

XXX Fonetikan päivät

Oulu 13.–14.5.2016

Julkaisut – Papers



Pentti Körkkö & Riikka Ylitalo (toim.)



XXX Fonetiiikan päivät

Oulu 13.–14.5.2016

Julkaisut – Papers

Pentti Kärkkö ja Riikka Ylitalo (toim.)

Oulun yliopisto

HuTK, Fonetikka 2016

ISBN 978-952-62-1479-5

Sisälyys

Esipuhe

Artikkelit

Paula Laine & Michael O'Dell

Suomen kielen alveolaarisen frikatiivin soinnillisuudesta.....1

Michael O'Dell & Tommi Nieminen

Suomen konsonanttiyhtymässä olevan lateraalin akustinen variaatio ja murteiden epenteettinen vokaali.....10

Lauri Tavi & Stefan Werner

Automaattisesta kohdistamisesta apua katoavan murteen nimikointiin?.....20

Joonas Vakkilainen

Puheenkierrätystutkimus suomessa.....29

Riikka Ylitalo

On the prosody of CV.CV words in the Eastern Finnish dialects.....36

Abstraktit

Immonen & Peltola.....46

Jähi, Alku & Peltola.....47

Kakouros & Räsänen.....48

Kallio.....49

Kalvik & Piits.....50

Kautonen & Heinonen.....51

Körkkö.....53

Laine P. & O'Dell.....54

Laine U.....55

Lehtinen & Kallio.....56

Malmi, Meister & Lippus.....57

Niemi & Koponen.....58

O'Dell & Nieminen	59
Palo, Schaeffler & Scobbie	60
Palo, Schaeffler & Scobbie	62
Räsänen	63
Seshadri & Laine U.	64
Seshadri, Remes & Räsänen	65
Taimi, Alku & Peltola	66
Tavi	67
Tavi & Werner	68
Teras, Asu, Lippus & Pajusalu	69
Tergujeff, Kuronen & Kautonen	70
Vakkilainen	71
Virkkunen & Vainio	72
Ylitalo	73

Esipuhe

Fonetiikan päivillä on takanaan jo varsin pitkä historia, sillä Oulussa 13.–14.5.2016 järjestetyt päivät olivat jo kolmannetkymmenen. Mukana oli 15 esitelmää ja 10 posteria, ja päiville osallistujia oli 35. Esitelmien ja postereiden aiheet koskettelivat mm. akustiikkaa, prosodiikkaa, kielenoppimisen fonetiikkaa ja puherytmiä. Päivät pidettiin Oulun keskustassa, uuden kauppakeskus Valkean yhteydessä sijaitsevan Sokos Hotel Arinan kokoustiloissa, ja konferenssi-illallinen nautittiin samoilla kulmilla Vanhan Paloaseman Ullakolla.

Olemme iloisia, että saimme järjestää 30. Fonetiikan päivät Oulussa. Haluamme lämpimästi kiittää kaikkia osallistujia, esitelmien ja posterien tekijöitä sekä kirjoittajia. Seuraaville XXXI Fonetiikan päiville kokoonnumme Jyväskylään maaliskuussa 2017.

Oulussa 1. joulukuuta 2016

Järjestelytoimikunta

Pentti Körkkö

Riikka Ylitalo

Suomen kielen alveolaarisen frikatiivin soinnillisuudesta

Paula Laine & Michael O'Dell

Tampereen yliopisto

Abstrakti: Tämän tutkimuksen tarkoituksena oli selvittää äänneympäristön vaikutusta alveolaarisen frikatiivin esiintymiseen soinnillisena suomen kielessä. Tarkoituksena on myös selvittää miten frikatiivien soinnillisuutta voitaisiin mitata ja luokitella.

Aineiston perusteella näyttää selvältä, että soinnillista varianttia tästä äänneestä käytetään. Soinnillinen variantti näyttäisi olevan soinnittoman äänten kanssa vapaassa vaihtelussa. Tämän tutkimuksen tarkoituksena on selvittää, onko löydettävissä jotain selkeää äänneympäristön vaikutusta alveolaarisen frikatiivin soinnillisuuteen.

Selvää on myös, että soinnillisuudessa ei ole kyse joko–tai-ilmiöstä, vaan kyseessä on jatkumo täysin soinnillisen ja soinnittoman äänten välillä. Soinnin voimakkuus vaihtelee frikatiiveissa siten, että täysin soinnillisten ja täysin soinnittomien äänteiden välille jää suuri joukko tapauksia, joissa sointi ei ole täysin selvää. Välillä sointi on hyvinkin voimakasta ja selkeää ja toisinaan se taas saattaa olla hyvinkin heikkoa. Toisaalta soinnin määrä vaihtelee myös sen mukaan, esiintyykö sointia koko äänten ajan vai vain osassa äännettä. Myös tätä soinnin ja frikaation suhdetta pyrittiin tässä tutkimuksessa selventämään.

Äänneympäristön lisäksi myös muut tekijät voivat vaikuttaa soinnillisuuteen. Kyse saattaa olla esimerkiksi puhujakohtaisesta ilmiöstä siten, että toiset puhujat käyttävät soinnillista äännettä joskus ja toiset eivät koskaan. Puhujakohtaisen variaation tutkimiseen tarvittaisiin kuitenkin perehtymistä laajempaan, eri puhujilta koottuun aineistoon, eikä siihen tämän tutkimuksen puitteissa valitettavasti ole mahdollisuuksia.

Avainsanat: frikatiivi, soinnillisuus, suomen kieli

1 Johdanto

Tämän tutkimuksen tarkoituksena on selvittää äänneympäristön vaikutusta soinnillisen alveolaarisen frikatiivin esiintymiseen suomen kielessä. Lisäksi tarkoituksena on etsiä suomen kieleen soveltuvaa tapaa mitata frikatiivin soinnillisuutta ja jaotella esiintymiä soinnin mukaan. Frikatiivien soinnillisuutta ei ole juurikaan tutkittu suomen kielessä. Joitain tutkimuksia frikatiivien soinnillisuudesta löytyy muista kielistä, esimerkiksi espanjasta (ks. esim. Hualde ja Prieto, 2014), mutta ylipäättään soinnillisuutta on tutkittu enemmän klusiilien kohdalla.

Toisinaan kuulee väitettävän, että suomen kielessä esiintyvä alveolaarinen frikatiivi on

aina soinniton. Monissa suomen äännejärjestelmää esittelevissä oppikirjoissa mainitaan edelleen, että suomessa on vain soinniton /s/ vaikka samalla joidenkin muiden äänteiden erilaisia variaatioita esitellään enemmän tai vähemmän tarkasti. Tämän tutkimuksen tavoitteena on osaltaan tuoda esiin, että myös soinnillinen variantti /s/-äänteelle on mahdollinen ja lisäksi selvittää, miten äänneympäristö siihen vaikuttaa.

Soinnillisuuden mittaamiseen ja aineiston analyysiin käytettiin Praat-ohjelmaa (Boersma & Weenink, 2016). Pyrittiin myös selvittämään, miten yleisiä täysin soinnilliset alveolaariset frikatiivit ovat suomen kielessä. Tutkimuskysymykset olivat: miten äänneympäristö vaikuttaa /s/-äänteen soinnillisuuteen suomen kielessä, miten soinnillisuutta voisi mitata ja kuinka yleistä se on?

Aiempaa tutkimusta suomen kielen alveolaarisen frikatiivin soinnillisuudesta on hankalaa löytää. Soinnillisuuteen liittyvää tutkimusta löytyy enemmän klusiileista (ks. esim. Suomi, 1980). Frikatiiveihin ja niiden soinnillisuuteen liittyvää tuoreempaa tutkimusta löytyy lähinnä espanjan kielestä. Espanjan kieli onkin alveolaarisen frikatiivin suhteen siinä mielessä samanlainen kuin suomi, että siinäkin kyseisen äänteen soinnillisuus tai soinnittomuus ei vaikuta merkitykseen. Sen sijaan espanjan kielessä alveolaarisella frikatiivilla on myös soinnillinen allofoni. Gradoville (2011) on tutkinut erilaisia tapoja mitata frikatiivien soinnillisuutta Argentiinan espanjassa. Hän käsittelee tutkimuksessaan useita erilaisia laskennallisia tapoja mitata frikatiivien sointia. Hänen mukaansa Praatin pulsseihin perustuva (*pulse-based*) sointiraportti ja matalien taajuuksien voimakkuuksien vertaaminen koko spektrin voimakkuuteen (*low-frequency to total intensity ratio*) antavat parhaat tulokset verrattuna audiitiiviseen tarkkailuun sekä siihen, mitä spektrogrammissa voidaan nähdä. Espanjan kielestä tehdyn tutkimuksen perusteella ei voida tehdä selkeitä päätelmiä siitä, mikä mittaustapa sopii suomen kieleen, mutta tässä tutkimuksessa päädyttiin käyttämään näistä ensimmäistä, eli Praatin automaattisesti luotua sointiraporttia. Tämä mittaustapa valittiin siksi, että se on verrattain yksinkertainen tapa mitata soinnillisuutta.

2 Aineisto

Tutkimuksessa käytettiin aineistona yhtä tunnin mittaista haastattelua. Haastattelu on osa TAPU-aineistoa, jossa nauhoitettiin 30 tamperelaista puhujaa. Samoja puhujia nauhoitettiin vuosina 1977 ja 1997. Aineisto kerättiin aikanaan variaation ja puhekielen muutoksen tutkimusta varten. Kaikki aineiston puhujat ovat syntyperäisiä tamperelaisia ja asuneet suurimman osan elämästään Tampereella. Tässä tutkimuksessa käytetty haastattelu on nauhoitettu vuonna 1997. Nauhoitteella esiintyvä puhuja on nainen.

Juoksevassa puheessa tulee esiin erilaisia suprasegmentaalisia tekijöitä, kuten kuiskaus, narinaa ja intensiteetin vaihteluita. Nämä tekijät ovat toisaalta hyvin mielenkiintoisia, mutta toisaalta hankaloittivat mittausten tekemistä. Näiden tekijöiden huomioiminen tutkimuksessa vaatisi laajemman aineiston. Koska nauhoitus on alun perin tehty variaation-

tutkimuksen tarpeisiin, se ei ole kaikilta osiltaan kovin selkeä. Äänitteellä esiintyy jonkin verran taustahälyä ja toisinaan myös päällekkäispuhuntaa haastattelijan ja haastateltavan välillä. Myös nämä tekijät vaikeuttivat äänteiden segmentointia ja vaikuttavat mahdollisesti myös analyysin tuloksiin ja niiden luotettavuuteen. Tämä otettiin huomioon tilastollisessa analyysissä suurempana virhemarginaalina.

3 Menetelmät

3.1 Aineiston analyysi ja siinä esiinnousseet ongelmat

Tutkimuksessa otettiin mukaan ainoastaan vokaalienväliset, sanansisäiset ja fonologisesti lyhyet /s/-äänteet. Tähän rajaukseen päädyttiin, koska soinnillinen /s/-äänne esiintyy todennäköisemmin soinnillisessa kuin soinnittomassa ympäristössä. Ympäristö rajattiin pelkästään vokaaleihin ja mukaan otettiin vain sanansisäiset esiintymät, koska segmentointi oli työlästä ja tutkimuksen laajuus huomioon ottaen tällä rajauksella löydettiin riittävästi esiintymiä (385 kappaletta). Praatia käyttäen segmentoitiin kaikki rajaukseen sopivat /s/-äänteet.

Kaksi esiintymää jouduttiin hylkäämään päällekkäispuhunnan vuoksi. Aineistossa esiintyi usein sana semmosia. Segmentointia tehdessä kuitenkin huomattiin, että monesti tämä sana sai juoksevassa puheessa muodon semmsia, eikä siinä siis enää esiintynyt vokaalien välistä /s/-äännettä. Tällaisia tapauksia ei otettu mukaan. Luonnollisesti sellaiset tapaukset, joissa /o/-äänne säilyi, otettiin mukaan.

Segmentointiin tuo omat haasteensa se, että normaali puhe ei koostu selkeistä peräkkäin esiintyvistä ja toisistaan erillisistä äänteistä, vaan puheen aikana siirrytään äänteestä toiseen mahdollisimman sujuvasti. /s/-äänten kohdalla tämä tarkoittaa esimerkiksi sitä, että siirtymä vokaalista /s/-äänteeseen ja taas takaisin vokaaliin tapahtuu siten, että selkeää rajaa äänteiden välille on toisinaan hyvin vaikea vetää. Kahden vokaalin välissä olevan /s/-äänten aikana kieli pyrkii muodostamaan frikatiivin vaatiman sulkeuman hammasvallia vasten, mutta se, kuinka voimakas sulkeuma on ja kuinka pitkäksi aikaa se ehditään muodostaa, riippuu esimerkiksi puhenopeudesta. Esimerkiksi kahden pitkän vokaalin välissä oleva lyhyt /s/ saattaa nopeassa puheessa olla hyvin heikosti frikatiivinen, koska sulkeumaa ei ehditä muodostaa kunnolla.

3.2 Frikatiivin soinnillisuuden mittaamisesta

Frikatiivin soinnillisuuden mittaaminen ei ole niin yksiselitteistä kuin esimerkiksi klusiilien. Frikatiiveissa sulkeuma ei ole tiivis, vaan päästää osan ilmavirrasta lävitseen, joten ei ole kovin yksinkertaista päättää, milloin frikatiivi alkaa ja milloin se loppuu. Frikatiiville tyyppillinen hälyääni saattaa joskus jatkua heikompana jonkin aikaa myös vokaalien aikana. On ajateltu, että soinnittoman frikatiivin aikana ei esiinny äänihuulten värähtelyä lainkaan ja

toisaalta soinnillisen variantin aikana äänihuulten värähtelyä esiintyy ainakin jonkin aikaa myös sinä aikana, kun frikatiivin sulkeuma on riittävän pieni tuottaakseen frikaatiohälyä (Stevens, Blumstein, Glicksman, Burton & Kurowski, 1992).

