden soittimien jännemäinen osa, jota jousella, plektralla tai suoraan sormilla koskettamalla saadaan aikaan erilaisia ääniä.</p> <p>5 vrt. 2. kuv.</p> <p>6 sanoista ja niiden käyttöä säätelevistä säännöistä koostuva järjestelmä, jolla ihmiset viestivät suullisesti tai kirjallisesti.</p> <p>7 muista kieltä (6) muistuttavista symbolijärjestelmistä:</p>	<p>1 Kieli on elin, joka on suussa. Kielen avulla ihminen voi puhua ja syödä.</p> <p>2 Kun ihmiset puhuvat ja kirjoittavat, he käyttävät kieltä. Kieli sisältää sanoja ja kieliopin. Maailmassa on yli 6 000 eri kieltä, joista suomi on yksi.</p> <p>3 Monessa soittimessa on pitkät kielet, joista tulee ääni, kun niitä soitetaan sormilla.</p>
<i>luottaa</i>	<p>pitää varmana, uskoa jonkin pitävän paikkansa, toimivan, uskoa jonkun pitävän lupauksensa, hoitavan asian tms.</p>	<p>Kun luotat johonkin, uskot että se on totta. Kun luotat ihmiseen, voit esimerkiksi kertoa hänelle asioita etkä pelkää, mitä tapahtuu.</p>
<i>politiikka</i>	<p>1 hallituksen, kansanedustajalaitoksen, puolueiden ym. toiminta tiettyjen yhteiskunnallisten tavoitteiden saavuttamiseksi.</p> <p>2 tiettyyn tavoitteeseen pyrkivä menettelytapa.</p>	<p>Politiikka tarkoittaa sitä, että hoidetaan tärkeitä yhteisiä asioita. Poliitikkaa on esimerkiksi se, kun vanha laki halutaan muuttaa.</p>
<i>sanakirja</i>	<p>kielen tai murteen sanavaraston tai tietyn sanaryhmän kirjana julkaistu luettelo, sanasto, leksikko, vokabulaari.</p>	<p>Sanakirja on kirja, jossa on sanoja aakkosjärjestyksessä. Sanakirjasta voi katsoa, miten jokin sana kirjoitetaan tai mitä se tarkoittaa. Englannin sanakirjasta voi katsoa, mikä jokin sana on englannin kielellä.</p>
<i>vieras</i>	<p>1 <i>adj.</i> tuntematon, outo.</p> <p>2 <i>adj.</i> outo, epätavallinen, asiaankuuluton, sivullinen, ulkopuolinen.</p> <p>3 <i>adj.</i> ulkomainen, ulkomaan.</p> <p>4 <i>adj.</i> ei oma, toiselle kuuluva.</p> <p>5 <i>subst.</i> kyläilijä, juhlaan osallistuja tms.</p> <p>6 <i>subst.</i>, kuv.</p>	<p>1 Kun jokin on vieras, se on tuntematon ja outo.</p> <p>2 Vieras on myös henkilö, joka on kylässä jonkun luona.</p>

Kotuksen maahanmuuttajakielten sanakirjahanke

Työskentelen maahanmuuttajakielten sanakirjatyyppejä ja sanakirjantekoa koskevien kysymysten parissa hankkeessa, jonka tarkoituksena on luoda Suomessa elävien maahanmuuttajien kotoutumista tukevia kaksikielisiä sanakirjoja. ”Maahanmuuttajakielten sanakirjahanke” käynnistyi Kotuksessa vuoden 2011 alussa. Toimin itse hankkeen vastuuhenkilönä, ja käytännön leksikografi on Ulla Horstia. Lisäksi kulloisenkin maahanmuuttajakielen sanakirjaan tarvitaan kyseisen kielen äidinkiellisiä asiantuntijoita.

Sanakirjat tulevat olemaan yksisuuntaisia siten, että suomi on lähdekielenä. Hanke perustuu Kotimaisten kielten keskuksen (Kotus) työryhmän muistiossaan esittämään ehdotukseen, että Kotus alkaisi laatia maahanmuuttajakielten sanakirjoja (OKM 2010: 30). Hanke saa tällä hetkellä opetus- ja kulttuuriministeriöltä erillisrahoitusta (OKM 2008: 3; OKM 2012: 2), mikä kertoo, että myös päättäjät pitävät tällaisia sanakirjoja tarpeellisina. Rahoitus perustuu poliittiseen päätöksentekoon, jonka tarkoitus on edistää kotoutumista ja jota edeltää eettinen pohdinta. Hankkeella on ohjausryhmä, joka koostuu Kotuksen ja Opetushallituksen edustajista ja jonka puheenjohtaja on Birgitta Romppanen.

Maahanmuuttajakielten sanakirjahankkeen päätavoite on tukea maahanmuuttajien kotoutumista. Tavoite on sama kuin alun perin Ruotsissa ja sittemmin myös Tanskassa ja Norjassa käynnistettyjen LEXIN-hankkeiden. Nimellä LEXIN (lexikon för invandrare) ylläpidetään Ruotsissa 20 maahanmuuttajakielen verkkosanakirjaa (ruotsi–muu kieli), ja useimmat sanakirjat ovat ilmestyneet painettunakin (Ruotsin LEXIN). Myös Norjassa pidetään yllä vastaavaa verkkopalvelua (Norjan LEXIN), mutta painettuna on ilmestynyt vain muutama sanakirja. Tanskassa projekti keskittyy toistaiseksi kuvasanakirjoihin, joita on noin 30 (Tanskan LEXIN). Ruotsin ja Norjan verkkopalveluiden sanakirjoissa on noin 30 000 hakusanaa, Tanskan sanakirjat ovat huomattavasti pienempiä (noin 7 000 hakusanaa).

Ruotsin LEXIN-hankkeen tuloksena on tällä hetkellä sanakirjoja, jotka selittävät ruotsin sanoja muun muassa seuraavilla kielillä: albania, amhara, arabia, azeri, kreikka, kroatia, kurdi, paštu, persia, romani, serbia, somali, suomi, turkki ja venäjä. Maahanmuuttajakielten sanakirjoissa esiintyvät kieliparit edustavat lähtökohtaisesti kaukana toisistaan olevia

kulttuureja. Tästä syystä leksikografinen työ on haastavaa, vaikka sanakirjat ovatkin selvästi yksisuuntaisia. Haastavaa ei ole kuitenkaan pelkäämään kysymys hyvien vastineiden löytämisestä – tai vastineen puuttuessa hyvän selittävän vastineen antamisesta –, vaan myös eettiset kysymykset nousevat voimakkaasti esille.

Maahanmuuttajakielten sanakirjojen tehtävänä on ministeriön toimeksiannon (OKM 2008: 2) mukaan myös palvella maahanmuuttajakielten oppimista ja käyttöä Suomessa. Vaikka näissä sanakirjoissa ei ole tarkoitus selittää maahanmuuttajaryhmän kieli- ja kulttuurioloja heidän lähtömaissaan, kotoutuminen ymmärretään kuitenkin prosessina, joka ei tähtää maahanmuuttajan kulttuurisen ja kielellisen taustan hävittämiseen.

Maahanmuuttajakielten sanakirjahankkeen käynnistäminen ja sen kytkeminen kotoutumispolitiikkaan ei ole itsestään selvää. Sanakirjojen laadinta perustuu arvovalintoihin eikä niistä välttämättä vallitse yksimielisyyttä yhteiskunnassa. Sanakirjatyöhön ryhtyvän leksikografian on itsekin pohdittava asemaansa maahanmuuttajakielen sanakirjan tekijänä, koska hänen työllään on vaikutusta kotoutumisen luonteeseen ja onnistumiseen.

Kohdekielen valinta

Opetus- ja kulttuuriministeriön tukipäätökseen ei sisälly tieto, mihin kieliin suomalaiset maahanmuuttajasanakirjat tulisi laatia ja mitkä kielet olisi priorisoitava. Olemme lähestyneet asiaa pragmaattisesti, ja tärkeimmiksi kriteereiksi ovat nousseet seuraavat kaksi:

- Onko suomen sanoja selittävä sanakirja tiettyyn kieleen olemassa vai ei?
- Mikä on kieliyhteisön koko?

Ensimmäinen kriteeri on leksikografinen. Tutkitaan, mikä on tietyn kieliparin sanakirjatilanne. Sanakirjoja voi olla monta, niitä voi olla vanhoja tai uusia, hyviä tai puutteellisia, ja ne voivat olla erikokoisia ja erityyppisiä. Jälkimmäinen kriteeri on yhteiskunnallis-poliittinen. Sanakirjan tarve on sitä polttavampi, mitä suuremmasta kieliyhteisöstä on kyse. Sanakirjan laatimisen edellytys on tietysti lisäksi se, että hankkeeseen saadaan päteviä tekijöitä.

Tilastokeskus antaa kieliyhteisöjen koosta vuodelta 2012 seuraavat tiedot (Tilastokeskus 2012):

TAULUKKO 2. Suurimmat vähemmistökielet Suomessa.

venäjä	62 554
viro	38 364
somali	14 769
arabia	12 042
kurdi	9 280
kiina	8 820
albania	7 760
thai	6 926
vietnam	6 549
persia	6 422
turkki	6 097
saksa	5 792
espanja	5 470

Selvitystemme perusteella sanakirjatilanteen voi katsoa somalin, kurdin, albanian, thain, vietnamin, persian ja turkin sekä ehkä myös arabian kielissä muita selvästi heikommaksi.

Maahanmuuttajien kieliyhteisöjen koon sekä kielten leksikografisen tilanteen perusteella ensimmäiseksi kohdekieleksi päädyttiin valitsemaan somalin kieli. Somalin kielen asiantuntijoiksi hankkeeseen on palkattu Liban Ali Hersi ja Abdi Awil Musse. Sanakirjat laaditaan kotoutumisen tueksi, mutta myös kyseisten kielten kääntäjien ja tulkkien ammatillisia tarpeita silmällä pitäen. Keskityn tässä kirjoituksessa sanakirjojen kotoutumista tukevaan tehtävään.

Somalin kielen ja somalinkielisen yhteisön näkökulmasta sanakirja on kulttuuriteko. Somalin kieleen tai kielestä on olemassa hyvin vähän sanakirjoja, ja yleensä parina on englanti tai italia (esim. Zorc–Osman 1993; Hashi 2004; Mire 2008). Myös yksikielisiä somalin sanakirjoja on vähän (esim. Keenadiid 1976), tosin aivan äskettäin ilmestyi uusi laaja sanakirja (Puglielli–Mansuur 2012). Sanakirjan ilmestyminen Suomessa tekee kielen näkyväksi.

Somalin oikeinkirjoitus perustuu normiin, joka on peräisin 1980-luvulta. Kaikki sen jälkeen kieleen tulleet sanat kirjoitetaan tätä normia parhaan kyvyn mukaan soveltaen. Tämä koskee tietysti myös sanakirjojen laatijoita. Suomessa laadittavan sanakirjan kirjoitusasusta tulee siis hyvin normatiivinen.

Tätä kirjoittaessani työ Kotimaisten kielten keskuksessa on vielä alkutekijöissään, parhaillaan viimeistellään suomenkielistä hakusanastoa ja aloitetaan suomi–somali-sanakirjan laatimista. Tästä syystä seuraavien lukujen pohdinnat on ymmärrettävä varautumisena, yrityksenä ottaa etukäteen huomioon mahdollisesti esiin nousevat eettiset kysymykset. Niinpä pohdinnat eivät kohdistu mihinkään tiettyyn kieleen tai kulttuuriin.

Hakusanojen määrä ja valinta

Sanaston parissa työskenteleville tutkijoille sanat ovat tasa-arvoisia. Jokainen sana on leksikaalinen yksikkö, jolla on äänne- ja kirjoitusasu, mahdollisesti sanaluokka, taivutus, merkitys ja konteksti. Näitä asioita tutkimalla leksikologi voi tehdä päätelmiä siitä, mihin tyyliin sana kuuluu, missä tilanteissa se sopii käytettäväksi ja missä yhteydessä sen käyttö herättäisi huomiota.

Sanakirjan tekijä lähestyy sanastoa hieman eri tavalla. Olipa kysymys kuinka laajasta ja minkä tyyppisestä sanakirjasta tahansa, leksikografi joutuu pohtimaan, mitkä sanat otetaan mukaan ja mitkä jätetään pois, mitkä merkitykset huomioidaan ja mitkä jätetään kertomatta. Ulkoisista syistä (kustantajan tahto, resurssien määrä tms.) pohdinnan lähtökohta on useimmiten hyvin yksinkertainen: hakusanojen määrä. Konkreettinen valintaprosessi ei ole kuitenkaan ongelmaton. Kun pyritään tiettyyn määrään hakusanoja, on valinta pois jätettävien ja mukaan otettavien sanojen välillä vääjäämättä mielivaltainen.

On sovittu, että Maahanmuuttajakielten sanakirjahankkeessa tuotettavissa sanakirjoissa tulee olemaan noin 30 000 suomenkielistä hakusanaa. Hakusanoille on tarkoitus löytää vastineet ja/tai selittävät vastineet kyseisellä maahanmuuttajakielellä. Myös käyttöesimerkkejä ja niiden käännökset annetaan. Hakusanojen määrä noudattaa Ruotsin ja Norjan LEXIN-sanakirjojen kokoluokkaa. Hankkeessamme käytettävissä olevien resurssien ja erityisesti Norjassa saatujen kokemusten perusteella arvioimme, että kukin sanakirja voisi valmistua noin kolmessa vuodessa. Tällä mitoituksella sanakirjat siis saataisiin valmiiksi kohtuullisen nopeasti. Sanakirjat tullaan julkaisemaan sähköisesti verkossa, jolloin päivityksiä ja lisäyksiä on mahdollista tehdä painotuotetta huomattavasti helpommin.

Selitettäväksi on siis valikoitu suomenkielinen sanasto. Miten valinta tehdään? Jos rahaa ja aikaa olisi käytettävissä rajattomasti, luotaisiin kenties aineistotietokanta, johon kerättäisiin sanoja ja niiden käyttöyhteyksiä. Keruutyön lähtökohdaksi voitaisiin ottaa ne tilanteet, joissa maahanmuuttaja tyypillisesti käyttää suomen kieltä: koulutus, perhe-elämä, työelämä, asiointi viranomaisten kanssa, lasten, vanhusten ja sairaiden hoito, vapaa-ajan vietto, kulttuuri ja niin edelleen. Luettelo on hyvin kattava. Keruu olisi paitsi aikaa vievää myös turhaa, sillä aineistoa voitaisiin olettaa karttuvan niin paljon, että ainakin hakusanojen valinnan pohjana voidaan yhtä hyvin käyttää ajan tasalla olevaa nykykielen sanakirjaa kuten noin 100 000 hakusanan KS:aa.

Keruu voitaisiin myös tarkentaa esimerkiksi nimenomaan maahanmuuttajalle tyypillisiksi oletettuihin tilanteisiin, onhan sanakirjan fokuksessa nimenomaan kotoutuminen Suomeen. Tämä valinta olisi kuitenkin hyvin kyseenalainen, koska silloin maahanmuuttajaa tuettaisiin nimenomaan maahanmuuttajana ja vain maahanmuuttajana. Kotouttaminen tähtää kuitenkin pidemmälle. Tavoitteena on, että maahanmuuttaja sopeuttaa identiteettinsä uuteen tilanteeseen. Tarkennettu valinta ei siis tuekaan kotoutumista. Se ei ole sanakirjan tehtävän mukainen eikä näin ollen myöskään eettisesti hyväksyttävä.

KS kattavimpana nykysuomen yleiskielen sanaston kuvauksena on hyvä lähtökohta. Sanakirja on Suomen yhteiskunnan leksikografinen peili, josta heijastuu juuri se yhteiskunta, johon maahanmuuttajien on tarkoitus kotoutua. Niinpä KS:n valinta lähtökohdaksi on myös eettisesti perusteltua.

Lähtökohdan valinnan jälkeen jää vielä mietittäväksi, millä kriteereillä KS:sta valitaan 30 000 hakusanaa maahanmuuttajasanakirjaan. Valittu on niin sanottu kielen perussanasto ja lisäksi jonkin verran yhdyssanoja ja erilaisia johdoksia. Luotamme valinnassa omaan leksikografiseen asiantuntemukseemme eikä käytössä ole sanojen tilastollisia yleisyystietoja.

Tämän kokemukseen perustuvan menetelmän avulla esimerkiksi KS:n runsaasta 50 yhdyssanasta, joiden määriteosa on *paperi-*, on otettu mukaan *paperikauppa*, *paperikori*, *paperikuva*, *paperilappu*, *paperinenäliina*, *paperipyyhe*, *paperisota*, *paperitehdas* ja *paperiteollisuus*. Johdosten osalta huolehdimme siitä, että keskeiset sanavartalot ovat edustettuina vähintään yhdellä sanalla, jotta käyttäjällä on mahdollisuus päättellä oman kie-

liosamisensa perusteella muiden samasta vartalosta johdettujen sanojen merkitykset.

KS:sta puuttuvia varsinaisia puhekielen sanoja emme ota hakusanoiksi. Meillä ei ole resursseja niiden keräämiseen, ja tällaisten sanojen mahdollisesti lyhyt elinkaari on leksikografian kannalta muutenkin haasteellinen. Sen sijaan on tärkeämpää ottaa sanakirjaan koulutukseen, työelämään, sosiaali- ja terveystalouteen ja muuhun vastaavaan liittyviä hakusanoja sekä ensyklopedista tietoa välittäviä hakusanoja. Tällaisia ovat esimerkiksi *Kela* ja *VR*. Tämän valinnan taustalla on leksikografien käsitys siitä, mihin tähdätään kotouttamisella ja mikä on se, mihin maahanmuuttajaa autetaan leksikografisin keinoin kotoutumaan. Näemme sanakirjojemme käyttäjän henkilönä, joka haluaa koulutusta, pyrkii työelämään tai on työssä ja haluaa asioida viranomaisten kanssa.

Kuvattavana suomalainen yhteiskunta

Ruotsin LEXIN-hankkeessa (niin painosta ilmestyneet kuin verkossa julkaistut) sanakirjat perustuvat yhteen ja samaan tietokantaan. Hankkeen johtajan, Christian Mattssonin mukaan sanakirjojen hakusanavalikoima on tarkoitus pitää yhtenäisenä (suullisesti 23.11.2010; vrt. myös Ruotsin LEXIN: Bakgrund ja LEXIN romskt 2007: IV). Ratkaisu helpottaa monien sanakirjojen samanaikaista ylläpitoa. Myös Kotuksen hankkeessa pyrimme pitämään suomenkielisen hakusana- ja esimerkkipohjan samana kaikissa kielipareissa. Kun selitettävänä on suomen kielen sanasto ja suomalainen kulttuuri, hakusanojen valinnassa on mahdotonta ottaa huomioon eri maahanmuuttajaryhmien kielelliset ja kulttuuriset lähestymistavat suomen kieleen. Voidaan kuitenkin olettaa, että Suomen yhteiskunnan ilmiöt ovat eri ryhmille eri tavalla tuttuja, läheisempiä tai kaukaisia, oudompia. Tämän mukaan erilaisten ilmiöiden selittämistarve saattaa vaihdella eri sanakirjoissa. Sanakirjan tekijä joutuu miettimään ainakin seuraavia:

- mitkä ovat maassamme ne ilmiöt, joita muualta tulijan on mahdotonta hyväksyä?
- mitkä ovat maassamme ne ilmiöt, joita muualta tulijan on vaikeaa ymmärtää?
- mitkä ovat maassamme ne ilmiöt, jotka vastaavat vain näennäisesti tulijan omaa kulttuuria (ns. väärät ystävät)?

- minkälainen koulutustaso jollakin tulijaryhmällä pääsääntöisesti on?
- mikä on lukutaidottomuuden määrä?

Esimerkkeinä ensimmäisistä kohdista voi mainita mahdollisesti eriävät käsitykset perheestä, suvusta jne., mahdollisesti eriävän suhtautumisen lapsiin ja vanhuksiin, naisten pukeutumisen, sukupuolten välisen tasa-arvon, suhtautumisen alastomuuteen (saunominen, pukuhuoneet), sukupuolielämän jne. Myös alkoholin ja tupakan käyttö, ruoka-aineet ja ruoan valmistus voivat tulla esille samoin kuin diskurssi-strategiat (meneminen suoraan asiaan), ruumiilliset kosketukset tervehdittäessä, keskustellessa ja erottaessa jne. Luettelemani aihepiirit eivät muodosta pelkästään mahdollisten kulttuurierojen kirjoa, vaan tähän saattaa liittyä voimakkaitakin tabuja, olivatpa ne tiedostettuja tai alitajuisia, suomalaisten tai maahanmuuttajien.

Kahta jälkimmäistä on leksikografin hankalaa ottaa huomioon, koska sanakirjan käyttö edellyttää esimerkiksi lukutaitoa. Maahanmuuttajakielten sanakirjahankkeessa syntyvät sanakirjat auttanevat osaltaan edistämään kouluttautumista.

Oman kokonaisuutensa muodostavat uskontoon ja moraaliiin liittyvät kysymykset, esimerkiksi:

- mitkä ovat suurimmat erot uskonnossa ja moraalissa, onko tabuja?
- mistä saa puhua, mistä ei, mitä saa esittää kuvana, mitä ei?

Kysymyksiä on varmasti vielä monia muitakin, mutta jo yrittäessään vastata edes näihin kysymyksiin huomaa, että kotoutumista tukevan sanakirjan laatiminen ei ole mitenkään suoraviivaista. Yllä luetellut kysymykset tähtäävät tulijan taustan ymmärtämiseen. Kuitenkin heti ensimmäinen kohta eli kysymys ilmiöistä, joita tulija ei voi hyväksyä, sisältää jo pääongelman: kotouttaminen tähtää siihen, että tulija voi elää ja toimia tasavertaisena Suomessa. Niinpä kotouttamista tukevista sanakirjoista ei voida jättää pois Suomen oloissa tärkeitä hakusanoja, vaikka niiden tarkoitteiden olemassaolo olisikin hankalaa tulijalle.

Nämä kysymykset johtavat pohtimaan yleisemmin maahanmuuttajien edellytyksiä sanakirjan käytölle. Puhtaasti deskriptioon perustuva realistisen yhteiskuntakuvauksen välittävä sanakirja saattaa sisältönsä vuoksi joutua maahanmuuttajien keskuudessa tai tietyssä maahanmuuttajaryhmässä epäsuosioon. Mahdotonta ei olisi sekään, että vaikutusvaltaiset henkilöt yrittäisivät kieltää sanakirjan käytön. (Kenties vain muutaman

sanan takia) käyttämättä jäävä sanakirja ei todellakaan täytä tehtävänsä eikä tue maahanmuuttajien kotoutumista.

Tähän reseptio-ongelmaan voidaan vastata verkostoitumalla heti hankkeen alussa. Tekeillä olevan suomi–somali-sanakirjan somalin kielien toimittajat käyttävät omia verkostojaan paitsi sanavastineiden etsimiseen myös sanakirjahankkeen tunnetuksi tekemiseen somalinkielisessä yhteisössä. Toistaiseksi tietoomme ei ole tullut hakusanojen valintaan liittyvää kritiikkiä.

Maahanmuuttajakielten sanakirjojen toimittamisperiaatteiden pohdinta johtaa kotouttamisen periaatteiden ja tavoitteiden pohdiskeluun. Jos kotoutumisen tavoitteena on, että maahanmuuttaja voi elää ja toimia tasavertaisena Suomessa, kantaväestölle vieraat kiellot tai tabut eivät saa olla esteenä. Toisaalta kotoutujan ei odoteta unohtavan kokonaan omaa taustaansa ja historiaansa. Nämä ovat kuitenkin kysymyksiä, joista poliitikot, viranomaiset jne. eivät suinkaan ole yksimielisiä. Toisaalta uusi kulttuuri ja todellisuus ja toisaalta oma tausta ja historia eivät maahanmuuttajilla välttämättä edes yhdessä sukupolvessa sulaudu ehjäksi identiteetiksi.

Kotouttamista tukevissa sanakirjoissa on lähdeittävä liikkeelle Suomen yhteiskunnallisista oloista. Hakusanojen valinta ja vastineiden ja selitteiden löytäminen voi tällöin vaatia erityistä taitoa. Lisäksi on pohdittava, luonnehditaanko hakusanaa tai yksittäistä käyttöä tai merkitystä jollakin tavalla, esimerkiksi ylä- tai alatyylisiin kuuluvaksi, ja jos luonnehditaan, niin mitkä ovat kestävät kriteerit siihen. Maahanmuuttajakielten sanakirjaprojektissa tukeudutaan tältä osin KS:n luonnehdintoihin. Samoin esimerkkien valinnassa on paljon pohdittavaa. Luodaanko esimerkit itse toimittajan kielikompetenssiin nojaten vai poimitaanko niitä jostakin faktista kielenkäyttöä heijastavasta aineistosta? Näitä yleisperiaatteita joudutaan todennäköisesti puntaroimaan joidenkin sanojen kohdalla hyvin tarkkaan.

Maahanmuuttajasanakirjoissa ongelmallisia eivät välttämättä ole samat sanat, jotka missä tahansa yleiskielen sanakirjassa saattavat aiheuttaa pahennusta (esim. *paska*, *vittu* yms.). Sen sijaan ongelmallisia ovat esimerkiksi asenteellisuutta maahanmuuttoa ja maahanmuuttajia kohtaan ilmaisevat sanat, kuten *ankkurilapsi* ’alaikäinen turvapaikanhakija, jonka avulla myös muu perhe toivoo pääsevänsä maahan’, *elintasonpakolainen* ’siirtolaisista, joiden muuton synnä on paremman elintason tavoittelu’, *ketjumaahanmuutto* ’perheenyhdistämistä hyödyntävä maa-

hanmuutto’, *turvapaikkashoppailu* ’turvapaikka-anomuksen jättäminen moneen (EU-)maahan’ ja niin edelleen. Asiaa monimutkaistaa vielä se, että mainitunlaiset termit eivät esiinny vain arkisessa kielenkäytössä, vaan niitä saatetaan tavata viranomaisten ja päättäjien kielenkäytössä. Poliittisessakin diskurssissa tämäntyyppisiä sanoja esiintyy, ja sieltä ne leviävät yleiseen kielenkäyttöön.

Näiden sanojen jättämistä pois sanakirjasta voitaisiin perustella sillä, että niillä on vähättelevä tai asenteellinen sävy eikä niiden käyttöä tai leviämistä siksi haluta sanakirjan avulla edistää. Samasta syystä olisi tosin jätettävä pois moni muukin sana. Maahanmuuttaja saisi näin käyttöönsä siistin ja moraalisesti puhtaan sanakirjan. Mutta kuvastaako tällainen teos Suomen yhteiskuntaa ja tukeeko se kotoutumista?

Tässä idealistisessa vaihtoehdossa leksikografi ratkaisee asian sanakirjan käyttäjän puolesta. Käyttäjä saa ruusuisen kuvan suomalaisesta yhteiskunnasta eikä joudu kohtaamaan itseään vähättelevää ja asenteellista sanastoa. Kuva on epärealistinen eikä sanojen kieltäminen poista sanojen tarkoitteita tai sanojen esiintymistä todellisessa kielenkäytössä. Yhteiskunnassa tosin on hyvin laaja pyrkimys asenteellisen ja alentavan kielenkäytön karsimiseksi, mutta näin välitöntä yhteiskunnallista vaikutusta ei leksikografialla sentään ole. Sellaiset sanat kuin *neekeri*, *rampa*, *mykkä* ovat nykyään ihmistä tarkoittavina jo enemmän tai vähemmän paheksuttavia. Pitäisikö nämäkin sanat jättää pois? Kun neutraalin ja asenteellisen sanaston välinen raja on veteen piirretty viiva, ratkaisu varmasti heijastaa sanakirjantekijän omia asenteita. Ratkaisu on eettisesti holhoava. (Vrt. Grönros tk. s. 48.)

Mainittujen sanojen mukaanottoa voidaan taas perustella sillä, että sanakirjojen tehtävä on edistää maahanmuuttajien kotoutumista ja tukea heitä jokapäiväisissä kielenkäytön tilanteissa. Kuinka maahanmuuttaja voi arvioida tilanteen oikein, kun häntä kutsutaan *mamuksi*, hänen lastaan *ankkurilapseksi* tai hänen pyrkimystään yhdistää perheensä *ketjumaa-hanmuutoksi*, jos hän ei tunne sanojen merkitystä ja varsinkaan niiden takana olevaa asenteellisuutta? (Tosin *mamu*-nimitys voi olla neutraalikin. Ks. myös Tyysteri tk. s. 138.) Onko eettisesti oikein jättää sanakirjan käyttäjä pulaan ja omien kokemustensa varaan? Maahanmuuttaja törmää tällaisiin sanoihin lähes varmasti, ja olisi leksikografisesti väärin jättää nämä sanat pois sanakirjasta. Sanat ovat (ainakin jonkin aikaa) osa Suomen kielimaisemaa. Ne koskevat ilmiöitä, joiden kanssa maahanmuut-

taja tahtomattaankin joutuu tekemisiin, ja näin ollen ne on syytä selittää sanakirjassa.

Ratkaisu perustuu deskriptioon ja on sanaston tutkimuksen kannalta eettisesti ja tieteellisestikin hyvin perusteltu. Vastuu kielenkäytöstä ei ole leksikografilla, vaan se on aina kielenkäyttäjillä. Sanakirjassa voidaan maahanmuuttoon ja maahanmuuttajiin asenteellisesti suhtautuvat sanat merkitä erikseen ja varustaa ne kohdekieliselällä selitteellä.¹ Tämä realismiin johdattama vaihtoehto välittää ohdakkeisen kuvan ja pyrkii olemassa olevan todellisuuden kuvaamiseen. Sanakirjan käyttäjä saa tietää, mitä edellä mainitut sanat tarkoittavat ja pystyy arvioimaan näiden sanojen käyttäjän asennetta. Hän voi opetella näiden sanojen merkityksen, mutta niiden käyttö jää sanakirjan käyttäjän omaan harkintaan. (Vrt. Tyysteri tk. s. 148–149.)

Sievistelevän vaihtoehdon mahdottomuuden huomaa myös siitä, että leksikografi joutuisi tekemään aivan epäloogisia hakusanavalintoja. Jos esimerkiksi katsottaisiin, että *rampa* ja *mykkä* ovat vältettäviä sanoja eikä niitä sovi ottaa mukaan sanakirjaan, silloin myös sanat *rampauttaa* tai *mykistää* jouduttaisiin jättämään pois, vaikka näihin sanoihin ei liitykään perussanojensa väheksyvää tyyliä. Leksikografisesti olisi nimittäin outoa ottaa mukaan selvät johdokset ja jättää perussanat pois.

Kotuksen maahanmuuttajakielten sanakirjahankkeessa toimitaan niin, että kiistanalaisia mutta relevantteja sanoja ei jätetä pois, vaikka ne voitaisiin kokea sanakirjan käyttäjää, hänen taustaansa, kulttuuriaan ja uskontoaan vähätteleviksi ja vaikka sanat liittyisivät suomalaisen yhteiskunnan ilmiöihin, jotka ovat maahanmuuttajalle esimerkiksi kielteisiä,

1 Vuonna 2007 ilmestynyt saksan kielen Universalwörterbuch (DDUW) on esimerkiksi siitä, että leksikografit eivät jätä harkintaa kokonaan käyttäjälle. Sanakirjan toimitus murtaa paikoin sanakirjan rakenteen antaakseen käyttäjälle käytännön neuvoja tietyistä hakusanasta. Enimmäkseen nämä toimituksen antamat tiedot koskevat hakusanoja, joita ei voinut jättää pois sanakirjasta, koska näitä sanoja on todettu käytettävän, mutta joita toimitus ei suosittele käytettäväksi; esim. *Fräulein* 'neiti [diminutiivipäätteen takia halventavaksi koettu sana]; *Neger* 'neekeri', *türken* 'väärentää [muodostettu turkkilaista tarkoittavasta sanasta]'. Myös hakusana *Erziehungsurlaub* "kasvatysloma", 'hoitovapaa' saa huomautuksen: "Vuodesta 2001 lainsäätäjän käyttämä virallinen nimitys on *Elternzeit* 'vanhempain aika', jotta kasvatustyötä ei väheksyttäisi lomana". Dudenin toimitus on valinnut edellä kuvaamani realistisen vaihtoehdon ja katsonut lisäksi tarpeelliseksi ohjata sanakirjan käyttäjiä tiettyjen sanojen käytössä. Ratkaisua tarkastelee kriittisesti Ickler 2006. Oppijan sanakirjoista poliittisen korrektiuden kannalta vrt. myös Busse 2000.

kiellettyjä tai vaikeita. Pidän tätä rehellisenä ja eettisenä ratkaisuna. Tosin nämä pohdinnat liikkuvat yleisellä tasolla. Konkreettinen työ, jossa suomalainen leksikografi ja – Kotimaisten kielten keskuksen tapauksessa ensimmäiseksi – somalin kielen asiantuntijat tekevät yhdessä työtä, näyttää, miten prosessi etenee.

Esimerkkinä konkreettisen sanakirjatyön hankaluudesta voidaan mainita suomalais-venäläisenä yhteistyönä laadittu Venäjä–suomi-suursanakirja (Kuusinen ym. 1997). Venäläinen osapuoli kieltäytyi ottamasta sanakirjaan muutamia edellä mainitun laisia ”rumia sanoja”, jotka suomalainen osapuoli kuitenkin halusi sisällyttää sanakirjaan. Sanakirjan Suomessa ilmestyneessä laitoksessa ns. rumat sanat ovat mukana, myöhemmin Venäjällä ilmestyneessä laitoksessa niitä ei ole (Kuusinen ym. 1999). Tämäntapainen kompromissi ei maahanmuuttajakielten sanakirjojen kohdalla ole mahdollinen.

En voi tässä käsitellä pragmatiikkaan ja kulttuuriin liittyviä asioita syvällisemmin. Ne vaativat omaa käsittelyä, eikä minulla ole vaadittavaa asiantuntemusta. Otan kuitenkin esiin vielä joitakin muita kieleen liittyviä kysymyksiä.

Kieliopin etiikka

Edellä on käsitelty hakusanojen valintaan ja selittämiseen liittyviä eettisiä kysymyksiä. Nämä ovat tärkeitä, mutta leksikografisen työn edetessä tulevat eteen vielä monet muut kysymykset. Maahanmuuttajasanakirjojen kannalta tärkeitä kielen rakenteeseen liittyviä pohdittavia kysymyksiä ovat muun muassa seuraavat:

- mitkä ovat suomen ja maahanmuuttajakielen väliset leksikografisesti relevantit rakenteelliset erot?
- onko maahanmuuttajakielessä kieliopillista sukua tai vastaavaa kategoriaa?
- onko maahanmuuttajakielessä muita kieliopillisia tai pragmaattisia kategorioita, joita suomessa ei ole (esim. prepositiot, duaali, aikamuotojärjestelmä ym. tai erilainen kielenkäyttö eri sukupuolilla, ikä- ja statusryhmillä)?
- minkälainen on tulijan kulttuuripiirin leksikografinen traditio (sanakirjan käyttötaito)?

Kysymykset ovat monikielisessä leksikografiassa muutoinkin hanka-

lia. Enimmäkseen on kyse kielten välisistä rakenne-eroista, kategorioiden vastaamattomuudesta. Tässä yhteydessä kiinnitän huomiota maahanmuuttajakielten sanakirjoissa tehtävien ratkaisujen eettiseen puoleen, joka nousee toki esiin muissakin sanakirjatyypeissä. Keskityn tässä kieliopilliseen sukuun.

Suomen kielessä ei ole kieliopillista sukua kieliopillisena kategoriana. Luonnollista sukua voidaan toki ilmaista esimerkiksi johto-opin ja yhdysanojen avulla: *laulajatar*, *miesopettaja*. Monessa kielessä kieliopillisen suvun kategoria kuitenkin on, ja myös sanoilla, joiden tarkoitteilla ei ole luonnollista sukua, on kieliopillinen suku. Näin on laita myös somalin kielessä, jossa on kaksi kieliopillista sukua: maskuliineja ovat esimerkiksi *dhagax* 'kivi', *biyo* 'vesi' ja *dayax* 'kuu', feminiinejä *qorrax* 'aurinko', *dabayl* 'tuuli' ja *daruur* 'pilvi'.

Esimerkiksi ammattinimikkeissä ja yleensä tekijännimissä sekä mahdollisesti myös adjektiiveissa (ja jopa verbeissä) joutuu pohtimaan, millä periaatteella vastine maahanmuuttajakielellä annetaan. Esimerkiksi hakusanojen *pappi*, *kokki*, *sotilas* ja *poliisi* vastineiden pohdinnassa konkretisoituvat edellä olevat teoreettiset näkökohdat. Entä *räättäli* ja *ompelija*? Minkälainen kuva annetaan suomalaisesta yhteiskunnasta? Kuinka hankalaa on antaa toisaalta sukupuolineutraali vastine sekä maskuliini että feminiini (mahdollisesti jopa neutrikin) vastine tai toisaalta sukupuolineutraali selite toisella kielellä?

TAULUKKO 3. Luonnollisen suvun näkyminen suomen ja somalin kielen ammattinimikkeissä.

suomi	pappi	kokki	sotilas	poliisi	räättäli	ompelija
somali	baadari baadariyad	cuntokariye cuntokariso	askari askariyad	boolis boolisad	dawaarle	dawaarle

Suomeksi nainen voi sen enempiä oikean muodon valitsemista miettimättä kertoa olevansa pappi, kokki, sotilas, poliisi, räättäli tai ompelija. Somalin kielessä tilanne on moniulotteisempi, ja joissain tapauksissa naisen on luontevampaa käyttää geneeristä maskuliinia (räättäli, ompelija), joissakin taas feminiiniä.

Kieliopillisen suvun tuoma erikoisongelma ovat persoonapronominut. Esimerkiksi arabian kielessä 2. ja 3. persoonan pronomineilla on

kaksi eri muotoa, maskuliini ja feminiini (Wikipedia FI s.v. *Arabian kieli: Persoonapronominit*). Esimerkkilauseiden käännöksissä täytyy valita jompikumpi. Minkälaiset käännökset saavat vaikkapa sellaiset esimerkit kuin *sinä ajat autoa*, kun esimerkiksi Saudi-Arabiassa naiset eivät saa ajaa (Wikipedia FI s.v. *Saudi-Arabia: Naisen asema*), *hän siunaa uskovaiset, hän laittaa ruokaa* jne.? Ratkaistavana on taas kysymys, minkälainen kuva annetaan suomalaisesta elämänmenosta. Maahanmuuttajakielen rakenteesta riippuen voidaan harkita myös sekä maskuliinin että feminiinin käyttöä (esim.: ”x/y ajaa autoa”), mutta selkeys kärsii silloin helposti. Suomenkielinen esimerkki *hän näki ystävänsä* saisi tällöin esimerkiksi tälläisen somalinnoksen *saaxiibkii/saaxiibaddii/saaxiibkeed ayuu/ayey arkay/aragtay*. Käännökset ovat purettuina taulukossa 4.

TAULUKKO 4. Lauseen *hän näki ystävänsä* somalinkieliset käännösmahdollisuudet.

suomi ♂ tai ♀	somali
hän ♂ näki ystävänsä ♂	<i>saaxiibkii ayuu arkay</i>
hän ♂ näki ystävänsä ♀	<i>saaxiibaddii ayuu arkay</i>
hän ♀ näki ystävänsä ♀	<i>saaxibaddeed ayey aragtay</i>
hän ♀ näki ystävänsä ♂	<i>saaxiibkeed ayey aragtay</i>

Eräänä ratkaisuna voidaan pitää sellaisten esimerkkien välttämistä, jotka esimerkiksi kieliopillisen suvun takia ovat hankalia. Eettisyyden asteikolla tällainen ratkaisu ei voi sijoittua kovinkaan korkealle. Persoonapronomineja voisi kiertää käyttämällä niitä mahdollisimman vähän esimerkeissä. Niiden sijaan voitaisiin käyttää esimerkiksi nimiä tai infiniittisiä lausekkeita (ks. myös Pilke & Puskala tk. s. 166). Nimien käyttö vaikuttaa helposti lapselliselta, usein toistuvat nimet turruttavat, suuri nimien kirjo taas hämää ja vie huomion esimerkissä väärään suuntaan. Infiniittisiin lausekkeisiin liittyy se suomen kielen erikoispiirre, että niissä objektin sija ei näy kunnolla. Tämän takia juuri infiniittisiä lausekkeita pitää esimerkeissä välttää. Toisaalta persoonapronominit ovat kielen käytetyimpiä sanoja, joten niiden pitää näkyä esimerkeissä. Ratkaisuna

käytämme esimerkkien käänöksissä vuoroin maskuliineja ja feminiinejä persoonapronomineja.

Lopuksi

Pyrin artikkelissani tuomaan esille joitakin käytännöllisiä ja eettisiä näkökulmia, jotka liittyvät leksikografiseen työhön maahanmuuttajakielten parissa. Tässä yhteydessä olen voinut käsitellä vain pientä osaa työssä eteen tulevista ongelmista ja ratkaisuista, koska työmme maahanmuuttajakielten sanakirjojen parissa on vasta alussa. Olemme alkaneet miettiä näihin sanakirjahankkeisiin liittyviä eettisiä seikkoja jo ennen konkreettisen työn aloittamista, ja tilaisuus osallistua tähän kokoomateokseen on siivittänyt pohdintaa ilahduttavasti.

SIRKKA SAARINEN

Suomalais-ugrilaisten vähemmistökielten sanakirjojen laatimisesta

Tarkastelen artikkelissani yhtä vanhaan sanastomateriaaliin perustuvaa sanakirjaa (Tscheremissisches Wörterbuch, 2008), kahta kaksikielistä sanakirjaa, joissa toisena kielenä on udmurtti (Udmurttilais-suomalainen sanakirja, 2008, ja Suomalais-udmurttilainen sanakirja, 2013) sekä tutkijoiden käyttöön tarkoitettua käänteissanakirjaa (Reverse Dictionary of Mari (Cheremis), 2002), joita olen ollut laatimassa. Kuvaan artikkelissani Venäjän suomalais-ugrilaisia kieliä koskevien sanakirjojen laatimiseen liittyviä ongelmia ja mahdollisia ratkaisuja niihin; osa kysymyksistä painottuu tieteellisiin tai leksikografisiin valintoihin, osa eettisiin. Esittelemäni sanakirjat ovat keskenään erityyppisiä. Niitä yhdistää suomalais-ugrilaisuus: kuvattavat kielet ovat suomen sukukieliä. Lisäksi ne ovat omalla alueellaan vähemmistökieliä, joiden yhteiskunnallinen asema ei ole vahva. Jaottelen tekstini sanakirjojen mukaan. Etenkin Tscheremissisches Wörterbuchin taustaa valottaakseni kerron tässä johdanto-osassa lyhyesti suomalais-ugrilaisen kielentutkimuksen eli fennougristiikan alkuvaiheista Suomessa ja niihin sisältyneistä suurista keruu- ja julkaisuhankkeista.

Suomalais-ugrilaisten kielten sukulaisuus osoitettiin 1700-luvun aikana, ja etenkin Henrik Gabriel Porthanin ansiosta tieto tästä levisi suomalaistenkin tutkijoiden pariin 1800-luvun alkuun mennessä. Suomessa suomalais-ugrilaisten kielten tutkimus alkoi alan suurten pioneerien, Anders Johan Sjögrenin (1794–1855) ja Matthias Alexander Castrénin (1813–1852), sukukieltä puhuvien kansojen pariin tekemistä tutkimusmatkoista, joilla he keräsivät paljon teksti- ja sanastoainesta. Castrén totesi ensimmäisenä suomalais-ugrilaisten ja samojedikielten sukulaisuuden. (Korhonen 1986: 28–33, 41–61.)

