

Tieteen matkamiehen siivellä – Lauri Kettusen *Suomen murteet* tutkimusaineistona ennen ja nyt

Outi Vesakoski, Lotta Aarikka, Jenni Santaharju

Tieteen matkamiehen taival on ollut pitkä, raskas, yksinäinen – mutta tuloksellinen. Sen vuoksi hän tiesi, mistä puhui, kun hän sanoi, että niitä matkoja tekemään ei ole toista. Hän tiesi oman arvonsa myös suomen murteiden tutkijana. Erityisen suuressa arvossa hän piti murrekartastoaan „Suomen murteet. III A. Murrekartasto“. Teoksensa hän esitteli yleensä myös luennoillaan, mutta muistan myös, miten hän paitahihassillaan ryhtiään kohentaen sanoi huvilansa olohuoneen keskilattialla viitaten tähän työhönsä: „Vielä kerran he antavat minulle arvoa.“

Lauri Kettusen oppilas Sirkka Pietikäinen (1991: 81)

Tässä artikkelissa käsittelemme Lauri Kettusen *Suomen murteet* -teossarjan syntyhistoriaa (luku 1) ja teossarjan neljättä osaa *III. B. Murrekartasto*. Luvussa 2 esittelemme kartaston digitointia tutkimusaineistoksi (luku 2) ja digitaalisen aineiston tarjoamia tutkimusmahdollisuuksia (luku 3). Lopuksi pohdimme kartaston arvoa ja siihen kohdistuvan tutkimuksen tulevaisuutta (luku 4).

1. Lauri Kettunen ja *Suomen murteet* -sarjan syntyhistoria

Lauri Kettunen (10.9.1885–26.2.1963) oli savolaislähtöinen suomen ja sen sukukielten tuottelijan tutkija, joka toimi mm. Tarton yliopiston itämerensuomalaisen kielten professorina (1919–1925) ja Helsingin yliopistossa ensin itämerensuomalaisen kielten professorina (1929–1929) sekä myöhemmin viron ja sen lähisukukielten professorina (1938–1953) (Suhonen 1985: 253).

Lauri Kettunen oli kiistatta yksi keskeisimmistä hahmoista suomen ja sen sukukielten tutkimusalalla 1900-luvun ensimmäisellä puoliskolla. Kettunen teki kenttätöitä käytännössä kaikkien suomen sukukielten kieliyhteisöissä – myös Virossa – ja tutki pitkän ja tuotteliaan uransa aikana useita itämeren-suomalaisia kieliä (Suhonen 1985). Kettusen kenttätöistään kirjoittamat laajamittaiset muistelmat (Kettunen 1945, 1948, 1960a, 1960b) ovat täynnä kiinnostavia tarinoita hänen monilta tutkimusmatkoiltaan.

1920-luvulla Kettunen tarttui myös suomen murteiden tutkimukseen. Työn tuloksena ilmestyi *Suomen murteet* -niminen teossarja, joka koostuu neljästä osasta: Ensimmäinen (*I. Murrenäytteitä*, 1930a) on kielennäyte-kokoelma, joka sisältää fonografi- ja parlografitallenteilta¹ tekstimuotoon litteroituja katkelmia Kettusen kenttäretkillään puhuttamien kieltenoppaiden puheesta. Murrenäytteet on järjestetty samana vuonna ilmestyneessä toisessa osassa (*II. Murrealueet*, 1930b) esiteltyjen murrealueiden mukaan ja näytteitä on noin 170 silloisesta pitäjistä (Räisänen 1991: 54). Teossarjan kolmas osa (*III. A. Murrekartasto*, 1940a) sisältää 213 kielenpiirteen varianttien levikkejä kuvaavaa karttaa. Teossarjan neljäs osa (*III. B. Selityksiä murrekartastoon*, 1940b) on kartastoa täydentävä selitysosa, jossa Kettunen selittää kartoissa kuvaamiaan ilmiöitä ja niiden kielihistoriallista taustaa sekä tekee huomiota myös varianttien rinnakkaisedustuksesta. Toisin sanoen *Suomen murteet* -teossarjan osien välillä on vuorovaikutuksellinen suhde: *Murrenäytteet* (1930a) sisältävät Kettusen aineiston siltä osin, mitä hän on tallentanut parlografiilla ja fonografiilla. Aineiston perusteella Kettunen on hahmotellut eri kielenpiirteistä ja niiden varianteista levikkikarttoja (keruussa n. 400 kpl, julkaistussa kartastossa 213 kpl) ja tuottanut näiden synteessinä esityksen suomen murteiden aluejaosta (1930b).

Kettusen *Suomen murteet* -teossarjalla, joka on ensimmäinen ja toistaiseksi ainoa suomen murteita käsittelevä murreatlas, on kansainvälisten tutkimussuuntauksien virtauksiin asettuva kontekstinsa. 1800-luvun puoliväliin saakka kielentutkijoiden tekemät murrejaot perustuivat pääasiassa intuitioon (Aarikka 2023: 281–283, Chambers, Trudgill 2009: 16). 1800-luvun loppupuolella sai alkunsa äännehistoriallinen (myös *nuorgrammaattinen*, *neogrammaattinen* ja *junggrammaattinen*) tutkimusperinne, joka vaikutti myös

¹ Lauri Kettusen *Suomen murteet* -teossarjaan liittyvät fonografi- ja parlografitallenteet eli vahaliერიöt ovat säilytteillä Kotimaisten kielten keskuksessa ja niistä toistokelpoiset on digitoitu.

murteiden tutkimukseen (Chambers, Trudgill 2009: 16–17). Suomen kontekstissa tämä merkitsi keskittymistä pääasiassa äänne- ja muoto-opillisiin kielenilmiöihin sekä äänne- ja muoto-opillisiin kielenilmiöihin sekä äänne- ja muoto-opillisiin kielenilmiöihin sekä äänne- ja muoto-opillisiin kielenilmiöihin (Hovdhaugen ym. 2000: 422, Kelomäki 2009, Aarikka 2023: 212–227). Äännehistoriallisen tutkimuksen piirissä esitetty äännelakiteoria² johti puolestaan murremaantieteelliseen tutkimukseen, jonka tavoitteena oli tarkastella empiirisesti ja systemaattisesti murteiden välisiä eroja (Chambers, Trudgill 2009: 17). Murteiden välisten erojen systemaattinen tarkastelu edellytti laajamittaisten kenttätöiden tekemistä ympäri Eurooppaa: esimerkiksi Saksassa (Georg Wenker 1895: *Sprachatlas des Deutschen Reichs*) ja Ranskassa (Jule Gillié, Edmond Edmont 1902–1910: *Atlas linguistique de la France*). Kettusen *Suomen murteet* -teossarja aineistoinen, samoin kuin Andrus Saaresten *Eesti murdeatlas* aineistoinen, asettuivat siis osaksi kansainvälistä murremaantieteellistä ja kielihistoriallista tutkimusperinnettä.

Siitä huolimatta, että Kettunen oli äännehistorioitsija, Kettusen *Suomen murteet* -sarjan aineistoa ja tarkastelun näkökulmaa voi kuvata synkroniseksi, sillä suurin osa Kettusen aineistosta on hänen itsensä vuosina 1925–1939 keräämää³, vaikka kielenilmiöitä kartaston selitysosassa käsitellessään Kettunen keskittyy kielenpiirteiden ja varianttien kielihistorialliseen tarkasteluun. Kettunen (1930a: XXIX, 1940b: VI) myös kertoo täydentäneensä itse keräämäänsä aineistoa muiden tutkijoiden⁴ muistiinpanoilla.

Lauri Kettunen (1930a: XXI) mainitsee teossarjan ensimmäisen osan esipuheessa, että syyt *Suomen murteet* -sarjan kokoamiseen liittyivät paitsi tutkimukseen myös opetukseen, jota hän piti Tarton yliopistossa:

² *Äännelaki* tarkoittaa historiallis-vertailevassa kielentutkimuksessa säännöllistä, samassa ympäristössä samalla tavoin tapahtunutta äänne- ja muoto-opillisten muutosten (ks. esim. Hamunen, Konstenius 2020: 125, Saarikivi 2020: 691–693).

³ Teossarjan ensimmäistä osaa varten Kettunen teki kenttätöitä vuosina 1925–1930, mutta kartastoa varten Kettunen on suorittanut täydennyskeräyksiä ainakin Keski-Suomessa, Pohjois-Savossa ja Kainuussa (v. 1930), Hämeessä ja Satakunnassa (v. 1932), Satakunnassa (v. 1935), Laatokan Karjalassa (v. 1938) sekä Pielisen Karjalassa, Kainuussa ja Lapissa (v. 1939) (Räisänen 1991: 58–60).

