

## KIELITAUSTAN VAIKUTUS VOKAALEJA MATKITTAESSA

Lotta Alivuotila, Turun yliopisto, Suu- ja leukakirurgia,  
Hammaslääketieteen laitos

Janne Savela, Turun yliopisto, Informaatioteknologian  
laitos

Olli Aaltonen, Helsingin yliopisto, Puhetieteiden laitos

Tutkimus tarkastelee puheen havaitsemisen ja matkimisen välistä suhdetta kielitaustan näkökulmasta. Tavoitteena on selvittää, matkivatko suomen- ja espanjankieliset koehenkilöt samoja synteettisesti tuotetun vokaalijatkumon ärsykeitä eri tavoin kielitaustastaan riippuen, eli pystyvätkö he matkimaan pelkkää akustista signaalia vai tukeutuvatko he matkimistehtävässä varhaislapsuudessa kehittyneisiin puheäänteiden aivomalleihin. Koehenkilöt suorittivat synteettisten vokaalien havaitsemis- ja matkimiskokeen. Tulokset osoittavat, että kielitausta vaikuttaa ihmisen tapaan havaita ja matkia kuulemiaan äänteitä. Samat synteettiset ärsykkeet aktivoivat erilaiset kielelliset mallit suomalaisten ja espanjalaisten aivoissa, ja vastaavasti he myös matkivat samat vokaalit eri tavoin. Kokemukseen perustuvat äidinkielen äänteiden muistijäljet aivoissa eivät siis ole helposti ohitettavissa, vaikka tietoisesti siihen pyrittäisiinkin.

**Avainsanat:** Matkiminen, vokaalit, formantit.

### JOHDANTO

Perinteisen näkemyksen mukaan puheen havaitseminen ja tuottaminen ovat kaksi erillistä prosessia, joiden välillä ei ole suoranaista yh-

---

Lotta Alivuotila  
Suu- ja leukakirurgia, Hammaslääketieteen laitos  
20014 Turun yliopisto,  
Sähköpostiosoite: lotta.alivuotila@utu.fi

Janne Savela  
Informaatioteknologian laitos  
20014 Turun yliopisto,  
Sähköpostiosoite: jansav@utu.fi

Olli Aaltonen  
Puhetieteiden laitos, Helsingin yliopisto  
PL 9, 00014 Helsingin yliopisto,  
Sähköpostiosoite: olli.aaltonen@helsinki.fi

teyttä (esim. Liberman & Mattingly, 1985; Liberman & Whalen, 2000). Yhteys syntyy vasta kognitiivisella tasolla, jolla varsinainen kielellinen prosessointi tapahtuu. Tämän *horisontaalisen näkemyksen* mukaan puheen tuottaminen ja havaitseminen eivät itsessään ole kielellisiä prosesseja, vaan ainoastaan keino välittää akustiseen signaaliin pakattu informaatio kognitiiviselle tasolle, jolla viesti puretaan ja ymmärretään. Puhe havaitaan vastaavanlaisena akustisena signaalina kuin mitkä tahansa muut ympäristömme ei-kielelliset äänet. Vasta kognitiivisella tasolla alun perin ei-kielellisen signaalin osat sovitetaan kielellisiin malleihin ja viesti ymmärretään. Näkemyksen mukaan foneettiset yksiköt eivät siis ole ensisijaisia ja välittömiä havaitsemisen kohteita,

vaan kognition tasolla esiintyviä abstrakteja äänemalleja, joihin akustisen signaalin osia verrataan (ks. myös Aaltonen, 1997).

Sen sijaan *vertikaalisen näkemyksen* mukaan puheen havaitseminen ja tuottaminen ovat tiiviisti kytkeytyneet toisiinsa ja puheen prosessointi on kokonaisuudessaan kielellistä jo perifeerisellä tasolla (esim. Liberman, 1991; Liberman & Whalen, 2000). Ihminen on biologisesti erikoistunut puhumiseen. Puheen prosessoinnin yksiköjä eivät ole puheäänteet itsessään (vrt. horisontaalinen näkemys) vaan pikemminkin äänteellisten ilmiöiden gesturaaliset mallit. Studdert-Kennedyn (2000, 2002; Studdert-Kennedy & Goldstein, 2003) mukaan puhetta tuotetaan yhdistelemällä ääntöelimestön itsenäisten osien (huulet, kielen kärki, kielen selkä, kielen juuri, pehmeä kitalaki, kurkunpää) diskreettejä foneettisia toimintamalleja. Nämä foneettiset toimintamallit ovat jo itsessään kielellisiä, eli tavallaan ne ovat nimenneet itse itsensä evoluution aikana. Vertikaalista ajattelutapaa tukee *embodiment*-näkemys (Fischer & Zwaan, 2008). Sen mukaan kognitio perustuu sekä havaitsemiseen että toimintaan. Kognition on oltava kytöksissä havaitsemis- ja motorisiin mekanismeihin, sillä niitä prosessoidaan myös kognition tasolla. Kyseiset osa-alueet ovat kytkeytyneet toisiinsa myös kielellisessä toiminnassa. Esimerkiksi *motorisen resonanssin* mukaan toiminnan verbaalisen kuvauksen ymmärtämiseen sisältyy kyseisen toiminnan sisäinen simulaatio. Tässä prosessissa peilineurosystemillä voi olla keskeinen tehtävä.

Peilineuronit ovat soluja, jotka ovat toisten yksilöiden toiminnan ymmärtämisen perusta (Rizzolatti & Arbib, 1998). Ne löydettiin alun perin kädellisten aivojen F5-alueelta, joka vastaa ihmisaivojen Brocan aluetta. Peilineuronit aktivoituvat apinan tehdessä tarttumisliikkeen, mutta myös silloin, kun apina havaitsee vastaavan liikkeen. Peilineuronisysteemi esiintyy myös ihmisillä. Sekä

apinoiden aivojen F5-alueella että ihmisai-vojen Brocan alueella on soluja, jotka kontrolloivat käsien, suun, kasvojen ja kurkunpään toimintaa. Ihmisen käsien ja suun liikkeiden, myös puheentuotolle merkittävien orolaryngaalisten eleiden, on todettu olevan tiiviisti kytöksissä. Kuunnellessamme verbaalisia ärsykeitä myös aivojen puheeseen erikoistuneet motoriset alueet aktivoituvat (Rizzolatti & Craighero, 2004).

