

KIELTEN VOKAALIT KUULOANALOGISESSA VOKAALIKARTASSA

Antti Iivonen, Helsingin yliopisto, Puhetieteet

Tavoitteena on yhdistää tutkimustuloksia kokonaiskuvaksi kielten vokaalilaatujen osoittamiseksi. Ortografiset vokaalimerkit ja fonologian foneemisymbolit ovat epätarkkoja kielten vertailuun. Kansainvälisissä foneettisissa vokaalisymboleissakin on puutteita: tarpeellisia perussymboleja puuttuu ja toisinaan niillä symboloidaan foneemeja, vaikka merkki ei vastaa todellista ääntämistä (etenkin englantia koskevissa tutkimuksissa). Vokaalien havaitseminen perustuu äänen akustisiin formantteihin. Audittiivinen ERB-asteikkoinen F1/F2-formanttikartta simuloi kuulohavaintoa. Siten vokaalien suhteellisia sijainteja ja audittiivisia etäisyyksiä voidaan arvioida. Minimietäisyys vaihtelee kahden lähimmän vokaalikeskiarvon välillä miespuhujilla suuren foneemiparadigman kielissä 0,45 – 1,47 ERB (keskiarvo 1,06 ERB). Etäisyys pienenee, jos puheen selvyys vähenee tai jos aineisto käsittää kokonaisia lauseita, lukutekstiä tai spontaania puhetta. Lisäksi vokaalin foneemisella lyhyt/pitkä -vastakohtaisuudella on vaikutusta: lyhyet keskistyvät, pitkät perifeeristyvät. Kun reduktion ja pitkän vokaalikeston vaikutus poistetaan, minimietäisyys kasvaa keskiarvoon 1,44 ERB. Valtaosin suomen etuvokaalien yksittäisesiintymät mahtuvat 1,3 ERBin suuruiseen, keskiarvoa kuvaavan ympyrän sisään. Takavokaalien hajonta on suurempi. Ääntöväylän pituudesta johtuvat miesten ja naisten erot voidaan normaalistaa keskenään vertailukelpoisiksi.

Avainsanat: vokaalien audittiivinen laatu, ERB-asteikkoinen formanttikartta, vokaalien minimietäisyys, vokaalien laadun variaatio, puhujan normalisointi

JOHDANTO

Etenkin suomenkielinen puhuja ajattelee vokaaleja ja konsonanteja kirjaimina ja on vakiinnuttanut mieleensä käsityksen, että – enimmäkseen latinalaisperäisten – kirjainmerkkien ääntäminen suomen kielen mukaan on universaalisesti oikein. Kieltenopiskelija tosin tietää, että opittavan kielen kirjainmer-

kit voivat ilmaista eri vokaaleja kuin suomes-
sa. Ruotsissa kirjain <o> äännetään tiettyissä sanoissa kuten suomen [o] mutta usein [u:] kuten *bo* 'asua'. Englannin <u> on vain tiettyissä sanoissa lähellä suomen [u]:ta kuten *butcher* 'teurastaja'. Sanassa *butter* 'voi' se kuulostaa suomen *a*-vokaalilta. Ranskan <u> ei ole suomen [u] vaan [y] kuten *chacune* 'jokainen'.

Tiedon lähteet vokaalilaatujen ominaisuuksista ovat oma introspektiivinen tietoisuus, puheen (toisten ja oman puheen) havaitseminen, puhujan visuaalinen tarkkailu sekä kokeelliset tulokset artikulaatiosta, akusttiikasta ja aivovasteista. Vasta 1800-luvun lop-

Antti Iivonen
prof., emeritus
Ristihaantie 8A 18
02750 ESPOO
GSM 050-5331333
antti.iivonen@kolumbus.fi

pupuolella syntyneet kansainvälinen foneettinen kirjoitus (IPA) sekä kokeellinen tutkimus alkoivat tuottaa aikaisempia luotettavampia tuloksia. Miten vokaalilaatuja pitäisi tutkia objektiivisesti? Yhtenäisen viitekehysten ja metodiikan etsiminen on tärkeää kielten vertailun kannalta. Tämän kirjoituksen tavoitteena on punoa useita aiheeseen liittyviä näkökulmia yhteen laajemmaksi kokonaisuudeksi. Näkökulmat koskevat fonologian käsitteitä, foneettisia symboleita, puheen akustiikkaa ja formantteja vokaalien havaitsemisen perustana, kuuloanalogisten asteikkojen kehitystä ja ERB-asteikkoista formanttikarttaa. Erilaiset vokaalifoneemien yksittäisiin tyyppien vaihtelutekijät johtavat kysymyksiin. Mikä on foneemin ja sen realisaation suhde? Mikä on vokaalien minimaalinen auditiivinen ero? Onko mies- ja naispuhujien vokaaleissa eroja? Pysykö kielen vokaalijärjestelmä samana sukupolvesta toiselle? Suomen osalta kirjoituksessa käsitellään ennen julkaisematonta aineistoa. Vertailuissa on käytetty muita kieliä koskevia julkaisutietoja uudella tavalla. Kohteena ovat vain perusvokaalit, joten esim. nasaaliset ja äänenlaatuun liittyvät piirteet jäävät käsittelemättä. Kirjaimet merkitään kulmasulkein < a A >, foneemit vinoviivoin /a/, toisinaan *a*-vokaali ja foneettiset laadut, ääntäminen, hakasulkein ja foneettisin symbolein [a a a:].

FONEETTINEN KIRJOITUS JA FONOLOGIA

Oppikirjoissa ja sanakirjoissa oikeinkirjoituksen ja ääntämisen erot selitetään nykyisin yhä useammin käyttäen kansainvälisen foneettisen kirjoituksen symboleja. Erityisesti saksan osalta ei kuitenkaan näin tehdä ehkä siksi, että käyttäjä tietää kirjoitusasun ja ääntämisen suhteen olevan johdonmukaista ja päätteellisen perusteella, miten äännetään. Saksan tiukkojen ja höllien vokaalien erottelu poikkeaa

suomesta, mikä aiheuttaa foneettisia eroja: vrt. esimerkiksi pareja *bieten/bitten, fühlen/füllen* (Iivonen, 1997a; Iivonen & Tella, 2009: 278). Alempana esitetyn brittienglannin esimerkin perusteella näyttää puolestaan siltä, että englannin oppikirjojen vokaalimerkit ovat ensisijaisesti foneemimerkkejä, eivätkä eräät niistä enää täysin ilmaise foneettista nykytodellisuutta.

Saksan ja ranskan puolisuuspuoleiden vokaalien foneettisiksi symboleiksi IPA:n perusmerkit [e ø o] ovat osuvia. Toisaalta suomen kannalta IPA:ssa ei ole perussymbolia välivokaaleille *e, ö ja o* (jotka tässä on esitetty ortografisesti). Nämä välivokaalit ovat kuitenkin tavallisempia kielissä kuin IPA:n puolisuuspuolet ja puoliväljät. Esimerkiksi IPA:n symboli [ø] suomen /ö/:lle ei siis ole täysin korrekti (tarkemmin Iivonen & Huhe, 2005; Suomi, Toivanen & Ylitalo 2006: 96). -- On erikoista, että sekä ranskan sanan *rouge* että englannin sanan *rouble* vokaali ilmaistaan tavallisesti symbolilla [u:]. Ranskassa, suomessa ja ruotsissa se onkin suppea takavokaali mutta brittienglannissa suppea pyöreä keskivokaali (ks. kuvia 9–10 jäljempänä). -- Kaksiulotteisessa vokaalinelikulmiossa ei voida erottaa vokaalien pyöreysulottuvuutta, labiaalisuutta. Niinpä IPA osoittaa pyöreiden vokaalien symbolit samoissa pisteissä kuin vastaavat laiveat vokaalit. IPA tarjoaa väljimmälle asteelle symbolit vain kahdelle vokaalille, etiselle [a]:lle ja takaiselle [ɑ]:lle. Tämä on sikäli omituista, että maailman kielissä epäilemättä keskinen ”a” on yleisin väljä vokaali. Kansainvälisen foneettisen seuran aikakauslehdessä tästä ongelmasta onkin käyty äsken keskustelua ja pohdittu omaa symbolia keskiselle tyyppille (Barry & Trouvain, 2009). Väljien vokaalien symbolit voisivat olla etinen [æ], keskinen [a] ja takainen [ɑ]. Jäljempänä kuva 6 näyttää osoittavan, että kolmea selvästi kuulohavainnossa toisistaan erottuvaa vokaalityyppiä ei voi olla. Tietyissä kielessä esiintyvä keskinen

”a” olisi näin ollen ainoa, tosin vaihtelevaa-
tuinen vokaali.

Fonologiassa kielikohtaisia **foneemeja** ilmaisevat symbolit, esimerkiksi /e/, puolestaan ovat usein liian karkeita kielten foneettiseen vertailuun, koska niiden tarkoitus on vain erottaa abstraktit vokaalityypit toisistaan kielen sisäisessä järjestelmässä. Kielen järjestelmän foneemi on ensisijaisesti abstrakti: kielenkäyttäjällä on vakaa käsitys foneemin samuudesta sen esiintymisvarianttien suu-
restakin vaihtelusta huolimatta. Esiintymisvariantin käsite on sama kuin foneemin toteuma, realisaatio, (yksittäis)esiintymä (vrt. englannin käsitteitä *type/token*). Fonologian foneemeilla ei varsinaisesti ole muuta ”aikaa” kuin niiden fonotaktinen peräkkäisyys sanoissa, mutta niiden akustisilla realisaatioilla on mitattava kesto.

Foneemillakin on kuitenkin jokin laatu, jota kuulija voi verrata toisen foneemin laatuun. Varsinkin englannin fonologisissa kuvauksissa vokaalifoneemit ilmaistaankin foneettisilla symboleilla. **Allofonin** kuuluminen fonologian käsitteisiin sisältää tunnustuksen foneemin tosiasiallisesta laadullisesta vaihtelusta. Fonologiassa foneemeja kuvataan binaaristen **distinktiivisten piirteiden** teorian mukaisilla piirteillä. Ne eivät ole riittävän konkreettisia kielten vertailuun. Foneemi ja distinktiivinen piirre ovat luokittelevia, diskreettejä, vailla jatkumollista ominaisuutta. Foneettisten kirjoitusjärjestelmien artikulaatioon perustuvat **vokaalisymbolit** ovat lähempänä todellisuutta, mutta niilläkin on rajoituksia. Koska konkreettisen artikulaation tutkiminen on etenkin anatomisen kielen asennon osalta hankalaa, litteroija valitsee käytännössä symbolit kuulon avulla (vrt. Suomi, Toivanen & Ylitalo 2006: 89).

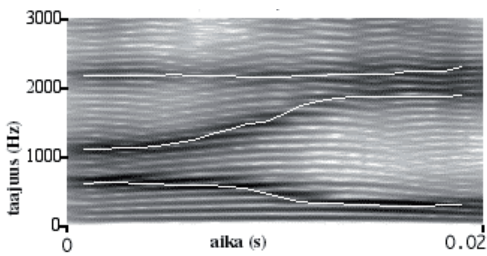
ÄÄNTÖVÄYLÄ, YLÄSÄVELET JA FORMANTIT

Artikulaatio on mentaalihistoriallisesti lähempänä ihmisen tietoisuutta kuin akustiikka. Siksi on ymmärrettävää, että kielitieteessä vokaalit määritellään artikulaatiotermien avulla. Käytännössä fonologit eivät kuitenkaan tutki konkreettista ääntämistä vaan luottavat introspektioon ja välittömään auditiiviseen havaintoon.

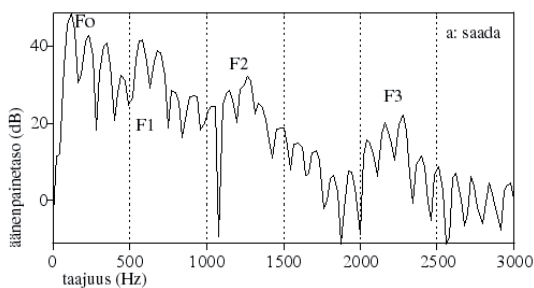
Artikulaatiotermien määritely vokaali ei ylipäänsä ole äänenä vielä kuultavissa. Artikulaation lisäksi ääntöväylän läpi pitää virrata ilmaa, joka tuottaa soinnin tai soinnittomuuden. Artikulaatiotermi ovat siten luokittelevia, eivätkä ne tavoita täysin vokaalien muodostusperiaatetta. Todellisudessa (soinnilliset) vokaalit ovat äänihuulissa tuotetun osasävelsarjan muunnelmia (kuva 1). Äänihuulten tuottama sointivärehely ei tuota vain äänen peruskorkeutta vaan myös kokonaisen yläsävelsarjan (kuva 2). Vokaalien muuntelu saadaan aikaan kielen ja huulten kulloistakin asentoa vaihtelemalla ja nasaalivokaaleissa avaamalla lisäksi nenäväylä. Tämä ääntöväylän muuntelu tuottaa vaihtelevia resonansseja, mikä johtaa kuulohavainnossa vokaalilaatujen vaihteluun (Fant, 2004: 29; Vainio, Palo, Aalto & Laine, 2009). Resonanssien vaihtelu aiheuttaa vokaaleille tyypillisiä yläsävelten voimakkuustasojen vaihtelua. Lyhytaikaisspektri (LPC) ilmaisee valittuna ajankohtana esiintyvien yläsävelten keskinäiset voimakkuuserot (kuva 2). Vain harvat ihmiset havaitsevat ja tiedostavat yksittäisiä yläsäveliä. Voimakkaimmat yläsävelet yhdessä niiden lähimpien naapurien kanssa muodostavat formantteja, joiden merkitys vokaalien havaitsemisen perustana on todistettu paitsi formanttisyntetisaattoreilla myös luonnollisen puheen vokaalien suodatuskokeilla.

