

KAISA HÄKKINEN

Vokaalisynteesiin perustuva tutkimus suomen ja unkarin vokaalifoneemien rajoista

Yleensä arvellaan, että suomalaisen on melko vaivatonta oppia ääntämään unkarin vokaaleja idiomaattisesti oikein; samoin katsotaan myös unkarilaisen oppivan helposti tuottamaan suomen vokaalit korrektisti. Oppikirjojen ääntämisohjeissa kiinnitetään useimmiten huomiota vain *a*- ja *e*-äänteisiin, joiden asema on suomen ja unkarin foneemisysteemeissä erilainen. Muuten katsotaan foneemien reaalistamismien olevan näissä kielissä niin samankaltaiset, ettei erityisiä ohjeita tarvita. Sekä suomen että unkarin vokaalisto on tutkittu melko runsaasti spektrografianalyysin avulla, ja mm. Antti Sovijärvi on myös vertaillut näitä kahta kieltä keskenään. Nykyiset akustiset tutkimusmenetelmät tekevät mahdolliseksi kuitenkin myös uudenlaisen lähestymistavan: puhesyntetisaattorin avulla voidaan tuottaa täysin luonnollisen puheen kaltaista ääntä, joka on kontrolloitavissa kaikilta osiltaan. Vokaalien tuottamiseksi tarvitaan ainoastaan kaksi formanttivariaabelia, jotka vastaavat luonnollisen puheen kahta alinta formanttia (= F1 ja F2). F0:aa sekä ylempiä formantteja vastaavat komponentit voidaan säilyttää vakioina. Näin voidaan tuottaa kaikki luonnollisessa puheessa mahdollisesti esiintyvät vokaaliäänteet. Koska fyysikaalinen rakenne on näin yksinkertaistettu, eivät luonnollisten ja syntetististen vokaalien formanttiarvot ole keskenään täysin verrannollisia. Analysoitaessa on välttämättä otettava huomioon useampia muuttujia. Ihminen ei siis tuota puhetta yhtä taloudellisesti kuin syntetisaattori. Tämä verrannollisuuden puute ei kuitenkaan haittaa kahden synteesikokeen tulosten vertailua.

Tutkimuksessa käytetty syntetististen vokaaliärsykkeitten sarja koostui 190 erilaisesta äänneestä, joitten tuottamiseen oli käytetty Turun yliopiston OVE 3 b -syntetisaattoria. F1:n ääriarvot olivat 250 Hz ja 900 Hz,

vaihteluväli 50 Hz, ja F2:n arvot vaihtelivat 100 hertsin välein ääriarvoina 600 Hz ja 2700 Hz. Mainitut raja-arvot katsottiin aikaisempien analyysien perusteella riittäviksi. Tosin tulosten käsittelyvaiheessa kävi ilmi, että erityisesti unkarin kannalta olisi tarkistuksen vuoksi ollut hyvä ottaa mukaan vielä F1:n arvo 200 sekä joitakin F2:n arvoja 2700 Hz:n yläpuolelta. Kaikki foneemirajat tulivat kuitenkin näkyviin jo tässäkin tutkimuksessa käytettyjen arvojen perusteella. Koemateriaali oli äänitetty magnetofoninauhalle satunnaisessa järjestyksessä. Koehenkilöt kuuntelivat nauhan kahteen kertaan viikon välein, ja kummallakin kerralla he merkitsivät numeroituun listaan arvionsa juuri kuullun vokaalin kvaliteetista. Merkitsemiseen käytettiin koehenkilöitten äidinkielen käytössä olevan grafeemijärjestelmän merkkejä, tarkkeitten käyttö oli kielletty.