Peilineuronit ovat neuraalinen perusta mekanismeille, joka luo suoran yhteyden viestin lähettäjän ja vastaanottajan välille (*pariteetti-periaate*, ks. esim. Liberman & Mattingly, 1989). Pariteetti on kaiken kommunikaation edellytys. Peilineuronimekanismin ansiosta toisten yksilöiden toiminta ymmärretään eräänlaisina viesteinä, ja tämä tapahtuu automaattisesti, eikä edellytä kognitiivista prosessointia (Rizzolatti & Craighero, 2004). Peilineuronimekanismi luo perustan myös matkimiselle, joka puolestaan on kielen omaksumisen edellytys (ks. esim. Ravila, 1937; Korhonen, 1993). Arbib (2003) on esittänyt hypoteesin, jonka mukaan puhekommunikaatio on kehittynyt nimenomaan peilineuronimekanismin ja matkimiskyvyn pohjalta. Hän esittää, että edeltäjillemme kehittyi ensin tarttumisen peilisysteemi, jonka pohjalta edeltäjämme oppivat matkimaan yhä monimutkaisempia kädenliikkeitä. Vähitellen liikkeet saivat yhä kommunikatiivisempia merkityksiä; imitaatio muuttui eräänlaiseksi pantomiimiksi. Edeltäjillemme kehittyi alkeellinen elekieli, joka myöhemmin korvautui protopuheella ja lopulta nykykielellä (Arbib, 2002, 2003). Myös Studdert-Kennedy (2002) korostaa peilineuronimekanismin merkitystä puheen evoluutiossa (ks. myös Studdert-Kennedy & Goldstein, 2003). Peilineuronit aktivoituvat, kun tietyt motoriset elimet (esim. käsi, huulet) toimivat tietyn mallin mukaisesti (esim. tarttuminen, maiskaus). Tämä periaate pätee myös puheeseen: *foneettisia eleitä* tuotetaan

yhdistämällä artikulaatioelimen tietynlainen liike fonaatioon tietyssä kohtaa ääntöväylää. Foneettiset toimintamallit voidaan nähdä puheen perusyksikköinä (vrt. *motorinen teoria*, esim. Liberman, 1996), ja niiden yhdistelymahdollisuudet laajenivat vähitellen evolution kuluessa. Peilineuroneilla ja matkimisella oli keskeinen merkitys fonologisen systeemin kehityksessä.

Paavo Ravila (1937) korosti matkimisen merkitystä jo 1930-luvulla. Hänen mukaansa matkiminen on perustavanlaatuisin mekanismi kielen omaksumisessa. Jokelteleva vauva matkii ensin itseään ennen kuin alkaa matkia toisia. Matkiminen on aluksi tiedostamatonta ja tahdosta riippumatonta. Ravilan mukaan merkittävä edistysaskel kielellisessä kehityksessä otetaan, kun matkiminen muuttuu tietoiseksi. Tietoinen matkimisen oppiminen on kielen oppimisen välttämätön edellytys, tosin ei ainoa. Myös Skoylesin (1998) mukaan kyky matkia on kielen olemassaolon edellytys. Ihmisillä on ainutlaatuinen kyky matkia puhetta, ja matkiminen on nopeaa ja sujuvaa. Sen voi suorittaa joko välittömästi (*shadowing*; ks. myös Chistovich, Fant, Serpa-Leitao & Tjernlund, 1966) tai lyhyt- ja pitkäkestoiseen muistiin varastoimisen jälkeen. Matkiminen ei ole riippuvainen kielellisistä kyvyistä tai älykkyydestä, ja matkimisen on todettu olevan myös esi-kielellinen ominaisuus (Skoyles, 1998). Fraser, Bellugi ja Brown (1963) ovat esittäneet, että matkiminen ei edellytä merkityksen ymmärtämistä. Vastaavasti Hurme (1975) vertaa artikkelissaan outoa kieltä matkivia ihmisiä papukaijoihin, jotka taitavasti matkivat puheen akustista rakennetta ymmärtämättä mitään sen sisällöstä.

Tässä tutkimuksessa käytettiin metodina matkimista, jonka keinoin haluttiin selvittää, missä määrin kielellinen tausta heijastuu ihmisen kykyyn havaita ja matkia äänteellisiä ilmiöitä. Mikäli horisontaalisen näkemyksen mukaisesti puheen havaitseminen on asteit-

tainen prosessi, jossa akustinen puhesignaali muuttuu kielelliseksi vasta kognitiivisella tasolla, eri kielitaustaisten koehenkilöiden matkimistulosten voidaan olettaa olevan samankaltaiset – edellyttäen, että matkimiskyky on spontaani, kielellisistä kyvyistä riippumaton ominaisuus (Skoyles, 1998). Jos taas puheen havaitseminen on vertikaalisen näkemyksen mukaisesti automaattista ja kielellistä jo perifeerisellä tasolla, voidaan olettaa, että erot koehenkilöiden fonologisissa järjestelmissä heijastuvat myös matkimistuloksiin.

### *Vokaaleiden havaitsemis- ja matkimistutkimuksista*

Synteettisten vokaalien matkimista ja matkimisen suhdetta vokaalien havaitsemiseen on tutkittu vain vähän. Ensimmäiset tutkimukset ovat peräisin 1960-luvulta, jolloin Chistovich ym. (1966) tutkivat venäjän vokaalien matkimista yhdellä naispuolisella äidinkieleltään venäläisellä koehenkilöllä. Tulokset osoittivat, että venäjän vokaalien matkiminen jatkumolla [a-ε-i] oli kategorista. Matkimistehtävässä ilmenneet kategoriat eivät kuitenkaan vastanneet venäjän foneemijärjestelmää, vaan matkimistehtävässä kategorioita tuli esiin useampia. Mahdollisesti osa näistä matkimiskategorioista vastasi koehenkilön tuottamia allofoneja (Chistovich ym., 1966). Kent (1973) puolestaan tutki matkimista amerikanenglannin vokaalijatkumoilla [æ-i] ja [u-i]. Matkimiskokeen suoritti neljä amerikanenglantia äidinkielenään puhuvaa henkilöä. Kentin tulokset osoittivat, että jatkumon [æ-i] matkiminen oli kategorisempaa kuin jatkumon [u-i] matkiminen. Kent esitti johtopäätöksen, että koehenkilöiden äidinkielen foneemisysteemi vaikutti heidän kykyynsä matkia jatkumoa [æ-i], koska se on amerikanenglannissa jaettavissa viiteen vokaaliin ([i], [ɪ], [e], [ɛ] ja [æ]). Sen sijaan jatkumolla [u-i] esiintyvät ainoastaan kysei-