Akustisen kuvauksen mahdollisuus reaalistui laajemmin vasta 1940-luvulta lähtien

äänispektrografian edistymisen vuoksi. Formantteja voidaan visualisoida spektrogrammeina (kuva 1) ja lyhytaikaisspektreinä (kuva 2). Kuvat on laadittu Praat 5.0-puheenkäsittelyohjelmalla (Boersma & Weenink). Spektrogrammi kuvaa äänivärähtelyä ajan (s) ja taajuuden (Hz) koordinaatistossa. Tummuusasteet ovat verrannollisia yläsävelten voimakkuusasteisiin eri taajuuksilla. Spektrogrammeissa formantit näkyvät vaakasuorina pidemmillä

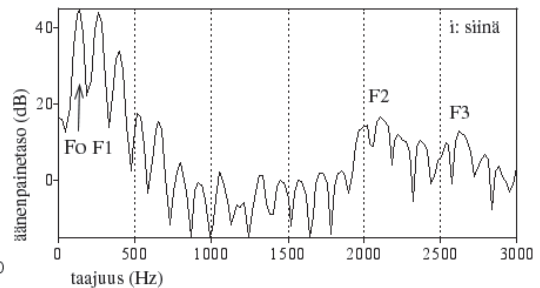


Kuva 1. Spektrogrammi (aika; sekuntia/taajuus; Hz) miespuhujasta TP:n ääntästä diftongista *ai* sanassa *aisti*. Vaakasuorat tummat juovat ovat osasäveliä. Niistä tummemmat eli vahvistuneet osasävelet lähimpine naapureineen ilmaisevat formanttien sijainnin. Kolme valkoista juovaa spektrogrammissa ovat Praat-ohjelmalla lasketut formantit F1, F2 ja F3.

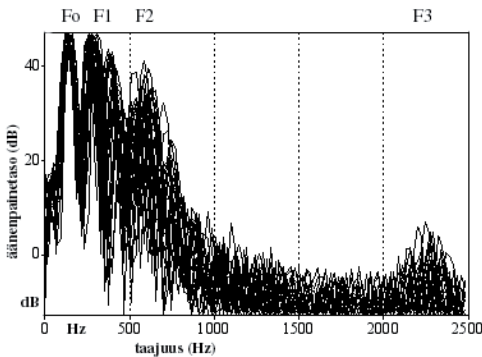


aikavälillä, kun taas lyhytaikaisspektreissä vokaalin paineallostta otetaan lyhyt noin 5–30 millisekuntia käsittävä otos, jolloin formantit erottuvat spektrin yläreunassa vuorijonon huippuja muistuttavana kuviona tai vahvistuneina osasävelryhminä. Analyysi esitetään taajuuden (Hz) ja äänenpainetason (dB) koordinaatistossa (Vainio ym., 2009). Automaattinen mittaustarkkuus on parantunut, mutta tarkistavat vertailumenetelmät ovat aiheellisia (vrt. jäljempänä). Osasävelten kuullun äänekkyyden kannalta taajuusalue noin 500 – 4000 Hz on herkintä. Vastaavasti psykoakustinen voimakkuus alueilla 100 – 500 Hz ja yli 4000 Hz on heikompaa.

Formanttien mittaamisessa ilmenee toisinaan hankalia tapauksia. Usein [u:]n esiintymisissä näyttää aluksi siltä kuin alin osasävel eli perusvärähtely tai perustaajuus F_0 ("af nolla") ja formantit F1 ja F2 olisivat sulautuneet yhdeksi formantiksi. Kuvassa 3 näkyy esimerkki suomen [u:]-vokaalin alimpien formanttien paikantamisesta. Kuvassa on asetettu irrallaan äännettyjen 30 sanan painollisen [u:]-vokaalin tavoitevaiheen spektrit päällekkäin. Äänihuulten tuottama perussävel F_0 ja sitä seuraava osasävel näkyvät voimakkaimpina osasävelinä. Perussävel on niistä useimmiten voimakkain,



Kuva 2. Vokaaleista [a:] ja [i:] Praat-ohjelmalla lasketut lyhytaikaisspektrit (taajuus; Hz/äänepainetaso; dB). Koesanat *saada* ja *siinä*. Näkyvissä ensimmäiset noin 20 osasäveltä, joista ensimmäinen on äänihuulivärähtelyn aikaansaama perussävel F_0 . Formantit F1, F2 ja F3 näkyvät yläsävelten vahvennuksina. Miespuhujasta TP.



Kuva 3. Kolmenkymmenen eri sanan painollisista [u:]-vokaaleista analysoidut lyhytaikaisspektrit päällekkäin asetettuina. Analyysit Praat-ohjelmalla. Aikaikkuna 33 ms. Miespuhujia TP. Perussävelen (Fo) ja formanttien (F1, F2, F3) sijainnit osoitettu.

mikä aiheutuu äänihuulten värähtelyvoimasta. Formantit F1 ja F2 sijoittuvat neljän seuraavan osasävelen alueelle ja niiden synty selittyy puolestaan ääntöväylän resonansseista. Tarkkaan ottaen formantti F1 sijaitsee siten toisen ja kolmannen osasävelen välissä, lähempänä toista osasäveltä. Ilmiötä käsittelevät jo Peterson ja Barney (1952: 123). F2 sijoittuu noin viidennen osasävelen kohdalle.

Formantit on paikannettu joko spektrogrammeja käsin mittaamalla (Wiik, 1965), myöhemmin osin automaattisia keinoja tai täysin automaattisia ohjelmia käyttäen. Ohjelmien mahdollisuudet edistää mittaus-tarkkuutta vaihtelevat. Yhteistä on se, että on määriteltävä mittauksen ajankohta, joka parhaiten edustaa vokaalin laatua. Vokaalin siirtymäkohdissa ja vokaalin aikanakin laatu voi muuttua ympäristökonsonanttien vaikutuksesta. Seurauksena on formanttien alku- ja loppuliikumia alas- tai ylöspäin. Siten varianttien määrä kasvaa.

Duckworth, McDougall, de Jong ja Shockey (2011) vertasivat kolmen brittilaboratorion tuloksia kolmen ensimmäisen formantin mittauksista, jotka kohdistuivat samaan ää-

nitettyyn puhemateriaaliin. Miespuhujia oli 40 ja analyysit tehtiin edellä mainitulla Praat-ohjelmalla. Kaikilla oli käytössä sama versio. Puhujat toistivat kuusi kertaa lauseet, joista jokaisessa esiintyi suuraakkosin jokin kuudesta *bVd*-sanasta (esim. *HARD*). Samoja analyysiperiaatteita sovellettiin. Osoittautui, että kahden laboratorion tulosten välillä oli huomattava korrelaatio. Formantti F2:n Pearsonin korrelaatiokertoimet vaihtelivat kuuden koesan vokaalien välillä 0,883–0,956, F1:n 0,765–0,936. Heikoin tulos oli F3:n mittaustuloksissa: 0,568–0,890. Yllättävää oli, että osalla puhujista tulokset aiheuttivat selvästi suurempaa hajontaa. Eroihin lienee vaikuttanut lopullinen mittaustapa. Kolmannen laboratorion tulokset poikkesivat enemmän.

KUULOANALOGISTEN FORMANTTIKARTTOJEN KEHITYKSESTÄ

Kuulijalle ”todellinen” vokaali on sen auditiivinen laatu, jonka hän tulkitsee joksikin kielten foneemiksi. Vokaalispektrin auditiivinen vaste on kuulojärjestelmän ja aivojen muunnosprosessien aiheuttama. Tätä muunnosprosessia voidaan simuloida ja lähestyä siten kuulijan havaintoa. Kehitys johti logaritmisesta ja mel-asteikon kautta Bark- ja ERB-asteikkoon.

Kuuloanaloginen formanttikartta tavoittelee artikulatorista vokaalinelikulmiota muistuttavaa graafista esitystä formanttien havaitsemista simuloivien psykoakustisten asteikoiden avulla (taustahistoriasta ks. Iivonen, 1994). Äänispektrografian keksimisen (1946) jälkeen useat tutkijat ehdottivat kahden formantin karttaa. Seuraten ajan kuulotutkimusta Peterson ja Barney (1952) sovelsivat F1/F2-kartassa Koenigin asteikkoa ja täyttivät logaritmisesta asteikkoa. Keskeiseksi ongelmaksi on muodostunut asteikkojen taajuusalue noin 0–500 Hz. Varhaisemmissa asteikoissa sekin tulkittiin ilman selvää näyttöä logaritmisiksi.

Uudempi tutkimus osoitti, että mel-asteikko kuvastaa tämän taajuusalueen osalta todellisuutta paremmin vastaavaksi ja että 100 melin kriittinen kaista eli yksi Bark on aiheellista ottaa käyttöön (Zwicker, 1961; Fant, 1973: 48; 1978; Suomi, 1984). Useita vokaalikarttoja on julkaistu käyttäen Bark-asteikkoa (esim. Lindblom, 1986; Iivonen, 1994, 1997c; van Bergem, 1995; Kuronen, 2000; Eek 2008).

Logaritminen asteikko laajentaa kohtuuttomasti suppeiden ja puolisuppeiden vokaalien aluetta väljempien vokaalien kustannuksella. Mel- ja Bark-asteikko luovat perustan käsitykselle, että formanttien F1:n ja F2:n alueet edustavat samaa psykoakustista mitta-kaavaa ja että vokaalien laadullinen erottelu on tarkempaa F2:n kuin F1:n alueella. Itse asiassa kokeneet foneetikot ovatkin saattaneet laajentaa julkaistuissa F1/F2 kartoissa F1-asteikkoa ilmoittamatta perustelua tähän (Ladefoged & Maddieson, 1990: 96, jossa 100 %:n laajennus; van Bergem 1995: 18). Perusteltu muunnelma Bark-asteikkoisesta kartasta on formantti F1:n asteikon laajennus 60 %:lla F2:n suhteen (Iivonen, 1994, 1997a; Kuronen, 2000; Iivonen & Huhe, 2005). Myöhemmässä kirjassaan Ladefoged ja Maddieson (1996) esittävät useita asteikoiltaan vaihtelevia karttoja. ERB-asteikko (*equivalent rectangular bandwidth*) poistaa Bark-asteikon puutteita (Moore, 1997; ks. myös Hawkins & Midgley, 2005). ERB-asteikko eroaa Bark-asteikosta siten, että ERBin arvo on traditionaalista kriittistä kaistanleveyttä kapeampi. Härmä, Karjalainen, Savioja, Välimäki, Laine & Huopaniemi (2000: 5) arvioivat ERB-asteikon Bark-asteikkoa tarkemmaksi, koska psykoakustiset testit osuvat yksiiin fysiologisten ja neuraalisten mittausten kanssa. Muunnoskaava hertseistä ERBeiksi (Moore, 1997; Hawkins & Midgley, 2005) on:

(ERB) = $21.4 \cdot \log(10)[0.00437 \cdot f + 1]$, jossa f = formantin taajuus (Hz)

Tavallisin formanttikartta on laadittu kahden ensimmäisen formantin taajuusarvojen mukaan. Kielellisesti ja foneettisesti vokaaleista kiinnostuneet sijoittavat origopisteen oikealle ylös, koska näin saadaan aikaan osittainen mutta ilmeinen analogisuus artikulatorisen vokaalinelikulmion kanssa. Joissakin julkaisuissa kolmas formantti on saanut erillisen F2/F3-kartan (esim. Wiik, 1965). Parempaa kuuloanalogisuutta tavoittelevat ovat korvanneet formantin F2 toisen ja kolmannen formantin painotetulla keskiarvolla F2':lla ("äf kaksi pilkku"). Selityksenä on, että F2:n ja F3:n sijaitessa lähekkäin ne muodostavat auditiivisesti toisiinsa integroituneen kokonaisuuden (Fant, 1978, 2004; Kuronen, 2000:72; Eek, 2008: 87). Joissakin julkaisuissa on käytetty 3-ulotteista karttaa (F1/F2/F3) (esim. Lindblom, 1986: 18). Muitakin ehdotuksia esiintyy julkaisuissa (ks. aiheesta enemmän Iivonen, 1994). Jäljempänä esitetyissä kartoissa näkyy myös ERB-asteikkoon suhteutettu hertsiasteikko oikealla (F1) ja alhaalla (F2).

Jäljempänä tarkastellaan vokaalin formanttien ERB-muunnoksen mukaisia auditiivisia karttoja. Ansaittua uutta eloa saavat sen myötä kirjallisuudessa julkaistut formanttiaineistot, mitkä näin mahdollistavat kielten vertailun. Hyötynä on myös mahdollisuus tutkia erilaisia variaatiota aiheuttavia tekijöitä. Menetelmän toimivuutta ilmaisee erityisen vakuuttavasti diftongien liikeradan osoittaminen kartassa (kuva 8).

Auditiivinen formanttikarttakaan ei ole riippumaton käytetystä puheaineistosta osoittamaan kielikohtaisten "standardivokaalien" sijainnit kartalla, koska äännöksen satunnaisvaihtelu, formanttien mittaustekniikka, puhujan ääntöväylän pituus, puheen selvyysaste ja konsonanttikonteksti vaikuttavat esiintymien sijaintiin. Vaikka julkaistuisa tutkimuksissa käytetty englannin sanasarja *heed, hid, head* jne. onkin tuottanut hyödyllisiä tuloksia ja poistanut kontekstuaalisen

vaihtelun, sanojen toistojen määrä tuntuu liian pieneltä vokaalikeskusten luotettavan tuloksen selville saamiseksi. Olisi tarpeellista tutkia toistojen hajontaa ja konsonanttikon-tekstien vaikutusta suuremmilla aineistoilla.

KUULON EROTTELUKYKY VOKAALIEN HAVAITSEMISESSA JA TUNNISTAMISESSA

Synteettisten ärsykkeiden erotettavuus

Kun ärsykeaineiston vaihtelukijät minimoidaan, saadaan vokaalinkaltaisille, synteettisille ärsykkeille **juuri-ja-juuri-havaittava ero** (*the just noticeable difference, JND*). Flanaganin (1955) mukaan kahden formantin synteettisillä vokaaliärsykkeillä, jotka erosivat vain formanttien korkeuksien suhteen, saatiin erotuskynnukseksi 3–5 % formantin korkeudesta. Nord ja Sventelius (1979) saivat samanlaisella koejärjestelyllä lähes saman tuloksen. Tästä seuraa, että vokaalien mahdollisessa F1/F2-avaruudessa olisi 400–500 toisistaan erottuvaa vokaalia avaruuden rajoista riippuen. Edellisiä entistä sofistikoitumalla koejärjestelyllä Hawks (1994) päätyi siihen yllättävään tulokseen, että formanttien F1 ja F2 erottelutarkkuus on paljon tarkempaa: niinkin pieni kuin noin 1,9 % formantin taajuudesta.

