

# Maailman kielellinen diversiteetti ja sen tutkimus

MATTI MIESTAMO

## Professorin juhlaluento 3. joulukuuta 2014 Helsingin yliopistossa

Kieli on ajattelun, ihmistenvälisen kanssakäymisen ja yhteiskunnallisen toiminnan keskeinen elementti. Ymmärtääksemme ajattelun prosesseja ja ihmisen toimintaa sekä voidaksemme ratkoa niihin liittyviä ongelmia tarvitsemme tietoa ja ymmärrystä kielestä ja kielistä. Tarvitsemme kielitieteellistä tutkimustietoa.

Yleinen kielitiede on nimensä mukaisesti yleistä, ei mihinkään yksittäiseen kieleen erityisesti liittyvää kielentutkimusta. Sitä voi pitää sateenvarjotieteenä, jonka piiriin kuuluu periaatteessa mikä tahansa kielentutkimus, olipa kyse kielen rakenteesta tai käytöstä, kielihistoriasta, kieleen liittyvistä mentaalisista prosesseista tai kielen suhteesta kulttuuriin ja yhteiskunnallisiin ilmiöihin. Yleisen kielitieteen kysymyksenasettelut koskevat kieltä ja kieliä yleensä, ei mitään yksittäistä kieltä. Ytimeltään se on kielitieteen teoria- ja menetelmäoppia. Siinä siis kehitetään kieltä koskevia teorioita ja metodeita: teorioita, jotka auttavat meitä ymmärtämään ihmiskielen olemusta, rakennetta, toimintaa ja käyttöä, ja metodeita, joita voidaan soveltaa kielestä riippumatta minkä tahansa kielen tutkimukseen. Yleisen kielitieteen kiinnostuksen kohteena on maailman kielten koko diversiteetti, ja sen on näin pystyttävä teorianmuodostukseensa kattamaan tämä monimuotoisuus kokonaisuudessaan. Pyrkinessään kieltä ja kieliä koskeviin yleistyksiin sen on siis nojattava laajaan kieltenväliseen empiiriseen taustaan.

Arviot maailmassa puhuttavien kielten määrästä liikkuvat nykyisin 7 000:n paremmalla puolella. Kielten määrää on toki vaikea arvioida, sillä edelleen on alueita, joiden kieliolot tunnetaan puutteellisesti, eikä kielen ja murteen välinen rajanveto ole yksiselitteinen: milloin meidän tulisi laskea kaksi kielimuotoa saman kielen eri murteiksi ja milloin voimme laskea ne kahdeksi erilliseksi kieleksi? Joka tapauksessa eri kieliä on tuhansia, ja kun jokaisen kielen sisällä on vaihteleva määrä alueellisia murteita sekä eri sosiaalisissa tilanteissa ja sosiaalisten ryhmien sisällä käytettäviä kielimuotoja, sosiolektejä, ei maailman kielellistä monimuotoisuutta voine pitää ainakaan vähäisenä. Tämä diversiteetti on luonnollisesti jatkuvassa muutoksessa – kielet muuttuvat, kieli-

muotoja syntyy ja katoaa. On kuitenkin selvää, että näistä jälkimmäistä – katoamista – tapahtuu meidän ajassamme paljon syntymistä enemmän. Maailman kielistä merkittävä osa on viime vuosisatojen ja -kymmenten aikana kadonnut, ja vielä suurempi määrä on vaarassa kadota. Tämän kehityksen taustalla on ensisijaisesti eurooppalaisesta kolonisaatiosta alkanut globalisaatiokehitys ja siihen liittyvät taloudelliset mullistukset, erityisesti kaupungistuminen, joka johtaa perinteisten yhteisöjen hajoamiseen. On esitetty arvioita, että seuraavien sukupolvien aikana jopa puolet maailman kielistä katoaisi ja sadan vuoden päästä vain kymmenen prosenttia nyt puhutuista kielistä olisi jäljellä.