⁴ Kettunen (1930a: XXIX): Maist. Inkeri Tuovinen on antanut muutamia keräämiään näytteitä Tornion murteesta [---]. Kettunen (1940b: VI): „Kiitän [---] tiedonannoillaan auttajia, joista mainitsen opettaja J. Kemilän, prof. E. A. Tunkelon, maisterit L. Lounelan, Veikko Ruoppilan, Arvo Inkilän, Hannes Tepon, Eeva Lindénin sekä tohtorit Kustaa Vilkun ja Lauri Hakulisen. Prof. Martti Rapola suoritti antamani luettelon mukaan v. 1935 kyselyjä Jyykeän vuonon murteesta, ja tiedot sieltä perustuvat hänen havaintoihinsa.“

Tarton yliopistossa olin pitänyt edistyneimmille kuulijoille luentoja suomen murteista ja sitä varten omien vähäisten kokoelmiäni lisäksi myös kirjallisuudesta kerännyt ja systematisoinut saatavissa olevat tiedot. Siirryttyäni takaisin suomen kielen dosentiksi Helsingin yliopistoon (kevätlukukauden alussa 1925) päätin sekä luentoja että omia tutkimuksiani varten erikoisesti ryhtyä syventymään suomen murteisiin, ja miltei koko kesän 1925 kiertelin itä-Suomessa sekä tein muistiinpanoja myös Kuopion, Mikkelin, Viipurin ja Lappeenrannan vankiloissa.

Alpo Räisänen (1991: 56) arvelee, että Kettunen olisi saanut idean kartastoon opettajansa ja kollegansa Heikki Ojansuun johtamasta murreseminaarista, kun taas Seppo Suhonen (1985: 257) puolestaan esittää, että Kettunen sai herätteen murreatlaksensa kokoamiseen Tartossa Andrus Saarestelta, joka aloitti Viron murrekartastoon liittyvän aineistonkeruun vuonna 1915 (Hakulinen 1952).

Kettunen (1930a: XXI–XXIII) kuvaa *Suomen murteet* -sarjan ensimmäisen osan esipuheessa kenttätöitään (v. 1925–1930) sekä käyttämiään keruumenetelmiä: hän on pyrkinyt systemaattiseen kyselyyn, mutta tallentanut myös murrenäytteitä sekä teknisin apuvälinein että kirjoittamalla kontrolloidakseen kysely- ja keskustelumenetelmällä saatuja vastauksia sekä päästäkseen käsiksi kielenilmiöihin, joita kyselymenetelmä ei paljasta:

Käytyäni itsekustakin murteesta läpi yksityiskohtaisen, tutkimusluonnoksen tapaisen hakemistoni (jonka jo Tartossa olin laatinut tuntemani savon murteen pohjalle) koetin mikäli mahdollista saada kielenoppaalta myös murrenäytteitä, joita kokosin etupäässä fonografilla sekä teknillisesti täydellisemmällä n.s. parlografilla. Näillä samoin kuin välittömästi sanelun mukaan kirjoitetuilla näytteillä ei aluksi ollut juuri muuta tarkoitusta kuin kontrolloida kyselyn mukaan tulleiden vastausten tarkkuutta sekä saada esiin piirteitä, joita kysellen ja keskustellen ei ollut tullut näkyviin. Seuraavana kesänä (1926) jatkoin tutkimuksia kaksi viikkoa metsäsuomalaisten luona (jolloin fonografoin heidän kielennäytteitään) sekä toista kuukautta Uudellamaalla ja Hämeessä (myös Katajanokan ja Hämeenlinnan vankiloissa). Tutkimustapani oli kuitenkin muuttunut sikäli, että minulla oli n. 300 kartaketta (myöhemmin niiden lukua lisäsin n. sadalla), joihin eri värein merkitsin huomattavimpien äänne- ja muoto-opillisten seikkain edustustavat sekä myös eräitä sanastollisia eroavaisuuksia, käyttäen täydennykseksi entistä hakemistoani ja koettaen tietenkin edelleen saada aina kielennäytteitä kyselyjen päälle.

Kiinnostavaa on, että Kettunen kuvaa käyttäneensä keruussaan noin 400 kartaketta ja huomioineensa myös sanastollisia eroavaisuuksia. Julkaistussa kartastossa (1940a) karttoja on siis vain noin puolet (213) keruussa käytetyistä. Kartaston selitysosan esipuheessa Kettunen (1940b: V) kertoo halunneensa ottaa kartastoonsa mukaan vielä parikymmentä sanastollisia eroja kuvaavaa karttaa, mutta karttoja oli jo enemmän, kuin kustantaja oli lupautunut painamaan.

Suomen murteet -sarjan ensimmäisen osan esipuheessa Kettunen (1930a: XXIV–XXIX) kuvaa aineistonsa litterointiprosessia ja noudattamaansa transkriptiosysteemiä. Kiitoksien yhteydessä hän (1930a: XXIX) mainitsee, että eräät tutkijat tai muut henkilöt⁵ ovat „tarkastaneet tuntemansa murteen tekstejä ja hyväntahtoisesti tehneet huomautuksiaan“. Saattaa olla, että toiset tutkijat ovat siis „korjailleet“ *Suomen murteet* -sarjan ensimmäiseen osaan (*Murrenäytteitä*) painettuja litteraatteja, mikä lienee ollut ennen tallennusteknologian yleistymistä tavanomaista (tarkemmin, ks. Aarikka 2023: 106, 145–146). Toisaalta Kettunen (1930b: XXVII) korostaa, että koska kielenäytteistä on olemassa parlografitallenteet, kuulohavainnon voi toistaa ja kontrolloida. Chambers ja Trudgillin (2009) esittelemistä murreatlaista vasta vuonna 1968 käynnistyneessä *Linguistic Atlas of the Gulf States* -hankkeen aineistokeruussa käytettiin nauhoitteita. Se, että Kettunen todella tallensi varhaisella tallennusteknologialla puhetta, tekee hänestä edelläkävijän: puheen koneellinen tallennus yleistyi vasta 1950-luvulla kelanauhureiden myötä ja korvasi puhetilanteessa tehdyn muistiinpanolitteroinnin lopullisesti vasta 1980-luvulla (tarkemmin, ks. Aarikka 2023: 98–106).

Suomen murteet -teossarjan kartastossa (Kettunen 1940b) on 213 karttaa, jotka on järjestetty samankaltaisesti kuin suomen ja fennougristiikan alalla tehtyjen äännehistorioiden dispositiot eli *konsonantismiin* ja *vokalismiin* (tarkemmin, ks. Aarikka 2023: 220–227, 332). Tarkkailemistaan kielenilmiöistä Kettunen (1930b: III) itse toteaa:

Eri murrealueiden luonteenomaisten piirteiden deskriptiiviseen luetteloon teki mieli lisäillä tietoja piirteiden alkuperästä ja levenemisestä, mikä taas vei murrealueiden keskinäisen suhteen, jopa niiden synnyinkin määräälemisiin

⁵ Kettunen mainitsee seuraavat henkilöt: E. A. Saarimaa, V. Myrsky, Toivo Lehtisalo, Lauri Hakulinen, Elsa Haavio, Aarne Anttila, Niilo Ikola, Pekka Katara, Anna Salava, Martti Rapola, Artturi Kannisto, E. A. Tunkelo, Selma Patajoki, Väinö Tupala, Aino Oksanen, Eeva Lindén, Hannes Teppo, J. Beronka ja Väinö Pajari.

tahi sitä koskeviin viittauksiin. Houkutus oli suuri painua yhä syvemmälle eri alueiden asutushistoriaan, mutta puutteelliset tiedot omankin tieteen alalla (sanasto, henkilö- ja paikannimistö) pakoittivat pidättyväisyyteen. Murteiden erikoisuudet luonnehtuvat kuitenkin varsin oleellisesti äänne- ja muotorakenteesta, ja sen perusteella tehdyt päätelmät eivät koskaan liene aivan yksipuolisia.

Vaikka Kettusen kartastossa (1940a) käsiteltävät kielenilmiöt ovat äänne- ja muoto-opillisia, on hyvä muistaa, että hän on kiinnittänyt keruussa huomiota myös sanastoon (Kettunen 1930a: XXI–XXII). Kettusella on hänen ilmoituksensa mukaan (1940b: V) ollut keruutyön tuloksena ainakin parikymmentä leksikaalisia eroja valaisevaa karttaa – todennäköisesti enemmän, kun otamme huomioon, että Kettunen kertoo kartakkeita olleen keruun tukena noin 400.

Kettusen keskittyminen äänne- ja muoto-opillisiin piirteisiin selittyy äännehistoriallisella tutkimusperinteellä, joka oli tutkimuksen valtavirtaa suomalaisessa fennistiikassa 1900-luvun alkupuoliskolla (tarkemmin, ks. Aarikka 2023: 212–227 ja siinä mainitut lähteet). Suomen murteiden sanastonkeruu oli aloitettu jo 1800-luvun puolella ja *Suomen murteiden sana-arkiston* aineistosta tehtiin myös erilaisia sanastokarttoja (Leino ym. 2006, Hyvönen ym. 2007, Aarikka 2023: 79–83, 311–318). Sekä Kettunen että esimerkiksi *Johdatus suomen murteisiin* -teoksen (1947) julkaissut Martti Rapola⁶ pitivät kuitenkin äänne- ja muotopiirteitä sanastoa merkittävämpänä asiانا murrealueita- ja rajoja määriteltäessä (Aarikka 2023: 319–344). Myös modernissa dialektologiassa esitetään usein sanastolliset erot kielen „pintarakenteena“, jonka perusteella murrerajoja ei kannata määritellä (Chambers, Trudgill 2009: 112–116, Mielikäinen 2001: 378–382).