Kuvatut erottelutarkkuudet vaihtelevat siis välillä 1,9–5 %. Niiden vaikutus vokaalien etäisyyksiin ilmenee seuraavasta vertailusta. Esimerkkinä on etuvokaali [e]:n variantti: [e1] = F1 500, F2 = 2000. Pienimmällä erottelutarkkuusarviolla siitä erottuvan vokaalin [e2] formantit ovat: F1 = 510, F2 = 2038 Hz ja vastaavat ERB-erot ovat 0,13 (F1) ja 0,16 (F2). Suurimmalla erotteluarviolla saadaan eroiksi 0,31 ja 0,41 ERB.

Jäljempänä osoittautuu, että 5 %:n tarkkuudellakaan ei päästä luonnollisten vokaalien minimaalisten etäisyyksien vaatimalle tasolle. Mainituissa kolmessa tutkimuksessa tarkasteltiin formanttien arvoja hertsiasteikolla. Tulokset heijastanevat joka tapauksessa kuulolle ominaista erottelutarkkuutta äänieristetyssä tilassa kielellisestä ja foneettisesta havaitsemisesta riippumattomalla aistimustasolla silloin, kun ärsykkeiden muita muuttujia kuin kahta ensimmäistä formanttia ei varioida. Hawksilla oli mukana formantti F3, mutta sen taajuutta ei muunneltu. Kun ärsykkeiden kompleksisuus lisääntyy, erotuskynnys kasvaa (van Wieringen, Boersma & Pols, 1993). Kielten foneemi-inventaarin kokoon vaikuttaakin nimenomaan se, että vokaalit on erotettava foneettisesti ja foneemisesti myös tavanomaisessa puheessa huolimatta kontekstuaalisesta ja prosodisesta vaihtelusta sekä ympäristön kohtuullisesta kohinasasta.

Puhuja tähtää artikulaatioon, jonka äänellinen tuotos vastaisi hänen kuulomuistissaan olevaa tavoitetta. Puheessa kuitenkin foneemia edustavat yksittäisesiintymät varioivat. Kuhl (1991) on tutkinut etenkin puhesynteesillä vokaalien **prototyyppjejä**. Niiden parhaat variantit ovat foneettisten kategorioiden (foneemien?) poikkeuksellisen hyviä edustumia (*exceptionally good instances of phonetic categories*). Ne synnyttävät metaforisesti ilmaisten **magneettisia efektejä**. Itse asiassa Liljencrants ja Lindblom (1972: 856) käyttivät jo magneettivertausta (*“vowel magnets”*), tosin jonkin verran eri merkityksessä. Kuhlin kokeessa ilmeni, että kuulijoiden aivovasteissa ei ilmennyt eroja pienten vokaaliärsykkeiden poikkeamien välillä, kun ne sijaitivat lähellä prototyyppiä. Käyttämällä 32 synteettistä *i*-vokaalivarianttia non-prototyyppi löytyi,

vokaalipari	F1	F2 Hz	erottelutarkkuus	F1 Hz	ERB	F2 Hz	ERB
e1 – e2	500	2000	1,9 %	510	0,13	2038	0,16
e1 – e2	500	2000	5 %	525	0,31	2100	0,41

kun aivovasteissa ilmeni havainnon heikentyminen prototyypistä. Prototyypin ja non-prototyypin eroksi tuli mel-asteikolla noin 100 meliä eli 1 Bark. ERB-asteikolle muutettuna tämä on 1,37 ERB. Tutkimuksessa koehenkilöinä oli aikuisten ohella 6-kuukautisia vauvoja. Tulosten mukaan vauvojen erottelukyky oli parempi. Tämä ei ole nähdäkseeni odottamatonta. Lasten **foneeminen erottelukyky** on lapsilla siinä iässä vielä heikko. Foneemisuuteen Kuhl ei kuitenkaan viittaa vaan käyttää sen sijaan kategorian ja kategorisaation käsitteitä.

Soveltaen Kuhl'n ym. metodeja Aaltonen, Eerola, Hellström, Uusipaikka ja Lang (1997) tutkivat suomen *i*- ja *y*-vokaalien tunnistettavuutta sekä prototyypisyyttä varioimalla synteettisten ärsykkeiden F2:n taajuutta ja ”hyvyyden” astetta. Kuuntelukokeiden ja aivovasteiden tutkimus osoitti, että 13 suomenkielistä koehenkilöä erosivat tunnistettavuuden osalta toisistaan jakaantuen hyviin ja huonoihin luokittelijoihin. Kummankin ryhmän sisällä oli lisäksi vaihtelua. Bark-asteikolla mitattuna paremmin tunnistaneiden erotuskynnys oli alle 1 Bark. Huonommalla ryhmällä vastaava arvo oli suurempi kuin 1 Bark. Kuulijoiden erottelutarkkuudessa on siis eroja.

Raimo, Savela ja Aaltonen (2005) ovat tutkineet vokaalijärjestelmien prototyypisiä laatuja antamalla kuuntelijoiden arvioida synteettisten vokaalien auditiivisen laadun ”hyvyyttä” asteikolla 1–7. Eri kielten vaihtelevansuuruisten vokaalijärjestelmien parhaat vokaalipisteet on kuvattu mel-asteikkoisessa F1/F2-formanttikartassa. Rajoituksina menetelmälle tekijät pitävät nasaalisten, höllien ja painottomien vokaalien esittämisen ongelmallisuutta (mts. 174). Lyhyiden ja pitkien vokaalien eroja ei tutkittu. Se olisi tärkeä esimerkiksi saksan ja englannin tiukka/höllä-vastakohtaisuuden, esim. parin /i:/-/ɪ/ kannalta (vrt. saksan *bieten-bitten*; Iivonen,

1997a; englannin *beat-bit*; Peltola, 2003; Lintunen, 2004: 126).

Kansainväliset vokaalisymbolit ja auditiivinen erotettavuus

IPA:n **perusvokaalisymbolien** (ks. lähteet: IPA) lukumääräksi on noin sadan vuoden kuluessa vakiintunut 28. Mukana ei siis ole nasaali- sekä ym. muihin lisäpiirteisiin perustuvia vokaaleja. Tiukka/höllä -erotteluun perustuvat vokaalit ovat kuitenkin mukana. Suomalais-ugrilaisessa tarkekirjoituksessa (ks. lähteet: SUT) vastaava luku on 24. Nämä tiedot todistavat sitä, että kokenut foneetikko voi — syrjäyttäen pelkän foneemisen kuulemisen — kuulla tällaisen määrän eri laatuja. Toisaalta luvut ovat huomattavan pieniä synteettisten kokeiden avulla saatuu erottelukykyyn verrattuina. Vaikka ottaisimme mukaan mahdollisten foneettisten tarkkeiden käytön, jää lukumäärä pieneksi.

Liljencrants ja Lindblom (1972) pyrkivät universaaliennustukseen kielten vokaalien keskinäisistä sijainneista formanttikartalla, kun vokaalien fonologinen lukumäärä kasvaa. Lindblom (1986) päätyi korjaamaan mainitun artikkelin ennusteita. Artikkelin taulukko 2.2. käsittää IPA:n vokaaleja lähellä olevien kvasikardinaalivokaalien neljän ensimmäisen formantin teoreettiset arvot. Niiden laskemisessa Lindblom on soveltanut erittäin sofistikoitua psykoakustista taustatietoa vakiovaiheen vokaalispektrien havaitsemiseen vaikuttavista tekijöistä. Myös formanttien äänekkyyden soniasteikoilla oli laskelmissa mukana. Tavoitteena oli osoittaa 19 kvasikardinaalivokaalille formanttien arvot psykoakustisesti tasavälisinä (*equidistant*). Tekijä korostaa tulosten teoreettisuutta ja likiarvoisuutta useiden vaihtelutekijöiden vuoksi (mts. 26). Vokaalien F4 on vakio 3594 Hz, F3 vaihtelee välillä 2331 [u] – 3112 [i], mutta ilman [i]-vokaalia vain välillä 2331 [u] – 2482

[y]. Erotus on vain 151 Hz. Luonnollisissa vokaaleissa ilmenee samankaltainen tilanne: F3 ja etenkin F4 vaihtelevat eri vokaaleissa korkeudeltaan suhteellisen vähän, ja siten päätehtävä vokaalien tunnistamisessa lankeaa formanttien F1 ja F2 osalle.

Lindblomin aineistossa [i] ja [y] ovat saaneet lähes saman F1-arvon (i=255, y=263 Hz) ja F2 täsmälleen saman arvon (2191 Hz). Kuvassa 2.6. (mts. 28) [i] ja [y] saavat kuitenkin formanttien mel- ja soni-arvojen mukaan selvästi eri asemat. Ero saattaa siten perustua [i]:n korkeaan F3:n arvoon. Toisaalta nämä kaksi vokaalia ovat F1/F2-kartassa hyvin lähellä toisiaan ruotsinruotsissa (Kuronen, 2000: 119, 128), eikä ero suurene juurikaan, vaikka F2:n sijaan käytetään F2'-muuttujaa, joka simuloi F2:n ja F3:n psykoakustista integroitumista (mts. 72).

Lindblomilla on väljimmässä sarjassa vain keskivokaali, jolle hän antaa symbolin [a]. Tällainen käytäntö puuttuu kardinaalivokaaleista ja IPA-vokaalisymboleista. Molemmissa [a] on väljä etuvokaali ja [ɑ:] väljä takavokaali. Näitä suppeammassa sarjassa Lindblom antaa arvot vokaaleille [æ] ja [ɑ]. Olisiko jälleen ruotsi vaikuttanut, koska suomenkielisen kuulijan on vaikeahkoa erottaa toisistaan sanat *tala* ja *tåla*? Ruotsinruotsin pitkä [ɑ:] on lievästi labiaalistunut ja suppeutunut lähestyen [ɒ:]:ta. Kuronen (2000: 55 ja mts. muualla) merkitseekin mainittujen sanojen vokaalit [ɒ:] ja [o:]. Lindblomin symboleissa on joitakin kardinaalivokaaleihin kuulumattomia, kuten [æ]. Tämä merkki on kylläkin IPA:n symboli. Se on väljintä etuvokaalia hieman suppeampi, mutta kirjallisuudessa sillä on merkitty myös suomen *ä*-vokaalia ja englannin väljää etistä vokaalia esim. sanassa *mad*. Nykyisessä brittitutkimuksessa se on saanut käyttöä väljää [ɑ:]:takin väljimmässä vokaalilaadussa (kuvat 9 ja 10 jäljempänä).

Lindblomin 19 vokaalia on sijoitettu ERB-asteikkoiseen formanttikarttaan kuvassa 4

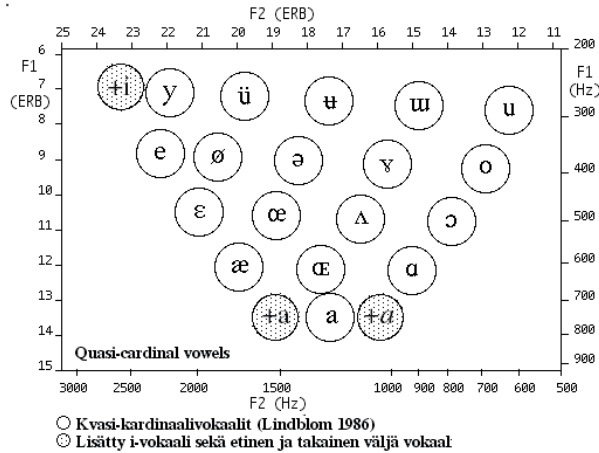
ympyröinä. Tasaetäisyydet toteutuvat varsin hyvin F1/F2-kartassakin. Lindblomin ilmoittamat kahden lähimpänä toisiaan olevien vokaalien [æ] ja [a] formanttiarvot ovat: [æ]: F1=616 ja F2=1293 Hz sekä [a]: F1=750 ja 1250 Hz.

Vastaavat ERB-arvot ovat:

[æ]: **F1=12,14** ja F2=17,61 ERB sekä [a] F1=**13,51** ja F2=17,34 ERB.

Vokaalipisteiden etäisyysero on 13,51–12,14 = 1,37 ERB, koska ero on suurempi F1-asteikolla. Ympyrän koko on kuvassa 1,3 ERB, koska suuremmalla halkaisijalla näiden kahden ympyrät leikkaisivat. Tätä voidaan pitää alustavana likiarvoisena todisteena minimaalisesta havaintoerosta kahden foneettisen vokaalilaadun kesken. Luonnollisten kielten yhteydessä tähän palataan tarkemmin jäljempänä.

Jotta kuvaus tekisi oikeutta enemmän kardinaalivokaaleille, etinen ja takainen väljä vokaali sekä [y]:stä eroava [i]-vokaali on lisätty kuvaan rasteroinnilla ja +-merkillä. Lisäys on oikeutettu myös useiden luonnollisten kielten F1/F2-sijaintien nojalla. Väljimpään sarjaan mahtuu kolme vokaalia. Kardinaalivokaalien artikulatorisessa nelikulmiossa pyöreät vokaalit saavat saman pisteen kuin laveat, mitä puutetta formanttikartalla ei ole, vaan pyöritys siirtää vokaalia kartan F2-akselilla hertseinä alemmaksi, kuvassa erbeinä oikealle. Lindblomin vokaalit voidaan rinnastaa irrallaan sana- ja lausekontekstista äännettyihin vokaaleihin. IPA:nkin 28 vokaalia mahtuisivat karttaan, koska ympyröiden väliin jää niille riittävästi tilaa.



Kuva 4. Lindblomin (1986) 19 kvasi-kardinaalivokaalia sijoitettuina ERB-asteikkoiseen formanttikarttaan. Rasteroinnilla ja +-merkillä merkityt lisätty Lindblomin vokaaleihin. Ympyrän koko = 1,3 ERB.

Luonnollisten kielten vokaalien inventaarit ja erotettavuus

Voidaanko sitten puhua abstraktien vokaalifoneemien minimaalisesta etäisyydestä? Fonologisesti looginen vastaus olisi: yhden erottavan piirteen ero riittää. Foneettisesti minimiero on määriteltävä akustis-auditivisesti. Kielenkäyttötilanteisiin kytkeytyvä erottelukyky ja siten minimietäisyyden vaatimus vaikuttavat rajoittavasti kielten perusvokaalien (foneemien) inventaarin laajuuteen. Kielissä, joissa on suuri vokaaliparadigma, ovat vokaalien vaihtelualueet pienempiä, kun taas kielissä joissa on hyvin suppea paradigma, on kontekstuaalisten varianttien määrä suuri. Tämä tunnetaan esimerkiksi australialaisesta kahden vokaalin itäaurenten kielestä ja kolmen vokaalin kaukaasialaisesta kabardista (Ladefoged ja Maddieson, 1996: 287–288). Eräissä eurooppalaisissa kielissä on suuri pääpainollisten monoftongien määrä kuten ranskassa (n = 11 oraalivokaalia), englannissa (11), ruotsissa ja virossa (9 lyhyttä) sekä suomessa (8 lyhyttä). Saksassa on 8 pitkä ja 7

lyhyttä painollista vokaalia. Pitkien ja lyhyiden vokaalien laatueroit ilmenevät selvästi saksan formanttikartassa suomea suurempina (Iivonen, 1997a; Iivonen & Tella, 2009: 278). Mainituissa kielissä lukumäärä on siis suurimmillaan alle puolet IPA:n ja SUT:n perusvokaaleista (28 ja 24).