set kaksi vokaalia ja nekin jatkumon päissä. Myös myöhemmässä tutkimuksessa Kent (1974) havaitsi äidinkielen fonologisen järjestelmän vaikutuksen matkimiseen: identifikaatiokokeen perusteella epämääräiseksi luokitellut vokaalit olivat vaikeammat matkia kuin vokaalit, jotka selkeästi kuuluivat äidinkielen vokaalijärjestelmään. Myös kaksikielisillä puhujilla matkimisen on havaittu olevan jossakin määrin kategorista ja sidoksissa molempien kielten vokaalijärjestelmiin (esim. Schouten, 1977). Repp ja Williams (1985) toistivat pienin muutoksin edellä esitetyn Kentin matkimiskokeen vuodelta 1973. He saivat tutkimuksissaan samansuuntaisia tuloksia, joskin matkittujen vokaalien formanttiarvojen muutokset olivat heidän mukaansa pikemminkin epälineaarisia kuin selkeästi kategorisia. Toisin kuin Kent, he havaitsivat myös [u-i] -jatkumon matkimistuloksissa selviä epälineaarisuuksia.

Puheen havaitsemisen ja matkimisen on esitetty olevan haasteellisempaa lapsille kuin aikuisille (ks. Kent, 1979, joka viittaa Eguchiin, 1976). Lasten heikompi matkimiskyky ei johdu ainoastaan artikulaation epätarkkuudesta, vaan he myös havaitsevat heikommin äänteellisiä eroja. Lisäksi lapset ovat alttiimpia äidinkielen vokaalijärjestelmän vaikutukselle (Kent, 1978; 1979). Myös harjoittelu vaikuttaa sekä äänteellisten ilmiöiden havaitsemiseen että matkimiseen (Alivuotila, Hakokari, Savela, Happonen & Aaltonen, 2007). Alivuotilan ym. (2007) tutkimuksessa vertailtiin lasten, tavallisten aikuisten kielenkäyttäjien sekä matkimiseen harjaantuneiden foneetikojen kykyä havaita ja matkia pieniä akustisia eroja synteettisissä vokaaleissa jatkumolla /æ-ɑ/. Vokaalien laadun arviointi oli lapsille haastavampaa kuin aikuisille, ja harjaantuneet foneetikot erottelivat vokaalien laatua parhaiten. Äidinkieli vaikutti matkimiseen sekä lapsilla että aikuisilla, lapsilla selvästi enemmän. Suurin osa lapsista tuotti matkimistehtävässä

joko vokaalin [æ] tai [ɑ], eivätkä juuri lainkaan epämääräisiä vokaalilaatuja; heidän matkimisensa oli selvästi kategorista. Aikuiset pyrkivät tuottamaan myös epämääräisiä vokaalilaatuja, mutta matkittujen vokaalien formanttiarvoissa oli havaittavissa lievää epälineaarisuutta. Harjaantuneet foneetikot sen sijaan kykenivät matkimaan vokaalien laadullisia eroja erittäin tarkasti ilman äidinkielen vaikutusta. Foneetikot ovat siis siinä mielessä poikkeuksellisia, että he ovat oppineet sulkemaan pois kielellisen kokemuksensa vaikutuksen äänteellisiä ilmiöitä matkiesseen, kun taas tavallisten kielenkäyttäjien matkiminen on kiinteämmin sidoksissa heidän käyttämäänsä kieleen, lapsilla aikuisia enemmän.

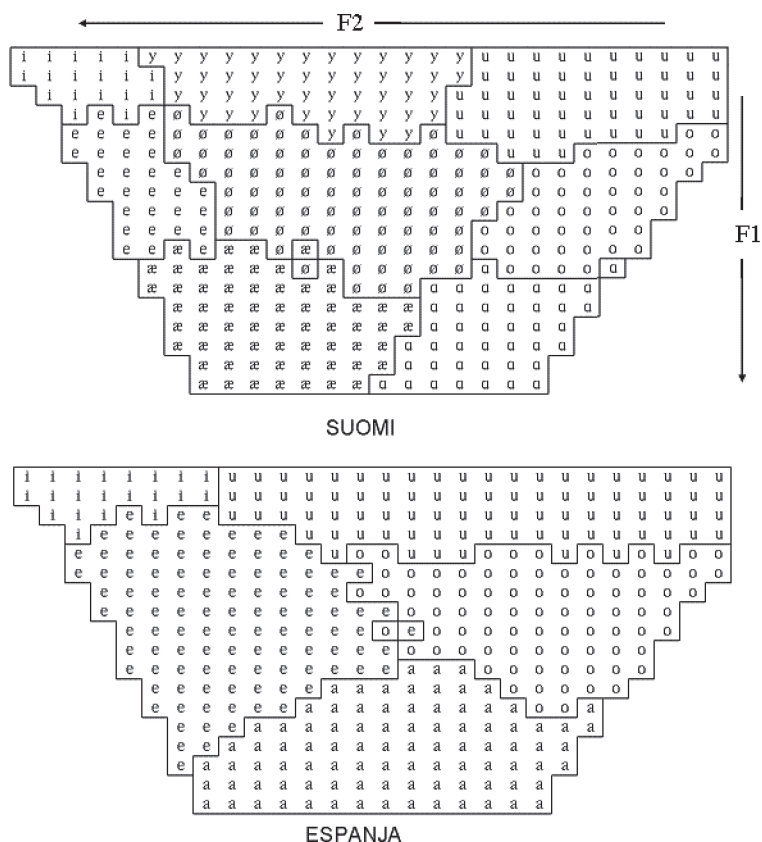
Kuten edellä esitetyistä tutkimuksista voidaan todeta, matkimiskokeista saadut tulokset eivät ole yksiselitteisiä ja aihealue vaatii lisää perehtymistä. Kuitenkin yleistendensi näyttäisi olevan, että vokaalien matkiminen on jossakin määrin kategorista ja sidoksissa äidinkielen foneemimalleihin. Tulosten voidaan tulkita tukevan motorista näkemystä, jonka mukaan puheen tuottamisen ja havaitsemisen välillä on tiivis neuraalinen yhteys (esim. Liberman & Mattingly, 1985; Studdert-Kennedy, 2002). Matkimme sitä mitä kuulemme, ja kuuntelemme puhetta automaattisesti äidinkielemme kautta.

