

Toivola, M., P. Lintunen & L. M. Heikkola (toim.) 2024. Puheen tutkimuksen uusia suuntia – Aineistona vapaasti tuotettu puhe. *New directions in speech research – Freely produced speech as data.* AFinLA-teema / n:o 17, 5–21.

**Minnaleena Toivola**

Helsingin yliopisto

**Leena-Maria Heikkola**

UiT Norjan arktinen yliopisto  
Åbo Akademi

**Pekka Lintunen**

Turun yliopisto

## Vapaasti tuotettu puhe tutkimuksen kohteena

### Nostot

- Vapaasti tuotettua puhetta käytetään usein aineistona puheentutkimuksessa.
- Vapaasti tuotetun puheen ja kontrolloidun puheen välillä on jatkumo.
- Tässä johdannossa esittelemme teemanumeron artikkeleiden teemoja.



VERTAISARVIOITU  
KOLLEGIALT GRANSKAD  
PEER-REVIEWED  
[www.tsv.fi/tunnus](http://www.tsv.fi/tunnus)

## Abstract

Freely produced speech is a common source of data for speech research. This article is an introduction to this thematic volume and discusses freely produced speech from different perspectives. Freely produced speech is often seen as the opposite of controlled, often phonetic experiments in laboratory conditions. However, it is very difficult to determine which data collection methods are sufficient for freely produced speech to be considered authentic and natural. The topics of the articles included vary from dialectological research in Finland and the phonetic features of Namibian English to second language fluency studies as well as proficiency or speaker assessments in a second language. As this thematic issue will show, it is essential to conduct research on freely produced speech in a comprehensive way, as well as to understand the importance of the results obtained. The way freely produced speech is assessed can have a significant impact on an individual's life, for example, in gatekeeper evaluations.

**Keywords:** speech, speech research, language learning, language use

**Asiasanat:** puhe, puheentutkimus, kielen oppiminen, kielen käyttö

## 1 Johdanto

Mitä on vapaasti tuotettu puhe? Sen voi määritellä ainakin osittain olevan suunnittelematonta, dynaamista ja tilanteista. Ääneen luettuun tekstiin verrattuna vapaasti tuotettuun puheeseen liittyy enemmän kielellisestä prosessoinnista johtuvia tekijöitä kuten epäröintiä tai sujuvuuden vaihtelua. Kun keskustelemme toistemme kanssa arkipäivän tilanteissa, usein dialogisesti ja multimodaalisesti, puhumme yleensä kiinnittämättä tietoisesti huomiota esimerkiksi siihen, miten tuotamme ääntä fysiologisesti. Emme myöskään välttämättä ole tietoisia siitä, miten alitajuisesti arvioimme keskustelukumppanin kielitaitoa ja miten olemme vuorovaikutuksessa sellaisten puhujien kanssa, joiden tuotoksessa kuuluu omasta puheestamme poikkeavia piirteitä kuten vierasta aksenttia. Tutkimuksen näkökulmasta on erityisen kiinnostavaa tarttua juuri tähän suunnitteleamattomuuteen ja pyrkiä vangitsemaan ohimenevä hetki, puhetilanne, tarkastelun kohteeksi.

Tässä julkaisussa tuomme esiin näkökulmia vapaasti tuotettuun puhuttuun kieleen tutkimuskohteena. Pyysimme tutkijoita Puheen ja puheentutkimuksen teemaverkoston kirjoituskutsussa pohtimaan vapaasti tuotetun puheen tutkimuksen mahdollisuuksia. Halusimme rajata teemaksi puheen erityisesti siinä muodossa, että tutkittava aineisto olisi mahdollisimman vähän kontrolloitua, jolloin päästään lähemmäksi autenttista kielenkäyttöä, kuten spontaania vuorovaikutusta, puhekielisiä ilmaisuja ja keskustelun dynaamista luonnetta. Kontrolloiduilla koeasetelmilla on puheen- ja kielentutkimuksessa kuitenkin nyt ja jatkossa oma merkityksensä, joka on luonnollisesti yhteydessä tutkimuksen tarkoitukseen. Käsittelemme eri aineistonkeruumenetelmien käyttöä myöhemmin tässä johdannossa.

Puheen tutkimusta, erityisesti suullista kielitaitoa ja ääntämistä, on tarkasteltu aiemmassa Kurosen ym. (2017a) toimittamassa AFinLA-teeman numerossa 10. Tässä teemanumerossa keskityimme osin samoihin aihepiireihin, kuten toisen tai vieraan kielen taitoon ja sen arviointiin, puheen sujuvuuteen, kielenvaihtoon sekä ensikielen vaikutukseen toisen kielen tuottamisessa. Kurosen ym. (2017a, ks. myös Kuronen ym. 2017b) julkaisussa keskityttiin erityisesti tutkimusasetelmiin, joissa kohteena oli toisen tai vieraan kielen oppiminen. Myös monet tässä teemanumerossa esitetyt tulokset ovat sovellettavissa kielenoppimiseen ja -opetukseen.

Käsillä oleva teemanumero sisältää tämän johdannon lisäksi kahdeksan tutkimusartikkelia, joissa vapaasti tuotettu puhe on ollut tutkittavana aineistona. Eri tilanteissa äänitettyä puheaineistoa ovat tutkineet seuraavat kirjoittajat:

- Haapanen, Saloranta, Peltola, Tamminen, Uwu-khaeb ja Peltola
- Lehtilä, Peltonen ja Lintunen
- Olkkonen, Mutta ja Lintunen,
- Peltonen ja Lintunen
- Ullakonoja

Aarikka sen sijaan käsittelee aihepiiriä toisesta näkökulmasta: hän tarkastelee, miten käsitteeseen vapaa puhe on suhtauduttu murteentutkimuksen perinteessä. Heikkolan, Toivolan, Kimin ja Salorannan tutkimuksen kohteena ovat puolestaan käsitykset tilanteista, joissa tuotetaan vapaata puhetta keskusteluissa ei-ensikielisen suomenpuhujan kanssa. Lisäksi von Zansen ja Kallio esittelevät kielenopetukseen suunnitellun DigiTala-työvälineen, jonka avulla on mahdollista arvioida puhujan kielitaitoa suomeksi tai ruotsiksi vapaasti tuotetussa puheessa.

Edellä kuvatut, käytettyjen tutkimusaineistojen mukaan jaotellut artikkelit muodostavat kolme kokonaisuutta, joita esittelemme tässä johdannossa seuraavasti: Esittelemme luvussa 2 Aarikan tutkimuksen, joka käsittelee puhuttujen aineistojen määritelmiä julkaisun osassa I sekä käymme läpi suhtautumista eri tavoin tuotettuun puheeseen, puheaineistojen laatuun ja määrään sekä siihen, miten hyvin niiden ajatellaan kuvaavaan tutkimuksessa ”aitoa” puhetta. Luvussa 3 esittelemme artikkelit seuraavilta kirjoittajilta:

- Ullakonoja
- Haapanen, Saloranta, Peltola, Tamminen, Uwu-khaeb ja Peltola
- Zansen ja Kallio
- Heikkola, Toivola, Kim ja Saloranta.

Nämä muodostavat julkaisun osan II, jonka yhdistävänä teemana on vapaasti tuotettu puhe kielenoppimisessa ja sen arvioinnissa sekä vuorovaikutuksessa.

Näiden artikkelien yhdistävänä teemana on vapaasti tuotettu puhe kielenoppimisessa ja sen arvioinnissa sekä vuorovaikutuksessa.

