

# Kielitilasto maahanmuuttajien väestöosuuden mittarina

SIRKKU LATOMAA

## Johdanto

Pari vuotta sitten tarkastelin pohjoismaisia maahanmuuttotilastoja erästä konferenssiesitelmää varten ja huomioni kiinnittyi terminologisiin valintoihin. Muissa Pohjoismaissa käytettiin käsitteitä *immigrant population*, *immigrants and their descendants*, *foreign-born population*. Suomalaiset tilastot eivät sen sijaan luonnehtineet tätä väestönosaa näillä käsitteillä, vaan englanniksi tietoa hakiessani päädyin esimerkiksi termiin *foreigners*. Lisäksi havaitsin, että muissa Pohjoismaissa oli mahdollista saada tietoa sekä maahan muuttaneista että heidän jälkeläisistään, kun taas suomalaisten tilastointitapojen – syntymävaltio, kansalaisuus, äidinkieli – perusteella tietoa oli saatavilla lähinnä ensimmäisestä sukupolvesta. Ilmiötä nimeltä maahanmuutto kuvataan siis Suomessa erilaisin käsittein ja mittarein kuin muissa Pohjoismaissa.

Koulutuksen ja tutkimuksen kehittämissuunnitelmassa (OKM 2011, 8) esitetään, että peruskoulun edellytyksiä toimia kaikkien lasten oppivelvollisuuskouluna parannetaan. Tavoitteena on, että perusopetuksen rahoitus perustuu nykyistä enemmän koulujen toimintaympäristöä kuvaaviin indikaattoreihin, muun muassa kunnassa asuvien maahanmuuttajien väestöosuuteen. Indikaattorina ”maahanmuuttajien väestöosuus” kuulostaa yksinkertaiselta määriteltävältä, mutta mikäli sitä arvioidaan nykyiseen tapaan syntymämaan, kansalaisuuden tai rekisteröidyn äidinkielen perusteella, kuva koulujen todellisesta toimintaympäristöstä jää auttamatta vajaaksi.

Väitettäni havainnollistaa seuraava esimerkki: Opetushallituksen mukaan perusopetuksessa oli vuonna 2009 noin 19 200 vieraskielistä – toisin sanoen muuta kuin suomea, ruotsia tai saamea äidinkielenään puhuvaa – oppilasta (Kumpulai-

nen 2010, 14). Jos lähtökohdaksi otetaan oppilaiden rekisteröidyn äidinkielen sijaan vanhempien taustat, luku on melkein kaksinkertainen: 36 150. Tämä luku sisältää ne 7–15-vuotiaat oppilaat, joiden vanhemmista vähintään toinen on syntynyt Suomen ulkopuolella ja puhuu muuta kieltä kuin suomea tai ruotsia äidinkielenään (Saari 2011). Luku sisältää siten myös niin sanotun 2,5-sukupolven (Rumbaut 2004) eli maahanmuuttajan ja kantaväestöön kuuluvan yhteiset lapset.

Pohdin tässä artikkelissa, miten tilastointia voisi kehittää siihen suuntaan, että kuva esimerkiksi koulujen ja työpaikkojen kulttuurisesta ja kielellisestä moninaisuudesta olisi nykyistä tarkempi. Keskityn esityksessäni erityisesti kielten tilastointiin, koska viime vuosina kielitilastoa on käytetty yhä useammin maahanmuuttajien väestöosuuden mittarina. Kuvaan aluksi väestönlaskennan historiaa kielten tilastoinnin näkökulmasta ja luonnehdin suomalaisen yhteiskunnan monikielisyttä nykytietojen valossa. Esittelen myös kielten tilastointiin liittyviä haasteita sekä kielitilastosta käytyä keskustelua. Sen jälkeen tarkastelen eri maissa saatavilla olevia tietoja asukkaiden kielitaustasta sekä perusteita kielten tilastoinnin tarpeelle. Lopuksi pohdin, mihin suuntaan kielitilastointia voitaisiin kehittää tulevaisuudessa.

## Suomalaisen kielitilastoinnin taustaa

Kielten tilastoinnilla on ollut keskeinen asema Suomen väestötiedoissa, sillä maan viralliseen kaksikielisyyteen liittyvät oikeudet ja velvollisuudet perustuvat suomen- ja ruosinpuhujien määrään. Tietoa tarvitaan, jotta kunnat voidaan määrittellä kaksi- tai yksikielisiksi sen mukaan, mikä kielenpuhujien suhteellinen määrä on kunnassa.

Tähän statukseen perustuu sitten se, millä kielellä asukas voi hoitaa asioitaan kunnan viranomaisissa (ks. esim. OM 2009).

Väestötietoja on saatavilla vuodesta 1749, ja kielen mukaan väestöä on tilastoitu vuodesta 1880 (McRae 1999, 83). Vuonna 1870 aloitettiin suurten kaupunkien väestönlaskennat, mutta vasta vuonna 1950 järjestettiin ensimmäinen yleinen väestönlaskenta. Siitä lähtien lomakkeisiin perustuvia valtakunnallisia väestönlaskentoja tehtiin kymmenen vuoden välein. (Allardt & Starck 1981, 124.) Vuodesta 1990 väestötiedot ovat perustuneet erilaisiin tietokantoihin.

Eri aikoina asukkaiden kielistä on kysytty eri tavoin. Vuonna 1880 väestönlaskennassa kysyttiin ”äidinkieltä”, joka määriteltiin koulutuksen ja jokapäiväisen kanssakäymisen kieleksi. Vuonna 1890 määritelmään liitettiin lisäohje: ”Epäselvissä tapauksissa tulee kristinuskon opetuskielen olla määräävä.” (Mt., 124.) 1900-luvun mittaan tilastoitavan kielen kriteerit muuttuivat. Vuoteen 1930 asti tilastoitiin ”kieli, jota tavallisesti puhuu”. Vuonna 1940 kriteeriksi tuli ”äidinkieli”, joka ymmärrettiin parhaiten osatuksi kieleksi. (McRae 1999, 82–83.) Vuoden 1950 väestönlaskennassa kysyttiin ”parhaiten hallussa olevaa kieltä”, ja lisäohjeena annettiin: ”Ilmoittakaa epäselvissä tapauksissa se kieliryhmä, johon katsotte lähinnä kuuluvan.” Vuosina 1960 ja 1970 tiedusteltiin saman lisäohjeen kera ”pääkieltä”. (Allardt & Starck 1981, 124–125.)

Vuodesta 1980 Tilastokeskus on käyttänyt Väestörekisterikeskuksen keräämiä tietoja. Tuolloin kriteeriksi muodostui ”äidinkieli”, mutta tällä kertaa sillä tarkoitettiin vuoden 1922 kielilaissa määriteltyä ”omaa kieltä”. (McRae 1999, 83–84.) Vuosikymmenten mittaan asukkaiden kieliä on rekisteröity siis ainakin kolmen erilaisen kriteerin avulla: käyttö, taito ja samastuminen (vrt. äidinkielen määritelmät Skutnabb-Kangas 1981, 21–26). Toisinaan käytössä on ollut yksi kriteereistä, toisinaan useampi.

