

AFinLA-e

Soveltavan kielitieteen tutkimuksia 2017/n:o 10

# Näkökulmia toisen kielen puheeseen Insights into Second Language Speech

Toimittaneet / Edited by

Mikko Kuronen, Pekka Lintunen & Tommi Nieminen



Suomen soveltavan kielitieteen yhdistys AFinLA  
Jyväskylä



# SISÄLTÖ

Esipuhe . . . . . 1

*Mikko Kuronen, Pekka Lintunen & Tommi Nieminen*

Suullisen kielitaidon ja ääntämisen tutkimuksesta  
soveltavan kielentutkimuksen alalla Suomessa . . . . . 3

## Osa I

*Leena Maria Heikkola & Jenni Alisaari*

Laulun sanoja ääntämällä taitavaksi ääntäjäksi? . . . . . 18

*Martina Huhtamäki & Elisabeth Zetterholm*

Uttalets plats i undervisningen av svenska som andraspråk . . . . . 45

*Leila Kääntä*

In search of proper pronunciation:  
students' practices of soliciting help during read-aloud . . . . . 61

*Irina Spita & Martina Björklund*

Teaching emotive language use in a beginner's course of Russian . . . . . 82

## Osa II

*Henna Heinonen*

Begripligheten av finskspråkiga gymnasisters svenska uttal:  
lyssnarvärderingar . . . . . 98

*Pauliina Peltonen*

L2 fluency in spoken interaction: a case study on the use of  
other-repetition and collaborative completions . . . . . 118

*Iina Salminen*

Erään venäjänopiskelijän harjoittelu ja kehitys  
venäjän kielen sibilanttien ja affrikaattojen parissa . . . . . 139

## Osa III

*Mari Honko*

Sadutettu sanasto: puhutun kielen  
leksikaalinen diversiteetti arviointikohteena . . . . . 163

*Heini Kallio, Juraj Šimko, Ari Huhta, Reima Karhila*

*Martti Vainio, Erik Lindroos, Raili Hildén & Mikko Kurimo*

Towards the phonetic basis of spoken second language assessment:  
temporal features as indicators of perceived proficiency level . . . . . 193

## II

*Mikko Kuronen & Elina Tergujeff*

L1 listeners' perception of L2 pronunciation:  
effect of prosody on accentedness ratings in Swedish . . . . . 214

*Sanna Olkkonen & Pauliina Peltonen*

Mitä on toisen kielen sujuvuus?  
näkökulmia kognitiivisen ja puhetuotoksen sujuvuuden tutkimuksesta . . . . . 234

*Minnaleena Toivola & Riikka Ullakonoja*

Identification of Russian accented Finnish by native and non-native  
with and without Finnish proficiency . . . . . 258

*Melike Uzal, Erkki Komulainen & Olli Aaltonen*

The effect of some listener background factors and task type  
that contribute to degree of perceived accent ratings in L2 Finnish . . . . . 277

## ESIPUHE

”Haa”, huudahti rosvo murteellisella hollannin kielellä ja katosi Lontoon suumuun.<sup>1</sup> Jos ei juuri tätä niin tämänkin kaltaisia havaintoja puhutun kielen viivahteikkaasta luonteesta käsitellään tässä AFinLA-e (n:o 10) julkaisussa. Julkaisun taustalla on Suomen soveltavan kielitieteen yhdistyksen, AFinLAN, alaisuuteen vuonna 2014 perustettu *Puheen ja puheentutkimuksen teemaverkosto*. Teemaverkoston tavoitteena on AFinLAN toiminnan puitteissa, erityisesti yhdistyksen syysseminariumien yhteydessä, tarjota suullisesta kielitaidosta ja ääntämisestä kiinnostuneille tutkijoille foorumi ajatustenvaihtoon alan tutkimuksesta kotimaassa ja kansainvälisesti. Verkostolla on tällä hetkellä liki 50 jäsentä. Käsillä oleva julkaisu on teemaverkoston toiminnan toistaiseksi näkyvin tulos. Ajatus teemanumerosta syntyi keskusteluissa syysseminariumien yhteydessä. Näissä teemanumerolle koettiin olevan tarve, koska sellainen antaisi lukijalle käsityksen teemaverkoston alalla tehtävästä tutkimuksesta maassamme ja tekisi soveltavan puheen tutkimusta tunnetummaksi laajemminkin suomalaisen kielentutkimuksen piirissä.

Kirjoituskutsussa toivottiin ehdotuksia tutkimuksista, joissa tarkastellaan empiirisesti jotain puhutun kielen, suullisen kielitaidon tai ääntämisen osaluuetta tavalla, joka on mielenkiintoinen suomalaisessa kontekstissa. Hyviä ehdotuksia tuli runsaasti, ja teemanumeromme tutkimukset edustavatkin ilahduttavasti lähestymistavoiltaan ja menetelmiltään suuntauksia, jotka ovat viime vuosina olleet vahvoja alalla myös kansainvälisesti. Julkaisumme voikin sanoa antavan lukijalle hyvän käsityksen soveltavan puheen tutkimuksen nykytuulista. Käsittelemme tätä kysymystä tarkemmin julkaisun ensimmäisessä artikkelissa, jossa tarkastelemme myös puhutun kielen tutkimusta AFinLAN tuotannossa aikavälillä 1977–2016.

Kiitämme lämpimästi kaikkia teemanumeron kirjoittajia, tehtävänsä huolellisesti paneutuneita mutta nimettömäksi jääviä vertaisarvioita sekä AFinLAN julkaisutoimikuntaa: ilman panostanne tätä julkaisua ei olisi. Kiitämme myös Koneen säätiötä ja Tieteellisten seurain valtuuskuntaa julkaisun taloudellisesta tukemisesta sekä Sinikka Lampista kannen laatimisesta.

Tampereella, Turussa ja Nokialla 28. 2. 2018

Teemaverkoston vetäjät:

Mikko Kuronen, Pekka Lintunen ja Tommi Nieminen

---

<sup>1</sup> Tämän ilmauksen on professori emeritus Kari Suomi välittänyt toimittaja Kuroselle vuosia sitten. Suomi ei muista, mistä hän on sen oppinut.

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 3–17.

**Mikko Kuronen<sup>1</sup>, Pekka Lintunen<sup>2</sup> & Tommi Nieminen<sup>3</sup>**

<sup>1</sup>Jyväskylän yliopisto, <sup>2</sup>Turun yliopisto, <sup>3</sup>Itä-Suomen yliopisto

## **Suullisen kielitaidon ja ääntämisen tutkimuksesta soveltavan kielitutkimuksen alalla Suomessa**

AFinLA's research network for Speech and Spoken Language Research was established in 2014. This is the first thematic publication from the network. This introductory article introduces and classifies the other articles in this collection. In addition, we examine earlier AFinLA publications from the preceding four decades and evaluate the extent of pronunciation or spoken language research in AFinLA's publication series. In our analysis, we noticed that studies related to pronunciation or spoken language research formed ca. 12% of the total volume of published articles. There has been a slight proportional decrease in number during the decades. Spoken language studies have been more common than studies focusing on pronunciation. The target languages have been most commonly either Finnish or English. The main focus has often been on learning or assessing foreign language skills, but studies focusing on the teaching of pronunciation or spoken language have so far been very rare in Finland.

**Avainsanat:** ääntäminen, suullinen kielitaito, oppiminen, opetus

**Keywords:** pronunciation, spoken communication, learning, teaching

## 1 Johdanto

*Puheen ja puheentutkimuksen teemaverkosto* perustettiin AFinLAN alaisuuteen keväällä 2014. Kuvauksen mukaisesti teemaverkosto on tarkoitettu puheesta, suullisesta kielitaidosta (puheen tuottamisesta ja ymmärtämisestä), prosodiasta, fonetiikasta, ääntämisestä tai muista puhutun kielen osa-alueista kiinnostuneille soveltavan kielitieteen tutkijoille. Käsillä oleva artikkelikokoelma on ensimmäinen teemaverkoston toiminnan perusteella aikaan saatu julkaisu. Kirjoituskutsussa pyydettiin kirjoittajia tarttumaan empiirisesti aiheeseen, joka liittyisi puheeseen tai suullisen kielitaidon tai ääntämisen oppimiseen ja joka olisi kiinnostava suomalaisessa kontekstissa. Artikkeleita saapui runsaasti, ja ne vastasivat hyvin niin teemaverkoston kuin kirjoituskutsummekin teemaa. Monet tämän kokoelman artikkeleista pohjautuvat AFinLAN syys-symposiumeissa pidettyihin temaattisiin työpajoihin.

Tässä katsausartikkelissamme esittelemme aluksi kartoituksen Suomen soveltavan kielitieteen yhdistyksen, AFinLAN, vuosikirjojen ja erikoisjulkaisujen puheen tutkimusta, tarkemmin suullista kielitaitoa ja ääntämistä, käsittelevistä tutkimuksista vuosivälillä 1977–2016. Katsauksen tarkoituksena on antaa kuva siitä, paljonko ja millaista tutkimusta vuosikirjoissa ja erikoisjulkaisuissa on ollut suullista kielitaitoa ja ääntämistä koskien. AFinLAN julkaisut ovat vain yksi monista kanavista, joissa puheen tutkimusta Suomessa julkaistaan, eikä katsauksemme siten anna tyhjentävää kuvaa maassamme tehdystä puhutun kielen tutkimuksesta. Otoksemme riittää kuitenkin antamaan käsityksen valinneista tutkimusperinteistä ja painopisteistä soveltavan kielentutkimuksen alalla.

Artikkelimme jälkimmäisessä osassa esittelemme käsillä olevan erikoisjulkaisun tutkimukset. Jälkimmäinen osa antaa kuvan siitä, millaisia kysymyksiä soveltava puheentutkimus nykyisin käsittelee.

## 2 Suullisen kielitaidon ja ääntämisen tutkimus AFinLAN julkaisuissa 1977–2016

Artikkeliämme varten kartoitimme AFinLAN vuosikirjat ja erikoisjulkaisut vuosina 1977–2016. Tarkasteltu 40 vuoden ajanjakso mahdollistaa tutkimuspainotusten muutosten analyysin. Kartoituksemme perustuu ensisijaisesti otsikotason analyysiin, mutta tarvittaessa perehdyimme myös itse tutkimuksiin. Sekä ensikielen (L1) että toisen (= vieraan) kielen (L2) puheeseen keskittyvät tutkimukset luokiteltiin alaluokkiin suullinen kielitaito tai ääntäminen. Määrittely 'puheeseen keskittyvä tutkimus' ei ollut aina helppo, koska lukuisissa tutkimuksissa käsitellään eri tavoin vuorovaikutusta, viestintää tai diskursia (engl. *interaction, communication, discourse*). Kriteerimme puhutun kielen

tutkimukselle olivat, että (i) sen aineisto on (vähintäänkin osin) puhuttua kieltä ja (ii) aineistoa käsitellään (vähintäänkin osin) puhutun kielen näkökulmasta ja sen erityisominaisuuksiin keskittyen. Emme luokitelleet tutkimusta puhutun kielen tutkimukseksi, ellei tutkimuksessa tehty eroa kirjallisen ja suullisen kielen välille (eli tutkimus esimerkiksi käsitteli jotakin vuorovaikutuksellista ilmiötä mutta ei suoranaisesti puhetta) tai tutkimuksen pääpaino oli kirjoitettussa kielessä. Rajauksemme ei toki sulkenut pois esimerkiksi vuorovaikutukseen keskittyviä tutkimuksia, mutta niiden luokittelu puhetta käsitteleviksi tutkimuksiksi edellytti, että vuorovaikutuksen tarkastelu kohdistui nimenomaan johonkin puhutun kielen piirteeseen, kuten Leila Käännän tutkimuksessa (ks. luku 3 alla) ääntämisen opetukseen luokkahuonetilanteessa.

Puhutun kielen tutkimukset määriteltiin tarkemmin opetusta, oppimista ja arviointia tai kielitaidon mittaamista koskeviksi. Myös alaluokittelu tuotti joissain tapauksissa vaikeuksia, koska samassa tutkimuksessa saatettiin käsitellä kahta tai (harvemmin) kaikkia kolmea alaluokkaa. Tällaisissa tapauksissa annoimme tutkimukselle alaluokan sen sisällöllisen painopisteen mukaan. Luokitteluun liittyvien rajanvetojen vuoksi kartoituksemme tuloksia ei tule tarkastella artikkelin eikä edes prosenttien tarkkuudella, vaikka sen käytännön syistä teemmekin alla.

Suullista kielitaitoa tai ääntämistä käsitteleviä tutkimuksia oli 89 kaikkiaan 724:sta AFinLAN julkaisuissa raportoiduista tutkimuksista, mikä tekee 12,3 % tutkimuksen kokonaismäärästä (taulukko 1). Kielentutkimus jaetaan joskus kahteen pääluokkaan, kirjalliseen ja suulliseen. Tämän lisäksi on tutkimusta, joka käsittelee kumpaakin pääluokkaa tai ei erityisesti kumpaakaan; myös tämä kartoitus perustuu näihin luokkiin. Täten ei voi odottaakaan, että suullisen kielitaidon ja ääntämisen tutkimus muodostaisi puolta tutkimuksen kokonaismäärästä. Silti vain reilun kymmeneksen osuutta voi pitää pienenä, koska suullisen kielitaidon merkitys on esimerkiksi kielen oppimisessa suuri. Suullisen kielitaidon vähäinen tutkimus AFinLAN julkaisuissa heijastelee näkemyksemme mukaan perinnettä, joka on leimannut kielitiedettä yleisemminkin: lähtökohtana on useimmiten ollut kirjallinen tuotos, kieliopin tai rakenteiden käyttö ja hallinta kirjoitetussa kielessä (vrt. Linell 2005: *the written language bias*). Vaikka tarkasteluvälimme onkin hieman eri, osuutta voi varovasti verrata esimerkiksi Karlssonin (1998, 2003) selvitykseen kielitieteiden tohtorinväitöskirjoista, jonka mukaan Suomessa valmistui 413 kielitieteellistä väitöskirjaa vuosina 1978–2001. Alakohtaisissa jaotteluissa todetaan vuosina 1972–2001<sup>1</sup> ilmestyneen 19 fonetiikan, puheviestinnän tai logopedian väitöskirjaa, joiden teeman voi ajatella suoraan liittyvän puhuttuun kieleen (19/453 eli 4 %). Suomi (2004) tarkensi lukumäärää huomauttamalla, että monet suulliseen kielitaitoon liittyvät väitöskirjat tehdään vieraiden kielten oppiaineisiin tai tutkin- to suoritetaan ulkomaisissa yliopistoissa. Suomi lisää 12 foneettista väitöskir-

<sup>1</sup> Alakohtaisissa erittelyissä on lähteissä eri vuosijako kuin kokonaissummissa.



TAULUKKO 1. Suullisen kielitaidon ja ääntämisen tutkimus AFinLAN julkaisuissa 1977–2016. Taulukossa annetaan tutkimusten määrä pää- ja alaluokissa. Kullekin luokalle annetaan myös prosenttiosuus tutkimuksen kokonaismäärästä.

	opetus		oppiminen		arviointi / mittaaminen		yhteensä	
	n	%	n	%	n	%	n	%
suullinen kielitaito	10	1,4	38	5,2	20	2,8	68	9,4
ääntäminen	2	0,3	14	1,9	5	0,7	21	2,9
yhteensä	12	1,7	52	7,2	25	3,5	89	12,3

jaa tarkastelemallemme ajanjaksolle. Koska ääntämiseen tai suulliseen kieli-taitoon liittyviä väitöskirjoja tehdään myös muihin oppiaineisiin ja yhteismää-rä on siis oletettavasti hieman näitä numeroita korkeampi, AFinLAN julkaisuissa ilmestyneiden aihepiirin tutkimusten suhteellisen osuuden voidaan tulki-ta vastaavan hyvin ääntämisen ja suullisen kielitaidon osuutta suomalaisessa kielitieteellisessä tutkimuksessa laajemminkin, jos lukumäärää verrataan esi-merkiksi väitöskirjojen teemoihin.

Ääntämistä on tutkittu vain vähän (2,9 % kaikista AFinLAN julkaisujen tut-kimuksista), ja aivan erityisen vähän on tutkittu sen opetusta: 0,3 % kaikista tutkimuksista (taulukko 1). Mitätöntä osuutta selittää osaltaan näin rajatun alueen kapeus, sillä ääntäminen on itsessään vain eräs suullisen kielitaidon osa-alue. Siltikin osuus on hämmästyttävän pieni, joten tutkimusalueetta ei il-meisesti ole pidetty tärkeänä. Tulos heijastelee toisaalta sitä, että ääntämi-sen opetuksen tutkimus - esimerkiksi eri menetelmien vertailu - on haastavaa (vrt. Derwing & Munro 2015: 77–108). Se edellyttää seurantaa, kontrolliryh-mää tai -ryhmiä, oppimiseen vaikuttavien tekijöiden kontrollia sekä mahdolli-sesti myös foneettisten menetelmien hallintaa. Lisäksi ääntämiseen keskitty-vää opetusta annetaan vähemmän kuin kirjalliseen ja yleiseen suulliseen kieli-taitoon keskittyvää opetusta, mikä vaikeuttaa opetusryhmien ja -menetelmien vertailua.

Suomalaisena erityispiirteenä tilanteeseen on voinut vaikuttaa puheen il-miöihin erikoistuneen fonetiikan oppiaineen jatkuvasti horjahdellut asema yli-opistoissa. Foneettiset tutkimuspainotukset ja menetelmien hallinta eivät aina ole olleet valtavirtaa. Fonetiikka lakkautettiin oppiaineena Jyväskylän yliopis-tosta 1990-luvun alkupuolella, Tampereen ja Oulun yliopistoissa sillä on kaiken aikaa ollut henkilömäärältään niukat resurssit ja Turun yliopistossa sen tilan-etta ovat haastaneet erilaiset hallinnolliset muutokset kuten katkokset pro-fessoritasoisissa työtehtävissä. Käytännössä vain Helsingin yliopistossa fone-tiikalla on säilynyt pysyvä ja suhteellisen vakaa asema koko tarkastellun ajan.

TAULUKKO 2. Suullisen kielitaidon ja ääntämisen tutkimusten kohdekielet julkaisuissa.

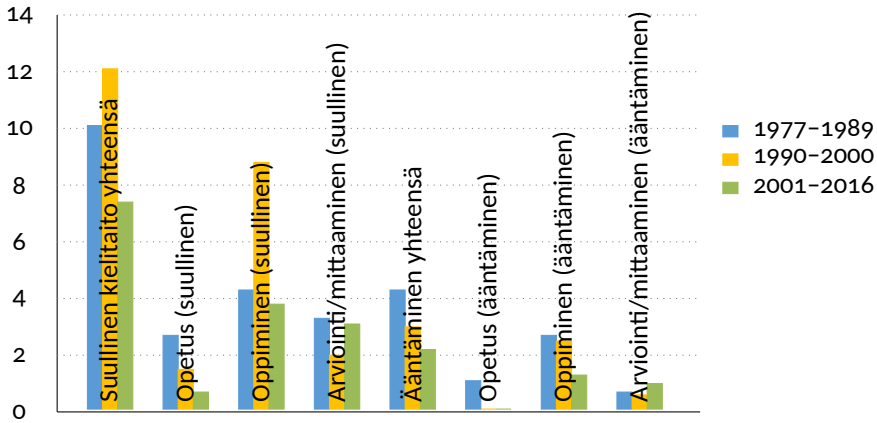
kohdekieli	lkm
englanti	18
suomi	17
ranska	3
saksa	3
venäjä	3
ruotsi	2
kielestä riippumaton	43

Toisaalta ääntämisen opetuksen tutkiminen on kansainvälisestikin ollut vähäistä (mm. Norris & Ortega 2000; Derwing & Munro 2005; Tergujeff 2013), vaikkakin se paraikaa vaikuttaa olevan hienoisessa nousussa (Derwing 2010). Ääntämisen opettamisen ja oppimisen vähäistä tutkimusta selittänee myös se, että foneettisia menetelmiä hallitsevien tutkijoiden mielenkiinto on ollut paremminkin ensikielissä ja sen murteissa kuin toisen kielen puheessa. Derwing ja Munro (2005: 379) toteavat ääntämisen tutkimisesta kansainvälisesti ja erityisesti englannin kieltä koskien, että se on ollut marginalisoitua soveltavan kielitieteen kentällä. Kartoituksemme näyttää, että näin voimme todeta tilanteesta Suomessakin.

Myös suullisen kielitaidon opetuksen tutkimus on ollut vähäistä (1,4 %, taulukko 1), ja syyt siihen ovat luultavasti samansuuntaiset kuin ääntämisen opetuksen osalta. Suullisen kielitaidon oppimista on tutkittu selvästi enemmän (5,2 %). Arviointia ja mittaamista on tutkittu hiukan enemmän kuin opetusta sekä ääntämisen että suullisen kielitaidon osalta (taulukko 1). Yhteenlaskettuna tavallisinta on ollut oppimislähtöinen tutkimus (7,2 %), sen jälkeen arviointi ja mittaaminen (3,5 %) ja viimeisenä opetus (1,7 %).

Jaomme tutkimukset myös kohdekielen mukaan (taulukko 2). Kielikohtaisia tutkimuksia oli 46, kun taas 43 tutkimusta oli kielestä riippumattomia. Huomattava osa tutkimuksista on ollut siis sellaisia, joissa ei käsitellä suoraan jotain tiettyä kieltä. Kielikohtaisista tutkimuksista valtaosa on käsitellyt englantia tai suomea.

Mikäli jaamme kartoituksemme kolmeen liki yhtä suureen aikaväliin 1977–89, 1990–2000 ja 2001–16, näemme, että suullisen kielitaidon osuus AFinLAN tutkimusten kokonaismäärästä ei ole ajan saatossa kasvanut vaan päinvastoin hienoisesti laskenut (kuvio 1). Kaikissa luokissa alin lukema on vuosivälillä 2001–2016 arviointia ja mittaamista lukuun ottamatta. Kun otetaan huomioon kartoituksen tarkkuuteen liittyvät rajoitukset, vähäisille muutoksille ei pidä antaa liian suurta painoa. Selvää kuitenkin on, että suullisen kielitaidon ja ääntämisen osuus tutkimuksen kokonaismäärästä ei ole ainakaan kasvanut



KUVIO 1. Suullisen kielitaidon ja ääntämisen tutkimus AFInLAN julkaisuissa vuosivälillä 1977-1989, 1990-2000 ja 2001-2016.

tarkasteltavana ajanjaksona. Ääntämisen opetuksen tutkimisesta ei ole julkaisuja lähes koko aikana. Olematonta osuutta 1990-2000 ja 2001-2016 ei selitä ainakaan se, että aihetta olisi tutkittu tyhjentävästi jo aiemmin. Myös suullisen kielitaidon opettamisen tutkimus on ollut hyvin vähäistä ja pienintä viimeisellä ajanjaksolla. Suullisen kielitaidon oppimisen tutkimuksessa tapahtui selvä nousu 1990-luvulla, joka nyttemmin on taittunut. Suullisen kielitaidon painoarvon kasvu opetuksessa näkyy opetussuunnitelmissa 2000-luvulla (vrt. LOPS 2003, 2015) ja myös siten, että lukiossa on nykyisin suulliseen viestintään keskittyviä toisen kotimaisen/vieraan kielen kurseja (vrt. Muutos LOP-Siin 2009; ks. myös Ruoho 2011). Ainakaan vielä suullisen kielitaidon painoarvon kasvu opetuksessa ei kuitenkaan tunnu näkyvän tutkimuksessa määrällisesti. Sen kartoitus kuitenkin näyttää, että opetuksen, oppimisen ja arvioinnin kysymykset ovat kiinnostaneet soveltavan kielitieteen tutkijoita jossain määrin koko kartoituksen ajan myös suullisen kielitaidon alueella (vrt. Kalaja ym. 2013).

Tiivistäen kartoituksemme osoittaa, että (i) suullinen kielitaito ja ääntäminen eivät ole olleet vahvasti edustettuina AFInLAN tuotannossa, (ii) suullisen kielitaidon ja ääntämisen tutkimukset eivät ole suhteelliselta määrältään nousussa vaan päinvastoin ne vaikuttavat olevan hienoisessa laskussa ja (iii) erityisen vähän on ollut tutkimusta suullisen kielitaidon ja ääntämisen opetuksesta. On perusteltua olettaa, että suullisen kielitaidon ja ääntämisen opettamisen tutkimuksen vähyyks heijastuu puutteina kieltenopettajien koulutuksessa: Koska aihetta on tutkittu vähän, on yliopistojen opetushenkilökunnanakin sitä koskeva tieto puutteellista. Puutteiden vuoksi merkittävä osa kielten opettajista kokee, ettei heillä ole riittäviä valmiuksia opettaa ääntämistä. He eivät yksinkertaisesti tiedä, millaisin menetelmin sitä voi opettaa (mm. Huh-

tamäki & Zetterholm tässä julkaisussa; Moilanen 2014). Tämä saattaa johtaa siihen, että ääntämisen opetusta laiminlyödään tai opetus jää implisiittiseksi eikä kehitä oppijan ääntämistietoisuutta tai -motoriikkaa eteenpäin (Lintunen 2005, 2014). Pahimmillaan ääntämisharjoitukset saatetaan ohittaa kokonaan. Opetusmenetelmiä vertailevan ja oppimistuloksiin keskittyvän tutkimuksen vähyyks vaikuttaa leimaavan oppimisen tutkimusta laajemminkin kuin vain kielennoppimisen alueella (mm. Eklund 2000). Myös tieto siitä, mikä ääntämisessä on ymmärtämisen ja vieraan korostuksen kannalta tärkeää, on puutteellista tutkimuksen vähyyden vuoksi (vrt. Kuronen & Tergujeff tässä julkaisussa).

AFinLAN *Puheen ja puheentutkimuksen teemaverkosto* perustettiin edistämään puhutun kielen ja suullisen kielitaidon ja ääntämisen tutkimusta sekä tukemaan tutkijoiden verkostoitumista. Kartoituksemme perusteella teemaverkostolle on ollut selvä tarve. Uskomme, että käsillä oleva julkaisu auttaa osaltaan edistämään puheen tutkimusta soveltavan kielentutkimuksen alalla ja tässä teemanumerossa julkaistavat tutkimukset saavat jälleen uusia tutkijoita tarttumaan aihepiiriin haasteisiin.

### 3 Käsillä olevan julkaisun tutkimukset

Tämä AFinLAN erikoisjulkaisu keskittyy puhutun kielen tutkimukseen. Julkaisun kaikki tutkimukset käsittelevät joko suullista kielitaitoa tai ääntämistä opettamisen, oppimisen tai arvioinnin näkökulmasta. Julkaisu koostuu 13 tutkimuksesta, joissa kaikissa pääpaino on toisessa kielessä vähintäänkin niin, että toista kieltä tarkastellaan suhteessa ensikieleen. Tutkimuksista neljä keskittyy suulliseen kielitaitoon ja yhdeksän ääntämiseen (taulukko 3). Kuten taulukko 3 osoittaa, tutkimukset lisäävät tietoa jaottelun jokaisessa kategoriassa. Useat tutkimuksista ovat kielikohtaisia, vaikka tulokset voivatkin olla yleistettävissä muihin kieliin. Neljä artikkelia voidaan laskea kielestä riippumattomiksi. Kielikohtaiset tutkimukset jakautuvat siten, että suomeen kohdistuu neljä, ruotsiin kolme ja venäjään kaksi artikkelia.

Kun artikkeleissa käytettyjä metodeja tarkastelee päämenetelmän mukaan, viidessä tutkimuksessa on käytetty kuuntelukokeita (syntyperäiset tai

TAULUKKO 3. Tämän teemajulkaisun artikkelit tutkimusalueittain.

	opetus	oppiminen	arviointi / mittaaminen	yhteensä
suullinen kielitaito	1	1	2	4
ääntäminen	3	2	4	9
yhteensä	4	3	6	13

ei-syntyperäiset kuulijat) ja kolmessa on mitattu oppimista (akustisesti, kuulonvaraisesti tai sanaston kannalta). Lisäksi teemajulkaisumme artikkelien menetelminä on opetusmenetelmän testaus, opetusmenetelmien vertailu, opetuksen havainnointi sekä opettajien haastattelu tai kysely, ja yksi artikkeli on lähestymistavaltaan teoreettinen.

Seuraavassa esittelemme tutkimukset ryhmitellen ne kartoituksemme luokkiin. Esittely noudattaa julkaisumme sisällysluetteloa järjestäen tutkimukset ryhmiin (i) suullisen kielitaidon/ääntämisen opettaminen, (ii) suullisen kielitaidon/ääntämisen oppiminen ja (iii) suullisen kielitaidon/ääntämisen arviointi ja mittaaminen.

### 3.1 Suullisen kielitaidon ja ääntämisen opetuksen tutkimukset

Tutkimuksista neljä käsittelee suullisen kielitaidon tai ääntämisen opettamista (*Heikkola & Alisaari, Huhtamäki & Zetterholm, Kääntä, Spita & Björklund*). Juuri opettamista on tutkittu aiemmin hyvin vähän, ja julkaisun tutkimukset ovatkin arvokas lisä tälle alitutkitulle alueelle. Tutkimukset ovat metodologisesti erilaisia: opetusmenetelmiä verrataan keskenään eri opetusryhmissä, opetusmenetelmää testataan, opetusta havainnoidaan tai opettajien kokemuksia kartoitetaan kyselyin ja haastatteluin.

**Leena Maria Heikkola ja Jenni Alisaari** selvittävät tutkimuksessaan, kuinka kolme erilaista ääntämisen opetuksessa käytettävää menetelmää – laulaminen, laulujen kuuntelu ja laulun sanojen rytmikäs toistaminen – vaikuttavat suomenoppijoiden ääntämisen kehittymiseen. Oppijat osallistuivat neljä viikkoa kestäneelle suomen kurssille, jonka aikana ääntämistä opetettiin yllä mainituilla menetelmillä. Oppijoiden ääntämistä kurssin alussa ja lopussa arvioi kaksi kuulijaraatia: tulevat S2- eli suomi toisena kielenä -opettajat sekä tulevat luokanopettajat. Tulosten mukaan alkeiskurssin oppijoilla tapahtui merkittävää kehitystä ääntämisessä, kun puolestaan hiukan edistyneempien oppijoiden kohdalla edistyminen ei ollut yhtä suoraviivaista. Laulun sanojen rytmikäs toistaminen vaikutti olevan menetelmistä tehokkain: se oppijaryhmä, jonka opetuksessa käytettiin tätä menetelmää, kehittyi eniten kohdekielisen ääntämisen suuntaan. S2-opettajat arvioivat vierasta aksenttia lievemmin kuin luokanopettajat. Kirjoittajat arvelevat tämän johtuvan siitä, että S2-opettajilla on enemmän kokemuksia vieraalla aksentilla puhutusta suomesta.

**Martina Huhtamäki ja Elisabeth Zetterholm** keskittyvät tapaustutkimuksessaan ääntämisopetukseen ruotsi vieraana kielenä -tilanteissa. He vertailivat aikuisopiskelijoille suunnattua ääntämisopetusta Suomessa ja Ruotsissa. Aineisto kerättiin kahdeksaa opettajaa haastattelemalla ja oppituntia seuraamalla. Tavoitteena oli myös verrata, kuinka se, mitä opettajat sanovat ja mitä he tunneilla tekevät, vastaavat toisiaan. Molemmissa maissa opettajat korostivat ääntämisen opettamisen keskeistä asemaa ruotsin opetuksessa. Opetta-

jat kuitenkin kokivat, etteivät he tienneet monia opetusmetodeja ääntämisen opetukseen. Suurin osa havaitusta ääntämisopetuksesta luokkahuonetilanteissa oli implisiittistä ja integroitu esimerkiksi kielipoin tai sanaston opetukseen. Ruotsalaisissa luokkahuonetilanteissa havaittiin enemmän eksplisiittistä ääntämisen opetusta kuin suomalaisissa. Maiden välillä on ymmärrettävästi yleisestikin eroja, koska Suomessa ruotsin vähemmistöasema tekee opettamisesta erilaista ja usein haasteellisempaa kuin Ruotsissa. Opettajien haastattelussa esiin tuomat mielipiteet vastasivat hyvin heidän toimintaansa luokkahuonetilanteissa. Kirjoittajat korostavat, että opettajankoulutuksessa on syytä kiinnittää huomiota ääntämisopetuksen metodeihin ja opettaa tulevia opettajia vertaamaan opettamaansa kieltä oppilaiden erilaisiin äidinkieliin.

**Leila Kääntä** keskittyy tutkimuksessaan suomenkielisten englanninoppijoiden ääntämisongelmiin luokkahuonetilanteissa. Hän tutki keskustelututkimuksen menetelmin, miten oppilaat ilmaisevat kohdanneensa ääntämisongelmia lukiessaan englantia ääneen. Tehtävän tarkoituksena ei ollut keskittyä suoraan ääntämiseen vaan tarjota mahdollisuus ääntämisen harjoitteluun muun toiminnan rinnalla. Laajemmasta aineistosta valikoitui kymmenen esimerkkiä, joissa oppilas kohtaa ääntämisongelman. Kääntä keskittyy analyysissaan erityisesti korjausten aloitukseen: suoriin kysymyksiin, erilaisten ääntämistapojen yritykseen ja lukemisen keskeyttämiseen. Näillä keinoilla oppilaat koettivat saada opettajalta apua oikeaan ääntämistapaan. Opettajan ja oppilaan interaktion tutkiminen selventää, millaista ääntämisopetukseen liittyvä luokkahuoneinteraktio on. Havainnoista on hyötyä opettajankoulutuksessa. Ne myös osoittavat, kuinka luokkahuoneharjoituksia voidaan käyttää vieraan kielen ääntämisen parantamiseen.

**Irina Spita ja Martina Björklund** tarkastelevat opetuskokeiluun perustuvassa tutkimuksessaan, kuinka venäjän kielen L2-puhujat oppivat nk. emotiivista morfologiaa, tarkemmin diminutiivisuffikseja, alkeiskursseilla, joka perinteisestä opetuksesta poiketen eteni pragmaattis-semanttisten taitojen hallinnasta sananhallinnan suuntaan. Tämä tarkoittaa sitä, että diminutiivisia sanamuotoja esiteltiin ja harjoiteltiin mm. toiminnallisissa dialogeissa, jotka osoittivat konkreettisesti, millaisia asenteita ja tunteita eri sanamuotojen käyttö välittää. Diminutiivimuotoja harjoiteltiin esimerkiksi siten, että opiskelijat perehtyivät venäläisten ravintoloiden ruokalistoihin esiintyviin diminutiivimuotoihin, ja tämän jälkeen harjoittelivat näiden muotojen käyttöä dialogeissa ja roolileikeissä. Kun opiskelijat tunnustivat diminutiivimuodot ja niiden merkityksen, kursseilla tarkasteltiin, millainen merkitys diminutiivimuotojen poistamisella ja/tai korvaamisella on erilaisissa sanoissa. Kirjoittajien mukaan opetuskokeilun tulokset olivat hyviä: emotiivisten muotojen käyttö edisti positiivisen ilmapiirin syntymistä luokkahuoneessa ja oppimistuloksetkin olivat hyviä. Oppijoiden motivaatiota vaikutti edistävän myös se, että opetuksen lähtökohtana oli paremminkin luonnollinen, arjessa käytettävä kieli kuin kankeahko oppikirjakieli.

### 3.2 Suullisen kielitaidon ja ääntämisen oppimisen tutkimukset

Kirjamme tutkimuksista kolme keskittyy suullisen kielitaidon ja ääntämisen oppimiseen (*Heinonen, Peltonen, Salminen*). Näissä keskitytään joko ymmärrettävyyteen, sujuvuuteen tai vieraaseen korostukseen.

**Henna Heinosen** tutkimuksessa tarkastellaan ensikieleltään suomenkielisten lukiolaisten puhutun ruotsin ymmärrettävyyttä sekä mahdollisten kuulijaa tai puhujaa koskevien taustatekijöiden vaikutusta arvioon. Tuloksista käy ilmi, että kuulijan taustalla (esim. kielitaustalla, sukupuoliella ja iällä) ei näyttäisi olevan vaikutusta. Yleisesti ottaen lukiolaisten tuottamaa ruotsia saattoi ymmärtää, mutta se saattoi vaatia kuulijalta keskittymistä. Puhujan ruotsin kielen käyttö vapaa-aikana vaikutti myönteisesti puheen ymmärrettävyyteen, kun taas yllättäen puhujan viimeisimmän ruotsin arvosanan ja puheen ymmärrettävyyden välinen korrelaatio oli negatiivinen. Jälkimmäisen Heinonen arvioi voivan olla tekemisissä ääntämistä koskevan tietoisuuden kanssa: ääntämistä koskevan tiedon lisääntyminen voi vaikuttaa haitallisesti sujuvan puheen tuottamiseen.

**Pauliina Peltonen** yrittää tutkimuksessaan murtaa vieraalla kielellä puhumisen sujuvuutta koskevan tutkimuksen perinteiset rajat ja tarkastella puhetta yhteistoiminnallisen vuorovaikutuksen kautta. Hän tarkastelee kahden yläkoululaisparin puhetta ongelmanratkaisutehtävässä ja analysoi tuloksia. Tulosten mukaan toistot (engl. *other-repetitions*) mahdollistavat puhujan sujuvuutta antamalla hänelle lisää aikaa kielen käsittelyyn. Toistoilla voi myös hakea puhekumppanin vahvistusta omalle ilmaisulle, jolloin ne lisäävät viestinnän koheesiota. Samoin koheesiota kohentavat täydennykset (engl. *collaborative completions*), mutta niiden avulla myös ylläpidetään sujuvuutta, mikä ei aiemmissa tutkimuksissa ole noussut esiin.

**Iina Salmisen** artikkelin kohteena ovat suomalaisille oppijoille perinteisesti vaikeina mainostetut venäjän ässät, jotka foneettisesti jakautuvat sibilantteihin ja affrikaattoihin. Hän tarkastelee eksplisiittisen ääntämisopetuksen vaikutusta oppimistuloksiin niin tuottamisen kuin havaitsemisen osalta sekä kontrastiivisen kielitieteen mahdollisuuksia ääntämisopetuksen kehittämisessä. Hänen tuloksistaan käy ilmi, että reseptiivisen ja produktiivisen osaamisen välillä näkyy vallitsevan yhteys joidenkin äänteiden kohdalla muttei läheskään kaikkien. Eri testikertojen tulosten kirjavuus ja aineiston pienuus vaikeuttaa lopullisten johtopäätösten tekoa – mikä viittaa lisätutkimuksen tarpeeseen – mutta eksplisiittisen ääntämisopetuksen hyödyn tekijä katsoo käyneen jo tässäkin aineistossa ilmi. Jatkoa ajatellen hän korostaa, että tuottamisen harjoittamisen lisäksi aikaa olisi varattava puhujalle vieraiden äänteiden havaitsemiseen ja tunnistamiseen.

### 3.3 Suullisen kielitaidon ja ääntämisen arviointi tai mittaaminen

Pääosa kirjamme tutkimuksista – kuusi artikkelia – käsittelee suullisen kielitaidon/ääntämisen arviointia tai mittaamista (*Honko, Olkkonen & Peltonen, Kallio ym., Kuronen & Tergujeff, Toivola & Ullakonoja, Uzal ym.*). Lähtökohtana näissä tutkimuksissa on – Hongon sekä Olkkosen ja Peltosen tutkimuksia lukuun ottamatta – kuulijoiden arviot toisen kielen ääntämisen hyvyydestä tai ymmärrettävyydestä. Muutamissa tutkimuksista arvioita tarkastellaan suhteessa puheen akustisiin tai rakenteellisiin piirteisiin.

**Mari Honko** tarkastelee tutkimuksessaan puhutun kielen leksikaalista diversiteettiä kouluikäisten lasten kielessä. Tutkimuksessa selvitetään, poikkeako puhutun kielen sadutettujen kertomusten leksikaalinen diversiteetti kirjoitetun kielen kertomusten leksikaalisesta diversiteetistä. Lisäksi Honko selvittää, onko puhuttujen kertomusten leksikaalisessa diversiteetissä systemaattisia eroja riippuen lapsen sanaston hallinnan yleistasosta, kielitaidon yleistasosta, ensikielestä, sukupuolesta tai iästä. Aiempien tutkimusten tapaan Hongonkin tutkimuksessa leksikaalinen diversiteetti oli korkeampaa kirjoitetussa kuin puhutussa kielessä. Kielitaidon yleistason ja sanaston hallinnan yleistason yhteys leksikaaliseen diversiteettiin osoittautui kuitenkin heikoksi ja epälineaariseksi. Honko toteaaakin eräänä johtopäätöksensä, että leksikaalista diversiteettiä ei voi tarkastellun aineiston perusteella pitää riittävän tarkkana menetelmänä silloin, kun tarkoituksena on arvioida yksilötason kielitaidon tai tarkemmin sanastollisten taitojen tasoa.

**Heini Kallio, Juraj Šimko, Ari Huhta, Reima Karhila, Martti Vainio, Erik Lindroos, Raili Hildén ja Mikko Kurimo** tarkastelevat artikkelissaan temporaalisten piirteiden (artikulaationopeus, hiljaisten ja täytettyjen taukojen määrä ja korjaukset ja/tai toistot) suhdetta arvioituun taitotasoon. Puhujina tutkimuksessa on suomalaisia lukiolaisia, jotka puhuvat ruotsia toisena kielellään. Verrokkipuhujina on ensikieleltään (suomen)ruotsalaisia lukiolaisia. Puhenäytteet ovat lukupuhunnasta ja puolispontaanista puheesta, ja arvioijina on seitsemän koulutettua arvioijaa. Tulosten mukaan tietyillä temporaalisilla piirteillä on yhteys puhujan arvioituun taitotasoon, mutta piirteiden merkitys riippuu puhemateriaalista: korjauksia, toistoja ja lyhyitä taukoja siedetään paremmin puolispontaanissa puheesta kuin lukupuhunnassa. Korkea artikulaationopeus ennusti korkeampaa taitotason arviointia kummassakin materiaalityypissä, kun taas täytettyjen taukojen korkea suhteellinen määrä ennusti alhaista arviota taitotasosta. Arvioijien välillä oli kuitenkin eroja siinä, millainen merkitys temporaalisilla piirteillä oli taitotason arvioinnille. Tutkimuksen eräs johtopäätös onkin, että tulisi kehittää menetelmiä arvioijien välisen vaihtelun vähentämiseksi erityisesti nk. korkean panoksen testeissä. Arvioijien profilointi ja siihen perustuva palaute arvioijille saattaisivat edistää arvioijien tietoisuutta eri foneettisista piirteistä ja parantaa heidän arviointisuoritustaan.



**Mikko Kuronen ja Elina Tergujeff** tutkivat artikkelissaan prosodian vaikutusta aksentin vahvuuden arviointiin. He pyysivät syntyperäisiä ruotsin kielen puhujia arvioimaan kielenoppijoiden tuottamaa ruotsia. Tutkimus keskittyi ruotsin tooniaksentteihin eli sävelkorkoihin ja puherytmin muutoksiin. Aineistona oli 30 ääneen luettua virkettä, joissa pitäisi olla tietty tooniaksentti ja pääpaino virkkeen loppupuolella. Virkkeet tuotti kuusi kielenoppijaa ja kaksi syntyperäistä ruotsin puhujaa. Arvioijina oli 53 syntyperäistä ruotsin puhujaa. Tulosten mukaan normista poikkeava puherytmi aiheutti vaikutelman vahvemmassa vierasperäisestä aksentista. Tooniaksenttien taas ei havaittu vaikuttavan merkittävästi aksenttiarvioihin. Kuronen ja Tergujeff tulivat siihen johtopäätökseen, että prosodian opetuksessa kannattaa painottaa erityisesti puhe-rytmiä ja muita prosodisia piirteitä tooniaksenttien sijaan.

**Sanna Olkkonen ja Pauliina Peltonen** tarkastelevat toisen kielen sujuvuudelle annettuja määritelmiä kirjallisuuden valossa ja pyrkivät yhdistämään kaksi tähän asti varsin erillisinä pysytellyttä näkökulmaa, kognitiivisen ja puhetuotoksen sujuvuuden. Näkökulmat liittyvät erilaisiin tutkimusperinteisiin eikä niitä siksi ole tavattu vertailla keskenään. Näkökulmat yhdistämällä kirjoittajat katsovat voivansa taata aiempaa kattavamman kuvan aiheeseen liittyvästä tutkimuksesta, erityisesti auttaa selittämään, miksei joidenkin sujuvuuteen liitettyjen piirteiden välille ole kyetty löytämään yhteyttä. Artikkelillaan kirjoittajat pyrkivät kannustamaan muita tutkijoita ylittämään perinteiset rajat ja suuntautumaan tutkimuskysymykseen monialaisesti.

**Minnaleena Toivola ja Riikka Ullakonoja** keskittyvät artikkelissaan venäläisen aksentin tunnistamiseen suomessa vieraana kielenä. Näytteet (n = 80) olivat ääneen luettuja lauseita äidinkieleltään suomalaisilta puhujilta ja venäläisiltä suomen kielen oppijoilta. Oppijat tuottivat lauseita eri vaiheissa kuuden kuukauden oleskeluaan Suomessa. Kuuntelutestin arvioijat olivat syntyperäisiä suomen puhujia, ei-syntyperäisiä suomen opiskelijoita ja ei-syntyperäisiä kuuntelijoita, jotka eivät opiskelleet suomea. Tulosten perusteella arvioijaryhmät erosivat toisistaan. Paremmiin suomea osaavat ryhmät erottivat venäjänkieliset puhujat helpommin syntyperäisistä. Arvioijaryhmät sanoivat kiinnittäneensä puherytmiä enemmän huomiota muihin prosodisiin ja segmentaalisiin piirteisiin. Kirjoittajat toivovat, että tutkijat testaisivat tulevaisuudessa enemmän vieraan aksentin tunnistamista pitkäkestoisissa tutkimuksissa.

**Melike Uzal, Erkki Komulainen ja Olli Aaltonen** tutkivat kuulijan taustatekijöiden ja tehtävätyypin vaikutusta aksentin vahvuuden arvioinnissa. Kohdekielenä heillä oli suomi vieraana kielenä. Kuulijoiden taustamuuttujista kontrolloitiin sukupuolta, ikää, opetuskokemusta, aiempia kontakteja vieraskieliin suomen puhujiin ja asumista pääkaupunkiseudulla. Tehtävätyypeinä olivat sanaparit, virkkeiden ääneen lukeminen ja spontaani puhe. Tehtävät tuottivat 31 syntyperäistä suomen puhujaa ja 40 kielenoppijaa. Kuuntelutestiin osallistui 66 syntyperäistä suomen puhujaa, jotka arvioivat kuulemansa vieraan ak-

sentin. Tulosten mukaan kuulijoiden taustamuuttujilla ei näyttänyt olevan vaikutusta aksentin arviointiin. Analyysissa havaittiin myös, että näytteen pituus vaikutti aksenttiarvioon: pidempi näyte arvioitiin yleensä heikommin. Samoin spontaanit näytteet arvioitiin heikommiksi kuin kontrolloidummat näytteet.

## 4 Loppupäätelmiä

Kartoituksemme osoittaa, että suullisen kielitaidon ja ääntämisen tutkimus ei ole ollut runsasta AFinLAN tuotannossa aikavälillä 1977–2016: vain reilu kymmenes julkaistun tutkimuksen kokonaismäärästä on käsitellyt näitä teemoja. Suullisen kielitaidon ja ääntämisen osuus ei ole myöskään kasvanut aikavälillä, vaan päinvastoin jonkin verran laskenut. Kartoituksen tuloksia tarkasteltaessa tulee tosin muistaa, että huomattava osa suullisen kielitaidon ja ääntämisen tutkimuksista julkaistaan muilla foorumeilla kuin AFinLAN julkaisuissa, eikä kartoitus siten anna kattavaa kuvaa suullisen kielitaidon ja ääntämisen tutkimuksesta maassamme (jälkimmäisestä ks. tarkemmin Kautonen ym. 2016). Se antaa kuitenkin käsityksen tutkimusalueen suhteellisesta osuudesta AFinLAN julkaisuissa ja laajemmin soveltavan kielitutkimuksen alalla Suomessa.

Soveltavan puheentutkimuksen painopiste on ollut ja on yhä toisen kielin puheessa: kaikki tämänkin julkaisun artikkeleista käsittelevät L2-puhetta. L2-tutkimus antaa kuitenkin samalla tietoa myös L1-puheesta, koska ensikieli toimii lähes aina vertailukohtana L2-puheelle. L2-ääntämisen tutkimuksessa on siirrytty yhä selvemmin ääntämisen oppimisen mittaamisesta siihen, miten kuulijat arvioivat L2-ääntämistä ymmärrettävyyden, sujuvuuden tai vieraan korostuksen kannalta (mm. Levis & LeVelle 2010a; Munro 2010). Painopisteen siirtyminen kuulijahavaintoon näkyy tämänkin julkaisun tutkimuksissa. Ymmärrettävyyden ja sujuvuuden tutkimista puoltaa se, että kommunikatiivisesti toimivaa ääntämistä pidetään nykyään tärkeämpänä kuin mahdollisimman vähäistä vierasta korostusta (mm. Levis 2005; Thorén 2009; Heinonen tässä julkaisussa). Suullisen kielitaidon testaus on yleistynyt ja tekee tuloaan mm. suomalaisen ylioppilastutkintoon (Huhta & Hildén 2016), ja tämänkin vuoksi tarvitaan tietoa siitä, miten kuulija ymmärtää ja kokee L2-puhetta. Kommunikatiivisesti toimivan ääntämisen painotukseen voi kuitenkin suhtautua myös kriittisesti, mikäli se tarkoittaa sitä, että ääntämisen opettamisen tavoitteita tietoisesti lasketaan (vrt. Kjellin 2002). Ääntämisen opetusta koskien opettajan ja oppijan tavoitteet eivät aina tunnukaan kohtaavan (Porter 1999), mistä saattaa olla haitallisia seurauksia opettamisen tehokkuuden näkökulmasta.

Riittävän hyvä suullinen kielitaito vieraissa kielissä on tärkeää lukuisissa ammateissa (EK-tiedustelu 2010), ja se on myös keskeistä tai ainakin mielleltään keskeiseksi mahdollisimman sujuvan maahanmuuttajien kotoutumisen kannalta (mm. Boyd 2004; Tarnanen & Pöyhönen 2011). Näistäkin syistä ään-

tämisen ja suullisen kielitaidon tutkiminen ja tutkimustulosten hyödyntäminen opettajakoulutuksessa ja opetuksessa on tärkeää.

Vaikka kartoituksemme näyttää, että suullisen kielitaidon ja ääntämisen tutkimus ei ole ollut AFinLAN tuotannossa määrällisesti runsasta, erikoisjulkaisumme kuitenkin osoittaa, että tutkimusta tehdään monissa yliopistoissa, tutkijat ovat verkostoituneet kiitettävästi ja Suomessa tehtävä tutkimus seuraa yllä kuvattuja kansainvälisen tutkimuksen painopisteitä.

## Lähteet

- Boyd, S. 2004. Utländska lärare i Sverige: attityder till brytning. Teoksessa K. Hyltens-tam & I. Lindberg (toim.) *Svenska som andraspråk: i forskning, undervisning och samhälle*. Lund: Studentlitteratur, 419–436.
- Derwing, T. M. 2010. Utopian goals for pronunciation teaching. Teoksessa J. Levis & K. LeVelle (toim.) *Proceedings of the 1st pronunciation in second language learning and teaching conference: Iowa State University, Sept. 2009*. Ames (Iowa): Iowa State University, 24–37.
- Derwing, T. M. & M. J. Munro 2005. Second language accent and pronunciation teaching: a research-based approach. *TESOL Quarterly*, 39 (3), 379–397. DOI: 10.2307/3588486.
- 2015. *Pronunciation fundamentals: evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins.
- EK-tiedustelu 2010. *Työelämässä tarvitaan yhä useampia kieliä: EK:n henkilöstö- ja koulutustiedustelu 2009*. Helsinki: Elinkeinoelämän keskusliitto.
- Eklund, H. 2000. Vart är pedagogikforskningen på väg? ämnesområden och forskningsmönster i svenska doktorsavhandlingar under en femårsperiod. *Pedagogisk Forskning i Sverige*, 5 (2), 131–150.
- Huhta, A. & R. Hildén 2016. Kielitutkinnot ja muu laajamittainen kielitaidon arviointi Suomessa. Teoksessa A. Huhta & R. Hildén (toim.) *Kielitaidon arviointitutkimus 2000-luvun Suomessa*. AFinLA-e: soveltavan kielitieteen tutkimuksia 9. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 3–26. <https://journal.fi/afinla/article/view/60844>.
- Kalaja, P., H. Dufva, K. Mäntylä & A. Pitkänen-Huhta 2013. AFinLA: the Finnish association of applied linguistics. *European Journal of Applied Linguistics*, 1 (1), 162–173. DOI: 10.1515/eujal-2013-0008.
- Karlsson, F. 1998. *Kielitieteiden tohtorinväitöskirjat Suomessa 1840–1997*. Helsinki: Helsingin yliopisto.
- 2003. Kielitieteiden tohtorinväitöskirjat Suomessa 1902–2001: määrällisiä suuntauksia. *Virittäjä*, 107 (1), 68–76. <https://journal.fi/virittaja/article/view/40238>.
- Kautonen, M., M. Kuronen & R. Ullakonoja 2016. Studier i uttalsinläring i finska, svenska och engelska: litteraturöversikt. *Puhe ja kieli*, 36 (3), 197–220. <https://journal.fi/pk/article/view/59010>.
- Kjellin, O. 2002. *Uttalet, språket och hjärnan: teori och metodik för språkundervisningen*. Uppsala: Hallgren & Fallgren.
- Levis, J. 2005. Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39 (3), 369–377. DOI: 10.2307/3588485.

- Levis, J. & K. LeVelle 2010a. Pronunciation and intelligibility: an overview of the conference. Teoksessa J. Levis & K. LeVelle (toim.) *Pronunciation and intelligibility: issues in research and practice: proceedings of the 2nd annual pronunciation in second language learning and teaching conference*. Ames (Iowa): Iowa State University, 1–6.
- Levis, J. & K. LeVelle (toim.) 2010b. *Pronunciation and intelligibility: issues in research and practice: proceedings of the 2nd annual pronunciation in second language learning and teaching conference*. Ames (Iowa): Iowa State University.
- Linell, P. 2005. *The written language bias in linguistics: its nature, origins and transformations*. London: Routledge.
- Lintunen, P. 2005. Ääntäminen, tietoisuus ja oppiminen. Teoksessa K. Virta, K. Merenluoto & P. Pöyhönen (toim.) *Ainedidaktiikan ja oppimistutkimuksen haasteet opettajankoulutukselle*. Turku: Turun yliopisto, 342–349.
- 2014. Ääntämisen oppiminen ja opettaminen. Teoksessa P. Pietilä & P. Lintunen (toim.) *Kuinka kieltä opitaan: opas vieraan kielen opettajalle ja opiskelijalle*. Helsinki: Gaudeamus, 165–187.
- LOPS 2003. *Lukion opetussuunnitelman perusteet*. Helsinki: Opetushallitus.
- 2015. *Lukion opetussuunnitelman perusteet*. Helsinki: Opetushallitus.
- Moilanen, J. 2014. *Auffassungen der Deutsch- und Schwedischlehrer über Ausspracheunterricht*. Pro gradu -tutkielma. Jyväskylä: Jyväskylän yliopisto. <http://urn.fi/URN:NBN:fi:jyu-201403241394>.
- Munro, M. J. 2010. Intelligibility: buzzword or buzzworthy? Teoksessa J. Levis & K. LeVelle (toim.) *Pronunciation and intelligibility: issues in research and practice: proceedings of the 2nd annual pronunciation in second language learning and teaching conference*. Ames (Iowa): Iowa State University, 7–16.
- Muutos LOPSiin 2009. *Nuorille tarkoitettun lukiokoulutuksen opetussuunnitelman perusteiden 2003 muuttaminen*. Helsinki: Opetushallitus.
- Norris, J. M. & L. Ortega 2000. Effectiveness of L2 instruction: a research synthesis and quantitative meta-analysis. *Language Learning*, 50 (3), 417–528. DOI: 10.1111/0023-8333.00136.
- Porter, D. 1999. Pronunciation. Teoksessa B. Spolsky (toim.) *Concise encyclopedia of educational linguistics*. Oxford: Pergamon, 647–652.
- Ruoho, M. 2011. *Svensklärares åsikter om, attityder till och erfarenhet av den muntliga kursen Tala och förstå bättre och övning av muntlig språkfärdighet*. Pro gradu -tutkielma. Jyväskylä universitet. <http://urn.fi/URN:NBN:fi:jyu-2011052510918>.
- Suomi, K. 2004. Foneettiset väitökset Suomessa. *Virittäjä*, 108 (4), 587–590. <http://urn.fi/URN:NBN:fi:ELE-1027889>.
- Tarnanen, M. & S. Pöyhönen 2011. Maahanmuuttajien suomen kielen taidon riittävyys ja työllistymisen mahdollisuudet. *Puhe ja kieli*, 31 (4), 139–152. <https://journal.fi/pk/article/view/4750>.
- Tergujeff, E. 2013. *English pronunciation teaching in Finland*. Jyväskylä Studies in Humanities 207. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-5322-5>.
- Thorén, B. 2009. Rätt betoning gör dig förstådd. *Språktidningen*, 1/2009.

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 18–44.

## Leena Maria Heikkola<sup>1</sup> & Jenni Alisaari<sup>2</sup>

<sup>1</sup>Senter for Flerspråkighet (Multiling), Universitetet i Oslo & Åbo Akademi,

<sup>2</sup>Turun yliopisto

## Laulun sanoja lausumalla taitavaksi ääntäjäksi?

The aim of the study was to investigate how different teaching methods, singing, listening to songs and reciting song lyrics, affect the development of pronunciation of Finnish as a foreign language pronunciation. The second objective of the study was to investigate whether future classroom teachers and future Finnish as a second language teachers evaluate the strength of foreign accent differently. The results indicate that reciting song lyrics is the most beneficial for the development of pronunciation. This teaching method is especially useful for beginners' level language learners. The future class teachers gave stricter evaluations of foreign accent than the future Finnish as a second language teachers. Based on the results of the study, it could be argued that reciting song lyrics could be used for teaching pronunciation. Further, the results support previous findings that more experienced listeners give milder evaluations than more unexperienced. Thus, it would be valuable for future teachers to gain experience in listening diverse accents.

**Keywords:** pronunciation, teaching method, second language acquisition

**Asiasanat:** ääntäminen, opetusmenetelmä, toisen kielen oppiminen

## 1 Johdanto

Hyvä ääntämistaito on tärkeä osa toimivia kieli- ja viestintätaitoja (Lintunen 2014; Lord 2005; Morley 1991). Erityisesti kielen suprasegmentaaliset piirteet, eli paino, intonaatio ja rytmi, vaikuttavat usein siihen, millaiseksi äidinkieliiset arvioivat toisen kielen puhujan kielitaidon (Kendrick 1997). Myös puheen temporaaliset tekijät, esimerkiksi taukojen määrät ja kestot, puhe- ja artikulaationopeudet sekä korjaukset ja uudelleen aloitukset, voivat kuulijan mielessä vaikuttaa vieraan aksentin vaikutelmaan (Munro & Derwing 2001).

Burgessin ja Spencerin (2000) mukaan laulaminen ja runojen lausuminen ovat erityisen hyödyllisiä menetelmiä, kun halutaan harjoitella kielen ääntämistä. Tutkimusten mukaan fonologinen erottelukyky saa tukea sävelkorkeuden muutoksista tavujen rajoilla (Schön ym. 2008). Myös yleinen musikaalisuus parantaa äänteellistä hahmottamista (Milovanov ym. 2008, 2010). Sen sijaan ei ole olemassa tutkimustietoa siitä, millä tavalla laulaminen tai rytmikäs runojen lausuminen ilman melodian tukea vaikuttaa erityisesti ääntämisen oppimiseen. Tässä tutkimuksessa tarkastelemme kolmen opetusmenetelmän, laulamisen, laulujen kuuntelun ja laulujen sanojen rytmikkään lausumisen, vaikutusta suomenoppijoiden ääntämiseen. Selvitämme myös, arvioivatko tulevat suomi toisena kielenä -opettajat ja luokanopettajat suomenoppijoiden ääntämistä samalla tavalla.

### 1.1 Ääntäminen osana kielitaitoa

Ääntämisen oppiminen ja opettaminen ovat tärkeä osa kielenopetusta, koska ääntäminen vaikuttaa merkittävästi puheen ymmärrettävyyteen (Munro & Derwing 2001). Kohdekielestä poikkeava ääntäminen saattaa lisäksi vaikuttaa negatiivisesti kuulijan asenteeseen puhujaa ja tämän kielitaitoa kohtaan (Toivola 2011). Leinosen (2015: 138) mukaan vahva vieras aksentti voidaan yhdistää ”epäsuomalaisuuteen” eli siihen, että vieraalla aksentilla suomea puhuva ei välttämättä ole omaksunut suomalaisia arvoja työssäkäymisestä, koulutuksesta, luotettavuudesta, rehellisyydestä ja ahkeruudesta.

Kielenopetuksen tavoitteena ei useinkaan ole natiivinomaisen vaan ymmärrettävän ääntämisen oppiminen (Morley 1991). Yksittäinen opiskelija saattaa kuitenkin pitää täydellistä ääntämistä tärkeänä osana kielitaitoaan (Toivola 2011). Yleisesti ajatellaan, että varhainen toisen kielen oppiminen johtaa varmemmin natiivinomaiseen ääntämiseen. Natiivinomainen ääntämys on kuitenkin mahdollista saavuttaa myös aikuisiässä (Abu-rabia & Kehat 2004; Bongaerts ym. 2000).

Ääntämisen opettamisesta ja oppimisesta on suhteellisen vähän tutkimuksia (Derwing & Munro 2005; Lord 2005; Tergujeff 2013), eikä varsinkaan laulujen opetus käytön yhteyttä ääntämisen oppimiseen ole juuri tutkittu. Tiede-

tään kuitenkin, että ääntämiseen on kiinnitettävä huomiota, jotta ääntämisen oppiminen olisi tehokasta (Derwing & Munro 2005; Schmidt 1990). Eriytyisen tehokkaita ovat harjoitukset, joissa huomio suunnataan kohdekielen ja oman äidinkielen välisiin äänne-eroihin. (Bongaerts 1999.) Mikäli kielenoppija havainnoi ja analysoi tietoisesti omaa ääntämystään, hänen ääntämistaitonsa kehittyvät enemmän natiivinomaiseksi (Lord 2005). Tärkeää on myös, että oppimisilmapiiri on kannustava ja turvallinen (Morley 1991).

Kendrick (1997) on osoittanut, että puheen prosodiset piirteet vaikuttavat siihen, miten kuulijat arvioivat kohdekielen oppijoiden ääntämistaitoa, ja siksi erityisesti intonaatiota, painoa ja rytmiä on syytä harjoitella kielen opetuksen yhteydessä. Tämän voi tehdä muun muassa sanoja rytmikkäästi taputtamalla, melodiaa hyräilemällä, laulamalla ja laulun sanoja rytmikkäästi lausumalla (Burgess & Spencer 2000). Burgessin ja Spencerin (2000) mukaan englannin kielen ääntämisen opettamisessa hyödynnetään paljon erilaisia tähän tarkoitukseen tehtyjä rytmikkäitä tekstejä (engl. *chants*), mutta suomen kielen ääntämisen harjoitteluun tällaisia ei juuri ole tarjolla. Kendrickin (1997) mukaan ääntämisen harjoittelussa tärkeintä on rohkaista oppijoita tuottamaan kieltä suullisesti niin paljon kuin mahdollista. Tässä laulaminen ja erilaisten tekstien rytmikäs lausuminen ovat hyviä menetelmiä.

## 1.2 Laulaminen, laulujen kuunteleminen ja laulun sanojen lausuminen kielen oppimisen ja ääntämisen tukena

Kieltä opitaan parhaiten, kun oppijan on mahdollista saada ympäristöstään käyttöönsä ymmärrettäviä kielellisiä aineksia eli tarjoumia. Oppiminen on optimaalista, kun oppija lisäksi kohdistaa huomionsa näihin tarjoumiin sekä tuottaa kieltä itse. (Van Lier 1998.) Lauluja kuunneltaessa näistä elementeistä läsnä on sanoitusten tarjoumat ja mahdollisesti myös aktiivinen huomion kohdistaminen kielen piirteisiin. Laulettaessa tai lausuttaessa rytmikkäästi laulun sanoja oppija saa tarjoumia oppimiselleen laulujen sanoituksista, ja hän tuottaa itse kieltä ja kohdistaa samalla aktiivisesti huomiota kielen sanoihin ja rakenteisiin (Alisaari 2016). Kaikissa tapauksissa kielenoppija voi valita laulun sanoista itselleen merkityksellisiä, oppimista edistäviä tarjoumia (vrt. van Lier 1998).

Laulamalla tai lausumalla sanoja voidaan kasvattaa kielen tuottamisen mahdollisuuksia (Alisaari 2016), mikä on Swainin (2000) mukaan keskeistä kielenoppimisessa. Mikäli oppija samanaikaisesti kuulee joko laulun laulettuna tallenteena tai opettajan laulamana tai runon tai laulun sanoituksen rytmikkäästi opettajan lausumana, hän saa myös tukea ääntämiselleen. Schönin ja muiden (2008) mukaan laulujen sävelkorkeus lisää kuulijan tarkkaavaisuuden tasoa. Laulamalla voidaan lisäksi vähentää kielenoppijoiden tuntemaa ahdistusta kielen oppimisesta ja tuottamisesta (ks. Medina 2002), sillä laulaminen on todettu vähentävän ahdistusta (Grape ym. 2003; Huron 2006). Sama vaiku-

tus voisi olla myös sanoitusten rytmikkäällä lausumisella. Mikäli laulut ja lorut ovat kielenoppijan mielestä mielenkiintoisia (vrt. Krashen 1982), kielen tuottamisen herättämä ahdistus saattaa olla lievempi. Ahdistusta lievittänee lisäksi se, että lauletaessa tai sanoituksia tai loruja lausuttaessa kenenkään ääni ja kielen tuottaminen eivät ole yksin toisten huomion kohteena (vrt. Phillips 1992) ja painotus on enemmän merkityksessä kuin muodossa (vrt. VanPatten & Glass 1998).

Kansainvälisten tutkimusten mukaan opettajat käyttävät laulamista ja rytmikstä lausumista jonkin verran ääntämisen opettamiseen (esim. Breikreutz ym. 2001; Burgess & Spencer 2000; Hismanoglu & Hismanoglu 2010). Suomen kielen opettajat taas pitävät laulamista, laulujen kuuntelua ja runojen lausumista hyödyllisinä keinoina ääntämisen opettamisessa, mutta käytännössä he käyttävät niitä kuitenkin vain harvoin (Alisaari & Heikkola 2017). Vaikka laulujen käytön on todettu tukevan kielenoppimista, ei laulamisen, laulujen kuuntelun ja laulun sanojen rytmisen lausumisen yhteyttä ääntämiseen ole kuitenkaan aiemmin tutkittu.

### 1.3 Vieras aksentti

Kielenoppijoiden ääntämistä tutkittaessa on usein tarkasteltu vierasta aksenttia. Vieras aksentti määritellään usein puheen poikkeaviksi piirteiksi, jotka eroavat äidinkielen puhujan ääntämisestä (Scovel 1969). Arviot vieraasta aksentista ovat kuitenkin aina subjektiivisia, koska niihin vaikuttavat kuuntelijoiden arvot ja asenteet sekä käsitykset itsestään (ks. esim. Munro 2003). Aikuisilla kielenoppijoilla vieras aksentti on yleistä ja normaalia (Munro 2008: 194), eikä nykyään kielenoppimisessa pyritäkään natiivinomaiseen ääntämiseen.

Kielenoppijoiden suhtautuminen ääntämiseen ja vieraaseen aksenttiin vaihtelee. Vieraan aksentin säilyttäminen voi olla tietoinen valinta, ja esimerkiksi maahanmuuttajanuorten on osoitettu lainaavan vieraan aksentin piirteitä toiselta kieliryhmältä ryhmään kuulumisen ja oman identiteetin korostamiseksi (ks. esim. Nortier & Dorleijnb 2008). Toisaalta kielenoppijat voivat kokea oman vieraan aksenttinsa haittaavan ymmärretyksi tulemistä ja siten myös kommunikaatiota. Vieras aksentti ei ole vain kielenoppijan haaste, vaan sillä on myös yhteiskunnallista merkitystä: vieras aksentti voi usein olla leimavaa, mikä voi vaikuttaa kielenoppijan asemaan uudessa maassa. Ääntämisen harjoittelussa tulisi lähtökohtana pitää kielenoppijan omaa näkemystä ääntämisestään (Derwing 2003).

Äidinkielen puhuja erottaa hyvinkin lyhyistä puhunnoksista mahdollisen vieraan aksentin. Tämän vuoksi vierasta aksenttia on tutkittu lähinnä kuuntelukokeiden avulla, minkä lisäksi on tarkasteltu myös vieraaseen aksenttiin liittyviä akustisia piirteitä. (Toivola 2011.) Vierasta aksenttia on tutkittu sekä yksittäisten äänteiden (ks. esim. Riney ym. 2005) että laajempien kokonaisuuk-



sien suhteen (ks. esim. Kang ym. 2010).

Aiemmissa tutkimuksissa ei useinkaan ole kiinnitetty huomiota siihen, millä perusteella arvioitsija tekee arvionsa (ks. kuitenkin Toivola 2011). Vierasta aksenttia määrittelevät siihen liittyvät segmenttaaliset ja suprasegmenttaaliset kielen piirteet. Vieraan aksentin arvioon voivat vaikuttaa myös sujuvuuden muuttajat, kuten tauot, epäröinti, toistaminen, väärät aloitukset sekä puhenopeus. (Piske ym. 2001.) Sujuvuuden lisäksi myös ymmärrettävyys (engl. *intelligibility*, *comprehensibility*) voi olla yhteydessä vieraaseen aksenttiin (Anderson-Hsieh ym. 1992). Munron ja Derwingin (2001) mukaan vieras aksentti, sujuvuus ja ymmärrettävyys ovat erillisiä tekijöitä (Munro & Derwing 2001), vaikkakin näiden kolmen käsitteen on myös osoitettu tilastollisesti korreloivan keskenään. Sujuvuuden ja ymmärrettävyyden sekä sujuvuuden ja vieraan aksentin vahvuuden välillä on osoitettu korrelaatio (Derwing ym. 2004). Toivolan (2011: 16) mukaan näitä kolmea toisen kielen oppimisen osa-aluetta tulisi kuitenkin tutkia erillään. Tässä tutkimuksessa tarkastellaan edellä mainittua vierasta aksenttia.

#### 1.4 Vieraan aksentin arviointi

Vieraan aksentin arvioinnissa on yleisesti käytetty aineistonkeruumenetelmänä kuuntelukokeita. Seuraavaksi esitellään vieraan aksentin arviointia puhujien, kuuntelijoiden sekä kuuntelukokeiden ja niiden arvioinnin osalta.

Monissa tutkimuksissa on osoitettu, että nuoremmat maahanmuuttajat oppivat vieraan kielen ääntämisen paremmin kuin vanhemmat maahanmuuttajat (ks. esim. Flege 1991; Flege ym. 1995). Kuitenkin myös aikuisina toista kieltä oppivat ovat tutkimusten mukaan voineet saavuttaa natiivinomaisen ääntämyksen (Bongaerts ym. 2000; Muñoz & Singleton 2007). Kuten jo aiemmin totesimme, erittäin hyvän, natiivinomaisen ääntämisen tason taustalla vaikuttavat kielenoppijan hyvä motivaatio, intensiiviset havaitsemis- ja tuottoharjoitukset sekä erittäin suuri määrä kohdekielen syötettä (Bongaerts 1999). Myös imitoimiskyvyllä (Markham 1997) ja musikaalisuudella (Milovanov ym. 2008, 2010) vaikuttaisi olevan merkitystä hyvän ääntämisen oppimisessa. Lisäksi kielenoppimisen alkuvaiheessa myös maassaoloajalla on osoitettu olevan merkitystä (Piske ym. 2001).

Vieraan aksentin tutkimuksissa erilaisten kuuntelijoiden arviot eivät ole eronneet toisistaan ratkaisevasti. On kuitenkin ehdotettu, että arvioinneissa kannattaa käyttää erilaisia kuuntelijaryhmiä (Flege & Fletcher 1992; Piske ym. 2001), jolloin taustatekijät eivät vaikuta arvioon voimakkaasti. Kuuntelukokeiden arvioijien määrät ovat vaihdelleet yhdestä satoihin (ks. esim. Piske ym. 2001).

Tutkimustulokset kuuntelijoiden arvioiden ja vieraan aksentin yhteyksistä ovat ristiriitaisia. Kokeneet kuulijat hyväksyvät kuitenkin enemmän virhei-

tä ja antavat lievempiä aksenttiarvioita (Thompson 1991: 198). Toisaalta on esitetty, että kokemattomat kuuntelijat tunnistavat vieraan aksentin yhtä luotettavasti kuin asiantuntija-arvioijat (Kennedy & Trofimovich 2008). Tutkimusten valossa vaikuttaa myös siltä, että omassa äidinkielessä kuultua vierasta aksenttia arvioidaan ankarammin kuin muissa kielissä (Munro ym. 2006). Vieraiden aksenttien kuunteluun tottuneet arvioivat aksenttia lempeämmin kuin kokemattomimmat kuulijat (Ahola 2016). Puhujan äidinkielen tunnistaminen vaatii kuitenkin paljon kokemusta kyseisestä vieraasta aksentista (Flege 1984).

### 1.5 Suomi toisena kielenä -oppijoiden ääntäminen

Viime vuosina on yhä enemmän tutkittu S2-oppijoiden ääntämistä (ks. esim. Leinonen 2015; Ullakonoja & Kuronen 2015). Suurin yksittäinen empiirinen tutkimusprojekti S2-oppijoiden ääntämisestä on tehty Proof-projektissa (Aho ym. 2016), jossa on tutkittu Suomen suurimpia maahanmuuttajakieliä ja jonka tavoitteena on ollut tuottaa tietoa S2-opetusta varten. Esimerkiksi Toivola (2011) tutki väitöskirjassaan venäjänkielisten S2-oppijoiden vierasta aksenttia lukupuhunnassa ja spontaanissa puheessa. Lukupuhunnalle annettiin raatiarviossa enemmän vahvan vieraan aksentin arvioita kuin spontaanille puheelle. Akustisessa analyysissä nousi esille, että venäjänkielisten puhujien vahvaa aksenttia selittävät osittain yksittäiset poikkeavat äänneet ja osittain temporaaliset tekijät, jotka usein liitetään puheen sujuvuuteen. Leinonen (2015) tutki väitöskirjassaan nuorten arvioita vieraalla aksentilla tuotetusta suomesta. Ääntämistä arvioitiin raatiarvioinnilla (n = 177). Kaksi eri raatiarvioi S2-puhujien vierasta aksenttia lukupuhunnasta ja spontaanista puheesta. Yllättävästi S2-puhujien spontaanin puheen vieras aksentti arvioitiin pääosin voimakkaamaksi kuin lukupuhunnan aksentti. Tähän on voinut vaikuttaa myös kielitaitotasoon viittaavat seikat, kuten pidemmät tauot ja epäröinnit.

Lisäksi on tutkittu suomenruotsalaisten suomen ääntämistä raatiarviolla sekä akustisella analyysillä (Kuronen 2016). Tutkimus tehtiin osana projektia ”Fokus på uttalsinlärningen med svenska som mål- och källspråk” (<https://www.jyu.fi/fokus>). Kurosen tutkimuksessa analysoitiin suomea natiivinomaisesti puhuvien suomenruotsalaisten ja muunkielisten ääntämistä (n = 11). Koehenkilöiden puheesta teetettiin ensin raatiarvio, jossa 93 suomenkielistä henkilöä arvioi puhujien suomen ääntämistä. Raatiarvion perusteella S2-puhujat erosivat äidinkielistä puhujista. Ryhmänä suomenruotsalaiset olivat ääntämisensä perusteella tunnistettavan aksentin rajalla, mutta yksilöllinen variaatio oli suurta. Muunkielisillä S2-puhujilla oli selkeämpi, tunnistettavampi vieras aksentti kuin suomenruotsalaisilla. Tämä ero ryhmien välillä selittyi osittain suomenoppimisillä.

Kuronen (2016) vertasi S2-puhujien puhetta myös äidinkielisten suomen puhujien puheeseen akustisesti keskittyen erityisesti vokaaleihin. S2-puhujilla

esiintyi vain vähän segmentaalisia aksenttipiirteitä, mutta heillä oli ongelmia tavupituuksien ja intonaation suhteen. Natiivinkaltaisten S2-puhujien suurimmat ongelmat olivat siis prosodiassa. Erityisesti pitkien sanojen suhteen S2-puhujat erosivat akustisen analyysin perusteella äidinkielistä suomen puhujista. Useimmiten suomenruotsalaiset käyttivät kullekin kielelle tyypillisiä vokaaleita puheessa.

## 2 Tutkimuskysymykset ja tutkimuksen tarkoitus

Tutkimuksemme tarkoituksena on selvittää, miten kolme eri opetusmenetelmää, laulaminen, laulujen kuuntelu ja laulun sanojen rytmien lausuminen, vaikuttavat siihen, miten suomen kielen ääntäminen kehittyy. Lisäksi haluamme tarkastella, hyötyvätkö alkeistason kielenoppijat näiden menetelmien käytöstä enemmän kuin alkeisjatkotason opiskelijat. Ääntämisen kehittymistä tutkitaan raatiarvioinnilla. Näillä kolmella eri opetusmenetelmällä opettujen opiskelijoiden (n = 63) suullisia kertomuksia kuuntelutettiin kahdella arviointiraadilla (n = 109), jotka arvioivat koehenkilöiden vieraan aksentin vahvuutta. Toinen raati koostui tulevista suomi toisena kielenä -opettajista (S2-raati) ja toinen luokanopettajaopiskelijoista (OKL-raati). Tutkimuksessa tarkastellaan myös sitä, miten nämä kaksi raatia eroavat arvioissaan.

Tutkimus on siis kaksiosainen: ensimmäisessä vaiheessa on kerätty puheaineistoa kurssin osallistujilta alku- ja lopputestissä; toisessa vaiheessa kaksi raatia arvioi kurssilaisten ääntämistä alku- ja lopputestistä irrotetuista puhenäytteistä. Pyrimme vastaamaan seuraaviin tutkimuskysymyksiin:

- 1a) Miten eri opetusmenetelmät, laulaminen, laulujen kuuntelu ja laulujen sanojen rytmikäs lausuminen, vaikuttavat suomenoppijoiden ääntämisen kehittymiseen kurssin aikana, alkutestistä lopputestiin?
- 1b) Hyötyvätkö alkeistason ja alkeisjatkotason suomenoppijat eri tavoin kolmen tutkimuksessa tarkastellun menetelmän käytöstä?
- 2) Eroavatko S2-raadin ja OKL-raadin arviot suomenoppijoiden ääntämisestä?

Olemme aiemmin osoittaneet, että laulaminen kehittää suomen kielen kirjoittamisen sujuvuutta eniten kolmesta tarkastellusta opetusmenetelmästä (Alisaari & Heikkola 2016). Tämän pohjalta odotamme, että lauluryhmällä suomen kielen ääntäminen arvioidaan lähemmäksi natiivinomaista suomen ääntämistä kuin muilla ryhmillä. Erityisesti alkeistason kielenoppijat hyötyvät laulujen käytöstä opetuksessa (Alisaari & Heikkola tekeillä). Oletamme, että ääntäminen kehittyy enemmän alkeistason kielenoppijoilla kuin alkeisjatkotason oppijoilla.

TAULUKKO 1. Opetusmenetelmäryhmien opiskelijamäärät (n = 63).

	Laulu	Kuuntelu	Loruilu
Kurssi I	10	9	10
Kurssi IIA	11	11	12
Yhteensä	21	20	22

Oletamme myös, että OKL-raati arvioi suomenoppijoiden vieraan aksenttin vahvemmaksi kuin S2-raati. Oletuksemme perustuu siihen, että kokoneempien arvioijien on tutkimuksessa todettu olevan kokemattompia lempeämpiä arvioissaan (Ahola 2016; Weigle 1999). Kuuntelijat, jotka ovat tottuneet kuulemaan maahanmuuttajien puhetta, hyväksyvät enemmän virheitä ja antavat aksentista lievempiä arvioita kuin kuuntelijat, jotka eivät ole tottuneet maahanmuuttajien kielen kuunteluun (Thompson 1991).

### 3 Tutkimusaineisto ja -menetelmät

Tutkimuksen ensimmäisen vaiheen aineisto kerättiin CIMOn (Kansainvälisen liikkuvuuden ja yhteistyön keskus) kahdelta suomen kielen intensiivikurssilta (kurssit I ja IIA). Kurssit järjestettiin Suomessa, ja ne kestivät noin neljä viikkoa. Kurssit olivat suunnattuja alkeistason suomen oppijoille, ja osallistujat opiskelivat suomea ulkomaisissa yliopistoissa. Opiskelijat olivat suurimmaksi osaksi A2-tasolla (Euroopan neuvosto 2012) olevia nuoria aikuisia. Kurssilaisista 63 osallistui käsillä olevaan tutkimukseen. Olemme kuvanneet kurseja ja niille osallistuvia opiskelijoita tarkemmin aiemmassa artikkelissamme (Alisaari & Heikkola 2016).

Kursseilla opetettiin kaikkiaan kuutta eri opetusryhmää (lauluryhmä I, kuunteluryhmä I, rytmiryhmä I, lauluryhmä IIA, kuunteluryhmä IIA ja rytmiryhmä IIA), joiden opetus ei eronnut toisistaan muilta kuin tutkimukseen liittyvin osin (ks. taulukko 1). Tutkimusta varten ryhmille opetettiin suomea tasaisin väliajoin neljän viikon kurssin aikana seitsemänä päivänä, 15 minuuttia kerrallaan. Lauluryhmien kanssa suomea opeteltiin laulaen, kuunteluryhmien kanssa lauluja kuunnellen ja rytmiryhmien kanssa laulujen tekstejä rytmikkäästi lausuen. Kaikkien ryhmien kanssa käytettiin samoja lauluja, eli erot ryhmien välillä syntyivät vain laulujen käyttötavasta. Tutkimuksen ensimmäisen vaiheen aineisto koostuu opiskelijoiden tutkijoille kertomista, kuvasarjaan perustuvista tarinoista. Samaa kuvasarjaa käytettiin sekä alku- että loppumittauksessa kurssien ensimmäisenä ja viimeisenä päivänä. Kertomukset tallennettiin digitaalisesti ja litteroitiin.

Raatiarviointia varten ensimmäisen vaiheen osallistujien kertomuksista irrotettiin sekä alku- että loppumittauksen ensimmäiset koherentit puhunnokset. Aiemmissa tutkimuksissa kuuntelukokeiden puhenäytteinä on käytetty luettuja yksittäisiä sanoja ja virkkeitä tai pitempiä yhtenäisiä tekstejä. Puhenäytteinä on käytetty myös spontaania puhetta haastatteluilta, kuvan tai sarjakuvan kerronnasta tai monologeista. Spontaanista puheesta varsinaiseen kuuntelukokeeseen valitaan useimmiten 10–20 sekunnin mittaisia näytteitä (ks. katsaus Jesney 2004). Tässä tutkimuksessa puhunnokset olivat kestoaltaan 10 sekuntia, ja ärsykkeiden välinen aika (engl. ISI = *inter stimulus interval*) oli 5 sekuntia. Ärsykkeet jaettiin neljään osaan, jotka olivat kahdeksan minuutin mittaisia. Kaikissa osissa oli 32 arvioitavaa puhunnosta. Osakokeessa 3 puhunnoksia oli kuitenkin vain 31, koska yhdellä ensimmäisen vaiheen osallistujalla (IL10) ei ollut suullista lopputestiä. Kaiken kaikkiaan raatilaiset arvioivat siis yhteensä 63 suomen kielen kesäkurssilaisen loppu- ja alkutestin puhunnoksia.

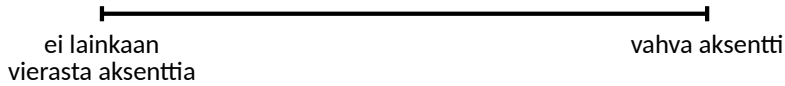
Tutkimuksen toisen vaiheen raatioarviointia varten kuuntelutettiin ensimmäisessä vaiheessa kerättyä aineistoa kahdella arviointiraadilla (n = 109). Raadit arvioivat kielenoppijoiden kurssin alun ja lopun puhenäytteistä vierasta aksenttia. Toinen raati koostui tulevista suomi toisena kielenä -opettajista (n = 39) ja toinen luokanopettajaopiskelijoista (n = 70). S2-opettajat kuulevat työssään paljon suomenoppijoiden puhetta. Myös yhä useampi luokanopettaja tulee kohtaamaan monikielisiä opiskelijoita, koska Suomen väestönkasvu on maahanmuuton varassa (Suomen virallinen tilasto 2017). Koska ääntämisen arvio saattaa vaikuttaa yleisemmin koko kielitaidon arviointiin, on opettajien koulutuksen vuoksi tärkeää saada tutkimustietoa siitä, miten nämä ryhmät arvioivat suomenoppijoiden ääntämistä. Raadeiksi valittiin suomi toisena kielenä -kurssin ja OKL-kurssin opiskelijat. Molemmilla kursseilla oli noin 70 opiskelijaa, mutta S2-kurssin osallistumisprosentti raatiarviointikerralla oli alhainen. Tulosten luotettavuuden parantamiseksi sisällytimme kaikki raatiarviointeihin osallistuneet tutkimukseen (ks. Markham 1997).

S2-raatilaista kaksitoista ilmoitti olevansa jo S2-opettajia. Ryhmien erottamiseksi raatilaiset on koodattu etuosalla S2 ja OKL. Kaikki paitsi kaksi raatilaista suostuivat osallistumaan tutkimukseen. Kolmelta raatilaistelta puuttuu kuuntelukokeen ensimmäinen osa. Näiden raatilaisten muut arviot ovat kuitenkin mukana tarkastelussa. Kaikkiaan tässä tutkimuksessa tarkastellaan siis 107 raatilaisten, 39 S2-raatilaisten ja 68 OKL-raatilaisten arvioita. Raatien taustatiedot kerättiin esitietolomakkeella, ja ne on esitetty taulukossa 2. Ryhmät eivät eronneet taustatietojen suhteen toisistaan. Kaikki raatilaiset ilmoittivat olevansa normaalikuuloisia. Raatien arviot esitetään tässä tutkimuksessa ryhmätasolla, eikä raatilaista siis yksilöidä tulosten raportoinnissa.

Raatilaisille kerrottiin tutkimuksesta suullisesti ja kirjallisesti, ja heitä pyydettiin täyttämään esitietolomake. Ennen kuuntelua kuuntelukokeen ohjeistus käytiin läpi suullisesti ja kirjallisesti raatilaisten kanssa. Koska kuunteluko-

TAULUKKO 2. Raatien taustatiedot.

	S2-raati (n = 39)		OKL-raati (n = 68)		t-testin p-arvo
	n	%	n	%	
Ikäryhmä					0,30
20–24 vuotta	24	61,5	43	62,3	
25–29 vuotta	9	23,1	19	27,9	
30–34 vuotta	1	2,6	3	4,4	
35–39 vuotta	3	7,7	3	4,4	
40–44 vuotta	2	5,1	0	0,0	
Sukupuoli					0,08
Nainen	32	82,1	56	82,4	
Mies	7	17,9	11	16,2	
Muu	0	0,0	1	1,5	
Äidinkieli					0,12
Suomi	36	92,3	67	98,5	
Ruotsi	0	0,0	1	1,5	
Muu (serbia, ukraina, venäjä)	3	7,7	0	0,0	
Fonetikan tai yleisen kielitieteen opiskelija					0,09
Kyllä	4	10,3	1	1,5	
Ei	35	89,7	67	98,5	
Päivittäin yhteydessä maahanmuuttajataustaisiin					0,39
Kyllä	5	12,8	5	7,4	
Ei	34	87,2	63	92,6	
Puhuu muita kieliä kuin äidinkieltään					0,45
Kyllä	39	100,0	67	98,5	
Ei	0	0,0	1	1,5	
On asunut muualla kuin Suomessa					0,38
Kyllä	9	23,1	11	16,2	
Ei	30	76,9	57	83,8	



Mitä arvelet puhujan äidinkieleksi?

Miksi arvelet näin?

KUVIO 1. Kuuntelukokeen VAS-arviointiasteikko ja kysymykset.

keessa annetuilla ohjeilla ja muilla tilanteeseen liittyvillä tekijöillä voi olla vaikutusta arvioihin, nämä olivat samat molemmille raadeille. Kaikkia tekijöitä on kuitenkin mahdotonta täydellisesti kontrolloida. Kuuntelukoe soitettiin raatilaisille miellyttävällä äänenvoimakkuudella isossa luentosalissa, ja raatilaiset merkitsivät arvionsa vastauspaperiin. Osien välillä raatilaiset saivat pitää lyhyen tauon. Kuuntelukokeen kaikki kuunneltavat puhunnokset oli järjestykseltään satunnaistettu ja ärsykkeet esitettiin raadeille osissa (engl. *blocks*), koska samassa järjestyksessä esitettyjen puhenäytteiden arvioinnissa väsyminen voi vaikuttaa arviointeihin (Major 2007). Lisäksi neljä osaa esitettiin eri järjestyksessä kahdelle raadille, jotta tuloksista voidaan sulkea pois väsymyksen vaikutus (ks. Munro & Mann 2005). Puhenäytteiden kuunteluttamiseen luentosalissa päädyttiin osittain suurten raatikokojen varmistamiseksi ja osittain siksi, että raatiarvioinnin kautta toisen kielen opetus saatiin tuotua osaksi sekä tulevien S2-opettajien että OKL:n opiskelijoiden meneillään olevia kursseja. Luentosalien äänentoisto oli erinomainen, ja se testattiin ennen raatiarviointia.

Raatilaiset arvioivat kurssille osallistujien ääntämistä VAS-asteikolla (ks. kuvio 1), koska sillä voidaan mitata erittäin pieniä eroja (Wewers & Lowe 1990). Lisäksi raatilaisia pyydettiin kertomaan, minkä kieliseksi he puhujan arvelivat ja miksi. Koska tutkimuksen päätavoitteena on tarkastella suomenoppijoiden ääntämisen kehittymistä, raatilaisia pyydettiin keskittymään pääasiassa vieraan aksentin arviointiin VAS-asteikolla ja vastaamaan kysymyksiin puhujan äidinkielestä vain, jos heillä jäi siihen aikaa. Raatilaisia pyydettiin merkitsemään arvionsa vieraasta aksentista VAS-asteikolla, vaikka he olisivat olleet arviostaan epävarmoja. Joissain tapauksissa raatilaiset olivat kuitenkin jättäneet joitain kohtia tyhjiksi, mikä voi johtua siitä, että kuuntelu tapahtui luentosalissa. Paperin säästämiseksi vastauslomakkeet tulostettiin 4 sivua yhdelle A4-paperille, mikä on voinut vaikuttaa mittausten tarkkuuteen. Arviot mitattiin viivoittimella, minkä jälkeen tulokset muokattiin vastaamaan yleisesti käytettyä 0–100 millimetrin kokoista asteikkoa tulosten vertailemiseksi aiempiin tutkimustuloksiin.

Raatiarvioinnin tuloksia testattiin tilastollisesti. Opetusryhmien (laulaminen, laulujen kuuntelu ja laulujen sanojen rytminen lausuminen) sisäisiä ja niiden välisiä eroja vieraan aksentin muutoksessa kurssin alusta kurssiin lop-

puun tarkasteltiin t-testein ja efektikoin (Cohenin d). Lisäksi t-testein tutkittiin opetusryhmien sisäisiä ja niiden välisiä eroja vieraan aksentin muutoksissa eri kursseilla (I ja IIA). Myös raatien välisiä eroja vieraan aksentin arvioinneissa eri opetusryhmien ja eri kurssien välillä vertailtiin t-testien ja efektikokojen avulla. t-testit kertovat, eroavatko ryhmät tilastollisesti merkitsevästi, ja efektikoot kertovat, kuinka suuri tämä ero on. Lisäksi raatien välisiä eroja analysoitiin koko aineiston sekä kahden kurssin osalta erikseen toistomittausten ANOVA-testeillä, joilla voidaan selvittää, ovatko raati ja vieraan aksentin arviot merkitsevästi yhteydessä toisiinsa. Myös ANOVA-testien yhteydessä tarkasteltiin efektikokoja (ositettu eetan neliö), jotka kertovat näiden yhteyksien vahvuudesta.

## 4 Tulokset

### 4.1 Ääntämisen kehittyminen eri opetusmenetelmäryhmissä

Aluksi käsitellään kaikkien raatilaisten arvioita kolmen opetusmenetelmäryhmän vieraan aksentin voimakkuudesta (0–100), erottelematta kahta raatia toisistaan. Tällä pyritään vastaamaan tutkimuskysymykseen 1a: Miten eri opetusmenetelmät vaikuttavat suomenoppijoiden ääntämisen kehittymiseen kurssin aikana alkutestistä lopputestiin? Ensin tutkitaan mahdollisia eroja vieraan aksentin arvioissa eri opetusmenetelmäryhmien välillä erottelematta kurssitasoja toisistaan.

Vieras aksentti oli arvioitu lauluryhmällä kaikista voimakkaimmaksi ja kuunteluryhmällä toiseksi voimakkaimmaksi sekä alku- että lopputestissä. Rytmiryhmällä ääntäminen oli raportoitu kaikista lähimmäksi natiivinomaista. Ääntämisen muutos natiivinomaisemmaksi oli merkitsevä ainoastaan rytmiryhmällä. Rytmiryhmällä keskisuuri efektikoko viittasi myös siihen, että ääntämisen kehittyminen paremmaksi on huomattavaa. Tulokset on esitetty taulukossa 3.

Seuraavaksi analysoitiin eri opetusmenetelmien välisiä eroja vieraan aksentin muutoksessa. Kuunteluryhmä ja rytmiryhmä sekä lauluryhmä ja rytmiryhmä erosivat toisistaan merkitsevästi vieraan aksentin muutoksen suhteen, koska rytmiryhmällä ääntämisen muutos oli kurssin aikana huomattavasti voimakkaampaa kuin muilla ryhmillä. Näiden ryhmien väliset erot vieraan aksentin muutoksessa olivat huomattavat, koska efektikoot olivat suuret. Lauluryhmä ja kuunteluryhmä eivät kuitenkaan eronneet toisistaan vieraan aksentin muutoksessa. Tulokset on esitetty taulukossa 4.

Seuraavaksi kuvataan opetusmenetelmäryhmien välisiä eroja vieraan aksentin arvioissa kursseilla I ja IIA tutkimuskysymykseen 1b vastaamiseksi: Hyötyvätkö alkeistason ja alkeisjatkotason suomenoppijat eri tavoin kolmen tutkimuksessa tarkastellun menetelmän käytöstä? I-kurssin opetusmenetelmä-



TAULUKKO 3. Kolmen opetusmenetelmäryhmän vieraan aksentin arviot (0–100: 0 = ei lainkaan vierasta aksenttia, 100 = vahva aksentti) ja erot alku- ja lopputestauksen välillä.

Opetusmenetelmä: testaus	n	KA	KH	t-testin p-arvo	Cohenin d
Laulu: alkutestaus	74	64,8	9,8		
Laulu: lopputestaus	74	64,1	9,3	0,31	0,07
Kuuntelu: alkutestaus	85	55,8	10,6		
Kuuntelu: lopputestaus	85	55,5	10,2	0,59	0,03
Rytmi: alkutestaus	88	52,9	10,3		
Rytmi: lopputestaus	88	46,8	11,2	< 0,01 **	0,57

\*\* Tilastollisesti erittäin merkitsevä ero.

TAULUKKO 4. Kolmen opetusmenetelmäryhmän väliset erot vieraan aksentin muutoksessa alkutestauksesta lopputestaukseen.

Opetusmenetelmä / vieraan aksentin muutos	t-testin p-arvo	Cohenin d
Lauluryhmä ja kuunteluryhmä	0,69	-0,06
Kuunteluryhmä ja rytmiryhmä	< 0,01 **	1,07
Rytmiryhmä ja lauluryhmä	< 0,01 **	0,93

\*\* Tilastollisesti erittäin merkitsevä ero.

TAULUKKO 5. Kurssien I ja IIA kolmen opetusmenetelmän vieraan aksentin arviot ja erot alku- ja lopputestauksen välillä.

	n	KA	KH	t-testin p-arvo	Cohenin d
I-kurssi: laulu					
alkutestaus	76	65,6	10,8		
lopputestaus	76	63,7	10,1	0,04*	0,18
I-kurssi: kuuntelu					
alkutestaus	88	62,8	11,1		
lopputestaus	88	59,8	10,6	< 0,01**	0,28
I-kurssi: rytmi					
alkutestaus	89	61,6	11,0		
lopputestaus	89	49,9	11,9	< 0,01**	1,02
IIA-kurssi: laulu					
alkutestaus	93	63,9	10,6		
lopputestaus	93	63,9	10,7	1,00	< 0,01
IIA-kurssi: kuuntelu					
alkutestaus	97	50,1	11,9		
lopputestaus	97	51,6	11,6	0,03*	-0,13
IIA-kurssi: rytmi					
alkutestaus	99	44,9	11,8		
lopputestaus	99	43,7	12,1	0,08	0,10

\* Tilastollisesti merkitsevä ero.

\*\* Tilastollisesti erittäin merkitsevä ero.

ryhmissä kaikkien ryhmien vieras aksentti oli kurssin aikana muuttunut natiivinomaisemmaksi. Lauluryhmällä vieras aksentti oli merkitsevästi vähentynyt kurssin aikana. Kuuntelu- ja rytmiryhmillä ääntäminen oli muuttunut erittäin merkitsevästi natiivinomaisemmaksi kurssin aikana. Ääntämisen kehittyminen natiivinomaisemmaksi oli erittäin suurta erityisesti rytmiryhmällä. Tähän viittaa suuri efektikoko. IIA-kurssilla kuunteluryhmän vieras aksentti oli arvioitu hieman voimakkaammaksi kurssin lopussa. Muutos oli merkitsevä. Rytmiryhmällä ääntäminen muuttui hieman natiivinomaisemmaksi, mihin viittasi myös tilastollinen tendenssi. Vieraan aksentin muutos oli kuitenkin pientä sekä kuunteluryhmässä että rytmiryhmässä, sillä efektikoot olivat pienet. Lauluryhmällä vieraan aksentin vahvuus ei muuttunut kurssin aikana. Tulokset on esitetty taulukossa 5.

Tämän jälkeen tutkittiin kurssien I ja IIA välisiä eroja opetusmenetelmitäin (ks. taulukko 6). Kurssit erosivat merkitsevästi toisistaan kaikkien opetus-

TAULUKKO 6. Kurssien välisten ja sisäisten kolmen opetusmenetelmän vieraan aksentin arviot.

<b>Kurssi: Opetusmenetelmä: vieraan aksentin muutos</b>	n	t-testin p-arvo	Cohenin d
Lauluryhmät (kurssi I ja IIA)	74	0,03 <sup>*</sup>	0,28
Kuunteluryhmät (kurssi I ja IIA)	85	< 0,01 <sup>**</sup>	-0,65
Rytmiryhmät (kurssi I ja IIA)	88	< 0,01 <sup>**</sup>	-1,25
Kurssi I			
Lauluryhmä ja kuunteluryhmä	72	0,19	0,19
Kuunteluryhmä ja rytmiryhmä	82	< 0,01 <sup>**</sup>	-1,00
Rytmiryhmä ja lauluryhmä	73	< 0,01 <sup>**</sup>	1,18
Kurssi IIA			
Lauluryhmä ja kuunteluryhmä	91	0,18	-0,21
Kuunteluryhmä ja rytmiryhmä	95	< 0,01 <sup>**</sup>	0,38
Rytmiryhmä ja lauluryhmä	91	0,30	0,16

\* Tilastollisesti merkitsevä ero.

\*\* Tilastollisesti erittäin merkitsevä ero.

menetelmäryhmien välillä. I-kurssilla lauluryhmän ääntäminen muuttui natiivinomaisemmaksi, IIA-kurssilla ei; I-kurssilla myös kuunteluryhmän ääntäminen muuttui natiivinomaisemmaksi, IIA-kurssilla vieras aksentti jopa voimistui; rytmiryhmissä ääntäminen muuttui natiivinomaisemmaksi, mutta muutos oli suurempi I-kurssilla. Efektikoot viittasivat siihen, että erityisesti rytmiryhmien ero kurssien välillä oli erittäin suuri.

Kun tarkasteltiin eroja kolmen opetusmenetelmäryhmän välillä kurssilla I, eroja vieraan aksentin muutoksessa oli kuunteluryhmän ja rytmiryhmän sekä rytmiryhmän ja lauluryhmän välillä. Kurssin I rytmiryhmässä vieras aksentti muuttui natiivinomaisemmaksi kaikista eniten ja kuunteluryhmällä toiseksi eniten kurssin aikana. Näillä ryhmillä suuret efektikoot tukevat tulkintaa ääntämisen kehittymisestä. Vaikka myös kurssin I lauluryhmällä ääntäminen muuttui natiivinomaisemmaksi kurssin aikana, lauluryhmä ei eronnut kuunteluryhmästä vieraan aksentin muutoksen suhteen. Kurssilla IIA vain kuunteluryhmä ja rytmiryhmä erosivat toisistaan vieraan aksentin muutoksen suhteen, vaikka muutos ei efektikoon perusteella ollutkaan suuri. Rytmiryhmän vieras aksentti muuttui kurssin aikana eniten natiivinomaiseen suuntaan. Tulokset on esitetty taulukossa 6.