Päädettiin käyttämään soinnillisuuden mittaamisessa Praatin automaattisesti luomaa sointiraporttia. Äänitiedostosta on mahdollista Praatin avulla luoda sävelkorkeus-objekti (Pitch). Tämän avulla voidaan tarkastella, missä kohden äänitiedostossa on sointia ja missä kohden ei. Lisäksi tästä sointiraportista voidaan luoda äänitiedoston yhteyteen uusi tekstirivi (TextGrid), jolle soinnilliset ja soinnittomat kohdat on merkitty kirjaimin V (*voiced*) ja U (*unvoiced*). Sointiraporttia tutkittiin yhdessä äänitiedoston ja tekstirivin kanssa. Pian kuitenkin huomattiin, että automaattisesti luotu sointiraportti ei ollut kovin tarkka ja oli oikeastaan useassa kohdin väärässä soinnin suhteen. Tämän vuoksi päätettiin tarkastamaan tämä sointiraportti manuaalisesti tarkastelemalla spektogrammia kaikkien /s/-äänteiden ja niitä ympäröivien vokaalien kohdalta. Lopputuloksena oli TextGrid, jonka toisella tasolla (Tier) oli merkitty /s/:t ja viereiset vokaalit ja toisella tasolla oli merkittynä soinnilliset ja soinnittomat kohdat käsin tarkasteltuna /s/:ien ja niiden ympäristön osalta.

3.3 Tehdyt mittaukset

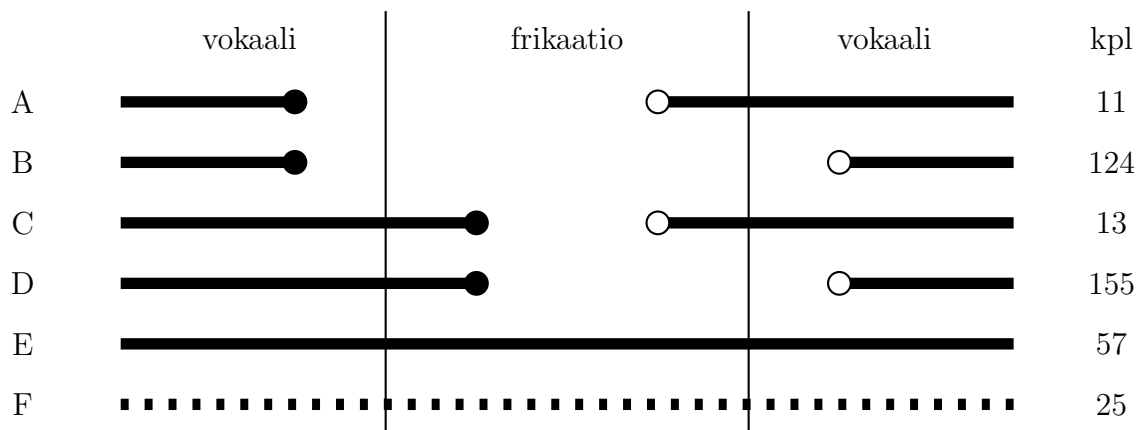
Praatilla luotiin skripti, joka etsii jokaisen /s/-äänteen alku- ja loppukohdat ja tarkistaa, onko kohta soinnillinen vai soinniton. Jos /s/-äänteen alkukohta on soinniton, skripti etsii myös edellisen soinnin loppukohdan ja jos se on soinnillinen, seuraavan soinnin loppukohdan. Samoin tehtiin /s/-äänteen loppukohdille. Näin saatiin selville, missä kohden s-äännettä — tai sen ympäristössä — soinniton osio kussakin tapauksessa alkaa ja loppuu. Soinnittomalla osiolla tarkoitetaan tässä sellaista ajanjaksoa, jolloin sointia ei esiinny, ainakaan siten, että se olisi spektogrammissa havaittavissa.

4 Havainnot

4.1 Soinnin ja frikaation suhde

Sointi ja frikaatio voivat esiintyä erilaisissa suhteissa toisiinsa. Kuva 1 havainnollistaa soinnin ja frikaation mahdollisia suhteita. Pystyviivat osoittavat frikaation alkua ja loppua. Tienkään frikatiivit eivät ole aineistossa samanpituisia, mutta tässä kuvassa näin on esitetty selkeyden vuoksi. Vaakasuurat viivat puolestaan osoittavat soinnin kulkua. Musta pallo viivan päässä tarkoittaa soinnin loppumista ja valkoinen pallo soinnin alkamista.

Sointi voi siis loppua joko ennen tai jälkeen frikaation alkamisen. Samoin sointi voi alkaa uudelleen ennen tai jälkeen frikaation loppumisen. Mahdollista on myös, että sointi ei frikaation aikana lainkaan lopu (tapaus E). Näitä täysin soinnillisia tapauksia oli aineistossa 57 kappaletta, eli noin 16 %. Jonkin verran aineistossa oli myös sellaisia esiintymiä, joiden aikana sointia ei lainkaan esiintynyt (tapaus F, 25 kpl), toisin sanoen myös viereiset vokaalit



Kuva 1: Soinnin ja frikaation suhde

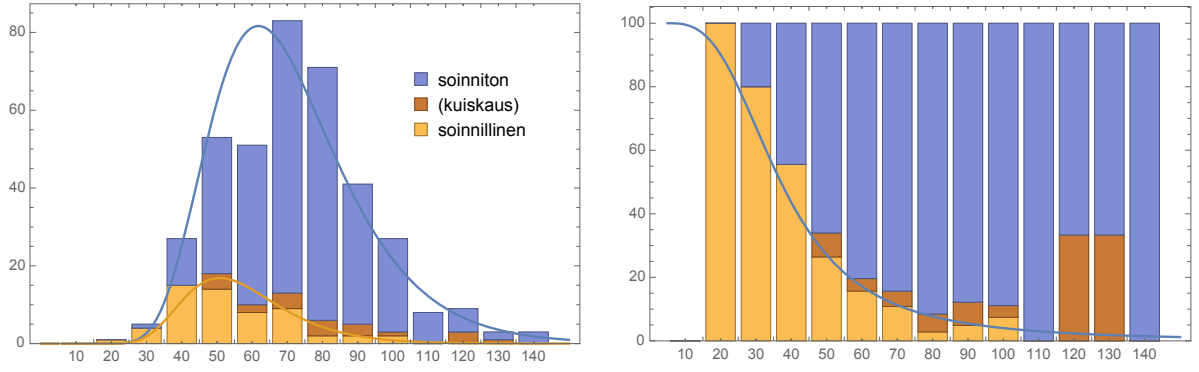
olivat täysin soinnittomia. Tällöin on kyse yleensä kuiskauksesta. Näitä tapauksia käsiteltiin tilastollisten menetelmien yhteydessä niiden erikoisen luonteen vuoksi siten, että sointia pidettiin puuttuvana tietona (frikaation kesto on tiedossa, muttei tiedetä millainen sointi olisi ollut ympäristön ollessa soinnillinen).

Kuvasta 1 nähdään, että selvästi yleisin tapaus tässä aineistossa on sellainen, jossa sointi loppuu frikaation alkamisen jälkeen ja alkaa uudelleen sen loppumisen jälkeen (155 kpl). Soinniton osio siis kulkee hieman frikaation jäljessä. Tämä voisi viitata passiiviseen ääntötapaan, jossa sointia ei tietoisesti katkaista, vaan se frikaation jatkuessa katkeaa itsestään. Toisaalta, myös päinvastaista tapausta, eli sellaista, jossa sointi loppuu ennen frikaation alkamista ja alkaa uudelleen ennen frikaation loppumista, löytyi aineistosta jonkin verran (11 kpl). Melko paljon (124 kpl) oli myös sellaisia tapauksia, joissa soinniton osio jatkuu jonkin verran molempien vokaalien puolelle. Tämä oli aineistossa toiseksi yleisin tapaus. Passiivisella ääntämisellä ei siis voida selittää läheskään kaikkia erilaisia tapauksia.

4.2 Täysin soinnillisten tapausten esiintyminen

Kuva 2 esittää soinnillisten (keltainen) ja soinnittomien (sininen) tapausten suhteen erikes-
toisissa frikatiiveissa. Oranssilla merkityt tapaukset ovat sellaisia, joissa koko ympäristö oli soinniton (kuiskattu), joten tieto frikatiivin soinnillisuudesta puuttuu.

Vasemmanpuoleisessa histogrammissa esitetään esiintymien todelliset määrät keston (ms) mukaan lajiteltuna. Vaaka-akselilla siis näkyy frikaation kesto ja pystyakselilla esiintymien määrä. Oikeanpuoleisessa kuvassa taas samat esiintymät on esitetty suhteellisen määrän mukaan. Tässä siis pystyakselilla on absoluuttisten määrien sijaan prosenttiosuus. Tästä kuvasta nähdään, kuinka frikaation kesto selvästi vaikuttaa soinnillisuuteen. Mitä lyhempi frikatiivi on, sen todennäköisemmin se on täysin soinnillinen. Tämäkin voisi viitata soinnittomuuden passiivisuuteen. Koska /s/:n ei tarvitse olla soinniton merkityksen vuoksi, ei lyhyessä äänteessä aina ehditä katkaista sointia. Toisaalta, täysin soinnillisia tapauksia on myös pi-

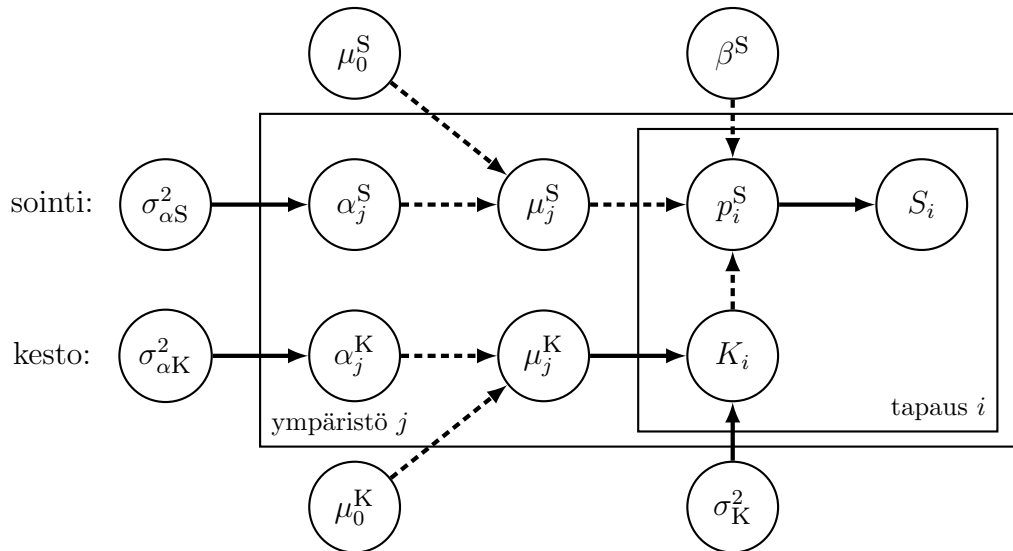


Kuva 2: Täysin soinnillisten tapausten suhde frikaation kestoon

dempien frikatiivien joukossa, joten kaikkia soinnillisia tapauksia passiivinen ääntäminen ei selitä.

4.3 Tilastollinen analyysi

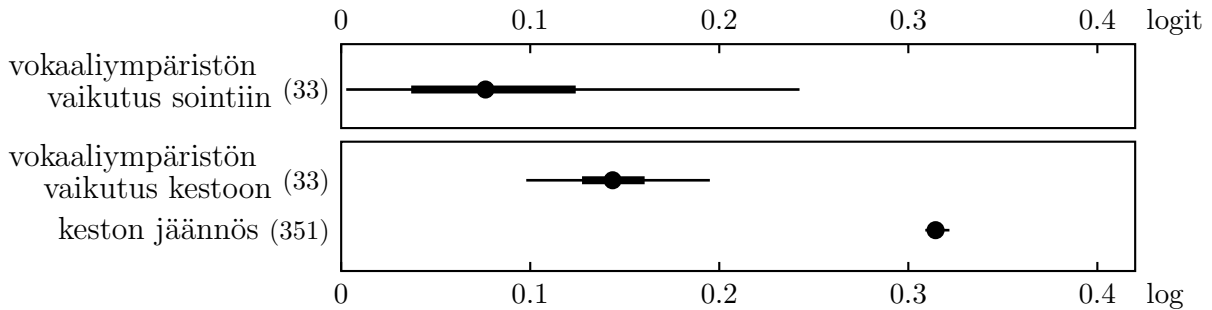
Bayesilaisessa analyysissä käytetty hierarkkinen malli (Gelman & Hill, 2007) esitetään kuvassa 3. Kuvassa yhtenäinen nuoli edustaa stokastista yhteyttä, katkoviivanuoli loogista (laskennallista) yhteyttä. Mallissa havaittu muuttuja S_i kertoo, oliko tapaus (kokonaan) soinnillinen, havaittu muuttuja K_i on taas frikaation kesto. Soinnillisuuden todennäköisyys p_i^S määräytyy keston funktiona: $p_i^S = \text{logit}^{-1}(\beta^S(\ln K_i - \mu_j^S))$. Parametri μ_j^S kertoo mikä kesto vastaa soinnillisuuden 50 % todennäköisyyttä, kulmakerroin β^S taas kertoo kuinka nopeasti soinnillisuuden todennäköisyys vähenee keston kasvaessa. Frikaation kesto noudattaa tilas-



Kuva 3: Tilastollinen malli: Soinnillisuuden S_i todennäköisyys $p_i^S = \text{logit}^{-1}(\beta^S(\ln K_i - \mu_j^S))$, frikaation kesto $K_i \sim \text{LogNorm}(\mu_j^K, \sigma_K^2)$, ANOVA-tyyppiset ympäristön efektit α_j^S ja α_j^K .

tollista lognormaalijakaumaa $K_i \sim \text{LogNorm}(\mu_j^K, \sigma_K^2)$. Lisäksi mallissa on ANOVA-tyyppiset vokaaliympäristön efektit α_j^S ja α_j^K , jotka vastaavat eri vokaaliympäristöjen mahdollisia vaikutuksia soinnillisuuteen ($\mu_j^S = \mu_0^S + \alpha_j^S$) ja frikaation kestoon ($\mu_j^K = \mu_0^K + \alpha_j^K$).