Tänä päivänä kieli rekisteröidään väestötietoihin seuraavasti: Maistraateissa käytössä olevissa lomakkeissa kysytään äidinkieltä. Sitä ei erikseen määritellä. Äidinkieleksi voi valita suomen, ruotsin tai muun. Kohdassa ”muu” on valikko, jossa on 180 kielivaihtoehtoa. Jos omaa äidinkieltä ei löydy valmiiksi annetuissa vaihtoehtoissa, on mahdollista valita kohta ”kirjoitettuna alla” se-

kä kirjoittaa kyseiseen kohtaan jokin muu kieli.<sup>1</sup> Vanhemmat ilmoittavat syntyneelle lapselle kielen, ja se säilyy samana, ellei sitä erikseen muuteta. Jos lapsi syntyy vuoden lopulla eikä hänelle ole ilmoitettu tammikuun aikana kieltä, vuodenvaihteen tilastoon viedään kieleksi äidille rekisteröity kieli. Väestörekisterin tiedot alle 18-vuotiaista lapsista perustuvat huoltajan ilmoitukseen (Laki lapsen huollosta ja tapaamisoikeudesta 4 §). Äidinkielen lisäksi väestötietojärjestelmään voi merkitä asiointikieleksi suomen tai ruotsin.

## Suomen monet kielet

Viime vuosikymmenten aikana suomalainen yhteiskunta on monikielistynyt merkittävästi. Ilmiönä monikielisyys ei kuitenkaan ole uusi, sillä esimerkiksi jo 1870-luvulla Helsinki oli huomattavan monikielinen. Väestöstä oli tuolloin suomenkielisiä 25,8, ruotsinkielisiä 57,0, venäjänkielisiä 12,1 ja saksankielisiä 1,7 prosenttia. Viipuri oli vielä Helsinkiäkin monikielisempi: muita kuin suomea tai ruotsia puhuvia henkilöitä asui kaupungissa tuolloin 32,4 prosenttia. (Tilastokeskus 2009.) Venäjänkielisten määrä oli vuonna 1870 yksistään Helsingissä 3 878 (Leitzinger 2008, 78), kun se vuonna 1990, satakaksikymmentä vuotta myöhemmin, oli koko maassa lähes samansuuruinen, 3 884 (SVT 2012).

Miltä sitten näyttää Suomen kielellinen moninaisuus tänä päivänä kielitilaston kuvaamana? Kaikkiaan niin sanottuja vieraskielisiä asukkaita oli vuoden 2011 lopussa 244 827 eli 4,5 prosenttia koko väestöstä. Väestön monikielisyys on kuitenkin erityisesti isojen asutuskeskusten piirre. Pääkaupunkiseutu on monikielistynyt muuta maata nopeammin. Vuoteen 2030 mennessä muuta kuin suomea tai ruotsia äidinkielenään puhuvan väestön osuuden on ennustettu kasvavan 15 prosenttiin ja määrän kasvavan 100 000:sta 250 000:een (Helsingin kaupungin tietokeskus 2010).

Taulukkoon 1 on ryhmitelty tilaston rekisteröimät kielet puhujamäärän mukaan. Tilastokeskuksen tietoja on muutettu siten, että kielten kirjoitusasu on Suomen Standardoimisliiton suosi-

<sup>1</sup> Maistraatin lomakkeessa sanotaan: ”Yhdenvertaisena rekisteriin voidaan merkitä muun itsenäisen kielijärjestelmän kieli henkilön äidinkieleksi.”

*Taulukko 1. Kielet ryhmiteltynä puhujamäärän mukaan suuruusjärjestyksessä 31.12.2011 (SVT 2012).*

| Puhujamäärä   | Kieli  |
|---------------|--|
| yli 4 800 000 | suomi (1)  |
| yli 290 000   | ruotsi (1)   |
| yli 50 000    | venäjä (1)   |
| yli 10 000    | viro, somali, englantia, arabia (4)  |
| yli 1 000     | kurdi, kiina, albania, thai, vietnam, turkki, farsi, saksa, espanja, ranska, puola, unkari, bengali, romania, saame, bosnia, italia, portugali, tagalog, ukraina, nepali, urdu, serbokroatia, bulgaria, hindi, hollanti, japani, tamil, amhara (29)  |
| 500–1 000     | swahili, latvia, liettua, kreikka, pandžabi, lingala, telugu, norja, paštu, burma (10)   |
| 100–499       | yoruba, tanska, tšekki, ruanda, korea, heprea, igbo, tšetšeeni, indonesia, azeri, slovakki, serbia, malajalam, tvi, akan, singali, khmer, turkmeeni, armenia, tigrinja, kikuju, marathi, makedonia, wolof, kongo, tataari, kroatia, kannada, islanti, georgia, valkovenäjä, uzbekki, uiguuri (33)                |
| 10–99         | oromo, malaiji, sloveeni, hausa, fulfulde, gudžarati, kazakki, ganda, afrikaans, katalaani, ewe, orija, moldavia, mongoli, luba, rundi, sindhi, tadžikki, lao, shona, komi, ndonga, malta, osseetti, kirgiisi, baski, iiri, kašmiri, luxemburg, fääri, malagasi, njandža, zulu (33)                              |
| 1–9           | sunda, avaari, tiibet, baškiiri, pohjoisdebele, sango, tongantonga, ido, bambara, cree, tswana, assami, tšamorro, jaava, gaeli, aimara, ketšua, swazi, venda, xhosa, dzonkha, bihari, esperanto, samoa, sotho, tšuvassi, abhaasi, bislama, fidži, friisi, galicia, grönlandi, herero, kanura, kymri, tahiti (36) |

tusten mukainen (SFS 2008). Rinnakkaismuodoista on valittu Standardoimisliiton suosittelumista nimistä ensimmäinen.

Kymmenessä vuodessa kaikki kieliryhmät ovat kasvaneet ja uusia kieliä on ilmaantunut suomalaiseseen kielimaisemaan. Kun vuonna 2001 yli tuhannen puhujan kieliä oli 22, vuonna 2011 niitä oli 36. Suuntaus on se, että yhä useampi aasia-

lais- tai afrikkalaisperäinen kieli on kielitilaston kärjessä. Eurooppalaisetkin kieliryhmät kasvavat, mutta hitaammin. Suomen ja ruotsin jälkeen seuraavien kymmenen kielen järjestys on pysynyt pitkälti samana parin viime vuosikymmenen ajan.<sup>2</sup> Venäjä on kuitenkin muita kieliryhmiä huomattavasti suurempi. Vuoden 2011 lopussa Suomessa asui 58 331 venäjän äidinkielekseen rekisteröinyttä henkilöä, toisin sanoen lähes neljäsosa kaikista muuta kuin suomea tai ruotsia puhuvista asukkaista. Yli kymmentuhannen puhujan ryhmiä oli venäjän ohella neljä: vironkielisiä oli 33 076, somalinkielisiä 14 045, englanninkielisiä 13 804 ja arabiankielisiä 11 252.