Suomen murteet -sarjaan liittyvää keruutyötä Kettunen teki apumoottorilla varustetulla pyörällä, moottoripyörällä ja autolla, välillä matkakumppanien kanssa ja lomamatkojen yhteydessä (tarkemmin, ks. Räisänen 1991, Kettunen 1960b). Poikkeuksellista ja kiinnostavaa on, että osa Kettusen kielenoppaista oli vankeja, sillä hän teki kenttätöitä ainakin Hämeenlinnan,

⁶ Rapola (1932) arvioi Virittäjässä Kettusen (1930b) *Murrealueet*-teoksen. Yksi Rapolan (1932) keskeinen kritiikki on, ettei Kettunen ota jaossaan huomioon sanastollisia eroja. Tästä huolimatta Rapola (1947) ei *Johdatus suomen murteisiin* -teoksessaan, jossa hän esittää oman – Kettusen aluejaon kanssa lähes identtisen – aluejakonsa, käsittele juurikaan sanastollisia eroja (tarkemmin, ks. Aarikka 2023: 319–344).

Katajanokan, Kuopion, Lappeenrannan, Mikkelin, Oulun ja Viipurin vankiloissa (Kettunen 1930a: XXI–XXII, Räisänen 1991: 57). Ratkaisun takana oli sattuma, jota Kettunen (1930a: XXI) kuvaa seuraavalla tavalla:

Saanen huvin vuoksi mainita, että vankila-ajatuksen minulle toi pelkkä onnettomuus. Polkupyöräni apumoottori joutui näet kerran ensi matkallani Kuopion seudulla auttamattomasti epäkuntoon, ja asemalle mennessäni juolahti mieleen yrittää vierailua lähellä olevassa vankilassa. Kun kokeilu antoi erittäin hyviä tuloksia, pyysin ja sain vankeinhoitohallitukselta luvan työskennellä kaikissa vankiloissa. Välistä oli kyllä turvauttava oppaisiin, jotka eivät olleet täysin luotettavia, mutta heidän tiedonantojaan on enimmäkseen ollut tilaisuus myöhemmillä matkoilla kontrolloida.

Suomen murteet -sarjan kakkososan (*Murrealueet*) Virittäjässä arvioinut Martti Rapola (1932: 80) kiinnittää Kettusen toimintatapaan huomiota: „Jo se tapa, jolla Kettunen on huomattavan osan ainestaan koonnut vankilain asukkailta, kehoittaa varovaisuuteen“. Lisäksi Kaisa Häkkinen (1993: 28–29) on esittänyt, että Suomen sisällissodan jälkeen „sivistyneistö oli pettynyt kapinoivaan roskajoukkoon ja halusi pitää varsinaisena Suomen kansana sitä kurissa ja Herran nuhteessa elävää väestönosaa, joka pysyi uskollisena vallitsevalle yhteiskuntajärjestykselle“ ja että „kielentutkijan ihanteeksi ja tutkimuskohteeksi nousi vapaa suomalainen talonpoika, joka syrjäisellä saloseudulla peltoaan raivaten vietti vakaan rauhallista elämää“. Ainakaan Kettusen tapauksessa Häkkisen (1993: 28–29) näkemys ei pidä paikkaansa: tärkeintä oli murre, ei kielenoppaan „kunnollisuus“.

Suomen murteet -sarjan, kuten myös muiden kielialueiden murrekartastojen tapauksessa ei ole tarkalleen tiedossa, miten montaa kielenopasta Kettunen haastatteli kunkin pitäjän murre-edustuksen kuvaamiseksi, ja keitä nämä ihmiset olivat. Kettunen (1930a: XXIX–XXX) kertoo kielenoppaistaan seuraavaa:

Kiitollisena muistelen myöskin avulialita kunnalliskotien emäntiä ja isäntiä sekä vankiloiden johtajia, mutta ennenkaikkea satoja kielenoppaitani ja tietomiehiäni, jotka viettävät vanhuutensa päiviä kunnalliskodeissa tahi askartelevat mökki-pahaisissaan maanteiden varsilla, – sekä niitä, jotka viettivät pitkiä päiviään vankiloissa ja osasivat muiden tavalla kanssani iloita siitä, että voivat tyhjänarvoisina pitämillään tiedoilla hyödyttää tutkimusta.

Todennäköisesti suurin osa kielenoppaista oli Kettusen kuvauksen mukaisesti kouluttamattomia vanhuksia, jotka olivat asuneet kotipitäjässään koko ikänsä. On kuitenkin tärkeää todeta, että kielenoppaiden valinnassa oleellisinta Kettuselle ei vaikuttaisi olleen puhujien ikä tai sukupuoli vaan heidän murteensa (kielenoppaiden kieliperustaisesta valikoimisesta fennistisessä murteentutkimuksessa, ks. Aarikka 2023: 168–177). Kettusen kielenoppaiden joukossa on myös koulutettuja ja nuoria, mikä käy ilmi *Suomen murteet* -sarjan ensimmäisestä osasta *Murrenäytteitä* (ks. esim. Kettunen 1930a: 58, 64, 75, 203). Tämä on huomion arvoista siksi, että esimerkiksi Ranskan murreatlasaineistossa 700 kielenoppaasta vain 60 oli naisia ja vain 2/7 oli koulutettu kansan keskiarvoa pitemmälle (Chambers, Trudgill 2009: 17). Miehiä haastateltiin murrekartastoihin naisia enemmän, sillä naisten oletettiin olevan tietoisempia puhumastaan kielestä ja miesten puhuvan murretta tilanteessa kuin tilanteessa (Chambers, Trudgill 2009: 30). Suomalaisen murteentutkimuksen traditiossa naisia ei kielenoppaina kuitenkaan sivuutettu (Aarikka 2023: 159–160) ja Kettusen kielenoppaissa vaikuttaisi *Murrenäytteitä*-teoksen perusteella olleen paljon naisia.

Teossarjan ensimmäisen, *Murrenäytteitä*-osan arvioi Virittäjässä Aarni Penttilä (1930) ja arvio oli perin myötäsukainen. *Murrealueet*-osa, joka oli ensimmäinen suomen murrealueiden kokonaisesitys, puolestaan sai Virittäjässä murskakritiikin tieteenalan auktoriteetilta, Helsingin yliopiston suomen kielen professorilta Martti Rapolalta⁷ (1932). Rapola (1932: 82) päätti arvionsa sanoihin:

En katso Kettusen kirjan täyttävän sellaisia vaatimuksia, joita yliopistolliselle kurssikirjalle on asetettava. Arvostelukykyinen käyttäjä huomaa esittämäni puutteellisuudet ja mahdottomuudet yhtä helposti kuin minäkin, ja hän voi teoksesta poimia joukon tietoja ja viittauksia hyödykseenkin, mutta opiskelijamme kaipaavat tosiaan paremmin jäsenneltyjä, valmiimmaksi ajateltuja, luotettavammin perustelevia ja selvemmin kirjoitettuja opastajia suomen murteisiin, kuin on kirja, jonka prof. Kettunen on heitä varten kirjoittanut.

Rapolan (1932) erittäin kriittinen arvio lienee osaltaan vaikuttanut siihen, ettei Kettusen (1930b) murrejaottelu saavuttanut tieteenalalla kanonisoitunutta

⁷ Lauri Kettusen ja Martti Rapolan vaikeasta suhteesta sekä Rapolan (1932) esittämästä kriittistä tarkemmin, ks. Aarikka 2023: 323–331.

asemaa⁸, vaikka muita kirjasarjan osia – erityisesti kartastoa (1940a) – on hyödynnetty fennistiikassa laajasti. On myös syytä epäillä, että Kettusen *Suomen murteet* -sarjaan liittyvää työtä on hyödynnetty osittain ilman asianmukaisia viitteitä myöhemmissä suomen murteiden aluejakoesityksissä (Aarikka 2023: 319–344).

Suomen murteet -sarjan kartasto-osa sekä sen selitysosa ilmestyivät kymmenen vuotta myöhemmin kuin teossarjan kaksi ensimmäistä osaa. Syynä viivästymiseen lieni muihin töihin ja professuurin hoitamiseen liittyvät kiireet ja taloudelliset hankaluudet kartaston julkaisuun ja painattamiseen liittyen, mutta kenties myös Rapolan (1932) kritiikki (Kettunen 1932: 186, Kettunen 1940a: III–VI). Teossarjan kartasto-osan ja sen selitysoosan arvioi Virittäjässä jälleen Aarni Penttilä (1941: 250, 252) hyvin arvostavaan sävyyn: „Prof. Kettusen työkyky ei ole tavallisin mitoin mitattavissa. [---] Tutkimuksen onkin täysi syy olla lähinnä kiitollinen tekijälle hänen uudesta suurteoksestaan.“

Kettusen *Suomen murteet* -teossarjan valmistumisen kynnyksellä Suomi oli ajautumassa talvisotaan ja Kettunen joutui viimeistelemään kartasto-osat kiireellä, eikä voinut tehdä kaikkia haluamiaan tarkistuksia niihin. Kartat painettiin Tampereella ja kartaston selitysosa Helsingissä sodan jo ollessa käynnissä. (Kettunen 1940a: III–VI, Räisänen 1991: 55.)