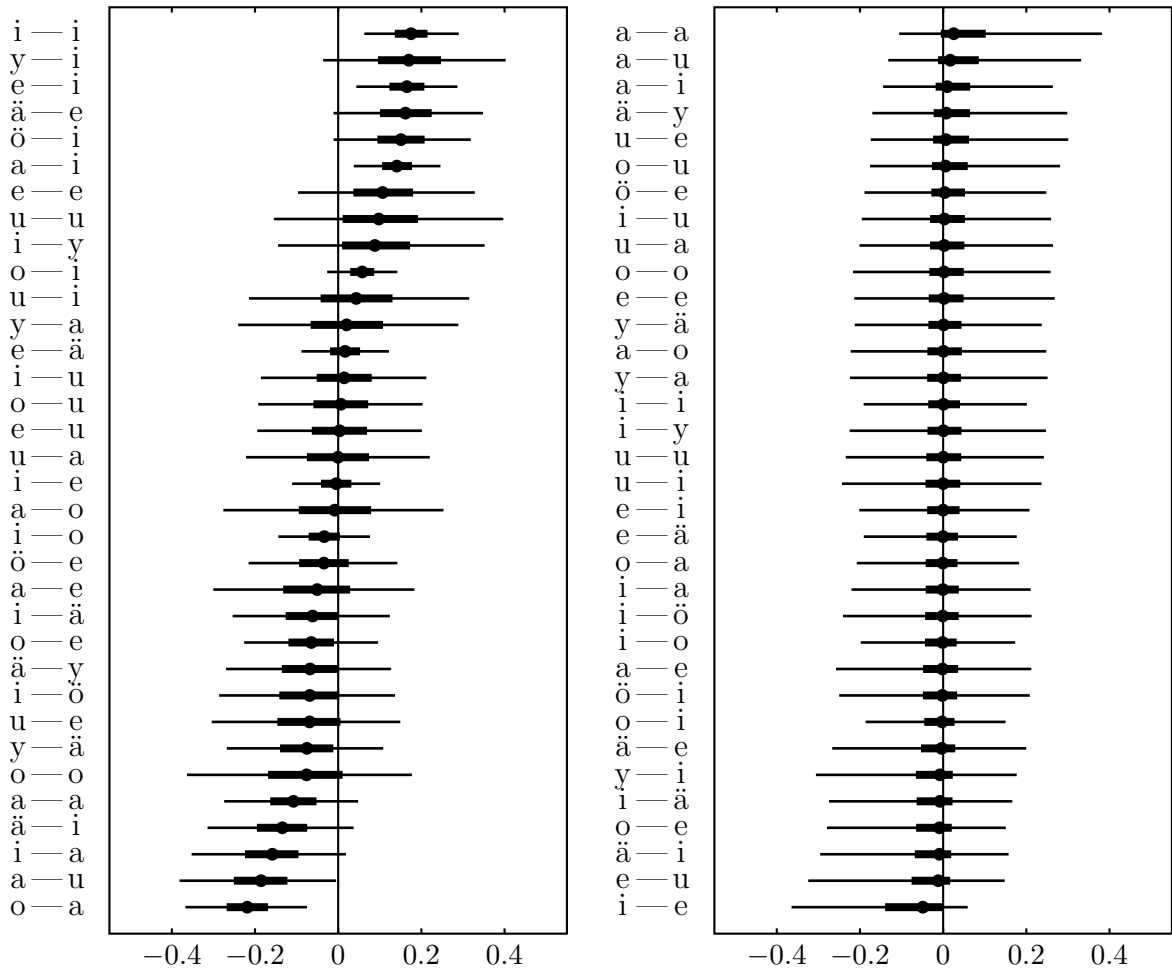
Kuten huomaa jo kuvasta 2, frikaation kestolla oli erittäin selvä vaikutus soinnillisuuteen: kulmakertoimen β^S 95 % luottamusväli (CI), $(-3.82, -2.05)$, osoittaa, että kulmakerroin on melko varmasti negatiivinen, eli soinnillisuuden todennäköisyys laskee keston kasvaessa. Käyrä oikeanpuoleisessa kuviossa kuvassa 2 näyttää soinnillisuuden todennäköisyysfunktion (ilman vokaaliympäristön mahdollisia vaikutuksia) posteriorimediaaniarvoja käyttäen: $p^S = \text{logit}^{-1}(-2.92(\ln K + 3.44))$. Kuvan 2 vasemmanpuoleisissa histogrammissa olevat käyrät taas näyttävät kestojen todennäköisyysjakauman posteriorimediaaneja käyttäen. Tässäkin vokaaliympäristön mahdolliset vaikutukset on jätetty pois, mikä on mahdollinen syy siihen, että käyrät eivät täysin vastaa histogrammin muotoa (lievää bimodaalisuutta).



Kuva 4: Efektien estimoidut keskihajonnat (vapausasteet suluissa)

Kuvassa 4 näytetään eri efektien suuruutta keskihajontojen posteriorijakaumina. Kuvassa 5 näytetään eri vokaaliympäristöjen vaikutukset keston (vasemmalla) ja soinnillisuuteen (oikealla). Kuvassa ovat eri vokaaliympäristöt järjestettyinä vaikutuksen mukaiseen järjestykseen. Näissä kuvissa pitempi ohut viiva edustaa posteriorijakauman 95 % CI, lyhyempi paksumpi viiva edustaa 50 % CI, ja keskellä oleva pallo edustaa posteriorin mediaania. Näyttää siltä, että vokaaliympäristöllä on selvää vaikutusta keston (keskihajonta on selvästi suurempi kuin 0, ks. kuva 4), vaikka se ei ole kovin suuri (selvästi pienempi kuin keston jäännessä- eli selittämätön hajonta). Myös kuvasta 5 (vasemmalla) näkyy, että ympäristöllä on vaikutusta frikatiivin keston.

Sen sijaan vokaaliympäristön mahdollinen vaikutus suoraan soinnillisuuteen on hyvin epävarma, kuten kuvassa 4 näkyy. Myös kuvasta 5 (oikealla) nähdään, että vaikutus soinnillisuuteen on hyvin epävarmaa. Kaikissa ympäristöissä luottamusvälit ovat melko pitkiä ja mediaanit osuvat hyvin lähelle nollaa. Mistään ympäristöstä ei siis tämän aineiston pohjalta voida varmasti sanoa, vaikuttaako se soinnillisuuden todennäköisyyteen, eikä edes kumpan suuntaan mahdollinen vaikutus olisi. Tietenkin vokaaliympäristö vaikuttaa soinnillisuuteen epäsuorasti, koska se vaikuttaa keston, ja, kuten aiemmin todettiin, frikatiivin kesto näyttäisi vaikuttavan selvästi soinnillisuuteen.



Kuva 5: Estimoidut efektit: ympäristön vaikutus kestoon (vasemmalla) ja soinnillisuuteen (oikealla)

4.4 Muita soinnillisuuteen vaikuttavia tekijöitä

Soinnillisuuteen voivat vaikuttaa myös muut tekijät, kuin äänneympäristö. Tämän tutkimuksen perusteella vaikuttaisi siltä, että frikatiivin kestolla on enemmän vaikutusta soinnillisuuteen, kuin ympäristöllä. Tosin tässä on tutkittu vain vokaalien välisiä ympäristöjä ja olisikin mielenkiintoista tietää, voiko frikatiivi olla soinnillinen myös konsonanttien vieressä. On mahdollista, että soinnillisen /s/-äänteen esiintyminen on jossain määrin puhujakohtaista, mutta siihen ei tämän tutkimuksen perusteella voida ottaa kantaa, koska aineistossa oli vain yksi puhuja. Muita mahdollisesti frikatiivin soinnillisuuteen vaikuttavia tekijöitä voivat olla muun muassa: puhujan vireystila, muut suprasegmentaaliset tekijät, alueellinen variaatio ja sosiolingvistiset tekijät. Näiden tekijöiden mahdollisen vaikutuksen selvittämiseksi tarvitaan jatkotutkimusta.

5 Lopuksi

Koska soinnillisuus tai soinnittomuus ei suomen kielessä erottele merkityksiä, niin olisikin luontevaa ajatella, että /s/ esiintyy vain soinnittomana. Soinnillisen frikatiivin tuottamiseen tarvitaan enemmän energiaa, koska tarvitaan riittävän voimakas ilmavirta tuottamaan sekä sointia kurkunpäässä, että frikaatiota sulkeuman kohdalla.

Tässä tutkimuksessa pyrittiin selvittämään, miten ympäristö vaikuttaa /s/-äänteen soinnillisuuteen suomen kielessä ja miten yleistä soinnillisuus on. Täysin soinnillisia /s/-äänteitä oli aineistossa 57 kappaletta eli lähes 16 %. On siis selvästi mahdollista, että /s/-äänne voi esiintyä myös täysin soinnillisena. Aineiston perusteella voidaan sanoa, että soinnillisuus on yleisempää lyhemmissä frikatiiveissa. Tämä voisi osin selittyä passiivisella ääntämyksellä, koska lyhyessä frikatiivissa sointi ei vokaalien välillä välttämättä ehdi loppua, kun taas pidemmässä frikatiivissa siihen on hyvin aikaa. Toisaalta täysin soinnillisia tapauksia oli myös pidempien frikatiivien joukossa, joten passiivinen ääntäminen ei selitä kaikkea soinnillisuutta.

Avoimeksi jäi vielä kysymys siitä, mitkä muut tekijät ympäristön lisäksi voivat vaikuttaa alveolaarisen frikatiivin soinnillisuuteen. Tämän tutkimuksen perusteella voidaan sanoa, että alveolaarisesta frikatiivista voi esiintyä myös soinnillinen versio, mutta olisi mielenkiintoista selvittää mitkä muut tekijät siihen vaikuttavat.

Viitteet

- Boersma, P. & Weenink, D. (2016). Praat: doing phonetics by computer (version 6.0.12) [tietokoneohjelma]. Haettu 31.1.2016 osoitteesta <http://www.praat.org/>.
- Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
- Gradoville, M. S. (2011). Validity in measurements of fricative voicing: evidence from Argentine Spanish. Teoksessa S. M. Alvord (Toim.), *Selected proceedings of the 5th conference on laboratory approaches to romance phonology* (s. 59–74). Somerville: Cascadilla proceedings project.
- Hualde, J. I. & Prieto, P. (2014). Lenition of intervocalic alveolar fricatives in Catalan and Spanish. *Phonetica*, 71, 108–126.
- Stevens, K. N., Blumstein, S. E., Glicksman, L., Burton, M. & Kurowski, K. (1992). Acoustics and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America*, 91(5), 2979–3000.
- Suomi, K. (1980). *Voicing in English and Finnish stops: a typological comparison with an interlanguage study of the two languages in contact*. Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja. University of Turku.

Suomen konsonanttiyhtymässä olevan lateraalin akustinen variaatio ja murteiden epenteettinen vokaali

Michael O'Dell & Tommi Nieminen

Tampereen yliopisto, Itä-Suomen yliopisto

1 Tutkimuksen taustaa

Suomen murteissa esiintyy tietyissä konsonanttiyhtymissä laajalti epenteettistä vokaalia eli švaata. Kvaliteetiltaan švaa on nähty foneemirepertuaariin kuuluvien vokaalien kaltaiseksi, missä suhteessa se eroaa esim. englannin kuvauksessa käytettävästä švaan (engl. *schwa*) käsitteestä, joka kuvaa kvaliteetiltaan neutraalistunutta vokaalia. Suomen švaan keston on nähty niin murteittain kuin kussakin murteessa puhujittainkin vaihtelevan hyvin lyhyestä (SUT-transkription termein yli- tai vajaalyhyestä) täyslyhyeen tai jopa huomattavasti sitä pidempään. Kvaliteetiltaan švaa on tulkittu joko aina identtiseksi švaan sisältävää konsonanttiyhtymää edeltävän vokaalin kanssa (*silmä* → *silimä*) tai joissakin murteissa yhtymää edeltävän ja sitä seuraavan vokaalin välimuodoksi (*silemä*). Kvaliteetin arviointi on aina tehty kuulonvaraisesti eli siinä mielessä švaa on ”pakotettu” foneemirepertuaariin. Švaavokaalin fonologista tulkintaa vaikeuttaa, ettei se välttämättä osallistu murteessa muuten vallitseviin (morfo)fonologisiin prosesseihin kuten yleis- tai erikoisgeminatioon.

Švaasiintymät – tai paremmin švaatulkinnat – voi jakaa eri tasoihin sen mukaan, kuinka usein švaa tapauksessa esiintyy ja kuinka laajalti. Perustapauksena voi pitää ensi ja toisen tavun rajalla sijaitsevia *l:n* ja eripaikkaisen konsonantin yhtymiä *lp(p)*, *lk(k)*, *lv*, *lj*, *lh*, *lm*, joissa švaan yleisesti katsotaan toteutuvan kaikissa murteissa, joissa sitä ylipäänsä esiintyy. Kaikki muut tapaukset ovat astetta kiistanalaisempia, etenkin ne, joissa jompanakumpana yhtymän konsonanteista on *h*.