Tässä tutkimuksessa puheen tuottamisen ja havaitsemisen välistä yhteyttä tarkasteltiin ennen kaikkea kielitaustan näkökulmasta. Halusimme selvittää, matkivatko kielitaustaltaan erilaiset henkilöt samoja ärsykeitä eri tavoin, eli pystyvätkö he matkimaan pelkkää akustista signaalia vai tukeutuvatko he matkimistehtävässä sisäisiin, kokemukseen perustuviin fonologisiin muistijälkiin. Tietääksemme vastaavanlaista vertailevaa tutkimusta ei ole aikaisemmin tehty.

## MENETELMÄT

Tutkimuksessa vertailtiin synteettisten vokaalien matkimista suomalaisten ja espanjalaisten välillä. Peruslähtökohtana oli Turun vokaalitestistä (*Turvotes*) kerätty aineisto, jonka mukaan suomalaiset ja espanjalaiset havaitsevat eri tavoin kuulemiaan synteettisiä vokaaleja (ks. esim. Raimo, Savela & Aaltonen, 2002). Vertasimme suomalaisten ja espanjalaisten matkimia vokaaleja väljällä jatkumolla, jonka suomalaiset jakavat kahteen äännekategoriaan [æ] ja [ɑ] ja espanjalaiset vain yhteen kategoriaan [a] (ks. kuva 1.). Käytimme havainto- ja matkimiskokeessa samoja synteettisiä vokaaleja, joita käytettiin Turvotes-vokaalitestissä, joten tulokset ovat verrannollisia Turvotes-aineistoon.

Yleisesti vokaalit tunnistetaan kahden ensimmäisen formantin, F1:n ja F2:n, Hertz-arvojen (Hz) perusteella (ks. esim. Dew & Jensen, 1977). F1:n Hz-arvon muutos korreloi suun avoimuuden kanssa (*suppea vs. väljä*). F2 vastaavasti korreloi kielen horisontaalisen asennon kanssa (*etinen vs. takainen*). F3:n muutos puolestaan liittyy vokaalien pyöreyyteen (*pyöreys vs. laveus*) (Stevens, 2000). Tässä tutkimuksessa käytetyissä synteettisissä vokaaleissa F1 ja F3 olivat vakiot, ainoastaan F2:n arvo muuttui asteittain.



Kuva 1. Suomen ja Espanjan vokaalikategoriat Turvotestuksen mukaan (ks. esim. Raimo ym. 2002).

### Koehenkilöt

Koehenkilöinä oli viisi suomenkielistä ja viisi espanjankielistä naista. Suomenkieliset naiset asuivat kaikki Turun seudulla ja olivat iältään 24–32 -vuotiaita. Espanjankieliset naiset olivat iältään 21–28 -vuotiaita Turun yliopiston vaihto-opiskelijoita, jotka olivat oleskelleet Suomessa kolme kuukautta, eivätkä osanneet suomea. Kaikilla koehenkilöillä oli normaali kuulo.

### Ärsykkeet

Kokeessa käytettiin 14 synteettistä vokaalia [æ-ɑ] -jatkumolta. Vokaalit syntetisoitiin HlSynin Klatt-ohjelman versiolla 1.0 (Klatt, 1980), ja niiden Hz-arvot mitattiin Praat-analyysiohjelmalla matkimisaineistoon vertaamista varten (Boersma & Weenink, 2007). Jatkumo luotiin muuttamalla 40 melin askelin toisen formantin (F2) arvoja 1995 Hertsiä (1500 mel) 969 Hertsiin (980 mel). Ensimmäisen (F1) ja kolmannen (F3) formantin arvot olivat kaikissa ärsykkeissä vakiot: 756 Hz (830 mel) ja 2517 Hz (1700 mel). Ärsykkeiden kesto oli 350 ms. Niiden perustaajuus ( $F_0$ ) nousi ensin asteittain saavuttaen 120 Hz:n arvon 120 millisekunnin kohdalla ja sitten vähitellen laski ollen vokaalin lopussa 80 Hz.

### Koasetelma

Koe toteutettiin PxLab-ohjelmalla (Irtel, 2007), ja koehenkilöt suorittivat sen kannettavalla tietokoneella, johon oli liitetty kuulokemikrofoni. Koe koostui havainto- ja matkimiskokeesta. Molemmissa tehtävissä koehenkilö kuuli yhteensä 70 vokaalia satunnaisessa järjestyksessä. F2-arvoiltaan 14 erilaista vokaaliärsykettä toistui kokeen aikana yhteensä viisi kertaa. Koe kesti ohjeistuksineen noin 15 minuuttia. Ohjeet annettiin koehenkilöille

sekä suullisesti että kirjallisesti, suomalaisille suomeksi ja espanjalaisille englanniksi. Koehenkilöt suorittivat kokeen itsenäisesti rauhallisessa suljetussa huoneessa.

### Havaintokoe

Koska Turvotes-aineiston perusteella tiedettiin, etteivät espanjalaiset suomalaisten tavoin jaa kokeen vokaalijatkumoa kahteen kategoriaan vaan ainoastaan yhteen, espanjalaiset eivät suorittaneet identifikaatiotehtävää (ks. alla) vaan pelkän hyvydenarviointitehtävän (ks. kuva 1.). Heitä pyydettiin arvioimaan asteikolla 1–7 (1 = *erittäin huono* ja 7 = *erittäin hyvä*), kuinka hyvä kategorian [a] edustaja kukin heidän kuulemansa vokaali oli. Arviointi suoritettiin painamalla jotakin näppäimistä <1–7>. Suomalaisia pyydettiin sen sijaan nimeämään ensin ärsykkeet joko vokaaliksi [æ] tai [ɑ] suomen vokaalijärjestelmän mukaisesti. Nimeäminen tapahtui painamalla näppäintä <Ä> tai <A>. Tällä tavoin selvitettiin yksittäisten koehenkilöiden kategorioiden rajat. Tämän jälkeen suomalaisia pyydettiin arvioimaan kunkin testivokaalin hyvyys joko [æ]-tai [ɑ]-kategorian edustajana asteikolla 1–7.

### Matkimiskoe

Havaintokokeen jälkeen koehenkilöt kuulivat samat vokaalit uudestaan sattumanvaraisessa järjestyksessä. Matkimiskoe oli molemmilla ryhmillä samanlainen. Koehenkilöitä pyydettiin matkimaan kunkin vokaalin akustista laatua mahdollisimman tarkasti. Kun koehenkilöt olivat matkineet kuulemansa ärsykkeen, he painoivat näppäintä <seuraava/next>, jonka jälkeen he kuulivat seuraavan ärsykkeen. Vastausaikaa ei rajoitettu. Halutessaan koehenkilöt saattoivat kuunnella ärsykkeen uudestaan painamalla näppäintä <toisto/repeat>. Vastaukset nauhoitettiin akustista analyysia varten.