Tasavälisen jakaumateorian (*dispersion theory*) mukaan kielten vokaalit sijoittuvat vokaaliavaruuteen tasaisin etäisyyksin (Liljencrants & Lindblom, 1972). Disner (1984) on tarkastellut 317 kielen otoksessa sellaisia kieliä, joissa teoria ei äkkiseltään näyttynyt pitävän paikkaansa. Tällaisia löytyi 43. Esimerkiksi hopi-kielessä symmetrian vaillinaisuuden syynä on puuttuva takainen suppea vokaali. Tämä näyttää puuttuvan myös brittienglannista (vrt. kuvia 9–10 jäljempänä). Laajan pohdinnan jälkeen Disner päättää kuitenkin, että 86 % kielistä noudattaa teoriaa selvästi, 10 % on lähentymässä tasavälisyyttä ja vain 6 % jää ulkopuolelle.

Formanttikartalla asiaa voidaan tarkastella puhujakohtaisten vokaalien kertaesiintymien ja keskiarvojen sekä useiden puhujien

Taulukko 1. Kirjoituksessa käytetyt formanttiaineistot. M= mies-, N= nais-, L= lapsipuhujia. RP = Received Pronunciation. Ilmaus "irraliset sanat" = ilman lausekontekstia.

Lähde	Kieli	Aineisto	Puhujat
Wiik 1965	suomi	irraliset sanat ja lausekehys;	5 M
	englanti	2–17 sanaa/16 x suomen vokaalit; N=756	5 M
Kukkonen 1990	suomi	pääosin 6–9 kaksitavuista sanaa/vokaali; puhujat afasiatutkimuksen kontrolliryhmä	4 M yli 50 v.
Kuronen 2000	suomi (kaksikieliset)	vapaamuotoisissa lauseissa painolliset pitkät 480+ lyhyet 480 = 960; Tampere	4 M
livonen; käsillä oleva artikkeli	suomi	20 x 16 x suomen vokaalit = 320; 20 x suomen diftongit /ai/ ja /äi/	1 M
Kuronen 2000	suomenruotsi	vapaamuotoisissa lauseissa painolliset pitkät 200 + lyhyet 160 = 360; Tampere	4 M
Kuronen 2000	ruotsinruotsi	pitkät 616 + lyhyet 400 = 1016; Helsinki ja Nyköping	4 M
Pols et al. 1973	hollanti	<i>hVt</i> -sanat	50 M
Koopmans-van Beinum	hollanti	useita konteksteja, mm. irraliset yksitavuiset sanat; tarkka puhujien valinta	5 M 5 N
Vihanta 1978	ranska	yksitavuiset sanat; vapaamuotoiset lauseet;	4 M
		kaikki puhujat pariisilaisia; ikä 19–27 v.	4 N
Deterding 1997	brittieng. RP	väitöskirjan 1990 irraliset <i>hVd</i> -sanat	8 M 8 N
Deterding 1997	brittieng. RP	BBC:n puheohjelmia (MARSEC) vähintään 5-10 esiintymää/vokaali	5 M 5 N
Deterding 2006	brittieng. RP	lukutekstit Pohjatuuli ja aurinko ja Poika ja susi; puhujat englannin lehtoreita	3 M
Hawkins & Midgley 2005	brittieng. RP	20 x 4 x 11 <i>hVd</i> -sanat 20 puhujaa, 4 ikäryhmää	20 M
Peterson & Barney 1952	amer.engl.	2 x 10 irralista <i>hVd</i> -sanaa; miespuhujat	33 M
		laajalta regionaaliselta <i>General American</i> , naispuhujat keskiatlanttiselta alueelta	28 N 15 L
Hillenbrand ym 1995	amer.engl.	12 <i>General American English hVd</i> -sanaa	45 M
		puhujat Michiganin osavaltiosta	48 N
			46 L
livonen 1997a	saksa	irraliset sanat; 5 esiintymää/vokaali	5 M

keskiarvoistettujen aineistojen avulla. Absoluuttisesta minimietäisyydestä ei voi olla kyse erilaisten vaihtelutekijöiden vuoksi, joista puhujien vaihteleva ääntöväylän pituus on suurin. Puhesyntetisaattorikin edustaa edellä kuvatuissa kokeissa yhtä ääntöväyläpituutta. Taulukkoon 1 on koottu vertailussa käytettyjen tutkimusten aineistotiedot. Aineistojen määrissä ja laaduissa sekä analyysimeteodeissa on ollut eroja.

Pienimmän mahdollisen etäisyyden likiarvoa voidaan etsiä sellaisista kielistä, joissa foneemiparadigma on suuri. Kuuden kielen julkaistujen formanttiaineistojen perusteella kahden toisiaan lähimpänä olevan vokaalikeskiarvon etäisyys on tällaisissa pareissa miespuhujilla keskimäärin 1,06 ERB ja keskihajonta 0,31 ERB (taulukko 2). Hollannissa ei-keskistyneiden vokaalien määrä on vain viisi (kuva 5), ja ne erottuvatkin hyvin toisis-

taan. Yleensä kielissä keskiarvojen etäisyydet ovat minimietäisyyttä suurempia (vrt. kuviin 6 ja 9–10), mikä tietenkin takaa suuremman erotettavuuden. Etäisyyden ollessa pienempi sekä F1:n että F2:n suhteen on syytä epäillä niin suurta laadullista samankaltaistumista, että kyseessä voi olla sulautuminen, ellei muuta akustista eroa ole.

Foneemiparadigman koon lisäksi etäisyyksiin vaikuttavat kaksi tekijää: vokaalin foneeminen pituusvastakohta ja puheen reduktioaste. Kumpikin vaikuttaa vokaalin perifeeristymis/keskistymisasteeseen F1/F2-kartalla. Ranskaa lukuun ottamatta mainituissa kielissä on joko suoranainen vokaalipituuden vastakohtaisuus (vrt. suomen *tuli/tuuli*; ruotsin *ful/full*), virossa jopa kolmen pituusasteen kesken. Viron vokaalien kohdalla on tehtävä huomautus Eekin yhdessä Meisterin (Eek, 2008: 99) kanssa tekemien mittausten perusteella. Viron erikoisuus on kolme foneemista kvantiteettiastetta. Vokaaliparissa /e-ö/ ero on kolmessa kvantiteettiasteessa 0,88 (Q1), 1,29 (Q2) ja 2,26 (Q3) ERB. Viron vokaalit /e ö õ/ ovat laadullisesti lähellä toisiaan F1/F2-kartassa. Olisiko tässä selityksenä hienoinen reduktio, koska aineisto oli lukupuhuntaa: kun puhenopeus kasvaa kesto vähenee ja siten vokaalien tavoitevaihe lyhenee. Selvyysaste siten vähenee. Bark-asteikkoisessa F2'/F1- kartassa mainitut vokaalit erottuvat paremmin (mts. 100).

Englannissa pitkäksi merkitty vokaali ei aina toteudu kestoerona (vrt. Peltola, 2003; Lintunen, 2004), mutta useiden foneemien transkriptiossa käytetään pituuden symbolia (*beat* /i:/). Reduktio saa aikaan vokaalin keskistymistä ja laadun heikkenemistä puhenopeuden lisääntyessä ja aineiston ollessa kokonaisia lauseita tai lukutekstejä. Kun aineistosta poistetaan tutkimukset, joissa on käytetty lauseita tai lukutekstejä (vrt. taulukot 1 ja 2), tilastolliset tunnusluvut muuttuvat vain vähän: keskiarvo = 1,11 ERB ja hajonta

= 0,30. Kun poistetaan lisäksi tutkimukset, joissa voidaan epäillä vokaalin keston vaikuttaneen keskiarvojen etäisyyksiin, jää jäljelle vain kolme: suomen lyhyet vokaalit (Kukkonen, 1990 ja Iivonen, 2011; tässä julkaisussa) sekä Iivonen (saksa, 1997a). Niissä minimietäisyyksien keskiarvo on 1,44 ERB ja hajonta 0,04 ERB. Kaikissa näissä puhujat tuottivat irrallisia sanoja. Riittävään tilastolliseen luotettavuuteen ei tämä riittäne, mutta suuntaantavaan hypoteesiin lienee perusteita: läheisimmän vokaaliparin etäisyys on riippuvainen vokaalien kestoista.

Brittienglannin osalta yksi käytetty vertailupari on /ʌ/-/ɑ:/ kuten sanoissa *bud* – *hard*. Erottavana tekijänä on laadun lisäksi pituus. Vertailu on kuitenkin kiintoisa useammasta syystä. Ensinnäkin /ʌ/ ei symbolina edusta väljää vokaalia vaan IPA:n mukaan puoliväljää laveaa takavokaalia. Toiseksi se näyttää formanttikartan todistuksen mukaan (jäljempänä kuvassa 10) olevan englannin RP-standardissa itse asiassa väljä vokaali ja lähellä takavokaalia /ɑ:/. Taulukon 2 mukaan ero on 0,45–0,99 ERB (Deterding 1997; 2006) ja eri ikäryhmissä 0,68–1,19 ERB (Hawkins & Midgley, 2005). Ikään sidonnaista tendenssiä etäisyyden muuttumiseen ei ilmennyt. Deterdingin (2006) aineisto perustui luettuun tekstiin, josta poimittiin selvästi äännettyjä sanoja, joissa kontekstikonsonanti vaihteli. Taulukossa 2 paria /ʌ/-/ɑ:/ edustaa sanapari *but* – *bath*.

Edellä kuvattujen synteettisten vokaalien erottelutarkkuus vaihteli välillä 1,9–5 % (vas-taten suurimmillaankin vain 0,41 ERB:n etäisyyttä), kun koejärjestelyn ärsykkeet olivat luonnollisiin vokaaleihin nähden yksinkertaistettuja. Lindblomin aineistossa minimaalinen vokaalietäisyys on 1,3 ERB. Kuhlin (1991) ja Aaltosen ym. (1997) synteetikokeissa erottelutarkkuus oli noin 1 Bark (= 1,37 ERB). Kun kunkin luonnollisen kielten aineistosta laskettiin kahden lähimmän

vokaalifoneemin keskiarvon pienin etäisyys, se vaihteli välillä 0,9–1,61 ERB, kun pituuseroa ollut. Kontekstista irrallaan äännettyjen sanojen minimaalinen vokaalietäisyys oli 1,44 ERB. Näin erottelutarkkuutta erilaisilla ehdoilla voidaan kuvata likiarvoisesti vaihteluvälillä 0,16 ERB (= yksinkertaisten synteettisten formanttiärsykkeiden 1,9 %:n erottelu) –1,44 ERB (= irrallaan äännettyjen sanojen painolliset vokaalit keskimäärin). Ehdotonta minimaalista erottelutarkkuutta ei siis voi ilmoittaa, vaan on puhuttava likiarvosta, joksi sopisi taulukon 2 luonnollisten vokaalien keskiarvo 1,06 ERB, jolloin mukana on vokaaliparin kestoeron vaikutus. Jos kestoeroa ei sallita, on likiarvo noin 1,4 ERB. Jäljempänä kuvissa sovelletaan 1 ja 1,3 ERB:n kokoisia

ympyrähalkaisijoita etäisyyksien havainnollistamiseksi.

Mikäli minimaalinen ero ei toteudu F1/F2-kartassa selvästi äännettyihin sanoihin perustuvassa aineistossa, on aiheellista epäillä kahden vokaalikategorian laadullista samankaltaistumista tai sitä, että ero on olemassa mutta perustuu yhteen tai useampaan lisäpiirteeseen. Kuvan 5 hollannin yksitavuisista, 50 miespuhujan /hVt/-kontekstissa ääntämissä vokaaleissa näyttää siltä, että pareissa /ε/-/i/, /œ/-/ø/ ja /o/-/ɔ/ minimietäisyys ei toteudu (numeerinen formanttiaineisto: Pols, Tromp & Plomp, 1973). Näissä pareissa ensimmäinen jäsen onkin pitkä vokaali, joka kirjoituksessa merkitään kahdella vokaalimerkillä. **Fonologinen lyhyt/pitkä** -vastakohtaisuus on

Taulukko 2. Vertailussa käytetyistä julkaisuista laskettuja minimaalisesti erottuvia vokaalipareja. Etäisyydet laskettu formanttien keskiarvojen perusteella. Suuremman eron aiheuttanut formantti mainittu.

kieli	tekijät	pari	F	ero
ranska	Vihanta 1978	œ – ø	F1	0,9
suomi	Kukkonen 1990	u – o	F1	1,45
suomi	Kuronen 2000	a – ä	F2	1,03
suomi	livonen tässä julk.	e – ö	F2	1,39
saksa	livonen 1997	Y – oe	F2	1,47
viro Q1	Eek 2008	e – ö	F2	0,88
ruots.ruotsi	Kuronen 2000	l – Y	F2	1,3
englanti	Deterding 1997	hud – hard	F2	0,99
englanti	Deterding 1997	heed – hit	F1	1,06
englanti	Deterding 2006	but – bath	F2	0,45
englanti 20–25	Hawkins ym. 2005	hood – heard	F1	1,1
englanti 35–40	Hawkins ym. 2005	hud – hard	F2	0,68
englanti 35–40	Hawkins ym. 2005	heed – hid	F1	0,75
englanti 35–40	Hawkins ym. 2005	head – had	F2	1,5
englanti 50–55	Hawkins ym. 2005	hud – hard	F2	1,19
englanti 50–55	Hawkins ym. 2005	heed – hit	F1	1,26
englanti 50–55	Hawkins ym. 2005	hard – hod	F2	1,19
englanti yli 65	Hawkins ym. 2005	who'd – hood	F1	1,23
englanti yli 65	Hawkins ym. 2005	hud – hard	F2	0,86
keskiarvo				1,06
keskihajonta				0,31

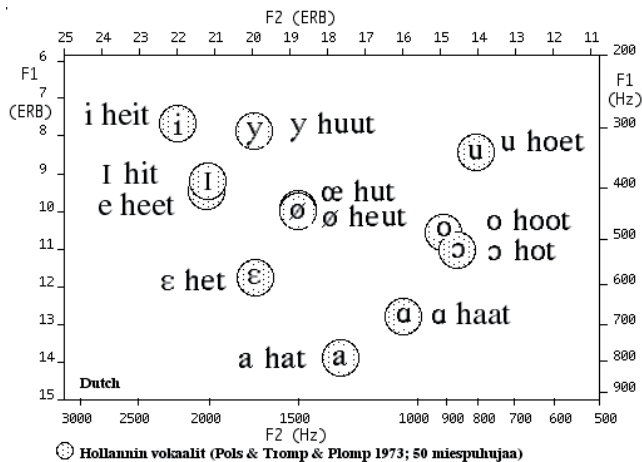
luokitteleva, ei jatkumollinen kuten mitattava kesto. Siirrettyinä ERB-asteikkoiseen karttaan myös Koopmans-van Beinumin (1980: 35) vokaalikeskiarvot limittyvät vastaavasti, mutta myös parissa /a/-/a/ ne leikkaavat toisiaan (vrt. kirjoitusasuja *beet-bid*, *baat-bad*, *beuk-buk*, *boot-bot*). Sitä vastoin pareissa /i/-/i/ ja /e/-/ε/ (kirjoituksessa *biet-bid*, *beet-bed*) ero toteutuu minimietäisyyttä suurempana. Hollannin 12 vokaalia ovat sikäli poikkeavia esimerkiksi suomeen verrattuna, että aikoinaan pitkät vokaalit /i: y: u:/ ovat muuttuneet lyhyiksi (Koopmans-van Beinum (1980: 7), mutta erottuvat silti vastaavista lyhyistä laadullisesti. Parissa /i/-/i/ (*heit-hit*) F1:n ero on 1,54 ERB.