TAULUKKO 7. S2- ja OKL-raadin arviot kolmen opetusmenetelmäryhmän vieraasta aksentista ja erot arvioiden välillä alku- ja lopputestauksessa. (Taulukossa alku = alku-testaus, loppu = lopputestaus, p = t-testin p-arvo; d = Cohenin d.)

	n	S2-raati			d	n	OKL-raati			d
		KA	KH	p			KA	KH	p	
Laulu										
alku	15	60,6	7,6			59	65,8	10,1		
loppu	15	59,3	7,0	0,34	0,18	59	65,3	9,5	0,49	0,05
Kuuntelu										
alku	27	52,0	11,1			58	57,5	9,9		
loppu	27	53,5	10,3	0,19	-0,14	58	56,4	10,0	0,08	0,11
Rytmi										
alku	26	48,4	9,8			62	54,7	10,0		
loppu	26	45,5	9,5	0,02*	0,30	62	47,4	11,8	< 0,01**	0,67

\* Tilastollisesti merkitsevä ero.

\*\* Tilastollisesti erittäin merkitsevä ero.

## 4.2 Kahden raadin arviot vieraasta aksentista

Seuraavaksi esitellään tulevista suomi toisena kielenä -opettajista koostuvan raadin (S2-raadin) ja tulevista luokanopettajista koostuvan raadin (OKL-raadin) mahdollisia eroja vieraan aksentin arvioinnissa. Tarkastelemalla raatien antamien vieraan aksentin arvioiden eroja pyritään vastaamaan toiseen tutkimuskysymykseen: Eroavatko S2-raadin ja OKL-raadin arviot suomenoppijoiden ääntämisestä? Raatien välisiä arvioiden eroja käsitellään ensin opetusmenetelmien välillä erottelematta kurssitasoja. Tämän jälkeen raatien arvioiden eroja kuvataan opetusmenetelmien välillä kurseilla I ja IIA.

S2- ja OKL-raatien arvioissa vain rytmiryhmien vieras aksentti muuttui kurssin aikana natiivinomaisemmaksi tilastollisesti merkitsevästi, kun tutkittiin opetusmenetelmäryhmiä erottelematta kurssiryhmiä. Vain OKL-raadin arvioiden mukaan muutos oli huomattava keskisuuren efektikoon perusteella. Molempien raatien arvioissa lauluryhmien vieras aksentti muuttui natiivinomaisempaan suuntaan, mutta muutos ei ollut merkitsevä. Myös efektikoon perusteella tämä muutos oli pieni. Raadit erosivat kuunteluryhmien vieraan aksentin arvioissa. Kuunteluryhmän vieras aksentti väheni OKL-raadin mukaan, mitä tuki tilastollinen tendenssi. Muutos oli kuitenkin pieni, koska efektikoko oli pieni. S2-raadilla kuunteluryhmän vieras aksentti vahvistui, mutta muutos ei ollut merkitsevä. Tulokset on esitetty taulukossa 7.

Raatien eroja analysoitiin lisäksi toistomittausten ANOValla. ANOVA-testillä tarkasteltiin tässä, eroavatko vieraan aksentin arviot kurssin alusta

TAULUKKO 8. Toistomittausten ANOVA: S2- ja OKL-raatien arviot lauluryhmien vieraan aksenttin muutoksesta. (Laulu: S2: n = 15; OKL: n = 59; Kuuntelu: S2: n = 27; OKL: n = 58; Loruilu: S2: n = 26; OKL: n = 62.)

	F	df	p	ositettu $\eta^2$
Lauluryhmät: alku/lopputestaus				
vieras aksentti	1,125	1,72	0,29	0,02
vieras aksentti × raati (S2/OKL)	0,173	1,72	0,68	< 0,01
raati	4,766	1,72	0,03*	0,06
Kuunteluryhmät: alku/lopputestaus				
vieras aksentti	0,80	1,83	0,78	< 0,01
vieras aksentti × raati (S2/OKL)	4,728	1,83	0,03*	0,05
ryhmä	3,270	1,83	0,07	0,04
Rytmiyhmät: alku/lopputestaus				
vieras aksentti	59,210	1,86	< 0,01**	0,41
vieras aksentti × raati (S2/OKL)	11,619	1,86	< 0,01**	0,12
raati	2,996	1,86	0,09	0,03

\* Tilastollisesti merkitsevä ero.

\*\* Tilastollisesti erittäin merkitsevä ero.

loppuun sekä antoivatko raadit erilaisia arvioita vieraasta aksentista. Lisäksi ANOVA-testillä tutkittiin vieraan aksenttin ja raadin välistä yhdysvaikutusta. Yhdysvaikutuksella vertailtiin, antoiko jompikumpi raati systemaattisesti esimerkiksi vahvempia vieraan aksenttin arvioita.

Lauluryhmän suhteen S2- ja OKL-raati erosivat arvioissaan tilastollisesti merkitsevästi, mitä tukee myös keskisuuri efektikoko. OKL-raati arvioi vieraan aksenttin voimakkaammaksi kuin S2-raati. Kuunteluryhmän suhteen vieraan aksenttin muutoksen ja raadin välinen yhdysvaikutus oli merkitsevä. OKL-raati arvioi myös kuunteluryhmien ääntämistä selvästi heikommaksi kuin S2-raati. Tätä tuki myös tilastollinen tendenssi raatien välisestä erosta. Rytmiryhmän suhteen sekä vieraan aksenttin muutos että vieraan aksenttin muutoksen ja raadin yhdysvaikutus oli erittäin merkitsevä. Erityisesti vieraan aksenttin muutos oli huomattava keskisuuren efektikoon perusteella. Vieraan aksenttin muutoksen ja raadin yhdysvaikutus vaikutti myös huomattavalta keskisuuren efektikoon perusteella. OKL-raati arvioi myös rytmiryhmän vieraan aksenttin voimakkaammaksi kuin S2-raati. Tulokset tarkoittavat sitä, että molempien raatien arvioissa ääntäminen muuttui natiivinomaisempaan suuntaan ja että OKL-raati arvioi parannuksen suuremmaksi. Tilastollinen tendenssi viittaa myös raatien väliseen eroon. Molempien raatien arvioissa rytmiryhmien vieras aksentti muuttui kurssin aikana eniten natiivinomaisempaan suuntaan. Tu-

TAULUKKO 9. S2- ja OKL-raadin arviot kurssien I ja IIA kolmen opetusmenetelmäryhmän vieraan aksentin muutoksesta ja muutosten erot alku- ja lopputestauksessa.

	S2-raati				OKL-raati			
	n	KA	KH	p	n	KA	KH	p
I-kurssi: laulu								
alkutestaus	16	59,7	7,8		60	67,1	11,0	
lopputestaus	35	60,4	10,2	0,40	65	65,4	9,9	0,06
I-kurssi: kuuntelu								
alkutestaus	29	57,9	11,5		60	65,4	10,2	
lopputestaus	39	57,7	11,2	0,75	64	61,3	9,8	< 0,01 **
I-kurssi: rytmi								
alkutestaus	32	54,1	11,0		64	64,6	10,2	
lopputestaus	29	49,1	9,5	> 0,01 **	66	50,2	12,5	< 0,01 **
IIA-kurssi: laulu								
alkutestaus	32	62,4	11,1		64	64,6	10,2	
lopputestaus	35	61,3	11,9	0,21	63	65,2	9,7	0,43
IIA-kurssi: kuuntelu								
alkutestaus	38	45,7	11,4		64	52,3	11,5	
lopputestaus	34	49,0	11,4	0,02 *	64	53,0	11,4	0,39
IIA-kurssi: rytmi								
alkutestaus	38	41,5	11,7		64	46,9	11,2	
lopputestaus	36	41,2	11,2	0,82	65	45,5	12,6	0,50

\* Tilastollisesti merkitsevä ero.

\*\* Tilastollisesti erittäin merkitsevä ero.

lokset on esitetty taulukossa 8.

Seuraavaksi esitellään kurssien I ja IIA vieraan aksentin arvioita kolmen opetusmenetelmäryhmän osalta. Sekä S2- että OKL-raati arvioi kurssin I rytmiryhmän vieraan aksentin muuttuneen tilastollisesti merkitsevästi natiivinomaisemmaksi. Vain OKL-raati arvioi myös kurssin I kuunteluryhmän aksentin lähentyneen natiivinomaista ääntämystä kurssin aikana. S2-raati ei arvioinut I-kurssin lauluryhmän ääntämisen muuttuneen merkitsevästi. OKL-raadilla vieraan aksentin muutosta tuki tilastollinen tendenssi. IIA-kurssin osalta ainoastaan S2-raati arvioi kuunteluryhmän vieraan aksentin jopa vahvistuneen kurssin aikana merkitsevästi. Laulu- ja kuunteluryhmän vieras aksentti ei muuttunut merkitsevästi kurssin aikana kummankaan raadin arvioissa. Tulokset on esitetty taulukossa 9.

Seuraavaksi tarkastellaan kahden raadin eroja opetusmenetelmäryhmien välillä kurssilla I ja IIA (ks. taulukko 10). Kurssilla I lauluryhmän arvioissa raadit

TAULUKKO 10. Toistomittausten ANOVA: S2- ja OKL-raatien arviot kurssin I ja IIA lauluryhmien alku- ja lopputestauksen vieraan aksenttin arvioista. (Kurssi I: Laulu: S2: n = 16; OKL: n = 60; Kurssi 1: Kuuntelu: S2: n = 29; OKL: n = 59; Loruilu: S2: n = 27; OKL: n = 62; Kurssi IIA: Laulu: S2: n = 30; OKL: n = 63; Kurssi 1: Kuuntelu: S2: n = 34; OKL: n = 63; Loruilu: S2: n = 35; OKL: n = 64.)

	F	df	p	ositettu $\eta^2$
<b>KURSSI I</b>				
Lauluryhmä: alku/lopputestaus				
vieras aksentti	2,592	1,74	0,11	0,03
vieras aksentti × raati (S2/OKL)	0,028	1,74	0,87	< 0,01
raati	7,605	1,74	< 0,01**	0,09
Kuunteluryhmä: alku/lopputestaus				
vieras aksentti	6,260	1,86	0,01*	0,07
vieras aksentti × raati (S2/OKL)	3,662	1,96	0,06	0,04
raati	6,266	1,86	0,01*	0,07
Rytmiyhmä: alku/lopputestaus				
vieras aksentti	95,355	1,87	< 0,01**	0,52
vieras aksentti × raati (S2/OKL)	16,768	1,87	< 0,01**	0,16
raati	4,801	1,87	0,03*	0,05
<b>KURSSI IIA</b>				
Lauluryhmä: alku/lopputestaus				
vieras aksentti	0,267	1,91	0,61	< 0,01
vieras aksentti × raati (S2/OKL)	2,073	1,91	0,15	0,02
raati	1,717	1,91	0,19	0,02
Kuunteluryhmä: alku/lopputestaus				
vieras aksentti	6,318	1,95	0,01*	0,06
vieras aksentti × raati (S2/OKL)	2,181	1,95	0,14	0,02
raati	4,747	1,95	0,03*	0,05
Rytmiyhmä: alku/lopputestaus				
vieras aksentti	1,926	1,97	0,17	0,02
vieras aksentti × raati (S2/OKL)	1,044	1,97	0,31	0,01
raati	4,138	1,97	0,05*	0,04

\* Tilastollisesti merkitsevä ero.

\*\* Tilastollisesti erittäin merkitsevä ero.

erosivat toisistaan merkitsevästi: OKL-raati arvioi kurssilaisten vieraan aksenttin vahvemmaksi kuin S2-raati. Tätä tuki myös keskisuuri efektikoko. Vieraan aksenttin päävaikutus tai vieraan aksenttin ja raadin yhdysvaikutus ei ollut merkitsevä. Vieraan aksenttin arviot eivät siis muuttuneet alkutestin ja lopputestin välillä. Kuunteluryhmällä sekä vieraan aksenttin muutos että raadit erosivat merkitsevästi toisistaan, mitä myös keskisuuri efektikoko tuki. OKL-raati arvioi vieraan aksenttin siis vahvemmaksi kuin S2-raati. Tilastollinen tendenssi viittasi raadin ja aksenttin muutoksen välisen yhdysvaikutuksen lähenevän merkitsevää, mutta vaikutus ei ollut suuri, koska efektikoko oli pieni. Rytmiryhmässä vieraan aksenttin arvion muutos, raatien antamien arvioiden ero sekä näiden yhdysvaikutus olivat merkitsevät. Vieraan aksenttin arvion muutos sekä muutoksen ja raatien yhdysvaikutus olivat huomattavat suuren efektikoon perusteella. OKL-raati arvioi rytmiryhmän vieraan aksenttin siis huomattavasti vahvemmaksi kuin S2-raati.

Kurssin IIA lauluryhmän ANOVA-testin tulokset eivät olleet merkitseviä. Kuunteluryhmällä sekä vieraan aksenttin muutos että raatien erot olivat merkitsevät, mutta vain vieraan aksenttin muutos vaikutti huomattavalta keskisuuren efektikoon perusteella. Rytmiryhmällä raadit erosivat merkitsevästi toisistaan. Tulokset on esitetty taulukossa 10.

## 5 Pohdinta

Tämän tutkimuksen tavoitteena oli selvittää, miten suomenoppijoiden ääntäminen kehittyi, kun he laulavat, kuuntelevat lauluja tai lausuvat laulun sanoja rytmikkäästi. Erityisesti haluttiin tarkastella, ovatko opetusmenetelmät hyödyllisempiä alkeistason vai alkeisjatkotason suomenoppijoille ääntämisen kehittämisessä. Lisäksi tutkimuksessa haluttiin selvittää, arvioivatko tulevat suomen kielen opettajat ja tulevat luokanopettajat kielenoppijoiden vierasta aksenttia eri tavoin.

Tutkimuksen tulokset osoittivat, että kolmesta opetusryhmästä rytmiryhmän suomen kielen ääntäminen oli lähinnä natiivinomaista alku- ja loppumittauksessa sekä silloin kun molempien kurssien opiskelijoiden aksentteja tarkasteltiin yhdessä että kun niitä tarkasteltiin kurssitasoittain (I ja IIA). Rytmiryhmien ääntäminen myös kehittyi eniten kohti natiivinomaista kurssin aikana, silloin kun kaikkia opiskelijoita tarkasteltiin yhdessä ja myös silloin kun kurseja I ja IIA tarkasteltiin erikseen. Eri opetusryhmät erosivat toisistaan suomen ääntämisen kehittymisen suhteen. Ainoastaan laulu- ja kuunteluryhmä eivät eronneet toisistaan merkitsevästi. Tämä voi johtua siitä, että sekä laulu- että kuunteluryhmässä suomen kielen ääntämisen kehittyminen oli erittäin vähäistä, mutta rytmiryhmässä kehitys oli huomattavasti suurempaa.

Nämä tulokset ovat hypoteesimme vastaisia. Aiemman tutkimuksemme



(Alisaari & Heikkola 2016) perusteella oletimme lauluryhmien ääntämisen kehittyvän lähimmäksi natiivinomaista. Tulosten tulkinnassa täytyy kuitenkin ottaa huomioon se, että rytmiryhmät olivat jo kurssin alkutestissä lähempänä natiivinomaista ääntämistä kuin laulu- ja kuunteluryhmät. Vaikuttaa siis siltä, että jos suomen kielen ääntäminen on jo lähempänä natiivinomaista, sitä on helpompi kehittää vahvemmin tähän suuntaan. Jatkotutkimuksessa ääntämisen lähtötasot olisikin hyvä ottaa huomioon.

Kun opetusryhmien suomen kielen ääntämisen kehittymistä tarkasteltiin kurssitasoittain (kurssit I ja IIA), kävi ilmi, että kurssien välillä oli huomattavia eroja. Kurssin I eli alkeiskurssin kaikissa opetusryhmissä suomen kielen ääntäminen muuttui natiivinomaisemmaksi. Kuunteluryhmällä ja rytmiryhmällä ääntäminen kehittyi natiivinomaisemmaksi jopa huomattavasti. Kurssin IIA eli alkeisjatkokurssin opetusryhmillä ääntämisen kehitys ei ollut yhtä suoraviivaista. Kurssin IIA opetusryhmissä rytmikäs lausuminen kehitti ääntämistä eniten, kun taas lauluryhmällä ääntämys ei muuttunut kurssin aikana. Yllättävästi kuunteluryhmän vieras aksentti arvioitiin jopa vahvemmaksi kurssin lopussa. Näiden tulosten pohjalta vaikuttaa siltä, että erityisesti kielen oppimisen alkuvaiheessa olevilla oppijoilla tapahtui merkitsevää kehitystä ääntämisessä intensiivikurssien aikana. Jatkotutkimuksessa olisikin tärkeää keskittyä tarkastellun kolmen opetusmenetelmän, laulamisen, laulujen kuuntelun ja laulujen sanojen rytmisen laulamisen, vaikutuksiin eri kielitaitotasoilla. Nämä tulokset eroavat aiemmista tutkimustuloksista, joissa laulamisen on todettu olevan tehokkain kielenoppimisen opetusmenetelmä tarkastelluista kolmesta menetelmästä, erityisesti kun tutkittiin kirjoittamisen sujuvuuden kehittymistä (Alisaari & Heikkola 2016). Laulamisen, laulujen kuuntelun ja laulun sanojen rytmikään lausumisen yhteydestä erityisesti ääntämiseen ei ole kuitenkaan aiemmin tehty tutkimusta. Tämän vuoksi ei ole mahdollista verrata tämän tutkimuksen tuloksia aiempaan tutkimukseen.

Tutkimuksessa haluttiin myös selvittää, arvioivatko tulevat suomi toisena kielenä -opettajat (S2-raati) ja tulevat luokanopettajat (OKL-raati) kielenoppijoiden vierasta aksenttia eri tavoin. Tulokset osoittavat, että raatien arviot vieraasta aksentista olivat erilaiset. Kuten oletimme aiemman tutkimuksen perusteella, OKL-raati arvioi vieraan aksentin kaikissa opetusmenetelmäryhmissä vahvemmaksi kuin S2-raati. Oletamme tämän johtuvan siitä, että tulevat suomen kielen opettajat ovat opintojensa vuoksi tottuneempia kuulemaan erilaisia aksentteja suomen kielessä, ja siksi he arvioivat ääntämistä lempeämmin (vrt. Ahola 2016). Aiemman tutkimuksen mukaan erilaiset arvioijat eivät yleensä eroa toisistaan, mutta kokeneet arvioijat antaisivat lievempiä aksenttiarvioita (Thompson 1991: 198). Toisaalta aiemmassa tutkimuksessa on myös esitetty, että S2-opettajat eivät arvioi ääntämistä eri tavalla kuin muut (Toivola 2011).

Tutkimuksemme tulosten perusteella voidaan olettaa, että S2-opettajik-

si opiskelevilla on enemmän kokemuksia vieraan aksentin kuuntelusta. Toisaalta luokanopettaja-arvioitsijat olivat pääsääntöisesti kolmannen vuosikurssin opiskelijoita, jotka olivat suorittaneet noin kolmen kuukauden verran opetusharjoitteluja monikielisessä koulussa, mikä on voinut vaikuttaa heidän arviointeihinsa. Tulevilla luokanopettajilla on voinut olla kokemusta monikielisistä opiskelijoista, eli he ovat voineet olla tottuneempia kuulemaan vieraita aksentteja kuin tutkimuksessa oletettiin. Toisaalta OKL-raadin ja S2-raadin välillä ei kuitenkaan ollut eroja siinä, kuinka paljon he viettivät aikaa maahanmuuttajataustaisten henkilöiden kanssa. Toisaalta tutkimuksen tulosten tulokinnassa täytyy ottaa myös huomioon, että S2-opiskelijat ovat myös voineet tehdä pedagogisiin opintoihin kuuluvan harjoittelunsa monikielisessä koulussa. Nämä tekijät tulisi kontrolloida jatkotutkimuksessa.

Kun tarkastellaan yksittäisiä opetusmenetelmäryhmiä, OKL-raati arvioi kaikkien kolmen opetusmenetelmän vaikuttaneen myönteisesti ääntämiseen. Kuitenkin vain laulun sanojen rytmikäs lausuminen paransi ääntämistä merkittävästi. OKL-raadin arvioissa laulaminen vaikuttaa parantavan ääntämistä jonkin verran. Sen sijaan kuunteluryhmän ääntäminen ei parantunut kurssin aikana. Myös S2-raati arvioi rytmiryhmän ääntämisen parantuneen kurssin aikana eniten. Lisäksi S2-raadin arvioissa lauluryhmä paransi ääntämystään, vaikka ei merkittävästi. Toisin kuin oletimme, kuunteluryhmän vieraan aksentin arvioitiin jopa vahventuneen, vaikka tämäkään muutos ei ollut merkittävä. Kuten aiemmassa tutkimuksessa on esitetty (ks. esim. Ahola 2016), kokemattomampi OKL-raati arvioi suomenoppijoiden ääntämistä tiukemmin kuin S2-raati.

Tulosten tarkastelussa täytyy ottaa huomioon tutkimusta rajoittavia tekijöitä. Raatilaiset merkitsivät vieraan aksentin arvionsa puhenäytteistä VAS-asteikoille. Heitä kehoitettiin merkitsemään arvio, vaikka he olisivat olleet epävarmoja. Jotkut raatilaisista jättivät kuitenkin tyhjiä vastauksia, mikä vaikuttaa vieraan aksentin tarkasteluun yksittäisen kielenoppijan osalta. Tyhjiin vastausten jättämiseen on voinut vaikuttaa se, että puhenäytteet kuuntelutettiin isossa luentosalissa, minkä vuoksi joidenkin puhenäytteiden kuuluvuus ei välttämättä ollut paras mahdollinen. Tarkempia vieraan aksentin arvioita voitaisiin saada siten, että raatilaiset kuuntelisivat puhenäytteitä kuulokkeilla. Lisäksi jotkut arvioijat ovat käyttäneet VAS-asteikkoa kapeasti, toiset arvioijat leveämmin. Jatkotutkimuksissa tätä voisi kompensoida tilastollisin menetelmin (ks. esim. Toivola 2011).

Tutkimuksemme tuloksia tarkasteltaessa tulee ottaa huomioon myös se, että kielenoppijoiden kielitaitotasot vaihtelevat. Vaikka suurin osa kielenoppijoista on kielitaitotasolla A2, voivat erot kielitaitotasossa vaikuttaa siihen, kuinka he voivat käyttää hyödykseen opetusmenetelmän tarjoumia kielenoppimiseen. Jatkotutkimuksissa voitaisiin keskittyä esimerkiksi vain A2-tasoisten kielenoppijoiden ääntämykseen, jolloin kielitaitotason vaihtelu ei vaikuttaisi tuloksiin.

Tämän tutkimuksen perusteella vaikuttaa siltä, että kaikki kolme opetusmenetelmää, laulaminen, laulujen kuuntelu ja laulujen sanojen rytmikäs lausuminen, vaikuttavat myönteisesti erityisesti alkeistason suomenoppijoiden ääntämiseen. Laulun sanojen rytmikäs lausuminen vaikuttaa kuitenkin tuottavan parhaita tuloksia ääntämisen harjoittelussa. Myös alkeisjatkokursin opiskelijoiden ääntäminen kehittyi eniten rytmikkään lausumisen kautta. Tämä tutkimus on uusi avaus opetusmenetelmien vaikutuksesta ääntämiseen. Tulokset viittaavat siihen, että ääntämisen opetukseen voidaan suositella erityisesti rytmistä lausumista. Toisaalta jatkotutkimus, jossa on tarkemmin kontrolloitu mahdollisia tuloksiin liittyviä tekijöitä, on tarpeen vahvojen suositusten taustalle.

## Lähteet

- Abu-rabia, S. & S. Kehat 2004. The critical period for second language pronunciation: is there such a thing? Ten case studies of late starters who attained a native-like Hebrew accent. *Educational Psychology*, 24 (1), 77–97. DOI: 10.1080/0144341032000146467.
- Aho, E., M. Toivola, F. Karlsson & M. Lennes 2016. Aikuisten maahanmuuttajien suomen ääntämisestä. *Puhe ja kieli*, 36 (2), 77–96. <http://hdl.handle.net/10138/168735>.
- Ahola, S. 2016. Puhetta arvioinnista: yleisten kielitutkintojen arvioijien käsityksiä arvioinnista. Teoksessa A. Huhta & R. Hildén (toim.) *Kielitaidon arviointitutkimus 2000-luvun Suomessa*. AFinLA-e: soveltavan kielitieteen tutkimuksia 9. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 89–109. <https://journal.fi/afinla/article/view/60848>.
- Alisaari, J. 2016. *Songs and poems in the second language classroom: the hidden potential of singing for developing writing fluency*. Turku: University of Turku. <http://urn.fi/URN:ISBN:978-951-29-6673-8>.
- Alisaari, J. & L. M. Heikkola tekeillä. *Laulut S2-alkeisopetuksessa*. Käsikirjoitus.
- 2016. Increasing fluency in L2 writing with singing. *Studies in Second Language Learning and Teaching*, 6 (2), 271–292. DOI: 10.14746/ssl.t.2016.6.2.5.
- 2017. Songs and poems in the language classroom: teachers' beliefs and practices. *Teaching and Teacher Education*, 63, 231–242. DOI: 10.1016/j.tate.2016.12.021.
- Anderson-Hsieh, J., R. Johnson & K. Koehler 1992. The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, 42 (4), 529–555.
- Bongaerts, T. 1999. Ultimate attainment in L2 pronunciation: the case of very advanced late L2 learners. Teoksessa D. Birdsong (toim.) *Second language acquisition and the critical period hypothesis*. Mahwah (N. J.): Lawrence Erlbaum, 133–159.
- Bongaerts, T., S. Mennen & F. van der Slik 2000. Authenticity of pronunciation in naturalistic second language acquisition: the case of very advanced late learners of

- Dutch as a second language. *Studia Linguistica*, 54 (2), 298–308. DOI: 10.1111/1467-9582.00069.
- Breitkreutz, J. A., T. M. Derwing & M. J. Rossiter 2001. Pronunciation teaching practices in Canada. *TESL Canada Journal*, 19 (1), 51–61.
- Burgess, J. & S. Spencer 2000. Phonology and pronunciation in integrated language teaching and teacher education. *System*, 28 (2), 191–215. DOI: 10.1016/S0346-251X(00)00007-5.
- Derwing, T. M. 2003. What do ESL students say about their accents? *Canadian Modern Language Review*, 59 (4), 545–564.
- Derwing, T. M. & M. J. Munro 2005. Second language accent and pronunciation teaching: a research-based approach. *TESOL Quarterly*, 39 (3), 379–397. DOI: 10.2307/3588486.
- Derwing, T. M., M. J. Rossiter, M. J. Munro & R. I. Thomson 2004. Second language fluency: judgments on different tasks. *Language Learning*, 54 (4), 655–679.
- Euroopan neuvosto 2012. *Eurooppalainen viitekehys: kielten oppimisen, opettamisen ja arvioinnin yhteinen eurooppalainen viitekehys* Käänt. I. Huttunen & H. Jaakkola. Helsinki: Sanoma Pro.
- Flege, J. E. 1984. The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, 76 (3), 692–707. DOI: 10.1121/1.391256.
- 1991. Age of learning affects the authenticity of voice-onset-time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, 89 (1), 395–411.
- Flege, J. E. & K. L. Fletcher 1992. Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America*, 91 (1), 370–389.
- Flege, J. E., M. J. Munro & I. R. A. MacKay 1995. Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97 (5), 3125–3134. DOI: 10.1121/1.413041.
- Grape, C., M. Sandgren, L. O. Hansson, M. Ericson & T. Theorell 2003. Does singing promote well-being? An empirical study of professional and amateur singers during a singing lesson. *Integrative Physiological & Behavioral Science*, 38 (1), 65–74. DOI: 10.1007/BF02734261.
- Hismanoglu, M. & S. Hismanoglu 2010. Language teachers' preferences of pronunciation teaching techniques: traditional or modern? *Procedia Social and Behavioral Sciences*, 2 (2), 983–989. DOI: 10.1016/j.sbspro.2010.03.138.
- Huron, D. 2006. Is music an evolutionary adaptation? *Annals of the New York Academy of Science*, 930, 43–61. DOI: 10.1111/j.1749-6632.2001.tb05724.x.
- Jesney, K. 2004. *The use of global foreign accent rating in studies of L2 acquisition*. Calgary: Language Research Centre, University of Calgary. <http://arts.ucalgary.ca/lrc/research/reports> [luettu 30. 5. 2017].
- Kang, O., D. L. Rubin & I. Pickering 2010. Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, 94 (4), 554–566.
- Kendrick, H. 1997. Keep them talking! A project for improving students' L2 pronunciation. *System*, 25 (4), 545–560. DOI: 10.1016/S0346-251X(97)00043-2.
- Kennedy, S. & P. Trofimovich 2008. Intelligibility, comprehensibility, and accentedness of L2 speech: the role of listener experience and semantic context. *The Canadian Modern Language Review*, 64 (3), 459–489. DOI: 10.3138/cmlr.64.3.459.

- Krashen, S. 1982. *Principles and practice in second language acquisition*. Oxford: Pergamon Press. [http://www.sdkrashen.com/content/books/principles\\_and\\_practice.pdf](http://www.sdkrashen.com/content/books/principles_and_practice.pdf).
- Kuronen, M. 2016. Uttal av S2-finska med fokus på svenskspråkiga talare. *Puhe ja kieli*, 36 (3), 147–174.
- Leinonen, A. 2015. "Riittää kun saa selvää": vieraalla aksentilla tuotettu suomi nuorten arvioimana. *Jyväskylä Studies in Humanities* 275. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-6448-1>.
- Lintunen, P. 2014. Ääntämisen oppiminen ja opettaminen. Teoksessa P. Pietilä & P. Lintunen (toim.) *Kuinka kieltä opitaan: opas vieraan kielen opettajalle ja opiskelijalle*. Helsinki: Gaudeamus, 165–187.
- Lord, G. 2005. (How) can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania*, 88 (3), 557–567.
- Major, R. C. 2007. Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition*, 29 (4), 539–556.
- Markham, D. 1997. *Phonetic imitation, accent, and the learner*. Lund: Lund University.
- Medina, S. L. 2002. Using music to enhance second language acquisition: from theory to practice. Teoksessa J. Lalas & S. Lee (toim.) *Language, literacy, and academic development for English language learners*. London: Pearson Educational Publishing. <http://www.forefrontpublishers.com/eslmusic/articles/238-using-music-to-enhance-second-languageacquisition-from-theory-to-practice/>.
- Milovanov, R., M. Huotilainen, V. Välimäki, P. A. A. Esquef & M. Tervaniemi 2008. Musical aptitude and second language pronunciation skills in school-aged children: neural and behavioral evidence. *Brain Research*, 1194, 81–89. DOI: 10.1016/j.brainres.2007.11.042.
- Milovanov, R., P. Pietilä, M. Tervaniemi & P. A. A. Esquef 2010. Foreign language pronunciation skills and musical aptitude: a study of Finnish adults with higher education. *Learning and Individual Differences*, 20 (1), 56–60. DOI: 10.1016/j.lindif.2009.11.003.
- Morley, J. 1991. The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25 (3), 481–520. DOI: 10.2307/3586981.
- Muñoz, C. & D. Singleton 2007. Foreign accent in advanced learners: two successful profiles. Teoksessa L. Roberts, G. Ayse, T. Sibel & L. Marti (toim.) *EUROSLA Yearbook 7: annual conference of the European second language association*. Amsterdam: John Benjamins, 171–190.
- Munro, M. J. 2003. A primer on accent discrimination in the Canadian context. *TESL Canada Journal*, 20 (2), 38–51.
- 2008. Foreign accent and speech intelligibility. Teoksessa J. G. Hansen Edwards & M. L. Zampini (toim.) *Phonology and second language acquisition*. Amsterdam: John Benjamins, 193–218.
- Munro, M. J. & T. M. Derwing 2001. Modeling perceptions of the accentedness and comprehensibility of L2 speech: the role of speaking rate. *Studies in Second Language Acquisition*, 23 (4), 451–468.
- Munro, M. J., T. M. Derwing & S. Morton 2006. Mutual intelligibility of L2-speech: language learning. *Studies in Second Language Acquisition*, 28 (1), 111–131. DOI: 10.1017/S0272263106060049.

- Munro, M. J. & V. Mann 2005. Age of immersion as a predictor of foreign accent. *Applied Psycholinguistics*, 26 (3), 311–341. DOI: 10.1017/S0142716405050198.
- Nortier, J. & M. Dorleijnb 2008. A Moroccan accent in Dutch: a sociocultural style restricted to the Moroccan community? *International Journal of Bilingualism*, 12 (1–2), 125–142.
- Phillips, E. M. 1992. The effects of language anxiety on student's oral test performance and attitudes. *Modern Language Journal*, 76 (1), 14–26. DOI: 10.1111/j.1540-4781.1992.tb02573.x.
- Piske, T., I. R. A. MacKay & J. E. Flege 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29 (2), 191–215. DOI: 10.1006/jpho.2001.0134.
- Riney, T. J., N. Takagi & K. Inutsuka 2005. Phonetic parameters and perceptual judgments of accent in English by American and Japanese listener. *TESOL Quarterly*, 39 (3), 441–466. DOI: 10.2307/3588489.
- Schmidt, R. 1990. The role of consciousness in second language learning. *Applied Linguistics*, 11 (2), 129–158. DOI: 10.1093/applin/11.2.129.
- Schön, D., M. Boyer, S. Moreno, M. Besson, I. Peretz & R. Kolinsky 2008. Songs as an aid for language acquisition. *Cognition*, 106 (2), 975–983. DOI: 10.1016/j.cognition.2007.03.005.
- Scovel, T. 1969. Foreign accents, language acquisition, and cerebral dominance. *Language Learning*, 19 (3–4), 245–253. DOI: 10.1111/j.1467-1770.1969.tb00466.x.
- Suomen virallinen tilasto 2017. Muuttoliike. Helsinki: Tilastokeskus. <http://www.stat.fi/til/muutl/index.html> [luettu 30. 9. 2017].
- Swain, M. 2000. The output hypothesis and beyond: mediating acquisition through collaborative dialogue. Teoksessa J. P. Lantolf (toim.) *Sociocultural theory and second language learning*. Oxford: Oxford University Press, 97–114.
- Tergujeff, E. 2013. *English pronunciation teaching in Finland*. Jyväskylä Studies in Humanities 207. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-5322-5>.
- Thompson, I. 1991. Foreign accents revisited: the English pronunciation of Russian immigrants. *Language Learning*, 41 (2), 177–204. DOI: 10.1111/j.1467-1770.1991.tb00683.x.
- Toivola, M. 2011. *Vieraan aksentin arviointi ja mittaaminen suomessa*. Helsinki: Helsingin yliopisto. <http://urn.fi/URN:ISBN:978-952-10-7217-8>.
- Ullakonoja, R. & M. Kuronen 2015. Young Russian immigrants' segmental duration and length in Finnish. Teoksessa The Scottish Consortium for ICPhS 2015 (toim.) *Proceedings of ICPhS 2015: 18th international congress of phonetic sciences, 10–14 Aug 2015 Secc Glasgow Scotland UK*. Glasgow: University of Glasgow. <https://www.internationalphoneticassociation.org/icphs-proceedings/ICP>.
- van Lier, L. 1998. The relationship between consciousness, interaction and language learning. *Language Awareness*, 7 (2–3), 128–145. DOI: 10.1080/09658419808667105.
- VanPatten, B. & W. R. Glass 1998. Grammar learning as a source of language anxiety: a discussion. Teoksessa D. J. Young (toim.) *Affect in foreign language and second language learning: a practical guide to creating a low-anxiety classroom atmosphere*. Boston (Mass.) McGraw-Hill College, 89–105.

Weigle, S. 1999. Investigating rater/prompt interactions in writing assessment: quantitative and qualitative approaches. *Assessing Writing*, 7 (2), 145-178.

Wewers, M. E. & N. K. Lowe 1990. A critical review of visual analogue scales in the measurement of clinical phenomena. *Research in Nursing & Health*, 13 (4), 227-236.

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 45–60.

**Martina Huhtamäki<sup>1</sup> & Elisabeth Zetterholm<sup>2</sup>**

<sup>1</sup>Helsingfors universitet, <sup>2</sup>Stockholms universitet

## **Uttalets plats i undervisningen av svenska som andraspråk**

This is a case study about pronunciation teaching in Swedish as a second language. The aim of this study is to compare pronunciation teaching for adults in Finland and Sweden. Data were gathered via interviews with eight teachers and observations of eight Swedish classes. The theoretical framework is language teacher cognition, which implies that we compare what teachers say with what they actually do in the classrooms. The study shows that teachers in both countries find pronunciation a central field of Swedish teaching. However, they think that they lack methods for pronunciation teaching. Most of the pronunciation teaching observed is implicit and integrated in the teaching of other elements, such as vocabulary and grammar. In Sweden, there were more examples of explicit pronunciation teaching than in Finland. In Finland, the minority position of Swedish implies certain challenges for teaching pronunciation and other aspects of the language. Overall, there was much correspondence between the interviews and the observed classes in the two countries.

**Keywords:** pronunciation teaching, Swedish as a second language, teacher cognition

**Nyckelord:** uttalsundervisning, andraspråksinläring, lärarens kognition



## 1 Inledning

Uttal är en grundläggande aspekt av att behärska ett språk, men för vuxna inlärare av ett nytt språk kan det vara en lång process att lära sig artikulera nya språkljud och använda en prosodi som inte alltid överensstämmer med modersmålet. Enligt teorin om fonologisk transfer (Jarvis & Pavlenko 2007) finns det en koppling mellan inlärarens modersmål och målspråket. Det innebär att man generellt kan säga att inlärare med olika modersmål kan ha helt olika svårigheter vad gäller just uttalet i andraspråksinläringen (Zetterholm & Tronnier 2017). Undervisningen och lärarens kunskap om den fonologiska relationen mellan inlärarens förstaspråk och målspråket kan då ha betydelse och skapa goda förutsättningar för inlärare att erövra ett lyssnarvänligt svenskt uttal.

Syftet med den här studien är att få mera kunskap om hur lärare resonerar om uttalsundervisning för vuxna inlärare av svenska som andraspråk samt om hur lärare undervisar i uttal. Studien har således ett didaktiskt perspektiv med fokus på läraren. Fyra lärarintervjuer och fyra observationer av lektioner har gjorts i Finland och i Sverige för att kunna jämföra uttalsundervisningen i de båda länderna. Studien och dess resultat ska ligga till grund för fortsatta studier om uttalsundervisning i en nordisk kontext.

Både äldre och nyare studier har visat att uttalet hos en andraspråksinlärare ofta påverkar målspråkstales attityder till inläraren (Boyd & Bredänge 2013; LeVelle & Levis 2014; Derwing & Munro 2015). Boyd & Bredänge (2013) ger exempel på hur brister i uttalet kan tolkas som att inläraren har bristande yrkeskompetens. Vidare kan en alltför kraftig brytning leda till missförstånd i kommunikationen och ha betydelse för inlärarens sociala relationer. Detta kan ha inverkan på självkänslan och identiteten hos inläraren och påverka möjligheterna att få ett arbete (Moyer 2013, 2014; LeVelle & Levis 2014; Derwing & Munro 2015). Det är dock ganska ovanligt att framför allt vuxna inlärare talar ett nytt språk helt utan brytning (Flege m. fl. 1995; Moyer 2013). Ändå finns det inte något tydligt och självklart samband mellan graden av en tales brytning och hur lätt eller svårt det är att förstå talaren (Derwing & Munro 1997).

Thomson & Derwing (2014) har gått igenom 75 olika forskningsstudier med fokus på uttalsundervisning. Samtliga studier var publicerade på engelska och de flesta studierna (74 %) berörde engelska som andraspråk eller främmandespråk. 13 procent av studierna handlade om spanska inlärare, 7 procent franska inlärare och i övriga publikationer hade inlärarna arabiska, holländska, japanska eller mandarin som förstaspråk. Thomson & Derwing kan efter noggrann genomgång av studierna konstatera att praktiska och tydliga instruktioner rörande uttalet i de allra flesta fall har en positiv effekt på inlärarens uttal. För att erövra ett funktionellt uttal behöver inlärare såväl instruktioner som tillfällen att öva sitt uttal i olika kontexter, och detta är något som bör integreras i undervisningen redan i ett inledande skede av språkinläringen (Zielinski

& Yates 2014; Yates 2017). Det finns annars en risk att inläraren erövrar ett uttal som är felaktigt och kan leda till missförstånd, vilket kan vara svårt att förändra senare. I kursmålen för vuxna inlärare av svenska som andraspråk i den svenska kontexten finns en tydlig progression vad gäller den muntliga produktionen, och målet är att inläraren, med viss anpassning, ska kunna kommunicera i olika situationer (Skolverket 2012). Därför bör målet med uttalsundervisningen vara att inläraren får ett begripligt och funktionellt uttal, inte ett brytningsfritt uttal, vilket även poängteras av bl. a. Levis (2005).

## 2 Bakgrund

Studier i en kanadensisk kontext har visat att uttalet inte alltid har en central plats i undervisningen i engelska som andraspråk (Foote m. fl. 2011). Verksamma lärare (cirka 200 stycken) svarade på frågor om sin attityd till uttalsundervisning, hur de bedrev sin undervisning och huruvida metoder för uttalsundervisning funnits med i deras egen utbildning. Resultaten visar att lärarna anser att uttalsundervisning är viktigt och att de integrerar det i den övriga undervisningen, exempelvis med grammatik och ordinlärning. Forskarna menar att en slutsats av studien är att metoder för uttalsundervisning i högre grad borde finnas i lärarnas utbildning.

En liknande internetbaserad enkätundersökning har gjorts i Sverige avseende uttalsundervisning i svenska som andraspråk. 92 verksamma lärare i svenska för invandrare, sfi-lärare, svarade på ett 10-tal frågor om sin undervisning i svenskt uttal (Zetterholm 2017). De flesta lärare svarar att de lär ut uttalet av svenska ord i samband med att de undervisar i grammatik eller vid läs- och skrivinlärning. Endast 12 procent uppger att de har särskilda lektioner eller övningstillfällen för uttal, ibland i mindre grupper då klassen bl. a. läser högt i kör med många repetitioner. Rent generellt tycker de flesta lärare att prosodi är viktigare än enskilda språkljud. Knappt hälften av de deltagande lärarna säger att de fått någon explicit undervisning i uttal under sin lärarutbildning.

Vad vi vet har det inte gjorts någon undersökning om uttalets plats i sfi-undervisning i Finland, men Ruoho (2011) har utfört en enkätundersökning om vilka åsikter lärare som undervisar finskspråkiga gymnasieelever i svenska har till uttalsundervisning. I Ruohos undersökning deltog 76 svensklärare. Resultaten visade att lärarna uppfattar uttalsundervisning som viktigt, men att de på grund av tidsbrist inte har tid att tillverka egna övningar. Därtill uppfattar de det som svårt att bedöma elevernas muntliga kunskap. Likaså påpekar lärarna att de gärna vill betona elevernas kommunikativa färdigheter i undervisningen, men att dessa står i kontrast till att det är grammatisk kunskap som betonas i studentexamen. Dessa studier, speciellt de två sistnämnda eftersom

de handlar om svenskundervisning, ligger nu till grund för en jämförande studie med klassrumsobservationer och intervjuer med lärare i en svensk och en finsk kontext.

För att ta reda på hur lärare resonerar om uttalsundervisning för vuxna inlärare av svenska som andraspråk har vi använt teorin om *language teacher cognition* (Borg 2006) som utgångspunkt. Detta används för att förstå samband mellan lärares kunskap om exempelvis språk och metodologiska och pedagogiska kunskaper samt hur de applicerar det i klassrumsundervisningen. I studien jämförs lärarens tankar om uttalets plats i undervisningen och lärarens praktik i klassrummet. Teorin har framför allt använts inom områden som grammatik samt läs- och skrivundervisning i ett andraspråksperspektiv. Studier som presenterar klassrumsforskning med fokus på uttalsundervisning är begränsade och handlar till övervägande del om engelska som målspråk (Baker 2011; Derwing & Munro 2015). Forskning som relaterar till teorin om *language teacher cognition* i ett andraspråksperspektiv och uttalsundervisning är i dagsläget nästan obefintlig (Baker 2011). Baker nämner ett par studier med fokus på engelska som delvis behandlat uttalsundervisning. Resultaten i dessa studier är baserade på vad som sagts om undervisningen utan någon förankring i relationen till den praktiska undervisningssituationen med observationer. Det framkommer i dessa studier att lärare saknar verktyg för att undervisa i uttal, både vad gäller undervisningsmaterial och bedömningsmatriser. Baker skriver även att lärare som har viss kunskap om uttalspedagogik känner sig ofta osäkra i undervisningssituationen när det handlar om uttal.

Teorin om transfer är relevant när det gäller just uttalet eftersom en talarers brytning på andraspråket ofta kan härledas till inlärarens förstaspråk (Jarvis & Pavlenko 2007; Moyer 2013; Zetterholm & Tronnier 2017). Lärare som har kunskap om likheter och skillnader mellan inlärarens förstaspråk och målspråket kan i högre grad individanpassa uttalsundervisningen. En sådan kunskap ger också en större förståelse för att inlärare kan ha olika typer av svårigheter när det gäller såväl uttal av olika språkljud som prosodi i målspråket. I klassrummen finns ofta många olika förstaspråk representerade och genom att uppmärksamma relationen mellan dessa språk och målspråket, i det här fallet svenska, kan man öka den metaspråkliga kompetensen hos inlärare.

### 3 Undersökningen

De lärare som ingår i studien undervisar vuxna inlärare av svenska som andraspråk i Finland och i Sverige. Vi har exkluderat lärare som undervisar elever med finska som förstaspråk, eftersom många finskspråkiga i Finland och Sverige troligen har fått undervisning i svenska under en längre tid och även i övrigt kommit i kontakt med det svenska språket. Inlärarna i de klassrum som obser-

verats har olika förstaspråk, vilket är vanligt i vuxenutbildningen då alla som inte har svenska som sitt modersmål har rätt till undervisning i svenska som andraspråk. Vilka språk som är mest representerade varierar över tid beroende på invandring från olika delar av världen.

### 3.1 Material och metoder

Detta är en etnografisk studie som omfattar deltagande observationer och kvalitativa semistrukturerade intervjuer. Undersökningsmaterialet samlades in i skolor med undervisning i svenska för vuxna andraspråksinlärare i södra Finland och i huvudstadsregionen i Sverige. Undersökningen i Finland utfördes av Huhtamäki och undersökningen i Sverige av Zetterholm. Dels intervjuade vi fyra lärare i vardera landet, dels observerade vi fyra lektioner i svenska i vardera landet. I Finland gjordes inte något speciellt urval av lärare eller lektioner på grund av att det inte finns många kurser i svenska för invandrare, utan författaren intervjuade de lärare som ställde upp och observerade de lektioner dessa erbjöd. En av lektionerna hölls av en vikarie, eftersom den ordinarie läraren fick ett plötsligt förhinder. Det innebär att lektionen inte var planerad länge på förhand. Läraren brukar dock undervisa samma grupp i svenska. I Sverige gjordes urvalet utifrån kontakter på olika sfi-utbildningar, men endast en av lärarna var bekant med forskaren sedan tidigare.

Vi har en gemensam stomme för undersökningen med möjlighet att notera sådant som är aktuellt i varje situation, exempelvis gällande fonetisk kunskap som tas upp och stämningen på lektionerna. Vi har sedan jämfört innehållet i lärarnas svar och observationerna i Finland och Sverige med varandra.

Vi informerade lärarna om syftet med studien och fick då deras samtycke att besöka skolan för att observera deras lektioner och använda intervjuerna i vår undersökning. I all presentation av resultat har vi anonymiserat lärarna.

Lärintervjuerna hade en semistrukturerad ansats där vi hade gemensamma frågor som utgångspunkt. Förutom några personliga bakgrundsfrågor om lärarens och elevernas språkliga och utbildningsmässiga bakgrund tog vi upp frågor om lärarnas syn på uttalsundervisning, vad de tyckte var viktigt, vilka metoder de använde samt hur de bedömde inlärarnas progression. Frågorna bygger på tidigare studier om uttalsundervisning (Foote m. fl. 2011; Zetterholm 2017). Då intervjuerna hade formen av semistrukturerade frågor hade lärarna möjlighet att svara fritt och det fanns utrymme för spontana följdfrågor. De kunde också lägga till sådant som de ansåg vara viktigt i sammanhanget, men som inte kommit fram i frågorna. Syftet med intervjuerna har varit att få en inblick i hur lärarna resonerar om svenskundervisningen, och då framför allt om uttalsundervisningen.

Lärintervjuerna, som var 20–40 minuter långa, spelades in och sedan har en grov transkription gjorts. Därtill har vi utfört en noggrannare transkription

av de lärarsvar som vi vill presentera som exempel i denna artikel enligt valda delar av den transkriptionsmetod som används inom samtalsanalys (se Lindström 2008: 309). Detta innebär att tvekljud och upprepningar har inkluderats i transkriptionen, t.ex. *d-de*, *ähm* och *Det* (.) *det det*. Tecknet (.) betecknar en paus. Varje ny prosodisk enhet skrivs på ny rad med inledande versal. Utdragen är för läsbarhetens skull transkriberade enligt skriftspråkets normer, även om de intervjuade använde sig av uttalslättnader som hör till informellt talspråk.

Klassrumsobservationerna har skett utifrån samma observationsschema i Finland och i Sverige. Därtill har vi skrivit ner fritt formulerade iakttagelser under lektionerna, till exempel om hur läraren korrigerar elevernas uttal. Lektionerna räckte antingen en och en halv eller två och en halv timme (med pauser) i Finland och en till två timmar i Sverige. Observationerna utfördes av Huhtamäki i Finland och Zetterholm i Sverige. Observationsschemat inkluderar frågor om metoder för uttalsundervisning, undervisningsmaterial, hurvida läraren uppmärksammar segmentella eller prosodiska drag, om läraren ger respons på uttal och om kamratrespons förekommer samt hur talutrymmet i klassen fördelas, t.ex. vem som tar initiativ till samtal och vilken typ av respons eleverna ger på lärarens frågor. Sammanlagt inkluderar schemat ett tjugotal frågor, både med färdiga svarsalternativ och öppna svar. Lektionerna bandades inte, eftersom vi ville störa lektionen så lite som möjligt.

## 4 Analys och resultat – lärarintervjuer

Vi presenterar här vad lärarna berättar i intervjuerna om sin undervisning och hur de förhåller sig till just uttalets roll i andraspråksundervisningen för vuxna inlärare. Först presenteras intervjuerna som gjordes i Finland, därefter intervjuerna från Sverige. För att exemplifiera det vi senare diskuterar har vi inkluderat transkriptioner av några lärarsvar från den finländska kontexten.

### 4.1 Lärarintervjuer i Finland

De lärare som intervjuades i Finland undervisar i svenska som andraspråk på CEFR-nivåerna A1 till B1. Ett par av lärarna har examen i något språk och språkläraryrke i sin utbildning och har undervisat i 15 till 30 år<sup>1</sup>. Ingen av dem har utbildning i att undervisa svenska för invandrare. Alla lärare har svenska som förstaspråk. En del av eleverna studerar svenska som en del av sin integrationsutbildning<sup>2</sup>, medan andra studerar språket frivilligt. Det är bara på kurserna som hör till integrationsutbildningen som eleverna bedöms. Eleverna har olika språk som förstaspråk, men de stora grupperna är arabis-

<sup>1</sup> För att lärarna ska få vara anonyma kan vi inte ge noggrannare uppgifter om dem här.

<sup>2</sup> Invandrare kan integreras på finska eller svenska i Finland, se Finlex 2010, speciellt 20 §.

ka, engelska och ryska. Elevernas utbildningsnivå varierar mycket och några av eleverna använder svenska även utanför undervisningen, speciellt de som har en finlandssvensk partner. Ett problem som flera lärare nämner är att det finns alltför få tillfällen för eleverna att använda svenskan i samhället.

På frågan om uttalsundervisning är viktig svarar lärarna jakande. Lärarna framhåller dessutom att uttalsundervisning är ett svårt område inom språkundervisningen. De förklarar att svårigheterna bland annat beror på att de själva sällan har någon systematisk utbildning i uttal, till skillnad från exempelvis grammatik. En av lärarna säger att hen borde ha mer kunskap om fonetik för att kunna undervisa i uttal. Lärarna säger sig också sakna metoder för uttalsundervisning. Som en följd av detta uppger lärarna att uttalsundervisningen ofta sker spontant under lektionerna, exempelvis när eleverna läser högt och fastnar på något visst ljud. De metoder som enligt lärarna främst används är högläsning (enskilt eller i grupp), upprepning efter läraren och genomgång av uttalsregler. Undervisningen i uttal handlar oftast om enstaka ljud, speciellt vokaler.

Enligt lärarna borde uttalsundervisningen helst ske individuellt, men det skulle kräva mer tid och resurser att handleda eleverna individuellt. Någon av lärarna berättar att hen ibland har löst problemet genom att låta gruppen arbeta självständigt med en uppgift och under tiden gått igenom uttal med eleverna en och en utanför klassen. Kursprogrammet sätter enligt lärarna begränsningar för hur mycket klassen kan öva uttal individuellt.

Ytterligare en svårighet med uttalsundervisningen som samtliga lärare lyfter fram är att läromedlen ofta är från Sverige. Detta innebär att böckerna tar upp sverigesvenskt uttal samt Sveriges geografi och kultur. De läroböcker som används på skolan hör till *Rivstart*-serien (se Levy Scherrer & Lindemalm 2015). Lärarna menar att de då anpassar materialet till finländska förhållanden. Ett par lärare gör det genom att läsa in styckena själva för eleverna. Läraren kan också hoppa över sådant i läroboken som inte känns relevant med tanke på finlandssvenskt uttal. Lärarna berättar att det finns extra finlandssvenskt material på nätet, men det är frivilligt för eleverna att använda (se Kotisuomessa.fi, Utbildningsstyrelsen 2017).

Det mål för uttalsundervisningen som lärarna säger sig sikta mot är att tallet ska vara begripligt. Ett perfekt uttal är däremot inget som de eftersträvar. Enligt styrdokumentet för integrationsutbildningen hör det till elevens kommunikativa färdigheter på nivån A1 att eleven "har ett uttal som är någorlunda begripligt" (Utbildningsinstitutionen i Finland (anonymiserad) 2012: 11). På nivån B1.1. ska eleven "ha ett någorlunda bra uttal". Det sker alltså en viss ökning i kraven på uttal mellan de olika nivåerna, men utifrån beskrivningen är det svårt att avgöra exakt hur stor skillnaden är mellan nivåerna. Lärarna ansluter sig alltså närmast till uttalsmålen för A1 i läroplanen. Nedan syns en lärares svar på intervjuarens fråga om uttalsundervisning är viktigt (Exempel 1).

- (1) Att förstå och göra sig förstådd (Lärare B, Finland)

Nå (.) d-de viktiga är kanske för för många av dem är det viktigaste att klara språkexamen.

Och det är en kommunikativ examen alltså man måste visa att man förstår och att man kan göra sig förstådd och då kommer ju uttalet in så att ähm

Jag tycker ju inte att det behöver

Det (.) det det

Det helt primära är inte att de har ett helt perfekt uttal men det måste vara så pass bra att det inte blir obegripligt (...)

I lärarens svar betonas att uttalet har en instrumentell betydelse i en provsituation: eleverna studerar för att klara språkexamen och då är det viktigt att ha ett tillräckligt bra uttal för att examinatorn ska förstå dem. För att eleverna ska uppnå detta mål behöver uttalet emellertid inte vara perfekt. Även en annan lärare påpekar att uttalet inte har ett egenvärde, utan att det handlar om förståelse (Exempel 2).

(2) Uttalet en del av allt det andra (Lärare C, Finland)

Och sen (.) sen är det ju klart det att uttalet om du inte kan uttala så kanske du inte heller kan bli förstådd

(.)

Men att hittills tycker jag ju nog att alla i den här gruppen att oberoende så förstår man vad dom säger att att

Jag tycker att där ska det finnas den där ribban att att att

Om man inte förstår så då måste man jobba mer med uttalet men annars

Uttalet i sig ska ju inte ha ett liksom tycker jag ett egenvärde utan det ska vara en del av allt det där andra

Lärarnas syn på uttal påverkar hur de förhåller sig till att korrigera elevernas uttal. En lärare säger att när uttalet är centralt, såsom vid högläsning, då rättar hen uttal, speciellt om uttalet är så felaktigt att det blir obegripligt. När uttalet inte är det centrala, såsom vid grammatikundervisning, rättar läraren inte uttalet.

En lärare framhåller att uttalsundervisningen måste ske i en kontext.

(3) Uttal hand i hand med innehållet (Lärare C, Finland)

Nå alltså det det måste ju gå hand i hand med det som alltså på något sätt med det där innehållet

Att det kan ju inte vara uttal för uttalets skull

Utan det måste ju gå hand i hand så att man liksom får en kontext till allting att allting måste finnas i ett sammanhang

(.)

Alltså det måste kännas att det här att varför gör vi det här

Att det kan inte vara små delar här och där utan det måste vara ett sammanhang att man måste kunna säga att att det här gör vi därför att

Enligt denna lärare borde uttalsundervisningen integreras med det hen kallar "innehållet". Med stor sannolikhet tänker läraren då på undervisningen i exempelvis grammatik och ordinlärning.

De intervjuade lärarna anser alltså att uttal har en instrumentell betydelse för eleverna och uttal borde integreras med andra moment i undervisningen.

## 4.2 Lärarintervjuer i Sverige

De intervjuade lärarna i Sverige undervisar på nybörjarnivå på kurser inom sfi, svenska för invandrare, dock inte inom alfabetiseringsundervisningen. Det innebär att alla inlärare har en viss skolbakgrund från sitt hemland även om den varierar. Det är stor variation bland eleverna i klassrummen både vad gäller förstaspråk, yrkesutbildning och motivation. Samtliga lärare har en lärarutbildning och behörighet att undervisa i svenska som andraspråk på sfi. Inte någon av lärarna säger sig ha fått explicit utbildning i uttalsundervisning, men de medverkande lärarna har ett stort intresse för uttalet och lägger därför in moment i sin undervisning där de helt fokuserar på uttalet. Endast en av lärarna har inte svenska som sitt förstaspråk, men behärskar svenska språkets grammatik och uttal på en nivå som ligger nära en modersmålstalare.

Samtliga lärare är helt eniga om att uttalsundervisning är en viktig del när man lär sig svenska och de är mycket positiva till att undervisa i det svenska uttalet. Det är viktigt att eleverna kan kommunicera på svenska, är en kommentar som alla lärare ger vid en fråga om målet med uttalsundervisningen. Den svenska prosodin med såväl ordbetoning som frasintonation liksom segmentens duration framhåller lärarna som mycket viktigt för att göra sig förstådd på svenska. Ett helt brytningsfritt uttal är inte målsättningen med undervisningen. Att lära ut uttalet i kombination med annan undervisning i svenska som exempelvis grammatik eller ordinlärning är bra, menar lärarna när vi samtalar om undervisningen.

Det finns många olika modersmål representerade i klassrummet och det kommer nya elever varje månad. Så snart en elev klarat proven på en nivå får hen gå vidare till nästa och då byta klass. Detta gör att det kan vara svårt att ha någon kontinuitet i undervisningen. De intervjuade lärarna säger att en individanpassad undervisning inte är möjlig att genomföra även om det vore önskvärt, framför allt vad gäller elevernas svenska uttal.

Ett par av lärarna gör helt eget material för att kunna anpassa sin undervisning utifrån elevernas behov. På en av skolorna finns ett omfattande undervisningsmaterial med fokus på svenskt uttal att tillgå, men lärarna använder det



ganska sporadiskt. Här finns även dataprogram där eleverna kan lyssna och träna sitt eget uttal, men inte heller det används särskilt ofta, berättar lärarna. En annan skola har inte tillgång till sådant material på samma sätt.

### 4.3 Jämförelse mellan lärarintervjuerna i Finland och Sverige

Intervjuerna visar stora likheter mellan lärarnas inställning till uttalsundervisning i Finland och i Sverige. Alla intervjuade lärare är positiva till uttalsundervisning, men säger att de har för lite utbildning om uttalsmetodik. Lärarna är också överens om att det är viktigt att inlärarna erövrar ett begripligt och lyssnarvänligt uttal, inte ett helt brytningsfritt. Det framkommer också att uttalet företrädesvis bör integreras med den övriga undervisningen i svenska språket.

## 5 Analys och resultat – klassrumsobservationer

Nedan kommer en presentation av våra klassrumsobservationer i Finland och i Sverige. Vi redogör här för det vi observerat med fokus på uttal, i vilka sammanhang lärare talar om och förklarar uttalet samt eventuellt korrigerar elevernas uttal under lektionerna. Några exempel på metoder och lärares kommunikation med eleverna lyfts fram.

### 5.1 Klassrumsobservationer i Finland

De fyra lektioner som observerades i Finland var både på nybörjarnivå och för mera avancerade inlärare. Två av lektionerna var en del av integrationsutbildningen och två av lektionerna hörde till frivilliga kurser i svenska för invandrare. Grupperna varierar från 5 till 18 elever.

Det allmänna intrycket av lektionerna är att det råder en positiv stämning mellan lärare och elever. Läraren har god kontakt med eleverna och uppmuntar dem, och eleverna är motiverade. Det förekommer dock ibland när programmet är mer fritt att eleverna börjar göra annat, såsom att titta på sina telefoner och prata med varandra. Undervisningen är i de allra flesta fall lärarcentrerad, så att läraren styr genomgången av lärostoff till exempel genom frågor. Vid ett tillfälle går klassen igenom olika texter och övningar i läroboken *Rivstart* (Levy Scherrer & Lindemalm 2015) under hela lektionen, vid ett annat tillfälle kombineras läroboken med material och övningar som läraren själv har gjort, exempelvis spel, samt material på nätet, såsom sånger på YouTube. Bara i en av de observerade grupperna är lektionen upplagd så att eleverna ska prata mycket svenska och formulera egna yttranden och därtill prata med varandra.

Uttalsmålet under de fyra lektionerna som observeras är i huvudsak finlandssvenska, då lärarna talar en finlandssvenska som inte har några tydliga

regionala särdrag. Klassen lyssnar på texter eller sånger på sverigesvenska. Vid några tillfällen förklarar läraren skillnader i uttal mellan finlandssvenska och sverigesvenska. I en grupp för de mer avancerade inlärarna säger läraren när de diskuterar *sje*-ljudet följande: *Maria tjänar, med tje- som vi säger i finlandssvenska. Hur skulle du säga på rikssvenska?* En lärare tar också upp skillnader i ords betydelse, exempelvis säger läraren när hen visar en bild på en melon att man också kan säga *arbus* på finlandssvenska (jfr Reuter m. fl. 2016).

När uttal behandlas på lektionerna handlar det i huvudsak om enskilda ljud samt ordbetoning och frasbetoning. Generellt kan dock sägas att det inte förekommer mycket explicit uttalsundervisning under lektionerna. Lärarna tar inte upp något specifikt om fonetik eller fonologi under de observerade lektionerna. Det förekommer att läraren går igenom artikulationen för att uttala så kallade svåra ljud i svenskan, vilket aktualiseras av något ord i en text. Där emot kommer uttal fram implicit i många situationer där klassen går igenom ord och texter. Metoderna som används är främst gemensam högläsning efter läraren, t.ex. får de lära sig att utelägna /t/ genom att upprepa *de här lilla huse*. Läraren kan också vid genomgången av nya ord uttala ordet mycket tydligt och så får eleverna upprepa. Korrigeringar av uttal sker oftast implicit, till skillnad från grammatik och ordval som lärarna rättar explicit. Om eleverna har problem med att uttala ett ord kan läraren korrigera genom att upprepa det övertydligt och med rätt betoning. Det är svårt att veta om eleverna faktiskt uppfattar korrigeringen. Ett par gånger upprepar eleven svaret med rätt uttal efter läraren. Man kan anta att den elev som korrigerades på följande sätt förstod att hen blev korrigerad, eftersom hen upprepade ordet med korrekt vokal (se Ex. 4: uttalet av det aktuella ordet är fonetiskt återgivet).

(4) Korrigering av vokalen i ordet *bord* (Lärare C, Finland)

Elev: ett långt [bord]  
 Lärare: [burd]  
 Elev: [burd]

Uttal finns alltså med i undervisningen i Finland, men tas sällan upp explicit. Eleverna ges rätt få tillfällen att öva uttal i spontant tal. I Ruohos (2011) undersökning om svensklärares åsikter om uttalsundervisning i finska gymnasier kommer fram att lärarna anser det motstridigt att de i undervisningen borde betona kommunikation, medan grammatiken är viktigast i studentexamen. Också i sfi-undervisningen kan det hända att det prov som många av eleverna siktar in sig på påverkar att lärarna måste fokusera på sådant som är centralt i provet, och då hör uttalet inte till dessa aspekter.

## 5.2 Klassrumsobservationer i Sverige

Under de lektioner som observerats i de svenska klassrummen har uttalet fått olika stor plats och undervisningen har varierat. Alla observationerna har gjorts i klassrum med inlärare på nybörjarnivå, vilket betyder att eleverna befinner sig på ungefär samma nivå. Grupperna varierar mellan 10–15 elever. Samtliga lärare har en mycket god kontakt med eleverna och en uppmuntrande attityd som bidrar till interaktion mellan lärare och elev. Som målspråk används en relativt neutral svensk varietet. Undervisningen är till största delen lärarledd men några inslag av parövningar förekommer.

Endast en av lärarna använde ett speciellt dataprogram där eleverna fick lyssna på ord och meningar för att sedan härma uttalet, såväl i grupp som individuellt. Det var ett stort fokus på de enskilda segmenten och läraren rättade ofta uttalet, men utan att göra någon koppling till svårigheter som kunde bero på uttalet av olika språkljud i inlärarens förstaspråk. Eleverna fick också identifiera ord de lyssnade på genom att välja ett av två förslag (minimala par), t.ex. *skor-skur*, *mor-mur*, *gata-gåta*. Man övade också på kvantitet i minimala par som *bar-barr*, *gratis-grattis*. Läraren fick förklara flera av orden som ibland förekom helt enskilt, ibland i kortare fraser.

De övriga lärarna inkluderade uttalet i andra delar av undervisningen, t.ex. undervisning om grammatik. Här använde läraren olika metoder för inläringen – eleverna upprepade orden i kör, de klappade ordets stavelser för att förtydliga betoningen, ibland förekom parövningar där elever med olika förstaspråk fick läsa texter eller svara på frågor tillsammans. Då var det den svenska prosodin med ordbetoning som stod i fokus för undervisningen. I de allra flesta fallen användes material som läraren själv tagit fram. I dessa klassrum var såväl aktiviteten som elevernas motivation för att lära hög.

Någon lärare förklarade artikulationen och bad eleverna titta på munnen vid läpprundning. En lärare var noga med uttalet och kunde på ett enkelt och naturligt sätt rätta elevernas uttal genom att göra en jämförelse med elevens förstaspråk. Ett exempel är då läraren påtalar att */l*-ljudet inte låter svenskt då det bör artikuleras längre fram i munnen (jämfört med liknande språkljud i inlärarens förstaspråk) i ord som *landar* och *spelade*. Eleven fick efter instruktioner upprepa ordet flera gånger och läraren uppmuntrade även de andra eleverna att lyssna och härma. Detta var positivt för hela gruppen då flera omedelbart jämförde sina språkkunskaper och på detta sätt fick möjlighet att erövra en metalingvistik förståelse för uttal av liknande språkljud i olika språk.

## 5.3 Jämförelse mellan klassrumsobservationerna i Finland och Sverige

Gemensamt för klassrumsobservationerna är att de flesta lärare inkluderar kommentarer om uttalet i kombination med undervisning om exempelvis

grammatik. I de finska klassrummen förekommer inte någon explicit uttalsundervisning, vilket kan observeras i ett par av klassrummen i Sverige. Lärarna i de finländska grupperna har ett större fokus på uttalet av de enskilda språkljuden, medan det i flera av de svenska klassrummen ofta är en kombination av såväl språkljud som de prosodiska egenskaperna i det svenska språket. Muntlig interaktion mellan eleverna, exempelvis i form av parövningar, förekommer speciellt i de svenska klassrummen. Ett särdrag i Finland är att eleverna hör både finlandssvenska och sverigesvenska under lektionerna, och ibland jämför läraren de olika varieteterna med varandra i de finska klassrummen.

## 6 Sammanfattande diskussion

Syftet har varit att studera hur lärare tänker omkring uttal och uttalsundervisning i svenska som andraspråk och hur de undervisar i uttal. Elevernas lärande har inte varit i fokus för denna studie, men det vore mycket intressant att ta upp i en uppföljande studie. Då kan man även studera hur inlärares modersmål påverkar inläringen och om det kan relateras till moment i undervisningen.

Det som är gemensamt för de åtta lärare som intervjuats är att de anser att uttalet är viktigt vid andraspråksinläring och att det borde få mera utrymme i undervisningen. Deras intervju svar kan dock vara färgade av att de känner till vårt forskningsfält, som presenterades i samband med att vi fick kontakt med lärarna. En annan gemensam nämnare är att de inte anser sig ha fått undervisning i uttalsmetodik och därför känner sig osäkra på hur de ska undervisa, vilket resulterar i att uttalsundervisning inte får någon högre prioritet för flera av lärarna. De här resultaten stämmer väl överens med studien i Kanada (Foote m.fl. 2011) samt enkät svar som sfi-lärare lämnat i en svensk studie (Zetterholm 2017). Likaså visar en enkätundersökning av svensklärares åsikter om uttalsundervisning i finska gymnasier att lärarna uppfattade att de inte har tid att undervisa i uttal, liksom brist på undervisningsmaterial (Ruoho 2011). En svårighet som lärarna i Finland lyfter fram i vår studie är att läromedlen är framtagna för inläring av sverigesvenska.

Det är ovanligt med explicit uttalsundervisning under de observerade lektionerna. I stället är uttalsundervisningen i båda länderna oftast integrerad i andra moment, som exempelvis grammatik. Detta är också något som kommer fram i den kanadensiska undersökningen (se Foote m.fl. 2011). Det behöver inte vara någon nackdel då man kan kombinera flera moment och även inkludera ordinläring. De metoder som lärarna använder, i den mån det sker någon uttalsundervisning, tycks ganska lika. De använder gemensam högläsning och upprepning av lärarens uttal. Det tycks dock förekomma mer uttalsundervisning i den svenska kontexten jämfört med den finländska, åtminstone i några klassrum. I de finska grupperna handlar uttalsundervisningen i första

hand om segment, medan även prosodiska drag tas upp i de svenska grupperna. Lärarna i Sverige korrigerade också inlärnarnas uttal i högre grad än lärarna i Finland. För att inlärnarna ska erövra ett uttal som blir funktionellt i kommunikation är det värdefullt att lärare kan rätta ett uttal som är svårt att förstå (jfr Moyer 2013, 2014; LeVelle & Levis 2014; Derwing & Munro 2015).

Om läraren rättar en inlärares uttal och gör jämförelser med inlärares förstaspråk kan det underlätta inläringen och ger dessutom möjlighet att öka den metalingvistiska förståelsen för alla elever i gruppen. Den typen av kommentarer avseende uttalet implicerar att läraren har kunskap om att det ofta förekommer transfer mellan inlärares förstaspråk och målspråk när det gäller fonologi och uttal (Jarvis & Pavlenko 2007; Moyer 2013; Zetterholm & Tronnier 2017). Sådana samtal i klassrummet blir med stor sannolikhet så nära en individuell uttalsundervisning som är möjlig att genomföra i en klassrumskontext, men naturligtvis måste lärarna rätta uttalet med hänsyn till varje enskild inlärare.

När vi jämför våra intervjuer med observationerna enligt teorin om *language teacher cognition* (Borg 2006) är det tydligt att lärarna har en medvetenhet och positiv attityd till uttalsundervisning och att flera av lärarna tar upp uttalet på olika sätt i klassrummet. En av lärarna i Finland uttrycker att uttalsundervisningen ska ske i en kontext, vilket kan relateras till att de undervisar i uttal i samband med exempelvis grammatik eller ordinläring. Vi kan alltså se en samstämmighet mellan intervjuer och observationer i klassrummet. Lärarna i Sverige uttrycker att uttalsundervisning är mycket viktigt, men de genomför sin undervisning på olika sätt. Lärarnas egen erfarenhet av uttalsundervisning och inläring av ett främmande språk, med fokus på att göra sig förstådd på ett nytt språk, har enligt Borg (2006) med stor sannolikhet betydelse för lärarens egen undervisning. Resultaten visar inte på någon direkt koppling i det avseendet vad gäller de deltagande lärarna i denna studie.

Det är inte möjligt att bedöma resultatet av själva undervisningen då det enbart är lärarna som står i fokus för den här studien. För att utöka studien bör man göra flera klassrumsobservationer och lärarintervjuer, men även inkludera elevernas inläring och genom intervjuer komplettera med hur eleverna ser på undervisningen och sitt eget lärande. Lärarna i Sverige poängterar att eleverna ska kunna kommunicera på svenska, men några sådana moment kan inte observeras i klassrummet. Huruvida lärarnas egen erfarenhet av språk-inläring och uttalsundervisning avspeglas i deras undervisning är en annan aspekt som bör tas i beaktande i en kommande studie.

Resultatet av studien väcker frågor om utbildningen av lärare i svenska som andraspråk och vad lärare har för kunskap gällande fonologi och uttalsmetodik när de ska undervisa. Det är tydligt att flera av de intervjuade lärarna inte anser sig ha kompetens att undervisa i det svenska uttalet eller att förklara uttalsvariationer och göra jämförelser med elevernas förstaspråk. Detta ger

oss anledning att gå vidare med liknande studier och att även koppla resultaten om lärarnas kunskaper till elevernas inläring och eventuella relation till deras modersmål, men även till kursutveckling på lärarutbildningen.

## Tackord

Tack till de lärare som gett oss både tid och tillträde till klassrummen. Tack också till Humanistiska fakulteterna vid Stockholms och vid Helsingfors universitet för bidrag till studien.

## Referenser

- Baker, A. A. 2011. *Pronunciation pedagogy: second language teacher cognition and practice*. Doktorsavhandling. Georgia State University. [https://scholarworks.gsu.edu/alesl\\_diss/16](https://scholarworks.gsu.edu/alesl_diss/16).
- Borg, S. 2006. *Teacher cognition and language education: research and practice*. London: Bloomsbury.
- Boyd, S. & G. Bredänge 2013. Attityder till brytning: exemplet utländska lärare i svenska skolor. I K. Hyltenstam & I. Lindberg (red.) *Svenska som andraspråk: i forskning, undervisning och samhälle*. 2. uppl. Lund: Studentlitteratur, 437–458.
- Derwing, T. M. & M. J. Munro 1997. Accent, intelligibility and comprehensibility: evidence from four L1s. *Studies in Second Language Acquisition*, 20 (1), 1–16. <https://www.cambridge.org/core/journals/studies-in-second-language-acquisition/article/div-classtitleaccent-intelligibility-and-comprehensibilitydiv/729C15F62F9EC9A51A33EAB5C2D05ED0>.
- 2015. *Pronunciation fundamentals: evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins.
- Finlex 2010. Lag om främjande av integration. Finlex 30.12.2010/1386. <http://www.finlex.fi/sv/laki/ajantasa/2010/20101386> [hämtad 2017-5-30].
- Flege, J. E., M. J. Munro & I. R. A. MacKay 1995. Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97 (5), 3125–3134. DOI: 10.1121/1.413041.
- Foote, J. A., A. K. Holtby & T. M. Derwing 2011. Survey of teaching pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal*, 29 (1), 1–22. DOI: 10.18806/tesl.v29i1.1086.
- Jarvis, S. & A. Pavlenko 2007. *Crosslinguistic influence in language and cognition*. New York (N. Y.): Routledge.
- LeVelle, K. & J. Levis 2014. Understanding the impact of social factors on L2 pronunciation: insights from learners. I J. M. Levis & A. Moyer (red.) *Social dynamics in second language accent*. Berlin: DeGruyter Mouton, 97–118.
- Levis, J. 2005. Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39 (3), 369–377. DOI: 10.2307/3588485.
- Levy Scherrer, P. & K. Lindemalm 2015. *Rivstart: textbok, övningsbok samt lärarhandledning*. 2. uppl. Stockholm: Natur & Kultur.

- Lindström, J. 2008. *Tur och ordning: introduktion till svensk samtalsgrammatik*. Stockholm: Norstedts akademiska förlag.
- Moyer, A. 2013. *Foreign accent*. Cambridge: Cambridge University Press.
- 2014. The social nature of L2 pronunciation. I J. M. Levis & A. Moyer (red.) *Social dynamics in second language accent*. Berlin: DeGruyter Mouton, 11–29.
- Reuter, M., C. af Hällström-Reijonen & B. Holmberg (red.) 2016. *Finlandssvensk ordbok*. Helsingfors: Institutet för de inhemska språken. <http://kaino.kotus.fi/fsob/?p=main> [hämtad 2017-5-30].
- Ruoho, M. 2011. *Svensklärares åsikter om, attityder till och erfarenhet av den muntliga kursen Tala och förstå bättre och övning av muntlig språkfärdighet*. Magisteravhandling. Jyväskylä universitet. <http://urn.fi/URN:NBN:fi:ju-2011052510918>.
- Skolverket 2012. *Utbildning i svenska för invandrare: kursplan och kommentarer*. Stockholm: Skolverket.
- Thomson, R. I. & T. M. Derwing 2014. The effectiveness of L2 pronunciation instruction: a narrative review. *Applied Linguistics*, 36 (3), 326–344. DOI: 10.1093/applin/amu076.
- Utbildningsinstitutionen i Finland (anonymiserad) 2012. *Läroplan för integrationsutbildningen vid denna institution*.
- Utbildningsstyrelsen 2017. Koti Suomessa. <http://www.kotisuomessa.fi/> [hämtad 2017-5-30].
- Yates, L. 2017. Learning how to speak: pronunciation, pragmatics and practicalities in the classroom and beyond. *Language Teaching*, 50 (2), 227–246. DOI: 10.1017/S0261444814000238.
- Zetterholm, E. 2017. Swedish for immigrants: teachers' opinion on the teaching of pronunciation. I E. Papatsouli (red.) *Proceedings of the International symposium on monolingual and bilingual speech 2017*. Chania: Institute of Monolingual och Bilingual Speech, 308–312.
- Zetterholm, E. & M. Tronnier 2017. *Perspektiv på svenskt uttal: fonologi, brytning och didaktik*. Lund: Studentlitteratur.
- Zielinski, B. & L. Yates 2014. Pronunciation instruction is not appropriate for beginning-level learners. I L. Grant (red.) *Pronunciation myths: applying second language research to classroom teaching*. Ann Arbor (Mich.): Michigan University Press, 56–79.

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 61–81.

**Leila Kääntä**

University of Jyväskylä

## **In search of proper pronunciation: students' practices of soliciting help during read-aloud**

This article examines Finnish L2 learners' interactional practices of flagging trouble in pronouncing words when reading aloud texts in English. Using conversation analysis, it describes how students employ three repair initiation techniques – direct requests, trying out, and aborting the reading – as methods through which they mobilize teachers' help in the form of a model pronunciation of the target word. By describing the sequential and temporal unfolding of read-aloud, the article presents an empirical way of tracing those classroom practices that students employ to develop their pronunciation skills of English in Finland. CA-based methodology that focuses on the interactional details of how classroom activities are organized provides new insights on what happens in classroom interaction in terms of pronunciation instruction. The findings not only have local relevance to teachers' pedagogical training in Finland, but also more broadly in showing L2 teachers how classroom activities can be organized to promote practicing of pronunciation skills.

**Keywords:** pronunciation; reading aloud; repair initiation; conversation analysis.



## 1 Introduction

Reading aloud is a method used to practice pronunciation in second or foreign language teaching and learning (e.g. Celce-Murcia et al. 2010; Sicola & Darcy 2015). In fact, a recent survey reports that it is among the most used methods in different second language (L2) classrooms across Europe, including Finland, the context of the present study (Henderson et al. 2012, 2015; Tergujeff 2013). So, even though read-aloud falls under traditional methods of pronunciation instruction (Hismanoglu & Hismanoglu 2010), it continues to be a key classroom activity that “offers a frequent and consistent opportunity for **the teacher** to draw students’ attention to pronunciation” (Sicola & Darcy 2015: 481, emphasis added). In this article, I argue that reading aloud also provides opportunities **for students** to orient to pronunciation.

When learning to read aloud in a L2, students need to learn to recognize written words, i.e. identify their orthographic form (Grabe 2009: 21–38), and to phonologically decode them, i.e. to pronounce them (Koda 2004: 33–34). Research on L2 reading suggests that decoding of written words is easier in the L2 when its orthographic system bears resemblance to that of students’ first language (L1; Koda 2004: 37–43). However, the orthographic system of English, the target language, differs considerably from that of Finnish, the students’ mother tongue. Thus, the ability to decode words in Finnish is not sufficient: students require instruction and practice in oral reading and pronunciation in English to improve these skills. For this reason, it is important to investigate what kinds of problems students encounter in decoding words and identifying their orthographic forms when they read aloud texts in English, and therefore the practices with which they seek pronunciation help from teachers. The knowledge gained by studying students’ ways of practicing English oral reading and learning how to pronounce words in L2 classroom interaction is of essence for pre- and in-service teachers. To that end, the article presents the findings of a small-scale study on the interactional practices that students employ to develop their pronunciation skills of English in Finland and suggests an activity in which these practices can be applied in teaching pronunciation through read-aloud.

The research question the study answers is what kind of repair initiation techniques students employ to flag trouble in pronouncing the next word in the text read aloud, and thereby seek teachers’ help. To answer the question, the theoretical and methodological framework of conversation analysis (CA) is used. It is argued that CA with its focus on the interactional details of classroom activities, and particularly on teachers’ and students’ methods of making sense of what they are doing moment-to-moment, can shed light on what happens in praxis in the classroom with respect to pronunciation instruction. By describing the sequential and temporal unfolding of read-aloud activities

in naturally occurring classroom interaction, this study offers new insights on how repair initiations are used to seek pronunciation help in the L2 classroom and how such practices are accomplished in and through instructional interaction. Overall, the use of CA in the field has slowly gained ground through, for instance, studies on the prosodic and rhythmical features of turns-at-talk and how this knowledge can be used in teaching and learning L2 pronunciation (e.g. Szczepek Reed 2015).

## 2 Corrective practices in pronunciation instruction in classroom interaction

According to Murphy & Baker (2015: 36), research on pronunciation instruction that examines teachers' and students' actual practices in L2 classrooms is still in its infancy. However, a focal topic of analysis in different areas of L1/L2 classroom research on pronunciation instruction is teachers' practices to correct student errors during classroom interaction. Studies on the topic have mainly employed quantitative, and to an extent experimental methods, by means of coding student errors and teacher corrective moves into different categories and investigating their distribution across lessons, teachers and learner groups (e.g. Allington 1980; Lyster 1998; Saito & Lyster 2012). For instance, Allington (1980) found that teachers corrected primary school children's L1 oral reading, including pronunciation errors, either during or right after the error was produced. Lyster's (1998) study in a L2 immersion classroom context conveyed that teachers mainly employ recasts to correct students' phonological errors, both decoding errors during read-aloud activities and mispronunciation errors. Similarly, Foote, Trofimovich, Collins and Soler Urzúa's (2016) classroom observation study showed that teachers correct pupils' reading errors or miscues via recasts, while explicit corrections and prompts are used to a lesser extent.

In Finland, Tergujeff (2012) observed Finnish teachers' methods of teaching English pronunciation to Finnish students. She analyzed the focal lessons with a pre-prepared observation form and identified ten different methods, among them activities like 'listen and repeat' and 'read aloud'. However, there were also more specific methods, e.g. correcting students' pronunciation and pointing out errors or typical pronunciation-related issues. Her findings differ from the studies cited above in that she did not consider recasts as addressing pronunciation-related problems and thus she excluded them from the data. In contrast, her findings underline the teachers' frequent use of explicit corrections, while other methods were used less. Interestingly, she did not observe whether during the read-aloud, teachers corrected students' pronunciation errors.

Overall, while there are studies on teacher corrections in pronunciation instruction, research on students' role in seeking help in relation to oral reading and pronunciation issues in L2 classroom interaction is nonexistent. Learning more about how students can become agents of their own learning processes is crucial for the development of all kinds of pedagogical practices, including pronunciation instruction (Celce-Murcia et al. 2010: 362).

### **3 Word searches and 'doing pronunciation' as forms of interactional repair**

In everyday conversations, repair is an interactional phenomenon that deals with participants' problems of hearing, speaking or understanding talk which can compromise participants' establishment of mutual understanding of that talk, i.e. the achievement of intersubjectivity (Schegloff et al. 1977). Repair can be initiated by the speaker of the trouble source (self-initiation) or by its recipient (other-initiation) and it can be solved by self (self-repair) or other (other-repair). The problems are referred to as 'trouble sources', which can be anything interactants deem in need of repairing so that intersubjectivity is maintained (Schegloff et al. 1977: 363). Two types of trouble sources are related to problems of speaking in interaction: word finding difficulties and pronunciation problems. The former concerns situations where speakers try to find a word to incorporate into their talk but are momentarily unable to do so (e.g. Schegloff et al. 1977; Goodwin & Goodwin 1986), while the latter refers to situations where speakers know the word but do not know how to pronounce it (Brouwer 2004; Koshik & Seo 2012). In both cases, the current speaker performs a repair initiation to display trouble in producing the emerging turn. Depending on the situation, the trouble is solved through self- or other-repair.

For the current study, previous findings on self-initiated repair techniques, which indicate to co-participants that their help is sought during word searches, are of relevance. In general, speakers do a great deal of interactional work to show that a word search is in progress and where they are in their search: whether resolution is achieved or not (Goodwin & Goodwin 1986; Hayashi 2003). When speakers are not able to resolve the problem, they initiate repair to seek recipients' help. The mobilization of co-participants' help happens through both verbal and nonverbal means. Among the verbal techniques are repetitions, revisions, and other explicit word search markers (Goodwin & Goodwin 1986; Brouwer 2003; Hayashi 2003). Speakers also often ask co-participants to provide the searched-for item by *wh*-questions (Oelschlaeger 1999; Brouwer 2003; Radford 2009). Speaker's gaze is in such situations directed toward the recipient, whereby it also mobilizes joint reso-

lution to the problem (Goodwin & Goodwin 1986; Oelschlaeger 1999; Hayashi 2003; Radford 2009, 2010). On other occasions, merely the speaker's gaze, without accompanying verbal indicators, is effective in inviting help. On the other hand, in interactions that involve the use of books, speakers do not necessarily employ their gaze when seeking recipient's help (Oelschlaeger 1999; Radford 2009). Instead, speaker's verbal indicators, wh-questions and self-cues, suffice in drawing a candidate solution from co-participants.

To my knowledge, the first CA study on pronunciation is Brouwer's (2004) on participants' interactional practices of 'doing pronunciation' in everyday L2 conversations. For her, 'doing pronunciation' represents a type of repair activity, on account of which she has identified three self-initiation techniques that L2 speakers employ to signal difficulty in producing, and pronouncing, a Danish word, thereby inviting help from the L1 speaker. The first technique entails the use of speech perturbations that include intra-turn pauses, word cut-offs, vocalizations (e.g. *uhh*, *euhh*) and sound stretches (see Schegloff et al. 1977; Goodwin & Goodwin 1986; Hayashi 2003; Radford 2009 for different L1 contexts). She shows that these phenomena signal trouble with the progressivity of the emerging turn. The second technique involves the use of rising intonation that locates the trouble source, while the third technique includes the repetition of the trouble item with or without framing practices. These remedial techniques help display that the speaker is initiating repair. The three techniques are used in different combinations and sequential constructions that clearly establish that 'doing pronunciation' is in play. In a more recent study, Koshik & Seo (2012) investigated ESL tutoring sessions and the tutees' practices of eliciting help during word searches. With respect to pronunciation problems, the findings show that the tutees employ rising intonation and interrogatives to seek confirmation for the way they pronounce words.

While Brouwer's (2004) and Koshik and Seo's (2012) studies shed light on the intricate interactional work participants accomplish in achieving shared understanding of the action they are performing in and through their turns-at-talk, this study illustrates how the institutional setting and the ongoing activity framework set boundaries for the range of actions students can perform. Since students read pieces of text aloud, pronunciation practice is established at the start of the task as a goal (also Tergujeff 2012). Pronunciation problems are thus potential trouble sources students encounter during the read-aloud activity. In contrast, word search troubles are not amongst them as students have all the words in the text. Despite these differences, this study underlines the similarity of the techniques used by L2 speakers in ordinary conversations (Brouwer 2004) and L2 learners in classroom interaction to solicit pronunciation help from co-participants.

## 4 Method and data

This study draws on the theoretical and methodological underpinnings of CA, which examines everyday social interaction and participants' methods of making sense of the interactions they are part of (see e.g. Sidnell & Stivers 2012). CA describes how participants produce different social actions (e.g. questions, requests for help, and instructions) and display to each other their understanding of what is happening at any moment in interaction. CA adopts an emic perspective – a participant perspective – into analyzing interaction by examining the audible and visible (i.e. talk and embodiment) means participants utilize in designing, for example, requests for help. Since participants' own understandings of the actions they perform both form the loci of the analysis and drive interaction forward, research can unveil those interactional practices related to pronunciation that participants themselves orient to as interactionally meaningful and consequential as interaction unfolds. For the analysts to be able to describe the details of the verbal and embodied resources participants utilize, the data comprise video-recordings of naturally occurring interactions that enable the repeated viewing and scrutiny of participants' interactional practices. The reported findings are thus based on the rigorous analysis of the data and the description of interactional events from the participants' viewpoint.

The data come from a classroom data corpus collected in co-operation by the Department of Languages and the Center for Applied Language Studies in the University of Jyväskylä, Finland. It consists of 50 video-recorded lessons that range from Year 6 in Elementary school to Year 12 in Upper Secondary School. Both English-as-Foreign-Language (EFL) and Content-and-Language-Integrated-Learning (CLIL) lessons have been recorded. The CLIL lessons include such subjects as history, physics, biology, religion, chemistry, physical education, and English. However, in all the lessons, English is the main medium of instruction and a target of learning. Due to the wide range of lessons, the students' level of English varies a great deal, and thus the level of English used differs. Except for two native-English speaking students in the CLIL biology and religion lessons, the students are native speakers of Finnish. Of the 14 teachers who taught the lessons, three are native speakers of English and the rest are native speakers of Finnish. The participants in the analyzed data extracts are all Finnish-speaking teachers and students, and students' names are pseudonyms.

For closer analysis, classroom tasks in which students read aloud texts written in English have been chosen. The tasks range from checking and doing exercises with the whole class to group work situations, where students report on their written product, often by reading aloud the text to the rest of the class. The length of the piece of text read aloud, therefore, varies from

short clauses to longer paragraphs. What is common to all tasks is that they are teacher-assigned and pedagogically have a dual-focus: there is a focus on accomplishing the ongoing task, whatever that is, and a focus on practicing oral reading. The main aim is not to develop students' oral reading per se, but rather to provide them with opportunities to read aloud and simultaneously practice pronunciation (also Tergujeff 2012) while another main activity is accomplished. Within the tasks selected for closer examination, the analysis has centered on instances where students audibly and/or visibly flag trouble in reading the text aloud. The resulting collection includes 14 instances across which ten students flag pronunciation trouble, i.e. it is a small collection. However, in most instances, as the analysis will show, there is an agglomeration of techniques which have not been taken into consideration in counting the instances. In addition, all the instances occur in two data sets: in Year 12 EFL lessons and Year 8 CLIL History lessons.<sup>1</sup> Excluded from the collection are teachers' corrections of students' pronunciation errors and instances where students do self-repair (see Extract 5, l. 6). Likewise, excluded are students' recognition problems, for example not being able to identify and/or pronounce roman numerals (e.g. *Henry VIII*).

## 5 Three self-initiating repair techniques

The analysis shows that students employ three techniques to flag trouble in relation to reading aloud the next item due in a text. They range from (a) requests of how a word is pronounced, (b) to trying out by phonological cluing and producing different types of speech perturbations, (c) to visibly aborting the oral reading that manifests in the form of a prolonged silence. Through these techniques, students initiate a help seeking sequence, a side sequence (Brouwer 2004), which consists of an adjacency pair: the student's repair initiation and the teacher's other-repair. The request for help is primarily addressed to the teacher, and thus the teacher is positioned as the more knowledgeable party language-wise, thereby being entitled to provide the pronunciation model (also Brouwer 2004; Koshik & Seo 2012).

Next, I shall provide illustrative data extracts of each technique and how they figure into the subsidiary activity of read-aloud and thereby into the primary activity of accomplishing the ongoing task. Although each analytic chapter focuses on a technique, the extracts demonstrate how several techniques are in play in a help seeking sequence, thus explicitly manifesting what a student's problem is. The analysis also delineates how the techniques include both retrospective and prospective practices (Schegloff 1979; Streeck 1995;

<sup>1</sup> Although the read-aloud activity was practiced in several subject lessons in the corpus, there were no student-produced repair-initiations viz. pronunciation in the other lessons.

Brouwer 2004) and how the resolution of the trouble momentarily delays the progressivity of the reading activity, after which it is resumed.

### 5.1 Request for help

The request for help is performed verbally through an interrogative (also Koshik & Seo 2012). The interrogative form both locates and indicates the nature of the student's problem, i.e. that there is a pronunciation problem with the next item due. The request can be performed in Finnish (Ex. 1) or in English (Ex. 2), although the text is in English and the ongoing activity is conducted mainly in English.

Extract 1 is from a Year 12 EFL lesson, from a whole class activity of checking a homework exercise on numerals, in which the students had to fill a text in English according to Finnish prompts. At the beginning of the activity, the teacher has instructed the nominated students to read aloud the whole sentence instead of just the target form, i.e. practice oral reading and thus pronunciation.

#### (1) EFL English\_crucial

1	T	*an' the last ↑one? *T GAZE AT HER BOOK
2		(19.6) T LOOKING AT HER BOOK GLANCING AT CLASS LOOKING AROUND AT CLASS
3	T	anyone?= <i>Katja</i> *T GAZE SHIFT TOWARDS TRANSPARENCY
4	<i>Katja</i>	»since then rock an' roll has been a« (0.4)
5	→	miten tuo lausutaa¿= how that say+PASS <i>how do you say that</i>
6	T	*= <i>crucial</i> . *T GLANCES TOWARD CLASS/KATJA
7	<i>Katja</i>	*» <i>crucial</i> part in musical experience in (x) *T GAZE DOWN AT TRANSPARENCY
8		twenty first century (x) remains to been seen«
9	T	hm↑m

Although the teacher waits for a long time for the next respondent (l. 2), she is able to select *Katja*<sup>2</sup> (l. 3), who begins to read the sentence from her book (l. 4). However, shortly after, she stops in the middle of the sentence and a silence emerges. It is followed by the request of how the next item due is said, produced in Finnish (l. 5). The teacher immediately provides the model

<sup>2</sup> *Katja* is not in either of the cameras that were used to record the lesson, so it is difficult to say whether she raises her hand to volunteer.

(l. 6) so that her repair turn latches Katja's (see Appendix for transcription conventions). Katja repeats the item by incorporating it into the sentence as she continues the reading (l. 7).

The student here pre-emptively invites help from the teacher on the proper pronunciation of the word 'crucial' before she has tried to say it herself. By producing the request, and by doing it in Finnish, she not only signals trouble but also locates the trouble source to be the next item in the sentence through the demonstrative pronoun *tuo* (Eng. *that*). Although the 0.4 s pause in line 4 can be seen to indicate potential trouble in terms of the progressivity of her reading, it does not yet serve to specify the nature of the trouble, or that there is trouble, while the request does this explicitly.

Extract 2 differs from Extract 1 slightly as the interrogative is produced in English and the student first tries to say the word before she seeks help. It comes from a Year 8 CLIL history lesson from a quiz activity on Stuart period in Britain.

(2) CLIL History9\_puritans

1	T	↑okay (1.3) and (.) we can conti↑nue
2		(0.4) T GAZE DOWN AT DOCUMENT CAMERA ESTERI GAZE AT BOOKLET
3	Esteri	»the rise of the (1.1) (pur:)<« (0.5)
4	→	>ho+w do you s[ay (that)< +ESTERI GAZE SHIFT TOWARDS T
5	T	[p <u>u</u> ritans.
6	Esteri	»puritans. (0.3) during James' reign
7		<radical (.) pro-testing> (.) groups called (0.3)
8		Puritans began to gain a sizable following. (.)
9		Puritans were- (.) Puritans wanted to pur- pur- (.)
10		p <u>u</u> rify the church by ↓paring down church riche- r <u>i</u> tual. (0.3)
11		educating (0.6) the (0.3) cler-ç (0.3) gy: (0.4) cler« (1.4)
12	T	mhm

Prior to the extract, the teacher has elaborated at length on a historical event related to Guy Fawkes and the Gunpowder Plot. In line 1, the teacher indicates that she is done with the explanation and that the activity can continue. Esteri, who has been assigned to read the text, recommences from where she left off before the teacher's explanation (l. 3). Shortly after, however, she stops and a silence emerges (l. 3). The silence is followed by a try from Esteri to pronounce the next item due, after which another silence follows. It is at this point that Esteri requests for help (l. 4). Again, the pronoun *that* indicates that the next item due is the trouble source. Her request is also visibly directed to the teacher as Esteri raises her gaze from the text toward the teacher. The teacher provides the model partly in overlap with Esteri's interrogative (l. 5). As Esteri resumes the reading activity, she incorporates the trouble item into her reading.



A crucial difference between the extracts is that in Extract 1 the request is forward-oriented since Katja does not try to pronounce the word beforehand, while in Extract 2 it is backward-oriented as Esteri first tries out the word before she initiates repair (see also Schegloff 1979; Streeck 1995; Brouwer 2004).<sup>3</sup> Despite the difference in the temporal orientation of the repair initiations, the sequence unfolds similarly in both extracts as the teacher's turn is produced immediately after or partly overlapping the request and both students incorporate the trouble source item into their reading as they resume the activity. However, in Extract 2, the fact that the teacher produces the model partly in overlap with the end of Esteri's request suggests her orientation to the rather lengthy silences and the trying out (l. 3) as indices of pronunciation trouble. Yet, she provides the model only after Esteri has begun to request help.

Both extracts reveal the importance of the institutional context and the goals of the ongoing activity framework with respect to how trouble is flagged and help is requested during read-aloud in L2 classroom interaction. In Brouwer's study (2004) no such requests were deployed, while in Koshik and Seo's (2012) study they were used as the last resource to indicate a pronunciation problem. The trouble was then dealt with through an extended repair sequence. Since here the ongoing activity is related to practicing pronunciation through read-aloud, it is natural that the most likely trouble students encounter is related to decoding the target words. In such instances, the participants' interactional work and the disruption of the ongoing activity is minimal, an adjacency pair, which is produced quickly, after which the main line of activity is resumed. The requests are thus quite an efficient way to solve the problem.

Interestingly, there is a difference in the language with which students initiate the repairs. While Katja uses Finnish in the EFL lesson, Esteri requests help in English in the CLIL lesson. The use of the two languages may reflect the English-only policy that the CLIL teacher imposes in her lessons (see Jakonen 2016), while English and Finnish are both legitimate languages in the EFL lessons. However, more empirical evidence would be needed to argue whether this really is the case.

## 5.2 Trying out

The second technique involves a process of trying out, i.e. attempting to pronounce the word. The technique resembles what Radford (2009, 2010) has

<sup>3</sup> According to Schegloff (1979: 273–275), different turn design features indicate whether repair is forward or backward-oriented. Features like pauses and 'uhh's generally precede the repair initiation and are thus forward-oriented, while cut-offs manifest backward-oriented repair. This is because the trouble source item has been / is being produced already.

identified as phonological self-clue strategy used by speakers with language impairments. In phonological cluing the speaker produces the first sound of the troublesome item (Radford 2010: 85). In the present data, the trying out through phonological cluing can range from the first sound of the word (Ex. 3) to the first syllable of the word (Ex. 4) to almost the whole word (Ex. 3). Likewise, different speech perturbations such as cut-offs, vocalizations and sound stretches along with silences occur when students try to pronounce the trouble item, but fail. Extracts 3 and 4 illustrate how the teachers orient to the trying out as an indication of pronunciation trouble. Hence, the cluing and the speech perturbations suffice in displaying to the teacher the nature of the student's problem.

Extract 3 comes from an EFL lesson from a whole class activity in which the teacher says a piece of text in Finnish and the students need to find an English translation from the textbook and to read it aloud.