Tulokset koottiin tavallisille koordinaatioille, joissa abskissana oli F1:n arvo ja ordinaattana F2:n arvo. Jo tällaisen alustavan taulukoinnin perusteella oli mahdollista havaita suurin piirtein, millaiseksi lopullinen foneemirajakuvio tulisi muodostumaan. Lopullisten foneemirajojen määrittäminen tapahtui kahdessa vaiheessa: aluksi etsittiin rajaviivan sijaintialuetta osoittavien pisteitten joukko valitsemalla vuorotellen vakioksi kukin testissä käytetty F1:n ja F2:n arvo erikseen ja määrittämällä kutakin vakiota vastaava toisen formantin arvo koordinaatistosta, jossa y-akselina oli eri vokaalitulokintojen suhteellinen frekvenssi ja x-akselina muuttuvan formantin arvo. Rajapiste oli tällöin tulkintojen suhteellisen frekvenssin kuvaajien leikkauspisteen projektio x-akselilla. Aina ei rajapisteitten määrittäminen kuitenkaan ollut näin yksinkertaista, muutamissa tapauksissa projektioita oli enemmän kuin yksi. Näitten tapausten käsittelyä olen selvittänyt tarkemmin toisaalla.

Koska saatujen pistejoukkojen kuvaajaksi ei näyttänyt sopivan mikään käyrä erityisen hyvin, on lopullinen foneemirajakuvio laadittu mukailemalla lineaarisen regression periaatetta siten, että jokainen pistejoukko on puolitettu ja näille osajoukoille on laskettu aritmeettinen keskiarvo sekä X:n että Y:n suhteen ($X = \text{kaikki } x\text{:n arvot}$, $Y = \text{kaikki } y\text{:n arvot}$). Tämä menetelmä on valittu sen vuoksi, että laskutoimitukset saadaan huomattavasti yksinkertaisemmiksi kuin regressiosuoraa määrittäessä ja toisaalta kulmakerroin korostuu. Olen toisaalla verrannut kahden mainitun menetelmän antamia tuloksia, ja niiden perusteella laskettujen toisensa leikkaavien suorien ero näyttää jäävän hyvin pieneksi. On huomattava, että regressiosuorakin voidaan laskea samasta aineistosta sekä Y:n että X:n suhteen, eivätkä nämä suorat yleensä ole identtiset. Tässä tapauksessa on lisäksi muistettava, et-

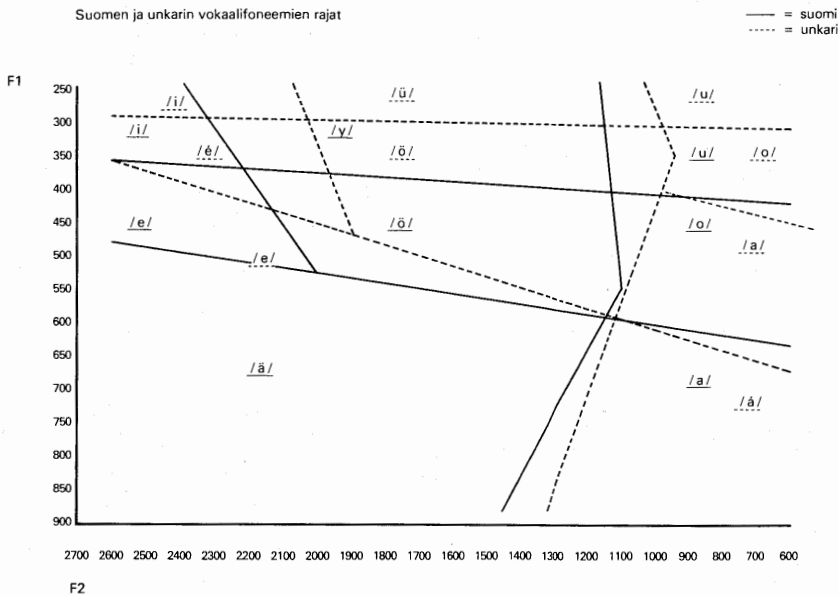
tä näkyviin piirretty rajaviiva kuvaa oikeastaan sitä foneemikartan kohtaa, jossa minkäänlaista foneemirajaa ei käytännössä ole. Tälle alueelle sattuvat vokaaliärsykkeet tulkitaan yhtä suurella todennäköisyydellä rajan molemmille puolille sattuviksi foneemeiksi. Se on kuitenkin raja siinä mielessä, että se osoittaa kohdan, jonka yli mentäessä tietyn tulokinnan todennäköisyys kasvaa. Mikäli koehenkilöitä olisi riittävästi, voitaisiin vastaavan testin avulla myös löytää alue, jossa tulkinta pysyy kaikkien koehenkilöitten kohdalla samana. Tätä aluetta voitaisiin nimittää foneemikeskukseksi. Koska tutkimuksessani oli käytettävissä vain yksi edustaja kummastakin kielestä, ei foneemikeskuksen määrittäminen tullut kysymykseen.