Maakunnat ja kunnat eroavat kieliprofileiltaan jossain määrin. Maakunnittain tarkasteltuna venäjänkieliset ovat suurin ryhmä kaikissa muissa paitsi Kanta-Hämeen ja Ahvenanmaan maakunnissa. Niissä suurimpina ryhminä ovat viron- ja latviankieliset. Venäjänkieliset ovat suurin ryhmä 177 kunnassa ja vironkieliset 117 kunnassa. (Rapo 2011.) Suurten kaupunkien kieliprofiili muistuttaa pitkälti koko maan kieliprofiilia, mutta eroakin on. Helsingissä, Espoossa ja Vantaalla kolme suurinta kieltä ovat venäjä, viro ja somali. Vantaalle on keskittynyt enemmän venäjän-, somalin-, vietnamin-, turkin- ja albaniankielisiä, kun taas Espoossa asuu Vantaata enemmän viron-, englannin-, kiinan- ja saksankielisiä. (Saukkonen 2010, 22.) Tampereen suurimpiin kieliin lukeutuu puolestaan heti venäjän jälkeen arabia. Sitä vastoin somali ei tässä kaupungissa kuulu kymmenen suurimman kielen joukkoon. (Latomaa 2009, 240.)

Myös eri maat eroavat kieliprofileiltaan. Valuing All Languages in Europe -hankkeessa kartoitettiin Euroopassa puhuttuja maahanmuuttaja- ja vähemmistökieliä (ks. tarkemmin McPake & al. 2007). Kielistä laadittiin kymmenen kärjessä -lista sen mukaan, kuinka monen maan kartoituksessa kieli mainittiin (ks. taulukko 2). Lista kuvastaa, mitkä kieliryhmät ovat levittäytyneet muita laajemmalle alueelle Euroopassa. Esimerkiksi kärjessä olevat puola ja saksa raportoitiin 17:ssä mukana olleista 21 maasta ja kaksi

*2 Vuoden 1991 tilastossa mukana olivat muuten samat kielet, mutta kurdi, albania ja thai puuttuivat. Vuonna 2001 kurdi ja albania olivat nousseet kärkikymmenikköön, ja vuonna 2011 mukana oli myös thai. Kärkikymmeniköstä ovat poistuneet 20 vuoden aikana saksa, turkki ja espanja.*

Taulukko 2. Tyypilliset maahanmuuttaja- ja vähemmistökielet Euroopassa ja Suomessa (McPake ym. 2007, 27; SVT 2012).

| Eurooppa         | Suomi    |
|------------------|----------|
| puola            | venäjä   |
| saksa            | viro     |
| ranska           | somali   |
| arabia           | englanti |
| venäjä           | arabia   |
| espanja          | kurdi    |
| turkki           | kiina    |
| romani           | albania  |
| englanti         | thai     |
| mandariini kiina | vietnam  |

viimeistä eli englanti ja mandariini kiina 13 maassa. Taulukko 2 havainnollistaa Suomen ja koko Euroopan kieliprofilien eroja kymmenen laajimmin käytössä olevan maahanmuuttaja- ja vähemmistökielen perusteella.

Suomea koskevassa sarakkeessa ovat *kursivoituina* ne kielet, jotka kuuluvat Euroopan tyypillisiin maahanmuuttajakieliin. Lihavoidut kielet – viro, somali, kurdi, albania, thai, vietnam – kuvastavat taas Suomen maahanmuuton erityispiirteitä suhteessa yleiseurooppalaiseen muuttoliikkeeseen. Suomalainen maahanmuutto painottui alkuvaiheessa pakolaisuuteen, ja nämä jäljet näkyvät vielä tänään kielitilaston kärkipäässä. Vironpuhujien suuri määrä perustuu puolestaan maiden läheisyyteen. Työperusteisen maahanmuuton myöhäisyys näkyy esimerkiksi siten, että puola ja turkki eivät ole Suomessa kymmenen kärjessä -listalla, kun taas monissa muissa maissa ne kuuluvat suurimpiin kieliryhmiin.

Vaikka Tilastokeskuksen tarjoaman tiedon avulla selviääkin paljon Suomen kielellisestä moninaisuudesta, tilastoa tarkastellessa herää myös lukuisia kysymyksiä. Ensinnäkin voi kysyä, millaiseen kielen määrittelmään se perustuu. Lisäksi voi pohtia, mitä seurauksia siitä on, että kukin voi ilmoittaa vain yhden (äidin)kielen. Tarkastelen seuraavassa näitä kysymyksiä tarkemmin.

## Kielitilastoinnin haasteita: mikä on kieli?

On arvioitu, että maailmassa on kieliä noin 6 000. On monia syitä siihen, että täsmällistä lukumäärää on mahdotonta esittää (ks. Anhava 2002, 11–16). Yksi keskeisimmistä syistä on se, ettei kielelle ole olemassa yksiselitteistä määrittelmää. Määrittelytapa riippuu tarkoituksesta ja siitä, kuka kielen määrittelee. On mahdollista tehdä erottelu kielen ja murteen välillä puhtaasti lingvistisin perustein, kriteerinä ymmärrettävyys. Jos kahden toisistaan eroavan kielimuodon puhujat ymmärtävät toisiaan, kyse on eri murteista; jos eivät, kyse on eri kielistä. Jos kieli määriteltäisiin puhtaasti kielitieteellisin perustein, esimerkiksi ruotsia ja norjaa voisi pitää saman kielen eri muotoina. Koska niitä kuitenkin pidetään erillisinä kielinä, on selvää, että kielen määrittelyssä ovat tärkeitä myös kielenulkoiset tekijät. Ruotsin ja norjan – kuten monien muidenkin kielisten jatkumoiden – tapauksessa kielen ja murteen välinen rajanveto perustuu historiallisiin ja poliittisiin tekijöihin. (Karlsson 1994, 241–242; Anhava 2002, 12–13.)

Kielenpuhujillakin voi olla omia näkemyksiä siitä, mikä tekee heidän kielestään kielen eikä murteen. Nämä näkemykset perustuvat usein identiteettiin enemmän kuin kielen lingvistisiin piirteisiin. Tästä esimerkkejä ovat Ruotsissa meänkieli ja Norjassa kveeni, jotka usein esitetään osana suomen kielen murteita, kun taas kielimuodon käyttäjistä suuri osa katsoo puhuvansa omaa kieltä (ks. esim. Lane 2011; Söderholm 1998; Winsa 1993). Sekä meänkielillä että kveenillä on tätä nykyä virallisen vähemmistökielen asema.

Tilastoinnissa kielen määrittelee siitä vastuussa oleva viranomainen. Taulukon 1 perusteella kielen lainsäädännöllinen asema ei ilmeisesti vaikuta siihen, mitä kieliä tilastoinnissa otetaan huomioon, sillä tilastosta näytävät puuttuvan niin viittomakieli, romani, inarinsaame kuin koltansaamekin. Saamen eri kielten sijaan tilastossa esiintyy yleisnimitys saame. Muita useita kieliä kattavia nimityksiä tilastossa ovat esimerkiksi kiina, arabia ja farsi. Huomio kiinnittyy myös siihen, että serbian ja kroatian ohessa on tilastoitu kieli nimeltä serbokroatia. Seuraavassa kuvaan, millaiseen järjestelmään nykyinen tilastointimme perustuu.

## Kielitilastoinnin haasteita: miten kielet koodataan?

Kielten tilastointitapaa kehitettäessä on päädytty käyttämään kansainvälistä standardia. Aiemmin Väestötietokeskus koodasi kielet omilla luokituksillaan. Ohjeistuksessa vuodelta 1982 sanottiin seuraavasti: ”Mikäli lapsen äidinkielenä merkitään muu kuin suomi tai ruotsi, ilmoitetaan tämä selväkielisenä sekä ruutuun kirjoitetaan 3 saame, 4 venäjä, 5 tanska, norja tai islanti, 6 englanti, 7 saksa, 8 ranska, 0 muu kuin edellä mainittu.” (Nikander 2010.)