Luvussa 4 käsitellään vieraan kielen oppimiseen liittyvää puheen sujuvuutta eri näkökulmista. Se muodostaa julkaisun osan III, joka sisältää Lehtilän, Peltosen ja Lintusen; Olkkosen, Muttan ja Lintusen sekä Peltosen ja Lintusen artikkelit.

## 2 Tutkimusta puheesta

### 2.1 Vapaasti tuotettu puhe aineistona

Puheaineistoja käytetään monilla, erityisesti humanistisilla, mutta myös muilla tieteenoaloilla, joilla työskennellään esimerkiksi haastattelujen ja kenttätöiden parissa erilaisten kieleen ja puheeseen liittyvien ilmiöiden dokumentoimiseksi. Kielitieteilijöistä murteentutkijoiden ja foneetikkojen voidaan sanoa olleen ensimmäisiä, jotka ovat käyttäneet äänitettä aineistoja. (ks. Mielikäinen 2017). Murteentutkimuksessa pyrittiin dokumentoimaan katoavaa kansankielistä puhetta, ja foneetikot tarvitsivat puheaineistoja voidakseen analysoida puhetta ja tehdä akustisia analyysejä myös näönvaraisesti. (Calamai & Frontini 2018.) Kurki ja Mustajoki (2019) nostavat keskiöön puhutun kielen ja variaation tutkimuksen fennistisessä kielentutkimuksessa. Jo 1800-luvun lopusta lähtien suomalaisessa murteentutkimuksessa on kerätty puhuttuja aineistoja haastattelumenetelmin (Kurki & Mustanoja 2020). He kuitenkin muistuttavat, että haastattelemalla on kerätty vuosikymmenten aikana hyvin monenlaisia murreaineistoja ja sanaa *haastattelu* on käytetty yleisnimityksenä, vaikka haastatteluiden sisältämässä keskustelupuheessa on voinut olla hyvin monitasoista variaatiota.

Tämän teemanumeron ensimmäisessä tutkimusartikkelissa Lotta Aarikka käsittelee vapaasti tuotetun puheen käsitettä ja sen määrittelyä rajatussa ja historiallisessa kontekstissa. Hänen vuonna 2023 ilmestyneen väitöstutkimuksensa aineistosta nousi esiin, että Suomessa vuosina 1871–2017 julkaistuissa murreaiheisissa väitöskirjoissa tavoiteltavaa puheaineistoa kuvattiin ”luonnolliseksi, luontevaksi, tuttavalliseksi tai spontaaniksi arkikieleksi”. Puheaineistojen kuvailun tarkempi käsittely rajautui pois hänen väitöskirjatutkimuksestaan (Aarikka 2023), mutta tämän teemanumeron julkaisussa Aarikka paneutuu aiheeseen syvemmin.

Murteentutkimuksessa aineistonkeruu on tyypillisesti perustunut tilanteisiin, joissa puhujan on toivottu tuottavan puhetta mahdollisimman vapaasti. Vaikka aineistoa on aiemmin pidetty vapaasti tuotettuna puheena, sen aitoutta ja institutionaalisuutta on myöhemmässä tutkimuksessa alettu pohtia tarkemmin. Analyysissään suomen murteita käsittelevistä väitöskirjoista Aarikka pohtii vapaan puheen käsitettä ja sen määritelmiä muun muassa seuraavista näkökulmista: Onko puheen tuottotapa provosoitu vai provosoimaton? Miten kielentutkimuksen menetelmät soveltuvat tiettyyn puhetilanteeseen? Vapaan puheen määrittely ei ole yksinkertaista, kuten Aarikka artikkelissaan toteaa. Selkeimpinä vastakohtina vapaalle puheelle hän nostaa esiin

lukupuhunnan ja kehyslauseisiin asetetut minimiparit. Kirjoittaja pohtii myös, miten vapaan puheen määritelmä vaikuttaa tutkimuskysymyksiin, tutkimuskohteiden määrittelyyn ja siten tutkimustuloksiin. Toisaalta Aarikka nostaa esiin, että pelkkä vapaaseen ja epämuodolliseen puheeseen keskittyminen ei ole riittävää, vaan myös muodollisia puhetilanteita tulisi tutkia lisää. Lisäksi Aarikka ehdottaa, että tulevaisuudessa tutkimus voisi kohdistua erityisesti arjessa käytettävään kieleen.

Suomenkielisen keskustelupuheen fonetiikkaa on ensimmäisten joukossa tutkinut Ogden (2001), joka käytti tutkimusmateriaalina radiosta äänitettyä puhetta. Ohjelman kuuntelijat soittivat lähetykseen ja keskustelivat juontajan kanssa, joten käytetty kieli vastasi hyvin luonnollista keskustelupuhetta. Haittapuolena tällaisissa äänityksissä on aineiston huono laatu, mikä vaikeuttaa erityisesti akustisen analyysin tekemistä. Lennes ym. (2006) sen sijaan tutkivat erittäin korkealaatua aineistoa, joka oli äänitetty Teknillisen korkeakoulun akustiikan ja audiosignaalien käsittelyn laboratorion suuressa, kaiuttomassa tilassa. Tilanteessa jäljiteltiin puhelinkeskustelua, sillä puhujat kuulivat sekä oman että toisen puhujan äänen kuulokkeilla, mutta eivät nähneet toisiaan. Kukin osallistujapari tunsu toisensa hyvin ja sai keskustella tilanteessa vapaasti. Voidaan kuitenkin kysyä, minkälainen merkitys tutkimusjärjestelyillä ja poikkeavalla tilalla oli puheen tuottoon. Lenneksen ym. (2006) mukaan aineisto oli runsasta ja puhujat keskustelivat vilkkaasti äänityksen aikana, joten on mahdollista, että oudossakin tilanteessa voidaan tuottaa spontaania puhetta, joka ulkopuolisen korvaan kuulostaa ”normaalilta” tai ”aidolta”. Teknisesti korkealaatuisen aineiston suurin hyöty näkyy luotettavuuden paranemisena, kun puheesta voidaan tehdä tarkkoja akustisia analyysejä myös automaattisia menetelmiä käyttämällä. Huonompilaatuisessa aineistossa esiintyvää taustahälyä poistettaessa saatetaan samalla hävittää puhe-signaaliin oleellisesti kuuluvia piirteitä, mikä mahdollisesti vaikuttaa heikentävästi analyysin onnistumiseen. Puheaineiston laatu voi olla hyvinkin vaihtelevaa, mutta oleellista on, että tutkijat ovat aineistoa hankkiessaan selvillä siitä, kuinka hyvin se sopii tutkimustarkoitukseen.

Erityisesti sosiofoneettisessa tutkimuksessa ollaan kiinnostuneita erilaisista tutkimustarkoituksiin suunnitelluista keskustelutehtävistä, joiden avulla puheesta saadaan elisitoitua tiettyjä tutkimuksen kohteeksi haluttavia piirteitä muutoin kuin ääneen lukemalla. Elisitoidun puheen voidaan ajatella olevan ainakin lähellä spontaania keskustelupuhetta. Esimerkiksi yleisesti käytetyssä Brownin ym. (1983) luomassa karttatehtävässä (*maptask*) saadaan elisitoitua tietynlaista keskustelua, kun toinen osallistuja antaa ohjeita hänellä karttaan merkityn reitin perusteella ja toinen henkilö pyrkii löytämään omasta kartastaan saman reitin ja määränpään. Sittemmin kyseistä karttatehtävää on muokattu monella tavalla. Sen uusimmat versiot ovat digitaalisia, ja sen tutkimuskäyttöä on laajennettu myös kirjoitetun viestinnän puolelle, kun osallistujat ovat voineet kommunikoida toistensa kanssa myös muuten kuin puhumalla (ks. esim. Berríos ym. 2023).