Tunnistamistestin suorittamisen edellytyksenä on tietoisuus kielen foneemijärjestelmästä. On varmistuttava siitä, että koehenkilö merkitsee kaikki eri foneemit eri merkeillä; sen sijaan ei ole haitaksi, vaikka koehenkilö käyttäisi samasta foneemista kahta tai jopa useampaa eri merkkiä, koska ne voidaan yhdistää tuloksia käsiteltäessä, edellyttäen tietenkin, ettei samoja merkkejä ole käytetty muista foneemeista. Tähän seikkaan ei suomea ja unkaria tutkittaessa tarvinnut kiinnittää erityistä huomiota, koska molempien kielten grafeemijärjestelmät kuvaavat erinomaisesti myös vokaalien fonemaattisia suhteita. Kvantiteetin kohdalla joudutaan kuitenkin näennäiseen ristiriitaan. Koehenkilöt käyttivät ärsykeitä tunnistessaan seuraavia merkkejä:

suomi	i	y	u	unkari	i	ü	u
	e	ö	o		é	ö	o
	ä		a		e		a
							á

Unkarilainen koehenkilö oletti kuulevansa ärsykeitten joukossa myös kaksi pitkää vokaalia. Kokeen järjestelyissä oli kuitenkin kaikki vokaalit laadittu fyysikaalisesti täsmälleen samanpituusiksi, joten pituusvastakohtasta ei missään tapauksessa voinut olla kysymys, vaikka koehenkilö vielä jälkekinpäin väitti ehdottomasti sellaisen havainneensa. Kvantiteettioppositioksi käsitetty vastakohta /e/—/é/ ja /a/—/á/ on siis itse asiassa kvalitatiivinen. Kielen kannalta on kuitenkin taloudellisinta käsitellä vastakohtaa kvantiteettioppositiona, koska muitten vokaalien kohdalla sellainen täytyy olettaa ja toisaalta diftongien puuttuminen osoittaa, että kahden vokaalisegmentin liittäminen samaan tavuun on kielenvastaista. Suomessa taas ei ole mitään estettä käsittää pitkiä vokaaleja kahden identtisen lyhyen segmentin muodostamaksi sekvenssiksi, joten pituus jäisi vain foneetti-

sen tason ilmiöksi. Vaikka suomen pitkät vokaalit poikkeavatkin laadullisesti jonkin verran vastaavista lyhyistä vokaaleista, ei ero ole niin suuri, että distinktio voitaisiin tehdä pelkästään sen perusteella. Näin ollen laadun muutos täytyy selittää ympäristöstä johtuvaksi. Näyttää siis siltä, että pituusopposition asema on näissä kielissä eritasoinen: suomessa se rajoittuu foneettiselle tasolle, unkarissa taas ilmiö on fonologinen. Tämä kiintoisa seikka ei kuitenkaan ole käsillä olevan tutkimuksen kannalta relevantti, koska käytetyn kaltaisten testiärsykkeitten avulla voi joka tapauksessa saada näkyviin vain kvalitatiiviset oppositiot. Laadullisesti kontrastoivia vokaaleja on suomessa kahdeksan, unkarissa yhdeksän. Jo lukumäärien perusteella on selvää, etteivät foneemirajat voi olla identtiset. Tämä käy havainnollisesti ilmi oheisesta foneemikartasta, johon suomen vokaalifoneemien rajat on merkitty yhtenäisellä viivalla, unkarin rajat katkoviivalla.