VOKAALITYYPIN YKSITTÄISESIINTYMIEN HAJONTA

Yksinkertaiset vokaalit

Vokaalin keston lyheneminen ja painollisuuden väheneminen johtavat vokaalien keskittymiseen formanttikartassa eli selvyysasteen

heikkenemiseen, reduktioon. Jo suhteellisesti selvästi äännettyssä aineistossa laatu redusoi-tuu, kun verrataan pääpainollisen tavun ja sitä seuraavan painottoman tavun formanttiarvoja (Wiik 1965: 133; yleisistä selvyysasteen taustatekijöistä, ks. Koopmans-van Beinum, 1980; van Bergem, 1995; Deterding, 2006; spontaanin suomen osalta: Lennes, 2008). Painottomien vokaalien sijaintipisteet keskistyvät formanttikartalla. Kun koeaineistona ovat koesanat semanttisesti luonteissa lauseissa, tapahtuu myös hienoista redusoitumista (ranskan vokaalit: Vihanta, 1977; suomen vokaalit: Kuronen, 2000). Niissä konteksti edistää *top down* -ilmiön takia ymmärrettävyyden säilymistä. Akustiset kontrastit heikkenevät lisää vapaassa keskustelupuheessa. Kahden mies- ja kahden naispuhujan vapaasta keskustelusta leikatut painottomat vokaalit tunnistettiin hyvin heikosti keskustelun yleisymmärrettävyyden silti säilyessä (Koopmans-van Beinum, 1980). Puheen reduktioilmiöt ovat rinnastettavissa käsinkirjoitettuun tekstiin: ympäristöstään rajatut kirjaimet ovat usein vaikeasti tunnistettavissa.



Kuva 5. Hollannin 12 vokaalia Polsin, Trompin ja Plompin (1977) formanttiarvojen perusteella. Tutkimuksessa käytettiin *hVt*-sanoja. Ympyrän halkaisija 1 ERB.

Vieruskonsonanttien vaikutus vokaalin formantteihin on ilmeinen (Wiik 1965: 72, 76 – 77, 79; Kukkonen, 1990: 163; Deterding, 1997: 49; Vihanta, 1978). Tämä näkyi hyvin suomen osalta tutkimuksessa, jossa käytettiin painollisen vokaalin symmetristä konsonantitikehystä (/p_p/, /m_m/, /v_v/ jne.; Iivonen ja Laukkanen 1993).

Seuraavassa kohteena ovat pääpainollisten vokaalifoneemien realisaatioiden vaihtelu. Sana-aineisto valittiin Kotimaisten kielten tutkimuskeskukselta saadusta suomen kielen perussanojen aakkosellisesta listasta. Listasta poimittiin 10.000 sanaa. Suomen yleispuhekieltä edustava ammattimainen (mies)puhuj TP äänsi sanat äänieristetyssä studiossa ja ne digitointiin DAT-nauhoittimella. Koesanat esitettiin hänelle tietokoneen näyttöruudulta, joka sijaitsi äänitystudion ikkunan toisella puolella. Koehenkilö harjoitteli etukäteen ääntämään sanoja ilman luettelointonaatiota

ja siten ikään kuin vastauksina esimerkiksi kysymykseen ”Mikä sana tämä on?”. Selvyyden kannalta aineisto edustaa selvää mutta luontevaa ääntämistä. Tästä tietokannasta poimittiin $2 \times 8 \times 20 = 320$ suomen kaksitai poikkeuksellisesti kolmitavuista sanaa, joiden pääpainollisessa ensi tavussa esiintyy jokin suomen kahdeksasta lyhyestä ja kahdeksasta pitkästä vokaalista. Vokaalien kontekstit valittiin niin, että vokaalialkuisten sanojen ohella painollista vokaalia edelsi ja seurasi eri konsonantteja nasaaleja ja tremulantteja lukuun ottamatta.

Kaikki 320 vokaalia analysoitiin Praat 5.0 -puheen käsittelyohjelmalla, jolla kolme alinta formanttia mitattiin vokaalin tavoitevaihtetta parhaiten edustavasta aikapisteestä soveltaen Burg-menetelmää. 3000 Hz:n alueelta etsittiin kolmea formanttia. Mittaus hyväksyttiin, kun spektrogrammissa näkyvä formantti ja automaattisen ohjelman laskema formantti-

Taulukko 3. Suomen kielen lyhyiden ja pitkien vokaalien formanttien F1, F2 ja F3 keskiarvot hertseinä. Formanttien F1 ja F2 keskiarvot ja hajonnat myös ERB-arvoina. Jokainen keskiarvo edustaa mittauksia 20 eri sanasta. Miespuhuj TP. Vastaava formanttikartta kuvassa 6.

V	Formantit Hz			ERB		Hajonta	
	F1	F2	F3	F1	F2	F1	F2
i	321	1967	2380	8,1	21	0,3	0,3
e	443	1722	2265	10	20	0,4	0,4
ä	600	1535	2268	12	19	0,4	0,2
y	330	1601	2188	8,3	19,3	0,3	0,5
ö	436	1451	2188	10	18,5	0,4	0,2
a	609	1256	2188	12,1	17,4	0,4	0,6
o	433	958	2163	9,9	15,3	0,6	0,9
u	332	690	2207	8,3	12,9	0,5	0,7
i:	294	2039	2536	8	21,4	0,3	0,2
e:	408	1822	2270	9,5	20,4	0,3	0,4
ä:	577	1540	2260	11,7	19	0,4	0,3
y:	296	1609	2081	7,7	19,4	0,3	0,4
ö:	458	1455	2129	10,2	18,5	0,7	0,3
a:	590	1229	2198	11,8	17,2	0,5	0,5
o:	442	858	2169	10	14,5	0,4	0,7
u:	333	610	2325	8,3	12,1	0,5	0,4

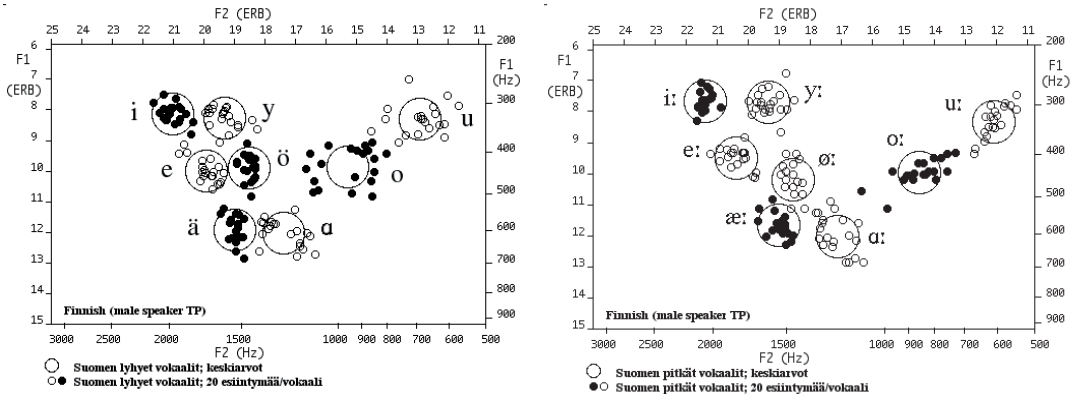
juova osoittivat yhtäpitävyyttä (kuva 1). Tarkistuksissa käytettiin hyväksi myös 33 millisekunnin aikaikkunalla tuotettuja spektrejä (kuva 2). Formanttien F1–F3 mittausarvot on esitetty Hz- ja formanttien F1 ja F2 osalta ERB-arvoina taulukossa 3. Mittausarvot sijoitettiin formanttikarttaan, jonka ohjelmoinnissa olen käyttänyt FutureBasic II -ohjelmointikieltä. Ohjelma lukee Praatin valikon vaihtoehdolla ”extract visible spectrogram” tuotetut *short text* -tiedostot ja valitsee jatkoon vain formanttien Hz-arvot. Nämä ohjelma muuttaa ERB-arvoiksi ja piirtää vokaalipisteet tai -ympyrät ERB-asteikkoiselle kartalle (kuvat 4–10). Foneettisten symbolien osalta Photoshop-kuvankäsittelyohjelma oli tarpeellinen viimeistelyvaiheessa.

Kuvassa 6 vertaillaan vokaalien kertaäänöksiä ja keskiarvoja F1/F2-kartassa. Kertaäänökset eli **yksittäisesiintymät** esitetään pisteinä ja **keskiarvot** ympyröinä. Kirjallisuudessa esiintyy tapaa ilmaista vain yksittäistapausten hajonnan suuruus keskiarvosta sekä F1:n että F2:n suhteen. Tämä tilastoiva tapa jättää käsittelemättä havaintoon liittyviä aspekteja. Kuvassa 6 ympyröiden halkaisija on 1,3 ERB edellä perustellun minimietäisyyksien mukaan likimääräinen alue, jonka sisäpuolelle sattuvia esiintymiä on vaikea erottaa toisistaan foneettisen erottelukyvyn avulla. Ympyrä ei siis kuvaa sitä, että saman foneemin toteutumien tulisi aina sijoittua sen sisään. Päinvastoin ilmeisesti aina vallitsee se tilanne, että osa varianteista sijoittuu sen ulkopuolelle. Tulos osoittaa, että sekä lyhyissä että pitkissä etuvokaaleissa puhuja on saanut vokaalien tavoitevaiheet ympyrän sisään tai sen lähelle, kun taas /o/:n ja /u/:n variantit hajaantuvat laajemmin. Etuvokaalien suurempi määrä ja auditiivinen lähekkäisyys vaativat siis suurempaa tarkkuutta, kun taas etuvokaalien ja takavokaalien väliin jäävä ”tyhjä” tila voi jäädä takavokaalien toteutumisalueeksi. Tämä johtune osin labiaalistumisen ja takistumi-

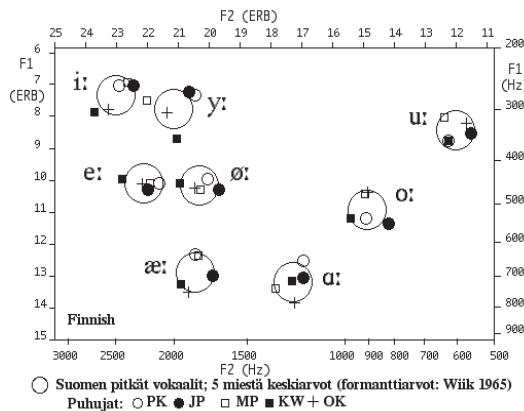
sen vähenemisestä ja joidenkin konsonantti-kontekstien aiheuttamasta koartikulaatiosta. Pitkien /o/:n ja /u/:n varianttien osuminen ”maaliin” on onnistunut paremmin kuin lyhyiden, ja [u:]-variantit ovat saavuttaneet suuremman pyöreäyden/takaisuuden. Oheinen taulukko näyttää vielä keskihajontojen keskiarvot etu- ja takavokaaleille, erikseen lyhyille ja pitkille. Siitä ilmenee, että molempien formanttien takavokaalien hajonnat ovat suuremmat sekä lyhyissä että pitkissä vokaaleissa.

	Hajontojen keskiarvot (ERB)	
	F1	F2
Lyhyet etuvokaalit	0,4	0,3
Lyhyet takavokaalit	0,5	0,7
Pitkät etuvokaalit	0,4	0,3
Pitkät takavokaalit	0,5	0,5

Yllättävästi Wiikin (1965: 57) tutkimuksen mukaiset viiden miespuhujan suomen kielen pitkien vokaalien yksilölliset keskiarvot sijoittuvat myös niin lähelle toisiaan, että nekin miltei mahtuvat 1,3 kokoisen ympyrän sisään (kuva 7). Näyttää vaikealta löytää yksilöllisten keskiarvopisteiden sijainnin vaihtelulle selvää systemaattista selitystä. Mitattujen esiintymien määrä vaihteli kuitenkin eri vokaaleissa ja eri puhujien keskiarvoissa suuresti. Puhujilla esiintymiä oli 2 – 17 vokaalia ja 128 – 201 yhteiskeskiarvoa kohti. Eniten mitattuja tapauksia oli puhujilla KW (166; musta neliö) ja OK (201; risti). Puhujat olivat kotoisin eri puolelta Suomea, mutta murrevaikutusta on kuvasta 7 vaikea osoittaa. Pikemminkin näyttää ilmeiseltä, että useiden puhujien yhteiskeskiarvo häivyttää yksilöllistä vaihtelua ja antaa kielen systeemisestä vokaalikonstel- laatiosta paremman käsityksen.