Artikulatorisen fonologian (AP) näkökulmasta resonanttien kuten *l:n* voi yleisesti olettaa säilyttävän ympäröivän vokaalikontekstin artikulatorisen position. Mikäli konsonantin laadussa keskeisen artikulaattorin (kielen kärjen ja lavan) asema ei kohtuuttomasti häiritse vokaalin laadussa keskeisen artikulaattorin (kielen selän) asemaa, jälkimmäisen säilyttäminen konsonantin yli on ääntämisen kannalta ”taloudellista”. Toisaalta myös *vilikas* ~ *vilekas* -tyyppinen vaihtelu on helppo selittää tältä kannalta, koska ”vokaaliartikulaattorin” eli kielen selän on ennen pitkää valmistauduttava toisen tavun vokaalin vaatimaan asemaan. Suomen *l* on ainakin vokaalien välissä usein lyhyt ja ballistinen, minkä olemme jo aiemmin (Nieminen & O'Dell, 2010) arvelleet voivan päteä myös konsonanttiyhtymän alussa, jolloin AP:n mallin mukaisesti variaatio *l:n* ja seuraavan konsonantin eleen suhteellisessa ajoituksessa synnyttäisi helposti švaan. Tässä tutkimme *lv*-yhtymän akustista variaatiota juoksevassa puheessa erityisesti ajallisten suhteiden kannalta ja vertaamme tuloksia myös tietoon puhujien alueellisesta taustasta.



Švaaton murre	Švaa-murre
<i>jalka</i>	<i>jalaka</i>
<i>velka</i>	<i>veleka</i>
<i>silmä</i>	<i>silimä</i> (~ <i>silemä</i>)
<i>polvi</i>	<i>polovi</i>
<i>tulva</i>	<i>tuluva</i>
<i>kylmä</i>	<i>kylymä</i> (~ <i>kylömä</i>)
<i>nälkä</i>	<i>näläkä</i>
<i>hölmö</i>	<i>hölömö</i>

Kuva 1: Vasemmalla: *jal(a)ka*-tyyppisen švaan levikki karkeasti Kettusen mukaan (Kettunen, 1940, kartta 199); oikealla: esimerkkisanoja.

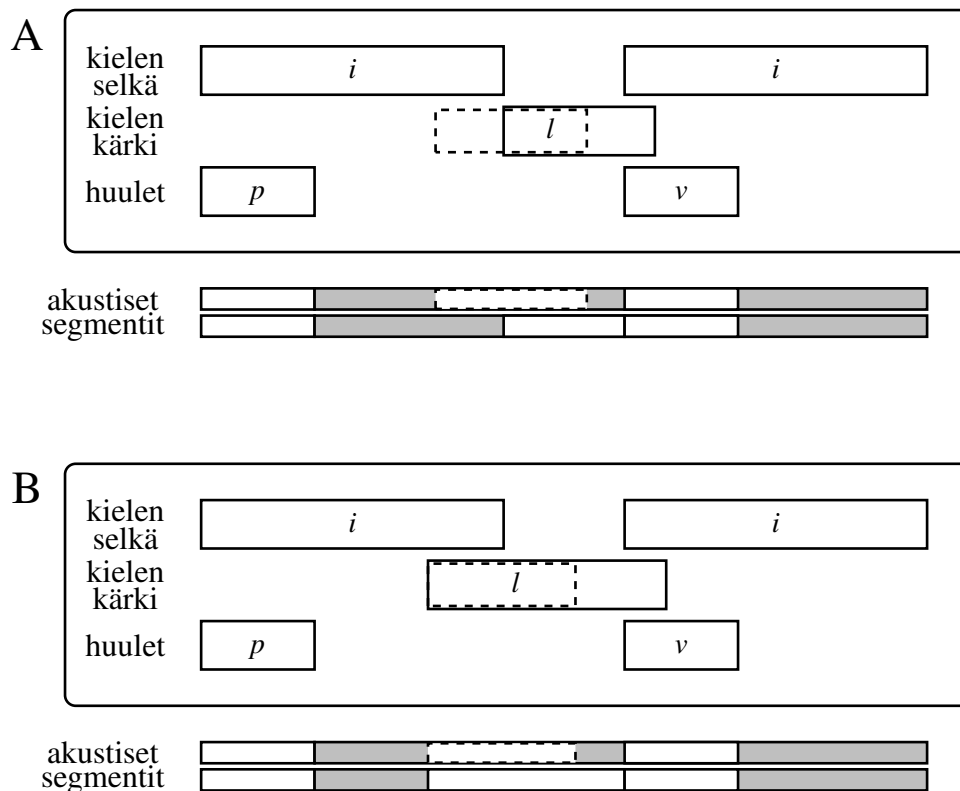
1.1 Švaan perinteiset kuvaukset ja foneettinen tutkimus

Vanhastaan švaailmiötä on pidetty suomen murteita erottavana piirteenä (ks. kuvio 1). Useimmissa perinteisissä kuvauksissa švaata esiintyy ainoastaan soinnillisen konsonantin aloittamassa heterorogaanisessa yhtymässä, etenkin *l*:n jäljessä. Švaata ei yleisesti ole katsottu esiintyvän homorgaanisissa yhtymissä eikä ainoassakaan klusiili–likvida-yhtymässä: *valta* ~ **valata*, *katras* ~ **kataras*, *täplä* ~ **täpälä* jne. Monet näistä yhtymistä ovat kielessä nuoria, minkä merkitystä on kuitenkin vaikea arvioida tietämättä, kuinka voimassa olevaksi ja leviäväksi (tai harveneväksi) prosessiksi tai äänne­muutokseksi švaa nykymurteissa katsotaan.

Toisaalta on kiinnostavaa, että joissakin yksittäisissä tutkimuksissa švaata on voitu havaita varsin yllättävissäkin yhtymissä. Esimerkiksi Laurosela (1922) on löytänyt aineistostaan švaatapauksia myös yhtymistä *kl*, *pr*, *kn*, *tr*, *fr*. Švaan kvaliteetiksi hän merkitsee neutraalivokaalin ə ([ə]), esim. *kikəla*, *kupəru*, *pukəna*, *peltəru*, *sifəra*. Lauroselan švaakontekstit ovat muutenkin laajemmat kuin yleensä. Hänellä on esim. tapauksia myös toisen ja kolmannen sekä kolmannen ja neljännen tavun rajalta (*katteləkä*, *kihelēmörä*, *tippurehərat*).

Sovijärvi (1937) mittasi epenteettisen vokaalin kestoja useista murteista: lyhimmillään se oli n. 20 % lyhyen ensitavun vokaalin kestosta, pisimmillään jopa 260 %. Sovijärvi myös kuvaa švaavokaalin foneettisen laadun mainiten sen säilyttävän tarkasti konsonanttiyhtymää edeltävän vokaalin kvaliteetin.

Vokaaliepenteesin syyksi – jos sitä kuvauksissa lainkaan mainitaan – esitetään yleensä ”ään­tä­misen helpottaminen”. Selitys ontuu etenkin siksi, että švaa esiintyy pikemmin artikulaatioltaan kohtuullisen helppoissa yhtymissä eikä esimerkiksi sellaisissa kuin *tj* (*patja* ~ **pataja*), *ts* (*patsas* ~ **patasas*) tai *rs* (*tursas* ~ **turusas*); kuitenkin esim. viimeksi mainitussa yhtymässä niin *r* kuin *s*:kin ovat artikulaatioltaan haastavimmat ja tulevat äidinkielisille puhujille yleensä kielenkehityksessä ään­teistä viimeisinä. Itkosen (1972, s. 29) mukaan epenteesi etenee niin, että ensi vaiheessa švaa on hyvin lyhyt (SUT ylilyhyt) vokaali, seuraavassa vaiheessa lyhyen vokaalin mittainen muodostamatta ”painotavua” ja lopulta täysi (lyhyt) vokaali, jolloin sanaan on syntynyt yksi lisätavu.



Kuva 2: Teoreettinen variaatio *l*-eleen vaiheessa (A, ylhäällä; aikaisempi ele merkitty katkoviivalla) ja *l*-eleen kestossa (B, alhaalla; lyhyempi ele merkitty katkoviivalla)

1.2 Švaa artikulatorisen fonologian näkökulmasta

Artikulatorisessa fonologiassa fonologiset yksiköt eivät ole segmenttejä vaan synkronoituja artikulatorisia ”eleitä” (”elepartituuri”, engl. *gesture*, *gestural score*). Nämä eleet eivät esiinny tiukasti peräkkäin, vaan limittyminen on normaalia. Näin ollen akustinen variaatio ei välttämättä ole seurausta fonologisen yksikön (eleen) lisäyksestä (tai poistosta). Hyvin tavallisia ovat myös jatkumoluonteiset ajoitukselliset erot eleiden koordinoinnissa (limittymisessä), jotka eleiden dynaamisen luonteen takia voivat synnyttää uusia akustisia segmenttejä.

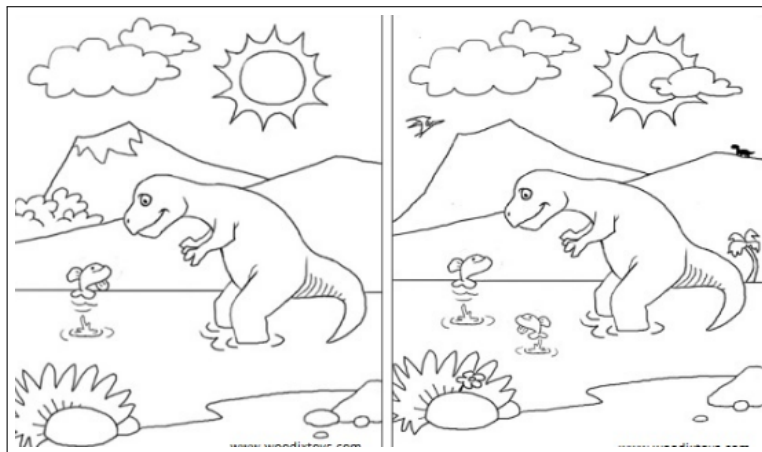
”Thus it seems that many allophonic differences are just quantitative differences that are large enough that phoneticians/phonologists have been able to notice them, and to relate them to distinctive differences found in other languages.” (Browman & Goldstein, 1992)

Švaan kaltaisissa tilanteissa muissa kielissä AP:ssa on tulkittu, että akustinen švaavokaali ei aina edusta omaa vokaalielettä, vaan voi olla seurausta siitä, että vierekkäiset konsonanttieleet ovat vähemmän limittäin (Browman & Goldstein, 1990, 1992; Steriade, 1990; Davidson & Stone, 2003; Hall, 2006). Tilannetta havainnollistetaan kuvassa 2 pelkistettyjen partituurien avulla. Ylemmässä kuviossa (A) *l*-eleen aikaistuminen (merkitty katkoviivalla) johtaa siihen, että akustinen *l* loppuu ennen kuin akustinen *v* alkaa, jolloin välissä kuullaan ”ylimääräinen” vokaali. Alemmassa kuviossa (B) sama ilmiö aiheutuu *l*-eleen lyhenemisestä.

2 Aineisto

Mittauksia varten mahdollisia švaatapauksia poimittiin Prosovar-aineistosta. Prosovar eli *Suomen kielen alueellinen ja sosiaalinen variaatio* oli vuosina 2013–15 toiminut, Koneen säätiön rahoittama tutkimushanke, jolla pyrittiin keräämään prosodian tutkimuksen tarpeisiin sopiva ja alueellisesti edustava puhekorpus uusien menetelmin verkossa. Hankkeen perustivat 2013 hankkeen vastuullisena johtajana toiminut Tommi Kurki Turun yliopistosta ja Tommi Nieminen Itä-Suomen yliopistosta. Myöhemmin tutkimusryhmään liittyivät projektitutkija Hamid Behravan Itä-Suomen yliopistosta (hankekaudella Turun yliopistosta) sekä apurahatutkijana toiminut Heini Kallio Helsingin yliopistosta.

Hankkeen toimintaperiaatteena oli joukkoistaa puheaineiston keruuta niin, että aineistoa tuottaisivat maallikkokäyttäjät suhteellisen omatoimisesti. Keruusivusto <https://puhu.utu.fi/> avattiin yleisölle 1. 5. 2014, ja keruu jatkuu yhä. Samalla kun keruusovellus toimii keruun välineenä, se on auttanut myös kokeilemaan ja testaamaan eri tehtävämuotojen toimivuutta verkkokeruussa. Käyttäjät äänittävät kielennäytteensä kotitietokoneellaan tai muulta vastaavalla verkkoon kytketyllä laitteella. Keruussa käytetään elisitoititehtäviä, joita käyttäjä saa aina sivustolle kirjautuessaan uuden sarjan. Joissakin tehtävissä informantin tulee lukea ruutuun tuleva teksti, toisissa taas ruutuun tuleva kehote vain opastaa, mitä informantin tulee omin sanoin tehdä (ks. esim. kuvan 3). Jälkimmäisen muotoisissa tehtävissä kehoitteen ja informantin tuottamaksi toivotun puhunnoksen etäisyyttä pidentämällä on pyritty simuloimaan tilannetta, jossa informantti pystyisi käyttämään omaa kieltään. Verkon yli äänittämiseen liittyy ilmeisiä vaikeuksia, jotka voivat heikentää aineiston teknistä laatua, mitä on tasapainottanut se, että kohteina olevat prosodiset piirteet ovat varsin robusteja.