2. Kettusen (1940b) kartasto digitaaliseksi ja avoimeksi

1800-luvun lopun ja 1900-luvun alun murrekartastot tähtäsivät kielen sisäisen vaihtelun alueellisen jakautumisen selvittämiseen. Periaattessa murrekartastot mahdollistivatkin alueellisen vaihtelun kuvaamisen, mutta käytännössä esitietokoneaikana aineistojen käsittely jäi yksittäisten piirteiden tai ihmisäivoin käsiteltävissä olevien piirryhmien tarkasteluun. Murrekartastojen maantieteellisen vaihtelun kaoottisuus tuskastutti tutkijoita ja oli vaarassa haastaa kielen kehittymiseen liittyvät hypoteesit (Bailey 1973: 86) Dialektologian päämääräksi asetettiin tarve hallita sitä kaaosta, jota murrekartastot edustivat (Goebel 2010: 435), ja laskennallisten kaaoksenhallintamenetelmien kehittyessä kehittyi myös uusi murretutkimuksen ala, dialektometria. Yksi

⁸ Kotimaisten kielten keskuksen sivuilla (sähköinen lähde, Kotus) suomen murteiden jaoteltuna esitetään Terho Itkosen (1989: 342–343) murrealuejako, joka perustuu Martti Rapolan (1947) esitykselle.

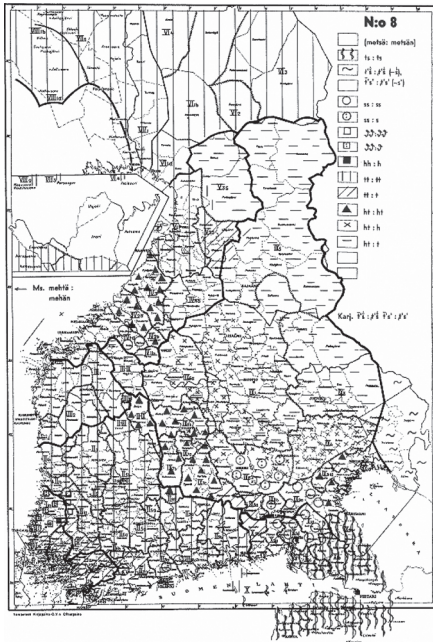
ensimmäisistä kehitetyistä mittareista oli Jean Séguy (1973) laskenta alu-
eiden kielellisen etäisyyden kuvaamiseksi, ja yksi uusimmista on tässäkin
työssä esitelty populaatiogeneettinen lähestyminen. Kretzschmar, Schneider
ja Johnson (1989) ennustivat ihan oikein, että tietokoneet tulevat auttamaan
laajojen murreaineistojen käsittelyssä.

Murreaineistojen laskennallinen käsittely tietenkin vaati, että aineistot
ovat digitaalisessa muodossa. Suomen murteiden digitaalisen tutkimuksen
kannalta oli onnekasta, että Yorkin yliopiston Sheila Embleton ja Eric S.
Wheeler sekä Kotimaisten kielten keskus (Kotus) näkivät Kettusen murre-
kartaston (1940b) mahdollisuudet digitaalisen kielentutkimuksen aineistona.
Kartaston digitalisointi aloitettiin Yorkin yliopiston kurssityönä 1990-luvulla,
ja ensimmäiset analyysit aineistosta julkaistiin jo vuosituhatlupulla
(Embleton, Wheeler 1997, 2000). Karttojen digitalisointi oli askel kohti Ket-
tusen isoglossikkaoksen tietokoneavusteista hallintaa. BEDLAN-työryhmä
(*Biological evolution and diversification of languages*) sekä Kotus jatkoivat
digitalisoidun murrekartaston käsittelyä vuonna 2009: aineiston virheet kor-
jattiin ja siihen sovellettiin evolutiivisen kielentutkimuksen viitekehystä sekä
maantieteen ja populaatiogenetiikan menetelmiä vaihtelun kuvaamiseksi ja
hallitsemiseksi. Samoja menetelmiä on sittemmin sovellettu myös sveitsin-
saksan (Romano ym. 2022) ja fidžin murteiden tutkimiseen (Murawaki 2020).

Kotus sijoitti korjatun aineiston Etsin.Fairdata.fi-sivustolle⁹. Fairdatan
versio on vaikeaselkoisessa muodossa asiaa tuntemattomalle ja ei-suomen-
kieliselle, ja siitä on pudonnut pois 26 varianttia. Helpommin lähestyttävä,
edelleen korjattu ja koodattu aineisto sekä sen kuvaus englanniksi löytyvät
julkaisusta Santaharju ym. (2024, lähetetty revisio), joka esittelee Kettusen
kartaston BEDLAN-töissä käytetyn aineistoversion¹⁰. Aineisto on nyt avattu
uudestaan FAIR-periaatteisiin pyrkien: Se on helposti löydettävissä Zeno-
dosta (*Findable*) ja vapaasti ladattavissa ja käytettävissä (*Accessible*). Sitä
voi käyttää yhdessä muiden spatiaalisten datojen kanssa (*Interoperable*),
sillä aineisto sisältää kuntien nimet, kuntakoodit ja koordinaatit. Se on myös
laajemmin, myös ei-suomenkielisten tutkijoiden käytettävissä (*Reusable*).
Korkeatasoiset kuvat Kettusen *Suomen murteet* -murrekartastosta ovat saa-
tavilla Juha Kuokkalan digitoimina sivustolta: <http://kettunen.fnhost.org>
(Kuva 1).

⁹ Saatavilla: <http://urn.fi/urn:nbn:fi:csc-kata20151130145346403821>

¹⁰ Saatavilla: doi.org/10.5281/zenodo.10078078



Kuva 1. Murrekartaston kartta 8: *metsä, messä, mehtä* jne. (tavunloppuisten klusiilien historiaa, soinnittomien konsonanttien edessä). Kyseisellä kielenpiirteellä on 12 erilaista varianttia, joista pitäjissä (n=525) edustuu yleensä 1, joskus kaksi ja hyvin harvoin 3 tai 4 (Syrjänen ym. 2016)

3. Murteiden muodostumisen tutkimusta

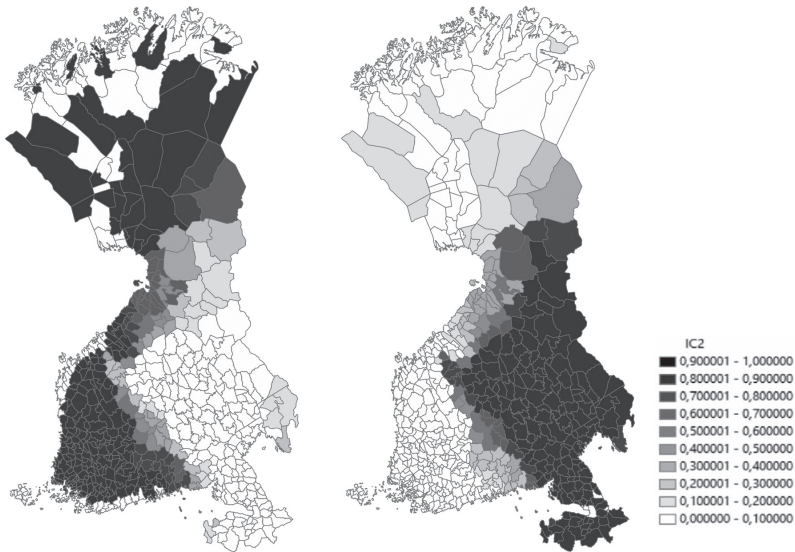
3.1. Murrejaot ja pitäjien kuuluminen moniin murteisiin

Murteiden jaottelu ei ole evolutiivisen kielentutkimuksen päätarkoitus, sillä usein laadullisesti tehdyt tutkimukset ovat tuottaneet hyvin pitävät murreajat¹¹. Laskennallisen lähestymistavan tarkoitus on kuvata aineistoa ja sen vaihtelua uusin tavoin, jotta kielen vaihtelun kehittymistä on mahdollista tutkia. Populaatiogenetiikan menetelmäpakki tarjoaa mittareita, joilla laajoista murrekartastoaineistoista saadaan näkyviin ilmiöitä, joita ei yksittäisiä piirteitä ja niiden variaatiota tarkastelemalla tavoiteta. Mallipohjaisilla, bayesilaisilla klusterointitekniikoilla voidaan jakaa aineisto esimerkiksi K-määrään klustereita, ja ohjelma arvioi erilaisia mittareita käyttäen, mikä jakotapa parhaiten sopii aineistoon. Esittelemme ensin Kettusen aineistolle saatuja jakoja, ja sen jälkeen tuloksista tuotettuja uusia lukuja.