Yleisen kielitieteen osa-aluetta, joka pyrkii ottamaan aineistokseen koko maailmanlaajuisen kielellisen diversiteetin, kutsutaan kielitypologiaksi. Kielitypologia on teoreettista kielitiedettä, jonka empiirisenä aineistona toimii maailmanlaajuinen kieltenvälinen vertailu. Kieliä laajasti vertailemalla kielitypologit pyrkivät kohti ymmärrystä ihmiskielen luonteesta: Millaiset rakenteet ja ilmiöt ovat mahdollisia kielessä ja millaisia ei esiinny? Mikä on yleistä ja mikä harvinaista? Miten havaitut kielenpiirteet jakautuvat eri kielikuntiin ja maantieteellisille alueille? Näiden kaikkien kysymysten kohdalla pyrimme luonnollisesti vastaamaan myös miksi-kysymyksiin: miksi tietyt rakenteet ovat mahdollisia tai eivät esiinny, miksi jotkin kielenpiirteet ovat yleisiä ja toiset eivät ja niin edelleen? Kielitypologisesti orientoituneelle yleislingvistille on itsestään selvää, että aitojen yleiskielitieteellisten yleistysten ja kaikista kielistä pätevien teorioiden on pohjauduttava mahdollisimman laajaan kielten vertailuun.

Matthew Dryerin (2013) laatima kartta kertoo, millaisia eri tapoja muodostaa kyllä/ei-kysymyslauseita maailman kielissä esiintyy. Tavallisin tapa on kysymyspartikkeli, jollaiseksi myös suomen *-ko/-kö* (esim. *tulenko, syönkö*) voidaan luokitella. Tätä tyyppiä on kartalla merkitty sinisellä. Karttaa katsottaessa huomataan, että keltaista väriä esiintyy lähes yksinomaan läntisessä Euroopassa. Keltainen symboloi kieliä, joissa kysymys tyyppillisesti muodostetaan vain sanajärjestyksestä muuttamalla: esimerkiksi ruotsin *hunden springer iväg / springer hunden iväg?* Tämä on esimerkki siitä, kuinka jokin tuttujen eurooppalaisten kielten näkökulmasta yleisenä pidettävä kielenpiirre onkin globaalissa tarkastelussa harvinainen ja eksoottinen. Esimerkki korostaa tarvetta perustaa kaikkia kieliä koskevat yleistykset ja yleiskielitieteelliset teoriat laajaan kieltenväliseen vertailuun.

Kielten laaja vertailu edellyttää tietysti, että tutkimustietoa maailman kielistä on saatavilla. Mitä suurempi määrä kieliä on dokumentoitu, sitä enemmän kielitypologisella tutkimuksella on aineistoa käytettävissään. Maailman yli 7 000 kielestä valitettavasti vain pieni osa on kuvattu edes tyydyttävällä perustasolla. Parin viime vuosikymmenen aikana kielitieteilijät ovat heränneet siihen, että valtaosa maailman kielellisestä diversiteetistä on dokumentoimatta, ja nykyisin kielten dokumentaatiotyötä tehdään ilahduttavan paljon eri puolilla maailmaa. Tekijöitä ja rahoitusta ei kuitenkaan ole lähimainkaan tarpeeksi, jotta edes kaikkein uhanalaisimmat katoamassa olevat kielet saataisiin kuvatuksi ja toivottavasti myös säilytetyksi. Uhanalaisten kielten dokumentoinnilla on kiire paitsi puhujayhteisöjen itsensä ja kielten revitalisaatiopyrkimysten kannalta myös, mikäli haluamme varmistaa riittävän laajan empiirisen perustan vertailevalle kielitieteelle ja sitä kautta yleiselle kielitieteelle ja koko ymmärryksellemme ihmiskielen luonteesta.

Kielitypologia ja kielten dokumentaatiota harrastava deskriptiivinen lingvistiikka tarvitsevat toisiaan. Kielitypologit ovat riippuvaisia yksittäiskielten tutkimuksesta, josta he

saavat aineistoa omaan vertailevaan tutkimukseensa. Toisaalta yksittäiskielten kuvausta tekevät tutkijat tarvitsevat kielitypologista tietoa ymmärtääkseen tutkimiensa kielten ilmiöitä paremmin. Voidaan sanoa, että kielitypologia ja deskriptiivinen lingvistiikka elävät hedelmällisessä symbioosissa. Ne muodostavat yhteisen pyrkimyksen dokumentoida, kuvata, kartoittaa ja ymmärtää maailman kielellistä diversiteettiä. Saksalainen kielitieteilijä Martin Haspelmath (2012) on ottanut käyttöön termin *diversiteettilingvistiikka* kuvatakseen näiden kahden tutkimussuuntauksen muodostamaa kokonaisuutta.