Kuva 3: Tehtävä, jossa puhuja vertailee kuvia ja kertoo omin sanoin, miten ne eroavat. Oikeammaisessa kuvassa on esim. ylimääräinen pilvi, joka tuotti tämän tutkimuksen aineiston.

2.1 Mittaukset

Tässä tutkimuksessa mitattavat sanat tulivat kuvan 3 esittämästä vertailutehtävästä. Jos informantti teki tehtävän oikein, hän joutui ennemmin tai myöhemmin käyttämään sanaa *pilvi*. Mitattavaksi

poimittiin henkilön ensimmäinen nominatiivimuotoinen *pilvi*-sana, jos sellainen löytyi. Nominatiivin puuttuessa poimittiin ensimmäinen muu muoto, esim. *pilvee*. Koska signaalin laatu oli vaihteleva esimerkiksi lateraaliin liittyvät spektraaliset mitat osoittautuivat mahdottomiksi. Niiden sijaan tyydyttiin pelkästään kestromittauksiin. Praatin avulla pyrittiin akustisesta signaalista löytämään seuraavat ajankohdat: *p*:n eksploosio, *i*:n ja *l*:n välinen raja, *l*:n akustinen loppu (ts. *l*:n ja švaan välinen raja švaallisessa tapauksessa, *l*:n ja *v*:n välinen raja muuten), *v*:n alku mikäli eri kun *l*:n loppu (švaan ja *v*:n välinen raja), sekä *v*:n loppu (*v*:n ja seuraavan vokaalin välinen raja).

Sopiiko švaattomien ja švaallisten kestojakaumat yhteen? Näin odotettaisiin, jos ero ei johdu ylimääräisestä tai puuttuvasta segmentistä, vaan on seurausta normaalista ajallisesta variaatiosta.

2.2 Kestojen tilastollinen mallinta

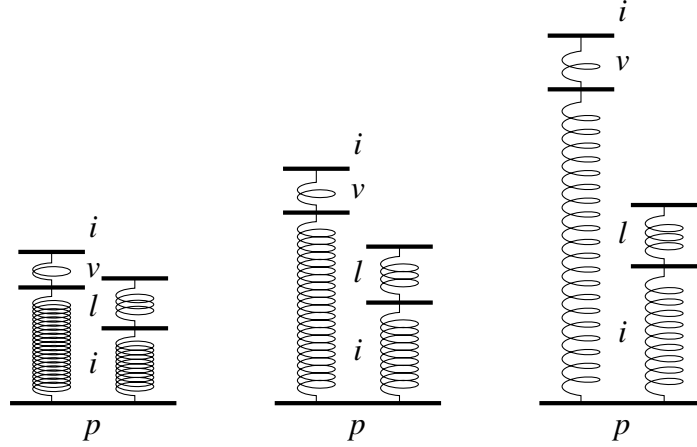
Partituurimallin mukaisesti švaaton tapaus tulkitaan sellaiseksi, että *v* alkaa ennen kuin *l* loppuu. Silloin *l* ja *v* ovat siis artikulaatiossa hieman päällekkäin, joten niiden välissä ei kuulu mitään vokaalin tapaista ääntä. Tämä tarkoittaa, että *l*:n jälkimmäisen rajan ajallinen sijainti on mitattavissa ainoastaan švaallisissa tapauksissa; švaattomissa tapauksissa se on ”piilossa” *v*:n takia, jolloin tiedetään vain, että se sijaitsee jossain *v*:n rajojen sisällä.

Sanan *pilvi* ääntämyksessä on variaatiota ensimmäisen *i*-vokaalin kestossa, jota tässä tulkitaan merkiksi siitä, että *l*-ele on alkanut aiemmin tai myöhemmin, ts. variaatioksi *l*-eleen vaiheessa. Tällainen variaatio voisi myös aiheuttaa eroja švaa-ilmiossö: *ceteris paribus* mitä aikaisemmin *l* alkaisi, sitä varmemmin švaan kaltainen vokaaliaines kuuluisi *l*:n ja *v*:n välissä (ks. kuva 2 A). Toisaalta sanan toteutuksessa on myös variaatiota *l*-eleen kestossa, ja *ceteris paribus* lyhyempi (ballistisempi) *l*-ele johtaisi varmemmin švaa-ääntämykseen (ks. kuva 2 B).

Vaikka selvää variaatiota esiintyy aineistossa ensimmäisen *i*:n ja *l*:n kestossa (kuten kaikkien jaksojen kestossa), tätä variaatiota ei pysty kuitenkaan suoraan tulkitsemaan, koska puhenopeus (tempo) oletettavasti vaikuttaa myös: kuinka paljon jokin havaittu kestoero on seurausta (vain) tempon vaihtelusta? Tempon vaikutus aineistossa on selvästi erittäin suuri (pisin *pilvi* on noin 3 kertaa pitempi kuin lyhin *pilvi*!), joten sitä ei voi vain jättää huomioimatta.

Tempon normalisointi esimerkiksi suhteuttamalla kaikki kestot sanan kokonaiskestoön ei ole tyydyttävä ratkaisu mm. sen takia, että tempon hidastuessa oletettavasti toiset segmentit ”venyvät” enemmän ja toiset vähemmän. Tästä syystä tilastollisessa analyysissä tempon vaikutus otettiin malliin mukaan yhtenä muuttujana soveltamalla ns. *elastisuus*-mallintaa (engl. *elasticity* vrt. W. N. Campbell (1992), N. Campbell (2000), Chung (1997)). Termiä *elastisuus* (suomeksi myös *jousito*) käytetään monella alalla (mm. mekaniikka, taloustiede, jne) kuvaamaan eri suureiden (tässä tapauksessa kestojen) kasvuvauhtien keskinäisiä suhteita: kuinka paljon toinen suure kasvaa jos toinen suure esimerkiksi tuplaantuu? Matemaattisesti asian voi esittää seuraavasti: tietyn jakson kesto $\delta_i = \exp(\mu_i + k s_i)$, jossa jakson ominaiskesto on $\exp(\mu_i)$, jakson jousto on s_i , ja k vastaa tilanteen (tässä tempon) vaikutusta.

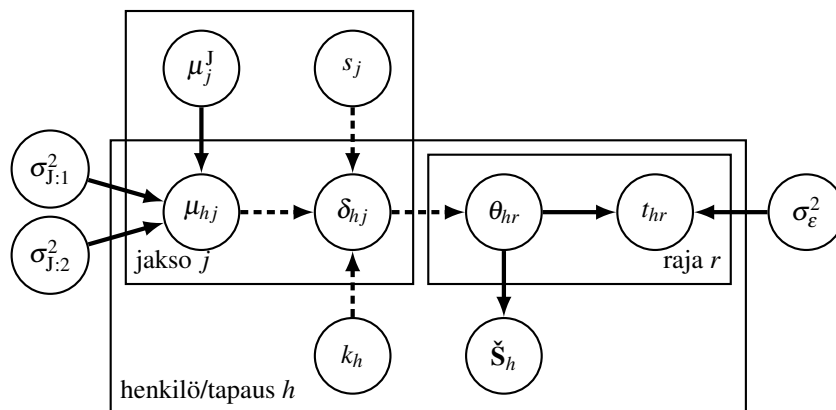
Kuvassa 4 näytetään miten jousto voi mahdollisesti vaikuttaa *pilvi*-sanankestoihin. Vasemmalta oikealle tempo hidastuu ja kaikki kestot kasvavat, mutta eri vauhtia jouhtuen eri jaksojen eri



Kuva 4: Elastisuusmallintaa: *pilvi*-esimerkki

joustoista. Tässä esimerkissä kuvitellaan, että l -ele ei ole kovin joustava, mutta vokaaliele on hyvin joustava. Kuten näkyy, näissä olosuhteissa ääntämys on švaaton (l ja v hieman päällekkäin) nopeassa tempossa (vasemmalla), mutta švaallinen kun tempo hidastuu (keskellä ja oikealla). Teoriassa švaallisuus vs. švaattomuus voisi siis johtua paitsi variaatiosta l -eleen vaiheessa ja kestossa (ks. kuva 2), myös tempon vaihtelusta.

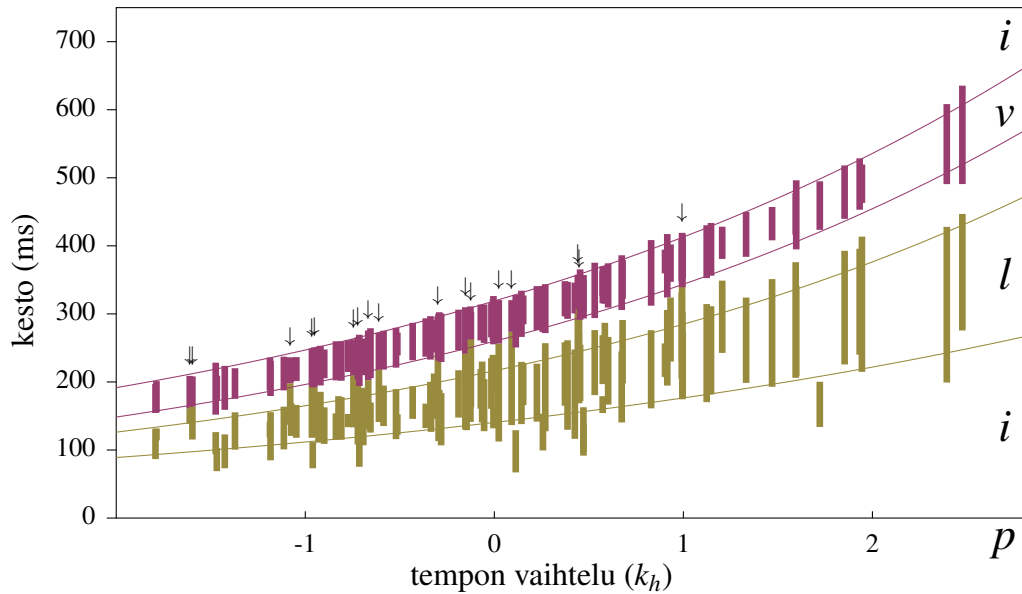
Bayesilaisessa analyysissä käytetty malli esitetään kuvassa 5. Kuvassa yhtenäinen nuoli edustaa stokastista yhteyttä, katkoviivanuoli loogista (laskennallista) yhteyttä. Havaittuina muuttujina ovat t_{hr} , rajan r mitattu ajallinen sijainti tapauksessa h (tai henkilöllä h , koska jokaiselta henkilöltä kerättiin vain yksi tapaus) ja \check{S}_h , joka kertoo havaittiinko mittauksessa švaa vai ei. Rajan mitattu aika on normaalijakutunut teoreettisen rajan θ_{hr} ympärillä: $t_{hr} \sim \text{Norm}(\theta_{hr}, \sigma_\epsilon^2)$. Mittauksissa havaittavan švaan \check{S}_h todennäköisyyttä mallinnetaan logistisella regressiolla funktiona l :n lopun (raja 2) ja v :n alun (raja 3) erotuksesta ($\theta_{h,3} - \theta_{h,2}$): mitä suurempi erotus sitä todennäköisemmin mittauksessa havaitaan švaa. Itse rajojen teoreettiset ajat θ_{hr} määräytyvät vastaavien teoreettisten kestojen δ_{hj} mukaan, esim. v :n loppu on v :n alku ynnä v :n kesto: $\theta_{h,4} = \theta_{h,3} + \delta_{h,4}$. Elastisuusmallinnassa jakson kesto taas riippuu sen ominaiskestosta μ_{hj} , joustosta s_j , ja tapauksen tempostä k_h :



Kuva 5: Tilastollinen malli (symbolien selitykset tekstissä)

$\delta_{hj} = \exp(\mu_{hj} + k_h s_j)$. Tempoparametrin keskihajonta mallissa on yksi ($k_h \sim \text{Norm}(0, 1)$), joten jousto s_j kertoo suoraan kuinka paljon jakson j kestovariaatiosta on peräisin tempon vaihtelusta ($k_h s_j \sim \text{Norm}(0, s_j^2)$).

Tärkeä piirre mallissa on se, että i -jakson ominaiskesto $\mu_{h,1}$, joka vastaa l -eleen vaihetta, ja l :n ominaiskesto $\mu_{h,2}$, joka vastaa l -eleen kestoa, saavat myös mallissa vaihdella tapauksesta (eli henkilöstä) toiseen: $\mu_{h,1} \sim \text{Norm}(\mu_1^J, \sigma_{J,1}^2)$, $\mu_{h,2} \sim \text{Norm}(\mu_2^J, \sigma_{J,2}^2)$. Näiden parametrien arvioitu variaatio (posteriorijakaumassa) kertoo jotakin švaavariaation mekanismista.



Kuva 6: Temposovitus: mitatut kestot tempon poikkeaman mukaan (posteriori mediaani). Švaattomat tapaukset osoitettu nuolella.