STRUCTURE-ohjelmalla laskettuna Kettusen kartastoaineistoon parhaiten sopiva jako tuotti kaksi klusteria (Kuva 2, itäinen ja läntinen

¹¹ Suomen murrealuejajoista tarkemmin, ks. Aarikka 2023: 271–375 ja laadullisen ja määrällisen murrejaottelun eroista Syrjänen ym. 2016.

päämurrealue), mutta klustereiden lisääminen aina 14:än asti paransi hie-
man mallin sopivuutta pitäjien väliseen kielelliseen vaihteluun kuvaamisessa
(Syrjänen ym. 2016). Toinen mallipohjainen klusterointiohjelma on BAPS,
jonka käytännön ero STRUCTURE:en on siinä, että se löytää usein hieno-
piirteisempiä eroja klustereiden välillä. BAPS antaa suoraan parhaan murre-
jaon, eikä STRUCTURE:n lailla anna käyttäjän arvioida itse erilaisia jakoja
(ks. myös Honkola ym. 2018). Kun BAPS etsi itse sopivinta murrejakoa, oli
lopputuloksena 16 aluetta (Santaharju ym. 2024, Kuva 3). Kanonisoituneena
suomen murrejakona voi pitää on Terho Itkosen (1964, 1989) jakoa kah-
deksaan murrealueeseen (ks. Aarikka 2023: 345–353), mutta tällaista jakoa
mikään BEDLAN-työryhmän analyysi ei ole tukenut.



Kuva 2. Kettusen murrekartaston aineiston jako kahteen päämurrealueeseen
BAPS-ohjelmistolla ja bayesilaisella klusteroinnilla (Santaharju ym. 2024),
kun tulos on yhdistetty paikkatieto-ohjelmaan rakennettuun kuntajakoon.
STRUCTURE-ohjelmistolla tehty jako on tätä vastaava. Värin tummuus kuvaa,
miten suurelta osin kukin pitäjä kuuluu kyseiseen murreklusteriin: musta pitäjä
kuuluu klusteriin 90.00001–100 prosenttisesti, toiseksi tummin väri 80.00001–90
prosenttisesti ja niin edelleen. Valkoisen värin pitäjä ei kuulu ao. klusteriin
laisinkaan. Tästä ja muista kuvista on jätetty Värmlanti pois

Mallipohjaisten klusterointimentelmien etu matemaattisiin klusterointeihin on, että tulokseen jää näkyviin isoglossien todellinen vaihtelu, eli murrealueille ei pakoteta tarkkoja rajoja (ks. Syrjänen ym. 2016). Klusterointiohjelmaa pyydetään jakamaan aineisto haluttuun määrään klustereita: esimerkiksi kuvassa 2 on tulos kahteen klusteriin jakamisesta ja kuvassa 3 kuuteentoista klusteriin jakamisesta. Kukin kielenpiirre (213 karttalehteä) kuvaa kielen alueellista vaihtelua hieman eri tavoin eli asettaa varianttien isoglossit hieman eri paikkoihin. Näin ollen harvempi pitäjä on joka karttalehdellä aina samassa klusterissa, vaan pitäjissä on vaikutteita useammasta klusterista (*affinity*). Esimerkiksi jaossa kahteen ($K=2$, Kuva 2) kullakin pitäjällä on itäisyysfrekvenssi ja läntisyysfrekvenssi, ja vaikkapa meänkielen alue on yli 90-prosenttisesti läntisen päämurteen aluetta, mutta siellä on myös itäisiä variantteja. Kuvasta 3 puolestaan käy ilmi, että itäisen Lapin pitäjät kuuluvat sekä keski- ja pohjoispohjalaismurteiden ryhmään että peräpohjalaismurteiden ryhmään (pitäjillä on siis „pohjalaisfrekvenssi” ja “peräpohjalaisfrekvenssi”).

Frekvenssit ovat käyttökelpoisia muuttujia, sillä ne voi yhdistää tai niitä voi verrata muihin pitäjäkohtaisiin aineistoihin. Olemmekin verranneet pitäjän kielellistä itäisyyttä samalla lähestymistavalla laskettuun pitäjän geneettiseen itäisyyteen (Kerminen ym. 2021) sekä itse laskemiimme pitäjien kansankulttuuripiirteiden itäisyyttä kuvaaviin frekvensseihin (Rantanen ym., tekeillä). Vertasimme *Mutual Information* -laskennalla pitäjien itäisyyttä eri aineistoissa tutkien, osuuko Suomen itä-länsijako kieli-, geeni- ja kulttuuriaineistojen perusteella samalle alueelle. Geeni- ja kulttuurieroja kuvaavien aineistojen rajat ovat samankaltaisemmat kuin murreaineistojen ja geeni- sekä kulttuuriaineistojen perusteella muodostetut rajat. Pitäjien frekvenssejä käyttämällä voimme yhdistää kielellisen tiedon muiden alojen spatiaalisista ilmiöistä kertoviin aineistoihin ja tarkastella esimerkiksi sitä, näkyvätkö Suomen historialliset hallinnolliset rajat – kuten Pähkinäsaaren rauhan raja – Kettusen aineiston itä-länsi-jaossa (Rantanen ym., tekeillä).

3.2. Hierarkinen murrejako

Kuten kieliaineistojen myös biologisten lajien rakenne on usein hierarkinen, ja hierarkisen rakenteen selvittämiseen on kehitetty omia laskentamenetelmiä. Santaharju ym. (2024) selvittivät populaatiogenetiikan viitekehukseen



Kuva 3. Kettusen murrekartaston aineiston jako 16 murrealueeseen BAPS-ohjelmistolla (Santaharju ym. 2024). Pohjoisimpien pitäjien yhdistyminen hämäläismurteiden kanssa samaan klusteriin liittyyne siihen, että Lapin pitäjistä Kettunen keräsi vain vähän aineistoa (joitain kymmeniä kielenilmiöitä 213 sijaan; Honkola ym. 2018). Värien koodit, ks. kuva 2

kuuluvalla AMOVA-analyysillä (*Analysis of Variance*) Suomen murteiden hierarkista rakennetta. GENALEX-ohjelmalle kerrottiin, että Kettusen aineisto koostuu itä- ja länsimurteista, jotka koostuvat paikallismurteista, jotka koostuvat puolestaan hieman naapuristaan poikkeavista pitäjänmurteista. Miten paljon vaihtelusta kukin taso selittää? Kettusen murrekartaston (1940b) esittämästä vaihtelusta jako itään ja länteen selitti yksinään 20 %. Jako 16 murteeseen – huomioiden että kukin paikallismurre kuuluu joko itään tai länteen – selitti 36 % vaihtelusta. Pitäjien välillä on kuitenkin paljon eroa, sillä jopa 44 % vaihtelusta selittyi pitäjien erilaisuudella. Sama on nähtävillä usein myös geneettisessä aineistossa: yksilöt muodostavat kyllä populaatioita, mutta pääosa vaihtelusta liittyy yksilöiden eroavaisuuksiin (tarkemmin, ks. Santaharju ym. 2024).

3.3. Kuinka erilaisia murteet ovat ja miksi?

Populaatiobiologiassa populaatioiden geneettistä erilaistumista pystytään mittaamaan monin tavoin: Syrjänen ym. (2016) sovelsivat fiksaatioindeksiä F_{ST} ja Santaharju ym. (2024) ϕ . Fiksaatioindeksit mittaavat populaatioiden erilaistumisen määrää pitkällä aikavälillä, eli populaatorakenteen merkitystä lajin yksilöiden välisen variaation selittäjänä. Santaharju ym. (2024) käyttävät fiksaatioindeksiä mittana murteiden välisistä geneaologisista yhteyksistä tai geneaologisesta etäisyydestä. Fiksaatioindeksit lasketaan murreklusterien välille, mutta pitäjien välille voi laskea kielellisen etäisyyden Séguyn menetelmällä, jossa lasketaan oikeastaan otostettujen yksiköiden eroavaisuuksia (Chambers, Trudgill 2009).

Honkola ym. (2018) laskivat F_{ST} -luvut 14 murrepopulaation kielieroille ja Séguyn distanssit pitäjien välisille kielellisille eroille. He laskivat alueiden väliset erot muunlaisista aineistoista: esimerkiksi erot pitäjien ja murrealueiden lämpötiloissa, lapsikuolleisuudessa ja savupirttien määrässä. Aineistojen vertaaminen oli mahdollista, koska Kettusen murteiden otostusyksikkö oli pitäjä. Samat pitäjät löytyvät monesta muustakin aineistosta, kuten tilastollisista vuosikirjoista, kirkonkirja-aineistoista tai Karjalan siirto-laisten lähtö- ja tulopitäjien tiedoista. Pitäjäkohtaista tietoa on siis saatavilla monesta eri muuttujasta.