Myös muunlaista kuin vapaasti tuotettua aineistoa tarvitaan osana foneettista tutkimusta, sillä intra- ja intersubjektivinen vertailu puheen segmentaalisisissa ja toisaalta myös prosodisissa piirteissä olisi käytännössä erittäin vaikeaa, jos aineistona olisi vain spontaania puhetta. Erilaisia luettuja tekstejä käytetäänkin edelleen foneettisessa tutkimuksessa, sillä niiden etuna on hyvä kontrolloitavuus sekä vertailtavuus. Yhtenäisten materiaalien käyttö lukupuhunnassa myös kansainvälisesti (ks. esim. Deterding 2006; Baird ym. 2022) tekee tutkimustulosten vertailusta parempaa. Kontrolloidun laboratoriopuheen puolesta on puhunut erityisesti Xu (2010). Hänen mielestään manipuloimalla puheesta tutkittavia tekijöitä ja pitämällä muut tekijät vakiona voidaan keskittyä tiettyihin ja ennalta määrättyihin puheen piirteiden havaitsemiseen. Vaihtelu puheessa voidaan tällöin suoraan liittää manipuloituihin tekijöihin. Xu (2010) esittää omana näkemyksenään, että voidaksemme edistyä sen ymmärtämisessä, mitä puhe puhe on ja mistä se koostuu, pitää kaiken foneettisen puheetutkimuksen nojata vahvasti kontrolloituun puheeseen, jonka avulla voidaan testata hypoteeseja kiinnostavista puheen ilmiöistä.

Foneettisen tutkimuksen kohteena ovat usein puheen piirteet, jotka esiintyvät segmentaalisisella tai suprasegmentaalisisella tasolla, mutta voidaan kysyä, riittävätkö ne kuvaamaan vapaasti tuotettua puhetta ja siihen liittyvää tilanteisuutta riittävästi. Muun muassa Klaus J. Kohler (2000) on tuonut jo yli kaksi vuosikymmentä sitten voimakkaasti esiin, että foneettisten peruspiirteiden tutkimisen lisäksi tulisi huomioida myös erilaiset puhetyylit sekä puhujien mukautuminen viestintätilanteeseen. Lisäksi yksittäisen piirteiden tutkiminen osana kehysanaa ei välttämättä ole sovellettavissa tuottamiseen ja havaitsemiseen edes lausetasolle, saati laajemmin (Kohler 2000). Tällä hetkellä foneettisessa tutkimuksessa ollaan puheentuoton ja -havaitsemisen lisäksi kiinnostuneita myös puheeseen liittyvistä multimodaalisista elementeistä, kuten eleistä, ilmeistä ja kehon asennosta osana viestintää (ks. esim. Danner ym. 2018).

Käsitys siitä, nimitetäänkö eri tavoin elisitoitua aineistoa osittain vapaasti tuotetuksi, semispontaaniksi tai jollain muulla tavoin määritellyksi puheeksi, on pitkälti mielipidekysymys. Voidaan myös kyseenalaistaa keskustelu siitä, onko tiukasti kontrolloitu ja laboratorio-olosuhteissa tuotettu puhe huonompaa vai parempaa kuin vapaasti tuotettu puhe, sillä niiden välinen vertailu on ongelmallista. Jossain määrin tätä vastakkainasettelua voidaan tutkia tietyn menetelmin tuotetuista puheaineistoista. Uusinta puheaineistojen vertailevaa tutkimusta puheen prosodian alueelta ovat tehneet de Ruiter (2015) sekä vokaaleista Hall-Lew ja Boyd (2020). De Ruiter (2015) on tutkinut saksankielistä kuvakerrontaa (jota hän tutkimuksessaan nimittää spontaaniksi puheeksi) ja luettua puhetta. Luettava teksti oli muodostettu samojen puhujien tuottaman, ensimmäisen tutkimusvaiheen kuvakerronnan perusteella. Intonaatio oli tulosten mukaan kuvakerronnassa ja sitä vastaavassa luetussa puheessa selvästi erilaista. (De Ruiter 2015.) Hall-Lew ja Boyd (2020) vertailivat, millaisia akustisia mittaustuloksia saadaan englannin vokaaleista, kun aineisto kerättiin samoilta puhujilta haastatteleamalla, elisitoitintehävän avulla sekä tekstejä ääneen lukemalla.

Tulosten mukaan vokaalien laadussa oli vain vähän eroja erityisesti elisointitehtävän ja haastattelussa tuotetun puheen välillä. Lisäksi haastattelupuheesta ja ääneen luetusta puheesta mitatut vokaalien formanttien erot olivat pienempiä kuin tutkijat olivat odottaneet. (Hall-Lew & Boyd 2020.) Tuloksiin on kuitenkin syytä suhtautua varauksella, sillä aineisto kerättiin vain muutamalta puhujalta ja se on sovellettavissa lähinnä yksittäisten äänteiden tasolla.

Akustisten analyysien osoittamien mittaustulosten lisäksi eri tavoin tuotetusta puheesta voidaan tehdä myös kuulonvaraista, vertailevaa tutkimusta. Vakkilaisen ja O'Dellin (2022) tutkimus osoitti, että suomenkielisen sanan lopussa olevan vokaalin ei tarvitse olla yhtä pitkä spontaanissa puheessa kuin lukupuhunnassa tullakseen havaituksi pitkänä. Tulosta pitää tarkastella suuntaa antavasti, sillä tutkimuksessa ei varsinaisesti pyritty vertaamaan, miten kuulija erottaa lukupuheen spontaanista puheesta. Tutkimusta eri puheaineistojen välisistä eroista tarvitaankin lisää, mutta sitä rajoittavat jo vertailtavuuteen itseensä liittyvät haasteet. Wagner ym. (2015) ovat ehdottaneet, että puheilmioita tutkittaessa olisi tärkeää lisätä tietoa erilaisista menetelmistä, kerätä runsaasti aineistoa erilaisilla menetelmillä ja pyrkiä ymmärtämään syvemmin niiden vaikutus saatujen tulosten tulkinnassa. Tämä edistäisi paremmin ymmärrystä puheesta ja sen tutkimisesta kuin argumentointi tietynlaisen puheaineiston puolesta tai sitä vastaan.

Nykyään spontaanin keskustelupuheen äänittäminen on teknisesti huomattavasti helpompaa kuin vuosikymmeniä sitten, mutta aineistojen kerääminen on yhä edelleen vaativaa ja siihen liittyy myös monia eettisiä kysymyksiä. Euroopan Unionin tietosuoja-asetus (General Data Protection Regulation; Yleinen tietosuoja-asetus (GDPR) 2016) rinnastaa puheäänän suoraksi tunnustetiedoksi, kuten nimen tai osoitteen, jolloin sen kerääminen on tarkkaan säädeltyä. Eettisesti voidaan äänittää vain luvanvaraista puhetta, joten voidaankin kysyä, onko tutkimukseen saatava aineisto koskaan täysin spontaania, vapaata puhetta vai liittyykö tilanteeseen aina ns. ”tutkimuksen maku”. Käytännössä puhuja voi kuitenkin nopeasti unohtaa, että häntä äänitetään.