Kuva 6. Suomen lyhyiden ja pitkien vokaalien yksittäisesiintymien (n = 2 x 8 x 20 = 320) keskiarvot kuvattuina ympyröinä, joiden halkaisija on 1,3 ERB. Kunkin vokaalin 20 esiintymää pieninä mustina tai valkoisina ympyröinä. Vasemmalla lyhyet, oikealla pitkät vokaalit. Miespuhujaja TP.

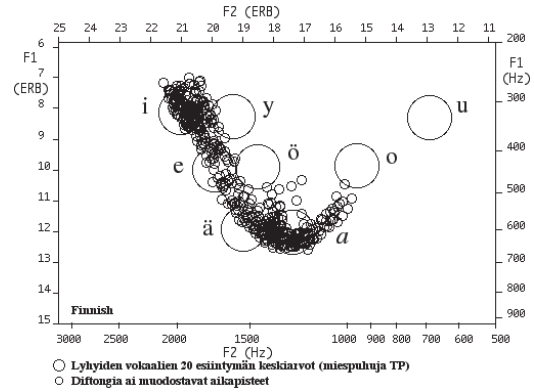
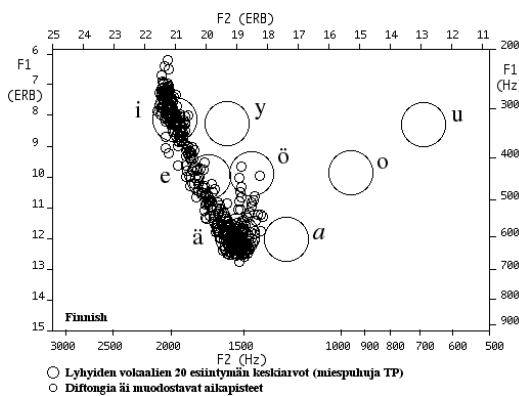


Kuva 7. Suomen pitkät vokaalit viiden miespuhujan PK, JP, MP, KW ja OK yksilöllisinä keskiarvoina (= pienet merkit) Kukin suuri 1,3 ERB:in kokoinen ympyrä piirretty viiden puhujan keskiarvon ympärille (formanttiarvot: Wiik 1965).

Diftongit

Diftongit antavat hyvän todisteen kuuloanalogisen formanttikartan toimivuudesta (vrt. Iivonen, 1997b). Analysoitu sana-aineisto valittiin edellä mainitusta miespuhujaja TP:n digitoidusta tietokannasta. Siitä poimittiin 2 x 20 sanaa, joissa on pääpainollinen *ai-* tai *äi-*diftongi. Formanttiliukumat mitattiin em.

Praat-ohjelmalla. Nämä asetettiin puhujan lyhyiden monoftongien keskiarvoympyröiden kanssa yhteiseen kuvaan (kuva 8). Yksittäisten varianttien muodostama liukuma-alue ei ylitä 1,3 ERB:in suuruista vaihtelua. Diftongissa *äi* (kuvassa vasemmalla) alkutavoite sattuu täysin samaan kohtaan lyhyen *ä*-monoftongin keskiarvon kanssa, johon muodostuu tumma tihentymä, diftongin alun tavoitevaihe. Sitä



Kuva 8. Miespuhuja TP:n diftongien *äi* (vasemmalla) ja *ai* 20 kertaäänöksen F1/F2-liu'ut sijoitettuina lyhyiden vokaalien keskiarvoja kuvaavien suurten ympyröiden karttaan, koko 1,3 ERB. Vokaalimerkit suomen ortografian mukaan.

edeltää koartikulaatiovaikutuksia alkusiirtymää suppeammalta alueelta eli F1:n hieman alhaisemmista arvoista. Lopputavoite sattuu suunnilleen *i*-vokaalin keskiarvoympyrän kohdalle. Alku- ja lopputavoitteen tihentymät osoittavat vokaalin laadun pysähtymistä pidemmäksi vakiovaiheeksi, mikä mahdollistaa vokaalilaadun selemmän tunnistamisen. Nopean siirtymävaiheen aikana ilmenee vaaleampi tihentymä.

Kuvassa oikealla *ai*-diftongin alkusiirtymäliuku alkaa *a*-monofoongin keskiarvoa suppeammasta asemasta, ja alkutavoite on [ɑ]:n keskiarvotaajuutta lievästi korkeampi. Alkusiirtymä on koartikulaatioiden aiheuttama. Lopputavoitteen tummin tihentymä muodostuu lyhyen /i/:n keskiarvoympyrän sisään mutta ei aivan keskelle. Tavoitevaiheiden välinen vaaleampi tihentymä osoittaa puolestaan nopeaa siirtymävaihetta, jonka aikana ei eroteta vokaalilaatuja (ei siis kuulla [aei]). Diftongien vähäinen ero näyttää olevan alkutavoitteen lisäksi siinä, että /ai/:n liuku pysyttelee kokonaisuudessaan keskemällä karttaa eli liuku on lievästi sentraalistunut.

MIESTEN JA NAISTEN ÄÄNTÖVÄYLÄN JA FORMANTTIEN ERO

Vaikuttavimmat erot vokaalien formanteissa ilmenevät lasten, naisten ja miesten aineistoissa (vrt. Peterson & Barney, 1952). Lyhyempi ääniraon etäisyys huulista eli lyhyempi ääntöväylä tuottaa korkeampia formanttiarvoja. Kirjallisuudessa on esitetty laskelmia miesten ja naisten formanttien prosentuaalisesta erosta pyrkien sen jälkeen normalisoimaan aineisto. Laskelmat on tehty hertsiarvoista. Fantin (1973: 29) mukaan naisten ero miehiin nähden olisi 20 %. Nierop, Pols ja Plomp (1973) sovelsivat logaritmistasta asteikkoa ja päätyivät tulokseen, että ero on 10 % ja että vokaaleissa /u y i/ eroa ei ole.

Normaalistamisen lähtökohdaksi voidaan ottaa alkuperäisten Hz-arvojen perusteella lasketut ERB-arvot. Kuvassa 9 vasemmalla näkyvät englannin monofoongien keskiarvot ympyröinä. Aineistona on käytetty Deterdingin (1997) väitöskirjaansa jo 1990 varten mitaamia Cambridgen alueella asuvien miesten ja naisten formanttiarvoja. Tekijä käytti lähes samoja sanarakenteita kuin Peterson ja Bar-

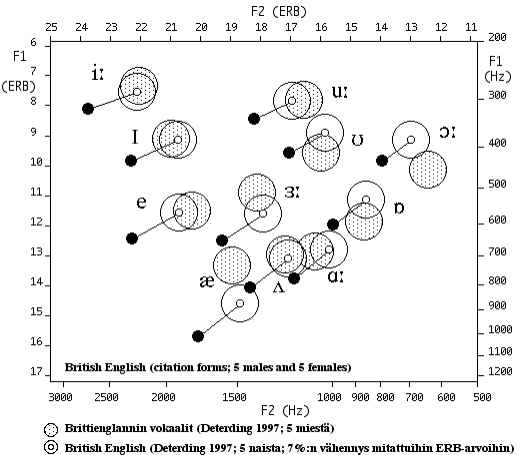
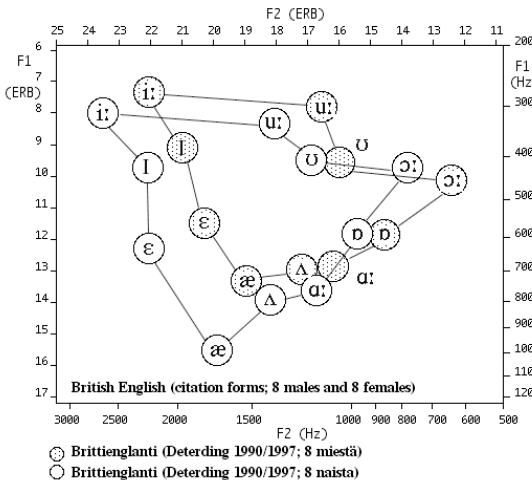
ney (1952): *heed* [i:], *hid* [ɪ], *head* [ɛ], *had* [æ], *hard* [ɑ:], *hud* [ʌ], *hod* [ɒ], *hoard* [ɔ:], *hood* [u], *whod* [u:], *herd* [ɜ:]. Aineisto edustaa tekijän mukaan eteläisen brittienglannin sanojen ns. sitaattimuotoista, *Received Pronunciation (RP)* -normin mukaista standardiääntämistä.

Kuvasta 9 vasemmalla voi nähdä, että erot noudattavat kaavaa, jonka mukaan naisten ja miesten vokaalikonstellaatio on samankaltainen, mutta naisten vokaalien formanttiarvot ovat korkeampia. Vokaalikonstellaatiot eivät kuitenkaan noudata täysin samaa kaavaa. Vokaalin [ɔ:] F1 on naisilla suhteellisesti matalampi kuin miehillä. Vokaali [æ] osoittaa naispuhujilla poikkeuksellisen korkeaa F1:n arvoa. Kun [ɔ:] ja [æ] jätetään ulkopuolelle, ja lasketaan muiden vokaalien formanttien F1 ja F2 keskiarvoiset ERB-erotukset sekä erotusten keskiarvo, saadaan keskimääräiseksi eroksi 7 %. Kuvassa 9 oikeanpuoleinen kart-

ta esittää naispuhujien alkuperäiset vokaalit mustina pisteinä ja muut vokaalit suurina (halkaisija 1 ERB) ympyröinä 7 %:n muunnoksen jälkeen. Miespuhujien vokaaliympyrät ovat tummennettuja.

Muunnoksen jälkeen naisten muut vokaalit kuin [ɔ: æ] osuvat yksiin osin tai kokonaan vastaavien miesten vokaalien kanssa. Mainitut kaksi vokaalia näyttävät vaihtelevan yleensäkin laadultaan brittienglannin aineistoissa, mikä ei siten liity ääntöväylän eroihin (vrt. kuva 10: Deterding, 1997; Hawkins & Midgley, 2005). Naisten ja miesten mahdollisiin foneettisiin eroihin viittaa myös Koopmansvan Beinum (1980:71).

Kaksi muuta tutkimusta vahvistavat muunnoksen onnistuvan 7 %:n siirrolla ERB-asteikossa. Petersonin ja Barneyen (1952) amerikanenglantia koskevassa tutkimuksessa naispuhujia oli 28 ja miespuhujia 33. Useimmat naisista edustivat keskiatlantista



Kuva 9. Ääntöväylän pituudesta aiheutuvan eron normalistaminen. Vasemmalla: Brittienglannin viiden miespuhujan (= tummennetut ympyrät) ja viiden naispuhujan (valkoiset ympyrät) keskiarvoiset vokaalisijainnit. Alkuperäiset Hz-arvot Deterding (1997). Keskivokaali [ɜ:] jätetty selvytyden vuoksi pois. Oikealla: Samat miesten vokaalit (= tummennetut ympyrät). Naisten alkuperäiset keskiarvopisteet (= mustat pisteet) ja lisäksi 7 %:lla normalisoituna (= suuret valkoiset ympyrät).

puhealuetta, enemmistö miehistä puolestaan paljon laajempaa regionaalista otosta (*General American*). Sovellus käyttäen samaa 7 %:n muunnosta Petersonin ja Barney'n mittamiin formanttiarvoihin tuottaa samantyyppisen tuloksen ERB-asteikolla kuin kuvassa 9. Myös Hillenbrandin, Gettyn, Clarkin ja Wheelerin (1995) formanttiaineisto tuottaa saman tuloksen. Siinä 45 miestä ja 48 naista valittiin huolellisesti varmistaen, että he edustivat *General American English* -puhetyyppiä. He olivat kotoisin Michiganin eteläosista ja rajaosavaltioista. Kultakin puhujalta pyrittiin saamaan keskiarvolaskelmia varten vaihtelevassa järjestyksessä irrallaan äännettyjen /hVd/-sanojen listasta kolme eri versiota. Kuuntelukokeiden perusteella todettiin foneemisia virhetunnistuksia, ja niitä aiheuttaneita sanoja poistettiin 11,5 % lopullisesta mittausaineistosta. Koeaineistossa ja -järjestyksessä on pieniä eroja Petersonin ja Barney'n tutkimukseen verrattuna.

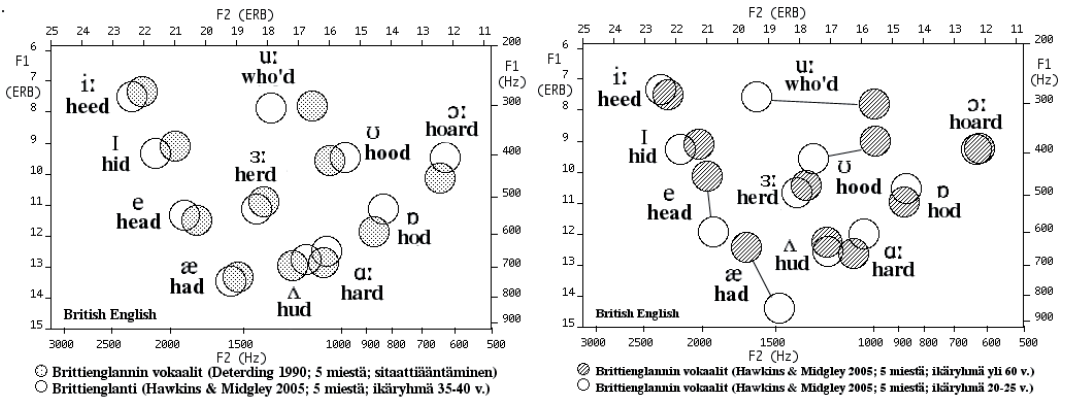
Erona Petersonin ja Barney'n (1952) tutkimukseen Hillenbrandilla ym. ovat mukana vokaalit /e/ ja /o/. Etuvokaalien sarja on / i: i e ε æ / (myös Lintusen taulukossa (2005: 165). Siten sanat *hayed* ja *head* sisältävät eri vokaalin: / e ε /. Parin *hid - hayed* / i e / ero ylittää minimieron sekä F1:ssä (1,29) että F2:ssa (1,11). Samaa suuruusluokkaa ovat ERB-erot parissa *hayed-head*. Parissa *head-had* /ε/-/æ/ erot ovat kuitenkin minimiero pienempiä: 0,09 ja 0,68 ERB. Miesten F1:n keskiarvo on /ε /:llä 580 ja /æ/:llä 588 Hz. Siten /æ/ ei olekaan selvä väljä vokaali kuten brittienglannissa. Voisiko kyseessä olla edellä mainittu laadullinen sulautuminen? Taulukko (mts. 3108) osoittaa kuitenkin, että vokaalien / i: i e ε æ / tunnistus onnistui 94–98,8 %:sesti. Mukana olivat miesten, naisten ja lasten vokaalit. Tunnistusvirheistä oli valtaosa parin /ε/-/æ/ keskinäisiä. Selityksenä tunnistusprosenttien suuruuteen tekijät toteavat, että kuullut ärsykkeet sisälsivät myös vokaalin siirtymävaiheet.