Kielimuotojen erilaisuuden taustalla oletetaan olevan maantieteellinen etäisyys – mitä kauempana alueet ovat toisistaan, sitä erilaisempia kielimuodot

todennäköisesti ovat, koska puhujaryhmien välillä on vähemmän kontakteja (Chambers, Trudgill 2009, Hinskens ym. 2009). Toisaalta isolaatiota voivat tuottaa myös esimerkiksi hallinnolliset rajat (Diener, Hagen 2012) ja kulttuuriset alueet, esimerkiksi sosiaalisten asenteiden kautta (Thomason 2008). Honkola ym. (2018) testasivat edellä mainittuja hypoteeseja ja lisäksi ajatusta, että luonnonympäristöjen erilaisuus aiheuttaa ihmisryhmien isolaatiota. Esiteollisen ajan ihmisryhmät ovat olleet nykyistä enemmän sidoksissa ympäristöönsä – tietynlainen maanviljelykulttuuri levittäytyi ehkä sellaiselle alueelle, johon se parhaiten soveltui.

Honkola ym. (2018) tutkivat, oliko pitäjien murteiden välisillä etäisyyksillä tekemistä alueiden maantieteellisten, hallinnollisten, kulttuuristen ja ympäristöllisten etäisyyksien yhdistelmän kanssa. Murteiden etäisyyksiä sen sijaan selitettiin maantieteellisellä, kulttuurisella ja ympäristöllisellä etäisyydellä. Pitäjien välistä kielellistä erilaisuutta ja samanlaisuutta selitti hyvin maantieteellinen etäisyys, joka tosin yhdistyi vahvasti hallinnollisiin eroihin. Sen sijaan maantieteellinen etäisyys ei juurikaan liittynyt murteiden välisiin kielellisiin eroihin. Murteiden väliset erot sen sijaan olivat sitä suurempia, mitä erilaisempia ne olivat luonnonympäristöltään tai kulttuuritekijöiltään (tarkemmin, ks. Honkola ym. 2018).

Maantieteellisen etäisyyden merkitys oli yllättävä. Se ei ollut tärkeä tekijä paikallismurteiden isolaatiossa, mutta liittyi kyllä pitäjämurteiden erilaistumiseen. Tulos antoi viitteitä siitä, että paikallismurrealueiden eriytymiskehitykseen liittyvät kielen ulkoiset tekijät: erilaiset kulttuuriset kontekstit ja näihin mahdollisesti liittyvät erilaiset luonnonympäristöt.

3.4. Murteiden konvergenssi

Murteiden ja kielten alueelliseen erilaistumisen, *divergenssin*, vastavoimana on *konvergenssi*, eli kielimuotojen samankaltaistuminen (Hinskens ym. 2009, Braunmüller, House 2009). Kielten eriytymiskehitys eli kielen vertikaalinen evoluutio on usein kielievoluution tutkimuksen keskiössä, mutta uusia menetelmiä kehitetään myös kontaktien tuottaman muutoksen eli horisontaalin siirtymän tutkimukseen. Murteet voivat kehittyä sukukieliksi ja edelleen kieliksi, jos *divergenssivoimat* ovat *konvergenssia* suuremmat. Murteiden välisen *konvergenssin* ymmärtäminen *divergenssin* ohella on siis välttämätön osa kielihistoriallisten mekanismien tutkimusta.

Sekä divergenssin että konvergenssin taustalla ovat kielikontaktit, jotka esiteollisessa yhteiskunnassa tarkoittivat lähinnä ihmisten fyysisiä kohtaamisia. Nyky-yhteiskunnassa kontaktit voivat liittyä tietenkin myös muunlaisiin kohtaamisiin, esimerkiksi sähköisen median välityksellä. Ihmisryhmien kontaktien vahvuuden vaikutus kielen divergenssiin tai konvergenssiin ei kuitenkaan ole suoraviivaista. Periaatteessa eri alueilla puhuttaviin kielimuotoihin kumuloituu innovaatioiden tai kontaktien kautta erilaisuuksia: erilaisia ääntenmuutoksia, rakenteellisia ilmauksia ja sanastoa (Thomason, Kaufman 1988). Kielimuotojen väliset kontaktit sen sijaan voivat estää innovaatioiden vakiintumista tai levittää kumpaankin ryhmään saman innovaation. Toisaalta on havaittu, että kielikontaktit voivat myös tehostaa kielimuotojen välisen erojen kehittymistä, jos puhujaryhmät haluavat korostaa erilaisuuttaan (Evans 2019, Ellison, Miceli 2017, Mansfield ym. 2023).

Santaharju ym. (2024) kokeilivat uutta tapaa mitata kielellistä konvergenssia. He mittasivat kielivirtaa murteiden välillä samaan tapaan kuin populaatioiden välillä on mitattu geenivirtaa. Geenivirta kuvaa, kuinka paljon geneettistä materiaalia siirtyy yksilöiden tai vaikkapa siitepölyn avulla populaatiosta toiseen. Kettusen murrekartastoaineistosta laskettiin, kuinka paljon vaikkapa lounaismurteissa on todennäköisesti hämäläismurteista tullutta vaikutetta ja päinvastoin. Murteiden välisen kielivirtamittarin avulla Santaharju ym. (2024) pystyivät testaamaan hypoteeseja siitä, että kielikontaktin vahvuus riippuisi maantieteellisestä etäisyydestä tai siitä, kuinka läheistä sukua kielimuodot ovat ϕ -indeksillä mitattuna. Vaikka konvergenssin eli samankaltaistumisen määrä oli suurempi maantieteellisesti läheisten murteiden välillä, oli kontaktissa olevien murteiden kielellinen erilaistumisen aste vielä tärkeämpi konvergenssin vaihtelua selittävä tekijä: lähimpien sukulaisten välillä oli enemmän konvergenssia.

Edellä mainitut tutkimustulokset viittaavat siihen, että alueiden välisen kontaktien vahvuus ei liity pelkästään maantieteelliseen etäisyyteen, vaan myös kulttuuriseen ja kielelliseen samanlaisuuteen. Lynch ym. (2022) tutkivat talvisodan aikaisen Karjalan siirtoväen muuttoliikettä. Aineisto kertoi, mistä Karjalan pitäjästä siirtolainen lähti ja mihin päätyi, ja lisäksi sen, palasiko hän takaisin Karjalaan välirauhan aikana. Päätöstä jäädä tai lähteä takaisin selitettiin sosiaalisilla, ympäristöön liittyvillä ja kielellis-kulttuurisilla tekijöillä. Tulosten mukaan siirtokarjalainen pysyi todennäköisemmin uudella alueella, jos paikkakunnan kieli oli samankaltainen kuin lähtöalueen

kieli. Murre-erot mitattiin Sequyn distanssilla Kettusen murreaineistosta, ja aineistojen vertaaminen onnistui, koska kaikki tiedot olivat pitäjakohtaisia.

4. Lopuksi

Lauri Kettusta luonnehditaan useissa häntä ja hänen toimintaansa käsittelevissä teksteissä poleemiseksi hahmoksi, jonka kriittisyys, auktoriteettien uhmaaminen ja väittelyinnokkuus saattoivat osaltaan vaikuttaa siihen, ettei hän saanut aikanaan ansaitsemaansa arvostusta (ks. esim. Kelomäki 2009: 85, Vares, Häkkinen 2001: 132, Hänninen, Koponen 1991: 7–8). Suomen kielen alalla tehdyt bibliometriset sitaatioita tarkastelevat tutkimukset kuitenkin osoittavat, että *Suomen murteet* -sarja ja erityisesti sen kartasto-osa selityksineen kuuluvat viitatuimpiin teoksiin sekä fennistisessä aikakauslehdessä *Virittäjässä* (Karlsson 1994) että suomen kielen alan murteita käsittelevissä väitöskirjoissa (Aarikka 2018). Kettusen mittavaa työtä on siis hyödynnetty fennistisessä, pääasiassa laadullisessa kielentutkimuksessa jo kohta sata vuotta. *Suomen murteet* -sarjan kartasto-osaa on käytetty myös pitkään suomen kielen opetuksessa: siitä ilmestyi vielä vuonna 1981 neljäs, supistettu painos.

Lauri Kettunen ehdotti murrekartastonsa (1940a) alkulauseessa, että isoglossien rajoja voisi verrata kartastoon liitettyyn kalvoon Suomen vesistöistä. Lisäksi hän kaipasi murremaiseman tarkasteluun apukarttoja vanhoista maanteistä, pitäjien rajoista ja asutustiheydestä. Kartastoaineiston digitaalisen version ja siitä laskettujen uusien muuttujien yhdistäminen muiden tieteenalojen aineistoihin on nyt mahdollista ja Kettusen aineisto avaa uusia mahdollisuuksia monenlaiselle evolutiiviselle kielentutkimukselle. Evoluutiota – myös kielen evoluutiota – voi tutkia seuraamalla muutosta ajassa, mutta koska tällaisia aineistoja on varsin vähän, on vaihtoehtona hyödyntää spatiaalista vaihtelua kuvaavia aineistoja ja selvittää siitä variaatiota ajavia ja vähentäviä tekijöitä (Vesakoski 2009). Dialektometrisesti Kettusen aineistoa on toistaiseksi käsitelty yhtenä kokonaisuutena, mutta digitaalisen aineiston kielen tasojen eroja ja yhtenäisyyksiä pystyisi laskennallisilla menetelmin helposti vertaamaan (uralilaisten kielten äänne-, muoto- ja rakennetasojen vertailusta, ks. Norvik ym. 2022).