## 2.2 Vapaasti tuotetun puheen määrä tutkimuksessa

Vaikka eri tavoin tuotettua puhetta on nykyään saatavilla myös erilaisten avointen puhekorpusten (ks. esim. Kielipankki) kautta helpommin kuin aiemmin, on silti epäselvää, kuinka paljon vapaasti tuotettua puhetta tutkitaan. Tähän ei ole löydettävissä kovin paljon tutkimustietoa, mutta esimerkiksi Wagner ym. (2015) tarkastelivat vuosien 2007 (Trouvain ja Barry 2007) ja 2011 (Lee ja Zee 2011) fonetiikan alan suurimman kansainvälisen konferenssin (International Congress of Phonetic Sciences, ICPhS) aiheita. Tarkastelu osoitti, että kontrolloimattomaan puheaineistoon perustuvien tutkimusten määrä oli kasvanut vuodesta 2007 vuoteen 2011, mutta se oli yhä huomattavasti pienempi kuin kontrolloitujen aineistojen määrä.

Tarkastelimme vastaavalla tavalla, vaikkakin suppeammin, kuinka paljon ja minkälaisia puheentutkimukseen liittyviä aineistoja oli mukana AFinLAN syysseminaarissa vuonna 2022, jossa myös esiteltiin suurin osa tämän julkaisun artikkeleiden aiheista. Kävimme ensin läpi kaikki symposiumin tiivistelmien otsikot (145) puheen- ja kielentutkimukseen liittyvillä suomen-, englannin- ja ruotsinkielisillä lyhennetyillä hakusanoilla: *puh, kesk, vuor, suull, ään, haas, spe, con, flu, inter, dia, språk* ja *rör*. Otsikoista 34 valikoitui sellaisiksi, joissa esiintyi jokin lyhennetyistä hakusanoista. Tämän jälkeen tarkistimme tiivistelmistä, oliko kyseisten tutkimusten aineistona ollut puhetta. Näistä osa jäi abstraktin perusteella epäselväksi tai aineisto perustui johonkin muuhun kuin puheeseen (mutta saattoi olla myös metapuhetta puheesta). Poistimme epäselvät tapaukset, jolloin jäljelle jäi 24 tutkimusta, joissa aineistona oli puhetta, joka oli tuotettu muutoin kuin luettuna. Tiivistelmissä tuotiin selvennyksenä myös lähes aina esiin, että puheaineisto oli jotenkin tallennettua. Tämän jälkeen kävimme kaikki loput symposiumin abstraktit läpi, ja niiden joukosta löytyi vielä 41 muuta tutkimusta, joissa oli mainittu puheaineistot ja menetelmänä oli käytetty esimerkiksi etnografista tutkimusta. Lisäksi aineistona oli muuta havainnointia ja keskusteluiden seuraamista sekä YouTubea katsottuja videoita. Osassa näistä tiivistelmistä tuotiin selvästi esiin, että tutkimuksessa käytetty puhuttu aineisto oli äänitetty tai videoitu tutkimusta varten. Lisäksi yhdestätoista muusta abstraktista ei saanut täysin selvää, mikä aineisto oli kyseessä, mutta tutkimus perustui mahdollisesti puheeseen.

Kaikissa tutkimuksissa, joissa tuotiin selvästi esille puheaineiston käyttö, koostui aineisto selkeässä enemmistössä tutkimuksista vapaasti tuotetusta puheesta. AFinLAN vuoden 2022 syysseminaarissa esityksiä tarkastellessa voi ainakin 65 (45 %) esitelmän sanoa perustuneen eri tavoin tuotettuun puheaineistoon. Saatu tulos osoittaa, että erilaisten puheaineistojen käyttö, erityisesti vapaasti tuotettu puhe, on yleistä soveltavan kielitieteen tutkimuksissa Suomessa.

### 3 Kielitaidon vai puhujan arviointia?

Usein kielitaidon arvioinnin tarkoitus on varmistua siitä, että kielenoppija on saavuttanut vaadittavan taidon johonkin ammattiin tai opiskeluun päästäkseen. Tällaista kielitaidon arviointia voidaan kutsua portinvartija-arvioinniksi (ks. tarkemmin Huhta ja Hilden 2016), josta esimerkkinä Suomessa ovat YKI-kokeet eli yleiset kielitutkinnot. Vaikka portinvartija-arvioinnissa pyritään objektiivisuuteen, yksittäisten arvioijien näkemyksillä voi olla yllättävän suuri rooli. Näkemyksiin voivat vaikuttaa myös muut kuin varsinaiseen kielitaitoon liittyvän arvioinnin ulkopuoliset tekijät, kuten puhujan ensikielen tunnistaminen. Aholan (2020) tutkimuksessa havaittiin, että yleiseen kielitutkintoon osallistuvien thainkielisten puhujien suorituksia arvioitiin ankarammin ääntämisen ja rakenteiden osalta, jos arvioijat tunnistivat puhujien olevan thainkielisiä.



Toisaalta thainkieliseksi tunnistaminen vaikutti siihen, että heille annettiin parempia arvioita sujuvuudesta ja yleiskriteerin hallinnasta (Ahola 2020).

Riikka Ullakonoja puolestaan tutkii artikkelissaan yleisen kielitutkinnon arvioiteja ääntämisen näkökulmasta. Tutkimuksen aineisto on kerätty kyseisen tutkinnon suomen kielen kokeeseen osallistuneilta venäläistaustaisilta puhujilta ja heidän puhutuoksiaan kuuntelevilta yleisen kielitutkinnon koulutetuilta arvioijilta. Ullakonoja keskittyy erityisesti siihen, millä perusteella puhujan ensikieli tunnistetaan pelkästään puhujan toisen kielen puheesta. Tällöin tunnistamiseen voivat vaikuttaa hyvinkin pienet vierasta aksenttia indikoivat puheen ominaisuudet. Tutkimukseen osallistuneet kuuntelijat tunnistivat puhujan ensikielen tiettyjen suomen kielen ääntämiseen liittyvien piirteiden perusteella. Tulos vahvistaa aiempia havaintoja siitä, että venäjänkielisille puhujille on haastavaa tuottaa suomen vokaalien kestoja sekä erityisesti suomen konsonantteja /l/ ja /h/. Lisäksi venäjän kieleen kuuluva palatalisaatio ja suomen kielestä poikkeava intonaatio liitetään tutkimuksessa siihen, että puhujan äidinkielenä tunnistetaan venäjä. Ullakononon tutkimus ei kuitenkaan vahvista aikaisempien tutkimusten havaintoja, joiden mukaan painottomien vokaalien redusoituminen ja narinan puuttuminen olisivat erityisesti venäläisen aksentin piirteitä. Sen sijaan tutkimuksessa nousee esiin mahdollisia uusia venäläisen aksentin erityispiirteitä sibilanteissa ja tremulanteissa, jotka saivat arvioijilta mainintoja niiden suomesta poikkeavan artikulaation vuoksi. Ullakonoja myös muistuttaa, että lyhyetkin spontaanit puhenäytteet riittävät vieraan aksentin kuulonvaraiseen arvioimiseen, vaikka syvälliseen ja monitahoiseen foneettiseen analyysiin ne saattavat olla liian suppeita. Tutkimuksesta saatuja tuloksia voidaan hyödyntää esimerkiksi venäjän- ja ukrainankielisten oppijoiden suomen kielen opetuksessa.