Vuonna 1999 Väestötietokeskus siirtyi kielikoodauksessa kansainväliseen ISO 639-1 -standardiin. Suomen Standardoimisliitto vahvisti vastaavan kansallisen standardin vuonna 2005 (SFS ISO 639-1). Se sisältää kaksikirjaimiset tunnukset 182 kielelle. Kaksikirjainkoodisto on laadittu alun perin bibliografiseen ja terminologiseen työhön. Koodisto kattaa suurimman osan sellaisista kielistä, jotka ovat useimmin edustettuina maailmankirjallisuudessa ja joilla on laadittu huomattava määrä erikoiskielisiä termistöjä.

Käytäntö on se, että Tilastokeskus saa Väestötietokeskukselta kielet valmiiksi koodattuina kansallisen standardin mukaisesti. Ne kielet, joille ei ole tunnusta, Väestötietokeskus toimittaa Tilastokeskukselle selväkielisinä. Tilastokeskuksessa käydään selväkieliset nimet läpi. Usein on osoittautunut, että selväkielisiin nimiin on merkitty useita koodautuvia kieliä vähän eri kirjoitusmuodossa kuin mitä standardissa on mainittu. Nämä kielet koodataan manuaalisesti, ja sen jälkeen lopullinen tilasto on valmis. (Rapo 2010.)

Useissa yhteyksissä on esitetty, että Suomessa puhutaan noin 150:tä kieltä. Luku sisältää kuitenkin vain ne kielet, joille on standardissa oma koodi. Muut kielet jäävät virallisen tilaston ulkopuolelle, joten osuvampi luonnehdinta olisi se, että Suomessa on tilastoitu noin 150 kieltä. Suomessa puhuttujen kielten kokonaismäärää ei siis ole tiedossa. Tuoreimman tilaston mukaan Suomessa on käytössä kansalliskielet mukaan lukien yhteensä 148 kieltä (ks. taulukko 1). Lisäksi tilastoon on merkitty kolme kieltä (divehi, latina<sup>3</sup>,

tsonga), joilla ei ole yhtään puhujaa, sekä kategoriari ”muu kieli” ja ”tuntematon”. Nollapuhujien kielten näkyminen tilastossa perustuu siihen, että kielellä on jossain vaiheessa ollut yksi puhuja (divehi: 2004–2009; latina: 2000, 2008; tsonga: 2001–2002) (Rapo 2010). Kaikkiaan ISO 639-1 -standardin kieliä, joilla ei toistaiseksi ole yhtään puhujaa Suomessa, on noin kolmekymmentä. Esimerkkejä tällaisista kielistä ovat kirkkoslaavi, bretoni, korni, valloni, norjan kirjanorja, volapük ja interlingua.

Vuoden 2011 tilastossa luokkaan ”muu kieli” oli merkitty yhteensä 4 388 puhujaa. Näillä henkilöillä on äidinkielenä sellainen kieli, jolla ei ole koodia standardissa. Luokka ”tuntematon” puolestaan tarkoittaa, että kielitieto puuttuu kokonaan. Vuoden 2011 tilastossa kieli puuttui 285 henkilön tiedoista.

Luokituksesta on poikettu saamen tapauksessa. Kansallinen standardi antaa koodin (se) pohjoissaamelle, mutta sekä Väestötietokeskus että Tilastokeskus käyttävät tämän koodin yhteydessä nimitystä saame.

Poikkeus on tehty myös serbokroatian (sh) tapauksessa. Kieli poistettiin ISO 639-1 -koodistosta vuonna 2000, ja sen korvasivat kroatia, serbia ja bosnia. Väestötietokeskuksen tiedoissa on kuitenkin yhä suurehko määrä henkilöitä, jotka ovat ilmoittaneet äidinkielenkseen serbokroatian. Koska näiden henkilöiden äidinkieltä ei voi uudelleen määrittellä tilastoinnin yhteydessä, kieli on säilytetty kieliluokituksessa. (SVT 2012.)

Väestötietokeskuksen käyttämä luokitus poikkeaa kansallisesta standardista myös kielten nimityksissä. Käytössä on sellaisia nimityksiä, joita ei ole standardissa tai joita ei enää pidetä suositeltavina (ks. SFS 2008), esimerkiksi ambo (po. ndonga), ivrit (po. heprea), grönlandineskimo (po. kalaallisut), skotti (po. gaeli), lubakatanga (po. luba) ja tonga (po. tongantonga). Myös virheellisiä kirjoitusasuja on lukuisia, esimerkiksi punjabi (po. pandžabi), chamorro (po. tšamorro), inupiak (po. inupiaq), manx (po. manksi), walloon (po. valloni), ojibwa (po. odžibwa), marsallese (po. marshall), navaho (po. navajo), pali (po. paali), tsuang (po. zhuang) ja pasto (po. paštu). Osa näistä virheistä toistuu Tilastokeskuksen tiedoissa, mutta osa on korjattu: esimerkiksi punjabin sijaan Tilastokeskus käyttää nimitystä pandžabi.

<sup>3</sup> Tämän taustalla lieene sekaannus: koodistossa on latina (la), mutta henkilö on luultavimmin tarkoittanut latin tai latinon kieltä.

## Kielitilasto maahanmuuttajataustaisen väestön mittarina

Viime aikoina on useissa yhteyksissä keskusteltu kielitilaston sopivuudesta maahanmuuttajataustaisen väestön määrän arviointiin. Tilastokeskuksessa työskentelevät Leena Kartovaara ja Markus Rapo puoltavat kieliperusteista tilastointia. Kartovaaran (2007) mukaan kielitilaston perusteella saadaan melko tarkka kuva tästä väestönosasta. Hän osoittaa maahanmuuttajataustaisia lapsia koskevassa artikkelissaan, että ”mutkikas syntymämaaluokittelu” ja äidinkielen perustuva luokittelu toisivat esiin 89-prosenttisesti samat lapset. Rapo (2011) puolestaan toteaa, että ”äidinkieli on ehdottomasti paras muuttuja ulkomaalaisten määrää kartoitettaessa”.

Tuomas Martikainen (2010) on sen sijaan kritisoinut kielitilaston käyttöä. Hän pitää kielitie-toja lähtökohtaisesti epäluotettavina, koska kielin rekisteröinti perustuu subjektiiviseen ilmoitukseen. Hän huomauttaa, ettei ole tutkimustietoa siitä, mikä on ilmoitetun kielen, kielitaidon ja kieli-identiteetin välinen suhde. Martikainen ehdottaa, että kielitiedon sijaan käytettäisiin syntymämaata, sillä se ei ole muutettavissa, kuten kieli ja kansalaisuus ovat. Hän painottaa, että erityisesti toista sukupolvea ei voi määrittellä pelkän kielitiedon perusteella, vaan heitä tutkittaessa olisi selvitettävä myös vanhempien syntymämaa.