Suomen murteet -sarjan syntyvaiheita ei ole tutkittu laadullisesti. Kuten luvussa 1 toimme esiin, Kettunen on käyttänyt keruutyönsä tukena n. 400 kartaketta, vaikka julkaistussa kartastossa karttoja on 213. On siis mahdollista,

että lisäkartakkeet on arkistoitu jonnekin ja että ne sisältävät tietoa sanastollisesta vaihtelusta. Lisäksi emme tiedä tarkalleen, keitä Kettunen haastatteli. Kartaston selitysosan esipuheessa Kettunen (1940a: VI) mainitsee: „Suurimmalla kiittollisuudella muistelen satoja vaatimattomia kielenoppaitani, joiden nimet henkilötietoineen jäävät luetteloihini.“ Missä tämä luettelo on? Myös sitä, keneltä Kettunen sai sytykkeen *Suomen murteet* -sarjan kokoamiseen ja miten hänen käyttämänsä aineiston keruun ja käsittelyn menetelmät vertautuvat muihin eurooppalaisiin murreatlaksiin, pitäisi tutkia syvällisemmin. Jotta ymmärtäisimme paremmin Kettusen kartastoaineiston taustaa, olisi tarpeen tarkastella systemaattisesti koko *Suomen murteet* -sarjaa, Kettusen laajamittaisia kenttätömuistelmia sekä erilaisia Kettuseen liittyviä arkistoaineistoja¹².

Määrällinen tutkimus kielen muutoksesta ja kieli- ja väestöhistoriasta jatkuu Turun yliopiston *Ihmisen monimuotoisuus* -konsortion työssä (www.humandiversity.fi) ja laadullinen tutkimus Kettuseen ja erityisesti Suomen murteet -sarjan syntyvaiheisiin liittyen käynnistyy Emil Aaltosen säätiön apurahan turvin syksyllä 2025.

Kiitokset

Kiitämme Koneen Säätiötä, Kotusta ja Toni Suutaria: Vuonna 2008 Suutari esitteli Kotuksen digitoidun murreaineiston Koneen Säätiön seminaarissa „Evoluutionäkökulmia ihmistieteisiin“; Urho Määttä, Kalle Korhosta, Niclas Wahlbergiä, Minna Palander-Collinia, Carita Klippiä ja Unni Leinoa, joiden kanssa ensimmäinen BEDLAN-rahoitus saatiin Koneen Säätiöltä kyseisen seminaarin innostamana. Erityiskiitos menee Jyri Lehtiselle, joka kävi läpi digitoidun aineiston ja muokkasi sen kauttaaltaan. Populaatiogeneettisen lähestymistavan adoptoimisesta merkittävän kokoinen kiitos kuuluu Kaj Syrjäselle sekä Terhi Honkolalle ja avusta aineistomuotojen sekä analyysien kanssa Perttu Sepälle. Kiitos myös Emil Aaltosen säätiölle, jonka apurahan turvin laadullinen tutkimustyö Kettusen *Suomen murteet* -sarjan parissa jatkuu.

¹² Kettusen tutkimusaineistoa on säilytteinä Suomalais-Ugrilaisen Seuran tutkimusarkistossa Kansallisarkistossa. Kettusen luetteloimaton henkilöarkisto on arkistoitu Suomalaisen Kirjallisuuden Seuran arkistoon. Lisäksi Tarton Akateemisen Heimoklubin arkistossa Eesti Ajaloo Arhiivissa Tartossa ja Viron Kansallisarkistossa. Emakeele Seltsin arkistossa on Kettusen seuratoiminta-arkistoa, Kirjandusmuuseumin Kulttuurihistoriallisessa arkistossa puolestaan Kettusen kirjeenvaihtoa.

Lähteet

- Aarikka, Lotta 2018. Murteiden tutkimus väitöskirjojen viitteiden valossa. – *Virittäjä* 3, 387–422. <https://doi.org/10.23982/vir.65118>
- Aarikka, Lotta 2023. Murre ja sen tutkimus. Näkökulmia fennistisen murteentutkimuksen historiaan ja kieli-ideologioihin 1871–2017. Turku: Turun yliopisto. <https://urn.fi/URN:ISBN:978-951-29-9301-7>
- Braunmüller, Kurt, Juliane House 2009. Introduction. – *Convergence and divergence in language contact situation*. Eds. Kurt Braunmüller, Juliane House. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/hsm.8>
- Chambers, J. K., Peter Trudgill 2009. *Dialectology*. 2nd Edition. Cambridge Textbook in Linguistics. New York: Cambridge University Press.
- Diener, Alexander C., Joshua Hagen 2012. Borders and territory in the ancient world. – *Borders: A Very Short Introduction*. Eds. Alexander C. Diener, Joshua Hagen. New York: Oxford University Press. <https://doi.org/10.1093/actrade/9780199731503.003.0002>
- Ellison, Mark, T., Luisa Miceli 2017. Language monitoring in bilinguals as a mechanism for rapid lexical divergence. – *Language* 2, 255–287.
- Evans, Nicholas 2019. Linguistic divergence under contact. – *Historical Linguistics 2015. Selected papers from the 22nd International Conference on Historical Linguistics, Naples, 27–31 July 2015*. Eds. Michela Cennamo, Claudia Fabrizio. Amsterdam: John Benjamins, 563–592. (Current Issues in Linguistic Theory 348.) <https://doi.org/10.1075/cilt.348.26eva>
- Goebel, Hans 2010. Dialectometry and quantitative mapping. – *Language Mapping: An International Handbook of Linguistic Variation*. Volume 2. Eds. Alfred Lameli, Roland Kehrein, Stefan Rabanus. Berlin, New York: De Gruyter Mouton, 433–464. <https://doi.org/10.1515/9783110219166.1.433>
- Hakulinen Lauri 1952. Andrus Saareste 60-vuotias. – *Virittäjä* 3, 207–208.
- Hamunen, Markus, Reetta Konstenius 2020. Monimenetelmäisen kielentutkimuksen metateoreettisia lähtökohtia. – *Kielentutkimuksen menetelmiä I*. Toim. Milla Luodonpää-Manni, Markus Hamunen, Reetta Konstenius, Matti Miestamo, Urpo Nikanne, Kaius Sinnemäki. Helsinki: Suomalaisen Kirjallisuuden Seura, 105–157.
- Hinskens, Frans, Peter Auer, Paul Kerswill 2009. The study of dialect convergence and divergence: Conceptual and methodological considerations. – *Dialect change: Convergence and divergence in European languages*. Eds. Peter Auer, Frans Hinskens, Paul Kerswill. Cambridge: Cambridge University Press, 1–48. <https://doi.org/10.1017/CBO9780511486623.003>

- Honkola ym. 2018. Terhi Honkola, Kalle Ruokolainen, Kaj Syrjänen, Unni-Päivä Leino, Ilpo Tammi, Niklas Wahlberg, Outi Vesakoski. Evolution within a language: Environmental differences contribute to divergence of dialect groups. – *BMC Evolutionary Biology* 18, 132. <https://doi.org/10.1186/s12862-018-1238-6>
- Hovdhaugen ym. 2000 = Even Hovdhaugen, Fred Karlsson, Carol Henriksen, Bengt Sigurd. The history of linguistics in the Nordic countries. Jyväskylä: Gummerus.
- Hyvönen ym. 2007 = Saara Hyvönen, Antti Leino, Marko Salmenkivi. Multivariate analyses of Finnish dialect data: an overview of lexical variation. – *Literary and Linguistic Computing* 22, 271–290.
- Häkkinen, Kaisa 1993. Suomen kielen tutkimus kansallisen diktatuurin kourissa. – *Vaikea totuus. Vuosi 1918 ja kansallinen tie*. Toim Heikki Ylikangas. Helsinki: Suomalaisen Kirjallisuuden Seura, 18–45.
- Hänninen, Osmo, Lauri Koponen 1991. Alkulause. – *Palava henki ja kieli*. Professori Lauri Kettunen. Murteiden tutkija ja Eestin kansallisen kulttuurin esitaiteilija. Toim Osmo Hänninen, Lauri Koponen. Kuopio: Snellman-instituutti, 7–9. (Snellman-instituutin B-sarja 26.)
- Itkonen, Terho 1964. Proto-finnic final consonants. Their history in the Finnic languages with particular reference to the Finnish dialects. I:1. Introduction. The history of *-k* in Finnish. Helsinki: Suomalais-Ugrilainen Seura.
- Itkonen, Terho 1989. Nurmijärven murrekirja. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Karlsson, Fred 1994. Linguistics in the light of citation analysis. Helsinki: Helsingin yliopisto. (Yleisen kielitieteen laitoksen julkaisuja 23.)
- Kelomäki, Tapani 2009. Nuorgrammaattisuuden myytti suomen kielen tutkimuksessa. Helsinki: Helsingin yliopiston suomen kielen ja kotimaisen kirjallisuuden laitos.
- Kerminen ym. 2021 = Sini Kerminen, Nicola Cerioli, Darius Pacauskas, Aki S. Havulinna, Markus Perola, Pekka Jousilahti, Veikko Salomaa, Mark J. Daly, Rupesh Vyas, Samuli Ripatti, Matti Pirinen. Changes in the fine-scale genetic structure of Finland through the 20th century. – *PLOS Genetics* 17(3). <https://doi.org/10.1371/journal.pgen.1009347>
- Kettunen, Lauri 1930a. Suomen murteet. I. Murrenäytteitä. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Kettunen, Lauri 1930b. Suomen murteet. II. Murrealueet. Helsinki: Suomalaisen Kirjallisuuden Seura.