Kielitypologia on siis maailmanlaajuisia vertailevaa kielitiedettä. Kun verrataan kymmeniä tai satoja kieliä keskenään, tutkija ei toki voi itse osata kaikkia vertailemiaan kieliä. Perinteisesti typologien aineistona ovat olleet yksittäiskielten tutkijoiden tuottamat deskriptiiviset materiaalit, kuten kieliopit, sanakirjat ja muut julkaistut tai julkaisemattomat kielenkuvaukset. Olen itse tutkinut esimerkiksi kieltolauseiden rakennetta maailman kielissä nimenomaan tällaiselta aineistopohjalta. Viime aikoina typologiassa on ruvettu käyttämään yhä enemmän esimerkiksi erilaisia kyselytutkimusmenetelmiä ja jopa laajoja tekstiaineistoja eli korpuksia. Joitakin tekstejä, lähinnä Raamatun osia, on saatavilla sadoille eri kielille käännettyinä, ja tällaisia massiivisia rinnakkaiskorpuksia on viime aikojen teknisen kehityksen myötä opittu hyödyntämään kielitypologiassa. On kuitenkin muistettava, että vertailtavien kielten määrä asettaa omat rajoituksensa korpus- ja kyselytutkimusmetodien käytölle ja niillä lähestyttävillä tutkimuskysymyksille. Joka tapauksessa uusien metodien ja aineistojen käyttöönotto laajentaa kielitypologian näköaloja. Typologisen tutkimuksen aihealueita ovat perinteisesti olleet kieliopin ydinalueet: peruslauseen rakenne, lauseenjäsenet, tempus, aspekti ja modus, ja vaikkapa negaatio. Myös äänneopin – fonetiikan ja fonologian – kysymyksiä on käsitelty laajasti vertailevasta perspektiivistä. Tämä on ymmärrettävää, sillä nämä ovat aihealueita, joista on mahdollista saada tietoa kielten peruskuvauksista. Syvällisempää semanttista analyysia vaativat aihepiirit tai laajan diskurssiaineiston läpikäymistä edellyttävät tutkimuskysymykset ovat olleet metodologisista syistä harvemmin typologien tutkimuskohteena. Viime aikoina on kuitenkin yhä enemmän pyritty laajentamaan aihepiirejä, joista typologista tutkimusta voidaan tehdä. Leksikaalinen typologia tutkii kielten sanaston rakennetta: miten eri kielet kielentävät eri semanttisia alueita, kuten vaikkapa liikkeen ilmaisemista tai lämpötilaa? Myös mahdollisuuksia diskurssi-ilmiöiden laajaan kieltenväliseen tutkimukseen on alettu kartoittaa.

Kielten dokumentaatiotyötä tehdään kenttälingvistin metodein. Elisitaatioissa puhujilta kerätään systemaattisesti kielen sanoja, sanojen muotoja ja lauseita. Aitoa kieliainesta saadaan nauhoittamalla tai videoimalla esimerkiksi tarinoita ja dialogeja, joita sitten analysoidaan: litteroidaan, käännetään ja analysoidaan kieliopillisesti. Dokumentaatiotyön tavoitteena on toisaalta kerätä kieliainesta ja tuoda se kieliyhteisön itsensä sekä tutkijoiden saataville ja toisaalta tuottaa tämän dokumentaatioaineiston pohjalta kielen kuvausta, esimerkiksi kielioppi ja sanakirja. Toisinaan dokumentaatio ja deskriptio halutaan erottaa toisistaan siten, että dokumentaatiolla tarkoitetaan nimenomaan aineiston keräämistä ja käyttökelpoiseen muotoon saattamista ja deskriptiolla tuota toista vaihetta eli analyysia ja kieliopillista kuvausta. Helsingin yliopistossa kielten dokumentaatiotyö kytkeytyy kansainvälisesti merkittävään, jo 1800-luvun alkupuolella syntyneeseen etnolingvistiseen kenttätöperinteeseen, jonka uranuurtaja oli M. A. Castrén.