Martikaisen kritiikki osuu siinä mielessä oikeaan, että tilastossa oleva kieli voi merkitä monenlaisia asioita jo siitä yksinkertaisesta syystä, ettei äidinkieltä ole määritelty mitenkään. Suurimmalla osalla väestöstä tieto on varmasti paikkansapitävä määrittelytavasta riippumatta, sillä suomen kieli on heidän ensiksi oppimansa, parhaiten osaamansa ja eniten käyttämänsä kieli ja lisäksi se edustaa kieliryhmää, johon puhujat mieltävät kuuluvansa (ks. Skutnabb-Kangas 1981: 21–26). Monikielisten henkilöiden tilanne on kuitenkin toinen: tilastointitapa pakottaa valitsemaan monista kielistä yhden. Valinnan takana voi tällöin olla monenlaista taitotasoa, käyttöasetta ja identifioitumista. Koska Suomessa ei saa tilastoida etnistä taustaa, kielen rekisteröinnillä annetaan samalla viesti siitä, mihin ryhmään halutaan identifioitua.

Kielivalinnan taustalla voi olla myös taktikointia, ja suomen kieli on valittu väestötietoihin muun muassa opiskeluun liittyvistä syistä. Vanhemmat voivat esimerkiksi arvella, että kou-

lutuksessa lapselle on eduksi, että rekisteriin on merkitty suomi eikä kotona puhuttu kieli. Vanhempien pelko ei ole aivan perusteeton, sillä kielirekisteriä on toisinaan käytetty kyseenalaisella tavalla. Toisen asteen ammatillisiin opintoihin hakiessaan kokelaan on osallistuttava kieliko-keeseen, jos väestötiedoissa on äidinkielenä muu kuin suomi. Jos kokelas sen sijaan häivyttää todellisen kielitaustansa rekisteristä, hän on samalla viivalla niin sanottujen syntyperäisten suomenpuhujien kanssa. Oppilaitosten käytännöstä tehtiin kirjallinen kysymys eduskunnalle helmikuussa 2009 (KK 4/2009 vp).

Kartovaara (2007) ja Rapo (2011) puhuvat kielitaustasta ”ulkomaalaisuuden” mittarina. Sen vastaparilla, suomalaisuudella, he eivät selvästikään tarkoita kansalaisuutta, vaan he näyttävät viittaavan sillä symboliseen suomalaiseen kansakuntaan, johon kuulumisen kriteerinä on suomen (tai ruotsin) kieli. Kun kielitilastossa on toinen kansalliskielistä, eivät asukkaat ole heidän määritelmänsä mukaan enää ”ulkomaalaistaustaisia”. Saman logiikan perusteella Suomessa ei myöskään olisi saamelastaustaisia enempää kuin 1 870, jonka verran vuoden 2011 kielitilasto ilmoittaa Suomessa asuvan (jonkin) saamen kielen puhujia. Kielitilaston tulkitseminen tällä tavoin yksikielistä ja -kulttuurista suuren osan Suomessa asuvista monikielistä henkilöistä. Muiden kielten osaaminenhan ei lopu siihen hetkeen, kun he ilmoittavat rekisteriin kielekseen suomen (tai ruotsin) kielen.

## Kielitilastot vertailussa

Sekä Kartovaara (2007) että Rapo (2011) esittävät, että äidinkieli-tieto puuttuu muista maista. Todellisuudessa Suomen tapa tilastoida asukkaiden äidinkieliä ei kuitenkaan ole poikkeuksellinen. UNEEC:n (YK:n Euroopan talouskomissio) tekemässä maahanmuuttotilastoja koskevassa vertailussa tutkituista 46 maasta 35 sisällytti väestönlaskentaan vähintään yhden kielitaustaa koskevan kysymyksen. Kahdessakymmenessäyhdeksässä maassa kysyttiin äidinkieltä, kolmessa pääkieltä, yhdessätoista eniten käytettyä kieltä, ja yhdeksässätoista maassa tiedusteltiin muiden kielten taitoa. (Aspinall 2007, 92; Poulain 2008, 50.)

Väestöä ei kuitenkaan tilastoida kielen perusteella kaikkialla. Kieltä koskevia kysymyksiä pidetään sensitiivisinä joissakin maissa, ja kielitau-

tan rekisteröimättä jättämiseen voi olla muitakin syitä. Yleisesti ottaen kieleen liittyviä kysymyksiä esitetään useammin Itä- kuin Länsi-Euroopassa, mutta kuten Jean Azar työtovereineen (2009) nostaa esiin, ei ole selvää yhteyttä maan poliittisen järjestelmän tai maan harjoittaman kielipoliitiikan ja kielitilastoinnin välillä. Kielitaustasta ei kysytä esimerkiksi Albaniassa, Bosnia ja Herzegovinassa, Saksassa, Italiassa, Iranissa ja Zimbabwessa. (Azar & al. 2009; Simon 2007, 113–117.) Pohjoismaista Norja, Ruotsi ja Tanska eivät tilastoi asukkaiden kielitaustaa, mutta Ruotsissa ja Tanskassa kerätään tietoja esimerkiksi perusopetusikäisten lasten äidinkielistä (Azar & al. 2009; Risager 2006; Simon 2007, 114).<sup>4</sup>

Kuten Erik Allardt ja Christian Starck (1981, 123) ovat huomauttaneet, kielten tilastoinnissa käytetyt määritelmät eivät ole poliittisesti viattomia, vaan niiden valinta vaikuttaa suoraan siihen, miten suurelta tai pieneltä kielivähemmistö näyttää. Kielitilasto on tässä mielessä sosiaalisen kontrollin väline. UNECE onkin antanut suosituksia siitä, miten kielitaustasta pitäisi kysyä. Asukkailta pitäisi kysyä vähintään kahta asiaa, esimerkiksi lapsuudessa ensin opittua kieltä, parhaiten hallittua kieltä, kyselyhetkellä eniten käytettyä kieltä tai yleisesti kiel(t)en taitoa eriteltyinä osa-alueittain. Valmiiksi tarjottujen vaihtoehtojen lisäksi tulisi antaa mahdollisuus avoimiin vastauksiin. Valitut käsitteet ja määritelmät tulisi selittää, ja kielten luokittelutavat pitäisi dokumentoida raporteissa. Luokittelujen tulisi olla kattavia: niiden kuuluisi sisältää erillisten kielten lisäksi myös alueellisia murteita sekä keinotekoisia kieliä ja viittomakieliä. (Poulain 2008, 48.)

## Mihin kielten tilastointia tarvitaan?

Kielten tilastoinnilla on symbolista merkitystä. Valtio, joka tunnustaa kielten rinnakkaiselon alueellaan, ei yritäkään esittää olevansa yksikielinen. Jos kieliä ei rekisteröidä, se ei muuta sitä tosiasiaa, että kielivähemmistöjä on olemassa ja että niiden määrästä kaivataan usein tietoa. Kun väestön kielitaustaa ei tilastoida, vähemmistöjen

määrät perustuvat arvioihin, jotka voivat vaihdella huomattavasti. Esimerkiksi meänkieltä puhuvia arvioidaan olevan Ruotsissa 20 000–80 000 ja suomenkielisiä 70 000–450 000 (Azar & al. 2009). Kuten Jean Azar & al. (2009) tähdentävät kirjoituksessaan *Sverige behöver en språkstatistik som bygger på fakta – inte på gissningar*, on haasteellista arvioida, miten vähemmistöpoliittiset toimet toteutuvat, kun vähemmistökielen puhujien määrästä ei ole tietoa.