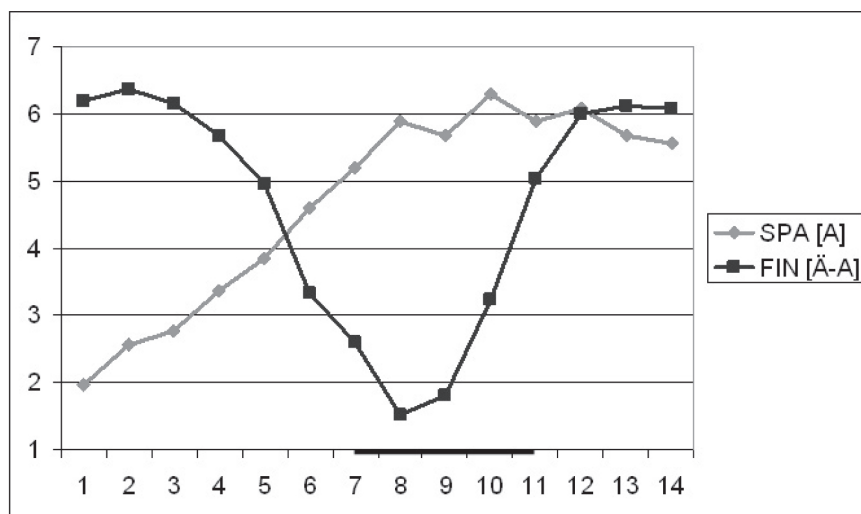
*Akustinen analyysi*

Akustinen analyysi tehtiin Praat-ohjelman versiolla 5.0 (Boersma & Weenink, 2007). Koska mittauspaikan erot saattavat aiheuttaa huomattaviakin taajuusarvojen eroja, formanttien keskiarvot laskettiin mahdollisimman suurelta alueelta vokaalia. On esitetty, että mitattujen formanttien virhemarginaali on suoraan verrannollinen puhujan äänen perustaajuuteen ja sen suuruus on neljäsosa  $F_0$ :sta (ks. Kent, 1979, joka viittaa Lindblomiin, 1972). Tämän perusteella oletettu virhemarginaali koehenkilöillä oli enimmillään noin 50 Hz ( $F_0$ :n keskiarvo  $\sim 190$  Hz). Matkituista vokaaleista mitattiin F1-, F2- ja F3-taajuudet. Maksimi asetettiin 5500 Hz:iin, ja alueelta etsittiin viisi formanttihuippua. Mikäli kyseiset asetukset eivät antaneet luotettavia tuloksia, arvot mitattiin manuaalisesti spektrin rakennetta tarkastelemalla. Jokaisen 14 vokaalin viidestä toistosta laskettiin kolmen ensimmäisen formantin keskiarvot ja keskihajonnat.

## TULOKSET

Suomalaiset nimesivät kuulemansa vokaalit ongelmitta. Kuten kuvasta 2 käy ilmi, alue, jossa [æ] vaihtui vokaaliksi [ɑ], osui keskimäärin ärsykkeiden 7 ja 11 välille. Kategoriarajalla vokaalit odotetusti arvioitiin laadullisesti huonoimmiksi. Suomalaisten [æ]-vokaaleista keskimääräisesti parhaaksi arvioitiin ärsyke 2 ja [ɑ]-vokaaleista keskimääräisesti parhaaksi arvioitiin ärsyke 13. Espanjalaisten vokaalijatkumon keskimääräisesti paras [a] oli ärsyke 10 ja huonoin oli ärsyke 1.

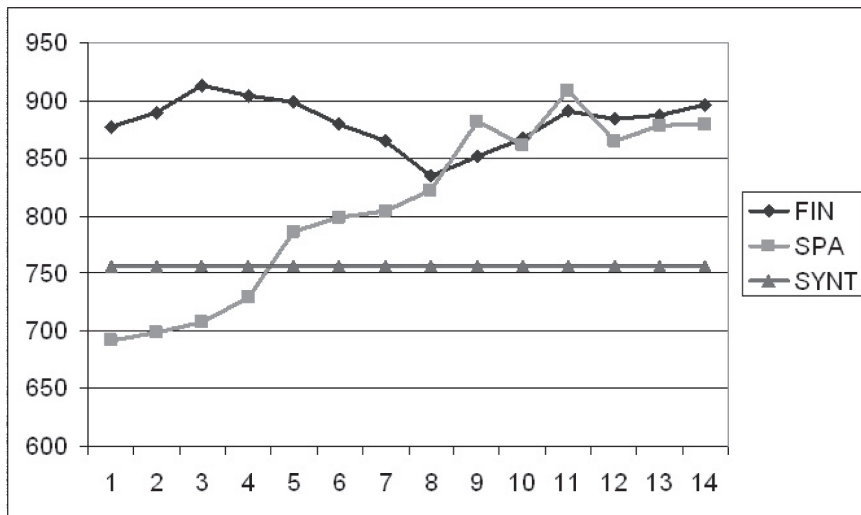
Matkimistuloksissa kielitaustan vaikutus näkyi ennen kaikkea F1-arvoissa, kuten kuvasta 3 käy ilmi. Suomalaisilla [æ]:n ja [ɑ]:n kategoriarajalla F1 selvästi laski. Sen sijaan espanjalaisilla F1-arvot systemaattisesti nousivat jatkumon loppua kohden. Jatkumon alkupäässä espanjalaisten F1-arvot olivat huomattavan alhaisia, lähes 200 Hz matalampia kuin suomalaisilla. Kuten kuvasta 4 voi nähdä, F2-arvoissa ei ryhmien välillä ollut nähtävissä