Dokumentaatio tuottaa dokumentoituja kielimuotoja, joita voidaan kutsua termillä *dokulekti* – samaan tapaan kun voimme puhua esimerkiksi dialektista, sosiolektista tai idiolektista. Esimerkiksi jossakin deskriptiivisessä kieliopissa, sanakirjassa tai korpuksessa esitetty kielimuoto on yksi dokulekti. Kun kielitypologi vertailee kieliä, hän itse asiassa vertailee dokulekteja – asiaankuuluvaa lähdekritiikkiä harjoittaen – ja tekee niiden pohjalta yleistyksiä kielestä ja kielistä. Keskeisiä metodologisia kysymyksiä on toisaalta kielten välinen vertailtavuus – miten tunnistaa sama ilmiö hyvin erilaisissa kielissä – ja toisaalta vertailtavien kielten valinta kielitokseen – yleisenä pyrkimyksenä on varmistaa maailman eri kieliryhmien ja maantieteellisten alueiden tasapuolinen edustus otoksessa.

Yksi suhteellisen vähän tutkittu kieli, josta tarvitaan dokumentaatioaineistoa, on Suomen Inarissa puhuttu koltansaame. Sitä puhuu äidinkielenään noin 300 ihmistä, joista valtaosa kuuluu jo iäkkäämpiin sukupolviin. Joitakin poikkeuksia lukuun ottamatta kieli ei ole siirtynyt lapsille viime vuosikymmeninä, ja äidinkielenään kolttaa puhuu vain muutama lapsi. Kieltä voidaan siis pitää vakavasti uhanalaisena. Viime vuosina kuitenkin koltansaamelaisessa yhteisössä on virinnyt uudella tavalla halua oman kielen säilyttämiseen ja on perustettu esimerkiksi kielipesiä, joissa pienet lapset oppivat kieltä isovanhempiensa sukupolvelta luonnollisessa kanssakäymisessä leikinomaisissa olosuhteissa. Kielipesien rahoitus on ollut uhattuna, mutta siihen on onneksi löytymässä ratkaisu. Koltansaamen dokumentaatiohankkeessa tuotamme dokumentaatioaineistoa, joka tallennetaan Suomen Kielipankin palvelimelle, ja se on sitä kautta sekä koltankielisen kieliyhteisön että tiedeyhteisön käytössä. Saamelaiskielten tutkijoiden ja kielitypologien lisäksi aineisto hyödyttää myös kielenpuhujia itseään kielen revitalisaatio-yrkimyksissä tarjoamalla välineitä kielinopetuksen ja kielenhuollon tarpeisiin.

Viime aikoina on puhuttu paljon humanististen tieteiden digitaalisesta vallankumouksesta, digitaalisesta humanismista. Kielitieteet ovat olleet eturintamassa digitaalisten aineistojen ja menetelmien kehityksessä, ja digitalisaatiokehitys on luonut uusia mahdollisuuksia myös diversiteettilingvistiikalle. Erilaiset kielten dokumentaation tarpeisiin kehitetyt tietokoneohjelmistot ja työkalut ovat mullistaneet alan metodologiaa ja helpottaneet kielten kuvausta tekevän lingvistin työskentelyä. Erityisen merkittävää digitalisaatio on aineistojen saatavuuden ja käytettävyyden näkökulmasta. Kielten dokumentaatioaineistoja voidaan nykyisin tallettaa sähköisiin arkistoihin, joissa ne ovat aivan uudella tavalla sekä kielen puhujien että tutkijayhteisön saatavilla. Kun digitaalisessa korpuksessa ääniraitaan ja litteraatioon yhdistetään myös käänös ja kieliopillinen analyysi, arkistoissa olevat digitaaliset aineistot ovat myös laajemmin vertailevaa tutkimusta tekevien lingvistien, kielitypologien, käytettävissä. Kielitypologia hyötyy digitalisaatiosta toki muutenkin. Typologiset tietokannat tuovat typologisen datan laajalti tutkijoiden ulottuville ja mahdollistavat kielen piirteiden vertailun ja korreloinnin aiempaa tehokkaammin. Karttasovellukset tuovat uusia metodologisia mahdollisuuksia kielenpiirteiden maantieteellisten jakaumien tutkimiselle ja areaalitypologian kehitykselle. Kehittyneet tilastolliset menetelmät ovat vieneet kvantitatiivisen kielitypologian uudelle tasolle.