Pietarin keisarillinen tiedeakatemia kustansi sekä Sjögrenin että Castrénin matkat Venäjän suomalais-ugrilaisen kansojen pariin. Vuonna 1883 perustettu Suomalais-Ugrilainen Seura asetti päätehtäväkseen M. A. Castrénin työn jatkamisen: oli kerättävä aineistoa suomen sukukielistä ja niitä puhuvista kansoista. Tarkoituksena oli edistää tiedettä mutta samalla selvittää suomen kielen ja suomalaisten alkuperää. Kuten jo Castrén aikanaan myös Seuran puuhamiehet ymmärsivät Suomen kansan juurien tutkimuksen vahvistavan kansallista identiteettiä. Fennougristiikka kuului alusta asti ns. kansallisiin tieteisiin, joiden tuottamalla tiedolla katsottiin olevan merkitystä kansakunnan muodostamisessa. (Ks. Salminen 2008: 10–25; Setälä 1912: 9–16.)

Alkoi yksi Suomen tieteenhistorian suurimmista projekteista: 30 vuoden aikana Suomalais-Ugrilainen Seura lähetti 30 stipendiaattia lähi- ja etäsuokieliiä puhuvien kansojen luo kenttätööhön. Kysymys ei ollut lyhyistä käynneistä vaan useiden kerääjien kohdalla monen vuoden yhtämittaisesta työrupeamasta, esimerkiksi Siperiassa toiminut Artturi Kannisto viipyi vogulien eli mansien parissa kuusi vuotta, tapasi heidät kaikki ja teki samalla heistä väestönlaskun. Seuran suurhanke päättyi ensimmäiseen maailmansotaan ja sitä seuranneeseen lokakuun vallankumoukseen, joka teki Venäjälle suuntautuvat tutkimusmatkat mahdottomiksi. (Eiras 2010: 139–151; Korhonen 1986: 144–167; Korhonen et al. 1983: 16–27.)

Keruumatkat tuottivat valtavan kielitieteellisen materiaalin, ja kaikki seuraavat fennougristipolvet 1900-luvun lopulle saakka työskentelivät sen toimittamisessa painokuntoon, sillä kerääjät eivät itse ehtineet tätä suurtyötä saada päätökseen. Monet aineistot, etenkin sanakirjat, joiden laatiminen keruulippujen pohjalta on työteliäämpää kuin esimerkiksi yhtenäisten folkloretekstien julkaiseminen, joutuivat odottamaan vuosikymmeniä, joissakin tapauksissa yli sata vuottakin, ennen kuin ne ilmesivät.

Tscheremissisches Wörterbuch

Vuonna 2008 minun ja Arto Moision toimittamana ilmestynyt marin eli tšeremissin murre-sanakirja oli yksi viimeisistä Suomalais-Ugrilaisen Seuran käynnistämän kansallisen suurhankkeen lopputuloksista. Se perus-

tuu kuuden kerääjän sanastomateriaaleihin: varhaisimmat ovat peräisin 1880-luvulta (Volmari Porkka [1854–1889], Arvid Genetz [1848–1915]), seuraavat 1900-luvun alkupuolelta (Yrjö Wichmann [1868–1932], Martti Räsänen [1893–1976]) ja uusimmat toisen maailmansodan aikana Suomessa olleilta marilaisilta sotavangeilta (T. E. Uotila [1897–1947], Erkki Itkonen [1913–1992]). Tässä se poikkeaa monista edeltäjistään, jotka on koostettu yhden kerääjän materiaalin pohjalta. Sana-aines mahdollisine esimerkkeineen on kirjoitettu musteella lipuille informanteilta, joita on fennougristiikan piirissä ollut tapana nimittää kielimestareiksi. Käännökset on kirjattu vaihtelevasti suomeksi, venäjäksi, saksaksi, joskus ruotsiksikin. Wichmann on kirjoittanut samalle sanalipulle kaikkien kuuden murteen muodot, kun taas Itkonen on merkinnyt keräämiensä neljän murteen sanat eri lipuille. Muut kerääjät ovat kukin merkinneet muistiin sanastoa vain yhdestä murteesta.

Kun aloitin tšeremissin murteiden sanakirjan toimittamisen Kotimaisten kielten tutkimuskeskuksessa (Kotus) vuonna 1981, oli niin suomalaisessa kuin ulkomaisessakin fennougristiikassa vielä käytössä ulkopuolisten, useimmiten naapurikansojen, kansalle ja kielelle antama nimi eli eksonyymi *tšeremissi*, johon ei alallamme sinänsä liittynyt pejoratiivista tai vähättelevää konnotaatiota. *Mari* alkoi tulla Suomessa käyttöön 1980-luvun loppupuolella toisaalta globaalien, vähemmistökansojen asemaa koskevan tietoisuuden kasvamisen myötä, toisaalta sen vuoksi, että rautaesirippu alkoi murtua ja suomalaisilla fennougristeilla oli yli 60 vuoden tauon jälkeen mahdollisuus tavata sukukielten puhujia, myös mareja. Esimerkiksi unkarilaisessa ja saksalaisessa fennougristiikassa suomalais-ugrilaiden kielten eksonyymeja on käytetty pitempään, osittain yhä. Kotuksessaakin nostettiin 1990-luvulla esiin kysymys sanakirjan nimen muuttamisesta, koska eksonyymit olivat yleisen yhteiskunnallisen keskustelun myötä väistymässä endonyymien tieltä. Koska materiaalin kerääjät käyttivät yksinomaisesti *tšeremissi*-nimitystä ja kyse on historiallisesta aineksesta – ei nykykielestä –, *mari* tuntui tässä kontekstissa anakroniselta ja siksi mahdottomalta. Syntyperäiset marilaiset tutkijat eivät ole vierastaneet sanakirjan nimeä, sillä he hyväksyvät *tšeremissi*-nimen esiintymisen historiallisissa ja kielitieteellisissä yhteyksissä.

Useimmat kaksikieliset sanakirjat ovat työryhmän työstämiä samoin kuin esimerkiksi etymologiset sanakirjat. Edellisissä tarvitaan molempien kielten asiantuntijaa, ja jälkimmäisissä monien kielten historian tun-

temusta. Sukukielten tieteellisistä sanakirjoista suurin osa on yhden ihmisen toimittamia, mikä johtuu toisaalta tiettyyn kieleen erikoistuneiden tutkijoiden vähyydestä, toisaalta hankkeisiin osoitetun rahan niukkuudesta. Olin pitkään tšeremissin sanakirjan päätoimittaja ja ainoa työn tekijä (tosin minulla oli sanakirjan lisäksi muitakin työtehtäviä). Koska työn alkutaipaleella oli lukuisia käytännöllisiä ja periaatteellisia kysymyksiä ratkaistavana, sanakirjalle luotiin aluksi ”tausta- ja tukiryhmä”, johon kuuluivat leksikografian ja marin tutkimuksen silloiset huiput eli Tuomo Tuomi, Alho Alhoniemi ja Paavo Siro. Arto Moisio jatkoi toimistustyötä vuodesta 1999 alkaen siirryttyäni pois Kotuksesta. Sen jälkeen osallistuin sanakirjan tekoon tarkistamalla hänen laatimansa artikkelit; loppuvaiheessa kävimme yhdessä läpi taittovedokset.

Tšeremissisches Wörterbuchissa sanat on järjestetty etymologisen periaatteen mukaisesti, mikä tarkoittaa yksinkertaisesti, että samassa sana-artikkelissa esitetään kantasanan lisäksi kaikki siitä muodostetut johdokset ja yhdyssanat. Näin rakennettuja sanakirjoja on fennougristiikassa nimitetty tieteellisiksi sanakirjoiksi¹ erotukseksi tavallisista kaksikielisistä sanakirjoista. (Viimeksi mainituissa johdokset ja yhdyssanat esiintyvät yleensä omina sana-artikkeleinaan, jotka aakkostuksessa saattavat joutua etäälle kantasanastaan. Tämä on tietenkin maallikkolukijan kannalta yksinkertainen tapa esittää sanat.) Etymologisin periaattein koostettujen artikkelien hahmottamista voidaan typografisesti helpottaa erilaisin keinoin, sisentämällä, lihavoimalla kantasana, merkitsemällä johdokset vaikkapa niiden eteen pantavalla ajatusviivalla tai pallerolla jne. Jotta samakantaiset sanat voidaan sijoittaa samaan sana-artikkeliin, niiden etymologia täytyy tuntea. Marin kaltaisessa agglutinoivassa kielessä lähes kaikki johdokset ovat läpinäkyviä, eikä ongelmaa yhteen kuuluvien sanojen havaitsemisessa muutamia poikkeuksia lukuun ottamatta ole. Murre-erotkin ovat melko säännöllisiä, joten eri murteiden yhteiset sanat on marintuntijan melko helppo havaita ja yhdistää yhdeksi sanueksi, mikäli kerääjä ei jo ole sitä tehnyt.

Ulkoisen etymologian eli lainasanojen kohdalla pitää etymologisen periaatteen mukaan eri lähteistä lainatut sanat erottaa omiksi sana-artikkeleikseen. Vuokko Eiras (1996: 284–288) on ansiokkaasti kuvannut

1 Toisaalla tässä teoksessa termiä *tieteellinen sanakirja* on käytetty viittaamaan ylipäänsä tieteelliseen tutkimukseen perustuviin sanakirjoihin.

etymologisen periaatteen soveltamisen ongelmia ja karikkoja vähemmän tutkitun suomalais-ugrilaisen kielen mansin eli vogulin murresanakirjan teossa. Marissa on paljon tšuvassilaisia ja tataarilaisia lainoja. Onneksi ne on tutkimuskirjallisuudessa hyvin kartoitettu samoin kuin vanhat venäläiset lainat; nuoret venäläiset lainat puolestaan ovat helposti tunnistettavissa. (Marin lainasanoista, ks. Fedotov 1965, 1968, Isanbaev 1989, 1994, Räsänen 1920, 1923, Saarinen 1997, Savatkova 1969.) Tšuvassilaiset lainat esiintyvät yleensä koko marilaisella kielialueella, koska ne on omaksuttu jo ennen marilaisen asutuksen pirstoutumista 1600-luvulta alkaen, kun taas toisesta turkkilaisesta kielestä, tataarista, on lainattu sanoja etenkin itäisiin murteisiin 1900-luvulle saakka. Jotkin tataarilaiset sanat ovat tulleet mariin venäjän välityksellä ja jotkin venäläiset puolestaan tataarin kautta. (Marin sanaston alkuperästä yleisesti, ks. Saarinen 2010a: 335–341.)

Useimpiin etymologisen periaatteen mukaan laadittuihin sanakirjoihin ei ole merkitty näkyviin sanojen etymologioita. Leksikografi on nähnyt vaivaa löytääkseen etymologian tutkimuskirjallisuudesta, tai hän on peräti itse etymologi ja rakentaa sana-artikkelin omaan sanan alkuperää koskevaan tutkimukseensa perustuen. Etymologia näkyy implisiittisesti vain artikkelin rakenteesta eli siitä, mitkä sanat tai muodot siihen on sisällytetty. Kuten Eiras (1996: 284–288) kuvaa, eri sanakirjoissa jotkin mansin kielen sanat on sijoitettu eri artikkeleihin, koska sanakirjojen tekijät ovat olleet eri mieltä sanan etymologiasta. Koska sanakirjaan ei mahdu pitkiä etymologisia selvityksiä, sanakirjan käyttäjä ei voi tietää, johtuuko erilainen ratkaisu tekijän tiedon puutteesta vai toisenlaisesta sanan alkuperän tulkinnasta. Joissakin tieteellisissä sanakirjoissa – esimerkiksi Heikki Paasosen (1990–1999) ainekseen perustuvassa mordvan murresanakirjassa – on lainasanojen alkuperä merkitty, mutta omaperäiset on jätetty merkinnättä; ei siis ole otettu kantaa siihen, onko sana esimerkiksi suomalais-volgalainen vai uralilainen.

Tscheremissisches Wörterbuchin kohdalla päätettiin etymologiat merkitä näkyviin lyhyesti, esim. [FU] tai [< Tschuw.]. Tämän ratkaisun mahdollistivat edellä luetellut lainasanatutkimukset ja toisaalta myös Uralisches etymologisches Wörterbuchin (UEW, 1988) ilmestyminen. Viimeisten parinkymmenen vuoden etymologinen tutkimus on tuonut uusia ajatuksia ja vaikuttanut käsityksiin kantakielten ikäyksestä ja jakautumisesta (vrt. esim. J. Häkkinen 2009: 45–51 ja siinä mainittua kirjalli-

suutta). UEW:ia voidaan kuitenkin yhä pitää etymologisen tutkimuksen perusteoksena. Sen ilmestyminen helpotti huomattavasti tšeremissin sanakirjan toimittamista ja erityisesti omaperäisen sanaston esittämisestä siinä. Toista näin kattavaa yleisesitystä uralilaisten kielten yhteisestä sanastosta ei ole (vertailevista etymologisista sanakirjoista, ks. Kulonen 1996: 40–65). Myös marin omaperäisen sanaston osalta UEW on laajin ja paras lähde, koska mariin keskittyvää etymologista sanakirjaa ei ole olemassa. (F. I. Gordeev alkoi vuonna 1979 julkaista marin etymologista sanakirjaa, mutta siitä ilmestyi vain kaksi osaa [1979, 1983], jotka kattavat neljällä ensimmäisellä aakkosella alkavat sanat. Teoksen etymologiat eivät perustu suomalais-ugrilaiseen äännehistoriaan, joten se ei kelpaa lähteeksi.) Ainoat kohdat, joissa olemme poikenneet UEW:n merkintätavasta, koskevat sanoja, jotka esiintyvät marissa ja permiläiskielissä tai mordvassa ja marissa. Edelliset olemme merkinneet lyhenteellä WP (= volgalais-permiläinen) ja jälkimmäiset lyhenteellä W (= volgalainen). UEW:ssa vastaavat merkinnät ovat FP (= suomalais-permiläinen) ja FW (= suomalais-volgalainen), koska teoksen tekijät ovat lähteneet mekaniinisesta sukupuun tulkinnasta, jonka mukaan näissä kielissä esiintyvät sanat ovat automaattisesti suomalais-permiläisiä tai suomalais-volgalaisia, vaikka ne puuttuisivatkin tyystin suomesta ja saamesta.

Joissakin tapauksessa tšuvassin ja tataarin sanat ovat äänteellisesti niin lähellä toisiaan, että on mahdotonta varmasti todeta, kummasta kielestä ne on omaksuttu; silloin sanakirjassa käytetään merkintää [< Tschuw./Tat.]. Usein naapurikielystä on mariin lainattu kantasanan lisäksi johdoksia. Perinteisen toimitustavan mukaan nämäkin olisi pitänyt esittää erillisissä sanakirja-artikkeleissa, koska ne on lainattu erikseen. Ongelmallisimpia ovat tapaukset, joissa sekä marissa että lainanantajakielessä on olemassa samannäköinen johdin: on mahdotonta aivan varmasti tietää, onko johdos lainattu sellaisenaan vai johdettu vasta marissa. Sanakirjassa on päädytty ratkaisuun, jossa johdoksen eteen sijoitetulla <-merkillä osoitetaan, että johtaminen on tapahtunut jo lainanantajakielessä. Näin kantasana ja johdotukset saadaan sijoitetuksi samaan sanakirja-artikkeliin myös lainasanoissa ja samalla informoidaan lukijaa siitä, että johtaminen ei näissä tapauksissa välttämättä noudata marin sääntöjä. Esimerkiksi karjan kaitsemista tarkoittava verbi *kütem* voi äänteellisen asunsa puolesta olla laina joko tataarista tai tšuvassista, siis sanakirjassa on sen jäljessä merkintä [< Tat./Tschuw.]. Siitä on marissa johdettu säännönmukaisesti

momentaanijohdos *kütäläš*. Samasta kannasta on käytössä karjalaumaa tarkoittava muoto *«kütü ~ kütö*, joka on voitu lainata sellaisenaan sekä tataarista että tšuvassista. Paimenta merkitsevistä johdoksista *«kütüčö* ja *«kütüžö* ensin mainitun katsotaan lainautuneen tataarista ja jälkimmäisen tšuvassista. (Galkin 1966: 27–29, 51; Räsänen 1923: 44; Saarinen 1997: 390–391.) Kun tieto sanojen alkuperästä on välttämätön sanakirjaa laadittaessa, on järkevää tarjota se myös sanakirjan käyttäjälle. Esipuheessa on lueteltu tärkeimmät etymologiset lähteet, joiden avulla käyttäjä voi halutessaan päästä sanan alkuperäselityksen jäljille. Tilanpuutteen vuoksi lähdeviitteitä ei ole voitu merkitä sanakirja-artikkeliin.

Kun kyse on murre-sanakirjasta, oli ratkaistava, missä järjestyksessä murremuodot esitetään. Monien suomalais-ugrilaisten kielten kohdalla on tutkimuksessa vakiintunut tietty järjestys, joka voi perustua esimerkiksi maantieteellisten puhuma-alueiden keskinäiseen suhteeseen, murteiden laatuun jatkumolla konservatiivinen–radikaalisti muuttunut tai murteiden puhujien määrään. Marissa on nykykäsitysten perusteella neljä päämurretta: itäinen murre, jota puhutaan Marin tasavallan itäpuolella, keski- eli niittymurre Marin tasavallan keski- ja itäosassa, länsi- eli vuorimurre tasavallan lounaiskulmassa ja luoteismurre tasavallasta luoteeseen sijaitsevalla alueella. Itä- ja keskimurre muodostavat itäisen murreryhmän, länsi- ja luoteismurre länsiryhmän. (Ivanov 1981.) Marissa on kaksi normitettua kirjakieltä: itäinen kirjakieli, johon on yhdistetty piirteitä sekä itä- että keskimurteista, ja länsikirjakieli, jonka pohjana on melko homogeeninen länsimurre. Luoteismurteella ei ole vakiinnutettua kirjallista normia. Etymologisessa kirjallisuudessa länsi- eli vuorimurre on ollut keskeinen, koska se on äänteellisesti itäistä kielimuotoa konservatiivisempi. Koska läntisen kirjakielen puhujia kuitenkin on vain muutama kymmenentuhatta ja itäisen kirjakielen puhujia marien suuri enemmistö eli yli 400 000 henkeä, päädyttiin itäryhmän edustuksen valitsemiseen lähtökohdaksi.

Olisi ollut mahdollista luetella murteiden edustukset idästä länteen ja luoteeseen, mutta tällöin ensimmäiseksi eli hakusanaksi olisi tullut itämurteen muoto, mikä olisi ollut epäkäytännöllistä, sillä itämurteen äänjärjestelmässä on paljon myöhäistä tataarin vaikutusta, minkä vuoksi se poikkeaa Marin tasavallassa puhutuista keskimurteista. Koska marissa on kaksi vakiintunutta kirjakieltä, päätimme antaa hakusanaksi itäisen kirjakielen muodon (lihavoidulla kursivilla) ja sen jälkeen murteiden

muodot idästä länteen kursivoituina. Jos sana esiintyy vain läntisessä murreryhmässä, ensimmäiseksi annetaan sen länsikirjakielen mukainen muoto. Jos sana on tuntematon molemmissa kirjakielissä, artikkeliin ei tule lainkaan lihavoitua hakusanaa vaan vain murremuodot peräkkäin kursivoituina. Päätös kirjakielen muodon käyttämisestä hakusanana noudattaa tapaa, jota on käytetty mm. Suomen murteiden sanakirjassa, toisin sanoen ensin annetaan normatiivinen muoto ja vasta sen jälkeen murreesiintymät. Olemassa olevien mallien lisäksi ratkaisu vaikutti myös eettisesti oikealta, sillä marilaiset kielentutkijat olivat keskusteluissa tuoneet esiin ihmetyksensä siitä, että länsimaisessa tutkimuksessa esiintyi heidän kirjakielensä asemesta vain murremuotoja. Fennougristisessa perinteessä ei 1980-luvulle asti ollut juurikaan otettu huomioon Venäjän suomalais-ugrilaiden nuoria kirjakieliä, vaan kutakin kieltä oli opiskeltu ja tutkittu murteiden kautta 1970-luvulle asti. Sitä edeltävä fennougristiikka oli siis lähtökohtaisesti murteiden tutkimusta. Tähän oli syynä toisaalta kirjakielten nuoruus tai vakiintumattomuus, mutta ennen kaikkea vuosisadan alun tutkijoiden keräämät valtavat kielimateriaalit, joita käytettiin sekä opetuksessa että tutkimuksessa. Ne edustivat murteita, ja ne oli kirjoitettu suomalais-ugrilaisella tarkekirjoituksella. Tšeremissin sanakirjassa haluttiin tietoisesti osoittaa kirjakielen olemassaolo. Kirjakielen muotoa ei kuitenkaan merkitty kyrillisin kirjaimin, sillä suomalaisessa fennougristiikassa niitä ei kolmekymmentä vuotta sitten paljонkaan käytetty sukukielten kirjoituksessa. Edelleenkin monet länsimaiset kielentutkijat eivät tunne kyrillistä kirjaimistoa, joten sen käyttö olisi vähentänyt sanakirjan käyttökelpoisuutta.

Murteiden järjestyksen lisäksi oli ratkaistava myös murre-signumit eli kirjainlyhenteet, joilla murteita merkitään. Suomalaisten ja unkarilaisten fennougristien keräämään vanhaan materiaaliin perustuvissa sanakirjoissa on yleensä säilytetty kerääjien käyttämät murre-signumit, jotka viittaavat informanttien kotikylään ja/tai siihen kuvernementtiin, jossa kylä on sijainnut. Näin on tehty vielä esimerkiksi unkarilaisen Ödön Beken ensimmäisen maailmansodan aikana Unkarissa olleilta marilaisilta sotavangeilta keräämään aineistoon perustuvassa murre-sanakirjassa, jota toimitettiin Unkarissa osittain samanaikaisesti kuin Tšeremissisches Wörterbuchia (Beke 1997–2001). Kerääjien antamia murre-signumeja on käytetty heidän kirjoittamissaan tutkimuksissa ja heidän julkaisemissaan folklorekokoelmissa. (Useimmat kenttätutkijat keräsivät myös

tekstejä, joko folklorea tai kertomuksia paikallisista tavoista ja oloista. Ne on yleensä julkaistu ennen sanakirjoja, koska niiden toimittaminen on sujunut nopeammin.) Toisin sanoen kulloisenkin kielen tutkijalle signumit ovat jo entuudestaan tuttuja. Muille ne eivät varsinaisesti kerro mitään, sillä kuvernementteja, joihin niillä viitataan, ei ole 1920-luvun jälkeen enää ollut olemassa, eikä pieniä kyliä löydy etenkään Neuvostoliiton aikaisista kartoista. Koska marin kielen murteista oli jo 1980-luvulla olemassa yksityiskohtainen selvitys ja jaottelu, tuntui välttämättömältä ottaa se jotenkin huomioon (Ivanov 1981). Niinpä noudatimme murteiden nykyistä nelijakoa (itä-, keski-, länsi- ja luoteismurre; ks. murrejakoa edellä), joiden mukaan muodostimme murresignumin (saksankielisen) alkuosan (O, M, W, NW) ja lisäsimme siihen alamurteen tai keruukylän alkukirjaimen pienellä kirjaimella. Näin marin nykyisen murrejaon tuntevat ymmärtävät jo signumin ensimmäisestä osasta, minkä alueen murteista on kyse. Vanhan materiaalin suhteuttaminen nykytietoon helpottaa sanakirjan käyttöä.

Tscheremissisches Wörterbuch on tehty ennen kaikkea tutkijoille, sekä fennougristeille että muille marista kiinnostuneille. Aineksen kerääjät ovat merkinneet sanojen foneettisen asun muistiin tarkekirjoituksella (keruuajankohtanahan ei foneemin käsitettä vielä ollut olemassa) eli mahdollisimman tarkasti, niin kuin he ovat sen kuulleet. Tämä tekee sanakirjan käyttämisen hankalaksi marilaiselle maallikolle, joka on tottunut kirjoittamaan kieltään kyrillisin kirjaimin. Marilaiselle kielentutkijalle sanakirja kyllä avautuu, koska hän hallitsee suomalais-ugrilaisen tarkekirjoituksen, jota käytetään Marin tasavallassakin fonetiikkaa kuvattaessa. Sanakirjan materiaali, jonka kerääminen oli osa suomalaisten kansallista projektia, hyödyttää nyt mareja heidän luodessaan uudissanvoja ja tutkiessaan omaa kieltään. Marista kuten muiltakin Venäjän suomalais-ugrilaisilta alueilta puuttuvat laajat vanhat murrekeräelmät. Siksi suomalaisten ja unkarilaisten fennougristien kokoamaan materiaaliin pohjautuvat julkaisut ovat heille hyvin tärkeitä.

Eettiseksi voidaan katsoa kysymys työpanoksen ja sen hinnan suhteesta potentiaaliin käyttäjiin. Jokainen kieli kätkee sisäänsä oman maailmankatsomuksensa, oman tapansa jäsentää olevaa. Nykyään ymmärretään, että kieliä pitää suojella, kuten uhanalaisia kasveja ja eläimiäkin, sillä niiden katoaminen köyhdyttää ihmiskunnan henkistä perimää. Venäjän suomalais-ugrilaiset vähemmistökielet ovat vaarantuneita, koska kielen-

vaihto venäjään on voimakasta eikä oma kieli välity nuorille sukupolville (kielen uhanalaisuuden kriteereistä ja suomalais-ugrialaisten kielten tilanteesta, ks. esim. Pasanen 2008: 47–70). Koska suomalais-ugrilaiset tieteelliset sanakirjat eivät ole laajojen kollektiivien laatimia vaan yleensä yksittäisen henkilön aikaansaamia, niiden kustannukset eivät ole kohtuuttomia. Humanistinen tutkimus tuottaa usein julkaisuja, joiden arvoa on ekonomistille vaikea selvittää, koska ne eivät ole suoraan muunnettavissa rahaksi. Sanakirjan merkitys kielentutkimukselle ja marin kielen kehitykselle on kuitenkin kiistaton. (Suomalais-ugrialaisten sanakirjojen asemasta kielenuudistuksessa, ks. Saarinen 2005: 82–96.)

Selityskielenä sanakirjassa on fennougriistiikan perinteiden mukaisesti saksa. Saksalla on pitkä perinne ”fennougriistiikan äidinkielenä”. Kun sanakirjaa alettiin 1980-luvulla laatia, saksan asema oli sekä fennougriistiikassa että yleisemminkin tieteessä vielä huomattavasti vahvempi kuin nyt. Unkarilaiset alan tutkijat osasivat yleensä erinomaisesti saksaa, ja myös monet Venäjän suomalais-ugrialaisten kielten syntyperäiset tutkijat pystyivät ainakin lukemaan saksankielistä tieteellistä kirjallisuutta tai jopa puhumaan sitä. Englanti oli saksaan verrattuna perifeerinen. Esimerkiksi fennougriстикongresseissa saksa oli pitkään valtakieli, vaikka englantiakin voitiin käyttää. Lähes kaikissa aikaisemmissa tieteellisissä etäsukukielten sanakirjoissa saksa on selityskielenä, muutamissa Suomessa ilmestyneissä suomi ja vastaavasti Unkarissa unkari. 1990-luvulla tilanne alkoi nopeasti muuttua englannin hyväksi.

Jos sanakirjan tärkein kohdeyleisö olisi ollut marit, selityskieleksi olisi pitänyt valita venäjä, mutta silloin (läntinen) tiedeyhteisö ei olisi voinut hyödyntää sitä. Paasosen mordvan sanakirjassa (1990–1999) tehtiin kompromissi: selityskielenä on saksa, mutta sanojen merkitys annetaan sen lisäksi myös venäjäksi, esimerkkien käännökset vain saksaksi. Ratkaisu vie tilaa (mitä yhdessäkään sanakirjassa ei ole riittävästi) ja käy myös kalliiksi.

Saksa on selityskielenä englantia parempi, sillä englannissa sanalla saattaa olla hyvinkin monia erilaisia merkityksiä, joista vain kontekstin avulla voi valita kulloisessakin tapauksessa oikean. Kontekstin kuvaaminen vie tilaa eikä aina onnistu riittävän yksiselitteisesti. Koska saksalla on pitkä perinne fennougriistisissa julkaisuissa, mm. monille suomalais-ugrialaisten kansankulttuurin tarkoitteille on vakiintunut saksankielinen käännösvastine. Mekaanista ei kääntäminen liene koskaan, eivätkä kään-

nösvastineet välttämättä sisällä samoja konnotaatioita kuin alkuperäinen sana. On selvää, että kääntäminen kielestä toiseen voi hämärtää sanan alkuperäistä merkitystä. Suomalais-ugrilaiden sanakirjantekijöiden keskuudessa kiertävä tyyppillinen esimerkki merkityksen muuttumisesta kääntämisprosessissa on suomen sanan *tuokkonen* kääntäminen saksaksi *Schachtel aus Birkenrinde*, joka takaisin suomeen käännettäessä muuttuu *koivuntuohesta tehdyksi rasiaksi*, ellei kääntäjä tunne riittävästi suomaista todellisuutta.

Tavallisten kaksikielisten sanakirjojen kohdalla paljon aikaa ja energiaa kuluttava työvaihe on käännettävien sanojen valitseminen. Tscheremissisches Wörterbuchissa ei tällaista ongelmaa ollut: koska se perustuu vanhaa materiaaliin, mukaan otettiin kaikki sanat, jotka kerääjät ovat muistiin merkinneet ja myös kaikki heidän antamansa esimerkit. Kerääjät ovat kuvailleet työtään ja informanttejaan Suomalais-Ugrilaiselle Seuralle antamissaan selvityksissä. Informantit ovat poikkeuksetta miehiä, mikä on ymmärrettävää keruuajan olosuhteet ja asenteet huomioon ottaen. Naisten olisi ollut sopimatonta työskennellä ulkomaalaisten muukalaisten kanssa. (Tšeremissin sanakirjan informanteista, ks. esipuhe, s. II–VII.)

Viettäessään vuoden marien luona 1905–1906 Yrjö Wichmann käytti yhtenä informanttinaan nuorta opettajaa Timofej Jevsevjeviä. Jevsevjev kiinnostui työstä niin, että hän alkoi itsekin kerätä marilaista folklorea ja lähetti Suomalais-Ugrilaiselle Seuralle yli kahdenkymmenen vuoden aikana valtavan tekstikokoelman ja Seuran kansatieteilijöille valokuvia ja kansatieteellisiä kuvauksia. Neuvostoaikana Jevsevjev toimi marien kansallismuseon johtajana jonkin aikaa. Hänet tapettiin v. 1937 Stalinin ajan vainoissa, kuten kaikki muutkin marilaisen sivistyneistön edustajat. (Lehtinen 1985, 8–45; Saarinen 2010b: 107–122.) Koska marin murre-sanakirjan esimerkkiaineisto oli paikoin niukka, sitä päätettiin täydentää Jevsevjevin aineksella. Käsikirjoitusmateriaali oli kuitenkin ensin julkaisutava, jotta siihen voitiin viitata asianmukaisesti. Niinpä rinnan sanakirjan kanssa toimitettiin painokuntoon neljä nidettä Jevsevjevin keräämää marilaista folklorea, tarinoita, satuja, pienfolklorea, rukouksia ja loitsuja sekä kansanlauluja (Jevsevjev 1983, 1989, 1992, 1994).

Sanakirjaan sisältyvä sanasto kuvastaa 1900-luvun vaihteen agraaris-ta marilaista kulttuuria. Toimitusvaiheessa ei sanoja ole karsittu. Mikäli sanakirjasta puuttuu jonkin alan sanastoa, se johtuu kerääjien tietoisista tai tiedostamattomista valinnoista, joita he eivät ole kirjanneet kenttä-

työkuvauksiinsa. Voisi ajatella, että kerääjät eivät ole halunneet merkitä muistiin nuoria, äänteellisesti mariin mukautumattomia venäläisiä lainoja, mutta aineisto ei oikeastaan tue tällaista luuloa. Länsimurteista, jotka ovat pisimpään olleet venäläisen vaikutuksen alla, kerääjät ovat tallentaneet eniten venäläisiä lainoja, mutta yksittäisiä venäläisiä lainasanoja näyttää löytyvän kaikista murteista. Yllättävästi varhaisimman kerääjän Volmari Porkan materiaalissa niiden frekvenssi on suurin, mutta ei siinäkään mitenkään merkittävä. Usein sopimattomana pidettyä seksuaaliterminologiaa ei sanakirjasta myöskään puutu, mutta informanttien sukupuolen vuoksi se edustaa miehistä näkökulmaa.

Reverse Dictionary of Mari (Cheremis) **– Обратный словарь мари́йского языка**

Marin käänteissanakirja ilmestyi vuonna 2002. Se on julkaistu Suomalais-Ugrilaisen Seuran Lexica-sarjassa, kuten murre-sanakirjakin. Verrattuna edellä kuvaamiini ratkaisuihin jo teoksen nimessä huomaa ympäröivän tiedeyhteisön muutoksen: kirja on englanninkielinen, ja kielen primaarimimityksenä on *Mari*, vaikka suluissa varmuuden vuoksi yhä mainitaan myös *Cheremis*. Sanakirja on nimensä tavoin kaksikielinen, toisena kielinä on venäjä. Kaksikielisyyteen vaikutti tekijäkollektiivi: teos on tehty Turun yliopiston Volgan alueen kielten tutkimusyksikössä yhteistyössä marilaisten tutkijoiden Oleg Sergeevin ja Lidija Matrosovan kanssa. Suomalaisia tekijöitä on kolme: Jorma Luutonen, joka vastasi koko työn organisoimisesta, minä ja Arto Moisio. Sanakirjan alaksi otettiin koko marin kieli, siis sekä itäisen että läntisen kirjakielen sanat ja jonkin verran myös murre-sanastoa, sillä kirjasta haluttiin tehdä mahdollisimman kattava marin sanaston kuvaus.

Volgan alueen kielten tutkimusyksikössä on laadittu monia kaksikielisiä sanakirjoja, joissa toisena kielinä on jokin Volgan alueella puhuttu suomalais-ugrilainen kieli ja toisena suomi, sekä useita käänteissanakirjoja joko painettuina tai sähköisinä listoina. Sanakirjat ovat olleet hyviä yhteistyöprojekteja toteutettaviksi syntyperäisten tutkijoiden kanssa. Kaksikieliset sanakirjat ovat tarpeellisia sekä kielten (joko suomen tai vastaavasti marin/mordvan/udmurtin) opiskelijoille että näiden suomalais-ugrilaisien kielten tutkijoille. Etenkin suomi–suomalais-ugrilainen

kieli -sanakirjoja laadittaessa on myös jouduttu luomaan uutta sanastoa mariin/mordvaan/udmurttiin suomen käsitteiden ilmaisemiseksi. Näin sanakirjat ovat edistäneet sukukielten kehitystä (useimmista on otettu Venäjällä toisintopainos). Käänteissanakirjat on laadittu samaa tarkoitusta varten. Suomen kielen käänteissanakirjan ilmestyminen 1970-luvulla siivitti johto-opin tutkimusta meillä, ja samaa olemme toivoneet tapahtuvan myös Venäjän suomalais-ugrilaisilla alueilla. Käänteissanakirja on hyvä apu uudissanojen luomisessa, koska se paljastaa kielen sananmuodostuskeinot, etenkin johdinsuffiksit ja niiden frekvenssin.

Maallikot olettavat usein, että sanakirjan tekeminen on yksinkertaista: pannaan vain sanat aakkosjärjestykseen ja käännetään niiden merkitykset. Monet kielentutkijatkin ajattelevat, että käänteissanakirja vaatii vain yhden tietokoneajon, jossa tavallisen sanakirjan sanat aakkostetaan lopusta päin. Miksi marin käänteissanakirjan tekemiseen siis tarvittiin viiden ihmisen kollektiivi? Marin käänteissanakirja perustuu moniin sanakirjoihin, sekä kaksikielisiin että erikoisalojen sanakirjoihin. Mukana on myös uudissanojen sanakirja (Ivanov 2001). Marissa pitkään tekeillä ollut kymmenosainen marin kielen sanakirja *Словарь мари́йского языка* (Slovar', 1990–2005), joka on tarkoitettu Nykysuomen sanakirjan (1951–1961) tyyppiseksi kielen sanavarat kattavaksi teokseksi mutta jossa selityskielenä on olosuhteiden pakosta venäjä, ei ollut vielä kokonaisuudessaan valmistunut. Teosta ei voitu siksi ottaa mukaan, kuten ei myöskään tuolloin vielä keskeneräistä *Tscheremissisches Wörterbuchia*.

Käänteissanakirjan pohjaksi ei valittu ainoastaan nykysanakirjoja vaan myös Heikki Paasosen *Ost-tscheremissisches Wörterbuch* (1948) ja Valerian Vasiljevin *Мари́й мутэр* (1926). Edellinen oli ongelmallinen, koska se on laadittu tarkekirjoituksella, jälkimmäinen siksi, että siinä on käytetty varhaisempaa kyrillistä ortografiaa, joka poikkeaa nykyisestä kirjoitustavasta etenkin vokaalien osalta. Jotta sama sana ei tulisi sanastoon kahta kertaa vain erilaisen kirjoitusjärjestelmän takia ja jotta näiden teosten sanat saataisiin integroitua uudempaan materiaaliin, siten helpotettua tietokoneajoja ja yhdenmukaistettua sanaluettelot, sanat oli muutettava nykyisen kyrillisen ortografian mukaisiksi. Sitä ei voi tehdä automaattisen ohjelman avulla. (Samaa käytäntöä noudatettiin myöhemmin myös mordvan käänteissanakirjan kohdalla (RDM, 2004).)

Käänteissanakirjat eivät sisällä sanojen merkityksiä. Jotta sanakirja olisi hyödyllisempi ja informatiivisempi sen käyttäjille, halusimme mer-

kitä sanoista näkyviin sanaluokan ja lähteen. Työläin osa marin käänteisanakirjassa oli sanaluokkakoodin antaminen kaikille sanoille. Kooditusta ei saanut automaattisesti lähteinä olleista sanakirjoista. Lisäksi marissa sana voi kuulua samanaikaisesti kolmeenkin sanaluokkaan, esim. substantiivi/adjektiivi/adverbi. Sanaluokkakoodin antamiseksi jokainen sana oli käytävä manuaalisesti läpi. Oma osuuteni käänteissanakirjan toimituksessa koostui juuri länsimarin sanaluokkakoodituksesta.

Udmurttilais-suomalainen sanakirja ja Suomalais-udmurttilainen sanakirja

Sergej Maksimov oli aloittanut udmurttilais-suomalaisen sanakirjan tekemisen itsekseen; tulin hankkeeseen mukaan melko myöhään, jolloin aika paljon materiaalia oli jo koossa. Maksimov on murteiden ja niiden sanaston tutkija, hän osallistuu muun muassa udmurtin kielikartaston laatimiseen. Suuri ongelma udmurttilais-suomalaisen sanakirjan laadinnassa oli venäläisen ja suomalaisen sanakirjatradition erilaisuus. Udmurttilaista sanakirjatraditiota ei juurikaan ole olemassa. Syntyperäiset udmurttitutkijatkin katsovat usein omaa kieltään venäjän kielen läpi. Vaikka on kyse udmurttilais-suomalaisesta sanakirjasta, sanat halutaan esittää, niin kuin ne esitettäisiin udmurttilais-venäläisessä sanakirjassa.

Udmurtissa ja suomessa lähes jokainen substantiivi voi toimia myös toisen substantiivin määritteenä/attribuuttina (suomessa on tällöin useimmiten kyse yhdyssanasta, udmurtissa sanaliitosta). Se on itsestäänselvyys, jota ei tarvitse osoittaa sanakirjassa. Venäjässä näin ei ole, vaan substantiivista muodostetaan attribuuttiasemassa vastaava adjektiivi, esimerkiksi udmurtin *пунь* (*piń*) 'hammas' käännetään venäjäksi sekä substantiivilla että adjektiivilla: 1. *зуб* (*zub*), 2. *зубной* (*zubnoj*). Tämä kaava tunkeutuu suomenkieliseenkin sanakirjaan, niin että *piń*-sanan käännösvastineina annetaan 1. hammas, 2. hammas=, jossa jälkimmäinen kohta osoittaa, että sana voi esiintyä yhdyssanan alkuosana, esim. *hammassärky*. Kysymys on leksikografian käytäntöön liittyvä mutta myös eettinen, sillä vähemmistökielen rakennetta katsotaan venäjän läpi silloinkin, kun ei olla kääntämässä venäjään tai venäjystä.

Vaikka sen paremmin suomessa kuin udmurtissakaan ei ole kieliopillista sukua, venäjän vaikutus tunkeutuu helposti merkityksenselityksiin,

kun on kyse ihmistä tarkoittavista sanoista. Esimerkiksi kansallisuuden nimitysten kohdalla udmurttilainen sanakirjantekijä haluaa suomessakin antaa sanalle *удмурт* (*udmurt*) selityksen: 'udmurtti, udmurttilainen mies, udmurttilainen nainen' tai, jos merkityksenkäännökseksi sovitaan pelkkä 'udmurtti', erikseen ainakin esimerkin *удмурт кышно* (*udmurt kišno*) 'udmurttilainen nainen', koska prototyypillisesti pelkkä *удмурт* (*udmurt*) tarkoittaa miestä. Osittain tässä on kyse myös samasta substantiivin attribuuttiasemaan liittyvästä ongelmasta, jota jo edellä kuvasin. Venäjässä sanalla *удмурт* (*udmurt*) olisi kaksi erillistä substantiivikäännöstä, jotka osoittavat molempia sukupuolia: 1. *удмурт*, 2. *удмуртка* (= 1. udmurttimies, 2. udmurttinainen). Sama koskee kaikkia kansallisuudennimityksiä, myös sanaa *финн* (*finn*) 'suomalainen'. Venäjästä on myös lainattu udmurttiin *ka*-johdin, joka esiintyy ennen kaikkea venäläisissä lainoissa. Suurin osa näistä on naispuolisen kansallisuuden nimiä, esimerkiksi *финка* (*finka*) 'suomalainen nainen'. (Vrt. Obratnyj 1992: 10–14.) Kieleen on luotu venäjänmukainen tapa erottaa sukupuolet kansallisuuden nimityksissä. Alkuperäisessä udmurttilaisessa systeemissä tätä eroa ei ole ollut kuten ei suomessakaan.

Kaksikielisessä sanakirjassa suuren osan työprosessista vie sanojen valitseminen. Esikuvana käytetään yleensä aikaisempia sanakirjoja työn helpottamiseksi, esimerkiksi suomessa Suomen kielen perussanakirjaa tai Kielitoimiston sanakirjaa. Kielen sanojen frekvenssiluettelo kertoo, mitkä sanat ovat yleisiä ja siis välttämättömiä, mitkä taas harvinaisempia. Valittavasti monista kielistä kuten udmurtistakaan ei frekvenssilaskelmia ole tehty. Sanoja valittaessa täytyy miettiä, kenelle sanakirja on tarkoitettu. Tietenkään käyttäjäryhmää ei sovi valita kovin kapeasti. Esimerkiksi udmurttilais-suomalainen sanakirja on ensi sijassa tarkoitettu udmurtin opiskelijoille ja udmurtin tutkijoille mutta yhtä lailla myös udmurttilaisille suomen opiskelijoille. Sanakirjassa kartoitettavaksi on valittava neutraali yleiskieli: kun tekeillä on mahdollisesti kautta aikojen ainoa udmurttilais-suomalainen tai suomalais-udmurttilainen sanakirja, ei ole mielekästä ottaa mukaan esimerkiksi hetkellistä puhekielistä sanastoa vaan pysyväämpää, kaikille ymmärrettävää normatiivista kieltä. Yleiskielikin muuttuu, mutta sen muutostahti on hitaampaa kuin puhekielen. Lisäksi se on homogeenisempaa, eli siinä ei ole niin paljon paikallista tai sosiaalista vaihtelua.