### (3) EFL English\_mirth

1 T \*<twelve?> (.)  
\*T GAZE DOWN AT HER BOOK

2 »joka on mutkikas mielikuvituksesta ja hilpeydestä punottu  
*which is complicated fantasy and mirth woven*  
juo↓ni«  
*plot*

3 (7.0) T GAZE DOWN AT HER BOOK  
AT 2.9s GAZE SHIFT TO CLASS

4 T \*Moo\*na  
\*T GAZE TOWARDS MOONA  
\*T GAZE SHIFT DOWN AT BOOK

5 Moona »which is a (.) complicated (0.2) plot.  
6 → woven of fantasy an' m- (.) mir-«=  
7 T =mirth.=  
8 Moona =>mirth«=  
9 T =hm↑m (0.5) that's ↑right.

The teacher selects Moona as the next respondent in line 4. As Moona reads the text, she tries out the last word of the phrase twice (l. 6). She first utters only the first sound of the word, which is followed by a brief pause. She then retries to pronounce the word, but aborts. At this juncture, the teacher produces the model so that her turn latches Moona's (l. 7). Moona resumes the reading by repeating the word, and thus finishing the phrase. Neither of them gaze toward each other during the reading; rather they gaze down at their books.

Extract 4 comes from the same CLIL history lesson as Extract 2, i.e. from the quiz activity.

## (4) CLIL History9\_anonymous

1 Esteri »the (.) Gunpowder Plot. (.)  
 2 James was a firm protest- (.) protestant  
 3 and in six<sup>↑</sup>teen o': f:our,  
 4 he expelled all catholic priests from the island. (0.3)  
 5 this was one of the factors, (.)  
 6 which led to the Gunpowder Plot of sixteen o' five. (.)  
 7 a group <sup>↑</sup>of catholic plotters (0.3)  
 8 planned to blow up parliament when it opened of November fifth  
 9 → (0.6) >ho'ever< an: (1.2) an:[:]:«  
 10 T [uh anonymous  
 11 Esteri »anonymous letter betrayed the plot, (.)  
 12 and one of the plotters (.) Guy F:awkes (.)  
 13 was captured in the cellars of the house (.) of parliament  
 14 wi- with enough (0.6) enough to blow the place sky ↓high.«

Akin to Extract 3, Esteri tries to pronounce the next item due in line 9. Already when she reads the indefinite article *an* she slows down and stretches it. After this, a long pause unfolds, during which the participants gaze down at their texts. Esteri then tries out the target word by stretching the beginning of the second syllable as if it was the first. The teacher provides the model in line 10, partly overlapping Esteri's try. Next, Esteri continues the reading by incorporating the word into the sentence (l. 11). Throughout the sequence, both the teacher and Esteri gaze down at their texts.

Both extracts bring forth interesting phenomena related to dealing with pronunciation trouble in classroom interaction as the phonological cluing, and the sound stretches, cut-offs and silences clearly display that the students are experiencing trouble with the next word due. Consequently, these indices also serve to locate the trouble source to be the tried-out item and indicate that the problem is related to pronunciation. In this respect, they are also specific examples of backward-oriented techniques to flag trouble (Schegloff 1979; Brouwer 2004). Moreover, the teachers' other-repairs are performed in latching or in overlap with the second try of the word. Thus, the resolution of the trouble is quickly dealt with when the teacher provides the pronunciation model. No further explanation or interactional work occurs at this point, and the student can resume the reading.

### 5.3 Silence after aborting the reading

The third technique consists of an emerging, prolonged silence that audibly manifests that a student has aborted the reading. The silence is what teachers seem to orient to as a primary indicator of a problem, although a range of other features such as students' embodied actions (Ex. 5) and vocalizations (Ex. 6) can further the interpretation. Extract 5 comes from an EFL lesson from

an activity where the class is beginning to discuss a piece of poetry by Edgar Louie Masters and a student, Eeva, is requested to read the introduction to the theme.

(5) EFL English\_equivalent

1 T \*Eeva \*could you give it a try?  
\*T GAZE AT EEVA  
\*T GAZE SHIFT DOWN AT HER BOOK

2 (1.2) T GAZE SHIFT UP TOWARDS EEVA  
EEVA GAZE DOWN AT HER BOOK

3 Eeva \*mm »by the way of introduction.  
\*T GAZE SHIFT DOWN AT HER BOOK (UNTIL L. 10)

4 American Edgar Louie Masters was:: originally a lawyer  
5 but after practising law for several (0.6) years he  
established  
6 his repsta- repu[tation as a] poet.« (0.7)  
7 T [reputa↑tion]

8 Eeva »he is best known for the spoon river anthology  
9 nineteen fifteen (0.5)  
10 → which he intended as a modern«  
(- -+ - - \*- -)  
+EEVA RAISES HEAD SLIGHTLY UP  
\*T RAISES GAZE TOWARD EEVA

11 T equival<sup>u</sup>lent.  
12 (0.5) T GAZE SHIFT DOWN AT HER BOOK  
EEVA GAZE SHIFT DOWN AT HER BOOK

13 Eeva »equivalent (0.3) of old Greek epitaph. (0.8)  
14 it is a series of poetic monologues by the (0.9)  
15 two hundred an' forty-four inhabitants of spoon river.«

As Eeva reads, we can see that in line 10 she suddenly stops, after which a silence of 0.6 seconds emerges. During it, Eeva raises her head slightly up from the text, but does not shift her gaze toward the teacher (also Radford 2009). The teacher, in contrast, raises her head from her book and directs gaze toward Eeva. Next, she produces the model (l. 11). It is followed by a short silence, during which both the teacher and Eeva lower their gazes at their books. When Eeva resumes the reading (l. 13), she incorporates the trouble source item into it. It seems that in addition to the emerging silence, Eeva's slightly raised head invites the model from the teacher, as it visibly manifests that Eeva has aborted the reading activity and is not oriented to the book as intently as before.

Extract 6 comes from a CLIL history lesson from an activity, in which the class is checking a quiz the students have done on Queen Victoria and her reign.

## (6) CLIL history\_jubilee

1	T	.hh okay. Inkaʒ
2		(2.0) INKA GAZE AT HER ANSWER SHEET T GAZE SHIFT DOWN AT HER ANSWER SHEET
3	T	number sixteen.
4		(1.0) INKA GAZE AT HER ANSWER SHEET T GAZE AT HER ANSWER SHEET
5	Inka	↑uhh (0.2) »when Victoria had been in power for fifty years,
6	→	she held her« +(1.2) ↓euhh+ (0.3) +INKA RAISES EYEBROWS+
7	T	jubilee.
8		(0.5) INKA GAZE TOWARD ANSWER SHEET T GAZE AT HER ANSWER SHEET
9	Inka	»£jubilee£ (0.3) wearing a red dress an' drove through
10		London where people cheered her.«

As Inka reads the true-or-false sentence from the quiz sheet, she suddenly stops (l. 6). A silence emerges, during which she continues to gaze at the sheet. Although her face is visible only diagonally (the camera is behind her), it appears that she raises her eyebrows, as if to display surprise of what the next word is. The facial expression is accompanied by the vocalization that is produced in lower pitch. However, at this point the teacher gazes toward the answer sheet in front of her and does not see Inka's facial expression. This demonstrates that she orients to the silences and the vocalization as indication of trouble and produces the model pronunciation accordingly (l. 7). After another silence, Inka smilingly pronounces the target item and resumes the reading (l. 9).

In Extracts 5 and 6, the silence is a strong indicator for the teacher to realize that the student is experiencing pronunciation trouble in relation to the next item due, i.e. that the silence is not only a momentary break from reading. The extracts, however, differ from one another in that in Extract 5, the activity sequence contains a prior instance of pronunciation trouble, where the teacher performs an other-correction by modelling the word *reputation* (l. 7) in overlap with Eeva's self-repair (l. 6). In addition, the silence that emerges during Eeva's reading (l. 10) is notably shorter than in Extract 6, where there is no such prior trouble. In Extract 6, the prosodically marked vocalization that Inka produces further underlines the next item as a trouble source.

## 5.4 Summarizing discussion of analysis

Out of the three repair-initiation techniques, the request for help explicitly seeks the teacher's involvement in the resolution of the trouble through an other-repair (cf. Oelschlaeger 1999; Brouwer 2003; Radford 2009). Unlike in

Seo and Koshik's (2012) study, the request is not used as the last resort (although see Ex. 2). The main reason for this is that the request provides the most precise way of indicating, locating and resolving the trouble, and thus can be conducive to the progressivity of the reading activity when the side sequence remains short. Overall, the request is used in five cases out of the 14 in the current collection. Curiously, trying out, which comprises phonological cluing alongside different speech perturbations, is the most used practice (7/14). Potential explanation is that participants prefer self-repair over other-repair, in a similar manner as Schegloff et al. (1977) have proposed for everyday conversations.<sup>4</sup> However, in these instances, the students are not able to perform the repair, so the teachers provide the model, thereby helping students achieve their 'zone of proximal development' (Vygotsky 1978). In contrast to the request, trying out indicates the nature of the trouble through its backward-oriented nature and locates the trouble by disrupting the progressivity of the emergent reading. The prolonged silence, on the other hand, manifests audibly as the abortion of the reading that locates the trouble, but does not specify the nature of the trouble. Perhaps due to this, the prolonged silence by itself is the least used practice in the current data (2/14). The understanding that the silence indicates pronunciation trouble is invoked mostly by the larger activity framework of the read-aloud activity (Ex. 6) or by an earlier occasion of pronunciation trouble (Ex. 5). Additionally, the silence can be accompanied by different embodied actions, such as raising one's head slightly in an expectant manner that provides a visible clue to teachers that their help is needed. As far as gaze is concerned, the analysis highlights that students' help seeking practices in the analyzed context do not involve participants' gaze contact (also Radford 2009). Although one of the participants can shift their gaze toward the other (Ex. 5), participants' orientation is mostly directed to the texts. The intensity with which the text is oriented to can then function as a resource, which mobilizes co-participants' help.

Overall, all the audible and visual resources used display participants' finely tuned orientation to the interactional relevance of the analyzed techniques for organizing classroom interaction. So, even though only 14 instances were found in all the read-aloud activities, the techniques form a 'practice' that teachers recognize as interactionally consequential since they model the pronunciation. Moreover, it is a practice that is used in two different data sets that were recorded seven years apart in different cities in Central Finland.

Although the analysis introduced the three techniques separately, they are generally produced in a range of combinations, akin to Brouwer's study

<sup>4</sup> The preference for self-repair is further attested by the vast number of self-repairs that students perform during read-aloud and that teachers do not orient to as requiring their help. However, I have not counted the students' self-repairs as they outnumber dozen, most likely more, times the number of other-repairs.

(2004). For instance, in several extracts, a silence precedes either the request (Ex. 1) or the trying out (Ex. 4). This indicates the interactional work participants do to establish that there is a trouble source and that the trouble source in this sequential and temporal position manifests a pronunciation problem. However, since the subsidiary goal of the ongoing task is to practice pronunciation, participants' orientation is more readily and demonstrably directed towards pronunciation problems, unlike in other settings (e.g. Brouwer 2004; Radford 2009; Koshik & Seo 2012). Consequently, the participants understand the nature of the trouble quickly and orient to solving it as efficiently as possible. This is visible in that the repair is resolved through an adjacency pair (i.e. student request for help and the teacher modelling). Although there is a clear disruption in the read-aloud activity, it is only momentary. Immediately after the trouble has been solved, the reading is resumed.

When the reading is resumed, most often than not students incorporate the trouble source into their reading. The incorporation manifests that students imitate the teacher's pronunciation model, whereby they practice its pronunciation (also Koshik & Seo 2012). An interesting question is why in some cases students can incorporate the trouble source immediately into their reading, while in others a silence emerges before the reading is resumed. In Extract 5, Eeva lowers her head first, which may be the cause of the delay. However, in Extract 6, Inka gazes at the sheet for 0.5 seconds and then pronounces the word with a smiling voice. This raises the question whether *jubilee* is an unfamiliar word, which she can neither decode phonologically nor identify orthographically, nor above all, understand its meaning.<sup>5</sup> An opposite example is Esteri's pronunciation of *anonymous* (Ex. 4), which is produced immediately after the teacher's modelling. Its pronunciation embodies familiarity with and recognition of the word and its meaning. Whether there is something to these trouble sources, in terms of (un)familiarity with word meanings, when students resume the reading is a topic for future research as the current data collection is too limited. But what can be claimed is that, to an extent, the analyzed student practices create particular kinds of learning opportunities for students in how to decode and pronounce English words. These opportunities are something the students have themselves instantiated by seeking help from the teacher. The findings thus highlight that teachers need to be made aware of the importance of such practices for students' learning, whereby they can become more sensitively tuned to students' divergent actions and help create a sense of agency for students in their own learning process (e.g. Celce-Murcia et al. 2010: 362).

<sup>5</sup> Although notice how the teacher does not orient to the word as being unfamiliar since she does not explain it.

## 6 Conclusion

Recent studies on pronunciation instruction have provided overviews of the methods with which pronunciation is taught and addressed in the L2 classroom (e.g. Tergujeff 2012; Henderson et al. 2015; Foote et al. 2016). This article has approached the topic by adopting a qualitative, descriptive perspective to examining students' and teachers' classroom practices in non-elicited data, whereby it answers the call for more empirical-based research (e.g. Tergujeff 2012; Baker 2014; Derwing & Munro 2015; Foote et al. 2016) that helps "expand the knowledge base of L2 pronunciation instruction" (Baker & Murphy 2011: 31).

To that end, the study set out to show how CA with its focus on the temporal and sequential organization of naturally occurring classroom activities can enhance our understanding of how students take an active role in practicing English pronunciation through read-aloud. It argues that by analyzing what students do in and through classroom interaction, we can become more aware of how many of the actions teachers perform are, in fact, occasioned by students' actions. By concentrating on the interactional practices students employ to seek pronunciation help from the teacher, the findings evince that seemingly simple repair initiation techniques that indicate trouble in producing emerging speech are effective in mobilizing teachers' response in a similar manner as various word search techniques are effective in inviting recipients' help in other settings (Goodwin & Goodwin 1986; Oelschlaeger 1999; Brouwer 2003; Hayashi 2003; Radford 2009, *inter alia*).

The knowledge gained by this study is of value when providing pedagogical training for pre-service teachers or further education for in-service teachers in Finland, but also elsewhere. In Finnish, the pronunciation and reading of words is based on the close letter-to-sound correspondence (Suomi et al. 2008: 141). When learning English, students need to learn how to decode, and therefore to pronounce, words since there is no one-to-one correspondence between the letters and sounds of English. Learning to decode can be done via explicit instruction on phonetics, but also through recurrent practicing of reading aloud texts in English that renders a more meaningful context for training one's pronunciation skills. That is, it caters for a broader approach to learning pronunciation, beyond the segmental focus (e.g. Celce-Murcia et al. 2010; Tergujeff 2012).

The findings thus provide valuable insights on the kinds of pronunciation issues that teachers could address not only during the read-aloud activity, but also more explicitly after the activity. This is important, as students have brought these issues to teachers' attention instead of teachers choosing to address specific issues beforehand (Sicola & Darcy 2015: 475). Furthermore, the practicing of English pronunciation in this study was accomplished as a



by-product of another classroom activity. Students in EFL lessons in Finland are often requested to translate textbook chapters to Finnish in pairs and to read the texts aloud. During such activities, students can be instructed to tune in to one another's repair initiation techniques and identify when their partner is facing pronunciation trouble. Afterwards they can discuss the specific problems together and thus help one another. To that end, the study offers an insight on how teachers can design their classroom activities in ways that bring off multiple goals. Preferably this ought to be done in a manner that affords students different occasions to negotiate how particular words are pronounced, whereby they can benefit from the read-aloud activity the most. These kinds of 'multiple goal' activities can be used alongside more focused instructional activities on pronunciation and the phenomena therein (e.g. phonetics) to provide a more rounded approach to learn how to pronounce English words.

## Acknowledgements

I would like to thank Elina Tergujeff and the anonymous reviewers of the earlier versions of the manuscript for their critical comments that helped me improve the article. The study was funded by the Academy of Finland as part of the project *The construction of knowledge and competences in CLIL classroom interaction* (2014–2017, no. 274539).

## References

- Allington, R. L. 1980. Teacher interruption behaviors during primary-grade oral reading. *Journal of Educational Psychology*, 72 (3), 371–377. DOI: 10.1037/0022-0663.72.3.371.
- Baker, A. 2014. Exploring teachers' knowledge of second language pronunciation techniques: teacher cognitions, observed classroom practices, and student perceptions. *TESOL Quarterly*, 48 (1), 136–163. DOI: 10.1002/tesq.99.
- Baker, A. & J. Murphy 2011. Knowledge base of pronunciation teaching: staking out the territory. *TESL Canada Journal*, 28 (2), 29–50. DOI: 10.18806/tesl.v28i2.1071.
- Brouwer, C. E. 2003. Word searches in NNS–NS interaction: opportunities for language learning? *The Modern Language Journal*, 87 (4), 534–545. DOI: 10.1111/1540-4781.00206.
- 2004. Doing pronunciation: a specific type of repair sequence. In R. Gardner & J. Wagner (eds) *Second language conversations*. London: Continuum, 93–113.
- Celce-Murcia, M., D. Brinton & J. Goodwin 2010. *Teaching pronunciation: a course book and reference guide*. 2nd ed. Cambridge: Cambridge University Press.
- Derwing, T. M. & M. J. Munro 2015. *Pronunciation fundamentals: evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins.

- Foote, J. A., P. Trofimovich, L. Collins & F. Soler Urzúa 2016. Pronunciation teaching practices in communicative second language classes. *The Language Learning Journal*, 44 (2), 181–196. DOI: 10.1080/09571736.2013.784345.
- Goodwin, M. H. & C. Goodwin 1986. Gesture and coparticipation in the activity of searching for a word. *Semiotica*, 62, 51–76. DOI: 10.1515/semi.1986.62.1-2.51.
- Grabe, W. 2009. *Reading in a second language: moving from theory to practice*. Cambridge: Cambridge University Press.
- Hayashi, M. 2003. Language and the body as resources for collaborative action: a study of word searches in Japanese conversation. *Research on Language & Social Interaction*, 36 (2), 109–141. DOI: 10.1207/S15327973RLSI3602\_2.
- Henderson, A., L. Curnick, D. Frost, A. Kautzsch, A. Kirkova-Naskova, D. Levey, E. Tergujeff & E. Waniek-Klimczak 2015. English pronunciation teaching in Europe survey: factors inside and outside the classroom. In J. A. Mompean & J. Fouz-González (eds) *Investigating English pronunciation: current trends and directions*. Hampshire: Palgrave Macmillan, 260–291.
- Henderson, A., D. Frost, E. Tergujeff, A. Kautzsch, D. Murphy, A. Kirkova-Naskova, E. Waniek-Klimczak, D. Levey, U. Cunningham & L. Curnick 2012. The English pronunciation teaching in Europe survey: selected results. *Research in Language*, 10 (1), 5–27. DOI: 10.2478/v10015-011-0047-4.
- Hismanoglu, M. & S. Hismanoglu 2010. Language teachers' preferences of pronunciation teaching techniques: traditional or modern? *Procedia Social and Behavioral Sciences*, 2 (2), 983–989. DOI: 10.1016/j.sbspro.2010.03.138.
- Jakonen, T. 2016. Managing multiple normativities in classroom interaction: student responses to teacher reproaches for inappropriate language choice in a bilingual classroom. *Linguistics and Education*, 33, 14–27. DOI: 10.1016/j.linged.2015.11.003.
- Koda, K. 2004. *Insights into second language reading: a cross-linguistic approach*. Cambridge: Cambridge University Press.
- Koshik, I. & M. S. Seo 2012. Word (and other) search sequences initiated by language learners. *Text and Talk*, 32 (2), 167–189. DOI: 10.1515/text-2012-0009.
- Lyster, R. 1998. Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48 (2), 183–218. DOI: 10.1111/j.1467-1770.2001.tb00019.x.
- Murphy, J. & A. Baker 2015. History of ESL pronunciation teaching. In M. Reed & J. Levis (eds) *The handbook of English pronunciation*. Chichester: Wiley-Blackwell, 36–65.
- Oelschlaeger, M. L. 1999. Participation of a conversation partner in the word searches of a person with aphasia. *American Journal of Speech-Language Pathology*, 8 (1), 62–71. DOI: 10.1044/1058-0360.0801.62.
- Radford, J. 2009. Word searches: on the use of verbal and non-verbal resources during classroom talk. *Clinical Linguistics & Phonetics*, 23 (8), 598–610. DOI: 10.1080/02699200902997491.
- 2010. Adult participation in children's word searches: on the use of prompting, hinting, and supplying a model. *Clinical Linguistics & Phonetics*, 24 (2), 83–100. DOI: 10.3109/02699200903407149.
- Saito, K. & R. Lyster 2012. Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning*, 62 (2), 595–633. DOI: 10.1111/j.1467-9922.2011.00639.x.

- Schegloff, E. A. 1979. The relevance of repair to syntax-for-conversation. In T. Givón (ed.) *Syntax and semantics*. New York (N. Y.): Academic Press, 261–286.
- Schegloff, E. A., G. Jefferson & H. Sacks 1977. The preference for self-correction in the organization of repair in conversation. *Language*, 53, 361–382. DOI: 10.2307/413107.
- Sicola, L. & I. Darcy 2015. Integrating pronunciation into the language classroom. In M. Reed & J. Levis (eds) *The handbook of English pronunciation*. Chichester: Wiley-Blackwell, 471–487.
- Sidnell, J. & T. Stivers 2012. *The handbook of conversation analysis*. Malden (Mass.): Wiley-Blackwell.
- Streeck, J. 1995. On projection. In E. N. Goody (ed.) *Social intelligence and interaction: expressions and implications of the social bias in human intelligence*. Cambridge: Cambridge University Press, 87–110.
- Suomi, K., J. Toivanen & R. Ylitalo 2008. *Finnish sound structure: phonetics, phonology, phonotactics and prosody*. Oulu: University of Oulu. <http://jultika.oulu.fi/files/isbn9789514289842.pdf>.
- Szczepek Reed, B. 2015. Pronunciation and the analysis of discourse. In M. Reed & J. Levis (eds) *The handbook of English pronunciation*. Chichester: Wiley-Blackwell, 190–208.
- Tergujeff, E. 2012. English pronunciation teaching: four case studies from Finland. *Journal of Language Teaching and Research*, 3 (4), 599–607. DOI: 10.4304/jltr.3.4.599-607.
- 2013. *English pronunciation teaching in Finland*. Jyväskylä Studies in Humanities 207. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-5322-5>.
- Vygotsky, L. S. 1978. *Mind in society: the development of higher psychological processes*. Cambridge (Mass.): Harvard University Press.

## Appendix: Transcription conventions

Following CA methodology, the participants' talk has been transcribed according to the Jeffersonian transcription notations.

---

.	downward/stopping intonation
,	continuing intonation
?	interrogative intonation
¿	slightly interrogative intonation
↑	rising intonation
↓	falling intonation
wh <sub>a</sub> t	word emphasis
>wh <sub>a</sub> t<	quick speech
<wh <sub>a</sub> t>	slow speech
°wh <sub>a</sub> t°	quiet speech
(1.9)	silence (approximately)
(.)	micro pause
(-)	one tenth of a silence
(wh <sub>a</sub> t)	dubious hearing
(x)	unidentifiable item
ye-	a cut-off word
[wh <sub>a</sub> t]	overlapping speech
[wh <sub>a</sub> t]	
=	contiguous utterances or units of talk
£wh <sub>a</sub> t£	smiley voice
»wh <sub>a</sub> t«	piece of text read aloud

---

In addition to verbal annotations, participants' focal embodied actions have been transcribed in capital letters underneath the spoken representation.

Teacher's embodied actions are indicated by an asterisk (\*) and students' by the plus (+) sign (when they have been captured in the camera view).

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. *Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 82–97.*

**Irina Spita & Martina Björklund**

Åbo Akademi University

## **Teaching emotive language use in a beginners' course of Russian**

The Russian language abounds with affixes for indexing speakers' age, sex, social status, relations, and the attitude to the topic of conversation, but despite their frequent use, most textbooks and exercises neglect this crucial aspect. The present article reports the empirical experience of teaching an experimental course design of Russian for beginners. The textbook material was enhanced with in-depth work on emotive morphology in text analyses and creative speaking exercises, aiming at skills of interpreting and responding to speech situations with an emotional layer. The results are solid and show that this course design can promote communicative skills and creative handling of vocabulary without excessive lexical burden. By the end of the course, the students distinguished the use of diminutives in baby talk, companionship, emotional manipulation, flirting, politeness, and sales promotion, and could adjust their speech to the circumstances and purpose of the situation.

**Keywords:** Russian, emotivity, diminutive, pragmatic functional situation

## 1 Introduction

The question of what has variously been discussed in terms of affect, emotivity, emotions, and feelings has gained increased interest in several disciplines during the last few decades. In linguistics, for instance, affect / emotivity / expressivity<sup>1</sup> for a long time was treated as a kind of rarity on the margins of the linguistic system (Volek 1987: 1). Reports and results of recent research and theorizing on the emotional side of language can be found in, among others, the volumes edited by Foolen et al. (2012) and Lüdke (2015). As argued by Foolen (2012: 350–354) emotions can both be conceptualized in words that denote them (e.g. *love, anger, joy*) and expressed directly in emotive/expressive language, which occurs on all linguistic levels: prosody, morphology, interjections, emotion-laden lexemes, syntax, and constructions. Among the morphological means for direct expression of emotion, diminutives play an important part in Russian.

An affective turn has also reached the field of language acquisition and learning (Pavlenko 2013: 5–6), which can be witnessed in increased research on the role of affect in second language acquisition (SLA) in the last 10–15 years (e.g. Dewaele 2010; Dörnyei & Ushioda 2011). As Dewaele (2015: 13) notes, until the current century even the word *emotion* itself was not used in the literature on SLA, and the only emotion studied was foreign language anxiety. Pavlenko (2013: 5–6) discusses the affective turn in terms of three lines of inquiry, viz. linguistic, psychological, and social aspects of the language learning process, and proposes that they need to be merged to reach an understanding of the affective dimensions of language learning, “placing embodied subjects in their linguistic and social contexts”. This urge rhymes well with Foolen’s (2012: 349) contention that not only motion but also emotion is part of the embodied grounding of language, and that emotion thus is “one of the preconditions for the functioning of language”, which entails that second language (L2) teaching should allow emotional involvement (Foolen 2012: 364). Swain (2013) likewise maintains that emotions no longer should be separated from cognition in L2 learning – they “are at the heart of the foreign language process” as Dewaele (2015: 13) puts it.

The teaching of the experimental course design that is the topic of our article was conceived in the same vein as the ideas discussed above. The students should not only learn and acquire the cognitive content of the linguistic material presented in the textbook, but in addition, they were to get acquainted with some emotive morphological variants of words (diminutives) and typical

<sup>1</sup> These terms are sometimes used basically synonymously, different authors preferring one or the other term, while other authors provide different definitions. In the present article, we use the terms *emotive* and *emotivity*, except when referring to authors using another term.

situational contexts of use in various pragmatic functions. Moreover, emotivity should play a dual role: it should be both an object of study and an important component of the process of teaching. Introducing positive emotive linguistic material could contribute to the creation of a friendly atmosphere in the classroom and positive “vibes” towards the teacher, the fellow students, and, finally towards the Russian language and culture (on “language learning vibes”, see Gregersen 2013). Since foreign languages are mostly studied to be able to communicate in a cultural environment with a different set of emotional and emotive norms, students should be prepared for how differently their foreign interlocutors may perceive and evaluate events, statements, and behavior. Russians are often accused of lack of courtesy, as they use the words *thank you* or *please* rarely in comparison with British people, whereas an Englishman’s speech may seem cold and unemotional for a Russian, especially in communication with family and close friends (Ovčinnikov 2008: 347). In their study of the experience of loving a partner in a foreign language, Dewaele & Salomidou (2017: 124) report on a participant who linked her difficulties in expressing her emotional state to the absence of the diminutives of her native Croatian in English, the language she was communicating with her partner in, which Dewaele & Salomidou (2017: 126) liken with “the lack of required sugar in a cup of tea”. In Russian, diminutives are of paramount importance for the communication of a large spectrum of emotive values.

The objective of this article is to report the above-mentioned teaching experiment conducted by the first author. The results reported are exclusively based on participant observation by the teacher. In the next section, we briefly discuss diminutives and their formation, semantics, and pragmatic functions. We then move over to describing the course, its additional experimental design and teaching, and, finally, some results.

## 2 Word formation, diminutives, and emotivity

Along with paralinguistic and other linguistic means of expressing emotivity, Russian abounds with suffixes with emotive functions, among them diminutives, which poses problems to those whose native languages do not use diminutives to the same extent and in the same functions as Russian. Russian diminutives can be formed with the help of suffixes placed after the base stem but before the ending (if there is one), most productively from nouns and in various degrees from all other word classes (for Finnish diminutives, see Laalo 2007, for Swedish, Åkerblom 2013). Referring to Dokulil’s three onomasiological categories, Volek (1987: 45) maintains that the word-formation value of diminutives is generally one of modification, i.e. adding a diminutive suffix does not change the part of speech of the word, nor does it change its

TABLE 1. The characteristics of the word-formation values.

Compared words	Word-formation value	Part of speech same/different	Lexical meaning same/different
<i>holodn-yj</i> ADJ – <i>holodn-o</i> ADV	Transposition	Different	Same
<i>holod</i> N – <i>holod-ec</i> <sup>2</sup> N	Mutation	Same	Different
<i>holodn-yj</i> ADJ – <i>holodn-en'k-ij</i> ADJ-Dim	Modification	Same	Same + additional modifier

lexical meaning – it adds a modifier, usually including the senses of ‘smallness’ and ‘endearment’. Table 1 lists and compares Dokulil’s three types of word-formation values as described by Uluhanov (1996: 149) and illustrated by the following pairs of words: *holodnyj* (‘cold’) – *holodno* (‘coldly’); *holod* (‘coldness’) – *holodec* (‘jellied meat’); *holodnyj* (‘cold’) – *holodnen'kij* (‘cold and cute’).

As transposition and mutation concern general lexical and grammatical meanings of words, the use and meaning of the resulting lexemes is mostly not as context-dependent and “fluid” as in the case of the additional modifiers of diminutives, which makes diminutives more challenging for foreign language (FL) learners.

Diminutives have been the object of numerous linguistic studies, both universally (e.g. Jurafsky 1996) and in individual languages. Detailed studies of different aspects of Russian diminutives are presented in, e.g., Volek (1987), Steriopolo (2008), and Makarova (2014), whereas a short overview can be found in Protassova & Voeikova (2007). In Table 3 we list the suffixes that the teacher (author 1) deemed relevant for and possible to include in the beginners’ course design. In spite of the meaning of the term diminutive, diminutives, for the most part, do not designate only ‘smallness’, but smallness in combination with different positive or negative attitudes and pragmatic implications, even only loosely associated with smallness. There are suffixes mainly conveying either positive or negative attitude and ambivalent suffixes, which may express either positive or negative attitude. In the experimental course design for beginners, only usage in positive meanings was included to keep the topic manageable within the frames of the course. Support for this deci-

<sup>2</sup> Many Russian suffixes have different word-formation values in different words. The suffix *-ec* can also form diminutives, see Table 3.



sion can also be found in Volek's (1987: 79) calculations of the frequencies of positive and negative emotive nuances, according to which the positive ones prevail.

Diminutives are used in various pragmatic functions, apparently first minutely described by Volek (1987: 175). She starts from the specific ability of diminutives to express the emotive attitude of the speaker both toward the phenomenon designated by the base stem of the word in the diminutive and toward other phenomena, notably the addressee (Volek 1987: 149–150). Below we list only the different functional situations which are included in the course design. The attachment of the emotive attitude toward the referent of the base stem is largely predictable in the following cases: intensification of the emotive meaning, presence of the literal meaning of "smallness", children's discourse, and when the referent is a person (Volek 1987: 158). The emotive attitude is directed towards the addressee rather than the referent of the word in the diminutive in the following situations: command or prompting when the relation between speaker and addressee is close, the speaker wants to create an atmosphere of concern, or the speaker wants the addressee to have a positive attitude toward an offered object (Volek 1987: 161). There are also cases where the emotive attitude towards the referent of the word in the diminutive radiates from this object to something else mentioned in the utterance or to the addressee. This may happen in "[j]ocular powerplay with the addressee", speech directed to children, and intimate advice (Volek 1987: 169). Russian diminutives are thus conspicuously present in many other types of discourse than children's speech and child-directed speech. Although in adult speech, diminutives are "a non-obligatory speech refinement" (Protašova & Voeikova 2007: 50), they form such an essential part of Russian colloquial everyday communication that FL learners of Russian have much to gain from starting to get used to them right from the beginning.

### 3 The course: basic facts

The experimental course design was taught in the framework of a one-term (56 hours) Open University course of Russian 1 (level A1 according to the Common European Framework of Reference for Languages), with classes (90 minutes) twice a week in the evenings. The language of instruction was English. A group of 14 students enrolled<sup>3</sup>, and to complete and pass the course they were required to score no less than 70% in the final test and make a presentation based on the studied material. 9 students successfully finished the course.

<sup>3</sup> When enrolling the students were informed about the experimental nature of the course, and they gave their permission to utilize anonymized student data, in addition to oral and written responses, for scientific purposes.

TABLE 2. Background facts about the students.

---

<b>Age</b>	21 (n = 2), 23 (n = 2), 28 (n = 1), 29 (n = 1), 40 (n = 1), 42 (n = 1), 55 (n = 1) years
<b>Occupation</b>	student (n = 6), Ph. D. student (n = 1), post-doctoral researcher (n = 1), freelancer (n = 1)
<b>Native language</b>	Finnish (n = 3), Swedish (n = 5), Czech (n = 1)
<b>Other languages</b>	English (n = 9), German (n = 6), Finnish (n = 5), Swedish (n = 4), French (n = 4), Danish (n = 1), Norwegian (n = 1), Spanish (n = 1)
<b>Amount of Russian studies after the course</b>	1 academic term (n = 6), 1 term + private lessons and self-study (n = 2), 1 term + 3 years at school in childhood (38 years ago, n = 1)
<b>Reasons for taking the course</b>	studies (n = 5), travelling (n = 3), business (n = 2), culture (n = 2), friends and family (n = 2), personal interest in languages (n = 2), for fun (n = 2), reading the news (n = 1)
<b>Contact with the Russian language (use/hear/speak) outside the lessons</b>	the Internet (n = 4), TV & radio (n = 3), friends (n = 2), family (n = 1), travelling (n = 1), reading (n = 1), with a classmate when doing homework (n = 1); almost never (n = 1)

---

The drop-out was due to “external” reasons, such as moving to another city and lack of time because of the workload in other academic subjects. The final written test contained a compulsory part based on the textbook and another part concerning the experimental part of the course.

As the course was given by an Open University, which forms part of the adult education available in Finland regardless of age and educational background, the group of students was highly heterogeneous: the age of the students ranged from 20 to 55 years, the level of education from Bachelor to doctoral student up to post-doctoral researcher, with Swedish, Finnish, Czech and Romanian as their L1 (the Romanian speaking students, however, moved away towards the end of the course). The students’ second and additional languages were English, Swedish, Finnish, French, German, Danish, Spanish, and Norwegian. This variety of languages enabled interesting comparisons and discovery of both commonalities and differences between the various linguistic systems with their morphology and expression of emotions. Table 2 subsumes background facts about the students given in a questionnaire that was conducted in addition to the final test.

As can be seen from Table 2, Russian was the third or further additional language to the students as they all had knowledge of English. Most of them

also had knowledge of German and Finnish/Swedish besides their L1. Most of the students had no or little knowledge of Russian prior to the course. Studies and travelling were the most common reasons for taking the course. All students but one seem to have used the language also outside the lessons.

The classroom was spacious enough for work in groups or pairs and allowed rearranging the furniture for different kinds of activities. It was equipped with all the necessary devices of audio and visual aids, as well as an Internet connection, which enabled quick and easy access to information, pictures, illustrative examples, and videos to clarify the meaning of unknown words.

The course used the beginners' textbook *Kafe Piter 1* (*Кафе Питер 1*) by Marjatta Alestalo (2009; 5th edition, 2015), widely used in Finland, and also in the Open University where the course was taught. The positive aspects of the textbook are that it contains sufficiently detailed explanations of the grammatical material in Finnish, which made it possible to leave much of the theoretical part for self-study. There are relevant explanations of phonetics, phonetical exercises and online audio support. The author of the textbook tried to introduce some youth slang and make the dialogues less formal, which brings the book closer to real colloquial speech, but there are only a handful of diminutives in the whole book. Some of the examples do not seem relevant and up-to-date (for a critical analysis of the *Kafe Piter* series of textbooks, see Mironenko 2016). The weak points of the book are the lack of exercises for revising grammar and vocabulary, and the fact that all the recurring characters are young people of the same age. Exercises and dialogues had to be reworked to demonstrate various conversational tactics according to the level of politeness, familiarity and age disparity.

As *Kafe Piter 1* was originally designed for a full academic year course, to adapt it to the length of the course, the order of lessons and presentation of new grammar structures had to be reworked. This also prompted the need for additional materials. Firstly, the grammar material from the textbook was reproduced in bright PowerPoint presentations that contained further examples related to the relevant lexical topic of the lessons, their dialogues and speaking tasks. Second, drilling exercises to supplement the textbook were added. These exercises combined the new grammar and vocabulary with revision of knowledge from previous lessons.

Since the textbook vocabulary is restricted (about 1,000 words), the students were encouraged to enhance it to their discretion by compiling their own "My vocabulary", using ten designed sheets where they could record 15–20 new words on the topic studied in addition to the words used in the textbook. The aim of this task was to motivate the students to boost their vocabulary with the words they felt necessary in order to tailor the material to their own needs. The new words became the object of grammar drilling exercises, but more importantly for the topic of this article, the students discussed each

new page of “My vocabulary” in groups and could share their ideas of why they had chosen a certain word, which turned into good speech practice as well, see (1):

- (1) I found the word *gulât*’ which means ‘to walk’. I like walking, so *â lûblû gulât*’.

The additional materials also included videos with songs and cartoons, so that every new topic would be associated with positive emotions and in order to further interest. These materials promoted the connection between language and culture and helped master pronunciation. They were enjoyed by the students and gave rise to fruitful conversations enlivening the lessons (classroom observation by the teacher). Other speech enhancing activities included: Revising the studied material and mastering new skills by playing cards, board games, and role-playing games. The speed of “linguistic reaction” was exercised using a dice that was thrown to students one at a time with an oral prompt that required the student to produce a new word to be memorized or a certain grammatical form of a word.

All the extra materials were available on Moodle and open to print, watch and listen at any convenient time.

## 4 The experimental course design and its implementation

The innovation of the experimental course design lies in the immediate immersion of beginner students in a colloquial environment with characteristic emotive diminutives in typical situational contexts. This allows students in each individual case to “feel” by themselves the pragmatic meaning of these words in concrete situations, much as it happens in child language acquisition. Thus, the order of acquainting the students with diminutives starts from pragmatics, further working through semantics and morphological composition. This is the reverse order as compared to more traditional curricula, where, apart from occasional instances, students are more systematically acquainted with diminutives at a much later stage in courses of grammar and/or word formation. Such courses usually start and concentrate the explanation on the formation of diminutives, which leaves much less time for understanding and practicing their utilization. The two opposing types of approach are illustrated in Figure 1.

In-depth work with affixes in combination with speech thus began from the stage of meeting with the students at the first lesson, and the first pragmatic function encountered was talking to people of the same age and companionship.<sup>4</sup> The teacher (author 1) had a list of the students in advance and

<sup>4</sup> The course design paid attention to word formation in general, helping the students to see

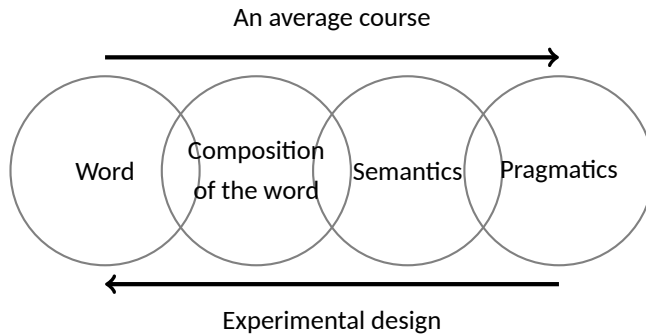


FIGURE 1. Teaching diminutives in the experimental design compared to the traditional approach.

could select a few appropriate hypocoristic nicknames derived from the students' names using Russian diminutive suffixes, e. g.: *Milka – Milkočka*, *Jonas – Jonasik*, *Dina – Dinočka*<sup>5</sup>. (In the textbook hypocoristic forms of Russian names are introduced only in lesson 3, and the first diminutive form of a name appears in lesson 8.) The students first familiarized themselves with the Cyrillic alphabet writing their names and choosing a nickname that they wanted to be used during the lessons in order to create a close friendly classroom atmosphere, where one of the basic affective functions of diminutives – informal communication could be felt, understood, and acted out. Based on classroom observation, this atmosphere and the Russian nicknames formed from their own names made the students feel comfortable and free to ask questions, participate in role games and dialogues, as well as to make and discuss mistakes. It was also the first step in recognizing the affixes in a word and starting to work with them.

Due to the level of the course, the number of diminutive suffixes introduced had to be quite limited. The productive suffixes taught are listed in Table 3.

Diminutive word forms were systematically introduced in additional dialogues on the topics of the textbook lessons (see examples 2–9 below), staging concrete examples of functional situations discussed in Volek (1987) (see section 2). Thus the situation of the speaker wanting the addressee to have a positive attitude toward an offered object was manifested in conversations involving food, where diminutives were met with and practiced in dialogues about ordering and serving food:

---

a new word not as a line of letters but as formed by meaningful components. In this article, however, we concentrate on the diminutives. Talking to strangers was also exercised, but is likewise beyond the scope of this article.

<sup>5</sup> These are not the real names of any of the students.

TABLE 3. Diminutive suffixes taught in the beginners' course.

N-Masc	-ik	nož-ik ('knife-Dim')
	-čik	žurnal'-čik ('magazine-Dim')
	-ec	brat-ec ('brother-Dim'; 'little brother'; 'pal')
N-Fem	-k-	spin-k-a ('back-Dim')
	-očk-	koft-očk-a ('cardigan-Dim')
	-ečk-	nož-ečk-a ('leg/foot-Dim')
	-onk-	ruč-onk-a ('arm/hand-Dim')
N-Neut	-k-	okoš-k-o ('window-Dim')
	-išk-	pal't-išk-o ('coat-Dim')
ADJ/ADV	-on'k-	tih-on'k-ij / tih-on'k-o ('quiet/quietly-Dim')
	-en'k-	horoš-en'k-ij / horoš-en'k-o ('good/well-Dim')

- (2) Jonasik, budeš' salat ili sup-čik? 'Jonasik, would you like some salad or soup-Dim'

Taking this topic a step further, the students were asked to use real Russian websites to study the menus of some popular pubs and cafeterias in Moscow and St. Petersburg. These menus often contain the diminutive forms of the names of dishes: *supčik* (*dnâ*) ('soup-Dim (of the day)'), *kotletka* ('patty-Dim'), *salatik* ('salad-Dim). Such examples taken from the Internet made the students aware of the sales promotion pragmatic function of diminutives used in menus to create an aura of homemade cooking around the dishes. The students further worked with online cafeteria websites using the menus in dialogues and role-playing. Through children's menus, the functions of child-directed discourse and children's speech were practised in role games of families with children.

The sales promotion function is also relevant to the topic of clothes and fashion, where the positive emotive attitude towards the referent of the word in the diminutive radiates from this object to the addressee:

- (3) Sejčas v mode takie sum-očk-i. 'This kind of bag-Dim-s are the craze now'

Here the additional meaning of cuteness reflects the speaker's positive attitude towards the object. As the positive attitude also extends to the addressee, diminutives of clothing and fashion vocabulary were used for making compliments and flirting. The students used online stores to find pictures of clothing and the relevant words to form diminutives from, e.g.:

- (4) O! kakaâ šap-očk-a! Tebe očen' idët! 'O, what a hat-Dim! It suits you very well!'

The emotive attitude is directed towards the addressee rather than the referent of the word in the diminutive in polite requests, which were practiced for different types of situations. In imperative clauses following clauses with the short adjective *nužen* ('need'), the request is softened by the diminutive, which even makes it possible to drop the word *požalujsta* ('please') without being rude, as shown in example (5):

- (5) Mne nužen karandaš. Daj mne *karandaš-ik!* 'I need a pencil. Give me a pencil-Dim'

Using the diminutive of proper names makes it possible to exert flattering emotional manipulation in requests, which makes it harder for the addressee to decline the request:

- (6) *Jonas-ik*, pomogi mne, požalujsta! 'Jonas-Dim, help me, please'

With words denoting an absolute value the diminutive can be used to decrease the importance of the object requested (by making it seem small), and therefore easier for the addressee to part with.

- (7) Odolži mne *million-čik!* 'Could you lend me a million-Dim'

Diminutives used to decrease the importance of an absolute entity, like a week, can also be used to temper a problem, as in (8):

- (8) Â uedu vsego na *nedel'-k-u*. 'I'm going away only for a week-Dim'

Irony and sarcasm, finally, were practised in situations where an obviously large object was named using a diminutive:

- (9) *I èto ty nazyvaeš' sum-očk-oj?! 'And you call this a handbag-Dim'* (Talking about a large handbag)

After plunging the students into a functional situation, where already familiar words occurred with diminutive suffixes, and helping them to sense the emotive attitude expressed, it was time to analyze the composition of the words, comparing words with similar suffixes, and to replace suffixes and subsequently observe the change in the meaning of the word. Bright visual aids were used to minimize the use of translation in explanations and to illustrate different senses of words. After discussing some suffix/es, students were asked to predict or guess the meaning of an unknown word with the same affix. Usually, the students could come up with a plausible meaning of a word based on their familiarity with the affixes and roots, or by analogy with familiar words. These exercises produced many "wow effects" and inspired the students to look for new words to recognize and memorize easily (classroom

observations by the teacher). Dealing with the words like with a Lego brick set significantly affected the interest in new words and their acquisition.

As the aim of the experimental course was not only to recognize and understand the emotive attitudes expressed by the diminutives, but also to internalize this mode of emotional expression<sup>6</sup>, the students got ample opportunity to practice the use of diminutives through a rich array of activities both in class and at home. These activities usually started with tasks to read and enact dialogues, and later to compose their own dialogues containing words with emotive elements. The students could imagine themselves of a different age, or occupation, communicating in different places, such as a classroom, a café, a business trip, etc. Playing such role games allowed the students to experience and implement conversational tactics in various circumstances and emotional contexts. The positive impact observed by the teacher was that the students felt more and more confident using Russian in different circumstances.

The students actively worked with online resources in order to bring the language practice beyond the classroom. With the help of smartphones and computers, they consulted online maps to find out a route around the city and give directions, practiced mock ordering of food and trips for the Christmas vacation. The students found using real Russian websites highly motivating since they could utilize knowledge from the pragmatic situational functions exercised and gain new interesting experience. They could observe the emotive meanings of diminutives in online stores and in dialogues between friends and family members on chats and social networks.

The most challenging part of the work with emotive suffixation was for the students to grasp the emotive attitude present in each new functional situation, for the simple reason that non-Slavic languages do not have and use morphological diminutives to the same extent as Russian. It was interesting to observe the students' first reaction to some diminutive forms, such as those derived from numerals ("Does it mean fewer?") or dishes in a menu ("Do they mean there will be a little portion of soup?"). Quite expectedly, the student whose L1 was Czech mastered the topic better and faster than the others since diminutives perform similar functions in Czech. A second difficulty concerned the fact that there are cases where a suffix added to a base stem can yield both a modification, i. e. a diminutive, and a mutation, i. e. a homonym with a different meaning cf.: *ruka* ('hand') – *ručka* ('hand-Dim/little hand') and homonyms that mean 'handle', 'pen', and 'ear of a cup', among others. This part of the course required close attention, especially when selecting examples.

<sup>6</sup> Cf. "L2 learners may face new modes of emotional expression, which they have to internalize or, at least, decode to avoid misunderstanding and miscommunication." (Pavlenko 2013: 10).



## 5 Final assessment and results

As already noted, the final assessment included a compulsory part and an optional part. The compulsory part consisted of: a personal oral presentation, 10 topics of “My vocabulary” filled during the course, and the final grammar test. The oral presentations were on the theme “For my trip to X, I need” and tested the use of numbers in combination with articles of personal belonging. On the whole, the students boldly spoke in front of the whole group, some of them practically without looking into their manuscripts, and some using PowerPoint presentations. The final test was based on the material studied in the textbook and reminded of the homework tasks: answering questions, completing grammar charts, a cloze test, and a short essay about themselves. The questionnaire (mentioned in section 3) and a morphology test were optional, but all the students were willing to complete them. The tasks of the morphology test were: to match proper names with their nickname forms; to form a few hypocoristic name forms from the student’s own name; to paraphrase the sentences of a conversation, adjusting it to talking with a six-year-old nephew, and, finally, to define the pragmatic situations and emotive attitudes of a few typical utterances containing diminutives.

The results of the final assessment were most positive. In addition to sufficiently good knowledge of Russian grammar and vocabulary, the students showed their ability to recognize and use diminutives in the functional situation types covered by the course design:

- (10) *Vot moj novyj telefončik.* ‘This is my new phone-Dim’ – positive attitude towards an object
- (11) *Kakie u tebâ botinočki!* ‘What shoes-Dim(=beautiful) you have’ – complimenting
- (12) *Kakie nožki!* ‘What legs-Dim’ – flirting
- (13) *Podoždi menâ odnu minutku!* ‘Wait for me a minute-Dim’ – strong request and emotional manipulation
- (14) *Počitaem knižečku.* ‘Let’s read a book-Dim’ – child-directed discourse
- (15) *Privet, Tanečka!* ‘Hello, Tania-Dim’ – talking to a person of the same age, chumminess, intimacy
- (16) *Odnú minutku!* ‘Just a minute-Dim’ – politeness
- (17) *Èto vsego na časik.* ‘It will only take a minute-Dim’ – emotional manipulation through flattery
- (18) *Supčik dnâ* ‘Soup-Dim of the day’ – sales promotion

However, recognizing and interpreting the diminutives was easier for the students than forming the preferred variant for a given situation, when, for instance, there were three masculine suffixes to choose between (see Table

3) which may express slightly different emotive nuances with different words and in different contexts. Base stems are also susceptible to diminutive formation in different degrees, so that some stems allow a whole range of suffixes, while others are more restricted. A further problem arises from the fact that not all diminutives of nouns of a certain semantic group express all the same pragmatic and emotive nuances. Thus, in the final test to name a desired Christmas present for a little boy, one creative student formed the diminutive *traktor-ec* in analogy with *mašin-k-a* ('car-Dim', 'toy car'). Although fully possible (but somewhat odd sounding to a Russian ear), the word *traktor-ec* would not be used about a toy tractor.

## 6 Conclusion

In this article, we have touched upon the emotive/affective turn in linguistics and second/foreign language learning and acquisition. After a brief sketch of the category of diminutives in Russian, their formation, semantics, and pragmatic functions, we turned to the main topic of the article: the experimental beginners' course of Russian, aiming at accustoming the students to the Russian mode of emotional expression through the use of diminutives, so characteristic of colloquial Russian. The experience of the experimental course design demonstrated that inclusion of emotive morphology into the course had a positive impact in several respects: first, from the very first lesson when the students chose their "Russian" nicknames, stress levels were very low in class; second, student motivation and interest did not decrease, but rather increased, as the course progressed; third, in-depth work with morphology helped to memorize and boosted vocabulary; and, finally, speaking skills remarkably developed on various levels of speech, and in particular on those levels that are absent from regular textbooks – in several contexts students could understand and produce native-like utterances, even jokes.

Thus, the results show that the inclusion of emotive language use into a beginners' course of Russian is something that students can cope with. The students of this particular course learnt both the textbook material and the additional part, and with great pleasure. The experimental course design thus managed to bridge the abyss between the barren style of a textbook and the living speech. This first experience definitely encourages continuation and future refinements of the course design.

## Abbreviations

ADJ	adjective
ADV	adverb
Dim	diminutive
Fem	feminine
FL	foreign language
L1	native language
L2	second language
Masc	masculine
N	noun
Neut	neuter
SLA	second language acquisition

## References

- Åkerblom, S. 2013. Om betydelsekategorin 'diminutiv' i svenskan. In *Nordlund 32: perspektiv på genus*, 102–127. <http://projekt.ht.lu.se/nordlund/nordlund-32/> [accessed 5/24/2017].
- Alestalo, M. 2015. *Kafe Piter 1*. Helsinki: Finn Lectura.
- Dewaele, J.-M. 2010. *Emotions in multiple languages*. Basingstoke: Palgrave Macmillan. DOI: 10.1057/9780230289505.
- 2015. On emotions in foreign language learning and use. *The Language Teacher*, 39 (3), 13–15. <http://jalt-publications.org/tlt/articles/4467-jalt2015-conference-article-emotions-foreign-language-learning-and-use>.
- Dewaele, J.-M. & L. Salomidou 2017. Loving a partner in a foreign language. *Journal of Pragmatics*, 108, 116–130. DOI: <https://doi.org/10.1016/j.pragma.2016.12.009>.
- Dörnyei, Z. & E. Ushioda 2011. *Teaching and researching motivation*. 2nd ed. Harlow: Longman. DOI: 10.4324/9781315833750.
- Foolen, A. 2012. The relevance of emotion for language and linguistics. In A. Foolen, U. M. Lüdke, T. P. Racine & J. Zlatev (eds) *Moving ourselves, moving others: motion and emotion in intersubjectivity, consciousness and language*. Amsterdam: John Benjamins, 349–368. DOI: 10.1075/ceb.6.
- Foolen, A., U. M. Lüdke, T. P. Racine & J. Zlatev (eds) 2012. *Moving ourselves, moving others: motion and emotion in intersubjectivity, consciousness and language*. Amsterdam: John Benjamins. DOI: 10.1075/ceb.6.
- Gregersen, T. 2013. Language learning vibes: what, why and how to capitalize for positive affect. In D. Gabrys-Baker & J. Bielska (eds) *The affective dimension in second language acquisition*. Bristol: Multilingual Matters, 89–98.
- Jurafsky, D. 1996. Universal tendencies in the semantics of the diminutive. *Language*, 72 (3), 533–578. DOI: 10.2307/416278.
- Laalo, K. 2007. Diminutives in Finnish child-directed and child speech. In I. Savickiene & W. Dressler (eds) *Acquisition of diminutives*. Amsterdam: John Benjamins, 263–278. DOI: 10.1075/lald.43.
- Lüdke, U. M. (ed.) 2015. *Emotion in language: theory – research – application*. Amsterdam: John Benjamins. DOI: 10.1075/ceb.10.

- Makarova, A. 2014. *Rethinking diminutives: a case of study of Russian verbs*. Tromsø: University of Tromsø, The Arctic University of Norway.
- Mironenko, O. 2016. *Discerning cultural depictions and mistakes in Kafe Piter, a Russian language textbook for Finnish students*. MA thesis. University of Turku.
- Ovčinnikov, V. 2008. *Sakura i dub*. Moskva: AST.
- Pavlenko, A. 2013. The affective turn in SLA: from 'affective factors' to 'language desire' and 'commodification of affect'. In D. Gabrys-Baker & J. Bielska (eds) *The affective dimension in second language acquisition*. Bristol: Multilingual Matters, 3–28.
- Protassova, E. & M. D. Voeikova 2007. Diminutives in russian at the early stages of acquisition. In I. Savickiene & W. Dressler (eds) *Acquisition of diminutives*. Amsterdam: John Benjamins, 43–72. DOI: 10.1075/lald.43.
- Steriopolo, O. 2008. *Form and function of expressive morphology: a case study of Russian*. Vancouver: The University of British Columbia. <http://hdl.handle.net/2429/424>.
- Swain, M. 2013. The inseparability of cognition and emotion in second language learning. *Language Teaching*, 46 (2), 195–207. DOI: 10.1017/S0261444811000486.
- Uluhanov, I. S. 1996. *Edinicy slovoobrazovatel'noj sistemy russkogo âzyka i ih lek-sičeskaâ realizaciâ*. Moskva: Russkie slovari.
- Volek, B. 1987. *Emotive signs in language and semantic functioning of derived nouns in Russian*. Linguistic & literary study in Eastern Europe 24. Amsterdam: John Benjamins. DOI: 10.1075/llsee.24.

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 98–117.

**Henna Heinonen**

Jyväskylän yliopisto

## **Begripligheten av finskspråkiga gymnasisters svenska uttal: lyssnarvärderingar**

The aim of this study is to investigate how comprehensible L1 Finnish speakers' L2 Swedish is considered to be and whether any listener- or speaker-specific background factors relating to the comprehensibility judgements can be found. The research material consists of Swedish read-aloud texts by 21 Finnish-speaking upper secondary school students. Judgements are given by 64 native Swedish-speakers (Finland-Swedish  $n = 35$ , Sweden-Swedish  $n = 29$ ) using a verbalized scale from 1 to 7. The study shows that L2 Swedish, pronounced by L1 Finnish speakers, is on average considered to be *comprehensible but the listener has to concentrate closely*. Both Finland-Swedish and Sweden-Swedish listeners gave similar ratings. Only two background factors correlated with the comprehensibility ratings: the speaker's latest school grade in Swedish had a negative correlation, whereas the speaker's usage of Swedish on his/her spare time correlated positively with the comprehensibility ratings.

**Keywords:** comprehensibility, pronunciation, L2 Swedish

**Nyckelord:** begriplighet, uttal, svenska som andraspråk

## 1 Inledning

Begriplighet har varit ett av de viktigaste målen för uttalsinläringen redan i årtionden (se t.ex. Thorén 2008; GLGY 2003; Abercrombie 1949 citerat i Munro & Derwing 1995a). Målet är relevant. Ett begripligt uttal är nödvändigt för lyckad kommunikation: det har större betydelse för lyssnarperceptionen än t.ex. grammatisk korrekthet (t.ex. Kjellin 2002; Fayer & Krasinski 1987). Det som gör uttalets roll avgörande för begripligheten är prosodiska drag såsom satsbetoning, rytm och melodi. De hjälper lyssnaren att uppfatta gränserna mellan orden och fraserna och upptäcka den viktigaste informationen (t.ex. Saito m.fl. 2016; Field 2005; Kjellin 2002; Munro & Derwing 1995a; Bannert 1984: 211). Uttalet och prosodin har dock förbisetts i den finländska språkundervisningen (Tergujeff 2013; Iivonen & Tellä 2009), som ändå syftar till ett lättbegripligt uttal (GLGY 2003, 2015). Denna diskrepans gör finskspråkiga talaras uttal och dess begriplighet till ett relevant undersökningsobjekt.

Syftet med den här undersökningen är att kartlägga begripligheten av finskspråkiga gymnasisters svenska uttal. Målet är att utreda begriplighetsnivån genom skalbedömning och undersöka hur olika bakgrundsfaktorer hos talare och lyssnare påverkar begriplighetsbedömningarna. Undersökningen besvarar följande forskningsfrågor: 1) Hur begripligt är finskspråkiga gymnasisters svenska uttal? 2) Avviker begriplighetsvärderingarna givna av finlandssvenska lyssnare och sverigesvenska lyssnare från varandra? 3) Finns det några faktorer i lyssnarnas och talarnas bakgrund såsom ålder, kön eller språklig bakgrund, som relaterar till begriplighetsvärderingen?

Enligt läroplanerna har begripligheten stått som mål för muntlig färdighet relativt länge (Opetushallitus 1994). Trots det har finskspråkiga gymnasisters svenska uttal och begriplighet studerats bara i en tidigare undersökning. Uttalet beskrevs likna finlandssvenska och det tänktes inte vålla några problem för begripligheten, som i sin tur ansågs förknippas främst med lexikala faktorer (Hildén 2000). Uttalet och begripligheten värderades av L2-lyssnare, medan L1-lyssnarens uppfattning förblev outforskad i studien (Hildén 2000: 245–256).

Talarna i den här undersökningen har studerat svenska i flera år och sålunda antas ett begripligt uttal inte vålla stora problem för dem (jfr Hildén 2000: 246). Samtidigt har språkundervisningen varit inriktad på skriftliga färdigheter redan länge, och därför kan talarnas färdigheter i uttal vara sämre än deras språkfärdighet i övrigt. Den andra forskningsfrågan gäller variantens inverkan på begripligheten. Svenskstudierna i gymnasium syftar till interaktion i Norden (GLGY 2015), vilket innebär kommunikation med både finlandssvenska och sverigesvenska talare. Då lyssnarens variant kan antingen likna fonetiskt talarens L1, såsom finlandssvenska gör i finskspråkiga talaras fall (Kuronen & Leinonen 2010), eller ha ett tydligt annorlunda uttal, såsom sverigesvenska (se t.ex. Garlén 1988), kan begripligheten bero på lyssnarens variant. På grund av

fonetiska likheter mellan finlandssvenska och finska samt finlandssvenska lyssnarens vana vid finskspråkiga talares svenska förväntas uttalet vara lättare för finlandssvenska än för sverigesvenska lyssnare att begripa (se kap. 2). Hypotesen gällande den tredje forskningsfrågan baserar sig på liknande grund: jag antar att kunskaper i finska och mångsidiga språkfärdigheter gör det lättare att förstå brutet uttal. Jag antar vidare att begriplighetsvärderingarna följer talarens framgång i svenskstudierna (se kap. 2).

Kontexten av undersökningen inramas av finländsk gymnasieutbildning. Gymnasister är en relevant talargrupp eftersom ett muntligt delprov införs i studentexamen inom den närmaste framtiden. Grunden för gymnasieutbildning utgörs av en oftast nioårig läroplikt som det finländska skolsystemet baserar sig på (Opetushallitus 2017a,b). Svenska språket är ett av de obligatoriska läroämnena och det kan börja studeras tidigast i årskurs 2. Då kallas lärokursen för en lång lärokurs A1. Vanligast är dock att eleven läser en medellång lärokurs B1 i svenska (Ylioppilastutkintolautakunta 2016), vilket ända till hösten 2016 betydde att studierna i svenska började i årskurs 7. Fr.o.m. hösten 2016 trädde en ny läroplan i kraft och sedan dess har studierna i svenska börjat i årskurs 6 som 12-åring. Den medellånga lärokursen i gymnasiet omfattar fem obligatoriska kurser och två valfria kurser (GLGY 2003, 2015).

## 2 Förståelighet och begriplighet: tal, talare och lyssnare

Vid forskning som syftar till att analysera talarens förmåga att producera tal eller lyssnarens förmåga att förstå det har det oftast gjorts en skillnad mellan begreppen förståelighet (eng. *intelligibility*), begriplighet (eng. *comprehensibility*) och brytning (eng. *accentedness*) (t.ex. Saito m.fl. 2016; Field 2005; Munro & Derwing 1995a). Begreppen är etablerade i uttalsforskning, speciellt med engelska som målspråk, men fastställda motsvarigheter i svenska saknas (se dock Sturesson 2016: 8). Orsaken till separata begrepp är olika analysmetoder. Företeelsen har nämligen undersökts inom lingvistik med olika metoder och uppgifter. Om lyssnarens uppgift är att identifiera vad talaren säger (t.ex. genom transkribering) används begreppet förståelighet (eng. *intelligibility*) (Munro & Derwing 1995a). Fokuset ligger då på talaren och förmågan att producera tal. Testet är då objektiva (Fayer & Krasinski 1987). Alternativt kan undersökningen fokusera på lyssnarens uppfattning om talets begriplighet. Lyssnaren värderar lättheten att förstå talet genom skalbedömning och i så fall talas det om begriplighet (eng. *comprehensibility*) (Munro & Derwing 1995a). Begriplighetstest är lyssnarbaserat och subjektivt (Fayer & Krasinski 1987). Med begreppet brytning (eng. *accentedness*) syftas på lyssnarens uppfattning om hur mycket uttalet avviker från målspråket, vilket analyseras ge-

nom skalbedömning (Munro & Derwing 1995a).

Fast tudelning i begriplighet och förståelighet är mycket vanlig inom förståelighetsundersökningar finns det dock studier som inte gör någon skillnad mellan begreppen och också studier som har annorlunda definitioner (t.ex. Smith 1992). Därför är det på plats att definiera hur begreppen används i denna artikel. Begreppet förståelighet används om den objektiva förståeligheten (*intelligibility*), dvs. när metoden har varit identifiering eller transkription. Härutöver används detta begrepp då förståelighetsforskningen diskuteras på den allmänna nivån och då det inte har framgått i den hänvisade studien om den undersöker förståelighet eller begriplighet i den betydelse som beskrivits ovan. Begreppet begriplighet används då metoden har varit skalbedömning (*comprehensibility*) och begreppet brytning hänvisar till uttalsmässiga avvikelser från målspråket (*accentedness*). Metodologiskt sett undersöker den här studien begriplighet; lyssnarna har värderat begripligheten på en skala.

Förståeligheten har stått i fokus i flera studier och konstaterats relatera till flera delområden inom muntlig kommunikation, men starkast till uttalet (t.ex. Saito m.fl. 2016; Crowther m.fl. 2015; Field 2005; Kjellin 2002; Munro & Derwing 1995a,b; Fayer & Krasinski 1987; Albrechtsen m.fl. 1980). Att fastställa en generaliserbar hierarki över de mest kritiska uttalsfelen har varit utmanande eller till och med omöjligt för de undersökta språken och forskningsmetoderna har varit varierande (se Jesney 2004: 2; Pihko 1994). Därtill kan de påverkande faktorerna variera enligt talarens L1 (Crowther m.fl. 2015). Field (2005: 399) föreslår att prosodiska drag har universellt störst inverkan på förståeligheten. Ett förklarande exempel på detta gäller ordbetoningen: då man processar talet fäster man störst uppmärksamhet på den betonade stavelsen som förknippas med ett visst ord och med den viktigaste informationen i satsen. Således kan ett felaktigt betoningsmönster fördröja processen att uppfatta budskapet och vålla problem i förståeligheten (t.ex. Field 2005; Kjellin 2002; Munro & Derwing 1995b; Bannert 1984). Prosodins betydelse för förståeligheten har också märkts av Lintunen (2014: 167) och Derwing & Munro (1997): Felfriheten i prosodi (och också i grammatik) överväger segmental korrekthet. Också i svenska har prosodin en central roll. För nästan 40 år sedan listade Bannert (1980: 97) de prosodiska dragen som viktigaste inlärningsmål i svenskt uttal – och detsamma gäller fortfarande (Thorén 2008: 31; Bannert 2004: 64, 66). Drag som rytm, ord- och satsbetoning samt kvantitet är drag som har tänkts vara de mest kritiska för förståeligheten, övervägande segment (se Thorén 2008: 121; Kjellin 2002: 96). I stället för förståeligheten förknippas segmentfelen vanligen med brytning (Saito m.fl. 2016; Thorén 2008), som i sin tur också påverkar begripligheten. Ett starkt brutet uttal kan få lägre begriplighetsvärderingar (t.ex. Abelin & Boyd 2000), men lyssnaren kan dock transkribera det talaren sagt helt korrekt (Munro & Derwing 1995a). Med andra ord kan talaren förmedla budskapet, men det är svårt att uppfatta för lyssna-



ren. Den låga begriplighetsvärderingen kan bero på processtiden hos lyssnaren vilken ökar om talet har stark brytning (Munro m.fl. 2006: 112). Å andra sidan finns det studier där korrelationen mellan begriplighet och brytning inte påträffats (t.ex. Munro & Derwing 1995a). Brytningen kan också relatera till lyssnarens attityd, och den eventuella negativa attityden mot en talare med stark brytning eller begränsade språkkunskaper kan försämra begriplighetsvärderingen – lyssnaren vill inte ens förstå (Kjellin 2002: 15; Fayer & Krasinski 1987).

Förutom uttal har andra lingvistiska och också icke-lingvistiska faktorer en inverkan på förståeligheten. Grammatik och lexikon, speciellt lexikal mångfald, relaterar till värderingen (Kennedy & Trofimovich 2008; Albrechtsen m.fl. 1980), men relationen är dock inte så stark som till uttalet (Saito m.fl. 2016; Fayer & Krasinski 1987). Därtill kan talhastighet (Munro & Derwing 2001: t.ex.), tonhöjd (Abelin & Boyd 2000), tvekan (Fayer & Krasinski 1987), självkorrigering (Albrechtsen m.fl. 1980), lyssnarens tidigare kunskap om temat (se Kennedy & Trofimovich 2008) och kännedom av genren och kulturen (Pihko 1994: 14) ha inverkan på förståeligheten. Vidare kan irritationen vara en orsak till – eller en följd av – svårforståeligheten: Lyssnaren kan bli irriterad till följd av konstanta fel och därför inte förstå, eller så kan lyssnaren bli irriterad då hen inte förstår (Fayer & Krasinski 1987).

Vid sidan av talet i sig relaterar förståeligheten till vissa talar- och lyssnarmässiga bakgrundsfaktorer. Således är det väsentligt att bestämma vem som talar och vem som ska förstå talet då man undersöker förståeligheten. I största delen av tidigare förståelighetsundersökningar har målspråket varit engelska (för en översikt se Jesney 2004), som kan konstateras vara *lingua franca* – ett språk som man kommunicerar på med både L1- och L2-talare. Svenska studeras däremot för att man ska kunna kommunicera i Norden (GLGY 2003, 2015) och i första hand använda språket med och bli förstådd av L1-talare. Lyssnarperspektivet i förståeligheten, begripligheten och brytningen samt deras värdering innebär lyssnarens tidigare erfarenheter av talares L1 (Kennedy & Trofimovich 2008; Wingstedt & Schulman 1984) och av fonetik och uttal. L1-lyssnare med tidigare kunskaper om talarens L1 kan ha lättare att förstå talare med stark brytning eller låga kunskaper i målspråket (Kennedy & Trofimovich 2008; Wingstedt & Schulman 1984) men å andra sidan är lyssnarvärderare som har samma L1 som talare mera kritiska i sina värderingar (Fayer & Krasinski 1987). Strängare värderingar kan också ges av en oerfaren eller fonetiskt utbildad lyssnare (Jesney 2004: 8; Munro & Derwing 1995b) om än inte alltid (t.ex. Saito m.fl. 2017; Kennedy & Trofimovich 2008).

Uttalsfärdigheterna är således av central betydelse i muntlig kommunikation. Uttals- och språkinläringen relaterar till ett flertal faktorer i talarbakgrunder, såsom exponeringsåldern och -tiden, det sociala umgänget, talares L1 osv. (se Ullakonoja & Dufva 2016). Talarens erfarenheter av och kontakter

med målspråket är betydelsefulla. Till exempel ett längre uppehåll i en målspråkig miljö stödjer uttalsinläringen (t.ex. Ullakonoja 2011). Om uppehållet är svårt att arrangera eller om kontakterna med målspråket är svårtillgängliga har fritiden och skolan en stor betydelse för uttalsinläringen (se Derwing & Munro 2013). Speciellt målspråkiga fritidsaktiviteters betydelse för språkinläringen har påvisats vara markant (t.ex. Huhta m.fl. 2016; Mustila 1990), men intresset för att använda svenska språket på fritiden är inte stort (Hildén & Rautopuro 2014: 118–119). Framgången i svenskstudier verkar vara könsfördelad: flickor är mer motiverade och når bättre resultat i alla språkliga delområden än pojkar (t.ex. Hildén & Rautopuro 2014: 77; Juurakko-Paavola & Palviainen 2011; Tuokko 2002). Det är dock viktigt att minnas att de ovannämnda resultaten grundar sig på vitsordet, självbedömningen eller lärarbedömningen, inte på autentiska situationer eller L1-lyssnarbedömningar.

### 3 Material, metoder och genomförande

Materialet i den här undersökningen utgörs av upplästa talprov av finskspråkiga talare som har blivit begriplighetsvärderade av svenskspråkiga lyssnare. **Talarna** i undersökningen är 21 finskspråkiga gymnasister i åldrarna mellan 17 och 19 år. 16 av talarna är kvinnor och 5 är män, vilket kan sägas spegla könsfördelningen i kurserna i svenska. Alla talare har läst den medellånga lärokursen i svenska, dvs. börjat sina studier i svenska vid 13 års ålder. Talarna går i ett vanligt gymnasium i Finland. Vid insamlingstiden gick talarna på en valfritt muntlig kurs (RUB6) som genomförs i slutskedet av studierna i svenska. Talarnas senaste vitsord i svenska varierade mellan 6 och 9 på en skala 4–10 så att fem talare hade vitsord 6, sju talare vitsord 7, sju talare vitsord 8 och en talare vitsord 9. En av talarna uppgav inte sitt vitsord. Ungefär hälften av talarna sade sig använda svenska på fritiden, dvs. de skriver, läser, talar och/eller lyssnar på svenska utanför skolan: Två talare använder svenska mer än en gång på två veckor och nio talare 1–2 gånger på två månader. Tio talare använder svenska inte alls på fritiden. Ingen av talarna sade sig ha tillbringat någon längre tid (mer än 2 månader) i något annat land än Finland.

I materialinsamlingen deltog sammanlagt 60 gymnasister. De som sade sig ha lässvårigheter e.d. bortvaldes och 21 finskspråkiga talare valdes med i undersökningen genom ett slumpmässigt urval. För att kunna försäkra lyssnarvärderingarnas reliabilitet användes fyra kontrolltalare som har svenska som hemspråk eller modersmål. Om en lyssnarvärderare inte hade gjort någon skillnad mellan målgruppen och kontrolltalarna tänktes värderingen inte ha gjorts med omsorg och värderingen uteslöts.

Som **talmaterial** i undersökningen användes en uppläst text (bilaga 1). Textens lexikala innehåll var avsett att vara okomplicerat och det innehåller inga

tunga konstruktioner. Textens struktur och innehåll var tänkta att likna den nivå som en gymnasist sannolikt skulle använda i verklig kommunikation. Texten har utformats för undersökningsändamål och den innehåller alla svenska ljud och varierande stavelsetyper samt ord- och satsbetoningsmönster. Genom valet av textuppläsning som undersökningsmaterial uteslöts faktorer som inte har med uttalet att göra, såsom grammatik eller lexikon. Jämfört med textuppläsning är spontant tal dock vanligare i förståelighets- och begriplighetsundersökningar (t.ex. Saito m.fl. 2016; Derwing & Munro 1997; Munro & Derwing 1995a; Fayer & Krasinski 1987), men att samla in material med spontant tal som skulle vara uttalsmässigt lika omfattande som en uppläst text kunde leda till långa talprov och ändå kan man inte garantera att alla aspekter kommer med i undersökningen. Man har inte heller kunnat se att textuppläsning skulle ge annorlunda resultat än spontant tal (Jesney 2004: 5). Det är dock viktigt att minnas att också textuppläsning kan innebära faktorer som påverkar värderingen men inte har någonting med uttalet att göra, t.ex. omtagningar (se Albrechtsen m.fl. 1980) och tvekande (Fayer & Krasinski 1987).

Inspelningarna utfördes med fyra olika gymnasistgrupper varav tre (41 gymnasister) inspelades i ett språklaboratorium med utrustningen Sanako Lab 100 och en grupp (19 gymnasister) med Roland -inspelningsapparater. Kontrolltalarna var med i de inspelade grupperna förutom en kontrolltalare som inspelades separat i en tyst lokal med en digital Zoom H2 -inspelningsapparat. Talarna fick ca 3 minuter på sig att bekanta sig med texten och ställa eventuella frågor om den före inläsningen. Om det blev felsägningar eller om talarerna kände sig missnöjda med uttalet, fick hen läsa om hela meningen, och den bästa varianten valdes med i undersökningen. Att inläsa texten tog från 80 till 130 sekunder. Talproven i lyssnartestet gjordes genom att indela den upplästa texten i sju delar så att längden på talproven blev från 22 till 31 ord (ca 10–20 sekunder). Varje talares talprov valdes slumpmässigt.

**Lyssnarna** i undersökningen är sammanlagt 64 finlandssvenska och sverigesvenska deltagare. Den finlandssvenska lyssnargruppen (senare i texten FISV-lyssnare) utgörs av 35 lyssnare (30 kvinnor, 5 män). FISV-lyssnarnas ålder fördelar sig relativt jämnt men största delen var 36–65-åriga ( $n = 24$ ). Över hälften av FISV-lyssnarna har utbildat sig vid universitet ( $n = 21$ ). Av FISV-lyssnarna använde 31 endast svenska hemma som barn och 4 lyssnare använde sig svenska och finska som barn. Förutom svenska har största delen kunskaper i engelska ( $n = 33$ ), tyska ( $n = 29$ ) och finska ( $n = 25$ ). Den språkliga mångfalden, dvs. hur många språk lyssnarna behärskar förutom modersmålet, är relativt varierande men för det mesta har FISV-lyssnarna kunskaper i två språk förutom svenska. Dialekt talar 14 lyssnare medan 21 sade sig inte tala någon dialekt. Lyssnarnas kontakter med finskan utreddes med en skala från 0 (= inga) till 5 (= mycket) och nästan hälften (48,6 %) av FISV-lyssnarna har mycket kontakter (5) med finska.

Den sverigesvenska lyssnargruppen (senare i texten SVSV-lyssnarna) består av 29 personer (18 kvinnor, 9 män, 2 ville inte uppge sitt kön). Jämfört med FISV-lyssnargruppen är SVSV-lyssnarna lite yngre. Största delen är 25-35-åringar ( $n = 13$ ), annars är åldersfördelningen jämn. Såsom FISV-lyssnarna uppgav nästan alla ( $n = 26$ ) SVSV-lyssnare ha använt endast svenska hemma som barn, men tre lyssnare hade också haft något annat hemspråk. Alla SVSV-lyssnare har universitetsutbildning. Alla SVSV-lyssnare har kunskaper i engelska ( $n = 29$ ) och största delen också i tyska ( $n = 19$ ) och franska ( $n = 15$ ). Den språkliga mångfalden liknar FISV-lyssnarna: största delen av SVSV-lyssnarna har kunskaper i två språk förutom svenska. SVSV-lyssnarnas kontakter med finska språket är mindre än FISV-lyssnarnas: 79 % av SVSV-lyssnarna har inga eller nästan inga kontakter med finskan (0-2 på en skala från 0 (= inga) till 5 (= mycket)). 14 lyssnare sade sig tala någon svensk dialekt, 15 lyssnare uppgav att de inte talar någon dialekt.

Lyssnarna önskades representera potentiella åhörare för gymnasisterna, dvs. vanliga svenskspråkiga personer. De rekryterades med ett sätt som liknar snöbollsurval: begriplighetsvärderingen utfördes på nätet och länken till lyssnartestet skickades till personer som tänktes ha möjligheten att sprida ut länken till eventuella lyssnare. I lyssnartestet deltog sammanlagt 101 lyssnare, varav 61 använder finlandssvensk och 40 sverigesvensk variant. Från lyssnargruppen bortvaldes de som angav sig ha använt något annat språk än svenska hemma som barn eller ha något annat primärt vardagsspråk än svenska. Bara lyssnare som sade sig ha normal hörsel togs med i undersökningen. Några av lyssnarna visade sig inte göra någon skillnad mellan målgruppen och kontrolltalarna och deras svar ansågs opålitliga och bortvaldes. Ett lyssnartest genomfördes separat med en större grupp ( $n = 12$ ) i en sal med högtalare, men denna grupp uteslöts ur undersökningen på grund av avvikande omständigheter. Efter urvalsprocessen blev det slutliga lyssnarantalet i undersökningen det ovannämnda 64 (35 FISV och 29 SVSV).

Talproven laddades in i ett nätbaserat program SoGo Survey där de presenterades i ett firsidigt formulär. Varje sida hade 6 eller 7 talprov med olika textinnehåll, inklusive kontrolltalarnas talprov. Således upprepades talprov med samma textinnehåll 3-4 gånger under lyssnartestet, vilket kunde göra värderingen något onaturlig och därför blev lyssnarna informerade om upprepandet inför testet. Talproven presenterades i en slumpmässig ordning på var sin sida. I anslutning till lyssnartestet svarade lyssnarna på bakgrundsfrågor och lyssnartestet tog sammantaget ca 20 minuter. Alla lyssnare ombads att använda hörlurar och göra testet i en tyst lokal. Alla lyssnare värderade begripligheten av alla 25 talprov (varav 4 var kontrolltalare) i testet så att undersökningen omfattar sammanlagt 1600 enskilda värderingar (1344 målgruppstalare + 256 kontrolltalare).

Begriplighetsvärderingen genomfördes med en verbaliserad skala på 1-7.

Detta passar bra i undersökningar med fonetiskt utbildade lyssnare och är lättare att använda för lyssnare och forskare än en niogradig skala, som dock oftast används i uttalsundersökningar (Isaacs & Thomson 2013; Jesney 2004: 2). För att underlätta både värderingen och analysen var skalan verbaliserad:

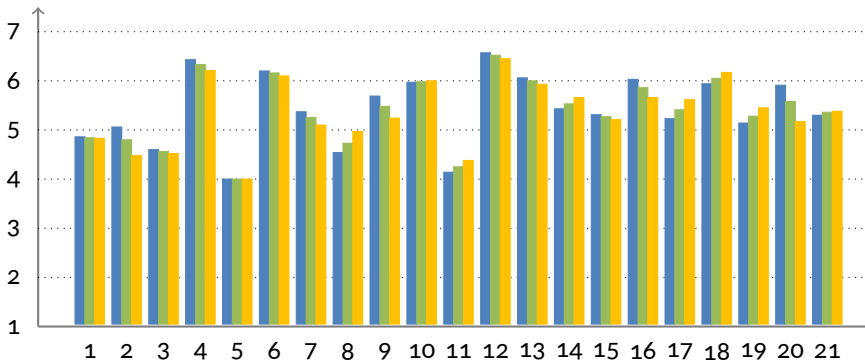
- 1 – förstod ingenting
- 2 – förstod ett par ord
- 3 – förstod några saker men helheten blev oklar
- 4 – förstod helheten men några saker blev oklara
- 5 – förstod allt men behövde koncentrera mig noga
- 6 – förstod allt
- 7 – förstod allt med lätthet
- \_ – vet ej

Forskningsfrågorna besvarades genom statistisk undersökning av lyssnarsvaren. Medelvärde (M), medianen (md) och typtalet (Mo) räknades med alla värderingar sammanlagt. På så sätt utreddes de talarmässiga värderingarna och fördelningen av dem. Den andra och den tredje forskningsfrågan analyserades genom att jämföra värderingarna som lyssnare med olika bakgrundsfaktorer gett och talare med olika bakgrundsfaktorer fått. Vid sidan av att gruppera talarna enligt de bakgrundsfaktorer (kön, vitsord, kunskaper i enskilda språk, fritidsaktiviteter osv.) som testades och att jämföra medelvärden mellan grupperna, analyserades det också om relationen mellan värderingarna och den ifrågavarande bakgrundsfaktorn är statistiskt signifikant. Relationsanalyserna genomfördes med t-test och korrelationstest (Pearson correlation). Ytterligare analyserades relationen mellan talarmässiga bakgrundsfaktorer och begriplighetsvärderingar statistiskt med Friedman test.

Resultaten är statistiskt signifikanta om värdet är mindre än  $p = 0,05$  och korrelationen mycket stark om värdet är nära  $r = 1,0$ . Förkortningen t syftar till t-testvariabeln och förkortningen F till storleken på testvärdet. Frihetsgraden har markerats med förkortningen df.

## 4 Resultat

I det här kapitlet presenteras resultaten gällande begriplighetsvärderingar samt deras relation till lyssnarnas och talarnas bakgrunder. Hur begripligt gymnasisters uttal har ansetts vara av alla lyssnare sammanlagt presenteras i avsnitt 4.1. Begriplighetsvärderingarna behandlas enligt lyssnarnas variant i avsnitt 4.2. I avsnitt 4.3 tas upp begriplighetsvärderingarnas relation till andra



FIGUR 1. Talarmässiga medelvärden av FISV- och SVSV-lyssnarnas samt alla lyssnarnas begriplighetsvärderingar på skala 1–7. Värderingarna givna av FISV-lyssnarna har markerats till vänster med blå färg, av SVSV-lyssnarna till höger med gul färg och alla lyssnares genomsnitt har markerats i mitten med grön färg.

bakgrundsfaktorer gällande såväl lyssnarna som talarna och i avsnitt 4.4 behandlas upprepadet av samma textinnehåll och bakgrundsbruset.

#### 4.1 Begripligheten

Syftet med den första forskningsfrågan var att analysera hur begripligt finskspråkiga gymnasister uttalar svenska. Resultaten (båda lyssnargrupperna sammanlagt) visar att uttalet är förståeligt men lyssnaren behöver koncentrera sig noga. Medelvärdet för värderingarna är 5,39 ( $md = 5$ ,  $Mo = 5$ ) på skalan 1–7.

Medelvärden som talarna fick varierar mellan 4,00–6,52 ( $md = 4$ –7,  $Mo = 4$ –7). Värderingen som största delen av talarna ( $n = 10$ ) fick var 5 – förstod allt men behövde koncentrera mig noga ( $M 5,28$ – $5,98$ ). Sex talare värderades med 4 – förstod helheten, men några saker blev oklara ( $M 4,00$ – $4,84$ ) och fem talare nådde värderingen 6 – förstod allt ( $M 6,00$ – $6,52$ ) (Figur 1). Resultaten skiljer sig delvis från Hildéns (2000: 246) antagande om att uttalet knappast vållar förståelighetsproblem: begripligt uttal var ingen självklarhet i alla talprov, men största delen blev dock någorlunda förstådda.

#### 4.2 Variantens påverkan

Syftet med den andra forskningsfrågan var att analysera om det finns någon skillnad mellan begriplighetsvärderingarna givna av finlandssvenska respektive sverigesvenska lyssnare. FISV-lyssnarna värderade talprovets begriplighet med medelvärdet 5,42 ( $md = 5$ ,  $Mo = 5$ ). De talarmässiga medelvärdena varierar mellan 4,00–6,57 ( $md = 4$ –7,  $Mo = 4$ –7). Talare som fick värderingen 5 – förstod allt men behövde koncentrera mig noga ( $M = 5,06$ – $5,97$ ) utgjorde den största gruppen ( $n = 11$ ). Värderingen 4 – förstod helheten, men några saker

blev oklara (4,00–4,86) fick 5 talare. Värderingen 6 – förstod allt (6,03–6,57) eller högre fick 5 talare (se Figur 1).

Medelvärde för begriplighetsvärderingarna som SVSV-lyssnarna gav var bara en aning lägre än FISV-lyssnarnas: 5,36 (md = 5, Mo = 5). Medelvärdena varierade mellan 4,00–6,45 (md = 4–7, Mo 4–7). I likhet med FISV-lyssnarna gav SVSV-lyssnarna värderingen 5 – förstod allt men behövde koncentrera mig noga (M 5,10–5,93) åt största delen av talarna (n = 10). Talare med värderingen 4 – förstod helheten, men några saker blev oklara (M 4,00–4,97) blev 6, en fler jämfört med FISV-talarna. Värderingen 6 – förstod allt (M 6,00–6,45) eller högre fick 5 talare. (Figur 1.)

Undersökningsresultaten visar alltså att finskspråkiga gymnasisters uttal av svenska är genomsnittligt lika begripligt för en FISV- och SVSV-lyssnare. Mellan medelvärdena finns det praktiskt taget ingen skillnad: 5,42 (FISV-lyssnarna) respektive 5,36 (SVSV-lyssnarna) utgör en skillnad på 0,06. Begriplighetsvärderingens relation till varianten är inte statistiskt signifikant ( $t = -0,467$ ,  $df = 62$ ,  $p = 0,642$ ), och mellan värderingarna givna av finlandssvenska lyssnare och sverigesvenska lyssnare finns en stark korrelation: talarmässigt är korrelationen  $r = 0,901$  och genomsnittligt  $r = 0,948$ , vilket visar att värderingarna är mycket eniga.

Båda lyssnargrupperna gav den högsta värderingen åt samma talare (talare 12) (M FISV 6,57/ SVSV 6,45) och den lägsta värderingen åt samma talare (talare 5) (M FISV 4,00/ SVSV 4,00). Annars varierade ordningen mellan talarnas begriplighetsvärderingar i viss mån. Vissa talare värderades vara begripligare för lyssnare av ena varianten än den andra. Två talare (talare 2 och 20) värderades statistiskt signifikant lättare att förstå av FISV-lyssnarna än av SVSV-lyssnarna ( $t = -2,956$ ,  $df = 62$ ,  $p = 0,004$  och  $t = -2,934$ ,  $df = 62$ ,  $p = 0,005$ ).

### 4.3 Andra bakgrundsfaktorer påverkan

Målet med den tredje forskningsfrågan var att utreda om det finns övriga lyssnar- eller talarmässiga bakgrundsfaktorer (se även kap. 2) som relaterar till värderingarna. Eftersom värderingarna av FISV-lyssnarna och SVSV-lyssnarna visade sig vara likriktade kunde lyssnargruppernas begriplighetsvärderingar förenas och analysen av bakgrundsfaktorernas påverkan göras med sammanlagda lyssnargrupper. Vid analysen av talarmässiga bakgrundsfaktorer (avsnitt 4.3.2) är det viktigt att minnas att samplet var litet, i synnerhet då talargruppen har indelats i mindre grupper enligt kön, vitsord och fritidsaktiviteter. Resultaten är således inte generaliserbara men de kan dock tänkas vara riktgivande.

#### 4.3.1 Lyssnarmässiga bakgrundsfaktorer

I analysen hittades anmärkningsvärt få lyssnarmässiga bakgrundsfaktorer som relaterade till begriplighetsvärderingen. Värderingarna relaterade inte till lyssnarens ålder ( $r = 0,201$ ,  $p = 0,110$ ) eller kön ( $t = 0,586$ ,  $df = 60$ ,  $p = 0,560$ ), bara utbildningens relation till värderingen var statistiskt signifikant ( $r = -0,251$ ,  $p = 0,046$ ). Enligt resultaten gav lyssnare med universitetsutbildning strängare värderingar, men det är väsentligt att beakta att representanter av andra utbildningar var få.

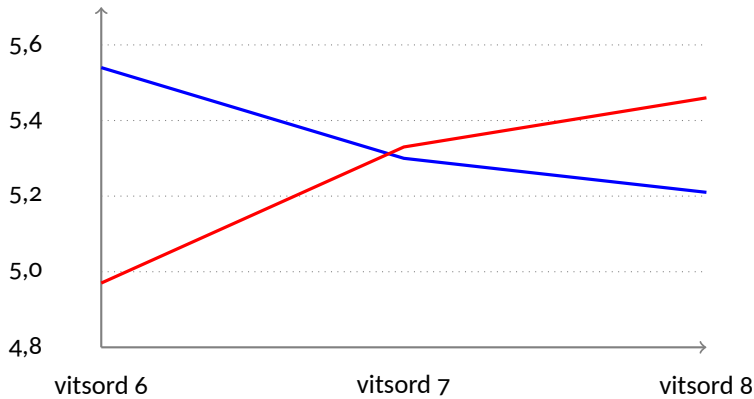
Ett oväntat resultat var att lyssnarens kunskaper i finska inte hade något samband med värderingen ( $t = -0,377$ ,  $df = 62$ ,  $p = 0,707$ ). FISV-lyssnarna med kunskaper i finska verkade värdera begripligheten till och med lite lägre än andra FISV-lyssnare (M kan finska 5,34/kan inte finska 5,61), medan med SVSV-lyssnarna var det tvärtom (M kan finska 5,48/kan inte finska 5,38). Lyssnarnas kontakter med finska ( $r = 0,036$ ,  $p = 0,779$ ) och SVSV-lyssnarnas kontakter med finlandssvenska ( $r = -0,003$ ,  $p = 0,986$ ) hade inte heller något samband med begriplighetsvärderingen. Härutöver relaterade mångsidiga språkkunskaper inte till begripligheten ( $r = 0,150$ ,  $p = 0,236$ ), inte heller kunskaper i enstaka språk ( $t = 1,300$ ,  $df = 62$ ,  $p = 0,198$  –  $t = -0,279$ ,  $df = 62$ ,  $p = 0,781$ ). Kunskaper i tyska tycks visserligen ha hjälpt svsv-lyssnarna att förstå talarnas svenska (M kan tyska 5,44/kan inte tyska 5,20), men resultatet var inte statistiskt signifikant. Om lyssnaren talar dialekt eller inte relaterade inte heller till begriplighetsvärderingen ( $t = 0,348$ ,  $df = 62$ ,  $p = 0,729$ ).

#### 4.3.2 Talarmässiga bakgrundsfaktorer

Såsom lyssnarens kön relaterade inte heller talares kön till begriplighetsvärderingen ( $t = -0,467$ ,  $df = 62$ ,  $p = 0,642$ ). Medelvärde 5,39 som kvinnliga talare fick är ytterst nära värdet 5,40 som manliga talare fick. Resultatet avviker från tidigare iakttagelser om könsfördelningen gällande framgången i svenskstudier (se kap. 2).

En intressant korrelation hittades mellan talarens senaste vitsord i svenska och begriplighetsvärderingen: korrelationen är negativ, dvs. talarna med ett lägre vitsord fick en högre begriplighetsvärdering och vice versa (Figur 2). Den enda talare som hade vitsord 9 fick dock den högsta begriplighetsvärderingen. Resultaten är statistiskt signifikanta ( $F = 23,543$ ,  $df = 2$ ,  $p = 0,000$ ) förutom värderingarna mellan vitsord 7 och 8 ( $p = 0,279$ ). Detta stämmer med Hildéns (2000) iakttagelser – studerande med vitsord 8 gör mera fel i svenskt uttal och talar mindre flytande. Talare med berömliga vitsord (9–10) hade dock bäst uttal både i Hildéns och i denna undersökning. Med vitsordet i andra språk verkade dock finnas en motsatt tendens. Talarna vars medelvärde av vitsord som de senast fått i andra språk är 6 ( $n = 4$ ) fick medelvärdet 4,97, vitsord 7 ( $n = 5$ ) 5,30, vitsord 8 ( $n = 7$ ) 5,45 och vitsord 9 ( $n = 5$ ) medelvärdet 5,69. Resultaten





FIGUR 2. Begriplighetsvärderingarnas relation till det vitsord som talaren senast fått. Med blå färg vitsord i svenska, med röd färg vitsord i andra språk.

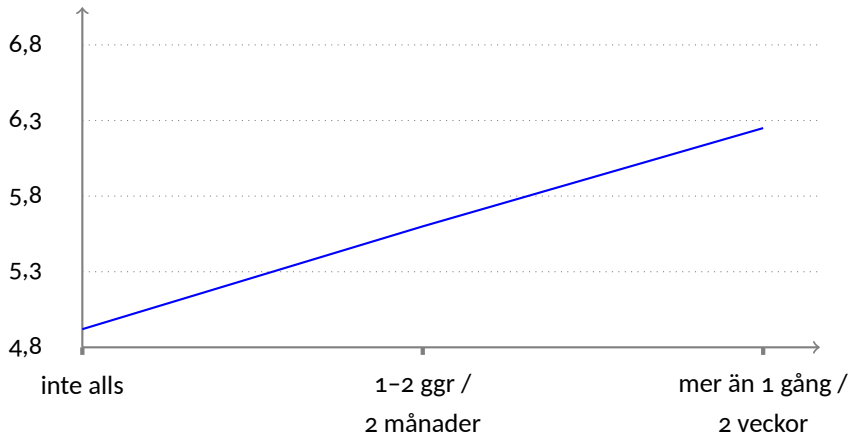
är statistiskt signifikanta ( $F = 51,687$ ,  $df = 3$ ,  $p = 0,000$ ) förutom mellan vitsord 7–8 ( $p = 0,485$ ) och 8–9 ( $p = 0,131$ ) (Figur 2). Detta kan tyda på att vitsordet i svenska inte mäter färdigheter i uttal.

Svenskspråkiga fritidsaktiviteter verkade förbättra begripligheten: korrelationen med begriplighetsvärderingen var positiv (jfr kap 2). Om talare använder svenska på fritiden hade hen värderats vara begripligare (M 5,71) än den som inte gör det (M 5,04). Till och med få fritidsaktiviteter med svenska såg ut att förbättra begripligheten. De talare som använde svenska mer än en gång per vecka fick medelvärden 6,25, 1–2 gånger på 2 månader medelvärdet 5,59 och de som inte alls sysslade med svenska språket medelvärdet fick 5,04 (Figur 3). Resultatet är statistiskt signifikant ( $F = 114,361$ ,  $df = 2$ ,  $p = 0,000$ ).

#### 4.4 Övriga kommentarer

Som väntat fick kontrolltalarna de högsta begriplighetsvärderingarna (M 6,66–7,00). Ändå fanns det några skillnader mellan dessa värderingar. Kontrolltalare som hade fått den högsta värderingen av FISV-lyssnarna använde den finlandssvenska varianten (M FISV 7/SVSV 6,86), medan SVSV-lyssnarna gav kontrolltalaren med den sverigesvenska varianten den högsta värderingen (M SVSV 6,90/ FISV 6,89).

Oavsett påpekandet i anvisningarna misstänkte några lyssnare i slutkommentarerna att upprepandet av samma textinnehåll hade påverkat begriplighetsvärderingen. Påverkan kunde dock inte konstateras vara kraftig. I Figur 1 visas talproven med värderingarna i den ordning som de blev uppspelade (dock slumpmässigt på sidorna). Värderingarna av de sista talproven avviker inte stort från de första talproven. Därtill listades begriplighetsvärderingarna av talproven med samma textinnehåll i spelordningen och det kunde märkas



FIGUR 3. Begriplighetsvärderingarnas relation till talarnas fritidsaktiviteter på svenska.

att värderingarna av de senare spelade talproven inte var konsekvent högre. Några lyssnare kommenterade att bakgrundsbruset också kan ha påverkat begripligheten. Påverkan kunde inte heller konstateras ha varit stor. Talproven med något bakgrundsbrus (1, 6, 8, 9, 10, 11, 13, 14, 15, 17, 21) värderades inte lägre än andra (se Figur 1).

## 5 Sammanfattande diskussion

Den här undersökningen har visat att finskspråkiga talares svenska uttal är begripligt men lyssnaren behöver koncentrera sig noga. Begriplighetsvärderingarna är inte beroende av lyssnarens variant, kön, ålder eller bakgrundsfaktorer som berör språket. De enda två bakgrundsfaktorer som relaterar till talprovets begriplighetsvärdering gäller talarna. Vitsord som talaren senast fått i svenska korrelerar negativt med begriplighetsvärderingen medan talarens fritidsaktiviteter på svenska har en positiv korrelation med begripligheten.

I motsats till vad som förväntades är lyssnarvärderingarna givna av den finlandssvenska och den sverigesvenska lyssnargruppen samstämmiga. Eftersom finska och finlandssvenska liknar varandra fonetiskt såväl på segmentell som på prosodisk nivå (Kuronen & Leinonen 2010) antogs FISV-lyssnarna ge högre värderingar än SVSV-lyssnarna. FISV-lyssnarnas värderingar är bara en aning högre än SVSV-lyssnarnas. Inte heller relaterar lyssnarnas erfarenheter av talarnas L1, dvs. finska, till begripligheten. FISV-lyssnare med kunskaper i finska verkar till och med ge lägre värderingar än lyssnare med inga kunskaper i finska. Detta kan antas anknyta till toleransen (jfr Munro m.fl. 2006; Fayer & Krasinski 1987), vilket uppmärksammas i värderingar av fyra utslutna FISV-

lyssnare som uppgav att de i stället för svenska använde främst finska i vardagen: de värderade begripligheten strängare (M 5,14) än de som använde mest svenska i vardagen (M 5,42). Om talaren däremot har stark brytning eller låga kunskaper i L2-språk förstås hen bättre av lyssnare med erfarenheter av talarens L1 (Kennedy & Trofimovich 2008; Wingstedt & Schulman 1984), vilket syns svagt i värderingarna av SVSV-lyssnarna: lyssnare med kunskaper i finska värderade begripligheten en aning högre än de som saknade kunskaper i finska.

De faktorer som ser ut att relatera till begriplighetsvärderingen gäller talarnas språkliga bakgrund, närmare sagt vitsord och intresse för svenska språket. Ett bra vitsord i svenska verkar inte garantera ett begripligt uttal – snarare tvärtom. Begripligheten värderades allra lägst hos de talare vars senaste vitsord i svenska var 8, medan hos talare med vitsord 6 värderades begripligheten vara allra högst (bortsett från den enda talaren med vitsord 9). Hildén (2000) antar att företeelsen sammanhänger med artikulatorisk medvetenhet: medvetenheten om fonologin kan försämra det fonologiska flytet (de Bot 1992; Krashen 1982 i Hildén 2000: 279–280). Den fonologiska medvetenheten kan verkligen skapa viss osäkerhet hos avancerade inlärare, men inlärare som fått uttalsundervisning betraktar sitt uttal vara bra och de är också villiga att lära sig mera om uttalet (Lintunen & Mäkilähde 2018). Antagandet att gymnasisterna lider av fonologisk medvetenhet strider delvis mot tidigare undersökningar enligt vilka muntlig kommunikation (Takala 1993) och speciellt uttal (Lintunen & Mäkilähde 2018; Tergujeff 2013; Iivonen & Tella 2009) förbises i den finska språkundervisningen, och således skulle inte heller talarna i den här studien vara fonologiskt medvetna. En tänkbar faktor är att språkinläringen är subtraktiv. Då språkundervisningen i gymnasiet koncentrerar sig mest på skriftliga kunskaper får de gymnasister som absorberar dem ett bättre vitsord. Med andra ord fungerar vitsordet som mätare för skriftliga kunskaper. Om språkundervisningen betonar kraftfullt skriftliga kunskaper och muntliga kunskaper inte övas eller används blir det obalans mellan de språkliga delområdena. Detta kan leda till att de skriftliga kunskaperna börjar dominera kognitivt vilket kan resultera i försämrat uttal. Baserat på de ovannämnda undersökningarna kan det antas att orsaken till det högre vitsordets negativa korrelation med begriplighetsvärderingen snarare ligger i obalansen mellan de språkliga delområdena i språkundervisningen än i den fonologiska medvetenheten. En undersökning som ger mer detaljerad kunskap om företeelsen behövs.