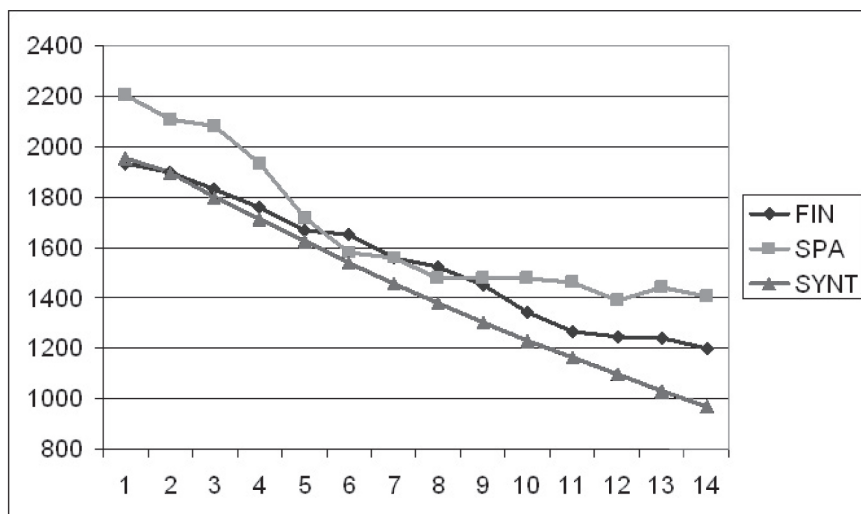
Kuva 2. Suomalaisten (FIN) ja espanjalaisten (SPA) hyvydenarviointitulosten keskiarvot syntheettisellä vokaalijatkumolla (x-akseli). Hyvyys arvioitiin asteikolla 1–7 (y-akseli). Suomalaisten keskimääräinen kategoriaraja on merkitty x-akselille lihavoidulla viivalla.

selvää eroa. Suomalaisilla tosin F2-arvot olivat pääosin lähempänä synteettisten mallivokaalien arvoja kuin espanjalaisilla. Kuvan 5 mukaan taas F3-arvot olivat espanjalaisilla selvästi korkeammat kuin suomalaisilla, mutta molemmilla ryhmillä ne olivat hieman korkeammat jatkumon alku- kuin loppupäässä.

Matkimistuloksien keskihajonta oli espanjalaisten tuloksissa systemaattisesti suomalaisten tulosten hajontaa suurempi. Espanjalaisilla kolmen ensimmäisen formantin keskihajontojen keskiarvot olivat 55 Hz, 90 Hz ja 83 Hz ja suomalaisilla vastaavasti 39 Hz, 68 Hz ja 80 Hz. Espanjalaisilla havaintokokeen tulok-

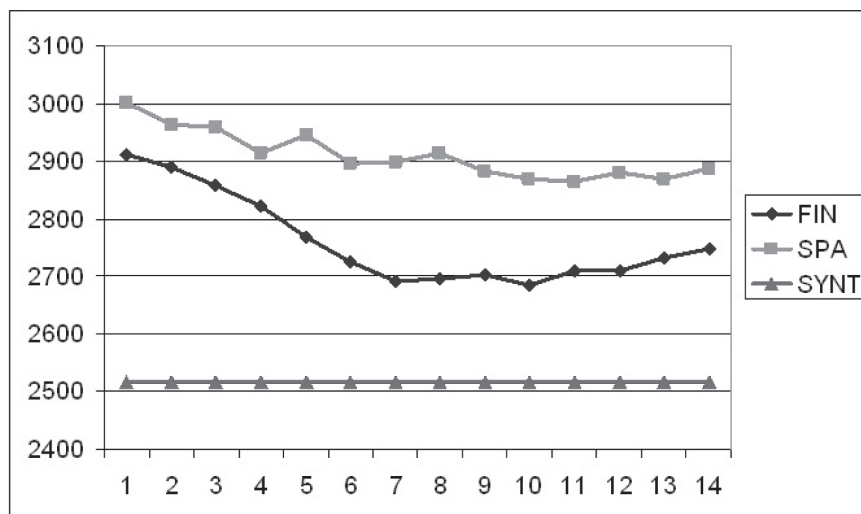


Kuva 3. Suomalaisen (FIN) ja espanjalaisen (SPA) matkimien vokaalien F1-arvot (Hz, y-akseli) verrattuna synteettisten mallivokaalien (SYNT) F1-arvoihin 14 vokaalin jatkumolla (x-akseli).



Kuva 4. Suomalaisen (FIN) ja espanjalaisen (SPA) matkimien vokaalien F2-arvot (Hz, y-akseli) verrattuna synteettisten mallivokaalien (SYNT) F2-arvoihin 14 vokaalin jatkumolla (x-akseli).





Kuva 5. Suomalaisen (FIN) ja espanjalaisen (SPA) matkimien vokaalien F3-arvot (Hz, y-akseli) verrattuna synteettisten mallivokaalien (SYNT) F3-arvoihin 14 vokaalin jatkumolla (x-akseli).

set heijastuivat jossakin määrin matkittujen vokaalien formanttiarvojen keskihajontoihin. Mallivokaalien hyvyysarvioiden ja matkittujen vokaalien F2-arvojen keskihajontojen välinen riippuvuussuhde osoittautui *Spearmanin korrelaatiotestillä* tilastollisesti merkitseväksi: F2-arvojen keskihajonnat olivat pienemmät hyväiksi arvioituja vokaaleja matkittaessa, kun taas huonon arvosanan saaneita vokaaleja matkittaessa hajonta oli suurempi ( $r = -.767, p = 0.01$ ). F1- ja F3-arvoissa ei tilastollisesti merkitsevää korrelaatiota esiintynyt. Suomalaisilla sen sijaan havaintokokeen tulokset eivät vaikuttaneet matkittujen vokaalien F1-, F2- tai F3-arvojen keskihajontoihin, eli vokaalin huonous/hyvyys ei merkitsevästi vaikuttanut minkään formanttiarvon keskihajontaan.

## POHDINTA

Suomalaisen matkimien vokaalien F1-arvojen lasku todennäköisesti johtuu siitä, että raja-alueen epämääräisten (hyvyydenarviointitehtävässä huonoimmiksi arvioitujen)

vokaalien matkiminen oli suomalaisille vaikeaa. He tukeutuivat epämääräisellä alueella suomen vokaalijärjestelmässä lähellä olevaan vokaaliin [ø], minkä seurauksena vokaalit myös usein tuotettiin [ø]:n kaltaisina eli pyöreähköinä ja suppeampina. Tämä näkyi spektrogrammeissa sekä F1-arvojen että F3-arvojen laskuna. Espanjalaisen matkimien vokaalien F1-arvojen nousu puolestaan selittyy sillä, että espanjalaiset kuuluivat etiset, väljät vokaalit epämääräisinä (hyvyydenarviointikokeessa ne arvioitiin huonoimmiksi). Espanjan vokaalijärjestelmässä ainoa etinen väljäkö vokaali on [e], johon espanjalaiset koehenkilöt tukeutuivat pyrkivänsä matkimaan heille epämääräisiä vokaaleja. Näin ollen vokaalit tuotettiin usein [e]:n kaltaisina, ja suun suppeampi avoimuusaste aiheutti matalamman F1-arvon. Yksinkertaistaen voidaan todeta, että espanjalaiset pyrkivät kompensoimaan vokaalien akustisia eroja ennen kaikkea ääntöväylän vertikaalisella akselilla (F1, kielen suppeus vs. väljyys), mikä todennäköisesti johtui siitä, että väljissä vokaaleissa horisontaalisen (F2, kielen etisyys vs.

takaisuus) akselin hyödyntäminen on heille vierasta. Myös suomalaiset turvautuivat vertikaaliseen ulottuvuuteen kategoriarajalla.