Sanakirja heijastaa aina kieltä puhuvan kansan kulttuuria ja maailmankuvaa, vaikka siihen ei erityisesti pyrittäisi. Udmurttilais-suomalainen

sanakirja kuvastaa maaseutukulttuuria eli sitä ympäristöä, jossa udmurtin kieli elävimmillään on yhä käytössä. Sanakirja on täynnä sellaisten esineiden, ilmiöiden ja käsitteiden nimityksiä, joille ei ole suomessa vastineita ja jotka täytyy sen vuoksi suomessa selittää määritelmänomaisesti. Esimerkiksi udmurtin *мудор* (*mudor*) 'rukouskoden pyhä nurkkaus' tai *чук* (*čuk*) 1. terttu, 2. tupsu, 3. tilkku; *чук шуக்கыны* (*čuk šukkinj*) 'lyöä tilkku (kolikolla seinään t. laipionkannattimeen armeijaan saattamismenojen osana)'. Jälkimmäinen esimerkki saattaa suomalaisesta tuntua perifeeriseltä ja oudolta, mutta udmurttilaisella maaseudulla ei tarvitse käydä monessakaan tuvassa, kun ilmiöön törmää, eli näkee seinänraossa kolikon tilkkuineen. Koska se on jokapäiväinen ilmiö Udmurtiassa, on järkevää ottaa se myös sanakirjaan. Näin sanakirja palvelee sekä udmurttia, joka etsii suomenkielistä käännöstä oman kulttuurinsa tarkoitteille, että suomalaista, joka kääntää udmurttilaisia tekstejä suomeksi.

Suomalais-udmurttilainen sanakirja puolestaan kuvaa suomalaista nykykulttuuria, jossa tietokone on paljon auraa useammin käytetty esine ja sana. Udmurttikulttuurille ominaisten tarkoitteiden nimet on enimmäkseen jätetty pois teoksesta. Ratkaisu johtuu toisaalta tilanpuutteesta, toisaalta siitä, ettei niille ole suomessa terminomaisia vastineita vaan ne pitää useimmiten kääntää selittävällä ilmauksella.

Udmurtin kielessä on vähemmän sanoja kuin suomessa, tosin niiden määrää on tarkempien tutkimusten puutteessa vaikea arvioida. Vuonna 2008 ilmestyneessä udmurttilais-venäläisessä sanakirjassa (URS) on n. 50 000 sanaa. Udmurttilaisen sanaston suppeuteen suomalaisen nähdessä vaikuttaa muun muassa yhdyssanojen harvalukuisuus. Suomessahan esimerkiksi Nykysuomen sanakirjan yli 200 000 hakusanasta runsaasti yli puolet on yhdyssanoja (Tuomi 1989: 28). Suomalais-udmurttilaiseen sanakirjaan on otettu vain sellaisia yhdyssanoja, jotka eivät ole läpinäkyviä, eli ymmärrettävissä osiensa perusteella, mutta kuitenkin keskeisiä merkitykseltään (esim. *tietokone*).

URS:ssa udmurtin sanamäärää lisää tapa antaa lähes kaikista verbeistä produktiivinen teonnimi erillisenä hakusanana, mikä vastaisi sitä, että suomalaiseen sanakirjaan merkittäisiin kaikki verbikantaiset *-minen*-johdokset. Udmurtin teonnimet ovat muodostukseltaan identtisiä partisiippien kanssa. Tästä käytännöstä on sekä udmurttilais-suomalaisen että suomalais-udmurttilaisen sanakirjan laatimisessa keskusteltu toistuvasti. Mielestäni lähes mekaanisesti eli täysin säännönmukaisesti tuo-

tettavat *-minen*-teonnimet ovat produktiivisuudeltaan rinnastettavissa infiniittisiin muotoihin eikä niiden paikka ole sanakirjassa. Valitettavasti Suomalais-udmurttilaiseen sanakirjaan ei mahtunut myöskään monia muilla johtimilla muodostettuja teonnimiä. Mukaan otettiin lähinnä leksikaalistuneita tapauksia. Ratkaisu voi olla udmurttien kannalta hankala, koska astevaihtelun vuoksi teonnimet voivat erota verbin hakumuodosta melko paljon (esim. *lukea* : *luenta*). (Astevaihtelutaulukot kylläkin esitetään kirjan lopussa.) Sähköiseen sanakirjaan ne tietenkin voitaisiin sisällyttää, koska tila ei ole sellaisessa niin rajallinen. URS:ssa kuten udmurttilais-suomalaisessa sanakirjassakin on myös otettu mukaan suuri määrä frekventatiivijohdoksia. Suomalaisissa sanakirjoissahan ne poimitaan mukaan lähinnä silloin, kun niiden frekvenssi on suuri tai ne ovat jotenkin leksikaalistuneita. Suomalais-udmurttilaisesta sanakirjasta niitä on karsittu.

Nykyudmurtissa on paljon nuoria venäläisiä lainasanoja, jotka eivät ole mukautuneet udmurtin äännejärjestelmään. Uudissanat kiellettiin nationalistisina Stalinin kaudella 1930-luvulla. Niitä alettiin kehitellä uudelleen vasta Neuvostoliiton hajottua Venäjän alkaessa demokratisoitua. Sergej Maksimov muodosti suomalaisia käsitteitä varten uusia udmurtin sanoja pääasiassa johtamalla. Uudismuodosteet on merkitty sanan jälkeen seuraavalla asteriskilla. Ymmärtämisen helpottamiseksi myös käytössä oleva venäläinen laina on annettu sana-artikkelissa. Sivistyssanat, jotka ovat suomessa ja udmurtissa (tai oikeammin venäjässä) lähes samannäköisiä (esim. *informaatio* – *informacija*), jätettiin pois, koska ne ovat tunnistettavissa sellaisinaan.

Olen usein kuullut väitettävän, että kaksikielisestä sanakirjasta on helppo tehdä toisinpäin kaksikielinen yksinkertaisesti antamalla tietokoneen hoitaa sanojen järjestely päinvastaiseen järjestykseen. Systemi voi toimia tietokoneinsinöörin aivoissa mutta ei käytännössä juuri edellä esitetyistä syistä. Kielenulkoinen maailma eli kulttuuri, jossa kieli elää, on otettava huomioon sanoja sanakirjaan valittaessa, ja valinta on tehtävä pitkälti lähtökielen ehdoilla. Muuten sanakirja kuvastaa jotakin muuta kuin kuvattavan kielen todellisuutta.

Lopuksi

Tieteellisen, vanhaan ainekseen perustuvan sanakirjan toimittamisessa leksikografin kädet ovat pitkälti sidotut: valintoja voi tehdä vain materiaalin puitteissa ja sen ehdoilla. Tieteenalan konventiot vaikuttavat ratkaisuihin voimakkaasti. Koska tällaisia sanakirjoja tehdään yleensä pitkään muun muassa niiden edellyttämän etymologisen taustatyön ja niukkojen työvoimaresurssien takia, toimitusperiaatteet ja -käytänteet on hiottava valmiiksi jo ennen työn alkua, sillä kurssin muuttaminen kesken työn aiheuttaisi aina valtavan korjaustyön jo toimitetussa aineistossa.

Käänteissanakirja on oikeastaan sanaluettelo, johon on voitu lisätä myös tieto esiintymästä (sanakirjalähteestä) ja sanaluokasta. Käänteissanakirjat ovat hyviä instrumentteja sanastontutkijoille ja myös uudissanojen luojille, koska ne antavat nopeasti tiedon sananmuodostuskeinoista ja niiden produktiivisuudesta.

Kaksikielisessä sanakirjassa keskeinen työvaihe on sanakirjaan otettavien sanojen valinta. Kielen sanapaljoudesta on seulottava esiin ne, jotka ovat tarpeellisia.

Sanakirjatyössä tehtävät valinnat perustellaan yleensä joko tieteellisellä ja käytännöllisellä kannalta, mutta monet niistä sivuavat myös eettisiä kysymyksiä. Usein on kyse paljolti siitä, miten paljon tietoa ja missä muodossa voidaan ja halutaan tarjota sanakirjan käyttäjälle – tämäkin on pohjimmiltaan eettinen ongelma.

LAURA TYYSSTERI

Arvosidonnaiset sanat yleiskielen selittävässä sanakirjoissa

Nykyään useimmissa sanakirjoissa pyritään kuvaamaan kieltä neutraalista näkökulmasta niin, että mitään arvorakennelmaa ei nosteta toista paremmaksi. Leksikografit eivät kuitenkaan voi toimia kulttuurisessa tyhjiössä. Koska kieli ei ole ideologiatonta, kielen kuvaus heijastaa aina väistämättä myös kieliyhteisön arvoja ja välittää niitä eteenpäin. Vaikka varsinkin monet käyttäjät tuntuvat uskovan sanakirjojen täydelliseen puolueettomuuteen ja objektiivisuuteen, sanakirjojen toimitustyössä joudutaan tekemään paljon valintoja, jotka heijastavat myös tekijöiden omaa arvomaailmaa. (Sarfati 2002; Grigoryan 2007: 2; Lehtosalo–Tyysteri 2011: 213–214.) Moniarvoisessa yhteiskunnassa on mahdotonta, että kaikki olisivat kaikesta samaa mieltä. Kenen käsitystä kielestä sanakirjassa sitten tulisi kuvata?

Koska sanakirjojen tarkoitus on olla mahdollisimman tarkkoja kielenkuvauksia, niitä tutkimalla voidaan onnistua paljastamaan kieleen heijastuvia yhteiskunnallisia vinoumia. Sanakirjoja on kuitenkin tutkittu tästä näkökulmasta vain vähän. Eniten aihetta on tarkasteltu sukupuolten välisen tasa-arvon kannalta (mm. Kram 1998; Epple 2000; Braun–Kitzunger 2001; Mattisson 2006; Grigoryan 2007; suomea sisältävien sanakirjojen osalta mm. Karlsson 1974; Siirainen 1988; Mantila 1998; Grönros 2007; Pilke 2009, 2010; Tyysteri 2011), mutta myös muutamiin muihin teemoihin on tartuttu. Esimerkiksi Georges-Elia Sarfati (2002) käsittelee tutkimuksessaan eri-ikäisten ranskalaisten sanakirjojen välittämää ja rakentamaa kuvaa juutalaisista. Michael Carr (1994) puolestaan vertailee japanilaisten ja muunmaalaisten sanakirjojen kuvauksia japaninkielisestä, erityisesti 1900-luvun sotien aikana politisoituneesta japanilaisesta ”henkeen” (*spirit/soul*) viittaavasta ilmauksesta *yamato-damashii*. Myös

sanastoa vaihtoehtoisista näkökulmista kuvaavat sanakirjat ovat useimmiten saaneet innoituksensa sukupuolten epätasa-arvosta. Pelkästään englanninkielisiä naisnäkökulmasta sanastoa kuvaavia sanakirjoja on julkaistu vuosina 1970–2006 yli tusina (Russell 2012: 1). Muita näkökulmia edustavista sanakirjoista mainittakoon Matti Wibergin (2011) teos Poliittinen sanakirja – Inhorealistin versio.

Sanakirjojen arvosidonnaisuutta käsitteleville tutkimuksille on tyyppillistä diskurssianalyttinen tutkimusote (diskurssianalyysin teoriasta esim. Fairclough 1992). Tutkimuksissa tehdään näkyviksi kielellisten valintojen takana olevia (usein tiedostamattomia) ideologioita sekä havainnoidaan sitä, miten kielen avulla tuetaan ja uhmataan yhteiskunnassa vallitsevia valtasuhteita (Fairclough 1992: 12; Kalliokoski 1996: 8). Myös tämän artikkelin tavoite on huomion kiinnittäminen sanakirjojen arvosidonnaisuuteen. En kuitenkaan tarkastele sanakirjojen välittämiä ideologioita tai maailmankuvaa, vaan keskityn arvosidonnaisen sanaston kuvaamisessa vastaan tuleviin valintatilanteisiin. Sanakirjojen tekijöiden ja käyttäjien kannalta on olennaista pohtia, 1) millainen sanasto ylipäänsä voi olla arvosidonnaista, 2) miten arvosidonnaista sanastoa kuvataan sanakirjoissa ja 3) millaisia syitä ja seurauksia valituilla kuvaustavoilla voi olla. Tarkastelen eri kysymyksiä omissa luvuissaan, mutta aiheet kiehtoutuvat myös vahvasti toisiinsa. Läpi artikkelin kulkee lisäksi pohdinta siitä, onko arvosidonnaisen sanaston objektiivinen kuvaaminen ylipäänsä mahdollista tai edes tarpeellista.

Arvosidonnaisen sanaston kuvaamiseen liittyvät kysymykset ovat erilaisissa sanakirjoissa osin erilaisia. Käsittelem aiheetta tässä artikkelissa ensisijaisesti yksikielisten yleiskielen selittävien sanakirjojen näkökulmasta. Niihin verrattuna kaksi- ja monikielisissä sanakirjoissa lisähaastetta arvosidonnaisen sanaston käsittelyyn tuovat esimerkiksi lähde- ja kohdekielen vastaamattomuudet. Vaikkapa politiikan sanastoon kuuluu paljon laajalle levinneitä kielestä toiseen lainautuneita kulkusanoja, jotka ovat eri kielissä ulkoisesti hyvin samankaltaisia mutta merkitykseltään tai vähintään merkitysvivahteiltaan erilaisia (esim. *propaganda*-sanan merkitykset eri kielissä). Erikoiskielten sanakirjoista monet paitsi kuvaavat olemassa olevaa sanastoa, myös antavat suosituksia uusiksi termeiksi. Sanakirjatyön ja sanastotyön yhdistyessä myös sanaston arvosidonnaisuutta on pohdittava yhtä aikaa kahdesta näkökulmasta: millaisille uusille sanoille on tarvetta ja miten sanat tulee kuvata sanakirjassa (esim.

lääketieteen sanastosta Saarela 2012: 8–9).

Arvosidonnaisen sanaston kuvaamisen kannalta on olennaista, tekevätkö sanakirjaa harrastelijat vai ammattilaiset. Esimerkiksi internetissä kaikkien täydennettävissä oleville sanakirjoille (esim. US, 2007–) ei voida asettaa samoja vaatimuksia kuin ammattimaisten leksikografien toimittamille sanakirjoille. Ammattimaisessa sanakirjatyössä sanakirjan tavoitteet ja esimerkiksi deskriptiivisyyden ja normatiivisuuden painotukset pyritään määrittelemään jo hankkeen suunnitteluvaiheessa (sanakirjan suunnittelusta ks. Kudashev tk.; deskriptiivisyydestä ja normatiivisuudesta ks. Grönros tk.). Suurissa hankkeissa toteutetut suomen kielen selittävät sanakirjat ovat luonteeltaan sekä deskriptiivisiä että normatiivisia (Eronen 2004; ks. myös Grönros tk.), vaikka Nykysuomen sanakirjan (NS, 1951–1961) laatimiseen alkusysäyksen antoikin tarve nimenomaan normatiiviselle sanakirjalle (Vesikansa 1977: 323–326). Harrastelijaleksikografille sanakirjan tavoitteiden määrittely ei välttämättä tule mieleen, ja ammattimaisissakaan hankkeissa tieto teoksen päämääristä ei välttämättä välity teoksen käyttäjälle asti (Vesikansa mts. 330).

NS ja sen seuraajat Suomen kielen perussanakirja (PS, 1990–1994) ja Kielitoimiston sanakirja (KS, 2004–2012¹) ovat monille suomalaisille vahvoja auktoriteetteja, joista katsotaan lopullinen totuus kielestä (Grönros tk. s. 50). Tämän vuoksi olen poiminut artikkeliini eniten esimerkkiaineistoa juuri suurissa kansallisissa hankkeissa tuotetuista suomen yleiskielen selittävästä sanakirjoista NS:sta, PS:sta ja KS:sta. Mukana on näyttöä myös muunlaisista sanakirjoista. Erilaisista sanakirjoista, eri aikakausilta ja eri kulttuureista poimituilla esimerkeillä voidaan havainnollistaa sitä, millaisia arvosidonnaisia valintoja myös nykypäivän suomalainen leksikografi joutuu tekemään.

Millainen sanasto on arvosidonnaista?

Aluksi on pohdittava, millainen sanasto ylipäänsä voi olla arvosidonnaista. *Arvolla* tarkoitan tässä tekstissä yksilön tai yhteisön käsityksiä tavoiteltavista asiantiloista ja hyväksytyistä toimintatavoista (myönteiset arvot) sekä

1 KS:aa on päivitetty versioiden välillä. KS:sta poimitujen esimerkkien kohdalla on mainittu, mistä versiosta esimerkki on peräisin.

vältettävistä asiantiloista ja ei-hyväksytyistä toimintatavoista (kielteiset arvot) (Häyry–Häyry 1997: 11). *Arvosidonnaisuudella* puolestaan tarkoitan jonkin asian riippuvuutta yksilön tai yhteisön arvoista. Näin ollen arvosidonnaiseen sanastoon kuuluvat sanat, joiden käyttö, merkitys tai konnotaatiot ovat jollakin tavalla sidoksissa kulloisenkin käyttäjän arvoihin.

Suurin osa sanastosta on perusmerkitykseltään neutraalia² ja melko helposti määriteltävää. Vaikka neutraaleillakin sanoilla on monille kielenkäyttäjille hieman eri merkitys, niiden käyttöä kuulijan omasta näkemyksestä poikkeavalla tavalla ei yleensä koeta loukkaavana. Sen sijaan neutraaliksikin tarkoitettu puhe itselle tärkeästä asiasta voi herättää närää ihmisissä, joilla on oma vahva mielipiteensä käytettyjen sanojen merkityksestä. Arvosidonnaista sanastoa kertyy erityisesti tiettyihin aihepiireihin. Attraktiokeskukset ovat samoja, joiden yhteydessä puhutaan usein poliittisesta korrektiudesta (vrt. Hughes 2009). Esimerkiksi monet politiikkaan, uskontoon ja muihin ideologioihin sekä sukupuoleen, syntyperään, ulkonäköön, yhteiskunnalliseen asemaan, ikään ja terveyteen liittyvät ihmisten ja ihmisryhmien nimitykset ovatkin leksikografille erityinen haaste, sillä niiden selittäminen kaikkia tyydyttävällä tavalla on usein vaikeampaa kuin neutraalin sanaston. Poliittinen korrektius ei sanakirjatyössä riitä, sillä selittävien sanakirjojen pitäisi pystyä kuvaamaan todellista kielenkäyttöä epäkorrektiuksineen kaikkineen niin, että sanakirjan käyttäjät saavat tietoa esimerkiksi ilmausten tyyliarvoista.

Sanojen arvosidonnaisuus korostuu silloin, kun vastakkainasettelut ovat vahvoja. Esimerkiksi Suomessa vuonna 1918 käytyyn sotaan liittyvät vahvat tunteet ja erilaiset tulkinnat heijastuvat myös sanakirjoissa. NS, PS ja KS kuvaavat sanaa *vapaussota* seuraavasti (esimerkit 1, 2 ja 3; vrt. myös *kansalaissota* ja *punakapina*):

2 Käytän tässä artikkelissa termiä *neutraali* viittaamassa sellaiseen sanastoon, jonka käyttö, merkitys tai konnotaatiot eivät riipu käyttäjän arvomaailmasta.

- (1) **vapaussota** sota, jonka jokin kansa käy vapautuakseen toisen kansan ylivallasta; **vars.** Suomessa vuonna 1918 käydystä sodasta. *Alan-komaiden vapaussota. Hallituksen joukkojen runkona olivat vapaussodassa suojeluskunnat.* (NS)³
- (2) **vapaussota** sota jota käydään valtiollisen vapauden saavuttamiseksi; **us.** Suomessa vuonna 1918 käydystä sodasta. *Suomen, Viron vapaussodan veteraanit.* (PS, KS 2004–2008)
- (3) **vapaussota** sota jota käydään valtiollisen vapauden saavuttamiseksi. *Viron vapaussota.* **Erik.** valkoisten näkökulmaa korostava nimitys Suomen kansalaissodasta. (KS 2012)

Leksikografit saavat tietoa yhteisön käsityksistä esimerkiksi yhteiskunnallisia keskusteluja seuraamalla. Koska vuoden 1918 sota on tiedetty tulenaraksi aiheeksi, sodasta käytettyjen nimitysten sanakirjaselitteisiin on kiinnitetty toimituksissa varmasti erityistä huomiota. (Sodan vaikutuksesta suomalaisen kielitieteeseen ks. Häkkinen 1993.) Sanan perusmerkitys on kuvattu kaikissa kolmessa sanakirjassa samansuuntaisesti. Sitä, että sana ei ole Suomessa usein käytetyssä merkityksessään kaikkien hyväksymä, on osoitettu NS:ssa, PS:ssa ja KS:ssa vuoteen 2008 asti elatiivilla. PS:n ja KS:n alkusanojen mukaan elatiivilla merkitään sanan ”epävirallisuutta”, NS:n esipuheessa mainintaa käyttötarkoituksesta ei ole. KS:n vuoden 2012 versioon on tehty muutos, jonka mukaan *vapaussota*-nimityksellä on valkoisen osapuolen näkökulmaa korostava erikoismerkitys.

Useimpia helposti arvolatauksia saavia sanoja voidaan käyttää myös neutraalisti. Arvosidonnaisiksi ne muuttuvat vasta, kun niihin liittyy jokin arvoja heijastava konnotaatio tai tarkoituserä. Silti jotkin sanat voivat olla tabusanoja, joiden käyttö ylipäänsä loukkaa. Esimerkki tästä on muutamien vuosien takainen keskustelu, jossa osa suomalaisista tahoi kaikenlaisen työn arvostamisen nimissä hylätä tai jopa kieltää työn ja

3 Vertailun helpottamiseksi olen muokannut kaikki esimerkit muotoon, jossa hakusana on lihavoitu, esimerkkilauseet kursivoitu ja mahdolliset tyyliarvomerkinnät yms. lihavoitu ja kirjoitettu muuta tekstiä pienemmällä fontilla. Esimerkin lähde on merkitty sulkuihin esimerkin perään. Tilan säästämiseksi käsiteltävän asian kannalta irrelevantteja tietoja esimerkiksi hakusanan sanaluokasta on jätetty pois. Osassa esimerkeistä olen käyttänyt harvennusta helpottamaan huomion kohdentamista. Joissakin esimerkeissä on esitetty tietoja useista samakantaisista tai -merkityksistä sanoista.

työntekijöiden arvoa alentavaksi kokemansa *paskaduuni*-sanana (ks. esim. Lindholm 2006; Ahoniemi 2010). KS:ssa (4) tämänkin sanan ”epävirallisuus” on huomioitu tyyliarvomerkintöjen lisäksi elatiivilla (ks. myös Kauppinen 2007).

- (4) **paskaduuni** ark., halv. = paskatyö. **paskatyö** ark., halv. vähäarvoisena pidetystä, huonosti palkatusta työstä, paskaduuni. (KS 2008–2012)

Samaan tapaan kuin melkein mitä tahansa sanaa voidaan käyttää neutraalisti, melkein mitä tahansa sanaa voidaan käyttää myös arvosidonnaisesti. Sana voi saada arvosidonnaisia konnotaatioita esimerkiksi metaforisessa käytössä. Pelkkä halu loukata on kuitenkin erotettava arvosidonnaisuudesta. Vaikkapa sana *keisari* päällepäsmäristä käytettynä voi loukata, mutta käyttötavalla ei välttämättä ole mitään tekemistä puhujan arvojen kanssa. Sen sijaan esimerkiksi sana *akka(mainen)* mieheen viittaavana haukkumasanana heijastaa myös naiseen ja miehiin liittyviä arvoja. Käyttötapa voi toki olla omaksuttu sellaisenaan ilman arvojen tiedostamista, mutta silloinkin se uusintaa arvosidonnaisia näkemyksiä siitä, että miehen naimaisuus ei ole hyväksyttävää ja että naiseus on alempiarvoista kuin mieheys. Ilmaukseen liittyy tietysti myös stereotypia naisista ja miehistä tietynlaisina.

Ihmisten nimitykset saavat helposti arvosidonnaisia merkityksiä, sillä niillä tehdään usein jakoa ”meihin” ja ”muihin”. Toimittaja Johanna Korhonen (2010) pahoittelee Helsingin Sanomien kolumnissaan sitä, että sanaa *suomalainen* käytetään nykyään, jollei eniten niin ainakin äänekäimmin, merkityksessä ’Suomessa syntynyt, valkoihoinen, suomea äidinkielenään puhuva ihminen, jonka sukujuuret ovat syvällä Suomessa’. Korhonen itse tahtois sisällyttää sanan merkitykseen myös muualla syntyneet, jotka syystä tai toisesta asuvat Suomessa ja haluavat olla suomalaisia. Suomen kielen sanakirja maahanmuuttajille (SSM 2010) ottaa kantaa asiaan *suomalainen*-hakusanan kohdalla seuraavasti (5):

- (5) **suomalainen** suomalainen on henkilö, joka asuu tai on asunut Suomessa ja ajattelee, että hän on suomalainen. *Minä olen suomalainen.* (SSM)

SSM antaa selitteessään lukijoille ikään kuin luvan laskea itsensä – niin halutessaan – suomalaisten joukkoon. Samankaltaiselle kohderyhmälle suunnatussa teoksessa Suomen kielen sanakirja ulkomaalaisille (SKSU, 1999) maiden nimistä muodostetut ihmistennimitykset selitetään kysei-

sessä maassa asumisen kautta (6 ja 7). Esimerkkilauseet kuitenkin kuvaavat myös toisenlaisia merkityksiä.

- (6) **suomalainen** 1. Suomen asukas *Nykyiset suomalaiset eivät kaikki muista edes talvisodan kärsimyksiä. Tapasimme Amerikan-matkallemme muutamia suomalaisia. Kyllä tavallisen suomalaisen tunnistaa ulkomailla jo kaukaa.* – – (SKSU)
- (7) **venäläinen** 1. Venäjän asukas *Suomessa asuu paljon venäläisiä.* – – (SKSU)

NS:ssa (8) *suomalaisen* ihmisen selite jaetaan kahtia: ihminen voi olla suomalainen joko kansalaisuudeltaan tai kansallisuudeltaan. NS:n mukaan Suomen kansalaisina suomalaisia voivat olla ihmiset syntyperään katsomatta, mutta etninen kansallisuus määrätty ensisijassa kielen kautta.

- (8) **suomalainen** 1. kansalaisuudesta: Suomen valtioon kuuluva. *Suomalainen mies. Suomalaisten urheilijain saavutuksen kansainvälisillä kilpailuilla. Suomalaista maatalousväkeä retkellä Tanskassa. Kaikki suomalaiset syntyperään ja kieleen katsomatta.* – – 2. syntyperää, rotua, kansallisuutta ajatellen: Suomen suomea puhuvaan pääväestöön kuuluva tai siitä lähtöisin oleva; suomenkielinen. *Maamme suomalainen ja ruotsalainen väestö. Lapin pitäjien suomalaiset ja lappalaiset asukkaat. Alkuaan suomalainen, mutta ruotsalaistunut suku. Mieleltään ja kieleltään suomalainen perhe.* (NS)

Myös KS:ssa (9) suomalaisiksi selitetään sekä kansalaisuudeltaan että kansallisuudeltaan suomalaiset, mutta kantaa siihen, kuka voi laskea itsensä suomalaiseksi, ei esimerkiksi SSM:n tapaan oteta.

- (9) **suomalainen** Suomeen kuuluva, Suomesta kotoisin tai lähtöisin oleva, Suomen asukkaille, oloille tms. ominainen; kansallisuudeltaan Suomen väestöön kuuluva; Suomen kansalainen. – – (KS 2004–2012)

Arvosidonnaisen sanaston kuvaamisessa on olennaista pohtia, kenen näkökulmasta sanastoa kuvataan ja kenen käsitykset sanojen merkityksistä otetaan vakavasti. Ihmisten nimitysten kuvaamiseen lisää haastetta tuo myös se, että vähemmistöryhmien alun perin neutraaleihinkin ni-

mityksiin liittyvät konnotaatiot muuttuvat nopeasti ja nimitykset pejoratiivistuvat herkästi (Grönros 1994: 13; Maamies 1995: 331–332; yleisesti pejoratiivistumisesta Kleparski 1986; Williamson 2009). Semanttiseksi arvonalennukseksikin kutsutulle pejoratiivistumiselle alttiita ovat myös naisten nimitykset (Maamies 1995: 332; Cameron 1996: 135–136). Viime vuosina vilkasta keskustelua on käyty esimerkiksi *maahanmuuttaja*-sanan ja alun perin sen neutraalina lyhenteenä koulutusohjelmien nimissä käytetyn *mamun* halventavuudesta. Varsinkin *mamun* sävy on joidenkin mielestä muuttunut halventavaksi, eikä sitä tulisi heidän mielestään enää käyttää neutraalissa yleiskielessä (ks. esim. internetartikkeliin Räsänen 2010 liittyvät kommentit). Ihmisoikeusliiton asiantuntijan Saido Mohamedin mukaan myös *maahanmuuttaja*-sanasta on tullut halventava, sillä ”se on alkanut tarkoittaa jotakin, mitä pitää sääliä tai hävetä” (Pakkanen 2011: 13).

Sanakirjoissa kynnys *maahanmuuttajan* kaltaisten läpinäkyvistä ja siinänsä neutraaleista elementeistä muodostettujen sanojen merkitsemiseen tyyliarvoltaan halventaviksi on korkea (vrt. myös *nais*-alkuiset yhdyssanat, ks. tarkemmin Grönros 2007: 15; Tyysteri 2011: 175, 177). Jos viittauksen kohteena olevat itse kuitenkin kokevat sanan halventavaksi, asian sivuuttaminen sanakirjassa voi olla arvosidonnainen viesti. Viime kädessä on kyse siitä, kenen kokemus kielen muutoksesta ja sanoihin liittyvistä uusista konnotaatioista otetaan huomioon, ja siitä, kenellä on oikeus päättää tietyistä ihmisryhmästä käytettävistä nimityksistä: ryhmään kuuluvilla itsellään vai jollakin toisella taholla. Esimerkiksi eduskunnassa vuonna 2011 (STT 2011) annetun kielenkäyttöohjeen mukaan istuntosalissa tulee puhutella puolueita ja niihin kuuluvia henkilöitä ”kirjakielen termein tai sellaisin lyhentein, joita ryhmät itse käyttävät”. Esimerkiksi nimityksiä *persu* ja *kepu* ei näin ollen hyväksytä eduskunnan virallisissa keskusteluissa. KS:ssa *kepulaisen* tyyliarvoksi on merkitty *leik.*, *halv.* ja uutena hakusana vuoden 2012 versioon otetun *persun* tyyliarvoksi *ark.*, *us. halv.*

Raja arvosidonnaisen ja neutraalin sanaston välillä on veteen piirretty, eikä toisen ihmisen henkilökohtaisiin poliittisiin, uskonnollisiin tai muihin vastaaviin kokemuksiin liittyvän sanaston tunnistaminen ole aina helppoa. Vanhan feministien iskulauseen mukaan myös henkilökohtainen on poliittista, ja joidenkin radikaalifeminististen suuntausten mukaan koko kieli on miesten luomaa ja hallitsemaa ja siten myös politisoitunutta (ks. Cameron 1996: 158–191). Äärimmäisiä esimerkke-

jä siitä, kuinka arkipäiväisiä ihmisten arvosidonnaisiksi kokemat sanat voivat olla, tarjoaa vaihtoehtoinen sanakirja *A Feminist Dictionary – In Our Own Words* (FD, 1985). Teos politisoi neutraalina pidettyä sanastoa tarkoituksenaan tehdä näkyväksi naisten kokemuksia. Esimerkkilauseina on käytetty lainauksia naisten kirjoittamista teksteistä. Kuten esimerkiksi 10 huomataan, sanojen arvosidonnaisuus tulee joskus esiin vasta, kun niitä tarkastellaan aivan uudenlaisesta näkökulmasta.

- (10) **face** A body part which with all its features and expressions forms a site for sex differentiation. *Males tend to use their faces in power plays; females use theirs in submissive acts. Many women perform such cosmetic acts on their faces as plucking eyebrows, adding artificial eyelashes, and wearing eye shadow, eye-liner, and mascara... because society defines so much on [women's] worth in terms of looks.* (FD)

Tietyinä aikana elävät ihmiset eivät aina edes voi havaita arvosidonnaisuutta, vaan se tulee näkyväksi vasta arvojen muuttuessa. Muiden painotuotteiden tapaan myös sanakirjat voivat kertoa meille monenlaisia asioita ajasta, jona ne on julkaistu ja jona niiden aineisto on kerätty. (Historiallisista sanakirjoista ks. Lehtosalo–Kuutti tk.)

Arvosidonnaisen sanaston kuvaaminen sanakirjoissa

Sanakirjan tulisi pystyä kertomaan sanakirjan kohderyhmään kuuluvalla käyttäjälle olennaiset tiedot sanojen merkityksistä. Perinteisesti tähän on pyritty selittämällä merkitykset esimerkiksi yläkäsitteen ja erottelevien käsittepiirteiden sekä synonyymien avulla. Muita käytettyjä keinoja ovat muun muassa alakäsitteiden luetteleminen sekä merkityksen selittäminen esimerkkilauseen tai vaikkapa kuvan avulla. Koska sanakirjahankkeissa käytetään usein pohjana aiempia sanakirjoja, samat merkityksen selitykset ja esimerkit saattavat toistua sanakirjasta toiseen.

Kielenkäyttäjät pystyvät soveltamaan prototyyppisiä kielenkäyttötapoja koskevaa tietoaan myös marginaalisiin tapauksiin. Varsinkin äidinkielisille aikuisille suunnatuissa yksikielisissä selittävässä sanakirjoissa riittävät näin ollen yleensä kuvaukset sanojen prototyyppisistä merkityksistä. Lisäksi on mahdollisuuksien mukaan tarjottava keinoja tulkita marginaalisia merkityksiä. Kaikkia mahdollisia merkityksiä sanakirjassa ei kuitenkaan ole tarpeen kuvata. (Atkins–Rundell 2008: 413–452.) Ar-

vosidonnaisen sanaston erityisen herkkyyden vuoksi sen eri merkitysten kuvaamista on pohdittava tarkasti. Myös kielen muutokseen on kiinnitettävä erityistä huomiota, sillä aiemmasta sanakirjasta poimitut ja päivittämättä jääneet arvosidonnaisten sanojen selitteet voivat vaikuttaa hyvin tökeröiltä. Muita arvosidonnaisen sanan selitteen rakentamisessa oleellisia seikkoja ovat esimerkiksi samaan kategoriaan kuuluvien sanojen johdonmukainen kuvaaminen sekä tasapainoilu todellista kielenkäyttöä kuvaavien mutta kuitenkin arvovapaiden esimerkkien välillä.

Esimerkiksi leimallisesti arvosidonnaiseen sanastoon kuuluvien poliittisten ideologioiden nimitysten selittäminen tiiviissä sanakirjaformaattissa ei ole helppoa. Yksi keino sanan merkitysten kuvaamiseen on määrittellä se vastakohtien tai lähikäsitteiden kautta. Englannissa julkaistu ensisijaisesti edistyneille kielenoppijoille suunnattu yleiskielen selittävä sanakirja Collins Cobuild English Dictionary (COBUILD, 1995) selittää sanat *capitalism* ja *socialism* seuraavasti (11 ja 12):

- (11) **capitalism** capitalism is an economic and political system in which property, business and industry are owned by private individuals and not by the state ...*the two fundamentally opposed social systems, capitalism and socialism.* (COBUILD 1995)
- (12) **socialism** socialism is a set of left-wing political principles whose general aim is to create a system in which everyone has an equal opportunity to benefit from a country's wealth. Under socialism the country's main industries are usually owned by the state. (COBUILD 1995)

COBUILD (1995) selittää kapitalismin järjestelmäksi, jossa tuotantovälineet ovat yksityisessä omistuksessa. Esimerkkilauseessa kapitalismin vastakohtaksi mainitaan sosialismi. Sosialismista puolestaan kerrotaan, että sen tavoite on luoda järjestelmä, jossa kaikilla on tasavertaiset mahdollisuudet hyötyä maansa vauraudesta. Tuotantovälineiden mainitaan olevan yleensä valtion omistuksessa. Kun COBUILDissa (1995) annetut kapitalismin ja sosialismin selitykset esimerkkilauseineen luetaan rinnakkain, syntyy käsitys, jonka mukaan sosialismi pyrkii ihmisten tasavertaisuuteen mutta kapitalismi sen vastakohtana tavoittelee päinvastaita. Tästä varsinkin kapitalismin kannattajat saattavat huomauttaa, että kapitalistiseen järjestelmään käsitetään yleensä kuuluvaksi esimerkiksi kaikkien ihmisten yhdenvertaisuus lain edessä sekä jokaisen tasavertaiset

oikeudet osallistua kaupankäyntiin vapailla markkinoilla (esim. Weber 2009 [1927]). COBUILDin (1995) selitteet kapitalismista ja sosialismista näyttävätkin kompastuvan tasavertaisuuden määrittelemättömyyteen.

COBUILDin (1995) selitteiden ongelma on myös niiden epäsymmetrisyys. Vaikka kapitalismi ja sosialismi mainitaan esimerkkilauseessa toistensa vastakohtiksi, niistä nostetaan esiin erilaisia asioita: tuotantovälineiden omistajuuden lisäksi sosialismista mainitaan myös sen oletettu päämäärä, kapitalismista sen sijaan ei. Toistensa vastakohtina esitettävien sanojen symmetriseen määrittelyyn on arvosidonnaista sanastoa käsiteltäessä kiinnitettävä erityistä huomiota. Näin sanakirjasta saadaan myös kokonaisuudessaan johdonmukaisempi. (Systemaattisuuden vaatimuksesta ks. myös Kudashev tk.)

Merkitysten selittäminen lähikäsitteiden kautta voi toisinaan johtaa kehäpäätelmiin ja sanakirjan käyttäjän vastuun kohtuuttomaan lisääntymiseen. Myös synonyymien kautta selittäminen sysää usein vastuun eteenpäin joko lukijalle tai synonyyminä mainitun sanan selitteelle. Käyttäjiltä on tietysti voitava odottaa sekä maailmantietoa että tietoa kielestä, mutta sanakirjaa toimitettaessa on kuitenkin pohdittava, minkä asioiden voidaan olettaa olevan valmiiksi tavallisen kielenkäyttäjän tiedossa ja minkä ei. Jos sanakirjantekijä ei itsekään varmasti tiedä, mistä on kyse, hän ei voi piiloutua epämääräisen selitteen taa ja vedota siihen, että muiden tulisi pystyä päättämään asia maailmantietonsa nojalla.

Esimerkiksi PS:n (13) selite *liberaalinen*-sanana merkityksestä politiikan kielessä vaatii lukijalta tietoa sosialismin ja konservatismien käsitteistä. Jos tietoa ei valmiiksi ole, todennäköisin paikka sen etsimiseen on valmiiksi kädessä oleva sanakirja (15 ja 16). Vaikka lukija voi etsiä tulkinnoilleen tukea myös muiden samakantaisten hakusanojen kohdalta (14), *liberaalisen* epätarkka ja ympäripyöreä selite jättää liikaa sanakirjankäyttäjän oman päättelyn varaan eikä täytä laadukkaana sanakirjaselitteen tavoitteita.

(13) **liberaalinen** 1. liberaali (2). *Liberaalinen teologia* liberaaliteologia. 2. pol. sosialismin ja konservatismien välistä linjaa noudattava. *Liberaalinen puolue*. (PS)

(14) **liberalismi** liberaali(su)us aatteena; yksilön vapauksien ja vapaan kilpailun merkitystä korostava pol. ja taloudellinen suuntaus. *Yhteiskunnallinen, taloudellinen liberalismi*. **liberaali** 1. subst. vapaamielinen

henkilö, liberalismien kannattaja; **pol.** liberaalisen puolueen (Suomessa Liberaalisen Kansanpuolueen) kannattaja. *Kulttuuriliberaali. Liberaalien puoluekokous. 2. adj. (rinn. liberaalinen) vapaamielinen, suvaitseva(inen). Kasvattajana liberaali. Liberaalit mielipiteet.* (PS)

- (15) **konservatiivi** konservatiivinen henkilö; **pol.** konservatismien kannattaja. *Liberaalit ja konservatiivit. konservatiivinen* vallitsevien olojen säilyttämistä kannattava, uudistuksia vierova, vanhoillinen. *Konservatiivinen kirjailija. Konservatiiviset mielipiteet. Pol.* konservatismia edustava. *Konservatiivinen poliitikko, puolue. konservatismi pol.* konservatiivinen aatesuunta, konservatiivisuus. (PS)

- (16) **socialismi** aatesuunta, joka pyrkii luomaan tuotantovälineiden yhteisomistukseen perustavan yhteiskuntajärjestyksen; sen mukainen järjestelmä. *Valtiosocialismi.* (PS)

Aatesuuntien määrittely on usein monitahoinen yhteiskunnallinen ja tieteellinen prosessi, joka ei välttämättä pitkänkään keskustelun jälkeen johda yksimielisyyteen. Suuntauksen kannattajat tai ylipäänsä sen hyvin tuntevat näkevät ideologiassa usein enemmän vivahteita kuin ulkopuoliset. Tyypillistä on, että esimerkiksi liberalismista tai feminismistä puhumisen sijaan peräänkuulutetaan puhetta useista liberalismeista tai feminismeistä (Gray 1989; Feminisms 1997). Yleiskielen sanakirjoihin tulevien selitteiden laatiminen ei vaadi yhtä tarkkaa käsitteiden määrittelyä kuin tieteelliset tutkimukset (erikoiskielten sanastotyö ja alan tieteellinen tutkimus sen sijaan voivat olla keskenään hyvin tiiviissä vuoro-vaikutuksessa). Kuitenkin myös yleiskielen sanakirjojen laadinnassa on otettava huomioon aiheita koskeva yhteiskunnallinen ja mahdollisuuksien mukaan myös tieteellinen keskustelu.

Internetissä kenen tahansa vapaasti muokattavissa olevan Wikipedia-tietosanakirjan virallisen käytännön mukaan sivuston artikkelit tulee kirjoittaa neutraalista näkökulmasta (*neutral point of view*). Neutraaliin näkökulmaan pyrkiminen ei tarkoita, että tietosanakirjan artikkelit voitaisiin tai tulisi kirjoittaa yhdestä puolueettomasta ja objektiivisesta näkökulmasta, vaan sitä, että asiaa koskeville kilpaileville näkemyksille annetaan tilaa suhteessa niiden merkittävyyteen luotettavissa lähteissä. Erilaisten näkemysten esittely jättää tietosanakirjan käyttäjälle tilaa omaan ajatteluun ja osoittaa, että kirjoittajat luottavat lukijoiden kykyyn muodostaa omat mielipiteensä. (Wikipedia EN: *Neutral Point of View*.)