Myös Isossa-Britanniassa käytiin pitkään keskustelua siitä, että väestönlaskentaan tulisi sisällyttää kielitaustaan liittyviä kysymyksiä. Esimerkiksi vuoden 1991 laskennassa kieli-indikaattoriksi jäi arvio, joka perustui iän ja syntymävaltion yhdistelmään. Seuraavassa väestönlaskennassa 2001 kieleen liittyvät kysymykset ohitettiin edelleen asian perusteellisesta pohjustuksesta huolimatta. (Aspinall 2007, 92.) Tuoreimpaan eli vuoden 2011 väestönlaskentaan sitä vastoin sisällytettiin kaksi kysymystä: asukkailta kysyttiin ensinnäkin pääkieltä (englanti vs. muu, mikä) ja lisäksi englannin puhumistaitoa neliportaisella asteikolla (oikein hyvin, hyvin, ei kovin hyvin, ei lainkaan). Väestönlaskennan tuloksia aletaan julkaista syksystä 2012 lähtien.

Nykyistä tarkempi tieto asukkaiden monikielisydestä hyödyttää yhteiskuntaa monella tavalla. Kielivaranto on yhteiskunnalle pääomaa, sillä yhä useammalla alalla tarvitaan monikielitaitoisia työntekijöitä, kuten asiaa koskevat selvitykset (esim. CILT 2006; EK 2010) ovat osoittaneet:

Pelkkä englannin osaaminen ei kuitenkaan riitä, vaan myös maailman muiden kielten osaamiselle on kasvavaa kysyntää. Globaalissa kilpailussa menestymiseen tarvitaan monipuolista kielitaitoa. (EK 2010, 15.)

Tietoa asukkaiden kielitaustasta tarvitaan myös useista käytännön syistä (ks. esim. Aspinall 2007; Azar & al. 2009; Extra & Gorter 2008). Tieto on tarpeen muun muassa monikielistä tiedottamista ja erilaisten kielipalvelujen kehittämistä ja järjestämistä varten. Vaikka kotouttamispolitiikan yhtenä tavoitteena on saada muualta muuttaneet normaalipalvelujen piiriin mahdollisimman varhain, on selvää, että erityisesti maahanmuuton alkuvaiheessa tarvitaan käännös- ja tulkkauspalveluja sekä omakielistä neuvontaa.

Esimerkiksi terveyden- ja vanhustenhuollossa on merkittävästi hyötyä asiakkaiden kielitaustan tuntemisesta. Potilaan asemasta ja oikeuksista annetussa laissa (785/1992) määrätään, että poti-

<sup>4</sup> Pohjoismaiden ministerineuvoston (Nordiska Ministerrådet 2007, 41) mukaan Pohjoismaissa puhutaan pohjoismaisten kielten lisäksi noin kahtasataa muuta kieltä. Arvio perustune juuri koululaisilta kerättyihin tietoihin.

laan tulisi saada selvitys terveydentilastaan sekä hoito-ohjeet niin, että hän ymmärtää asian sisällön. Laissa todetaan myös, että jos terveydenhuollon ammattilainen ei osaa potilaan kieltä, on huolehdittava tulkkaamisesta. Jotta on mahdollista järjestää hoitotilanteissa tarvittava tulkkaus, on selvitettävä asiakkaan äidinkieli. On ilmeistä, että hoidon taso ja potilasturvallisuus kärsivät, jos potilas ja terveydenhuollon ammattilainen eivät ymmärrä toisiaan. Erityisesti mielenterveyspalveluja on mahdotonta kuvitella hoidettavaksi yksistään suomen kielellä (ks. Halla 2007). Maahanmuuttajaväestön ikääntyessä tulee välttämättömäksi huolehtia monikielisten palvelujen saatavuudesta myös vanhustenhuollossa.

Sosiaali- ja terveyspalvelujen mitoittaminen vastaamaan tulevaisuuden tarpeita edellyttää, että jo suunnitteluvaiheessa otetaan huomioon yhteiskunnan alati lisääntyvä monikielisuus. Asiakastyön tärkeimmän työvälineen, kielen, tulisi olla suunnittelun keskiössä. Sisäasiainministeriön (2010) teettämässä selvityksessä todettiin, että monikielisissä palveluissa on vielä runsaasti kehitettävää. Kyselyyn vastanneilla valtaväestön edustajilla oli selvästi maahanmuuttajia myönteisempi kuva palvelujen laajuudesta ja laadusta. Toisaalta saatiin myös näyttöä siitä, että omakielinen neuvonta on toiminut hyvin niillä seuduilla, joissa sitä on järjestetty. Tulos sai vahvistusta toisessa selvityksessä (Vainio 2012), jossa havaittiin, että omakielinen neuvonta auttaa maahanmuuttajia kotoutumaan ja hyödyttää samalla myös viranomaisia, koska heidän työaikansa ei kulu maahanmuuttajataustaisten asiakkaiden hoitamiseen samassa määrin kuin ilman omakielistä neuvontaa. Monikielisistä palveluista sekä monikielisen henkilöstön rekrytoinnista on hyötyä myös viranomaisille, sillä toimet täydentävät heidän voimavarojaan. Palvelujen ja tiedottamisen kehittämisessä aiempaa monikielisemmiksi ei siis ole kyse vain asiakkaiden edusta.

## Lopuksi

Kielten tilastoinnilla on Suomessa pitkät perinteet. Aiempiin luokituksiin verrattuna kansainvälisen järjestelmän käyttöönotto on tehnyt vuosittaisista kielitilastoista vertailukelpoisempia, koska standardin koodit ovat yksikäsitteisiä. Toisaalta vain pienelle osalle maailman kielistä on mää-

ritelty ISO 639-1 -koodi, ja siksi luokitus jättää tilaston ulkopuolelle lukuisia kieliä. Ulkopuolelle jäävät esimerkiksi Afganistanin toinen virallinen kieli dari ja Etelä-Sudanissa puhuttava dinka, joita kumpaakin opetetaan suomalaisessa koulussa omana äidinkielenä. Sitä vastoin standardi sisältää sellaisia kieliä, joiden äidinkielisiä puhujia ei Suomessa ole nyt eikä luultavasti edes lähitulevaisuudessa (kirkkoslaavi, volapük).

Koska luokitukseen on tehty serbokroatiata koskeva poikkeus, ei ole perusteltua, että Suomen lainsäädännön tunnustamat kielet (pohjoissaame, inarinsaame, koltansaame, romani, viittomakieli) jäävät vaille rekisteröintiä. Nykyisellään näiden kielten käyttäjät voivat siis merkitä kielen nimen henkilö tietoihinsa selväkielisenä, mutta koska kielille ei ole omaa koodia, niiden puhujat jäävät kategorian ”muu kieli” alle.<sup>5</sup> Poikkeusjärjestelyjen sijaan olisi kuitenkin mielekkäämpää ottaa käyttöön kansainvälisen standardin laajennettu versio, ISO 639-2 -koodisto, joka sisältää noin neljäsataa kieltä. Esimerkiksi saamen kielillä ja viittomakielillä on siinä omat koodit, ja laajennetun koodiston ansiosta rekisteröinnin piiriin pääsisivät myös lukuisat muut nykyään sen ulkopuolelle jäävät kielet.