Matkimistuloksien keskihajontojen erot selittyvät myös todennäköisesti kielitaustan erolla. Vieraat äänneet ja ääntämistapa aiheuttavat espanjalaisissa epävarmuutta. Lisäksi suurempi foneemikategoria sallii suuremman hajonnan. F3-arvojen hajonnan ero ryhmien välillä oli hyvin pieni, mikä johtuu todennäköisesti siitä, ettei kolmannella formantilla ole yhtä merkittävää roolia vokaalien tunnistamisessa kuin kahdella ensimmäisellä formantilla. Vaikka espanjalaisilla matkittujen vokaalien toisen formantin keskihajonnat korreloivat vokaalien hyvyysarvioiden kanssa tilastollisesti merkitsevästi, suomalaisilla vastaavaa ilmiötä ei ollut havaittavissa. Alun perin oletimme, että molempien ryhmien tulokset olisivat tässä suhteessa samankaltaiset, joten havainto vaatii vielä lisää tutkimusta.

Vaikka synteettiset mallivokaalit olivat siinä mielessä stabiileja, että vokaaleissa F1- ja F3-formantit olivat vakiot läpi jatkumon ja ainoastaan F2 muuttui, vastaavaa ilmiötä ei ollut nähtävissä matkimistuloksissa. Matkituissa vokaaleissa kaikissa kolmessa formantissa oli havaittavissa vaihtelua. Jotta matkitut vokaalit olisivat vastanneet synteettisiä mallivokaaleja, koehenkilöiden olisi täytyntä pitää suun avoimuus ja laveus vakiona läpi jatkumon ja muuttaa ainoastaan kielen horisontaalista asentoa. Kun kyseessä on formantteja säätelevän puhesynteesiohjelman sijasta tavallinen kielenkäyttäjät, tämä on käytännössä mahdotonta. Huomattava tosin on, että vastaavuus formanttiarvojen ja artikulaatioelinten (lähinnä kielen) liikkeiden välillä ei ole täysin yksiselitteinen, sillä tosiasiaa formantit ovat koko ääntöväylän konfiguraation vaikutuksen alaisena.

Tuloksemme osoittavat, että kielitausta vaikuttaa selvästi ihmisen tapaan havaita ja matkia kuulemiaan irrallisia äänneitä. Myös

aikaisemmat tutkimukset ovat osoittaneet, että matkiminen on ainakin jossakin määrin kategorista, vaikkakaan tulokset eivät ole olleet yksiselitteisiä. Vaikka koeasetelmamme vokaalit olivat synteettisiä ja näin ollen matkimisen kannalta neutraaleja, ja vaikka koehenkilöitä pyydettiin matkimaan nimenomaan akustisia eroja ärsykeissä, äidinkielen vaikutukselta ei välttytty. Varhaislapsuudessa kehittyneet fonologiset muistijäljet eivät ole helposti ohitettavissa, vaikka siihen tietoisesti pyrittiäsiinkin. Tässä tapauksessa samat synteettiset ärsykkeet aktivoivat erilaiset kielelliset mallit suomalaisten ja espanjalaisten aivoissa, ja vastaavasti he myös matkivat samat ärsykkeet eri tavoin.

Tulosten voidaan tulkita tukevan vertikaalista näkemystä, jonka mukaan puheen havaitseminen ja tuottaminen ovat tiiviisti sidoksissa toisiinsa (ks. esim. Liberman & Whalen, 2000). Yhteys toteutuu neuraalisella tasolla mahdollisesti peilineuroneiden toiminnan kautta (Studdert-Kennedy, 2002). Synteettisiä vokaaleja matkiessamme emme matki niinkään vokaalien akustista edustumaa, vaan aivoissamme sijaitsevia gesturaalisia malleja, joiden perusteella sekä tuotamme että havaitsemme puhetta. Nämä gesturaaliset mallit ovat olemassa puheessa (tässä tapauksessa irrallisissa synteettisissä vokaaleissa) itsessään ja määräytyvät äidinkielemme mukaan, jolloin omaksumamme fonologinen systeemi väistämättä vaikuttaa kielelliseen prosessointiin. Näin ollen myös matkitut vokaalit olivat erilaiset espanjalaisilla ja suomalaisilla. Mikäli puheen havaitseminen ja tuottaminen ovat kaksi toisistaan riippumatonta prosessia ja yhdistyvät vasta kognitiivisella tasolla, jolla viesti saa kielellisen merkityksen ja tulee ymmärretyksi (ks. esim. Ryalls, 1996), voidaan olettaa, että tämänkaltaisessa kokeessa synteettisten vokaalien havaitseminen ja tuottaminen eivät olisi niin kiinteästi yhteydessä toisiinsa. Koehenkilöt pystyisivät matkimaan

puheen akustiikkaa, koska matkiminen on automaattinen, merkityksistä ja kielellisistä kyvystä sekä älykkyydestä riippumaton ominaisuus (Fraser ym. 1963; Skoyles, 1998). Tulostemme valossa vertikaalinen näkemys tuntuu luonnollisemmalta vaihtoehdolta. Toisin kuin horisontaalinen näkemys esittää, puhesignaali on kielellinen jo itsessään, eikä ainoastaan informaation välittäjä. Sen havaitseminen ja tuottaminen on automaattista, ja nämä kaksi ilmiötä näyttävät vahvasti olevan sidoksissa toisiinsa. Me matkimme sitä mitä havaitsemme ja havaitsemme parhaiten sitä, mitä osaamme matkia.

## KIITOKSET

Kiitämme filosofian tohtori Juha-Pertti Laaksoista (Albertan yliopisto, Edmonton) tilastotieteellisestä asiantuntemuksesta. Lisäksi kiitos valtakunnalliselle kielentutkimuksen tutkijakoulu Langnetille.