Kieltä kuvaavat sanakirjat eivät ole ensyklopedioita, eikä niissä voida kuvailla yhtä sanaa loputtoman pitkästi. Toisin kuin tietosanakirjoissa, deskriptiivisissä sanakirjoissa pyritään yleensä kuvaamaan jonkinlaista yleistä käsitystä hakusanojen merkityksestä. Myös normatiivisten sanakirjojen lähtökohdan on oltava kieliyhteisön käsitys sen omasta kielestä, sillä normittamisen kohde on tunnettava ennen normien asettamista. Sanakirjatyö nojaa kieliaineistojen ja kielenkäytön havainnointiin, mutta kieliyhteisön yhteisen näkemyksen etsintää ohjaavat myös sanakirjan tekijöiden persoona ja ammattitaidon taso sekä ympäröivässä yhteiskunnassa vallitsevat käsitykset oikeasta ja väärästä. Koska ihmiset eivät ole koskaan samaa mieltä kaikista asioista, konsensuksen hakeminen ei välttämättä tee oikeutta kaikille kilpaileville näkemyksille. Myös leksikografiassa kannattaisi pohtia edellä esitellyn neutraalin näkökulman soveltamismahdollisuuksia varsinkin arvosidonnaisen sanaston kuvaamisessa.

Elektronisuus avaa monia uudenlaisia mahdollisuuksia sanastoa useita eri näkökulmista kuvaavien deskriptiivisten sanakirjojen luomiseen. Elektronisten sanakirjojen suurimmat edut lienevät rajoittamaton tila, päivitettävyys, mahdollisuus linkittämiseen sekä talkoistamisen tekninen helppous (talkoistamisesta esim. OED: Appeals; muista elektronisten sanakirjojen eduista ja mahdollisuuksista esim. Varantola 2001: 217, 235–236; Vanhatalo 2002; Schryver 2003; Shaposhnikova 2007; Dziemianko 2010). Sähköisiin sanakirjoihin voidaan painettuja paremmin sisällyttää muun muassa tarkkoja tietoja sanojen käyttötavoista ja tyyliarvoista. Oxford English Dictionary (OED, 2012) kuvaa elektronisessa versiossaan esimerkiksi sanan *nigger* tyyliarvoa tarkasti (17) (suomen *neekeri*-sanaa ovat käsitelleet mm. Maamies 1995 ja Kaartinen 2004).

- (17) **nigger** A. I. Senses referring to people 1. A dark-skinned person of sub-Saharan African origin or descent; This term is strongly racially offensive when used by a white person in reference to a black person. In written Black English and written representations of spoken Black English, however, there are usually not the same negative connotations. Recently the term has been reclaimed by some black speakers and used with positive connotations in various senses (esp. in the form *nigga*: see note in etymology, and senses A. 1c, A. 4, and A. 5). However, even among black speakers, use of the word is problematic because of its potential to give offence, as is clear from the following – – (OED)

Tyyliarvoa koskevien tietojen lisäksi OED:ssä esitellään *nigger*-sanan useita erilaisia käyttötapoja ja annetaan niistä lukuisia esimerkkejä tarkkoine lähdetietoineen. Esimerkkilauseet ovatkin yleensä olennainen osa hakusanojen kuvausta sanakirjoissa. Esimerkkien tulisi kuvata sanojen tyypillistä käyttöä ja tavallisimpia käyttöyhteyksiä. Hyvinä esimerkkilauseina on pidetty muun muassa kielelle tyypillisiä sanontoja ja vakiintuneita sanaliittoja (Zgusta 1971: 263–268; Kivekäs 1972: 467). Kielessä pitkään säilyvät sanonnat ja sananparret saattavat kuitenkin heijastaa sellaista maailmankuvaa, joka ei välttämättä vastaa enää nyky-yhteiskunnan arvoja. Esimerkiksi käsitykset sukupuolten tasa-arvosta ovat muuttuneet lyhyessä ajassa niin, että aihepiiriin liittyvien selitteiden ja esimerkkilauseiden tarkistaminen erityisen huolellisesti on paikallaan aina, kun sanakirjaa päivitetään (vrt. ns. kummitussanat ja -merkitykset, ks. Pilke–Puskala tk. s. 159–162). Myös kaunokirjallisuudesta poimittujen esimerkkien käyttö edellyttää varovaisuutta, sillä monet niistä asettavat turhaan lukijalle vaatimuksia erilaisten yhteiskunnallisten diskurssien tai jopa yksittäisten teosten sisällön tuntemisesta (18 ja 19).

(18) **kapitalisti** pääoman omistaja, varakas henkilö, rikas, porvari; kapitalismin kannattaja – – *Täällä hän kuuli – –, että kaikki kapitalistit olivat rosvoja ja että rikkaudet olivat kansan hiellä kootut.* LEINO. (NS)

(19) **juutalainen** 1. *Juutalaiset* muinaisista israelilaisista polveutuva, omaleimaisena säilynyt kansa, jolla on yhteinen uskonto. – – *Jeesus Nasaretilainen, juutalaisten kuningas* UT. – – (KS 2004–2012)

Esimerkkien lähdetietojen merkitseminen sanakirjaan parantaa käyttäjän mahdollisuuksia arvioida lauseiden arvosidonnaisuutta. Usein lähteiden merkitseminen ei kuitenkaan ole käytännössä mahdollista, sillä leksikografit myös muokkaavat esimerkkejä ja keksivät niitä itse (Svensén 2004: 345–346).

Arvosidonnaista sanastoa voidaan pyrkiä kuvaamaan neutraalisti kiisittomista faktoista kertovilla lauseilla. Esimerkiksi konflikteja kuvattaessa neutraaliutta voidaan vahvistaa vaikkapa kertomalla maantieteellisesti tai ajallisesti kaukaisista tapahtumista tai jopa muinaisista myyteistä (esimerkit 20 ja 21). Kaikkein neutraaleimmiksi esimerkit saadaan, jos niistä häivytetään konfliktin osapuolet kokonaan (jälkimmäinen esi-

merkkilause esimerkissä 20). Samalla esimerkit voivat kuitenkin muuttua niin abstrakteiksi, että ne eivät enää kuvaa kovin hyvin sanan merkitystä tai tyypillistä käyttöä (vrt. pelkästään välttämättömistä kieliopillisista elementeistä koostuvat ns. kuolleet esimerkit, Svensén 2004: 345–346; Pilke–Puskala tk. s. 166).

(20) **vallutussõda** teiste maade ja rahvaste alistamiseks peetav sõda. *Napoleoni, Rooma impeeriumi vallutussõjad. Riik pidas valduste suurendamiseks vallutussõdu.* (EKSS 2009)

(21) **sotajuoni** vars. kuv. juoni. *Troiija kukistettiin ovelan sotajuonen avulla. Punoivat sotajuonia vaalien voittamiseksi.* (KS 2004–2012)

Neutraalien esimerkkilauseiden löytäminen ei aina ole mahdollista, mutta onko ehdoton neutraalius edes tavoiteltavaa? Jos sanakirjassa käytetään vain neutraaleja esimerkkejä ja selitteitä, sanan käytöstä annetaan helposti vääränlainen ja todellisuutta paljon harmaampi kuva. Yksi kielen keskeinen tehtävä on erilaisista arvoista kertominen, joten oleellisempaa olisi tiedostaa sanakirjojen olevan arvosidonnaisia siinä missä yhteiskunnan muutkin tekstituotteet.

Arvosidonnaisen sanaston kuvaustapojen syyt ja seuraukset

Selittävien yleiskielen sanakirjojen päämäärä on kuvata kieltä sellaisena kuin kieltä käyttävän yhteisön jäsenet sen näkevät. Samalla esimerkiksi NS:n, PS:n ja KS:n tarkoitus on antaa myös normatiivisia ohjeita kielenkäytöstä. (Deskriptiivisyydestä ja normatiivisuudesta ks. Grönros tk.). Oikeinkirjoituksen lisäksi sanakirjat normittavat myös merkityksiä ja rajaavat mahdollisina pidettyjä merkitystulkintoja. Kolmas usein sanakirjatyön tärkeäksi päämääräksi mainittu tehtävä on kielen sanavarojen dokumentointi. Aikojen saatossa sanakirjoille on asetettu myös muita tavoitteita.

Se, tehdäänkö kielestä ylipäänsä sanakirjoja ja millaisia tehdyt sanakirjat ovat, heijastaa kielen arvostusta. Kansallisvaltioajatuksen nousun aikaan 1800-luvulla suomen kielen sanakirjat olivat tärkeä osa kielen kehitystyötä ja samalla väline kielen ja sitä kautta myös kansakunnan

arvostuksen nostamiseen. Oman kielen sanavaroja ja murteita pidettiin tärkeänä osana sitä pohjaa, jolle suomalaisten kansallista identiteettiä tahdottiin rakentaa. Sanakirjatyöllä oli siis selvä poliittinen päämäärä. Suomalaiselle sanastotyölle on pidetty ominaisena purismia, ja esimerkiksi Elias Lönnrot suhtautui lainasanoihin kielteisesti (Pitkänen 2008: 78–82, 275–276). Lönnrotin ja monien hänen aikalaistensa purismi ei liittynyt vain heidän oikeakielisyyksityksiinsä, vaan ensisijaisesti kyseessä oli arvosidonnainen näkemys, jonka mukaan suomen kielen arvon nostaminen on yhteiskunnallisesti tärkeää. Jokaisen (uuden) suomen kielen sanan ottaminen sanakirjaan oli näin ollen arvosidonnainen valinta.

Tärkeä tavoite Lönnrotin ja hänen aikalaistensa purismin taustalla oli myös kielellisen yhdenvertaisuuden lisääminen. Lönnrotin mukaan kieleen luotavan uuden sanaston tuli olla ymmärrettävää myös niille, jotka eivät taitaneet vieraita kieliä. (Pitkänen 2008: 78–82, 275–278.) Nykyäänkin suomalaisessa sanastotyössä pyritään omakielisyyteen aina, kun se vain on mahdollista. Tekemällä suomen kieltä kuvaavia sanakirjoja ja kehittämällä omakielistä termistöä voidaan osaltaan huolehtia siitä, että tieteellisen tiedon ääreen päässeiden ja ulkopuolelle jääneiden välille ei synny kuilua (Suomen kielen tulevaisuus 2009: 132–138). Kielellisen yhdenvertaisuuden takaaminen on myös yksi nykyaikaisen suomalaisen sanakirja- ja sanastotyön kielipoliittisista päämääristä.

Yleiskielen selittävässä sanakirjoissa kuvataan kielen keskeisten sanavarojen lisäksi usein myös sellaista erikoiskielen sanastoa, joka on taajaan käytössä joukkoviestimissä, sekä sellaista harvinaisempaa sanastoa, josta kielenkäyttäjät erityisesti tarvitsevat tietoa (esim. KS: alkusanat). Kuitenkaan minkään kielen kaikkien sanojen kaikki merkitykset eivät mahdu yhteen sanakirjaan. Erilaisia käyttäjäryhmiä varten tarvitaankin erilaisia sanakirjoja. Esimerkiksi maahanmuuttajille suunnatut sanakirjat parantavat kielenoppijan mahdollisuuksia toimia yhteiskunnan jäsenenä. Erikoiskielten sanakirjoja puolestaan tarvitsevat sekä alan ammattilaiset että termien merkityksestä tietoa etsivät maallikot. Kaikki sanakirjat voivat osaltaan lisätä ja pitää yllä kielellistä yhdenvertaisuutta tarjoamalla ihmisille relevanttia ja huolella tuotettua tietoa kielestä. Se, että erilaisten sanakirjojen tarve sekä eri käyttäjäryhmien erilaiset tarpeet otetaan vakavasti, on kielipoliittinen valinta.

Sanakirjatyön tavoitteiden lisäksi sanakirjasta heijastuviin arvosidonnaisuuksiin vaikuttavat ympäriöivässä yhteiskunnassa vallitsevat arvot

sekä leksikografian oma persoona. Omien käsitysten mukaista arvosidonnaista selitettä on vaikeampi huomata kuin omien arvojen kanssa ristiriidassa olevaa. Vähemmistöön kuuluva kielenkäyttäjä huomaa itseensä kohdistuvan syrjivän kielenkäytön enemmistöön kuuluvia helpommin. Naiset todennäköisesti huomaavat naisia sukupuolen perusteella syrjivän kielenkäytön helpommin kuin miehet.

Esimerkiksi yhteiskuntamme valtakulttuurissa vaikuttava automaattinen oletus ihmisen heteroseksuaalisuudesta (enemmän aiheesta esim. Lehtonen 2003) näkyy myös suomalaisissa sanakirjoissa, joissa vaikkapa seurustelu selitetään yleensä miesten ja naisten väliseksi. Hetero-olettamuksen siirtyminen sanakirjoihin on luonnollista, sillä useimmat sanakirjat pyrkivät kuvamaan juuri kaikkein tyypillisimpiä tapoja käyttä kieltä. Sanojen tyypillisten merkitysten ja käyttötapojen kuvaaminen on kuitenkin huomattava erottaa puhetavasta, joka sulkee osan ihmisistä kielen ulkopuolelle (ks. myös Tyysteri 2011: 176–177).

NS:n julkaisemisen aikaan homoseksuaalisuus oli Suomessa rikos, ja seurustelu kuvataankin NS:ssa (22) vain joko miehen ja naisen tai tyttöjen ja poikien väliseksi. Monissa uudemmissa sanakirjoissa myös muunlaiset seurustelusuhteet on sisällytetty määritelmään joko jättämällä seurustelijoiden sukupuoli tarkentamatta (23) tai mainitsemalla seurustelijoina myös ei-heterot (24).

- (22) **seurustella** – 2. erik. a. miehen ja naisen välisistä läheisistä, tav. eroottisista suhteista. *Tyttö, joka seurustelee monen pojan kanssa. Heikkiläkin, vanha mies, rupesi juomaan ja seurustelemaan kaikenlaisten tyttöjen kanssa* TALVIO. *Siihen aikaan kun Niemisen Kalle jo julkisesti seurusteli tulevan vaimonsa kanssa* SILL. – – **seurustelu** < ed. – – Tyttöjen ja poikien välinen seurustelu (NS)
- (23) **seurustella** – 2. olla vakituksessa [vars.] rakkauteen perustuvassa kanssakäymisessä jkn kanssa (ennen kihlausta tai avio- tai asuinliittoa [tai rekisteröityä parisuhdetta]) *Seurustella vakituksesti. Seurustelee erään kurssitoverinsa kanssa. Seurustelimme pari vuotta ennen kihlausta.* – – (PS, KS 2004–2008 [hakasulkeissa olevat lisäykset KS 2012])
- (24) **seurustella** – 2. (miehen ja naisen) eroottisesta suhteesta: *Pekka ja Liisa ovat ruvenneet julkisesti seurustelemaan.* **seurustelu** – 2. eroottisesta suhteesta: *Tyttöjen ja poikien välinen seurustelu alkaa puberteetti-iässä. Katselin ravintolassa homojen seurustelua.* (SKSU)

PS:ssa ja KS:ssa (23) seurustelun syyksi tai taustavaikuttimeksi mainitaan rakkaus. KS:n vuoden 2012 versio jättää myös muiden motiivien mahdollisuuden avoimeksi, sillä siihen tehty lisäys selittää seurustelun olevan ”varsinkin rakkauteen perustuvaa”. NS:ssa (22) seurustelu on määritelty tavallisesti eroottiseksi. Myös SKSU (24) esittää seurustelun perusteena tai siihen olennaisesti liittyvänä asiana eroottisuuden. Eroottisuuden mainitseminen on keino erottaa seurustelun tämä merkitys toisesta, ylipäänsä kanssakäymistä tarkoittavasta merkityksestä. Erityisesti SKSU:n selitteessä huomattavaa on, että teos on suomalaisen kulttuuri-
piirin ulkopuolelta tuleville suunnattu sanakirja. Teoksen kuvaus seurustelusta vastaa kenties tavanomaista suomalaista seurustelua (tosin ks. esim. Salomäki 2004; Aseksuaalit.net), mutta ei jätä avioliittoa edeltäviin suhteisiin kuuluvan eroottisuuden kieltävistä kulttuurista tuleville ihmisille tilaa kertoa suomeksi omasta elämästään.

Ammattieettisesti toimivan ja laadukkaaseen sanakirjaan tähtäävän leksikografian tulee taustastaan riippumatta varmistaa, että hänen omat arvonsa eivät pääse tahattomasti heijastumaan lopputulokseen. Erilaisten arvomaailmojen huomioimisen kannalta paras vaihtoehto on sanakirjan toimittaminen ryhmässä. Varsinkin suurissa hankkeissa sanakirjoja tehdään usein kollektiivisesti. Kommentteja on mahdollista pyytää myös varsinaisen työryhmän ulkopuolelta (esim. Nivanka–Airila–Soini–Penttilä 1930).

Kaikille kielentutkijoille ovat varmasti tuttuja maallikoiden kysymykset siitä, onko jotakin tiettyä sanaa ”olemassa”. Monille ihmisille sanan puuttuminen sanakirjasta riittää todisteeksi siitä, että sanaa ei ole tai se ei ole osa ”virallista” kieltä. Joillekin sanan puuttuminen sanakirjasta riittää todisteeksi myös siitä, että sanalla kuvattua ilmiötä ei ole olemassa (Kolehmainen 2005). Sanakirjojen ja muiden hakuteosten auktoriteetti-
asema on huomattu myös totalitaarisissa hallintomalleissa, joissa sanakirjoja on pyritty käyttämään jopa propagandan välineenä. Vaikka sanan puuttuminen sanakirjasta ei tietenkään poista sanalla kuvattua ilmiötä maailmasta, ei-toivottujen ilmiöiden hävittämistä on ajateltu voitavan edesauttaa kieltämällä aiheesta puhuminen ja jättämällä siihen liittyvät sanat pois sanakirjasta. Esimerkiksi lännessä totalitaristisena pidetyn Neuvostoliiton aikaisissa venäläisissä tietosanakirjoissa ja poliittisissa sanakirjoissa ei ollut 1930- ja 1940-luvuilla lainkaan mainintoja sanasta *totalitarismi*. 1950-luvulta alkaen sanan vastineeksi esitettiin *fasismi*.

(Bergman⁴ 1998: 250.) Tiettyihin aihepiireihin liittyvän sanaston tuomista kieleen on voitu myös pitää niin vaarallisena, että se on päätetty kieltää. Esimerkiksi Kongon demokraattisessa tasavallassa tasa-arvosanastoa tsiluban kielelle kääntävä sanasto *Terminologie de l'élimination de la discrimination à l'égard de la femme* (2001) on hallituksen kieltämä. Sanaston tekijä, kielentutkija-kääntäjä Gilbert Adi Kadima on joutunut pakenemaan kotimaastaan. (Nikkanen 2007, 2009.) Totalitaarisissa hallintomalleissa tuotetuissa sanakirjoissa sanastoa on voitu myös selittää vahvan värityneesti. Esimerkiksi neuvostoaikaisessa venäjän kielen uudissanakirjassa *Новое в русской лексике* (*Novoe v russkoj leksike*, 1980) ilmausta *узники совести* (*uzniki sovesti* 'mielipidevangit') ei ole selitetty suoraan, mutta yksi esimerkeistä esittää mielipidevangit Amerikan valanpitäjien uhreina, joita vainotaan heidän poliittisten näkemystensä tai edistykseellisen julkisen toimintansa vuoksi (Rothstein 1985: 467).

Sanakirjoilla on vaikutusvaltaa myös suomalaisessa yhteiskunnassa (ks. esim. Vanhatalo 2005a: 2). KS:n päätoimittajan Eija-Riitta Grönrosin mukaan (2011) sanakirjat ovat joillekin käyttäjille melkeinpä lakikirjaan vertautuvia auktoriteetteja (ks. myös Varantola 2001: 218–221). Hakuteosten auktoriteettiaseman vuoksi niiden viesteille saatetaan olla jopa alttiimpia kuin esimerkiksi median vaikutukselle (Smith 1985: 37; Cameron 1996: 140–145, 171). Sanakirjojen antamia merkityksenselityksiä käytetään useissa tieteellisissä tutkimuksissa, opinnäytteissä ja oppikirjoissa (muiden lähteiden ja oman pohdinnan rinnalla) varsinkin lähtökohtana käsiteltävän aiheen määrittelylle. Myös monet Wikipedian artikkelit alkavat sanakirjasta poimitulla määrittelmällä (esim. Wikipedia FI s.v. *muilutus*, *fasismi*, *totalitarismi*, *uussuomalainen*). Sanakirjojen sisältöä siis toistetaan osana muuta tiedontuottamista. Sanakirjoihin vedotaan myös lausunnoissa ja muissa virallisissa teksteissä (esim. Grafia; Arktinen keskus) sekä toisinaan varsin vakavissakin henkilökohtaisissa asioissa. Seuraava lainaus (25) on internetissä vapaasti luettavissa olevasta Jehovan todistajien haaratoimistokomitealle osoitetusta, seurakunnan jäsenen erottamista koskevasta kirjeestä (Johanneksenpoika.net):

4 Bergman ei tee artikkelissaan eroa kieltä kuvaavien sanakirjojen ja ensyklopedisten tietosanakirjojen välille.

(25) Syyte oli Jehovan todistajien herjaaminen, joka kuitenkin osoittautui perusteettomaksi, koska kukaan ei pystynyt osoittamaan yhtään kohtaa, jossa olisin esittänyt perättömiä ja valheellisia tietoja Jehovan todistajien opetuksista, jota herjaaminen *Nykysuomen sanakirjan* mukaan tarkoittaisi. Niinpä syyte muutettiin Jehovan todistajien mustamaalaamiseksi.

Leksikografin näkökulmasta sitaatissa on asian vakavuuden lisäksi huomattavaa se, että tavalliset kielenkäyttäjät eivät välttämättä tiedosta kielen muutosta ja sanakirjojen päivitystarvetta. Varsinkin NS:n uudempiin mutta muuttamattomiin painoksiin vedotaan edelleen nykykieltä kuvaavina auktoriteetteina (ks. myös Grönros 2011).

Lopuksi: Sanakirjojen arvosidonnaisuudesta on syytä keskustella

Olen tässä artikkelissa käsitellyt arvosidonnaista sanastoa ja sen kuvaamista yleiskielen selittävässä sanakirjoissa. Olen lähestynyt aihetta kolmen pääkysymyksen kautta: millainen sanasto ylipäänsä voi olla arvosidonnaista, miten arvosidonnaista sanastoa kuvataan sanakirjoissa ja millaisia syitä ja seurauksia valituilla kuvaustavoilla voi olla. Tarkastelussa keskeisiksi seikoiksi ovat nousseet yhteiskunnan todellinen moniarvoisuus, kielen nopea muutos sekä leksikografin eettinen ja yhteiskunnallinen vastuu.

Monien muiden tekstituotteiden tapaan sanakirjat ottavat kantaa yhteiskunnallisiin, poliittisiin ja muihin arvosidonnaisiin aiheisiin ja välittävät eteenpäin tietynlaista maailmankuvaa. Siksi arvosidonnaiseksi arvellun sanaston kuvaamiseen tulee kiinnittää sanakirjatyössä erityistä huomiota. Yksi keskeisistä ongelmista on, että toisen ihmisen arvoihin liittyvän sanaston tunnistaminen on toisinaan vaikeaa. Leksikografeilla ja muilla kielen asiantuntijoilla tulisi kuitenkin olla erityistä herkkyyttä huomata tällaiset sanat (vrt. esim. Wierzbicka 2010). Nykykieltä kuvaava leksikografi katsoo kieltä tietysti aina oman aikansa näkökulmasta, ja toisinaan sanojen ja niiden selitteiden arvosidonnaisuus voi paljastua kie-liasantuntijallekin vasta ajan kuluessa ja arvojen ja yleisten näkemysten muuttuessa.

Perinteinen sanakirjaformaatti tiivistää valtavasti tietoa pieneen tilaan, ja eri näkökulmien selittämisen sijaan sanakirjoissa on yleensä

pyrityt jonkinlaisen yleisen ja yhteisen totuuden etsintään. Aidosti moniarvoisessa yhteiskunnassa leksikografian tärkeä tehtävä erityisesti arvosidonnaisen sanaston kohdalla on valita, kenen käsitys kielestä lopulta asetetaan normiksi. Varsinkin elektronisten sanakirjojen tarjoamia mahdollisuuksia selittää arvosidonnaista sanastoa yhtäaikaisesti useista eri näkökulmista olisi punnittava tarkoin. Päätettäessä sanaston kuvaustavoista on kuitenkin otettava huomioon ensisijaisesti käyttäjien tarpeet (Atkins–Rundell 2008: 413–452). Monet ihmiset odottavat sanakirjoilta normittavuutta, eivätkä useat vaihtoehtoiset selitteet välttämättä vastaa käyttäjien toiveita.

Mahdollisimman hyvän kielenkuvauksen lisäksi sanakirjatyölle nykyään yleisesti asetettu päämäärä on kielellisen yhdenvertaisuuden lisääminen ja takaaminen. Kielellistä yhdenvertaisuutta on esimerkiksi se, että ihmisillä on yhtäläiset mahdollisuudet saada tietoa kielestä. Kielelliseen yhdenvertaisuuteen kuuluu myös syrjivän kielenkäytön välttäminen. Koska osa todellisesta kielenkäytöstä kuitenkin on syrjivää, kielenkuvauksen tarkkuus ja kielellinen yhdenvertaisuus voivat tuntua ristiriitaisilta päämääriltä.

Sanakirjatyössä joudutaan aina pohtimaan myös tulevaisuutta (mm. uusien sanojen vakiintuminen kieleen). Muutoksessa kohti yhdenvertaisempaa ja tasa-arvoisempaa yhteiskuntaa on keskeistä toisenlaisista mahdollisuuksista tarjottava tieto (Smith 1985: 169–181). Yksi tapa reagoida kielen muutokseen voi olla pyrkimys kielellisen yhdenvertaisuuden edistämiseen vaihtoehtoisista kielenkäyttötavoista kertomalla. Sanakirjoissa on mahdollista esimerkiksi kertoa, miten voidaan välttää syrjiviä kielenkäyttötapoja ja tukea alistettujen ja syrjittyjen ryhmien omaa voimaantumista (vrt. syrjintä ja positiiviset erityistoimenpiteet, esim. Syrjinnän vastainen käsikirja 2003: 10–12, 72–74). Esimerkki tällaisesta voisi olla vaikkapa aiemmin geneerisinä pidettyjen *mies*-loppuisten yhdyssanojen rinnalle spontaanisti ilmestyneiden *nais*-loppuisten yhdyssanojen (esim. *lakimies*, *lakinainen*) ottaminen yleiskielen sanakirjaan. Kaiken kaikkiaan arvosidonnaista sanastoa käsiteltäessä olennaista on leksikografien kyky reagoida kielen muutokseen. Käytännössä reagointimahdollisuuksiin vaikuttavat esimerkiksi sanakirjahankkeen kesto ja sanakirjan julkaisuformaatti.

Monet artikkelissa kuvatuista ongelmista voitaisiin välttää, jos leksikografit ja sanakirjojen käyttäjät ymmärtäisivät sanakirjojen funktion sa-

malla tavalla. Käyttäjäkoulutuksista on todettu olevan selvää hyötyä käyttäjien sanakirjatietoisuuden parantamisessa (esim. Varantola 2001: 233), ja nähdäkseni leksikografien olisikin panostettava entistä enemmän erilaisista sanakirjoista ja niiden funktioista tiedottamiseen sekä käyttäjien opastukseen ja koulutukseen. Hyvä olisi pyrkiä kertomaan yleistajuisesti myös kielen muutoksesta. Erityinen haaste tässä työssä on kieliyhteisön mielenkiinnon herättäminen, ja tärkeää olisikin esimerkiksi tulevien opettajien tavoittaminen jo opiskeluvaiheessa. Arvosidonnaisen sanaston kannalta olennaista on, että leksikografit ovat valmiita kertomaan avoimesti työssään tekemistään valinnoista ja niiden taustoista. Esimerkiksi sanakirjojen johdannoissa olisi syytä kertoa eksplisiittisesti, onko kyse deskriptiivisestä vai normatiivisesta sanakirjasta ja mille kaikille kielen alueille normatiivisuus teoksessa ulottuu (ks. Grönros tk.).

Sanakirjojen arvosidonnaisuutta ei ole toistaiseksi tutkittu paljon, ja varsinkin leksikografisesta näkökulmasta julkinen keskustelu aiheesta on vasta alussa. Tässä artikkelissa aihetta on käsitelty ensisijaisesti nykykieltä kuvaavien yksikielisten yleiskielen selittävien sanakirjojen näkökulmasta. Monet arvosidonnaisen sanaston kuvaamiseen liittyvät kysymykset ovat samantyyppisiä erilaisissa sanakirjoissa. Kuitenkin myös eroja on. Esimerkiksi lapsille suunnatut sanakirjat (kuvasanakirjoista lähtien) olisivat mielenkiintoisia tutkimuskohteita aivan erityisen auktoriteettiasemansa vuoksi. Ylipäänsä erilaisille kohderyhmille suunnattujen sanakirjojen vertailu sekä käyttäjälähtöisestä että yhteiskunnallisesta näkökulmasta avaa monia mielenkiintoisia tutkimusnäkökulmia. Myös sanakirjojen yhteiskunnalliset vaikutukset ovat vielä pitkälti tutkimatta, vaikka niistä osoituksia onkin. Erilaisten sanakirjojen välittämän maailmankuvan tutkimuksessa on jo päästy alkuun, mutta uusia aiheita riittää vielä pitkäksi aikaa. Yksittäisten sanakirjojen maailmankuvaa tutkittaessa on muistettava ottaa huomioon, onko kyse deskriptiivisestä vai normatiivisesta sanakirjasta, kuinka laajasta tai suppeasta sanakirjasta on kyse, milloin sanakirja on julkaistu, minkä ajan sanastoa teos kuvaa (Grigoryan 2007: 4, 6–7), millaiselle kohderyhmälle teos on suunnattu ja ketkä sen tekijöitä ovat.

Arvot herättävät suuria tunteita. Siksi niihin liittyvän sanaston selittäminen on herkkä ja vaativa tehtävä. Osa artikkelissa esitetyistä ajatuksista voi tuntua kohtuuttomilta vaatimuksilta leksikografeja kohtaan. Artikkelin tarkoitus ei kuitenkaan ole vaatia, syyllistää tai antaa valmiita vastauksia, vaan herättää ajatuksia ja keskustelua. Toivon tämän artikkelin myös

havahduttavan aiempaa useammat sanakirjojen käyttäjät huomaamaan, että sanakirjat eivät edusta lopullista totuutta, josta ei olisi oikeutta olla eri mieltä. Muille kuin leksikografeille tämä artikkeli toivottavasti konkretisoi myös merkitysten selittämisten vaikeutta ja lisää sitä kautta huolella tehdyn leksikografisen työn arvostusta.

NINA PILKE & JAANA PUSKALA

Kummitussanat ja kummalliset esimerkit käyttäjänkin kontolla? Sanakirjan käyttäjän etiikka osana kieliasiantuntijuutta

Artikkelimme käsittelee sanakirjoja ja etiikkaa sanakirjojen käyttäjien näkökulmasta. Aloitamme pohtimalla sanakirjan ja sanakirjan käyttäjän käsitteitä sekä tarkastelemalla niitä suhteessa tutkimaamme aineistoon. Siirrymme käsiteanalyysin kautta tarkastelemaan sana-artikkelien sisältöä erityisesti esimerkkien kautta ja lopuksi peilaamme havaintojamme käyttäjän etiikkaan. Artikkelimme perustuu yksikielisiin suomen- ja ruotsinkielisiin sanakirjoihin sekä kaksikielisiin suomi–ruotsi–suomi-sanakirjoihin, jotka ovat meille ruotsin kielen tutkijoina ja opettajina tärkeimpiä työkalujamme. Pohdintojamme voi soveltaa kuitenkin myös muunkielisiin sanakirjoihin.

Sanakirjat ja niiden käyttäjät

Kun suomalaiselta kysyy sanakirjan määritelmää, useimmat vastaavat sanakirjan olevan yhtä kuin kaksikielinen painettu sanakirja. Tämä on tyypillistä suomalaisille, koska meillä ei juuri ole perinteenä hyödyntää oman kielen yksikielisiä selittäviä sanakirjoja sanojen merkityksen tai kirjoitustavan selvittämiseen kuten monen muun kielen puhujat tekevät (ks. esim. Kipfer 1987: 45). Lisäksi monet satunnaiset sanakirjan käyttäjät uskovat, että etenkin laajemmat kaksikieliset sanakirjat kattavat molempien kielten koko sanavaraston ja niissä annetut hakusanojen vastineet ovat hyväksytty tapa kääntää hakusana toiselle kielelle tilanteessa kuin

tilanteessa. (Snell–Hornby 1987: 159–160; vrt. Grönros 2011.) Aiemmat tutkimuksemme (Pilke 2008) tukevat osaltaan tätä kuvaa: yliopisto-opiskelijat sanovat luottavansa eniten painettuun sanakirjaan.

Käyttäjät: aloittelijoista ammattilaisiin

Sanakirjojen käytön ja käyttäjien empiirinen tutkimus on saanut jaloisijaa leksikografiassa vasta parin viime vuosikymmenen aikana (ks. esimerkiksi Tarp 2008: 5–7). Aiemmin sanakirjojen laatimista ohjasivat lähinnä sanakirjan toimittajien omat käsitykset sanakirjan käyttäjästä, hänen taidoistaan ja tarpeistaan, mutta nykyisin pidetään itsestään selvänä, että sanakirjoja käyttävät monet erilaiset käyttäjät monin eri tavoin moniin erilaisiin tarkoituksiin monissa eri tilanteissa (ks. Svensén 2004: 533).

Sanakirjojen käyttäjät voidaan luokitella eri tavoin. Yksi käyttäjäryhmä käsittää tulevat kieliasiantuntijat eli kieliaineiden yliopisto-opiskelijat, jotka opintojensa alussa ovat käytännössä samalla tasolla kuin lukiolaiset. Suurin osa heistä on ennen opintojen alkua käyttänyt pelkästään kaksikielisiä sanakirjoja lähinnä vieraskielisen tekstin ymmärtämiseen tai tuottamiseen. Jotkut ovat selvinneet ehkä pelkästään oppikirjojen sanastojen avulla. Tälle ryhmälle on tyypillistä melko yhtenäinen käsitys kääntämisestä yksittäisten lähtökielisten sanojen kohdekielisten vastineiden hakemisena ja niiden sijoittamisena peräkkäin kohdekieliseen tekstiin.

Toisen käyttäjäryhmän muodostavat työelämässä toimivat kieliasiantuntijat, jotka ovat tottuneita sanakirjojen käyttäjiä ja tiedonhakijoita. He käyttävät myös yksikielisiä sanakirjoja. Ammattilaisuuden yksi tärkeä piirre on kriittinen suhtautuminen sanakirjoihin ja muihin apuvälineisiin ja kasvaminen siihen, että kielissä on harvoin täydellisiä vastineita.

Aikaisemmissa tutkimuksissamme olemme vertailleet, kuinka nämä kaksi käyttäjäryhmää sekä sanovat käyttävänsä että käyttävät erilaisia apuvälineitä kääntäessään (Pilke 2008; Pilke–Puskala 2010; Puskala 2011). Tässä artikkelissa keskitymme pohtimaan sanakirjojen rakennetta ja sisältöä sekä käyttäjän etiikkaa työelämässä toimivien kieliasiantuntijoiden silmin.

Sanakirjat: painetusta sähköiseen, toimitetusta wikiin

Nykyään sanakirjan käsite kattaa painetun teoksen lisäksi laajan kirjon elektronisessa muodossa olevia apuvälineitä (ks. Romppanen 2001: 311; vrt. Nordisk leksikografisk ordbok 1997: 21–24). Osa on kaupallisten painettujen sanakirjojen maksullisia cd- tai verkkoversioita, osa taas verkossa wikiperiaatteella julkaistavia ilmaisia sanakirjoja, joita kuka tahansa saa täydentää ja joiden laatu voi vaihdella.

Elektronisten sanakirjojen suosio kasvaa niiden yleistymisen myötä. Kääntämisessä käytettävien apuvälineiden todellista käyttöä koskeneessa tutkimuksessamme (Pilke–Puskala 2010) opiskelijat turvautuvat lähes yksinomaan elektronisiin sanakirjoihin, sekä kaupallisiin että ilmaisiin. Tätä selittää osaltaan niiden käytön nopeus ja vaivattomuus – vaikka tutkimustilanteessa opiskelijoilla oli käytössään myös painettuja sanakirjoja, vain murto-osa sanahauista kohdistui niihin.

Sanakirjaksi on katsottava myös terminologisin periaattein koostetut eri erikoisalojen sanastot. Tällaisia ovat esimerkiksi Sanastokeskuksen Laajakaistan sanasto (TSK 43) ja Työsuojelusanasto (TSK 35) sekä ruotsalaisen TNC:n Betongteknisk ordlista (TNC 103) ja Skogsordlista (TNC 96). Verkossa on paljon erilaisia suppeahkoja ammattikielisiä sanastoja, jotka yksinkertaisimmillaan ovat mallia sana lähtökielellä ja vastine toisella kielellä tai toisilla kielillä. Ne voivat olla yksi- tai monikielisiä sanastoja, joissa hakusana on määritelty lähtökielellä. Määritelminä voidaan käyttää perinteisen sanakirjoissa esiintyvän yläkäsitteen ja erottelevien käsitepiirteiden lisäksi esimerkiksi synonyymejä, lyhenteiden auki kirjoittamista, pelkkää yläkäsitettä, alakäsitteiden luetteloa, esimerkkilauseita, joissa hakusana esiintyy kontekstissa, kuvia jne. Tällaiset sanastot ovat usein ei-leksikografien tekemiä, mutta ne edustavat ajantasaista kieltä, jota ei yleensä vielä ole hakusanoina perinteisissä sanakirjoissa. (Törnroos 2007.)

Sanakirjojen ohella ammatissa toimivat kieliasiantuntijat hyödyntävät saman aihealueen erikielisiä rinnakkaistekstejä kaksikielisen sanakirjan tapaan. He käyttävät myös hakukoneita ja vertaavat niiden avulla löytämiään erilaisissa korpukissa ja muissa autenttisissa verkkoteksteissä käytettäviä ilmauksia omaan kielitaitoonsa pohjautuviin ilmauksiin. Tällainen tiedonhankinta voidaan rinnastaa yksikielisen sanakirjan käyttöön. (Ks. esim. Hulstijn–Atkins 1998: 10; Puskala 2011.) Se vaatii käyttäjältä todella hyvää kielitaitoa sekä kykyä arvioida näiden tekstien laatua ja luotettavuutta (vrt. Törnqvist 2010: 492).

Artikkelissa lähemmin tarkasteltavat sanakirjat on koottu taulukkoon 1. Taulukosta käyvät ilmi sanakirjan täydellinen nimi ja julkaisuvuosi (jos tiedossa), sanakirjasta tässä artikkelissa käytettävä lyhenne sekä muuta tutkimuksen kannalta oleellista tietoa.

TAULUKKO 1. Artikkelissa tarkasteltavat sanakirjat.

Sanakirja	Lyhenne	Muuta
Yksikieliset sanakirjat		
MOT Kielitoimiston sanakirja (2006)	MOT KS	Osa Kielikone Oy:n maksullista digitaalista sanakirjapalvelua, perustuu painettuun Kielitoimiston sanakirjaan
MOT Norstedts Svensk ordbok 1.0 (1999)	MOT Norstedts	Osa Kielikone Oy:n maksullista digitaalista sanakirjapalvelua, perustuu painetun Svensk ordbokin 3. painokseen
Norstedts Svenskt språkbruk (2003)	Svenskt språkbruk	Idiomeihin ja sanojen rektioon keskittyvä sanakirja
Kaksikieliset sanakirjat		
MOT Suomi-ruotsi-suomi	MOT Ruotsi	Osa Kielikone Oy:n maksullista digitaalista sanakirjapalvelua, yrityksen ”yhdessä ruotsin kielen asiantuntijoiden kanssa kokoama laaja kaksikielinen yleiskielen suursanakirja”
Iso ruotsalais-suomalainen sanakirja (1982–87)	Ru-su	Laajin käytettävissä oleva painettu ruotsi–suomi-sanakirja
Suuri suomi–ruotsi-sanakirja (1997)	Su-ru	Laajin käytettävissä oleva painettu suomi–ruotsi-sanakirja
Suuri suomi–ruotsi cd-rom-sanakirja (1998)	Cd-rom	Perustuu WSOY:n painettuun suomi–ruotsi-suursanakirjaan

Vaikka artikkelimme pohdinnat perustuvat suurelta osin taulukossa 1 esiteltyihin painettuihin tai elektronisiin kaupallisiin sanakirjoihin, katsomme sanakirjan käsitteen piiriin kuuluvan myös erilaiset vapaasti saatavilla olevat sanastot ja wikiperiaatteella täydennettävät sanakirjat niiden toimitusperiaatteista huolimatta.

Sanakirja ja sen sisältö

Tässä luvussa pohdimme sanakirjan sisältöä. Tarkastelemme ensin hakusanojen merkityksiä yhden esimerkkitapauksen pohjalta. Sana-artikkelien esimerkkejä käsittelemme luvussa *Sana-artikkelien esimerkit*. Luvussa *Esimerkin käsite kolmessa sanakirjassa* käymme läpi MOT KS:n ja MOT Norstedtsin sekä Su-ru:n sisältämät näkemykset esimerkin käsitteestä ja pohdimme, mitä seurauksia käyttäjälle käsite-eroavaisuuksista saattaa aiheutua. Tämän jälkeen kuvaamme esimerkin käsitettä käsitejärjestelmän avulla luvussa *Esimerkkien käsitejärjestelmä*. Luvussa *Esimerkkien sisältö: tasa-arvoa vai kummallista vinoumaa?* käsittelemme esimerkkien sisältöä kahden *hon-* ja *han-*pronomineja koskevan tapaustutkimuksen pohjalta.

Sanakirjan johdannossa tai esipuheessa käyttäjälle kerrotaan sanakirjan sisällöstä sekä ohjeistetaan häntä sanakirjan käytössä ja annetun tiedon oikeassa tulkinnassa esimerkiksi selvittämällä hakusanojen ryhmittelyyn, vastineiden järjestykseen yms. liittyvät periaatteet. Leksikografisessa tutkimuksessa niihin ei ole juurikaan kiinnitetty huomiota, vaikka Svensén (2004: 449–455) toteaaakin esipuheen olevan tärkeä.

Harva sanakirjan käyttäjä aloittaa silti sanakirjan käytön perehtymällä johdantoon (Svensén 2004: 450, 542). Tavallisinta on, että käyttäjä turvautuu ohjeisiin vasta siinä vaiheessa, kun hän törmää sana-artikkelissa johonkin vaikeasti tulkittavaan merkkiin tai lyhenteeseen. Näin ollen käyttäjälle jäänee epäselväksi, miksi esimerkiksi *Saksalais-suomalaisessa suursanakirjassa* (Hirvensalo 1980) hakusanan *Mädchen* ensimmäisenä esitetty suomenkielinen vastine on ’pieni (juusto)mato’. Tosin käyttäjä saattaisi alkusanoihin tutustumisen jälkeenkin olla yhtä ymmällään, sillä niissä todetaan, että hakusanan ensimmäiseksi merkitykseksi on ”pyrityt sijoittamaan joko alkuperäinen tai nykyisin yleisin merkitys”. Jokainen vähänkin saksaa osaava käyttäjä pystyisi juuri *Mädchen*-sanan osalta päättelemään, että vasta toinen merkitys ’tyttö’ on nykyisin yleisin merkitys. Jos kyseessä oleva sana olisi kuitenkin merkitykseltään harvinaisempi, ei sanakirjan käyttäjän kielitaju välttämättä riittäisi oikeaan ratkaisuun päätymiseen. Etymologisessa sanakirjassa vastaavat tapaukset olisi ratkaistu sijoittamalla homonyymit eri sana-artikkeleihin.