Ääntämisen arviointi on tärkeä osa kielitaidon portinvartija-arvioinneissa. Vieraaseen aksenttiin suhtautumisella voi olla myös sosiaalisia vaikutuksia, erityisesti silloin, jos kielenoppija kokee, että hänen ääntämiseensä suhtaudutaan kielteisesti. Ääntämistä pidetään osana identiteettiä, ja siksi se onkin kielen oppimisessa monimutkainen tekijä (Setter & Jenkins 2005). Jos kielenoppijan tavoitteena on saavuttaa kohdekieleessä hyvä ääntämistaito, on oleellista tietää, mihin piirteisiin kielenopetuksessa tulisi keskittyä. Tämän selvittämiseksi tarvitaan tutkimusta eri kohdekielten ja lähtökielten puhujien tuottamasta ääntämisestä. Katja Haapanen, Antti Saloranta, Kimmo Peltola, Henna Tamminen, Lannie Uwu-khaeb ja Maija Peltola tutkivat artikkelissaan, kuinka khoekhoegowabia ensikielenään puhuvat tuottavat soinnillisuutta sananalkuisissa klusiilikonsonanteissa vapaasti tuotetussa Namibian englannissa. Soinnillisuutta tarkastellaan akustisesti mitattavien VOT-arvojen perusteella vertaamalla niitä tyypillisiin englannin ja afrikaansin tuottomalleihin. Haapanen ym. osoittavat, että khoekhoegowabia ensikielenään puhuvat tuottavat Namibian englannissa selkeitä, britti-englannin kielen mallin mukaisia klusiileja, joiden sointikontrastit eivät heijasta afrikaansin ja khoekheongowabin malleja. Jos puhuja ei pysty tuottamaan soinnillisia ja soinnittomia klusiileja englannin mallin mukaisesti, se voi johtaa vääri-

nymmäryksiin yksittäisten foneemien tasolla, koska soinnillisuus voi erottaa sanojen merkityksiä toisistaan. Kirjoittajat ehdottavat, että tutkimuksen tulokset voivat auttaa kehittämään foneettisia harjoitusmalleja englannin kielen opetukseen Namibiassa. Koska tutkimuksessa on tarkasteltu sointia vain sananalkuisesti, nostavat Haapanen ym. esiin, että sitä tulisi tarkastella myös muissa foneettisissa konteksteissa ja eri ensikielten puhujilla.

Koska kielitaidon arviointiin voivat vaikuttaa monet muutkin tekijät kuin varsinainen puhetuotos, on tärkeää kehittää ihmisten tekemien arvioiden tueksi uusia menetelmiä, jotta arvioinnista saataisiin neutraalimpaa sekä tarvittaessa nopeampaa. Anna von Zansenin ja Heini Kallion artikkelissa esitellään puheentunnistus- ja arviointiteknologiaa käyttävää DigiTala-työkalua, jota voidaan käyttää apuna kielitaidon arvioinnissa. Työkalu tarjoaa suomen ja suomenruotsin oppijoille mahdollisuuden harjoitella ääneen lukemista sekä spontaania puhetta, joista se tekee automaattista arviointia. Oppijalta äänitetään kohdekielellä puhetta, josta saadaan työkalun avulla palautetta kielitaidon eri osa-alueista sekä kootusti yleisen kielitaitotason arvio. Palaute voi auttaa kielenoppijaa tunnistamaan omat kehitystarpeensa. Von Zansen ja Kallio pohtivat automaattisen arvioinnin mukanaan tuomia hyötyjä erityisesti portinvartijatilanteissa, kuten kansalaisuuteen ja työnhakuun liittyvissä arvioinneissa, ja tuovat myös esiin tekoälyyn pohjautuvan arvioinnin rajoitukset erityisesti pienissä kielissä kuten suomessa ja suomenruotsissa. Tekoälypohjaisen arvioinnin tueksi tarvitaan laaja puhetietokanta, joka edustaisi kaikkia kielitaitotasoja ja kielenoppijoiden taustoja, eikä aineisto olisi vinoutunut. Tällaisen aineiston kerääminen on haasteellista. Lisäksi ohjelmistoa tulisi kehittää, ja eri käyttäjäryhmät tulisi osallistaa kehitystyöhön. Kirjoittajien mukaan olisi myös tärkeää pohtia arvioiden luotettavuutta, validiteettia ja oikeudenmukaisuutta automaattista arviointia käytettäessä. Von Zansen ja Kallio ehdottavat tuleviksi kehittämiskohteiksi dialogipuheen ja ei-sanallisen kommunikation sisällyttämisen automaattiseen kielitaidonarviointiin.

Spontaaneissa keskusteluissa tapahtuu monia asioita samanaikaisesti. Oman puheen monitoroinnin lisäksi kiinnitämme keskusteluissa hyvin vaihtelevasti huomiota muihin seikkoihin. Esimerkiksi voimme tehdä oletuksia siitä, mitä kieltä keskustelukumppanimme haluaa puhua tai voimme muuttaa omaa tapaamme puhua tekemiemme havaintojen tai oletusten mukaan. Leena Maria Heikkolan, Minnaleena Toivolan, Jeongdo Kimin ja Antti Salorannan artikkelissa tarkastellaan ensikielisten suomenpuhujien näkemyksiä vapaasti tuotetusta puheesta keskusteluissa puhujan kanssa, jonka ensikieli on joku muu kuin suomi. Tutkimuksen tavoitteena on selvittää, mitä kieltä ei-ensikielisen suomenpuhujan kanssa puhutaan ja vaihtuuko keskustelun kieli englanniksi, vaikka aloitus olisi tapahtunut suomen kielellä. Kyselytutkimuksessa saatujen tulosten mukaan ensikieliset suomenpuhujat ilmoittivat aloittavansa keskustelut oletetun ei-ensikielisen suomenpuhujan kanssa suomeksi, jos keskustelu käytiin suomenkielisessä ympäristössä. Heikkolan ym. artikkelissa kiinnitetään huomiota siihen, että heidän tuloksensa poikkeaa aiemmasta tanskalaisesta tutkimuksesta:

Jespersen ja Hejnán (2021) mukaan ei-ensikieliset tanskanpuhujat ilmoittivat ensikielisten tanskanpuhujien aloittavan keskustelun hyvin usein englanniksi. Heidän tutkimuksessaan ei-ensikieliset tanskanpuhujat pitivät keskustelun aloittamista englanniksi tai kielenvaihtamista englanniksi toiseuttavana. Samalla he kokivat tulevansa luokitelluksi ”toisiksi” vieraan aksenttinsa tai ulkonäkönsä vuoksi. Heikkolan ym. tutkimuksessa ensikieliset suomenpuhujat olettivat, että ei-ensikieliset suomenpuhujat osaavat englantia paremmin kuin suomea, ja yli puolet kyselyyn vastaajista ilmoitti vaihtavansa kielen suomesta Englantiin kommunikaation sujuvoittamiseksi. Ensikieliset suomenpuhujat myös olivat sitä mieltä, että ei-ensikieliset suomenpuhujat reagoivat kielenvaihtamiseen positiivisesti ja näkivät sen kohteliaana ja hyödyllisenä. Tutkimuksen tulokset korostavat tarvetta lisätä tietoisuutta omien kielivalintojen ja kielenvaihtamisen seurauksista sekä suomen kielen puhumisen tärkeydestä kielenoppimisen ja osallisuuden mahdollistajana. Heikkola ym. esittävät, että jatkossa kielivalintoja ja kielenvaihtamista sekä näiden syitä tulisi tutkia myös ei-ensikielisten suomenpuhujien näkökulmasta.