I stället för vitsordet visar sig begripligheten korrelera positivt med användning av svenska utanför studierna. De talare som uppgav att de använder svenska på fritiden fick bättre resultat än de som inte alls använde svenska utanför studierna. Till och med ett litet antal aktiviteter på svenska på fritiden tycks ha en inverkan. Motsvarande resultat har fått också tidigare: studerande som intresserar sig för målspråket utanför skolan har bättre kunskaper (Huhta

m.fl. 2016; Hildén & Rautopuro 2014; Mustila 1990). Eftersom studenter sällan styrs självständigt in i informella inlärningssituationer (Hildén & Rautopuro 2014) är det viktigt att svensklärarna uppmuntrar sina studenter att använda svenska också på fritiden.

Ett resultat som avviker från den stereotypa uppfattningen gäller talarens kön. Fastän pojkarnas framgång i svenskstudier har bedömts vara lägre än flickornas (se t.ex. Hildén & Rautopuro 2014: 77; Juurakko-Paavola & Palviainen 2011; Sajavaara 2006; Tuokko 2002) och pojkar ansetts vara omotiverade och mindre framgångsrika inlärare av svenska (se t.ex. Sajavaara 2006), visar resultatet av denna undersökning att manliga talare uttalar svenska lika begripligt som kvinnliga talare, till och med en aning begripligare.

## Tackord

Den här undersökningen har gett mångsidig kunskap om finskspråkiga talares uttal av svenska, och den erbjuder flera forskningsfrågor till fortsatta undersökningar. För att kunna genomföra undersökningen fick jag ett arbetsstipendium av Svenska kulturfonden vilket jag är tacksam för. Värdefullt stöd har jag fått av projektet FOKUS vid Jyväskylä universitet, vilket jag också vill rikta ett stort tack till.

## Referenser

- Abelin, Å. & S. Boyd 2000. Voice quality, foreign accent and attitudes to speakers. I A. Botinis & N. Torstensson (red.) *Proceedings of Fonetik 2000*. Skövde: Högskolan i Skövde, 21–24.
- Abercrombie, D. 1949. Teaching pronunciation. *ELT Journal*, 3 (5), 113–122.
- Albrechtsen, D., B. Henriksen & C. Faerch 1980. Native speaker reactions to learners' spoken interlanguage. *Language Learning*, 30 (2), 365–396.
- Bannert, R. 1980. *Svårigheter med svenskt uttal: inventering och prioritering*. Praktisk lingvistik 5. Lund: Institutionen för lingvistik, Lunds universitet.
- 1984. Problems in learning Swedish pronunciation and in understanding foreign accent. *Folia Linguistica*, 18 (1–2), 193–222.
- 2004. *På väg mot svenskt uttal*. Lund: Studentlitteratur.
- Crowther, D., P. Trofimovich, K. Saito & T. Isaacs 2015. Second language comprehensibility revisited: investigating the effects of learner background. *TESOL Quarterly*, 49 (4), 814–837. DOI: 10.1002/tesq.203.
- de Bot, K. 1992. A bilingual production model: Levelt's "speaking" model adapted. *Applied Linguistics*, 13 (1), 1–24.

- Derwing, T. M. & M. J. Munro 1997. Accent, intelligibility and comprehensibility: evidence from four L1s. *Studies in Second Language Acquisition*, 20 (1), 1–16. <https://www.cambridge.org/core/journals/studies-in-second-language-acquisition/article/div-classtileaccent-intelligibility-and-comprehensibilitydiv/729C15F62F9EC9A51A33EAB5C2D05ED0>.
- 2013. The development of L2 oral language skills in two L1 groups: a 7-year study. *Language Learning*, 63 (2), 163–185. DOI: 10.1111/lang.12000.
- Fayer, J. & E. Krasinski 1987. Native and nonnative judgments of intelligibility and irritation. *Language Learning*, 37 (3), 313–326.
- Field, J. 2005. Intelligibility and the listener: the role of lexical stress. *TESOL Quarterly*, 39 (3), 399–423. DOI: 10.2307/3588487.
- Garlén, C. 1988. *Svenskans fonologi*. Lund: Studentlitteratur.
- GLGY 2003. Grunderna för gymnasiets läroplan 2003. Helsingfors: Utbildningsstyrelsen. [http://www.oph.fi/lagar\\_och\\_anvisningar/laroplans-\\_och\\_examensgrunder/gymnasiet](http://www.oph.fi/lagar_och_anvisningar/laroplans-_och_examensgrunder/gymnasiet) [hämtad 2016-11-29].
- 2015. Grunderna för gymnasiets läroplan 2015. Helsingfors: Utbildningsstyrelsen. [http://www.oph.fi/lagar\\_och\\_anvisningar/laroplans-\\_och\\_examensgrunder/gymnasiet](http://www.oph.fi/lagar_och_anvisningar/laroplans-_och_examensgrunder/gymnasiet) [hämtad 2016-11-29].
- Hildén, R. 2000. *Att tala bra, bättre och bäst: suomenkielisten abiturienttien ruotsin kielen suullinen taito testisuoritusten valossa*. Helsinki: Helsingin yliopisto.
- Hildén, R. & J. Rautopuro 2014. *Ruotsin kielen A-oppimäärän oppimistulokset perusopetuksen päättövaiheessa 2013*. Helsinki: Opetushallitus.
- Huhta, A., J. Alderson, L. Nieminen & R. Ullakonoja 2016. The role of background factors in the diagnosis of SFL reading ability. I D. Tsagari & J. Banerjee (red.) *Contemporary Second Language Assessment*. Contemporary applied linguistics 4. London: Bloomsbury, 125–146.
- Iivonen, A. & S. Tella 2009. Vieraan kielen ääntämisen ja kuulemisen opetus ja harjoittelu. I O. Aaltonen, R. Aulanko, A. Iivonen, A. Klippi & M. Vainio (red.) *Puhuva ihminen: puhetieteiden perusteet*. Helsinki: Otava, 269–281.
- Isaacs, T. & R. Thomson 2013. Rater experience, rating scale length, and judgments of L2 pronunciation. *Language Assessment Quarterly*, 10 (2), 135–159. DOI: 10.1080/15434303.2013.769545.
- Jesney, K. 2004. *The use of global foreign accent rating in studies of L2 acquisition*. Calgary: Language Research Centre, University of Calgary. <http://arts.ucalgary.ca/lrc/research/reports> [hämtad 2017-5-30].
- Juurakko-Paavola, T. & Å. Palviainen 2011. *Svenskan i den finska skolan och högre utbildningen: om kunskaper och motivation genom olika utbildningsstadier*. HAM-Kin julkaisu 11/2011. Hämeenlinna: Hämeen ammattikorkeakoulu. <http://urn.fi/URN:NBN:fi:juu-201502241365>.
- Kennedy, S. & P. Trofimovich 2008. Intelligibility, comprehensibility, and accentedness of L2 speech: the role of listener experience and semantic context. *The Canadian Modern Language Review*, 64 (3), 459–489. DOI: 10.3138/cmlr.64.3.459.
- Kjellin, O. 2002. *Uttalet, språket och hjärnan: teori och metodik för språkundervisningen*. Uppsala: Hallgren & Fallgren.
- Krashen, S. 1982. *Principles and practice in second language acquisition*. Oxford: Pergamon Press. [http://www.sdkrashen.com/content/books/principles\\_and\\_practice.pdf](http://www.sdkrashen.com/content/books/principles_and_practice.pdf).

- Kuronen, M. & K. Leinonen 2010. *Svenskt uttal för finskspråkiga: teori och övningar i finlandssvenskt och rikssvenskt uttal*. Tampere: Juvenes Print.
- Lintunen, P. 2014. Ääntämisen oppiminen ja opettaminen. I P. Pietilä & P. Lintunen (red.) *Kuinka kieltä opitaan: opas vieraan kielen opettajalle ja opiskelijalle*. Helsinki: Gaudeamus, 165–187.
- Lintunen, P. & A. Mäkilähde 2018. Short- and long-term effects of pronunciation teaching: EFL learners' views. I J. Volin & R. Skarnitzl (red.) *Pronunciation of English by speakers of other languages*. Newcastle upon Tyne: Cambridge Scholars Publishing, 46–72.
- Munro, M. J. & T. M. Derwing 1995a. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45 (1), 73–97.
- 1995b. Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, 38 (3), 289–306.
- 2001. Modeling perceptions of the accentedness and comprehensibility of L2 speech: the role of speaking rate. *Studies in Second Language Acquisition*, 23 (4), 451–468.
- Munro, M. J., T. M. Derwing & S. Morton 2006. Mutual intelligibility of L2-speech: language learning. *Studies in Second Language Acquisition*, 28 (1), 111–131. DOI: 10.1017/S0272263106060049.
- Mustila, E. 1990. *Finska gymnasieelevers motivation för, attityder till och intresse för svenska*. Tutkimuksia 125. Helsinki: Helsingin yliopiston kasvatustieteen laitos.
- Opetushallitus 1994. *Lukion opetussuunnitelman perusteet*. Helsinki: Opetushallitus.
- 2017a. Gymnasiet. [http://www.oph.fi/utbildning\\_och\\_examina/gymnasiet](http://www.oph.fi/utbildning_och_examina/gymnasiet) [hämtad 2017-1-9].
- 2017b. Läroplikt och skola. [http://www.oph.fi/utbildning\\_och\\_examen/grundlaggande\\_utbildning/laroplikt\\_och\\_skola](http://www.oph.fi/utbildning_och_examen/grundlaggande_utbildning/laroplikt_och_skola) [hämtad 2017-1-9].
- Pihko, M. 1994. *Cross-linguistic intelligibility of native and non-native L2 speech varieties*. Jyväskylä: Korkeakoulujen kielikeskus, Jyväskylän yliopisto.
- Saito, K., P. Trofimovich & T. Isaacs 2016. Second language speech production: investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, 37 (2), 217–240. DOI: 10.1017/S0142716414000502.
- 2017. Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: a validation and generalization study. *Applied Linguistics*, 38 (4), 1–25. DOI: 10.1093/applin/amv047.
- Sajavaara, K. 2006. Kielivalinnat ja kielten opiskelu. I R. Alanen, H. Dufva & K. Mäntylä (red.) *Kielen päällä: näkökulmia kieleen ja kielenkäyttöön*. Jyväskylä: Soveltavan kielentutkimuksen keskus, Jyväskylän yliopisto, 223–254.
- Smith, L. E. 1992. Spread of English and issues of intelligibility. I B. B. Kachru (red.) *The other tongue: English across cultures*. 2. uppl. Urbana (Ill.): University of Illinois Press, 75–90.
- Sturesson, L.-O. 2016. *Bruten svenska: brytning och förståelighet i svenskt uttal*. Kalmar Växjö: Linne universitet. <http://urn.kb.se/resolve?urn=urn%3Anbn%3Ase%3Alnu%3Adiva-53464>.

- Takala, S. (red.) 1993. *Suullinen kielitaito ja sen arviointi*. Kasvatustieteiden tutkimuslaitoksen julkaisusarja B: Teoriaa ja käytäntöä 77. Jyväskylä: Kasvatustieteiden tutkimuslaitos.
- Tergujeff, E. 2013. *English pronunciation teaching in Finland*. Jyväskylä Studies in Humanities 207. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-5322-5>.
- Thorén, B. 2008. *The priority of temporal aspects in L2-Swedish prosody: studies in perception and production*. Stockholm: University of Stockholm. <http://urn.kb.se/resolve?urn=urn%3Anbn%3Ase%3Asu%3Adiva-7448>.
- Tuokko, E. 2002. *Perusopetuksen päättövaiheen ruotsin kielen oppimistulosten kansallinen arviointi 2001*. Helsinki: Opetushallitus.
- Ullakonoja, R. 2011. *Da. Eto vopros! Prosodic development of Finnish students' read-aloud Russian during study in Russia*. Jyväskylä Studies in Humanities 151. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-4209-0>.
- Ullakonoja, R. & H. Dufva 2016. Toisen ja vieraan kielen ääntämisen oppimisen haasteet. *NMI-Bulletin*, 26 (2), 4–18.
- Wingstedt, M. & R. Schulman 1984. Comprehension of foreign accents. I W. Dressler (red.) *Phonologica 1984: proceedings of the fifth international phonology meeting*. New York (N. Y.): Cambridge University Press, 339–344.
- Ylioppilastutkintolautakunta 2016. Ylioppilastutkintoon osallistujat kokeittain. [https://www.ylioppilastutkinto.fi/images/sivuston\\_tiedostot/stat/FB2016KT2001.pdf](https://www.ylioppilastutkinto.fi/images/sivuston_tiedostot/stat/FB2016KT2001.pdf) [hämtad 2016-5-19].

## Bilaga 1

Detta är ett röstmeddelande som du ska tala in i din väns telefon. Vännen heter Peter, och meddelandet handlar om hans syster Eva som har råkat ut för en bilolycka, men har inte skadats allvarligt.

Titta igenom texten först, och läs sedan in den på band.

Om du säger / uttalar någonting fel, ingen fara – ta bara om hela meningen!

---

Hej Peter!

Jag försökte ringa dig, men din mobil var avstängd. Var är du?! Hoppas att du hör mitt meddelande snart. Eva ligger på sjukhus. Hon råkade ut för en bilolycka i morse, men det är ingen fara med henne!

Eva åkte till jobbet med min bil. I den stora korsningen på Vasa-gatan kom en buss som körde för fort och kunde inte stanna vid rödljuset. Som du kanske vet var det kyligt och jättehalt på morgonen. Eva hann inte stoppa sin bil utan körde rakt in i bussen. Hon kände sig yr, hon hade ont i huvudet, ryggen och ena benet. Därför fördes hon till sjukhus.

Till all lycka mår Eva ganska bra nu! Läkaren tror att hon kommer att skrivas ut i övermorgon eller kanske på fredag. Ska vi hälsa på henne på sjukhuset? Hon är i rum nummer 28B, på fjärde våningen. Om du vill så följer jag gärna med. Jag kan visa rummet för dig sedan – det är svårt att hitta på sjukhuset! Min bil är okörbar just nu så det blir du som får skjutsa mig!

Men ring mig när du hör det här meddelandet!



Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 118–138.

**Pauliina Peltonen**

University of Turku

## **L2 fluency in spoken interaction: a case study on the use of other-repetitions and collaborative completions**

Second language (L2) speech fluency has usually been studied from an individual's perspective with monologue speech samples, whereas fluency studies examining dialogue data, especially with focus on collaborative practices, have been rare. In the present study, the aim was to examine how participants maintain fluency collaboratively. Four Finnish upper secondary school students of English completed a problem-solving task in pairs, and their spoken interactions were analyzed qualitatively with focus on collaborative completions and other-repetitions. The findings demonstrated that collaborative completions and other-repetitions contribute to interactional fluency by creating cohesion to the interaction. Collaborative completions were also used to help the interlocutor to overcome temporary (individual) disfluent phases. Overall, the findings suggest that individual and interactional fluency are intertwined in spoken interaction, which should be acknowledged in theoretical approaches to L2 fluency and in empirical studies examining L2 fluency in interactional contexts.

**Keywords:** L2 speech fluency, interaction, language learning, oral proficiency

**Asiasanat:** vieraan kielen puheen sujuvuus, interaktio, kielen oppiminen, suullinen kielitaito

## 1 Introduction

Spoken interactions are fundamentally collaborative: each participant contributes to maintaining the flow of speech and to minimizing silences (see e.g. Clark 1996). Participants link and adapt their utterances to previous turns, demonstrating “[t]he ability to vary one’s lexis while still saying more or less the same thing” (McCarthy 1998: 115) and in this way create cohesion to the interaction (see also Tannen 1989: 50–51). However, creating links across turns can be particularly challenging in second language (L2) interactions, since resources must also be allocated to turn-internal aspects, including grammatical, lexical and phonological processing, which are not as automatized as in first language (L1) speech production (see Dörnyei & Kormos 1998: 354–355; Kormos 2006). Since the participants have to attend to both higher, interactional level (between-turn) and lower level (within-turn) aspects during interaction, the differences in their skills in maintaining fluency may be reflected in *individual fluency* (how fluently they produce speech during their own turns) as well as *interactional fluency* (how fluently their discussion proceeds across turns) (see also Lauranto 2005). The present study examines the interplay between these two phenomena by focusing on the following question: how do learners keep the flow of talk going collaboratively in spoken interaction?

While L2 speech fluency in monologue settings has been widely studied (recently, e.g., by Götz 2013; Kahng 2014; Peltonen & Lintunen 2016), the concept of L2 fluency has rarely been applied to interactional contexts (for recent exceptions, see Witton-Davies 2014; Tavakoli 2016). Both in L2 speech fluency research and the mainstream second language acquisition (SLA) research tradition more generally, L2 competence has traditionally been approached from an individual’s perspective (for criticism of the individualistic approach in SLA, see e.g., Firth & Wagner 1997). Similarly, when L2 speech fluency has been examined in dialogue, the focus has often been on analyzing an individual’s performance, while less attention has been paid to the interaction between the participants (but see Riegenbach 1991; Hüttner 2009). However, for capturing fluency in an interactional context, it is necessary to extend the analysis from individual contributions to phenomena that relate to the collaborative construction of fluency. The present, exploratory case study approaches this under-researched area by focusing on two interactional practices, *other-repetitions* (words or longer stretches of interlocutor’s speech repeated without modification) and *collaborative completions* (contributions filling in an utterance that the previous speaker has started). The practices are examined as indicators of interactional L2 fluency. With a detailed qualitative analysis, the study shows how other-repetitions and collaborative completions can be used as a basis for examining fluency in interactional data and highlights the importance of incor-

porating collaborative aspects to conceptualizations of L2 fluency.

This article starts with a discussion on the approaches to fluency in interactional settings in section 2. In section 3, the methodology for the study is presented, including information about the participants (section 3.1) and data collection and analysis (section 3.2). The present study focuses on four upper secondary school students of English performing a problem-solving task in pairs. The findings of the qualitative analysis are presented in section 4: section 4.1 focuses on the use of other-repetitions and section 4.2 on collaborative completions. While the analysis focuses on how the participants maintain interactional fluency collaboratively, connections between individual and interactional fluency are also examined. The findings are discussed in section 5. In the discussion, particular attention is paid to the implications of the findings for L2 fluency research.

## 2 Approaches to interactional L2 fluency

Fluency is generally regarded as one aspect of (oral) L2 proficiency (on the Complexity-Accuracy-Fluency-framework, see e.g., Housen et al. 2012). The degree of smoothness and effortlessness in L2 speech (e.g. Chambers 1997) is commonly studied by quantifying fluency-related phenomena, such as pauses and other hesitations, that can be grouped into three main dimensions of fluency: speed (e.g., speech rate), pausing (the frequency, duration and location of pauses), and repair (false starts, repetitions, and reformulations; Skehan 2003, 2009, 2014; Tavakoli & Skehan 2005). As the measures are applied to L2 speech, the implicit point of comparison is maximally fluent native speech without pauses or other disfluencies. However, this idealized view has been criticized by several researchers (e.g., Fillmore 1979; Lennon 1990) and challenged in empirical studies; studies that have included a native speaker control group have demonstrated that also native speakers vary in their fluency (e.g., Götz 2013; Kahng 2014; Peltonen & Lintunen 2016).

As stated in the introduction, fluency has usually been examined from the individual's perspective with monologue speech samples. The focus on an individual's fluency is also often reflected in definitions of fluency: a case in point is Lennon's (2000: 26) oft-cited description of fluency as "the rapid, smooth, accurate, lucid, and efficient translation of thought or communicative intention into language under the temporal constraints of on-line processing". Viewing efficient processing as the basis for fluent speech, the definition is in line with the view of *utterance fluency*, i.e. measurable features in speech samples, reflecting an individual's underlying *cognitive fluency* (the efficiency and ease in processing; Segalowitz 2010).

To complement the view of fluency as an individual's ability with a so-

cial perspective, researchers have recently suggested that a social dimension should be incorporated to L2 fluency analysis (e.g., Segalowitz 2016; Wright & Tavakoli 2016), in line with calls to approach L2 learning and competence, as well as language use more broadly, from both social and cognitive perspectives (e.g., Clark 1996; Douglas Fir Group 2016). However, so far descriptions of L2 speech fluency have usually not included the social aspect (but see Lauranto 2005; Kirk & Carter 2010), and few studies have examined collaborative fluency in interactional contexts. A notable exception is Riggensbach's (1991) pioneering study on L2 fluency in an interactional setting, which included an examination of collaborative completions and *echoes* (other-repetitions), among other interactional phenomena, as potential "conversational fluency" indicators. More recently, Witton-Davies (2014) and Tavakoli (2016) compared L2 utterance fluency in monologue and dialogue settings; however, as the starting point in these studies was to apply monologue measures to dialogue data, the individual's perspective is more prominent than the collaborative aspects. Extending fluency analysis from monologue to dialogue contexts is the important first step in exploring interactional fluency, but more studies specifically on learners' joint efforts to maintain fluency are needed to achieve a comprehensive picture of fluency in interactional settings.

Despite the lack of studies focusing on interactional fluency within the field of L2 speech fluency research, the co-construction of interactional flow has been examined from other perspectives. For instance, CA-SLA/CA-for-SLA researchers applying conversation analysis to L2 data have examined learners' interactional practices as indicators of *interactional competence* (IC; for overviews, see e.g., Hall & Pekarek Doehler 2011; Kasper & Wagner 2011). Interactional practices related to creating cohesion in turn-taking and avoiding long pauses (*turn-taking management*, see e.g., He & Young 1998; Galaczi 2014) can also be considered as essential aspects of interactional fluency. For instance, Peltonen (2017) explored L2 fluency in an interactional setting and found that one of the main features distinguishing learner groups from different school levels (9th grade and upper secondary school) was the duration and frequency of between-turn pauses in addition to measures of individual fluency. Similarly, studies in the field of language testing have shown that fast-paced turn-taking, including overlaps and latches, along with mutual topic development and smoothness in topic transitions, characterize learners' IC at the highest proficiency levels (e.g., Galaczi 2014). Contributing to the co-construction of interaction, the two interactional practices examined in the present study have also been analyzed from the perspectives of *alignment* and *accommodation*, since both other-repetitions and collaborative completions can be treated as indicators of speakers acknowledging previous turns and adapting their contributions accordingly (e.g., Pickering & Garrod 2004; Cogo & Dewey 2006: 66–73; Dings 2014).

The essential aspect in the approaches presented above is the view of interaction as co-constructed (see e.g., Jacoby & Ochs 1995; He & Young 1998). As time is a shared resource in dialogue (affecting e.g., interactional rhythm; Auer et al. 1999), avoiding long pauses and maintaining the progressivity of talk are shared responsibilities for the participants (see also Clark 1996; on the universal tendency to minimize silences between turns, see Stivers et al. 2009). From the perspective of fluency, the collaborative nature of spoken interaction is captured particularly well in McCarthy's (2010: 7) concept *confluence*, which refers to the joint production of flow in a dialogue setting: in cohesive interaction, the participants maintain fluency not only within their own turns but also across turn boundaries. The notion of confluence (or interactional fluency) thus expands the perspective from an individual speaker's fluency to the interaction as a whole. Also Hüttner's (2009) examination of raters' perceptions of fluency in English as a lingua franca (ELF) interactions revealed that raters orient to the co-constructed nature of fluency. Therefore, while individual and interactional fluency are approached theoretically as related but separate notions in the present study (see also Sato 2014: 88; Tavakoli 2016: 147–148), in practice the two aspects are likely to interact and overlap.

The first interactional practice examined in the present study, *other-repetition*, refers to words or longer stretches of interlocutor's speech that are repeated without modification (cf. self-repetition, where the speaker's own speech is repeated). In the context of ELF interaction, other-repetitions have also been referred to as *Represents* (House 2002: 254), highlighting the fact that "re-presenting" the interlocutor's words helps in coping with processing time pressure (see also Hüttner 2009). In other words, from the perspective of fluency, other-repetitions can contribute to individual fluency by providing more planning time for the speaker; repeating words from the interlocutor's output does not require much attentional resources and thus processing resources can be allocated to planning the rest of the utterance (Dörnyei & Kormos 1998: 368–371; see also Tannen 1989: 48–49; Pickering & Garrod 2004: 181). In addition to facilitating processing, other-repetitions can demonstrate attentiveness to and agreement with the interlocutor's contribution (e.g., Cogo & Dewey 2006: 66–73)<sup>1</sup>. Finally, since other-repetitions link the speakers' contributions to the preceding turn(s), they function as devices for creating cohesion to the interaction (Tannen 1989: 50–51) and therefore also enhance interactional fluency.

Similarly, *collaborative completions*, also referred to as *anticipatory completions* (Lerner 1996) or *pre-emptive completions* (Lerner 2004), contribute to

<sup>1</sup> (Other-)repetitions can also have various other functions in interaction. For instance, CA and CA-SLA studies examining repair (in a fairly broad sense, referring to practices for addressing various problems in interaction; Schegloff et al. 1977) have focused on the repair-initiating functions of repetitions (e.g., Kurhila 2006; Lilja 2010; Kurhila & Lilja 2017).

interactional fluency, cohesion and the co-construction of interaction. Collaborative completions fill in an utterance that the previous speaker has started (e.g., Dings 2014). While collaborative completions have been studied in different L1s, including conversations in English (e.g., Lerner 1991, 1996, 2004; Rühlemann 2007), Finnish (Helasvuo 2004), Japanese (e.g., Ono & Yoshida 1996; Hayashi 2014) and Korean (Kim 2002), fewer studies have been conducted on collaborative completions in L2 interactions (but see Dings 2014; Taguchi 2014). From the perspective of the present study, the central functions of collaborative completions identified in previous studies (e.g., Hayashi 2014; Taguchi 2014) include achieving a shared perspective with the interlocutor by demonstrating agreement with their viewpoint and, commonly found especially in L2 interactions, solving communication problems for instance in conjunction with word searches<sup>2</sup>.

While many opportunities for turn completion are created with certain two-part syntactic constructions (e.g. the “if X – then y”-construction, forming a *compound turn-constructive unit*; Lerner 1991), for the present study, unprojected opportunities for completion are more central: they occur in interaction due to “occasional halting of a turn’s progressivity” (Lerner 1996: 257), including laughter, within-turn pauses, word searches, word-cut-offs, repetitions, and non-verbal behavior (Lerner 1996: 256–267). For instance, the opportunity for completion arising from a within-turn pause is demonstrated in Lerner’s (1996: 260) example, where the first speaker produces the beginning of the turn (“Did they do that old trick with the basketball where they putta”) and the second speaker completes it after a brief 0.4 second pause (“string around it”). In addition to examining mid-turn pauses, identifying instances of word searches is particularly relevant for the present study, as collaborative completions have often been found to occur in conjunction with word searches in L2 data (Taguchi 2014). Typically, completions related to word searches are short and only include the expression that is being searched for (Lerner 1996: 261). Furthermore, word searches often occur near the end of the turn unit (*terminal item completions*; Lerner 1996: 256, 262). In terminal item completions, the final word(s) in a turn can be produced by the interlocutor, or they can be co-produced with the current speaker (Lerner 1996: 256).

To summarize, the study examines L2 speech fluency in an interactional setting. Two main senses of fluency can be distinguished: *individual fluency* refers to maintaining the flow of speech within one’s own turn, while *interactional fluency* refers to collaboration in maintaining flow across turn boundaries (cf. McCarthy’s 2010 confluence and the discussion above). The study addresses the following research question: How do collaborative completions and other-repetitions contribute to maintaining fluency in L2 interaction?

<sup>2</sup> Word searches (and collaborative completions) have also been examined within the broader framework of repair organization in CA-SLA studies (see e.g., Kurhila 2006).

Both other-repetitions and collaborative completions involve acknowledgment of the interlocutor's contributions and are therefore regarded primarily as potential indicators of interactional fluency.

### 3 Methodology

#### 3.1 Participants

The participants in the present study were four 17-year-old, Finnish-speaking second year upper secondary school students of English (pairs: Eero and Timo, both male; Kati and Anni, both female). The participants' names have been replaced with pseudonyms to secure their anonymity. The students took part in a larger research project examining L2 speech fluency from different perspectives (see Peltonen 2017) and were chosen for the present study since they were among the pairs that produced the most collaborative completions and other-repetitions. The participants had studied English for approximately eight years; the results from a vocabulary test aimed at estimating their overall proficiency in English (LexTALE, see Lemhöfer & Broersma 2012) suggest that Kati (score 75.00%), Eero (score 72.50%) and Anni (score 63.75%) represent level B2 (scores between 60% and 80%) in the Common European Framework of Reference (Council of Europe 2001), while Timo (score 97.50%) represents level C1–C2 (scores between 80% and 100%; see Lemhöfer & Broersma 2012: 341).

#### 3.2 Data collection and analysis

The participants completed a communicative problem-solving task in pairs. The participants were given pictures of sixteen items and asked to discuss them in English and to rank them in the order of their potential usefulness for survival on a desert island (after Klippel 1984: 63–64 and Ur 1990: 70–72). The task was piloted for suitability with a learner group that did not participate in the actual study. The participants completed the task one pair at a time in a quiet space during regular school days. Before the task, the participants were given a maximum of two minutes of individual preparation time. After the preparation, the pairs were given six minutes to complete the task together. The researcher was present in the room only to record the dialogue and did not participate in the interaction.

In contrast to the mostly quantitative fluency studies, the approach employed in the present study was qualitative. The approach was influenced by the qualitative components included in previous mixed methods L2 (monologue) fluency studies (e.g., Ejzenberg 2000; Hilton 2008; Peltonen & Lintunen 2016), as well as conversation analysis. After all instances of collabo-

rative completions and other-repetitions had been identified from the dialogue transcriptions (for transcription conventions, see Appendix), the functions and contexts of these two interactional practices were analyzed in detail. More specifically, for collaborative completions, the following aspects were coded: the type of completion (unprojected or forming a compound turn-constructional unit), the extent of completion (a single word or phrase or a longer contribution) and the location of completion (terminal item completion or other). The interplay between interactional and individual fluency was also examined. The data were analyzed twice by the researcher with a 5-month gap between the rounds to increase the reliability of the analysis.

During their relatively short but fast-paced discussion that lasted for approximately three minutes (173.79 seconds), Timo and Eero produced 674 syllables (Eero 59% of them). Their discussion included 14 instances of other-repetition (Eero 4, Timo 10) and three collaborative completions (Eero 1, Timo 2). Kati and Anni's discussion lasted for approximately seven minutes (414.26 seconds). During the discussion, Kati and Anni produced 995 syllables (Kati 61% of them). Their discussion contained four other-repetitions (Kati 3, Anni 1) and five collaborative completions (Kati 4, Anni 1). In the following, the findings for the use of other-repetitions and collaborative completions will be discussed in sections 4.1 and 4.2, respectively.

## 4 Findings

### 4.1 Other-repetitions

Example 1 illustrates a typical pattern in the use of other-repetitions: other-repetition followed by elaboration (being applicable approximately to 2/3 of other-repetitions in the data).

#### (1) Other-repetition with elaboration

66 →	E:	ää (.) maybe (0.39) the knife (0.28)
67		[we can-]
68 →	T:	[knife ] yeah <b>knife</b> so you can (.)
69	E:	[ah]
70	T:	[gut] the fish [so]
71	E:	[ y]es (0.27) or +hunt+
72		(0.27)
73 →	T:	<b>hunt</b> [yeah]
74	E:	[wah-]
75		(0.64)
76	E:	[but-]
77	T:	[°or°] (.) kill boars [with it *heh*]
78	E:	[*ahahahah* ] yes



In Example 1, Timo's other-repetition of *knife* in line 68 indicates acknowledgement of Eero's contribution and shows agreement with it (note also the following agreement token *yeah*), as well as creates cohesion to the interaction. From the perspective of individual fluency, the other-repetition also helps in coping with processing time pressure: picking the word from Eero's output frees resources for planning the rest of the turn. Timo's subsequent self-repetition of the word *knife* provides him additional planning time. In the rest of the turn (lines 68 and 70), Timo provides a reason why the knife is a good choice (*so you can gut the fish*) and in this way contributes to the progression of the discussion. Typically, the elaborations following other-repetitions contain justifications for the chosen items.

After Timo's elaboration, Eero continues in line 71, acknowledging Timo's contribution (*yes*) and providing another reason for choosing the knife (*or hunt*), creating cohesion to the interaction by linking his turn to Timo's utterance. After Eero's contribution, Timo uses the "other-repetition + elaboration" -pattern again with a repetition of the word *hunt* in line 73, followed by a more specific example of hunting with the knife in line 77 (*kill boars with it*). The interaction proceeds cohesively; both participants contribute to interactional fluency by linking their own contributions to the previous speaker's turn with other-repetitions and elaborations.

Example 2 demonstrates the use of other-repetition in conjunction with a word search, which occurred five times in the data. In Example 2, in addition to demonstrating acknowledgement and creating cohesion, the other-repetition confirms the result of the word search as the correct target language item. In similar cases in the data, the target item was also occasionally provided by the interlocutor, which was then repeated by the first speaker (as in Example 5, section 4.2). Especially the latter instance is potentially beneficial from the perspective of language learning, since the interlocutor can provide interactional adjustments above the other participant's level (see the *negotiation for meaning* framework in SLA; e.g., Long 1996). The role of *scaffolding* (Wood et al. 1976) or assistance by more advanced participants in completing activities beyond one's own capabilities has also been central in sociocultural approaches to L2 learning.

## (2) Other-repetition in conjunction with a word search

85 →	E:	and the- what is (.) that called +hammer+
86		(0.27)
87 →	T:	<b>hammer</b> yeah
88	E:	[so-]
89	T:	[so ] you can build stuff (.)
90	E:	ye[s ] [yes ]
91	T:	[out of] the palms and [so on]
92		(0.46)
93	E:	and we can (0.35) ahm umm {*pt*_0.45}

94 E: set up the (.) camp an (.)  
 95 T: yeah  
 96 (0.36)  
 97 E: tent

Example 2 begins with Eero's word search in line 85, indicated by the wh-question *what is that called*. The question is, however, not directed at Timo: Eero gazes towards the paper (indicating engagement in a "solitary word search"; Goodwin & Goodwin 1986) and immediately provides the result of the word search (*hammer*) himself. Followed by a brief silent pause in line 86, Timo confirms the result of the word search with a repetition of the word *hammer* in line 87. Following the pattern discussed earlier, Timo continues his turn by providing a reason for choosing the hammer (lines 89 and 91). During Timo's turn, Eero indicates agreement with two backchannels in line 90. After Timo's elaboration, the discussion proceeds cohesively, as Eero builds on Timo's turn and continues with an alternative use for the hammer in lines 93–94 and 97 (*set up the camp and tent*). Due to this linking of turns, the discussion can be characterized as interactionally fluent.

## 4.2 Collaborative completions

None of the eight collaborative completions in the data occurred as parts of compound turn-constructural units (Lerner 1991, 1996); that is, they did not complete particular two-part syntactic structures that include a preliminary component and project a final component. In most cases, the completions were clearly of the "unprojected" type and occurred mid-clause, most commonly being preceded by silent pauses that provided the opportunity for turn entry for the interlocutor. In some cases, a shorter micropause (of less than 0.25 seconds in duration) occurred in combination with another cue (filled pause *uh*, FP, as in Example 5, or rising intonation). Overall, half of the completions produced by the participants were terminal item completions: that is, they occurred towards the end of the turn. The majority of the completions were also relatively short, usually consisting of a single word (Example 5) or a single phrase (Examples 3 and 4).

Example 3 illustrates a collaborative completion occurring in conjunction with a word search.

### (3) Collaborative completion in conjunction with word search 1

98 A: okay (.) a::nd (.) then would be: {\*pt\*\_2.76}  
 99 K: mm  
 100 (2.70)  
 101 K: \* RAISES GAZE FROM PAPER, LOOKS AT A  
 102 K: [oh the- ]  
 103 → A: [maybe the]n the uh fishing- (0.26)

104	A:	* RAISES GAZE FROM PAPER, LOOKS AT K
105	A:	* RIGHT HAND MOVES FROM CENTER TO RIGHT
106 →	K:	yeah (.) [ <b>fishing</b> ] gear[s ]
107	A:	[°thing°] [ye]ah
108	K:	and maybe the: (.) emergency (0.62)
109	K:	um (0.38) r- uh \$ro[cket\$ *hah*]
110	A:	[*hahaha* ]

Typically for a word search completion, Kati's collaborative completion in line 106 is relatively short (a single noun phrase) and preceded by a hesitation cluster in line 103 that provides the unprojected opportunity for mid-clause turn entry. The cluster consists of a FP, followed by a cut-off word *fishing* and a brief silent pause, suggesting that Anni is engaged in a word search. In addition, Anni's non-verbal behavior in lines 104–5 indicates that she is not able to retrieve the word by herself: she raises her gaze towards Kati during the word search, which can be interpreted as a request for Kati to participate in it, supported also by Anni's simultaneous gesturing with her right hand (see also Goodwin & Goodwin 1986). At the same time when Kati offers her help and suggests a solution to the word search (*fishing gears*) in line 106, Anni completes the word search herself with an all-purpose word *thing* in line 107 (note that it is spoken softly, possibly signaling her awareness of it not being the correct target word). This results in slightly overlapping contributions (see also Example 4). Anni acknowledges Kati's suggestion immediately (*yeah*) in line 107, and Kati proceeds with the turn (as in Example 4) by suggesting another item. Thus, a potential problem is solved collaboratively and the flow of the discussion is maintained through mutual effort (see also Taguchi 2014).

Another example of a collaborative completion is illustrated in Example 4. Note that in this example, Kati's collaborative completion in line 26 overlaps with Anni's turn (cf. partially overlapping contributions in Example 3). While Eero and Timo did not produce overlapping completions, four of the five completions produced by Kati and Anni occurred partly or completely in overlap with the interlocutor's speech.

#### (4) Overlapping collaborative completion

19	A:	okay (.) the second would (0.46) be::: (0.41)
20		mm: (2.78)
21		I would say matches (0.45) because=
22	K:	=yeah
23		(0.37)
24 →	A:	then we can (0.62)
25 →		[light up a fi]re (.) [yeah]
26 →	K:	[ <b>make fire</b> ] (.) [yeah]
27		(0.45)
28	K:	if it's cold [then it]'s °ke-° getting warmer
29	A:	[yeah ]

In Example 4, Kati's collaborative completion (*make fire* in line 26) matches the content in Anni's turn (*light up a fire* in line 25), although the lexical choices differ slightly. The utterances are co-produced, and the overlap continues after the collaborative completion, when both acknowledge each other's contributions (*yeah*) at the same time in lines 25–26. Despite not uttering exactly the same words and the overlap lasting only for a short while, the co-ordination of their talk is reminiscent of so-called *choral co-production*, where contributions are produced at the same time and matched for e.g. lexis and tempo (Lerner 2002: 226).

The opportunity for turn entry in line 26 is created by the preceding silent pause in line 24<sup>3</sup>. However, compared to Example 3, it is not entirely clear whether the completion here relates to a word search. Similarly to typical word search completions, the “unprojected” opportunity for turn entry emerges mid-clause, and the completion is in itself short (a verb phrase). However, Kati does not yield the floor after the completion and continues with the turn herself in line 28, developing the idea of choosing matches further (*if it's cold then it's getting warmer*). Kati's completion, followed by the elaboration, ensures the smooth progression of the discussion and thus contributes to interactional fluency. Although the interaction between Anni and Kati was relatively well balanced (Kati produced 61% of the syllables during the interaction), overall Kati was the more active participant in taking responsibility of ensuring the progression of the discussion.

The final example of a collaborative completion is illustrated in Example 5. Here, as in Example 3, the completion is clearly related to a word search.

##### (5) Collaborative completion in conjunction with word search 2

154	T:	the sleeping bag might be (0.25)
155		like (.) useless
156		(0.43)
157	E:	yes but (0.44)
158		it's nice if (.) [it's if ] it's get- (0.41)
159	T:	[it's nice]
160	E:	\$if it gets\$ really (.) u:h
161 →	T:	cold
162 →	E:	cold (.) [during] the nights ah (.)
163	T:	[yeah ]
164	E:	you might need it

In Example 5, Timo provides a collaborative solution to Eero's word search with his completion (*cold*) in line 161 (notably during a silence, not overlapping with the interlocutor's speech as in Examples 3 and 4). Since *cold* is a high-frequency word, the word search is most likely due to a retrieval problem and

<sup>3</sup> As pointed out by an anonymous reviewer, Anni's turn also requires continuation from a grammatical perspective, as it only contains the finite part of the verb chain (on grammatical and interactional projection, see Auer 2005).

not due to a lexical gap. The completion can be regarded as a typical word search completion, since it is very short, only a single word, and occurs towards the end of the utterance (terminal item completion). Also characteristically for a word search completion, Timo yields the floor to Eero after the completion (cf. Examples 3 and 4). First, Eero accepts and acknowledges Timo's completion by repeating it in line 162 and then continues with the turn in lines 162 and 164.

A closer examination of the context preceding the completion reveals that Eero's turn (lines 157–160) is not entirely fluent (from the perspective of individual fluency): he stumbles over his words slightly (and most likely also notices this himself, indicated by the laughing voice in line 160), as indicated by the multiple false starts in line 158. The disfluencies occur already before the immediate context for turn entry in line 160, where a combination of a micropause and a FP (that in itself also contains a drawl, lengthening of the initial vowel) precede the completion. This "halting of a turn's progressivity" (Lerner 1996: 257) indicates a solitary word search (Eero's gaze directed away from Timo, not inviting his participation), but at the same time provides the unprojected opportunity for Timo's turn entry. In contrast to the clearly noticeable silent pauses that often precede completions, Timo still appears attuned to these fairly subtle cues. At least two reasons for this can be postulated: first, Eero's struggles with formulating his turn (the false starts) foreshadow the possibility for completion even before the combination of a micropause and a FP in line 160. Second, the overall fast-paced turn-taking during the interaction (the average turn pause duration being only 0.44 seconds) reflects the participants' mutual orientation to overcoming even relatively unnoticeable "hitches" in turn progressivity and efficiently minimizing very short silences.

From the perspective of interactional and individual fluency, Example 5 illustrates particularly well how the orientation to keeping the flow of talk going does not only occur collaboratively across turns (interactional fluency) or individually within the speaker's own turns (individual fluency), but can also reach to the other speaker's turn. In other words, minimizing pauses and maintaining fluency within the turns, which is perhaps more commonly thought of as being an individual speaker's responsibility (at least from an L2 fluency perspective), can also be accomplished collaboratively with the help of the interlocutor. Timo's collaborative completion could thus be characterized as him helping Eero to get over a temporary, local disfluency, while simultaneously maintaining interactional fluency.

## 5 Discussion and conclusion

This study examined the role of other-repetitions and collaborative completions in maintaining fluency in L2 interaction. The analysis of other-repetitions demonstrated that their main functions were to acknowledge and confirm the interlocutor's contribution and to create cohesion to the interaction (see also Tannen 1989; Cogo & Dewey 2006). The other-repetitions were often followed by elaborations, which further contributed to the smooth progression of the discussion and therefore also to interactional fluency. In addition to contributing to interactional fluency by linking utterances across turn boundaries, other-repetitions also helped to maintain individual fluency during the speaker's own turn. The speaker can use other-repetitions to cope with processing time pressure, since picking words from the interlocutor's output frees processing resources and provides more planning time (e.g., Tannen 1989: 48–49; Dörnyei & Kormos 1998: 368–371; on individual fluency resources, see also Peltonen 2017). To summarize, other-repetitions simultaneously contribute to both individual and interactional fluency (see also Figure 1).

Collaborative completions were also found to contribute to the flow of the interaction by creating cohesion to the talk across individual contributions (see also Rühlemann 2007: 101). In line with other studies on L1 (Rühlemann 2007: 101–102; Hayashi 2014) and L2 collaborative completions (Riggenbach 1991: 437; Dings 2014: 752; Taguchi 2014), they were found to be important in demonstrating a shared perspective with the interlocutor, as well as showing agreement with and acknowledgement of the interlocutor's contribution. Furthermore, similarly to Taguchi's (2014) findings, some of the collaborative completions in the present study occurred in conjunction with word searches and were used to solve potential communication problems collaboratively.

However, the present study also showed how collaborative completions functioned as devices for maintaining fluency, which has not been highlighted in previous studies. While fluency in an interactional setting was theoretically approached from the perspectives of individual within-turn fluency and interactional between-turns fluency (see also Figure 1), the analysis of collaborative completions showed that this distinction was not clear-cut. Example 5 was particularly illustrative: Timo's one-word contribution in the middle of Eero's turn helped Eero to cope with his (individual) within-turn disfluency. The example demonstrated that in addition to minimizing gaps collaboratively between turns and individually within turns, the participants engaged in within-turn collaboration to maintain fluency (cf. Lerner 1996: 267). Figure 1 illustrates how the element of "collaborative within-turn fluency" relates to individual and interactional fluency. These findings highlight the importance of acknowledging the intertwined nature of individual and interactional fluency both in theoretical and empirical approaches to fluency in interactional settings.

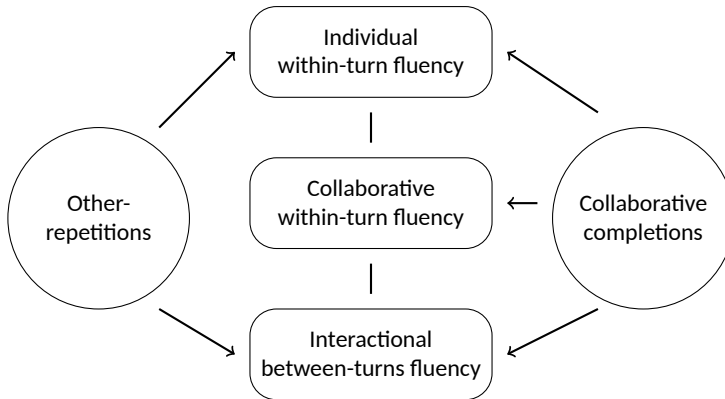


FIGURE 1. Resources for maintaining L2 fluency in interaction.

Methodologically, the present study demonstrated the usefulness of examining other-repetitions and collaborative completions as indicators of fluency in an interactional context. For multifunctional phenomena, such as the two practices examined in the present study, a fine-grained qualitative analysis is particularly revealing. However, acknowledging the limitations to generalizability based on the findings of this small-scale, exploratory case study and the potential effects of the task instructions on the interaction, more studies on interactional fluency in different settings are needed to confirm the tendencies demonstrated here.

In addition to collaborative completions and other-repetitions, future research could also examine other potential resources for the co-construction of fluency, such as choral co-production (see Lerner 2002), different means for recycling vocabulary (e.g. *relexicalization*, or the use of near-synonyms in linking turns, McCarthy 1998: 112–116; see also Galaczi 2014 on topic development as an indicator of IC) and non-verbal behavior (including *embodied completions*, see e.g., Olsher 2004; Mori & Hayashi 2006). However, while broadening the scope of L2 fluency analysis to the collaboration between the participants provides a new viewpoint to L2 fluency, it also poses new challenges: to narrow the focus of interactional fluency analysis and to avoid interactional fluency becoming an all-encompassing term, it is important to explicate how interactional fluency relates to other, similar concepts (such as alignment or IC, cf. discussion in section 2). Therefore, in the future, more interdisciplinary dialogue among CA-SLA, language testing and L2 fluency researchers is needed.

Acknowledging the interconnected nature of individual and interactional fluency has also important implications for language teaching and language testing. In language testing contexts, in addition to assessing the individual's

competence, collaboration between the interlocutors should also be taken into account (see also Hüttner 2009; Sato 2014). Similarly, Taguchi (2014: 531) points out that due to the development of the use of collaborative completions during study abroad, they could be assessed as indicators of IC. Furthermore, from a pedagogical perspective, pair activities encouraging collaboration and providing opportunities for joint problem-solving facilitate the development of learners' interactional competence (see also Long 1996; He & Young 1998) and prepare the learners for interactions taking place outside the classroom.

In conclusion, the study has illustrated how participants maintain fluency in interaction not only individually, but also collaboratively. Instead of focusing solely on their individual contributions by presenting two monologues to each other (cf. House's 2002: 251 observations of "parallel monologues" in ELF interaction), the participants displayed mutual efforts to maintain the progressivity of talk. Therefore, focusing on the participants' individual performances when analyzing fluency in an interactional setting provides only a partial view of L2 fluency; it is equally important to acknowledge the jointly constructed nature of fluency (or confluence, McCarthy 2010; see also Lauranto 2005; Hüttner 2009). When analyzing interactional fluency, in addition to examining how cohesive links are created across turns with other-repetitions and collaborative completions (as in the present study), it is also important to pay attention to turn pauses, as they can provide an additional perspective to examining how smoothly the interaction proceeds (see Peltonen 2017). While this approach can provide a starting point for studying the collaborative aspect of L2 fluency, exploring also other ways of establishing interactional flow is essential to reach a comprehensive account of L2 fluency in interaction.

## References

- Auer, P. 2005. Projection in interaction and projection in grammar. *Text & Talk*, 25, 7–36.
- Auer, P., E. Couper-Kuhlen & F. Müller 1999. *Language in time: the rhythm and tempo of spoken interaction*. Oxford: Oxford University Press.
- Chambers, F. 1997. What do we mean by fluency? *System*, 25, 535–544. DOI: 10.1016/S0346-251X(97)00046-8.
- Clark, H. H. 1996. *Using language*. Cambridge: Cambridge University Press.
- Cogo, A. & M. Dewey 2006. Efficiency in ELF communication: from pragmatic motives to lexico-grammatical innovation. *Nordic Journal of English Studies*, 5 (2), 59–93. <http://ub016045.ub.gu.se/ojs/index.php/njes/article/view/65>.
- Council of Europe 2001. *Common European framework of reference for languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.



- Dings, A. 2014. Interactional competence and the development of alignment activity. *The Modern Language Journal*, 98, 742–756. DOI: 10.1111/j.1540-4781.2014.12120.x.
- Dörnyei, Z. & J. Kormos 1998. Problem-solving mechanisms in L2 communication: a psycholinguistic perspective. *Studies in Second Language Acquisition*, 20, 349–385.
- Douglas Fir Group, The 2016. A transdisciplinary framework for SLA in a multilingual world. *The Modern Language Journal*, 100, 19–47. DOI: 10.1111/modl.12301.
- Ejzenberg, R. 2000. The juggling act of oral fluency: a psycho-sociolinguistic metaphor. In H. Riggenbach (ed.) *Perspectives on fluency*. Ann Arbor (Mich.): University of Michigan Press, 287–313.
- Fillmore, C. J. 1979. On fluency. In C. J. Fillmore, D. Kempler & W. S-Y. Wang (eds) *Individual differences in language ability and language behavior*. London: Academic Press, 85–102.
- Firth, A. & J. Wagner 1997. On discourse, communication, and (some) fundamental concepts in SLA research. *The Modern Language Journal*, 81, 285–300. DOI: 10.1111/j.1540-4781.1997.tb05480.x.
- Galaczi, E. D. 2014. Interactional competence across proficiency levels: how do learners manage interaction in paired speaking tests? *Applied Linguistics*, 35, 553–574. DOI: 10.1093/applin/amt017.
- Goodwin, M. H. & C. Goodwin 1986. Gesture and coparticipation in the activity of searching for a word. *Semiotica*, 62, 51–76. DOI: 10.1515/semi.1986.62.1-2.51.
- Götz, S. 2013. *Fluency in native and nonnative English speech*. Amsterdam: John Benjamins.
- Hall, J. K. & S. Pekarek Doehler 2011. L2 interactional competence and development. In J. K. Hall, J. Hellermann & S. Pekarek Doehler (eds) *L2 interactional competence and development*. Bristol: Multilingual Matters, 1–15.
- Hayashi, M. 2014. Activity, participation, and joint turn construction: a conversation analytic exploration of ‘grammar-in-action’. In K. Kabata & T. Ono (eds) *Usage-based approaches to Japanese grammar: towards the understanding of human language*. Amsterdam: John Benjamins, 223–258.
- He, A. W. & R. Young 1998. Language proficiency interviews: a discourse approach. In R. Young & A. W. He (eds) *Talking and testing: discourse approaches to the assessment of oral proficiency*. Amsterdam: John Benjamins, 1–24.
- Helasvuo, M.-L. 2004. Shared syntax: the grammar of co-constructions. *Journal of Pragmatics*, 36, 1315–1336. DOI: 10.1016/j.pragma.2004.05.007.
- Hilton, H. 2008. The link between vocabulary knowledge and spoken L2 fluency. *Language Learning Journal*, 36 (2), 153–166. DOI: 10.1080/09571730802389983.
- House, J. 2002. Developing pragmatic competence in English as a lingua franca. In K. Knapp & C. Meierkord (eds) *Lingua franca communication*. Frankfurt am Main: Peter Lang, 145–268.
- Housen, A., F. Kuiken & I. Vedder (eds) 2012. *Dimensions of L2 performance and proficiency: complexity, accuracy and fluency in SLA*. Amsterdam: John Benjamins.
- Hüttner, J. 2009. Fluent speakers—fluent interactions: on the creation of (co-)fluency in English as a lingua franca. In A. Mauranen & E. Ranta (eds) *English as a lingua franca: studies and findings*. Newcastle upon Tyne: Cambridge Scholars Publishing, 274–297.

- Jacoby, S. & E. Ochs 1995. Co-construction: an introduction. *Research on Language and Social Interaction*, 28 (3), 171–183. DOI: 10.1207/s15327973rlsi2803\_1.
- Kahng, J. 2014. Exploring utterance and cognitive fluency of L1 and L2 English speakers: temporal measures and stimulated recall. *Language Learning*, 64, 809–854. DOI: 10.1111/lang.12084.
- Kasper, G. & J. Wagner 2011. A conversation-analytic approach to second language acquisition. In D. Atkinson (ed.) *Alternative approaches to second language acquisition*. New York (N. Y.): Routledge, 117–142.
- Kim, H. 2002. Collaborative turn completion in Korean conversation. *Language Research*, 38 (4), 1281–1316.
- Kirk, S. & R. Carter 2010. Fluency and spoken English. In M. Moreno Jaén, F. Serrano Valverde & M. Calzada Pérez (eds) *Exploring new paths in language pedagogy: lexis and corpus-based language teaching*. London: Equinox, 25–38.
- Klippel, F. 1984. *Keep talking: communicative fluency activities for language teaching*. Cambridge: Cambridge University Press.
- Kormos, J. 2006. *Speech production and second language acquisition*. Mahwah (N. J.): Lawrence Erlbaum.
- Kurhila, S. 2006. *Second language interaction*. Amsterdam: John Benjamins.
- Kurhila, S. & N. Lilja 2017. Toisto ja korjauksen rajat. *Virittäjä*, 121, 213–243. DOI: 10.23982/vir.56632.
- Lauranto, Y. 2005. Sujuvuuden mittoja. In L. Kuure, E. Kärkkäinen & M. Saarenkunnas (eds) *Kieli ja sosiaalinen toiminta – language and social action*. AFinLA yearbook 2005. Jyväskylä: Finnish Association of Applied Linguistics AFinLA, 127–147.
- Lemhöfer, K. & M. Broersma 2012. Introducing LexTALE: a quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44, 325–343. DOI: 10.3758/s13428-011-0146-0.
- Lennon, P. 1990. Investigating fluency in EFL: a quantitative approach. *Language Learning*, 40, 387–417. DOI: 10.1111/j.1467-1770.1990.tb00669.x.
- 2000. The lexical element in spoken second language fluency. In H. Riggenbach (ed.) *Perspectives on fluency*. Ann Arbor (Mich.): University of Michigan Press, 25–42.
- Lerner, G. H. 1991. On the syntax of sentences-in-progress. *Language in Society*, 20, 441–458. DOI: 10.1017/S0047404500016572.
- 1996. On the “semi-permeable” character of grammatical units in conversation: conditional entry into the turn space of another speaker. In E. Ochs, E. A. Schegloff & S. A. Thompson (eds) *Interaction and grammar*. Cambridge: Cambridge University Press, 238–276.
- 2002. Turn-sharing: the choral co-production of talk-in-interaction. In C. E. Ford, B. A. Fox & S. A. Thompson (eds) *The language of turn and sequence*. Oxford: Oxford University Press, 225–256.
- 2004. Collaborative turn sequences. In G. H. Lerner (ed.) *Conversation analysis: studies from the first generation*. Amsterdam: John Benjamins, 225–256.
- Lilja, N. 2010. *Ongelmista oppimiseen: toisen aloittamat korjausjaksot kakkoskielisesä keskustelussa*. Jyväskylä Studies in Humanities 146. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-4116-1>.

- Long, M. H. 1996. The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (eds) *Handbook of second language acquisition*. New York (N. Y.): Academic Press, 413–468.
- McCarthy, M. 1998. *Spoken language and applied linguistics*. Cambridge: Cambridge University Press.
- 2010. Spoken fluency revisited. *English Profile Journal*, 1 (1), 1–15. DOI: 10.1017/S2041536210000012.
- Mori, J. & M. Hayashi 2006. The achievement of intersubjectivity through embodied completions: a study of interactions between first and second language speakers. *Applied Linguistics*, 27, 195–219. DOI: 10.1093/applin/amlo14.
- Olsher, D. 2004. Talk and gesture: the embodied completion of sequential actions in spoken interaction. In R. Gardner & J. Wagner (eds) *Second language conversations*. London: Continuum, 221–245.
- Ono, T. & E. Yoshida 1996. A study of co-construction in Japanese: we don't finish each other's sentences. In N. Akatsuka, S. Iwasaki & S. Strauss (eds) *Japanese/Korean linguistics, vol. 5*. Stanford (Cal.): CSLI Publications, 115–129.
- Peltonen, P. 2017. Temporal fluency and problem-solving in interaction: an exploratory study of fluency resources in L2 dialogue. *System*, 70, 1–13. DOI: 10.1016/j.system.2017.08.009.
- Peltonen, P. & P. Lintunen 2016. Integrating quantitative and qualitative approaches in L2 fluency analysis: a study of Finnish-speaking and Swedish-speaking learners of English at two school levels. *European Journal of Applied Linguistics*, 4, 209–238. DOI: 10.1515/eujal-2014-0018.
- Pickering, M. J. & S. Garrod 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–226. DOI: 10.1017/S0140525X04000056.
- Riggenbach, H. 1991. Toward an understanding of fluency: a microanalysis of nonnative speaker conversations. *Discourse Processes*, 14 (4), 423–441. DOI: 10.1080/01638539109544795.
- Rühlemann, C. 2007. *Conversation in context: a corpus-driven approach*. London: Continuum.
- Sato, M. 2014. Exploring the construct of interactional oral fluency: second language acquisition and language testing approaches. *System*, 45, 79–91. DOI: 10.1016/j.system.2014.05.004.
- Schegloff, E. A., G. Jefferson & H. Sacks 1977. The preference for self-correction in the organization of repair in conversation. *Language*, 53, 361–382. DOI: 10.2307/413107.
- Segalowitz, N. 2010. *The cognitive bases of second language fluency*. New York (N. Y.): Routledge.
- 2016. Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics in Language Teaching*, 54, 79–95. DOI: 10.1515/iral-2016-9991.
- Skehan, P. 2003. Task-based instruction. *Language Teaching*, 36, 1–14. DOI: 10.1017/S026144480200188X.
- 2009. Modelling second language performance: integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30, 510–532. DOI: 10.1093/applin/amp047.

- Skehan, P. 2014. The context for researching a processing perspective on task performance. In P. Skehan (ed.) *Processing perspectives on task performance*. Amsterdam: John Benjamins, 1–26.
- Stivers, T., N. J. Enfield, P. Brown, C. Englert, M. Hayashi, T. Heinemann, G. Hoymann, F. Rossano, J. P. de Ruiter, K.-E. Yoon & S. C. Levinson 2009. Universals and cultural variation in turn taking in conversation. *PNAS (Proceedings of the National Academy of Sciences)*, 106 (26), 10587–10592. DOI: 10.1073/pnas.0903616106.
- Taguchi, N. 2014. Development of interactional competence in Japanese as a second language: use of incomplete sentences as interactional resources. *The Modern Language Journal*, 98, 518–535. DOI: 10.1111/j.1540-4781.2014.12087.x.
- Tannen, D. 1989. *Talking voices: repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.
- Tavakoli, P. 2016. Fluency in monologic and dialogic task performance: challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics in Language Teaching*, 54, 133–150. DOI: 10.1515/iral-2016-9994.
- Tavakoli, P. & P. Skehan 2005. Strategic planning, task structure, and performance testing. In R. Ellis (ed.) *Planning and task performance in a second language*. Amsterdam: John Benjamins, 239–273.
- Ur, P. 1990. *Discussions that work: task-centered fluency practice*. Cambridge: Cambridge University Press.
- Witton-Davies, G. 2014. *The study of fluency and its development in monologue and dialogue*. PhD thesis. Lancaster University. [http://www.forex.ntu.edu.tw/en/files/writing/4092\\_dco088cd.pdf](http://www.forex.ntu.edu.tw/en/files/writing/4092_dco088cd.pdf).
- Wood, D., J. S. Bruner & G. Ross 1976. The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 17, 89–100. DOI: 10.1111/j.1469-7610.1976.tb00381.x.
- Wright, C. & P. Tavakoli 2016. New directions and developments in defining, analyzing and measuring L2 speech fluency. *International Review of Applied Linguistics in Language Teaching*, 54, 73–77. DOI: 10.1515/iral-2016-9990.

## Appendix: Transcription conventions

- : Extended or stretched sound, syllable or word.
- + Vocalic emphasis.
- ( . ) Micropause. A pause of less than 0.25 seconds.
- ( 1 . 21 ) Timed pause. A pause of 0.25 seconds or longer.
- um, uh A non-lexicalized filled pause.
- \*heh\* A separate laugh syllable (cf. chuckling talk below).
- \$ \$ Laughing/chuckling talk between markers.
- \*pt\* Lip smack. Included in silent pause time measures, marked with curly brackets { \*pt\* \_0.87 }.
- ◦ A passage of talk noticeably softer than surrounding talk.
- Halting, abrupt cut off of sound or word.
- = Latching of contiguous utterances, with no interval or overlap. (No clear pause between speakers' utterances.)
- [ ] Speech overlap.
- \* CAPS Non-verbal behavior, e.g. gestures, gaze.

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 139–162.

**Iina Salminen**

Tampereen yliopisto

## **Erään venäjänopiskelijan harjoittelu ja kehitys venäjän kielen sibilanttien ja affrikaattojen parissa**

The paper in question reports an experiment that examines the difficulties of acquiring Russian sibilant and affricate phonemes faced by a Finnish L2 learner of Russian. The experiment is based on the theoretical framework of explicit pronunciation teaching and that of contrastive analysis. The experiment itself consists of three parts: 1) phonetic testing in three phases, including tasks of both receptive and productive spoken language skills, 2) a short lesson about the features and pronunciation of the sibilants and affricates of Russian, conducted by the author, 3) independent practice by the experimentee. The analysis concentrates on finding similarities and differences between a) the pronunciation target and the sounds produced by the experimentee and b) the sounds produced in different phases. The results show that there is little improvement in the reception skills of the experimentee (6 % increase between the first and the last test), but more so in their production skills (16 %). Interestingly, while most of the mistakes made by the experimentee were anticipated on the basis of a contrastive analysis between Russian and Finnish, some tasks proved to be significantly easier than expected, for example the recognition of the word-initial *s /z/*.

**Keywords:** Russian as second language, pronunciation teaching, Russian phonetics, contrastive analysis

**Asiasanat:** Venäjä vieraana kielenä, ääntämisen opettaminen, venäjän fonetiikka, kontrastiivinen analyysi

## 1 Johdanto

Suomalaisen koulun kontekstissa vieraan kielen tunneilla pääpaino on tavallimmin kirjoitetun kielen taidoissa: kieliopissa, sanastossa, syntaksissa, lukemisessa. Formaalin opetuksen ohella painotus voidaan havaita myös suomalaisten taidoissa. Pekka Lintunen (2005: 375) kokoaa yhteen muiden tutkijoiden tuloksia todeten, että erinomaisesta englannin kielen osaamisesta huolimatta suullinen kielitaito, ääntäminen eritoten, koetaan kaikkein heikoimmaksi kielen osa-alueeksi. Suullisen kielitaidon täytyisi kuitenkin olla paljon merkittävämmässä roolissa opetuksessa, sillä esimerkiksi venäjäksi tapahtuvasta tosielämän viestinnästä 42 % on kuulemista, 26 % puhumista ja vasta jäljelle jäävät 32 % jakautuvat lukemisen ja kirjoittamisen kesken (Fedotova 2013: 78).

Tässä artikkelissa raportoitu tutkimus syventyy suullisen kielitaidon perusrakennuspalikoihin: äänneiden tuottamiseen ja havaitsemiseen eli suullisen kielitaidon teknisiin taitoihin. Artikkelissa tutkitaan venäjän kielen sibilanttien ja affrikaattojen (taulukko 1) eksplisiittisen opettamisen vaikutuksia oppimistuloksiin sekä kontrastiivisen kielitieteen tarjoamia mahdollisuuksia eksplisiittisessä ääntämisopetuksessa. Termillä ääntämisopetus viitataan niin tuottamiseen kuin havaitsemiseen, ts. suullisen kielitaidon teknisiin taitoihin. Tutkimuksella halutaan myös selvittää, ovatko venäjän sibilantit ja affrikaatat todella niin vaikeita suomalaiselle oppijalle, kuin usein kuulee sanottavan. Vaikka kyseiset äänneet yleensä nostetaan esiin yhtenä venäjän ääntämisen vaikeimmista osa-alueista suomenkieliselle oppijalle, on empiiristä tutkimusta aiheesta tehty Suomessa vähän. Kontrastiivisen venäjän ääntämisen tutkimuksen saralla on tutkittu mm. suomenkielisten venäjänoppijoiden prosodisten taitojen kehitystä (Ullakonoja 2011) sekä venäläismaahanmuuttajien suomen ääntämisestä (Aho & Toivola 2008; Aho ym. 2016b).

Artikkelin tutkimuskysymykset ovat:

1. Ovatko jotkin venäjän sibilanteista tai affrikaatoista selkeästi muita vaikeampia tai helpompia havaita ja/tai tuottaa?
2. Onko reseptiivisen ja produktiivisen osaamisen välillä yhteys, ts. ilmenekö foneemin helppous tai vaikeus sekä reseptiivisessä että produktiivisessa tehtävässä vai vain toisessa?
3. Vaikuttaako ääntämisen eksplisiittinen opetus ja harjoittelu venäjän si-

TAULUKKO 1. Venäjän sibilantit ja affrikaatat.

grafeemi	с	з	ш	ж	ч	ц	щ
foneemi	s	z	ʃ	ʒ	t͡ʃ	t͡s	ʃ:

bilanttien ja affrikaattojen tuottamis- ja havaitsemistarkkuuteen?

## 2 Ääntämisen opetus ja sen asema suomalaisessa koulussa

Venäläiset vieraan kielen didaktiikan tuntijat Moskovkin & Šukin (2012: 302, 306–307) sanovat osuvasti, että kielen äänteiden tuottamisessa ja tunnistamisessa on kyse yhden ilmiön kahdesta eri puolesta, joita ei ole mielekästä erottaa opetuksessa, eiväthän ne ole erillään todellisissa kielenkäyttötilanteissa. On tärkeää harjoitella molempia taitoja samanaikaisesti, jotta reseptiiviset ja produktiiviset taidot tukisivat toinen toistaan. Lisäksi on tärkeä huomata, että havainnoinnin taidot edeltävät lähes aina tuottamisen taitoja. (Moskovkin & Šukin 2012: 306–307; Bernštejn 1975: 11; Kiparsky & Menn 1977: 49.) Silti monesti todetaan, että etenkin kuullun ymmärtämistä harjoitellaan liian vähän tai riittämättömän äänimateriaalin avulla – tämä siitäkkin huolimatta, että kielen kuuleminen on välttämätöntä sen oppimiseksi. Yksin opettajan kohdekielinen puhe ei kata kielen kuulemisen tarpeita, ja usein tavallisissa kielenoppimistilanteissa oppija ei yksinkertaisesti kuule tai erota kuulemastaan opettajan puheesta yksittäisten äänteiden tunnuspiirteitä. (Akišina & Kagan 2014; Fedotova 2013: 79; Lintunen 2014b: 172; Iivonen 2005: 49.) Myös äänteiden havainnoinnin ja tuottamisen mekanismeja ja oppimista käsitteleviin teorioihin (Best 1995: ks. Flege 1995; Kuhl 2000) nojaten vieraiden kielten opetuksessa on sijaa eksplisiittiselle suullisen kielitaidon opetukselle. Näiden teorioiden pohjalta voi yleistäen sanoa äidinkielen vaikutuksen olevan sekä produktiivisella että reseptiivisellä osa-alueella niin suuri, että ilman kohdekielen ääntämisen erikoisuuksiin paneutumista oppija ei välttämättä tiedosta näitä erikoisuuksia, mistä johtuen puheen oikein tunnistamisen ja tuottamisen taso jää pyrittyä alhaisemmaksi.

Kommunikatiivisen kieltenopetuksen suuntauksen noustua hallitsevaksi paradigmaksi 1970-luvulla se jätti ääntämisen opetuksen ”tarpeettomaksi”, olihan kielenkäytön päätavoitteena pelkistäen sanoen selvitä viestintätilanteesta miettimättä sitä, oliko käytetty kieli virheetöntä ja norminmukaista vai ei (Tergujeff 2016: 221). Tosin Harjanteen ja Tellan (2010) kiinnostavan meta-analyysin mukaan yli 30 vuoden aikana muutoksia kommunikatiiviseen suuntaan on opetuksessa ollut vain vähän huolimatta siitä, että Suomessakin alettiin virallisissa dokumenteissa painottaa kommunikatiivista näkökulmaa. Esimerkiksi kirjallisia tehtäviä suositaan edelleen suullisten kustannuksella ja tavallisesti opetuskieli ja opetettava kieli tarkoittavat eri asiaa (mp.).

Kommunikatiivinen suuntaus kielenopetuksessa on tietenkin näkynyt kansallisissa opetussuunnitelmissa, kaikkein selkeimmin viimeisimmissä. Esimerkiksi lukion B3-kielen kolmen ensimmäisen kurssin keskeisiä tavoitteita ovat



jokapäiväisissä suullisissa viestintätilanteissa selviytyminen sekä tottuminen kohdekielen ääntämiseen (Opetushallitus 2015: 115, 117). Tässä on yksi väkevä peruste suullisen kielitaidon huomioimiseen opetuksessa, onhan opetussuunnitelma tärkein opettajan työtä ohjaava asiakirja. Opetussuunnitelman ohella on myös muita painavia syitä. Vaikka suullinen kommunikaatio sisältääkin universaaleja, helpommin opittavia osa-alueita kuten elekielen, ovat kielen foneemirepertuaari ja sen tuottaminen aina tyypillisiä piirteitä vain kullekin yksittäiselle kielelle, mistä syystä sen hallintaa on tarpeen tietoisesti opetella (Lintunen 2014b: 165). Venäjän fonetiikan kohdalla yksi merkittävimpiä kohdekielen omintakeisia piirteitä ovat suomesta puuttuvat konsonanttien liudennus sekä yksittäiset äänneet kuten affrikaatat. Lisäksi, koska suomi on foneeminen kieli eli sen kirjoitus- ja ääntötavalla on korkea vastaavuus, suomalainen oppija hyötyy merkittävästi eksplisiittisestä opetuksesta, jos kohteena on kieli, jonka kirjoitus- ja ääntötavan vastaavuus ei ole yhtä korkea (Flege 1988: 279, 311; Lintunen 2014a: 1–3, 2014b: 175). Foneemigrafeemivastaavuuden osalta venäjän sanapainon aiheuttamat erot ääntö- ja kirjoitustavan välillä sekä assimilaatiosäännöt ovat suurimpia kompastuskiviä.

Sopivia ääntämisen opetusmetodeja valittaessa tulee huomioida oppijan ikä. Siinä missä pienet lapset oppivat parhaiten jäljittelemällä ja imitoimisella, varttuneemmat oppijat hyötyvät analyttisemmästä tavasta, esimerkiksi lähtö- ja kohdekielen foneemien ääntöpaikkojen- ja tapojen vertailusta (Aho ym. 2016a: 48; Lintunen 2014b: 184; Temple 2005: 193) – toisin sanoen eksplisiittisestä opetuksesta. On myös jyrkästi esitetty, ettei imitaatio voi johtaa tyydyttäviin oppimistuloksiin, ellei oppijalla sitten ole ”erityisiä foneettisia lahjoja” (Bernštejn 1975: 7–10).

Tässä yhteydessä en voi olla kertomatta omaan opettajakokemukseeni pohjaavaa esimerkkiä, joka korostaa ääntämisopetuksen roolia. Lukion venäjän ensimmäisen kurssin opiskelijat suorittivat tehtävää, jossa joukko adjektiiveja piti yhdistää substantiiveihin sukukategorioiden mukaan. Eräs pari yhdisti sanan *нос* [nos] ’nenä’, maskuliini, feminiinipäätteeseen adjektiiviin. Kysyessäni perustelua he vastasivat: ”Eikö tämä sana tarkoitakin ’yötä’? Se on feminiini.” Venäjäksi ’yö’ todella on feminiini, mutta sana on *ночь* [nočʲ]. Se siis eroaa sanasta ’nenä’ vain yhden äänneen suhteen, mutta opiskelijoiden sekoittaessa kirjoitetun kielen symbolit *c-y* ja sen takia myös äänneet /s-ʧ/ syntyi virhe myös kieliopin alueella. Vain yhden esimerkin perusteella ei voi tehdä joihinkin päätöksiä, mutta voisi varovaisesti spekuloida, että koska venäjän kohdalla merkitykset voivat helposti sekoittua, jos kirjainsymbolia ei osata yhdistää oikeaan äänneeseen, myös kirjallisen kielitaidon kohdalla foneettiset taidot ovat oleellisessa asemassa.

Ääntäminen, etenkin sen eksplisiittinen käsittely, näyttää yleensä jäävän opetuksessa vähälle huomiolle. Missä lienee syy tähän painotukseen? Peruskoulun opetussuunnitelmassa ääntämiseen puututaan melko pintapuolisilla

lauseilla, vaikkakin tilaa ääntämiselle ja sen eksplisiittiselle opetukselle toki annetaan: ”[kiinnitetään] huomiota myös ääntämiseen”, ”havainnoidaan ja harjoitellaan runsaasti ääntämistä sekä sana- ja lausepainoa, puherytimiä ja intonaatiota”, ”havainnoidaan kohdekielen rytmiä, intonaatiota ja muita ääntämisen piirteitä ja harjoitellaan ääntämään luontevasti” (Opetushallitus 2014: 361, 225). Lukion opetussuunnitelma puolestaan viittaa tavoitteissa eurooppalaisen viitekehyksen (The Common European Framework of Reference for Languages, CEFR) taitotasokuvauksiin, mutta nekin ovat melko tulkinnanvaraisia. Esimerkiksi B3-oppimäärän tavoitteena on puhe, jonka ääntäminen on ymmärrettävää, jossa vierasperäinen korostus voi olla selkeä ja jossa voi olla ääntämisvirheitä. (Opetushallitus 2015.) Jää siis opettajan ammattitaidon varaan tulkita, minkälaista ääntämistä opetussuunnitelmassa tarkoitetaan.

Lopuksi on mainittava eräs lukiokontekstissa merkittävä tekijä, joka ennen kaikkea vaikuttaa suullisen kielitaidon ja sitä kautta sen eksplisiittisen opettamisen vähäiseen osuuteen: ylioppilaskirjoitukset. Vaikka ylioppilaskoe sisältää kuullun ymmärtämistä testaavan osion, ei koe (toistaiseksi) testaa produktiivista suullista kielitaitoa ollenkaan. Päinvastoin, koe edelleen painottaa kirjallisen kielen, etenkin kieliopin, hallintaa. Koska kokeesta luonnollisesti tavoitellaan korkeaa arvosanaa ja koska kokelaan jatko-opintomahdollisuudet usein riippuvat osittain myös ylioppilaskokeista (tulevaisuudessa viimeisimmän keskustelun mukaan jopa enenevässä määrin), opettaja joutuu taipumaan realiteettien edessä: Jos produktiivista suullista kielitaitoa ei testata, siihen ei kannata tuhata muutenkin rajallisia resursseja – tärkeämpää on keskittyä kokeen painotamiin asioihin hyvien arvosanojen tavoittelemiseksi. Etenkin ylioppilaskokeen rakenteeseen liittyen on herkkua pitää mielessä aiemmin mainittu kielitaidon eri osa-alueiden jakautuminen venäjänkielisissä viestintätilanteissa, joissa suullista kielitaitoa tarvitaan 68 % tilanteista (Fedotova 2013: 78). Näitä lukuja vasten ylioppilaskokeen rakenne ja painotukset ovat kaukana siitä, että ne vastaisivat tosielämän kielenkäyttötarpeita. On enemmän kuin aiheellista kysyä, miksi näin on. Suullisen kielitaidon osuutta pitäisi kielen tunneilla ja vieraan kielen ylioppilaskokeessa nostaa merkittävästi<sup>1</sup>, jotta saatu lukiokoulutus todella hyödyttäisi opiskelijoita ja siten laajemmin myös yhteiskuntaa.

### 3 Kontrastiivinen ote kielenopiskelussa ja empiirisen tutkimuksen hypoteesit

Kontrastiivisella kielitieteellä on aina ollut soveltava painotus, ja jo Robert Laddon yksi kontrastiivisen tutkimuksen pioneeriteoksista *Linguistics across cultures* (1957) on kielenopettajille suunnattu toimintatapojen muutokseen täh-

<sup>1</sup> Kuitenkin kielen kirjallisen hallinnan kustannuksella niin, ettei kokelaan kokonaistyömäärä (jo kurssienkin aikana) kasva kohtuuttomaksi.

TAULUKKO 2. Suomen ja venäjän sibilantit ja affrikaatat ääntötavan ja -paikan mukaan. Suomen äänneet erotettu **sinisellä**.

	dentaalit	alveolaarit	post- alveolaarit	palato- alveolaarit
affrikaatat	ʈʂ			ʈʂ
frikatiivit	s z	s	ʃ	ʃ ʂ ʐ

täävä opas opettamiseen ja testaamiseen. Kuten jo aiemmin todettu, lapsuusiän ohittaneet oppijat hyötyvät matkimisen ja toistamisen ohella myös analyttisemmästä lähestymistavasta. Kontrastiivinen kielitiede tarjoaa mainioita työkaluja analyttiseen työskentelyyn, ja monet tutkijat ovatkin kirjoittaneet sen sovellutuksista vieraan kielen opetuksessa ja opiskelussa. On todettu muun muassa, että se voi auttaa äänneiden samankaltaisuuksien ja eroavaisuuksien hahmottamisessa, oman vieraskielisen puheen kontrolloinnissa ja fonologisen kuulon kehityksessä sekä ääntämisessä tehtävien virheiden selittämisessä ja ymmärtämisessä (Aho ym. 2016a: 48; Vasil'ev 1975: 104; Lado 1957: 2-4; James 1980: 148). Kohdekielen äänneiden kohdalla kontrastiivinen analyysi, jossa lähtö- ja kohdekielen äänneitä vertaillaan ääntöpaikan ja -tavan mukaan, voi osoittautua hyvinkin hedelmälliseksi. Esimerkiksi venäjän tapauksessa analyysin avulla voidaan oppijoiden kanssa todeta, ettei yksikään suomen ja venäjän foneemirepertuaarien sibilanteista täysin vastaa toisiaan, minä kautta näiden oletettavasti vaikeasti opittavien äänneiden harjoitteluun tulee kiinnittäneeksi huomiota jo alusta alkaen. Tällä tavoin eksplisiittisesti opettaen oppijat saattavat myös ymmärtää, että kyseessä on tärkeä kielen osa-alue siinä missä sanasto ja kielioppikin ja että näitä kaikkia tulee tietoisesti harjoitella.

Vertailemalla suomen ja venäjän sibilantteja ja affrikaattoja (taulukko 2) voidaan tehdä seuraavia havaintoja:

1. Kummassakin kielessä on frikatiiveja eli saman ääntötavan äänneitä...
2. ...kun taas ääntöpaikan suhteen samankaltaisia äänneitä ei ole.
3. Näin ollen voidaan todeta, ettei kielillä ole keskenään identtisiä äänneitä.
4. Suomessa ei ole lainkaan affrikaattoja.
5. Toisiinsa usein rinnastetut suomen ja venäjän /s/:t sekä suomen /ʃ/ ja venäjän /ʃ/ eivät täysin vastaa toisiaan artikulatorisilta ominaisuuksiltaan.

Havainto 3 on tärkeässä roolissa. Kontrastiivisen analyysin vahvaan versioon nojaten voidaan ennustaa, että kaikki venäjän sibilantit ja affrikaatat tulevat

olemaan vaikeita oppia. Analyysin heikon eli vähemmän jyrkän version mukaan havainto 3 selittää mahdollisia äänneissä tehtäviä virheitä: kun kieltenväliset äännevastineet puuttuvat, oppija ei pysty hyödyntämään transferenssia eli siirtovaikutusta eli suoraan käyttämään suomen äänneiden /s, ʃ/ tietämystään venäjän parissa.

Lyhyesti sanoen kontrastiivisen analyysin pohjalta hypoteesina on, että jokainen tutkimuksen äänne voi osoittautua hankalaksi. Kuitenkin, vaikka mikäään venäjän sibilanteista tai affrikatoista ei olekaan identtinen suomen äänneiden kanssa, voidaan odottaa, että ”perusässäksi” yleensä nimitettävä, terävä dentaalinen /s/ sekä soinnilliset /z, zʃ/ tunnistetaan muista äänneistä helpoiten, sillä ne poikkeavat akustisilta ominaisuuksiltaan eniten: /s/ on ”siisti”, terävä äänne ilman soinnillisuutta tai alveolaarisen ääntötavan tuomaa suhinaa, /z/ ja /zʃ/ ovat molemmat soinnillisia eli suomalaiselle vieraan kuuloisia, ja niitä erottaa dentaalisen ja palatoalveolaarisen ääntötavan terävyyden ja suhuäänteen vastakkainasettelu. Lisäksi affrikaatat /tʃ, tʃs/ erotettaneen frikatiiveista helposti erilaisen ääntötapansa vuoksi. Sen sijaan ominaisuuksiltaan toisiaan muistuttavien affrikaattojen /tʃ, tʃs/ keskinäisessä tunnistamisessa voi ilmetä paljon virheitä, samoin palatoalveolaaristen äänneiden /ʃ, ʒ, zʃ/ erotamisessa toisistaan. Tuottamisessa puolestaan odotetaan tapahtuvan enemmän virheitä kuin tunnistamisessa, sillä vaikka seitsemästä venäjän sibilantista ja affrikaatasta peräti kuusi löytyy koehenkilölle tutuista länsi-eurooppalaisista kielistä, on niiden tuottaminen jokseenkin vierasta – varsinkin kyrillistä aakkostoa luettaessa. Lisäksi produktiivista suullista kielitaitoa on harjoiteltu koehenkilön kurssilla vähänlaisesti. Tuottamisessa affrikaattojen /tʃ, tʃs/ välille odotetaan sekaannuksia, sillä kokemus osoittaa näiden kahden äänneen ääntämisen hahmottamisen olevan vaikeaa suomalaiselle venäjänoppijalle. Lisäksi on odotettavissa, että soinnillisten äänneiden /z, zʃ/ soinnillisuus jää puuttumaan, jolloin ne muistuttavat soinnittomia parejaan /s, ʃ/. Myös suomelle vieras dentaalinen ääntöpaikka voi osoittautua toistuvaksi kompastuskiveksi tuottamisessa.

Pitkittäistutkimuksen näkökulmasta hypoteesina on, että äänneiden eksplisiittinen opetus ja harjoittelu parantavat tuottamisen ja tunnistamisen tarkkuutta, mikä puoltasi eksplisiittistä ääntämisopetusta.

## 4 Aineisto ja analyysimenetelmät

Artikkelin aineisto on kerätty yhdeltä lukion B3-venäjän opiskelijalta. Aineisto koostuu kolmesta osasta: 1) Alku-, väli- ja lopputestauksesta, joilla mitattiin äänneiden tuottamis- ja havainnointitarkkuutta, 2) opetustuokiosta ja 3) itsenäisestä harjoittelusta. Aineistonkeruu toteutettiin vuoden 2016 aikana. Toukokuussa, jolloin tutkittava oli lukion ensimmäisellä vuosikurssilla, suori-

tettiin alkutestaus. Jo samana päivänä pidettiin opetustuokio, jossa kirjoittajan johdolla perehdyttiin äänteiden artikulatorisiin ominaisuuksiin ja niiden muun kielisiin vastineisiin. Lisäksi harjoiteltiin äänteiden tuottamista erilaisin harjoituksin. Opetustuokio pidettiin vain kerran, eikä koehenkilön ollut tarkoitus kerrasta oppia ääntämään kaikki tutkimuksen äänteet; ennemminkin tarkoituksena oli tietoisesti rakentaa pohja koehenkilön itsenäiselle harjoittelulle. Lopuksi koehenkilölle annettiin itsenäistä harjoittelua varten pieni harjoitusehdotuksia sisältänyt kirjoittajan kokoama vihkonen, venäjän sibilanttien ja affrikaattojen grafeemit korteille kirjoitettuna toiminnallisempia harjoituksia varten sekä itsenäisen harjoittelun seurantalomakkeet. Tuottamista ehdotettiin harjoiteltavan mm. lukemalla suomenkielistä tekstiä venäläisiä sibilantteja ja affrikaattoja käyttäen, omaa ääntämistä äänittämällä ja kuuntelemalla sekä opettamalla äänteitä jollekulle venäjää taitamattomalle. Havainnoinnin harjoittelua varten annettiin mm. musiikkivinkkauksia sekä linkki kirjoittajan tekemälle verkkosivulle (ks. liite), jolla eri äänteiden akustisia ominaisuuksia pääsi vertailemaan pseudo- ja oikeista sanoista koostuvia foneettisia minimipareja kuuntelemalla.

Itsenäinen harjoittelu sijoittui kesälomalle ja uuden lukuvuoden alkuun. Tänä aikana koehenkilö piti kirjaa harjoittelustaan kuvaten lyhyesti harjoituksen sisältöä ja arvioitua pituutta. Vaikka kirjanpito ”välillä unohtui”, kuten koehenkilö välitestauksessa kertoi, antoivat seurantalomakkeet riittävän kuvan harjoittelun luonteesta. Kesäloman alusta koulun ensimmäisille viikoille koehenkilö harjoitteli 16 päivänä. Tämän lisäksi hän kuukauden mittaisen Yhdysvaltoihin suuntautuneen englannin kielikurssin aikana tutustui venäjänkielisiin opiskelijoihin, joilta hän aina itselleen uuden venäjän sanan kohdatessaan kysyi, mikä sibilantti tai affrikaatta sanassa kuuluu. Harjoitukset kestivät 5–30 minuuttia kerrallaan ja koostuivat pääasiassa äänteiden itsekseen toistelusta. Koehenkilö oli myös hyödyntänyt foneettisten minimiparien kuuntelua ja suomenkielisen tekstin lukemista venäläisiä äänteitä käyttäen. Lisäksi hän oli opettanut äänteet tuttavalleen ja kertoikin välitestauksen yhteydessä oppivansa parhaiten silloin, kun toimii itse opettajan roolissa. Vapaaehtoisuuteensa nähden koehenkilö työskenteli itsenäisen harjoittelun aikana varsin tunnollisesti ja ansiokkaasti, mikä kieli sisäisestä oppimismotivaatiosta.

Välitestaus suoritettiin elokuun lopussa ja lopputestaus joulukuussa. Alku- ja välitestin aikana venäjä lukeutui tutkittavan lukujärjestykseen (3 × 75 min oppituntia viikossa) ja hän käytti kieltä jonkin verran vapaa-ajallaankin, kun taas väli- ja lopputestin väliin sijoittui muutaman kuukauden venäjätön ajanjakso niin opintojen kuin vapaa-ajan puolella.

Testeillä mitattiin sekä reseptiivistä että produktiivista foneettista tarkkuutta tunnistus- ja tuottamistehtävillä. Tunnistus- ja tuottamistehtäviä oli vuorotellen tutkittavan mielenkiinnon ylläpitämiseksi. Tunnistamista testattiin foneettisten minimiparien havaintotehtävällä. Tutkittava kuunteli yksittäisiä

sanoja äänitteeltä ja merkitsi vastauslomakkeelle, kumman lomakkeella lukevan minimiparin sanoista hän kuuli. Äänitteen lukijana toimi syntyperäinen venäjän puhuja. Sanat luettiin kahdesti peräkkäin, yhteensä sanoja oli 84, joista osa oli oikeita ja osa pseudosanoja. Jokaista minimiparia testattiin kahdesti siten, että eri tehtävissä joka minimiparin eri sanat (ja siten eri äänneet) tulivat testatuiksi. Jokainen testattu äänne esiintyi sekä sanan alussa, että keskellä. Suomen ja venäjän sibilanttien ja affrikaattojen välisen kontrastiivisen analyysin ennakoimien oppimisvaikeuksien ja hypoteesin pohjalta testattiin seuraavia minimipareja:

s - ʃ s - z z - ʒ ʦ - ʧ ʧ - ʃ ʃ - ʧ:

Tuottamista testattiin lukupuhuntatehtävällä, jossa tutkittava luki yhteensä 42 venäjänkielistä lausetta. Kaikki seitsemän testattavaa äännettä esiintyivät myös lukupuhuntatehtävän lauseissa sekä sananalkuisissa- että keskisissä asemissa. Lukupuhuntatehtävässä tuotetut äänneet analysoitiin Praat-ohjelman avulla annotoinnin ja oskillogrammin tarjoamaa visuaalista apua hyväksi käyttäen. Kun äänneitä luokiteltiin tarkempaa analyysiä varten, oleellisinta oli keskittyä eri äänneiden tunnistettavuuteen toisistaan eli Flegen (1988) termistöä käyttäen ääntämisen riittävyteen. Esimerkiksi affrikaatoille ominaisen täyden sulun esiintyminen sekä sibilanteille binäärinen ominaisuus soinnillisuus-soinnittomuus nousivat olennaiseen rooliin. Vaikka venäjässä myös liudentuneisuus-liudentumattomuus on hyvin keskeinen binäärinen ominaisuus, jätettiin sibilanttien liudentuneet parit tässä tutkimuksessa kokonaan huomiotta sillä perusteella, että liudennus fonologisine sääntöineen edustaa kirjoittajan mielestä jo sellaista edistyksestä venäjän kielen taitoa, johon B3-oppimäärän aikana on erittäin vaikea pyrkiä ääntämisen ollessa opeudessa muutenkin heikossa asemassa kirjalliseen kielitaitoon nähden. Poikkeuksen tähän luovat tietenkin tutkimukseen sisältyvät aina liudentuneet äänneet /ʦ, ʧ:/.

Tuloksista ja johtopäätöksistä lukiessa on syytä ottaa huomioon äänneiden /s, ʃ/ analysointia koskeva seikka. Kuten yllä on todettu, jonkinasteiset analogiat keskenään muodostavat suomen /s/ ja venäjän /s/ sekä suomen /ʃ/ ja venäjän /ʃ/ eivät ole keskenään täysin identtiset. Lukupuhuntatehtävässä koehenkilön tuottamia äänneitä luokitellessa ei kuitenkaan pystytty saavuttamaan niin tarkkaa analyysin tasoa, jolla nämä rinnastetut äänneet olisi kyetty erottaa toisistaan. Suomen ja venäjän /s/-äänneiden kohdalla tehtävää osaltaan vaikeutti se, että koehenkilön suomenkielisessä spontaanissa puheessa /s/ yleensä realisoitui alveolaarista äännettä terävämmin ja sijoittui jonnekin molempien kielten /s/-äänneiden välimaastoon. Näistä syistä myös sellaiset koehenkilön tuottamat äänneet, jotka kuulostivat suomen alveolaariselta äänneeltä /s/ tai sen ja venäjän /s/:n välimuodoilta, luokiteltiin venäjän dentaliseksi äänneeksi /s/. Myös suomalaisittain eli postalveolaarisesti tuotettu

/ʃ/ luokiteltiin venäjän palatoalveolaariseksi äänneeksi /ʃ/. Menettely osaltaan vääristää tuloksista näkyvää, Flegen termistöä käyttäen koehenkilön ääntämisen autenttisuutta eli venäjän foneemien ääntönormien noudattamisen tarkkuutta. Kommunikatiivisuutta korostavan opetussuunnitelman ja oppimisen tavoitteiden hengessä tärkeimpänä kriteerinä näidenkin äänneiden luokittelussa oli kuitenkin ääntämisen riittävyys eli se, että äänneen tunnistaa selkeästi muista kohdekielen äänneistä, mikä vähentää väärinymmärretyksi tulemisen vaaraa.

Analyysissä keskitytään eroihin ja yhtäläisyyksiin a) suhteessa tuotto- tai tunnistustavoitteeseen ja b) eri testikertojen tulosten välillä. Oleellisinta on tarkkailla, mitkä äänneet näyttäytyvät ongelmallisimpina ja muuttuvatko havaitsemis- ja tuottamistarkkuus testikertojen välillä, ts. vaikuttavatko eksplisiittinen opetus ja harjoittelu tuloksiin. Lisäksi mielenkiinnon kohteena on, ovatko vaikeat äänneet vaikeita sekä havainnoinnin että tuottamisen kannalta vai painottuvatko virheet vain toiseen osa-alueeseen.

## 5 Analyysi ja tulokset

Seuraavassa osassa esitellään havainto- ja tuottamistestien tulokset taulukkoina ja nostetaan jokaiselta testikerralta esiin mielenkiintoisimpia huomioita.

### 5.1 Kevät 2016 – ensimmäinen testaus

**Havaintotehtävässä** koehenkilön tuli valita kahdesta vastauslomakkeella olevasta foneettisesta miniparista se, jonka kuuli äänitteeltä. Testattavat minimiparit valittiin paitsi kontrastiivisen analyysin, myös äänneiden akustisten ominaisuuksien perusteella: toisistaan akustisesti suuresti poikkeavat äänneparit kuten /z-ʃs/ jätettiin tunnistustehtävän ulkopuolelle resurssien säästämiseksi.

Taulukon 3 tuloksista käy ilmi, että 81 äänneestä oikein tunnistettuja äänneitä oli yhteensä 63 eli 78 %. Sanan alussa olevista äänneistä tunnistettiin oikein 76 %, kun taas sanan keskellä olevista 80 %. Kaikkiaan kevään havainnointitarkkuuden voi sanoa olevan hyvällä tasolla, kun otetaan huomioon vähäinen ääntämisen huomiointi venäjän opetuksessa sekä siihenastisten venäjänopin-  
tojen määrä (testauksen aikana käynnissä kolmas B3-oppimäärän kurssi). Aina oikein tunnistettuja äänneitä oli kaikkiaan seitsemän kappaletta, kun taas aina väärin tunnistettuja äänneitä ei ollut ollenkaan.

Sananalkuinen palatoalveolaarinen /ʃ/ asetettiin testissä vastakkain usean eri äänneen kanssa, sillä hypoteesin mukaan se on todennäköisesti helppo sekoittaa muihin saman ääntöpaikan äänneisiin. Hypoteesi osoittautui jokseenkin oikeaksi, kun puolet /ʃ/-äänneistä tunnistettiin väärin.

Kahdessa tapauksessa kuudesta soinnillinen /z/ sekoitettiin sanan alussa soinnilliseen /z/-äänneeseen. Dentaalisen ja palatoalveolaarisen ääntöpaikan

TAULUKKO 3. Kevään havaintotehtävässä havaitut äänneet. 100 % tarkkuus eroteltu. Viiva tarkoittaa, ettei kyseistä äänneparia testattu (esim. /z-s/).

luettu		havaittu								oikein %
		n	s	ʃ	z	ʒ	ʃs	ʃɕ	ɕ:	
sanan alussa	s	6	6	0	0	-	-	-	-	100
	ʃ	12	1	6	-	3	-	0	2	50
	z	6	0	-	4	2	-	-	-	67
	ʒ	5	-	0	0	5	-	-	-	100
	ʃs	3	-	-	-	-	2	1	-	67
	ʃɕ	6	-	1	-	-	0	5	-	83
	ɕ:	3	-	0	-	-	-	-	3	100
sanan keskellä	s	6	6	0	0	-	-	-	-	100
	ʃ	12	0	12	-	0	-	0	0	100
	z	5	0	-	5	-	-	-	-	100
	ʒ	6	-	1	3	2	-	-	-	33
	ʃs	3	-	-	-	-	3	0	-	100
	ʃɕ	5	-	1	-	-	3	2	-	40
	ɕ:	3	-	1	-	-	-	-	2	67

tuottaman erilaisen akustiikan erottaminen ei siis ollutkaan niin helppoa, kuin hypoteesissa oletettiin. Toisaalta /z/ tunnistettiin aina oikein. Mitään toista äännettä ei myöskään virheellisesti tunnistettu /z/-äänneenä, joten tästä näkökulmasta tämän äänneen voisi ehdottaa olevan helposti erotettavissa muista äänneistä.

Sananalkuisten affrikaattojen tunnistuksen tulokset ovat mielenkiintoiset. Kolmesta kuullusta /ʃs/-äänneestä kaksi tunnistettiin oikein ja yksi virheellisesti /ʃɕ/-äänneenä, mutta sen sijaan kuudesta kuullusta /ʃɕ/-äänneestä yhtäkään ei tunnistettu toisena affrikaattana /ʃs/. Voi olla, että kyseessä on samanlainen tendenssi kuin äänneiden /z-z/ tunnistamisen välillä: Dentaalinen äänne (/z-ʃs/) tunnistettiin joitain kertoja palatoalveolaarisena (/z-ʃɕ/), mutta palatoalveolaarista äännettä ei kertaakaan dentaalisena. Sanan alussa soinnillinen palatoalveolaari /z/ tunnistettiin aina oikein, mutta sanan keskellä äänne osoittautui hankalaksi tunnistettavaksi.

Alkuoletukseen nähden /ɕ:/ tunnistettiin yllättävänkin hyvin, sillä kyseessä on ääntöpaikan ja liudentuneisuuden takia suomalaiselle hyvin vieras äänne.

**Lukupuhuntatehtävässä** koehenkilön tuli lukea kohdekielisiä lauseita, joissa eri äänneet esiintyivät sekä sanan alussa, että keskellä. Taulukon 4 sarakke "epätarkka" viittaa sellaisiin tuottotapoihin, joiden akustiset ominaisuudet ovat jonkinlainen yhdistelmä kahta eri äännettä, esim. suomen äänneiden /t, s/ peräkkäinen ääntäminen venäjän affrikaatan /ʃs/ sijasta. 83 äänneestä oikein tuotettiin 55 %. Tarkkuus ei liittynyt äänneeseen asemaan sanassa: sanan



TAULUKKO 4. Kevään lukupuhuntatehtävässä tuotetut äänteet. 100 % tarkkuus eroteltu.

lauseessa		n	s	ʃ	z	ʒ	tuotettu			epätarkka	oikein %
							ʃs	ʃʒ	ʒ:		
sanan alussa	s	12	11	1	-	-	-	-	-	-	92
	ʃ	4	-	4	-	-	-	-	-	-	100
	z	6	3	-	3	-	-	-	-	-	50
	ʒ	5	1	3	1	-	-	-	-	-	0
	ʃs	4	-	-	-	-	-	4	-	-	0
	ʃʒ	4	-	-	-	-	-	3	-	1	75
	ʒ:	3	-	1	-	-	-	-	-	2	0
sanan keskellä	s	10	10	-	-	-	-	-	-	-	100
	ʃ	9	-	9	-	-	-	0	0	-	100
	z	6	5	-	-	-	-	-	-	1	0
	ʒ	3	-	3	-	-	-	-	-	-	0
	ʃs	6	-	-	-	-	1	4	-	1	17
	ʃʒ	6	-	-	-	-	1	4	-	1	67
	ʒ:	5	-	2	-	-	-	1	1	1	20

alussa olevista äänneistä tuotettiin oikein 55 % ja sanan keskellä olevista 56 %.

Alkutestauksessa viittä äännettä ei tuotettu kertaakaan oikein, eikä ole yllättävää, että kyseessä olevien äänneiden ominaisuudet poikkeavat merkittävästi suomen ääntönormistosta. Joukossa ovat soinnilliset äänteet /z, ʒ/, dentaalinen affrikaatta /ʃs/ sekä venäjän sibilanteista eksoottisin, liudentunut suhuässä /ʒ:/.