## LÄHTEET

- Aaltonen, O. (1997). *Vowel perception: Behavioural and psychophysiological experiments*. Annales Universitatis Turkuensis, D. 261.
- Alivuotila, L., Hakokari J., Savela J., Happonen R-P., Aaltonen, O. (2007). Perception and imitation of Finnish open vowels among children, naïve adults and trained phoneticians. *Proceedings of the 16<sup>th</sup> International Congress of Phonetic Sciences* (s. 361–364). Saarbrücken, Universität des Saarlandes.
- Arbib, M.A. (2002). The mirror system, imitation, and the evolution of language. Teoksessa K. Dautenhahn & C. L. Nehaniv (toim.), *Imitation in animals and artifacts* (s. 229–280). Cambridge, Mass: MIT Press.
- Arbib, M.A. (2003). The evolving mirror system: A neural basis for language readiness. Teoksessa M. H. Christiansen & S. Kirby (toim.), *Language evolution* (s. 182–200). Oxford: Oxford University Press.
- Boersma, P. & Weenink, D (2007). *Praat: doing phonetics by computer* (Versio 5.0) [Tietokoneohjelma]. Haettu 11.12.2007 osoitteesta <http://www.praat.org/>.
- Chistovich, L., Fant, G., Serpa-Leitao, A., & Tjernlund, P. (1966). Mimicking of synthetic vowels. *Speech Transmission Laboratory - Quarterly Progress and Status Report 2*, Stockholm, 1–18.
- Dew, D. & Jensen, P.J. (1977). *Phonetic processing: The dynamics of speech*. Columbus, Ohio: Merrill.
- Eguchi, S. (1976). Difference limens for the formant frequencies: Normal adult values and their development in children. *Journal of the American Audiology Society*, 1, 145–149.
- Fischer, M. H., & Zwaan, R. A. (2008, painossa). Embodied language: A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology*.
- Fraser, C., Bellugi, U., & Brown, R. (1963). Control of grammar in imitation, comprehension, and production. *Journal of Verbal Learning and Verbal Behaviour*, 2, 121–135.
- Hurme, P. (1975). Oudon kielen matkimisesta: Ihmisen kyvystä ja rajoituksista imitoida oudon kielen lauseita. *Fonetiikan paperit – Helsinki 1975*. Helsingin yliopiston Fonetiikan laitoksen julkaisuja, 27, 19–42.
- Irtel, H. (2007). *PxLab: The psychological experiments Laboratory* [online]. Versio 2.1.11. Mannheim: University of Mannheim. Haettu 6.10.2006 osoitteesta <http://www.pxlab.de>.
- Kent, R. D. (1973). The imitation of synthetic vowels and some implications for speech memory. *Phonetica*, 28, 1–25.
- Kent, R. D. (1974). Auditory-motor formant tracking: A study of speech imitation. *Journal of Speech and Hearing Research*, 17, 203–222.
- Kent, R. D. (1978). The imitation of synthesized vowels by preschool children. *Journal of the Acoustical Society of America*, 63, 1193–1198.
- Kent, R. D. (1979). The imitation of synthesized English and nonEnglish vowels by children and adults. *Journal of Psycholinguistic Research*, 8, 43–60.
- Klatt, D. H. (1980). Software for cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, 67, 971–995.
- Korhonen, M. (1993). *Kielen synti*. Helsinki: WSOY.
- Lieberman, A. M., & Mattingly I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1–36.

- Lieberman, A. M., & Mattingly I. G. (1989). A specialization for speech perception. *Science*, 240 (4890), 489–494.
- Lieberman, A. M. (1991). Observations from the sidelines. *Reading and Writing: An Interdisciplinary Journal*, 3, 429–433.
- Lieberman, A.M. (1996). *Speech: A special code*. Cambridge, Mass.: MIT Press.
- Lieberman, A. M., & Whalen, D. G. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, 4, 187–196.
- Lindblom, B. (1972). Comments on Paper 15 – Development of speech sounds in children, by S. Eguchi and I. Hirch. Teoksessa G. Fant (toim.), *International Symposium on Speech Communication Ability and Deafness*, Washington D.C.
- Raimo, I., Savela, J., & Aaltonen, O. (2002). The Turku Vowel Test. Teoksessa P. Korhonen (toim.), *Fonetikan päivät 2002* (s. 45–52). Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Report 67.
- Ravila, P. (1937). Lasten kielen kehityksestä. *Länsi*, 3, 139–146.
- Repp, B. H., & Williams, D. R. (1985). Categorical tendencies in vowel imitation: preliminary observations from a replication experiment. *Speech Communication*, 4, 105–120.
- Rizzolatti G., & Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience* 2004, 27, 169–192.
- Rizzolatti, G., & Arbib M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21, 188–194.
- Ryalls, J. (1996). *A basic introduction to speech perception*. San Diego, CA.: Singular.
- Schouten, M. E. H. (1977). Imitation of synthetic vowels by bilinguals. *Journal of Phonetics*, 5, 273–283.
- Skoyles, J. R. (1998) Speech phones are a replication code. *Medical Hypotheses*, 50, 167–173.
- Stevens, K.N. (2000). *Acoustic phonetics*. Cambridge, Mass: MIT Press.
- Studdert-Kennedy, M. (2000). Imitation and the emergence of segments. *Phonetica*, 57, 275–283.
- Studdert-Kennedy, M. (2002). Mirror neurons, vocal imitation and the evolution of articulate speech. Teoksessa M. I. Stamenov (toim.), *Mirror neurons and the evolution of brain and language* (s. 217–238). Philadelphia, PA.: John Benjamins.
- Studdert-Kennedy, M., & Goldstein, L. (2003). Launching language: The gestural origin of discrete infinity. Teoksessa M. H. Christiansen & S. Kirby (toim.), *Language evolution* (s. 235–254). Oxford: Oxford University Press.

### EFFECT OF LANGUAGE BACKGROUND ON IMITATING VOWELS

Lotta Alivuotila, Department of Oral and Maxillofacial Surgery, Institute of Dentistry, University of Turku

Janne Savela, Department of Information Technology, University of Turku

Olli Aaltonen, Department of Speech Sciences, University of Helsinki

This study examines the relationship between speech perception and imitation from the point of view of language background. The aim of the study is to find out whether the different phonological systems of Finnish and Spanish speakers also induce differences in imitation results along the same acoustic continuum of synthetic vowel stimuli. The goal is to investigate whether speakers are able to imitate the acoustic signal alone or do they imitate the vowels on the basis of learned phonemic models in the brain. The results indicate that language background affects the way speech sounds are perceived and imitated. The same synthetic stimuli activated different linguistic models in the brains of Finnish and Spanish speakers. Respectively, the imitation results were different in the two groups. In conclusion, the experience-based memory traces for native-language speech sounds are not easy to ignore while imitating even if one consciously intends to do so.

**Keywords:** Imitation, vowels, formants.