Luvussa 3 esittelemämme tutkimukset osoittavat, että on tärkeää tutkia havaitsemiseen liittyviä, usein tiedostamattomia tekijöitä, jotka saattavat vaikuttaa toimintaamme erityisesti puhujien kanssa, jotka oppivat ensikieltämme. Vieraskielisyys on ilmiö, jonka ensikieliset puhujat havaitsevat herkästi, joskus hyvin pienistäkin ääntämisen piirteistä segmenttaalisella tasolla. Toisaalta yhtä tarkasti ei välttämättä tunnusteta puhujan ensikieltä tai sen tunnustaminen aiheuttaa ennako-oletuksia, millä taas voi olla laajempia merkityksiä esimerkiksi kielitaitoarvioinneissa. Onkin tärkeää tiedostaa, miten vieraskielisyys vaikuttaa ja miten suomenoppijoiden kielitaitoa arvioidaan tai olemmeko valmiita puhumaan omaa ensikieltämme vieraskieliseltä kuulostavan puhujan kanssa.

## 4 Vapaasti tuotettu puhe vieraan kielen sujuvuuden tutkimuksessa

Vieraita kieliä opiskeltaessa sujuvuutta pidetään tärkeänä tavoitteena: oppija haluaa usein itsekin tulla kohdekielen sujuvaksi käyttäjäksi. Käsitteenä sujuvuus on kuitenkin monitulkintainen ja voi viitata laajasti koko kielitaidon yleistasoon, puhutun kielen yleistasoon tai kapeammassa merkityksessä puheen ajallisesti mitattaviin ilmiöihin kuten taukoihin ja puhenopeuteen (Lennon 1990, 2000; Tavakoli & Hunter 2018). Tämän teemanumeron kolmannen kokonaisuuden artikkeleissa keskitytään sujuvuuteen erityisesti kapean määrittelyn näkökulmasta. Puheen sujuvuuden tutkimus alkoi varsinaisesti 1950-luvulla ja keskittyi aluksi erityisesti ensikieliin, kun taas vieraan kielen sujuvuutta ryhdyttiin tutkimaan enemmän Suomessa ja kansainvälisesti 1970-luvun lopulla (ks. Lintunen ym. 2020). Tätä ennen tutkijat olivat usein kiinnostuneita puheen

tarkkuudesta. Kiinnostus sujuvuuteen lisäsi myös kiinnostusta vapaasti tuotettuun puheeseen aineistona samaan aikaan, kun vieraan kielen oppimisen tutkijat kiinnostuivat oppijoiden kommunikatiivisesta kompetenssista.

Vieraan kielen sujuvuus on monitahoinen ilmiö, joka jaetaan usein Segalowitzin (2010) mukaan kolmeen ulottuvuuteen, eli puhetuotoksen sujuvuuteen (*utterance fluency*), kognitiiviseen eli prosessoinnin sujuvuuteen (*cognitive fluency*) sekä kuulijan havaitsemaan sujuvuuteen (*perceived fluency*). Tässä teemanumerossa jokaista ulottuvuutta käsittelee yksi artikkeli. Sujuvuusaineistoja kerätään eri metodein riippuen siitä, mikä sujuvuuden osa-alue on tarkastelun kohteena. Tavallista kuitenkin on, että puhetuotoksella on analyysissa jonkinlainen rooli. Tuotosta voidaan käyttää esimerkiksi havaitun sujuvuuden arvioinnin perusteena tai vertailukohtana kognitiivisen sujuvuuden testeihin.

Tuotettua puhetta voidaan kerätä eri tavoin. Ellis ja Barkhuizen (2005: 22–49) kuvaavat mahdollisuuksia kerätä suullista aineistoa kielenoppijoilta. Heidän mukaansa keräystilanteen vaatima kontrollointi vaikuttaa siihen, kuinka luonnollista tai vapaata puhetta tutkija pystyy keräämään (ks. myös osa 2 foneettisesta tutkimuksesta): vapain puhe esiintyy luonnollisissa kielenkäyttötilanteissa ja täyttää yleensä jonkin kommunikatiivisen tarpeen. Tutkija taas usein haluaa osallistujien tuottavan puhetta äänitystilanteessa melko luonnollisesti, mikäli tutkittavia piirteitä ei ole täysin määritelty etukäteen, tai hyvin kontrolloidusti, jos tavoitteena on saada esimerkkejä tietyistä kielen piirteistä. Puhetta saadaan tavallisesti jonkin tehtävän perusteella, mikä vaikuttaa siihen, kuinka tallennettu puhe sijoittuu vapaan ja kontrolloidun puheen jatkumolle. Jokaisessa äänitystilanteessa ns. tarkkailijan paradoksi (Labov 1972) vaikuttaa kuitenkin puheen luonnollisuuteen: audio- tai videolaitteet voivat vaikuttaa puhujan toimintaan ja tuotettuun kieleen (ks. myös ns. Hawthorne-ilmiö vieraan kielen tutkimuksessa esim. Mackey & Gass 2005:187–188). Kun kyseessä ovat kielenoppijat, täytyy myös huomioida osallistujien ikä ja kielitaidon taso: matalampi kielitaito saattaa esimerkiksi luokkahuoneahdistukseen yhdistettynä tehdä puhetilanteesta haastavan, eikä se välttämättä tuota edustavaa näytettä puhujan taidoista (ks. Horwitz ym. 1986). Sujuvuustutkimuksen aineistoiksi kerätään usein lyhyitä dialogisia ongelmanratkaisutehtäviä tai kuvakerrontaa, jossa osallistuja kertoo tarinan omin sanoin esimerkiksi videon, sarjakuvan tai erillisten kuvien pohjalta (ks. lisää Peltonen 2020: 40–41). Tällä tavoin koottujen aineistojen ongelma on, etteivät ne ole täysin vapaita vaan perustuvat tehtävänantoon, mutta toisaalta ne tarjoavat vertailukelpoista aineistoa eri puhujilta (ks. myös osa 2).

Tässä teemanumeron artikkelissaan Pauliina Peltonen ja Pekka Lintunen tutkivat artikkelissaan puhetuotoksen sujuvuutta. Tätä tarkasteltiin analysoimalla ensikielen (suomi) ja toisen kielen (englanti) sanahakuja ja näiden yhteyttä sujuvuuteen. Aineistona käytettiin ongelmanratkaisutehtävässä tuotettua vapaata puhetta. Tutkimuksen mukaan ensikielellä tehtiin enemmän yksittäisiä sanahakuja kuin toisella kielellä. Ensikielen ja toisen kielen sanahakujen välillä ei ollut merkittävää korrelaatiota, vaik-

ka laadullinen analyysi viittasikin siihen, että puhujien kommunikaatiotyylit olivat samanlaiset eri kielissä. Merkittävä uutta luova tulos artikkelissa on se, että erityisesti yhteistyötä edistävät sanahaut itse asiassa paransivat vuorovaikutuksen sujuvuutta, eivätkä heikentäneet sitä, kuten aiemmassa tutkimuksessa on esitetty. Kirjoittajat tarkastelivat myös ongelmanratkaisutaitoja ja niiden yhteyksiä sanahakuihin: jotkut käyttivät strategiana enemmän yksittäisiä sanahakuja, kun taas toiset käyttivät tukena vuorovaikutuskumppania. Kirjoittajat korostavatkin, että ongelmanratkaisustrategiat saattavat olla kielikohtaisia. Tuloksia voidaan käyttää apuna ongelmanratkaisustrategioiden harjoittelussa kielenopetuksessa.