Vertaillen Petersonin ja Barney'n tuloksia omiinsa he viittaavat erojen tärkeimpänä ja potentiaalisena selityksenä noin 40 vuoden aikana tapahtuneeseen puheen muutokseen. He viittaavat myös mahdolliseen suurempaan diftongoitumiseen omassa aineistossaan (mts. 3107-8).

Huolimatta naisten ja miesten vokaalien akustisista erilaisuuksista kuulijat tunnistavat vastaavat vokaalit samoiksi. Tunnistuksen selittänee se, että kuulija havaitsee vokaalilaadut suhteessa ääntöväylän pituuteen (vrt. Fant 1973: 84; 2004: 89, 158). Ilmiötä on selitetty puheen havaitsemisen teoriolla sekä viittaamalla puhujan perussävelen korkeuteen suhteessa ääntöväylän pituuteen. Irino ja Patterson (2002) selittävät, että ääntöväylän koon ja muodon aiheuttaman vaihtelun normalisointi tapahtuu jo sisäkorvassaan. Mellin-muutoksen avulla.

BRITTIENGLANNIN VOKAALISYMBOLIT JA FONEETTINEN TODELLISUUS

Tuottavatko kaksi eri tutkimusta samasta kielestä saman vokaalikonstellation formanttikartalla? Vaihteluilmiöiden vuoksi saman kielen tutkimuksia voidaan vertailla vain, jos aineiston taustatekijät ovat samat. Tämä ehto toteutuu suureksi osaksi kahdessa brittienglannin vokaalitutkimuksessa: Deterding (1997: 52) ja Hawkins & Midgley (2005: 195). Tarkat äänitystiedot puuttuvat Deterdingin osalta jo 1990 analysoidusta aineistosta. Se käsitti kahdeksan miespuhujaa, Hawkinsilla ja Midgleyllä neljässä ikäryhmässä yhteensä 4 x 5 = 20. Koesanoissa oli sama konsonanttikonteksti /hVd/: *heed* [i:], *bid* [ɪ], *head* [ε] jne. Kummassakin sovellettiin ns. sitaatti- eli selvää ääntämystä. Hawkinsin ja Midgley'n tutkimuksessa sanat lausuttiin neljänä sanalistana järjestystä vaihdellen. Neljä puhujaryhmää edustivat ikäjakautia 20–25,



Kuva 10. Vasemmalla: Brittienglannin vokaalit kahden eri tutkimuksen mukaan: Deterding 1990/1997 (tummennettu) ja Hawkins ja Midgley 2005 (ikäryhmä 35 – 40 v.). Kumpaakin edustaa viisi miespuhujaa ja sanalistat *heed*, *hid*, *head* jne. Oikealla: Brittienglannin vokaalien vertailu kahdessa ikäryhmässä: yli 65-vuotiaat (tummennettu) ja 20 – 25 -vuotiaat.

35–40, 50–55 ja yli 65-vuotiaat. Kaikki olivat Cambridgen yliopiston opiskelijoita, opettajia tai muita työntekijöitä. Kuvassa 10 vertaillaan vasemmalla Deterdingin tuloksia Hawkinsin ja Midgley'n 35–40 -vuotiaiden keskiarvoihin. Koska ilmeni, että jälkimmäisessä aineistossa formantti F1 on systemaattisesti alhaisempi tuntemattoman tekijän aiheuttamana, kaikkien vokaalien F1:n ERB-arvoja korotettiin käsillä olevaa julkaisua varten 4 %. Muunnoksen jälkeen kaikki vokaaliympyrät leikkaavat toisiaan paitsi vokaalissa /u:/ (kuva 10, vasemmalla; *who'd*). Kun vertailu tehdään Deterdingin ja Hawkinsin & Midgley'n 50–55 -vuotiaiden kesken saadaan 5 %:n muutoksella sama tulos, mutta vokaalissa /ɔ:/ on nyt selvä ero. Deterdingin koehenkilöt olivat 15 vuotta vanhempia vuonna 2005 ja mahdollisesti siten paremmin verrattavissa Hawkinsin & Midgley'n ikäryhmään 35–40 -vuotiaat. On aiheellista pitää vertailtujen tutkimusten tulosta likiarvoisesti samanlaisena.

Foneettisten symbolien osalta on erikoista, että englannin suppea pyöreä keskivokaali on symboloitu mainituissa julkaisuissa ja muu-

allakin [u:]lla, joka on varsinaisesti suppean pyöreän takavokaalin merkki. Käytännön on aloittanut Daniel Jones, joka röntgenkuvien perusteella asetti mahdollisesti englannin ko. vokaalin mukaisesti kardinaalivokaalien äärimmäiseksi suppeiden takavokaalien merkiksi. Ero Jonesin kardinaalivokaali [u:]n ja englannin [u:]n välillä on Jonesin oppikirjassaan (1960: 52, 64) hyvin pieni. Oliko englannin ”[u:]” Jonesin aikaan selvä takavokaali? Se on kuitenkin muuttanut paikkaansa edelleen kohti etuvokaaleja nuoremmassa sukupolvissa (ks. kuvaa 10, oikealla 20–25 -vuotiaat) ja lähenee samaa asemaa kuin suomenruotsin [ʉ:] (Kuronen, 2000: 141).

Oikealla vertaillaan yli 65- ja 20–25 -vuotiaiden vokaaleja. Erojen tulkinnessa tulee merkittävään asemaan kielellinen muuttuminen nuoremmassa ikäpolvissa ja etenkin ryhmässä 20–25 v. Näyttää ilmeiseltä, että ”[u:]”n siirtyä kohti etuvokaaleja on yhteensä noin 4 ERBin verran tasaisin välein yli 65 vuotiaista 20–25 vuotiaisiin. Vokaali /ʊ/ on siirtynyt kohti keskivokaalia /ɜ:/ (kuten sanassa *heard*). Merkittävältä näyttää myös

/æ/:n väljeneminen väljemmäksi kuin /ɑ:/. Tulkintaa hankaloittaa se, että kirjoittajat tunnustavat mahdollisuuden, että valittujen RP-puhujien taustalla saattaa vaikuttaa regionaalisia tekijöitä. Muut erot tekijät tulkitsevat huolellisemman ääntämisen tulokseksi, mikä formanttikartassa tarkoittaa keskiarvojen selvempää perifeerisyyssastetta.

Hawkins ja Midgley (mts. 194) kuvaavat ikämuutoksen vaikutusta vokaalien siirtymisenä kartassa vastapäivään, mutta rajoittavat sen koskemaan vain etisten ja suppeiden vokaalien laatujen muuttumista lukuun ottamatta vokaaleja /i: ɪ/. Siihen näyttäisivät kuitenkin sopivan myös /ɛ ʊ ɔ:/. Vokaali /ɔ:/ on laadultaan melko suppea sekä vanhimmilla että nuorimmilla puhujilla. Se kuuluu kartan mukaan pikemminkin puolisupeiden kuin puoliväljien vokaalien sarjaan, koska F1:n keskiarvo on alle 400 Hz (vrt. kuvaan 4). Sen merkiksi näyttäisi sopivan [o:].

Vuosisatoja sitten englannin foneemi /u/ alkoi tietyissä sanoissa illabialistua ja väljentyä [ʌ]:ksi, mikä ilmenee nykysanojen *butcher* ja *butter* ääntämisen erona. Formanttikartta osoittaa tosin, että nykyinen foneemi /ʌ/ on laadullisesti lähellä /ɑ:/:ta (kuva 9, oikealla; kuva 10). Foneettisen [ʌ]:n paikka on väljässä sarjassa etisen ja takaisen variantin välissä. Sille ei kuitenkaan ole siis IPA:ssa omaa merkkiä ellei [ɑ] olisi osuva (vrt. Lintunen, 2004: 70). Muutoinkin englannin vokaalisymbolit ovat kylläkin foneettisia mutta pikemminkin foneemisia (vrt. mts. 25, 69). Johtopäätös on, että olisi siis perusteita tarkistaa brittienglannin joidenkin vokaalien traditionaalista symbolointia.

TULOKSET JA PÄÄTELMÄT

Tavoitteena oli yhdistää eri näkökohtia kokonaiskuvaksi kielten vokaalilaatujen osoittamiseksi. Fonologia tyytyy abstraktioihin: osoittamaan vokaalien kielikohtaisen foneemi-inventaarin ja (ensisijaisesti artikulatori-

set) erottavat piirteet, toisinaan myös allofonit. Jatkumollisuuden huomiotta jättäminen estää vokaalien laadun ja yksittäisrealisaatioiden vaihtelun esittämisen vokaaliavaruudessa. Tästä seuraa, ettei kielten vokaalien todellinen laadullinen vertailu ole mahdollista ja kielen järjestelmän vaihtelu puheessa ja vähittäinen muutos jäävät näkymättä. Kielen foneemianalyysi on kuitenkin välttämätön.

Foneettisen kirjoituksen (esim. IPA:n) symbolit ovat tarkempia ja välttämättömiä kielten vertailulle. Symbolien käytössä ilmenee kuitenkin puutteita ja ristiriitoja symbolin ja todellisen ääntämisen kesken. Esitetyt esimerkit kansainvälisen foneettisen aakkoston (IPA:n) merkkien käytöstä osoittavat, että merkkien puutteiden lisäksi niitä sovelletaan usein pikemminkin foneemien symbolointiin kuin foneettisten laatujen ilmaisemiseen. Kielen muuttuessa entinen vokaalisymboli jää edelleen muuttuneenkin äänteen merkiksi. Artikkelissa käytetty ERB-asteikkoinen kuuloanalogninen formanttikartta esittää vokaalit formanttien F1 ja F2 auditiivisessa graafisessa kartassa, jolloin niiden todellinen laatu paljastuu. Kartan laatimisella on kuitenkin myös ehtoja ja koejärjestelyn edellytyksiä. Artikkelin peruskohteita ovat vokaalifoneemin yksittäisesiintymien suhde niiden keskiarvoon ja hajontaan F1/F2-kartalla, keskiarvojen etäisyydet muiden vokaalien keskiarvoihin ja keskiarvojen minimaaliset etäisyydet. Myös vokaalien universaaliin tasavälisestä jakaumatheoriaa (*dispersion theory*) käsitellään.

Julkaistujen tutkimusten formanttitietojen perusteella nyt lasketut erot ovat seuraavat. Kun sovellettiin kolmen julkaisun tulosten mukaisia erottelun ääriarvoja 1,9 – 5 % yksinkertaistettujen synteettisten vokaalien erottelutarkkuuden vaihteluväli oli 0,13 – 0,31 (F1) ja 0,16 – 0,41 (F2) ERB. Enemmän luonnollisia vokaaleja muistuttavat vokaalit tuottivat 1,3 ERB:in (= 1 Barkin) tarkkuuden (laskettu Kuhlin, 1991 ja Aaltosen ym.,

1997 julkaisutietojen perusteella). Teoreettisten formanttiarvojen (Lindblom, 1986) perusteella laskettuna pienin kahden kardinaalivokaalin etäisyys on 1,37 ERB.

Minimaalista etäisyyttä tutkittiin kunkin luonnollisen kielen aineistosta ”minimiparin” avulla, mikä viittaa kahden toisiaan lähimmän vokaalin etäisyyteen. Suomalaisella miespuhujalla TP:llä minimaalinen ero oli 1,39 ERB. Se ilmaisee TP:n lyhyiden vokaalien /e/ ja /ö/ realisaatioiden keskiarvon etäisyyttä. Kuutta kieltä edustavista julkaisuista formanttiedoista tätä julkaisua varten lasketut kahden toisiaan lähimpänä sijaitsevan vokaalikeskiarvon etäisyydet vaihtelivat välillä 0,45 – 1,47 (keskiarvo 1,06 ERB). Suurimpia lähietäisyyksiä kuvaavat ERB-arvot 2,19 (puhujan TP pari /a-o/) ja 2,65 (TP:n pari /a:-o:/). Käytettyjen julkaisuiden koejärjestelyt ovat vaihtelevia. Osassa tutkimuksista käytettiin vain yhtä sanarakennetta ilman vaihtelevaa konsonanttikontekstia tai kehyslauseetta. Toisissa aineistona olivat kontekstittomat sanat ja vaihteleva konsonanttikonteksti tai luonnollisissa lauseissa äännetty vaihtelevarakenteiset sanat. Lisäksi esiintyi luetusta tekstistä poimittuja selvästi äännettyjä sanoja. Puhujien ja sanojen määrät vaihtelivat.

Luonnollisten kielten vokaalietäisyyksiin vaikuttaa selvässä puheessa ensinnäkin paradigman koko: suuri vokaalifoneemien määrä edellyttää suurempaa ääntämistarkkuutta jolloin seurauksena on pienempiä keskiarvoetäisyyksiä formanttikartalla. Laadun lisäksi vaikuttaa vokaalien fonologinen pituus ja tuotettujen vokaalien kesto. Aineistona käytetyt kielet jakaantuvat karkeasti kolmeen ryhmään: (1) kielessä on selvä kvantiteettiero lyhyt-pitkä -vastakohtaisuus (suomi, ruotsi, saksa, hollanti; virossa kolme astetta Q1, Q2, Q3), (2) kielessä on jonkinasteinen keston vaikuttava vastakohtaisuus (englanti: tense/lax-vastakohtaisuus) tai (3) kielessä on sanarakenteellinen keston vaikuttava ominaisuus

(Vihanta, 1978: fonotaktinen pituus). Pitkien ja lyhyiden vokaalien laatueroit ilmenevät selvästi saksan formanttikartassa (Iivonen, 1997a; Iivonen & Tella, 2009: 278). Pitkät vokaalit perifeeristyvät kartalla, lyhyet sentraalistuvat. Pääpainolliset vokaalit ovat pidempiä kuin pääpainottomat.