Kielten välinen positiivinen siirtovaikutus eli transferenssi oletettavasti näkyy tarkimmin tuotettujen äänneiden /s, ʃ, ʃʒ/ kohdalla. Kaksi ensimmäistä äännettä vastaavat melko läheisesti suomen äänneitä /s, ʃ/, kun taas /ʃʒ/ on koehenkilölle tuttu englannista ja ruotsista. Sitä vastoin negatiivinen siirtovaikutus eli interferenssi saattaa selittää etenkin äänneiden /z, ʒ/ tuottamista: soinnilliset äänteet äännettiin lähes aina niiden soinnittomina vastapareina /s, ʃ/. Toisaalta voi argumentoida, että myös äänneiden /z, ʒ/ tulisi olla tuttuja englannin kautta, jolloin oletus transferenssista affrikaatan /ʃʒ/ kohdalla on epäpätevä. Edellä mainittujen vaikeuksien lisäksi suomen kielelle eksoottinen, terävä dentaalinen affrikaatta /ʃs/ sekä venäjälle omintakeinen liudentunut /ʒ:/ osoittautuivat vaikeiksi ääntää.

Kuten yllä todettiin, virheitä tarkastellessa soinnillisuuden ja soinnittomuuden oppositio on yksi selkeimmistä ongelmakohdista. Muita tällaisia ovat toisistaan vain ääntöpaikan perusteella eroavat äänteet kuten /s-ʃ/ ja /ʃʒ-ʃs/. Ainoa yllättävä, ennakoimaton virhe kevään testikerralla tapahtui liudentuneen suhuäänneen /ʒ:/ kohdalla, kun se tuotettiin affrikaattana /ʃʒ/.

viä tekijöitä voivat olla niin grafeemi-foneemivastaavuuden heikko tunteminen kuin puhdas lipsahduskin – tekeehän kielen natiivipuhujakin puheessaan äänteellisiä lipsahduksia.

Epätarkkojen äänneiden tapauksessa kyse oli yleensä kahden eri äänteen yhdistelmästä, jona aiottu äänne realisoitui. Näin tapahtui esimerkiksi äänteen /ç:/ kohdalla, kun se äännettiin äänneiden /ʃ-z/ ja /ç:-ʃ/ välimuotona, tai äänteen /z/ kohdalla, kun se äännettiin vaikeasti luokiteltavana, äänneiden /s-z/ välimuotona. Epätarkkoiksi luokitellut affrikaattojen ääntövirheet puolestaan syntyivät, kun venäjän /tʃ/ tai /tʂ/ äännettiin ikään kuin kahtena erillisiksi tulkittavina, peräkkäisinä foneemeina. Affrikaattojen ääntämisessä todennäköisesti näkyi interferenssi. Ne äännettiin pikemminkin suomalaisittain, ”laiskasti”, kahtena peräkkäisenä äänteenä (kuten esim. suomen sanassa ’itse’) affrikaatoille ominaisen nopean ääntötavan sijaan.

Yhteenvetona empiirisen kokeen alkutestauksen tuloksista voidaan todeta, että reseptiiviset taidot (78 % oikein) ovat selkeästi kehittyneemmät kuin produktiiviset (55 %). Mielenkiintoista on, että sanan alussa sijaitsevan /ʃ/:n ja keskellä sijaitsevan /tʃ/:n tapauksessa produktiiviset taidot olivat reseptiivisiä tarkemmat. Saattaa olla, että näiden äänneiden kohdalla grafeemi-foneemivastaavuus oli heikompa, mikä selittäisi virheet tunnistettaessa. Suurin osa tuloksista kuitenkin noudattaa universaalia tendenssiä: oikein havaitseminen edeltää oikein tuottamista.

## 5.2 Syksy 2016 – välitestaus

Tulokset taulukossa 5 osoittavat, etteivät opetustuokio ja itsenäinen harjoittelu vaikuttaneet tunnistustarkkuuteen: oikein tunnistettuja äännteitä on määrällisesti yhtä paljon kuin alkutestauksessa (78 %). Myöskään sanan alussa tai keskellä oikein havaittujen äänneiden määrä ei muuttunut merkittävästi.

Välitestausten tulosten voidaan kuitenkin sanoa jossain mielessä heikentyneen. Aina oikein tunnistettuja äännteitä oli enää neljä verrattuna alkutestauksen seitsemään. Toisin kuin keväällä, välitestauksessa ei ollut yhtäkään äännettä, joka olisi havaittu aina oikein sekä sanan alussa, että keskellä. Laadullisesti tuloksia heikentää myös sananalkuinen /ç:/, joka tunnistettiin aina väärin.

Alkutestauksessa jokseenkin hankalaksi osoittautuneen sananalkuisen /ʃ/:n tunnistustulosten kehityksessä ei ole johdonmukaisuutta, sillä vaikka virheiden määrä väheni, olivat virheet erilaisia kuin keväällä. Kevään ja syksyn toisistaan paljon poikkeavien tulosten pohjalta voisi /ʃ/:n tunnistamisen päätellä olevan hankalaa.

Sananalkuinen soinnillinen dentaali /z/ tunnistettiin tismalleen kuten keväällä: kaksi kertaa kuudesta se luokiteltiin soinnilliseksi palatoalveolaariksi /z/. Dentaalisen ääntöpaikan tuottaman terävyyden ja palatoalveolaarisen ääntöpaikan tuottama suhina ja väljyys eivät edelleenkään olleet helppoja

TAULUKKO 5. Syksyn havaintotehtävissä havaitut äänteet. 100 % tarkkuus eroteltu. Viiva tarkoittaa, ettei kyseistä foneemiparia testattu (esim. /z-s/).

luettu		havaittu								oikein %
		n	s	ʃ	z	ʒ	ʃs	ʃɕ	ɕ:	
sanan alussa	s	6	6	0	0	-	-	-	-	100
	ʃ	12	2	9	-	0	-	0	1	75
	z	6	0	-	4	2	-	-	-	67
	ʒ	5	-	0	0	5	-	-	-	100
	ʃs	3	-	-	-	-	2	1	-	67
	ʃɕ	6	-	1	-	-	1	4	-	67
	ɕ:	3	-	3	-	-	-	-	0	0
sanan keskellä	s	6	5	1	0	-	-	-	-	83
	ʃ	12	0	12	-	0	-	0	0	100
	z	5	0	-	5	-	-	-	-	100
	ʒ	6	-	0	2	4	-	-	-	67
	ʃs	3	-	-	-	-	2	1	-	67
	ʃɕ	5	-	0	-	-	1	4	-	80
	ɕ:	3	-	2	-	-	-	-	1	33

tunnistaa. Kontrastiivinen analyysi ehdottaisi, että koska molemmat ääntöpai-  
kat ja sen lisäksi soinnillisuuskin ovat sellaisia ominaisuuksia, joita suomen kie-  
lessä ei pääsääntöisesti ole, ei suomenkielisen oppijan korva ole (liian vähäisen  
harjoituksen takia) oppinut erottamaan niitä, mistä syntyy virheitä havaitse-  
misessa. Mielenkiintoista on kuitenkin huomata, että sekä alku- että välites-  
tauksessa näistä kahdesta äänneestä vain /z/:n kohdalla tehtiin virheitä: /z/  
tunnistettiin molemmilla testikerroilla aina oikein.

Kiinnostavaa pohdittavaa tarjoavat sananalkuisen liudentuneen suhu-  
äänteen /ɕ:/ tunnistustulokset, sillä ne ovat aivan päinvastaiset alkutes-  
taukseen verrattuna. Keväällä sanan alussa olevien äänneiden tunnistustark-  
kuus oli 100 %, mutta syksyllä 0 %. Tulos antaisi viitettä sille, että grafeemi-  
foneemivastaavuuden osaamisessa on horjahtelua tai että tunnistustehtävän  
täyttäminen on kenties tapahtunut vähemmän väärän vaihtoehdon valitse-  
malla, jolloin tuloksiin on tullut näinkin suurta vaihtelua.

Välitestauksen lukupuhuntatehtävissä (taulukko 6) tuottamistarkkuus oli  
66 % eli se nousi ensitestaukseen nähden 16 %. Äänteen sijainti sanassa ei  
ensitestauksen tavoin vaikuttanut tuottamistarkkuuteen: sanan alussa ole-  
vien äänneiden tuottamistarkkuus oli 66 % ja sanan keskellä olevien äänneiden  
67 %.

Aina väärin tuotettujen äänneiden määrä väheni viidestä kolmeen. Samo-  
ja äänneitä äännettiin aina väärin myös alkutestauksessa. Yllättäviä tuottamis-  
virheitä ei syksyn testikerralla ollut. Produktiivisten taitojen ongelmat keskit-

TAULUKKO 6. Syksyn lukupuhuntatehtävissä tuotetut äänneet. 100 % tarkkuus eroteltu.

lauseessa		n	s	ʃ	z	ʒ	tuotettu			epätarkka	oikein %
							ʃs	ʃʒ	ʒ:		
sanan alussa	s	12	12	-	-	-	-	-	-	-	100
	ʃ	4	-	4	-	-	-	-	-	-	100
	z	6	3	-	5	-	-	-	-	1	83
	ʒ	5	-	2	-	2	-	-	-	1	40
	ʃs	4	-	-	-	-	0	4	-	-	0
	ʃʒ	4	-	-	-	-	-	2	-	2	50
	ʒ:	3	-	3	-	-	-	-	0	-	0
sanan keskellä	s	10	10	-	-	-	-	-	-	-	100
	ʃ	1	8	-	-	-	-	-	-	-	89
	z	6	1	-	5	-	-	-	-	-	83
	ʒ	3	-	2	-	0	-	-	-	1	0
	ʃs	6	-	-	-	-	4	2	-	-	67
	ʃʒ	6	1	-	-	-	1	2	-	2	33
	ʒ:	5	-	3	-	1	-	-	1	-	20

tyvät hypoteesissa mainittujen ominaisuuksien, eli opposition soinnillisuussoinnittomuus sekä etenkin affrikaattojen tapauksessa ääntöpaikan, ympärille. Usein toistuneita epätarkkoja ääntötapoja olivat affrikaatan /ʃs/ realisointuminen äänneiden /t/ ja /s/ "suomalaistyyppisenä" yhdistelmänä sekä soinnillisen palatoalveolaarin /z/ realisointuminen äänneiden /ʃ, z/ yhdistelmänä.

Huomattavin parannus produktiivisissa taidoissa tapahtui soinnillisten äänneiden tuottamisessa. Kun alkutestauksessa sanan keskellä sijaitseva /z/ äännettiin soinnittomana äänneenä /s/ tai näiden kahden epätarkkana väli-  
muotona, välitestauksessa soinnillisuus oli mukana viidessä tuotoksessa kuu-  
desta. Muutoinkin soinnillisuuden tuottaminen näyttäisi kehittyneen sitten al-  
kutestauksen, mikä ilmenee äänneiden /z, z/ tulosten parantumisesta. Myös  
dentaalisen affrikaatan /ʃs/ ääntäminen lienee hahmottuneen koehenkilölle  
paremmin – tosin vain sanan keskellä sijaitessa. Suomen sibilanttien saman-  
kaltaisuus venäjän äänneisiin /s, ʃ/ nähden näkyy edelleen näiden äänneiden  
tuottamistarkkuudessa positiivisen transferenssin kautta.

Välitestauksen tuloksissa reseptiiviset taidot yhä edelsivät produktiivisia tai-  
toja, mutta tarkkuuksien välinen ero kaventui: syksyn tarkkuusprosentti oli re-  
septiivisissä taidoissa 78 (sama kuin alkutestauksessa) ja produktiivisissa taidois-  
sa 66 (alkutestauksessa 55). Reseptiiviset taidot olivat edelleen produktiivisia  
vahvemmat.

TAULUKKO 7. Talven havaintotehtävässä havaitut äänteet. 100 % tarkkuus eroteltu. Viiva tarkoittaa, ettei kyseistä äänneparia testattu (esim. /z-s/).

luettu		n	s	ʃ	z	havaittu				oikein %
						z	ʃs	ʃz	ɕ:	
sanan alussa	s	6	5	1	0	-	-	-	-	83
	ʃ	12	1	7	-	1	-	0	3	58
	z	6	0	-	5	1	-	-	-	83
	z	5	-	0	0	5	-	-	-	100
	ʃs	3	-	-	-	-	2	1	-	67
	ʃz	6	-	0	-	-	0	6	-	100
	ɕ:	3	-	0	-	-	-	-	3	100
sanan keskellä	s	6	5	1	0	-	-	-	-	83
	ʃ	12	0	12	-	0	-	0	0	100
	z	5	0	-	4	1	-	-	-	80
	z	6	-	0	1	5	-	-	-	83
	ʃs	3	-	-	-	-	3	0	-	100
	ʃz	5	-	2	-	-	0	3	-	60
	ɕ:	3	-	1	-	-	-	-	2	67

### 5.3 Talvi 2016 - lopputestaus

Lopputestauksen **tunnistustehtävässä** (taulukko 7) oikein tunnistettiin 83 %. Kokonaisuudessaan tunnistustarkkuus nousi alku- ja välitestauksesta (joissa tunnistustarkkuudet olivat samat) 6 %. Vaikka kokonaistarkkuus oli testikerroista korkein, aina oikein havaittuja äänteitä ei kuitenkaan ollut lopputestauksessa eniten. Lopputestauksessa, toisin kuin alkutestauksessa, ei myöskään ollut yhtäkään äännettä, joka olisi havaittu aina oikein sijainnistaan huolimatta. Toisaalta myöskään aina väärin tunnistettuja äänteitä ei ollut, toisin kuin välitestauksessa.

Sananalkuisen palatoalveolaarin /ʃ/ tunnistaminen säilyi kirjavana loppuun asti, sillä talven testikerta esittelee jälleen uudenlaisen havaintotaulukon erilaisine virhevastauksineen. Sananalkuisen dentaalin /z/ tarkkuus taas oli yhdellä virhetunnistuksella testikertojen parhain, samaten muilla testikerroilla vaikeaksi osoittautuneen sanan keskellä sijaitsevan soinnillisen suhuäänteen /z/ tarkkuus. Ylipäätään tarkkuus lopputestauksen havaintotehtävässä oli hyvällä tasolla. Vain kahden äänteen kohdalla tehtiin useampi kuin yksi virhe, ja etenkin ero soinnittomien ja soinnillisten äänteiden välillä näyttäisi hahmottuneen vielä edellisiä testikertoja paremmin. Myös affrikaattojen kokonaistulos (82 % oikein) oli testikertojen paras.

Lopputestauksen lukupuhuntatehtävässä (taulukko 8) oikein tuotettiin 63 %, mikä on 5 % vähemmän kuin välitestauksessa. Äänteen sijainti sanan

TAULUKKO 8. Talven lukupuhuntatehtävissä tuotetut äänteet. 100 % tarkkuus eroteltu.

lauseessa		n	s	ʃ	z	ʒ	tuotettu			epätarkka	oikein %
							ʃs	ʃɕ	ɕ:		
sanan alussa	s	12	10	1	-	-	-	-	-	1	83
	ʃ	4	-	4	-	-	-	-	-	-	100
	z	6	1	1	3	-	-	-	-	1	50
	ʒ	5	-	2	1	2	-	-	-	-	40
	ʃs	4	-	-	-	-	-	4	-	-	0
	ʃɕ	4	-	-	-	-	1	3	-	-	75
	ɕ:	3	-	1	-	-	-	-	1	1	33
sanan keskellä	s	10	10	-	-	-	-	-	-	-	100
	ʃ	9	-	9	-	-	-	-	-	-	100
	z	6	2	-	2	-	-	-	-	2	33
	ʒ	3	-	2	-	1	-	-	-	-	33
	ʃs	6	-	-	-	-	3	3	-	-	50
	ʃɕ	6	-	-	-	-	2	3	-	1	50
	ɕ:	5	-	2	-	-	-	-	1	2	50

alussa tai keskellä ei viimeiselläkään testikerralla ollut erityisen merkittäväsä asemassa: sanan alussa tuotettujen äänteiden kokonaistarkkuus oli 61 % ja sanan keskellä tuotettujen 64 %.

Aina väärin tuotettujen äänteiden määrä laski kevään viidestä ja syksyn kolmesta yhteen aina väärin tuotettuun äänneeseen (sananalkuinen affrikaatta /ʃs/). Tästä näkökulmasta kehitystä oli siis tapahtunut kiitettävästi. Yllättäviä ääntövirheitä ei ollut, vaan virheet liittyivät sekaannuksiin toisiaan läheisesti muistuttavien äänteiden kanssa. Koehenkilön tauko venäjän opiskelusta ja käyttämisestä vapaa-ajalla saattaa näkyä siinä, että ns. epätarkkoja ääntötapoja esiintyi lopputestauksen tuloksissa testikerroista eniten. Epätarkat ääntötavat edustavat jälleen kahden eri lähiaänteen yhdistelmiä, esimerkiksi tavoitteena ollut /ɕ:/ realisoitui äänteiden /ʃ, ɕ:/ yhdistelmänä, ja affrikaataksi /ʃs/ aiottu äänne muistutti ”suomalaistyyppistä” äänneyhdistelmää, jossa /t/ ja /s/ esiintyvät peräkkäin (kuten suomen sanassa *itse*).

#### 5.4 Yhteenveto havaintotehtävistä ja lukupuhuntatehtävistä

Taulukko 9 tiivistää tutkimuksen empiirisen osan tulokset ja havainnollistaa mm. havaintotehtävien tulosten kirjavuutta. Jokaisella testikerralla 100 % tarkkuudella tunnistettuja äänneitä olivat sananalkuinen /z/ ja sanan keskellä sijaitseva /ʃ/. Nousujohteisuutta on havaittavissa kahden äänteen (sananalkuinen /z/, sanan keskellä sijaitseva /z/) kohdalla, kun taas laskujohteisuutta

jopa kolmen äänteen (/s/ molemmissa asemissa, sanan keskellä sijaitseva /z/) kohdalla. Kuitenkaan minkään äänteen tunnistustarkkuus ei systemaattisesti laskenut testikerrasta toiseen.

Yllä mainittujen lisäksi joukkoon mahtuu myös äännteitä, joiden tunnistustarkkuus vaihteli tarkemmasta epätarkempaan ja takaisin (sananalkuiset /t̪/, ɕ:/ ja sanan keskellä sijaitsevat /t̪s, ɕ:/) sekä niitä, joiden tunnistus vaihteli epätarkasta tarkempaan ja takaisin (sananalkuinen /ʃ/ ja sanan keskellä sijaitseva /t̪ɕ/).

Kolmen eri testikerran tulosten perusteella myöskään sanan alussa ja keskellä esiintyvien äännteiden tunnistustarkkuudet eivät ole suoraan tai kääntäen verrannollisia keskenään. Lisäksi ei voida tehdä yleispätevää johtopäätöstä siitä, onko yksittäinen äänne helpompi tunnistaa sanan alussa vai keskellä. Jos sijaintia sanassa tutkitaan äänne kerrallaan, huomataan, että vain /ʃ/ tunnistetaan johdonmukaisesti aina paremmin sanan keskellä ja /z/ aina sanan alussa.

Suomen kielestä puuttuvan binääriominaisuuden soinnillisuus-soinnittomuus tunnistamista koskien ei myöskään voida muodostaa pitävää lainalaisuutta. Soinnittomat äännteet havaittiin alkutestauksessa keskimäärin tarkemmin (81 %) kuin soinnilliset (75 %), mutta väli- ja lopputestauksessa soinnilliset äännteet havaittiin tarkemmin (84 % syksyllä ja 87 % talvella) kuin soinnittomat (67 % syksyllä ja 82 % talvella). Täten suomen ja venäjän äännejärjestelmien kontrastoinnin pohjalta tehtävä oletus siitä, että suomenkieliselle oppijalle tummat soinnittomat äännteet ovat helpompia havaita, voi olla virheellinen, mikä tukisi sitä, että kontrastiivista analyysiä kannattaa käyttää virheitä selittävänä eikä niinkään ennustavana työkaluna.

Tunnistustarkkuus nousi alkutestauksesta lopputestaukseen vain 6 %, ja laadullisestikin reseptiivisten taitojen kehitys on pientä. Ailahtelevat testitulokset, joissa vain kahden äänteen kohdalla voidaan havaita nousujohteisuutta, eivät luo kuvaa tutkimuksen aikana kehittyneestä venäjän sibilanttien ja affrikaattojen reseptiivisestä hallinnasta.

Lukupuhunnassa sen sijaan havaitaan näkyvämpää muutosta, kun tuottamistarkkuus nousi alku- ja välitestauksen välillä 16 %. Jokaisella testikerralla aina oikein tuotettuja äännteitä olivat sananalkuinen /ʃ/ sekä sanan keskellä sijaitseva /s/. Minkään äänteen kohdalla ei ole havaittavissa systemaattista nousu- tai laskujohteisuutta. Lukupuhuntakokeiden tuloksia voidaan lyhyesti sanoen kuvailla kirjaviksi ja ailahteleviksi, mutta kokonaisuus huomioon ottaen kehitystä tapahtui alkutestaukseen verrattuna.

Kolmen eri testikerran perusteella ei voida sanoa, että yhden ja saman äänteen tuottaminen tietystä asemassa ennakoisi tiettyä tulosta toisessa asemassa tuottaessa. Tutkimuksen perusteella tutkittujen äännteiden tuottamistarkkuudet eri sijainneissa eivät siis ole suoraan tai kääntäen verrannollisia. Ei voida myöskään luoda yleistä lainalaisuutta siitä, ovatko äännteet helpompia tuottaa sanan alussa vai keskellä. Lukupuhuntatehtävien tuloksissa sanan

TAULUKKO 9. Yhteenvedo havainto- ja lukupuhuntatehtävien tuloksista. 100 % tarkkuus eroteltu.

lauseessa		kevät		syksy		talvi	
		havainto	tuotto	havainto	tuotto	havainto	tuotto
sanan alussa	s	100	92	100	100	83	83
	ʃ	50	100	75	100	58	100
	z	67	50	67	83	83	50
	ʒ	100	0	100	40	100	40
	/ʃs/	67	0	67	0	67	0
	/ʃʒ/	83	75	67	50	100	75
	ʒ:	100	0	0	0	100	33
	sanan keskellä	s	100	100	83	100	83
ʃ	100	100	100	89	100	100	
z	100	0	100	83	80	33	
ʒ	33	0	67	0	83	33	
/ʃs/	100	17	67	67	100	50	
/ʃʒ/	40	67	80	33	60	50	
ʒ:	67	20	33	20	67	20	

alussa ja keskellä tuotettujen äänteiden tarkkuudet dominoivat vuorotellen eikä tuloksista paljastu selvää linjaa asian suhteen. Jos äänteen sijaintia sanassa tutkitaan äänne kerrallaan, havaitaan, että vain /ʃʒ/ tunnustetaan johdonmukaisesti aina paremmin sanan alussa ja /ʃs/ aina sanan keskellä.

Toisin kuin reseptiivisten taitojen osalta, lukupuhuntatehtävän osalta voidaan todeta, että soinnittomat äänteet ovat keskimäärin helpompia tuottaa kuin soinnilliset. Siinä, missä soinnilliset äänteet tuotettiin ensin 13 %, sitten 56 % ja lopuksi 39 % tarkkuudella, olivat soinnittomien äänteiden vastaavat luvut 57 %, 52 % ja 61 %. Voidaan siis todeta, että soinnittomat äänteet ovat (varsinkin ilman opetuksen ja harjoittelun tuoreita muistijälkiä) helpompia tuottaa tarkasti. Yllä esitetyistä luvuista käy toisaalta mielenkiintoisesti ilmi, että välitestauksessa, jolloin opetustuokion ja harjoittelun tuoma taito todennäköisesti olivat tuoreessa muistissa, koehenkilö tuotti muutaman prosenttiyksikön erolla paremmin soinnillisia kuin soinnittomia äänteitä. Tilastoon vaikuttanevat tutkimuksessa haastaviksi osoittautuneet affrikaatat, jotka ovat soinnittomia äänteitä.

Tutkimalla taulukkoa 9 yksittäisten äänteiden tasolla huomataan, että havainnointi on tuottamista vahvempaa kahta tapausta lukuun ottamatta: sanan keskellä sijaitseva /s/ ja sanan alussa sijaitseva /ʃ/ tuotettiin tarkemmin kuin havaittiin. Yleisesti voidaan siis varmentaa muiden tutkijoiden tulokset siitä, että oikein havaitseminen yleensä aina edeltää oikein tuottamista.



## 6 Johtopäätökset

Empiirisestä kokeesta nousi esiin äänneitä, joita voidaan luonnehtia helpoimmiksi tai päinvastoin vaikeimmiksi joko reseptiivisellä, produktiivisella tai molemmilla suullisen kielitaidon osa-alueilla. Kaikki näistä äänneistä eivät kuitenkaan noudattaneet alkuperäistä hypoteesia. Helpoimmiksi äänneiksi osoittautuivat suomen *s*-äänneitä lähimpänä olevat /s, ʃ/, tosin /s/:n oikein havaitseminen sanan alussa tuotti vaikeuksia. Lisäksi sananalkuinen /z/ sekä sanan keskellä olevat /z, ʒ/ olivat helpoimpia äänneitä tunnistamisen kannalta. Vaikeimmiksi äänneiksi sekä tuottaa että tunnistaa osoittautuivat sananalkuiset /ʃs, ʒ:/ sekä sanan keskellä sijaitseva /z/. Toisaalta on mielenkiintoista, että /z/ oli helppo havaita sanan alussa. Äänne /z/ osoittaa, ettei jako helppoihiin ja vaikeisiin äänneisiin ole aina yksiselitteinen. Lisäksi erityisiä vaikeuksia tuotti äänneen /ʒ:/ tuottaminen sanan keskellä.

Reseptiivisen ja produktiivisen osaamisen välillä pystyttiin joidenkin äänneiden kohdalla löytämään yhteys, eli äänneen helppous tai vaikeus ilmeni sekä havaitsemisessa että tuottamisessa. Esimerkkejä tästä ovat sananalkuiset /s, ʃ/ sekä sanan keskellä sijaitsevat /s, ʃ/. Kuitenkin suurimmassa osassa äänneitä reseptiivinen ja produktiivinen äänneiden hallinta eivät kulkeneet käsi kädessä, vaan tarkkuudet näiden välillä vaihtelivat hyvinkin satunnaisesti. Tästä esimerkkinä on sanan keskellä sijaitseva /ʃʒ/. Kyseisen äänneen alhainen tunnistustarkkuus alkutestissä parani välitestissä ja jälleen heikkeni lopputestissä, kun taas hyvä tuottamistarkkuus heikkeni välitestissä ja jälleen parani lopputestissä. Reseptiivisten ja produktiivisten taitojen saman tasoisuus tai samankaltainen kehitys on saaduissa tuloksissa hyvin riippuvainen kyseessä olevasta äänneestä.

Vaikka sekä reseptiiviset että produktiiviset taidot paranivat alkutestausen tuloksiin nähden, ovat eri testikertojen tulokset keskenään niin kirjavat ja vaihtelevat, että on vaikea varmasti todeta eksplisiittisen opetustuokion ja harjoittelun auttaneen venäjän sibilanttien ja affrikaattojen omaksumisessa. Toisaalta äänneiden tuottamisen osalta kehitys on sen verran huomattavaa, että eksplisiittisellä työskentelyllä on mitä luultavimmin ollut vaikutusta. Loppujen lopuksi tutkimuksen tulokset osoittavat, että kyseessä on vaikea vieraana kielenä opiskeltavan venäjän osa-alue, johon on syytä kiinnittää systemaattisesti huomiota pysyvien tulosten aikaansaamiseksi.

Empiirisen kokeen suunnittelussa ja toteutuksessa on joitain seikkoja, jotka kannattaa jatkossa tehdä toisin luotettavampien ja monipuolisempien tutkimustulosten varmistamiseksi. Opetustilanteessa teorian ohella tehty harjoittelu rajoittui vain tuottamiseen sen sijaan, että myös tunnistamista olisi harjoiteltu. Tässä kohdin kirjoittaja siis kaivoi omaa kuoppaansa: ajatustasolla havainnoinnin ja tuottamisen yhdistämisen tärkeys on tunnustettu ja tärkeäksi katsottu, mutta käytäntöön asti teoria ei edes tutkimusasetelmassa päässyt.

Testitilanteen ulkopuolella tapahtuva tunnistamisen harjoittelu olisi antanut tutkijalle ja testattavalle itselleen tietoa testattavan lähtötasosta. Havainnointiharjoittelu olisi myös saattanut herätellä tutkittavaa reseptiivisten taitojen tärkeyteen. Lisäksi reseptiivisten taitojen yhteydessä olisi voitu varmistaa, että testattava osaa yhdistää venäjän grafeemit ja foneemit toisiinsa. Nykyisellään tutkimuksessa ei ole varmuutta grafeemien osaamisesta, joten osa tuloksista voi olla seurausta siitä, että testitilanteessa testattava on valinnut vastauslo-makkeelta väärän grafeemin kuin mitä tarkoitti valita. Grafeemien opettelua ja osaamisen arviointia on helppo toteuttaa venäjän oppitunneillakin esimerkiksi erilaisten pelillisten aktiviteettien myötä.

Opetustuokion lisäksi myös testien tehtävätyypeissä on parannettavaa. Jos havaintotehtävässä olisi minimiparien erotteluun sijaan annettu tehtäväksi valita kuuluu äänne venäjän koko sibilantti- ja affrikaattarepertuaarista, olisi saatu selville, millaisia (ehkä yllättäviäkin) virheitä äänteiden tunnistuksessa tehdään. Äänteiden tuottamisen testaukseen puolestaan olisi lukupuhunta-tehtävän ohelle voinut ottaa spontaania puheentuottoa esim. kuvantunnis-tustehtävän muodossa. Tämä toisi merkittävää lisäarvoa produktiivisten tai-tojen arviointiin kankean paperista lukemisen ohelle.

Lisäksi testitehtävien suunnittelu vaatii kehitystä, sillä monia äännteitä tes-tattiin todella vähän: esimerkiksi sananalkuinen /ç:/ esiintyi havaintotehtävissä yhteensä vain kolmesti. Täten yksittäisten äänteiden vaikeuteen tai help-pouteen kantaa ottavien johtopäätösten luotettavuus kärsii. Samasta syystä myös eri äännteitä koskevien lukujen keskinäinen vertailu ontuu: kun esimer-kiksi sananalkuista /ʃ/-äännettä testattiin havaintotehtävässä 12 kertaa ja /ç:/-äännettä kolme kertaa, ei voida luotettavasti todeta, kumpi näistä äännteistä todella oli haastavampi koehenkilölle. On myös huomioitava, että vaikka ään-teitä testattiinkin sekä havaitsemisen että tuottamisen osalta niin sanan alkui-sissa kuin keskisissä asemissa, äännteet esiintyvät erilaisissa foneemiympäris-töissä ja erilaisissa asemissa sanapainoon nähden. Analyysi on tässä suhteen suppea, sillä se ei ota huomioon foneemiympäristön tai sanapainon vaikutusta tuottamiseen tai tunnistamiseen. Oletettavaa kuitenkin on, että nämäkin tekijät saattavat vaikuttaa tunnistus- ja tuottamistarkkuuteen, joten myöskään tältä kantilta katsottuna johtopäätökset eivät ole lopullisia.

Lopuksi, tarkempiin johtopäätöksiin päästäkseen olisi aiheellista käyttää venäjän natiivipuhujien apua. Havaintotehtävässä käytetyt äänitteet voisi tes-tata teettämällä havaintotehtävät kontrolliryhmällä, samaten aineiston tulkin-nassa voisi hyödyntää venäjän natiivipuhujien apua. Jälkimmäinen olisi ollut arvokas lisä dentaalisen äänteen /s/ tuottamista analysoitaessa, sillä on oletet-tavaa, että joissain tapauksissa venäjänkielinen kuuliija saattaa havaita suomen alveolaarisen /s/:n niin väljänä, että tulkitsee sen palatoalveolaarisena venä-jän äänteenä /ʃ/ eikä suinkaan dentaalisena äänteenä /s/, kuten tutkimukses-sa tehtiin. Äänteiden autenttisuuden eli kohdekielen ääntönormien noudatta-

misen tulkitsemisen tukena voisi käyttää natiivipuhujilta äänitettyjä kontrollilauseita, joihin vertaamalla koehenkilön suorituksia voisi paremmin arvioida autenttisuuden näkökulmasta.

Aineiston koon vuoksi kyseessä on vain yksittäinen tapaustutkimus eivätkä sen tulokset ole yleistettävissä. Kuitenkin tutkimuksen aikana karttunut kokemus ääntämisen eksplisiittisestä opettamisesta ja sen testaamisesta, samoin kuin venäjän ja suomen kontrastoinnista, sekä tutkimuksen sivutuotteena syntynyt erittäin käyttökelpoinen opetus- ja harjoitusmateriaali tarjoavat monipuolisen työkaluvalikoiman niin opetustyötä kuin myös tulevia tutkimuksia ajatellen. Tarkemman koesuunnittelun lisäksi jatkossa tulisi keskittyä entistä monipuolisempiin harjoittelu- ja opetustapoihin. Mielenkiintoista olisi myös selvittää venäjän oppijoiden mielipide eksplisiittisestä ääntämisopetuksesta sekä heidän asenteitaan ääntämiseen ja sen tärkeyteen, opetteluun ja opettamiseen ylipäätään.

## Lähteet

- Aho, E., A. Huhtaniemi & M. Nikonen 2016a. *Fonetikkaa suomen kielen oppijoille*. Helsinki: Finn Lectura.
- Aho, E. & M. Toivola 2008. Venäläisten maahanmuuttajien suomen prosodiasta. *Virtittäjä*, 112 (1), 3–23. <https://journal.fi/virtittaja/article/view/40638>.
- Aho, E., M. Toivola, F. Karlsson & M. Lennes 2016b. Aikuisten maahanmuuttajien suomen ääntämisestä. *Puhe ja kieli*, 36 (2), 77–96. <http://hdl.handle.net/10138/168735>.
- Akišina, A. A. & O. E. Kagan 2014. *Utšimsâ utšit': dlâ prepodavatelej russkogo âzyka kak inostrannogo*. Moskva: Russkij âzyk.
- Bernštejn, S. I. 1975. Voprosy obučenie proiznošeniû (primenitel'no k prepodavaniû russkogo âzyka kak inostrannogo). Teoksessa A. A. Leont'ev & N. I. Samujlova (toim.) *Voprosy fonetiki i obučenie proiznošeniû*. Moskva: Izdatel'stvo moskovskogo universiteta, 5–49.
- Best, C. T. 1995. A direct realist view of cross-language speech perception. Teoksessa W. Strange (toim.) *Speech perception and linguistic experience*. Baltimore (Md.): York Press. <http://www.haskins.yale.edu/publications/pub-b.html> [luettu 1. 4. 2017].
- Fedotova, N. L. 2013. *Metodika prepodavaniâ russkogo âzyka kak inostrannogo: praktičeskij kurs*. Sankt-Peterburg: Zlatoust.
- Flege, J. E. 1988. The production and perception of foreign language speech sounds. Teoksessa H. Winitz (toim.) *Human communication and its disorders: a review 1988*. Norwood (N. J.): Ablex, 224–401.
- 1995. Second language speech learning: theory, findings, and problems. Teoksessa W. Strange (toim.) *Speech perception and linguistic experience: issues in cross-language research*. Timonium (Md.): York Press. [http://jimflege.com/files/Flege\\_in\\_Strange\\_1995.pdf](http://jimflege.com/files/Flege_in_Strange_1995.pdf) [luettu 16. 4. 2017].

- Harjanne, P. & S. Tella 2010. Vieraiden kielten opetus Suomessa: nykytilanteen tarkastelua KIELO-tutkimushankkeen taustaa ja teoriaa. Teoksessa E. Ropo, H. Silberberg & T. Soini (toim.) *Toisensa kohtaavat ainedidaktiikat: ainedidaktinen symposiumi 13. 2. 2009 Tampereella*. Tampereen yliopiston opettajankoulutuslaitoksen julkaisuja A 31. Tampere: Tampereen yliopiston opettajankoulutuslaitos, 35–48. <http://www.helsinki.fi/~tella/ads09.pdf> [luettu 18. 3. 2017].
- Iivonen, A. 2005. Fonetiiikan merkitys kielenomaksumisessa ja -opetuksessa. Teoksessa A. Iivonen, R. Aulanko & M. Vainio (toim.) *Monikäyttöinen fonetiikka*. Helsingin yliopiston fonetiikan laitoksen monisteita. Helsinki: Helsingin yliopisto, 45–64.
- James, C. 1980. *Contrastive analysis*. Harlow: Longman.
- Kiparsky, P. & L. Menn 1977. On the acquisition of phonology. Teoksessa J. Macnamara (toim.) *Language learning and thought*. New York (N. Y.): Academic Press, 47–84.
- Kuhl, P. K. 2000. A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, 97 (22), 11850–11857.
- Lado, R. 1957. *Linguistics across cultures: applied linguistics for language teachers*. Ann Arbor (Mich.): The University of Michigan Press.
- Lintunen, P. 2005. Aloittavat yliopisto-opiskelijat englannin kielen ääntäjinä. Teoksessa L. Kuure, E. Kärkkäinen & M. Saarenkunnas (toim.) *Kieli ja sosiaalinen toiminta – Language and social action*. AFinLAn vuosikirja 63. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 375–386.
- 2014a. Kuinka kielenoppija eroaa foneetikosta? Ääntämistaidon opettamisesta ja itsearvioinnista: Turku 25.–26. lokakuuta 2013. Teoksessa K. Jähi & L. Taimi (toim.) *XXVIII Fonetiiikan päivät: Turku 25.–26. lokakuuta 2013*. Turku: University of Turku, 1–9. [https://www.utu.fi/fi/sivustot/fp2013/Documents/XXVIII\\_Fonetiiikan\\_p%C3%A4iv%C3%A4t\\_konferenssijulkaisu.pdf](https://www.utu.fi/fi/sivustot/fp2013/Documents/XXVIII_Fonetiiikan_p%C3%A4iv%C3%A4t_konferenssijulkaisu.pdf) [luettu 19. 3. 2017].
- 2014b. Ääntämisen oppiminen ja opettaminen. Teoksessa P. Pietilä & P. Lintunen (toim.) *Kuinka kieltä opitaan: opas vieraan kielen opettajalle ja opiskelijalle*. Helsinki: Gaudeamus, 165–187.
- Moskovkin, L. V. & A. N. Šukin 2012. *Hrestomatîâ po metodike prepovadaniâ russkogo âzika kak inostrannogo*. 3. laitos. Moskva: Russkij âzyk.
- Opetushallitus 2014. Peruskoulun opetussuunnitelman perusteet 2014. [http://www.oph.fi/download/163777\\_perusopetuksen\\_opetussuunnitelman\\_perusteet\\_2014.pdf](http://www.oph.fi/download/163777_perusopetuksen_opetussuunnitelman_perusteet_2014.pdf) [luettu 9. 3. 2017].
- 2015. Lukion opetussuunnitelman perusteet 2015. [http://www.oph.fi/download/172124\\_lukion\\_opetussuunnitelman\\_perusteet\\_2015.pdf](http://www.oph.fi/download/172124_lukion_opetussuunnitelman_perusteet_2015.pdf) [luettu 29. 11. 2016].
- Temple, R. 2005. Linguistics. Teoksessa J. A. Coleman & J. Klapper (toim.) *Effective learning and teaching in modern languages*. London: Routledge, 192–200.
- Tergujeff, E. 2016. Wanted: tutkimusperustainen ääntämisen opetus. *Puhe ja kieli*, 36 (3), 221–225. <http://journal.fi/pk/article/view/59011/20419> [luettu 12. 3. 2017].
- Ullakonoja, R. 2011. *Da. Eto vopros! Prosodic development of Finnish students' read-aloud Russian during study in Russia*. Jyväskylä Studies in Humanities 151. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-4209-0>.
- Vasil'ev, A. I. 1975. O postroenii russkoj fonetiki dlâ nerusskih. Teoksessa A. A. Leont'ev & N. I. Samujlova (toim.) *Voprosy fonetiki i obučenje proiznošeniû*. Moskva: Izdatel'stvo moskovskogo universiteta, 89–111.

## Liite



KUVIO 1. Kuvakaappaus tutkimusta varten luodusta, venäjän sibilanttien ja affrikaattojen akustisten ominaisuuksien erottamisen harjoitteluun tarkoitetusta nettisivusta.

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 163–192.

**Mari Honko**

Jyväskylän yliopisto

## **Sadutettu sanasto: puhutun kielen leksikaalinen diversiteetti arviointikohteena**

This study analyses lexical diversity (sums of probabilities) in spoken narratives of L1 and L2 school age children (n = 99) and compares the results to the lexical diversity of written narratives of the group of comparison. The key research questions are: a) does the lexical diversity of the spoken narratives systematically differ from the lexical diversity of written narratives and b) does the lexical diversity of spoken narratives systematically differ depending on five individual variables: lexical skills, language proficiency, L1, gender and age of the speaker? All the narratives are produced in Finnish in storytelling events during the spring semester of the 2d and 3rd school year. Sums of probabilities is an index of lexical diversity. It is based on widely used D (Malvern & Richards 1997) and can be used with texts that differ in length and genre.

The results reveal a weak but significant difference between the lexical diversity of spoken and written narratives and weak but complex correlations between the lexical diversity of spoken narratives and the proficiency level as well as the lexical skills of the child. However, there is no correlation or other appreciable connection between the lexical diversity and language background (L1/L2), gender or the school grade of the child. In addition, in L2 group there is no connection between the lexical diversity and the length of residence in Finnish speaking environment or, between the lexical diversity and the specific language difficulties observed by their teacher, either. The results are discussed and compared with the individual differences in turns and turn-taking during the storytelling event tentatively.

**Keywords:** lexical diversity, vocabulary skills, language learning, school age

**Asiasanat:** leksikaalinen diversiteetti, kielen oppiminen, toinen kieli, kouluikäiset

## 1 Johdanto

Artikkelissa tarkastellaan puhutun kielen leksikaalista diversiteettiä (*lexical diversity*), josta on suomenkielisessä kirjallisuudessa käytetty myös nimityksiä sanaston monimuotoisuus ja rikkaus (Malin 2012), monipuolisuus (Honko 2013) ja vaihtelevuus (Taimisto 2014). Leksikaalisen diversiteetin on havaittu olevan yhteydessä mm. kielenkäyttäjän mentaalileksikon kehittyneisyyteen sekä kokonaiskielitaitoon ja vaikuttavan niin kielen ilmaisuvoimaan kuin kielenkäyttäjistä muodostettuihin tulkintoihinkin (ks. esim. Bradac & Wisegarver 1984; Burroughs 1991; Malvern ym. 2004; Jarvis 2013a,b). Leksikaalista diversiteettiä on lukuisten tutkimusten perusteella pidetty lupaavana leksikaalisen tiedon ja taidon indikaattorina, sillä sen on katsottu peilaavan paitsi yksilön sanavaraston laajuutta ja kompleksisuutta myös hänen kykyään käyttää leksikaalisia resursseja tehokkaasti (ks. esim. Malvern & Richards 2002: 85; Malvern ym. 2004; Jarvis 2013a,b). Lisäksi leksikaalisen diversiteetin mittaamista ja tutkimista on pidetty hyödyllisenä tapana kuvata erilaisten tekstien laatua laajemminkin (McCarthy & Jarvis 2007: 476, 482–483; Yu 2010).

Tämän artikkelin tehtävänä on selvittää, a) poikkeako kouluikäisiltä lapsilta kerättyjen puhuttujen kertomusten leksikaalinen diversiteetti aiemmin (Honko 2013) tutkittujen kirjoitettujen kertomusten diversiteetistä (luku 4.1) ja b) onko puhuttujen kertomusten leksikaalisessa diversiteetissä systemaattisia ryhmäkohtaisia eroja, jotka selittyvät lapsen kielitaidolla, ensikielellä, sukupuolella tai iällä (tarkemmin luku 4.2). Lisäksi tarkastellaan sitä, vaikuttaako aikuisen sadutustilanteessa tuottama sanasto lapsen kertomuksen leksikaaliseen diversiteettiin (luku 4.3). Tulosten kontekstoimiseksi kussakin analyysiosion alaluvussa esitellään myös aiemman tutkimuksen, erityisesti kirjoitelma-aineistoon pohjautuvan verrokkitutkimuksen (Honko 2013), tuloksia. Havaintojen pohjalta arvioidaan, voisiko leksikaalinen diversiteetti toimia kehityksellisenä mittarina vastaavia aineistoja analysoitaessa. Vaikka leksikaalista diversiteettiä on aiemmassa tutkimuksessa hyödynnetty jo melko laajalti, sen määritelmä ja mittaustapa eivät ole vakiintuneet, minkä vuoksi tämän tutkimuksen rajaukset perusteluineen esitellään yksityiskohtaisesti luvussa 3.

Tämän tutkimusartikkelin varsinaisena aineistona on sadan alakouluikäisen lapsen suomenkielinen sadutusaineisto, joka on kerätty yhtenevän puolistrukturoidun tehtävänannon avulla. Tutkitun ryhmän lapset ovat iältään 8–11-vuotiaita ja edustavat eri kieli- ja kulttuuriryhmiä. Mukana on sekä suomea ensimmäisenä että toisena kielenä omaksuvia lapsia. Kieliaineistojen ohella käytettävissä on kyselytietoa muun muassa tutkimukseen osallistuneen lapsen kielitaidosta sekä iästä, Suomessa asumisen kestosta ja varhaislapsuudessa käytetyistä kielistä. Lisäksi käytettävissä ovat tulokset sanastonhallintaa erikseen mittaavasta strukturoidusta testistä. Sadutusmenetelmä, aineiston eri osiot ja analyysimetodit esitellään tarkemmin luvussa 3. Kirjoitetun kie-

len aineisto mahdollistaa leksikaalisen diversiteetin vertailun modaliteettien eli puheen ja kirjoituksen välillä ja toisaalta pakottaa metodin kriittiseen arviointiin: esimerkiksi se, että puhekieli harvoin jakautuu siististi ehyiksi ja erillisiksi leksikaalisiksi yksiköiksi, asettaa aineiston analysoinnille kirjoitetun kielen teksteistä poikkeavia haasteita.

Alakouluiässä sanaston määrällinen ja laadullinen kehitys on kiivasta ja yksilölliset erot sanaston hallinnassa suuria (ks. Honko 2013). Erot lasten leksikaalisissa valmiuksissa ovat merkityksellisiä, sillä vahva sanastohallinta on laaja-alaisesti yhteydessä kielelliseen suoriutumiseen kuten luku- ja kirjoitustaitoon ja siten esimerkiksi kykyyn ottaa haltuun koulun oppisisältöjä (Saarela 1997; Alderson 2005; Tannenbaum ym. 2006; Milton 2009; Lervåg & Aukrust 2010). Sanastollinen osaaminen myös ennustaa menestymistä kielellisissä taidoissa myöhemmin: vahva sanasto helpottaa sekä kielellä toimimista että sen edelleen kehittämistä (Cain ym. 2004; Dockrell & Messer 2004; Mutter ym. 2004; Qian & Schedl 2004; Honko 2013). Lapsuusiän kielitaitotutkimuksen avulla voidaan tunnistaa sanastollisten valmiuksien puutteita ja pyrkiä ennaltaehkäisemään erojen kasvua paitsi suoran kielellisen tuen avulla myös esimerkiksi tukemalla lasten sosiaalisia suhteita ja monipuolisia kielenkäyttömahdollisuuksia (Verhoeven 1990: 106–107). Siksi leksikaalisten taitojen tutkiminen on perusteltua juuri koulunaloitusvaiheessa. Leksikaalisen diversiteetin kiinnostavuus arviointi- ja diagnosointivälineenä perustuu ennen kaikkea sen potentiaaliin hyvin monenlaisten tekstien arvioimisessa sekä selkeisiin, laskennallisia menetelmiä hyödyntäviin analyysimalleihin ja niiden toistettavuuteen.

Eroja leksikaalisessa diversiteetissä on aiemmissa tutkimuksissa havaittu muun muassa lasten ja aikuisten sekä eri-ikäisten lasten ja nuorten kirjoittamissa teksteissä (Berman & Verhoeven 2002; Johansson 2008) sekä jo varhaislapsuudessa eri-ikäisten lasten puheessa niin tyyppillisen (Durán ym. 2004) kuin epätyypillisen kielenkehityksen yhteydessä (Klee ym. 2004). Leksikaalista diversiteettiä on pidetty potentiaalisena kehityksellisenä mittarina, mutta näyttö on toistaiseksi vahvinta kielenkehityksen alkuvaiheessa eli ensikielen osalta varhaislapsuudessa ja toisen kielen osalta alimmilla taitotasoilla.

Aiempi tutkimus on tehty suureksi osaksi englantia toisena ja vieraana kielenä -kontekstissa (ks. kuitenkin esim. Castañeda-Jiménez & Jarvis 2014). Koska leksikaalinen diversiteetti on kytköksissä tarkasteltavan kielen morfologiseen ja syntaktiseen rakenteeseen, eri kieliä koskevat tutkimustulokset eivät ole suoraan vertailukelpoisia (Dewaele & Pavlenko 2003: 132–133; Strömqvist ym. 2002). Uutta tietoa tarvitaan sekä puhutun kielen leksikaalisesta diversiteetistä yleisesti että leksikaalisen diversiteetin soveltamisesta erityisesti suomenkieliseen puhutun kielen aineistoon.

Tämä tutkimus rajataan tyyppillisesti kehittyvien lasten puhutun kielen leksikaalisen diversiteetin tarkasteluun, sillä aiempien tutkimusten tulokset etenkin puhutun kielen aineistosta ovat ristiriitaisia (ks. Watkins ym. 1995; Scott &



Windsor 2000; Vermeer 2000; Wong ym. 2010; Ellis ym. 2015; Lai & Schwanenflugel 2016). Tuloksia verrataan lasten muuhun kielelliseen osaamiseen sekä aiemmassa kirjoitettuun kieleen kohdistuneessa tutkimuksessa saatuihin tuloksiin. Lisäksi tuloksia arvioidaan suhteessa niihin yksilöllisiin taustatekijöihin, joiden on aiemmassa tutkimuksessa havaittu vaikuttavan leksikaaliseen diversiteettiin.

## 2 Leksikaalinen diversiteetti kielitaidon arvioinnissa

Leksikaalisen diversiteetin hyödyntäminen kielitieteellisessä tutkimuksessa juontaa juurensa 1930-luvun loppupuolelle John Carrollin artikkeliin ”Diversity of vocabulary and the harmonic series law of word frequency distribution”, jossa Carroll (1938: 379) määritteli diversiteetin (*diversity*) sanaston suhteelliseksi toisteisuudeksi tai vaihteluksi tietyssä tekstissä (“the relative amount of repetitiveness or the relative variety in vocabulary”). Myöhempään määrittelyyn on vaihtelevasti sisällytetty myös sanojen sironta eli sijoittuminen tekstiin ja sanavalikoiman laatu, kuten käytettyjen sanojen harvinaisuus kielessä ja yksilöllisyys tarkastellussa useamman tekstin aineistossa (Malvern ym. 2004; Jarvis 2013b). Vaikka leksikaalista diversiteettiä on varhaislapsuuden jälkeen sovellettu enimmäkseen kirjoitetun kielen tutkimukseen, se koskee jo Carrollin mukaan sekä puhuttua että kirjoitettua kieltä ja on riippuvainen monista tekijöistä kuten tekstin tuottajan iästä, älykkyydestä ja taustasta.<sup>1</sup> Leksikaalisen diversiteetin määrittelytapa, sen tutkimisessa käytetty käsitteistö tai analyysimetodit eivät kuitenkaan ole vakiintuneet, mikä vaikeuttaa sekä tutkimustulosten vertailua että itse tutkimusmetodin arviointia.

Ajatus tietyn tekstin tai tekstikokoelman leksikaalisen diversiteetin yhteydestä tekstintuottajan ominaisuuksiin ja taustaan on tehnyt siitä monien tutkijoiden mielestä kiinnostavan metodin juuri kielitaidon arvioinnin näkökulmasta. Tyypillisesti on verrattu kielenoppijoiden ja äidinkielisten kielenpuhujien tekstejä, tarkasteltu eri-ikäisten ja eri kielitaitotasoa edustavien kielenoppijoiden tekstejä tai analysoitu eriasteisista kielihäiriöistä kärsivien tekstejä suhteessa tyypillisen kielenkehityksen ryhmään. Havaintojen mukaan korkea kielitaitotaso (sekä L1 että L2) ja tyypillinen kielenkehitys (vs. kielihäiriö) ovat tietyn varauksin yhteydessä runsaampaan leksikaaliseen diversiteettiin, ja puheen leksikaalisella diversiteetillä on arvioitu olevan vaikutusta myös siihen,

<sup>1</sup> Carroll itse (1938) pohjasi työnsä Zipfin (1935, 1937) aiempiin julkaisuihin, joissa diversiteetti määriteltiin kapeammin yksittäisten sanojen toiston ja esiintymisvälien kautta (“average rate of repetitiveness”). Toisaalta jo Zipf nosti esille muun muassa sanatoisteisuuden vaikutuksen tekstin vastaanottajan (lukijan tai kuulijan) kokemukseen, mikä ei myöhemmässä tutkimuksessa ole saanut juuri huomiota ennen kuin aivan viime vuosina (ks. myös Jarvis 2013b; Castañeda-Jiménez & Jarvis 2014).

millaisia tulkintoja kuulija tekee puhujasta ja kuinka tähän suhtautuu.<sup>2</sup> Miltonin (2009: 127) esittämän tiivistyksen mukaan kielellisen sujuvuuden lisääntyessä ja taitotason noustessa tuotetun kielen sanaston variaatiokin vähitellen kasvaa. Tuoreimmissa tutkimuksissa leksikaalista diversiteettiä on pidetty lupaavana mittarina myös kielellisen attrition osoittamisessa: attrition myötä leksikaalinen diversiteetti vähenee (Schmid & Jarvis 2014).

Leksikaalisen diversiteetin soveltaminen oppijankielentutkimukseen perustuu kahteen keskeiseen oletukseen: a) kielenoppimisen myötä osattujen sanojen määrä ja sanojen käyttämisessä tarvittava tieto karttuu ja b) karttuva osaaminen heijastuu suoraan henkilön tuottamien tekstien sanastoon: käytetystä mittarista riippuen joko pelkästään määrään ja vaihteluun tai määrään, vaihteluun ja laatuun. Tekstin leksikaalista diversiteettiä voidaan kasvattaa vain, jos siihen lisätään uutta sanastoa, ja siksi runsas diversiteetti jo kohtuullisen pitkissä teksteissä väistämättä edellyttää myös kielen harvinaisen sanaston käyttöä – ja siten laajaa aktiivista sanavarastoa.

### 3 Aineisto ja metodit

#### 3.1 Sadutusaineiston yleisesittely

Perusaineisto koostuu 103 puhutusta kertomuksesta, jotka on kerätty saduttamalla, litteroitu ja syötetty Excel-tietokannaksi (ks. tarkemmin luku 3.2). Sadutus on osallistava metodi, jossa sadutettavaa pyydetään kertomaan suullisesti satu tai tarina vapaasti haluamastaan aineesta tai joskus myös rajatummin tietystä aiheesta. Tässä tutkimuksessa käytössä on aihesadutus, ja aihe (*kerro satu tai tarina Unelmien päivästä*) sekä ohjeistus ovat samat kuin kirjoitettujen kertomusten leksikaalista diversiteettiä käsittelevässä verrokkitutkimuksessa (Honko 2013). Sadutusmetodia on tyypillisesti käytetty arjen työkaluna lasten parissa, mutta sitä on hyödynnetty myös tutkimuksessa ja aikuisten kanssa. (Ks. Karlsson 2013, 2014; Riihelä 2013; metodin varhaisvaiheista esim. Riihelä 1991.) Tukea juuri sadutusmetodin käyttämiseen diversiteettitutkimuksessa löytyy epäsuorasti myös aiemmasta lingvistiksestä tutkimuksesta, sillä Schmid ja Jarvis (2014) pitävät vapaasti tuotetun puheen leksikaalisen diversiteetin tarkastelua validimpana vaihtoehtona kuin muodollisten tehtävien tai kontrolloitujen kertomusten (*elicited narratives*) avulla kerätyn aineiston tarkastelua. Kaikkien sadutustuokioiden olosuhteet, tehtävänanto ja kerronnan aihe sekä aineiston käsittelytapa ovat olleet yhtenevät, mikä on kertomusten leksikaalisen diversiteetin vertailtavuuden kannalta tärkeää (Durán ym. 2004: 75).

<sup>2</sup> Ks. Bradac & Wisegarver 1984; Burroughs 1991; Watkins ym. 1995; Scott & Windsor 2000; Jarvis 2002; Dewaele & Pavlenko 2003; Wright ym. 2003; Durán ym. 2004: 234; Malvern ym. 2004; Unsworth 2004: 317–318, 322; Tidball & Treffers-Daller 2007; Yu 2010; Jarvis 2013a,b; Honko 2013; Gregori-Signes & Clavel-Arroitia 2015.

TAULUKKO 1. Sadutettujen kertomusten määrä. Taulukossa on esitetty kertomusten kokonaismäärät ryhmittäin. Sulkeissa on lisäksi esitetty asetetun pituusehdon täyttävien ja siten analyysiin nostettujen kertomusten kokonaismäärä.

	S1	S2	S1	S2	yht.
pojat	5 (4)	20 (19)	8	16	49 (47)
tytöt	4	25 (23)	15	10	54 (52)
yht.	9	45	23	26	103 (99)

TAULUKKO 2. Sadutettujen kertomusten kokonaissanamäärät.

	min	maks.	ka	kh	md
saneita	57	1380	305	245	227
eri sanoja	14	312	107	57,6	93

Sadutustuokion keskimääräinen kesto on noin 20 minuuttia.

Lyhyimmät, alle 50 saneen kertomukset ( $n = 4$ ) on poistettu aineistosta ennen analyysia. 50 saneen rajaa on pidetty tulosten luotettavuuden kannalta turvallisena vaihtoehtona, ja se vastaa kirjoitettujen kertomusten analysoinnissa tehtyä rajausta (ks. Durán ym. 2004: 228; Honko 2013: 363). Taulukossa 1 on eritelty tois- ja kolmasluokkalaisten, tyttöjen ja poikien sekä ensikielisten (S1) ja suomi toisena kielenä -oppijoiden (S2) aineisto. Kokonaismäärä (99) on tilastollisten analyysien kannalta riittävä. Alaryhmät (erityisesti ensikieliset suomenpuhujat,  $n = 32$ ) ovat kuitenkin pienet ja jokaiselta puhujalta on tutkittu vain yhden sadutuskerran aineisto, minkä vuoksi tulosten yleistämisessä täytyy noudattaa varovaisuutta. S2-ryhmän lapsissa ei ole maassaoloajan ( $\geq 2$  vuotta) ja koulunkäynnin vaiheen (yleisopetuksen 2.-3. lk.) perusteella aivan alkeistason suomenoppijoita.

Vaikka tehtävänanto ja saduttaja olivat kaikissa sadutustuokioissa samat, tuokioiden kulussa oli paljon vaihtelua: osa lapsista eteni kerronnassa yksittäisin sanoin ja lausekkein, osa kertoi vuolaasti ja pitkään. Sadutustuokion vuorojäsenystä havainnollistavat liitteiden 1 ja 4 esimerkkilitteraattit ja kertomusten pituuseroja taulukko 2, johon on koottu kertomusten sanemäärän osoittavat tunnusluvut (keskipituus, suurin ja pienin pituus, keskihajonta, mediaani). Lapsen puhe on lähtökohtaisesti sisällytetty kertomukseen kokonaissuudessaan (ks. aineiston käsittely ja poistot tarkemmin luku 3.2). Pisin tietokantaistettu kertomus on 1 380 ja lyhin 57 sanetta, keskipituus 305 ja keskihajonta 245 sanetta, eli hajonta on suuri. Aineiston kokonaissanemäärä on 30 641 ja eri sanojen määrä 2 229. Sadutusaineiston kertomukset ovat huomattavasti pitempiä kuin samanikäisten lasten samalla tehtävänannolla kirjoittamat

tekstit (2. vuosiluokan keskipituus 53 ja kolmannen 83 sanetta, kun kirjoitusai-  
ka oli n. 45 minuutin, ei tarkkaa aikarajausta).

### 3.2 Aineiston käsittely: kontekstina puhuttu kieli

Puhe ei lukupuhunutta lukuun ottamatta yleensä jakaudu siististi peräkkäin lausuttujen kokonaisten sanojen muodostamiksi lauseiksi, vaan keskustelu etenee toisiinsa lomittuvina vuoroina, joihin voi sisältyä kesken jääneitä ilmauksia, epäröintejä, toistoja ja korjauksia. Sanojen rajat voivat kadota tai olla häilyviä, jolloin syntyy kahden tai useamman sanan yhteensulautumia. Käytetyissä ilmauksissa ja niiden muodossa voi olla vaihtelua sekä eri yksilöiden välillä että samankin yksilön puheen eri kohdissa ja eri puhetilanteissa. Lisäksi sisältöjä ja merkityksiä voidaan rakentaa yhdessä tai kierrättää toisten puheesta.

Sadutus toisaalta poikkeaa arkisesta vuorovaikutustilanteesta epäsymmetrisyydellään: sadutuksessa on yleensä läsnä tilannetta ohjaava aikuinen sekä lapsi. Vuorovaikutuksessa ei pyritä tasaiseen vuorotteluun, vaan aikuisen perustehtävä on alkuorientaation ja -ohjeistuksen jälkeen tukea lapsen kertomista mutta antaa ensisijainen puhetila lapselle ja välttää kertomisen ohjailua. (Karlsson 2014.) Saduttaja välttää tuomasta kerrontaan omia aineksiaan kuten juonirakenteita tai valmiita ilmauksia, mikä vähentää sekä vuorojen kerrostuneisuutta (kierrättämistä, yhdessä rakennettuja lausekkeita) että päällekkäispuhunutta ja kesken jääneitä ilmauksia (liite 1; luku 4.3; myös Honko 2017).

Eri tekstien diversiteettiarvojen vertailu edellyttää saman mittarin käyttämistä sekä tietokannaksi syötettäessä aineiston samanlaista käsittelytapaa eli tekstin kirjoitusasun yhdenmukaistamista ja mittarissa käytettävän leksikaalisen yksikön määrittämistä (Durán ym. 2004: 228). Tässä tutkimuksessa litteroidut ja toisen kuuntelijan tarkistamat sadutustuokioidut on syötetty Excel-tietokannaksi seuraavien periaatteiden mukaan:

1. Lapsen tuottama puhe on tuotu Excel-tietokantaan (kukin sane omalle rivilleen) ja lemmattu kokonaisuudessaan sadutuksen orientaatio- ja lopetusvaihetta lukuun ottamatta.
2. Tietokannan perusyksikkö leksikaalisen diversiteetin tarkastelussa on perusmuotoinen sana (lemma), jota kirjoituksessa yleensä vastaa yksi sanavälein erotettu yksikkö (*pelata, nopeasti, hipi hiljaa*).
3. Taivutusmuodot ja foneettiset sekä murteelliset varieteetit (*pelata, be-laa, pelattii; simmottii, semmosii*) kuuluvat samaan sanaan yleiskielisen varieteetin kanssa ja on lemmattu samalla tavalla (**pelata, semmoinen**). Menettelytapa mahdollistaa tulosten vertaamisen lasten kirjoitetun kielien korpuksesta havaittuun.<sup>3</sup>

<sup>3</sup> On kuitenkin syytä huomata, että esimerkiksi perosoonapronominien tapauksessa puhe-

4. Epäröintiäännähdyksiä tai sanan toistettuja osia ei ole lemmattu (*s-sitten* > **sitten** *niin et ää-* > **niin, että**; *tul- mentiin* > **mennä, hää- juhla- vaatteet > **juhlavaate**). Toisinaan lapsen vuorosta voi päätellä, että epäroinnilla on myös sisällön rajoittamisen funktio (*Ara- Irakii; sitten se Ant- se poika otti*).**
5. Kokonaisten sanojen toistamisella on kertomuksissa usein selkeä sisällöllinen funktio (*Kuusi oli hyvin hyvin iloinen*). Siksi toistetut sanat on lemmattu erikseen, kuten kirjoitetun kielen aineistossakin.
6. Monisanaiset erisnimet ja kiteytyneet ilmaukset on lemmattu yhtenä sanana.
7. Sanojen yhteen sulautumat ( $f = 436$ , ks. myös ISK 2004: § 140) kasautuvat aineistossa samoille puhujille, ja ne on eroteltu yleiskielen mukaan kuten verrokkikorpuksessakin (*son* > **se, olla, mostettiin** > **me, ostaa**; *sille / sillai / sillei* > **sillä lailla, miksei** > **miksei**).
8. Epäröinti-ilmauksia kuten öö-, ä- ja dialogipartikkelia *mm* ei ole lemmattu.
9. Sanat on raakalemmattu aakkostetusta sanalistasta. Kaikki monitulkintaiset muodot (esim. *et* > **että** tai **ei**; *leikkii* > **leikki** tai **leikkiä**; *ankkuja* > **ankka, punkkoja** > **punkka**) on merkitty lemmauksen yhteydessä värikoodilla ja lemmattu lopullisesti juoksevasta tekstistä esiintymisympäristön perusteella.
10. Kokonaan tunnistamattomat sanat (esim. hyvin hiljaisesti äännetyt tai häiriöäänänen peittämät) on jätetty analyysin ulkopuolelle. Tällaisia sanahahmoja koko aineistossa on kuitenkin vain noin 20 ja osuus aineistosta marginaalinen.

### 3.3 Kirjoitetun kielen verrokkiaineisto ja kyselyaineistot

Artikkelin ensimmäisessä analyysiosiossa (luku 4.1) puhutun kielen aineiston leksikaalista diversiteettiä peilataan kirjoitetun kielen leksikaaliseen diversiteettiin. Kirjoitelma-aineisto on kerätty, lemmattu ja analysoitu jo aiemmin tämän tutkimuksen kanssa yhteneviä periaatteita noudattaen. Aineiston määrä on koottu taulukkoon 3 kirjoittajaryhmittäin.

Yksilöllisinä vertailumuuttujina aineiston tarkastelussa käytetään lapsen kielitaitoa, ensikieltä, sukupuolta tai ikää. Kielitaidon yleistaso on mitattu summamuuttujalla, joka koostuu kyselytietona kerätyistä kielitaitoarvioista

---

kieliset muodot ovat korpuksessa yleiskielisiä muotoja yleisempiä (esim. *mä*  $f = 960$ , *minä*  $f = 604$ ; *sä*  $f = 40$ , *sinä*  $f = 15$ ).

TAULUKKO 3. Kirjoitelma-aineiston kertomusten määrä. Taulukossa on esitetty pituusehdon täyttävien kertomusten kokonaismäärät ryhmittäin.

	S1	S2	S1	S2	yht.
pojat	20	10	37	19	86
tytöt	47	13	69	25	154
yht.	67	23	106	44	240

kolmella kielitaidon osa-alueella: puhuminen, kirjoittaminen ja sanasto. Summamuuuttajat on koostettu sekä lapsen itsearvioinneista että lapsen opettajalta pyydetyistä kielitaitoarvioista. Opettajan arviot ovat 5-portaisia (mahdolliset arvot 1–5, vaihteluväli summamuuuttujan arvoissa 2,67–5,0) lapsen arviot 3-portaisia (mahdolliset arvot 1–3). Sekä opettajat että lapset ovat arvioineet myös lapsen mahdollisia erityisiä kielellisiä vaikeuksia. Sanastollisia taitoja on lisäksi arvioitu erillisellä sanastotestillä, joka mittaa sanaston reseptiivistä ja produktiivista sanastonhallintaa eri yleisyystasoilla.<sup>4</sup>

### 3.4 Leksikaalisen diversiteetin mittari

Leksikaalisen diversiteetin arvioinnissa käytetyissä mittareissa on sekä yhteneväisyyksiä että eroja. Lähes kaikki mittarit rakentuvat tavalla tai toisella tarkasteltavan tekstin eri lekseemien ja saneiden suhteen varaan ja yksinkertaisimmillaan vain siihen kuten paljon käytetty TTR eli *type-token-ratio* (eri sanojen määrä jaettuna saneiden määrällä). TTR-pohjaiset mittarit ovat kuitenkin herkkiä verrattavien tekstien tai tekstiaineistojen kokoeroille ja siten epäluotettavia heterogeenisessä aineistossa.<sup>5</sup> Sen takia on pyritty yhä tarkempien ja tekstipituuden vaihtelun paremmin sietävien mittareiden kehittelyyn (esim. D: Malvern & Richards 1997; MTL: McCarthy 2005; McCarthy & Jarvis 2007) sekä näiden mittarien vertailevaan validointiin.<sup>6</sup>

Tässä tutkimuksessa leksikaalisen diversiteetin mittariksi valittiin sama SOP-indeksi (*sums of probabilities*), jota on käytetty myös aiemmin tehdystä kirjoitettujen kertomusten leksikaalisen diversiteetin arvioinnissa (Honko 2013). SOP:n Excel-version etuna on pidetty tarkkuutta: aineisto on analyysissa mukana kokonaisuudessaan. SOP pohjautuu Malvernin ja Richardsin kehittä-

<sup>4</sup> Ks. yksityiskohtaisempi selostus sekä taustatietolomakkeessa kysytyistä tiedoista että sanastotestistä Honko 2013.

<sup>5</sup> Pitkissä teksteissä sanatoisteisuus on yleensä luonnostaan lyhyitä suurempi, sillä jo sanatoisteisuuden pysyminen tasaisena edellyttäisi jatkuvasti uusien sanojen lisäämistä tekstiin samassa suhteessa kokonaissanemäärän lisääntymisen kanssa (Malvern ym. 2004: 124).

<sup>6</sup> Jarvis 2002; Malvern & Richards 2002; Koizumi & In’Nami 2012; Deboer 2014; Choi & Jeong 2016.

mään vocd-instrumenttiin, jonka tuottamaa D-arvoa (D-indeksiä) versioineen on hyödynnetty suurimmassa osassa 2000-luvun taitteessa leksikaalista diversiteettiä tarkastelleista tutkimuksista (ks. tarkemmin Malvern & Richards 1997, 2002; McKee ym. 2000; Malvern ym. 2004; Durán ym. 2004). SOP, kuten D:kin, mahdollistaa hyvinkin erimittaisten tekstien vertaamisen (mm. McCarthy & Jarvis 2007: 460).<sup>7</sup> SOP-indeksin saamiseksi tekstin kullekin eri sanalle laskeaan esiintymistodennäköisyys valitun teoreettisen otoskoon tai otoskokojen avulla ja analyysin tulos ilmoitetaan kaikkien todennäköisyyksien summana.

Sadutusaineiston tarkastelussa teoreettiseksi otoskooksi asetettiin verrokkitutkimuksen kirjoitettujen tekstien aineistoa vastaten 42 tekstisanaa. (Teoreettisen otoskoon määrittymisestä ja SOP:n sekä vocD:n vertailusta ks. myös Honko 2013: 175–178.) Excel-taulukon (liite 2) esimerkki tarkentaa SOP-indeksin laskentaperiaatteen. SOP laskettiin erikseen jokaiselle sadutuskerptomukselle SOP-riviarvojen summana. SOP-riviarvo puolestaan on määritelty tekstinäytteen jokaiselle leksikaaliselle yksikölle eli eri sanalle. SOP-riviarvo ( $p(F > 0)$ ) kertoo todennäköisyyden, jolla kyseinen leksikaalinen yksikkö esiintyi vähintään kerran teoreettisen otoskoon mittaisessa kyseisen tekstin näytteessä. Pitkässä tekstissä yksittäiset SOP-riviarvot ovat suhteellisen pieniä mutta yhteenlaskettavia arvoja saadaan enemmän. Liitteessä 2 on esitetty aineiston runsaimman (SOP  $\approx 34,61$ ) sekä vähäisimmän (SOP  $\approx 10,60$ ) leksikaalisen diversiteetin tekstien SOP-laskentataulukot. Jälkimmäisen esimerkin leksikaalinen diversiteetti on tutkitussa aineistossa kuitenkin poikkeuksellisen vähäinen (seuraavaksi pienin SOP-arvo on 21,24).

### 3.5 Tilastolliset analyysit ja hajonta-/sirontakuviot

Muuttujan arvojen normaalijakautuneisuuden testaamiseen on käytetty Shapiro-Wilkin testiä ja vertailtavien otosten varianssien yhtäsuuruuden testaamiseen F-testiä (*F test to compare two variances*). SOP-arvojen jakautuminen aineistossa ei noudata normaalijakaumaa (Shapiro-Wilk,  $p < 0,001$ ) vaan vinoutuu hieman oikealle, ja lisäksi vertailtavien otosten (kuten puhutun ja kirjoitetun kertomusaineiston) SOP-arvojen varianssien yhtäsuuruus ei toteudu (*ratio of variances*  $\approx 0,049$ ,  $p < 0,001$ ). Siksi analyyseissa on käytetty epäparametrisiä menetelmiä.

Koska parametrusten testien edellytykset eivät ole voimassa, muuttujien arvojen (lineaarista) riippuvuutta on tutkittu Spearmanin järjestyskorrelaatiolla. Eri otosten muuttujan arvojen vertailussa, kuten puhutun ja kirjoitetun aineiston SOP-arvojen vertailussa on käytetty Mann-Whitneyn U-testiä.

<sup>7</sup> Lupaavina mittareina on pidetty myös tasapitkien tekstikatkelmien analysointiin perustuvaa MSTTR-indeksiä sekä TTR:n vakiointiin perustuvaa MTLD-indeksiä, joilla kuitenkin on omat rajoituksensa ja joiden käyttäminen ei tässä tutkimuksessa olisi mahdollistanut vertailua kirjoitettujen kertomusten verrokkiaineistoon (ks. tarkemmin esim. Jarvis 2013a: 94).

Luokiteltujen muuttujan arvojen vertailuun perustuvassa testauksessa (SOP-arvojen jakaumien vertailu) on käytetty khiin neliö -testiä.

Tilastollisten analyysien merkitsevyytasot on asetettu seuraavasti:  $p \leq 0,05$  = tilastollisesti melkein merkitsevä,  $p \leq 0,01$  = tilastollisesti merkitsevä,  $p \leq 0,001$  tilastollisesti erittäin merkitsevä. (Ts. virheen todennäköisyys nolahypoteesin kumoamisessa on asetetuilla merkitsevyytasooilla korkeintaan 5 %, 1 % ja 0,1 %.) Tilastollisten analyysien ohella muuttujien välisiä suhteita on systemaattisesti arvioitu myös tarkastelemalla hajonta-/sironnakuvioita, jotka voivat paljastaa tarkasteltavien muuttujien epälineaarisen yhteyden ja viitata kompleksisiin, usean muuttujan yhteisvaikutuksesta muodostuviin riippuvuussuhteisiin aineistossa. Vain pieni osa kuvioista on kuitenkin tilan säästämiseksi nostettu artikkeliin.

## 4 Tulokset

Analyysiosio jakautuu kolmeen alalukuun. Ensimmäisessä luvussa käsitellään modaliteetin yhteyttä kertomuksen leksikaaliseen diversiteettiin ja verrataan toisiinsa lasten puhumalla ja kirjoittamalla tuottamien kertomusten leksikaalista diversiteettiä.<sup>8</sup> Toisessa alaluvussa tarkastellaan yksilöllisten tekijöiden yhteyttä puhuttujen kertomusten leksikaaliseen diversiteettiin. Keskeisiä vertailumuuttujia on viisi: sadutettavan lapsen sanastolliset taidot, yleinen kielitaitotaso, kielitausta (S1/S2), sukupuoli sekä ikä. Aikuisen läsnäolo sadutusvuorovaikutuksessa on lähtökohtaisesti kiinteämpi kuin pelkkään alkuohjeistukseen perustuvassa kirjoittamistilanteessa, mikä saattaa vaikuttaa sadutettavan lapsen käyttämiin kielellisiin ilmauksiin ja sitä kautta leksikaaliseen diversiteettiin. Siksi kolmannessa analyysiluvussa tarkastellaan tutkijan puheesta kierrätetyn sanaston mahdollista vaikutusta puhuttujen kertomusten leksikaaliseen diversiteettiin.

### 4.1 Modaliteetin vaikutus leksikaaliseen diversiteettiin

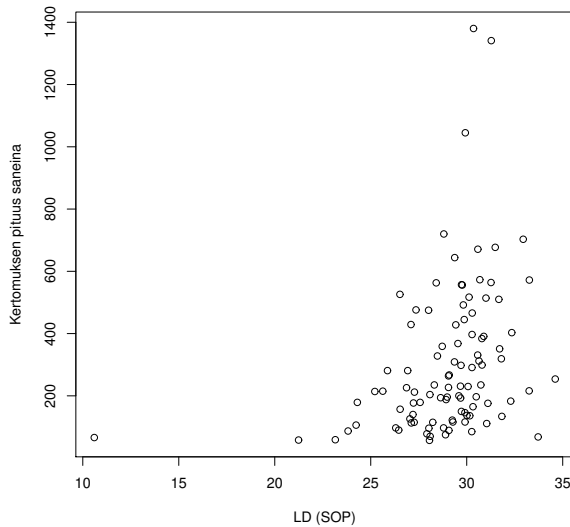
Vaikka puheen ja kirjoituksen erot ovat sähköisen viestinnän myötä kaventu-neet, kirjoitettu teksti on usein puhetta prosessoidumpaa ja sen leksikaalinen diversiteetti on todettu puhetta runsaammaksi (ks. Strömqvist ym. 2002; Kuiken & Vedder 2012: 368). Myös sadutusaineistossa puhutun kielen keskimääräinen leksikaalinen diversiteetti on jonkin verran kirjoitettujen kertomusten verokkiaineistoa vähäisempi (taulukko 4 ja liite 3). Ero on tilastollisesti erittäin merkitsevä (Mann-Whitneyn U-testi,  $W = 108$ ,  $p < 0,001$ ).

<sup>8</sup> Kirjoitetun kielen aineistoa koskevat tulokset on julkaistu aiemmin osana väitöstutkimustani (Honko 2013), jossa tutkimusasetelma (kertomisen ohjeistus, aineiston käsittely- ja analysointitapa) oli modaliteettia lukuun ottamatta sama kuin nyt analysoitavassa sadutusaineistossa.



TAULUKKO 4. Leksikaalisen diversiteetin SOP-jakauma.

	min.	maks.	ka	kh	md
sadutusaineisto (n = 99)	11,00	34,61	28,96	2,94	29,43
ikäverrokkien kirjoitelmat (n = 239)	20,48	38,37	30,22	2,88	30,46
kirjoitelmat, reaaliseurata (n = 7)	24,06	32,27	28,53	2,74	28,52



KUVIO 1. Kertomuksen leksikaalisen diversiteetin suhde sen kokonaissanamäärään.

Seitsemältä lapselta on käytettävissä sekä sadutusaineisto että saman lukukauden aikana kirjoitettu kertomus. Aineisto on hyvin pieni, mutta suuntaantava korrelaatio puhutun ja kirjoitetun kertomustekstin yksilöllisessä leksikaalisessa diversiteetissä on korkea ja tilastollisesti melkein merkitsevä:  $r = 0,821$ ,  $p = 0,023$ ,  $n = 7$  (käytössä Spearmanin järjestyskorrelaatioanalyysi  $r_s$ ). Tämä tarkoittaa, että leksikaalisen diversiteetin yksilöllinen taso näiden puhujien joukossa on suhteellisen pysyvä modaliteetista toiseen siirryttäessä. Jatkossa ilmiön tarkasteluun tarvittaisiin kuitenkin suurempi aineisto.

Puheaineistossa leksikaalisen diversiteetin ja kokonaissanamäärän välillä on tilastollisesti merkitsevä mutta matala korrelatiivinen yhteys ( $r_s = 0,422$ , saneet,  $p < 0,001$ ). Kirjoitetuissa teksteissä leksikaalisen diversiteetin ja kokonaissanamäärän välillä on alaluokilla havaittu ainoastaan heikko positiivinen yhteys (Honko 2013: 378–380). Puheaineistossa leksikaalisella diversiteetillä on yhteys myös sadutuksen eri sanojen määrään ( $r_s = 0,555$ ,  $p < 0,001$ ). Korrelaatioanalyysin tulos ja sirontakuviot (kuviot 1) kertovat, että leksikaalinen

diversiteetti ei kuitenkaan ole kokonaissanamäärän funktio: hyvin eripituiset kertomukset ovat saaneet suurita diversiteetti-arvoja.

## 4.2 Yksilöllisten tekijöiden vaikutus leksikaaliseen diversiteettiin

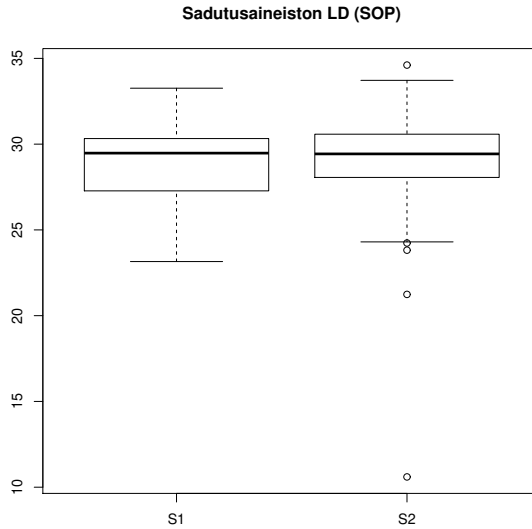
**a) sanastolliset taidot** Sanastollisia taitoja on arvioitu strukturoidulla testillä (ST), joka mittaa sekä produktiivista että reseptiivistä osaamista. Lisäksi lapsen opettajalta (S1/S2) on pyydetty holistinen arvio sanastollisista taidoista.

Sadutusaineistossa leksikaalinen diversiteetti ja sanastotestillä arvioidut sanastolliset taidot eivät korreloi ( $r_S = 0,146$ ,  $p = 0,1714$ ,  $n = 89$ ), eivät myöskään pelkästään produktiivis- tai reseptiivispainotteisten tehtävien avulla arvioituna. Sirontakuvion perusteella muuttujien välillä ei ole muutamaakaan säännöllistä yhteyttä. Leksikaalisella diversiteetillä on tosin tilastollisesti melkein merkitsevä yhteys opettajan antamaan holistiseen 5-portaiseen arvioon lapsen sanastollisista taidoista ( $r_S = 0,349$ ,  $p = 0,025$ ,  $n = 41$ ). Korrelaatiokertoimen arvo on kuitenkin jälleen niin matala, että yksilötason arvioinnin kannalta yhteyttä ei voi pitää merkityksellisenä.

Kirjoitelmakorpuksessa laskettu kertomuksen leksikaalinen diversiteetti sen sijaan korreloi kaikilla vuosiluokilla kirjoittajan sanastollisiin taitoihin (ST). Korrelatiivinen yhteys on vain kohtalainen ( $r_S = 0,433-0,537$ ) mutta suurehkoissa ryhmässä tilastollisesti erittäin merkitsevä ( $p < 0,001$ ) (Honko 2013: 383). Myös yhteys opettajan arvioimaan kielitaidon yleistason (sanastollisen taidon, kirjoitustaidon ja puhetaidon holististen arvioiden indeksi) on tilastollisesti merkitsevä, mutta muuttujien selitysvoima toisiinsa nähden hyvin alhainen ( $r_S = 0,302$ ,  $p = 0,002$ ,  $n = 104$ ).

Kiinnostavaa on, että sadutetun sanaston leksikaalisella diversiteetillä löytyy yllättävä yhteys sanastolliseen osaamiseen 3 vuotta myöhemmin ( $r_S = 0,419$ ,  $p = 0,008$ ,  $n = 39$ ). Käytetty testi (ST2) on mukautettu uuteen ikätasoon (5.-6. luokka) ja siinä painottuu aiempaa testiä (ST) enemmän sanavaraston laajuuden arviointi. Peruseriaatteet ja erottelukyky ovat kuitenkin samat (ks. Honko 2013). Alaluokilla lasten erot tekstitaidoissa voivat kuitenkin vaikuttaa sanastotestin tulokseen, mikä selittäisi yhteyden puuttumista juuri alaluokilla.

**b) yleinen kielitaitotaso** S2-ryhmässä sadutusaineistossa leksikaalisella diversiteetillä on heikosti merkitsevä matala korrelatiivinen yhteys sekä opettajan ( $r_S = 0,32$ ,  $p = 0,038$ ,  $n = 41$ ) että lapsen itsensä arvioimaan kielitaidon yleistason ( $r_S = 0,34$ ,  $p = 0,033$ ,  $n = 41$ ). Yhteyttä sen sijaan ei löydy muihin pyydettyihin taustatietoihin: edelliseen suomen kielen todistusarvosanaan (numeerinen tai numeeristettu sanallinen arvio,  $n = 28$ ), opettajan tai lapsen arvioimiin kielellisiin erityisvaikeuksiin (summamuuttuja  $n = 40$ ,  $n = 41$ ) tai Suomesa asumisen keston ( $n = 69$ ).



KUVIO 2. Kielitaidon yhteys sadutetun kertomuksen leksikaaliseen diversiteettiin.

Kielitaidon yleistason yhteys kirjoitettujen kertomusten leksikaaliseen diversiteettiin on opettajan arvioimana merkitsevä mutta heikompi kuin yhteys erillisellä testillä (ST) mitattuun sanastonhallintaan (Honko 2013). Lapsen itsearvioinnin tulos sen sijaan ei ole yhteydessä kirjoitetun kertomuksen leksikaaliseen diversiteettiin. Lisäksi kielitaidon yleisarvioinnin karkeana kuvajana on käytetty edellistä suomen kielen todistusarvosanaa (vaihteluväli 6–9) ja opettajan sekä lapsen arvioimien kielellisten erityisvaikeuksien määrää (vaihteluväli 0–9 ja 0–12) sekä maassaolon kestoa (2–11 vuotta). Maassaolon keston yhteys kielitaitoon tosin tiedetään tutkitussa ryhmässä kompleksiseksi (Honko 2013).

**c) ensikieli** Sadutusaineistossa leksikaalisessa diversiteetissä ei ilmene eroa kieliryhmien välillä (Mann-Whitney  $W = 1009$ ,  $p = 0,8461$ , ks. kuvio 2). Eroa S1- ja S2-oppilaiden leksikaalisessa diversiteetissä ei ole myöskään erikseen tyttöjen ja poikien tai tois- ja kolmasluokkalaisten sadutuksista arvioituna ( $W = 276$ ,  $p = 0,8047$ ;  $W = 233$ ,  $p = 0,9088$ ;  $W = 232$ ,  $p = 0,1249$ ;  $W = 210$ ,  $p = 0,1189$ ). Hajonta on S2-aineistossa hieman suurempi, ja ylin neljännes sijoittuu hieman ensikielisten ryhmää korkeammalle tasolle. Tulosta ei selitä ero tuottamisen runsaudessa, sillä sadutuksen sane- tai sanamäärä eivät eroa kieliryhmittäin.

Kirjoitelma-aineistossa S1-oppilaiden keskimääräinen leksikaalinen diversiteetti sen sijaan on ryhmätasolla kaikilla vuosiluokilla S2-oppilaiden tekstien diversiteettiä runsaampi (Honko 2013). Tasavälein ryhmitettyjen SOP-arvojen

ristiintaulukointi ja tarkastelu khiin neliö -testillä osoittavat, että eniten toisistaan poikkeavat jakauman ääripäät: S2-oppijoiden teksteissä on suhteessa enemmän vähäisen ja vähemmän runsaan leksikaalisen diversiteetin kertomuksia. Tähän tulokseen suhteutettuna on yllättävää, että puheaineistossa kieliryhmien leksikaalisessa diversiteetissä ei ilmene eroa.

**d) sukupuoli** Sadutusaineistossa poikien leksikaalisen diversiteetin mediaani on hieman tyttöjen leksikaalisen diversiteetin mediaania suurempi, mutta ero ryhmien välillä ei ole tilastollisesti merkitsevä (Mann-Whitney  $W = 1494$ ,  $p = 0,059$ ). Tätä selittää poikien aineiston suurempi hajonta: poikien aineistossa ovat sekä pienimmät että suurimmat diversiteettiarvot (liite 3).

Alakoululaisten kirjoittamistutkimuksissa tyttöjen kirjoittamien tekstien leksikaalinen diversiteetti on havaittu poikien kirjoittamien tekstien diversiteettiä runsaammaksi (Honko 2013; ks. myös Saarela 1997). Aikuisten puhutun kielen aineistoista tehtyjen tutkimusten tulokset ovat kuitenkin ristiriitaiset: Dewaelen ja Pavlenkon (2003: 134) tutkimuksessa naisten käyttämä sanasto on miesten sanastoa vaihtelevampaa, Singhin (2001: 260–261) ja Härnqvistin, Christiansonin, Ridingsin ja Tingsellin (2003: 191) tutkimuksessa tulos on päinvastainen.

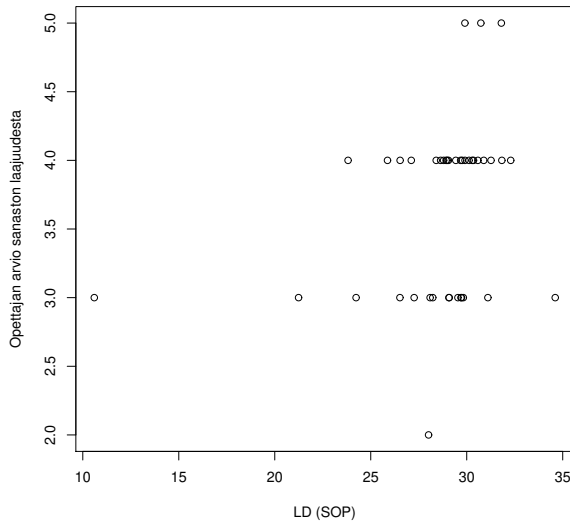
**e) ikä** Tyypillisessä kielenkehityksessä oppijan leksikko karttuu ja syvenee kouluvuosina huimasti, mikä heijastuu spontaanin tuottamisen sanastoon (Berman 2007; Pajunen 2012). Aiemman tutkimuksen perusteella leksikaalinen diversiteetti, myös SOP, erottelee kielitaidoltaan eritasoisia puhujia ja kirjoittajia paremmin varhaisemmassa kielenkehityksen ja kirjoitustaidon vaiheessa (Vermeer 2000; Jarvis 2002; Honko 2013: 370–372).

Tois- ja kolmasluokkalaisten sadutusaineistossa SOP-diversiteettiarvot eivät poikkea toisistaan (Mann-Whitney  $W = 1149$ ,  $p = 0,6019$ ). Ikäluokittain havaintoja ei ole riittävästi kaikissa luokissa (7–11 vuotta) Mann-Whitney U-testin suorittamiseen, mutta Spearmanin järjestyskorrelaatioanalyysi ja siron-takuvio paljastavat, että iän ja leksikaalisen diversiteetin välillä ei tutkimuksessa aineistossa ole korrelatiivista ( $r_s = 0,088$ ) tai muutenkaan havaittavaa säännönmukaista yhteyttä.

Kirjoitetuissa kertomuksissa leksikaalisen diversiteetin kehitys sen sijaan on kaikissa alaryhmissä vielä alakouluvaiheessa nousujohteista vaikkakaan ei täysin lineaarisesta luokkatasolta toiselle. Ero vuosiluokkien 2–3 välillä on t-testillä mitattuna tilastollisesti erittäin merkitsevä. (Honko 2013: 365–366.)

### 4.3 Kierrätetyn sanaston vaikutus leksikaaliseen diversiteettiin

Ne lapset, joiden sanastolliset taidot opettaja on arvioinut parhaiksi, ovat myös käyttäneet sadutuksessa verrattain vaihtelevaa sanastoa. Toiseen suun-



KUVIO 3. Opettajan arvioiman sanavaraston yhteys sadutetun kertomuksen leksikaaliseen diversiteettiin.

taan yhteys ei kuitenkaan päde (kuvio 3). Tulos saattaa tarkoittaa, että joidenkin oppilaiden sanastollinen osaaminen jää arjessa opettajalta piiloon ja todentamatta myös sanastotestissä (ks. luku 4.1 kohta b). On kuitenkin mahdollista, että sadutustilanteet eivät ole keskenään täysin vertailukelpoisia. Yksi selitys voisi piillä siinä, että heikompaan sanastonhallintaa on vuorovaikutustilanteessa mahdollista kielellisesti kompensoida vuorovaikutuskumppanilta kierrätetyllä sanastolla.

Toisen kielenkäyttäjän tekstien käyttäminen lähteenä voi kasvattaa leksikaalista diversiteettiä, mikä on aikaisemmassa tutkimuksessa todennettu kirjoitetun kielen kontekstissa (Gebril & Plakans 2016). Sadutusaineistossa tämä tarkoittaa sitä, että sadutettu lapsi voi omassa puheessaan kierrättää paitsi sadutustuokion ulkopuolella oppimiaan sanoja myös sen aikana aikuisen puheesta poimimaansa kielenainesta – kuten sanastoa. Sadutettujen kertomustekstien leksikaalisen diversiteetin ja kielitaustan sekä luokkatason ja muilla menetelmillä arvioidun sanastollisen osaamisen välisen riippuvuuden puuttuminen (luku 4.2) voisikin selittyä sillä, että sanastollisilta taidoiltaan heikoimmat lapset ovat vuorovaikutustilanteessa aktiivisimpia sanaston kierrättäjiä. Sen takia on tarpeen tarkastella erikseen sanaston mahdollista kierrättämistä sadutusaineistossa.

Jo sadutustuokioista kirjoitettujen keskustelulitteraattien alustava laadullinen tarkastelu osoittaa, että saduttajan rooli uuden sanaston tuojana on tutkitussa vuorovaikutusaineistossa hyvin vähäinen: alkuorientaation jälkeen

pääosa saduttajan vuoroista koostuu pelkästä dialogipartikkelista kuten *mm, joo, nii* (ks. myös liite 1) tai lyhyestä kehotuksesta tai kysymyksestä (*kerro lisää, mitä sitten tapahtui?*), johon lapsi reagoi esimerkiksi jatkamalla kertomista. Silloinkin, kun se olisi mahdollista, lapsi kierrättää aikuisen puheen kautta tarjoutuvaa sanastoa omaan puheeseensa vain harvoin. Saduttajalle (S) sen sijaan on tyypillisempää toistaa sanoja ja laajempia ilmauksia lapsen (L) puheesta (esimerkki 1).

## (1)

- 1 L: mä tulisin **kouluun** ja **tekisin** mitä opettaja olis sanonu.  
 2 (.)  
 3 S: joo.  
 4 (.)  
 5 S: sä tykkäät käydä **koulussa**.  
 6 (.)  
 7 S: kiva .hh  
 8 (..)  
 9 S: mitä koulussa tehtäis (.) semmosena päivänä.

Saduttajan rooli kielellisenä osallistujana korostuu tilanteissa, joissa lapsi tarvitsee tukea: aikuinen toistaa lapsen puhetta, toisinaan myös kokoaa tai jatkaa lapsen vuoroja tai esittää lisäkerrontaan rohkaisevia kysymyksiä (esimerkki 2).

## (2)

- 1 S: °mikäs olis sinulle **ihanin päivä**. =mitä siellä olis.°  
 2 (4 s.)  
 3 L: ai **ihanin päi**[vä].  
 4 S: [**ihanin päivä** maailmassa. =mitä siellä olis.

Kierrättämistä oli tarpeellista tutkia myös tarkemmin: sadutustuokioista kirjoitettujen litteraattien avulla etsittiin kaikki sellaiset sisältösanaluokkien sanat, joita lapsi käyttää sadutustuokiossa ensimmäisen kerran vasta aikuisen käytettyä sanaa aiemmin omassa puheessaan. Tarkastelu rajattiin eri diversiteettitasoilta poimittuun 20 kertomuksen otokseen (20 % koko aineistosta). Analyysissa huomioitiin sanan kaikki esiintymiskontekstit: esiintyminen osana yhdyssanoja ja erilaisia monisanaisia konstruktioita joko välittömästi aikuisen vuoron jälkeen tai vasta myöhemmin sadutustuokion aikana.

Näin laskettuna kierrätettyjä sanoja esiintyy lapsen puhutuissa kertomuksissa keskimäärin vähemmän kuin yksi esiintymä sadutusta kohti (vaihteluväli 0–5). Kierrätettyjen eri sanojen määrä jää vielä pienemmäksi (0–3). Kahdesatoista eli yli puolessa (60 %) tarkastelluista sadutustuokioista ei esiinny lainkaan aikuisen vuoroista kierrätettyä sanastoa. Havaintojen perusteella sanojen kierrättäminen aikuisen puheesta ei selitä aiemmissa luvuissa esitettyjä tuloksia.

Tuloksen voi olettaa olevan vahvasti sidoksissa sadutusvuorovaikutuksen luonteeseen ja heijastavan sadutusmetodin ohjeistusta: saduttajan ei kuulu tarjota lapselle kerronnan sisältöjä (Karlsson 2014). Myös tutkimuksessa aineistossa saduttajan puhe on karsittua; vuorot ovat lyhyitä ja rakentuvat niukan, pitkälti jo tehtävänannossa tai lapsen omissa ilmauksissaan käyttämän sanaston varaan. Vuorottelurakenteen tarkempi tarkastelu kuitenkin paljastaa, että sillä saattaa olla muunlaisia – leksikaaliseen diversiteettiin heijastuvia – vaikutuksia sadutettavan tuottamaan puheeseen. Sadutustuokion aikana eniten tukea tarvitsevien lasten vuorot ovat yleensä hyvin lyhyitä lausekkeita, joista puuttuu vapaalle kerronnalle tyypillinen sisällöllinen ja rakenteellinen yhtenäisyys (liite 4). Pidempiin yhtenäisiin vuoroihin perustuvassa kerronnassa sen sijaan yhtenäisyyttä luodaan muun muassa sisäisiä viittaussuhteita rakentavalla leksikaalisella toistolla, jolla voi olla myös puheen prosessointiin ja sadutustuokion vuorojäsennyksen muokkaamiseen liittyviä tehtäviä. Vuorojäsennyksen ja vuorojen rakenteen vaikutusta leksikaaliseen diversiteettiin ei ole aiemmin tutkittu, mutta systemaattinen tarkastelu on kiistatta tarpeen, mikäli puhutun kielen leksikaalista diversiteettiä jatkossakin tutkitaan keskusteluvuorovaikutusaineistoista.

## 5 Tulosten koonti ja pohdinta

Tutkimuksen tehtävänä oli selvittää, onko a) lasten sadutettujen kertomusten leksikaalisessa diversiteetissä systemaattisia eroja eri modaliteettien tai puhujaryhmien välillä ja b) voisiko leksikaalinen diversiteetti toimia kehityksellisenä mittarina vastaavia aineistoja analysoitaessa. Kirjoitetun kielen leksikaalinen diversiteetti osoittautui puheen diversiteettiä runsaammaksi, mikä tukee aiempien tutkimusten tuloksia. Myös yksilötasolla modaliteettien välinen yhteys on tarkastellussa pienessä ryhmässä havaittavissa. Kieliryhmien (L1/L2), tyttöjen ja poikien tai tois- ja kolmasluokkalaisten välillä leksikaalisessa diversiteetissä ei kuitenkaan ilmennyt eroja, ja sekä kielitaidon yleistason että sanaston hallinnan yleistason (testisuoritus tai opettajan arvio) yhteys leksikaaliseen diversiteettiin osoittautui epälineaariseksi ja heikoksi. Sadutettujen kertomusten leksikaalisella diversiteetillä ei tarkastellussa aineistossa ollut yhteyttä myöskään lapsen edelliseen suomen kielen kouluarvosanaan, opettajan listaamien kielellisten erityisvaikeuksien määrään tai Suomessa asumisen kestoon (S2-ryhmä). Tulosten perusteella leksikaalista diversiteettiä (SOP) ei voi pitää riittävän tarkkana menetelmänä yksilötason kielitaidon tai tarkemmin sanastollisten taitojen arvioimiseen puhevuorovaikutuksessa.

Myös aiemman puhevuorovaikutuksesta tehdyn tutkimuksen tulokset ovat jättäneet leksikaalisen diversiteetin arviointikäyttöön varauksia. Sadutustuokioiden vuorojäsennyksen laadullinen tarkastelu osoittaa, että eri diversi-

teettitasojen tekstit saattavat poiketa toistaan laadullisesti muutoinkin kuin sanaston osalta. Vaikka ryhmätasolla runsas leksikaalinen diversiteetti yhdistyykin tuottamisen runsauteen, yhteys ei ole lineaarinen: Runsaan leksikaalisen diversiteetin kertomukset näyttävät usein koostuvan lyhyistä ja rakenteeltaan yksinkertaisista vuoroista, jotka on tuotettu ikään kuin reaktiona saduttajan rohkaisevaan viestintään, vähäisen diversiteetin kertomuksissa puolestaan on paljon pitkiä, spontaanisti tuotettuja vuoroja. Yksilötasolla kertomisen niukkuus voi siksi selittää runsasta leksikaalista diversiteettiä ja yhtenäisen, vuolaan kertomuksen leksikaalinen diversiteetti puolestaan olla vähäinen. Vaikka sadun kertomisessa on monologimaisia piirteitä, sadutustuokio on kuitenkin vuorovaikutustilanne ja lapselle mahdollisesti myös uudenlainen ja jännittävä tilanne. Puhevuorovaikutukseen osallistumiseen sadutuksessa tarvitaan sanastollisen osaamisen lisäksi paljon muutakin – tehtävänantoon reagoimisen lisäksi esimerkiksi avoimuutta ja uskallusta sekä henkilökohtainen tarve osallisuuteen. Siksi sadutettu puhe väistämättä antaa kapean kuvan sadutettavan kokonaiskielitaidosta.