Sanna Olkkonen, Maarit Mutta ja Pekka Lintunen taas keskittyvät artikkelissaan kognitiiviseen sujuvuuteen. Tutkimuksen kohteena on edistyneiden kielenoppijoiden sujuvusprofiilit, joiden avulla tarkastellaan kognitiivisen sujuvuuden, toisen kielen kielitaidon ja puheen tuottamisen sujuvuuden välisiä suhteita. Osallistujien ensikieli on suomi ja kohdekielenä englanti. Kielenoppijoiden sujuvusprofileja tutkimalla kirjoittajat haluavat syventää tietoa puheen tuoton taustalla olevista mekanismeista. Sujuvusprofileita tarkastellaan monimenetelmäisesti käyttämällä kuvankerrontatehtävällä kerättyjä puhenäytteitä ja kognitiivisia sanankäyttötehtäviä ensi- ja toisessa kielessä. Odotusten vastaisesti kognitiivinen sujuvuus kertoi enemmän yksilöllisistä eroista kuin toisen kielen kielitaidosta, eikä näiden välinen korrelaatio ollut korkea. Kielitaito- ja sujuvustehtävien tulosten pohjalta Olkkonen ym. tekivät klusterianalyysin, jonka pohjalta luotiin sujuvusprofiilit. Profiilit paljastivat, että korkea kielitaito ei välttämättä ole aina yhteydessä korkeaan sujuvuuteen toisessa kielessä ja että ensikielen yksilölliset piirteet voivat vaikuttaa toisen kielen puheen tuottamiseen. Tutkimus osoittaa myös käänteisen suhteen tarkkuuden ja sujuvuuden välillä: toiset kielenoppijat antoivat nopeita mutta epätarkempia vastauksia, kun taas toiset puolestaan hitaampia mutta tarkempia vastauksia. Kirjoittajat esittävät, että erilaiset sujuvusprofiilit tulisi ottaa huomioon kielenopetuksessa räätälöimällä yksilöllisempiä tehtäviä sujuvuuden kehittämiseksi.

Elina Lehtilä, Pauliina Peltonen ja Pekka Lintunen tarkastelevat artikkelissaan kuulijan havaitsemaa sujuvuutta. Artikkelin tuo sujuvuustutkimukseen uutena näkökulmana sen, että sujuvuutta arvioidessa oli mahdollista tutustua myös arvioitavan puhujan ensikieliseen puhenäytteeseen. Arvioijilta kysyttiin, kuinka hyödylliseksi he kokivat puhujan ensikielisen (suomi) tuotoksen kuuntelun, kun arvioitavana oli toisen (englanti) ja kolmannen (ruotsi) kielen sujuvuus. Vaikka vain osa arvioijista koki ensikielen kuuntelun hyödylliseksi, korreloivat ensikielen koettu hyödyllisyys ja sen vaikutusaste sujuvuusarvioon keskenään. Laadullisen analyysin perusteella sujuvuuden arvioijat kokivat ensikielisen näytteen kuuntelun vapauttavan kognitiivisia resursseja toisen ja kolmannen kielen sujuvuuden arviointiin, vaikka sen kuuntelun koettu hyödyllisyys vaihtelikin arvioijien välillä. Ensikielen sujuvuutta arvioitaessa tärkeimmäksi piirteeksi nousi puhenopeus, kun taas epäroinnit ja korjaukset mainittiin harvemmin. Tämä voi viitata siihen, että näiden arviointi saattaa olla alitajuista tai että ne jäävät

arvioiden ulkopuolelle, koska arvioijilla ei ole juurikaan aiempaa arviointikokemusta. Lehtilä ym. nostavat esiin kiinnostavan tuloksen, että osa sujuvuuden arvioijista vertailee ensikielisen tuotoksen piirteitä toisen tai kolmannen kielen piirteisiin, kun taas osa arvioijista jättää tietoisesti ensikielen huomioimatta. Tutkijat korostavatkin, että on tärkeä huomata kielenoppijoiden yksilöllisten piirteiden mahdollinen, usein alitajuinenkin vaikutus sujuvuuden ja laajemmin kielitaidon arviointiin.

## 5 Lopuksi

Tässä teemanumerossa lähestytään vapaasti tuotettua puhetta eri näkökulmista ja pohditaan yhtäältä sitä, mitä vapaasti tuotettu puhe on, ja toisaalta sitä, kuinka yleisesti vapaasti tuotettua puhetta käytetään tutkimuksen aineistona. Pienoistutkimuksemme vuoden 2022 AFinLA-syysseminariumin esitelmistä osoittaa, että puheaineistojen käyttö soveltavan kielentutkimuksen aineistona on runsasta. Mielestämme tästäkin syystä on tärkeää nostaa esiin, mitä oikeastaan tarkoitetaan, kun tutkimukseen on valittu tietynlainen puheaineisto, erityisesti tämän teemanumeron aiheena oleva vapaasti tuotettu puhe. Ennen kaikkea on tärkeää kuvailla aineiston hankintaa mahdollisimman tarkasti, luotettavasti ja läpinäkyvästi. Puheaineistojen tekniseen laatuun vaikuttavat puolestaan usein melko yksinkertaiset asiat, kuten sopivan mikrofonin valinta. Hyvälaatuisia aineistoja on myös helpompi analysoida automaattisin menetelmin. Lisäksi kuulonvarainen arviointi on sekä miellyttävämpää että tehokkaampaa, jos aineistossa on vain vähän taustahälyä. Teemanumeron artikkeleiden tulokset antavat arvokasta uutta tietoa vapaasti tuotetusta puheesta. Lisäksi ne tuovat esiin uusia tutkimusaiheita, joihin tarttua tulevaisuudessa.

## Kiitokset

Tätä julkaisua on avustanut Suomen tiedekustantajien liitto. Julkaisu on myös saanut TSV:n kautta tieteellisen julkaisutoiminnan avustusta, jota opetus- ja kulttuuriministeriö myöntää.