Etäisyyteen vaikuttaa vokaalien selvyysasteen väheneminen eli reduktio, kun koearneisto käsittää kokonaisia lauseita tai lukutekstii. Kun reduktioherkkä aineisto ja vertailuparit, joissa vokaalin pitkällä kestolla voisi olla vaikutusta, poistettiin, tuli kontekstittomien sanojen minimaalisten vokaalietäisyyksien keskiarvoksi 1,44 ja keskihajonnaksi 0,24 ERB. Koko reduktioherkässä aineistossa minimikeskiarvo oli 1,05 ja hajonta 0,22 ERB. Kun aineistona on lukupuhuntaa ja spontaania puhetta, vokaalien laadullinen vaihtelu kasvaa lisää. On ymmärrettävää, että kahden foneemin foneettisten varianttien F1/F2-keskiarvojen auditiivinen minimietäisyys voidaan ilmaista siis vain likiarvoisesti. Se on noin välillä 1 – 1,4 ERB. Vain selvästi äännetty aineisto voi osoittaa kielen standardi- tai tyyppilaadut. Tilastollisen luotettavuuden lisäämiseksi ja suuremman kielten välisen vertailun parantamiseksi nykyisiä koejärjestelyjä pitäisi tarkistaa.

F1/F2-keskiarvojen absoluuttiseen sijoittumiseen vaikuttava puhujan anatomisen ääntöväylän suurempi pituus laskee formanttiarvoja, mikä ilmenee mm. naisten ja miesten formanttiarvoissa. Vokaalien karttasijoitusten muodostamat kokonaiskonstellaatiot säilyvät kuitenkin suhteissaan samankaltaisina ja kuulija ignoroii akustiset erot. Ääntöväyläerojen vaikutus voidaan poistaa normalisaatiolla (kuva 9). Naisten ERB-formanttiarvojen vähentäminen 7 %:lla mahdollistaa vertailun miesten F1/F2-konstellaation kanssa kartalla. Kokeilussa ilmeni, että pelkkä ääntöväyläero ei selitä jäljelle jääviä poikkeamia sukupuolten kesken. Mahdollisesti myös joissakin vokaaleissa on todellisia ääntämiseroja.

Käsiteltyjen luonnollisten kielten vokaalikeskiarvot eivät asetu tasavälisesti F1/F2-formanttikartalle. Esimerkiksi laveiden ja pyöreiden etuvokaalien ero on pieni. Tasavälisen jakauman teoria ei siten saa tukea käsitellyistä kielistä. Karkeamman arvion mukaan kielten vokaalijärjestelmien tietyissä osissa on tasavälisyyteen pyrkimistä (kuten brittienglannin etuvokaalit nuorimmassa ikäryhmässä; kuva 10).

VIITTEET

- Aaltonen, O. (1997). *Vowel Perception: Behavioral and Psychophysiological Experiments*. Turun yliopiston julkaisuja. Sarja D 261. Turku: Turun yliopisto.
- Aaltonen, O., Eerola, O., Hellström, Å., Uusipaikka, E. & Lang, H. (1997). Perceptual magnet effect in the light of behavioral and psychophysiological data. *Journal of the Acoustical Society of America*, 101(2), 1–16.
- Barry, W. J. & Trouvain, J. (2009). Do we need a symbol for a central open vowel? The discussion so far and a reply to Daniel Recasens and Martin Ball. *Journal of the International Phonetic Association*, 39, 365–366.
- Bergem van, D. (1995). *Acoustic and Lexical Vowel Reduction*. Amsterdam: IFOTT.
- Boersma & Weenink < <http://en.wikipedia.org/wiki/Praat> > (Luettu 26.10.2010).
- Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association*, 27, 47–55.
- Deterding, D. (2006). The Northwind versus Wolf: short texts for the description and measurement of English pronunciation. *Journal of the International Phonetic Association*, 36 (2), 187–196.
- Disner, S.F. (1984). Insights on vowel spacing. Luku 9 teoksessa I. Maddieson, *Patterns of Sounds*. Cambridge/London/jne.: Cambridge University Press.
- Duckworth, M., McDougall, K., de Jong, G. & Shockey, L. (2011). Improving the consistency of formant measurement. *The International Journal of Speech, Language and the Law*, 18.1, 35–51.
- EEK, A. (2008). *Eesti keele foneetika I*. (toim. Pire Teras). Tartu: Tartu Ylikooli Kirjastus.
- FANT, G. (1973). *Speech Sounds and Features*. Cambridge (Mass.)/London: The MIT Press.
- FANT, G. (1978). Vowel perception and specification. *Rivista Italiana di Acustica II*, 69–87. Julkaistu uudelleen: Fant, 2004, 201–215.
- FANT, G. (2004). *Speech Acoustics and Phonetics. Selected Writings*. Dordrecht/Boston/ London: Kluwer Academic Publishers.
- FLANAGAN, J.L. (1955). A difference limen for vowel formant frequency. *Journal of the Acoustical Society of America*, 27 (3), 613–617.
- HAWKS, J.W. (1994). Difference limens for formant patterns of vowel sounds. *Journal of the Acoustical Society of America*, 95 (2), 1074–1084.
- HAWKINS, S. & MIDGLEY, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association*, 35, 183–199.
- HILLENBRAND, J., GETTY, L., CLARK, M.J. & WHEELER, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97 (5), 3099–3011.
- HILLENBRAND, J. & CLARK, M.J. (2001). Effects of consonant environment on vowel formant patterns. *Journal of the Acoustical Society of America*, 109 (3), 748–763.
- HOLBROOK, A. & FAIRBANKS, G. (1962). Diphthong formants and their movements. *Journal of Speech and Hearing Research*, 5, 38–58.
- HÄRMÄ, A., KARJALAINEN, M., SAVIOJA, L., VÄLIMÄKI, V., LAINE, U. & HUOPANIEMI, J. (2000). Frequency-warped signal processing for audio applications. *Journal of Audio Engineering Society* 48 (11), 1011–1029.
- IIVONEN, A. (1994). A psychoacoustical explanation for the number of the major IPA vowels. *Journal of the International Phonetic Association* 24:2, 72–90.
- IIVONEN, A. (1997a). Die Monophthonge im Deutschen und im Finnischen. Paradigmatische Ähnlichkeit, syntagmatische und phonetische Differenz. J. Korhonen & G. Gimpl (toim.), *Kontrastiv. Der Ginkgo-Baum* 15, 66–83. Helsinki: Finn Lectura.
- IIVONEN, A. (1997b). Zum Begriff des Diphthongs und zur Qualität der Diphthonge des Deutschen. Teoksessa M. Haase & D. Meyer (toim.), *Von Sprechkunst & Normphonetik*, 81–91. Hahnau und Halle (Saale): Verlag Werner Dausien.

- Iivonen, A. (1997c). Number of possible basic vowel qualities and their psychoacoustical distance measure. K. Elenius & Branderud, P. (toim.), *Proceedings of the XIII International Congress of Phonetic Sciences, Stockholm, 13–19 Aug. 1995, Vol 1*, 404–407.
- Iivonen, A. & Laukkanen, A-M. (1993). Explanations for the qualitative variation of Finnish vowels. *Studies in Logopedics and Phonetics 4. Publications of the Department of Phonetics, University of Helsinki, Series B, 5*, 29–54.
- Iivonen, A. & Huhe, H. (2005). Acoustical comparison of the monophthong systems in Finnish, Mongolian and Udmurt. *Journal of the International Phonetic Association*, 35, 59–71.
- Iivonen, A. & Tella, S. (2009). Vieraan kielen ääntämisen ja kuulemisen opetus ja harjoittelu. O. Aaltonen, R. Aulanko, A. Iivonen, A. Klippi & M. Vainio (toim.) *Puhuva ihminen*, 269–281. Helsinki: Otava.
- IPA <<http://www.langsci.ucl.ac.uk/ipa/ipachart.html>> <<http://www.finnlectura.fi/verkko/esitely/sivu110.htm>> (21.10.2010).
- Irino, T. & Patterson, R.D. (2002). Segregating information about the size and shape of the vocal tract using a time-domain auditory model: the stabilised wavelet-Mellin transform. *Speech Communication*, 36 (3), 181–203.
- Jones, D. (1960). *An Outline of English Phonetics*. 9. laitos. Cambridge: Heffer.
- Koopmans-van Beinum, F.J. (1980). *Vowel Contrast Reduction. An Acoustic and Perceptual Study of Dutch Vowels in Various Speech Conditions*. Amsterdam: Academische Pers.
- Kuhl, P. (1991). Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perceptual Psychophysics*, 50, 93–107.
- Kukkonen, P. (1990). *Patterns of Phonological Disturbances in Adult Aphasia*. Helsinki: Suomalaisen kirjallisuuden seura.
- Kuronen, M. (2000). *Vokaluttalets akustik i sverigesvenska, finlandsvenska och finska*. *Studia philologica Jyväskyläensia* 49. Jyväskylä: University of Jyväskylä.
- Ladefoged, P. & Maddieson, I. (1990). Vowels of the world's languages. *Journal of Phonetics* 18, 93–122.
- Ladefoged, P. & Maddieson, I. (1996). *The Sounds of the World's Languages*. 2. painos 1997. Oxford/Malden: Blackwell Publishers.
- Lenne, M. (2008). Segmental features in spontaneous and read-aloud Finnish. Teoksessa V. Silva & R. Ullakonoja (toim.), *Phonetics of Russian and Finnish*, 145–166. Frankfurt am Main/Berlin/jne.: Peter Lang.
- Liljencrants, J. & Lindblom, B. (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language*, 48, 839–862.
- Lindblom, B. (1986). Phonetic universals in vowel systems. Teoksessa J. J. Ohala & J. J. Jaeger (toim.), *Experimental Phonology*, 13–44. Orlando: Academic Press.
- Lintunen, P. (2004). *Pronunciation and Phonetic Transcription: A Study of Advanced Finnish Learners of English*. Anglica Turkuensia 24.
- Lintunen, P. (2005). Do you speak American? Amerikan englannin fonetiikkaa suomalaisille. A. Iivonen (toim.) *Puheen salaisuudet*, 151–170. Helsinki: Gaudeamus.
- Moore, B. C. J. (1997). Aspects of auditory processing related to speech perception. Teoksessa W. Hardcastle & J. Laver (toim.), *The Handbook of Phonetic Sciences*, 539–565. Oxford: Blackwell.
- Nierop van, D., Pols, L.C.W. & Plomp, R. (1973). Frequency analysis of Dutch vowels from 25 female speakers. *Acustica*, 29 (2), 110–118.
- Nord, L. & Sventelius, E. (1979). Analysis and perception of difference limen data for formant frequencies. Experiments in Speech Perception, Phonetic Research Seminar 1978–1979, *PERILUS*, Report I, 24–37.
- Peltola, M.S. (2003). *The Attentive and Preattentive Perception of Native and Non-native Vowels: The Effect of Second Language Learning*. Turku: University of Turku, Department of Phonetics.
- Peterson, G.E. & Barney, H.L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 24, 175–184.
- Pols, L.C.W., Tromp, H.R.C. & Plomp, R. (1973). Frequency analysis of Dutch vowels from 50 male speakers. *Journal of the Acoustical Society of America*, 53 (4), 1093–1101.
- Raimo, I., Savela, J. & Aaltonen, O. (2005). Vokaalijärjestelmät testissä. Teoksessa Iivonen, A. (toim.), *Puheen salaisuudet. Fonetiikan uusia suuntia*, 171–181. Helsinki: Gaudeamus.
- Suomi, K. (1984). Vokaalien psykoakustisen laadun määrittämisestä: algoritmisen menetelmän kuvaus ja tuloksia suomen monofongeista. *Virtittäjä* 88, 308–328.

- Suomi, K., Toivanen, J. & Ylitalo, R. (2006). *Fonetikan ja suomen äänneopin perusteet*. Helsinki: Gaudeamus.
- SUT <<http://www.kotus.fi/index.phtml?s=2436>>; <<http://www.finnlectura.fi/verkko/esittely/sivu110.htm>> (Katsottu 27.9.2011)
- Vainio, M., Palo, P., Aalto, D. & Laine, U. (2009). Lähde ja suodin – puheentuoton akustiikasta ja mallintamisesta. Teoksessa O. Aaltonen, R. Aulanko, A. Iivonen, A. Klippi. & M. Vainio (toim.), *Puhuva ihminen*, 161–173. Helsinki: Otava.
- Wieringen van, A., Boersma, P. & Pols, L. C. W. (1993). Perceptual asymmetry between initial and final glides: Psychoacoustics and cochlear encoding. *Proceedings of the Institute of Phonetic Sciences, University of Amsterdam* 17, 79–96.
- Vihanta, V. (1978). *Les voyelles toniques du français et leur réalisation et perception par les étudiants finnophones*. *Studia philologica Jyväskyläensia* 12. Jyväskylä: L'Université de Jyväskylä.
- Wiik, K. (1965). *Finnish and English Vowels*. *Annales Universitatis Turkuensis. Series B*, Tom. 94. Turku: Turun yliopisto.
- Zwicker, E. (1961). Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *Journal of the Acoustical Society of America*, 33, 248.

VOWEL QUALITIES PRESENTED ON A PSYCHO-ACOUSTICALLY SCALED F1/F2 FORMANT CHART

Antti Iivonen, University of Helsinki, Speech Sciences

Letters from Latin alphabet, phoneme symbols and basic phonetic symbols of the International Phonetic Alphabet (IPA) all lack more or less some degree of accuracy. An auditive formant chart offers a more uniform method for comparisons between the vowel qualities of languages. The formants F1 and F2 in Hz are changed to ERB values and presented on an ERB chart. Experiments using new Finnish formant data and the data from literature revealed that the minimal distance needed between the averages of two vowel categories in male speakers varies between 0,45 and 1,47 ERB (mean 1,06). The shortest distances were found between the vowels occurring in the pairs like *but - bath* in English, where a qualitative blending might be possible. The distances get smaller, if the clarity of speech is lower (cf. read natural clauses, read and spontaneous speech). Also the shorter vowels reduce the quality and they are centralized on the chart. In the data including only single read words, the average minimal distance 1,44 ERB was observed. The major part of all 20 single occurrences of each Finnish front vowel phoneme spoken by a male speaker fit within a circle with a diameter of 1,3 ERB (Fig. 6). The distribution is considerably greater in the back vowels.

Key words: psycho-acoustical vowel qualities, ERB scaled formant chart, minimal vowel distance, vowel quality variation, speaker normalization