Tämän tutkimuksen perusteella näyttää siltä, että aikuisen tuen suora vaikutus leksikaaliseen diversiteettiin sadutustuokioissa on kuitenkin hyvin pieni: lasten kerronnassa esiintyy hyvin vähän aikuiselta kierrätettyjä sanoja, mikä johtuu osittain aikuisen pelkistetyistä vuoroista (Honko 2017). Koska saduttajan vuoroista kierrätettyä sanastoa esiintyy vähän, kierrättämisen vaikutus leksikaaliseen diversiteettiin on minimaalinen. Tukea tarvitsevien lasten niukka kerronta ja lyhyet vuorot sen sijaan johtavat usein välillisesti suhteellisen runsaaseen diversiteettiin, kun esimerkiksi tekstiä sidostava leksikaalinen toisto ja funktiosanojen käyttö on niukkaa.

Oletuksia siitä, että leksikaalinen diversiteetti voisi toimia kehityksellisenä mittarina ja jopa diagnostisena työkaluna on kritisoitu muun muassa kielinoppimisprosessin yksinkertaistamisesta ja kielenkäytön tilanteisuuden sivuuttamisesta: Kaikki sanastollinen osaaminen ei esimerkiksi ole luonteeltaan määrällistä eikä siis heijastu sanavaraston kasvuna. Kaikissa tilanteissa ja tekstilajeissa ei myöskään tarvitse tai edes kannata käyttää samanlaista kieltä – ja siten myöskään samalla tavalla varioivaa sanastoa. On selvää, että esimerkiksi ääneen prosessointi sanoja toistamalla yksittäistapauksissa laskee leksikaalista diversiteettiä (*mä olin- mä- mä- mä oon aina nii tehny*). Lisäksi puhutussa kielessä leksikaalisen tiivyyden haittapuolet korostuvat, minkä vuoksi hyvin runsas leksikaalinen diversiteetti ei aina ole vuorovaikutuksessa etu (ks. myös Broeder ym. 1993: 149). Kuten tässä artikkelissa aiemmin esillä olleiden tutkimusten perusteella voidaan todeta, pidemmällä edenneet kielinoppijat kuitenkin tyypillisesti toistavat sanastoa vähemmän ja heidän tuottamissaan teksteissä leksikaalinen diversiteetti on suhteessa runsaampi kuin verrokki-ryhmissä (ei-äidinkielliset tai alkeistason oppijat). Heterogeenisissa aineistossa leksikaalinen diversiteetti voikin toimia suuntaa-antavana kielitaidon mittari-



na, mikäli aineistoa on riittävästi ja se on tuotettu yhdenmukaisilla tavoilla. Jatkoissa tarkastelua on kuitenkin syvennettävä laajemmalla aineistolla ja tarpeen mukaan myös monimuuttuja-analyysia käyttäen sen selvittämiseksi, millaisia kerrannaisvaikutuksia tai mahdollisesti myös toisensa pois rajaavia vaikutuksia eri muuttujilla on suhteessa leksikaaliseen diversiteettiin.

Leksikaalisen diversiteetin menetelmällisessä tutkimuksessa on tyypillisesti keskitytty siihen, kuinka hyvin käytetty mittari ennustaa kielellistä suoriutumista verrattuna johonkin toisentyypiseen mittariin. Oletuksena on ollut, että leksikaalisen diversiteetin määritelmään joka tapauksessa kuuluu ainakin sanojen valikoiman laajuus (*range*) ja vaihtelu (*variety*) (McCarthy & Jarvis 2007: 459). Aivan viime vuosina on kuitenkin havahduttu huomaamaan, että edes diversiteettimittarin tilastollisesti merkitsevä erottelukyky ei takaa sen käytettävyyttä. Samalla on ryhdytty tarkemmin perehtymään leksikaaliseen diversiteettiin tutkittavana ilmiönä, toisin sanoen tutkimaan, mistä – ja erityisesti mistä muusta kuin sanatoisteisuudesta tai tarkasteltavan tekstin kokonaissanamäärästä – leksikaalinen diversiteetti mahdollisesti koostuu (Jarvis 2013b).

Näyttää siltä, että leksikaalisen diversiteetin määrittäminen pelkästään tekstin sisältämien sanojen vaihteluksi on pelkistys, joka jättää huomiotta muun muassa sanojen ominaislaadun. Tämä ja monet muut seikat saattavat kuitenkin olennaisesti vaikuttaa tekstin vastaanottajan kokemukseen leksikaalisesta diversiteetistä, sanojen määrän ja vaihtelevuuden ohella (Crossley ym. 2011a,b; Jarvis 2013b). Tämän tutkimuksen perusteella leksikaalisen diversiteetin mahdollisissa jatkotutkimuksissa vaaditaan aiemman diversiteettitutkimuksen perinteestä irrottautumista ja kehitystyötä kahdella tavalla: diversiteetin määrittelyssä ja puhevuorovaikutuksen ominaislaadun huomioon ottamisessa.

## Kiitokset

Tutkimus on osa myöhemmän kielenkehityksen tarkasteluun keskittyvää hanketta (ks. esim. Pajunen 2012). Kiitän Tampereen yliopistoa, Tampereen Yliopiston Tukisäätiötä, Jyväskylän yliopistoa, artikkelin kahta nimetöntä arvioijaa sekä aivan erityisesti tutkimusavustaja Minna Bogdanoffia tuesta tämän tutkimusartikkelin syntyprosessin eri vaiheissa aineistonkeruusta viimeistelyyn.

## Lähteet

- Alderson, J. C. 2005. *Diagnosing foreign language proficiency: the interface between learning and assessment*. London: Continuum.
- Berman, R. 2007. Developing language knowledge and language use across adolescence. Teoksessa E. Hoff & M. Shatz (toim.) *Handbook of language development*. London: Blackwell, 346–367.

- Berman, R. & L. Verhoeven 2002. Cross-linguistic perspectives on the development of text-production abilities: speech and writing. *Written Language and Literacy*, 5 (1), 1-43.
- Bradac, J. J. & R. Wisegarver 1984. Ascribed status, lexical diversity and accent: determinants of perceived status solidarity, and control of speech style. *Journal of Language and Social Psychology*, 3 (4), 239-255.
- Broeder, P., G. Extra & R. van Hout 1993. Richness and variety in the developing lexicon. Teoksessa C. Perdue (toim.) *Adult language acquisition: cross-linguistic perspectives, volume I: Field methods*. Cambridge: Cambridge University Press, 145-232.
- Burroughs, E. I. 1991. Lexical diversity in listeners' judgments of children. *Perception and Motor Skills*, 73 (1), 19-22.
- Cain, K., J. Oakhill & K. Lemmon 2004. Individual differences in the inference of word meanings from context: the influence of reading comprehension, vocabulary knowledge and memory capacity. *Journal of Educational Psychology*, 96 (4), 671-681.
- Carroll, J. B. 1938. Diversity of vocabulary and the harmonic series law of word-frequency distribution. *Psychological Record*, 2, 379-386.
- Castañeda-Jiménez, G. & S. Jarvis 2014. Exploring lexical diversity in second language Spanish. Teoksessa K. Geeslin (toim.) *The handbook of Spanish second language acquisition*. Hoboken (N. Y.): Wiley, 498-513.
- Choi, W. & H. Jeong 2016. Finding an appropriate lexical diversity measurement for a small-sized corpus and its application to a comparative study of L2 learners' writings. *Multimedia Tools and Applications*, 75 (21), 13015-13022.
- Crossley, S. A., T. Salsbury & D. S. McNamara 2011a. Predicting the proficiency level of language learners using lexical indices. *Language Testing*, 29 (2), 243-263.
- Crossley, S. A., T. Salsbury, D. S. McNamara & S. Jarvis 2011b. What is lexical proficiency? Some answers from computational models of speech data. *TESOL Quarterly*, 45 (1), 182-193.
- Deboer, F. 2014. Evaluating the comparability of two measures of lexical diversity. *System*, 47, 139-145.
- Dewaele, J.-M. & A. Pavlenko 2003. Productivity and lexical diversity in native and non-native speech: a study of cross-cultural effects. Teoksessa V. J. Cook (toim.) *L2 effects on the L1*. Clevedon: Multilingual Matters, 120-141.
- Dockrell, J. E. & D. Messer 2004. Lexical acquisition in the early school years. Teoksessa R. Berman (toim.) *Language development across childhood and adolescence: psycholinguistic and crosslinguistic perspectives*. Amsterdam: John Benjamins, 35-52.
- Durán, P., D. Malvern, B. Richards & N. Chipere 2004. Developmental trends in lexical diversity. *Applied Linguistics*, 25 (2), 220-242.
- Ellis, C., Y. F. Holt & T. West 2015. Lexical diversity in Parkinson's disease. *Journal of Clinical Movement Disorders*, 2 (5). DOI: 10.1186/s40734-015-0017-4. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4710975/>.
- Gebriel, A. & L. Plakans 2016. Source-based tasks in academic writing assessment: lexical diversity, textual borrowing and proficiency. *Journal of English for Academic Purposes*, 24, 78-88.

- Gregori-Signes, C. & B. Clavel-Arroitia 2015. Analysing lexical density and lexical diversity in university students' written discourse. *Procedia Social and Behavioral Sciences*, 24 (198), 546–556.
- Honko, M. 2013. *Alakouluikäisten leksikaalinen tieto ja taito: toisen sukupolven suomi ja S1-verrokki*. Acta Universitatis Tamperensis 1865. Tampere: Tampere University Press. <http://tampub.uta.fi/handle/10024/94544>.
- 2017. Kieli- ja kielitaitokäsitykset tutkivan opettajan kenttäpäiväkirjamerkinnöissä. *Puhe ja kieli*, 4/2017, 215–238.
- Härnqvist, K., U. Christianson, D. Ridings & J.-G. Tingsell 2003. Vocabulary in interviews as related to respondent characteristics. *Computers and the Humanities*, 37 (2), 179–204.
- ISK 2004 = Hakulinen, A., M. Vilkkuna, R. Korhonen, V. Koivisto, T.-R. Heinonen & I. Alho 2004. *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Jarvis, S. 2002. Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing*, 19 (1), 57–84.
- 2013a. Capturing the diversity in lexical diversity. *Language Learning*, 63 (1), 87–106.
- 2013b. Defining and measuring lexical diversity. Teoksessa S. Jarvis & M. Daller (toim.) *Vocabulary knowledge: human ratings and automated measures*. Amsterdam: John Benjamins, 13–44.
- Johansson, V. 2008. Lexical diversity and lexical density in speech and writing: a developmental perspective. Teoksessa *Department of Linguistics and Phonetics Working Papers* 53. Lund: Lund University, 61–79. <http://journals.lub.lu.se/index.php/LWPL/article/view/2273>.
- Karlsson, L. 2013. Storycrafting method: to share, participate, tell and listen in practice and research. *The European Journal of Social & Behavioural Sciences, Special Volumes VI Design in Mind*, 6 (3), 1109–1117.
- 2014. *Sadutus: avain osallistavan toimintakulttuuriin*. 3. laitos. Jyväskylä: PS-kustannus.
- Klee, T., S. F. Stokes, A. M.-Y. Wong, P. Fletcher & W. Gavin 2004. Utterance length and lexical diversity in Cantonese-speaking children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*, 47 (6), 1396–1410.
- Koizumi, R. & Y. In'Nami 2012. Effects of text length on lexical diversity measures: using short texts with less than 200 tokens. *System*, 40 (4), 554–564.
- Kuiken, F. & I. Vedder 2012. Speaking and writing tasks and their effects on second language performance. Teoksessa S. M. Gass & A. Mackey (toim.) *The Routledge handbook of second language acquisition*. London: Taylor ja Francis, 364–377.
- Lai, S. & P. J. Schwanenflugel 2016. Validating the use of “D” for measuring lexical diversity in low-income kindergarten children. *Language, Speech, and Hearing Services in Schools*, 47 (3), 225–235.
- Lervåg, A. & V. G. Aukrust 2010. Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry*, 51 (5), 612–620.
- Malin, E. 2012. *Suomi toisena kielenä -oppijoiden sanaston kehittyminen taitotasolta toiselle siirryttäessä*. Pro gradu -tutkielma. Jyväskylä: Jyväskylän yliopisto. <http://urn.fi/URN:NBN:fi:jyu-201211223060>.

- Malvern, D. & B. Richards 1997. A new measure of lexical diversity. Teoksessa A. Ryan & A. Wray (toim.) *Evolving models of language: papers from the annual meeting of the British Association of Applied Linguists held at the University of Wales, Swansea, September 1996*. Clevedon: Multilingual Matters, 58–71.
- 2002. Investigating accommodation in language proficiency interviews 426 using a new measure of lexical diversity. *Language Testing*, 19 (1), 85–104.
- Malvern, D., B. Richards, N. Chipere & P. Durán 2004. *Lexical diversity and language development: quantification and assessment*. Houndmills: Palgrave Macmillan.
- McCarthy, P. M. 2005. *An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual, lexical diversity (MTLD)*. Väitöskirja. University of Memphis.
- McCarthy, P. M. & S. Jarvis 2007. Vocd: a theoretical and empirical evaluation. *Language Testing*, 24 (4), 459–488.
- McKee, G., D. Malvern & B. Richards 2000. Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing*, 15 (3), 323–337.
- Milton, J. 2009. *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.
- Muter, V., C. Hulme, M. J. Snowling & J. Stevenson 2004. Phonemes, rimes, vocabulary and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental Psychology*, 40 (5), 665–681.
- Pajunen, A. 2012. Kirjoittamistaitojen kehitys 8–12-vuotiailla: alakoululaisten unelma- kirjoitelmat. *Virittäjä*, 114 (1), 481–501.
- Qian, D. D. & M. Schedl 2004. Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing*, 21 (1), 28–52.
- Riihelä, M. 1991. *Aikakortit: tie lasten ajatteluun*. Helsinki: VAPK-kustannus.
- 2013. Suomalaisten pienten lasten ajatuksia heijastava erittäin laaja aineisto kappaa tutkijoita! [http://www.edu.helsinki.fi/lapsetkertovat/lapset/Tutkimus/tutkimus\\_aineisto.htm](http://www.edu.helsinki.fi/lapsetkertovat/lapset/Tutkimus/tutkimus_aineisto.htm) [luettu 30. 9. 2017].
- Saarela, L. 1997. *Peruskoululaisten kirjoitelmien kehittyminen sanastotutkimuksen valossa*. Acta Universitatis Ouluensis B Humaniora 25. Oulu: Oulun yliopisto.
- Schmid, M. & S. Jarvis 2014. Lexical access and lexical diversity in first language attrition. *Bilingualism: Language and Cognition*, 17 (4), 729–748.
- Scott, C. M. & J. Windsor 2000. General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech, Language and Hearing Research*, 43 (2), 324–339.
- Singh, S. 2001. A pilot study on gender differences in conversational speech on lexical richness measures. *Literary and Linguistic Computing*, 16 (3), 251–264.
- Strömqvist, S., V. Johansson, S. Kriz, H. Ragnarsdóttir, R. Aisenman & D. Radvid 2002. Toward a cross-linguistic comparison of lexical quanta in speech and writing. *Written Language and Literacy*, 5 (1), 45–67.
- Taimisto, H. 2014. *Taitotasolta toiselle: korpuspohjainen tutkielma vironkielisten suomenoppijoiden verbisanaston kehittämisestä*. Pro gradu -tutkielma. Suomen kieli, Oulun yliopisto.
- Tannenbaum, K. R., J. K. Torgesen & R. K. Wagner 2006. Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*, 10 (4), 381–398.

- Tidball, F. & J. Treffers-Daller 2007. Exploring measures of vocabulary richness in semi-spontaneous French speech. Teoksessa H. Daller, J. Milton & J. Treffers-Daller (toim.) *Modelling and assessing vocabulary knowledge*. Cambridge: Cambridge University Press, 133–149.
- Unsworth, S. 2004. Comparing child L2 development with adult L2 development: how to measure L2 proficiency. Teoksessa B. Haznedar & E. Gavruseva (toim.) *Current trends in child second language acquisition: a generative perspective*. Amsterdam: John Benjamins, 301–333.
- Watkins, R., D. Kelly & H. Harbersh 1995. Measuring children's lexical diversity: differentiating typical and impaired language learners. *Journal of Speech and Hearing Research*, 38 (6), 1349–1355.
- Verhoeven, L. 1990. Acquisition of reading in a second language. *Reading Research Quarterly*, 25 (2), 90–114.
- Vermeer, A. 2000. Coming to grips with lexical richness in spontaneous speech data. *Language Testing*, 17 (1), 65–83.
- Wong, A., T. Klee, S. Stokes, P. Fletcher & L. Leonard 2010. Differentiating Cantonese-speaking preschool children with and without SLI using MLU and lexical diversity. *Journal of Speech, Language, and Hearing Research*, 53 (3), 794–799.
- Wright, H., S. Silverman & M. Newhoff 2003. Measures of lexical diversity in aphasia. *Aphasiology*, 17 (5), 443–452.
- Yu, G. 2010. Lexical diversity in writing and speaking task performances. *Applied Linguistics*, 31 (2), 236–259.
- Zipf, G. K. 1935. *The psycho-biology of language*. Boston (Mass.): Houghton-Mifflin.
- 1937. Observations of the possible effect of mental age upon the frequency-distribution of words from the viewpoint of dynamic philology. *Journal of Psychology*, 4, 239–244.

## Liite 1: Esimerkki sadutustuokiosta (3. lk tyttö S2)

- 01 L: olipa kerran,  
 02 (.)  
 03 S: mm?  
 04 (.)  
 05 L: mmh mun perhe?  
 06 S: =mm?  
 07 L: =ja minä?  
 08 S: =mm?  
 09 L: =mentiin matkustelemaan Vietnami:n,  
 10 (.)  
 11 S: mm,  
 12 L: =sit me leikittiin siel sit mentiin kirkkoon sit me tutkittiin mu äidin kaa ku mentiin  
 13 sen kaa .h sie:llä sitte .hh >siel oli< haus-tausjuttuja siel oli hautoja sellassii,  
 14 (.)  
 15 S: m[m?  
 16 L: [sit lähettiin kotiin sit mentiin ö- pelaamaan tietsikal siel meiän viereises  
 17 talossa siellä .hh  
 18 S: =joo?  
 19 L: =ku me asutaan siin mummin kaa,  
 20 S: (..) [joo?  
 21 L: [nii- siin viereises talossa .hh (jäätelöö) ja sit sellanen e- konepelejä sellassii  
 22 pelattiin siellä.  
 23 (..)  
 24 S: joo?  
 25 (..)  
 26 L: k-sitten mentiin öö- papalle (...) <sitten s-siellä nii me leikittiin>  
 27 .hh meidänmaalaisii leikkei sellassii .hh leikittii sitä.  
 28 (.)  
 29 L: °joo-o°  
 30 L: =mun kavereitten kaa ku siel on °myös° mun kavereita.  
 31 (.)  
 32 S: mm  
 33 (...)  
 34 L: sitten me mentiin tonne (..) tonne tivoliin hh [sinne- siellä .hh sit me leikittiin  
 35 sielki <sitten>  
 [mm?  
 36 L: .hh mentiin tonne (...) .h tonne tonne >mis oli nyt se oli< .hh >no se on< sinn-  
 37 tonne juna-asemalle .hh  
 38 S: =mm?  
 39 L: sit mentii jonnekki taas .hh em mä muista enää.



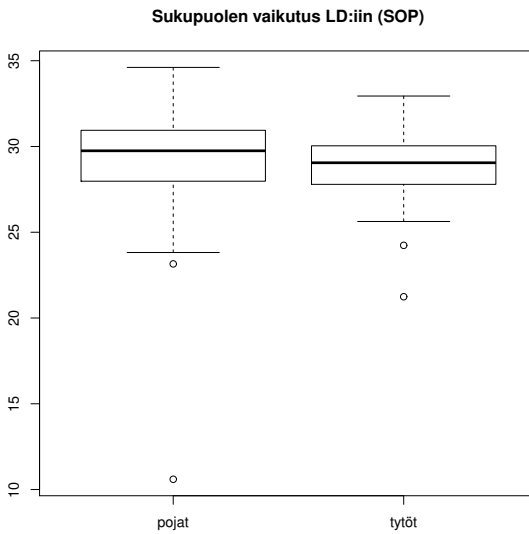
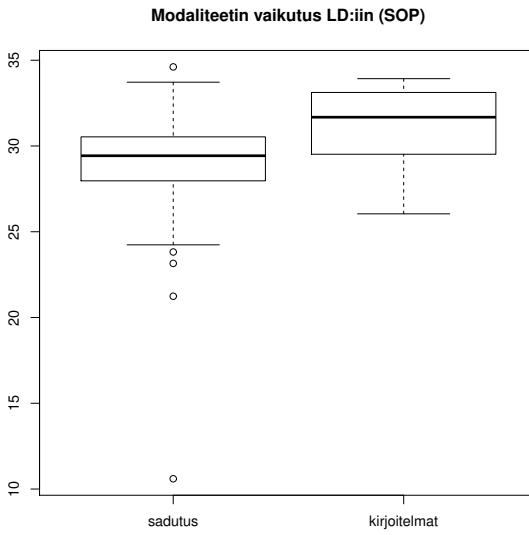
## PIENIN LEKSIKAALINEN DIVERSITEETTI

sana	F	x	r	saneita	p(F = 0)	p(F > 0)	KL	KTOT
ei	18	0	42	66	0,000000	1,000000	0,023810	0,023810
hyvä	1	0	42	66	0,363636	0,636364	0,023810	0,015152
joku	1	0	42	66	0,363636	0,636364	0,023810	0,015152
joo	10	0	42	66	0,000009	0,999991	0,023810	0,023809
kanssa	1	0	42	66	0,363636	0,636364	0,023810	0,015152
kaveri	1	0	42	66	0,363636	0,636364	0,023810	0,015152
keksiä	1	0	42	66	0,363636	0,636364	0,023810	0,015152
kärpänen	1	0	42	66	0,363636	0,636364	0,023810	0,015152
mennä	1	0	42	66	0,363636	0,636364	0,023810	0,015152
minä	12	0	42	66	0,000001	0,999999	0,023810	0,023810
olla	2	0	42	66	0,128671	0,871329	0,023810	0,020746
talvi	1	0	42	66	0,363636	0,636364	0,023810	0,015152
tietää	15	0	42	66	0,000000	1,000000	0,023810	0,023810
välitunti	1	0	42	66	0,363636	0,636364	0,023810	0,015152
yht. 14	66							10,59859

Todennäköisyyden  $F = 0$  laskemisessa on käytetty neljän muuttujan ( $x$ ,  $r$ ,  $F$ , saneita) hypergeometrista jakaumaa.  $KL = 1/r$  eli  $F$ :n kontribuutio lemmän TTR:ään näyttekoolla  $r$ .  $KTOT = p(F > 0) \times 1/r$  eli  $F$ :n kokonais-kontribuutio.



### Liite 3: Sukupuolen ja modaliteetin vaikutus leksikaaliseen diversiteettiin



## Liite 4: Esimerkit vuorojäsennyksestä

Esimerkkiteksti 1: melko runsas diversiteetti, paljon lyhyitä vuoroja

olipa kerran X ja minä → moltiin kavereita nii ja sit me aina leikittiin toisien kanssa → joskus meille tuli r-riitaa → ku joskus me tapeltii → em mä tiä enää → joskus → sit me sovittiin takas → pleikkoja → pleikka pleikat seedeet jalkapallo kakstuhattaseittemän → mitä → pelataa → tietsikalla → pelaa → fifa kakstuhattaseittemän → no pelataan jalkapalloa → me voitetaan → joo pokaalin → kultanen → em mä tiä → moisin myyny sitä pokaalii → sitte tehny nukkumaan → sille oli vähän- vähän kallis → saan paljon rahaa → sitten sillä rahalla ostan vaatteita ja kirjoja → oisin ainaki ostanu kirjan Risto Räppääjästä → mä oon luenu kaks Risto Räppääjä kirjaa → nyt mä haluu ostaa nuu- nuudelipää Risto Räppääjän → nuudelipää Risto Räppääjän → mul on kotonna Hilpuri Tilli → joskus luen oma kielellä -ki → ja eilen luen n-neljä oma kielellä → joskus mä paan (paperissa ja kynä) ja kirjotan → joskus piirrän autoja → kirjotan kaikkee → mä harrastan euoliigajalkapalloa → kaks vuotta (3. lk. poika S2)

Esimerkkiteksti 2: melko vähäinen diversiteetti, paljon pitkiä vuoroja, vähän tutkijan vuoroja

olipa kerran yksi nalle sen nimi oli Kalle sitten se oli ihan se oli päiväkodissa sit- kun hän- kun se meni kotiin sen äiti tuli hakee niin sit se meni kotiin se söi eka vähän ruokaa vaan sitten se meni pihaan sitten kun sen isä tuli ne meni samaa aikaa kotiin sisään nii sit sen synttärät oli nii sen isä ei oo viel kertonu mitä se otti lahjaks se anto niinku sit ku ne vieraat tuli sen kaverit ja nää nii sit se kerto isäki kerto sille sit se oli tosi iloinen sitten kun ne vieraat meni kotiin toi sen isovelki ja nalle- toi Kalle meni nuk- ne meni nukkuu sit toisena kertana ne meni see Kalle meni kouluun sit se tutustui näihin kaikkeihin kavereitten kaa sitten toisena päivänä ne men- ne meni taas kouluun sit sit ne oli se leikki tosi paljo niitten kaa ja → jaa niinkun se niinkun meni luokkaan niil oli käsiyötä ne teki ison niinku käsinuken kis- sellasen kissan ja sit- sitten ku se meni kotiin taas sit sen isä ja äiti oli kaupungissa sen äiti- sen isovelki meni oli vielä koulussakin sitten ku se tuli koulusta ne meni yhdessä sisään sit ne teki eka läksyt sit ne pelas korteilla → ne meni ihan ku uunoa sellasta → sit ku sen isä ja äiti tuli ne on ostanu niille vaatteita sit ne sai kokeilla niitä kun toi isä ja äiti ja nää toiset söi ruokaa sit yks niitten kaveri Kallen kaveri ja sen isä on pyytäny niitä että niinku me- niinku että ne menee sinne ne sais sit uudet vaatteet laittaa sit ne tuli takas kotiin ja nee otti

vaatteet uudet vaatteet pois ja sit meni nukkuun → em mä enää muista  
(2. lk. tyttö S2)

Yhdellä nuolella on merkitty pitkän mietintätauon tai tutkijan lyhyen vuoron tuoma katkos, kahdella nuolella tutkijan kertomiseen rohkaisu, usein kysymys kuten kerro siitä lisää, mitä sitten tapahtui. Kummassakaan kertomuksessa ei esiinny saduttajalta kierrätettyä leksikaalista ainesta.

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 193–213.

**Heini Kallio<sup>1</sup>, Juraj Šimko<sup>1</sup>, Ari Huhta<sup>2</sup>, Reima Karhila<sup>3</sup>, Martti Vainio<sup>1</sup>, Erik Lindroos<sup>1</sup>, Raili Hildén<sup>1</sup> & Mikko Kurimo<sup>3</sup>**

<sup>1</sup>University of Helsinki, <sup>2</sup>University of Jyväskylä, <sup>3</sup>Aalto University

## **Towards the phonetic basis of spoken second language assessment: temporal features as indicators of perceived proficiency level**

This study investigates whether temporal features in speech can predict the perceived proficiency level in Finnish learners of Swedish. In so doing, seven expert raters assessed speech samples produced by 60 upper secondary school students using the revised CEFR scale for phonological control. The effect of temporal features was studied with a cumulative link mixed model, and the assessments were further analyzed to study inter-rater variation. The results indicate that articulation rate and certain types of disfluencies in speech can predict the perceived proficiency level. Furthermore, assessors seem to weigh temporal features differently depending on the speech type and their individual focus.

**Keywords:** L2 proficiency assessment, L2 phonetics, prosody, temporal features

## 1 Introduction

Speaking a language fluently is often the ultimate goal of mastering a second or foreign language (L2), and fluency in some sense is frequently tested in high-stakes tests that have a tangible effect on learners' futures. In Finland, the Ministry of Education and Culture (MEC) have set a goal to include oral language skills as part of language tests at the end of upper secondary education by 2022 (Ministry of Education and Culture 2017). This increases the importance and relevance of research on spoken L2 skills and their assessment.

Despite the regular use of the term fluency in pedagogy, language testing and linguistics, its definition varies (Huhta 1993; Chambers 1997). Lennon (2000) presents two types of fluency definitions: a broad one and a narrow one. The broad sense corresponds to a higher-order, global oral proficiency, while the narrower definition of fluency refers to spoken performance, more closely to the temporal properties and "smoothness" of the speech. The present study approaches fluency from the latter, narrower perspective with the reliability and efficiency of the assessments in mind. Lennon (2000) states that an individual's fluency is acceptable as long as it engages the listener. L2 speakers' fluency is indeed very likely the primary measure that ordinary interlocutors assess in everyday interaction. It is therefore crucial to understand and study fluency and its relation to language proficiency.

### 1.1 Measuring fluency: an overview

While proficiency is generally the target of assessments, the assessment criteria are often based on assumptions on fluency. Reliable ways to measure spoken fluency are therefore important in language testing as well as future research in applied linguistics, but the variables behind the perception of fluency remain less studied (Koponen & Riegenbach 2000). Over the last decades, however, speech researchers have been increasingly interested in the prosodic (tonal, temporal and dynamic) features of L2 speech. In this paper we will focus on the temporal features of L2 speech fluency and their effect on language proficiency assessments. Speech fluency is generally related to temporal features that have also proven to be strong predictors of prosodic competence (see, e.g., Cheng 2011). Previous studies on L2 speech have used measures such as speech and articulation rate, mean length of run, phonation-time ratio, number of stressed words as well as number and duration of pauses to measure second language fluency (Cucchiari et al. 2002; Derwing et al. 2004; Kormos & Dénes 2004; Trofimovich & Baker 2006; Hönig et al. 2010; Bosker et al. 2013). Derwing et al. (2004) also stated that the measure of standardized pruned syllables (self-corrections, self-repetitions, false starts and non-lexical filled pauses) was a successful predictor of fluency judgements. Bosker et al.

(2013), in turn, found that filled and silent pauses and the mean length of pause together with mean length of syllable are significant predictors of fluency. Researchers have obtained slightly mixed results, however, depending on the target language, fluency measures taken and the design of the study. The relevance of temporal variables together with the measurement methods have been under discussion. For example, researchers lack consensus on the minimum length of a pause, despite the wide range of pause studies in both native (L1) and L2 speech. This is problematic, since all temporal measurement results (except speech rate) depend on the pause threshold used. Many studies of L1 and L2 speech follow Goldman-Eislers' (1968) proposal of a pause threshold of 250 ms based on distinguishing "articulatory" (< 250 ms) and "hesitation" (> 250 ms) pauses (see, e.g., Bosker et al. 2013; De Jong & Bosker 2013). Yet many pauses shorter than 250 ms cannot be attributed to articulation (Hieke et al. 1983; Campione & Véronis 2002) but have a psychological function. For example, speech rate and pausing are involved in expressing paralinguistic information, such as confidence or emotions (Scherer et al. 1973; Scherer 1986). In the current study, we consider also the relevance of pauses shorter than 250 ms as part of perceived fluency.

Another issue in fluency measures is the types and number of disfluencies. Silent pauses – periods of vocal inactivity during speech performance – are easy to detect automatically from speech, especially if pause threshold is set high enough. There can, however, be non-silent, non-lexical pauses in speech that affect the perception of fluency. These pauses are often referred to as filled or hesitation pauses. Other temporal disfluencies include phenomena such as self-corrections and self-repetitions (Derwing et al. 2004; Bosker et al. 2013).

## 1.2 The context-dependent disfluencies

Many studies concentrate on several disfluencies at a time (Lennon 1990; Foster & Skehan 1996; Towell et al. 1996; Derwing et al. 2004). Filled pauses are typical for spontaneous speech and rarely occur in read speech. The use of filled and unfilled pauses, however, is highly context-dependent. For example in French filled pauses seem to be frequent and long especially in conversational speech, and the distribution of silent pauses is related to the syntactic structure of the sentence (Duez 1982). In spontaneous Finnish speech filled pauses might be less common and silent pauses more acknowledged than in French, but there is no extensive study comparing pauses in these languages. However, Toivola et al. (2009) found that native Finnish speakers tend to have longer silent pauses in read speech than non-native speakers of Finnish. Their results indicate that long silent pauses do not necessarily affect the quality of L2 speech. Campione & Véronis (2002) noted that Italians make generally

shorter pauses and Spanish, in turn, longer pauses than French, English and German. These results prompt to take into account the language-specific temporal features and individual differences in prosody, when analyzing pauses as fluency measures. Whether a certain type of pause is considered as a speech disfluency or not is indeed language-specific. It is thus relevant to study the effect of various disfluencies in the perception of fluency. However, the use of too many variables can lead to unreliable results, since many analysis methods require uncorrelating variables, but many temporal features in speech depend on each other: for example, the amount of pauses affects strongly the measurement of speech rate. This leads to a high correlation between these variables. In the current study, we avoid intercorrelating variables and use revised, systematic measurements that are less used in fluency research.

### **1.3 Assessing fluency as a part of language proficiency**

Spoken language fluency studies have generally involved trained assessors (see, e.g., Wennerstrom 2000; Kormos & Dénes 2004), but also native speakers of the target language have proved to be reliable raters of L2 fluency (Derwing et al. 2004, 2006; Rossiter 2009). Studies on temporal features and fluency vary with regard to instructions and criteria given to assessors as well as speech type and sample length. Speech samples have generally been short, from less than 20 seconds to 2–3 minutes, and included either read sentences or narrative speech. The studies have commonly used Likert-type scales (see, e.g., Bosker et al. 2013; Pinget et al. 2014), but the wide descriptors of the Common European Framework of Reference for languages (CEFR, Council of Europe 2001) have also been used in fluency assessments (Préfontaine et al. 2016). The CEFR is a guideline used to describe achievements of L2 learners across Europe and in other countries. The present study is novel in that it uses updated descriptors of the CEFR scale for phonological control to investigate the perceived fluency of Finnish learners of Swedish. The revised phonological control scale consists of two descriptive subsections: prosodic features and articulation of sounds. In this study we examine the descriptor scale for prosodic features, since temporal properties of speech are considered as part of prosody. Additionally, our research reflects actual test contexts with regard to data collection methods and the assessment protocol.

### **1.4 The goal of this study**

The goal of this study is to make spoken L2 assessment more objective and reliable by scrutinizing the quantitative temporal features in speech that affect the assessments, even when the assessors are unaware of these features. This gives us valuable information about the L2 learning and assessing process. Our

study addresses the following research questions (RQ):

- a) Can objectively measured temporal features in speech be used to predict the proficiency level of the speaker?
- b) How do read and semi-spontaneous speech differ with regard to temporal features and assessments?
- c) How do the raters differ in their assessments?

To answer RQ(a), objective acoustic measurements of speech are related to subjective proficiency ratings of the same speech samples. A group of trained raters assessed the proficiency level of Finnish learners of Swedish. Based on the previous studies (Cucchiari *et al.* 2002; Derwing *et al.* 2004; Kormos & Dénes 2004; Bosker *et al.* 2013), we expect that articulation rate and pauses have stronger effect on proficiency ratings than disfluencies related to repairing (repetitions and corrections). To answer RQ(b), we study the two speech types separately. We consider the speech type elicited from different test tasks as a meaningful variable. To answer RQ(c), we study the effect of acoustic variables separately on each assessor. The assessors are also analyzed with respect to intra- and inter-rater consistency. We expect to find some systematic differences between the assessors.

## 2 Material and methods

### 2.1 Speech data

The data used in this study is part of a larger speech corpus, which has been collected while piloting a computer-aided oral language test for large-scale purposes (the project *DigiTala*<sup>1</sup>). The piloting was done in Finnish Swedish as a second language. Finnish Swedish is a variant of Swedish spoken in Finland and differs from standard Swedish with regard to e.g. sound production as well as prosodic features like sentence and word stress. Swedish is a compulsory subject in basic education in Finland, and the national matriculation examination test of L2 Swedish is taken by approximately 8,000 upper secondary school students yearly, which makes Swedish the second most tested L2 in Finland (Finnish Matriculation Examination Board 2017).

Seven upper secondary schools from six municipalities around Finland participated in the pilot tests, and speech data from approximately 760 voluntary pupils has been recorded and stored anonymously in a database. The pilot test works as a web-based application and includes four subtasks:

<sup>1</sup> <http://blogs.helsinki.fi/digitala-projekti/>



- a) A read-aloud task: newspaper headlines or a written phone message
- b) Situational reacting task: reacting to situations given in written L1 or with a picture and a written clue
- c) A simulated video phone call with pre-recorded replies from one native speaker of the target language
- d) A live dialogue task with a peer.

These tasks cover various dimensions of speaking proficiency, differing in discourse genre, formality and complexity. The use of computer-administered tasks helps to standardize the testing conditions for all participants as well as to reduce the effect of the behavior of other individuals in the communication, enabling to study language proficiency as an individual attribute. The subtasks were split to smaller sections and every task was timed. Instructions were in written Finnish or Swedish depending on the task.

A subset of 60 speech samples from the larger pilot data was used in this study. 50 samples were produced by native Finnish and 10 by native Finnish Swedish upper secondary school students, aged 16–18 years. The native samples were selected to elicitate higher proficiency scores and thus enable investigating all levels of the CEFR scale. The speech samples included read ( $n = 19$ ) and semi-spontaneous ( $n = 41$ ) utterances from subtasks a, b and c.

## 2.2 Assessments

Seven trained evaluators, Swedish language teachers or native speakers of Finnish Swedish (see Table 1) assessed each speech sample using the updated Common European Framework of Reference for Languages (CEFR) scales (Council of Europe 2017). The assessments were part of piloting the revised descriptor scales from a proposed version of the CEFR illustrative descriptors, authorized by the Council of Europe Language Policy Section. This study focuses on the assessments of prosodic features, which is a subsection of the revised CEFR descriptor scale for phonological control and pays attention to features such as word and sentence stress, rhythm and intonation with respect to the perceived intelligibility of the speech.

The descriptor scales for prosodic features were translated to Finnish. The assessors were familiarized with the new reference scale and some speech samples were assessed and discussed together before the actual rating task. No specific instructions to focus on the temporal features were given to assessors, unlike many previous studies have done (see, e.g., Derwing et al. 2004; Bosker et al. 2013; Pinget et al. 2014). The assessments were collected using a web interface, where the assessors could listen and assess the speech samples at their own pace. Before the actual assessment task the assessors were

TABLE 1. The background of the assessors.

Assessor	L1	Language teacher	Phonetic studies	Studied/taught spoken L2 skills
A1	Finnish	Swedish	Yes	Yes
A2	Finnish	Swedish	Yes	Yes
A3	Finnish	Swedish	Yes	Yes
A4	Swedish	No	Yes	No
A5	Finnish	Swedish	Yes	Yes
A6	Finnish	Swedish	Yes	Yes
A7	Swedish	Other language	Yes	No

asked to answer questions concerning their background. The background information of the assessors is presented in Table 1.

### 2.3 Acoustic analysis

Annotation and analysis were done using the Praat software (Boersma & Weenink 2010) and a script for large-scale systematic analysis of continuous prosodic events (Xu 2013). The acoustic variables measured for statistical analysis were articulation rate (AR), silent pause-time ratio (S), filled pause-time ratio (F), and corrections or repetitions-time ratio (CR). These four fluency variables were chosen because previous research suggests that similar measures are salient predictors of fluency. Intercorrelating variables, like speech rate and overall pause-time ratio, were avoided. Disfluency-time ratios are used as fluency measures instead of more commonly used number of disfluencies, since the sole frequency of disfluencies does not give information on the amount of time they take during speech. Disfluency-time ratio is therefore seen as more comprehensive measure, because it depends on both the number and the length of disfluencies. However, relative amount of disfluencies and distributions of disfluency durations are used to illustrate and reflect on the differences between read and semi-spontaneous speech.

All disfluencies of 50 ms or longer were marked manually and measured from every sample (see Figure 1 for annotation example). Articulatory pauses, such as plosive closure phases or prepausal lengthening, that were not part of hesitations, self-corrections or self-repetitions were excluded. Disfluency-time ratios for all disfluency types and articulation rate were calculated separately for read and semi-spontaneous speech samples. The variables were operationalized as follows:

- a) Articulation rate (AR): The number of syllables produced per second excluding all pauses longer than 50 ms. Number of syllables was used instead

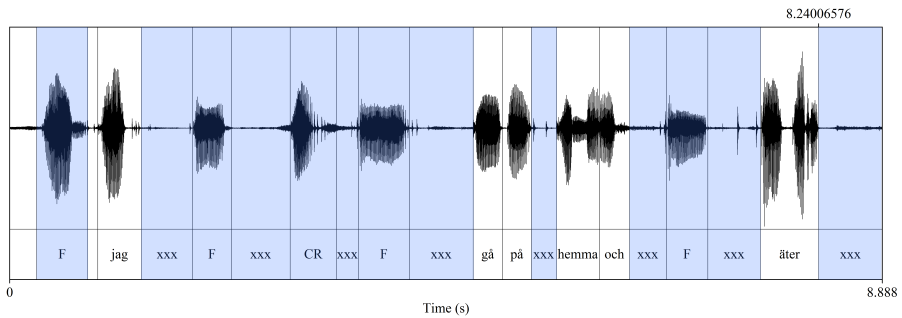


FIGURE 1. Example of pause annotation. F = filled pause, CR = self-correction or repetition, xxx = silent pause.

of phones, since the syllable is considered as the smallest speech unit that carries relevant prosodic and articulatory information (Collier 1992: 206).

- b) Silent pause-time ratio (S): The total duration of silent pauses above 50 ms (no vocal activity) divided by total duration of the given speech sample. Furthermore, silent pauses were grouped by their length into three categories: short (SS, < 200 ms), medium (MS, 200–1,000 ms) and long (LS, > 1,000 ms), following the trimodal distribution presented by Campione & Véronis (2002).
- c) Filled pause-time ratio (F): The total duration of filled pauses above 50 ms divided by total duration of the given speech sample. Pauses were considered as filled, when they included non-lexical vocal activity, often used as hesitation markers such as “umm”.
- d) Corrections or repetitions-time ratio (CR): The total duration of self-corrections and self-repetitions divided by total duration of the given speech sample.

Silent pause-time ratio is an equivalent of more commonly used phonation-time ratio, but silent pause-time ratio is easier to compare to filled pause-time ratio and corrections or repetitions-time ratio than phonation-time ratio, because all nominators are considered as types of disfluencies. A trimodal length distribution was used for silent pauses, but the number of other disfluencies were not sufficient for such grouping.

## 2.4 Statistical analysis

The relation between temporal features and proficiency assessments was studied using the R program (Baayen 2008). We analyzed our data with a cumulative link (also called as ordinal regression or proportional odds) mixed

(CLM) model implemented in the *clmm2* package in R and designed specifically for the analysis of responses measured on an ordinal scale (see, e.g., Christensen 2015). The CLM model gives improved estimates of regression coefficients compared to continuous models that do not take into account the ceiling and floor effects nor the possible skewness of the ordinal variable. In our analysis we treated grades given by assessors as an ordered response (categories A1-C2), and acoustic measurements as explanatory variables. Assessor was treated as a random effect, and analysis was done separately for read and semi-spontaneous speech samples. We also studied the effect of temporal measurements on each assessor with a separate cumulative link (CL) model.

The assessments were further analyzed with a multi-faceted Rasch measurement (MFRM) using the Facets-program first developed by Linacre (1989). Facets is a development of simpler Rasch models (Rasch 1960) which are psychometric models used in analyzing data from tests, questionnaires and other types of assessment. Facets is widely used in language assessment to analyze ratings of speaking and writing performances (Bachman et al. 1995; McNamara & Knoch 2012) because it can simultaneously model, and take into account, such factors as learner ability, rater severity, scale, and difficulty of the tasks and dimensions of language. The analysis produces a logit scale (an interval scale) against which all the facets (the components conceptualized to combine to produce the data, e.g., persons, items, judges, tasks) included in the analysis are directly comparable. In the present study, Facets was used to investigate the quality of the ratings (particularly consistency of the assessors) and of the new CEFR scale for phonological control.

### 3 Results

#### 3.1 Temporal features

Articulation rate, silent pauses, filled pauses, and corrections and repetitions were measured from each speech sample. For silent and filled pauses, and corrections and repetitions, disfluency-time ratio was calculated from each sample.

Some speech samples included no disfluencies, while most samples included more than one type of disfluencies. Figure 2 shows the relative amount of disfluencies in speech samples with respect to speech type. Silent pause of 200–1,000 ms was the most frequent for both speech types, and long silent pauses (> 1,000 ms) occurred almost exclusively in semi-spontaneous speech samples. Corrections and repetitions occurred more frequently in read than in semi-spontaneous speech samples, and filled pauses occurred more frequently in semi-spontaneous than in read speech samples.

Figure 3 illustrates the density of disfluencies with respect to disfluency

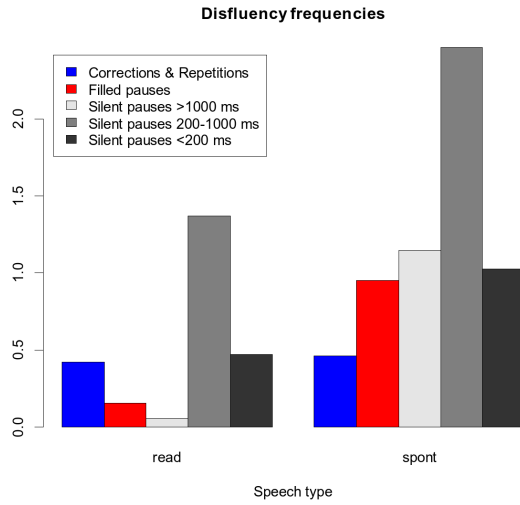


FIGURE 2. The relative amount of disfluencies in read and semi-spontaneous speech samples.

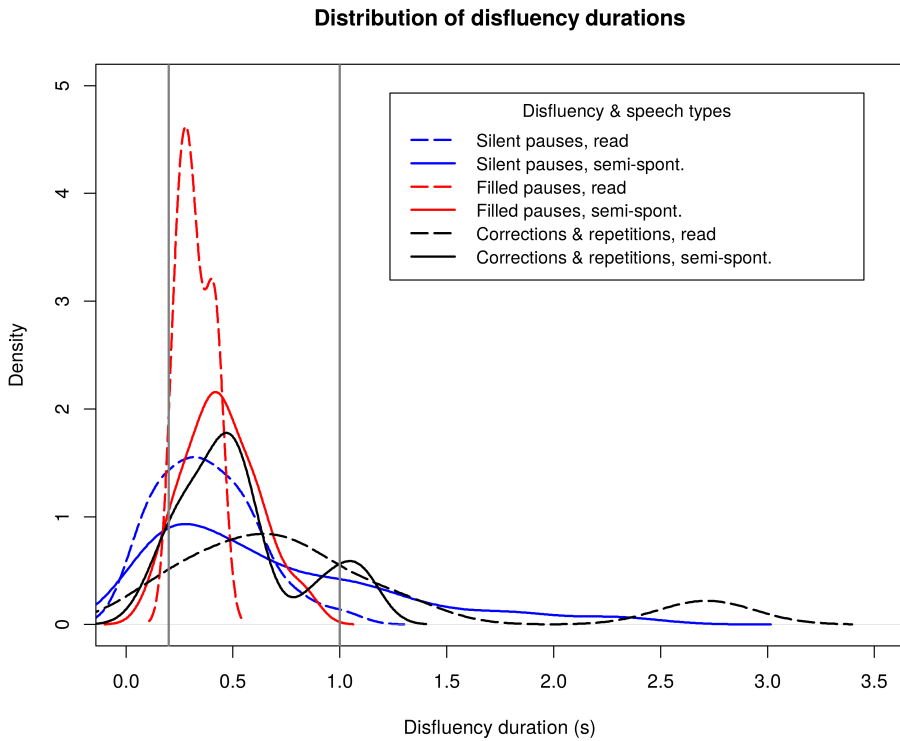


FIGURE 3. Distribution of disfluency durations. Dotted line: read speech, solid line: semi-spontaneous speech. Vertical lines are positioned at 200 ms and 1 s to illustrate the trimodal distribution of disfluency durations.

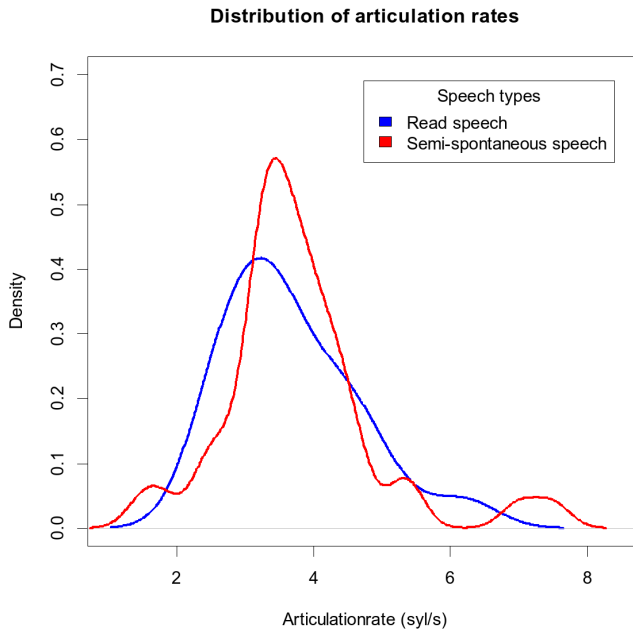


FIGURE 4. Distribution of articulation rates (syllables per second). Blue line: read speech, red line: semi-spontaneous speech.

duration, disfluency type, and speech type. The figure shows that most disfluencies are shorter than one second, but disfluencies even longer than three seconds were also detected. Filled pauses tend to be shorter than other disfluency types in both speech types, but read speech contains a considerable amount of filled pauses shorter than 500 milliseconds. Longer corrections and repetitions occurred more often in read speech than semi-spontaneous speech.

Articulation rate varied from 2.5 to 6.2 syllables per second in read speech and from 1.5 to 7.5 syllables per second in semi-spontaneous speech. The mean articulation rate was 3.7 syllables per second in read speech, and 3.8 syllables per second in semi-spontaneous speech. Figure 4 shows the distribution of articulation rates with respect to speech type.

### 3.2 Temporal feature effect on assessments

Each speech sample was assessed by seven expert raters using the revised CEFR scale for phonological control. In our analysis, we treated grades given by assessors as an ordered response (categories A1-C2). The data was first analyzed with a cumulative link mixed (CLM) model, where assessments were treated as dependent variables and acoustic measurements as explanatory

TABLE 2. The variables used in cumulative link mixed model analysis.

Type	Variable
Dependent	Assessment
Explanatory	Articulation rate (AR)
	Short (< 200 ms) silent pause-time ratio (SS)
	Medium (200–1,000 ms) silent pause-time ratio (SM)
	Long (> 1,000 ms) silent pause-time ratio (SL)
	Filled pause-time ratio (F)
	Corrections or repetitions-time ratio (CR)
Random	Assessor

variables (see Table 2 for list of variables). Assessor was treated as a random variable, and analysis was done separately for read and semi-spontaneous speech samples. We also examined the effect of articulation rate and disfluency types on each assessor (see tables 4 and 5 in section 3.3).

Figure 5 shows the assessment distribution for all samples. The most common proficiency grades for prosodic features were B1 ( $n = 136$ ) and A2 ( $n = 109$ ). Standard deviation of assessments between assessors varied from 0.38 to 2.9 proficiency scales, leaving the mean SD to 0.81. The assessment distribution is skewed clearly towards lower proficiency levels. Three samples were excluded from statistical analysis, since they were marked as non-analyzable: the assessors were unable to assess prosody from these speech samples.

Table 3 shows the results of our CLM-model. Estimate values and statistical significance of articulation rate and different disfluency types on assessments are introduced with regard to speech type. The model showed a positive effect for AR and negative effect for disfluencies: that is, faster articulation rate and small disfluency-time ratio indicate higher perceived proficiency level. AR and F proved to be extremely significant predictors of the assessed proficiency level in both read and spontaneous speech. Both SS and SL as well as CR were significant predictors in read speech, while in semi-spontaneous speech only SL was somewhat significant, but SS, SM and CR remained insignificant.

As the CLM-model itself does not provide a straightforward method to estimate the quality of fit, we approximate the multinomial logistic regression with a linear mixed effect regression models using the numerical assessment as a dependent variable (with the same independent variables and random effect structure as in the original logistic regression). The marginal  $r^2$ -value provides an estimate of the quality of fits in terms of variance explained by fixed effects in the linear mixed-effect models (Nakagawa & Schielzeth 2013).

The linear model yielded very similar effect estimates; in fact, the signifi-

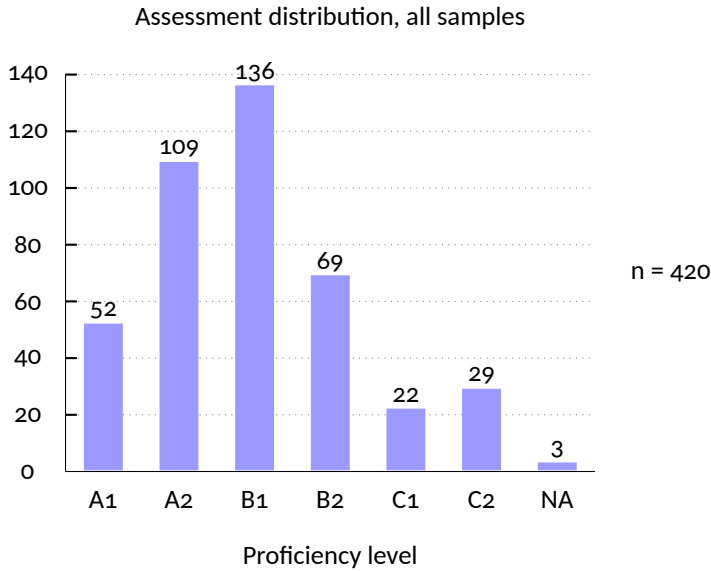


FIGURE 5. Assessment distributions computed from all assessed speech samples.

TABLE 3. The effect of articulation rate and disfluency-time ratios (estimate values and statistical significance) on prosodic features assessments. AR = articulation rate, SS = silent pauses < 200 ms, SM = silent pauses 200–1,000 ms, SL = silent pauses > 1,000 ms, F = filled pauses, CR = corrections and repetitions. p-values: 0.01–0.05 \*, 0.01–0.001 \*\*, < 0.001 \*\*\*.

Speech type	AR	SS	SM	SL	F	CR
Read	2.40***	–39.78**	–1.29	–34.96***	–41.26***	–4.48**
Spont	0.98***	–7.33	–0.65	–1.79*	–9.67***	–4.97



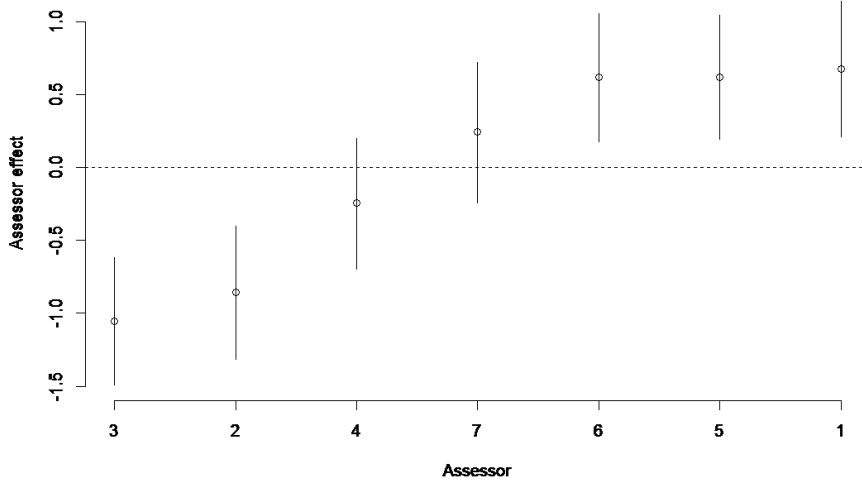


FIGURE 6. The assessor effect in prosody assessments.

cance of different effects was identical to the CLM-model as reported above. The marginal  $r^2$ -value was 0.59 for read speech and 0.34 for semi-spontaneous speech, showing that the dependent variables explain a substantial amount of assessment variance; these measures provide a lower bound for the quality of fit of the more elaborate CLM-models.

### 3.3 Inter-rater variation

The assessor effect was analyzed from the CLM-model, based on the conditional distribution of the random effects (Assessor, see Table 3). Figure 6 shows that assessors A06, A05 and A01 give generally higher proficiency scores for prosody than other assessors, and differ statistically significantly from assessors A03, A02, and A04, but not from A07. Assessors A03 and A02, in turn, are stricter than the other assessors. The distance between the strictest and most permissive assessor is 1.7 logits, which corresponds to about 0.8 proficiency levels. The average proficiency raters for prosodic features are the native Finnish Swedish speaking assessors A04 and A07.

The effect of articulation rate and disfluency types on each assessor was examined separately with a cumulative link (CL) model using the R software. The data was divided into subsets by assessors, and analysis was done separately for read and semi-spontaneous speech samples. Assessments were treated as dependent variables and acoustic measurements as explanatory variables.

TABLE 4. The effect of articulation rate and disfluency types (estimate values and statistical significance) on prosodic features assessments of the read speech samples. p-values: 0.01–0.05\*, 0.01–0.001\*\*, < 0.001\*\*\*. When the field is empty, the model found probabilities of 0 or 1.

Assessor	AR	SS	SM	SL	F	CR
A01	3.57**	–2.43	8.63	–1.58***	–6.68*	–2.22
A02	1.93*	1.79	2.29	–1.49***	–4.53*	2.37
A03	3.94**		2.89	–1.47***	–2.7	–2.36
A04		–6.29*	–8.54	–1.46***	–1.48	–3.49
A05	3.25**		3.02	–2.59	–61.29*	–16.04**
A06		–7.68*	–1.09	–1.46***	–1.11	–3.47
A07	2.99**	–9.7	–7.38	–1.53***	–2.79	–1.05*

TABLE 5. The effect of articulation rate and disfluency types (estimate values and statistical significance) on prosodic features assessments of the semi-spontaneous speech samples. p-values: 0.01–0.05\*, 0.01–0.001\*\*, < 0.001\*\*\*. When the field is empty, the model found probabilities of 0 or 1.

Assessor	AR	SS	SM	SL	F	CR
A01	1.09**	–9.58	–4.7	–4.26	–9.35	1.9
A02	0.73*	–9.19	0.31	–1.7	–5.71	–3.38
A03	0.91*	–13.84	0.12	–1.26	–14.53	–11.12
A04	1.54***	–6.74	1.25	0.18	–11.54	–13.44
A05	1.28**		3.57		–21.63*	–16.59*
A06	1.58***	–6.97	2.16	–1.4	–21.66*	–6.18
A07	1.10**	–12.29	–1.57	–2.56	–5.42	–1.28

Tables 4 and 5 show the estimate values and statistical significance of the examined temporal features on assessments with regard to speech type and individual assessors. Table 4 shows the effects for read speech and Table 5 for semi-spontaneous speech. In read speech, the negative effect of SL was extremely significant for all assessors but A05; instead, CR had a strong negative effect on A05's assessments. The positive effect of AR was very significant or significant to all assessors, unless the probability was 0 or 1. In semi-spontaneous speech, AR was significant or very significant for all assessors, but CR was significant only for assessor A05.

The assessments were further analyzed with the Facets program in order to measure the consistency of the assessments. Table 6 shows the infit and outfit values that indicate how systematic the raters' assessments are (Linacre 1989). Outfit values are sensitive to outliers in the data whereas infit values focus on the core set of ratings. Ideal infit / outfit mean square values are

TABLE 6. The Infit and Outfit values of assessors measured from prosody assessments.

Assessor	Infit		Outfit	
	MnSq	ZStd	MnSq	ZStd
A01	0.81	-0.9	0.8	-0.9
A02	0.67	-1.7	0.6	-1.8
A03	0.85	-0.6	0.71	-1.1
A04	1.09	0.4	1.08	0.4
A05	1.46	2.0	1.78	3.2
A06	0.73	-1.4	0.7	-1.5
A07	1.21	1.0	1.15	0.7

close to 1.0. Very high values (general cut-off point 1.5) are problematic as they indicate that the rater's assessments are unsystematic and unpredictable. Very low values usually indicate that the rater is behaving more systematically than could be modeled but it may also result from the rater not using the full scale but overusing, for example, the middle part of the rating scale. Assessor A05's infit mean square value is close to 1.5, which indicates some inconsistency in this assessor's assessments.

## 4 Discussion

This study investigated whether objectively measured temporal features in speech can be used in predicting the proficiency level assessed by expert raters. The effect of articulation rate, pause-time ratios of silent and filled pauses as well as corrections and repetitions-time ratio was studied with a cumulative link mixed model. The assessments were further analyzed with Facets to study inter-rater variation and reliability. Below, we discuss the effect of temporal features on assessments and differences between the assessors.

### 4.1 Temporal features as proficiency indicators

The effect of articulation rate (AR), silent (S) and filled (F) pause-time ratio, and corrections and repetitions-time ratio (CR) was examined. Our findings suggest that different speech types have different requirements when it comes to prosodic features. AR and F proved to be extremely significant predictors in both read and semi-spontaneous speech. Higher AR implicates higher perceived proficiency level for both speech types. Higher F, in turn, implicates lower proficiency level for both speech types. Although SL occurred almost exclusively in semi-spontaneous speech samples, SS, SL and CR were all significant for read speech. Only SL was somewhat significant for semi-spontaneous

speech. Interestingly, SM remained non-significant for both speech types, although silent pauses of 200–1,000 ms were the most common in our data. The significance of short silences in read speech suggest that silences shorter than 200 ms should be taken into account when studying fluency and the commonly used threshold of 250 ms for silent pauses could cause misleading results. The insignificance of CR in semi-spontaneous speech, in turn, support the findings of previous studies (Cucchiaroni et al. 2002; Bosker et al. 2013), and the general assumption that disfluencies such as self-corrections are more tolerated in conversational speech, which our semi-spontaneous speech samples reflect, than in read or formal speech. On the other hand, CR were on average longer in read speech than in semi-spontaneous speech, which could have caused the difference in effect. Our measures fail to distinguish between different types of self-corrections, remaining insensitive to the extent of these disfluencies, for example the length of the correction or repetition. This should be studied further, since several quick repetitions of single words or syllables may be perceived as less obstructive than long misrepresentations requiring notable backtracking. Disfluencies should be studied also by their length, but further research requires larger speech and assessment data. In this study the size of the speech data allowed scrutinizing only the length of silent pauses; the number of filled pauses and corrections and repetitions were insufficient for such grouping. Another interesting question is, which disfluencies are automatically detectable. New state-of-the-art prosody analysis methods (Suni et al. 2017) could be used to examine the possible automatization of fluency analysis. Automatic measurement of disfluencies could make the assessment procedure more efficient especially in large-scale testing, but many existing systems based on automatic speech recognition still struggle with the evaluation of spontaneous speech (see, e.g., Witt 2012).

## 4.2 Inter-rater variation

Many previous fluency studies have used the mean rating for each assessed item, thereby disregarding the individual differences between assessors. We studied each assessor individually, and the results revealed differences in the severity and consistency of the assessors as well as indications of the assessors' individual focuses. In our study the native assessors were ranked closer to the average assessor with respect to the severity of their assessments. The assessor AO5, whose Infit Mean Square value indicated inconsistency in assessments, differed from other assessors also with respect to the significance of temporal variables: only the estimate values for articulation rate reflected the ones of other assessors. The results indicate that the assessors weigh the various disfluencies differently. This issue has not been discussed much in previous studies and should be scrutinized in more detail in order to gain better

knowledge on the assessors' foci on temporal properties of speech. Modeling the variation between assessors enables profiling the assessors and development of more relevant training procedures, which can improve the reliability of the assessments especially in large-scale tests.

## 5 Conclusions

We studied whether objectively measured temporal features in speech can be used in predicting the proficiency level assessed by expert raters. Our results suggest that certain temporal features do have an effect on perceived proficiency level, but speech type needs to be taken into account: corrections, repetitions and short silences are more tolerated in semi-spontaneous than in read speech. For both speech types, however, higher articulation rate implicates higher proficiency level, whereas higher filled pause-time ratio implicates lower proficiency level. Larger data is yet recommended to study different types of disfluencies more closely, but objective phonetic measurements can serve as an anchor for assessments. Precise measurements related to assessments can also be used to develop automatic methods for more effective and reliable assessment protocol.

Furthermore, we found that assessors seem to weigh temporal features differently depending on the speech type and their individual focus. This issue should be scrutinized further and methods for improving inter-rater agreement should be examined. Profiling the assessors can help to improve their assessing performance with useful feedback that will increase their phonetic awareness. Moreover, implementing human-machine hybrid scoring can significantly improve inter-rater agreement and thus the reliability and objectivity of the assessments (see, e.g., Luo et al. 2016).

When integrating oral tests to large-scale high-stakes exams, such as the Finnish Matriculation Examination, there is an abiding need for an objective and effective assessment of students' proficiency. Inconsistency between or within assessors have a negative effect on assessments and assessed L2 learners. Detecting indicators of what assessors are reacting to when subjectively assessing speech fluency helps make L2 assessment more reliable in three ways: the assessment criteria can be clarified with more specific features related to fluency, the assessors can be profiled according to their individual focus, and the assessors can be better trained to recognize these prosodic features they subconsciously use while assessing L2 proficiency.

## References

- Baayen, R. H. 2008. *Analyzing linguistic data: a practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Bachman, L., B. Lynch & M. Mason 1995. Investigating variability in tasks and rater judgements in a performance test of foreign language speaking. *Language Testing*, 12 (2), 238–257.
- Boersma, P. & D. Weenink 2010. PRAAT. Amsterdam: University of Amsterdam. <http://www.fon.hum.uva.nl/praat/>.
- Bosker, H. R., A.-F. Pinget, H. Quené, T. Sanders & N. H. De Jong 2013. What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, 30 (2), 159–175. DOI: 10.1177/0265532212455394.
- Campione, E. & J. Véronis 2002. A large-scale multilingual study of silent pause duration. In B. Bel & I. Marlien (eds) *Speech prosody 2002*. Aix-en-Provence: Laboratoire Parole et Langage, 199–202.
- Chambers, F. 1997. What do we mean by fluency? *System*, 25, 535–544. DOI: 10.1016/S0346-251X(97)00046-8.
- Cheng, J. 2011. Automatic assessment of prosody in high-stakes English tests. In P. Cosi, R. De Mori, G. Di Fabbrizio & R. Pieraccini (eds) *INTERSPEECH 2011: 12th annual conference of the International Speech Communication Association, Florence, Italy, August 27–31, 2011*, 1589–1592. [http://www.isca-speech.org/archive/interspeech\\_2011/](http://www.isca-speech.org/archive/interspeech_2011/).
- Christensen, R. H. B. 2015. A tutorial on fitting cumulative link mixed models with `clmm2` from the ordinal package. [ftp://ftp.ussg.indiana.edu/pub/CRAN/web/packages/ordinal/vignettes/clmm2\\_tutorial.pdf](ftp://ftp.ussg.indiana.edu/pub/CRAN/web/packages/ordinal/vignettes/clmm2_tutorial.pdf).
- Collier, R. 1992. A comment on the prediction of prosody. In G. Bailly & C. Benôit (eds) *Talking machines*. Amsterdam: North-Holland, 205–208.
- Council of Europe 2001. *Common European framework of reference for languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.
- 2017. *Common European framework of reference for languages: learning, teaching, assessment*. Council of Europe. (Companion Volume with New Descriptors).
- Cucchiarini, C., H. Strik & L. Boves 2002. Quantitative assessment of second language learners' fluency: comparisons between read and spontaneous speech. *Journal of the Acoustical Society of America*, 111 (6), 2862–2873. DOI: 10.1121/1.428279.
- De Jong, N. H. & H. R. Bosker 2013. Choosing a threshold for silent pauses to measure second language fluency. In Robert Eklund (ed.) *Proceedings of DiSS 2013: the 6th workshop on disfluency in spontaneous speech (DiSS), KTH Royal Institute of Technology, Stockholm, Sweden, 21–23 August 2013*. Tal, musik, hörsel – Quarterly Progress and Status Report 54(1). Stockholm: Kungliga Tekniska Högskolan, 17–20. <http://liu.diva-portal.org/smash/record.jsf?pid=diva2%3A1079414%5C&dsid=647>.
- Derwing, T. M., M. J. Rossiter, M. J. Munro & R. I. Thomson 2004. Second language fluency: judgments on different tasks. *Language Learning*, 54 (4), 655–679.
- Derwing, T. M., R. I. Thomson & M. J. Munro 2006. English pronunciation and fluency development in Mandarin and Slavic speakers. *System*, 34 (2), 183–193. DOI: 10.1016/j.system.2006.01.005.

- Duez, D. 1982. Silent and non-silent pauses in three speech styles. *Language and Speech*, 25, 11–28.
- Finnish Matriculation Examination Board 2017. Kevään 2016 ja 2017 ylioppilastutkintoon ilmoittautuneiden määrät kokeittain. [https://www.ylioppilastutkinto.fi/images/sivuston\\_tiedostot/stat/FB2017KT2001.pdf](https://www.ylioppilastutkinto.fi/images/sivuston_tiedostot/stat/FB2017KT2001.pdf).
- Foster, P. & P. Skehan 1996. The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18 (3), 299–323.
- Goldman-Eisler, F. 1968. *Psycholinguistics: experiments in spontaneous speech*. London: Academic Press.
- Hieke, A. E., S. Kowal & D. C. O’Connell 1983. The trouble with “articulatory” pauses. *Language and Speech*, 26 (3), 203–214.
- Hönig, F., A. Batliner, K. Weillhammer & E. Nöth 2010. Automatic assessment of non-native prosody for English as L2. In *Speech prosody 2010 conference proceedings*. Vol. 100973. Chicago (Ill.): 1–4.
- Huhta, A. 1993. Suullisen kielitaidon arviointi. In S. Takala (ed.) *Suullinen kielitaito ja sen arviointi*. Kasvatustieteiden tutkimuslaitoksen julkaisusarja B: Teoriaa ja käytäntöä 77. Jyväskylä: Kasvatustieteiden tutkimuslaitos, 143–225. <https://journal.fi/afinla/article/view/60844/22613>.
- Koponen, M. & H. Riggenbach 2000. Overview: varying perspectives on fluency. In H. Riggenbach (ed.) *Perspectives on fluency*. Ann Arbor (Mich.): University of Michigan Press, 5–24.
- Kormos, J. & M. Dénes 2004. Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32 (2), 145–164. DOI: 10.1016/j.system.2004.01.001.
- Lennon, P. 1990. Investigating fluency in EFL: a quantitative approach. *Language Learning*, 40, 387–417. DOI: 10.1111/j.1467-1770.1990.tb00669.x.
- 2000. The lexical element in spoken second language fluency. In H. Riggenbach (ed.) *Perspectives on fluency*. Ann Arbor (Mich.): University of Michigan Press, 25–42.
- Linacre, J. 1989. *Many-facet Rasch measurement*. Chicago (Ill.): MESA Press.
- Luo, D., W. Gu, R. Luo & L. Wang 2016. Investigation of the effects of automatic scoring technology on human raters’ performances in L2 speech proficiency assessment. In T. Lee, L. Xei, J. Dang, H.-M. Wang, J. Wei, H. Weng, Q. Hou & Y. Wei (eds) *Chinese spoken language processing (ISCSLP): 2016 10th international symposium*. IEEE, 1–5.
- McNamara, T. & U. Knoch 2012. The rasch wars: the emergence of Rasch measurement in language testing. *Language Testing*, 29 (4), 555–576.
- Ministry of Education and Culture 2017. *Gaudeamus igitur: ylioppilastutkinnon kehittäminen*. Opetus- ja kulttuuriministeriön julkaisuja 2017:16. Helsinki: Opetus- ja kulttuuriministeriö.
- Nakagawa, S. & H. Schielzeth 2013. A general and simple method for obtaining R<sup>2</sup> from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4 (2), 133–142.
- Pinget, A.-F., H. R. Bosker, H. Quené & N. H. De Jong 2014. Native speakers’ perceptions of fluency and accent in L2 speech. *Language Testing*, 31 (3), 349–365. DOI: 10.1177/0265532214526177.

- Préfontaine, Y., J. Kormos & D. E. Johnson 2016. How do utterance measures predict raters' perceptions of fluency in French as a second language? *Language Testing*, 33 (1), 53–73. DOI: 10.1177/0265532215579530.
- Rasch, G. 1960. *Probabilistic models for some intelligence and achievement tests*. Chicago (Ill.): University of Chicago Press.
- Rossiter, M. J. 2009. Perceptions of L2 fluency by native and non-native speakers of English. *The Canadian Modern Language Review*, 65, 395–412. DOI: 10.3138/cmlr.65.3.395.
- Scherer, K. R. 1986. Vocal affect expression: a review and a model for future research. *Psychological Bulletin*, 99 (2), 143.
- Scherer, K. R., H. London & J. J. Wolf 1973. The voice of confidence: paralinguistic cues and audience evaluation. *Journal of Research in Personality*, 7 (1), 31–44.
- Suni, A., J. Šimko, D. Aalto & M. Vainio 2017. Hierarchical representation and estimation of prosody using continuous wavelet transform. *Computer Speech & Language*, 45, 123–136. DOI: 10.1016/j.csl.2016.11.001.
- Toivola, M., M. Lennes & E. Aho 2009. Speech rate and pauses in non-native Finnish. In M. Uther, R. Moore & S. Cox (eds) *Interspeech-2009*. Brighton: ISCA, 1707–1710.
- Towell, R., R. Hawkins & N. Bazergui 1996. The development of fluency in advanced learners of French. *Applied Linguistics*, 17, 84–119. DOI: 10.1093/applin/17.1.84.
- Trofimovich, P. & W. Baker 2006. Learning second language suprasegmentals: effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28 (1), 1–30. DOI: 10.1017/S0272263106060013P.
- Wennerstrom, A. 2000. The role of intonation in second language fluency. In H. Rigenbach (ed.) *Perspectives on fluency*. Ann Arbor (Mich.): University of Michigan Press, 102–127.
- Witt, S. M. 2012. Automatic error detection in pronunciation training: where we are and where we need to go. In O. Engwall (ed.) *Proceedings of the international symposium on automatic detection of errors in pronunciation training: Stockholm, Sweden*. Stockholm: KTH Royal Institute of Technology, 1–8.
- Xu, Y. 2013. ProsodyPro: a tool for large-scale systematic prosody analysis. In B. Bigi & D. Hirst (eds) *TRASP 2013: tools and resources for the analysis of speech prosody*. Aix-en-Provence: Laboratoire Parole et Langage, 7–10.



Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. *Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 214–233.*

**Mikko Kuronen & Elina Tergujeff**

University of Jyväskylä

## **L1 listeners' perception of L2 pronunciation: effect of prosody on accentedness ratings in Swedish**

This paper reports on a study in which L1 speakers (n = 53) of Swedish rated the accentedness of read-aloud sentences produced by L2 speakers (n = 6) of Swedish. The aim of the study was to investigate the effect of certain prosodic features on the accentedness ratings. These features include Swedish tone accents 1 and 2 and deviant speech rhythm (e.g. lack of primary stress, and realization of multiple stresses). The results reveal that utterances with deviant speech rhythm yielded perceptions of stronger foreign accent than utterances with target-like rhythm. As for the tone accents, their effect on the scores were non-existent or only marginal. As missing tone accents have been previously found not to compromise intelligibility, the results of the present study reinforce the conclusion according to which emphasis in teaching should be put on prosodic features other than tone accents, especially rhythm.

**Keywords:** pronunciation, prosody, accentedness, tone accents, rhythm

## 1 Introduction

Second language (L2) pronunciation is mainly studied by acoustic analysis and auditory tests. The acoustic characteristics of L2 Swedish have been studied concerning vowels (Hertteli 2015; Kuronen 2016), consonants (Kokkonen 2016; Kuronen 2016), tone accents (Kaiser 2011; Tronnier & Zetterholm 2014; Hed 2014; Kuronen et al. 2016), and rhythm and utterance intonation (Kuronen & Tergujeff 2018). Acoustic studies give us detailed information on what L2 pronunciation is like in comparison to first language (L1) pronunciation. In turn, auditory tests reveal how listeners experience L2 pronunciation: how comprehensible it is, how accented it is, and how pleasant it sounds. These issues are investigated by using scalar evaluations, whereas methods such as dictation are used to measure actual intelligibility (Derwing & Munro 2005). Most of the studies on L2 Swedish have been either acoustic-phonetic or they have focused on accentedness and/or comprehensibility. Thus, there is a lack of studies on relations between acoustic features and perception of accentedness.

In this study, focus is on the phonetic basis of L2 accentedness – not intelligibility or comprehensibility (cf. Derwing & Munro 1997). Our goal is to produce new knowledge on how prosodic features affect accentedness ratings by L1 listeners. We aim to analyse the perceptual importance of the main features of utterance prosody in Sweden-Swedish: tone accents, rhythm, and utterance intonation (Bruce 1977; Myrberg 2010; Riad 2014). These are difficult learning goals for L2 speakers, but there is very little knowledge on the perceptual relevance of the features. The results are applicable in teaching of pronunciation especially to advanced L2 learners of Swedish, irrespective of their L1. For example, if tone accents do not affect L1 listeners' perception of accentedness, not much effort needs to be spent by L2 learners to acquire them. Most of the previous studies on phonetic aspects of L2 perception have concerned English (see e.g. Jesney 2004; Derwing & Munro 2015). Since learning and perception of prosody is partly language-dependent, the present study on Swedish will hopefully elucidate new aspects on how L2 pronunciation is perceived.

## 2 Previous studies on perception of accentedness in L2 Swedish

There are numerous studies on how listeners perceive L2 English (e.g. Flege & Fletcher 1992; Bongaerts et al. 1997; Derwing & Munro 1997; Piske et al. 2001), and several of them are phonetically detailed (e.g. Flege 1984; Jilka 2000). The perception of L2 pronunciation in other languages has also been studied in some detail, such as in French (e.g. Flege 1987; Birdsong 2007) and

in German (e.g. Missaglia 2007). However, there are only a few studies on how Swedish listeners perceive and rate accentedness based on phonetic features. The main findings of these studies will be discussed below.

Cunningham-Andersson & Engstrand (1989) analysed the perceptual effects of different phonetic features in Finnish- and British-accented L2 Swedish, as imitated by the second author of the study. 14 phonetic features were combined to make 35 different imitations of a short text, which Swedish listeners ( $n = 35$ ) rated on a scale from 0 to 4. The focus of the analysis was on segments. Also, exaggerated quantity distinctions and mixing of tone accents were analysed. The most significant L2 features of Finnish-accented L2 Swedish were velarized /l/, unaspirated stops, trilled /r/, exaggerated quantity distinctions, and mixed tone accents (accent 2 pronounced as accent 1). Occurrence of one of these features was enough for 50% of the listeners to perceive a foreign accent, and two or more of the features increased the rating of a speech sample as foreign-accented. Identification of foreign accent as Finnish had a positive correlation with an increasing number of the above-mentioned features. Further, listeners were quite good at identifying the imitated accent as Finnish (60–100%), a finding also made by Cunningham-Andersson & Engstrand (1988a,b), Bijvoet (1996), and Boyd et al. (1999) with authentic speech material.

Kuronen & Zetterholm (2017) also used imitations. They varied concerning segmental qualities, tone accent production, utterance intonation, and rhythm. Swedish listeners ( $n = 30$ ) rated the accentedness of the imitations (a short text) on a scale from 1 to 6. Segmental deviations from L1 Swedish affected the ratings much more than deviations in utterance intonation, rhythm, and/or tone accents. Imitations with L2-like intonation and native-like segments were rated in the same way as imitations with native-like intonation and native-like segments. If segments were native-like, imitations with native-like rhythm and L2-like rhythm were rated similarly. Derwing & Munro (1997: 11) have made a similar finding concerning English: segments were the main cause for perception of foreign accent (92%), followed by grammar (46%), enunciation (mumble, 38%), prosody (23%), and speech rate (15%), as evaluated by the listeners themselves. Based on these two studies, segmental deviations seem more easily perceived as L2 features than prosodic deviations (at least in Swedish and English). This may be due to the fact that segment quality varies less audibly than prosody between different speaking styles, situations, and speakers. In other words, listeners may be more used to prosodic variation than segmental variation in their L1. Concerning Swedish, the dialectal variation might also be more noticeable in prosody than in segments (perhaps disregarding the dialects in Scania and Gotland). This too can make Swedish listeners more tolerant towards prosodic than segmental deviations when asked to rate accentedness of L2 pronunciation.

Abelin & Thorén (2015a,b) analysed the relative importance of incorrect word stress placement versus deviant tone accent. A male phonetician pronounced two-syllable words (e.g. *bilen* (Eng. *the car*), *sagan* (Eng. *the fairy tale*)) and non-words with varying stress placement and intonation. The listeners ( $n = 18$ ) were asked to decide under time pressure whether the word was a real word or a non-word in Swedish. Incorrect stress placement caused much more non-word answers than a mispronounced tone accent (accent 2 as accent 1, the opposite was not analysed). Thus, word stress placement was clearly more important for word recognition than tone accent, so it should also affect the perceived accentedness much more than tone accent.

To our knowledge, there are no other studies on the connection between phonetic features and perceived accentedness in Swedish than the ones summarised above. Taken together, the results suggest that segments are more crucial for perception of accentedness than prosody, as long as the deviations are at a phonetic level, i.e. non-phonemic. In other words, listeners seem to be more sensitive to segmental than prosodic deviations when rating accentedness. Of course, prosodic deviations at phonological level (such as incorrect word stress placement) are decisive for perception of accentedness. The speech materials in Cunningham-Andersson & Engstrand (1989) and Kuronen & Zetterholm (2017) were imitations, and since the findings concerning prosodic features in particular have not been investigated in other studies on L2 Swedish, we set our focus solely on prosody in this study. Authentic L2 speech was used as speech material, and the experimental setup was developed from Cunningham-Andersson & Engstrand (1989) and Kuronen & Zetterholm (2017).

Before we advance to the empirical part, it is appropriate to state that we think that a communicatively functional pronunciation (cf. Levis 2005) is the first priority for L2 speakers. However, the phonological aspects of L2 may not be sufficient learning goals for all advanced learners. If that was the case, the pronunciation of e.g. *kultur* (Eng. *culture*) as [kul'tu:ːr] and *kaka* (Eng. *biscuit*) as ['ka:k:a] (typical pronunciations for L1 Finnish-speaking learners of Swedish) would be adequate goals instead of [kəl'tʰɥ:ːɹ] and ['kʰɔ:k:aː] (L1 Swedish). Yet, [kul'tu:ːr] and ['ka:k:a] would certainly give an impression of a strong foreign accent. Learning goals vary depending on L2 speakers' ambitions and needs: [kul'tu:ːr] and ['ka:k:a] should be well intelligible in most contexts and can perhaps serve as basic learning goals for L2 speakers. Still, aspiration, vowel quality, post-vocalic consonant duration (in *kaka*), and/or tone accents must be adjusted/acquired if an L2 speaker aims higher than just to make him- or herself understood. This is true especially when dealing with L2 speakers who wish to acquire comfortable intelligibility or native-like pronunciation. Comfortable intelligibility is important for both professional and non-professional language users because a strong foreign accent will put more

strain on the listener, and can affect listeners' attitude toward the speaker (Mennen 2007; Torstensson 2010; Boyd & Bredänge 2013). Comfortable intelligibility can also improve an L2 speaker's self-confidence as a language user.

### 3 Empirical part

In the following sections, we will describe our aims and research questions (3.1), speech material (3.2), and methods (3.3).

#### 3.1 Aims

The present paper is based on acoustic studies conducted by the authors (Kuronen 2015; Kuronen & Tergujeff 2018). The results of these studies show that many L2 speakers of Swedish can develop prosodically during phonetic training: many participants learned to produce tone accents, their utterance intonation became more native-like, they learned to use duration as a stress correlate, and they acquired a native-like rhythm. Here, our aim is to find out how important the attainment of these features is for L1 Swedish listeners. Segmental qualities are not analysed since they have shown to be pivotal in previous studies (see Chapter 2). The following research questions are addressed:

1. Do tone accents affect foreign accentedness ratings by L1 Swedish listeners? Are tone accents 1 and 2 equally important or unimportant? Is native-like tone accent 2 more important in compounds than in simple words?
2. Does mixing up tone accents affect foreign accentedness ratings by L1 Swedish listeners?
3. Does lack of clear primary stress and tendency not to distress syllables outside the primary stress affect foreign accentedness ratings by L1 Swedish listeners?
4. Do multiple stresses affect foreign accentedness ratings by L1 Swedish listeners?

Earlier knowledge on the effect of tone accents is insufficient. According to Cunningham-Andersson & Engstrand (1989), mixing of tone accents has a negative perceptual effect, but lack of tone accents was not analysed in their study. Still, lack of tone accents is probably a more common trait in L2 Swedish than mixing of them (Kuronen 2015). In the present study, both lack of tone accents and mixing of them were analysed. Our assumption was that mixing of

tone accents has a negative perceptual effect, while lack of them has merely a marginal effect due to listeners' tolerance for dialectal and/or situational variation in prosody. However, one could expect that a native-like tone accent 2 would have some positive effect due to its distinct tonal gesture.

Tone accent 2 has a connective function in compounds (Elert 1981; see also Riad 2015). Compounds are pronounced in Standard Swedish (Central-Swedish, Riad 2014) with two stressed syllables and a high+low+high (H\*LH) contour with H in the stressed (= long) syllables, e.g. *lastbilschaufför* (Eng. *lorry driver*) ['la`s:tbi:lsʃafœ:´ɹ]. Due to two stressed syllables in compounds, tone accent 2 might be more important for native-like pronunciation in compounds than in simple words (e.g. *kappa*, Eng. *coat*). If listener ratings are similar irrespective of intonation in the primary stressed word, it would support the idea of tone accents not being a purposeful learning goal for L2 speakers of Swedish.

Research questions 3 and 4 have been studied earlier only by Kuronen & Zetterholm (2017) with imitations of a text sample (see Chapter 2). The L2 speaker was highly fluent in prosody, and therefore the imitations might not be representative for a more ordinary L1 Finnish-speaking learner of Swedish. Also, only native-like rhythm and rhythm with lack of stress(es) were studied. Here, three different rhythmic patterns are studied: (i) native-like rhythm with primary stress and distressing of syllables (= a tendency to shorten them) in the unstressed words in the utterance, (ii) lack of primary stress and no distressing of lexical (= phonological) stresses, and (iii) multiple stresses. Two last-mentioned deviations are typical for L2 speakers despite their L1 (Rasier & Hiligsmann 2007; on Finnish-Speakers' Finland-Swedish see Kautonen 2017). As a consequence of the rhythmic differences, the utterance intonation varies as follows: (i) in utterances with native-like rhythm, the contour is non-descending with the highest  $f_0$  peak in the primary-stressed syllable, (ii) in utterances with lacking primary stress, the contour is descending after the initial  $f_0$  peak, and (iii) in utterances with multiple stresses, there are several distinct  $f_0$  peaks (see Figures 1–5 in section 3.3).

Knowledge on how listeners rate different rhythmic patterns is important when teaching pronunciation to L2 speakers. Native-like rhythm is considered an important learning goal for L2 speakers of Swedish (e.g. Bannert 2004), but knowledge on perception of different deviating patterns is scarce.

### 3.2 Speech material

The speech material consisted of read-aloud sentences; mainly statements ( $n = 13$ ), but a few questions ( $n = 3$ ) were also included. The utterances had (or should have had) primary stress towards the end, and with tone accent 1 or 2 depending on the utterance. The utterances were 1.5 to 2.5 seconds long,

and they were prosodically quite simple, meaning they did not vary much regarding primary stress placement, tone accents, rhythm, and utterance intonation, as pronounced by native-speaker controls. Below, five example utterances from the speech material are given. The usual primary stress placement in L1 pronunciation is marked with blue.

- |  |  |
|--|--|
| 1. Ska vi åka <i>ut</i> ikväll?              | Are we going <i>out</i> tonight?             |
| 2. Jag ska köpa en <i>matta</i> till pappa.  | I will buy a <i>carpet</i> to my father.     |
| 3. Vi ska till <i>stugan</i> ikväll.         | We are going to the <i>cottage</i> tonight.  |
| 4. Jag vill köpa en <i>kappa</i> till mamma. | I want to buy a <i>coat</i> for (my) mother. |
| 5. Det kan <i>blåsa</i> ikväll.              | It might get <i>windy</i> tonight.           |

We preferred to use authentic L2 speech instead of synthesized speech for two reasons. Firstly, we had collected a large L2 speech corpus, which enabled us to find utterances with appropriate phonetic characteristics; tokens with prosody of the wanted types. Secondly, synthetic editing of rhythm and utterance intonation can easily make utterances sound unnatural. Use of read-aloud sentences instead of (longer) samples of free speech is based on the assumption that in read-aloud sentences listeners can more easily focus on the accentedness, because the contents have been controlled (e.g. no grammatical errors). Influence of the speaker's voice can be assumed weaker in short utterances than in longer speech samples. In longer samples, accentedness may be easier to rate, but phonetic factors behind the ratings are very difficult to distinguish.

The L2 speakers ( $n = 6$ , all female) who produced the test utterances were high-intermediate/advanced speakers of Swedish with Finnish as their L1. L1 speakers ( $n = 2$ , females from Stockholm) were used for reference. The L2 speakers were 21–36 years old, while the L1 speakers were 24 and 25 years old. The L2 speakers had studied Swedish for at least six years as part of their secondary education before their current studies of Swedish as a major or minor subject at a Finnish university. They took part in a course in Swedish pronunciation and oral skills, and their pronunciation was recorded in a language lab before and after the course. The recordings were saved in a sound file format WAV. The recordings of the L1 speakers were made in a recording studio and saved in a sound file format AIFF. Both file formats used in the collection of speech materials are uncompressed formats. The speech material for the present study was chosen from these recordings. The quality of the recordings for the L1 and L2 speakers was similar.

### 3.3 Auditory test

We chose 30 utterances for the auditory test from a larger material of ca. 250 utterances. 24 of the utterances were pronounced by L2 speakers, 3 utterances by L1 speakers, and 3 utterances were presented twice to the listeners

to test the reliability of the ratings. We used the same L2 speakers in as many of the utterances as possible: 17 of 24 L2 utterances were pronounced by only three speakers. All utterances in which the effect of tone accents was analysed were pronounced by these three speakers. Most of the utterances occurred twice in the auditory test representing different prosodic categories. Overall, the selection of L2 utterances was based on three criteria: (i) the segments are as native-like as possible so that defects in them would affect the ratings as little as possible, (ii) there are no pauses, hesitation, and/or corrections in the utterances, and (iii) in the utterances with tone accent, there is a clear primary stress (prominence level 2, Myrberg 2010), meaning that this stressed word should have either tone accent 1 (L\*H) or tone accent 2 (H\*LH), as was the case in the native speakers' pronunciation.

The L2 speakers have some segmental deviations from L1 pronunciation. These deviations may affect the ratings, but they are similar in all analysed prosodic categories, because (i) the speakers are the same in the vast majority of the utterances, (ii) the speakers have the same kind of segmental deviations, primarily caused by their L1 Finnish, and (iii) we have chosen the utterances so that segmental differences between them are as minor as possible. Thus, the outcome of the auditory test should substantially depend on the prosodic character of the utterances, not the segmental quality in them.

The utterances were chosen as follows: The first author listened to the larger material, and chose 40 suitable utterances, which were presented to the second author. Only when both authors categorised a certain utterance belonging to one of the eight categories below, the utterance was used in the test. In other words, the material selection was a result of careful negotiation between the authors. The prosodic categories that the chosen utterances demonstrate are:

1. correct primary stress and rhythm, no tone accent 1 (H\*L)
2. correct primary stress and rhythm, tone accent 1 (L\*H)
3. correct primary stress and rhythm, no tone accent 2 (H\*L)
4. correct primary stress and rhythm, tone accent 2 (H\*LH)
5. correct primary stress and rhythm, mix-up of tone accents (accent 1 pronounced as accent 2, and vice versa)
6. correct primary stress and rhythm, no tone accent 2 in a primary-stressed compound with four or more syllables
7. L2-like rhythm with lack of distinct primary stress and no distressing of the phonologically stressed syllables in the utterance



8. L2-like rhythm with multiple stresses, short stress groups, no distinct primary stress
9. L1 Swedish speakers' pronunciation.

We used three utterances per category in the auditory test (3 x 8 L2 prosodic categories + 3 L1 utterances + 3 control utterances by L2 speakers = total of 30 utterances). This was done so that minor differences in segment quality and other possible individual speaker qualities would not skew the results. No more than 30 utterances were used so that the auditory test would not take more than 15 minutes to finish.

In Figures 1–5, segment quality and reduction (cf. transcription), syllable durations (stress and rhythm), and intonation in five utterances belonging to five different prosodic categories are presented as examples. In these figures all the varied prosodic features are illustrated: L1- and L2-like utterance intonations (Figures 1 and 4, respectively), L1- and L2-like tone accent productions (Figures 2 and 3, respectively), and L1- and L2-like rhythmic patterns (Figures 1 and 5, respectively). The analyses have been conducted with Praat (Boersma & Weenink 2016).

The listeners were asked to rate the accentedness of each utterance on the following scale:

- 6 = native-like
- 5 = really good pronunciation, almost native-like
- 4 = quite good pronunciation with an audible foreign accent
- 3 = good pronunciation with a clear foreign accent
- 2 = quite poor pronunciation with a strong foreign accent
- 1 = very poor pronunciation with a very strong foreign accent

In most of the earlier studies, 5 to 9 point scales have been used (Jesney 2004). Here, we used a 6 point scale because it worked well in a similar study made earlier by the first author (Kuronen & Zetterholm 2017). Further, we wanted to obtain our results in a form that is easy to compare with the previous study. A scale with 7 to 10 points would probably bring forth even more subtle differences, but the description of the scale is more difficult with a larger scale, and we wanted to describe the scale points for precision. Because the L2 speakers had quite good pronunciation skills, we anticipated that the listeners would mainly use the upper part of the scale.

The auditory test was done with an online survey tool SoGoSurvey. Before the test, the listeners were told that (i) the context of the test was learning

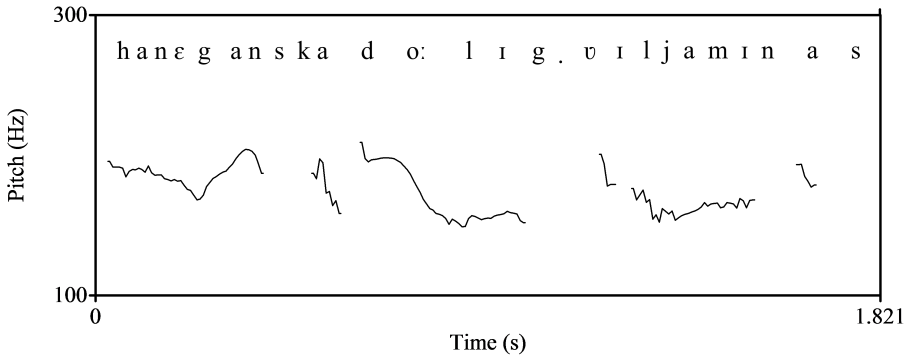


FIGURE 1. L2 speaker's pronunciation of *Han är ganska dålig, vill jag minnas* (Eng. *He is quite bad, I recall*). The utterance belongs to group 3: correct primary stress and rhythm, no tone accent 2 in primary stress (H\*LL instead of H\*LH). The most obvious prosodic L2 trait is lack of tone accent 2 despite a clear primary stress in *dålig* (Eng. *bad*).

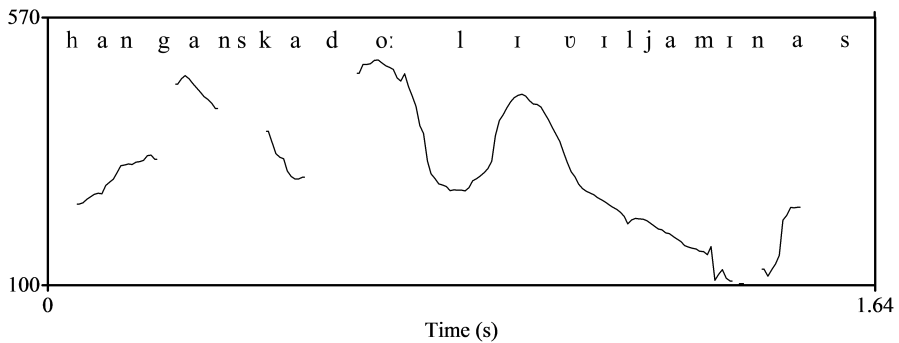


FIGURE 2. L1 speaker's pronunciation of *Han är ganska dålig, vill jag minnas* (Eng. *He is quite bad, I recall*). The most obvious prosodic difference compared with Figure 1 is tone accent 2 (H\*LH) in the primary stressed *dålig* (Eng. *bad*).

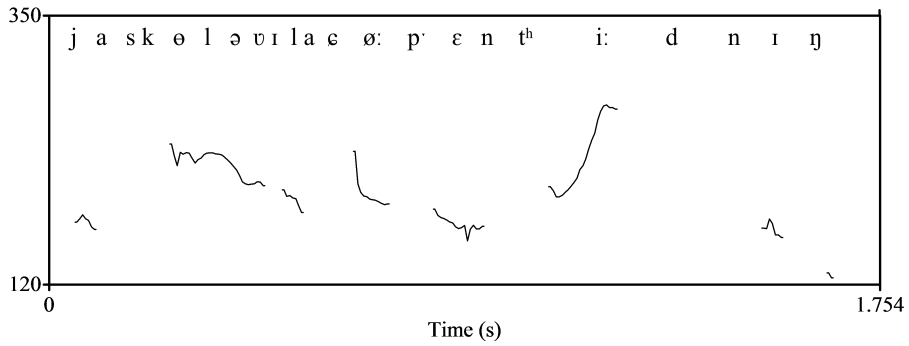


FIGURE 3. L2 speaker's pronunciation of *Jag skulle vilja köpa en tidning* (Eng. *I would like to buy a newspaper*). The utterance belongs to group 5: correct primary stress and rhythm, mixed tone accents (tone accent 2 realized as tone accent 1) in the primary stressed *tidning* (Eng. *newspaper*).

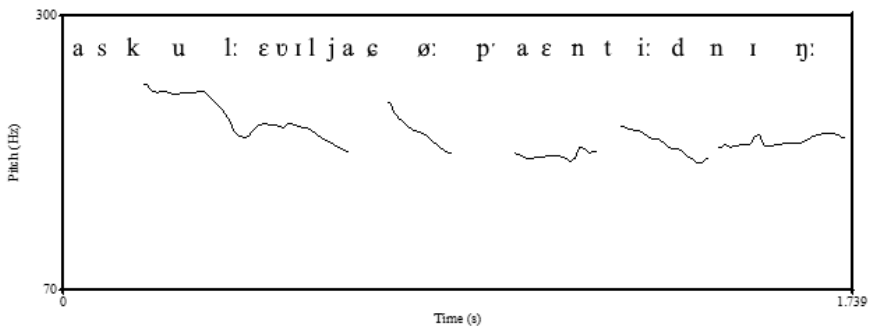


FIGURE 4. L2 speaker's pronunciation of *Jag skulle vilja köpa en tidning* (Eng. *I would like to buy a newspaper*). The utterance belongs to group 7: L2 like rhythm with lack of a distinct primary stress and no distressing of the phonologically stressed syllables in the utterance.

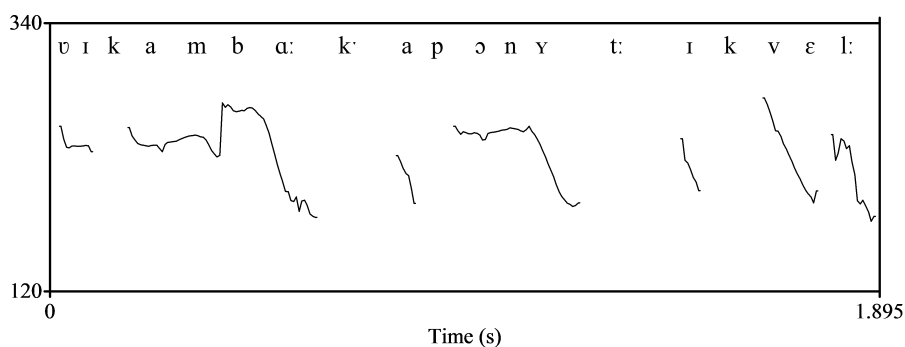


FIGURE 5. L2 speaker's pronunciation of *Vi kan baka på nytt i kväll* (Eng. *We can bake again tonight*). The utterance belongs to group 8: L2 like rhythm with multiple stresses, short stress groups, no distinct primary stress. Here, the stressed syllables are *baka* (Eng. *bake*), *pånytt* (Eng. *again*), and *ikväll* (Eng. *tonight*).

of Sweden-Swedish pronunciation, (ii) they will hear 30 utterances, (iii) their only task is to rate the accentedness of each utterance on a given scale, and (iv) the utterances have no major defects in pronunciation, e.g. completely incorrect speech sounds. The last-mentioned information was given to offer the listeners prior knowledge about the speakers' advanced proficiency in Swedish. The rating scale and the description of it were shown on the screen below the sound icon for each utterance. The utterances were presented to the listeners in a randomised order. The listeners were instructed to listen to each utterance as many times as they wanted. Statistical analyses were done using a reliability alpha (Cronbach's  $\alpha$ ) and independent samples t-tests. Effect sizes of the statistical significances were calculated using Cohen's  $d$ .