## Kirjallisuus

- Aarikka, L. 2023. *Murre ja sen tutkimus: Näkökulmia fennistisen murteen tutkimuksen historiaan ja kieli-ideologioihin 1871–2017*. Annales Universitatis Turkuensis, sarja B, osa 626. Turku: Turun yliopisto.
- Ahola, S. 2020. Yleisten kielitutkintojen arvioijien käsityksiä thainkieliseksi tunnistettujen suomenoppijoiden suullisesta kielitaidosta. *Puhe ja kieli*, 40 (4), 203–224. <https://doi.org/10.23997/pk.103307>
- Baird L., N. Evans & S. J. Greenhill 2022. Blowing in the wind: using ‘North Wind and the Sun’ texts to sample phoneme inventories. *Journal of the International Phonetic Association*, 52 (3), 453–494. <https://doi.org/10.1017/S002510032000033X>
- Berrios, J., A. Swain & M. Fricke 2023. Implementing the map task in applied linguistics research: what, how, and why. *Research Methods in Applied Linguistics*, 2 (3). <https://doi.org/10.1016/j.rmal.2023.100081>.
- Brown, G., A. Anderson, R. Shillcock & G. Yule 1983. *Teaching talk: strategies for production and assessment*. Cambridge: Cambridge University Press
- Calamai, S. & F. Frontini 2018. FAIR data principles and their application to speech and oral archives. *Journal of New Music Research*, 47 (4), 339–354. <https://doi:10.1080/09298215.2018.1473449>
- Danner S., A. Barbosa & L. Goldstein 2018. Quantitative analysis of multimodal speech data, *Journal of Phonetics*, 71, 268–283. <https://doi.org/10.1016/j.wocn.2018.09.007>.
- Deterding, D. 2006. The North Wind versus a Wolf: short texts for the description and measurement of English pronunciation. *Journal of the International Phonetic Association*, 36 (2), 187–196. <http://www.jstor.org/stable/44526411>
- Ellis, R. & G. Barkhuizen 2005. *Analysing learner language*. Oxford: Oxford University Press.
- Hall-Lew, L. & Z. Boyd 2020. Sociophonetic perspectives on stylistic diversity in speech research. *Linguistics Vanguard*, 6 (1). <https://doi.org/10.1515/lingvan-2018-0063>
- Horwitz, E. K., M. B. Horwitz & J. Cope 1986. Foreign language classroom anxiety. *The Modern Language Journal*, 70 (2), 125–132. <https://doi.org/10.2307/327317>
- Huhta, A. & R. Hilden 2017. Kielitutkinnot ja muu laajamittainen kielitaidon arviointi Suomessa. Teoksessa A. Huhta & R. Hilden (toim.) *Kielitaidon arviointitutkimus 2000-luvun Suomessa*. AFinLA-e Soveltavan kielitieteen tutkimuksia 9. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 3–26.
- Jespersen, A. B. & A. Hejná 2021. Second language learners of Danish as the linguistic other. *Otherness: Essays & Studies*, 8 (3), 129–161. [https://otherness.dk/fileadmin/www.othernessandthearts.org/Publications/Journal\\_Otherness/Otherness\\_Essays\\_and\\_Studies\\_8.3/Second\\_Language\\_Learners\\_of\\_Danish\\_as\\_the\\_Linguistic\\_Other\\_-\\_Jespersen\\_and\\_Hejna\\_\\_with\\_abstract\\_.pdf](https://otherness.dk/fileadmin/www.othernessandthearts.org/Publications/Journal_Otherness/Otherness_Essays_and_Studies_8.3/Second_Language_Learners_of_Danish_as_the_Linguistic_Other_-_Jespersen_and_Hejna__with_abstract_.pdf)
- Kielipankki [www.kielipankki.fi](http://www.kielipankki.fi). Luettu [28.3.2024]
- Kohler, K. J. 2000. Investigating unscripted speech: implications for phonetics and phonology. *Phonetica*, 57 (2–4), 85–94. <https://doi.org/10.1159/000028464>
- Kurki, T. & L. Mustanoja 2019. Käänteitä, aaltoja ja risteymiä fennistisessä variaation- tutkimuksessa. Teoksessa M. Hamunen, T. Nieminen, T. Kelomäki & H. Dufva (toim.) *Käänteitä ja käsitteitä: näkökulmia kielitieteelliseen keskusteluun*. Suomalaisen Kirjallisuuden Seuran Toimituksia 1454. Helsinki: Suomalaisen kirjallisuuden Seura, 87–120.
- Kurki, T. & L. Mustanoja 2020. Haastattelu aineistonkeruun metodina sosiolingvistiikassa. Teoksessa M. Luodonpää-Manni, M. Hamunen, R. Konstenius, M. Miestamo, U. Nikanne & K. Sinnemäki (toim.) *Kielitutkimuksen menetelmiä II*. Helsinki: Suomalaisen Kirjallisuuden Seura, 274–312. <https://doi.org/10.21435/skst.1457>

- Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017a. *Näkökulmia toisen kielen puheeseen – Insights into second language speech*. AFinLA-e. Soveltavan kielitieteen tutkimuksia n:o 10. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA. <https://journal.fi/afinla/issue/view/5195>
- Kuronen, M., P. Lintunen & T. Nieminen 2017b. Suullisen kielitaidon ja ääntämisen tutkimuksesta soveltavan kielentutkimuksen alalla Suomessa. Teoksessa M. Kuronen, P. Lintunen & T. Nieminen (toim.) *Näkökulmia toisen kielen puheeseen – Insights into Second Language Speech*. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 3–17. <https://journal.fi/afinla/article/view/73121>
- Labov, W. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania.
- Lee, W.-S. & E. Zee (toim.) 2011. *Proceedings of the 17th International congress of the phonetic sciences 2011*. Hong Kong: City University of Hong Kong.
- Lennes, M., E. Aho, M. Toivola & L. Wahlberg 2006. On the use of the glottal stop in Finnish conversational speech. Teoksessa Aulanko, R., Wahlberg, L., & Vainio, M. (toim.) *Fonetiikan päivät 2006: The phonetics symposium 2006*. Helsingin yliopisto, 93–102.
- Lennon, P. 1990. Investigating fluency in EFL: a quantitative approach. *Language Learning*, 40, 387–417. <http://dx.doi.org/10.1111/j.1467-1770.1990.tb00669.x>.
- Lennon, P. 2000. The lexical element in spoken second language fluency. Teoksessa H. Riggensbach (toim.) *Perspectives on fluency*. Ann Arbor, MI: The University of Michigan Press, 25–42.
- Lintunen, P., M. Mutta & P. Peltonen 2020. Sujuvuustutkimuksen käännteitä: katsaus sujuvuustutkimuksen menetelmiin. Teoksessa S. Grasz, T. Keisanen, F. Oloff, M. Rauniomaa, I. Rautiainen & M. Siromaa (toim.) *Menetelmällisiä käännteitä soveltavassa kielentutkimuksessa – Methodological Turns in Applied Language Studies*. AFinLAN vuosikirja 78. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 174–197. <https://journal.fi/afinlavk/article/view/89441>
- Mackey A. & S. Gass 2005. *Second language research: methodology and design*. Mahwah: Lawrence Erlbaum Associates.
- Mielikäinen, A. (2017). Ensimmäisten suomenkielisten murre tutkimusten metakieli. *Virittäjä*, 121 (3), 355–381. <https://doi.org/10.23982/vir.58386>
- Ogden R. 2001. Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association*, 31 (1), 139–152. <https://doi.org/10.1017/S0025100301001116>
- Peltonen, P. 2020. *Individual and interactional speech fluency in L2 English from a problem-solving perspective: a mixed-methods approach*. Turku: University of Turku. <https://urn.fi/URN:ISBN:978-951-29-8137-3>
- de Ruiter, L. 2015. Information status marking in spontaneous vs. read speech in story-telling tasks – Evidence from intonation analysis using GToBI. *Journal of Phonetics*, 48, 29–44. <https://doi.org/10.1016/j.wocn.2014.10.008>.
- Segalowitz, N. 2010. *The cognitive bases of second language fluency*. New York: Routledge
- Setter, J. & J. Jenkins 2005. Pronunciation. *Language Teaching*, 38, 1–17.
- Tavakoli, P. & A.-M. Hunter 2018. Is fluency being 'neglected' in the classroom? Teacher understanding of fluency and related classroom practices. *Language Teaching Research*, 22 (3), 330–349. <http://dx.doi.org/10.1177/1362168817708462>
- Trouvain, J. & W. J. Barry (toim.) 2007. *Proceedings of the 16th International congress of the phonetic sciences 2007*. Dudweiler: Pirrot GmbH.
- Vakkilainen, J. & M. O'Dell 2022. Suomen pitkien painottomien vokaalien havaitseminen spontaanissa ja luetussa puheessa. *Puhe ja kieli*, 41 (4), 247–264. <https://doi.org/10.23997/pk.113851>



- Wagner, P., J. Trouvain & F. Zimmerer 2015. In defense of stylistic diversity in speech research. *Journal of Phonetics*, 48, 1–12. <https://doi.org/10.1016/j.wocn.2014.11.001>.
- Xu, Y. 2010. In defense of lab speech. *Journal of Phonetics*, 38 (3), 329–336. <https://doi.org/10.1016/j.wocn.2010.04.003>.
- Yleinen tietosuoja-asetus (GDPR) 2016/679. <https://eur-lex.europa.eu/Fl/legal-content/summary/general-data-protection-regulation-gdpr.html>