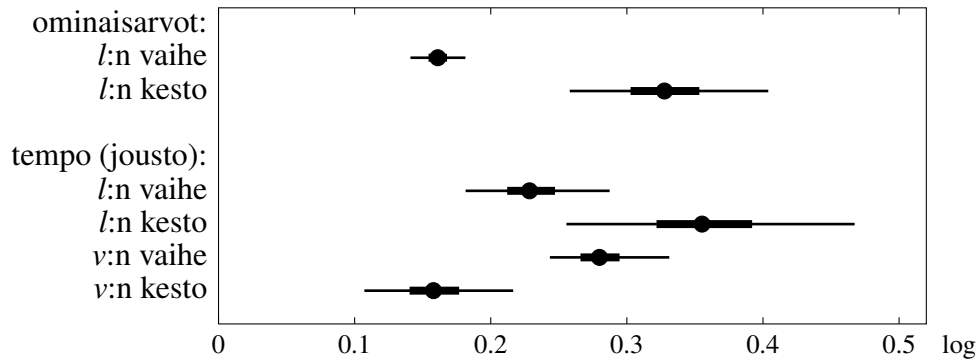
3 Tuloksia

3.1 Tempon vaikutus

Kuvassa 6 näytetään mitatut i :n, l :n, švaan ja v :n peräkkäiset kestot (pystyakselilla) järjestettyinä vaaka-akselilla estimoidun tempon poikkeaman mukaan (posteriori mediaani). Kuvassa näkyy ohuina nousevina käyrinä myös estimoidut tempon mukaiset kestot ilman variaatiota l -eleen vaiheessa ja kestossa (posteriori mediaani) tempon funktiona. Kuvan perusteella tempovaihtelulla näyttää olevan selvää vaikutusta yleisellä tasolla. Tempovaihtelu ei kuitenkaan selitä milloin švaa esiintyy ja milloin ei: švaattomat tapaukset (merkitty kuvassa nuolilla) esiintyvät laajalla tempoaueella eikä pelkästään esim. nopealla tempolla.

Eri efektejä voi arvioida ja vertailla myös katsomalla niiden keskihajontia. Kuva 7 esittää keskeisimpien efektien keskihajontien posteriorijakaumia (pitempi ohut viiva edustaa 95 % CI, lyhyempi paksumpi viiva 50 % CI ja piste edustaa posteriorijakauman mediaania). Tässä näkee vielä selkeämmin, että tempovaihtelut vaikuttivat sekä l -eleen että v -eleen vaiheeseen ja keston (tempon

vaikutuksen keskihajonta on mallissa sama kuin ns. joustoparametri, s_j). Nähtävästi tempovaihtelut vaikuttivat eniten l -eleen kestoon (95 % CI=(0,26–0,47)) ja todennäköisesti hiukan vähemmän l -eleen vaiheeseen (95 % CI=(0,26–0,47)).



Kuva 7: Efektien estimoidut (geometriset) keskihajonnat

3.2 Lateraاليةeen vaihe ja kesto

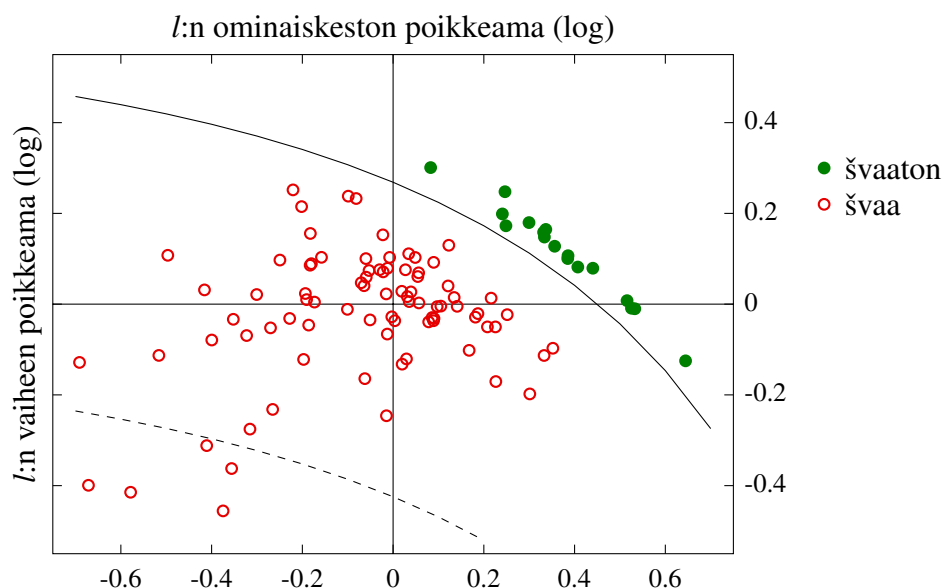
Kuten näkyy kuvassa 6 varsinkin l :n vaiheessa (i :n kestossa) ja l :n kestossa on suurehkoa variaatiota temposta riippumatta. Tämä näkyy kuvassa l -eleen vaiheen osalta, kun vertaa l :n kestoa kuvaavan janan alkua alimpaan ohueen käyrään, joka kuvaa l :n alkua pelkästään tempon funktiona. Myös l :n kestossa on selvää variaatiota, mikä näkyy kuvassa, kun vertaa l :n kestoa kuvaavan janan loppua seuraavaan ohueen käyrään, joka kuvaa l :n loppua pelkästään tempon funktiona.

Koska tilastollinen malli sallii vaihtelua l -eleen vaiheen ja keston ominaisarvoissa ($\mu_{h,1}$, $\mu_{h,2}$), voidaan myös tarkastella näiden muuttujien posteriorijakaumaa. Tempovaikutusten lisäksi kuva 7 esittää myös näiden efektien keskihajontien posteriorijakaumat. Selvää variaatiota näkyy l -eleen vaiheessa ($\mu_{h,1}$:n keskihajonnan 95 % CI=(0,14–0,18)) ja selvästi vielä suurempaa variaatiota l -eleen kestossa ($\mu_{h,2}$:n keskihajonnan 95 % CI=(0,26–0,40)).

Kuvassa 8 esitetään samanaikaisesti aineiston kaikista sanoista l -eleen molemmat arvioidut ominaisarvojen parametrit (vain posteriorijakauman mediaanit): sekä l :n vaihe (pysty-akseli) että l :n kesto (vaaka-akseli). Tässäkin näkyy, että on variaatio on suurempi kestossa kuin vaiheessa. Tapaukset, joista ei mitattu švaata (vihreät pisteet kuvassa 8), näyttävät myös sopivaan samaan jakaumaan kuin švaalliset. Voidaan ajatella, että švaattomuus on vain seurausta tavallisesta l :n vaiheen ja keston variaatiosta: silloin kun l loppuu vasta v :n alettua (vrt. yhtenäinen viiva kuvassa 8) akustista švaata ei synny.

3.3 Maantieteellinen jakauma

Kettusen kartaston mukaan (Kettunen, 1940) švaa esiintyy laajalla alueella Suomen keskellä, mutta ei etelässä tai pohjoisessa. Aineiston kaikista puhujista kerättiin myös tietoa asuinpaikkakunnasta, sekä lapsuuden että äänityshetken. Kuvassa 9 näytetään samat arvioidut parametrit kuin kuvassa 8 (l :n vaihe ja kesto), tällä kertaa värikoodattuina pisteinä, joita on sijoitettu kartalle puhujien

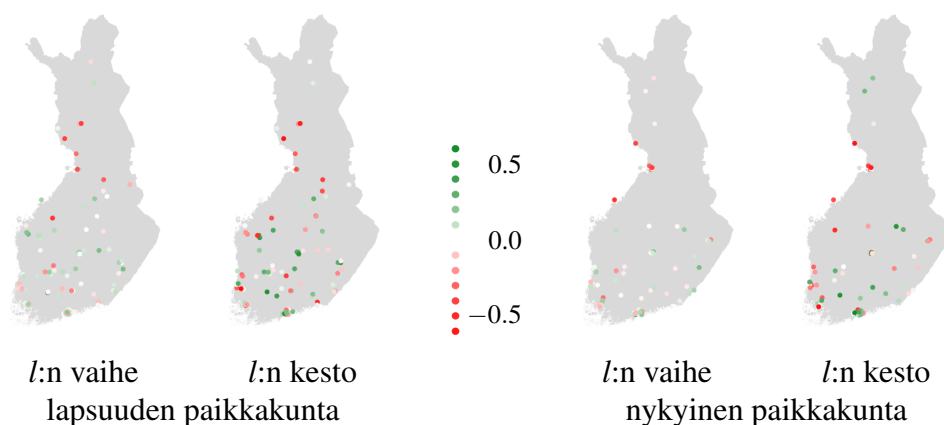


Kuva 8: Variaatio $l:n$ vaiheessa ja kestossa (posteriorijakauman mediaanit); yhtenäinen viiva: švaattomuuden teoreettinen raja ($l:n$ loppu = $v:n$ alku), katkoviiva: švaan kesto = $i:n$ kesto

paikkatietojen mukaan. Kartoissa on vaikeaa hahmottaa mitään selvää alueellista ryhmittymistä. Tämä tarkoittanee, että erot švaan ääntämisessä eivät olekaan (enää?) alueellisesti jakautuneet, ainakaan niin selvästi, että alueelliset rajat paljastuisi näin pienellä aineistolla. Tietenkin täytyy pitää mielessä, että aineisto on todella pieni, eikä siinä ole tietoa yksittäisten henkilöiden mahdollisesta variaatiosta.

4 Johtopäätöksiä

Yleisenä päätelmänä tutkimuksen tuloksista voi sanoa, että akustinen švaallisuus vaikuttaa olevan suomessa normaali (tai ainakin yleisin) tilanne. Vokaaliepenteesiä ei voi selittää yksin tempovaihtelulla. Selvää variaatiota on odotetusti $l:n$ alkamisvaiheessa mutta jopa enemmän sen suhteelli-



Kuva 9: Poikkeama $l:n$ vaiheessa ja kestossa puhujan lapsuuden ja nykyisen paikkakunnan mukaan. Positiiviset arvot (vihreät) vievät lähemmäs švaattomuutta.

sessä kestossa, mistä voidaan päätellä, että ilmeisesti *l* ei aina ole yhtä ”ballistinen” (itse asiassa muutama tapaus aineistossa oli pikemmin *pillivi* kuin *pilivi*). Kaikki tutkitut tapaukset – myös harvinaiset švaattomat ja voimakkaasti švaalliset – sopivat samaan jatkumoon.

Avoimia kysymyksiä on kuitenkin paljon: puhujien sisäinen variaatio, mahdollinen švaasegmentin laadun variaatio ([*pilivi*] ↔ [*piləvi*]), *l*:n ja *v*:n laadun variaatio, dynamiikan variaatio (esim. formanttisiirtymissä) jne. Samoin tämän akustisen epenteesin suhde havaittuun (tai ei-havaittuun) švaahan jää toistaiseksi auki – eli kysymys siitä, missä olosuhteissa ja mistä syistä švaata kussakin murteessa on transkriboitu. Myös suhde äänneympäristöön (mistä vokaalista tai kvantiteettihahmosta on kyse jne.) vaatisi jatkotutkimuksia.

Viitteet

- Browman, C. P. & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. Teoksessa J. Kingston & M. E. Beckman (Toim.), *Papers in laboratory phonology I: between the grammar and physics of speech* (s. 341–376). Cambridge University Press, Cambridge.
- Browman, C. P. & Goldstein, L. (1992). Articulatory Phonology: an overview. *Phonetica*, 49(3–4), 155–180.
- Campbell, N. (2000). Timing in speech: a multi-level process. Teoksessa M. Horne (Toim.), *Prosody: theory and experiment: studies presented to Gösta Bruce*. Kluwer Academic Publishers.
- Campbell, W. N. (1992). *Multi-level speech timing control* (tohtorinväitöskirja, University of Sussex).
- Chung, G. (1997). *Hierarchical duration modelling for a speech recognition system* (tutkielma, M.I.T.).
- Davidson, L. & Stone, M. (2003). Epenthesis versus gestural mistiming in consonant cluster production: an ultrasound study. Teoksessa M. Tsujimura & G. Garding (Toim.), *Wccfl 22 proceedings*. Somerville, MA: Cascadilla Press.
- Hall, N. (2006). Cross-linguistic patterns of vowel intrusion. *Phonology*, 23(3), 387–429.
- Itkonen, T. (1972). Kuoreveden ja Keuruun murretta: tekstejä ja sandhiseikkojen tarkastelua. *Suomi*, 117(1).
- Kettunen, L. (1940). *Suomen murteet III: murrekartasto*. Helsinki: SKS.
- Laurosela, J. (1922). *Foneettinen tutkimus Etelä-Pohjanmaan murteesta*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Nieminen, T. & O’Dell, M. L. (2010). *Epenthetic vowels in Finnish dialects: an Articulatory Phonology study*. (24th Scandinavian Conference of Linguistics, Joensuu).
- Sovijärvi, A. (1937). Foneettisia havaintoja Tornionjoen ja Pohjanmaan murteiden švaa-vokaaleista. *Suomi*, 5(19). (Deutsches Referat).
- Steriade, D. (1990). Gestures and autosegments: comments on Browman and Goldstein’s paper. Teoksessa J. Kingston & M. E. Beckman (Toim.), *Papers in laboratory phonology I: between the grammar and physics of speech* (s. 382–397). Cambridge University Press, Cambridge.

Automaattisesta kohdistamisesta apua katoavan murteen nimikointiin?

Lauri Tavi & Stefan Werner

Itä-Suomen yliopisto

Abstrakti

Koneen säätöön SILK-hankkeen (Suomen itäpuolisten lähialueiden kielikorpukset) yhteydessä tutkimme erilaisia *forced alignment* -järjestelmiä ja niiden mukautuvuutta katoaviin murteisiin. Tässä artikkelissa luomme katsauksen tutkituista kohdistinjärjestelmistä ja esitämme niistä käyttökokeuksiamme. SILK-työn pohjalta valitsimme tutkimukseemme automaattiseksi kohdistimiksi *AaltoASR*:n, puheanalyysiohjelma *Praatin* oman kohdistimen sekä *Prosodylab-Aligner*in. Käyttökokeuksiemme mukaan peruskäyttäjälle *Praatin* kohdistin on periaatteessa helpoin ja luontevin ratkaisu. Kyseisellä kohdistimella tulokset muuttuvat kuitenkin epätarkoiksi rajaamattomien pitkien jaksien kanssa, ja kohdistimen suurin ongelma on, ettei se ainakaan toistaiseksi tue suomen kielellä äänteiden kvantiteettierojen merkitsemistä. *AaltoASR* on erityisesti pitkien ja häiriöäännettömien tiedostojen kanssa nopea ja tarkka kohdistin, mutta sen tunnistamat merkit on rajattu suomen kirjakieleen. *Prosodylab-Aligner* edustaa puolestaan vähemmän käyttäjäystävällistä järjestelmää, joka ei välttämättä tarjoa oman akustisen mallin luomisen avullakaan parempaa tulosta. Alustavien tulosten mukaan automaattisten kohdistimien käytöllä on kuitenkin merkittävää hyötyä katoavien kielimuotojen nimikoinnissa, vaikka niiden suorituksissa esiintyy vielä toistuvaa epätarkkuutta.