Sana-artikkelit ovat aina jonkinlaisen valinnan tulos. Esipuheiden perusteella käyttäjä saa jonkinlaisen kuvan siitä, millä periaatteilla kunkin sanakirjan valinnat on tehty. MOT Norstedtsin, MOT KS:n, MOT su-

ru:n, Ru-su:n ja Su-ru:n esipuheissa todetaan, että ne ovat yleiskielen sanakirjoja, joihin sisältyy runsaasti erikoiskielen sanastoa useilta eri aloilta.

Yleiskielen ja erikoiskielen raja on kuitenkin häilyvä. Mitään tarkkaa rajaa niiden välille ei edes ole tarkoituksenmukaista yrittää vetää, sillä tekstit muodostavat jatkumon, jonka toisessa ääripäässä ovat niin erikoistuneet tekstit, että niiden tuottamiseen ja ymmärtämiseen tarvitaan alan substanssiosaamista (Puuronen–Koskela–Laurén–Nordman 1993).

Painettujen sanakirjojen ja niihin perustuvien elektronisten sanakirjojen ongelmana on etenkin tekniikan alan ja siihen liittyvän sanaston nopea kehitys. Hyvänä esimerkkinä tästä ovat monien nykyisin tuiki tavallisten, lähes yleiskielisten sanojen (*tabletti*, *lukulaite*, *muistitikku*) puuttuminen näistä ja toisaalta niiden sisältämät, jo pois käytöstä jääneet käsitteet kuten *kaksoisnäyttökello*: 'klocka med analog och digital visning' (Su-ru).

Hakusanan merkitykset: prototyypeistä kummituksiin

Yksittäisen hakusanan eri merkitykset ja niiden esitysjärjestys kertovat osaltaan sanakirjan laatijoiden maailmankuvasta. Merkityksethän voidaan esittää etymologisessa järjestyksessä (vanhin merkitys ensin), loogisessa järjestyksessä (ns. ydinmerkitykset ensin) tai tärkeys-/yleisyysjärjestyksessä (Svensén 2004: 429–431). Leksikografit pohtivat nykyisin, pitäisikö sanan prototyyppinen, siis tyypillinen merkitys nostaa lähtökohdaksi. Ongelmaksi muodostuu usein kielen kehitys: jonkin sanan tyypillinen merkitys ja käyttöala on saattanut 20–30 vuotta sitten, eli nykyisten painettujen sanakirjojen laatimisaikaan, olla jotakin muuta kuin nykyisin.

Yksi myös Svensénin (2004: 76) esille nostama sanakirjojen ongelma on niin sanotut kummitussanat (*spökord*), eli sanakirjoihin sisältyvät sanat, jotka eivät ole käytössä sanakirjan ulkopuolisessa maailmassa. Svensénin mukaan ongelmaan on syynä sanakirjatoimitusten kritiikitön aikaisempien sanakirjojen käyttö sanakirjatyön pohjana autenttisten lähteiden sijaan.

Havainnollistamme seuraavaksi sanakirjojen käyttäjien hakusanojen merkityksiin liittyviä haasteita konkreettisen esimerkin avulla. Aiempi tutkimuksemme (Pilke–Puskala 2012) on osoittanut, että hakusanan *förvarva* (*sig*) käyttöala on selvästi laajentunut viimeisten 30 vuoden aikana

ja että sanan eri sanakirjoissa esiintyvissä suomenkielisissä vastineissa on osittain kummitussanojen piirteitä.

Sekä ruotsalaisen Språkbankenin sanomalehtikorpuksessa vuosina 1997–2009 että sanoilla *förvärva*, *förvärvar* ja *förvärvade* tammikuussa 2012 tehtyjen Google-hakujen perusteella *förvärva*-sanana merkitys 'hankkia ostamalla' on nykyisin selvästi yleisin. Tällöin yleisin subjekti on kollektiivi tai yritys, mutta sanomalehtikorpuksessa esiintyy jonkin verran myös tapauksia, joissa yksittäiset ihmiset hankkivat omaisuutta.

Kuten taulukosta 2 käy ilmi, sanakirjoissa, joiden aineisto on kerätty ennen 1990-lukua (MOT Ruotsi, Ru-su), lähtökohdaksi on otettu lähinnä yksilöön liittyvä yleiskielinen merkitys 'hankkia' (esim. *hankkia tuloa*, *ystäviä*). Uusimmissa sanakirjoissa (Cd-rom, MOT Norstedts, Svenskt språkbruk) on puolestaan keskeinen liiketalouden diskurssiin kuuluva merkitys 'hankkia ostamalla', Svenskt språkbrukissa se on ainut. Näin ollen pelkästään kaksikielisiin painettuihin sanakirjoihin tai niihin perustuvaan elektroniseen MOT Ruotsiin tukeutuva käyttäjä ei saa ajantasaista kuvaa sanan prototyyppisestä merkityksestä.

MOT Ruotsissa refleksiivimuoto *förvärva sig* on lisäksi erillisenä hakusanana, jonka vastine on sanakirjan mukaan lakitieteen 'hankkia itselleen'. Refleksiivimuotoa ei tosin esiinny ainakaan Suomen nykykylainsäädännössä (Finlex). Se mainitaan Svenska Akademiens ordbokin *förvärva*-hakusana-artikkelissa (SAOB, 1928) transitiivimuodon rinnakkaismuotona, jolloin sillä on useita eri merkityksiä. Google-haun perusteella muoto esiintyy nykyään verkkoteksteissä lähinnä eri koulutusohjelmien esittelyissä, joissa kerrotaan, mitä tietoja ja taitoja opiskelija niissä hankkii (*förvärvar sig*). *Förvärva sig* -hakusanan oikeustieteen diskurssiin kuuluvan merkityksen voidaan siis Svenséniä (2004: 76) mukailleen katsoa olevan nykykielessä sanan kummitusmerkitys, jota ei esiinny sanakirjan ulkopuolella todellisessa elämässä.

Cd-rom:n *förvärva*-hakusanan eri merkitysten synonyymiset ilmaukset ja suomenkieliset määritelmät voidaan myös osittain katsoa kuuluvan kummituskäsitteisiin tai -ilmaisuihin. Suomenkielinen verbilauseke *saada toimeen* on kahden lausekkeen (*panna toimeen* ja *saada toimeksi / saada aikaan*) kontaminaatio. Merkitysvivahteen ruotsinkielinen synonyymi *utverka* määrittellään MOT Norstedtissa 'efter begäran erhålla' eli *saada anomuksesta*. *Utverka*-verbin ja sen suomenkielisen vastineen yhteisten merkityskomponenttien selvittäminen vaatisi tuekseen esimerk-

TAULUKKO 2. *Förvärva/förvärva sig* -hakusanojen merkitykset eri sanakirjoissa.

	hankkia	ansaita	saavuttaa, saada	hankkia itselleen	ostaa	toimittaa
Yksikieliset sanakirjat						
MOT Nor- stedts	skaffa genom köp, arbete e.d. (h. osta- malla, työllä tms). f. kunska-per f. fiender				f. byggnaden (kommunen)	
Svenskt språk-bruk					köpa upp (ostaa) f. företag (bo- laget)	
Kaksikieliset sanakirjat						
Ru-su	f. kunska-per f. väänner f. landet valuta	f. sitt up- pehälle	f. doktors- graden f. aktning f. nngns tack- samhet	f. (sig) kunska- per f. (sig) väänner		
MOT Ruotsi	hankkia	hankkia tuloa	saada osakseen	lakitiede		
Cd-rom	(laajemmista hankin- noista) anska- ffia		(saavuttaa, saada toimeen) utverka		(ostaa) köpa	(toimittaa) leverera

kejä, kuten myös neljäs sanakirjaan sisältyvä merkitysvivahte 'toimittaa/leverera'. Siinähan toiminnan suunta (subjektista pois päin) on päinvastainen kuin peruserkityksessä.

Verkossa saatavilla olevista ilmaisista sanakirjoista sanakirja.org ei tunnista *förvärva*-sanaa ollenkaan, mutta ”ammattilaisten ja käyttäjien kääntämiin sanoihin” perustuva Ilmainen sanakirja antaa sille neljä eri suomenkielistä vastinetta: 'hankkia', 'ostaa', 'värvätä' ja 'voittaa' ilman merkitysvaihteiden käyttöalaa selventäviä esimerkkejä. Käyttäjän on näin ollen valittava vastineista kulloiseenkin yhteyteen sopivin täysin omaan kielitaitoonsa tukeutuen.

Kun sanakirjan käyttäjä turvautuu näihin tällä hetkellä yleisimpiin, helpoiten saatavilla oleviin tai kattavimpiin kaksi- ja yksikielisiin sanakirjoihin tarkistaessaan *förvärva*-sanan merkityksiä, hän saa siis hieman eri vastauksia sanakirjasta riippuen. Pelkästään painettuihin tai sähköisessä muodossa oleviin maksullisiin kaksikielisiin sanakirjoihin turvautuva ei löydä kaikista sanan nykyisin yleisintä merkitysvaihdetta 'hankkia ostamalla'. Sen sijaan yksi ilmaisista verkkosanakirjoista antaa tämänkin vastineen. Myös yksikieliset sanakirjat nostavat tämän merkitysvaihteen keskiöön. Kaksikielisissä sanakirjoissa hakusanojen eri vastineiden merkitykset vaihteineen välittyisivät käyttäjälle paremmin esimerkkien kautta.

Sana-artikkelien esimerkit

Esimerkki voi tuntua selkeältä metaleksikografian käsitteeltä, jonka määritelmä on yksinkertainen. Svensén (2004: 341–345) käy läpi useita eri esimerkin määritelmiä, mutta toteaa itse käyttävänsä termiä *språkprov* (*kielinäyte*) suunnilleen samassa merkityksessä kuin moni käyttää termiä *exempel* (*esimerkki*): kielinäyte on hänen mukaansa sana-artikkelissa esiintyvä tekstinpätkä, joka sisältää hakusanan jossakin muodossa. Kielinäytteen tehtävä on valaista hakusanan käyttöä kontekstissa: sen syntaktisia ja semanttisia mutta myös pragmaattisia ja ensyklopedisia ominaisuuksia. Artikkelissamme määrittelimme esimerkin käsitteen Svensénin tapaan hakusanan merkitystä ja sen kontekstuaalista käyttöä havainnollistavaksi ilmaukseksi.

Eri sanakirjojen eksplisiittiset tai implisiittiset esimerkin määritelmät poikkeavat jonkin verran edellä esitetystä määritelmästä. Tähän voi olla

syynä leksikografian alan koulutuksen ja tutkimuksen rajallisuus Suomessa (ks. esim. Hiidenmaa 2007: 444).

Esimerkin käsite kolmessa sanakirjassa

MOT KS ei määrittele esimerkkiä Ohjeita sanakirjan käyttäjälle -osiossa erikseen vaan toteaa esimerkkien tehtävän olevan hakusanan merkityksen sekä tyypillisten lause- ja käyttöyhteyksien valaiseminen. MOT Norstedtsin ohjeissa esimerkki jaetaan kahteen alakäsitteeseen: *morfologiseen* ja *syntaktiseen esimerkkiin*. Molemmat valottavat sanan merkitysvivahteita, morfologiset esimerkit lisäksi hakusanan sananmuodostusta ja syntaktiset esimerkit hakusanan käyttöä lauseyhteyksissä. Ohjeissa käytetään *esimerkin* (*exempel*) ohella myös termiä *kielinäyte* (*språkprov*), todennäköisesti Svensénin tavoin synonyymisesti.

TAULUKKO 3. *Esimerkki-* ja *exempel-*termien käyttö Su-ru:n otoksessa.

Otoksen kohdat, joihin selitteessä viitataan <i>esimerkki</i> -sanalla	selite
(pystyviiva/lodstreck) aakko nen	erottaa sanan esimerkeissä muuttumatomana pysyvän alkuosan/avskiljer den del av uppslagsordet som förblir oförändrat i exemplen
♦ (vinoneliö/snedruta) aalto våg -en -or (myös fys), (runok) bölja -an or ♦ -jen pauhu vågbrus -et; dånet från vågorna	esimerkkiosa alkaa, tämän jälkeen esimerkki ja käänнос/ Exempeldelen börjar, efter detta exempel och översättning
aakkostaa ordna ¹ <i>ngt alfabetiskt</i>	Korvautuva pronomini pakollisen objektin paikalla vastineessa, rektiotiedoissa tai malli-esimerkissä / Pronomen i obligatoriskt objekt konstruktioner eller mönsterexempel
aarnialue 1 <i>urskogsområde -t -n</i> 2 (<i>metsät</i>): (Suom) <i>ödemarksreservat -et =</i> ; (Ruots) <i>domänreservat</i>	otsikonomainen selite (kaksoispiste)/ exempelrubrik (kolon)
aalto [- -] (<i>kuv</i>) <i>vihreä ~ gröna vågen</i>	otsikoitu esimerkki [ryhmä]/ exempel [grupp] med rubrik

Su-ru:ssa esimerkin käsite jää käyttäjälle epäselväksi. Käyttäjälle-luvussa termi esiintyy vain yhdessä virkkeessä, jossa todetaan sanakirjassa käytettävän puolipistettä erottamaan toisistaan paitsi vastineita, joiden merkitys ei ole täsmälleen sama, myös esimerkkejä sekä esimerkkien erilaisia käännösvaihtoehtoja. Tämän lisäksi sanakirjan etusisäkanteen sijoitetussa, käyttöohjeeksi tarkoitettussa 17 hakusanan otoksessa *esimerkki/exempel* mainitaan kuudessa kohdassa taulukon 3 mukaisesti.

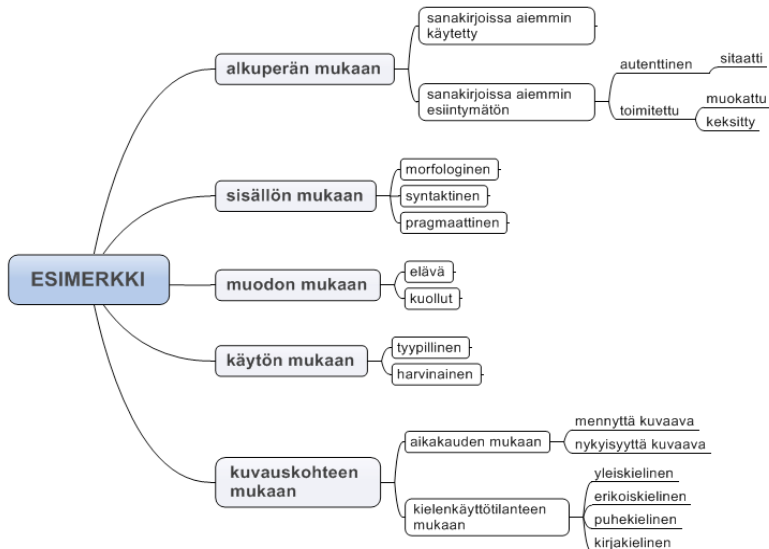
Selitteissä käytetyt termit *malliesimerkki (mönsterexempel)*, *otsikonomainen selite (exemplrubrik)* ja *otsikoitu esimerkki[ryhmä] (exempl[grupp] med rubrik)* eivät täysin avaudu sanakirjan käyttäjälle. Epäselvää on, viittaako *malliesimerkki* samaan käsitteeseen kuin *esimerkki* sekä liittyykö *otsikoitu esimerkkiryhmä* käyttöalapidonnaisuuteen ja *exemplrubrik* (suomeksi *otsikonomainen selite*) ensi sijassa merkitysten selitteisiin vai varsinaisiin esimerkkeihin. *Malliesimerkki*-termi voi käyttäjän mielessä saada epätarkoituksenmukaisen normatiivisen merkityksen.

Esimerkkien käsitejärjestelmä

Esimerkkejä voidaan tarkastella monesta eri näkökulmasta. Kuviossa 1 esittelemämme käsitejärjestelmä sisältää viisi eri näkökulmaa, jotka ovat alkuperä, sisältö, muoto, käyttö ja kuvauskohde.

Esimerkit voivat olla peräisin aiemmista sanakirjoista tai ne voivat olla uusia. Ne voivat olla autenttisuudeltaan puhtaita sitaatteja (lähinnä kaunokirjalliset lähteet) tai toimituksellisia. Korpusaineistoja (esim. lehtiaineistot) muokataan usein ainakin nimien osalta ja osa esimerkeistä voidaan luoda kokonaan toimituksessa. (Vrt. Svensén 2004: 345–346.)

Artikkelissamme tarkasteltavina olevien sanakirjojen johdantojen perusteella niiden sisältämissä esimerkeissä on sekä sanakirjoissa aiemmin käytettyjä että niissä aiemmin esiintymättömiä esimerkkejä. Sanakirjoista vanhimmassa, vuonna 1982 ilmestyneessä Ru-su:ssa pääasiallisena hakusanalähteenä on käytetty Bertil Molden toimittaman *Illustrerad svensk ordbokin 3*. painosta vuodelta 1964 sekä Svenska Akademiens *ordlistan 10*. painosta vuodelta 1973. Suomenkielinen aineisto on peräisin Nykysuomen sanakirjasta. Molempien kielten osalta sanakirjaa on täydennetty ammattisanakirjoihin ja erikoissanastoihin sisältyvillä sanoilla sekä sanakirjatoimituksen sanastamilla sanomalehtiteksteillä, ammattikirjallisuudella ja kaksikielisillä esitteillä, joista poimitut sanat edustavat sana-



KUVIO 1. Esimerkit viidestä näkökulmasta.

kirjoissa aiemmin esiintymättömiä sanoja. Sanakirjoista uusin, Svenskt språkbruk (2003) mainitsee takakannessa olevansa täysin uusi sanakirja. Johdannossa todetaan sen korvaavan suurilta osin vanhentuneen Svensk handordbokin vuodelta 1966. Epäselväksi jää, ovatko myös kaikki esimerkit aiemmin sanakirjoissa esiintymättömiä.

Sanakirjan käyttäjälle esimerkkien alkuperällä on merkitystä monelta-kin kannalta. Luvussa *Hakusanan merkitykset: prototyypeistä kummituksiin* mainitut kummitussanat ja -merkitykset saattavat kummitella myös esimerkeissä, jolloin käyttäjä saattaa harhautua luulemaan niiden olevan käypää kieltä. Myös painettujen sanakirjojen ”uusista” lähteistä poimitut esimerkit edustavat jo ilmestyessään eilispäivän kielenkäyttöä, joka Ru-su:n tapauksessa on vähintään 30 vuotta vanhaa.

Sisällön mukaan esimerkit voidaan jakaa *morfologisiin* eli sananmuodostukseen liittyviin, *syntaktisiin* eli lauseyhteyttä valaiseviin sekä *pragmaattisiin* eli käyttöyhteyttä ilmaiseviin esimerkkeihin. Näitä kaikkia tyyppisiä esiintyy kaikissa artikkelissamme tarkasteltavissa sanakirjoissa

idiomeihin ja sanojen rektioihin keskittyvää Svenskt språkbrukia lukuun ottamatta, jossa ei luonnollisestikaan ole morfologisia esimerkkejä. Käyttäjän on otettava selvää kunkin sanakirjan käyttötarkoituksesta tutustumalla sen esipuheeseen.

Muodon mukaan jaoteltuina esimerkit voivat olla myös joko *kuolleita* tai *eläviä*. Kuolleet esimerkit ovat Svensénin (2004: 345–346) mukaan ei-autenttisia, pelkästään välttämättömistä kieliopillisista elementeistä koostuvia lausekkeitä, joissa on käytetty indefiniittipronomineja ja verbin infiniittisiä muotoja, elävät esimerkit sisältävät joko persoonapronomineja tai tyypillisiä substantiivilausekkeitä ja verbin finiittimuotoja. Tyypillinen kuollut esimerkki on MOT KS:n sana-artikkelissa *osoittaa* lauseke *Osoittaa kädellään jhk suuntaan t. jtak* ja tyypillinen elävä esimerkki saman sana-artikkelin virke *Opiskelijat osoittivat mieltään hallituksen päätöstä vastaan*. Molempia tyyppisiä esiintyy kaikissa tarkastelemissamme sanakirjoissa. Kokenut käyttäjä osaa täydentää kuolleet esimerkit kontekstin mukaisesti, mutta sanakirjoja vähemmän käyttäneille *jhk-* ja *jtak-*tyyppiset lyhenteet saattavat tuottaa hankaluuksia, vaikka ne olisivatkin käyttäjän äidinkielellä. Elävissä esimerkeissä on toisaalta ratkaistava minkä tyyppinen ja kumpaa sukupuolta oleva toimija lauseeseen valitaan (ks. myös luku *Esimerkkien sisältö: tasa-arvoa vai kummallista vinoumaa?*)

Käytön mukaan esimerkit voidaan jakaa *tyypillisiin* ja *harvinaisiin*. Vaikka tyypillisuus (prototyypisyys) onkin nykyisin sekä merkitysten kuvausta että esimerkkien valintaa ohjaava tekijä, harvinainen merkitys ja siihen liittyvä esimerkki tuo käyttäjän kannalta usein lisäarvoa. Etenkin kieliasiantuntijat turvautuvat sanakirjaan halutessaan tarkistaa, olisiko jossakin kontekstissa esiintyvällä sanalla heidän tuntemastaan prototyypisestä merkityksestä poikkeava merkitys.

Kuvauskohteen mukaan luokiteltuna esimerkit voidaan jakaa aikakauden ja käyttötilanteen mukaan. Tietyissä mielessä ainakin painettujen sanakirjojen ja painettuihin sanakirjoihin perustuvien elektronisten sanakirjojen sisältö kuvaa suurimmaksi osaksi jo mennyttä kielenkäyttöä, koska kielenkäyttö elää jatkuvasti. Kieliasiantuntijan kuuluu sanakirjan käyttäjänä olla tietoinen tästä, vaikka sanakirjojen markkinoinnissa nostetaan esiin niiden ajantasaisuus. *Nykyinen–mennyt*-jako liitetään kuvion 1 jaottelussa toisaalta historialliseen, tietyn aikakauden kielenkäytön tai ajattelutavan kuvaukseen, toisaalta nykykieleen ja -käsitteisiin.

Käyttötilanteen mukaan esimerkit voidaan jaotella yleis- ja erikoiskielisiin sekä puhe- ja kirjakielisiin. Sekä kirjakielinen että puhekielinen esimerkki voivat kuvata yleiskielen tai erikoisalaviestinnän ilmiötä. Käyttäjän näkökulmasta esimerkkien yleis- ja erikoiskielen välinen raja on oleellinen hänen valitessaan tiettyyn asiayhteyteen sopivaa ilmausta. Jos jo aiemmin esiin tulleen MOT Ruotsin lakitieteen alaan kuuluvan *förvärva sig* -hakusanan ja sen suomenkielisen *hankkia (itselleen)* -vastineen yhteydessä olisi esimerkki, käyttäjän valintapäätös ehkä helpottuisi. Toisaalta sekä yleis- ja erikoiskielen että puhe- ja kirjakielen raja on häilyvä ja elää koko ajan. Sellaiset sanat ja ilmaukset, jotka muutama vuosikymmen sitten on luokiteltu puhekielisiksi tai erikoisalan kieleen kuuluviksi, voivat olla nykyisin täysin yleiskielisiä ja hyväksytyjä myös kirjoitetussa kielessä. Käyttäjän onkin syytä olla tietoinen myös näistä sanakirjojen rajoituksista.

Esimerkkien sisältö: tasa-arvoa vai kummallista vinoumaa?

Koska sanakirjat koetaan usein vahvasti normatiivisiksi, niiden sisällöllä voidaan olettaa olevan kielenkäyttöä ohjaava vaikutus (ks. Karlsson 1987: 52; Vanhatalo 2005a: 2; Edlund–Erson–Milles 2007: 192). Tämän vuoksi sanakirjan aikaa ja kulttuuria heijastavilla esimerkeillä voi olla vaikutusta myös käyttäjän näkemyksiin yhdenvertaisuudesta ja tasa-arvosta.

Englannin ja ruotsin kielen sukupuolisidonnaiset pronominit ovat herättäneet paljon keskustelua yleisten tasa-arvokysymysten myötä (ks. esim. Epple 2000; Språkriktighetsboken 2005: 75–79; Martyna 2006; Milles 2008: 53). Vaikka kielenhuolto nykyisin suositteleeikin sukupuolineutraalia kielenkäyttöä (esim. *han-* ja *hon-*pronomien rinnastaminen, *den*, *hen*), sanakirjan esimerkeissä ei ole tarkoituksenmukaista välttää kokonaan viittauksia sukupuoleen (vrt. Grönros 2007). Merkityksellistä on, missä kontekstissa ja tietyssä määrin myös missä määrällisessä suhteessa sukupuolet esitetään. Tutkimuksissa, jotka koskevat kaksi- ja yksikielisen sanakirjan *han-* ja *hon-*pronomien käyttöä esimerkeissä (Pilke 2009, 2010; Soukka 2010), on havaittu sekä määrällistä että laadullista vinoumaa sukupuolten välillä.

Taulukossa 4 on esitetty *han-* ja *hon-*pronomien esiintyminen Cdrom:ssa. *Hon* esiintyy 154 artikkelissa ja 163 esimerkissä ja *han* vastaavasti yli 3 500 artikkelissa ja 5 000 esimerkissä. Määrällinen ero on tosiasiasa vielä suurempi, koska tutkimusaineisto kattaa *hon*-pronominin osalta

koko sanakirjan mutta on *han*-pronominin osalta rajattu hakuteknisistä syistä kirjaimiin *a-s* (ks. Pilke 2010). Samassa artikkelissa pronominit esiintyvät 67 kertaa ja samassa esimerkissä 8 kertaa. Soukka (2010) on Svenskt språkbruk -sanakirjan esimerkkejä koskevassa tutkimuksessaan (kirjaimet *a* ja *m*) päätenyt tasaisempaan jakaumaan: *hon*-pronomini esiintyy 557 esimerkissä, *han* 807 esimerkissä ja molempia pronomeineja käytetään samassa sana-artikkelissa 179 kertaa.

TAULUKKO 4. *Hon* ja *han*-pronominien esiintyminen Cd-rom:ssa.

pronomini	sana-artikkeli	esimerkki
hon (a-ö)	154	163
han (a-s)	3 542	5 086
hon & han (a-ö)	67	8

Cd-rom:n hakusanaan *hyvä* liittyvä artikkeli esimerkkeineen kuvaa sanakirjassa esiintyvää pronominiinien määrällistä eroa. *Han* esiintyy artikkelissa 27 esimerkissä, *hon* vain yhdessä. *Han har ett gott namn om sig som vetenskapsman* ja *hon är av god familj* ovat ainoat merkitykseen 'hieno, arvostettu' liittyvät esimerkit. Sukupuolia kohdellaan näin määrällisesti samalla tavalla. Laadullisena erona on kuitenkin konteksti, jossa aktiivinen mies on saavuttanut arvostetun yhteiskunnallisen statuksen, mutta jossa passiivisen naisen viitekehys muodostuu suvusta ja perheestä. Myös Nikula (1997) on saanut tutkimuksessaan vastaavia tuloksia: nainen esiintyy objektin (*patient*) roolissa joko kokijana (*experiencer*) tai saajana (*receiver*). Esimerkkien lähempi tarkastelu osoittaa vastaavia laadullisia eroja molemmissa sanakirjoissa. Nainen kuvataan seksistisesti miehensä kautta (vrt. Müller-Fuchs 1993: 11) *alistaa*-hakusanan viimeisessä esimerkissä tämän nöyryytettynä ovimattona (esimerkki 1) ja *kaiku*-hakusanan esimerkissä vain miehensä kaikuna (esimerkki 2). Mies on puolestaan samoissa sana-artikkeleissa outoääninen kerskuri, joka ei alistu komenneltavaksi.

- (1) **alistaa** (Cd-rom)
 - han vill alltid spela översittare, han låter sig inte toppridas (underkuvas)
 - hon var en hemmafru som kuschats till att vara dörrmatta för sin man
- (2) **kaiku** (Cd-rom)
 - hon är bara ett eko av sin man
 - han lät konstig på rösten

Varsin usein mies saa esimerkeissä aktiivisen toimijan roolin, jossa hän voi yllättää kaikki tai selvittää helposti, kun nainen on passiivinen, vastaanottava objekti, joka voi muistuttaa kuuluisaa näyttelijää (hakusana *hämmästyttää*, esimerkki 3).

(3) **hämmästyttää** (Cd-rom)

- han förvånade alla med sin prestation
- han klarade sig förvånansvärt lätt
- hon är förvånande lik Greta Garbo

Kun naista arvioidaan ulkonäön tai epävakaa tunne-elämän pohjalta, kyse on Guentherodtin, Hellingerin, Trömel-Plötzin ja Puschin (1980: 20) mukaan seksistisestä kielenkäytöstä. Näiden aspektien osalta molemmissa sanakirjoissa on samankaltaista vinoumaa. Hakusanojen *alla/allt* yhteydessä (esimerkki 4) *han* on rakastettava ja tärkeä, päämäärätietoinen, vahvatahtoinen, lapsiaan rakastava mies. *Hon* haluaa sulautua joukkoon, on huonomainen, kohtaa vaikeuksia ja näkee huonosti. Tämän hakusanan yhteydessä on havaittavissa kuitenkin pyrkimys asettaa mies perhekontekstiin lastensa kautta sekä korostaa hänen jokseenkin rajallisia taitojaan (ainoastaan tanssii hyvin).

(4) **alla/allt** (Svenskt språkbruk)

- han är allt för henne
- han krävde allt eller ingenting
- han älskade sina barn över allt annat
- han dansar bra, men det är också allt
- hon vill se ut som alla andra
- hon är allt annat än den familjeflicka hon utmålas som
- den kamp hon har att utkämpa är allt annat än enkel
- hon började se allt sämre

Mattisson (1999, 2006) on tutkinut Svenska Akademiens Ordboken (SAOB) lähteitä historialliselta kannalta ja tullut siihen tulokseen, että ne ovat miesvaltaisia. Mattisson kysyykin tässä yhteydessä, onko SAOB:n noin 100 vuoden ajan täysin miesvaltaisena toiminut toimitus voinut tiedostamattaan vaikuttaa sisältöön niin, että ”ord och uttryck använda av kvinnor gått förlorade samt så att den kvinnliga verkligheten inte finns med i citaten.” (Mattisson 1999: 139, 2006: 66). Nikulan (1997: 200–201) yksikielistä ruotsin sanakirjaa koskevassa tutkimuksessa nainen esitetään

stereotyyppisessä roolissa esimerkiksi hakusanojen *byka* ja *städa* yhteydessä, kun hakusanan *tjänst* esimerkeissä subjekti on maskuliininen. (Ks. myös Pilke 2009.)

Sanakirjan käyttäjän etiikka: totuutta, vastuuta ja valtaa

Erilaiset sanakirjat ovat kieliasiantuntijan tärkeimpiä työkaluja. Siksi tämän työkalun kriittisen käytön ja kehittämisen tulisi olla erottamaton osa kieliasiantuntijan ammattietiikkaa. Etiikka liittyy Salmisen ja Ikola-Norrbackan (2009: 10) mukaan toimintoja ja käyttäytymistä ohjaavana säännöstönä moraalisiin valintoihin ja tavoiteltaviin eettisiin arvoihin. Eettisiä arvoja suojelevat käyttäytymissäännöt eli normit, jotka sallivat, kieltävät tai suosivat tietynlaista toimintaa (ks. esim. Kotkanvirta–Nyysönen 1997: 7–10). Kieliasiantuntijan ammattietiikan tulisi käsittää ainakin seuraavat kolme sanakirjan käyttöön ja kehittämiseen liittyvää keskeistä arvoperustaista asiaa: totuuteen pyrkiminen, vastuunkanto ja vallankäyttö.

Totuuteen pyrkiminen. Käyttäjä pyrkii aina löytämään asiayhteyteen sopivimman ilmauksen sanakirjoista tai muista apuvälineistä. Tämä edellyttää ensinnäkin tietoisuutta erilaisista sanakirjoista sekä niiden käyttötarkoituksesta, mahdollisuuksista ja rajoituksista. Toiseksi tämä edellyttää sanakirjojen ja muiden apuvälineiden käytön hallintaa, mutta myös kriittistä suhtautumista niiden sisältämään tietoon ja välittämään maailmankuvaan.

Esimerkiksi käännskurssin opettaja voi tietoisesti sallia myös verkosta löytyvien ilmaisten sanakirjojen käytön. Asettamalla tavoiteltavaksi arvoksi totuuteen pyrkimisen sekä edellyttämällä lopputuloksen tarkastelua tämän pohjalta opettaja ohjaa opiskelijoita eri sanakirjojen ja muiden tiedonlähteiden kriittiseen käyttöön. Tällaisen taidon harjoittelu pitäisi mielestämme aloittaa jo perusopetuksessa.

Vastuunkanto. Sanakirjan käyttäjä valitsee, miten hän käyttää sanakirjoja. Sanakirjaa ei koskaan lueta kirjana kannesta kanteen. Nykyisellään varsin harvoin luettuihin sanakirjojen johdantoteksteihin tutustuminen auttaisi kuitenkin käyttäjää arvioimaan ja tulkitsemaan sanakirjan sisältöä tekijöiden tarkoittamista lähtökohdista. Käyttäjien mielipiteet

johdantotekstien sisällöistä auttaisivat parantamaan sekä niiden luettavuutta että informaatioarvoa oleellisella tavalla.

Sanakirjojen kokonaisvaltaisen käytön tulisi johtaa entistä tiedostetumpaan vastuunkantoon myös niiden sisällöstä. Käyttäjän tulisi olla valmis antamaan palautetta havaitsemistaan puutteista, virheistä, kummitussanoista ja syystä tai toisesta johtuvista vinoumista hakusanojen tai esimerkkien valinnassa. Tämä edellyttää kuitenkin, että palautteenanto tehdään teknisesti vaivattomaksi huolimatta siitä, vaihtuuko esimerkiksi sanakirjan kustantaja. Tarjolla tulisi olla samanlainen sähköinen palautelomake kuin ilmaisilla verkkosanakirjoilla. Avoin kysymys on tietysti, tulisiko vastuullisuuden ulottua myös wikiperiaatteella toimiviin sanakirjoihin, eli voidaanko kieliasiantuntijoilta edellyttää myös niiden kommentointia ja sisällön kehittämistä?

Käyttäjät edellyttävät tulevaisuuden sanakirjoilta entistä enemmän sisältöä. Nykytekniikka mahdollistaa elektronisen interaktiivisen hybridisanakirjan, eräänlaisen sanakirjaportaalin, jossa yhdistyvät yksikielisen sanakirjan määritelmät ja runsaat kontekstisidonnaiset esimerkit sekä kaksi- ja monikielisen sanakirjan vastineet esimerkkeineen. Tällaisessa uudenmuotoisessa sanakirjassa käyttäjä pystyisi navigoimaan sana-artikkeleissa jatkumoilla prototyypinen–harvinainen, diakroninen–synkroninen, yleiskieli–erikoisalaviestintä ja niin edelleen. Tarvitessaan käyttäjä löytäisi esimerkiksi ääntämiseen, taivutukseen ja merkitykseen liittyvää tietoa. Kääntämisen tutkijoiden jo 1990-luvulla visioima sanakirja (Varantola 1998) olisi siis nykyään mahdollinen.

Askel tähän suuntaan on otettu MOT-sanakirjastossa, jossa käyttäjä voi koota oman sanakirjaportaalin kirjastoon kuuluvista sanakirjoista ja saada näin yhdellä haulilla näytölle eri sanakirjojen sisältämän tiedon. Tavallaan myös ammattikäytäntöjen käyttämien käännösmuistien voidaan katsoa olevan eräänlaisia yksilöllisiä hybridisanakirjoja.

Vallankäyttö. Yksinkertaisimmillaan yksilön vallankäytössä on kyse siitä, mitä apuvälineitä kieliasiantuntija käyttää sekä mistä hän on valmis maksamaan ja kuinka paljon. Erilaisissa konteksteissa ja työtehtävissä toimivilla kieliasiantuntijoilla eli kääntäjillä, tulkeilla, opettajilla, suunnittelijoilla, toimittajilla ja tutkijoilla on sekä yksilöinä että etenkin kollektiivisesti mahdollisuuksia vaikuttaa siihen, millaisia sanakirjoja Suomessa tehdään ja onko niitä varaa jättää tekemättä.

Kuten edeltä käy ilmi, Suomessa ei juuri ole tarjolla leksikografian alan

koulutusta. Hiidenmaa toteaa asian tiimoilta Kielikoulutuspoliittisen projektin loppuraportissa vuodelta 2007 seuraavaa:

- (5) Leksikografian tutkimusta ja koulutusta on vain yhdessä yliopistossa, mikä tarkoittaa sitä, että sanakirja-alalle ei riitä tarpeeksi sanaston tutkimuksen koulutuksen saaneita. On mahdollista myös, että sanakirjan teossa vaadittavaa koulutusta ei tunneta kustantamoissa riittävästi, vaan sanakirjoja voidaan laatia yleisen kielitaidon perusteella. (Hiidenmaa 2007: 442.)

Lisäksi suhteellisen harva suomalainen kielentutkija tutkii sanakirjoja ja niiden yhteiskunnallista merkitystä. Tutkimuksen ja koulutuksen perinteen puuttuminen voidaan näin ollen osaltaan nähdä syynä siihen, ettei sanakirja-alalla Grönrosin (2011) sanoin ”ole mennyt viime aikoina kovin hyvin”. Kaupallisissa kustantamoissa ei enää ole sanakirjantoyhtymittäjiä ja myös ainakin yhden Kotimaisten kielten keskuksen sanakirjahankkeen jatko on henkilöstövähennysten vuoksi ollut jonkin aikaa vaaka-tilassa (Kotus 2011). Grönros (2011) toteaa sanakirjatyön kaltaisen kulttuurialan asettamisen säästökohteeksi vaarantavan sekä kielitiedon että kulttuuriarvot. Vilppula (2006) korostaa sekä historiallisten että pienten sukukielten sanakirjojen kansallista ja kulttuurista merkitystä nykykielen, uhanalaisten kielten ja myös etymologisen tutkimuksen kulmakivinä. Tässä yhteydessä myös sanakirjan käyttäjän valinnat ja arvot nousevat merkittävään rooliin.

Kollektiivisessa vallankäytössä on kyse opetuksen ja tutkimuksen painopisteistä ja valinnoista. Yliopistot ovat avainasemassa suomalaisen leksikografian tulevaisuutta ajatellen. Tutkijoilla on mahdollisuus valita tai olla valitsematta leksikografisia aiheita tutkimuskohteiksi. Kun alaa tutkitaan, on mahdollista kehittää myös tutkimukseen perustuvaa opetusta, kuten esimerkiksi Göteborgin yliopistossa, jossa on leksikografiseen tutkimukseen keskittyvä Lexikaliska institutet ja leksikografiaan suuntautuva maisteriohjelma.

Kyseessä on samalla myös moraalinen valinta, johon liittyvät vastuun ja velvollisuuden käsitteet. Tehtävä on jakaantunut useamman toimijan vastuulle: sanakirjahankkeita toteuttavan Kotimaisten kielten keskuksen sanakirjaosaston, alan tutkimuksesta ja opetuksesta vastaavien yliopistojen ja erilaajuisia alan hankkeita rahoittavien säätiöiden.

Lopuksi

Artikkelissa olemme pohtineet keskeisiä leksikografian käsitteitä. Sanakirjan käsite sekä koko leksikografian ala on nähdäksemme murroksessa. Sanakirja kuvaa aina tietyssä mielessä menneen ajan kielenkäyttöä. Vaikka todellisuutta ei luoda sanakirjojen toimituksissa, on sisältöön liittyvillä valinnoilla merkitystä kieliyhteisön kollektiiviselle tietoisuudelle. Jo pelkästään *förvärva*-hakusanan tarkastelu kummitussanoineen ja -merkityksineen sekä sukupuolisidonnaisten *han*- ja *hon*-pronominien käyttö esimerkeissä osoittavat väittämät todeksi.

Sanakirjojen käyttöä helpottaisi sellaisten metaleksikografisten terminen kuin *selite* ja *esimerkki* selkeämpi käyttö, joka perustuisi alan yhteiseen näkemykseen käsitteiden sisällöstä. Artikkelissa olemme esitelleet yhden näkemyksen esimerkin käsitteen sisällöstä, joka voisi olla tulevaisuudessa pohjana käyttäjälle suunnatuissa ohjeissa. Vastaavanlaista käsiteanalyysiä on syytä tehdä laajemminkin.

Artikkelin lopuksi esittämiemme eettisten pohdintojen perusteella voimme todeta, että kieliasiantuntijat sanakirjan käyttäjänä ovat sekä yksilöinä että etenkin kollektiivina paljon vartijana. Kaikkien kieliasiantuntijoiden tulisi kantaa yhteisvastuuta alasta siten, että lopputuloksena olisi yhä innovatiivisempia sanakirjoja. Tämä edellyttää poikkitieteellistä tutkimusta, joka kattaa myös tekniset ratkaisut. Tarvitaan myös yhteiskunnallista keskustelua kielitiedon ja kulttuuriarvojen tallentamisen merkityksestä sekä tutkimuksen ja keskustelun pohjalta syntyvä kielipoliittinen strategia.

RAIMO JUSSILA

Sanakirjat ja tekijänoikeus

Artikkelissani pohdin, mitä tekijänoikeuslaki merkitsee sanakirjan näkökulmasta. Taustaksi selvitän lyhyesti aiheeni kannalta keskeisiä säädöksiä. Sitten tarkastelen tekijänoikeuksia sanan, määritelmän, esimerkistön ja sanakirjan kokonaisuuden kannalta. Luettelosuoja ansaitsee oman lukunsa. Tärkeä seikka on se, ylittääkö sanakirja niin sanotun teoskynnyksen. Käytännön sanakirjatyön kannalta sitaattioikeus on keskeinen oikeus; sen laajuus sekä lähteiden ilmoittaminen on myös tutkimuseettisesti ongelmallinen. Arkistoaineokset ovat useimmissa sanakirjoissa perustana. Mikä niiden suhde on tekijänoikeuteen? Tähän liittyy myös kysymys, mitä ammattimaista käyttöä varten voi kopioida ja kuinka paljon. Sanakirjan toimittajan oikeuksista käsittelen edelleenluovutus-, isyys- ja respektioikeuksia, joilla on myös eettisiä ulottuvuuksia.¹

Tekijyyden suoja-aita

Tekijänoikeuslain keskeiset säädökset on laadittu hyvin yleiselle tasolle osoittamaan henkisen omaisuuden suurpiirteisiä rajoja. Ne ovat rakennusohjeita, mihin ja miten tekijänoikeuden suoja-aita portteineen tulee pystyttää. Laissa ei mainita sanallakaan sanakirjoista, ei edes kirjoista ylipäätään.

Kuitenkin tekijänoikeus tulkintoineen ja oikeustapauksineen ulottuu käytännön elämässä varsin pitävänä aitana aivan ruohonjuuritasolle asti.

1 Kiitän Suomen tietokirjailijat ry:n lakimiestä Mari Wallgrenia käsikirjoituksen asiantuntevasta kommentoinnista.

Sen ali on vaikea livahtaa eikä sitä helposti pysty kiertämään. Umpimuuriksi sitä ei ole muurattu, vaan yksityishenkilö saa sen raoista nauttia toisten teosten hedelmistä omiksi tarpeikseen. Laillista käyttöä varten siihen on rakennettu erilaisilla avaimilla avautuvia portteja.