- Kettunen, Lauri 1932. „Suomen murteet II“. Prof. Martti Rapolan arvostelun johdosta. – *Virittäjä* 36, 185–198.
- Kettunen, Lauri 1940a. Suomen murteet. III. A. Murrekartasto. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Kettunen, Lauri 1940b. Suomen murteet. III. B. Selityksiä murrekartastoon. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Kettunen, Lauri 1945. Tieteen matkamiehenä. Kaksitoista ensimmäistä retkeä 1907–1918. Porvoo–Helsinki: Werner Söderström Osakeyhtiö.
- Kettunen, Lauri 1948. Tieteen matkamiehen uusia elämyksiä. Murrosvuodet 1918–1924. Porvoo–Helsinki: Werner Söderström Osakeyhtiö.
- Kettunen, Lauri 1960a. Kahdeksan matkaa Vermlannin metsäsuomalaisiin (1907–1937). Helsinki: Kauppakirjapaino.
- Kettunen, Lauri 1960b. Matkapakinoita ja muita muistelmia. 1925–1960. Helsinki: Jakaja.
- Kettunen, Lauri 1981. Suomen murteet. Murrekartasto. 4., supistettu painos. Helsinki: Suomalaisen Kirjallisuuden Seura. (Suomalaisen Kirjallisuuden Seuran toimituksia 188.) https://www.kotus.fi/kielitieto/murteet/suomen_murteet (23.5.2024)
- Kretzschmar, William, Edgar Schneider, Ellen Johnson (eds.) 1989. Computer methods in dialectology. – *Journal of English Linguistics* 21–22.
- Leino ym. 2006 = Antti Leino, Saara Hyvönen, Marko Salmenkivi. Mitä murteita suomessa onkaan? Murren sanaston levikin kvantitatiivista analyysistä. – *Virittäjä* 1, 26–45.
- Mansfield ym. 2023 = John Mansfield, Henry Leslie-O’Neill, Haoyi Li. Dialect differences and linguistic divergence. A crosslinguistic survey of grammatical variation. – *Language Dynamics and Change* 2, 232–276. <https://doi.org/10.1163/22105832-bja10026>
- Mielikäinen, Aila 2001. Murre-eroista ja isoglossien hierarkkisuudesta. – *Congressus nonus internationalis fenno-ugristarum 7.–13.8.2000 Tartu. Pars V. Dissertationes sectionum: Linguistica II*, 378–384.
- Murawaki, Yugo 2020. Latent geographical factors for analyzing the evolution of dialects in contact. – *Conference on Empirical Methods in Natural Language Processing*. November 16–20, 2020, 959–976. <http://doi.org/10.18653/v1/2020.emnlp-main.69>
- Norvik ym. 2022 = Miina Norvik, Yingqi Jing, Michael Dunn, Robert Forkel, Terhi Honkola, Gerson Klumpp, Richard Kowalik, Helle Metslang, Karl Pajusalu, Minerva Piha, Eva Saar, Sirkka Saarinen, Outi Vesakoski 2022. *Uralic*

- typology in the light of new comprehensive dataset. – *Journal of Uralic Linguistics* 1, 4–41.
- Penttilä, Aarni 1930. Ensimmäinen suomalainen murrelukemisto. – *Virittäjä* 34, 455–458.
- Penttilä, Aarni 1941. Suomalainen murrekartasto. – *Virittäjä* 45, 250–252.
- Pietikäinen, Sirkka 1991. Professori oppilaan näkökulmasta. – *Palava henki ja kieli*. Professori Lauri Kettunen. Murteiden tutkija ja Eestin kansallisen kulttuurin esitaistelija. Toim Osmo Hänninen, Lauri Koponen. Kuopio: Snellman-instituutti, 75–85. (Snellman-instituutin B-sarja 26.)
- Rantanen, Timo, Jenni Santaharju, Elina Salmela, Harri Tolvanen, Michael Dunn, Reima Välimäki, Unni Leino, Päivi Onkamo, Outi Vesakoski 2024. East-west division of genes, language and culture in Finland – drivers and hindrances. [Ilmumas.]
- Rapola, Martti 1932. Yleiskatsaus suomen murteisiin. – *Virittäjä* 32, 66–82.
- Rapola, Martti 1947. Johdatus suomen murteisiin. 1. painos. Helsinki: Suomalaisen Kirjallisuuden Seura. (Tietolipas 4.)
- Romano ym. 2022 = Noemi Romano, Peter Ranacher, Sandro Bachmann, Stéphane Joost. Linguistic traits as heritable units? Spatial Bayesian clustering reveals Swiss German dialect regions. – *Journal of Linguistic Geography* 1, 1–12. <https://doi.org/10.1017/jlg.2021.12>
- Räisänen, Alpo 1991. Murrekartaston tekijä kiertelemässä moottoripyörällä ympäri Suomea. – *Palava henki ja kieli*. Professori Lauri Kettunen. Murteiden tutkija ja Eestin kansallisen kulttuurin esitaistelija. Toim Osmo Hänninen, Lauri Koponen. Kuopio: Snellman-instituutti, 54–63. (Snellman-instituutin B-sarja 26.)
- Saarikivi, Janne 2020: Historiallinen ja vertaileva kielentutkimus. – *Kielentutkimuksen menetelmiä*. III. Toim Milla Luodonpää-Manni, Markus Hamunen, Reetta Konstenius, Matti Miestamo, Urpo Nikanne, Kaius Sinnemäki Helsinki: Suomalaisen Kirjallisuuden Seura, 679–720.
- Santaharju ym. 2024 = Jenni Santaharju, Terhi Honkola, Perttu Seppä, Kaj Syrjänen, Unni Leino, Outi Vesakoski. New version of the digitized Dialect Atlas of Finnish by Lauri Kettunen. – *Digital Humanities in the Nordic and Baltic Countries Publications*. [Ilmumas.]
- Séguy, Jean 1973. Atlas linguistique de la Gascogne. Vol. 6. Notice explicative. Paris: Centre national de la recherche scientifique.
- Suhonen, Seppo 1985. Hellittämätön tutkija ja tieteen matkamies. Sata vuotta Lauri Kettusen syntymästä. – *Virittäjä* 3, 253–269.

- Suhonen, Seppo 1991. Lauri Kettunen itämerensuomalaisten kielten tutkijana ja Tarton yliopiston opettajana. – Palava henki ja kieli. Professori Lauri Kettunen. Murteiden tutkija ja Eestin kansallisen kulttuurin esitaistelija. Toim Osmo Hänninen, Lauri Koponen. Kuopio: Snellman-instituutti, 40–53. (Snellman-instituutin B-sarja 26.)
- Syrjänen, Kaj 2021. Quantitative Language Evolution. Case Studies in Finnish Dialects and Uralic Languages. Tampere: Tampereen yliopisto.
- Syrjänen ym. 2016 = Kaj Syrjänen, Terhi Honkola, Jyri Lehtinen, Antti Leino, Outi Vesakoski. Applying population genetic approaches within languages: Finnish dialects as linguistic populations. – *Language Dynamics and Change* 6, 235–283. <https://doi.org/10.1163/22105832-00602002>
- Thomason, Sarah, Terrace Kaufman 1988. *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.
- Thomason, Sarah 2008. Social and linguistic factors as predictors of contact-induced change. – *Journal of Language Contact* 2, 42–56. <https://doi.org/10.1163/000000008792525381>
- Vares, Vesa, Kaisa Häkkinen 2001. *Sanan valta. E. N. Setälän poliittinen, yhteiskunnallinen ja tieteellinen toiminta*. Helsinki: WSOY.
- Vesakoski Outi 2009. Local and sexual divergence in host-use traits and anti-predator adaptations in the marine herbivore. *Idotea balthica*. Turku: Turun yliopisto.