The listeners ( $n = 53$ ) were adults with Sweden-Swedish as their L1 and with no reported hearing problems nor studies in the Finnish language. They were invited through the authors' contacts in Sweden to take the listening test using their personal computers and headphones. No previous experience in accentedness ratings nor phonetic expertise was required. The majority of the listeners had an East Central Swedish pronunciation (cf. Standard Swedish; Riad 2014). They were not asked to rate the utterances in regards to any specific variety of Sweden-Swedish, because such information probably would have confused them: if e.g. East Central Swedish or Standard Swedish had been mentioned in the instructions, some of the listeners might have thought that they are unable to define these (regional) varieties and/or unable to do the ratings for some other reason. We knew before the test that most of the listeners resided in areas where East Central Swedish is spoken, and we thought that it was natural to them to rate the utterances with regard to East Central Swedish, not e.g. the variety spoken in Scania. Further, L2 speakers had Stan-

ard Swedish as their target variety, which could be heard in segment qualities – a circumstance that also directed the listeners to rate the utterances with regard to Standard Swedish.

## 4 Results and analyses

In the following sections, we will report the results. Scale and listener reliability are commented in section 4.1. The ratings by prosodic feature and by speech sample are reported and analysed in section 4.2.

### 4.1 Scale and listener reliability

When rating the 24 test samples pronounced by L2 speakers, mean scores given by the raters ( $n = 53$ ) varied from 2.54 to 5.79. Thus, some raters were much stricter in their scores than others. However, the listeners were consistent in their ratings: a reliability alpha (Cronbach's  $\alpha$ ) for all ratings, including native-speaker controls and check samples, was excellent ( $\alpha = 0.949$ ), indicating very high rater consistency and reliability of the scale.

Native-speaker controls and check samples served well in their purpose. Raters made a clear distinction between L2 speakers and native speakers by scoring the native-speaker samples substantially higher than the L2 speaker samples. Mean score for L2 speakers was 4.18 (SD 1.10), whereas it was 5.98 (SD 0.14) for native speakers. Independent samples t-test yielded a statistically significant difference between scores received by L2 speakers and native speakers ( $p = 0.000$ ), and the effect size was very large ( $d = 2.91$ ). On the contrary, no statistically significant differences were found between the two rounds of ratings for the three check samples ( $p = 0.44$ ,  $p = 0.61$ ,  $p = 0.55$ ).

### 4.2 Accentedness ratings

The mean scores for each prosodic feature category are presented in Figure 6. The results reveal that mean scores for the six different tone accent categories were between 4.20 and 4.73, whereas deviant rhythm (lack of primary stress or multiple stresses) received mean scores closer to 3 (3.48 and 3.18, respectively). No major differences were found in mean scores for tone accent categories, but the mean score for target-like accent 2 was the highest (4.73). Mixing up accent 1 and accent 2 did not affect the ratings. The accent mix-up category received a mean score of 4.57, which is almost as high as the mean score for target-like accent 2 and higher than the mean score for target-like accent 1 (4.20). Standard deviation for target-like accent 2 was 0.89, whereas it was between 0.95 and 1.01 for the other categories.

When scores for samples presenting no tone accents and target-like accents were subjected to an independent samples t-test, no statistically significant difference was found between no accent 1 and target-like accent 1 ( $p = 0.53$ ). This was not the case with accent 2, for which the t-test resulted in a statistically significant difference ( $p = 0.009$ ) between missing tone accent and target-like productions. However, the effect size was small ( $d = 0.29$ ). When scores for target-like accent 2 were compared to scores of the accent mix-up category, no statistically significant difference was found ( $p = 0.14$ ). Similarly, the difference between target-like accent 2 and no accent 2 in compounds was not statistically significant at the 0.05 level ( $p = 0.06$ ).

The outcome that native-like tone accents have no or only a very marginal perceptual effect is in line with the findings in Kuronen & Zetterholm (2017). Still, it is surprising that not even missing accent 2 in compounds got lower ratings than correct tone accent 2. Further, it is curious that mixing up tone accents did not affect the ratings.

The results concerning tones might be caused by two factors: (i) listeners are very tolerant for variation in word intonation in both L1 and L2 Swedish, and/or (ii) they hear that the speakers are non-native in Swedish, and therefore ignore the tonal gesture in the primary-stressed word. The latter would mean that once an L1 listener hears that a speaker is an L2 speaker, the tonal gesture in the stressed word does not affect the rating. Further, the L2 speakers in this study may have slight differences in their productions of tone accents if compared to native speakers. Yet, these speakers had somewhat native-like pronunciation, and if tone accents have no effect as produced by them, this should apply to many other L2 speakers as well.

The results concerning tone accents could also depend on the fact that listeners perceived the stressed words in L2 speakers' utterances as pronounced as prominence level 1 instead of prominence level 2. As prominence level 1, tone accents are pronounced as HL\* (tone accent 1) and H\*L(L) (tone accent 2) – in other words, with a descending contour like they were produced by L2 speakers in some of the utterances. Even with this aspect taken into account, it remains to be answered why mixing of tone accents did not have a negative effect on the ratings.

All six prosodic categories with native-like rhythm received much higher ratings than the two categories with rhythmic deviations (Figure 6). The difference between native-like rhythm categories and the two deviant rhythm categories was statistically significant ( $p = 0.000$ ), and the effect size was large ( $d = 1.13$ ). It is not surprising that multiple stresses had the strongest negative effect on the ratings, because this pattern differs from native-like rhythm by adding a disturbing component and making stress groups short. Somewhat higher but still relatively low scores were given to lacking primary stress and no distressing of phonologically stressed syllables (which is a common pro-

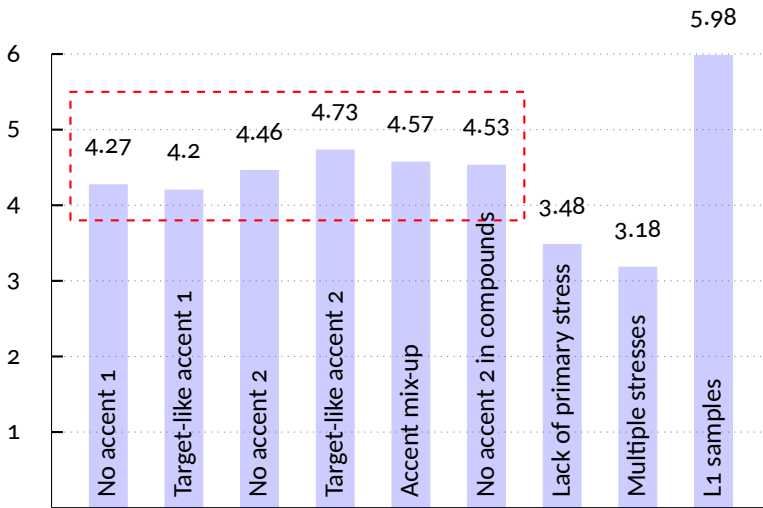


FIGURE 6. Mean scores by prosodic feature, including scores received by L1 samples. Utterances with native-like rhythm are marked with a rectangle.

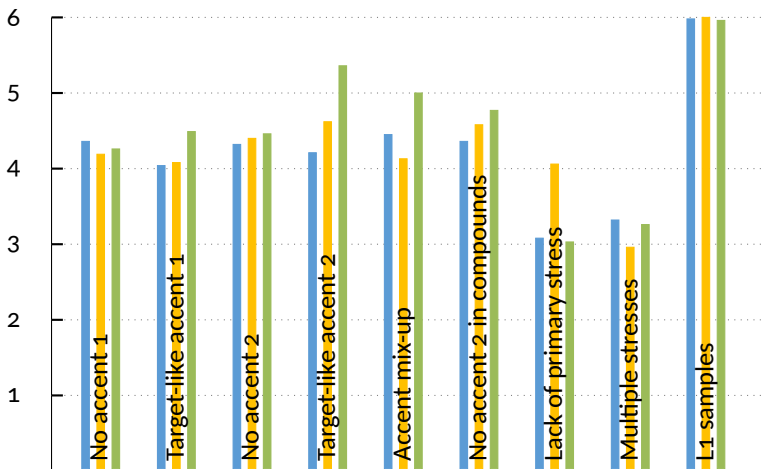


FIGURE 7. Mean scores by speech sample, including scores received by L1 samples.

cess in Swedish; Figure 6). The difference between multiple stresses and lack of primary stress is statistically significant ( $p = 0.009$ ). However, the effect size was small ( $d = 0.29$ ). In Kuronen & Zetterholm (2017), lack of primary stress had a very marginal negative effect on accentedness ratings. The difference between the findings here and in Kuronen & Zetterholm is most probably due to the fact that the speaker in Kuronen & Zetterholm was more native-like in other aspects of pronunciation than the L2 speakers in utterances with lacking primary stress in this study; his articulation rate and fluency were better. In that kind of pronunciation, lack of primary stress can be considered merely as a situational variation by L1 listeners, not as a trait of L2 pronunciation.

Ratings by feature were formed of scores received by three speech samples per feature, i.e. 159 scores in total. Figure 7 presents mean scores for each speech sample. The figure shows that not all speech samples within one category scored equally. Most variation was found within the following categories: target-like accent 2, no accent 2 in compounds, and lack of primary stress. This is probably due to the minor segmental differences between the utterances. In addition, a somewhat lower speech rate and  $f_0$  may have lowered some scores. Aside from three utterances, the coherence of the ratings is quite evident. Also, the native speakers were recognized as native speakers with a very high precision, presumably based on both prosody and segments in their utterances.

## 5 Summary and concluding remarks

The purpose of this study was to find out how important tone accents and different rhythmic patterns are for Swedish listeners' ( $n = 53$ ) ratings of L2 accentedness. The speech material consisted of 1.5 to 2.5 sec. long utterances with different prosodic characteristics. L1 speakers were used for reference. The results are summarized as follows:

- (i) Tone accents play no or only a very marginal role for L1 listeners' ratings.
- (ii) Not even mixing of tone accents had a negative perceptual effect.
- (iii) Utterances with compounds and without correct tone accent got ratings as high as utterances with correct tone accent 2.
- (iv) Rhythm with a clear primary stress and distressing of syllables outside primary stress got high ratings.
- (v) Rhythm with multiple stresses got lowest ratings of all.
- (vi) Lack of primary stress rendered also relatively low ratings.



Kuronen & Zetterholm (2017) found out that lack of tone accents did not have a negative perceptual effect even in longer speech samples (21–24 sec.). The speaker was a near native-like speaker of Swedish. The results here and in Kuronen & Zetterholm suggest that the acquisition of tone accents does not seem to make L2 pronunciation more native-like. Thus, even advanced L2 speakers of Swedish should focus in their learning on aspects of pronunciation other than tone accents.

In contrast to our results, Cunningham-Andersson & Engstrand (1989) found out that the mixing of tone accents had a clear negative perceptual effect. They used imitations by a native speaker. Further, the material consisted of a short text with multiple occurrences of mixed tones. Both of these factors can cause the difference between the findings here and in Cunningham-Andersson & Engstrand's study.

Concerning rhythm, our findings suggest that a clear primary stress and distressing of syllables outside the primary stress is an important feature for Swedish listeners. In Kuronen & Zetterholm (2017), lack of primary stress had a very marginal negative effect on accentedness ratings. The difference between the findings here and in Kuronen & Zetterholm is most probably due to the difference in articulation rate – a factor found to affect accentedness (Derwing & Munro 2015). Thus, the present study implies that articulation rate and mean *f*<sub>0</sub> may also play a role when L1 Swedish listeners rate accentedness in L2 Swedish. These factors should be addressed in future studies.

Segmental deviations were not a variable in our experimental setup. Yet, the results show indirectly that L1 listeners are very sensitive to them: the highest score given to an L2 utterance was 5.36; all other L2 utterances were rated below 5.0 despite the native-like prosody in some of them. Imitations with some deviating prosodic characteristics but native-like segments got considerably higher ratings (5.60–5.63) in Kuronen & Zetterholm (2017). These findings suggest that segments are very important if the goal of an L2 speaker is to acquire as unaccented pronunciation as possible.

The results concerning tone accents should not be interpreted as if they played *no role* in native-like pronunciation. Namely, there seems to be a connection between the acquisition of tone accents and other aspects of intonation such as overall contour, *f*<sub>0</sub> range, and mean *f*<sub>0</sub> (Kuronen & Tergujeff 2018). Thus, the learning of tones can facilitate the learning of other aspects of intonation, but the perceptual effect with native listeners still seems to be non-existing or at least very marginal.

## Acknowledgements

This study is part of a research project called *Fokus på uttalsinläringen med svenska som mål- och källspråk* (<https://www.jyu.fi/fokus>). The project is funded by The Society of Swedish Literature in Finland (SLS), which also supported Tergujeff's work with an additional research grant.

## References

- Abelin, Å. & B. Thorén 2015a. The relative perceptual weight of two Swedish prosodic contrasts. In E. Babatsouli & D. Ingram (eds) *Proceedings of the international symposium on monolingual and bilingual speech 2015*. Chania: Institute of Monolingual and Bilingual Speech, 1–7.
- 2015b. What affects recognition most: wrong word stress or wrong word accent? In M. Svensson Lundmark, G. Ambrazaitis & J. van de Weijer (eds) *Proceedings from Fonetik 2015: Lund, June 8–10 2015*. Working Papers 55. Lund: University of Lund, 7–10.
- Bannert, R. 2004. *På väg mot svenskt uttal*. Lund: Studentlitteratur.
- Bijvoet, E. 1996. Attityder till några varieteter av talad svenska. In M. Thelander, L. Elmevik, B.-L. Gunnarsson & B. Melander (eds) *Samspel och variation: språkliga studier tillägnade Bengt Nordberg på 60-årsdagen*. Uppsala: Department of Scandinavian Languages, University of Uppsala, 55–66.
- Birdsong, D. 2007. Nativelike pronunciation among late learners of French as second language. In O.-S. Bohn & M. J. Munro (eds) *Language experience in second language speech learning*. Amsterdam: John Benjamins, 99–116. DOI: 10.1075/lllt.17.12bir.
- Boersma, P. & D. Weenink 2016. PRAAT. <http://www.fon.hum.uva.nl/praat/>.
- Bongaerts, T., C. van Summeren, B. Planken & E. Schils 1997. Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19 (4), 447–465. DOI: 10.1017/S0272263197004026.
- Boyd, S., Å. Abelin & B. Dorriots 1999. Attitudes to foreign accents. In R. Andersson, Å. Abelin, J. Allwood & P. Lindblad (eds) *Fonetik 99: proceedings from the twelfth Swedish phonetics conference*. Gothenburg: University of Gothenburg, 31–35.
- Boyd, S. & G. Bredånge 2013. Attityder till brytning: exemplet utländska lärare i svenska skolor. In K. Hyltenstam & I. Lindberg (eds) *Svenska som andraspråk: i forskning, undervisning och samhälle*. 2nd ed. Lund: Studentlitteratur, 437–458.
- Bruce, G. 1977. *Swedish word accents in sentence perspective*. Lund: LiberLäromedel.
- Cunningham-Andersson, U. & O. Engstrand 1988a. Attitudes to immigrant Swedish: a literature review and preparatory experiments. In O. Engstrand, M. Dufberg & J. Stark (eds) *Perilus XIII*. Stockholm: Institute of Linguistics, University of Stockholm, 103–152.
- 1988b. Which foreign accents can Swedes identify? In *Papers from the second Swedish phonetics conference*. Lund: Department of Linguistics, University of Lund, 65–86.

- Cunningham-Andersson, U. & O. Engstrand 1989. Perceived strength and identity of foreign accent in Swedish. In O. Engstrand, M. Dufberg & C. Kylander (eds) *Perilus X: experiments in speech processes*. Stockholm: Institute of Linguistics, University of Stockholm, 65–86. <https://www.ling.su.se/perilus/perilus-x>.
- Derwing, T. M. & M. J. Munro 1997. Accent, intelligibility and comprehensibility: evidence from four L1s. *Studies in Second Language Acquisition*, 20 (1), 1–16. <https://www.cambridge.org/core/journals/studies-in-second-language-acquisition/article/div-classtileaccent-intelligibility-and-comprehensibilitydiv/729C15F62F9EC9A51A33EAB5C2D05ED0>.
- 2005. Second language accent and pronunciation teaching: a research-based approach. *TESOL Quarterly*, 39 (3), 379–397. DOI: 10.2307/3588486.
- 2015. *Pronunciation fundamentals: evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins.
- Elert, C.-C. 1981. *Ljud och ord i svenskan 2*. Uppsala: Almqvist & Wiksell.
- Flege, J. E. 1984. The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, 76 (3), 692–707. DOI: 10.1121/1.391256.
- 1987. The production of “new” and “similar” phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47–65.
- Flege, J. E. & K. L. Fletcher 1992. Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America*, 91 (1), 370–389.
- Hed, A. 2014. Perception and production of Swedish word accents by Somali L1 speakers. In M. Heldner (ed.) *Proceedings from FONETIK 2014: Stockholm, Sweden, 9–11 June 2014*. Stockholm: Department of Linguistics, University of Stockholm, 105–110.
- Hertteli, H. 2015. *Inlärnin g av sverigesvenska vokaler hos universitetsstudenter med finska som S1*. MA thesis. University of Jyväskylä. <http://urn.fi/URN:NBN:fi:jyu-201511213759>.
- Jesney, K. 2004. *The use of global foreign accent rating in studies of L2 acquisition*. Calgary: Language Research Centre, University of Calgary. <http://arts.ucalgary.ca/lrc/research/reports> [accessed 5/30/2017].
- Jilka, M. 2000. Testing the contribution of prosody to the perception of foreign accent. *New Sounds*, 4, 199–207.
- Kaiser, R. 2011. Do Germans produce and perceive the Swedish word accent contrast? a cross-language analysis. *Tal, musik, hörsel – Quarterly Progress and Status Report*, 51 (1), 93–96.
- Kautonen, M. 2017. Finskspråkiga talares intonation av finlandssvenska i påståendeyttranden i fritt tal. *Folkmålsstudier*, 55, 31–60. <http://urn.fi/URN:NBN:fi:jyu-201710314099>.
- Kokkonen, A.-M. 2016. *Avancerade finskspråkiga inlärares uttal av svenska och tyska frikativor*. MA thesis. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:NBN:fi:jyu-201612145097>.
- Kuronen, M. 2015. Tonaccenterna i avancerade finska inlärares svenska: en akustisk studie. *Nordand – Nordisk tidsskrift för andrespråksforskning*, 10 (1), 53–80. <http://urn.fi/URN:NBN:fi:jyu-201508042611>.
- 2016. Avancerade finskspråkiga inlärares uttal av segment i sverigesvenska. *Puhe ja kieli*, 36 (3), 175–196. <https://journal.fi/pk/article/view/59009/20417>.

- Kuronen, M. & E. Tergujeff 2018. Second language prosody and its development: connection between different aspects. *The Language Learning Journal*. (Online ahead of print). <https://www.tandfonline.com/doi/full/10.1080/09571736.2018.1434228>.
- Kuronen, M., R. Ullakonoja & M. Kautonen 2016. Inläringen av de svenska tonaccenterna hos finska S2-talare: automatiseras uttalet? *Språk och stil*, 26, 161–194. <http://urn.fi/URN:NBN:fi:jyu-201801221305>.
- Kuronen, M. & E. Zetterholm 2017. Olika fonetiska drags relativa betydelse för upplevd inföddlighet i svenska. *Nordand – Nordisk tidsskrift för andrespråksforskning*, 12 (2), 134–156. <http://urn.fi/URN:NBN:fi:jyu-201712194777>.
- Levis, J. 2005. Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39 (3), 369–377. DOI: 10.2307/3588485.
- Mennen, I. 2007. Phonological and phonetic influences in non-native intonation. In J. Trouvain & U. Gut (eds) *Non-native prosody: phonetic description and teaching practice*. Berlin: Mouton de Gruyter, 53–76.
- Missaglia, F. 2007. Prosodic training of Italian learners of German: the contrastive prosody method. In J. Trouvain & U. Gut (eds) *Non-native prosody: phonetic description and teaching practice*. Berlin: Mouton de Gruyter, 237–258.
- Myrberg, S. 2010. *The intonational phonology of Stockholm Swedish*. Stockholm: Department of Scandinavian Languages, University of Stockholm. <http://urn.kb.se/resolve?urn=urn%3Anbn%3Ase%3Asu%3Adiva-38697>.
- Piske, T., I. R. A. MacKay & J. E. Flege 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29 (2), 191–215. DOI: 10.1006/jpho.2001.0134.
- Rasier, L. & P. Hiligsmann 2007. Prosodic transfer from L1 to L2: theoretical and methodological issues. *Nouveaux cahiers de linguistique française*, 28, 41–66.
- Riad, T. 2014. *The phonology of Swedish*. Oxford: Oxford University Press.
- 2015. *Prosodin i svenskans morfologi*. Stockholm: Morfem.
- Torstensson, N. 2010. *Judging the immigrant: accents and attitudes*. Umeå: Department of Language Studies, University of Umeå.
- Tronnier, M. & E. Zetterholm 2014. Appropriate tone accent production in L2-swedish by L1-speakers of Somali? *Concordia Working Papers in Applied Linguistics*, 5, 722–736.

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 234–257.

**Sanna Olkkonen & Pauliina Peltonen**

Turun yliopisto

## **Mitä on toisen kielen sujuvuus? Näkökulmia kognitiivisen ja puhetuotoksen sujuvuuden tutkimuksesta**

In this review article, we combine two approaches to the study of second language (L2) fluency: cognitive fluency and utterance fluency. The former concentrates on cognitive processing and its limitations, whereas the latter involves analyzing fluency-related features from speech samples. Despite theoretical links, the research traditions associated with the approaches have been relatively separate. In addition to providing an overview of the two approaches, the similarities and differences between them are illustrated with results related to one aspect of (dis)fluency, repetitions. Bringing together results related to cognitive and utterance fluency highlights the usefulness of combining different viewpoints in L2 fluency research and demonstrates the need for further interdisciplinary dialogue to gain a comprehensive picture of L2 fluency. Together, the results of studies emphasizing different aspects of L2 fluency also have important implications for L2 fluency assessment.

**Keywords:** cognitive fluency, utterance fluency, second language acquisition

**Asiasanat:** kognitiivinen sujuvuus, puhetuotoksen sujuvuus, toisen kielen oppiminen

## 1 Johdanto

Suullista sujuvuutta käytetään usein yhtenä kielitaidon mittarina arvioinneissa, mutta haasteena on se, että sujuvuuden kriteerit ovat melko yleisluontoisia ja osittain epätarkkoja. Esimerkiksi Eurooppalaisessa viitekehelyssä (Euroopan neuvosto 2012) sujuvuus on yksi viidestä puhutun kielitaidon kriteeristä, ja sitä kuvaillaan toiseksi korkeimmalla, C1-tasolla seuraavasti: ”Pystyy ilmaisemaan ajatuksiaan sujuvasti ja spontaanisti, lähes vaivattomasti. Vain käsitteellisesti vaikeassa aiheessa voi esiintyä kielellistä epäröintiä.” (Euroopan neuvosto 2012: 53). Vastaavasti viitekehelyksen suomalaisessa sovelluksessa alimman, A1.1-tason sujuvuutta kuvataan seuraavasti: ”Puheessa voi olla paljon pitkiä taukoja, toistoja ja katkoksia.” (Opetushallitus 2004). Siinä missä Eurooppalaisen viitekehelyksen määritelmän ongelmana on konkreettisten puheen piirteiden puute, suomalainen kuvausasteikko sisältää melko epämääräisen viittauksen epäsujuvuuteen liitettyihin piirteisiin, joiden yhteys kielitaitoon on empiiristen tutkimustulosten valossa epäselvä. Esimerkiksi pitkien taukojen esiintymisen on ongelmallista, sillä joidenkin viimeaikaisten tutkimusten valossa taukojen pituus näyttäisi liittyvän enemmän yksilölliseen puhetyyliin kuin vieraan kielen taitotasoon (De Jong ym. 2015). Yksilöllisten erojen on esitetty näkyvän myös korjauksissa ja toistojen määrässä (esim. Kormos 1999; Kahng 2014). Sujuvuustutkimuksen yhtenä keskeisenä haasteena voidaan pitää sen määrittelyä, mitkä puheen piirteistä liittyvät kielitaidon tasoon ja mitkä puolestaan enemmän yksilölliseen puhetyyliin.

Tähän haasteeseen tarttumisessa voi auttaa sujuvuuden hahmottaminen moniulotteiseksi ilmiöksi. Esimerkiksi Segalowitzin (2010: 47) mukaan sujuvuustutkimuksessa on keskeistä kartoittaa erityisesti niitä piirteitä, jotka ”heijastavat kielenkäyttäjän prosessoinnin tehokkuutta kommunikatiivisesti hyväksyttävän puheen tuottamisessa”. Tähän määritelmään tiivistyvät sujuvuuden kolme ulottuvuutta: sujuvuuden piirteiden tarkastelu eli puhetuotoksen sujuvuus (*utterance fluency*), kognitiivinen eli prosessoinnin sujuvuus (*cognitive fluency*) sekä kuulijan tulkinta puhetuotoksen piirteiden luomasta kokonaiskuvasta eli havaittu sujuvuus (*perceived fluency*). Tavoitteenamme tässä artikkelissa on tarkastella toisen<sup>1</sup> kielen sujuvuutta kognitiivisen ja puhetuotoksen sujuvuuden näkökulmista (ks. myös De Jong 2016)<sup>2</sup>. Kognitiivisen sujuvuuden näkökulma keskittyy informaation prosessointiin ja sen rajoituksiin, kun taas puhetuotosta tarkastellessa mitataan sujuvuuteen liitettyjä piirteitä

<sup>1</sup> Käytämme termiä toisen kielen oppiminen (*second language acquisition*, SLA) kattamaan sekä toisen kielen omaksumisen että vieraan kielen oppimisen (ks. Pietilä & Lintunen 2014: 12–14).

<sup>2</sup> Havaittua sujuvuutta sivutaan alaluvussa 3 puhetuotoksen sujuvuuden tarkastelun yhteydessä. Kuulijoiden arviot perustuvat puheen piirteiden tulkintaan, joten ne tarjoavat tietoa kognitiivisesta sujuvuudesta vain epäsuorasti eikä niitä ole juurikaan käsitelty kognitiivisen sujuvuuden näkökulmasta.

puheesta. Näkökulmat ovat kuitenkin yhteydessä toisiinsa: puheen sujuvuuspiirteiden ajatellaan heijastavan taustalla olevien kognitiivisten prosessien tehokkuutta (Segalowitz 2010). Esimerkiksi laajalti toisen kielen oppimisen (SLA) sujuvuustutkimuksissa viitatus Lennonin (2000: 26) määritelmän mukaan sujuvuus on ”nopeaa, tarkkaa ja tehokasta ajatusten kielellistämistä reaaliaikaisen prosessoinnin rajoitusten puitteissa”. Toisin sanoen kognitiivista sujuvuutta tarkasteltaessa tutkimuskohteena ovat kognitiiviset prosessit itsessään, kun taas usein toisen kielen oppimisen sujuvuustutkimuksissa kognitiivinen sujuvuus on enemmän ollut sujuvuuskäsitykseen vaikuttava taustatekijä.

Teoreettisista yhteyksistä huolimatta kognitiivisen ja puhetuotoksen sujuvuuteen liittyvät tutkimusperinteet ovat käytännössä olleet verrattain erilliset. Eri näkökulmat yhdistävää teoreettista keskustelua on kansainvälisestikin vähän (mutta ks. Chambers 1997; Koponen & Riggenbach 2000; De Jong 2016), ja tarkastelun kohteena olevat aineistot ovat hyvin erityyppisiä näkökulmasta riippuen. Tämän katsausartikkelin tavoitteena on näkökulmia rinnakkain käsittelemällä analysoida näkökulmien yhtäläisyyksiä ja eroja sekä täydentää näistä perspektiiveistä suomenkielistä sujuvuuteen liittyvää kirjallisuutta (ks. Lauranto 2005; Lintunen & Peltonen 2016). Aiemmin sujuvuutta on käsitelty suomeksi esimerkiksi opetuksen (Nissilä ym. 2006), ääntämisen (Toivola 2011; Lintunen 2014), kirjoittamisen (Mutta 2006), arviointikäsitteiden (Ahola 2016), suullisen kielitaidon arvioinnin (Hildén 2000) sekä kognitiivisen suoritusnopeuden (Olkkonen 2012) näkökulmista. Sujuvuus mainitaan myös kielen oppimisen tutkimukseen johdatteluvissa teksteissä tarkkuuden ja kompleksisuuden ohella osana kielitaitoa (esim. Martin & Alanen 2011; Pietilä & Lintunen 2014).

Käsittelemällä katsauksessamme rinnakkain kognitiivisen ja puhetuotoksen näkökulmia artikkeli luo kattavamman kuvan toisen kielen sujuvuuteen liittyvästä tutkimuksesta. Alaluvussa 2 perehdymme siihen, millaiset tekijät vaikuttavat kognitiiviseen sujuvuuteen ja millaisia mittareita kognitiivisen sujuvuuden tarkasteluun voidaan käyttää. Alaluvussa 3 valotamme sujuvuustutkimusta puhetuotoksen näkökulmasta ja keskitymme erityisesti puhetuotoksen sujuvuuden määrittelyyn, mittareihin ja viimeaikaisiin tutkimustuloksiin. Alaluvussa 4 konkretisoimme kognitiivisen ja puhetuotoksen sujuvuuden näkökulmia perehtymällä yhteen sujuvuusmittariin, toistoihin, ja siihen liittyviin tuloksiin. Toistot valittiin tarkastelun kohteeksi, sillä vaikka ne yleensä yhdistetään epäsujuvuuteen, sekä kognitiivisen että puhetuotoksen sujuvuuteen keskittyneiden tutkimusten valossa toistojen ja epäsujuvuuden yhteys on epäselvä. Alaluvussa 5 laajennamme kognitiivisen ja puhetuotoksen sujuvuuden näkökulmien vertailua toistoista sujuvuuteen yleisesti. Alaluvussa 6 kokoamme artikkelin keskeisiä huomioita kognitiivisen ja puhetuotoksen sujuvuuden suhteesta.

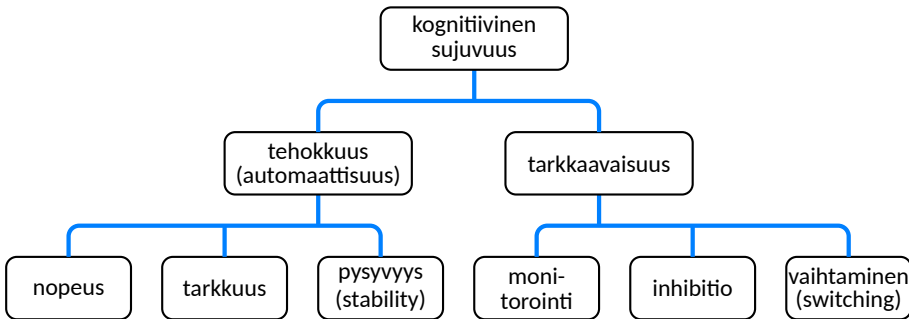
## 2 Kognitiivinen sujuvuus: prosessointinäkökulma sujuvuuteen

Kognitiivinen sujuvuus määritellään kaikkien kielenkäytön taustaprosessien tehokkaaksi toiminnaksi (Segalowitz 2000, 2010). Kognitiivisen sujuvuuden tärkeimpiä taustatekijöitä on informaatioprosessoinnin rajallinen kapasiteetti (Just & Carpenter 1992), johon vaikuttavat erityisesti työmuistin rajoitukset (ks. Baddeley 1996). Kognitiivisia prosesseja on kahdentasoisia, ja alemman tason prosessien tehokkuus määrittää sen, kuinka paljon resursseja riittää ylemmän tason käyttöön (LaBerge & Samuels 1974; Kirby ym. 2003). Kielenkäytössä alemman tason prosessit viittaavat esim. sanojen dekodeeraukseen ja sananhakuun, joiden automaattisuus on edellytys tiedon tehokkaalle prosessoinnille (Schmidt 1992; Segalowitz 2000, 2010: 91). Alemman tason prosessien tulee siis toimia nopeasti ja tarkasti, alitajuisesti ja vaivattomasti (Grabe 2009). Nämä prosessit vaikuttavat myös kielenkäytön sujuvuuteen: puhumisessa tämä tarkoittaa esimerkiksi, että ”normaali” puhenopeus edellyttää pääosin automaattista sananhakua (esim. Ehri 1991; Schmidt 1992; Segalowitz & Frenkiel-Fishman 2005). Servicen ym. (2002) tutkimus esimerkiksi osoitti, että vähemmän vierasta kieltä opiskelleet yliopisto-opiskelijat suoriutuivat samanaikaisesta lauseenymmärtämisen ja muistiinpainamisen tehtävästä huomattavasti huonommin kuin kieltä enemmän opiskelleet. Tämä kertoo siitä, että vähemmän kieltä opiskelleiden ryhmässä sanatason epäsujuvuus verotti työmuistin resursseja ylemmän tason toiminnoilta. Mitä automaattisemmin alemman tason prosessit toimivat, sitä enemmän resursseja ylemmän tason prosesseille riittää (Kormos 2000; Walczyk ym. 2004; Grabe 2009; Kahng 2014).

Ylemmän tason kontrollia vaativiin prosesseihin kuuluu esimerkiksi tarkkaavaisuuden ohjaus. Kontrollia vaativat prosessit ovat tahdonalaisia ja hitaita mutta mukautuvia, mikä on edellytys sujuvuuden ylläpitämiselle vuorovaiikutustilanteissa (Schmidt 1992). Tarkkaavaisuuden ohjaus mahdollistaa tehtävään keskittymisen ja vaihtelun eri tehtävien välillä (Segalowitz 2010: 93). Tarkkaavaisuuden ohjaus on joissain tutkimuksissa selittänyt jopa 59 prosenttia toisen kielen taitojen variaatiosta (Segalowitz & Frenkiel-Fishman 2005). Schmidt (1992) kuvaakin sujuvuutta ”pääosin automaattisten prosessien hallintana valikoivan tarkkaavaisuuden avulla”. Näiden ilmiöiden vaikutusta sujuvuuteen tarkastellaan seuraavaksi potentiaalisten kognitiivisen sujuvuuden mittareiden kautta.

Kognitiivisen sujuvuuden tutkimuksessa on käsitelty siihen vaikuttavien kognitiivisten osatekijöiden, kuten muistin (Georgiadou & Roehr-Brackin 2017), inhibition (Engelhardt ym. 2010) ja tehtävän tai kielen vaihtamisen (*switch cost*; Segalowitz & Frenkiel-Fishman 2005) suhdetta sujuvuuteen. Yhteisymmärrys siitä, kuinka kognitiivista sujuvuutta voitaisiin mitata kokonaisvaltaisemmin, kuitenkin puuttuu. Mittaamiseen käytetyt metodit ovat usein





KUVIO 1. Kognitiiviseen sujuvuuteen vaikuttavat tekijät (muokattu kuvioista Olkkonen 2017b: 18).

myös epäsuoria, kuten retrospektiiviset haastattelut, joissa osallistujia pyydetään jälkikäteen raportoimaan tehtävän suorittamiseen liittyvistä ajatuksistaan (Kahng 2014). Suuremmaksi kognitiivisen sujuvuuden mittariksi on esitetty sananhaun ja -tunnistuksen (*lexical access*) tehokkuutta, jota voidaan tarkastella esimerkiksi sanojen ääneen lukemisen ja nimeämisen testeillä (Segalowitz 2010: 75). Tällaiset hyvin rajatut, kontekstista irrotetut testit saattavat soveltua vähemmän rajattuja puhetehtäviä paremmin nimenomaan kielitaidon taustalla vaikuttavien kognitiivisten sujuvuusaspektien mittaamiseen. Sananhakua ja -tunnistusta on tutkittu erityisesti yhteydessä lukemisen kehitykseen, mutta tulokset ovat osin yleistettävissä myös muuhun kielenkäyttöön.

Segalowitzin viitekehysten (2010) mukaan kognitiiviseen sujuvuuteen vaikuttavat perusprosessien automaattisuus ja tarkkaavaisuuteen käytettävissä olevat resurssit. Näitä kognitiiviseen sujuvuuteen vaikuttavia tekijöitä havainnollistetaan kuviossa 1. Automatisoitumista kognitiivisen sujuvuuden indikaattorina on tutkittu erityisesti sananhaun nopeuden näkökulmasta, ja sananhaun nopeuden on osoitettu selittävän kielitaidon tasoa merkittävästi niin lukemisessa (esim. Kirby ym. 2003; Di Filippo ym. 2005; Protopapas ym. 2013) kuin kirjoittamisessakin (McCutchen 2000; Olkkonen 2012). Sanojen ääneen lukemisen ja nimeämisen on näytetty selittävän kielitaidon tasoa enemmän kuin esimerkiksi äännettömästi luettujen sanojen (Denckla & Cutting 1999; Di Filippo ym. 2005). Sananhaussa esiintyvät ongelmat ennustavat myöhemmin kehittyvää lukihäiriötä jo 3,5-vuotiailla (Puolakanaho ym. 2007), ja Leveltin (1989) puheen tuottamisen mallin mukaan äidinkielen puheen tuottaminen eroaa eniten vieraan kielen puheen tuottamisesta juuri sananhaussa.

Kognitiivisen sujuvuuden näkökulmasta suorituksen pitäisi olla paitsi nopeaa myös tarkkaa, sillä kielenoppimisen alkuvaiheessa nopeuden ja tarkkuuden oletetaan kilpailevan samoista resursseista (Grabe 2009: 292; vrt. alaluku 3, jossa sujuvuus ja tarkkuus nähdään erillisinä ulottuvuuksina). Esimerkiksi Walczyk ym. (2004) ovat havainneet, että huonot äidinkielen lukijat ovat yhtä

tarkkoja, mutta hitaampia kuin hyvät lukijat. Syyksi tähän esitetään nimenomaan lukemisen alemman tason prosessien tehottomuutta. Tämä vaikuttaisi erottavan myös toisen kielen lukijoita äidinkielistä lukijoista (Shaw & McMillion 2008). Muutamat tutkijat ovat myös löytäneet eroja erityyppisten kielenkäyttäjien välillä, joista toiset suosivat nopeutta ja toiset tarkkuutta (Salmi 2008; Jeon 2012), mutta tulokset ovat alustavia.

Siitä, mitä tarkkuudella sananhaussa varsinaisesti tarkoitetaan, ei ole kattavaa teoriaa. Useimmiten sananhaun tarkkuutta on tarkasteltu väärin luetujen sanojen määrällä (esim. Gholamain & Geva 1999; Salmi 2008; Olkkonen ym. tekeillä). Virheiden määrän sijaan informatiivisempaa voi kuitenkin olla eri virhetyyppien tarkasteleminen, sillä virhetyypit saattavat kertoa esimerkiksi suunnitteluvaiheen ongelmista enemmän kuin kielitaidon puutteesta (Ehri 1991; Engelhardt ym. 2010; Protopapas ym. 2013). Epäsujuva lukija voi esimerkiksi hypätä sanojen yli joko vahingossa tai tahallaan nopeuttaakseen lukuprosessia (Marian ym. 2013). Kielenkäyttäjät saattavat myös priorisoida nopeutta suhteessa tarkkuuteen arvailemalla sanoja; erityisesti harvinaisia sanoja korvataan helposti yleisemmällä sanoilla, jotka muodoltaan ja pituudeltaan muistuttavat alkuperäistä (Balota ym. 2006). Sama ilmiö on havaittu myös kuuntelemisessa, eli yleisemmät sanat tunnistetaan nopeammin ja väärät vastaukset haetaan yleisempien sanojen joukosta (Ellis 2002). Näin sananhaun hitauden lisäksi myös sananhaun tarkkuuden ongelmat vaikuttavat kielenkäytön sujuvuuteen. Kolmantena ulottuvuutena automatisoitumisessa sananhaun nopeuden ja tarkkuuden ohella Segalowitz (2010) mainitsee suorituksen pysyvyyden, joka kertoo pysyvämmistä muutoksista prosessoinnissa. Sitä on mitattu pääosin reaktioaikojen variaatiokertoimella (*coefficient of variation*; ks. esim. Akamatsu 2008), mutta mittarin luotettavuudesta ei ole kuitenkaan yksimielisyyttä (ks. Hulstijn ym. 2009).

Sujuvuuteen vaikuttaa automaattisuuden lisäksi tarkkaavaisuuden ohjaus. Sujuva kielenkäyttö vaatii resursseja estää epäolennaiset ärsykkeet (inhibitio), siirtää huomiota muuttuvan kielenkäyttölanteen vaatimusten mukaan (vaihtaminen, *switching*) sekä tarkkailla omaa suoritusta ja vastaanottajan ymmärrystä (monitorointi, Miyake ym. 2000; ks. kuvio 1). Oman puheen monitorointi ja korjaukset katkaisevat puheen eli periaatteessa tuottavat epäsujuvuutta, mutta toisaalta varmistavat oikean ja asianmukaisen tuotoksen (ks. Levelt 1989). Ne myös vaativat sekä riittävää kielitaitoa että kognitiivista kapasiteettia, sillä aloittelevan kielenoppijan resurssit kuluvat ääntämis- ja sananhakutasolla (Kormos 2000; Kahng 2014). Muutamissa tutkimuksissa on esitetty, että itsekorjauksia esiintyy enemmän edistyneempien kielenoppijoiden tuotoksissa (Lennon 1990; Kaukonen & Lanu 2005; Segalowitz 2010) tai paremmin hallitulla kielellä (Olkkonen 2017a). Myös Freedin (2000) tutkimuksessa ulkomaila opiskelleiden kielenoppijoiden havaittiin korjaavan puhettaan jonkin verran enemmän kuin luokkahuoneessa vierasta kieltä opiskelleiden, vaikkakaan ryh-

mien ero ei ollut tilastollisesti merkitsevä. Lisäksi inhibitiota ja tarkkaavaisuuden siirtämistä on tutkittu suhteessa sujuvuuteen esimerkiksi Stroopin testillä, jossa koehenkilön tulee nimetä värisanan fontin väri välittämättä sanasta. Reaktioajat ovat pidemmät silloin, kun väri ja sana ovat ristiriidassa keskenään (katsaus aiheeseen MacLeod 1991), ja myös äidinkielellä reaktioaika on usein pidempi kuin vieraalla kielellä (Marian ym. 2013). Tämän on selitetty johtuvan siitä, että äidinkielellä automaattisesti toimivan sanantunnistuksen torjuminen vaatii enemmän resursseja kuin vähemmän automaattisella vieraalla kielellä. Näin ollen tarkkaavaisuuden ohjaukseen liittyvät, epäsujuvuuksina näyttäytyvät ilmiöt voivat myös kertoa prosessoinnin automaattisuudesta kielitaidon puutteen sijaan, mikä voi selittää erilaisiin epäsujuvuusilmiöihin liittyviä ristiriitaisia tuloksia.

### 3 Toisen kielen sujuvuus puhetuotoksen näkökulmasta

Tässä alaluvussa käsittelemme toisen kielen sujuvuutta puhetuotoksen näkökulmasta. Sujuvuuden määrittelyn lisäksi tarkastelemme puhetuotoksen sujuvuuden mittareita sekä keskeisiä viimeaikaisia puhetuotoksen sujuvuutta analysoineita tutkimuksia. Puhetuotoksen analyysiin keskittyneissä sujuvuustutkimuksissa sujuvuus ymmärretään yleisesti ottaen nopeana, vaivattomana ja luonnollisena puheena (Chambers 1997: 535), jossa tauot ja muut epäroinnit ovat vähäisiä (Lennon 1990: 390). Kuten johdannossa totesimme, puhetuotoksen sujuvuuden ajatellaan heijastavan sen taustalla vaikuttavia kognitiivisia prosesseja. Kuitenkin erotuksena kognitiiviseen näkökulmaan, jossa sujuvuus koostuu tarkkuudesta ja nopeudesta, toisen kielen oppimisen puhetuotokseen keskittyneissä sujuvuustutkimuksissa on usein sovellettu CAF-viitekehystä, joka hahmottaa sujuvuuden (*fluency*) yhdeksi erilliseksi kielitaidon osa-alueeksi tarkkuuden (*accuracy*) ja kompleksisuuden (*complexity*) ohella (Housen & Kuiken 2009; Housen ym. 2012).

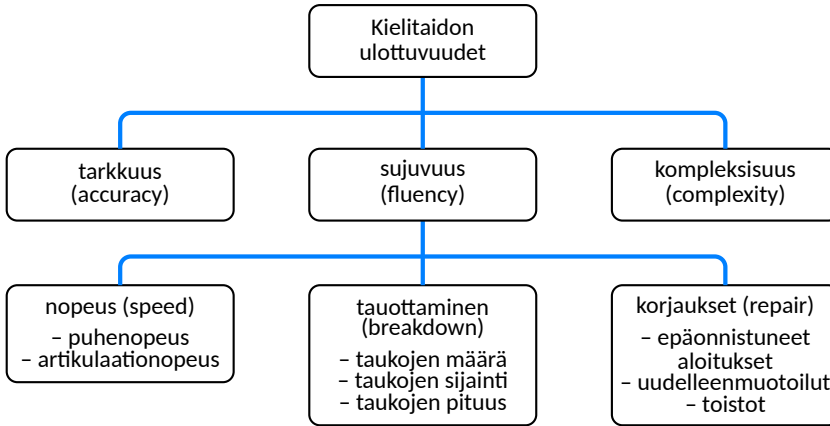
Joitakin kielitaidon osa-alueita voidaan painottaa tuotoksessa toisia enemmän, ja niiden keskinäisistä suhteista on esitetty kaksi hypoteesia. Skehanin (2009) vaihtoehtojen hypoteesin (*trade-off hypothesis*) mukaan osa-alueet ”kilpailevat” rajallisista resursseista, minkä seurauksena esimerkiksi keskittyminen tarkkuuteen saattaa vaikuttaa negatiivisesti sujuvuuteen (vrt. alaluku 2). Toisaalta Robinson (2001) on kognitiohypoteesissaan (*cognition hypothesis*) esittänyt, etteivät kaikki osa-alueet kilpaile keskenään samoista resursseista ja näin kognitiivisesti haastavat tehtävät voivat johtaa sekä tarkkaan että kompleksiseen tuotokseen. Tutkimustulokset eivät ole yksiselitteisesti vahvistaneet kumpaakaan hypoteesia, mikä saattaa johtua ainakin osittain siitä, että tutkimuksissa on tarkasteltu erilaisia tehtävätyyppejä. (Ks. myös Pietilä &

Lintunen 2014: 23.) Kielitaidon eri osa-alueiden sisällyttäminen yhteen tutkimukseen mahdollistaa niiden välisten yhteyksien ja interaktion tarkastelun, mutta vaarana on, että yksittäisiä osa-alueita käsitellään liian suppeasti (esim. Larsen-Freeman 2009: 582–583). Sujuvuuden näkökulmasta esimerkiksi pelkän puhenopeuden mittaaminen antaa melko kapean kuvan puhetuotoksen sujuvuudesta.

Toisinaan sujuvuudella saatetaan viitata yhden kielitaidon osa-alueen sijasta kielitaitoon laajemmin. Siinä missä tieteellisissä yhteyksissä sovelletaan yleensä sujuvuuden rajattua merkitystä (*narrow sense*) eli käsitystä sujuvuudesta yhtenä suullisen kielitaidon osa-alueena, erityisesti arkisessa kielenkäytössä käytetään sujuvuuden laajaa merkitystä (*broad sense*), jolloin sujuvuus rinnastetaan kielitaidon kokonaisvaltaiseen hallintaan, kuten virkkeessä ”Olen sujuva ranskassa, mutta saksasta osaan vain alkeet.” (Lennon 1990). Näiden kahden merkityksen välistä yhteyttä on tarkasteltu empiirisesti esimerkiksi tutkimuksissa, joissa on kartoitettu sujuvuuden roolia suullisen kielitaidon osatekijänä suhteessa muihin kielitaidon osa-alueisiin ja havaittu, että puhetuotoksen sujuvuus on merkittävä kielitaidon tasoja erotteleva tekijä (esim. Iwashita ym. 2008). Rajatun määritelmän mukainen sujuvuus on siis keskeinen osa laajempaa suullista kielitaitoa.

Lisäksi objektiivisten sujuvuuden mittareiden (puhetuotoksen sujuvuuden) yhteyttä subjektiivisiin kuulijoiden arvioihin sujuvuudesta (havaittuun sujuvuuteen) tarkastelleet tutkimukset ovat osoittaneet, että mittareiden ja arvioiden väliset korrelaatiot vaihtelevat yleensä kohtalaisista voimakkaisiin (esim. Bosker ym. 2013; ks. myös De Jong 2016). Korrelaatioihin näyttäisi vaikuttavan arvioijien ohjeistus: erityisesti tarkkojen kuvaimien puuttuessa kuulijat huomioivat usein myös ”kapeaan” sujuvuusmääritelmään kuulumattomia piirteitä, kuten tarkkuutta, sanaston monipuolisuutta, kieliopillista kompleksisuutta ja ääntämistä (Lennon 1990; Riggenbach 1991; Freed 2000; Kormos & Dénes 2004). Kun arvioijia puolestaan on pyydetty keskittymään temporaaliseen sujuvuuteen, arviot ovat olleet yhdenmukaisempia ja arvioijaryhmien väliset erot vähäisiä (esimerkiksi natiivien ja ei-natiivien välillä, Rossiter 2009). Arvioijien mahdollisten erilaisten sujuvuuskäsitysten vuoksi onkin tärkeää pyrkiä hyödyntämään empiiristä tutkimustietoa kielitaidon arvioinnissa. Arvioinnin luotettavuutta voidaan lisätä esimerkiksi sisällyttämällä sujuvuuden kuvaimiin konkreettisia puheen piirteitä, jotka perustuvat objektiivisiin kielitaidon mittareihin.

Yleisesti ottaen puhetuotoksia tarkastelleet sujuvuustutkimukset ovat osoittaneet, että sujuvuuden tarkasteluun käytettyjen mittarien luotettavuudessa on eroja. Puhetuotoksen sujuvuutta tarkasteltaessa analysoidaan puheen vaivattomuuteen liittyviä piirteitä useimmiten monologipuheesta (mutta ks. dialogipuheen sujuvuudesta Sato 2014; Tavakoli 2016; Peltonen 2017), ja sujuvuusmittarien valinnassa on erityisesti viimeaikaisissa sujuvuustutkimuk-



KUVIO 2. Puheen sujuvuuden ulottuvuudet suhteessa CAF-viitekehukseen.

sisä hyödynnetty Skehanin (2003, 2009, 2014) viitekehystä. Viitekehyksessä sujuvuuden nähdään koostuvan kolmesta ulottuvuudesta: puhenopeudesta (*speed fluency*), tauottamisesta (*breakdown fluency*<sup>3</sup>) ja korjauksista (*repair fluency*). Näiden ulottuvuuksien suhdetta CAF-viitekehukseen sekä esimerkkejä usein käytetyistä mittareista havainnollistetaan kuviossa 2.

Vaikka useimmat viimeaikaiset sujuvuustutkimukset ovat tarkastelleet kaikkia kolmea ulottuvuutta, on käytetyissä mittareissa sekä niiden operatiivisoinnissa jonkin verran vaihtelua. Esimerkiksi nopeuden tarkasteluun on käytetty sekä puhenopeutta (*speech rate*) että artikulaationopeutta (*articulation rate*). Artikulaationopeus on kuitenkin yksiselitteisempi nopeuden mittari, koska se suhteuttaa tuotettujen tavujen (tai sanojen) määränpuhe-aikaan (*speaking time* eli puheen kokonaiskestoon ilman taukoja), kun taas puhenopeudessa tavujen määrä suhteutetaan tuotoksen kokonaiskestoon, jolloin mittariin vaikuttavat tavujen tuottamisen nopeuden lisäksi myös tauot (De Jong 2016: 212). Toisinaan on sovellettu myös puhenopeuden mittaria, josta on karsittu pois erilaisia epäsujuvuuksia (*pruned speech rate*, esim. Lennon 1990): epäsujuvuuksien monimerkityksisyyden sekä mahdollisten sujuvuutta lisäävien funktioiden vuoksi tämä käytäntö on kuitenkin melko kyseenalainen.

Eri sujuvuusmittareita on sovellettu aiemmin käsiteltyjen puhetuotoksen ja havainnoitua sujuvuutta vertailevien tutkimusten lisäksi myös esimerkiksi sujuvuuden kehittymisen tarkasteluun. Tällaisia ovat olleet erityisesti ul-

<sup>3</sup> Termille *breakdown fluency* ei ole vakiintunutta suomennosta. Käsitteellä viitataan yleisesti erilaisiin taukoihin liittyviin mittareihin, joiden asema (epä)sujuvuutta lisäävinä piirteinä ei ole täysin yksiselitteinen. Erityisesti ns. täytetyt tauot (*filled pauses*, kuten suomen öö tai englannin *uh, um*) voidaan nähdä myös sujuvuutta lisäävinä mekanismeina, joita ei tulisi rinnastaa ”hiljaisiin” taukoihin (esim. Dörnyei & Kormos 1998; Kahng 2014; vastaavasti äidinkielisestä puheesta ks. esim. O’Connell & Kowal 2004).

komailta oleskelun vaikutuksiin keskittyneet, ns. *study abroad* -tutkimukset. Vaikka yleisesti ottaen ulkomailla ololla on todettu olevan positiivisia vaikutuksia sujuvuuteen (Lennon 1990; Towell ym. 1996; Freed 2000), kehittymiseen vaikuttavat myös monet taustatekijät, kuten kohdekielen käytön määrä ulkomailla oleskelun aikana (esim. Mora & Valls-Ferrer 2012). Pitkittäistutkimusten lisäksi eri taitotasojen edustavien oppijaryhmien sujuvuuden eroja on tarkasteltu poikittaistutkimuksissa (esim. Kormos & Dénes 2004), toisinaan sisällyttäen tutkimusasetelmaan oppijaryhmän tai -ryhmien tuotosten vertailun ”optimaalista” sujuvuutta edustavien natiivipuhujien kontrolliryhmään (esim. Hilton 2008; Götz 2013; Kahng 2014; Peltonen & Lintunen 2016). Tutkimukset ovat osoittaneet, että välttämättä edes yliopisto-opiskelijat eivät täysin yllä natiivien tasolle kaikilla sujuvuuden osa-alueilla. Esimerkiksi Götzin (2013) tutkimuksessa oppijat käyttivät natiiveja vähemmän erilaisia täytesanoja, kun taas täytettyjä taukoja käytettiin natiiveja enemmän. Natiivien rooliin toisen kielen oppimisen sujuvuustutkimuksissa palataan alaluvussa 5.

Vaikka sujuvuusmittarien luotettavuus on harvoin ollut eksplisiittisenä tutkimuskohteena, erityisesti puhe- ja artikulaationopeus sekä taukojen sijainti ovat osoittautuneet puhetuotosta tarkastelleiden tutkimusten perusteella luotettavimmiksi sujuvuuden indikaattoreiksi (mittareista tarkemmin ks. De Jong 2016). Ne ovat erotelleet eritasoisia oppijaryhmiä ja kielenoppijoiden suorituksia natiivipuheesta (esim. Kormos & Dénes 2004; Tavakoli 2011; Götz 2013; Kahng 2014) sekä osoittaneet sujuvuuden kehitystä (esim. Towell ym. 1996). Täytettyjen taukojen ja korjausten asema (epä)sujuvuuden mittareina ei kuitenkaan ole yhtä yksiselitteinen. Kyseisten mittarien ja epäsujuvuuden välille ei ole aina löydetty selkeää yhteyttä, ja niiden käyttöön liittyy usein runsaasti yksilöllisistä eroista kertovaa ryhmien sisäistä vaihtelua (Lennon 1990; Riggenbach 1991; Cucchiariini ym. 2002; Kormos & Dénes 2004). Mittarit eivät siis aina ole erotelleet eritasoisia ryhmiä toisistaan, ja osittain erojen vähäisyyttä saattaa selittää sujuvuustutkimusten määrällinen painotus: tarkastelu on keskittynyt piirteiden määriin, ei niiden funktioihin tai käyttöyhteyksiin (mutta määrällistä tarkastelua laadullisella analyysillä täydentävistä tutkimuksista ks. esim. Ejzenberg 2000; Hilton 2008; myös alaluku 4). Mahdollisten eri funktioiden vuoksi myös joissakin tutkimuksissa (Cucchiariini ym. 2002; Kormos & Dénes 2004; Iwashita ym. 2008) käytetty kaikki korjaukset yhdistävä, globaali epäsujuvuuden mittari ei välttämättä ole tarpeeksi hienosyinen sujuvuuden ja korjausten yhteyksien tarkasteluun. Kuten aiemmin alaluvussa 2 todettiin, kognitiivisen sujuvuuden näkökulmasta korjaukset on mahdollista tulkita merkeiksi edistyneiden oppijoiden kyvystä monitoroida puhetta sen sijaan, että niitä käsiteltäisiin epäsujuvuuksina. Seuraavassa luvussa perehdymme yhteen korjauksiin liittyvään mittariin, toistoihin, kognitiivisen ja puhetuoksen sujuvuuden näkökulmista.

## 4 Toistot sujuvuuden mittarina kognitiivisen ja puhetuotoksen sujuvuuden näkökulmista

Toistoja on tarkasteltu sekä kognitiiviseen että puhetuotoksen sujuvuuteen liittyvissä tutkimuksissa (epä)sujuvuuden mittarina, mutta ei ole kuitenkaan täysin selvää, mistä ne kertovat. Perehtymällä toistoihin eri näkökulmista tuomme esiin sekä niiden erilaisia painotuksia että yhtäläisyyksiä. Kognitiivisen sujuvuuden näkökulmaa toistoihin käsitellään suomalaisten oppijoiden sananhakua ja -tunnistusta tarkastelleen tutkimuksen valossa (Olkkonen 2017a). Puhetuotoksen näkökulmaa havainnollistamme puolestaan käsittelemällä suomenkielisten ja ruotsinkielisten oppijoiden englanninkielisiä monologipuhetuotoksia vertaillen tuoreen tutkimuksen (Peltonen & Lintunen 2016) tuloksia. Näkökulmat yhdistämällä pohdimme erilaisia selityksiä toistoihin liittyville tuloksille.

Toistoja (sanoja tai pidempiä yksiköitä, jotka toistetaan puheessa ilman muokkauksia, ks. Foster & Skehan 1999) on tarkasteltu sekä kognitiivisen että puhetuotoksen sujuvuuden näkökulmasta usein merkkeinä epäsujuvuudesta muiden korjausten tavoin: puhe on sitä epäsujuvampaa, mitä enemmän toistoja puheessa on. Tutkimusten tulokset eivät kuitenkaan ole yksiselitteisesti tukeneet tätä näkemystä, sillä toistojen ja epäsujuvuuden välille ei ole pystytty osoittamaan selkeää yhteyttä (esim. Cucchiarini ym. 2002; Kormos & Dénes 2004). Vähäiset erot oppijaryhmien välillä toistojen suhteen sekä toisaalta merkittävä ryhmien sisäinen vaihtelu saattaa osaltaan selittyä oppijoiden välisillä yksilöllisillä eroilla ja osittain toistojen vähäisellä määrällä aineistoissa (esim. Freed 2000; Cucchiarini ym. 2002; Götz 2013; Kahng 2014). Siinä missä kaikki puhetuotokset sisältävät aina taukoja ja puhetta, toistoja ei välttämättä esiinny (Grosjean 1980). Kognitiivisen sujuvuuden näkökulmasta toistojen käyttö voi kertoa mm. inhibition toimimattomuudesta eli ongelmista väärän tuotoksen estämisessä. Inhibition ongelmat ovat yksi ADHD:n oireista (Nigg 2001), joten Engelhardt ym. (2010) tutkivat toistojen määrää ADHD-ryhmässä sekä verrokeilla. Toistojen määrä lauseentuottamistehtävässä ei kuitenkaan hypoteesin vastaisesti erotellut ADHD-ryhmää kontrolliryhmästä. Kyseisessä tutkimuksessa tutkijat päätyivät luokittelemaan toistot suunnittelun ongelmista kertoviksi mittareiksi ja toistojen jakauma oli samanlainen kuin täytettyjen taukojen.

Suomenkielisten oppijoiden aineistossa toistoja kognitiivisesta näkökulmasta on tutkinut Olkkonen (2017a). Tutkimuksessa tarkasteltiin epäsujuvuuksia (yhteensä 8 mittaria, joista yksi toistot) sananhausta ja -tunnistuksessa suomenkielisillä englanninoppijoilla (n = 563, 4., 8. ja lukion 2. luokalta). Sanantunnistusta mitattiin sanalistan lukemisella minuutin ajan sekä suomeksi että englanniksi (Häyrinen ym. 1999). Listan alussa sanat olivat lyhyitä ja tuttuja, mutta muuttuivat vähitellen pidemmiksi ja harvinaisemmiksi. Sananhakua mitattiin *Rapid Alternating Stimulus (RAS)* -tehtävällä, joka mittaa hy-

vin tutun sanaston nimeämisnopeutta (numeroita, kirjaimia ja värejä; Ahonen ym. 2003; ks. myös Wolf 1986). Tulokset osoittivat, että toistojen määrä oli aineistossa hyvin vähäinen ja lisäksi ryhmien sisäinen vaihtelu erittäin suurta. Ainoastaan suomenkielisessä RAS-tehtävässä 4.-luokkalaiset käyttivät toistoja  $\chi^2$ -testin perusteella merkittävästi enemmän (keskiarvo 0,39 toistoa per oppilas) kuin toiset ryhmät (ryhmien keskiarvot 0,15 toistoa per oppilas). Muissa tehtävissä toistojen määrä väheni hiukan luokka-asteelta toiselle, mutta ero ei ollut tilastollisesti merkitsevä. Huomioitavaa on erityisesti se, että myöskään äidinkielen suoritus ei ollut täydellisen sujuvaa millään luokka-asteella. Tämä johtui osin tehtävien aikapaineesta, mutta on tärkeä huomio mietittäessä ylipäätään toistojen ja muiden tarkkuusilmiöiden käyttöä kielitaidon mittareina. Mielenkiintoista oli se, että eniten toistoja löytyi äidinkielellä tehdystä, sanastoltaan jo automatisoituneesta RAS-tehtävästä. Näin ollen toistojen käyttö ei kyseisessä tutkimuksessa ollut suoraan yhteydessä kielitaidon puutteeseen vaan voi olla tulkittavissa esimerkiksi strategiseksi keinoksi haastavassa tilanteessa (ks. Engelhardt ym. 2010; Bosker ym. 2013).

Puhetuotoksen sujuvuuteen keskittyneessä Peltosen ja Lintusen (2016) tutkimuksessa puolestaan tarkasteltiin äidinkielen ja koulutusasteen vaikutusta vieraskielisen puheen sujuvuuteen 14 mittarilla, joista yksi oli toistot. Tutkimuksessa analysoitiin 40 englanninkielistä puhenäytettä suomea ja ruotsia äidinkielenään puhuvilta englannin opiskelijoilta kahdelta koulutusasteelta, lukiosta ja yliopistosta. Lisäksi tarkasteltiin puhenäytteitä 10 englannin nativi-puhujalta. Puhenäytteet olivat suullisia monologeja, joissa osallistujat kertoivat tarinan sarjakuvaan perustuen (ks. myös esim. Goldman-Eisler 1968; Lennon 1990; Kormos & Dénes 2004). Tutkimuksen tulokset osoittivat, että toistoja esiintyi eri ryhmien tuotoksissa yleisesti ottaen melko vähän. Ruotsia äidinkielenään puhuvat yliopisto-opiskelijat (keskiarvo 2,68 toistoa minuutissa) ja natiivit (keskiarvo 1,66 toistoa minuutissa) käyttivät toistoja jonkin verran muita ryhmiä enemmän (ruotsinkieliset lukiolaiset 0,80 toistoa minuutissa, suomenkieliset lukiolaiset 0,55 toistoa minuutissa, suomenkieliset yliopisto-opiskelijat 0,37 toistoa minuutissa), mutta ryhmien väliset erot toistojen määrässä eivät olleet Kruskal-Wallis-testin mukaan tilastollisesti merkitseviä.

Toistojen vähäisten määrien lisäksi myös ryhmien sisäinen vaihtelu saattoi vaikuttaa siihen, että merkitseviä eroja ei löytynyt Peltosen ja Lintusen (2016) tutkimuksessa: useimmissa ryhmissä keskihajonnat olivat merkittäviä, mikä viittaa yksilöllisiin eroihin toistojen käytössä. Tendenssit olivat samankaltaisia myös muiden korjausten kohdalla sekä täytettyjen taukojen määrässä. Vaikka toistojen määrät eivät selkeästi erotelleet tutkimusryhmiä, tutkimuksessa määrällistä analyysiä täydentänyt laadullinen analyysi paljasti kuitenkin eroja yksilöiden välillä toistojen käytössä. Laadullisessa analyysissä tarkasteltiin lähemmin kolmen vähän korjauksia sekä kolmen paljon korjauksia käyttäneen koehenkilön puhetuotoksia, ja analyysi keskittyi erityisesti korjausten



funktioihin, käyttöyhteyksiin sekä koehenkilöiden yleisiin sujuvuusprofileihin. Toistojen osalta analyysi osoitti, että niitä voitiin käyttää strategisena keinona sujuvuuden lisäämiseen minimoimalla hiljaisuuksia ja toisaalta ylläpitämällä puheen jatkuvuutta. Toistoilla oli siis prosessointia helpottava funktio, eivätkä ne yksiselitteisesti liittyneet epäsujuvuuteen.

Lähestymistapojen ja aineistojen eroista huolimatta kognitiiviseen (Olkkonen 2017a) ja puhetuotoksen (Peltonen & Lintunen 2016) sujuvuuteen keskittyneissä tutkimuksissa oli toistojen suhteen havaittavissa samanlaisia tendenssejä: toistojen määrät eivät pääosin erotelleet eritasoisia ryhmiä toisistaan ja ryhmien sisäinen vaihtelu oli suurta. Tulokset vahvistavat myös aiempien tutkimusten tuloksia toistojen käytön yksilöllisistä eroista ja vähäisistä esiintymistä (esim. Cucchiariini ym. 2002; Götz 2013; Kahng 2014) sekä käsitystä siitä, että taukojen pituuden lailla (De Jong ym. 2015) myöskään toistot eivät näyttäisi liittyvän suoraan kielitaidon tasoon vaan enemmänkin yksilölliseen puhetyyliin. Tämä voisi osaltaan myös selittää toistojen käyttöön liittyvää vaihtelua.

Epäsujuvuuden ja erityisesti toistojen määrän välisen yhteyden osoittaminen saattaa olla haasteellista osittain myös siksi, että toistoilla voidaan nähdä olevan erilaisia funktioita. Esimerkiksi Dörnyein ja Kormosin (1998) viitekehyksessä, jossa käsitellään toisen kielen puheen ongelmia ja niiden ratkaisemista psykolingvistikseen näkökulmasta, toistot hahmotetaan epäsujuvuukseen sijaan ongelmanratkaisukeinoiksi, jotka helpottavat vieraskielisen puheen tuottamiseen liittyvää prosessointitaakkaa muun muassa täytettyjen taukojen, äänneiden venyttämisen (*drawls*) ja erilaisten täytesanojen (*fillers*, kuten *like, well, you know*) ohella. Tämän näkökulman mukaan toistot voivat toimia jopa sujuvuutta lisäävinä viivytysmekanismeina (*stalling mechanisms*), joita taitavat oppijat hyödyntävät välttääkseen pitkiä taukoja ja ylläpitääkseen puhetta suunnittelun aikana. (Dörnyei & Kormos 1998). Samaan tapaan Götz (2013) käsittelee toistoja puheen prosessointia helpottavina strategioina, jotka auttavat lisäämään puheen luonnollisuutta ja sujuvuutta (*fluency enhancement strategies*). Toistot voivatkin muiden vastaavien mekanismien ohella auttaa ylläpitämään puheen jatkuvuutta (esim. Hilton 2008) ja luoda vaikutelmaa sujuvan kuuloisesta puheesta (Ejzenberg 2000). Myös Peltosen ja Lintusen (2016) tulokset tukivat näkemystä toistojen strategisesta käytöstä. Toistoja voidaan siis tarkastella tulkitsemalla ne tämänkaltaisesta tietoisesta puheen organisoinnista eikä pelkästä taidon puutteesta johtuviksi.

Strategisen käytön lisäksi tehtävätyypillä näyttäisi olevan vaikutusta toistojen käyttöön. Esimerkiksi Olkkosen (2017a) tuloksissa sanalistatehtävässä toistoa esiintyi hieman enemmän vieraalla kielellä ja RAS-tehtävässä äidinkielellä. Automatisoitunutta sanastoa mitanneessa RAS-tehtävässä toistoja saatetaan käyttää taukojen täyttämiseen ja seuraavan tuotoksen suunnitteluun (vrt. Marian ym. 2013), ja sanalistatehtävässä taas mahdollisesti monitoroidaan vieraskielisten sanojen lukemisen oikeellisuutta. Tehtävätyypin vaikutusta toistoihin

olisi tärkeää tutkia enemmän myös esimerkiksi dialogipuheessa. Toistot saattavat olla dialogipuheessa vielä monologipuhetta keskeisempiä niiden vuorovaikutukseen liittyvien funktioiden, kuten vuoron säilyttämisen ja koheesion luomisen vuoksi (Tannen 1989: 50–52). Esimerkiksi Peltosen (2017) tutkimuksessa edistyneemmät, lukion 2. vuoden opiskelijat hyödynsivät englanninkielisissä dialogeissaan 9.-luokkalaisia tilastollisesti merkitsevästi enemmän toistoja (sekä oman puheen, *self-repetition*, että parin puheen, *other-repetition*; ks. myös Peltonen 2018 tässä teoksessa). On kuitenkin huomattava, että kuten usein monologiaineistoissa, myös Peltosen (2017) dialogiaineistossa luokkaasteiden sisäinen vaihtelu oli toistojen käytössä huomattavaa ja yleisesti ottaen toistojen käyttö oli vähäistä. Tulevaisuudessa olisikin hyödyllistä tarkastella toistoja muissa dialogiaineistoissa ja vertailla tuloksia monologiaineistoihin pohjaaviin tutkimuksiin.

## 5 Kognitiivisen ja puhetuotoksen sujuvuuden näkökulmien vertailua

Tässä luvussa vertailemme kognitiivisen ja puhetuotoksen sujuvuuden näkökulmia toistoja laajemmin. Keskitymme näkökulmien eroihin ja yhtäläisyyksiin erityisesti tarkkuuteen ja korjauksiin liittyen. Näkökulmien yhdistämiseen pohjautuen pohdimme erityisesti toistoihin liittyvien tutkimustulosten implikaatioita sujuvuuden arvioinnille sekä sitä, millaisia näkökulmia sujuvuustutkimuksessa kannattaisi tulevaisuudessa erityisesti huomioida.

Sujuvuustutkimuksen yhtenä haasteena on ollut termien eriävä ja osittain päällekkäinen käyttö eri tieteenaloilla (esim. Chambers 1997). Erityisesti terminologinen ero rajatun, tieteellisen mittaamisen pohjaksi soveltuvan sujuvuuden ja arkikielessä käytetyn, laajasti kielitaitoon viittaavan merkityksen välillä tulisi pitää selkeänä (ks. Lennon 1990 ja alaluku 3). Kuten Koponen ja Riggenbach (2000: 19–20) ovat huomauttaneet, ei ehkä ole realistista pyrkiä yhteen, kaikenkattavaan sujuvuusmäärittelmään, vaan pikemminkin kiinnittää huomiota eri ”sujuvuuksien” määrittelyyn. Tässä artikkelissa olemme osaltamme pyrkineet tähän erittelemällä kognitiivisen ja puhetuotoksen sujuvuuden käsitteitä. Eri lähestymistavat painottavat eri aspekteja ja hyödyntävät paljolti eri menetelmiä, mutta täydentävät toisiaan: esimerkiksi kognitiivisen sujuvuuden tarkka määrittely ja sen osa-alueiden ymmärtäminen voi auttaa puhetuotoksen sujuvuuden mittareiden määrittelyssä ja tiettyjen piirteiden käytön syiden ymmärtämisessä. Sujuvuutta eri perspektiiveistä hahmottavien määrittelmien käsitteleminen rinnakkain voi siis selkeyttää tutkimuskohteen luonnetta ja auttaa määrittelemään tutkimuskohdetta aiempaa täsmällisemmin.

Konkreettisenä esimerkkinä sujuvuuden määrittelmien vertailemisen hyödyistä haluamme erityisesti nostaa esiin tarkkuuden roolin eri näkökulmissa.

Kuten alaluvussa 3 todettiin, kognitiivisesta näkökulmasta sujuvuuden katsotaan kattavan sekä prosessoinnin nopeuden että tarkkuuden esimerkiksi lukemisessa. Useimmiten tarkkuuden on oletettu olevan erityisesti suomen tyyppisissä läpinäkyvän ortografian kielissä lähelle täydellistä jo hyvin nuorilla (Kirby ym. 2003; Di Filippo ym. 2005). Useat testit, kuten nimeämistehtävä RAS, perustuvatkin siihen, että tarkkuus pyritään vakioimaan, ja näin variaatio tuloksissa syntyy lähinnä nopeudesta. Tämä johtuu osittain siitä, että näitä testejä on käytetty pääosin äidinkielen diagnosoinnissa. Toisen kielen kohdalla vähäisempi automatisoituminen saattaa tuottaa kuitenkin eroja myös tarkkuuteen. Puhetuotoksen näkökulmasta tarkkuus on puolestaan hahmotettu perinteisesti omaksi kielitaidon ulottuvuudekseen (CAF-viitekehysten mukaisesti, ks. alaluku 3). Erottelun taustalla saattaa olla perinteinen, kielten opeutuksessa hyödynnetty jaottelu sujuvuuteen tai tarkkuuteen keskittyviin harjoituksiin (Brumfit 2000). Toisaalta CAF-viitekehystä sovellettaessa on myös esitetty, että kilpailua esiintyisi erityisesti muotoon (tarkkuus ja kompleksisuus) ja merkitykseen (sujuvuus) keskittymisen välillä (Skehan 2009). Kuten alaluvussa 2 todettiin, kiinnostavaa on se, että vaikka puheen tuotoksen katsotaan heijastavan kognitiivisia prosesseja, on tarkkuutta kuitenkin tuotosten analyysissä käsitelty useimmiten omana ulottuvuutenaan.

Näkökulmien erilainen suhtautuminen tarkkuuteen näkyy myös korjauksissa: erityisesti toisen kielen oppimisen sujuvuustutkimuksessa (CAF-viitekehystä seuraten) korjauksia on toisinaan käsitelty tarkkuuden ja toisinaan sujuvuuden indikaattoreina. Toisaalta korjaukset voidaan nähdä pyrkimyksinä tarkkuuteen kognitiivisen prosessoinnin näkökulmasta (esim. Levelt 1989; Kormos 2000), toisaalta taas puhetuotoksen näkökulmasta epäsujuvuudesta kerovina piirteinä (Skehanin 2003, 2009, 2014 viitekehys). Kognitiivisen sujuvuuden näkökulma voi auttaa selittämään sitä, miksi kielitaidon ja toistojen (tai laajemmin korjausten) määrän välille on ollut haastavaa osoittaa yksiselitteistä yhteyttä. Kuten toimme aiemmin esiin, pelkän epäsujuvuuden indikaattorien sijaan toistoilla saattaakin olla prosessointia helpottava funktio (Dörnyei & Kormos 1998), ja toisaalta toistot voivat liittyä kielitaidoltaan edistyneiden oppijoiden strategiaan kykyihin (Lennon 1990) tai taitoon monitoroida omaa puhettaan (Kormos 2000; Olkkonen 2017b).

Näkökulmien vertailu lisää myös tietoisuutta eri tutkimusperinteissä käytetyistä menetelmistä. Esimerkiksi kognitiivisen sujuvuuden tarkastelussa käytetyt kontrolloidut tutkimusasetelmat voivat auttaa selvittämään puhetuotoksen sujuvuuden ja kognitiivisen sujuvuuden välisiä yhteyksiä, joita on tutkittu vielä melko vähän (vrt. alaluku 2). Ylipäättänsä aineistot, jotka mahdollistavat sekä kognitiivisen että puhetuotoksen empiirisen tarkastelun samoilta koehenkilöiltä, ovat kansainvälisestikin harvinaisia. Toisaalta laadullista ja määrällistä analyysiä yhdistävää *mixed methods* -lähestymistapaa (esim. Hilton 2008; Peltonen & Lintunen 2016), jota on toisinaan hyödynnetty toisen kielen oppi-

misen sujuvuustutkimuksissa puhetuotosten tarkasteluun, voisi soveltaa määrälliseen tutkimukseen painottuneissa kognitiivista sujuvuutta tarkastelevissa tutkimuksissa laajemminkin (ks. Olkkonen 2017a). Eryityisesti toistoihin liittyvien tulosten valossa nähdäksemme metodologisesti mahdollisimman kattava ja monipuolinen kuva sujuvuudesta ja sujuvuuteen liittyvien piirteiden käytöstä saavutetaan laadullisen (piirteiden funktiot ja käyttöyhteydet huomioiva) ja määrälliseen analyysin yhdistelmällä.

Toistoja käsitellyt alaluku toi esiin myös tärkeitä implikaatioita kielitaidon arvioinnille. Kielitaidon arvioinnissa heijastuvat olennaisesti vallalla olevat näkemykset kielestä ja kielitaidosta (esim. Huhta & Hildén 2016: 19), joihin puolestaan vaikuttavat empiiriset tutkimustulokset. Koska toistot eivät näyttäisi yksiselitteisesti kertovan vieraan kielen taitotasosta, tulisi toistojen käyttöä arviointikriteerinä harkita. Kriteerien konkreettisuuden ja siten myös arvioiden luotettavuuden lisäämiseksi tulisi kielitaidon arviointikriteerien pohjautua empiirisiin havaintoihin. Näin ollen toistojen lisäksi myös muiden puheen sujuvuuden mittareiden luotettavuuteen liittyvää tutkimustietoa tulisi hyödyntää kriteerien suunnittelussa vielä nykyistä laajemmin (ks. myös esim. Van Moere 2012).

Sujuvuuden luotettavan arvioinnin kannalta on tärkeää myös selvittää äidinkielen puhetyylin vaikutusta toisen kielen sujuvuuteen (esim. Derwing ym. 2009). Viimeaikaiset äidinkielen ja toisen kielen sujuvuuden välisiä yhteyksiä kartoittaneet tutkimukset ovat osoittaneet, että äidinkielen sujuvuus selittää ainakin jonkin verran toisen kielen sujuvuuden vaihtelusta useimmilla sujuvuusmittareilla, vaikka selitysvoima vaihtelee mittarien välillä (De Jong ym. 2015; Huensch & Tracy-Ventura 2017). Vaikka aiheeseen liittyviä tutkimuksia on toistaiseksi verrattain vähän, jo tähänastiset tulokset osoittavat, että puhetyylin vaikutus ja yleisesti ottaen puhujien yksilölliset erot tulisi huomioida nykyistä laajemmin sujuvuustutkimuksessa. Tutkimusta tarvitaan lisää erityisesti eri kieliyhdistelmillä sekä eritasoisilla oppijoilla, jotta saadaan tarkempaa tietoa kielten välisten erojen sekä kielitaidon tason vaikutuksesta äidinkielen ja toisen kielen sujuvuuden yhteyksien voimakkuuteen (ks. myös Huensch & Tracy-Ventura 2017).

Oppijoiden äidinkielen puhetyylin huomioinnin lisäksi toisen kielen sujuvuuden arvioinnissa ja tutkimuksessa tulisi myös kiinnittää laajemmin huomiota siihen, mihin oppijoiden sujuvuutta verrataan ja mitä pidetään tavoitteena. Usein sujuvuustutkimusten taustalla on ollut implisiittinen oletus ”täydellisestä” natiivipuheesta, joka ei sisällä epäsujuvuuksia, ja oppijoiden sujuvuutta on tarkasteltu suhteessa tähän oletukseen. Todellisuudessa kuitenkin myös äidinkielen puheseen liittyy taukoja, korjauksia ja toistoja (esim. Fillmore 1979; Lennon 1990; Peltonen & Lintunen 2016; Olkkonen 2017a). Tämän vuoksi onkin syytä tarkastella kriittisesti myös idealisoituja käsityksiä äidinkielisistä kielenkäyttäjistä (Lennon 1990; Hulstijn 2011). Sujuvuuden näkökulmasta tähän

ovat vastanneet tutkimukset, joihin on sisällytetty natiivien kontrolliryhmä kienoppijoiden lisäksi (ks. myös alaluku 3). Tällainen asetelma mahdollistaa äidinkielen sujuvuuden empiirisen tarkastelun, mukaan lukien erot myös äidinkielisten puhujien sujuvuudessa (ks. myös Fillmore 1979), ja tarjoaa implisiittistä natiivinormia realistisemman vertailukohteen oppijoiden puheelle.

Sujuvuuden mittaamiseen ja arviointiin vaikuttaa keskeisesti myös tehtävätyyppi. Suurin osa erityisesti puhetuotoksen sujuvuuden tutkimuksesta on keskittynyt melko kontrolloiduissa oloissa tuotettuihin monologeihin. Edellisessä luvussa toimme esiin, että toistoja ja ylipäätään puhetuotoksen sujuvuuden mittareita on tutkittu vieraskielisestä dialogipuheesta toistaiseksi vain vähän (mutta ks. Riggenbach 1991; Sato 2014; Tavakoli 2016; Peltonen 2017). Sujuvuuden käsitteen soveltaminen vuorovaikutustilanteisiin on melko haastavaa ottaen huomioon, että sujuvuuskäsityksellä on kognitiivinen, yksilön taitoihin perustuva tausta (mutta ks. Pickering & Garrod 2004 prosessoinnista vuorovaikutustilanteessa). Keskustelujen analyysissä tulisi kuitenkin yksilöiden suorituksia ja puhujien yksittäisiä vuoroja laajemmin huomioida myös kuinka puhujat ylläpitävät sujuvuutta yhdessä (*confluence*, McCarthy 2010). Näin ollen jo lähtökohtaisesti moniulotteiseen sujuvuuden käsitteeseen liittyy keskustelutilanteissa yksilöllisen sujuvuuden lisäksi sosiaalinen, puhujien väliseen interaktioon liittyvä ulottuvuus (ks. esim. Lauranto 2005; Peltonen tässä erikoisnumerossa). Myös keskustelutilanteisiin perustuvassa arvioinnissa on siis yksilön suorituksen lisäksi tärkeää huomioida myös puhujien yhteistyö (esim. Sato 2014).

## 6 Lopuksi

Artikkelimme tavoitteena oli tarkastella toisen kielen sujuvuutta kognitiivisen sujuvuuden ja puhetuotoksen sujuvuuden näkökulmista. Näkökulmat liittyvät pääosin erillisiin tutkimusperinteisiin, ja niiden eksplisiittinen vertailu on ollut sujuvuustutkimuksessa harvinaista. Artikkelimme kuitenkin osoitti, että vaikka näkökulmat painottavat sujuvuuden eri puolia, on niillä myös useita yhtymäkohtia. Eri näkökulmien vertailu voi auttaa erityisesti selittämään, miksi joidenkin (epä)sujuvuuteen liitettyjen piirteiden, kuten korjausten, ja epäsujuvuuden välille ei ole löydetty selkeää yhteyttä. Toistoihin keskittynyt alaluku havainnollisti, kuinka erilaisista lähtökohdista toteutetut tutkimukset voivat päätyä samankaltaisiin päätelmiin: toistot eivät kerro yksiselitteisesti kielitaidon tasosta ja niiden käyttöön liittyy paljon yksilöllistä vaihtelua. Lisäksi pyrimme osoittamaan, että näkökulmien yhdistämisestä on hyötyä erityisesti sujuvuuden aiempaa täsmällisemmässä määrittelyssä, sujuvuustutkimuksen menetelmien monipuolistamisessa sekä sujuvuuden arvioinnin kehittämisessä. Tarkastelumme pohjalta kannustamme sujuvuutta eri näkökulmista lähes-

tyviä tutkijoita yhä laajempaan tieteenalaraajat ylittävään tutkimusyhteistyöhön. Erityisesti monitieteisillä tutkimusprojekteilla, jotka mahdollistavat sujuvuuden eri puolien empiirisen tarkastelun, voidaan saavuttaa aiempaa kattavampi kokonaiskuva toisen kielen sujuvuudesta.

## Lähteet

- Ahola, S. 2016. Puhetta arvioinnista: yleisten kielitutkintojen arvioijien käsityksiä arvioinnista. Teoksessa A. Huhta & R. Hildén (toim.) *Kielitaidon arviointitutkimus 2000-luvun Suomessa*. AFinLA-e: soveltavan kielitieteen tutkimuksia 9. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 89–109. <https://journal.fi/afinla/article/view/60848>.
- Ahonen, T., S. Tuovinen & T. Leppäsaari 2003. *Nopean sarjallisen nimeämisen testi*. Jyväskylä: Niilo Mäki Instituutti.
- Akamatsu, N. 2008. The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics*, 29, 175–193. DOI: 10.1017/S0142716408080089.
- Baddeley, A. 1996. Exploring the central executive. *The Quarterly Journal of Experimental Psychology Section A*, 49, 5–28. DOI: 10.1080/713755608.
- Balota, D. A., M. J. Yap & M. J. Cortese 2006. Visual word recognition: the journey from features to meaning (a travel update). Teoksessa M. Traxler & M. A. Gernsbacher (toim.) *Handbook of psycholinguistics*. 2. laitos. London: Academic Press.
- Bosker, H. R., A.-F. Pinget, H. Quené, T. Sanders & N. H. De Jong 2013. What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, 30 (2), 159–175. DOI: 10.1177/0265532212455394.
- Brumfit, C. 2000. Accuracy and fluency: the basic polarity. Teoksessa H. Riggenbach (toim.) *Perspectives on fluency*. Ann Arbor (Mich.): University of Michigan Press, 61–73.
- Chambers, F. 1997. What do we mean by fluency? *System*, 25, 535–544. DOI: 10.1016/S0346-251X(97)00046-8.
- Cucchiari, C., H. Strik & L. Boves 2002. Quantitative assessment of second language learners' fluency: comparisons between read and spontaneous speech. *Journal of the Acoustical Society of America*, 111 (6), 2862–2873. DOI: 10.1121/1.428279.
- De Jong, N. H. 2016. Fluency in second language assessment. Teoksessa D. Tsagari & J. Banerjee (toim.) *The handbook of second language assessment*. Berlin: De Gruyter, 203–218.
- De Jong, N. H., R. Groenhout, R. Schoonen & J. H. Hulstijn 2015. Second language fluency: speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*, 36, 223–243. DOI: 10.1016/S0346-251X(97)00046-8.
- Denckla, M. B. & L. E. Cutting 1999. History and significance of Rapid Automatized Naming. *Annals of Dyslexia*, 49, 29–42.
- Derwing, T. M., M. J. Munro, R. I. Thomson & M. J. Rossiter 2009. The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition*, 31, 533–557. DOI: 10.1017/S0272263109990015.

- Di Filippo, G., D. Brizzolara, A. Chilosi, M. De Luca, A. Judica, C. Pecini, D. Spinelli & P. Zoccolotti 2005. Rapid naming, not cancellation speed or articulation rate, predicts reading in an orthographically regular language (Italian). *Child Neuropsychology*, 11 (4), 349–361. DOI: 10.1080/09297040490916947.
- Dörnyei, Z. & J. Kormos 1998. Problem-solving mechanisms in L2 communication: a psycholinguistic perspective. *Studies in Second Language Acquisition*, 20, 349–385.
- Ehri, L. 1991. Development of the ability to read words. Teoksessa R. Barr, M. Kamil, P. Mosenthal & P. D. Pearson (toim.) *Handbook of reading research*. New York (N. Y.): Longman, 383–417.
- Ejzenberg, R. 2000. The juggling act of oral fluency: a psycho-sociolinguistic metaphor. Teoksessa H. Riggenbach (toim.) *Perspectives on fluency*. Ann Arbor (Mich.): University of Michigan Press, 287–313.
- Ellis, N. C. 2002. Frequency effects in language acquisition: a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143–188.
- Engelhardt, P. E., M. Corley, J. T. Nigg & F. Ferreira 2010. The role of inhibition in the production of disfluencies. *Memory & Cognition*, 38 (5), 617–628. DOI: 10.3758/MC.38.5.617.
- Euroopan neuvosto 2012. *Eurooppalainen viitekehys: kielten oppimisen, opettamisen ja arvioinnin yhteinen eurooppalainen viitekehys* Käänt. I. Huttunen & H. Jaakkola. Helsinki: Sanoma Pro.
- Fillmore, C. J. 1979. On fluency. Teoksessa C. J. Fillmore, D. Kempler & W. S-Y. Wang (toim.) *Individual differences in language ability and language behavior*. London: Academic Press, 85–102.
- Foster, P. & P. Skehan 1999. The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, 3, 215–247. DOI: 10.1177/136216889900300303.
- Freed, B. F. 2000. Is fluency, like beauty, in the eyes (and ears) of the beholder? Teoksessa H. Riggenbach (toim.) *Perspectives on fluency*. Ann Arbor (Mich.): University of Michigan Press, 243–265.
- Georgiadou, E. & K. Roehr-Brackin 2017. Investigating executive working memory and phonological short-term memory in relation to fluency and self-repair behavior in L2 speech. *Journal of Psycholinguistic Research*, 46, 877–895. DOI: 10.1007/s10936-016-9463-x.
- Gholamain, M. & E. Geva 1999. Orthographic and cognitive factors in the concurrent development of basic reading skills in English and Persian. *Language Learning*, 49, 183–217. DOI: 10.1111/0023-8333.00087.
- Goldman-Eisler, F. 1968. *Psycholinguistics: experiments in spontaneous speech*. London: Academic Press.
- Grabe, W. 2009. *Reading in a second language: moving from theory to practice*. Cambridge: Cambridge University Press.
- Grosjean, F. 1980. Temporal variables within and between languages. Teoksessa H. W. Dechert & M. Raupach (toim.) *Towards a cross-linguistic assessment of speech production*. Frankfurt am Main: Peter Lang, 39–53.
- Götz, S. 2013. *Fluency in native and nonnative English speech*. Amsterdam: John Benjamins.

- Hildén, R. 2000. *Att tala bra, bättre och bäst: suomenkielisten abiturienttien ruotsin kielen suullinen taito testisuoritusten valossa*. Helsinki: Helsingin yliopisto.
- Hilton, H. 2008. The link between vocabulary knowledge and spoken L2 fluency. *Language Learning Journal*, 36 (2), 153–166. DOI: 10.1080/09571730802389983.
- Housen, A. & F. Kuiken 2009. Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30, 461–473. DOI: 10.1093/applin/amp048.
- Housen, A., F. Kuiken & I. Vedder 2012. Complexity, accuracy and fluency: definitions, measurement and research. Teoksessa A. Housen, F. Kuiken & I. Vedder (toim.) *Dimensions of L2 performance and proficiency: complexity, accuracy and fluency in SLA*. Amsterdam: John Benjamins, 1–20.
- Huensch, A. & N. Tracy-Ventura 2017. Understanding second language fluency behavior: the effects of individual differences in first language fluency, cross-linguistic differences, and proficiency over time. *Applied Psycholinguistics*, 38, 755–785. DOI: 10.1017/S0142716416000424.
- Huhta, A. & R. Hildén 2016. Kielitutkinnot ja muu laajamittainen kielitaidon arviointi Suomessa. Teoksessa A. Huhta & R. Hildén (toim.) *Kielitaidon arviointitutkimus 2000-luvun Suomessa*. AFinLA-e: soveltavan kielitieteen tutkimuksia 9. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 3–26. <https://journal.fi/afinla/article/view/60844>.
- Hulstijn, J. H. 2011. Explanations of associations between L1 and L2 literacy skills. Teoksessa M. S. Schmid & W. Lowie (toim.) *Modeling bilingualism: from structure to chaos*. Amsterdam: John Benjamins, 85–111.
- Hulstijn, J. H., A. Van Gelderen & R. Schoonen 2009. Automatization in second-language acquisition: what does the coefficient of variation tell us? *Applied Psycholinguistics*, 30, 555–582. DOI: 10.1017/S0142716409990014.
- Häyrinen, T., S. Serenius-Sirve & M. Korkman 1999. *Lukilasse: lukemisen, kirjoittamisen ja laskemisen seulontatestistö peruskoulun ala-asteen luokille 1–6*. Helsinki: Psykologien Kustannus.
- Iwashita, N., A. Brown, T. McNamara & S. O'Hagan 2008. Assessed levels of second language speaking proficiency: how distinct? *Applied Linguistics*, 29, 24–49. DOI: 10.1093/applin/amm017.
- Jeon, E. H. 2012. Oral reading fluency in second language reading. *Reading in a Foreign Language*, 24 (2), 186–208.
- Just, M. A. & P. A. Carpenter 1992. A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, 99 (1), 122–149.
- Kahng, J. 2014. Exploring utterance and cognitive fluency of L1 and L2 English speakers: temporal measures and stimulated recall. *Language Learning*, 64, 809–854. DOI: 10.1111/lang.12084.
- Kaukonen, S. & A.-L. Lanu 2005. *Lukemisen osatekijöitä ja ongelmia*. Pro gradu -tutkielma. Jyväskylän yliopisto. <http://urn.fi/URN:NBN:fi:juu-2005325>.
- Kirby, J. R., R. K. Parrila & S. L. Pfeiffer 2003. Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology*, 95, 452–464. DOI: 10.1037/0022-0663.95.3.453.
- Koponen, M. & H. Riggenbach 2000. Overview: varying perspectives on fluency. Teoksessa H. Riggenbach (toim.) *Perspectives on fluency*. Ann Arbor (Mich.): University of Michigan Press, 5–24.



- Kormos, J. 1999. The effect of the speaker variables on the self-correction behaviour of L2 learners. *System*, 27 (2), 207–221. DOI: 10.1016/S0346-251X(99)00017-2.
- 2000. The role of attention in monitoring second language speech production. *Language Learning*, 50, 343–384. DOI: 10.1111/0023-8333.00120.
- Kormos, J. & M. Dénes 2004. Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32 (2), 145–164. DOI: 10.1016/j.system.2004.01.001.
- LaBerge, D. & S. J. Samuels 1974. Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293–323. DOI: 10.1016/0010-0285(74)90015-2.
- Larsen-Freeman, D. 2009. Adjusting expectations: the study of complexity, accuracy and fluency in second language acquisition. *Applied Linguistics*, 30, 579–589. DOI: 10.1093/applin/amp043.
- Lauranto, Y. 2005. Sujuvuuden mittoja. Teoksessa L. Kuure, E. Kärkkäinen & M. Saarenkunnas (toim.) *Kieli ja sosiaalinen toiminta – Language and social action*. AFinLAN vuosikirja 2005. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 127–147.
- Lennon, P. 1990. Investigating fluency in EFL: a quantitative approach. *Language Learning*, 40, 387–417. DOI: 10.1111/j.1467-1770.1990.tb00669.x.
- 2000. The lexical element in spoken second language fluency. Teoksessa H. Riggenbach (toim.) *Perspectives on fluency*. Ann Arbor (Mich.): University of Michigan Press, 25–42.
- Levelt, W. J. M. 1989. *Speaking: from intention to articulation*. Cambridge (Mass.): MIT Press.
- Lintunen, P. 2014. Ääntämisen oppiminen ja opettaminen. Teoksessa P. Pietilä & P. Lintunen (toim.) *Kuinka kieltä opitaan: opas vieraan kielen opettajalle ja opiskelijalle*. Helsinki: Gaudeamus, 165–187.
- Lintunen, P. & P. Peltonen 2016. Sujuvaa puhetta vieraalla kielellä. *Tempus*, 5, 18–19.
- MacLeod, C. 1991. Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin*, 109, 163–203. DOI: 10.1037/0033-2909.109.2.163.
- Marian, V., H. K. Blumenfeld, E. Mizrahi, U. Kania & A.-K. Cordes 2013. Multilingual Stroop performance: effects of trilingualism and proficiency on inhibitory control. *International Journal of Multilingualism*, 10 (1), 82–104. DOI: 10.1080/14790718.2012.708037.
- Martin, M. & R. Alanen 2011. Oppijankieli: kehitystä ja virheitä. Teoksessa P. Kalaja, R. Alanen & H. Dufva (toim.) *Kieltä tutkimassa: tutkielman laatijan opas*. Helsinki: Finn Lectura, 33–48.
- McCarthy, M. 2010. Spoken fluency revisited. *English Profile Journal*, 1 (1), 1–15. DOI: 10.1017/S2041536210000012.
- McCutchen, D. 2000. Knowledge, processing, and working memory: implications for a theory of writing. *Educational Psychologist*, 35 (1), 13–23. DOI: 10.1207/S15326985EP3501\_3.
- Miyake, A., N. P. Friedman, M. J. Emerson, A. H. Witzki, A. Howerter & T. D. Wager 2000. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cognitive Psychology*, 41, 49–100. DOI: 10.1006/cogp.1999.0734.

- Mora, J. C. & M. Valls-Ferrer 2012. Oral fluency, accuracy, and complexity in formal instruction and study abroad learning contexts. *TESOL Quarterly*, 46 (4), 610–641. DOI: 10.1002/tesq.34.
- Mutta, M. 2006. Yksilölliset taukoprofiilit vieraan kielen kirjoitusprosesseissa. Teoksessa P. Pietilä, P. Lintunen & H.-M. Järvinen (toim.) *Kielenoppija tänään – Language learners of today*. AFinLAN vuosikirja 2006. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 379–396.
- Nigg, J. T. 2001. Is ADHD a disinhibitory disorder? *Psychological Bulletin*, 127, 571–598. DOI: 10.1037/0033-2909.127.5.571.
- Nissilä, L., M. Martin, H. Vaarala & I. Kuukka 2006. *Saako olla suomea? Opas suomi toisena kielenä -opetukseen*. Helsinki: Opetushallitus.
- O’Connell, D. C. & S. Kowal 2004. The history of research on the filled pause as evidence of *The Written Language Bias in Linguistics* (Linell, 1982). *Journal of Psycholinguistic Research*, 33, 459–474.
- Olkkonen, S. 2012. Suoritusnopeus vieraan kielen taitojen automaattistumisen mitarina. Teoksessa L. Meriläinen, L. Kolehmainen & T. Nieminen (toim.) *AFinLA-e: soveltavan kielitieteen tutkimuksia 4*. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 36–46.
- 2017a. Processing limitations in L2 fluency: analysis of inaccuracies in lexical access. *APPLES: Journal of Applied Language Studies*, 11, 19–41. DOI: 10.17011/apples/urn.201704252081.
- 2017b. *Second and foreign language fluency from cognitive perspective: inefficiency and control of attention in lexical access*. Jyväskylä Studies in Humanities 314. Jyväskylä: Jyväskylän yliopisto. <http://urn.fi/URN:ISBN:978-951-39-7062-8>.
- Olkkonen, S., K. Eklund & P. H. T. Leppänen tekeillä. *Fluency of lexical access in L1 and L2 in relation to L2 literacy*.
- Opetushallitus 2004. Perusopetuksen opetussuunnitelman perusteet 2004. [http://www.oph.fi/saadokset\\_ja\\_ohjeet/opetussuunnitelmien\\_ja\\_tutkintojen\\_perusteet/perusopetus](http://www.oph.fi/saadokset_ja_ohjeet/opetussuunnitelmien_ja_tutkintojen_perusteet/perusopetus) [luettu 24. 5. 2017].
- Peltonen, P. 2017. Temporal fluency and problem-solving in interaction: an exploratory study of fluency resources in L2 dialogue. *System*, 70, 1–13. DOI: 10.1016/j.system.2017.08.009.
- 2018. L2 fluency in spoken interaction: a case study on the use of other-repetitions and collaborative completions. Teoksessa M. Kuronen, P. Lintunen & T. Nieminen (toim.) *Insights into second language speech*. AFinLA-e. Soveltavan kielitieteen tutkimuksia 10. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 118–138.
- Peltonen, P. & P. Lintunen 2016. Integrating quantitative and qualitative approaches in L2 fluency analysis: a study of Finnish-speaking and Swedish-speaking learners of English at two school levels. *European Journal of Applied Linguistics*, 4, 209–238. DOI: 10.1515/eujal-2014-0018.
- Pickering, M. J. & S. Garrod 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–226. DOI: 10.1017/S0140525X04000056.
- Pietilä, P. & P. Lintunen 2014. Kielen oppiminen ja opettaminen. Teoksessa P. Pietilä & P. Lintunen (toim.) *Kuinka kieltä opitaan: opas vieraan kielen opettajalle ja opiskelijalle*. Helsinki: Gaudeamus, 11–25.