Tekijänoikeuden lähtökohtana on sananvapaus, oikeus käyttää tietoa esteettä ymmärryksensä mukaan ja ilmaista itseään vapaasti. Tästä syystä idea, aihe ja tieto eivät saa tekijänoikeuden suoja; ne ovat lintuja, jotka lentävät aidan yli. Suoja koskee vain teoksen *muotoa*, tekijän itsenäisesti luomaa ilmaisutapaa. Ideat, aiheet ja tiedot ovat periaatteessa kenen tahansa vapaasti käytettävissä. Kuitenkin käytännössä esimerkiksi työsopimus usein rajoittaa työnantajan tietojen käyttöä hyvinkin tiukasti.

Tekijänoikeuden kohde on *teos*, jonka luova toiminta – kirjoittaminen, säveltäminen, maalaaminen, muotoileminen, laulaminen, soittaminen, näyttelemine, valokuvaaminen ja niin edelleen – on synnyttänyt. Sen konkreettinen ilmentymä on *teoksen kappale*. Teokset jaetaan vain kahteen pääläjiin, *kirjallisiin* ja *taiteellisiin*. Kirjallisuus kätkeytyy termin *kirjallinen teos* taakse ja tietokirjallisuus täsmentyy sen piirissä *selittäväksi kirjalliseksi esitykseksi*. Lisäksi laki tuntee teoksia tai teoksen osia yhdistämällä aikaansaadun *kirjallisen kokoomateoksen*.

Erikseen tekijänoikeuslaki säättää niin sanotusta *luettelosuojasta*, joka koskee luetteloa, tietokantaa tai muuta sellaista työtä, joissa kerääminen, yhdistäminen ja esittäminen vaativat huomattavaa panostusta ja jotka sisältävät suuren määrän tietoa. Oleellinen ero 70-vuotiseen tekijänoikeussuojaan on se, että luettelosuojan kesto on vain 15 vuotta teoksen ilmestymisestä. Myös luettelosuojan rajat niin suojattomiin aineistoihin kuin varsinaista tekijänoikeutta nauttiviin teoksiin päin ovat sumeat.

Tekijänoikeus syntyy samalla hetkellä kuin teoskin. Teos syntyy siis valmiiksi suoja-aitauksen sisään. Tätä oikeutta ei tarvitse rekisteröidä patentin tapaan. Se ei vaadi teokseen copyright-merkintöä, joka kuitenkin selventää sitä, kenelle tai keille oikeudet kuuluvat. Myös julkistamattomalla teoksella on tekijänoikeus.

Mikä tahansa hengenluoma tai näppäryyden näyte ei kuitenkaan nauti tekijänoikeuden suoja. Teoksen on ylitettävä eräänlainen omaperäisyyden ja luovuuden rima, *teoskynnys*, ennen kuin se saa suoja. Tekijänoikeuslaki ei sitä määrittele eikä edes suoraan mainitse. Sen ”korkeus” vaihtelee teoslajeista toiseen. Viime kädessä oikeusistuin asettaa kynnyksen oikeaksi arvioimalleen korkeudelle kiistakumppaneita ja asiantunti-

joita kuultuaan. Nyrkkisääntönä teoskynnyksen tulkinnassa on se, ettei kukaan olisi tehnyt kyseistä teosta täysin samalla tavalla. Aitavertausta käyttäen ”teos”, joka ei ylitä teoskynnystä, valuu suoja-aitauksen alariu’un ali vapaaseen käyttöön. Käytännössä teoskynnys on lähes maarajassa.

Tekijänoikeus syntyy vain yksittäiselle luonnolliselle henkilölle. Tätä moraalista *isysoikeutta* ei voi siirtää toiselle; yhteisteoksissa se kuitenkin jakautuu kaikkien tekijöiden kesken. Taloudelliset oikeudet tekijä voi luovuttaa edelleen esimerkiksi kustantajalle tai yhdistykselle. Palkkатыössä tehdyn teoksen osalta luovutuksesta on tavallisesti sovittu jo työ sopimuksessa työnantajan kanssa. Yritys, yhteisö, kustantaja tai muu laitos voi saada taloudelliset tekijänoikeudet vain erillisen sopimuksen, ei suoraan tekijänoikeuslain perusteella. Tavallisimmat sopimustavat ovat kustannussopimus ja työ sopimus.

Tekijyyden suoja-aidan portit avautuvat tavallisimmin taloudellisilla avaimilla, joita ovat kustannus- ja työ sopimus. Erikseen sopimalla avautuvat kustantajalle mahdollisuudet teoksen taloudelliseen jatkokäyttöön, kuten käänös-, muuntelu- ja edelleenluovutus oikeuksiin sekä oikeuteen julkaista teos muilla kuin alkuperäisellä tavalla. Muunteluoikeudet saaneen toimintaa rajoittaa respektioikeus, joka kieltää muuttamasta teosta tekijän kirjallista arvoa tai omalaatuisuutta loukatun.

Laissa on rakennettu myös muutamia yleisiä portteja, jotka mahdollistavat teoksen käytön ilman tekijän lupaa tietyn rajaehdoin. Näin laki antaa oikeuksia myös käyttäjille eli se rajoittaa määrätapauksissa tekijän oikeutta teokseensa. Tärkein niistä on *sitaattioikeus*, joka antaa jokaiselle luvan siiteerata (< latinan *citare* ’panna liikkeeseen, kutsua’) lain määräämissä rajoissa mitä tahansa julkistettua teosta. Tämän oikeuden ylittäminen voi johtaa plagiointiin (< latinan *plagiarius* ’ihmisryöstäjä, orjakauppias’), joka saattaa merkitä sakko- ja korvaus seuraamuksia. Tutkimusmaailmassa plagiointi on moraalisesti raskauttavaa ja saattaa katkaista tutkijanuran. Toinen tärkeä tekijänoikeuden rajoitus on *kopioinnin* sallittavuus yksityisiin tarkoituksiin. Tällöin on pohdittava rajanvetoa siihen, kuinka paljon saa kopioida ja levittää lähipiiriin, ja siihen, onko kopiointi yksityistä vai esimerkiksi työnantajan työhön liittyvää. *Sopimuslissenssit* ovat eräänlaisia yleisiä portteja, joissa portinvartijana on tekijänoikeuksien omistajia edustava järjestö. Kirjan alueella tällainen menettely on Kopioston valvoma kopiointi. Toinen kirja-alan lissenssijärjestö on Sanasto, joka kerää ja jakaa lainauskorvauksien lissenssimaksut tekijöille.

Nykyinen tekijänoikeuslakimme on peräisin vuodelta 1961 (Tekijänoikeuslaki 8.7.1961/404). Sitä on vuosien mittaan monin osin täsmennetty. Sitä uudistettiin merkittävästi 2000-kymmenluvun puolivälissä erityisesti teosten kopiointiin ja niitä suojaavien teknisten ratkaisujen ohittamiseen sekä jälleenmyyntiin liittyvissä kysymyksissä. Tiivis yleisesitys tietokirjailijan ja tutkijan tarvitsemista tekijänoikeustiedoista on esimerkiksi Heikki Jokisen kirjassa Tietokirjailijan eväät (2010).

Onko sanalla tekijänoikeutta?

Kieli on puhujien yhteistä omaisuutta. Kieleen ja sen yksiköihin, kuten sanaan, ei kenelläkään ole lähtökohtaisesti tekijänoikeutta. Tekijänoikeutta alkaa syntyä silloin, kun kieltä käytetään eli sanoja yhdistellään luovasti niin, että syntyy teoskynnyksen ylittävä dokumentoitu teos.

Eräänlainen poikkeus on tavaramerkki (Tavaramerkkilaki 10.1.1967/7), joka säätää tavaramerkkinä olevan sanan, nimen, kuvion tai muun sellaisen oikeuksista ja suojasta. Niiden suoja perustuu rekisteröintiin tai vakiintuneisuuteen. Tavaramerkin suoja ei tietenkään ulotu kielenkäyttöön. Kukaan ei voi kieltää käyttämästä tavaramerkkisanoja ja -nimiä kielen tasolla. Samantyyppinen on väestötietojärjestelmään merkittyjen ja muiden suojattujen sukunimien suoja. Niitä voi käyttää vapaasti tekstissä, mutta niitä ei voi omia omaksi nimekseen. (Nimilaki 9.8.1985/694: 12. §.)

On tietysti olemassa sanoja, joita vältetään tai kielletään käyttämästä (esimerkiksi uskonnollisia tunteita loukkaavat sanat tai rasistinen kielenkäyttö) ja joiden käytöstä pahimmillaan on joissakin kulttuureissa ja yhteiskunnissa menettänyt henkensä (esimerkiksi yksinvaltiaita solvaavat sanat). Kysymys ei tällöin ole kuitenkaan tekijänoikeudesta, vaan tapojen, moraalien, uskonnon, politiikan tai vallanpitäjien vaikutuksesta ja vallasta.

Tuntuu ehkä saivartelulta pohtia sitä, onko sanalla tekijänoikeutta. Yksittäisen sanan suojasta on kuitenkin käyty pitkä oikeusprosessi, joka ratkaistiin viimein vuonna 2005 Korkeimmassa oikeudessa. Kyseessä oli sana *turvaura* 'auton tuulilasiin hiottu ura, joka puhdistaa sen ylitse liikkuvan pyyhkimen sulan', jonka eräs yrittäjä lanseerasi kehittämänsä tuulilasinpyyhkimiä puhdistavan uran nimeksi ja rekisteröi myös yri-

tyksensä nimeksi. Kyseinen yritys vaati, että vastaajaa kielletään käyttämästä *turvaura*-sanaa ”missään muodossa ja hävittämään kaikki aineisto, johon on oikeudettomasti merkitty mainittu sana”. Lisäksi kantaja vaati vahingonkorvausta tavaramerkki- ja toiminimioikeuden loukkauksesta. Kaikki oikeusasteet hylkäsivät tavaramerkkilain perusteella kantajan vaatimukset. Korkeimman oikeuden perusteellisesti valmistellun päätöksen (KKO:2005:118) 39. kohdassa prosessi päättyy vastaajan eduksi: ”Sanan *turvaura* muunnuttua edellä kerrotuin tavoin yleiskielen sanaksi Suomen Autokatsastus Oy:llä on ollut oikeus käyttää sanaa liiketoimintansa laatua kuvaavana yleisenä ilmaisuna. Suomen Autokatsastus Oy ei ole siten markkinoidessaan tuotettaan nimellä *Prosecur-turvaura* ja käyttäessään toiminnassaan sanaa *turvaura* loukannut Turvaura Oy:n oikeutta toiminimeen.” Suomen kielen perussanakirjaan (PS, 1990–1994) *turvaura* pääsi vuonna 1994 arkistotietojen perusteella, joista ensimmäiset olivat 1970-luvulta. Oikeusprosessin aikana ilmestyneisiin sanakirjoihin sanaa ei otettu. Niinpä se puuttuu Kielitoimiston sanakirjan (KS, 2004–) vuonna 2006 ilmestyneestä painoksesta. (Kielitoimiston kirjeenvaihto 1999 ja 2000: dnot 45 ja 68/206.)

Tämä ennakkopäätös on kielenkäytön vapauden ja yleiskielen sanakirjojen kannalta merkittävä. Jos yleiskieliset sanat voisivat saada edellä mainitussa oikeustapauksessa vaaditun suojan vaikkapa vain tavaramerkin osanakin, kirjoittajien ja kustantajien olisi oltava kieli keskellä suuta, ettei tekstiin vain pujahtaisi suojattu sana, mikä pahimmillaan johtaisi korvauksiin ja teosten hävittämiseen. Lisäulottuvuuden tähän oikeustapaukseen toisivat mahdollisesti semanttisesti erilaiset käyttötavat. Voisihan kyseisellä sanalla olla teoriassa myös merkitykset ’turvallinen elämänura’ ja ’turvallinen toimintatapa’ (esim. ”hän valitsi leksikografian turvauran”; ”toimitus asettui turvauralle jättäessään sanan pois”). Kanne ei koskenut mahdollisia muita merkityksiä, joten Korkein oikeus ei ole ottanut kantaa *turvaura*-sanan semanttisiin ulottuvuuksiin.

Entä sitten uudissana, jonka todistettavasti joku on varta vasten seipittänyt tai kielentaitajana synnyttänyt, onko seipittäjällä siihen tekijänoikeuksia? Maallikkotulkintani mukaan uudissanojen luonti ei missään nimessä täytä tekijänoikeuslain ehtoja. Kuka tahansa kielenkäyttäjä pystyy periaatteessa luomaan oman kompetenssinsa avulla uusia sanoja ja käytännössä tekeekin näin. Tosin kovin harvat uudet sanat päätyvät loppujen lopuksi yleiseen kielenkäyttöön ja sitä tietä normatiivisiin sanakirjoihin.

Uudissanoillakin on kuitenkin tietty oikeutensa, lähinnä eettiseltä pohjalta lähtevä isyysoikeus. Aina ei uudissanan tekijää voi löytää tai kannata etsiäkään valtavasta tekstivirrasta, mutta varsinkin oppitekoisesti luoduilla termeillä on yksilöitävissä oleva keksijänsä. Esimerkiksi Helsingin yliopiston suomen kielen professori Lauri Hakulinen johti takaperoisesti *muovata*-verbistä sanan *muovi*, jota ei hevin enää uudissanaksi tunnistakaan. Kielitoimiston virkakielen tutkija Jussi Kallio puolestaan muodosti siirtolaisten sopeuttamista tarkoittavan sanan *kotouttaa* 1990-luvulla. Itse keksaisin muutama vuosi sitten suomen kielen vokaaliennätystä pitäneen *hääyöaie*-yhdyssanan päihittävän sanan *riiuuyöaieaatos*, jonka dokumentoin Hiidenkivi-lehdessä (Jussila 2012). Pidän itseäni sanan ”isänä” ainakin siihen saakka, kun se löytyy jostakin varhemmasta julkaisusta jonkun muun esittämänä.

Onko määritelmä suojattu?

On itsestään selvää, että kaksi- ja monikielisissä sanakirjoissa eivät erikieliset vastineet kuulu suojan piiriin. Kukaan ei luovasti ole keksinyt, että suomen *syksy* on ruotsin *höst*. Vasta siinä vaiheessa, kun näitä yhdistelmiä on koottu riittävä määrä, astuu voimaan kokonaisuuden – ei yksittäisten määritelmien – luettelosuoja. Ja kun näitä yhdistelmiä on järjestelty kyllin luovasti, alkaa teoskokonaisuuden varsinaisen tekijänoikeuden suoja koko laajuudessaan.

Yksikielisten sanakirjojen määritelmät eivät mielestäni luovalta muoltakaan nouse teoskynnyksen yli kuin erikoistapauksissa. Esimerkiksi Nykysuomen sanakirjan (NS, 1951–1961) *tuoli*-sanon määritelmä ’yhden henkilön istuttava jalallinen, tav. selkä-, us. myös käsinojallinen istuin’ ei nauttine suojaa varsinkaan, kun sen taustalla saattaa olla jo suomen- tai muunkielinen samanasuinen määritelmä. KS:aan sama määritelmä on siirtynyt ”vars.”-lyhenteen laajentamana. Ammattitaitoiset leksikografit päätyisivät todennäköisesti samaan tai samantyyppiseen määritelmään, mikä on osoitus siitä, ettei ole kyse tekijänoikeuslain tarkoittamasta luovuudesta.

Vaikka yksinkertaiset määritelmät eivät mielestäni voi saada tekijänoikeuden suojaa, sanakirjantoinittaja voi kärsiä tutkimuseettisiä omantunontuskia käyttäessään lähteitä sellaisenaan. Tämä johtaa suotta helposti

keinotekoiisiin muotoiluihin, jotka eivät ainakaan sanakirjan käyttäjän kannalta ole parannuksia.

Kyllä määritelmäkin saattaa kuulua suojan piiriin. Sellainen ryhmä ovat esimerkiksi termisanakirjat, joissa termin määritelmät ovat monimutkaisia ja muodoltaan pikemmin sisällöllistä asiantuntemusta vaativia itsenäisiä artikkeleita kuin tiiviitä muutaman sanan mittaisia leksikografisia määritelmiä. Niitä käytettäessä on syytä turvautua normaaliin sitaattitekniikkaan.

Luettelosuoja

Tekijänoikeuslaissa säädetään niin sanotusta luettelosuojasta, jonka saa se, ”joka on valmistanut 1) luettelon, taulukon, ohjelman tai muun sellaisen työn, jossa on yhdistelynä suuri määrä tietoa, taikka 2) tietokannan, jonka sisällön kerääminen, varmistaminen tai esittäminen on edellyttänyt huomattavaa panostusta” (TEO 49. §). Merkittävä ero tekijänoikeuteen on luetteloiden ja tietokantojen suoja-aika, joka on vain 15 vuotta luettelon valmistumisesta tai ilmestymisestä.

Luettelosuojan kynnyks on varsin matalalla. Esimerkiksi oppikirjaan sisältyvä sanaluettelo saa suojan. Korkeimman oikeuden päätöksessä (KKO:2005/43) todetaan: ”Kielen oppikirja oli käsittänyt erilliset tekstiosan ja sanaston. Yhtiö oli ilman oppikirjan tekijänoikeuden haltijoiden suostumusta kopioinut sanaston tietokoneella luettavissa oleville levykkeille ja ryhtynyt markkinoimaan levykkeitä. Korkeimman oikeuden tuomiossa mainituilla perusteilla yhtiön katsottiin menettelyllään loukanneen tekijänoikeuksia.” Vastaaja tuomittiin tuntuviin korvauksiin. Toisena esimerkkinä voisin mainita Helsingin yliopiston ylläpitämän nimipäiväluettelon. Se on todettu oikeusprosessissa tekijänoikeuslain suojaamaksi (KKO:2000:56). Listaa luvatta käyttänyt vastaaja tuomittiin korvauksiin. Yksittäiset nimipäivätiedot sen sijaan ovat vapaasti käytettävissä, koska niihin suoja ei yllä.

Tietokannan ja sen sisällöksi otetun sanakirjan suoja-ajoissa on sovitamattoman tuntuinen ristiriita. Suoja-ajoa ei voi tulkita pelkästään lyhyen luettelosuojan mukaisesti, sillä sisällöllä on teoksen pitkä suoja. Olisi kovin omituista, jos teoskynnyksen ylittävä teos menettäisi alkuperäisiä oikeuksiaan siinä lisäprosessissa, jossa siitä tuotetaan atk-pohjai-

nen käyttöliittymä. Tekninen muuntaminen ei siirrä lain mukaan sisältö-oikeuksia tekijöiltä muuntajalle.

Tiedossani ei ole oikeustapauksia, jotka koskisivat käänteis- ja taajuus-sanakirjoja. Äkkipäätä ajatellen periaatteessa kuka tahansa voisi tuottaa samanlaisen teoksen, jolloin niiden suoja olisi vain luettelosuojan tasoisia. Kuitenkin esimerkiksi otokseen, hakusanoitukseen ja sisältöanalyysiin vaikuttavat erilaiset ratkaisut tuottaisivat erilaisia tuloksia. Laajat käänteis- ja taajuus-sanastot ylittävät ainakin käyttäjiensä mielikuvissa teoskynnyksen. Lainkäyttäjän tulkinnasta ei voi olla etukäteen varma.

Ylittääkö sanakirja teoskynnyksen?

Sekä kustannus- että yliopistomaailmassa oli aikoinaan suullisena kansanperinteenä levinnyt käsitys, että sanakirjan tekeminen on arkistotietojen kopioimista aakkojärjestykseen. Tällainen teos nauttisi vain luettelosuojaa. Sanakirjojen kustantajat luopuivat tästä käsityksestä viimeistään siinä vaiheessa, kun huomasivat, että heidän suuria kustannuksia vaatineet sanakirjansa olisivat tällöin 15 vuoden jälkeen ilmestymisestä vapaita kenen tahansa julkaistaviksi. Yliopistopuolella käsitykset ovat sitä mukaa muuttuneet, kun sanakirjoja koskeva tieto ja tutkimus on lisääntynyt.

Oikeuskäytännön mukaan teoskynnys ylittyy silloin, kun kukaan toinen ei todennäköisesti tekisi teosta samalla tavoin. Tässä on ristiriita sanakirjojen toimitusten tavoitteisiin, jotka tähtäävät mahdollisimman johdonmukaisiin ja yhtenäisiin tuloksiin, kohteena olevan aineiston objektiiviseen esittämiseen. Monihenkisissä toimituksissa luovan työn keskeisiä ominaisuuksia, itsenäisyyttä ja omaperäisyyttä, pyritään itse asiaan suuitsimaan. Suuri sooloilija ei kirjoita ryhmässä hyvää sanakirjaa.

Vuonna 1982 Suomen murteiden sanakirjan (SMS, 1985–) toimituksessa päätettiin testata, onko toimitustyössä päästy ihanteeksi katsottuun yhdenmukaisuuteen. Järjestin kokeen, jossa koesanana oli *kuva*. Valitsin arkistoaineksista noin kymmenenneksen eli 120 numeroitua sanalippua edustamaan kaikkea *kuvan* käyttöä eri murteissa. Testissä SMS:n kymmenen toimittajan tuli laatia koeaineiston perusteella artikkelin luokitus ja merkityksenselitteet sekä merkitä numeroin, mitkä sanaliput kuuluivat mihinkin luokitusryhmään. Murteeseen kuulumattomat vieraat käytöt tuli karsia.

Testin tulos oli sekä testaajalle että toimitukselle yllätys. Toimitusohjeista ja lukemattomista toimituksen kokouksista huolimatta toimituksellista yhtenäisyyttä vastauksissa ei ollut. Tyrmäävän vastaanoton takia en milloinkaan analysoinut tuloksia sen tarkemmin julkaisemisesta puhumattakaan. Kommentoin niitä hieman nyt kolmenkymmenen vuoden jälkeen.

Yhtenäisintä oli konkreettisten merkitysten määrittäminen (esim. merkitykset 'viehe', 'pelikorttien kuvakortti', 'houkutuslintu'). Mitä abstraktisemmasta käytöstä oli kyse, sitä varioivampaa oli sekä luokittelu että määritelmät. Luokitusryhmiä oli vastauksissa kolmestakymmenestä viiteenkymmeneen ja niiden hierarkkinen rakennekin varioi. Eroja oli myös siinä, mitä tulkittiin fraasiksi, mitä yleiskieliseksi tai kontekstiltaan liian hämäräksi. Toimitus tulkitsi muistamani mukaan tuloksen ohjeiden, koulutuksen ja kykyjen puutteeksi, ja tulos aiheutti ilmeisesti joissakuissa suurta ahdistusta.

Myöhemmin olen alkanut tulkita tulosta aivan toisin. Testi kertoo siitä, että laajojen artikkeleiden kirjoittaminen on luovaa työtä. Kukaan ei todennäköisesti päädy täsmälleen samaan tulokseen kuin joku toinen. Vaikka leksikografisia välineitä ja menetelmiä yhdenmukaistetaan, niillä syntyy kuitenkin erilaista jälkeä. Leksikografia ei ole mekanisoitavissa, se on luovaa kirjoittamista.

Erilaisuutta aiheuttavaa luovaa työtä on ainakin seuraavissa leksikografisissa vaiheissa: 1) arkistoaineiston lähdekritiikki, 2) sanakirjaan tulevien sanojen valinta, 3) sanakirjaan tulevien esimerkkien valinta tai luominen, 4) aineiston leksikografinen luokittelu, 5) määritelmien laatiminen, 6) fraasiaineiston valinta ja tulkinta, 7) tyyllisävyjen arviointi. Tilanne on samantyyppinen kuin muussakin humanistisessa tutkimuksessa. Samasta aineistosta samoinakin metodein syntyy erilaisia tulkintoja.

Sanakirjat koostuvat usein kovin erikokoisista ja -tasoisista artikkeleista. Viiteartikkelit ovat luettelotyypisiä, ja niitä voidaan ja pitäisikin tehdä mekaanisesti. Lyhyet artikkelit näyttävät rutiininomaisilta, mutta niidenkin taakse kätkeytyy usein laajoja aineistoja ja harkintaa. Suuret artikkelit vaativat sekä ammattitaitoa että luovuutta. Kaikkien artikkeleiden kokonaisuudesta muodostuu teos, joka on tekijänoikeudella suojeltu.

Sanakirjojen lähteet

Sanakirjoja siteerataan, mutta sanakirjan – ainakin tietosanakirjan – toimittajien väitetään plagioivan. Siteeraus on lainmukaista: ”Julkistetusta teoksesta on lupa hyvän tavan mukaisesti ottaa lainauksia tarkoituksen edellyttämässä laajuudessa.” (TEO 22. §.) Sanakirjassa ”tarkoituksen edellyttämä laajuus” tuskin milloinkaan ylittyy, koska tavoitteena on mahdollisimman tiivis, mutta sanan käyttöä kyllin selvästi edellyttävä esimerkki. Voisi mieluummin moittia sanakirjoja liian lyhyestä siteeraamisesta, jolloin sitaattikatkelma ei anna täyttä kuvaa kirjailijan tarkoituksista.

Ongelmallisempi on käsite ”hyvä tapa”, sillä lähdeviitteissäkin sanakirjat pyrkivät selvyyteen ja lyhyteen luettavuuden takia. NS:ssa viitataan Raamatun, Kalevalan ja vastaavien yleisteosten lisäksi noin viiteenkymmenen kaunokirjailijaan sukunimellä tai sen lyhenteellä teosta yksilöimättä. Tapa oli ilmeisesti kirjailijoiden kannalta hyvä, koska se nosti listalle hyväksytyjen kirjailijoiden statusta. Yleiskielen myöhemmissä sanakirjaversioissa sitaattikäytäntöä on muutettu niin, että kirjallisuussitaatteja niistä pitää etsimällä etsiä. Käytäntöä ei olisi tarvinnut ainakaan tekijänoikeuslain takia muuttaa.

Sanakirjojen taustalla on tavallisesti laajoja arkistoaineistoja. Niitä ei ole tapana tulkita teoksiksi, joita tekijänoikeus koskisi. Kuitenkin on muistettava, että jos tekijänoikeuden suojaamaa aineistoa (esim. kirjeitä, sitaattikatkelmia) arkistoidaan, ne säilyttävät alkuperäisen tekijänoikeutensa. Ne ovat arkistonmuodostajan omaisuutta vain fyysisinä olioina, kuten kirja on ostajansa omaisuutta. Jos muodostaja on ollut yhteiskunnan tai säätiöiden rahoitusta saanut yhteisö, arkisto on jo alkuaankin tarkoitettu vapaaseen käyttöön ainakin tutkimustarkoituksia varten. Arkiston antama käyttö lupa ei vapauta noudattamasta tekijänoikeutta, vaan esimerkiksi siteerausta ja lähteen mainitsemista koskevia säädöksiä on noudatettava.

Otan esiin yhden esimerkkitapauksen arkistojen käyttöoikeusongelmista. Kotikielen seura järjesti 1970-luvulla laajan Helsingin slangisanaston keruun. Keruuvaiheessa ilmoitettiin, että aineisto on tutkijoiden vapaassa käytössä. Kun Kaarina Karttunen pyysi aineistoa Nykyslangin sanakirjaansa (1979) varten, seura kieltäytyi sitä luovuttamasta. Aineisto oli varattu Helsingin slangin sanakirjatyötä varten. Tämä slangiaineisto ei ole tekijänoikeuslain suojaamaa, vaan käyttörajoituksista päätti Koti-

kielen seura. Tutkimuseettisesti oli kuitenkin arveluttavaa, että keruu-
vaiheessa vapaaksi ilmoitettu aineisto kuitenkin myöhemmin suljettiin
muiden tutkijoiden saatavilta. Tekijänoikeuden näkökulmasta tässä ta-
pauksessa rajoitettiin tekijöiden eli slangikilpailuun osallistuneiden slan-
gintaitajien antamaa ”hiljaista” lupaa aineiston vapaaseen käyttöön kysy-
mättä jälkikäteen tekijöiltä tähän lupaa.

Tutkimuksen ja sanakirjojen aineistoina käytetään myös vapaasta
puheesta tallennettua kieltä. Kenellä on sen tekijänoikeus? Tavallises-
ti murreaineistoja käytetään kielennäytejulkaisujen pohjalta. Silloin
mahdollisia primaarien tekijänoikeuksien haltijoita voisi olla kolmekin:
1) kielenpuhuja, jonka kieltä on äänitetty (ja jonka oikeudet ovat jo ny-
kyään useimmiten perikunnalla), 2) haastattelija, joka on näyttteen äänit-
tänyt ja 3) litteroija, joka on kirjaintanut näyttteen ja luovuttanut sen jul-
kaistavaksi. Miten sanakirjassa voi käyttää näitä aineistoja tekijänoikeutta
rikkomatta?²

Käytännössä on tulkittu niin, että haastattelutilanteessa kielenpuhuja
myöntymällä haastatteluun on antanut luvan äänitteen arkistointiin ja
edelleen tutkimuskäyttöön. Haastattelija on tavallisesti saanut työstään
apurahan tai palkkaa, jonka edellytyksenä on ollut aineiston luovuttami-
nen arkistolle. Äänitteestä syntyy jo haastattelu- tai arkistointivaiheessa
teos, joka nauttii tekijänoikeutta; tämä ainoa teoskappale ja sen kopiot
ovat arkistonmuodostajan fyysistä (ei henkistä) omaisuutta ja sen hallin-
nassa. Litteroijan julkaisema kielennäyte on tekijänoikeuden alainen teos,
joka tavallisesti tulkitaan vain litteroijan hengentuotteeksi. Litterointi ei
kuitenkaan hävitä alkuperäisiä oikeuksia, joten kielennäytteissä tekijyys
on myös kielenpuhujalla ja haastattelijalla. Heidän tekijyytensä ei sisällä
enää taloudellisia oikeuksia, sillä he ovat luovuttaneet ne joko ilmaiseksi
(suostumalla haastatteluun ilman palkkiota) tai palkallisesti arkistolle.

Kun kirjallisuudesta poimitaan esimerkkejä arkistoon, syntyneen ar-
kistoaineuksen tekijänoikeudet eivät siis siirry arkistonmuodostajalle,
vaan oikeudet säilyvät teoksen alkuperäisellä tekijällä tai tekijöillä. Nämä
arkistomuistiinpanot (olivatpa ne sitten paperilla tai tietokannassa) on
tulkittava sitaateiksi, joita on käsiteltävä tekijänoikeuslain mukaan. Niitä
voi korvauksetta käyttää, mutta ei lähdeä mainitsematta.

2 Lauri Honko (2002) on pohtinut laajasti folkloren tekijänoikeuskysymyksiä artikkelis-
saan ”Kenen omaisuutta on suullinen perinne?”

SMS:ssa murre-esimerkkien lähdeviitteitä ei ole otettu mukaan. Osa aineistosta ei tarvitsekaan välttämättä lain perusteella viitettä, sillä se on (osittain anonyymiakin) tutkimuksen perusaineistoa, jonka käytöstä päättää arkistonmuodostaja. Mukana on kuitenkin myös kirjallisuudesta, esimerkiksi edellä mainituista kielennäytejulkaisuista poimittuja esimerkkejä, jotka ovat tekijäsuojan alaista aineistoa. Niiden siteeraaminen edellyttää lain mukaan lähteen mainitsemista.

SMS:ssa noudatettu ”hyvä tapa” on se, että kaikkien esimerkkien lähteet ovat tallella murrearkistossa ja siellä jäljitettävissä. Näin suuressa hankkeessa lähteiden painaminen olisi paisuttanut kohtuuttomasti sanakirjaa. Noudatettu käytäntö on tutkijakunnan hiljaisesti hyväksymä ja siinä mielessä varmastikin ”hyvän tavan” mukainen. Aikoinaan Suomen kielen etymologisessa sanakirjassa (SKES, 1955–1981) jätettiin kaikki lähdeviitteet pois, minkä vuoksi sitä kritisoitiin tiedeyhteisössä. Toki sen arkisto- ja kirjallisuuslähteet oli lueteltu johdannossa ja taustaaineisto löydettävissä arkistossa. Sen jatkajassa Suomen sanojen alkuperä-teoksessa (SSA, 1992–2000) tärkeimmät lähteet koottiin kunkin artikkelin perään, mikä on toimiva ja korrekti käytäntö tämäntyyppisessä sanakirjassa. Eettisesti etymologian ensimmäisen esittäjän mainitseminen lähteissä on aiheellista.

Sanakirjan tekijöiden oikeudet

Yhden tai muutaman tekijän lyhytkestoisissa sanakirjoissa tekijöiden oikeudet ovat varsin yksinkertaisesti määriteltävissä. Toisin on asia monikymmenvuotisissa hankkeissa, joissa vuosien mittaan ennättää työskennellä kymmeniä ihmisiä ja joissa teoksia päivitetään jatkuvasti korjatuin ja täydennetyin painoksin. Tyypillisin esimerkki on NS:sta alkava nykykielen yleissanakirjojen sarja.

Lähtökohtana on se, että kaikki oikeudet syntyvät tekijöille. Aiemmin ei tietääkseni sanakirjantöimittäjien työsopimuksiin ollut kirjattu mitään tekijänoikeuksien luovutuksesta. Kuitenkin jokainen toimittaja tiesi, että hänen tekstinsä oli tarkoitettu työnantajan julkaistavaksi, ja hän oli hiljaisesti hyväksynyt tämän eli hän oli näin tehnyt niin sanotun konkludenttisen sopimuksen. Nykyään on oikeuksien siirto tavallisesti kirjattu työsopimukseen.

Sanakirjatyö on tyypillisesti yhteistyötä. Toimittajan tekstiä ruoditaan toimituksen yhteiskokouksissa. Se käy läpi toimitussihteerin ja päätoimittajan tarkan seulan. Tekijänoikeuslaki säättää 7. pykälässään: ”Jos kaksi tai useammat ovat yhdessä luoneet teoksen heidän osuuksiensa muodostamatta itsenäisiä teoksia, on tekijänoikeus heillä yhteisesti. Kullakin heistä on kuitenkin valta vaatimusten esittämiseen oikeuden loukkauksen johdosta.” Tämä merkitsee, että kaikki teoksen toimittajat on merkittävä tosiasiallisesta työosuudestaan riippumatta tasavertaisiksi tekijöiksi toki päätoimittajuus ja toimitussihteriys erikseen mainiten.

Kun sanakirjasta otetaan muutettu painos, eivät pohjana olevan teostekstin oikeudet siirry uusille tekijöille. Heille kuuluu tekemiensä täydennyksien ja muutosten tekijyys. Lisäksi he saavat teoskokonaisuuden tekijyyttä siksi, että teoksesta poistetaan tarpeettomiksi katsottuja artikkeleita. Laki toteaa tästä tilanteesta (TEO 5. §): ”Sillä, joka yhdistämällä teoksia tai teosten osia on aikaansaanut kirjallisen tai taiteellisen kokoomateoksen, on siihen tekijänoikeus, mutta hänen oikeutensa ei rajoita oikeutta ensimmäintuihin teoksiin.”

Jos NS:n ja sen seuraajien tekijät luelteltaisiin uusissa kielitoimiston sanakirjoissa, tekijöiden listasta tulisi kovin pitkä. Sinänsä tällaisen luettelon julkaiseminen ei olisi mahdotonta, sillä mainitaanhan jokaisessa elokuvassa kymmenittäin tekijöitä aina viimeiseen pukuompelijaan asti. Tekijöiden ilmoittaminen on esimerkiksi KS:ssa (2006) hoidettu niin, että varsinaisen toimituksen jäsenten lisäksi on ilmoitettu sen sanakirjan toimitus, johon uusi versio perustuu. Tällainen menettely on mielestäni asiallinen.

Sanakirjan julkaisijan kannalta tärkeitä ovat muuttamis- ja muunteluoikeudet. Edellinen tuottaa tekijälle ”yksinomaisen oikeuden määrätä teoksesta – – muuttamattomana tai muutettuna, käännöksenä tai muunnelmana, toisessa kirjallisuus- tai taidelajissa taikka toista tekotapaa käyttäen” (TEO 2. §). Muunteluoikeus sisältyy lain neljänteen pykälään (TEO 4. §), joka säättää teoksen muuntamisesta toiseen esitysmuotoon tai formaattiin. Näiden molempien oikeuksien siirtymisestä julkaisijalle on syytä sopia työ sopimuksissa. Olisi sietämätöntä niin julkaisijan kuin toimituksen kannalta, jos yksi toimittaja voisi riitauttaa vaillinaisten sopimuksen takia sanakirjan uusien versioiden päivittämisen.

Tekijänoikeuslain lukuisista yksityiskohdista voisi mainita vielä respektioikeuden, joka kieltää muuttamasta teosta ”tekijän kirjallista tai

taiteellista arvoa tahi omalaatuisuutta loukkaavalla tavalla” (TEO 3. §). Sanakirjoissa tällainen loukkaus on kovin epätodennäköinen. Ehkä toimittajan tutkimusparadigmasta johtuvia näkemyksiä saatetaan uusissa versioissa loukata, mutta silloin ei ole kyse niinkään tekijänoikeudesta vaan tutkimusetiikasta.

Työsuhteisen sanakirjantoimittajan ei tarvitse piitata taloudellisista oikeuksista, sillä ne ovat siirtyneet hiljaisissa sopimuksissa tai työsopimuksissa työnantajalle. Omaa sanakirjaa tekevän on huolehdittava kaikista oikeuksistaan muiden kirjailijoiden tapaan. Apua saa kirjailijajärjestöistä ja erilaisista opaskirjoista (ks. esim. Jokinen 2010).

Summaus

Tekijänoikeus ei aseta sanakirjoja mihinkään erityisasemaan, ei edes mainitse niitä. Lain soveltamisessa sanakirjoihin on kuitenkin erityispiirteitä, jotka sekä julkaisijoiden että sanakirjojen tekijöiden on syytä tunnistaa. Leksikografien on syytä kiinnittää erityishuomiota siteeraukseen, lähteiden käyttöön ja merkitsemiseen sekä omiin tekijänoikeuksiinsa. Sanakirjojen julkaisijoiden puolestaan on huolehdittava, että sanakirjojen jatkokäyttö on sovittu kestäväällä ja leksikografien oikeuksia kunnioittavalla tavalla.

Tekijänoikeus suojelee paradoksaalisesti yhtä lailla huuhaata ja propagandatekstejä kuin sanakirjoja ja tutkimusartikkeleita, sillä sen suojeleuksessa on vain luova muoto. Sisällön ja tiedon se säätää nimenomaan vapaaksi. Tämä vapaus on vapaan tutkimuksen elinehto. Sen varassa tieto voi elää ja kasvaa.

Sisällön ja tiedon laatuun ja tuottamiseen vaikuttavat asiat ovat tutkimusetiikan alaa. Sitä varten tarvitaan sääntöjä ja käytäntöjä, joita tutkijakunnan tulee noudattaa. Tutkimusetiikka on tutkijoiden moraalinen varassa. Sillä on tukenaan tekijänoikeuslaki, jonka välineinä ovat lain sanktiot.