Maahanmuuttoa ja väestön moninaisuutta koskevat käytössä olevat käsitteet ja mittarit ratkaisevat, mitä tietoja meillä on nyt ja millaisiin tietoihin väestöennusteet perustuvat. Vaikka kielten tilastointia onkin kehitetty, nykyisenlainen kielitilasto ei yksistään riitä. Nykyisten mittareiden riittämättömyyttä kuvastaa seuraava koulu maailman esimerkki. Vantaan kaupungin opetuspäällikkö Eero Väätäisen mukaan Vantaalla yhdistetään useita tietoja, jotta saataisiin kuvaa maahanmuuttajataustaisten oppilaiden kokonaismäärästä. Väestörekisterikeskuksen tiedot oppilaiden kielitaustasta yhdistetään S2-opetukseen ja oman äidinkielen opetukseen osallistuvien määriin, ja sillä tavoin opetustoimella on käytettävissään kielitilastoa osuvammin luonnehtiva tieto oppilaiden kokonaismäärästä. (Väätäinen 2011.)

Kuten Tuomas Martikainen (2010) on ehdotanut, tulisi maahanmuuttajien tilastollista tut-

<sup>5</sup> Muuta kieltä puhuvien määrä (4 388) tuntuu melko pieneltä, jos se sisältää myös viittomakieliset suomalaiset. Vaikuttaakin siltä, että niin viittomakieliset kuin myös romanit rekisteröivät tällä hetkellä itsensä suomen- tai ruotsinkielisiksi.



kimusta ja seuranta lisätä ja monipuolistaa. Ollisi hyvä pohtia, miten väestön kielellinen ja kulttuurinen moninaisuus saataisiin nykyistä paremmin näkyviin. Yksi mahdollisuus olisi yhdistää kahden sukupolven tietoja ja siten tavoittaa myös niin sanottu toinen sukupolvi, joka Suomessa syntyneenä, Suomen kansalaisena ja usein suomenkieliseksi rekisteröitynä muutoin katoaa tilastojen ulottumattomiin. Lisäksi voisi harkita vaihtoehtoja nykykäytännölle, jonka mukaan kukin asukas saa rekisteröidä vain yhden kielen. Suomessakin kysyttiin 1950-luvulla väestönlaskennassa paitsi äidinkieltä (tai pääkieltä) myös toisen kansalliskielen taitoa (Allardt & Starck 1981, 141). Muiden maiden nykyiset käytänteet (ks. Aspinnall 2007, 92) kertovat siitä, että asukkaiden monikielisyttä voi kartoittaa usealla eri tavalla.

Suuressa osassa Eurooppaa, muun muassa Suomessa, maahanmuutto on vielä nuorta, ja se näkyy esimerkiksi siinä, miten maahanmuuttajataustaista väestönosaa on pyritty identifioimaan. Nuorissa maahanmuuttomaissa tavallisimmin

## TIIVISTELMÄ

*Sirkku Latomaa: Kielitilasto maahanmuuttajien väestöosuuden mittarina*

Maahanmuuttoa ja väestön moninaisuutta koskevat käsitteet ja mittarit ratkaisevat, millaisiin tietoihin väestöennusteet perustuvat. Suuressa osassa Eurooppaa, muun muassa Suomessa, maahanmuutto on vielä nuorta, ja se näkyy esimerkiksi siinä, miten maahanmuuttajataustaista väestönosaa on pyritty identifioimaan. Nuorissa maahanmuuttomaissa tavallisimmin käytettyjä mittareita ovat kansalaisuus ja syntymämaa. Maissa, joissa maahanmuutolla on pitkät perinteet, käytetään niiden lisäksi luokittelukriteereinä muun muassa henkilöiden omaa identifikaatiota ja kielellistä elämää.

Artikkelissa pohditaan, miten tilastointia voisi kehittää siihen suuntaan, että kuva suomalaisen yhteiskunnan kulttuurisesta ja kielellisestä moninaisuudesta olisi nykyistä tarkempi. Esityksessä keskitytään erityisesti kielten tilastointiin, koska viime vuosina kielitilastoa on käytetty yhä useammin Suomessa maahanmuuttajien väestöosuuden mittarina.

Kielten tilastoinnilla on Suomessa pitkät perinteet, sillä tietoja väestön kielitaustasta on olemassa jo vuodesta 1749. Vuonna 1999 Väestörekisterikeskus siirtyi kielikoodauksessa kansainväliseen ISO 639-1 -standardiin. Kansainvälisen järjestelmän käyttöönotto on tehnyt vuosittaisista kielitilastoista aiempaa vertailukelpoisempia, koska standardin koodit ovat yksikäsit-

käytettyjä kriteerejä ovat kansalaisuus ja syntymämaa. Maissa, joissa maahanmuutolla on pitkät perinteet, käytetään niiden lisäksi luokittelukriteereinä muun muassa henkilöiden omaa identifikaatiota ja kielellistä elämää. Esimerkiksi Kanadassa kysytään asukkaiden kielellisestä repertoaarista monella eri kysymyksellä: äidinkieli ("the language first learnt at home and still understood"), kotona puhuttu kieli, työssä puhuttu kieli sekä englannin kielen taito. (Extra & Gorter 2008, 21–23.) Ison-Britannian tuoreimmassa väestönlaskennassa oli puolestaan mahdollista antaa yhden etnisen luonnehdinnan sijaan useita (*multiple ethnicities*). Nämä kysymisen tavat kertovat siitä, että yhteiskunnan monimuotoisuudessa yksiuulotteiset dikotomiat eivät enää toimi: joko-tain tilalle on tullut sekä-että. Totta on kuitenkin, etteivät mitkään kategorisoinnit voi täydellisesti tavoittaa sitä monimuotoisuutta, joka yhteiskunnissa nykyään vallitsee. Hybridien identiteettien ja kielimuotojen todellisuutta vasten kategoriat näyttävät aina jäykiltä.

teisiä. Toisaalta vain pienelle osalle maailman kielistä on määritelty standardissa koodi. Siksi luokitus jättää tilaston ulkopuolelle lukuisia Suomessa käytössä olevia kieliä, muun muassa pohjoissaamen, inarinsaamen, koltansaamen, romanin ja viittomakielen.

Viime aikoina on keskusteltu paljon kieliperusteisen tilastoinnin ongelmallisuudesta. Kielitilastoa on pidetty lähtökohtaisesti epäluotettavana, koska kielen rekisteröinti perustuu henkilön omaan ilmoitukseen ja sitä voi muuttaa kuten kansalaisuutta. Yksi mahdollisuus saada väestön kielellinen moninaisuus nykyistä paremmin näkyviin olisi yhdistää kahden sukupolven tietoja ja siten tavoittaa myös niin sanottu toinen sukupolvi, joka Suomessa syntyneenä, Suomen kansalaisena ja usein suomenkieliseksi rekisteröitynä muutoin katoaa tilastojen ulottumattomiin. Lisäksi voisi harkita vaihtoehtoja nykykäytännölle, jonka mukaan kukin asukas saa rekisteröidä vain yhden kielen. Muiden maiden käytänteet todistavat siitä, että asukkaiden monikielisyttä voi kartoittaa usealla eri tavalla.