1 Johdanto

Viimeisen kolmen vuosikymmen aikana uhanalaisten kielten dokumentoinnista on tullut tärkeä tehtävä kielitieteelliselle tutkimukselle, ja varsinkin viime vuosina siihen käytettyjä resursseja on kasvatettu maailmalla (DiCanio et al. 2013). Tutkimuksissa hyödynnettävät puhekorpuksat vaativat kuitenkin usein nauhoitteiden äännetason nimikointia, ja jo vain muutaman minuutin pituisen nauhoitteen manuaalinen nimikointi voi vaatia useamman tunnin työmäärän.

Manuaalisen nimikoinnin sijaan puhekorpuksia on rakennettu viime vuosina kasvavissa määrin hyödyntäen erilaisia automaattisia kohdistamisjärjestelmiä. Automaattisella kohdistamisella (*forced alignment*) tarkoitetaan äänitetystä puheesta kirjoitetun transkription ajoittamista niin, että tämän transkription merkit voidaan esittää niitä vastaavien nauhoitteen äänisegmenttien kanssa kohdakkain. Kyseessä on esimerkiksi puheentunnistamisen tai syntetisoinnin kannalta hyödyllinen toi-

menpide, sillä sen avulla saadaan kerättyä äänisignaalia sanoihin tai foneemeihin liittyvä tieto selvästi manuaalista segmentointia nopeammin. Ongelmana on kuitenkin ollut se, että automaattisen kohdistamisen tulokset voivat olla hyvinkin heikkoja varsinkin silloin, kun puhesignaali ei ole täysin häiriötön.

Erilaisia automaattisia kohdistimia kehitetään jatkuvasti. Kirjallisuudessa esiintyviä automaattisia kohdistimia ovat *SPPAS* (Bigi 2015), *MAUS*, *P2FA* (Yuan et al. 2008), *AaltoASR* (Kurimo 2016), *Train&Align* (Brognaux et al. 2012), *ALISA* (Stan et al. 2016), *EasyAlign* (Goldman 2011), *Praatalign* (Lubbers et al. 2016) ja *Prosodylab-Aligner* (Gorman et al. 2011). Näiden kohdistimien avulla on mahdollista rakentaa suhteellisen nopeasti suuriakin puhekorpuksia erityisesti maailman puhutuimmille kielille. Kuitenkin uhanalaisten kielten kohdalla sopivien automaattisten kohdistimien määrä vähenee: kohdistimien harjoittamiseen tarvittavien kielivarojen eli äänitetyn puheen määrä on yleensä hyvin rajallinen (van Niekrik & Barnard 2009) eivätkä kaikki kohdistimet tarjoa loppukäyttäjälle ollenkaan mahdollisuutta uuden kielen lisäämiseen.

Harva uhanalainen kieli sisältyy valmiina vaihtoehtona edellä mainittuihin kohdistimiin. Uhanalaisten kielten automaattiseen kohdistamiseen jäljelle jääkin usein kaksi vaihtoehtoa: jos puheaineistoa on tarpeeksi ja kohdistin tarjoaa mahdollisuuden uuden kielen lisäämiseen, tutkija voi luoda uuden kielimallin itse. Jos aineistoa ei ole tarpeeksi tai kohdistin ei tarjoa tutkijalle mahdollisuutta uuden kielen lisäämiseen, kohdistamiseen käytetään eri kielelle tarkoitettua mallia. Tällöin kohdistustarkkuus ei ole kovinkaan hyvin ennakoitavissa. Yleensä kielten (lähi)sukulaisuus ennakoi parempia tuloksia, mutta esimerkiksi kielten samankaltaisuuden konsonanttien palatalisoitumisessa on huomattu vaikuttavan tuloksiin kielisukulaisuutta enemmän (DiCanio et al. 2013).

Tässä artikkelissa kerromme alustavia käyttökokemuksiamme tutkimuksen kohteena olevista automaattisista kohdistimista, kun aineistona on katoava kielimuoto. Luvussa 2 esittelemme tutkimuksessa käytettyä aineistoa ja luvussa 3 tutkimukseen valitut kohdistimet. Luvussa 4 kokoamme yhteen kohdistimista saadut käyttökokemukset. Luvussa 5 esittelemme tulosten pohdinnan lisäksi tutkimuksemme jatkoa.

2 Aineisto

Itä-Suomen yliopistossa koostettava Raja-Karjalan korpus sisältää noin 120 tuntia vanhan Raja-Karjalan alueen haastattelutallenteita, niiden transkriptiot sekä karkeasti lausetasolle kohdistetut Text-Grid-tiedostot (ks. Itä-Suomen yliopisto: SILK). Korpus valmistuu alkuvuonna 2017 ja tulee saataville Kielipankkiin.

Tutkimme kohdistimien tarkkuutta ja käytettävyyttä Raja-Karjalan Ilomantsin murteella.

Osa kohdistimien akustisista malleista on harjoitettu suomen kielellä, joten valitsimme tutkimuksemme aineistoksi Ilomantsin rajakarjalaismurteen, sillä sen fonologia on rajakarjalaismurteista lähimpänä suomen fonologiaa (tarkemmin, ks. S. Tavi 2015). Tallenteet on nauhoitettu 60- ja 70-luvuilla ja ne ovat äänenlaadultaan kohtalaisia. Tutkimukseemme valitun nauhan puhujana on iäkäs naishenkilö ja nauhalla on myös muita puhujia sekä taustamelua.

3 Kohdistimet

Tutkimustamme varten valitsimme automaattisiksi kohdistimiksi *AaltoASR*:n, puheanalyysiohjelma *Praatin* (Boersma & Weenink 2016) oman kohdistimen sekä *Prosodylab-Aligner*n. Päädyimme tutkimaan kolmea edellä mainittua kohdistinta, sillä kaksi ensiksi mainittua mahdollistaa suoraan loppukäyttäjälle suomensukuisen kielimuodon kohdistamisen suomen kielen akustisen mallin avulla (*AaltoASR* ja *Praatin* kohdistin), ja viimeisin oman akustisen mallin luomisen uudelle kielelle (*Prosodylab-Aligner*). Toinen syy valintoihin oli edellä mainittujen kohdistimien käytettävyyys ja saatavuus.

3.1 AaltoASR

AaltoASR on Aalto-yliopistossa kehitetty suomen yleiskielelle suunniteltu puheentunnistin. Tutkimuksessamme *AaltoASR*:ää käytettiin suoraan ssh-yhteydellä CSC:n Taito-palvelimella. Koska kohdistinta käytetään Taito-palvelimella, edellytyksenä on, että loppukäyttäjällä tulee olla CSC-tunnus sekä vähintään alkeistason taidot UNIX-ympäristön käytössä. Seuraavaksi on lueteltu huomioita *AaltoASR*:n käytöstä rajakarjalaiseen murteeseen.

AaltoASR noudattaa suomen sanojen tavusääntöjä. Esimerkiksi litteroidun sanan *marjjoo* etuosa aiheuttaa ongelmia. Jos sana ei ole tavusääntöjen mukainen, kohdistin voi kohdistaa sen kirjaimittain: esim. *hmm* -> *hoo äm äm*. Sääntöjen vastaisen tavun ilmetessä kohdistin ilmoittaa seuraavaa:

*huono tavu! : *marj* in marj-joo/marjjoo*

Edellä esitetyt tavut eivät kuitenkaan kaada kohdistusta, toisin kuin erikoismerkit (esim. lainausmerkit). Murteelliset muodot (esim. *uamusil*) eivät sen sijaan aiheuttane suuria ongelmia kohdistamisessa. Huomion arvoista on, että kohdistin muuttaa transkriptioiden *z*-merkit merkkijonoksi *ts*. *AaltoASR*:n suurin rajoitus kuitenkin on, että se hyväksyy vain suomen kirjakielen mukaiset merkit *a, b, d, e, f, g, h, i, j, k, l, m, n, o, p, r, s, t, u, v, y, ä* ja *ö*. Kohdistimen rajallinen foneemilista aiheut-

taakin ongelmia esimerkiksi affrikaatta- sekä η -foneemien kanssa.

Kohdistin luo TextGrid-tiedostoon phone, word ja utterance -kerrokset, joista ensimmäisellä on eniten epätarkkuutta. Epätarkkuutta esiintyy muun muassa sanojen alun ja lopun foneemien kohdistuksessa: erityisesti sananalkuisissa konsonanteissa (esim. glottaali *h*) annotaatoraja asettuu ennen aikaiseksi. Palatalisoituneet äänteet, joita esiintyy ilomantsissa, eivät aiheuta ongelmia vaan luokituvat samoiksi mutta palatalisoitumattomiksi äänteiksi. *AaltoASR* ei myöskään muodosta pitkiä äänteitä eikä diftongeja (tai triftongeja), vaan merkitsee jokaisen transkriboidun merkin erillisenä.

Kohdistin on vaihtelevan herkkä erilaisille häiriöäänille, kuten kolinalle. Häiriöääni saattaa kin aiheuttaa kohdistukseen epätarkkuutta muutaman lauseen ajaksi, jonka jälkeen kohdistin pystyy kuitenkin normaaliin tulokseen. Varsinkin alle viiden minuutin kestoissa tiedostoissa *AaltoASR* ylittää tarkkoihin tuloksiin jo muutamassa minuutissa erityisesti sanatason kohdistuksessa.

3.2 Praatin kohdistin

Nykyisissä *Praat*-versioissa on *Espeakin*¹ grafeemi-foneemi -muunnokseen perustuva *Align interval* -toiminto. Kohdistin sisältää useita kielikohtaisia vaihtoehtoja, ja käytimme aineistomme kohdistamiseen suomen mallia. Kohdistuksen esivaatimuksena on, että TextGrid-tiedosto tulee olla valmiiksi rajattu vähintään lausetason intervaleihin. Mitä enemmän puhumatonta jaksoa lausetason intervalli sisältää, sitä enemmän sana- ja foneemitason kohdistukseen tulee virheitä. *Praatin* oman kohdistimen kokeilemisessa hyödynsimmekin *AaltoASR*:n lausetason segmentointeja, sillä ne osoittautuivat suhteellisen tarkkoiksi (ks. luku 3.1).

Praatin Align interval -toiminnon vahvuus on sen käytettävyys. Äänisignaali ja TextGrid-tiedoston merkkitasot ovat käyttäjän nähtävissä, ja CTRL- ja D-näppäimillä valittu lausetason intervalli jakautuu sana- ja foneemitason intervaleiksi alle sekunnissa. Kohdistin käyttää foneemitason merkitsemiseen IPA-aakkosia ja se tukee diftongien merkitsemistä. Sen sijaan kohdistimen suuri ongelma on, ettei se ainakaan toistaiseksi tue suomen kielellä vokaalien kvantiteettierojen merkitsemistä, eli sanojen pitkät vokaalit merkitään lyhyinä vokaaleina. Tällä hetkellä *Praatin* kohdistin vaatiikin äänteiden kvantiteettieroja sisältävien kielten nimikointiin manuaalisen viimeistelyn, mutta verrattuna kokonaan manuaaliseen nimikointiin, kyseinen kohdistimen vajaa tulos säästää merkittävästi aikaa.