Lähteet

- Agricolan sanakirja [tulossa]. Suomen kieli ja suomalais-ugrilainen kielentutkimus, Turun yliopisto, Turku.
- Ahoniemi, Petri 2010: Työläisille kunnia ja paskaduuni sanana pannaan. Iltalehden blogit 12.11.2010. Saatavissa: <http://blogit.iltalehti.fi/petri-ahoniemi/2010/11/12/tyolaisille-kunnia-ja-paskaduuni-sanana-pannaan/>. [Viitattu 20.9.2012.]
- Airaksinen, Timo 1993: Ammattien etiikan filosofiset perusteet. *Ammattien ja ansaitsemisen etiikka*, s. 19–60. 3. painos. Toim. Timo Airaksinen. Yliopistopaino, Helsinki.
- Algeo, John 1989: Dictionaries as Seen by the Educated Public in Great Britain and the USA. *Dictionaries: an international encyclopedia of lexicography*, Volume 1, s. 28–34. Toim. F. J. Hausmann – Oskar Reichmann – Herbert E. Wiegand – Ladislav Zgusta. Walter de Gruyter, Berlin.
- Ammattietiikka yliopistoissa: Eväitä työelämään*. Toim. Liisa Myyry. Ammattieettisen opetuksen kartoittamis- ja kehittämisprojekti. Loppuraportti. Helsingin yliopisto – Helsingin kauppakorkeakoulu – Teknillinen korkeakoulu, Helsinki 1999.
- Anhava, Jaakko 1998: *Maailman kielet ja kielikunnat*. Gaudeamus, Tampere.
- Arktinen keskus = Asutus arktisilla alueilla: Alkuperäiskansan määrittäminen ja ILO-sopimus. Arktinen keskus, Lapin yliopisto. Saatavissa: www.arcticcentre.org/?DeptID=7767. [Viitattu 19.10.2012.]
- Aseksuaalit.net = Aseksuaalit.net. *Suomen aseksuaalikeidas*. Saatavissa: <http://aseksuaalit.net>. [Viitattu 20.9.2012.]
- Asetus lasten päivähoitosta. 16.3.1973/239, ks. Finlex.
- Atkins, B. T. Sue 2002: Bilingual Dictionaries: Past, Present and Future. *Lexicography and Natural Language Processing. A Festschrift in Honour of B. T. S. Atkins*, s. 1–29. Toim. Marie-Hélène Corréard. S. I., EURALEX.
- Atkins, B. T. Sue – Rundell, Michael 2008. *The Oxford Guide to Practical Lexicography*. Oxford University Press, Oxford.
- Baider, Fabienne H. 2007: The Death of the Author, the Birth of the Lexicographer: How French Historical Dictionaries Construct History. *International Journal of Lexicography* 20(1), s. 67–83.
- Baigent, Elizabeth – Brewer, Charlotte – Larminie, Vivienne 2005: Gender in the archive: women in the Oxford Dictionary of National Biography and the Ox-

- ford English Dictionary. Journal of the British Records Association, *Archives* 30(113), s. 13–35. Saatavissa myös: <http://oed.hertford.ox.ac.uk/main/images/stories/articles/baigent2005.pdf>.
- Bakker–Kyuchukov 2000 = *What is the Romani language? Centre de recherches tsiganes*. Toim. Peter Bakker – Hristo Kyuchukov. University of Hertfordshire Press, Hertfordshire 2000.
- Béjoint, Henri 1994: *Tradition and Innovation in Modern English Dictionaries*. Clarendon Press, Oxford.
- 2000: *Modern Lexicography: An Introduction*. Oxford University Press, Oxford.
- Beke, Ödön 1997–2001: *Mari nyelvjárási szótár (Tscheremissisches Dialektwörterbuch) I–IX*. Unter Mitarbeit von Zsófia M. Velenyák und József Erdödi. Neu redigiert von Gabor Bereczki. Bearbeitet von Margarita Kuznecova. Herausgegeben von János Pusztay. Bibliotheca ceremissica, tomus IV. Berzsényi Dániel Főiskola, Savariae.
- Bergenholtz, Henning – Malmgren, Sven-Göran 2006: Historiske ordbøger. *Lexico-Nordica* 13, s. 1–5. Saatavissa: www.nordisk-sprakrad.no/dokumenter/Lex-pdf/LexicoNordica%202006.pdf. [Viitattu 20.11.2012.]
- Bergenholtz – Tarp 1995 = *Manual of Specialised Lexicography. The Preparation of Specialised Dictionaries*. Toim. Henning Bergenholtz – Sven Tarp. Benjamins, Amsterdam 1995.
- Bergman, Jay 1998: Was the Soviet Union Totalitarian? The View of Soviet Dissidents and the Reformers of the Gorbachev Era. *Studies in East European Thought* 50(4), s. 247–281.
- Berkov 1971 = Берков, В. П. 1971: *Вопросы двуязычной лексикографии: Автореф. дис. ... д-ра филол. наук*. Ленинград.
- Betongteknisk ordlista*. TNC 103. Terminologikum TNC, Solna 2012.
- Blomstedt, Yrjö 1987: *Kuningas Kristoferin maanlaki Herra Martin suomentamana*. Valtion painatuskeskus, Helsinki.
- Braun, Virginia – Kitzunger, Celia 2001: Telling it straight? Dictionary definitions of women's genitals. *Journal of Sociolinguistics* 5(2), s. 214–232.
- Bugge, Sophus 1858: Vermischtes aus der Sprache der Zigeuner. *Beiträge zur vergleichenden Sprachforschung auf dem Gebiete der arischen, celtischen und slawischen Sprachen*, s. 143–155. Toim. A. Kuhn – A. Schleicher. Dümmlers, Berlin.
- Busse, Ulrich 2000: Recent English Learners' Dictionaries and their Treatment of Political Correctness. *Symposium on Lexicography IX. Proceedings of the Ninth International Symposium on Lexicography April 23–25, 1998 at the University of Copenhagen*, s. 165–201. Toim. Jens Erik Mogensen – Viggo Hjørnager Pedersen – Arne Zettersten. Lexicographica. Series Maior 103. Niemeyer, Tübingen.
- Cameron, Deborah 1996: *Sukupuoli ja kieli. Feminismi ja kielentutkimus*. Alkuteos *Feminism & Linguistic Theory* 1992 [1985]. Käännöstyöryhmää ohjannut Riitta Oittinen. Vastapaino, Tampere.
- Carr, Michael 1994: Yamato-Damashii 'Japanese Spirit' Definitions. *International Journal of Lexicography* 7(4), s. 279–305.
- CD-*Perussanakirja*. Päätoim. Risto Haarala – Marja Lehtinen. KTKJ 94. Edita, Hel-

- sinki 1997.
- Clarkeburn, Henriikka – Mustajoki, Arto 2007: *Tutkijan arkipäivän etiikka*. Vastapaino, Tampere.
- COBUILD 1987 = *Collins Cobuild English Language Dictionary*. Toim. John Sinclair ym. Collins, London 1987.
- COBUILD 1995 = *Collins Cobuild English Dictionary*. HarperCollins, London 1995.
- DDUW = *Duden. Deutsches Universalwörterbuch*. 6., überarbeitete und erweiterte Auflage. Mannheim 2007.
- Denisov 1976 = Денисов, П. Н. 1976: *Основные проблемы теории лексикографии: Автореф. дис. ... д-ра филол. наук*. Москва.
- Deutsches Wörterbuch = *Das Deutsche Wörterbuch von Jacob Grimm & Wilhelm Grimm*. 1852–1961. Neubearbeitung/herausgegeben Leipzig 1983 und Stuttgart 1998.
- Devkin 2000 = Девкин, В. Д. 2000: *Очерки по лексикографии*. Прометей, Москва.
- Dubičinskij 1994 = Дубичинский, В. В. 1994: *Искусство создания словарей: конспекты по лексикографии: Учебное пособие*. Харьковский политехнический университет, Харьков.
- 1995 = Дубичинский, В. В. 1995: Антиномии лексикографии. *Vocabulum et vocabularium: Сборник научных трудов по лексикографии*. Выпуск 2.
- Dubois, Jean – Dubois, Claude 1971: *Introduction à la lexicographie: le dictionnaire*. Larousse, Paris.
- DWB ks. *Deutsches Wörterbuch*.
- Dziemianko, Anna 2010: Paper or electronic? The role of dictionary form in language reception, production and the retention of meaning and collocations. *International Journal of Lexicography* 23(3), s. 257–273.
- Dölle, Sirkku 2004: [henkilökohtainen tiedonanto 19.1.2004.]
- Edlund, Ann-Catrine – Erson, Eva – Milles, Karin 2007: *Språk och kön*. Norstedts Akademiska Förlag, Stockholm.
- Edu.fi: Sanakirjoja = Sanakirjoja < Kielen opetus ja oppiminen < Maahanmuuttajien koulutus < Yleissivistävä koulutus < Edu.fi – Opettajan verkkopalvelu. Opetushallitus, Helsinki. Saatavissa: www.edu.fi/yleissivistava_koulutus/maahanmuuttajien_koulutus/kielen_opetus_ja_oppiminen/sanakirjoja. [Päivitetty 16.8.2010/viitattu 27.3.2012.]
- Eesti keele seletav sõnaraamat*. Eesti Keele Sihtasutus, Tallinn 2009. Saatavilla myös verkkosanakirjana: www.eki.ee/dict/ekss/. Eesti Keele Instituut, Tallinn 2009.
- Eesti õigekeelsussõnaraamat ÕS* 2006. Eesti Keele Sihtasutus, Tallinn.
- Eiras, Vuokko 1996: Etymologisesta periaatteesta ja sen soveltamisesta etäsuokukielten sanakirjatöissä. *Kielen kannoilla. Kotimaisten kielten tutkimuskeskus 20 vuotta*, s. 283–291. KKTK – Edita, Helsinki.
- 2010: Artturi Kanniston vogulilaisaineistot, niiden käyttö ja merkitys. *Kenttäretkistä tutkimustiedoksi*, s. 139–151. Toim. Paula Kokkonen – Anna Kurvinen. Uralica Helsingiensia 4, SUS, Helsinki.
- EKSS ks. *Eesti keele seletav sõnaraamat*.
- Epple, Barbara 2000: Sexismus in Wörterbüchern. *Proceedings of Euralex 2000*, s.

- 739–754. Saatavissa: www.euralex.org/elx_proceedings/Euralex2000/085_Barbara%20EPPL_ Sexismus%20in%20Worterbuchern.pdf. [Viitattu 15.9.2012.]
- Eronen, Riitta 2004: Sanat kirjoissa. *Kielikello* 4/2004, s. 3.
- Ethnologue.com = Ethnologue. Languages of the World. SIL International, Dallas. Saatavissa: www.ethnologue.com/. [Viitattu 4.4.2012.]
- Examining the OED.uk: Outline material = Outline material < Initial Results < Examining the OED. Charlotte Brewer. Saatavissa: <http://oed.hertford.ox.ac.uk/main/content/category/10/35/124/>. [Viitattu 4.4.2012.]
- Fairclough, Norman 1992: *Discourse and social change*. Polity Press, Cambridge.
- FD = Kramarae, Cherie – Treichler, Paula A. 1985: *A Feminist Dictionary. In Our Own Words*. Pandora Press, Boston.
- Fedotov 1965 = Федотов, М. Р. 1965: *Исторические связи чувашского языка с языками угро-финнов Поволжья и Перми. Часть I. Чувашско-марийские связи*. Чувашское книжное издательство, Чебоксары.
- 1968 = Федотов, М. Р. 1968: *Исторические связи чувашского языка с волжскими и пермскими финно-угорскими языками. Часть II*. Чувашское книжное издательство, Чебоксары.
- Feminisms = *Feminisms*. Toim. Sandra Kemp – Judith Squires. Oxford Readers. Oxford University Press, Oxford 1997.
- Finlex. Tietokanta. Oikeusministeriö – Edita Publishing Oy, Helsinki. Saatavissa: www.finlex.fi.
- Fischer, David Hackett 1971: *Historians' Fallacies. Toward a Logic of Historical Thought*. Routledge – Kegan Paul, Lontoo.
- Fjeld, Ruth Vatvedt 1994: Om ordbokskriminalitet og etikk i leksikografisk arbeid. *LexicoNordica* 1, s. 27–42.
- Florinus, Henricus 1976 [1678]: *Nomenclatura Rerum brevissima Latino-Sveco-Finonica*. Näköispainos. SKS, Helsinki.
- FO = *Ordbok över Finlands svenska folkmål*. KKTK – Svenska litteratursällskapet i Finland, Helsinki 1982–2007.
- Gak 1977 = Гак, В. Г. 1977: О некоторых закономерностях развития лексикографии (учебная и общая лексикография в историческом аспекте). *Актуальные проблемы учебной лексикографии*, s. 11–27. Русский язык, Москва.
- Galkin 1966 = Галкин, И. С. 1966: *Историческая грамматика марийского языка. Морфология, часть II*. Марийское книжное издательство, Йошкар-Ола.
- Ganander, Kristfrid 1780: *Undersökning om De så kallade TATTERE eller Zigeuner, Cingari, Bohemiens, Deras härkomst, Lefnadssätt, språk m.m. Samt om, när och hwarest några satt sig ner i Swerige?* Käsikirjoitus. Vitterhetsakademien, Stockholm.
- Gerd 2005 = Герд, А. С. 2005: Закон об авторском праве, понятие «плагиат» и академическая лексикография. *Ялик. Информационный бюллетень Санкт-Петербургского государственного университета*, № 64, s. 1–3.
- Gordeev 1979 = Гордеев, Ф. И. 1979: *Этимологический словарь марийского языка, том 1 А–Б*. Марийское книжное издательство, Йошкар-Ола.

- 1983 = Гордеев, Ф. И. 1983: *Этимологический словарь марийского языка, том 2 В–Д*. Марийское книжное издательство, Йошкар-Ола.
- Grafia = Tekijänoikeuspolitiikka 2012. Lausunto tekijänoikeuspolitiikka 2012 -keskustelumuistiosta. Grafia ry, Helsinki. Saatavissa: http://www.minedu.fi/export/sites/default/OPM/Tekijaenoikeus/tekijaenoikeuden_kehittaaminen/linjaukset/keskustelumuistiopalautteet/Grafia_ry.PDF. [Viitattu 27.8.2013.]
- Granger, Sylviane – Paquot, Magali 2009: Customising a General EAP Dictionary to Meet Learner Needs. *Proceedings of eLex 2009, Louvain-la-Neuve, 22–24 October 2009*, s. 87–96. Toim. Sylviane Granger – Magali Paquot. Presses universitaires de Louvain, Louvain.
- Granqvist, Kimmo 1997: Suomi–romani-sanaluettelo. [Moniste]. Helsinki.
- 1999: *Suomen romanikielen käänteissanasto. Reverse Lexicon of Finnish Romani*. KTKJ 111. KTKJ, Helsinki.
- 2012: Kuinka tarkastella Suomen romanikielen attritiota. Esitelmä Kielitieteen päivillä, Tallinnan yliopisto 17.5.2012.
- Granqvist, Kimmo – Pirttisaari, Helena 2003: Suomen romanin leksikaalisista kerrostumista ja niiden fonologisista ja morfologisista prosesseista. *Puhe ja kieli* 23(1), s. 25–41. Puheen ja kehityksen tutkimuksen yhdistys, Helsinki.
- Granqvist, Kimmo – Viljanen, Anna Maria 2002: Kielelliset tabut romani-identiteetin kuvaajana. *Moniääninen Suomi*, s. 109–125. Toim. Sirkka Laihala-Kankainen – Sari Pietikäinen – Hannele Dufva. Soveltavan kielentutkimuksen laitos, Jyväskylän yliopisto, Jyväskylä.
- Gray, John 1989: *Liberalisms: Essays in Political Philosophy*. Routledge, London.
- Grigoryan, Ashot A. 2007: On Gender Aspect in Lexicography. *Essays on Lexicon, Lexicography, Terminography in Russian, American and Other Cultures*, s. 2–11. Toim. Olga Karpova – Faina Kartashkova. Cambridge Scholars Publishing, Newcastle, UK.
- Grinev 1986 = Гринев, С. В. 1986: *Введение в терминологическую лексикографию: (Учебное пособие)*. Москва.
- 1995 = Гринев, С. В. 1995: *Введение в терминографию*. Изд. 2-е, испр. и доп. Московский педагогический университет, Москва.
- Grönfors, Martti – Virolainen, Kari – Åkerlund, Tuula – Lounela, Paavo 1997: Suomen romaniväestö: unohdettu kulttuuriryhmä. *Suomen kulttuurivähemmistöt*, s. 148–184. Toim. Juha Pentikäinen – Marja Hiltunen. Suomen Unesco-toimikunnan julkaisuja No 72. Yliopistopaino, Helsinki.
- Grönros, Eija-Riitta 1994: Sävyt ja vivahteet. Sanojen tyylilajin osoittaminen sanakirjassa. *Kielikello* 3/1994, s. 11–15.
- 2004: Kielitoimiston sanakirja. *Kielikello* 4/2004, s. 4–8.
- 2007: Mies ja nainen sanakirjassa. *Kielikello* 4/2007, s. 13–15.
- 2011: Sanakirjat ovat avain kielen ytimeen. *HS* 30.7.2011. Saatavissa myös: www.hs.fi/paakirjoitus/artikkeli/Sanakirjat+ovat+avain+kielen+ytimeen/HS20110730SI1MA01e9f.
- 2012a: Mitä uutta Kielitoimiston sanakirjassa? *Kielikello* 1/2012, 4–6.
- 2012b: Sanakirjatyo – leksikografia. *Genreanalyysi – tekstilajitutkimuksen kä-*

- sikirja, s. 631–635. Toim. Vesa Heikkinen – Eero Voutilainen – Petri Lauerma – Ulla Tiililä – Mikko Lounela. KKTKJ 169. Gaudeamus, Helsinki.
- Guentherodt, Ingrid – Hellinger, Marlis – Trömel-Plötz, Senta – Pusch, Luise F. 1980: Richtlinien zur Vermeidung sexistischen Sprachgebrauchs. *Linguistische Berichte* 69, s. 15–21.
- Gummeruksen suuri suomen kielen sanakirja. Kolmas, tarkistettu ja päivitetty painos. Toim. Timo Nurmi. Gummerus, Helsinki 2004.
- Gurevich, Aaron I. 1994: The Double Responsibility of the Historian. *The Social Responsibility of the Historian*, s. 65–83. Toim. François Bédarida. Berghahn, Providence.
- Haapanen, Minna 2008: Viherpesu, pelittää ja fajita – uudistunut Kielitoimiston sanakirja. *Kielikello* 3/2008, s. 14–17.
- Hakulinen, Lauri 1999 [1967]: Suomen sanakirjoista. *Sanat on kaikki sarvipäitä. Lauri Hakulisen kirjoituksia ja puheita vuosilta 1926–1974*, s. 134–150. Toim. Auli Hakulinen – Pentti Leino – Tiina Onikki. SKS, Helsinki.
- Hartmann, Reinhard Rudolf Karl 2001: *Teaching and Researching Lexicography*. Longman, New York.
- Hashi, Awil Ali 2004: *Essential English-Somali Dictionary*. 2. p. Fiqi Publishers.
- Hausmann, Franz Josef 1989: Dictionary Criminality. *Dictionaries: an international encyclopedia of lexicography*. Volume 1, s. 97–101. Toim. F. J. Hausmann – Oskar Reichmann – Herbert E. Wiegand – Ladislav Zgusta. Walter de Gruyter, Berlin.
- Hausmann, Franz Josef – Wiegand, Herbert Ernst 1989: Component Parts and Structures of General Monolingual Dictionaries: A Survey. *Dictionaries: an international encyclopedia of lexicography*. Volume 1, s. 328–360. Toim. F. J. Hausmann – Oskar Reichmann – Herbert E. Wiegand – Ladislav Zgusta. Walter de Gruyter, Berlin.
- Hedman, Henry 1996: *Sar me sikjavaa romanes. Romanikielen kieliooppiopas*. Opetushallitus, Jyväskylä.
- 2004: Suomen romanikieli salakielenä. *Papers from the 30th Finnish Conference of Linguistics, Joensuu, May 15–16, 2003*, s. 42–48. Toim. Marja Nenzen. Joensuun yliopisto, Joensuu.
- 2009: *Suomen romanikieli: sen asema yhteisössään, käyttö ja romanien kieliasenteet*. KKTK:n verkkojulkaisuja 8. KKTK, Helsinki. Saatavissa: <http://scripta.kotus.fi/www/verkkojulkaisut/julk8/>. [Viitattu 30.7.2010.]
- Heikkinen, Vesa 2012: Yleisiä genreongelmia. *Kielikello* 4/2012, s. 11.
- Heininen, Simo 1976: *Nuori Mikael Agricola*. Suomi 120: 3. SKS, Helsinki.
- 1999: *Mikael Agricola raamatunsuomentajana*. SKST 761. SKS, Helsinki.
- Hiidenmaa, Pirjo 2007: Ei vain luonnonlahjoilla: ohjailua, opetusta, apuneuvoja kielenkäytön tueksi. *Kohti tulevaisuuden kielikoulutusta. Kielikoulutuspoliittisen projektin loppuraportti*, s. 431–450. Toim. S. Pöyhönen – M-R. Luukka. Soveltavan kielen tutkimuksen keskus, Jyväskylä. Saatavissa: www.jyu.fi/hum/laitokset/solki/tutkimus/projektit/kielo/projektin_loppuraportti/KIEPOn_loppuraportti.pdf. [Viitattu 8.10.2012.]
- Hirvensalo, Lauri 1980: *Saksalais-suomalainen suursanakirja*. WSOY, Porvoo.

- HLEO 2008 = *Turun yliopiston historian laitoksen eettinen ohjeisto*. Turun yliopisto, Turku 2008. Saatavissa: www.hum.utu.fi/sivustot/historia/ajankohtaista/uutiset_etusivu/eettisetohjeet/Eettisetohjeet.pdf. [Viitattu 2.4.2012.]
- Holma, Juhani 2008: *Sangen ialo Rucous. Schwenckfeldiläisten rukouskirja Mikael Agricolan lähteenä*. Yliopistopaino, Helsinki.
- Holzinger, Daniel 1995: *Romanes (Sinte)*. Languages of the World/Materials 105. Lincom Europa, München.
- Honko, Lauri: Kenen omaisuutta on suullinen perinne? *Hüdenkivi* 2/2002, s. 6–9. SKS, Helsinki.
- Hormia, Osmo 1961: *Gananderin sanakirjan lähteet*. SKST 271. SKS, Helsinki.
- Hoyer, Karin 2012: *Dokumentation och beskrivning som språkplanering – perspektiv från arbete med tre teknade minoritetsspråk*. Nordica Helsingensia 29. Helsingfors universitet, Helsingfors. Saatavissa: <https://helda.helsinki.fi/bitstream/handle/10138/29202/dokument.pdf?sequence=1>.
- HS = *Helsingin Sanomat*.
- Hughes, Geoffrey 2009: *Political Correctness: A History of Semantics and Culture*. Wiley-Blackwell, Malden, MA.
- Hulstijn, Jan – Atkins, B. T. Sue 1998: *Empirical research on dictionary use in foreign language learning: survey and discussion*. Saatavissa: <http://dare.uva.nl/document/176766>. [Viitattu 15.10.2012.]
- Häkkinen, Jaakko 2009: Kantaurlin ajoitus ja paikannus: perustelut puntarissa. *Suomalais-Ugrialaisen Seuran Aikakauskirja* 92, s. 9–56. SUS, Helsinki.
- Häkkinen, Kaisa 1993: Suomen kielen tutkimus kansallisen diktatuurin kourissa. *Vaikea totuus. Vuosi 1918 ja kansallinen tiede*, s. 18–45. Toim. Heikki Ylikangas. SKST 603. SKS, Helsinki.
- 1994: *Agricolasta nykykieleen. Suomen kirjakielen historia*. WSOY, Helsinki.
- 2002: *Suomen kielen historia 1. Suomen kielen äänne- ja muotorakenteen historiallista taustaa*. Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 69. Kolmas, uudistettu painos. Turku.
- 2006: Laajojen aineistojen käyttö ja käyttöongelmat kielentutkimuksessa. *Etiikkaa ihmistieteille*, s. 203–219. Toim. Jaana Hallamaa – Veikko Launis – Salla Lötjönen – Irma Sorvali. SKS, Helsinki.
- 2007: *Agricolan eläintarha*. *SJ* 49, s. 7–21.
- Häkkinen, Kaisa – Lempiäinen, Terttu 2007: *Agricolan yrtit. Mikael Agricolan Rucouskirian terveyttä tuovat kasvit, niiden esiintyminen ja käyttö 1500-luvulla*. Kirja-Aurora, Turku.
- 2011: *Aaloesta öljypuuhun. Suomen kielellä mainittuja kasveja Agricolan aikaan*. Teos, Helsinki.
- Hämäläinen, Simo 1963: *Suomalainen sotilasslangi*. SKST 273. SKS, Helsinki.
- Häyry, Heta – Häyry, Matti 1997: Arvot ja niiden arviointi. *Hyvä, kaunis, tosi – arvojen filosofiaa*, s. 5–24. Yliopistopaino, Helsinki. (Artikkeli ilmestynyt ensimmäisen kerran teoksessa *Rakennamme tulevaisuutta. Utopiasta vai uhkakuvista totta?*, s. 18–36. Toim. A. Kajanto. Vapaan sivistystyön yhteisjärjestö, Kokkola 1988.)
- Ickler, Theodor 2006: Duden – politisch korrekt. Der „angemessene Gebrauch von

- Wörter“. *Schrift und Rede* 01.08.2006. Saatavissa myös:
www.fds-sprachforschung.de/ickler/index.php?show=news&id=577.
- Illustrerad svensk ordbok*. 3. painos. Toim. Bertil Molde. Natur och kultur, Stockholm 1964 [1958].
- Ilmainen sanakirja. Saatavissa: <http://ilmainensanakirja.fi/>.
- Isanbaev 1989 = Исанбаев, Н. И. 1989: *Марийско-тюркские языковые контакты. Часть первая*. Марийское книжное издательство, Йошкар-Ола.
- 1994 = Исанбаев, Н. И. 1994: *Марийско-тюркские языковые контакты. Часть вторая. Словарь татарских и бакирских заимствований*. Научный центр Финно-Угроведения, Йошкар-Ола.
- Iso ruotsalais-suomalainen sanakirja*. Toim. Göran Karlsson – Brita Löflund – Ulla Ekblom – Margareta Kurtén – Lea Westerlund – Marjatta Tammissalo. SKST 358. SKS, Helsinki 1982–1987.
- Ivanov 1981 = Иванов, И. Г. 1981: *Марий диалектологий*. Марий государственный университет, Йошкар-Ола.
- 2001 = Иванов, И. Г. 2001: *Мутвундынам пойдарена. У мут-влак мутер*. Марий книга савыктыш, Йошкар-Ола.
- Jevsevjev 1983 = *Timofej Jevsevjevs Folklore-Sammlungen aus dem Tscheremissischen I. Märchen, Sagen und Volkserzählungen*. Toim. Alho Alhoniemi – Sirkka Saarinen. SUST 184. SUS, Helsinki.
- 1989 = *Timofej Jevsevjevs Folklore-Sammlungen aus dem Tscheremissischen II. Vorzeichen, Traumdeutungen, Sprichwörter, Spottverse und Rätsel*. Toim. Sirkka Saarinen. SUST 199. SUS, Helsinki.
- 1992 = *Timofej Jevsevjevs Folklore-Sammlungen aus dem Tscheremissischen III. Gebete und Zaubersprüche*. Toim. Sirkka Saarinen. SUST 211. SUS, Helsinki.
- 1994 = *Timofej Jevsevjevs Folklore-Sammlungen aus dem Tscheremissischen IV. Lieder*. Toim. Sirkka Saarinen. SUST 219. SUS, Helsinki.
- Johanneksenpoika.net = [14.4.2012 päivätty otsikoimaton kirjoittajan itsensä internetissä julkaisema kirje.] Saatavissa: <http://johanneksenpoika.fi/kirja/Kirjeet%20JT%20RR.pdf>. [Viitattu 20.9.2012.]
- Johnson, Samuel 1747: *The Plan of a Dictionary*. S. l., London.
- Jokinen, Heikki 2010: *Tietokirjailijan eväät. Tekijänoikeus, sopimukset, apurahat*. Toim. uudistettu painos. Suomen tietokirjailijat ry – Edita Prima Oy, Helsinki.
- Jokinen, Ulla 1986: Ranskan kielen huolto ennen ja nyt. *Kielikello* 4, s. 4–13.
- Juslenius, Daniel 1968 [1745]: *Suomalaisen Sana-Lugun Coetus*. Näköispainos. SKS, Helsinki.
- Jussila, Raimo 1997: Den finländska romanin och dess ordböcker. *LexicoNordica* 4, s. 75–81.
- 2009: *Kalevalan sanakirja*. Otava, Helsinki.
- 2012: Riiuuyöaieaatos. *Hiidenkivi* 4/2012, s. 40. SKS, Helsinki.
- Jussila, Raimo – Länsimäki, Maija 1994: *Se siitä. Suomen kielen seksisanakirja*. WSOY, Helsinki.
- Kaartinen, Marjo 2004: *Neekerikammo. Kirjoituksia vieraan pelosta*. K&h, Kulttuurihistoria, Turun yliopisto, Turku.

- Kadima, Gilbert Adi 2001: *Terminologie de l'élimination de la discrimination à l'égard de la femme*. [Ei tietoa julkaisijasta.] Ks. myös *Lexique des termes relatifs à l'élimination de la discrimination à l'égard des femmes*, saatavissa: <http://nsangalubangu.webs.com/droitsdelafemme.htm>. [Viitattu 20.10.2012.]
- Kaino.fi = Vanhan kirjasuomen korpuksen hakulomake < Vanhan kirjasuomen korpus < Aineistot < Aineistopalvelu Kaino < Kotus.fi [tekstikorpus] – KKTK, Helsinki. Saatavissa: <http://kaino.kotus.fi/korpushaku/vks-haku.xml>. [Viitattu 23.10.2012.]
- Kalela, Jorma 2000: *Historiantutkimus ja historia*. Gaudeamus, Helsinki.
- Kalliokoski, Jyrki 1996: Johdanto. *Teksti ja ideologia. Kirjoituksia kielestä ja vallasta julkisessa kielenkäytössä*, s. 8–36. Toim. Jyrki Kalliokoski. Kieli 9. Helsingin yliopiston suomen kielen laitos, Helsinki.
- Karlsson, Fred 1974: Sukupuoliroolien kielellisistä heijastumista. *SJ* 16, s. 24–33.
- 1987: Om ordböcker och lexikografi. *Fennistica festiva in honorem Göran Karlsson septuagenarii*. Toim. Mauno Koski – Eeva Lähdemäki – Kaisa Häkkinen. Åbo Akademis förlag, Åbo.
- Karttunen, Kaarina 1979: *Nykyslangin sanakirja*. WSOY, Helsinki.
- Kauppinen, Perttu 2007: McSanoja jo parikymmentä Wikipediassa. *HS* 20.3.2007. Saatavissa myös: www.hs.fi/talous/artikkeli/McSanoja+jo+parikymment%C3%A4+Wikipediassa/1135225984257.
- Keenadiid, Yaasiin Cismaan 1976: *Qaamuuska Af Soomaaliga*. Le Monnier, Mogadishu–Florence.
- Kielikello = *Kielikello – kielenhuollon tiedotuslehti*. KKK, Helsinki.
- Kielitoimiston kirjeenvaihto 1999 ja 2000. Kielitoimiston arkisto. KKK, Helsinki.
- Kielitoimiston sanakirja*. Päätoim. Eija-Riitta Grönros. KKK, Helsinki 2004– (eri painosten tiedot alla, teosta päivitetty painosten välissä).
- 2004: (sähköinen) *Kielitoimiston sanakirja* 1.0. KKTKJ 132. KKTK – Kielikone Oy, Helsinki 2004.
- 2006: (painettu) *Kielitoimiston sanakirja, osat 1–3*. KKTKJ 140. KKTK, Helsinki 2006.
- 2008: (sähköinen) *Kielitoimiston sanakirja, osat 1–3*. KKTKJ 149. KKTK – Kielikone Oy, Helsinki 2008.
- 2012: (painettu ja sähköinen) *Kielitoimiston sanakirja, osat 1–3*. 3., uudistettu painos. KKKJ 170. KKK – Kielikone Oy, Helsinki 2012.
- Kipfer, Barbara Ann 1987: Dictionaries and the intermediate student: communicative needs and the development of user reference skills. *The dictionary and the Language Learner. Papers from the EURALEX Seminar at the University of Leeds, 1–3 April 1985*, s. 44–54. Toim. Anthony Cowie. Max Niemeyer Verlag, Tübingen.
- Kivekäs, Maija 1972: Leksikografian käsikirja. *Vir.*, 466–469.
- KKK = Kotimaisten kielten keskus.
- KKKJ = Kotimaisten kielten keskuksen julkaisuja.
- KKO = Korkeimman oikeuden ennakkopäätökset. Finlex-tietokanta, ks. Finlex. Saatavissa: www.finlex.fi/fi/oikeus/kko/kko/.

- KKTK = Kotimaisten kielten tutkimuskeskus.
 KKTKJ = Kotimaisten kielten tutkimuskeskuksen julkaisuja.
 Klemetilä, Hannele 2008: *Keskiajan julmuus*. Atena, Jyväskylä.
 Kleparski, Grzegorz 1986: *Semantic change and componential analysis. An inquiry into pejorative developments in English*. Eichstätter Materialien 9, Abteilung Sprache und Literatur 4. F. Pustet, Regensburg.
 Koivisto, Viljo 1994: *Romano-finitiko-angliko laavesko liin. Romani-suomi-englanti sanakirja. Romany-Finnish-English Dictionary*. KKTKJ 74. Painatuskeskus, Helsinki.
 ——— 2001: *Suomi-romani-sanakirja. Finitiko-romano laavesko liin*. SKST 811 – KKTKJ 117. SKS–KKTK, Helsinki.
 ——— 2005: *Romano-finitiko-angliko laavesko liin. Romani-suomi-englanti sanakirja. Romany-Finnish-English Dictionary*, 2. painos. Opetushallitus, Helsinki.
 ——— 2011: *Romano-finitiko laavesko liin. Romani-suomi-sanakirja*. Opetushallitus, Tampere.
 Kolehmainen, Taru 2005: Suomettuminen ennen ja nyt. *HS* 28.6.2005. Saatavissa myös: www.kotus.fi/index.phtml?s=904. [Viitattu 20.10.2012.]
 Korhonen, Johanna 2010: Haluan sanani takaisin. *HS* 29.12.2010. Saatavissa myös: www.hs.fi/kulttuuri/artikkeli/Haluan+sanani+takaisin/1135262687077.
 Korhonen, Mikko 1986: *Finno-Ugrian language studies in Finland 1828–1918*. Societas Scientiarum Fennica, Helsinki.
 Korhonen, Mikko – Suhonen, Seppo – Virtaranta, Pertti 1983: *Sata vuotta suomen sukua tutkimassa. 100-vuotias Suomalais-Ugrilainen Seura*. Weilin+Göös, Helsinki.
 Korkeimman oikeuden päätökset ks. KKO.
 Kotkanvirta, Jussi – Nyyssönen, Seppo 1997: *Ajatus: Etiikka*. WSOY, Porvoo.
 Kotus 2011: *Kotus sai Koneen Säätioiltä apurahan. Vanhan kirjasuomen sanakirja jatkuu*. KKK, Helsinki. Saatavissa: www.kotus.fi/files/2065/Tiedote_vks_jatkuu.pdf. [Viitattu 19.2.2013.]
 Kotus.fi: Suomen ruotsalaismurteiden sanakirja = *Suomen ruotsalaismurteiden sanakirja* < Suomenruotsin murteet < Murteet < Kotus.fi – KKK, Helsinki. [Verkkosanakirja.] Saatavissa: www.kotus.fi/index.phtml?s=163. [Päivitetty 6.10.2009/viitattu 16.10.2012.]
 Koukkunen, Kalevi 2012: Kaikilla pitää piletit oleman walmisna – All must have their tickets ready. Suomalais-englantilaisten sanakirjojen historiaa. *SJ* 54, s. 76–106.
 Kram, Kristina 1998: ”..blott rent undantagsvis om kvinnor”. Om representation av kvinnan i moderna ordböcker. *Språk & stil NF* 8, s. 159–171. Tidskrift för svensk språkforskning. Uppsala Universitet, Uppsala.
 Kromann, Hans-Peder – Riiber, Theis – Rosbach, Poul 1991: Principles of Bilingual Lexicography. *Dictionaries: an international encyclopedia of lexicography*. Volume 3, s. 2711–2728. Toim. F. J. Hausmann – Oskar Reichmann – Herbert E. Wiegand – Ladislav Zgusta. Walter de Gruyter, Berlin.
 KS ks. *Kielitoimiston sanakirja*.
 Kudashev, Igor 2000: *Suomalais-venäläinen sosiaalialan sanasto*. Venäjän ja Itä-Eu-

- roopan instituutti, Helsinki.
- Kudashev 2007 = Кудашев, И. 2007: *Проектирование переводческих словарей специальной лексики*. Helsinki University Translation Studies Monographs 3. Helsinki University Print, Helsinki.
- Kudashev, Igor – Carlson, Lauri – Kudasheva, Irina 2010: TermFactory: Collaborative Editing of Term Ontologies. *Terminology and Knowledge Engineering Conference 2010*, s. 479–500. Toim. Úna Bhreathnach – Fionnuala Barra-Cusack. Fiontar, Dublin City University, Dublin.
- Kudasheva, Irina – Kudashev, Igor 2008: *EU-Russia Project Co-operation Glossary (English-Finnish-Russian). About 1 300 terms*. University of Helsinki, Palmenia Centre for Continuing Education, Kouvola.
- Kulonen, Ulla-Maija 1996: *Sanojen alkuperä ja sen selittäminen. Etymologista leksikografiaa*. Suomi 181. SKS, Helsinki.
- Kuusi, Matti 1985: Kuka tutkisi sanakirjoja? *Vir.*, s. 485–490.
- Kuusinen ym. 1997 = *Venäjä-suomi-suursanakirja*. Laatineet Kuusinen, Martti – Ollikainen, Veera – Syrjäläinen, Julia. Toim. M. Kuusinen. WSOY, Helsinki 1997.
- Kuusinen ym. 1999 = *Новый большой русско-финский словарь* в 2-х томах. М. Э. Куусинен, В. М. Олликайнен, Ю. Э. Сюръялайнен. 1-2. Русский язык, Москва 1999.
- Kuutti, Pirkko 2010: Varhaiset lainsuomennokset. *Oikeuskieli ja säädöstieto. Suomenkielinen lakikirja 250 vuotta. Rättsspråk och författningsinformation. Den finskspråkiga lagboken 250 år*, s. 107–129. Toim. Heikki E. S. Mattila – Aino Piehl – Sari Pajula. Suomalainen lakimiesyhdistys, Helsinki.
- Laajakaistanasto*. TSK 43. Sanastokeskus TSK, Helsinki 2012. Saatavissa: www.tsk.fi/tiedostot/pdf/Laajakaistanasto.pdf.
- Laine, Esko M. – Laine, Tuija 2010: Kirkollinen kansanopetus. *Huoneentaulun maailma. Kasvatus ja koulutus Suomessa keskiajalta 1860-luvulle*, s. 258–306. Toim. Jussi Hanska – Kirsi Vainio-Korhonen. SKS, Helsinki.
- Laki kotoutumisen edistämiseksi. 1386/2010, ks. Finlex.
- Laki maahanmuuttajien kotouttamisesta ja turvapaikanhakijoiden vastaanotosta. 493/1999, ks. Finlex.
- Laki turvapaikan hakijoiden vastaanottoasemista ja keskuksista. 1465/1991, ks. Finlex.
- Landau, Sidney I. 2001: *Dictionaries. The Art and Craft of Lexicography*. 2nd edition. Cambridge University Press, Cambridge.
- Larsson, Lennart 2009: Hur märker användaren av SAOB de 110 åren mellan A och TYNA? *Språk & stil NF 19*, s. 35–52. Tidskrift för svensk språkforskning. Uppsala Universitet, Uppsala. Saatavissa myös: <http://uu.diva-portal.org/smash/get/diva2:282845/FULLTEXT01>. [Viitattu 10.6.2013.]
- 2012: Varför inte dub(b)lettformer? Om SAOB som normativ ordbok. *Nordiska studier i lexicografi 11. Rapport från Konferensen om lexicografi i Norden Lund 24–27 maj 2011*, s. 385–395. Toim. Birgit Eaker – Lennart Larsson – Anki Mattisson. Skrifter utgivna av Nordiska föreningen för lexicografi. Skrift nr 12. I samarbete med Språkrådet i Norge och Svenska Akademiens ordboksredaktion.

- Lauerma, Petri 2012: Kieli. *Genreanalyysi – tekstilajitutumuksen käsikirja*, s. 51–54. Toim. Vesa Heikkinen – Ero Voutilainen – Petri Lauerma – Ulla Tiililä – Mikko Lounela. KKKJ 169. Gaudeamus, Helsinki.
- Le Nouveau Petit Robert 2007 = *Dictionnaire alphabétique et analogique de la langue française*. Le Robert, Paris 2007.
- Lehtinen, Ildikó 1985: Timofej Jevsevjev – der grosse tscheremissische Sammler. *Timofej Jevsevjevs ethnographische Sammlungen über die Tscheremissen*, s. 8–45. Toim. Ildikó Lehtinen. SUS:n Kansatieteellisiä julkaisuja XII:1. SUS, Helsinki.
- Lehtinen, Marja 1994: Sanoja, merkityksiä, ohjeita. *Kielikello* 3/1994, s. 5–10.
- Lehtonen, Jukka 2003: *Seksuaalisuus ja sukupuoli koulussa. Näkökulmana heteronormatiivisuus ja ei-heteroseksuaalisten nuorten kertomukset*. Nuorisotutkimusverkoston ja Nuorisotutkimusseuran julkaisuja 31. Yliopistopaino, Helsinki. Saatavissa myös: <http://ethesis.helsinki.fi/julkaisut/val/sosio/vk/lehtonen/seksuaal.pdf>.
- Lehtonen, Tuomas M. S. 1999: Keskiajan kirjallinen kulttuuri. *Suomen kirjallisuushistoria* 1. *Hurskaista lauluista ilostelevaan romaaniin*, s. 12–35. Toim. Yrjö Varpio – Liisi Huhtala. SKST 724. SKS, Helsinki.
- Lehtosalo, Kaisa – Tyysteri, Laura 2011: Eettiset kysymykset sanakirjatyössä. *Kieli ja etiikka. VAKKI Symposium XXXI. Vaasan yliopisto 11.–12.2.2011*, s. 212–223. Toim. Niina Nissilä – Nestori Siponkoski. Vaasan yliopiston käännteorian, ammattikielten ja monikielisyyden tutkijaryhmän julkaisut N:o 38. Vaasan yliopisto, Vaasa. Saatavissa myös: www.vakki.net/publications/2011/VAKKI2011_Lehtosalo&Tyysteri.pdf.
- Lehtosalo, Kaisa – Vaittinen, Tanja 2009: Sai sampo valmihiksi, kirjokansi kaunihiksi. *SJ* 51, s. 205–211.
- Leino, Pentti 2002 [1996]: Miten pitkälle kieltä voidaan normittaa? *Mittoja, muotoja, merkityksiä*, s. 332–343. Toim. Liisa Tainio – Aki Ontermä – Tapani Kelomäki – Minna Jaakola. SKS, Helsinki.
- LexicoNordica*. Nordisk forening for leksikografi, Oslo.
- LEXIN romskt = LEXIN. *Svensk-romskt/arli lexikon. Švedikano-Romano/Arlikano leksikoni*. Käänt. Baki Hasan. Institutet för språk och folkminnen, Vällingby 2007.
- Lindholm, Arto 2006: Paskaduuni. *Prekaariruoska? Portfoliosukupolvi, perustulojakaussalaistoiminta*, s. 16–20. Toim. Tommi Hoikkala – Mikko Salasuo. Nuorisotutkimusverkoston verkkojulkaisusarja. Nuorisotutkimusseura/Nuorisotutkimusverkosto. Saatavissa: www.nuorisotutkimusseura.fi/julkaisuja/prekaariruoska.pdf. [Viitattu 20.9.2012.]
- Lipsonen, Leo 1990: *Vankilanslangin sanakirja*. Vankeinhoidon koulutuskeskuksen julkaisuja 2/1990. Valtion painatuskeskus, Helsinki.
- Ljungo Tuomaanpojan lainsuomennoksen sanasto. Matti Ulkuniemen postuumin käsikirjoituksen mukaan*. Toim. Antero Niemikorpi – Raimo Jussila – Anne Kattilakoski – Esko Koivusalo – Heikki Mattila – Erkki Rintala. SKST 1126. SKS, Helsinki 2007.
- LSA = Ethics Statement. Linguistic Society of America, Washington 2009. Saatavissa:

- www.lsadc.org/info/pdf_files/Ethics_Statement.pdf. [Viitattu 30.3.2012.]
- Länsimäki, Maija 2008: Miksi suuret sanakirjat pitää tehdä? *Vir.*, s. 565–570.
- Lönnrot, Elias 1958 [1866–1880]: *Suomalais-Ruotsalainen Sanakirja I–II*. Kolmas, manulmenetelmällä jäljennetty painos. WSOY, Porvoo–Helsinki.
- Maamies, Sari 1995: Neekeri ja musta. *Sana kiertää. Kirjoituksia suomen kielestä*. Toim. Maija Länsimäki. KKTKJ 80. KKTK, Helsinki.
- Mantila, Harri 1998: Jätkät ja jässikät, don juanit ja gentlemannit. *Tuppisuinen mies. Kirjoitelmia sukupuolesta, kielestä ja kulttuurista*, s. 15–32. Toim. Vesa Heikkinen – Harri Mantila – Markku Varis. Tietolipas 154. SKS, Helsinki.
- 2010: Suomalaisen kielenhuollon periaatekeskustelu 1990- ja 2000-luvulla. *Kielellä on merkitystä. Näkökulmia kielipolitiikkaan*, s. 179–205. Toim. Hanna Lappalainen – Marja-Leena Sorjonen – Maria Vilkkuna. SKS, Helsinki.
- Marčuk 1992 = Марчук, Ю. Н. 1992: *Основы терминографии: Методическое пособие*. Изд-во МГУ, Москва.
- Martyna, Wendy 2006: What Does ‘He’ Mean? Use of the Generic Masculine. *Journal of Communication*, Volume 28, Issue 1, s. 131–138.
- Matras, Yaron 2002: *Romani. A Linguistic Introduction*. Cambridge University Press, Cambridge.
- Mattisson, Anki 2006: Genusperspektiv på SAOB:s källor. *LexicoNordica* 13, s. 55–68. Saatavissa myös: www.nordisk-sprakrad.no/dokumenter/Lex-pdf/LexicoNordica%202006.pdf.
- Mattisson, Ann-Christin 1999: Verbet dänka och SAOB:s korpus och språkprovssamling. *LexicoNordica* 3, s. 91–103.
- Merimaa, Heidi 2007: Mikael Agricolan kielen muoto- ja lauserakenteen erikoispiirteitä. *Agricolan aika*, s. 103–117. Toim. Kaisa Häkkinen – Tanja Vaittinen. BTJ Kustannus, Helsinki.
- MES = Museotyön eettiset säännöt. ICOM Suomen komitea ry, Espoo 2005. Saatavissa: <http://finland.icom.museum/etiikka.html>. [Viitattu 2.4.2012.]
- Mikael Agricolan teokset* 1–4. Uudistettu näköispainos. WSOY, Porvoo 1987.
- Milles, Karin 2008: *Jämställt språk. En handbok i att skriva och tala jämställt. Skrifter utarbetade av Språkrådet*. Norstedts Akademiska förlag, Falun.
- Mire, Ahmed Hussein 2008: *Advanced English-Somali Dictionary*. Color Printing Co.
- MNS = Mustalaiskielen normatiivi sanasto. Eripainos Mustalaiskielen ortografiakomitean mietinnöstä 1971: A 27. Moniste. Helsinki.
- MOT Kielitoimiston sanakirja ks. *Kielitoimiston sanakirja*.
- MOT KS ks. *Kielitoimiston sanakirja*.
- MOT Norstedts = *MOT Norstedts Svensk ordbok* 1.0. [Verkkosanakirja.] Språkdata och Norstedts Akademiska Förlag.
- MOT Ruotsi = *MOT Suomi-ruotsi-suomi* 3.1. [Verkkosanakirja.] Kielikone Oy, Helsinki.
- Müller, Sigrid – Fuchs, Claudia 1993: *Handbuch zur nichtsexistischen Sprachverwendung in öffentlichen Texten. Die Frau in der Gesellschaft*. Fischer, Frankfurt am Main.
- Mäkinen, Kirsti 1989: Sanojen tyyliä. *Nykysuomen sanavarat*, s. 200–212. Toim.