- Protopapas, A., A. Fakou, S. Drakopoulou, C. Skaloumbakas & A. Mouzaki 2013. What do spelling errors tell us? Classification and analysis of errors made by Greek schoolchildren with and without dyslexia. *Reading & Writing: An Interdisciplinary Journal*, 26, 615–646. DOI: 10.1007/s11145-012-9378-3.
- Puolakanaho, A., T. Ahonen, M. Aro, K. Eklund, P. Leppänen, A-M. Poikkeus, A. Tolvanen, M. Torppa & H. Lyytinen 2007. Very early phonological and language skills: estimating individual risk of reading disability. *Journal of Child Psychology and Psychiatry*, 48, 923–931. DOI: 10.1111/j.1469-7610.2007.01763.x.
- Riggenbach, H. 1991. Toward an understanding of fluency: a microanalysis of nonnative speaker conversations. *Discourse Processes*, 14 (4), 423–441. DOI: 10.1080/01638539109544795.
- Robinson, P. 2001. Task complexity, cognitive resources, and syllabus design: a triadic framework for examining task influences on SLA. Teoksessa P. Robinson (toim.) *Cognition and second language instruction*. Cambridge: Cambridge University Press, 287–318.
- Rossiter, M. J. 2009. Perceptions of L2 fluency by native and non-native speakers of English. *The Canadian Modern Language Review*, 65, 395–412. DOI: 10.3138/cmlr.65.3.395.
- Salmi, P. 2008. *Nimeäminen ja lukemisvaikeus: kehityksen ja kuntoutuksen näkökulma*. Jyväskylä: Jyväskylän yliopisto. <http://urn.fi/URN:ISBN:978-951-39-3458-3>.
- Sato, M. 2014. Exploring the construct of interactional oral fluency: second language acquisition and language testing approaches. *System*, 45, 79–91. DOI: 10.1016/j.system.2014.05.004.
- Schmidt, R. 1992. Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 14, 357–385. DOI: 10.1017/S0272263100011189.
- Segalowitz, N. 2000. Automaticity and attentional skill in fluent performance. Teoksessa H. Riggenbach (toim.) *Perspectives on fluency*. Ann Arbor (Mich.): University of Michigan Press, 200–219.
- 2010. *The cognitive bases of second language fluency*. New York (N. Y.): Routledge.
- Segalowitz, N. & S. Frenkiel-Fishman 2005. Attention control and ability level in a complex cognitive skill: attention shifting and second-language proficiency. *Memory & Cognition*, 33, 644–653. DOI: 10.3758/BF03195331.
- Service, E., M. Simola, O. Metsänheimo & S. Maury 2002. Bilingual working memory span is affected by language skill. *European Journal of Cognitive Psychology*, 14, 383–408. DOI: 10.1080/09541440143000140.
- Shaw, P. & A. McMillion 2008. Proficiency effects and compensation in advanced second language reading. *Nordic Journal of English Studies*, 7 (3), 123–143.
- Skehan, P. 2003. Task-based instruction. *Language Teaching*, 36, 1–14. DOI: 10.1017/S026144480200188X.
- 2009. Modelling second language performance: integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30, 510–532. DOI: 10.1093/applin/amp047.
- 2014. The context for researching a processing perspective on task performance. Teoksessa P. Skehan (toim.) *Processing perspectives on task performance*. Amsterdam: John Benjamins, 1–26.

- Tannen, D. 1989. *Talking voices: repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.
- Tavakoli, P. 2011. Pausing patterns: differences between L2 learners and native speakers. *ELT Journal*, 65 (1), 71–79. DOI: 10.1093/elt/ccq020.
- 2016. Fluency in monologic and dialogic task performance: challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics in Language Teaching*, 54, 133–150. DOI: 10.1515/iral-2016-9994.
- Toivola, M. 2011. *Vieraan aksentin arviointi ja mittaaminen suomessa*. Helsinki: Helsingin yliopisto. <http://urn.fi/URN:ISBN:978-952-10-7217-8>.
- Towell, R., R. Hawkins & N. Bazergui 1996. The development of fluency in advanced learners of French. *Applied Linguistics*, 17, 84–119. DOI: 10.1093/applin/17.1.84.
- Walczyk, J. J., C. S. Marsiglia, A. K. Johns & K. S. Bryan 2004. Children's compensations for poorly automated reading skills. *Discourse Processes*, 37 (1), 47–66.
- Van Moere, A. 2012. A psycholinguistic approach to oral language assessment. *Language Testing*, 29 (3), 325–344. DOI: 10.1177/0265532211424478.
- Wolf, M. 1986. Rapid alternating stimulus naming in the developmental dyslexias. *Brain and Language*, 27, 360–379. DOI: 10.1016/0093-934X(86)90025-8.

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 258–276.

Minnaleena Toivola<sup>1</sup> & Riikka Ullakonoja<sup>2</sup>

<sup>1</sup>University of Helsinki, <sup>2</sup>University of Jyväskylä

## Identification of Russian accented Finnish by native and non-native listeners with and without Finnish proficiency

The study focuses on how different groups of listeners, that is native Finnish speakers (n = 18), non-native learners of Finnish (n = 12) and non-native non-learners of Finnish (n = 18), judge samples (n = 80) of Russian accented Finnish in an auditory experiment. The samples are read-aloud phrases of three Finnish speakers and three L1 Russian beginner learners of Finnish who were recorded three times at six-month intervals during their residence in Finland. The results show that the listener groups differ significantly from each other in the ability to identify the Russian accented samples. In addition, all the listener groups reported to have paid more attention to prosodic and segmental cues than speech rate. Finally, we conclude by considering implications of the study to further longitudinal studies on perception of foreign accented speech.

**Keywords:** Foreign accent, Finnish as a second language, non-native raters, pronunciation

## 1 Introduction

The small scale study reported in this paper aims to investigate how different listener groups, including naive listeners, successfully identify Russian speakers in an auditory experiment consisting of short read-aloud samples from native Finnish speakers and L1 Russian beginner learners. We explore the potential of such a non-standard experimental design in the study of foreign accent. In addition, we are interested in what phonetic aspects the listeners based their judgments on. The listener groups are: native Finnish speakers ( $n = 18$ ), non-native learners of Finnish ( $n = 12$ ) and non-native non-learners of Finnish ( $n = 18$ ), to be further referred to as G1, G2 and G3, respectively. The study aims to contribute to the almost non-existent literature on judgments of foreign accented Finnish and to offer suggestions for further studies on this topic.

We will now briefly describe the prosody of the languages under investigation. Russian prosody has a greater role in communication than Finnish prosody, because Russian uses intonation extensively for distinguishing questions from statements whereas Finnish does not (Bondarko 1998; Iivonen 1978). Finnish intonation has often been described as rather monotonous and produced with a narrow pitch range with creaky voice occurring frequently (Iivonen 1998, 2009a). Russian intonation, on the other hand, is more variable and lively and creaky voice typically does not occur (Volskaya 2009). In Finnish, word stress is fixed on the first syllable (Iivonen 2009b), whereas in Russian it can be placed on any syllable and even change position in different forms of the same syllable (Bondarko 2009). Finnish stressed vowels do not differ as greatly from unstressed ones in quality and quantity compared to Russian ones, which differ from their unstressed counterparts a great deal (Bondarko 2009; Iivonen 1998). Previous research on Russian accented Finnish shows, for example, that non-native like prosodic chunking, stress (exaggerated stress on the word-initial syllable) and pitch variation (rapid increase in pitch in word-final syllables or utterance final positions) are typical for Russian learners of Finnish (Aho & Toivola 2008).

In this paper, we first summarize previous studies on foreign accent, focusing on the perception of foreign accent and the role of listener's background in perception. A material and methods section, describing the auditory experiment in detail, follows this introduction. We describe the listener groups as well as the three L1 Russian speakers, beginner learners of Finnish in more detail as well as explain the statistical methods used. Next, we present our main findings and finish with a discussion.

## 2 Perception of foreign accent

Foreign accent is the term often used for non-native pronunciation that deviates from native speech (Scovel 1969; Flege 1981; Munro 2008). As mentioned by Munro & Derwing (2015), studying foreign accent can be a key to understanding speech processing, and the results from such studies can often be applied to pronunciation teaching.

The goal for pronunciation learning is often comprehensibility (Jenkins 2000; Walker 2011) rather than complete native-likeness or speaking without any foreign accent, at least for adult learners (e.g. Abrahamsson & Hyltenstam 2009). Nonetheless, foreign accent can in some cases hinder comprehensibility, whereas in other cases accented speech can be completely comprehensible (Munro & Derwing 1999). For immigrants, foreign accent has societal relevance in integrating into the host society, for example, because native speakers may value it negatively (Lippi-Green 2012; Leinonen 2015). As Russian speakers are the largest immigrant group in Finland ( $n = 75,444$  in 2016, Statistics Finland 2017), negative accent perception is of particular concern to this group in Finnish society. Attitudes towards Russian speakers have been rather negative; for example, over 60% of Finns reported they do not wish Russian speakers to move to Finland (Jaakkola 2006, 2009). Aho & Toivola (2008) found in their study that many Russian learners of Finnish wanted to speak Finnish without a Russian accent. For immigrants, sounding native, or at least not easily being recognized as a non-native speaker, can thus be a justifiable learning goal.

The present study focuses on factors contributing to the perception and identification of foreign accent. More precisely, the study investigates how successfully native and non-native listeners identify the foreign accent in Finnish spoken by native speakers of Russian and what phonetic aspects their judgments are based on. Non-learners of the language have been used as listeners only in a few previous studies measuring foreign accent, fluency or pronunciation (e.g. Major 2007; Weber & Pöllmann 2011; Wilkerson 2010). In the majority of previous studies the listeners have been native speakers or non-native learners of the language to be judged (Gonet & Pietroń 2004; MacKay et al. 2006; Munro et al. 2006). The studies have shown that different groups of listeners can differ in their evaluations. For example, in Weber & Pöllmann's (2011) study the judgments of non-native non-learners differed from native and non-native language learners. However, non-native non-learners are also able to judge the stimuli very reliably (Major 2007), especially the stimuli with a strong foreign accent. The studies speculate that non-native non-learner listeners may rely on their first language knowledge in their judgments. As one explanation for perception of foreign accent Major (2007) proposes the term *displaced foreign accent detection*, which refers to the fact that the listener

is paying attention to the features of their native language that can be heard in the unknown language. Pilot studies by Gilbert (1980) have indicated the strong role of prosody in recognizing the speaker's L1. Furthermore, Gupta (2005) showed that different accents of English are easier to understand if one is familiar with the accent in question.

Previous research has also focused on rater's expertise as a background factor influencing ratings and the results have been controversial (see review by Piske et al. 2001). In a study by Cunningham-Andersson & Engstrand (1989) it did not play a significant role as naive listeners were able to identify foreign accent from rather short samples as well as expert listeners. Similarly, in Kennedy & Trofimovich (2008) English as L2 teachers' foreign accent ratings did not differ from naive listeners' ratings. However, Bongaerts, Mennen and van der Slik's (2000) study showed that language teachers and non-teachers differed in their ratings. Thus, Piske et al. (2001) recommend in their review that listeners from different backgrounds should be used in studies investigating the perception of foreign accent.

Previous studies have identified a number of factors contributing to the perception of accentedness. Research focusing on fluency and the strength of foreign accent (e.g. Major 2007; Weber & Pöllmann 2011) suggest that the utterance duration as well as slower speech rate would particularly indicate that the speaker is a non-native speaker. However, the link has not always been this clear. Derwing & Munro (2001) found that speech that was too slow or too fast did not sound native-like. Major (2007) proposes that duration differences alone do not account for the foreign accent ratings. Trofimovich & Baker (2006) found that speech rate and duration of pauses were associated with foreign accent more than stress timing and peak alignment.

Pinget et al. (2014) studied both suprasegmental and segmental features and measured pitch alternation hypothesizing that monotonous pitch contributes to the perception of foreign accent. The results showed, however, that pitch alternation and the sound segments chosen for the study explained only a small proportion of the strength of foreign accent. Additionally McCullough (2013) studied multiple acoustic properties such as VOT, vowel quality, *f<sub>0</sub>* (fundamental frequency), vowel duration and sentence stress in English words produced by American English, Hindi, Korean and Mandarin speakers. In the study VOT and vowel quality were linked to the perception of foreign accent more than *f<sub>0</sub>* and vowel duration. Saito et al. (2017) studied 11 linguistic factors in the perception of foreign accent of Japanese learners of English. They found that native listeners' ratings of accentedness were associated with word stress and vowel/consonant errors.

In Riney, Takagi and Inutsuka's (2005) study the listeners disagreed on which of the non-native speakers were perceived the most and least native-like. They focused on Japanese and American raters listening to Japanese L2



speakers of English and showed that all listeners were able to identify the native speakers. However, non-native listeners paid more attention to intonation, fluency and speech rate as compared to native listeners, who based their judgments more on deviances in the segmental level (especially /r/ and /l/ sounds).

Previous literature on L2 Finnish is scarce both from perception and production points of view. In previous perception studies the raters have been native Finnish speakers (e.g. Toivola 2011; Leinonen 2015; Uzal et al. 2015). The perception study by Leinonen (2015) showed that Finnish adolescents living in different parts of Finland rated foreign accented Finnish of a number of different L1s consistently. Speakers were rated differently based on their L1. Swahili and Vietnamese L1 speakers were given the strongest foreign accent ratings, whereas Arabic L1 speakers were rated the weakest and English and Russian L1 speakers in between. There is one recent study (Ahola & Tossavainen 2016) on L2 Finnish pronunciation assessment focusing on Estonian accented Finnish. It found that the similarities between Estonian and Finnish pronunciation make Estonian accented Finnish rather easy to understand and to listen to. However, the study also found that raters perceived Finnish spoken by Estonian L1 speakers too fast and not containing enough pauses.

Features related to L2 Finnish pronunciation and fluency have been studied by Toivola and colleagues. Toivola et al. (2009) compared Finnish native speakers and Finnish L2 learners (from four different L1 backgrounds: Russian, Thai, Turkish and Vietnamese) and found that pauses were longer in read-aloud speech by Finnish native speakers than L2 learners. Toivola et al. (2011) also found that articulation rate becomes faster in a longitudinal setting, as the length of residence in Finland increases, both in read-aloud and conversational speech. The number of pauses in read-aloud speech also decreases, but the mean duration of pauses increases. They conclude that learning native-like pausing in spoken Finnish requires the ability to use context-dependent pause durations in speech. Toivola (2011) also studied the role of phonetic features in perception of foreign accent and found that the number of single deviant phonetic segments, filled pauses and articulation rate contributed to the perceived degree of accentedness in the spontaneous speech of experienced Finnish L2 learners (with Russian L1 background). In addition, in the read-aloud data the number of single deviant phonetic segments and their quality explained the majority of the strength of the foreign accent by Russian speakers (Toivola 2011).

Small scale studies on L2 Finnish spoken by Russians have also been conducted by Ullakonoja and colleagues. These studies have focused on how adult Russian native speakers with no prior knowledge of Finnish imitate short Finnish utterances (Ullakonoja et al. 2014a,b) and how young Russian immigrants produce Finnish segmental duration and length in a read-aloud task

(Ullakonoja & Kuronen 2015). The results showed, firstly, that subject's working memory is correlated with the comprehensibility of imitated utterances, and secondly, that it is challenging to imitate/to learn the Finnish durational contrasts. All studies revealed great interspeaker variation in the success of production.

### 3 Aims, data and methods

The main aim of the study is to investigate the identification of Russian accented Finnish by three different listener groups. In particular, the present study sought to answer the following questions:

1. How well were non-native, that is Russian L1, speakers identified by different listener groups (that is native Finnish speakers, native Russian speakers who are learning Finnish and native Russian speakers with no knowledge of Finnish)?
2. Are samples from consequent recording sessions rated differently, and if so, is the change due to perceived changes in pronunciation?

In contrast to most previous studies on L2 pronunciation, the present study focuses on pronunciation in a longitudinal research setting and includes a listener group with no knowledge of the language under investigation. We focus on three adult beginner learners of Finnish, with Russian as their L1, recorded three times within 6-month intervals. Extracts of the recordings were subjected to a perception experiment, through which we investigated the differences between rater groups in identification and if the perceptual judgments change over time and if so, whether the pronunciation also changes as reported by the raters. Seeing how non-native non-learners of Finnish judge the speech stimuli in different recording sessions is of particular interest, as such a listener group is very rarely included in L2 perception studies. The underlying hypotheses are that, firstly, the degree of proficiency in Finnish is linked with the ability to identify the Russian accented speech (the better the raters' proficiency in Finnish, the better their ability to identify Russian speakers' samples is) and, secondly, that all the different groups of raters judge the L2 speakers more often as native speakers of Finnish in the final recording session.

#### 3.1 Raters

Three different groups of raters ( $n = 48$ ), listened to the read-aloud stimuli by L1 and L2 speakers in the auditory experiment: L1 Finnish speakers (G1), Russian speakers with L2 Finnish (G2), L1 Russian speakers with no previous

knowledge or proficiency in Finnish (G3). None of the listeners reported hearing difficulties and they were all students. G1, the L1 Finnish speakers ( $n = 18$ , all female) were from the Helsinki or Jyväskylä regions aged 20–33 (mean age 23) with no knowledge or previous study of Russian. Their only L1 was Finnish, but they had studied at least English and Swedish at school (as is typical in Finland). All Finnish speakers can be considered at least somewhat familiar with Russian-accented Finnish, as Russian is one of largest immigrant minority languages in Finland. G2, the L1 Russian speakers ( $n = 12$ , all female) with L2 Finnish, aged 20–35 (mean age 28) were living in the Helsinki or Jyväskylä regions at the time of the experiment. They had a self-rated minimum of B1-level oral proficiency of Finnish on the CEFR scale (CEFR 2001), i.e. were at least at an intermediate level. They had studied Finnish 2–13 years (mean 6) and had lived in Finland from 1 to 13 years (mean 6). G3, the L1 Russian speakers ( $n = 18$ , 2 male, 16 female) with no Finnish proficiency came from the St. Petersburg region and were between 17 and 30 years of age (mean age 20).

### 3.2 Speakers

In the auditory experiment there were samples from two groups of speakers: non-native speakers of Finnish with Russian as their L1 and native speakers of Finnish. The non-native speakers were three female speakers (further Rus1, Rus2 and Rus3) from the ProoF-project corpus (ProoF 2012; see Aho et al. 2016, for a description of the corpus) aged 25–29 (mean age 27). They were recorded three times with 6-month intervals (further T1, T2 and T3) and they had lived in Finland for 8–12 months (mean 11 months) and had studied Finnish for 4–9 months (mean 7 months) before the first recording session. Two speakers continued their Finnish studies throughout the experiment and one (Rus3) interrupted her Finnish lessons after T2. Two speakers (Rus1 and Rus3) reported that they mostly spoke Finnish outside the home, and Rus3 also at home, whereas Rus1 spoke Russian at home. Rus2 said that she spoke some Finnish, but mostly English both at home and outside the home. The native speakers of Finnish were three female speakers from the Helsinki region aged 21–24 (mean 22), as well as contributors to the ProoF-project corpus.

### 3.3 Speech samples

The read-aloud speech samples were recorded in a sound-proof studio, using high quality audio equipment within the ProoF-project. Every speaker was recorded with a head-mounted microphone at a sample rate of 44.1 kHz and sample size of 16 bits. Speakers were asked to read aloud short sentences using an ordinary speech rate. As opposed to previous studies, which have used longer stimuli, we chose short extracts of read-aloud speech for the auditory

experiment. The reason for this is that, firstly, we hypothesized that producing Finnish intonation, which has a high beginning followed by declination, would be difficult for Russian speakers to produce. Secondly, longer samples would have contained more pauses and hesitation as well as a potentially slower speech rate in the L2 speech, which would have increased the likelihood of being identified as “Russian”. The following seven phrases were extracted from the beginning of longer read-aloud sentences from each speaker (for L2 speakers from three consequent recording sessions): *viime yönä* (last night), *kuten tiedät* (as you know), *hän sanoi* (s/he said), *en tiedä* (I don’t know), *eilen illalla* (last night), *mielestäni* (in my opinion), *viime kesänä* (last summer).

### 3.4 Listening test

The auditory experiment consisted of a total of 80 samples, 63 of which were obtained from three Russian native speakers from the three recording sessions over time, and the remaining 17 samples from three Finnish native speakers (5–6 samples per speaker). Thus, the total number of all ratings was 3840, of which 3024 were ratings of Russian speakers’ samples.

In the experiment (23 min), the listeners first heard seven practice stimuli, containing both L2 and L1 samples. Then, the actual stimuli were presented in two different randomized orders so that all listeners listened to the stimuli only once and had 10 seconds to respond after each stimulus. The experiments were all administered in a quiet room and the stimuli were presented either through headphones or loudspeakers (where the use of headphones was impossible for practical reasons). The listeners chose the language (Finnish/Russian) of the rating sheet. First, the listeners were asked to decide whether the speaker’s mother tongue was Russian. Second, the two listener groups (L1 Finnish speakers and L1 Russian speakers with L2 Finnish proficiency) who knew Finnish were asked to define the basis for their judgments for each stimulus in a multiple-choice question, (allowing for multiple responses per stimulus) where the alternatives were prosody (referring to intonation, stress and rhythm), speech rate and vowel and consonant quality/quantity. As we considered such a multiple-choice question too demanding for the participants for whom Finnish was an unknown language, the non-learners of Finnish group were asked to express the basis of their judgments across all the stimuli freely in their own words. This was a good decision, as we observed that for some listeners the fact that they had to participate in an experiment in a language they have never heard was enough to make them question their willingness and capability to participate.

### 3.5 Data analysis

The perceptual judgments were analyzed and processed quantitatively using MS Excel 14.0 and IBM SPSS Statistics 24.0 software. For investigating the differences between the listener groups, we used one-way between-groups ANOVA analysis of variance and reported the values of the Welch test as the assumption of the homogeneity of variance was violated. Further, we used the Tukey HSD post-hoc comparisons. In the further analysis of the data, we used non-parametric tests, as the parametric ones require normal distribution of the data, which was not the case here for the smaller sub-groups. When comparing listener groups we used the Kruskal-Wallis test, a non-parametric equivalent to one-way variance analysis allowing for comparison of three groups. In comparing the three time points, we used the Friedman Test, an alternative to the one-way ANOVA with repeated measures.

## 4 Results

### 4.1 Identification by the listener groups

First, we present the results considering the comparison of the three listener groups in the success of identification samples spoken by a native Russian speaker and by a native Finnish speaker. Figure 1 shows the differences in identification between the listener groups. Not surprisingly, the non-Russian, i.e. Finnish speakers' stimuli ( $n = 17$ ) were almost perfectly recognized by (G1) Finnish listeners ( $n = 18$ ) in the experiment, i.e. in only 2% of these cases Finnish listeners responded that the speaker's mother tongue was Russian. Russian listeners with Finnish proficiency (G2) ( $n = 12$ ) judged Finnish stimuli as native speakers of Russian in 7% of the cases. Non-learner listeners (G3) ( $n = 18$ ) falsely identified Finnish stimuli as native speakers of Russian in 15% of the cases. A one-way between-groups analysis of variance was conducted to explore the impact of listener group on identification. As the Levene's test indicated that the homogeneity of variance assumption of the ANOVA was violated, we are reporting the values of the Robust test of equality of means, the Welch test. There was a statistically significant difference at the  $p < .0001$  level in the three listener groups:  $F(2, 23.78) = 9.85, p = .001$ . Post-hoc comparisons using the Tukey HSD test indicated that the mean score for G3 ( $M = 14.44, SD = 2.15$ ) was significantly different from G1 ( $M = 16.61, SD = .61$ ) and G2 ( $M = 15.91, SD = .90$ ). However, G1 did not differ from G2. This indicates that proficiency in Finnish (whether native or language learner) helps to identify the Finnish native speakers' samples better than listeners with no knowledge of Finnish.

However, it is more interesting to investigate how successfully the Russian

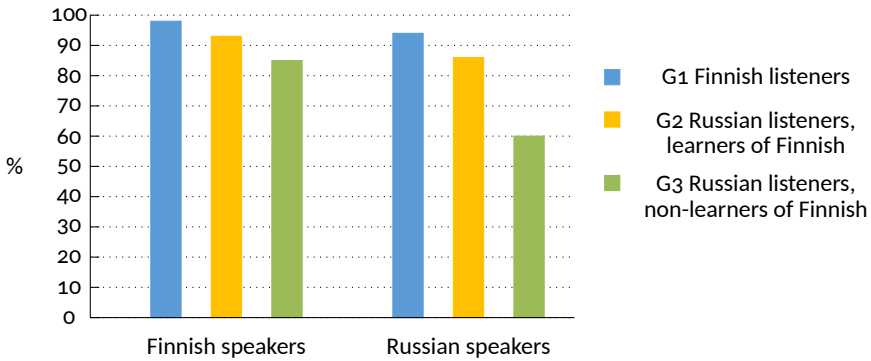


FIGURE 1. Identification rates of Finnish and Russian speakers' samples by the three listener groups.

speakers' stimuli ( $n = 63$ ) were identified. As Figure 1 shows, Finnish native listeners were in general better than Russian listeners at recognizing speakers with a Russian language background (94% recognition rate for all speakers). Also Russian listeners with Finnish proficiency (G2) were better than those with no proficiency in Finnish (G3) in recognizing the Russian speaker samples. G2 recognized 86% and G3 60% of the Russian stimuli. A one-way between-groups analysis of variance was conducted to explore the impact of listener group on identification. As the Levene's test indicated that the homogeneity of variance assumption of the ANOVA was violated, we are again reporting the values of the Robust test of equality of means, the Welch test. There was a statistically significant difference at the  $p < 0.0001$  level in the three listener groups:  $F(2, 23.43) = 126.24$ ,  $p = 0.0001$ . Post-hoc comparisons using the Tukey HSD test indicated that the mean score for G3 ( $M = 36.83$ ,  $SD = 4.89$ ) was significantly different from G1 ( $M = 59.28$ ,  $SD = 3.32$ ) and G2 ( $M = 53.17$ ,  $SD = 7.18$ ). G1 also differed significantly from G2 in mean score. Thus, the degree of proficiency in Finnish seems to be linked to the ability to identify Russian speakers' stimuli from the native Finnish ones.

#### 4.2 Identification across speakers and time

Differences between the listener groups for all the three speakers at three time points (recording sessions T1, T2 and T3) are shown in Figure 2. Kruskal-Wallis test  $\chi^2(2, 48) = 28.76$  (T1), 32.34 (T2) and 34.62 (T3),  $p = 0.001$  shows a statistically significant difference in ratings across time. Judgments by the native Finnish listeners (G1) (95%, 93% and 94%) and Russian listeners with Finnish proficiency (G2) (83%, 85% and 88%) were significantly different from Russian listeners with no proficiency in Finnish (G3) (67%, 52% and 66%) in

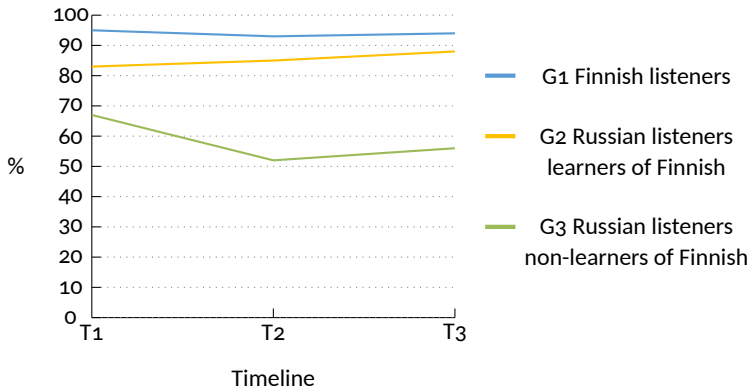


FIGURE 2. Judgments of all the Russian speakers at different recording sessions (T1, T2, T3) by all rater groups.

all the recording sessions (T1, T2 and T3). This was shown in pairwise comparisons at  $p = 0.001$  level between G1 and G3 ( $X^2 = 12.54$ ) and at  $p = 0.045$  level between G2 and G3 ( $X^2 = 24.75$ ) for T1. For the second recording session (T2) this was shown in pairwise comparisons at  $p = 0.001$  level between G1 and G3 ( $X^2 = 20.19$ ) and between G2 and G3 ( $X^2 = 25.28$ ) and for the third recording session (T3) at  $p = 0.001$  level between G1 and G3 ( $X^2 = 26.39$ ) and between G2 and G3 ( $X^2 = 19.75$ ).

As Figure 2 indicates, there is most variation across time in the judgments of the Russian listeners with no proficiency in Finnish (G3), and thus we will next focus on them and the differences in their judgments between the recording sessions. The results of the Friedman Test indicated that there was a statistically significant difference in the identification rates across the three time points for Russian listeners with no proficiency in Finnish (G3) (based on Friedman Test  $X^2(2) = 11.35$ ,  $p < .01$ ). For the other two listener groups, such a difference was not observed. Post-hoc testing was done using Wilcoxon Signed Ranks Test, which revealed that T1 differed from T2 significantly ( $z = -3.23$ ,  $p = .005$ ), but T2 did not differ from T3 in the ratings of G3.

Figure 3 shows only the ratings ( $n = 1134$ ) of G3, that is Russian listeners with no Finnish proficiency ( $n = 18$ ) across time and speakers. All Russian speakers were judged less often as Russian between the first two recording sessions (from T1 to T2) and Rus1 and Rus2 between the first and last recording sessions (from T1 to T3) by the Russian listeners with no proficiency in Finnish (G3) (Figure 3). For speakers Rus1 and Rus2, there was a tendency to be less often identified as Russian as the time they spent in Finland increased. However, there was a different tendency for speaker Rus3, who was identified as Russian in 74% of the samples at T1 and 48% at T2, but the identification

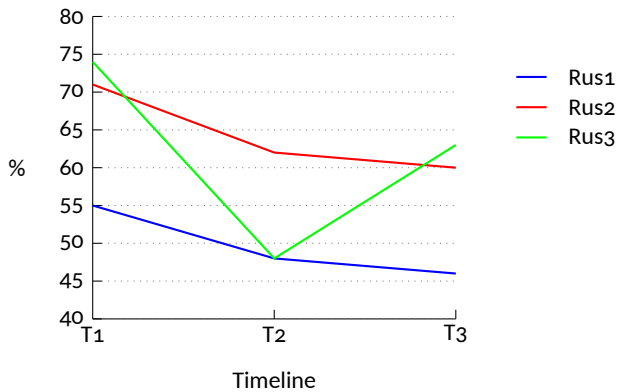


FIGURE 3. Success in identifying the three speakers' (Rus1, Rus2, Rus3) mother tongue as Russian at three time points (T1, T2, T3) by the Russian naive listener group (G3) only.

rate increased at T3 to 63%. This could be due to the fact that this speaker interrupted her Finnish course between T2 and T3.

### 4.3 Pronunciation features behind the judgments

The listeners also reported the basis of their judgments. Russian listeners with no Finnish proficiency (G3) answered an open-ended question for the whole task, whereas Finnish listeners (G1) and Russian listeners with Finnish proficiency (G2) were asked to choose between prosody, segmental features and speech rate for each stimulus they heard. A majority of the Russian listeners with no proficiency in Finnish (G3) mentioned one or two prosodic features as the basis of their judgments. Of these, intonation was mentioned most often, but also rhythm, stress and segment duration were reported. The Finnish listeners reported that they paid most attention to segmental features (48%), followed by prosody (37%), whereas speech rate was mentioned in only 15% of the cases.

Figure 4 shows the judgments of only one group, that is Russian listeners with Finnish proficiency (G2) at different recording sessions. This group mostly paid attention to segmental quality and duration at all the time points: 66% of the listeners mentioned this at T1, 65% at T2, and 66% at T3. At the first recording session (T1) prosody (62% of the listeners) and speech rate (44% of the listeners) were also more frequently mentioned than in the other recording sessions, where prosody was mentioned by 52% of the listeners both at T2 and T3 and speech rate by 28% at T2 and 26% at T3.



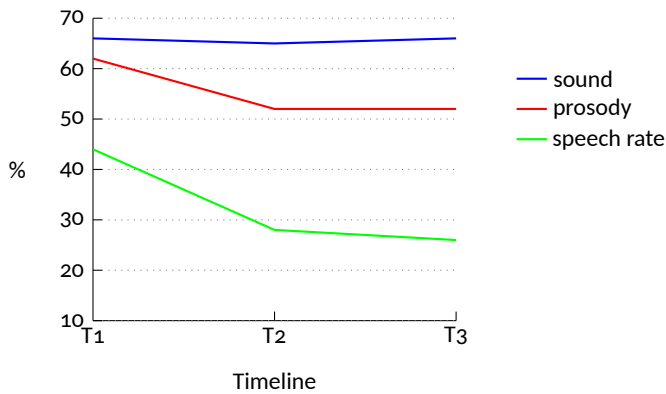


FIGURE 4. Responses of the Russian listeners with Finnish proficiency (G2) about the features of speech they based their judgments on at different recording sessions (T1, T2 and T3).

## 5 Discussion and conclusions

The purpose of the study was to investigate the potential differences between three different listener groups in evaluation of Russian accented Finnish. The study shows that proficiency in Finnish seems to be connected with the success of differentiation of Russian speakers' stimuli from the native Finnish ones in an auditory experiment where the listeners are asked to define whether each Finnish stimulus is spoken by a native Russian speaker or not. Finnish listeners (G1) identified both Finnish speakers' and Russian speakers' stimuli with very high accuracy and Russian listeners with Finnish proficiency (G2) were also rather accurate in their identification of both speaker groups. Similarly as in Weber & Pöllmann (2011) the ratings of non-native non-learners differed from ratings of other listener groups. The listeners with no proficiency in Finnish (G3) were the least successful in identifying the Russian speakers. The differences in the identification rates between the three listener groups were significant for Russian speakers' samples, but only between G1 and G2 for Finnish speakers' samples. Thus, we can conclude that some or native proficiency in Finnish makes a difference in identification of the Finnish speakers' stimuli as compared to the listeners without Finnish proficiency. For the main research question, that is, how well non-native speakers were identified by different listener groups, we can conclude that the ability to identify Russian accented samples is linked with the listener's degree of proficiency in Finnish.

When the listeners were asked to define the basis of their judgments overall, many mentioned prosodic features and segmental duration. Thus, as Russian and Finnish prosody are known to differ greatly, it is very likely that this

is explained by displaced foreign accent detection proposed by Major (2007), meaning that the listeners are making their judgments based on their knowledge of their mother tongue or language(s) they are learning. In addition, it is possible that the listener ratings were based on how proficient they were in Finnish, as could have been the case for G2 in our study. Our findings are in line with Riney et al. (2005), as native listeners (G1) paid the most attention to segmental features. Both Russian listener groups (G2 and G3), responded similarly to each other: they paid attention to prosody, but also to segmental quality and vowel /consonant duration. However, based on our findings it seems that both prosody and segments contributed equally to the judging of whether the speaker's mother tongue was Russian or not. Speech rate also received some attention at the first recording session by the listeners who were Finnish L2 learners.

As the study included only three speakers that were followed during one year, we can only draw tentative conclusions about pronunciation learning. We can say that we identified individual differences in "learning to sound less Russian" as perceived by naive listeners. Speakers Rus1 and Rus2 showed a tendency to have less Russian accent over the whole observation period, whereas Rus3 did so during the first six months (from T1 to T2), but then declined during the last six months (from T2 to T3). The decline may be explained by the fact that Rus3 interrupted her Finnish studies after T2 for family reasons and that her Finnish use changed at the same time, from rather frequent to a lot less frequent, and she started using more Russian at home.

From the pedagogic point of view, the tentative results of this study imply that participating in a language course while residing in the country where the language to be learnt is spoken can diminish features of foreign accent in speech. In this limited sample ( $n = 3$ ) of Russian learners of Finnish, formal instruction in Finnish seemed to make a difference in how successfully listeners recognized Russian speakers. As was mentioned above, speaker Rus3, who dropped out of the language course during the data collection, was more often recognized as Russian after dropping out than before.

There are a number of possible directions for the future longitudinal studies on assessment of foreign accented speech, as the extant literature covers only limited issues. First, it would be important to record a great number of speakers during the first phase of data collection, as it is impossible to control for background variables, e.g. participation in language courses, interaction with native speakers and it can be challenging to record the same speakers multiple times over time without having to sacrifice ideal conditions for phonetic data collection. In the present study only three out of the ten Russian speakers originally recorded at the first recording session were available for two other consequent recording sessions conducted in the sound proof studio over one year.

Second, acoustic measurements would complete the results from the auditory perception task. It would be interesting, for example, to acoustically measure the perceived pronunciation features reported by the raters that contribute to the perception of foreign accent. Similarly as for the auditory perception task, the acoustic analysis also requires more speakers in order to be able to draw generalizable conclusions. Third, more raters with no knowledge of Finnish could be recruited for such a study. Fourth, in the experiment, we did not ask the listeners with no proficiency in Finnish (G3) to determine the basis of their judgments for each stimulus, as we considered this too demanding a task in a language completely unknown to them. In future studies, however, it would be interesting to give the same multiple-choice questionnaire to all listener groups. Such an experiment would allow a more detailed comparison between the listener groups.

There is still a great deal of research to be undertaken in the longitudinal study of L2 pronunciation to improve our understanding of the potential general trends in pronunciation learning, as much of the scarce previous longitudinal studies (e.g. Ullakonoja 2011; Derwing & Munro 2013) have concluded that individual differences in learning paths are great.

## Acknowledgements

We wish to thank the Proof-project (funded by the Academy of Finland) for providing the data for the auditory experiment as well as the Tell Me project (funded by Tekes) and FOKUS-project (funded by Society of Swedish Literature in Finland) for enabling the authors to work on this paper. Moreover, we wish to acknowledge the help of the late docent Nina Volskaya (from St. Petersburg State University) for her help with data collection.

## References

- Abrahamsson, N. & K. Hyltenstam 2009. Age of onset and nativelikeness in a second language: listener perception versus linguistic scrutiny. *Language Learning*, 59 (2), 249–306. DOI: 10.1111/j.1467-9922.2009.00507.x.
- Aho, E. & M. Toivola 2008. Venäläisten maahanmuuttajien suomen prosodiasta. *Virtittäjä*, 112 (1), 3–23. <https://journal.fi/virtittaja/article/view/40638>.
- Aho, E., M. Toivola, F. Karlsson & M. Lennes 2016. Aikuisten maahanmuuttajien suomen ääntämisestä. *Puhe ja kieli*, 36 (2), 77–96. <http://hdl.handle.net/10138/168735>.
- Ahola, S. & H. Tossavainen 2016. Helppoa vai vaikeaa arvioida? Yleisten kielitutkintojen suomen kielen arvioijien käsityksiä vironkielisten arvioinnista. *Lähivördlusi. Lähivertailuja*, 26, 27–53. DOI: 10.5128/LV26.01.
- Bondarko, L. V. 1998. *Fonetika sovremennogo russkogo äzyka*. St Petersburg: Izdatel'stvo Sankt-Peterburgskogo universiteta.

- Bondarko, L. V. 2009. Short description of Russian sound system. In V. de Silva & R. Ullakonoja (eds) *Phonetics of Russian and Finnish: general description of phonetic systems, experimental studies on spontaneous and read-aloud speech*. Frankfurt am Main: Peter Lang, 23–35.
- Bongaerts, T., S. Mennen & F. van der Slik 2000. Authenticity of pronunciation in naturalistic second language acquisition: the case of very advanced late learners of Dutch as a second language. *Studia Linguistica*, 54 (2), 298–308. DOI: 10.1111/1467-9582.00069.
- CEFR 2001. *Common European framework of reference for languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Cunningham-Andersson, U. & O. Engstrand 1989. Perceived strength and identity of foreign accent in Swedish. *Phonetica*, 46 (4), 138–154.
- Derwing, T. M. & M. J. Munro 2001. What speaking rates do non-native listeners prefer? *Applied Linguistics*, 22 (3), 324–337. DOI: 10.1093/applin/22.3.324.
- 2013. The development of L2 oral language skills in two L1 groups: a 7-year study. *Language Learning*, 63 (2), 163–185. DOI: 10.1111/lang.12000.
- Flege, J. E. 1981. The phonological basis of foreign accent: a hypothesis. *TESOL Quarterly*, 15 (4), 443–455. DOI: 10.2307/3586485.
- Gilbert, J. 1980. Prosodic development: some pilot studies. In R. C. Scarella & S. D. Krashen (eds) *Research in second language acquisition: selected papers of the Los Angeles second language acquisition research forum*. Rowley (Mass.): Newbury House, 110–117.
- Gonet, W. & G. Pietroń 2004. The Polish tongue in the English ear. In E. Waniek-Klimczak & W. Sobkowiak (eds) *Zeszyty naukowe państwowej wyższej szkoły zawodowej w Koninie*. Konin: Wydawnictwo PWSZ, 56–65.
- Gupta, A. 2005. Inter-accent and inter-cultural intelligibility: a study of listeners in Singapore and Britain. In D. Deterding, A. Brown & E. Low (eds) *English in Singapore: phonetic research on a corpus*. New York (N. Y.): McGraw-Hill, 138–152.
- Iivonen, A. 1978. Is there interrogative intonation in Finnish? In E. Gårding, G. Bruce & R. Bannert (eds) *Nordic prosody: papers from a symposium*. Travaux de l'Institut de Linguistique de Lund XIII. Lund: Department of Linguistics, University of Lund, 43–53.
- 1998. Intonation in Finnish. In D. Hirst & A. Di Cristo (eds) *Intonation systems: a survey of twenty languages*. Cambridge: Cambridge University Press, 311–327.
- 2009a. Finnish sentence accent and intonation. In V. de Silva & R. Ullakonoja (eds) *Phonetics of Russian and Finnish: general description of phonetic systems, experimental studies on spontaneous and read-aloud speech*. Frankfurt am Main: Peter Lang, 67–73.
- 2009b. Major features of standard Finnish phonetics. In V. de Silva & R. Ullakonoja (eds) *Phonetics of Russian and Finnish: general description of phonetic systems, experimental studies on spontaneous and read-aloud speech*. Frankfurt am Main: Peter Lang, 47–65.
- Jaakkola, M. 2006. Suomalaisten suhtautuminen ulkomaiseen työvoimaan ja pakolaisiin. *Kuntoutus*, 29 (1), 5–21.
- 2009. *Maahanmuuttajat suomalaisten näkökulmasta: asennemuutokset 1987–2007*. Helsinki: Helsingin kaupungin tietokeskus.

- Jenkins, J. 2000. *The phonology of English as an international language*. Oxford: Oxford University Press.
- Kennedy, S. & P. Trofimovich 2008. Intelligibility, comprehensibility, and accentedness of L2 speech: the role of listener experience and semantic context. *The Canadian Modern Language Review*, 64 (3), 459–489. DOI: 10.3138/cmlr.64.3.459.
- Leinonen, A. 2015. "Riittää kun saa selvää": vieraalla aksentilla tuotettu suomi nuorten arvioimana. *Jyväskylä Studies in Humanities* 275. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-6448-1>.
- Lippi-Green, R. 2012. *English with an accent: language, ideology, and discrimination in the United States*. 2nd ed. New York (N. Y.): Routledge.
- MacKay, I. R. A., J. E. Flege & S. Imai 2006. Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics*, 27 (2), 157–183. DOI: 10.1017/S0142716406060231.
- Major, R. C. 2007. Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition*, 29 (4), 539–556.
- McCullough, E. 2013. Perceived foreign accent in three varieties of non-native English. *Ohio State University Working Papers in Linguistics*, 60, 51–66.
- Munro, M. J. 2008. Foreign accent and speech intelligibility. In J. G. Hansen Edwards & M. L. Zampini (eds) *Phonology and second language acquisition*. Amsterdam: John Benjamins, 193–218.
- Munro, M. J. & T. M. Derwing 1999. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49 (s1), 285–310. DOI: 10.1111/j.1467-1770.1995.tb00963.x.
- 2015. A prospectus for pronunciation research in the 21st century: a point of view. *Journal of Second Language Pronunciation*, 1 (1), 11–42. DOI: 10.1075/jslp.1.1.01mun.
- Munro, M. J., T. M. Derwing & S. Morton 2006. Mutual intelligibility of L2-speech: language learning. *Studies in Second Language Acquisition*, 28 (1), 111–131. DOI: 10.1017/S0272263106060049.
- Pinget, A.-F., H. R. Bosker, H. Quené & N. H. De Jong 2014. Native speakers' perceptions of fluency and accent in L2 speech. *Language Testing*, 31 (3), 349–365. DOI: 10.1177/0265532214526177.
- Piske, T., I. R. A. MacKay & J. E. Flege 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29 (2), 191–215. DOI: 10.1006/jpho.2001.0134.
- Proof 2012. Pronunciation of Finnish by immigrants in Finland, version 1.0. <http://urn.fi/urn:nbn:fi:lb-20140730130> [accessed 1/5/2018].
- Riney, T. J., N. Takagi & K. Inutsuka 2005. Phonetic parameters and perceptual judgments of accent in English by American and Japanese listener. *TESOL Quarterly*, 39 (3), 441–466. DOI: 10.2307/3588489.
- Saito, K., P. Trofimovich & T. Isaacs 2017. Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: a validation and generalization study. *Applied Linguistics*, 38 (4), 1–25. DOI: 10.1093/applin/amv047.
- Scovel, T. 1969. Foreign accents, language acquisition, and cerebral dominance. *Language Learning*, 19 (3–4), 245–253. DOI: 10.1111/j.1467-1770.1969.tb00466.x.
- Statistics Finland 2017. Statistics Finland's databases. Helsinki: Statistics Finland. [http://pxnet2.stat.fi/PXWeb/pxweb/en/StatFin/StatFin\\_\\_vrm\\_\\_vaerak/statfin\\_](http://pxnet2.stat.fi/PXWeb/pxweb/en/StatFin/StatFin__vrm__vaerak/statfin_)

vaerak\_pxt\_010.px/?rxid=22bbo37-029e-4doe-ac53-69dda024ae3c [accessed 5/4/2017].

- Toivola, M. 2011. *Vieraan aksentin arviointi ja mittaaminen suomessa*. Helsinki: Helsingin yliopisto. <http://urn.fi/URN:ISBN:978-952-10-7217-8>.
- Toivola, M., M. Lennes & E. Aho 2009. Speech rate and pauses in non-native Finnish. In M. Uther, R. Moore & S. Cox (eds) *Interspeech-2009*. Brighton: ISCA, 1707–1710.
- Toivola, M., M. Lennes, J. Korvala & E. Aho 2011. A longitudinal study of speech rate and pauses in non-native Finnish. In K. Dziubalska-Kotaczyk, M. Wrembel & M. Kul (eds) *New sounds 2010: proceedings of the sixth international symposium on the acquisition of second language speech*. Poznań: Adam Mickiewicz University, 499–504.
- Trofimovich, P. & W. Baker 2006. Learning second language suprasegmentals: effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28 (1), 1–30. DOI: 10.1017/S0272263106060013P.
- Ullakonoja, R. 2011. *Da. Eto vopros! Prosodic development of Finnish students' read-aloud Russian during study in Russia*. Jyväskylä Studies in Humanities 151. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-4209-0>.
- Ullakonoja, R., H. Dufva, M. Kuronen & P. Hurme 2014a. How to imitate an unknown language? Russians imitating Finnish. In K. Jähi & L. Taimi (eds) *XXVIII Fonetiikan päivät: Turku 25.–26. lokakuuta 2013*. Turku: University of Turku, 10–18. [https://www.utu.fi/fi/sivustot/fp2013/Documents/XXVIII\\_Fonetiikan\\_p%C3%A4iv%C3%A4t\\_konferenssijulkaisu.pdf](https://www.utu.fi/fi/sivustot/fp2013/Documents/XXVIII_Fonetiikan_p%C3%A4iv%C3%A4t_konferenssijulkaisu.pdf) [accessed 3/19/2017].
- Ullakonoja, R. & M. Kuronen 2015. Young Russian immigrants' segmental duration and length in Finnish. In The Scottish Consortium for ICPhS 2015 (ed.) *Proceedings of ICPhS 2015: 18th international congress of phonetic sciences, 10–14 Aug 2015 Secc Glasgow Scotland UK*. Glasgow: University of Glasgow. <https://www.internationalphoneticassociation.org/icphs-proceedings/ICP>.
- Ullakonoja, R., M. Kuronen, P. Hurme & H. Dufva 2014b. Segment duration in Finnish as imitated by Russians. In N. Campbell, D. Gibbon & D. Hirst (eds) *Social and linguistic speech prosody: proceedings of the 7th International conference on speech prosody*. Dublin: Science Foundation Ireland & International Speech Communication Association, 507–511.
- Uzal, M., T. Peltonen, M. Huottilainen & O. Aaltonen 2015. Degree of perceived accent in Finnish as a second language for Turkish children born in Finland. *Language Learning*, 65 (3), 477–503. DOI: 10.1111/lang.12126.
- Volskaya, N. 2009. Aspects of Russian intonation. In V. de Silva & R. Ullakonoja (eds) *Phonetics of Russian and Finnish: general description of phonetic systems, experimental studies on spontaneous and read-aloud speech*. Frankfurt am Main: Peter Lang, 34–46.
- Walker, R. 2011. *Teaching the pronunciation of English as a lingua franca*. Oxford: Oxford University Press.
- Weber, A. & K. Pöllmann 2011. Identifying foreign speakers with an unfamiliar accent or in an unfamiliar language: proceedings of the sixth international symposium on the acquisition of second language speech. In K. Dziubalska-Kotaczyk, M. Wrembel & M. Kul (eds) *New sounds 2010: proceedings of the sixth international symposium on the acquisition of second language speech*. Poznań: Adam Mickiewicz University, 536–541.

Wilkerson, M. E. 2010. Identifying accent in German: a comparison of native and non-native listeners. *Die Unterrichtspraxis/Teaching German*, 43 (2), 144-153.

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 277–300.

**Melike Uzal, Erkki Komulainen & Olli Aaltonen**

University of Helsinki

## **The effect of some listener background factors and task type that contribute to degree of perceived accent ratings in L2 Finnish**

This study evaluated the effect of some listener background factors—the listeners' gender, age, experience of teaching Finnish as a second language, frequency of contact with immigrants, and being a native of the Helsinki Metropolitan Area in Finland— and task type on their degree of perceived accent (DPA) ratings in L2 Finnish. The participants were 31 native-Finnish speakers and 40 nonnative speakers of Turkish origin who ranged in age from 7 to 66 as well as 61 Finnish listeners who evaluated the speech samples for a foreign accent using a 9-point scale. Three speech samples were administered (word pairs, the reading-aloud of sentences, and a spontaneous speech task). The results showed that no marked differences were observed, despite differences in the listeners' gender, age, Finnish as a L2 teaching experience, frequency of contact with immigrants speaking Finnish as an L2, and being a native of the Helsinki Metropolitan Area. The results also showed that ratings of native decreased with sample duration and extemporaneity.

**Keywords:** perceived accent; L2 phonology; listener background factors; task type effect



## 1 Introduction and previous research on task types used in foreign accent studies

Much research on second language (L2) phonological acquisition has revealed many factors that can affect the perception of foreign accent in a L2 (see Piske et al. 2001 and Moyer 2013 for a review). Of these factors, those that concern characteristics specific to speakers are termed **speaker-internal factors**, whereas those that concern characteristics not specific to speakers are termed **speaker-external factors**. Studies have found that the perception of foreign accent in L2 speech is affected by speaker-external factors, such as **elicitation techniques** and **listener characteristics** (Levi & Winters 2005). For instance, a few studies (Flege 1984; Piper & Cansin 1988; Munro & Derwing 1994) reported that the use of different elicitation techniques (extemporaneous versus read speech) was not a significant factor in relation to the degree of perceived foreign accent. In this respect, Munro & Derwing's (1994) study is revealing because they demonstrated that when extemporaneous narratives the speakers had produced before were compared to the read transcriptions of their extemporaneous narratives, no differences were found in their accentedness ratings. In contrast, Moyer (1999) detected a trend toward a higher number of native ratings for more isolated task items, so that reading isolated words received the lowest accent ratings, implying the highest native-like speech, which was followed by sentence reading, paragraph reading and free speech production. However, as no nonnative speaker scored significantly better on any particular task, this trend was not statistically significant.

Nevertheless, according to a study by Ishida (2004), this trend of more isolated task items receiving the highest number of native ratings was found to be statistically significant in that the discourse-length tasks (the reading-aloud of a paragraph and picture description) were rated to be more nonnative-like than the more isolated tasks (the reading-aloud of words and sentences). That is, based on previous studies by Ishida (2004) and Moyer (1999), one direction emerges: the more complex a task (from isolated words, to sentences, paragraphs, and picture description) is, the more nonnative-like learners' performances become. Again, however, contradictory findings to this generalization have also been reported. For instance, Oyama (1976), Snow & Hoefnagel-Höhle (1982), Thompson (1991), Munro & Mann (2005), and Toivola (2011) determined that spontaneous speech received the least accented ratings, implying the highest native-like speech when compared to read speech. The reason for the results of Munro & Mann (2005), however, might be their use of read speech samples that were artificially-created, pre-scripted, (words, sentences and a paragraph) and developed to be particularly difficult for Mandarin speakers of English. The same applies to the study by Thompson (1991), which incorporated the reading-aloud tasks of sentences and a passage seeded with

difficult sounds, and free oral production regarding the participants' daily routine. Thompson's (1991) results indicated that adult L2 learners were rated as most accented in the reading aloud task of sentences and least accented in the free oral production task. In short, when speech samples are seeded with difficult sounds for nonnative speakers, it inevitably results in a more pronounced accent (Munro & Derwing 1994). Moreover, as Larson-Hall (2006) argues, L2 speakers might purposely avoid the use of problematic sounds when they produce spontaneous speech samples. Another important point is that all of these studies except Snow & Hoefnagel-Höhle (1982) used reading tasks of paragraphs or sentences instead of a repetition task. When reading aloud, differences in reading ability cause the read speech of weaker readers to sound more accented and detection of their accent becomes easier (Piske et al. 2001). Reading aloud also relies on different skills than spontaneous speech and repetition, such as a higher degree of monitoring as well as literacy skills. Another factor may be related to orthographic differences in script or grapheme phoneme correspondence between the first language (L1) and the target language (Schmid & Hopp 2014). Consequently, another reason for the L2 speakers' accent being more marked for read speech samples in these studies might be differences in their reading ability. Munro & Mann (2005) identified read sentences as the most effective sampling type to rate accentedness, and considered extended speaking tasks as more representative of real communication than single words, but of course lexical, suprasegmental, and morpho-syntactic features inherent in such tasks could influence accentedness ratings. Overall, task-based effects for accentedness ratings in L2 speech have been documented in past research, but past research has provided conflicting results concerning the direction of these effects. This strongly suggests a need for additional research. Also, since there is no consensus as to the best sampling type that serves as a research instrument (Piske et al. 2001), this has resulted in the use of different speech samples in previous research to measure accentedness in L2.

### **1.1 Some listener background factors affecting degree of perceived accent ratings in an L2**

Previous studies have also reported that the perception of foreign accent is affected by listener background factors (Levi & Winters 2005). However, due to conflicting research results, there is no consensus on which listener background factors affect listeners' degree of perceived accent (DPA) ratings or on the strength of their impact (Kang 2012). This points to a need for additional research to address the degree to which some listener background factors affect DPA ratings. For instance, McDermott (1986) concluded that listener background factors were not associated with any significant differences in the

**overall** assessment of foreign accent. Of the many listener background factors McDermott (1986) tested, age, sex and ethnic diversity of the neighbourhood of the listeners predicted accentedness ratings, with younger, male raters who live in ethnically homogeneous all-English-speaking neighbourhoods giving the strictest ratings (Schmid & Hopp 2014). Likewise, Toivola (2011) also found that listener background factors (age, gender, studying, frequency of contact with immigrants, Finnish as a L2 teacher status, being a native of the Uusimaa region in Finland) had no effect on listeners' accent ratings. In contrast, Kang (2008) discovered that accentedness ratings were affected by listeners' native speaker status, exposure to nonnative speakers (NNSs), training status, prior teaching experience, and attitudes toward accented English.

The following listener background factors are within the scope of the present study: the listener's age, gender, experience of the target language (teacher experience), foreign accent experience in the form of frequency of contact with immigrants, and being a native of the Helsinki Metropolitan Area in Finland. Few studies have been conducted on the relationship between listeners' age and their foreign accent ratings (Toivola 2011). To date, it seems that only McDermott (1986), Scovel (1981) and Toivola (2011) have analyzed the effect of listeners' age on accent ratings. McDermott (1986) found listener age to predict accentedness ratings, Toivola (2011) reported no significant effect for listener age and Scovel (1981) discovered that 5-10-year-old children were unable to perceive foreign accent as well as older children and adults. Similarly, very few studies have been also conducted on the effect of listeners' gender on accent ratings. Currently, it seems that only McDermott (1986) and Toivola (2011) have investigated this factor, Toivola (2011) reporting that gender had no effect on ratings. However, several previous studies have analyzed listeners' experience of the target language (limited specifically to teacher experience), and foreign accent experience and the results of these studies are discussed below.

### **1.1.1 Listeners' experience of being a language teacher and listeners' foreign accent experience**

Studies on foreign accent vary in their operationalization of listener's **experience** of the target language. The literature offers various definitions of listeners' experiences of the target language, including formal training in language and linguistics (phonetics), familiarity with foreign-accented speech, or experience of language tutoring and teaching. As a consequence, when comparing the results of different studies, it is crucial to examine how the listeners' experience of the target language has been operationalized. In the context of this study, listener experience of the target language was operationalized as the effect of language tutoring and teaching experience. Hence, the construct

was measured using the dimension of teaching versus non-teaching. Let us now turn to discuss previous studies in which listeners' experience of the target language was similarly defined.

The literature on the effect of listeners' experience of the target language on accent ratings has produced study findings that have varied substantially. For instance, studies that have used both experienced teaching listeners and inexperienced non-teaching listeners have, on the one hand, determined that non-teaching listeners gave harsher accent ratings than teaching listeners (Barnwell 1989; Bongaerts et al. 2000; Kang & Rubin 2009; Hsieh 2011). For example, Hsieh's (2011) study found that when rating the foreign accents of international teaching assistants, the non-teaching American undergraduate listeners were more severe than experienced ESL teachers. On the other hand, Kang (2008) reported that teaching listeners with teaching experience in English as a second language were harsher in their accent ratings. In contrast to these findings, Bongaerts et al. (1997), Kennedy & Trofimovich (2008) and Toivola (2011) did not detect significant differences between the accentedness ratings of teaching and non-teaching listeners.

When it comes to the literature on the effect of listeners' foreign accent experience on accent ratings, Kennedy & Trofimovich (2008) demonstrated that the experience of nonnative speech did not result in harsher ratings of accentedness by seasoned ESL teachers and they did not rate accentedness differently from inexperienced non-teaching listeners who had had little or no contact with L2 speakers of English. Likewise, the teaching listeners in the study by Toivola (2011), who were teachers of Finnish as an L2 with frequent or daily contact with nonnative speakers of Finnish, did not rate accentedness differently from non-teaching listeners who had frequent, rare, or nonexistent contact with nonnative speakers of Finnish. Furthermore, in a study by Munro et al. (2010), native listeners' familiarity with foreign-accented speech through regular contact with nonnative speakers of English was not determined to have had a statistically significant effect on their accent ratings. In contrast, Thompson (1991) and Schmid & Hopp (2014) concluded that perception of a foreign accent depended on the listeners' familiarity with foreign-accented speech. Thompson (1991) found that listeners' linguistic experience had an effect on their ratings so that linguistically experienced listeners with frequent contact with nonnative speakers of English were more lenient in their ratings when compared to linguistically inexperienced listeners who had little or no contact with nonnative speakers of English. It should be noted, nonetheless, that Thompson's (1991) experienced listeners were not L2 teachers but language experts who had elected courses in linguistics. Moreover, Schmid and Hopp's (2014) findings showed that listeners' lower familiarity with foreign accents resulted in more variable and more strongly foreign-accented judgments.

## 2 The present study

### 2.1 The purpose of the study and research questions

The present article is part of a study by Uzal et al. (2015) and the purpose of that study was to explore the effect of some speaker background factors —the speakers' age of onset of extensive L2 acquisition (AO), L1 use, L2 use, home use of L1, L1 proficiency, L2 proficiency, the amount of L2 exposure indexed as length of residence in the target language country, age at the time of testing — on the degree of perceived accent ratings for child L2 learners of Finnish. The findings of Uzal et al. (2015) showed that AO was the main determiner of perceived accent, followed by home use of L1, and the amount of L2 and L1 use, confirming the salience of both age-related factor of AO and language experience factors in determining child L2 learners' foreign accent. The focus of this article is the effect of some listener background factors and task type. In short, the present study has two purposes. The first is to discover how some listener background factors affect listeners' perception of a foreign accent. The factors examined in the present study are the listeners' age, gender, Finnish-as-a-L2 teacher status, frequency of contact with immigrants, and being a native of the Helsinki Metropolitan Area in Finland. The second purpose is to identify the effect of three different speech samples (word pairs, the reading-aloud of sentences, and a spontaneous speech task) on the degree of perceived accent ratings.

The research questions of this article are the following:

1. What is the contribution of some listener background factors (listeners' gender, age, Finnish-as-a-L2 teacher status, frequency of contact with immigrants, and being a native of the Helsinki Metropolitan Area in Finland) to the degree of perceived accent ratings in L2 Finnish?
2. How do sampling effects (speech sample duration and speech sample extemporaneity) affect the degree of perceived accent ratings in L2 Finnish?

Regarding the first research question, due to the inconclusive research results in the literature, no stand is taken on the effects of listener background factors. For the second research question, the hypothesis is that different types of speech samples would cause listeners to give significantly different foreign accent ratings based on previous research. However, for the present analysis, it was preferred not to take a stand on how the effects of sample duration (sample duration scale: word pairs < single sentences < spontaneous speech) and extemporaneity of the speech samples (extemporaneity scale: scripted word pairs = scripted single sentences < extemporaneous sponta-

neous speech) would present themselves, as previous studies have reported conflicting findings.

## 2.2 Methodology

### 2.2.1 Data analysis

Reliability must always be estimated as it is a necessary (but not sufficient) condition to the validity of measurements. There are no valid measurements without a sufficient amount of reliability. It is a rule of thumb that reliability should be 0.7 or higher (Nunnally & Bernstein 1994). The problem is that there is a variety of reliability indices. Here, four were chosen:

1. parallel measures (items) approach
2. mean correlation between the listeners
3. intra-class-correlation (ICC, two-way mixed, single measurement) and
4. ICC two-way mixed, average measures

Differences in listener/speaker background factors were examined using ordinary one-way analysis of variance (ANOVA). Overall significance was expressed via F-ratio. Explanation of the variance in ratings analysis included initially the linear mixed model (LMM) approach. This analysis was expanded later to include the analysis of the generalized linear mixed model (GLMM) as well. Both methods allow to model the correlative structure of the data. The distribution of the response variable (foreign accent rating) was very strongly positively skewed (see Figure 2). Only GLMM fulfills both of the requirements: structure of the data and the way to treat the response variable distribution as ranked categories and use the link function  $-2\text{LogLog}$ . This model is usually called ordered categories mixed model regression. LMM and GLMM analyses produced almost the same results. This strengthens the reliability of the results. Also, Cohen's  $d$  was used as an index of effect size. It is simply the difference of two means divided by the pooled standard deviation of the groups compared.

### 2.2.2 Speakers

Speakers from four different groups were selected: child nonnative speakers (NNSs), adult NNSs, child native speakers (NSs) and adult NSs as controls. The part of the speaker data concerning the foreign accent ratings of 19 child NNSs and 20 child NSs as native control speakers comes from Uzal et al. (2015). Unlike Uzal et al. (2015), 32 more speakers were added to this study: 9 child NNSs, 11 adult NSs, and 12 adult NNSs. Thus, of the 71 speakers, 31 were NSs

of Finnish and 40 were NNSs of Finnish. All the native-Finnish controls were born and raised in Finland, lived in the Helsinki Metropolitan Area, and spoke standard Finnish. They consisted of 20 female speakers and 11 male speakers aged 7–39 ( $M = 17$  years). All 11 adult NSs were recruited students from the University of Helsinki. All 20 child NSs were pupils at a comprehensive school in the Helsinki Metropolitan Area. The 20 child NSs (age at the time of testing:  $M = 11$ ,  $SD = 3$ , range = 7–17) were matched to the 28 child NNSs according to their age at the time of testing ( $M = 11$ ,  $SD = 3$ , range = 7–17). The ANOVA result showed that there was a statistically significant difference for the mean age at time of testing between the 12 adult NNSs (range = 27–66) and the 11 adult NSs (range = 22–39) ( $F(1, 21) = 7.39$ ,  $p = 0.013$ ), the mean being 38 and 28 years, respectively. That is, adult NNSs were significantly older when compared to adult NSs.

The NNSs were all NSs of Turkish from a wide variety of Turkish cities, including 22 female speakers and 18 male speakers aged 7–66 ( $M = 19$  years). All 12 adult NNSs were acquaintances of the first author. All 28 child NNSs were pupils at various primary and secondary schools in the Helsinki Metropolitan Area, and most of them had received instruction in Turkish as a mother tongue two hours a week from the first author. To summarize, all 71 speakers resided in the Helsinki Metropolitan Area and spoke standard Finnish and they therefore fulfilled the most crucial requirement for speaker selection (standardized dialect; Long 1993).

### 2.2.3 Speech samples

The data sources, the speech sample collection and methods in the present study were predominantly the same as in Uzal et al. (2015). The differences were that the present study had 32 more speakers and an additional spontaneous speech sample task unlike Uzal et al. (2015). All speakers were assigned a task involving the repetition of eight sentences, from which five sentences and three word pairs were used for rating. The three word pairs were obtained from the remaining three sentences. The actual foreign accent rating task incorporated five sentences, three word pairs—obtained by extracting each word pair from the remaining three sentences—and spontaneous speech. Therefore, for the rating task, three types of speech samples were obtained from the adults (sentences, word pairs, spontaneous speech) and two types of speech samples (word pairs and sentences) from the children. All of these speech samples differed in duration: word pairs, sentences, and a 40-second passage of spontaneous speech on a topic selected by the speakers from three options. Also, all of these speech samples differed in extemporaneity: read word pairs, read sentences and the extemporaneous speech sample of spontaneous speech. Due to the children's short attention spans,

the scripted speech samples (word pairs and single sentences) were designed to be short and simple and contained the entire phonetic inventory of Finnish.

The speaker recordings for all children were made in empty school classrooms in the Helsinki Metropolitan Area, whereas for all adult speakers, the recordings were made in the soundproof recording studio at the University of Helsinki. A model voice of a female, monolingual, native Finnish adult was recorded in a soundproof recording studio beforehand, and this was presented to the speakers through a computer loudspeaker; this recording represented the spoken standard Finnish norm (Karlsson 2008). The speech materials of the five sentences and the three word pairs have been presented in Uzal et al. (2015). For instance, one of the five sentences presented to repeat was “Kotona on pöytä ja pöytävalaisin”, whereas one of the three sentences presented to repeat was “Kotona on porkkana ja tomaatti”. Then, “kotona on” part was deleted from this sentence and the remaining “porkkana ja tomaatti” word pair was presented for rating. Sentences were presented to speakers simultaneously in both written and aural form to reduce the potential of reading ability biasing accent ratings (Flege et al. 1995). Each sentence was presented once, followed by a silent six-second delay. The six-second delay was followed by a beep, after which the speakers repeated each sentence once. This six-second delay was intended to minimize the possible effect of direct imitation (Tench 1996). If the speakers could not produce a sentence or had forgotten the sentence altogether, the model was presented as many times as necessary to obtain a production without speech irregularities. Thus, the speakers had the opportunity to correct their productions; however, they were not allowed to practice beforehand. After this, unlike Uzal et al. (2015), an additional task of producing spontaneous speech was assigned to all the 11 NS adults and to 10 of the 12 NNS adults<sup>1</sup>. The spontaneous speech instructions have been presented in Uzal et al. (2017). The speakers were instructed to discuss one of the three subjects (or make a subject up themselves). The speakers’ spontaneous speeches were recorded for 1 minute and the first 40-second segments were presented for rating. One of the three options for them to discuss was to tell their weekend or their daily routine (e.g. what do you usually do, when, with whom, for how long, what is interesting about it, etc.?). All the recordings were made on a Marantz PMD 660 digital audio recorder with a power microphone. The duration of the recordings ranged from 15 to 25 minutes for each speaker.

---

<sup>1</sup> As two NNS adults reported that they could not produce spontaneous speech due to their very poor oral language skills in Finnish, these speakers’ spontaneous speech samples were missing. That is, these two speakers who were both beginner level Finnish speakers were not recorded because they informed that they do not trust their oral language skills in Finnish enough to produce spontaneous speech in Finnish.



### 2.3 Listeners

The source of the listener data concerning the foreign accent ratings of 19 child NNSs and 20 child NSs as native control speakers is from Uzal et al. (2015). The requirement for the 61 native listeners recruited was that they be monolingual native speakers of Finnish and residents of the Helsinki Metropolitan Area. These listeners were subdivided into two groups according to their background education and teaching experience: non-teaching listeners (34) and teaching listeners (27). This ensured a balance in listeners' sensitivity to accents (Moyer 2007). Thus, the non-teaching listeners were NSs of Finnish with no experience of linguistics and phonetics, while the teaching listeners were linguistically experienced NSs of Finnish and were all teachers of Finnish as an L2. In addition, all the non-teaching listeners were required to be monolingual native speakers of standard Finnish, natives of the Helsinki Metropolitan Area, and residents of the Helsinki Metropolitan Area. All teaching listeners were also required to be residents of the Helsinki Metropolitan Area, but it sufficed that they were monolingual native speakers of Finnish so they did not have to be natives of the Helsinki Metropolitan Area.<sup>2</sup> Therefore, 17 of the 27 teaching listeners grew up outside the Helsinki Metropolitan Area, and 9 stated that they spoke Finnish with a regional dialect. This meant that there were 17 listeners who were not natives of the Helsinki Metropolitan Area and 44 listeners who were natives of the Helsinki Metropolitan Area. A majority of the listeners (52 of 61) spoke standard Finnish. In addition, all 61 listeners spoke Finnish as their L1 and had studied Swedish and English as their L2s. All listeners also reported normal hearing on the preliminary information form.

The gender distribution of the listeners was rather uneven (46 females, 15 males), primarily because the majority of teaching listeners (26 out of 27) were female.<sup>3</sup> However, the 34 non-teaching listener group also had only 14 male listeners. The listeners' age distribution was also uneven: 13 listeners

<sup>2</sup> Speaking standard Finnish was not used as a selection criterion for the teaching listeners because they were recruited through an email sent to the Association of Teachers of Finnish as a Second Language and most importantly because it was very challenging to recruit such listeners in the first place. The ratings were all made in the University of Helsinki's soundproof recording studio, and due to this logistics of travelling, all teaching listeners were residents of the Helsinki Metropolitan Area. Consequently, 9 of the 27 teaching listeners reported that they spoke Finnish with a regional dialect. We were thus aware of the possibility that their regional accent background might affect their accent ratings. However, the ANOVA findings revealed that these 9 teaching listeners did not give statistically different accent ratings when compared to the other 18 teaching listeners (mean ratings = 3.0 and 2.5,  $F(1, 25) = 3.463$ ,  $p > 0.05$ ).

<sup>3</sup> The gender and age distributions were uneven because it was extremely challenging to recruit teaching listeners, whom were recruited by sending an email to the Association of Teachers of Finnish as a Second Language. As a result, we received answers from 26 female L2 Finnish teachers and one male teacher. This imbalance in the gender distribution of the teaching listeners also led to the age distribution being uneven.

were aged between 19 and 24, 16 listeners were aged between 25 and 29, 12 listeners were aged between 30 and 34, and 20 listeners were aged between 35 and 64. The ages of the listeners ranged from 19 to 64, and their mean age was 35. The mean age of the non-teaching listeners was 27 (range 19–44), whereas the mean age of the teaching listeners was 42 (range 25–64). In other words, the teaching listeners were older than the non-teaching listeners, and the ANOVA result showed that the difference was statistically significant ( $F(1, 59) = 27.11, p < 0.001$ ).

The listeners' foreign accent experience in the form of frequency of contact with immigrants was operationalized as the frequency of hearing the non-native Finnish speech of immigrants in their daily lives. The four reply options on the preliminary information form given to the listeners were as follows: never, rarely, often and daily. All the teaching listeners had contact with immigrants speaking Finnish as an L2 either daily ( $n = 15$ ) or often ( $n = 12$ ). A majority of the 34 non-teaching listeners reported having contact with immigrants either rarely ( $n = 24$ ) or never ( $n = 3$ ). Only a few non-teaching listeners stated that they had contact with immigrants either daily ( $n = 4$ ) or often ( $n = 3$ ). Moreover, as expected, the difference between the teaching listeners and the non-teaching listeners in terms of their frequency of contact with immigrants was statistically significant ( $F(1, 59) = 44.685, p < 0.001$ ).

### 2.3.1 Procedure

All listeners completed the rating task individually in a soundproof recording studio, where a total of 589 speech samples (71 speakers  $\times$  9 target recordings = 5 sentences + 3 word pairs + 21 adult speakers  $\times$  1 spontaneous speech recording) were presented through headsets.<sup>4</sup> Prior to the listeners' rating task, they were provided with a modified version of Toivola's (2011) preliminary information form (see Appendix A) and a short training session. First, they filled the preliminary information form and then read the instructions for the foreign accent listening test. To avoid unrelated linguistic factors affecting the DPA ratings, before the rating began, the listeners were instructed by the preliminary information form to ignore all nonphonological speech content and assess only DPA.

Since the main objective was to capture listeners' unguided holistic per-

<sup>4</sup> Two different softwares were used for the rating task because it was recognized that using the Praat software, some of the samples were accidentally rated before they were fully heard. These 170 ratings were excluded from the analyses. 35 listeners did the rating task using the Praat software (Boersma & Weenink 2013), whereas the remaining 26 listeners performed the same rating task using the Presentation software (<https://www.neurobs.com/>) to prevent such premature ratings. The Presentation software forced the listeners to listen to each sample until the end before giving their rating, which they were asked to confirm before advancing to the next sample.

ceptions of foreign accent, Scovel's (1969) definition of foreign accent was adopted to guide listeners in that direction (see Appendix A). This holistic definition of foreign accent was used to avoid listeners from becoming confused during the foreign-accent-rating task with specialist linguistic terminology for the aspects of speech they were requested to rate. The listeners were also instructed by the preliminary information form to use the entire scale when rating the samples and were told to guess if they were uncertain. To help familiarize the listeners with the rating process and the range of accents, there were 30 practice speech samples of sentences (13 from Finnish children, 17 from Turkish children). These sentences were not analyzed.

A nine-point scale was used to rate accent. The listeners were instructed that they would hear productions spoken by NNSs or NSs of Finnish. They were requested to rate each production for the degree of perceived accent by pushing one of nine buttons representing a scale from one (*no foreign accent*) to nine (*very strong foreign accent*). The ratings were given in a single session that lasted between 60 and 100 minutes in three separate blocks. The listeners were allowed to take a short break in between the blocks and midway through the blocks of sentence and word pair rating. In all three types of rating tasks, the listeners were able to adjust the volume before the rating started; the same sample could be played up to five times, and the ratings could not be changed once given. The different sample types were divided so that the rating task of each block involved one sample type: only word pairs, only sentences, or only spontaneous speech. Runs were randomized within each sampling type, including speech samples and speakers. To balance ordering effects, the three blocks were presented in a randomized order.

### 3 Results

#### 3.1 Reliability analysis

To investigate inter-rater consistency, four procedures were chosen. The results of reliability analyses are best summarized by producing a pictorial view of them. Since there was only one spontaneous speech sample per a speaker, split-approach could not be used. Instead, the 9 items were treated as (a) parallel measures. Correlation of the foreign accent ratings between the listeners gave the second way (b) mean correlation. It has the advantage that the reliability of individual listeners is available. Intra-class-correlation (ICC) is widely used in studies using ratings. Therefore, (c) ICC-single measure and (d) ICC-average measure were calculated. ICC-single measure was useful, but ICC-average measure gave unrealistically high values and did not discriminate between items. These features can be seen in Figure 1. As seen in Figure 1, spontaneous speech sample type as stimulus got the highest reliability, but the

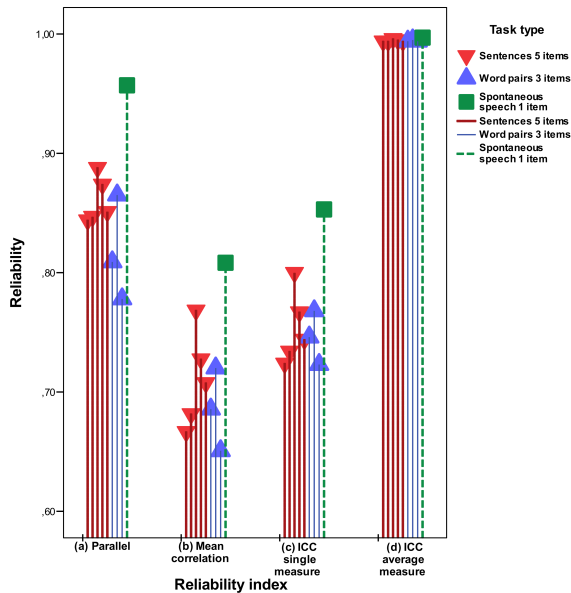


FIGURE 1. Four reliability indices of listeners' accent ratings according to three task types.

reliability differences among three task types were not big. The difference between word pairs and sentences was not very clear, but sentences had slightly higher reliability compared with word pairs. The listeners were more accurate, reliable and consistent in their rating of spontaneous speech samples compared with rating word pairs and sentences. The general conclusion was that all sample type measurements had high enough reliability to ensure validity in further analyses.

### 3.2 The effect of listener background factors on accent ratings

To assess the effects of the listeners' background factors (research question one) on accent ratings, both LMM and GLMM analyses were conducted. Initial analysis was LMM and final analysis was GLMM. Since GLMM analysis is slightly more correct approach to use with this kind of data at hand, GLMM findings were reported here. As shown in Table 1, the accent ratings were not affected by the listeners' gender, age, Finnish-as-a-L2 teacher status, foreign accent experience in the form of frequency of contact with immigrants or being a native of the Helsinki Metropolitan Area (all  $p$  values = n. s.). These five listener background factors' effect sizes (Cohen's  $d$ ) ranged from 0.01 to 0.15, and could be all characterized as very small (Cohen 1988; Sawilowsky 2003). It was therefore justified to ignore the distinction between teaching and non-

TABLE 1. Statistical significances and effect sizes of listener characteristics on accent ratings (n = 61).

Variable		M	SD	GLMM overall p	Post hoc pairs	Cohen's d of pairwise differences
Gender	Female	2.48	2.10	=0.868		-0.04
	Male	2.57	2.18			
Teaching	No	2.69	2.22	=0.109		0.14
	Yes	2.47	2.10			
Age	15-24	2.43	2.03	=0.152	1-2	-0.11
	25-34	2.67	2.24		1-3	0.04
	35-63	2.34	1.99		1-4	0.15
Contacts	Never	2.57	1.96	=0.720	1-2	0.02
	Rarely	2.52	2.10		1-3	0.07
	Often	2.41	2.06		1-4	0.01
	Every day	2.54	2.22		2-3	0.05
					3-4	-0.06
Native of H. M. A.	No	2.49	2.10	=0.414		0.02
	Yes	2.54	2.16			

Note. Contacts refers to the frequency of hearing the nonnative Finnish speech of immigrants in listeners' daily lives measured by the four reply options (never, rarely, often, daily) in the preliminary information form given to the listeners. Native of H.M.A. stands for being a native of the Helsinki Metropolitan Area in Finland.

teaching listeners, because it transpired that Finnish as a L2 teaching experience had no significant effect on foreign accent ratings.

### 3.3 The effect of speech sample types on accent ratings

To assess the effect of task type (research question two) on accent ratings, both LMM and GLMM analyses were conducted. All pairwise post hoc comparisons were Bonferroni corrected. GLMM findings were reported in Table 2. As can be seen in Table 2, two out of three pairwise LMM comparisons had statistically significant differences. The two task types differed statistically very significantly ( $p < 0.001$ ), but their effect sizes ( $d$ ) ranged from -0.43 (small) to 0.15 (very small) (Cohen 1988; Sawilowsky 2003). That is, the two task types had a very significant effect on the accent ratings. Regarding the effect of sample duration, Table 2 shows that the duration of the sample (sample duration scale: word pairs < single sentences < spontaneous speech) affected the mean accent ratings, as word pairs were rated statistically more nativelike than single sentences ( $p < 0.001$ ), though with a very small effect size. Besides, there

TABLE 2. Statistical significances and effect sizes of task type effects on accent ratings.

Task type	Focus	M	SD	p	pairs	comp.	d
1 Sentences	Dur.	2.60	2.16		1 < 2	< 0.001*	0.15
2 Word pairs	Dur., Ex.	2.28	1.95	< 0.001*	1 = 3	= 0.999	-0.25
3 Spontaneous	Dur., Ex.	3.15	2.73		2 < 3	< 0.001*	-0.43

\*. The mean difference is significant at the 0.001 level.

Note. "p" stands for GLLM overall p, "pairs" for post hoc pairs, "comp." for pairwise post hoc comparisons and "d" for Cohen's d of pairwise differences, "Dur." for duration, and "Ex." for extemporization.

was a durational trend to rate single sentences more nativelike than spontaneous speech. However, this trend of durational effect reached statistical significance only between shorter word pairs and longer spontaneous speech samples with a small effect size. Thus, there was a durational trend: the longer the sample type length, the less native the mean accent rating. The same statistical difference between spontaneous speech samples and word pairs meant that the extemporaneity of the sample (extemporaneity scale: scripted word pairs = scripted single sentences < spontaneous speech) affected the mean accent ratings as well. That is, the less extemporaneous the sampling type, the more native the mean accent rating: spontaneous speech samples were rated less nativelike than word pair samples ( $p < 0.001$ ) the effect size being small. Also, the purpose of Figure 2 was to show how extremely skewed foreign accent rating distributions were, which caused means and standard deviations to give false impressions.

## 4 Discussion

This study evaluated the effect of some listener background factors, such as gender, age, Finnish as a L2 teaching experience, frequency of contact with immigrants, and being a native of the Helsinki Metropolitan Area in Finland, and task type on the degree of perceived accent ratings in L2 Finnish. As for the first research question, the results revealed that none of these listener background factors had any statistically significant effect on accent ratings. Thus, it seems that when both speakers and listeners speak the same target language and the same variety—standard Finnish in this study—foreign accent detection is reliable and accurate, and no marked differences can be observed, despite differences in the listeners' gender, age, Finnish as a L2 teaching experience, frequency of contact with immigrants speaking Finnish as an L2, and being a native of the Helsinki Metropolitan Area. The findings established that native

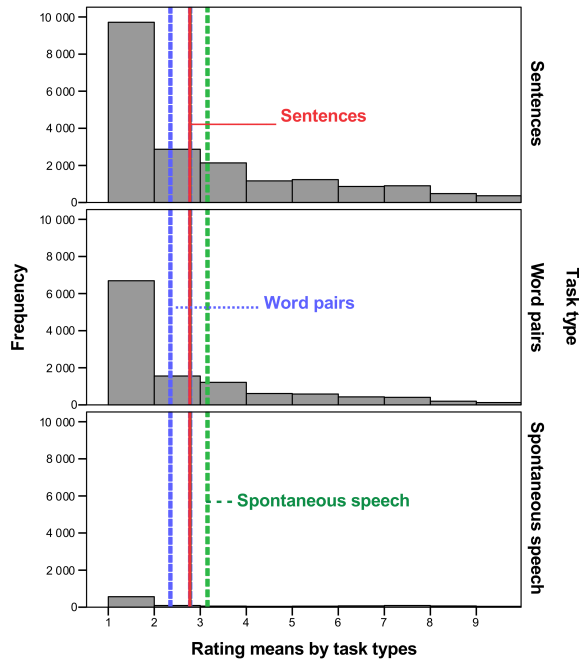


FIGURE 2. Rating distribution and mean foreign accent ratings according to three task types.

Finnish listeners were reliable and consistent in their rating of all three types of speech samples although spontaneous speech had the highest inter-rater reliability. It is interesting that the teaching listeners, who were all accustomed to hearing foreign accented Finnish either daily or often, gave similar ratings to those of the non-teaching listeners, most of whom reported hearing foreign accented Finnish either rarely or never. This finding is in line with some previous studies (Bongaerts et al. 1997; Kennedy & Trofimovich 2008; Toivola 2011, etc.). As the teachers of Finnish as an L2 transpired to be neither more compassionate nor harsher listeners than the non-teaching listeners, the results of this study contradict those of some earlier studies (e.g., for more lenient ratings, Barnwell 1989; Bongaerts et al. 2000; Kang & Rubin 2009; Hsieh 2011; for harsher ratings, Kang 2008). This finding suggests that the teaching and the non-teaching native listeners shared a similar perception regarding the degree of perceived accent despite their different experience of hearing foreign accented L2 speech and the non-teaching group having no experience of teaching Finnish as an L2.

The non-teaching listener group in the present study consisted of university students from different departments, whereas the teaching listener group

was comprised of teachers of Finnish as an L2. As no difference arose in the accent ratings between the non-teaching (student) listener group and the teaching listener group, it can be further concluded that studying per se had no effect on foreign accent ratings. Likewise, Toivola (2011) concluded that even though university students gave more ratings of strongly accented speech compared to non-studying listeners, most of whom were also L2 Finnish teachers, when rating read-aloud speech, this relationship was statistically insignificant. Furthermore, the results of the present study agree with those of Toivola (2011) in that no effect on foreign accent ratings was detected for listener's gender, age, Finnish as a L2 teaching experience, frequency of contact with immigrants, and being a native of the Helsinki Metropolitan Area as well.

As for the second research question, the results showed that perception of a foreign accent depended on the different types of speech samples presented for rating. Thus, it can be concluded that to analyze perceptions of foreign accent, the selection of speech materials is of the utmost importance due to a risk of instrumental bias. It is widely known that task differences affect findings still cited in the L2 phonology literature Moyer (2013). As Moyer (2013) rightly observes, unfortunately, the comparability of studies will continue to be an issue as long as researchers stick to using variable tasks for measuring the same construct, degree of perceived accent. This is a serious methodological problem and this task range problem documented in previous foreign accent studies was seen in this study as well: isolated, artificial and read word pairs were rated more nativelike than extemporaneous spontaneous speech samples which managed to reveal the L2 speakers' real oral performance resulting in more realistic (but harsher) foreign accent ratings. Spontaneous speech task ratings seem to reflect the L2 learners' more comprehensive assessment of oral performance in L2 when compared to word pair ratings. As Moyer (2013) points out, some late L2 learners are only expected to produce an authentic sound within an isolated task, yet they are far less convincing in real conversation where suprasegmentals and pragmatic skills come into play. Likewise, the results of this study revealed that all L2 learners' accent ratings were worse and less native on the spontaneous speech task when compared to the rather isolated task of read word pairs. This finding agrees with the results of several other studies (Moyer 1999; Ishida 2004), but conflicts with some research (as in Oyama 1976; Snow & Hoefnagel-Höhle 1982; Flege 1984; Piper & Cansin 1988; Munro & Derwing 1994; Thompson 1991; Munro & Mann 2005).

When the sample duration was examined, word pairs received significantly lower and better accent ratings than single sentences. Also, there was a trend (which did not reach significance) to rate single sentences more nativelike than spontaneous speech. All in all, there was a durational trend that the longer the sample type length was, the less native the mean accent rating



was. These durational results are consistent with the finding of the study by Ishida (2004) and Moyer (1999) in that the trend is toward a higher number of native ratings for the more isolated task items. Furthermore, this durational trend also agrees with Munro & Mann's (2005) durational effect finding that read paragraphs were the least native, followed by sentences, and then words. When the sample extemporaneity was examined, it was found out that there was a statistically systematic difference in ratings for spontaneous vs. read speech as well. That is, this study's findings showed that the less extemporaneous the speech sample was (extemporaneity scale: scripted word pairs = scripted single sentences < spontaneous speech), the more native the mean accent rating was. That is, the extemporaneous spontaneous speech sample was rated less native than scripted read word pairs. This meant that there was a significant difference between extemporaneous spontaneous speech and scripted read speech of word pairs. This effect of extemporaneity contradicts the results of Munro & Mann (2005) and Toivola (2011) on sample extemporaneity, as they reported that the more extemporaneous the sampling type is, the more native the mean accent rating (for instance, spontaneous picture narrations received more nativelike accent ratings than scripted paragraphs in the study by Munro & Mann 2005). To summarize, the findings of the present study support the general tendency in previous foreign accent studies: the longer, and less constrained the speech sample, the stronger the accent is rated (DeKeyser & Larson-Hall 2005).

## 5 Conclusion

When speaking informally with the listeners, they made it clear that even though they were instructed to ignore all other nonphonological speech content, it was very difficult for them to rate only purely phonological foreign accent while they rated spontaneous speech. The listeners told openly that they just could not simply ignore the mistakes in the choice of words, lexical errors, and grammatical errors and stated that these types of mistakes affected their accent ratings vastly. For instance, one listener provided the following comment: "If someone says 'valmistan lapset' (I make the kids ready), it is clear immediately that the speaker is nonnative". This informal observation agrees with the study by Moyer (1999), in which native listeners were influenced by structures beyond the L2 phonological production in their ratings of authenticity on a six-point scale of nativeness when the speech samples were more than mere word repetition. McDermott (1986) also noticed the same methodological effect pointing out that morphosyntactic or lexical errors can influence accentedness judgments when extemporaneous speech is used. This means that it is important that future researchers be aware of listeners' difficulties in

ignoring other linguistic factors when confronted with spontaneous speech, although it should be purely phonological variables determining accentedness ratings.

In line with Toivola's (2011) findings, teaching status did not affect the results in the present study either. However, a limitation of the present study (similar to Toivola 2011) was the fact that all 16 listeners not native to the Helsinki Metropolitan Area were teaching listeners. Out of 27 teaching listeners, 11 were natives of the Helsinki Metropolitan Area. In order to rule out possible interdependence between teaching and not being native to the Helsinki Metropolitan Area, future studies should recruit native listeners who are non-teaching and not native to the Helsinki Metropolitan Area. This would allow detection of whether not being native to or a resident of the Helsinki Metropolitan Area has an independent effect on foreign accent ratings. Moreover, although the results of both the present study and Toivola (2011) showed that frequency of contact with immigrants and being a teacher had no effect on accent ratings, it is possible that not being a native to the Helsinki Metropolitan Area has an effect on ratings when not combined with teacher status. As Toivola (2011) observes, future research should strive to discover whether native Finnish listeners' language variation affects their foreign accent rating behavior and their native speaker identification success, and thus listeners should be recruited from areas of Finland where native Finnish listeners are seldom in contact with foreigners and rarely hear foreign accented Finnish. For instance, much more heterogeneous reactions to foreign accents could be expected from Finns living in the north of the country, where native Finnish speakers are seldom in contact with nonnative speakers of Finnish. In this study even though most of the 34 nonteaching listeners reported rarely or never having contact with immigrants, the reality of being a resident of the Helsinki Metropolitan Area cannot be underestimated. The proof comes from statistics from public registers which show that at the beginning of 2016, 88,132 residents of Helsinki spoke a foreign language (other than Finnish, Swedish or Sami) as their mother tongue (Hiekkavuo et al. 2017). Almost half of Finnish residents with a foreign mother tongue live in the Helsinki Metropolitan Area, which is home to around 20 per cent of Finland's entire population. Consequently, the Helsinki Metropolitan Area is clearly the center of foreign language speakers in Finland (Hiekkavuo et al. 2017). Therefore, residents of the Helsinki Metropolitan Area are expected to be more used to hearing foreign-accented Finnish compared to native Finnish people from other regions of Finland. Also, as another future research suggestion, as one of the reviewers suggested, it would be interesting to compare the effect of listeners' highly educated status, whether educated people's and ordinary laymen's foreign accent ratings differ, on their perception of a foreign accent.

## References

- Barnwell, D. 1989. 'Naive' native speakers and judgements of oral proficiency in Spanish. *Language Testing*, 6 (2), 152–163. DOI: 10.1177/026553228900600203.
- Boersma, P. & D. Weenink 2013. PRAAT. Amsterdam: University of Amsterdam. <http://www.fon.hum.uva.nl/praat/>.
- Bongaerts, T., S. Mennen & F. van der Slik 2000. Authenticity of pronunciation in naturalistic second language acquisition: the case of very advanced late learners of Dutch as a second language. *Studia Linguistica*, 54 (2), 298–308. DOI: 10.1111/1467-9582.00069.
- Bongaerts, T., C. van Summeren, B. Planken & E. Schils 1997. Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19 (4), 447–465. DOI: 10.1017/S0272263197004026.
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale (Mich.): Lawrence Erlbaum.
- DeKeyser, R. & J. Larson-Hall 2005. What does the critical period really mean? In J. F. Kroll & A. M. B. DeGroot (eds) *Handbook of bilingualism: psycholinguistic approaches*. New York (N. Y.): Oxford University Press, 88–108.
- Flege, J. E. 1984. The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, 76 (3), 692–707. DOI: 10.1121/1.391256.
- Flege, J. E., M. J. Munro & I. R. A. MacKay 1995. Effects of age of second-language learning on the production of English consonants. *Speech Communication*, 16, 1–26. DOI: 10.1016/0167-6393(94)00044-B.
- Hiekkavuo, A., E. Haapamäki, S. Ranto & M. Salorinne 2017. *Population with foreign background in Helsinki 2016*. Helsinki: Helsingin kaupungin tietokeskus.
- Hsieh, C. 2011. Rater effects in ITA testing: ESL teachers' versus American undergraduates' judgments of accentedness, comprehensibility, and oral proficiency. *Spain Fellow Working Papers in Second or Foreign Language Assessment*, 9, 47–74.
- Ishida, S. 2004. *Taking pronunciation to task: exploring methodological issues in maturational constraints research*. A paper presented at the Annual Meeting of the American Association of Applied Linguistics, May, 1–4, Portland, Oregon, USA.
- Kang, O. 2008. Ratings of L2 oral performance in English: relative impact of rater characteristics and acoustic measures of accentedness. *Spain Fellow Working Papers in Second or Foreign Language Assessment*, 6, 181–205.
- 2012. Impact of rater characteristics and prosodic features of speaker accentedness on ratings of international teaching assistants' oral performance. *Language Assessment Quarterly*, 9 (3), 249–269. DOI: 10.1080/15434303.2011.642631.
- Kang, O. & D. L. Rubin 2009. Reverse linguistic stereotyping: measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology*, 28 (4), 441–456. DOI: 10.1177/0261927X09341950.
- Karlsson, F. 2008. *Finnish: an essential grammar*. New York (N. Y.): Routledge.
- Kennedy, S. & P. Trofimovich 2008. Intelligibility, comprehensibility, and accentedness of L2 speech: the role of listener experience and semantic context. *The Canadian Modern Language Review*, 64 (3), 459–489. DOI: 10.3138/cmlr.64.3.459.
- Larson-Hall, J. 2006. What does more time buy you? Another look at the effects of long-term residence on production accuracy of English /ɹ/ and /l/ by

- Japanese speakers. *Language and Speech*, 49 (4), 521–548. DOI: 10.1177/00238309060490040401.
- Levi, S. V. & S. J. Winters 2005. *External factors affecting the perception of foreign accent in L2 speech*. A poster presented at the 1st ASA Workshop on Second Language Speech Learning, May 14–15, 2005, Vancouver, Canada.
- Long, M. H. 1993. Second language acquisition as a function of age: research findings and methodological issues. In K. Hyltenstam & A. Viberg (eds) *Progression & regression in language: sociocultural, neuropsychological & linguistic perspectives*. Cambridge: Cambridge University Press, 196–221.
- McDermott, W. L. C. 1986. *The scalability of degrees of foreign accent*. PhD thesis. Ithaca (N. Y.): Cornell University.
- Moyer, A. 1999. Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition*, 21 (1), 81–108. DOI: 10.1017/S0272263199001035.
- 2007. Do language attitudes determine accent? A study of bilinguals in the USA. *Journal of Multilingual and Multicultural Development*, 28 (6), 502–518. DOI: 10.2167/jmmd514.0.
- 2013. *Foreign accent*. Cambridge: Cambridge University Press.
- Munro, M. J. & T. M. Derwing 1994. Evaluations of foreign accent in extemporaneous and read material. *Language Testing*, 11 (3), 253–266. DOI: 10.1177/026553229401100302.
- Munro, M. J., T. M. Derwing & C. S. Burgess 2010. Detection of nonnative speaker status from content-masked speech. *Speech Communication*, 52 (7–8), 626–637. DOI: 10.1016/j.specom.2010.02.013.
- Munro, M. J. & V. Mann 2005. Age of immersion as a predictor of foreign accent. *Applied Psycholinguistics*, 26 (3), 311–341. DOI: 10.1017/S0142716405050198.
- Nunnally, J. & I. Bernstein 1994. *Psychometric theory*. 3rd ed. New York (N. Y.): McGraw-Hill.
- Oyama, S. 1976. A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, 5 (3), 261–283.
- Piper, T. & D. Cansin 1988. Factors influencing the foreign accent. *Canadian Modern Language Review*, 44 (2), 334–342.
- Piske, T., I. R. A. MacKay & J. E. Flege 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29 (2), 191–215. DOI: 10.1006/jpho.2001.0134.
- Sawilowsky, S. S. 2003. A different future for social and behavioral science research. *Journal of Modern Applied Statistical Methods*, 2 (1), 128–132. DOI: 10.22237/jmasm/1051747860.
- Schmid, M. S. & H. Hopp 2014. Comparing foreign accent in L1 attrition and L2 acquisition: range and rater effects. *Language Testing*, 31 (3), 367–388. DOI: 10.1177/0265532214526175.
- Scovel, T. 1969. Foreign accents, language acquisition, and cerebral dominance. *Language Learning*, 19 (3–4), 245–253. DOI: 10.1111/j.1467-1770.1969.tb00466.x.
- 1981. The recognition of foreign accents in English and its implications for psycholinguistic theories of language acquisition. In J.-G. Savard & L. Laforge (eds) *Proceedings of the 5th congress of AILA*. Québec: Le Presses de L'Université Laval, 389–401.

- Snow, C. & M. Hoefnagel-Höhle 1982. The critical period for language acquisition: evidence from second language learning. In S. Krashen, R. Scarcella & M. Long (eds) *Child-adult differences in second language acquisition*. Rowley (Mass.): Newbury House, 93–111.
- Tench, P. 1996. Intonation and the differentiation of syntactic patterns in English and German. *International Journal of Applied Linguistics*, 6 (2), 223–256. DOI: 10.1111/j.1473-4192.1996.tb00096.x.
- Thompson, I. 1991. Foreign accents revisited: the English pronunciation of Russian immigrants. *Language Learning*, 41 (2), 177–204. DOI: 10.1111/j.1467-1770.1991.tb00683.x.
- Toivola, M. 2011. *Vieraan aksentin arviointi ja mittaaminen suomessa*. Helsinki: Helsingin yliopisto. <http://urn.fi/URN:ISBN:978-952-10-7217-8>.
- Uzal, M., E. Komulainen, M. A. Kılıç & O. Aaltonen 2017. Detection of non-native speech in a familiar language and in an unfamiliar language. *SKY Journal of Linguistics*, 20, 109–136. [http://www.linguistics.fi/julkaisut/SKY2017/SKYJol30\\_2017.pdf#page=111](http://www.linguistics.fi/julkaisut/SKY2017/SKYJol30_2017.pdf#page=111).
- Uzal, M., T. Peltonen, M. Huotilainen & O. Aaltonen 2015. Degree of perceived accent in Finnish as a second language for Turkish children born in Finland. *Language Learning*, 65 (3), 477–503. DOI: 10.1111/lang.12126.

## Appendix 1: Preliminary information form

1) Age	15-19 30-34 45-49 60-64	20-24 35-39 50-54 65-69	25-29 40-44 55-59
2) Gender	Female Male		
3) Mother tongue	Finnish Swedish Other		
4) Is your hearing normal?	Yes No		
5) Are you a native speaker of Finnish from the Helsinki Metropolitan Area?	Yes No		
6) Are you a Finnish-as-a-second-language teacher?	Yes No		
7) Are you a student?	Yes No		
8) Are you a linguistics/phonetics student?	Yes No		
9) Are you a linguistics/phonetics teacher/researcher?	Yes No		
10) Are you in daily contact with immigrants speaking Finnish as a second language?	never rarely often every day		
11) Dialect background: In my opinion, I speak the	Standard Finnish.		dialect.

The above information will be connected to the foreign accent judgments you give. Individual listener identity information will not be saved. I have read and accept the above information on the connection of the information.

Date \_\_\_\_\_

Signature and name in block capitals \_\_\_\_\_

#### Instructions for the Foreign Accent Listening Test

Read the instructions before starting the listening test!

Turn off your cell phone, please! Thank you!

Your task is to assess the kind of Finnish you hear pronounced by the speakers. Pay attention solely to the speaker's pronunciation. Avoid drawing attention to possible errors in grammar, syntactic errors, word-choice errors, and style errors. If you are unsure, make your best guess and use the whole scale. No speech impediments have been noted in the speech data of any speaker.

Foreign accent is a concept which has no generally accepted, uniform definition. A foreign accent means deviations in the standard pronunciation of second language as compared to native speaker pronunciation (Scovel 1969). Before the actual listening test, you will hear a total of 30 sentences from different speakers. The purpose is to briefly familiarize you with listening to speech samples, different pronunciation and giving ratings.

The actual listening test consists of three blocks:

- a word pairs listening test,
- a sentences listening test,
- a spontaneous speech listening test.

The blocks will come in random order. You can take a short break between blocks as well as in the middle of the word pairs and sentences listening tests. The spontaneous speech listening test is shorter than the others, and it includes no break.