Vaikka kuviot ovatkin samantyyppiset, ei yhtään täysin yhtenevää rajaviivaa ole havaittavissa, joten suomen ja unkarin vokaalifoneemien realisatiot vokaalijärjestelmien näennäisestä samankaltaisuudesta huolimatta poikkeavat toisistaan varsin selvästi. Kuvion perusteella on mahdollista ennustaa, missä kohdin suomalainen kohtaa vaikeuksia opiskellessaan unkaria. Takavokaaleita on unkarissa yksi enemmän, joten ongelmia on ilmeisesti odotettavissa niiden kohdalla. Kuviosta näkyy suoraan, kuinka

niistä olisi selvittävä: Suomalaisen olisi opittava jakamaan oma /u/-foneeminsa hyvin suppeaksi unkarin *u*:ksi sekä suppeaksi *o*:ksi. Suomen /o/ vastaa täysin unkarin /a/:n realisaatiota, ja suomen /a/ peittää täysin unkarin /á/:n alan. Radikaali muutos eli tässä tapauksessa yhden vokaalifoneemin jakaminen kahdeksi täytyy suorittaa kohdalla, joka pelkääntään foneemisysteemiä tarkasteltaessa näyttää olevan täsmälleen samanlainen sekä suomessa että unkarissa: molemmissa kielissä on toki /u/-foneemi. Mutta mikäli suomalainen käyttäisi unkarin /u/:ta tavoittellessaan vapaasti kaikkia oman /u/-foneeminsa allofoneja, olisi odotettavissa, että unkarilainen kuulija tulkitseisi osan *u*-äänteistä /o/:n allofoneiksi. Myös muitten suppeitten vokaalien kohdalla on odotettavissa samanlaisia ongelmia, sillä suomen allofonit näyttävät olevan huomattavasti avoimempia, ts. F1 on niissä korkeampi. Periaatteessa kaikki seuraavat tulkinnot näyttäisivät mahdollisilta.

1. Suomen vokaalifoneemi

/i/

/y/

/u/

/e/

/ö/

/o/

/ä/

/a/

Unkarilainen tulkinta

/i/, /è/

/ü/, /i/, /è/, /ö/

/u/, /ü/, /o/, /ö/

/e/, /è/

/ö/, /e/, /è/

/a/, /o/, /ö/

/e/

/ä/, /a/, /ä/

2. Unkarin vokaalifoneemi

/i/

/ü/

/u/

/è/

/ö/

/o/

/e/

/a/

/ä/

Suomalainen tulkinta

/i/, /y/

/y/, /u/

/u/

/i/, /y/, /e/, /ö/

/y/, /u/, /ö/, /o/

/u/, /o/

/e/, /ä/, /ö/

/o/, /a/

/a/

Tällaisen vertailun valossa näyttää siltä, ettei unkarilainen voi oman foneemisysteeminsä perusteella tulkita sataprosenttisen varmasti oikein yhtäkään suomen vokaaliäännettä. Tosin hän kuulee johdonmukaisesti