Moderni kielitypologia, joka alkoi amerikkalaisen Joseph Greenbergin tutkimuksista 1950–1960-luvuilla, oli alun perin kiinnostunut kielen universaaleista: siitä, mitä kielten laaja vertailu voi kertoa kielten yhteisistä piirteistä ja kieltenvälisen variaation rajoista. Se oli siis metodologisesti (ja teoreettisesti) vastakkainen lähestymistapa

chomskylaiselle universaalikieliopin tutkimukselle, joka yritti löytää kielen universaaleja yhden kielen yksityiskohtaisen analyysin pohjalta. Kielen universaalit ja kieltenvälisen variaation rajat ovat edelleen kielitypologien keskeinen tutkimuskysymys. Kielitypologit toimivat yleensä funktionalistisessa viitekehyksessä, jossa kielen funkioon liittyvät kielenulkoiset seikat nähdään selittävinä tekijöinä kielen ilmiöiden taustalla. Selityksiä universaaleille ja muille kieltenvälisille yleistyksille etsitään siis semantiikasta, pragmatiikasta, kielen prosessoinnista ja laajemmin kognitiivis-psykologisista ilmiöistä sekä sosiokulttuurisista tekijöistä.

Viime vuosina kielitypologiassa on kiinnostuttu yhä enemmän myös kielenpiirteiden maantieteellisestä jakaumasta ja niihin vaikuttavista historiallisista prosesseista. Sveitsiläisen kielitieteilijän Balthasar Bickelin (2007: 239) sanoin, sen sijaan että kysyttäisiin ”mikä on mahdollista?”, yhä useammin kysytään ”mitä on missä ja miksi?”. Areealigeneologisen näkökulman myötä myös kvantitatiivisten menetelmien rooli korostuu. Laajan kieltenvälisen variaatiodatan korreloiminen maantieteeseen, kielihistoriaan ja vaikkapa kulttuuriin ilmiöihin edellyttää myös laskennallisilta malleilta aiempaa enemmän. Typologisessa tutkimuksessa kieli nähdään nyt yhä enemmän kontekstissaan, joka on yhtä aikaa sociolingvistinen, maantieteellinen, biologinen, kulttuuris-yhteiskunnallinen ja historiallinen. Tällaista kokonaisvaltaista ympäristön huomioivaa lähestymistapaa kieleen kutsutaan ekolingvistiikaksi eli kielellisen ekologian tutkimukseksi. Tämä lähestymistapa avaa myös kiinnostavia tieteidenvälisiä näkökulmia kielen, kulttuurin, yhteiskunnan ja historian tutkimukseen.

Palaan vielä alussa esiin nostamaani kielitieteellisen tutkimustiedon tarpeeseen nimenomaan diversiteettilingvistiikan näkökulmasta. Diversiteettilingvistiikan yhteiskunnallinen relevanssi tulee suorimmin esiin sen tuen muodossa, jonka kielen dokumentaatiotyö voi antaa kielyhteisöjen pyrkimyksille säilyttää ja revitalisoida kieltään. Perustutkimuksen näkökulmasta voidaan yleisellä tasolla sanoa, että pyrkiessään kaikkia kieliä koskeviin yleistyksiin diversiteettilingvistiikka ja yleinen kielitiede laajemminkin on tärkeä osa sitä monimutkaista palapeliä, jota tieteentekijät kokoavat ymmärtääkseen paitsi kieltä ja kieliä myös ihmismieltä ja ihmisen toimintaa.

## Lähteet

- BICKEL, BALTHASAR 2007: Typology in the 21<sup>st</sup> century. Major current developments. – *Linguistic Typology* 11 s. 239-251.
- DRYER, MATTHEW 2013: Polar questions. – Matthew S. Dryer & Martin Haspelmath (toim.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/116>.
- HASPELMATH, MARTIN 2012: Diversity linguistics comment. – Kirjoitus blogissa *Diversity Linguistics Comment* 1. <http://dlc.hypotheses.org/1>.

Kirjoittajan yhteystiedot:  
etunimi.sukunimi@helsinki.fi