Perusteita kielitiedon tarpeellisuudelle on lukuisia. Tieto on tarpeen muun muassa monikielistä tiedottamista ja erilaisten kielipalvelujen kehittämistä ja järjestämistä varten. Esimerkiksi sosiaali- ja terveyspalvelujen mitoittaminen tulevaisuuden tarpeita vastaavaksi edellyttää, että jo suunnitteluvaiheessa otetaan huomioon yhteiskunnan lisääntyvä monikielisyys.

## KIRJALLISUUS

- Allardt, Erik & Starck, Christian: Språkgränser och samhällsstruktur. Finlandssvenskarna i ett jämförande perspektiv. Stockholm: Almqvist & Wiksell, 1981.
- Anhava, Jaakko: Maailman kielet ja kielikunnat. Helsinki: Gaudeamus, 2002.
- Aspinall, Peter J.: Language ability: A neglected dimension in the profiling of populations and health service users. *Health Education Journal* 66: 1 (2007), 90–106.
- Azar, Jean & Lasko, Lars-Nila & Lindblom, Per-Åke & Sener, Refik & Voionmaa, Kaarlo: Sverige behöver en språkstatistik som bygger på fakta – inte på gissningar. *Tidningen Kulturen* 13.8.2009.
- CILT: ELAN: Effects on the European Economy of Shortages of Foreign Language Skills in Enterprise. London: The National Centre for Languages, 2006.
- EK: Työelämässä tarvitaan yhä useampia kieliä. EK:n henkilöstö- ja koulutustiedustelu 2009. Helsinki: Elinkeinoelämän keskusliitto, 2010.
- Extra, Guus & Gorter, Durk: The Constellation of Languages: An Inclusive Approach. In Extra, Guus & Gorter, Durk (eds.) *Multilingual Europe: Facts and Policies*. Berlin, New York: Mouton de Gruyter, 2008.
- Halla, Tapio: Psykkisesti sairas maahanmuuttaja. *Duodecim* 123 (2007), 469–475.
- Helsingin kaupungin tietokeskus: Vieraskielisen väestön ennuste Helsingin seudulla 2010–2030. Helsinki: Helsingin kaupungin tietokeskus, 2010.
- Karlsson, Fred: Yleinen kielitiede. Helsinki: Yliopistopaino, 1994.
- Kartovaara, Leena: Kerro, kerro tilasto, ken on maahanmuuttajalapsi! *Hyvinvointikatsaus* 1/2007, 17–23.
- Kumpulainen, Timo (toim.): Koulutuksen määrälliset indikaattorit 2010. Koulutuksen seurantaraportit 2010:4. Helsinki: Opetushallitus, 2010.
- Laki potilaan asemasta ja oikeuksista 785/1992.
- Lane, Pia: The birth of the Kven language in Norway: emancipation through state recognition. *International Journal of the Sociology of Language* 209 (2011): 57–74.
- Latomaa, Sirkku: Tataarista teluguun. Tampereen monet kielet. Teoksessa: Lönnroth, Harry (toim.) *Tampere kieliyhteisönä*. Helsinki: Suomalaisen Kirjallisuuden Seura, 2009.
- Leitzinger, Antero: Ulkomaalaiset Suomessa 1812–1972. Helsinki: East-West Books, 2008.
- Martikainen, Tuomas: Maahanmuuton tilastointi puutteellista. *ETMU-blogi* 17.6.2010.
- McPake, Joanna & Tinsley, Teresa & Broeder, Peter & Mijares, Laura & Latomaa, Sirkku & Martyniuk, Waldemar: *Valuing All Languages in Europe*. ECML Research and Development reports series. Graz: ECML, 2007.
- McRae, Kenneth D.: *Conflict and Compromise in Multilingual Societies: Finland*. The Finnish Academy of Science and Letters & Wilfrid Laurier University Press, 1999.
- Nikander, Timo: sähköpostiviesti 20.12.2010.
- Nordiska ministerrådet: *Deklaration om nordisk språkpolitik*. Köpenhamn: Nordiska ministerrådet, 2007.
- OM: Valtioneuvoston kertomus kielilain soveltamisesta 2009. Helsinki: Oikeusministeriö, 2009.
- OKM: Koulutus ja tutkimus vuosina 2011–2016. *Kehtämässuunnitelma*. Helsinki: Opetus- ja kulttuuriministeriö, 2011.
- Poulain, Michel: European migration statistics. Definitions, data and challenges. In Barni, Monica & Extra, Guus (eds.) *Mapping Linguistic Diversity in Multicultural Contexts*. Berlin, New York: Mouton de Gruyter, 2008.
- Rapo, Markus: sähköpostiviesti 29.12.2010.
- Rapo, Markus: Vieraskielinen on yhä monessa kunnassa harvinaisuus. Tilastokeskus, 2011 [http://tilastokeskus.fi/tup/vl2010/art\\_2011-06-21\\_001.html](http://tilastokeskus.fi/tup/vl2010/art_2011-06-21_001.html) (luettu 15.4.2012).
- Risager, Karin: Hvilke sprog tales der i Danmark? *Sprogforum* 36 (2006), 13–14.
- Rumbaut, Ruben G.: Ages, Life Stages, and Generational Cohorts: Decomposing the Immigrant First and Second Generations in the United States. *International Migration Review* Vol. 38 (2004): 1160–1205.
- Saari, Matti: sähköpostiviesti 16.8.2011.
- Saukkonen, Pasi: Kotouttaminen ja kulttuuripolitiikka. Tutkimus maahanmuutosta ja monikulttuurisuudesta suomalaisella taiteen ja kulttuurin kentällä. Helsinki: Cupore, 2010.
- SFS: Kielten nimien tunnukset. Osa 1: kaksikirjaimiset tunnukset. Helsinki: Suomen Standardoimisliitto, 2008.
- Simon, Patrick: “Ethnic” statistics and data protection in the Council of Europe countries. Study Report. Strasbourg: Council of Europe, 2007.
- Sisäasiainministeriö: Maahanmuuttajien omakieliset palvelut. Selvityksen loppuraportti. Helsinki: Sisäasiainministeriö, 2010.
- Skutnabb-Kangas, Tove: *Tvåspråkighet*. Lund: Liber-Läromedel, 1981.
- SVT, Suomen virallinen tilasto. Väestörakenne. Tilastokeskus, 2012 <http://www.stat.fi/til/vaerak/index.html> (luettu 16.3.2012).
- Söderholm, Eira: Kieli, murre, varieteetti vai oma kieli? Pohjoiskalotin kielet ja kielimuodot. *Virttäjä* 102 (1998): 571–592.
- Tilastokeskus: Kulttuuritilasto 2007. Helsinki: Tilastokeskus, 2009.
- Vainio, Sirpa: Yli kielimuurin – omakielinen neuvonta osana julkisen sektorin palvelua, case Infotori. Liiketalouden koulutusohjelman opinnäytetyö. Satakunnan ammattikorkeakoulu, 2012.
- Vätäinen, Eero: henkilökohtainen tiedonanto 22.10.2011.
- Winsa, Birger: Meän kieli ja torniolaaksolaisitten kaksikielisyys: täälä plandathan sprookit. *Virttäjä* 97 (1993): 3–33.