suomen /ä/:n allofonit oman /e/:nsä vastineiksi, mutta toisaalta hän ei pysty erottamaan suomen /ä/:tä suomen /e/:stä. Tämä on helposti kuultavissa jokaisen suomea opiskelevan unkarilaisen puheesta. Suomalaisen kannalta on tilanne sikäli valoisampi, että hän voi olla aina varma oikeasta tulkinnasta unkarin /u/:n kohdalla, ja lisäksi unkarin /á/ yhdistyy aina suomen /a/:han. Tällöin tosin saattaa syntyä aivan erityisiä vaikeuksia kvantiteetin suhteen. Koska unkarilainen erottaa /a/:n ja /á/:n ensisijaisesti laadun perusteella, on luultavaa, että /á/realisoituu usein hyvin lyhyenä. Tällöin suomalainen katsoo sen ilman muuta lyhyeksi *a*-äänteeksi ja yhdistää sen unkarin /a/-foneemiin. Unkarin /e/:tä tavoitellessaan voi suomalainen käyttää mitä tahansa *e*- tai *ä*-äännettä, paitsi erittäin suppeita /e/:nsä allofoneja, ja lopputulos on unkarilaisen kannalta yksiselitteinen.

Kuvatun kaltainen kuuntelutesti on kuitenkin keinotekoinen siinä mielessä, että kaikki esitetyt testiärsykkeet oli tulkittava jollakin tavoin. Todellisuudessa vokaalikolmion ala ei jakaudu pelkästään vokaalien kesken, vaan samalle alueelle sijoittuu myös nasaalien ja likvidoitien formantteja. Lisäksi pelkän kuuntelutestin avulla on mahdotonta määritellä, mitä osaa kunkin foneemin rajojen sisäpuolelle jäävästä alueesta todellisuudessa käytetään. Vaikka luonnollisessa puheessa tuskin koskaan esiintyykään kahta täysin samanlaista äännettä, tuotoksessa pyritään kuitenkin aina kohti tiettyä normia, foneemikeskusta. Foneemikeskuksen käsite ei kuitenkaan ole määriteltävissä yksinkertaisesti esim. tiettyjen hertsiarvojen avulla. Äänteet esiintyvät näet irrallisina vain poikkeustapauksissa; esim. suomen kielessä ei ole yhtään ainoata yhden lyhyen vokaalifoneemin muodostamaa sanaa. Käytännössä on siis aina otettava huomioon myös ympäristön vaikutus, ja kuvatun kaltaisessa synteettisten vokaalien tunnistamistestissä tämä ei tullut kysymykseen. Koska irrallisten vokaalien tuottaminen kuitenkin on helppoa myös luonnollisessa puheessa, olisi tietysti mahdollista analysoida tällaisia äänneitä ja verrata niitä synteettisten antamiin tuloksiin. Mutta kuten jo yllä on käynyt ilmi, eivät luonnollisen ja synteettisen vokaalien formanttiarvot ole täysin verrannolliset keskenään. Lisäksi perusongelma jäisi edelleen selvittämättä. Irrallisten vokaalien ja puhevirran osina esiintyvien vokaaliäänteitten tuottamista säätelevät osittain eri tekijät: vokaaliääne, jonka tunnistaminen irrallisena saattaisi tuottaa suuriakin vaikeuksia, voi tietyissä ympäristöissä olla täysin korrekti ja yksiselitteisesti tulkittavissa. Absoluuttisia rajoja ei ole. Tämä käy selvästi ilmi, jos esim. analysoimme möreä-äänisen miehen ja pikkutyön puhetta. Mikäli tällaiset henkilöt lausuisivat vain yhden ainoan äänteen, esim. *a*-vokaalin, joka analysoitaisiin, olisivat mittaustulokset esim. formanttien