- Jouko Vesikansa. WSOY, Porvoo–Helsinki.
- Nieminen, Mauri 2004: *Siirtolaisuus V*. Tilastokeskus. Saatavissa: www.stat.fi/tup/tietoaika/tilaajat/ta_08_04_nieminen.html. [Viitattu 24.9.2012.]
- Nikkanen, Hanna 2007: Ulkomaalaisvirasto suosittaa: Kuole Kongossa. *Voima* 4/2007. Saatavissa myös: <http://fifi.voima.fi/voima-artikkeli/2007/numero-4/ulkomaalaisvirasto-suositaa-kuole-kongossa>.
- 2009: Kerrosten välissä: Kuinka paljon rahaa miehen pitää ansaita ollakseen hyvä isä? *Voima* 2/2009. Saatavissa myös: <http://fifi.voima.fi/voima-artikkeli/2009/numero-2/kerrosten-valissa>.
- Nikula, Kristina 1992: Deskriptiva ordböcker – finns dom? *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden* 28.–31. mai 1991, s. 43–52. Toim. Ruth Vatvedt Fjeld. Nordisk forening for leksikografi, Oslo.
- 1997: Språkprov i könspektiv. *Fackspråk och översättningsteori, Vakksymposium XVI*, s. 194–207. Toim. D. Wilske. Publikationer av Forskargruppen för översättningsteori och fackspråk vid Vasa universitet, Nr 22. Vaasan yliopisto, Vaasa.
- 2002: ”Användaren, normen och ordböckerna eller det är lätt att säga tulipanas...” *LexicoNordica* 9, s. 31–49.
- Nilsson, Pär 2012: Språkteoriens avtryck i SAOB. *Nordiska studier i lexicografi* 11. *Rapport från Konferensen om lexicografi i Norden Lund 24–27 maj 2011*, s. 479–487. Toim. Birgit Eaker – Lennart Larsson – Anki Mattisson. Skrifter utgivna av Nordiska föreningen för lexicografi. Skrift nr 12. I samarbete med Språkrådet i Norge och Svenska Akademiens ordboksredaktion.
- Nimiläki. 9.8.1985/694, ks. Finlex.
- Nivanka, E. – Airila, M. – Soini, Ester – Penttilä, Aarni 1930: Nykysuomen sanakirjan artikkelien luonnoksia. *Vir.*, s. 244–254.
- NODE = *The New Oxford Dictionary of English*. Oxford University Press, Oxford 2001.
- Nomenclator Latino Finnicus*. Serenissimo Etrurie Principi. Dicus & Consecratus a Martino Fogelio Hamburgensi, 1669. Käsikirjoitus. Kopio saatavissa: SKS–KKK, Helsinki.
- Nordisk leksikografisk ordbok*. Toim. Henning Bergenholtz – Ilse Cantell – Ruth Vatvedt Fjeld – Dag Gundersen – Jón Hilmar Jónson – Bo Svensén. Skrifter utgitt av Nordisk forening for lexicografi, skrift nr. 4. Universitetsforlaget, Oslo 1997.
- Nordman, Marianne 2010: Språkbruk och språkvård. *Oikeuskieli ja säädöstieto. Suomenkielinen lakikirja 250 vuotta. Rättsspråk och författningsinformation. Den finskspråkiga lagboken 250 år*, s. 73–93. Toim. Heikki E. S. Mattila – Aino Piehl – Sari Pajula. Suomalainen lakimiesyhdistys, Helsinki.
- Norjan LEXIN = Lexin – Hovedside. Utdanningsdirektoratet – Uni Computing, Oslo. Saatavissa: decentius.hit.uib.no/lexin.html?ui-lang=eng. [Viitattu 24.9.2012.]
- Norstedts Svenskt språkbruk = *Svenskt språkbruk: ordbok över konstruktioner och fraser*. Norstedts ordbok. Norstedts Akademiska Förlag, Stockholm 2003.
- Novoe v russkoj leksike = *Новое в русской лексике: словарные материалы -77*. Русский язык, Москва 1980.

NS ks. *Nyky-suomen sanakirja*.

Nummilla, Kirsi-Maria 2010: 1600-luvun sanakirjojemme suomalaissanasto ja sen alkuperä. *Sanoista kirjakieliin. Juhlakirja Kaisa Häkkiselle 17. marraskuuta 2010*, s. 189–202. Toim. Sirkka Saarinen – Kirsti Siitonen – Tanja Vaittinen. SUST 259. SUS, Helsinki.

— 2012: Martin Fogelin sanakirjakäsikirjoitus *Nomenclator Latino Finnicus 1669*. SJ 54, s. 163–169.

Nyky-suomen sanakirja I–VI. Päätoim. Matti Sadeniemi. Valtion toimeksiannosta teettänyt SKS. WSOY, Porvoo–Helsinki 1951–1961.

Obratnyj 1992 = *Обратный словарь удмуртского языка*. Издательство «Удмуртский университет», Ижевск.

OED = *Oxford English Dictionary* (1884–). Vuoteen 1933 asti nimellä *New English Dictionary*. Oxford University Press, Oxford. Saatavissa: www.oed.com, vaatii käyttäjätunnuksen. [Viitattu 20.9.2012.]

OED: Appeals = Oxford English Dictionary Online. OED Appeals: Help record the history of our language. Saatavissa: <http://public.oed.com/appeals/>. [Viitattu 10.10.2012.]

OED: FAQ = Oxford English Dictionary Online. Frequently asked questions: I've made up a word. Please add it to the OED. Saatavissa: <http://public.oed.com/about/frequently-asked-questions/#madeup>. [Viitattu 10.10.2012.]

OKM 2008 = Opetus- ja kulttuuriministeriön ja Kotimaisten kielten tutkimuskeskuksen tulossopimus vuosille 2011–2012 ja voimavarat vuodelle 2012. Saatavissa: www.minedu.fi/export/sites/default/OPM/Tiede/ohjaus_ja_rahoitus/tulosohjaus/tulossopimukset/2011/Kotus_10012011.pdf.

OKM 2010 = Kotimaisten kielten tutkimuskeskuksen tehtävät. Opetusministeriön työryhmämuistioita ja selvityksiä 2010:1. Yliopistopaino, Helsinki. Saatavissa: www.minedu.fi/export/sites/default/OPM/Julkaisut/2010/liitteet/tr01.pdf?lang=fi.

OKM 2012 = Opetus- ja kulttuuriministeriön ja Kotimaisten kielten keskuksen välinen tulossopimus vuosille 2013–2016 ja voimavarat vuodelle 2013. Saatavissa: http://www.minedu.fi/export/sites/default/OPM/Tiede/ohjaus_ja_rahoitus/tulosohjaus/tulossopimukset/2013/KOTUS_2013.pdf.

Ollila, Maija-Riitta 1997: *Moraalin tuolla puolen*. WSOY, Helsinki.

Opetusministeriön päätös ylimääräisen valtionavustuksen myöntämisestä maahanmuuttajien tukiovetukseen peruskoulussa ja lukiossa. 1609/1992, ks. Finlex.

Paasonen 1948 = *H. Paasonens Ost-tscheremissisches Wörterbuch bearbeitet und herausgegeben von Paavo Siro*. Lexica Societatis Fenno-Ugricae XI. SUS, Helsinki.

— 1990–1999 = *H. Paasonens Mordwinisches Wörterbuch*. Toim. Martti Kahla. Lexica Societatis Fenno-Ugricae XXIII. Helsinki.

Pakkanen, Lea 2011: Kaksi kulmaa maahanmuuttoon. [Saido Mohamedin haastettu.] *Vihreä lanka*, nro 29: 21.10.–3.11.2011, s. 11–15.

Pasanen, Annika 2008: Uralilaiset vähemmistökielet assimilaation ja revitalisaation ristipaineissa. *Murros. Suomalais-ugrilaiset kielet ja kulttuurit globalisaation paineissa*, s. 47–70. Toim. Sirkka Saarinen – Eeva Herrala. Uralica Helsingensia 3. Helsingin yliopiston suomalais-ugrilainen laitos – Suomalainen Tiede-

- akatemia – SUS, Helsinki.
- Paunonen, Heikki 1996: Suomen kielen ohjailun myytit ja stereotypiat. *Vir.*, s. 544–555.
- Pilke, Nina 2008: Ordboksanvändning hos blivande språkexperter och hos professionella översättare. *LexicoNordica* 15, s. 135–153.
- 2009: Hon arbetar som extrabiträde – han blev chef för avdelningen. Kvinnlig och manlig yrkesbild i ljuset av exempelmeningar i en tvåspråkig cd-romordbok. *Vakki-symposium XXIX, Språk och makt*, s. 332–342. Toim. Mona Enell-Nilsson – Niina Nissilä. Publikationer av Forskargruppen för översättnings-teori, fackspråk och flerspråkighet vid Vasa universitet, Nr 36. Vakki, Vaasa.
- 2010: Morrande män och skrikande kvinnor – om könsbundna beteckningar i en tvåspråkig cd-romordbok. *Svenskans beskrivning* 30, s. 291–302. Toim. Cecilia Falk – Andreas Nord – Rune Palm. Institutionen för nordiska språk, Stockholms universitet, Stockholm.
- Pilke, Nina – Puskala, Jaana 2010: Vad heter förvärv på finska? Användningen av hjälpmedel när blivande språkexperter översätter. *Nordiska studier i lexikografi* 10, s. 342–445. Toim. H. Lönnroth – K. Nikula. Skrifter utgivna av Nordiska föreningen för lexikografi Skrift nr 11. Tammerfors universitet – Språkrådet i Norge, Tammerfors.
- 2012: Vem förvärvar, vem forskar och vem liknar Greta Garbo? Om fackspråkighet och könsbundna beteckningar i betydelsebeskrivningar. *LexicoNordica* 19. *Betydelsebeskrivningar i nordiska ordböcker*, s. 141–161.
- Pirtisaari, Helena 2002: *Suomen romanin partisippien morfologiaa*. Pro gradu. Helsingin yliopisto, Helsinki.
- Pitkänen, Kaarina 2008: *Suomi kasvitieteen kieleksi. Elias Lönnrot termistön kehittäjänä*. SKST 1164. SKS, Helsinki.
- PS ks. *Suomen kielen perussanakirja*.
- Puglielli, Annarita – Mansuur, Cabdalla Cumar 2012: *Qaamuuska af-Soomaaliga*. Roma TrE-Press, Roma.
- Puskala, Jaana 2011: Att översätta ett pressmeddelande till finska. *Svenskan i Finland* 12, s. 215–223. Toim. S. Niemi – P. Söderholm. Itä-Suomen yliopisto, Joensuu.
- Puuronen, Nina – Koskela, Merja – Laurén, Christer – Nordman, Marianne 1993: Forskningsinformation – ett mångdimensionellt spektrum. *Fackspråk och översättningsteori, Vakki-symposium XIII*, s. 228–240. Toim. D. Wilske. Publikationer av Forskargruppen för översättningsteori och fackspråk vid Vasa universitet, Nr 18. Vakki, Vaasan yliopisto, Vaasa.
- Puusta katsoen. Metsätermit ja metsäsanakirjan laadinta*. Toim. Inkeri Vehmas-Lehto. Käännöstieteen laitoksen julkaisuja VII. Helsingin yliopisto, Käännöstieteen laitos, Helsinki 2009.
- Raatikainen, Panu 2006: Voiko ihmistiede olla arvovapaata? *Etiikkaa ihmistieteille*, s. 93–107. Toim. Jaana Hallamaa – Veikko Launis – Salla Lötjönen – Irma Sorvali. SKS, Helsinki.
- Ralph, Bo 2012: Biskopar, ordböcker och verkligheten. Vad är lexikografihistoria bra för? *Nordiska studier i lexikografi* 11. *Rapport från Konferensen om lexikografi*

- i Norden Lund 24–27 maj 2011, s. 30–54. Toim. Birgit Eaker – Lennart Larsson – Anki Mattisson. Skrifter utgivna av Nordiska föreningen för lexikografi. Skrift nr 12. I samarbete med Språkrådet i Norge och Svenska Akademiens ordboksredaktion.
- RDM = Luutonen, Jorma – Mosin, Mikhail – Shchankina, Valentina 2004: *Reverse Dictionary of Mordvin – Обратный словарь мордовских языков*. Lexica Societatis Fenno-Ugricae XXIX. Société Finno-Ougrienne, Helsinki.
- Rej – Delesal’ 1983 = Рей А. – Делесаль С. 1983: Проблемы и антиномии лексикографии. *Новое в зарубежной лингвистике*, s. 261–300. Вып. 14: Проблемы и методы лексикографии. Прогресс, Москва.
- Reverse Dictionary of Mari (Cheremis) – Обратный словарь марийского языка*. Compiled by Jorma Luutonen – Sirkka Saarinen – Arto Moisio – Oleg Sergeevyev – Lidiya Matrosova. Société Finno-Ougrienne, Helsinki 2002.
- Rintala, Päivi 1992: Suomen kielen normeista. *SJ* 34, s. 47–66.
- 2010: Näköalaton tiedepolitiikka. *Vir.*, s. 254–256.
- Romppanen, Birgitta 2001: *Från målspråk till källspråk. Utvecklingen av den tvåspråkiga finsk-svenska och den enspråkiga finska ordboken*. Acta Wasaensia nr 91. Vasa universitet, Vasa.
- Rothstein, Robert A. 1985: While Reading Soviet Dictionaries of Neologisms. *The Slavic and East European Journal*, vol. 29, No. 4, Winter, s. 461–470.
- Ruotsin LEXIN = Lexin. Institutet för språk och folkminnen – KTH, Stockholm. Saatavissa: lexin.nada.kth.se/lexin/. [Viitattu 24.9.2012.]
- Ruotsin LEXIN: Bakgrund = Lexin – bakgrund. Institutet för språk och folkminnen – KTH, Stockholm. Saatavissa: lexin.nada.kth.se/lexin/#about=2;main=3. [Viitattu 24.9.2012.]
- Russell, Lindsay Rose 2012: This is What a Dictionary Looks Like: The Lexicographical Contributions of Feminist Dictionaries. *International Journal of Lexicography* 25(1), s. 1–29.
- Ru-Su = *Iso ruotsalais-suomalainen sanakirja*. Toim. Göran Karlsson – Ulla Ekblom – Margareta Kurtén – Lea Westerlund. SKS, Helsinki 1982–87.
- Räsänen, Martti 1920: *Die tschuwassischen Lehnwörter im Tscheremissischen*. SUST XLVIII. SUS, Helsinki.
- 1923: *Die tatarischen Lehnwörter im Tscheremissischen*. SUST L. SUS, Helsinki.
- Räsänen, Saija 2010: Suomeen syntymässä taas uusi mamukriittinen puolue. *Vihreä lanka* 13.9.2010. Saatavissa: www.vihrealanka.fi/node/7803. [Viitattu 20.9.2012.]
- Saarela, Lauri 2012: Lääketieteen sanastolautakunta. Umpikalevalaista pakkosuomea vai lääkärilatinasta kaikille ymmärrettävää? *Kielikello* 3/2012, s. 7–10.
- Saarinen, Sirkka 1997: Language Contacts in the Volga Region: Loan Suffixes and Calques in Mari and Udmurt. *Language in Time and Space. Studies in Honour of Wolfgang Viereck on the Occasion of his 60th Birthday*, s. 388–396. Toim. Heinrich Ramisch – Kenneth Wynne. ZDL Beihefte 97. Franz Steiner Verlag, Stuttgart.
- 2005: Kielen uudistus sanakirjoissa – sanakirjat kielen uudistuksessa. *Suku-*

- kansaohjelman arki. Suomalais-ugrilainen perintö ja arkipäivä*, s. 82–96. Toim. Paula Kokkonen. Castrenianumin toimitteita 64. M. A. Castrénin seura – SUS – Helsingin yliopiston suomalais-ugrilainen laitos, Helsinki.
- 2010a: Marin sanaston alkuperästä. *Sanoista kirjakieliin. Juhlakirja Kaisa Häkikiselle 17. marraskuuta 2010*, s. 335–341. Toim. Sirkka Saarinen – Kirsti Siitonen – Tanja Vaittinen. SUST 259. SUS, Helsinki.
- 2010b: Timofej Jevsevjev – marilaisen perinteen tallentaja. *Kenttäretkistä tutkimustiedoksi*, s. 107–122. Toim. Paula Kokkonen – Anna Kurvinen. Uralica Helsingiensia 4. SUS, Helsinki.
- Sadeniemi, Matti 1952: Nykysuomen sanakirjan ilmestymisen johdosta. *Vir.*, s. 1–8.
- Saksa, Silja 2004: *Babelin perilliset. Kääntäjien ja kääntämisen historiaa*. Atena, Jyväskylä.
- Salmi, Eeva – Laakso, Mikko 2005: *Maahan lämpimään. Suomen viittomakielisten historia*. Kuurojen liitto ry, Helsinki.
- Salminen, Ari – Ikola-Norrbacka, Rinna 2009: *Kuullaanko meitä? Eettinen hallinto ja kansalaiset*. Vaasan yliopiston julkaisuja. Tutkimuksia 288. Hallintotiede 37. Vaasan yliopisto, Vaasa.
- Salminen, Timo 2008: *Aatteen tiede. Suomalais-Ugrilainen Seura 1883–2008*. SUST 1172. SUS, Helsinki.
- Salomäki, Hanna 2004: *Tosi rakkaus odottaa. Seksuaalieettinen nuortenliike Suomessa*. Nuorisotutkimusverkoston ja Nuorisotutkimusseuran julkaisuja 39. Yliopistopaino – Nuorisotutkimusverkosto – Nuorisotutkimusseura, Helsinki.
- Sanakirja.org*. Jonne Jyrylä. Saatavissa: www.sanakirja.org/.
- SAOB = *Svenska Akademiens ordbok*. Svenska Akademien, Lund 1893–. Saatavissa: <http://g3.spraakdata.gu.se/saob/> [Viitattu 25.2.2013.]
- Sarfati, Georges-Elia 2002: Cultural Study, Doxa, Dictionaries: The Case of Jewish Identity. *Poetics Today*, Vol. 23, Nr. 3, Fall 2002, s. 489–512. Duke University Press, Durham, US.
- Saussure, Ferdinand de 1959 [1916]: *Course in General Linguistics*. Käänt. Wade Baskin, toim. Charles Bally – Albert Sechehaye – Albert Reidlinger. Philosophical Library, New York.
- Savatkova 1969 = Саваткова, А. А. 1969: *Русские заимствования в марийском языке*. Марийское книжное издательство, Йошкар-Ола.
- Schroderus, Ericus 1941 [1637]: *Lexicon Latino-Scondicum*. Näköispainos. Toim. Bengt Hesselman. Långmanska kulturfonden, Uppsala.
- Schryver, Gilles-Maurice de 2003: Lexicographers' Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography* 16(2), s. 143–199.
- Šerba 1958 = Щерба Л. В. 1958: *Избранные работы по языкознанию и фонетике*. Том 1. Изд-во ЛГУ, Ленинград.
- Setälä, E. N. 1892: Lisiä Suomalais-ugrilaisen kielentutkimuksen historiaan. *Suomi* 3:5. SKS, Helsinki.
- 1912: Suomalais-Ugrilaisen Seuran tehtävät ja päämäärä. Alkajaispuhe seuran vuosikokouksessa 2. p. joulukuuta 1912. *Suomalais-Ugrilaisen Seuran Aikauskirja XXVIII*: 5, s. 9–16. Helsinki.

- Shaposhnikova, Ekaterina A. 2007: *New Intelligent Electronic Reference. Essays on Lexicon, Lexicography, Terminography in Russian, American and Other Cultures*, s. 197–205. Toim. Olga Karpova – Faina Kartashkova. Cambridge Scholars Publishing, Newcastle, UK.
- Siironen, Mari 1988: Aito naisellinen tunteenpurkaus eli Nykysuomen sanakirjan naiskuva. *Isosuinen nainen. Tutkielmia naisesta ja kielestä*, s. 41–46. Toim. Lea Laitinen. Yliopistopaino, Helsinki.
- Simpson, John 2011: What has the OED become? *Pleenaariesitelmä Lundin yliopistossa 26.5.2011*. 11. Konferenssen om lexicografi i Norden 24.–27.5.2011.
- SJ = *Sananjalka*. Suomen Kielen Seuran vuosikirja. Suomen Kielen Seura, Turku.
- SKES = *Suomen kielen etymologinen sanakirja 1–6*. Toim. Y. H. Toivonen – Erkki Itkonen – Aulis J. Joki – Reino Peltola. SUS, Helsinki 1955–1978.
- Skogsordlista TNC*. 96. Terminologisentrum TNC, Solna 1994.
- SKS = Suomalaisen Kirjallisuuden Seura.
- SKST = Suomalaisen Kirjallisuuden Seuran Toimituksia.
- SKSU = Nurmi, Timo 1999: *Suomen kielen sanakirja ulkomaalaisille*. Gummerus, Jyväskylä.
- Slovar’ = *Словарь марийского языка – Марий мутер I–X*. Марийское книжное издательство, Йошкар-Ола 1990–2005.
- Smith, Philip M. 1985: *Language, The Sexes and Society*. *Language in Society* 8. Basil Blackwell, Oxford, England – New York, USA.
- SMS ks. *Suomen murteiden sanakirja*.
- Snell-Hornby, Mary 1987: Towards a Learner’s bilingual dictionary. *The dictionary and the Language Learner. Papers from the EURALEX Seminar at the University of Leeds, 1–3 April 1985*, s. 159–170. Toim. Anthony Cowie. Max Niemeyer Verlag, Tübingen.
- SO = *Svensk ordbok utgiven av Svenska Akademien*. Osat A–L ja M–Ö. Svenska Akademien, Stockholm 2009.
- Sobkowiak, Włodzimierz 2002: What can be but is not (and why) in Learners’ MRDs. *Teaching English with Technology* 2(3). Saatavissa: http://iatefl.org.pl/call/j_article9.htm.
- Soukka, Anu 2010: *Pronomina han och hon i exempelmeningar i två ordböcker*. Opublicerad avhandling pro gradu i svenska språket. Pro gradu. Tiedekirjasto Tritonia, Vaasa.
- Språkbanken = Sanomalehtikorpus 1997–2009. Saatavissa: <http://spraakbanken.gu.se/>.
- Språkriktighetsboken 2005. Norstedts Akademiska Förlag, Stockholm.
- SSA = *Suomen sanojen alkuperä 1–3*. Toim. Erkki Itkonen – Ulla-Maija Kulonen ym. SKS, Helsinki 1992–2000.
- SSM = Saarikalle, Anne – Vilkuna, Johanna 2010: *Suomen kielen sanakirja maahanmuuttajille*. Gummerus, Helsinki.
- SSOU = Nurmi, Timo 2009: *Suomen sanakirja opiskelijoille ja ulkomaalaisille. Finnish Dictionary for Students and Foreign Learners*. 3., tarkistettu painos. Gummerus, Helsinki.

- SSS = *Taskusanasto suomi–somalı–suomi*. Toim. Mohamed Warsame – Maija Kajava. 2. tarkistettu painos. Suomi-Somalia seura, Helsinki 2010.
- ST = *Eettiset ohjeet*. Suomen tietokirjailijat ry, Helsinki. Saatavissa: www.suomontietokirjailijat.fi/jasenyyys/eettiset-ohjeet/. [Päivitetty 9.2.2012/ viitattu 6.3.2012.]
- Steiner, F. 2004: *Taboo*. Routledge, London.
- STT 2011 = *Persusta tuli kielletty sana eduskunnassa*. STT–Lehtikuva, Helsinki. (Julk. esim. HS 20.9.2011). Saatavissa: www.hs.fi/politiikka/artikkeli/Persusta+tuli+kielletty+sana+eduskunnassa/1135269477714. [Viitattu 20.9.2012.]
- Suomalais-udmurtilainen sanakirja = Maksimov, Sergej – Saarinen, Sirkka – Danilov, Vadim – Seliverstova, Ekaterina 2013: *Suomalais-udmurtilainen sanakirja. Apuneuvoja suomalais-ugrilaisten kielten opintoja varten XVI*. SUS, Helsinki.
- Suomalais-venäläinen metsäsanakirja. Noin 5000 termiä*. Laatineet Kudasheva, Irina – Kudashev, Igor. Toim. Inkeri Vehmas-Lehto – Aleksandr Gerd. Metsäakustanus, Helsinki 2008.
- Suomen kielen käännteissanakirja. Reverse dictionary of modern standard Finnish*. Koostanut Tuomo Tuomi. SKST 274. SKS, Helsinki 1972.
- Suomen kielen perussanakirja 1–3*. Päätoim. Risto Haarala. KKTKJ 55. KKTK – Valtion painatuskeskus, Helsinki 1990–94.
- Suomen kielen tulevaisuus 2009 = Suomen kielen tulevaisuus. Kielipoliittinen toimintaohjelma*. Hakulinen, Auli – Kalliokoski, Jyrki – Kankaanpää, Salli – Kanner, Antti – Koskeniemi, Kimmo – Laitinen, Lea – Maamies, Sari – Nuolijärvi, Pirkko. KKTK:n verkkojulkaisuja 7. KKTK, Helsinki 2009. Saatavissa: http://scripta.kotus.fi/www/verkkojulkaisut/julk7/suomen_kielen_tulevaisuus_kotus_verkkojulkaisuja_7.pdf.
- Suomen murteiden sanakirja 1–8*. Toim. Tuomo Tuomi – Matti Vilppula ym. KKTKJ 36. KKTK–SKS, Helsinki 1985–2008.
- Suomi24.fi, Romano-keskustelupalsta = <http://keskustelu.suomi24.fi/debate/4143>. [Viitattu 1.11.2012.]
- SUS = Suomalais-Ugrilainen Seura.
- SUST = Suomalais-Ugrilaisen Seuran Toimituksia.
- Suuri suomi–ruotsi cd-rom-sanakirja*. Toim. Ilse Cantell – Nina Martola – Birgitta Romppanen – Mats-Peter Sundström. KKTK–WSOY, Helsinki 1998.
- Suuri suomi–ruotsi -sanakirja*. Toim. Ilse Cantell – Nina Martola – Birgitta Romppanen – Mats-Peter Sundström. KKTK–WSOY, Helsinki 1997.
- Svensén, Bo 2004: *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. 2., uudistettu ja laajennettu painos. Norstedts Akademiska Förlag, Stockholm.
- Svensk handordbok*. Toim. Ture Johannisson – K.G. Ljunggren. Läromedelsförlagen & Språkförlaget, Stockholm 1966.
- Svenska Akademiens ordlista*. P.A. Norstedt – Söners förlag, Stockholm 1973.
- Svenskt språkbruk ks. *Norstedts Svenskt språkbruk*.
- Syrjinnän vastainen käsikirja*. IOM:lle tuottanut Timo Makkonen. Kansainvälinen

- siirtolaisuusjärjestö IOM, Baltian ja Pohjoismaiden aluetoimisto, Helsinki 2003.
- Tammilahti 1988 = *Русско-финский технический словарь: ок. 100 000 терминов / Вейкко Таммиллахти. Хельсинки 1988.*
- Tanskan LEXIN = EMU Lexin Billedtema. Ministeriet for Flygtninge, Indvandrere og Integration, København. Saatavissa: lexin-billedtema.emu.dk/billedtema/. [Viitattu 24.9.2012.]
- Tarp, Sven 2008: Kan brugerundersøgelser overhovedet afdække brugernes leksikografiske behov? *LexicoNordica* 12, s. 5–32.
- 2009: Beyond Lexicography: New Visions and Challenges in the Information Age. *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographic Tools Tomorrow*, s. 17–32. Toim. Henning Bergenholtz – Sandro Nielsen – Sven Tarp. Peter Lang, Bern.
- Tavaramerkkilaki. 10.1.1967/7, ks. Finlex.
- Tekijänoikeuslaki. 8.7.1961/404, ks. Finlex.
- TEM = Maahanmuuttajien kotouttaminen. Työ- ja elinkeinoministeriö. Saatavissa: www.tem.fi/index.phtml?s=4626. [Viitattu 24.9.2012.]
- TENK: käytäntö = Hyvä tieteellinen käytäntö ja sen loukkausepäilyjen käsitteleminen Suomessa. Tutkimuseettinen neuvottelukunta, Helsinki 2012. Saatavissa: http://www.tenk.fi/sites/tenk.fi/files/HTK_ohje_2012.pdf. [Viitattu 27.3.2013.]
- TEO ks. Tekijänoikeuslaki.
- Thesleff, A. 1901: *Wörterbuch des Dialekts der finnländischen Zigeuner*. Acta Societatis Scientiarum Fennicae XXIX:6. Finnische Literatur-Gesellschaft, Helsinki.
- Tilastokeskus 2012 = Tilastokeskuksen PX-Web-tietokannat. Taulukko: Kieli iän ja sukupuolen mukaan maakunnittain 1987–2012. Saatavissa: 193.166.171.75/Dialog/varval.asp?ma=030_vaerak_tau_102_fi&ti=Kieli+i%E4n+ja+sukupuolen+mukaan+maakunnittain+1987+%2D+2012&path=../Database/StatFin/vrm/vaerak/&lang=3&multilang=fi. [Viitattu 27.3.2013.]
- Tolkki, Toivo 1951: Poliisi ja mustalaiset. *Poliisimies* 11.
- Trap-Jensen, Lars 2002: Normering og deskription i Den Danske Ordbog – mere eller mindre? *LexicoNordica* 9, s. 63–78.
- 2010: One, Two, Many: Customization and User Profiles in Internet Dictionaries. *Proceedings of the XIV Euralex International Congress (Leeuwarden, 6–10 July 2010)*, s. 1133–1143. Toim. Anne Dykstra – Tanneke Schoonheim. Fryske Akademy, Afûk, Ljouwert.
- Tscheremissisches Wörterbuch = Moisio, Arto – Saarinen, Sirkka 2008: *Tscheremissisches Wörterbuch aufgezeichnet von Volmari Porkka, Arvid Genetz, Yrjö Wichmann, Martti Räsänen, T. E. Uotila und Erkki Itkonen*. Lexica Societatis Fenno-Ugricae XXXII. SUS–KKTK, Helsinki.
- Tuomi, Tuomo 1989: Yleiskielemme murrepohjainen sanasto. *Nykysuomen sanavarat*, s. 28–56. Toim. Jouko Vesikansa. Werner Söderström osakeyhtiö, Porvoo–Helsinki–Juva.
- Tuominen, Markku – Wihersaari, Jari 2006: *Ammattikasvatusfilosofia*. OKKA-säätiö, Helsinki.

- Turunen, Aimo 1949: *Kalevalan sanakirja*. SKST 235. SKS, Helsinki.
- Tyysteri, Laura 2011: Sukupuolen merkitseminen Nykysuomen sanakirjan ja Kielitoimiston sanakirjan *nais-* ja *mies-*alkuisissa yhdyssubstantiiveissa. *SJ* 53, s. 162–180.
- Työsuojelusanasto*. TSK 35. Sanastokeskus TSK, Helsinki 2006.
- Törnqvist, Lars 2010: Ordböcker på Internet och Internet som ordbok. *Nordiska studier i lexikografi* 10, s. 484–493. Toim. H. Lönnroth – K. Nikula. Skrifter utgivna av Nordiska föreningen för lexikografi Skrift nr 11. Tammerfors universitet – Språkrådet i Norge, Tammerfors.
- Törnroos, Lars 2007: Språklig och encyklopedisk information i fackordlistor på Internet. *LexikoNordica* 14, s. 35–47.
- Udmurttilais-suomalainen sanakirja = Maksimov, Sergej – Danilov, Vadim – Saari, Sirkka 2008: *Udmurttilais-suomalainen sanakirja*. Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 79. Turun yliopisto, Turku.
- UEW = Károly Rédei 1988: *Uralisches etymologisches Wörterbuch*. Akadémiai Kiadó, Budapest.
- Urbaani sanakirja*. Nixarn, Turku 2007–. Saatavissa: www.urbaanisankirja.com. [Viitattu 6.5.2013.]
- URS = *Удмурт-зүч кыллюкам – Удмуртско-русский словарь*. Ред. Л. Е. Кириллова. УИИЯЛ УрО РАН, Ижевск 2008.
- US ks. *Urbaani sanakirja*.
- Uspenski, Eduard 1977: *Krokotiili Gena ja hänen ystävänsä*. Otava, Helsinki.
- W3 = *Webster's Third New International Dictionary*. G. & C. Merriam Company, Springfield, Massachusetts 1961.
- Valtonen, Pertti 1968: *Suomen mustalaiskielen kehitys eri aikoina tehtyjen muistiinpanojen valossa*. Julkaisematon lisensiaattityö. Helsingin yliopisto, Helsinki.
- Valtonen, Pertti 1972: *Suomen mustalaiskielen etymologinen sanakirja*. Tietolipas 69. SKS, Helsinki.
- Vanhatalo, Ulla 2002: ”Naiset motkottaa aiheesta ja nalkuttaa syyttä”. Kyselytesti verbien semanttisten sisältöjen arvioinnissa. *Vir.*, s. 330–353.
- 2005a: Kyselytestit synonymian selvittämisessä – sanastotietoutta kielenpuhujilta sähköiseen sanakirjaan. *Vir.*, Verkkoilite, s. 1–5. Verkkoilite. Saatavissa: www.kotikielenseura.fi/virittaja/hakemistot/jutut/2005_3vanhatalo.pdf. [Viitattu 2.1.2013.]
- 2005b: *Kyselytestit synonymian selvittämisessä — sanastotietoutta kielenpuhujilta sähköiseen sanakirjaan*. Yliopistopaino, Helsinki. Saatavissa: <https://helda.helsinki.fi/bitstream/handle/10138/19638/kyselyte.pdf?sequence=2>.
- Varantola, Krista 1998: Translators and their Use of Dictionaries: User Needs and User Habits. *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*, s. 179–204. Toim. B. T. Sue Atkins. Max Niemeyer Verlag, Tübingen.
- 2001: Sanakirjoista ja sanakirjojen käyttäjästä. *Alussa oli käänös*, s. 215–238. Toim. Riitta Oittinen – Pirjo Mäkinen. Tampere University Press, Tampere.
- 2002: Use and Usability of Dictionaries: Common Sense and Context Sensibi-

- lity? *Lexicography and Natural Language Processing. A Festschrift in Honour of B. T. S. Atkins*, s. 30–44. Toim. Marie-Hélène Corréard. S. l., EURALEX.
- Variarum rerum vocabula Latina, cum Svetica et Finnonica Interpretatione*. Editio prioribus auctior et emendatior. Holmiæ, typis Ignatii Meureri 1644.
- Varpio, Yrjö 1999: Suomenkielisen kirjallisuuden varhaisvaiheita. *Suomen kirjallisuushistoria 1. Hurskaista lauluista ilostelevaan romaaniin*, s. 92–106. Toim. Yrjö Varpio – Liisi Huhtala. SKST 724. SKS, Helsinki.
- Vasiljev 1926 = Ёпымарий (Васильев, В. М.): *Марий мутэр. Тўрлў вэрэ ылымэ марийын мутшым тагастарэн нэргэлымэ кнага*. СССР калык-влак рўдў савыктыш, Моско 1926.
- Weber, Max 2009 [1927]: *General economic history*. 10. printing, Transaction Publishers, New Brunswick, New Jersey. Alkuteos *Wirtschaftsgeschichte* 1927. Duncker & Humblot, München 1923.
- Welker, Herbert Andreas 2010: *Dictionary Use. A General Survey of Empirical Studies*. Author's Edition, Brasília. Saatavissa: www.let.unb.br/hawelker/dictionary_use_research.pdf. [Viitattu 20.3.2012.]
- Verkuyl, Henk – Janssen, Maarten – Jansen, Frank 2003: The codification of usage by labels. *A Practical Guide to Lexicography*, s. 297–311. Toim. Piet van Sterkenburg. John Benjamins Publishing Company, Amsterdam.
- Vesikansa, Jouko 1977: Viisikymmentä vuotta Nykysuomen sanakirjan alullepanosta. *Vir.*, s. 323–330.
- Wiberg, Matti 2011: *Politiikan sanakirja. Inhorealistin versio*. Toim. Kalevi Koukkunen – Matti Wiberg. Kustannusosakeyhtiö Siltala, Helsinki.
- Wierzbicka, Anna 2010: *Experience, Evidence, and Sense: The Hidden Cultural Legacy of English*. Oxford University Press, Oxford.
- Vihonen, Sakari 1972: Ideologisen kielen semanttista analyysia. *Vir.*, s. 262–277.
- Wikipedia EN = *Wikipedia, The Free Encyclopedia*. Wikimedia Foundation – Creative Commons. Saatavissa: <http://en.wikipedia.org>. [Viitattu 6.5.2013.]
- Wikipedia FI = *Wikipedia, vapaa tietosanakirja*. Wikimedia Foundation – Creative Commons. Saatavissa: <http://fi.wikipedia.org>. [Viitattu 20.9.2012.]
- Wiktionary = Wikisanakirja. Vapaa sanakirja. Wikimedia Foundation – Creative Commons. Saatavissa: <http://fi.wiktionary.org/wiki/Etusivu>. [Viitattu 16.10.2012.]
- Viljanen, Anna Maria 1974: *Mary Douglasin symboliteorian sovellus Suomen mustalaisilta kerättyyn*. Pro gradu. Helsingin yliopisto, Helsinki.
- Viljanen-Saira, Anna Maria 1979: *Mustalaiskulttuuri ja kulttuurin muutos*. Julkaisen maton lisensiaattityö. Helsingin yliopisto, Helsinki.
- Viljanen, Anna Maria 1994: *Psykiatria ja kulttuuri. Tutkimus oikeuspsykiatrisesta argumentaatiosta*. Stakes – Suomen Antropologinen Seura, Helsinki.
- Williams, John 1992: The Question of Plagiarism and Breach of Copyright in the Dictionary-Making Process (with Particular Reference to the UK). *EURALEX '92 Proceedings I–II.: Papers Submitted to the 5th EURALEX International Congress on Lexicography in Tampere, Finland*, s. 561–570. Toim. Hannu Tommola – Krista Varantola – Tarja Salmi-Tolonen – Jürgen Schopp. University of

- Tampere, Tampere.
- Williamson, Timothy 2009: Reference, inference and the semantics of pejoratives. *The Philosophy of David Kaplan*, s. 137–159. Toim. Joseph Almog – Paolo Leonardi. Oxford University Press, Oxford.
- Vilppula, Matti 2006: Suuret sanakirjamme. Puhe Kotuksen 30-vuotisjuhlassa 3.3.2006. Saatavissa: www.kotus.fi/index.phtml?s=2763. [Päivitetty 1.7.2008/ viitattu 16.10.2012.]
- Winchester, Simon 1998: *The Professor and the Madman. A Tale of Murder, Insanity, and the Making of the Oxford English Dictionary*. HarperCollins Publishers, New York.
- Vir. = *Virittäjä*. Kotikielen Seura, Helsinki.
- Wittgenstein, Ludwig 1999 [1953]: *Filosofisia tutkimuksia*. Suom. Heikki Nyman. WSOY, Helsinki. Alkuteos *Philosophische Untersuchungen* postuumisti 1953.
- VKS = *Vanhan kirjasuomen sanakirja 1–2*. KKTKJ 33. KKTK – Valtion painatuskeskus, Helsinki 1985–1994.
- Woordenboek der Nederlandsche Taal*. Instituut voor Nederlandse Lexicologie 1863–2001. Saatavissa: <http://gtb.inl.nl/>.
- Voutilainen, Eero 2012: Rekisteri. *Genreanalyysi – tekstilajitutkimuksen käsikirja*, s. 70–76. Toim. Vesa Heikkinen – Eero Voutilainen – Petri Lauerma – Ulla Tiirilä – Mikko Lounela. KKTKJ 169. Gaudeamus, Helsinki.
- VTEO = *Viittomakielen tulkin eettiset ohjeet*. Kuurojen liitto, Espoo 1982.
- Zgusta, Ladislav 1971: *Manual of Lexicography*. Academia. Publishing House of the Czechoslovak Academy of Sciences, Pragua – Mouton, Hague–Paris.
- Zorc, R. David – Osman, Madina M. 1993: *Somali-English Dictionary with English Index*. 3. p. Dunwoody Pr, Kensington, MD.
- Åkerlund, Tuula 2002: Oman kielen merkitys romanien kulttuuri-identiteetille. *Moniääninen Suomi*, s. 126–129. Toim. Sirkka Laihiala-Kankainen – Sari Pietikäinen – Hannele Dufva. Soveltavan kielentutkimuksen laitos, Jyväskylän yliopisto, Jyväskylä.
- Östman, Jan-Ola 2000: Merkityksen ajankohtaiset haasteet yleiselle kielitieteelle. *Merkillinen merkitys*, s. 69–84. Toim. Anu Airola – Heikki J. Koskinen – Veera Mustonen. Gaudeamus, Helsinki. Saatavissa: <https://helda.helsinki.fi/bitstream/handle/10224/3610/ostman69-84.pdf?sequence=1>. [Viitattu 20.3.2012.]

Teoksen kirjoittajat

Kimmo Granqvist

Fil.dr, dos., yliopistonlehtori

Suomen kielen, suomalais-ugrialaisten ja pohjoismaisten kielten ja kirjallisuuksien laitos,
Helsingin yliopisto

Eija-Riitta Grönros

FM, Kielitoimiston sanakirjan päätoimittaja

Kotimaisten kielten keskus

Raimo Jussila

FL, toimittanut useita sanakirjoja ja toiminut Suomen tietokirjailijat ry:n toiminnanjohtajana

Igor Kudashev

FT, dos., erikoissuunnittelija (T&K)

Koulutus- ja kehittämiskeskus Palmenia (Kouvola), Helsingin yliopisto

Pirkko Kuutti

FM, Vanhan kirjasuomen sanakirjan päätoimittaja

Kotimaisten kielten keskus

Kaisa Lehtosalo

FM, toiminut Agricolan sanakirjan toimittajana

Suomen kieli, Turun yliopisto

Nina Pilke

FT, professori

Pohjoismaisten kielten yksikkö, Vaasan yliopisto

Jaana Puskala

FT, yliopistonlehtori

Pohjoismaisten kielten yksikkö, Vaasan yliopisto

Klaas Ruppel

FM, sanakirjaosaston johtaja

Kotimaisten kielten keskus

Sirkka Saarinen

FT, professori

Suomalais-ugrilainen kielentutkimus, Turun yliopisto

Laura Tyysteri

FM, väitöskirjatutkija

Suomen kieli, Turun yliopisto