frekvenssien osalta varsin erilaiset. Tästä huolimatta olisi mahdollista tunnistaa vokaali molemmissa tapauksissa /a/:ksi, vaikka minkäänlaisia havaintoja puhujien iästä, sukupuolesta ja puhettomuksista ei etukäteen olisi voitu tehdä. Vaikka foneemiraja onkin testitilanteessa mahdollista saada näkyviin melko jyrkkänä ja selkeänä, sen eksplisiittinen määrittäminen tuntuu olevan mahdotonta. Se ei ilmeisesti perustu pelkkiin fysikaalisiin realiteetteihin, vaan kuulohavainnon psykologinen puoli olisi otettava myös huomioon. Puhevirran sisältämä informaatio ei suinkaan ole kunkin erillisen segmentin sisältämän informaatiomäärän summa. Puheessa on aina redundanssia, joka auttaa meitä havaitsemaan sellaistaakin, mitä itse asiassa ei lainkaan ole olemassa: kuulemme sen, mitä odotamme kuulevamme. Esim. sanan viimeisen vokaalin informaatioarvo on usein hyvin pieni — poikkeuksena on tietysti esim. adessiivin ja allatiivin vastakohta —, eikä puheen ymmärrettävyyttä yleensä heikennä lainkaan, jos se jätetään kokonaan pois. Mikäli kuulijaa pyydetään yllättäen toistamaan sana, hän saattaa huomaamattaan täydentää sen. Tässä valossa ei ole lainkaan hämmästyttävää, että foneemien realisaatiot saattavat vaihdella hyvinkin paljon. Formantiarvojen käyttäminen foneemirajakuvioitten laadinnassa on vain apukeino, pelkillä luvuilla ei juuri ole merkitystä muuten kuin suhteessa toisiinsa. Sitä paitsi koko formantikäsité on melkoisen teoreettinen, fysikaalisesti tarkastellen vokaalit ovat toki paljon muitakin kuin pari kolme energiakasautumaa eri taajuuksilla. Mutta kuten erityisesti synteesikoheet osoittavat, on formantikäsitteen käyttöönotto erittäin onnistunut askel havaintopsykologian suuntaan. Ääni ja puhe ovat kaksi eri asiaa.

KAISA HÄKKINEN: *Untersuchung zur Abgrenzung der finnischen und ungarischen Vokalphoneme auf der Grundlage der Vokalsynthese*

Da das finnische und das ungarische Vokalsystem sehr ähnlich sind, hält man es im allgemeinen für selbstverständlich, dass der Finne die Artikulation der ungarischen Vokale leicht lernt. Jedoch lassen sich Lernschwierigkeiten nicht einfach aufgrund der Phonemsysteme voraussehen, denn die idiomatische Ausspracheweise dürfte auch dann verschieden sein, wenn die Phonemsysteme absolut identisch sind. Im Falle des Finnischen und Ungarischen sind auch die phonologischen Unterschiede beachtlich; so lässt sich z.B. die finnische Quantitätsopposition nur als phonetisches Phänomen erklären, während die ungarische Opposition unbedingt phonologischen Charakter hat. Die ungarische Opposition /a/ — /ä/ und /e/ — /é/ realisiert sich jedoch qualitativ, so dass besonders die Beziehungen zwischen den angeführten ungarischen Vokalen und dem finnischen /e/, /a/ und /ä/ sehr kompliziert sind.

Als Beobachtungsmaterial benützt die Studie eine Serie von 190 synthetischen Vokalen, wobei die Werte von F1 und F2 (F1 250-900 Hz und F2 600-2700 Hz) als Variable dienten. Aufgrund der Identifizierungen dieser Reihe wurde eine Vokalkarte sowohl des Finnischen als auch des Ungarischen erstellt, die die F1- und F2-Koordinaten der als Allophone des jeweiligen Phonems klassifizierten Laute verzeichnet. Aus der Position der Grenzen lässt sich u.a. schliessen, dass man im allgemeinen die Allophone des finnischen /o/ zur Realisierung des ungarischen /a/ verwenden kann; ebenso scheinen die Allophone des finnischen /a/ für den Phänotypus des ungarischen /á/ geeignet zu sein. Ausserdem lässt sich daraus schliessen, dass die ungarischen hohen und Zwischenvokale wesentlich enger gesprochen werden als die entsprechenden finnischen.