¹ <http://espeak.sourceforge.net/>

3.3 Prosodylab-Aligner

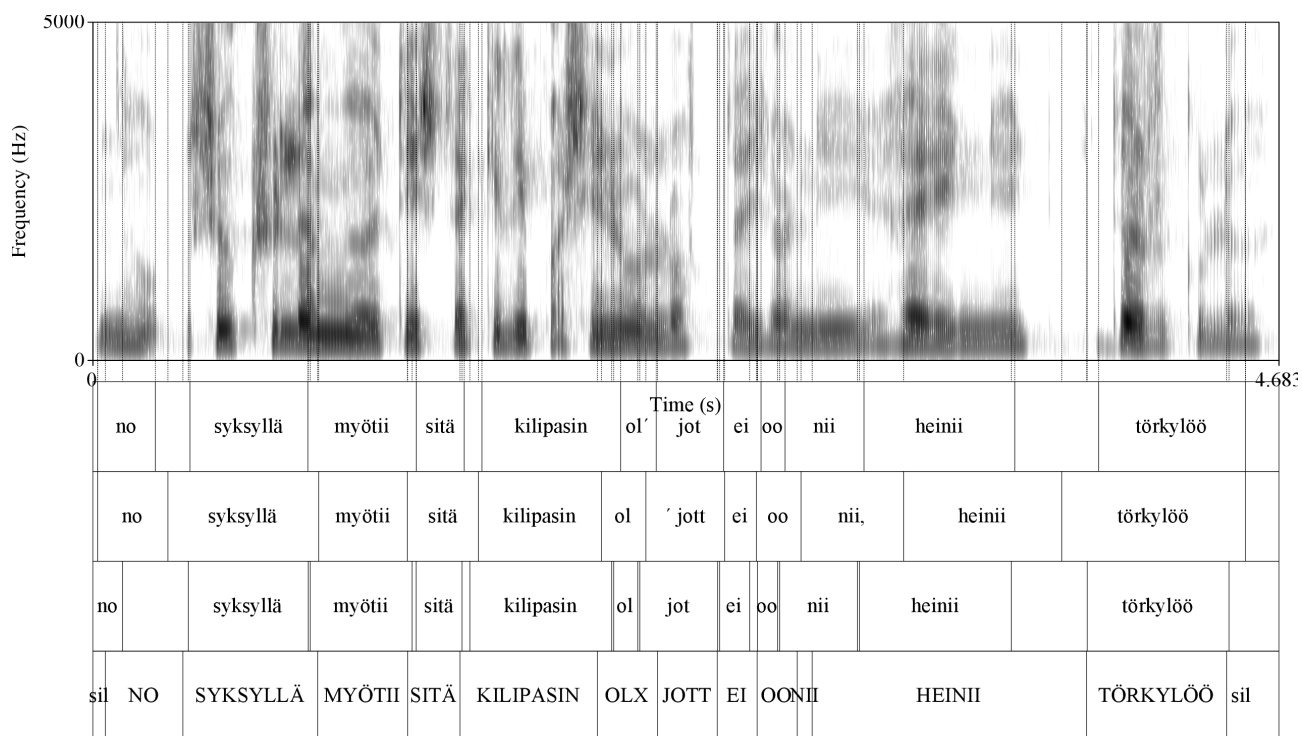
McGill-yliopistolla kehitetty *Prosodylab-Aligner* -järjestelmä koostuu Python- ja UNIX/Linux-shell -skripteistä, jotka automatisoivat Markovin piilomallien perustyökalun HTK:n² käyttöä. Muiden kuin englanninkielisten äänitysten kohdistamiseen tarvitaan joko oma sanakirja ja akustinen malli tai muutaman tunnin verran puhetta sanatason transkriptioineen (jälkimmäistä vaihtoehtoa emme ole vielä kokeilleet). Tarpeelliset valmistelut suomenkielisen puheen *Prosodylab-Aligner* -analyysia varten eivät sinänsä ole vaikeita, mutta edellyttävät shell-komentojen perustason tuntemusta kaikenlaisia muunnoksia ja tekstitiedostojen manipulointeja varten; toimivat ohjeet ja apuskriptit voi pyytää kirjoittajilta.

4 Tulokset

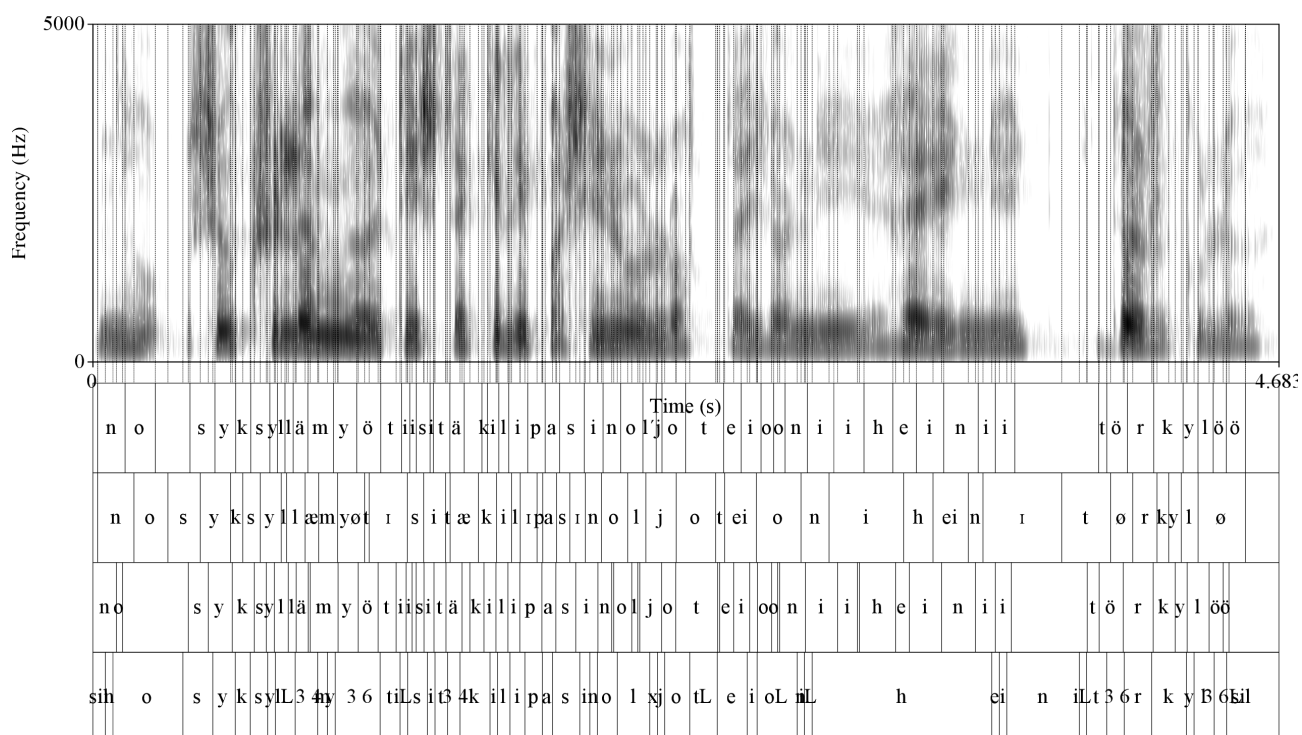
Tässä luvussa kerromme johtopäätöksiä luvuissa 3.1–3.3 esiteltyjen automaattisten kohdistimien tuloksista ja käytettävyyksistä. Ensimmäisenä esiteltyä *AaltoASR*:ää käytettiin CSC:n Taito-palvelimella, joka edellyttää ainakin toistaiseksi CSC-tunnusta. Kyseinen kohdistin on moduulin lataamisen jälkeen suoraan käytettävissä, eikä erilaisten konfiguraatio- ja skripti-tiedostojen ajaminen vienyt aikaa. Toisena esitelty *Praatin* oma kohdistin on suoraan käytettävissä *Praat*-ohjelman *align interval* -komennolla, mutta tarvitsee kohdistuksen aloittamiseksi valmiit lausetason intervallit. Viimeisenä kuvailimme *Prosodylab-Aligneria*, joka puolestaan tarvitsee toimiakseen useita erilaisia asennuksia, mutta mahdollistaa oman akustisen mallin luomisen.

Seuraavaksi esitellyt kuvat kertovat eri kohdistimien annotaatorajojen vaihteluista äänne- sekä sanatasoilla (Kuvat 1 ja 2). *AaltoASR*:n segmentoinnin tulos on suhteellisen lähellä manuaalista segmentointia. *Praatin* omalla kohdistimella sekä *Prosodylab-Alignerilla* on puolestaan vaikeuksia sanojen välissä olevien taukojen havaitsemisessa, jolloin jotkin sananrajaiset foneemit on merkitty kestoiltaan liian pitkiksi. Toisin kuin muut järjestelmät *Praatin* kohdistus ei toimi, jos signaali ei ole valmiiksi pilkottu yksittäisiin ilmauksiin ennen automaattista kohdistamista. Tämä nähdään Kuvassa 3, jossa vertaillaan rajatun ja rajaamattoman lausuman kohdistamista *Praatilla* (TextGridin alaosa; puhe jatkuu kuuden sekunnin jälkeen).

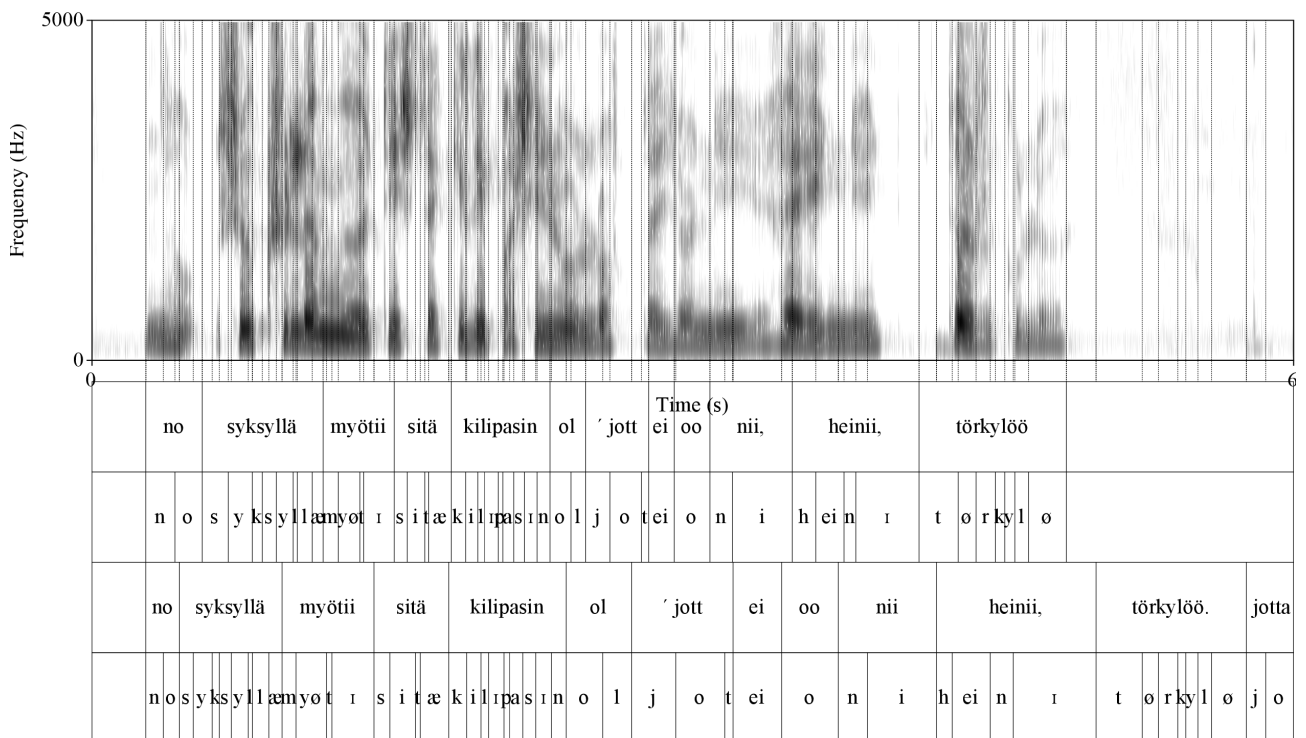
2 <http://htk.eng.cam.ac.uk/>



Kuva 1. Sanatason esimerkkiannotaatiot ylhäältä alas: manuaalinen, *Praat*, *AaltoASR*, *Prosodylab-Aligner*



Kuva 2. Foneemitason esimerkkiannotaatiot ylhäältä alas: manuaalinen, *Praat*, *AaltoASR* ja *ProsodyLab-Aligner*



Kuva 3. *Praatin* kohdistus valmiiksi pilkotussa signaalissa (ylhäällä) ja osana pitempää signaalia (alhaalla)

Tutkimuksemme mukaan peruskäyttäjälle *Praatin* kohdistin on helpoin ja luontevin ratkaisu, sillä se on käyttäjäystävällinen ja nopea. Toisaalta kyseisellä kohdistimella tulokset muuttuvat epä-tarkoiksi pitkissä jaksoissa eikä kohdistin tue vokaalien kvantiteettierojen merkitsemistä. *AaltoASR* on myös varsinkin pitkien ja häiriöäänettömien tiedostojen kanssa nopea ja tarkka kohdistin, mutta sen tunnistamat merkit on rajattu suomen kirjakieleen sekä sen käyttö edellyttää ainakin toistaiseksi CSC-tunnusta. *ProsodyLab-Aligner* edustaa puolestaan vähemmän käyttäjäystävällistä järjestelmää, joka ei välttämättä tarjoa oman akustisen mallin luomisen avullakaan parempaa tulosta. Alustavien tulosten mukaan automaattisten kohdistimien käytöllä on merkittävää hyötyä uusien kielimuotojen nimikoinnissa, mutta niiden suorituksissa esiintyy myös toistuvaa epätarkkuutta varsinkin foneemitasolla. Tutkimuksemme on kuitenkin vielä kesken, ja seuraava vaihe on tulosten tilastollinen vertailu (vrt. DiCanio et al. 2013).

5 Lopuksi

Tässä tutkimuksessa olemme esitelleet kokemuksiamme uhanalaisen kielimuodon nimikoinnista käyttäen erilaisia automaattisia kohdistimia. Rajasimme tutkimuksemme kolmeen automaattiseen kohdistimeen: *AaltoASR*:ään, *Praatin* omaan kohdistimeen sekä *Prosodylab-Aligneriin*. Alustavat käyttökokemukset osoittavat, että osa automaattisista kohdistimista tarjoaa merkittävän hyödyn

myös katoavan kielimuodon nimikoimisessa. Vaikka katoavien kielten kohdalla puheaineisto sisältää usein äänisignaalia häiritsevää taustamelua, eikä artikulointi ole haastattelutilanteissa aina selkeää (esim. kuiskaukset tai nauraen sanotut jaksot), kohdistimien tarkkuus varsinkin sanatasolla ylittää hyviin tuloksiin. Äännetason tarkkuus ei välttämättä ole riittävää esimerkiksi suoraan spektritiedon keräämiseen foneemeista, mutta kohdistimien tuottamien valmiiden intervallirajojen avulla myös manuaalinen nimikointi on huomattavasti nopeampaa. Automaattisia kohdistimia on syytä käyttää rakennettaessa puhekorpuksia katoavia kielistä, mutta nauhoiteaineisto saattaa vaatia editointia – esimerkiksi puhesignaalin rajaaminen vain puhuttuihin jaksoihin tai häiriöäänten editointi kokonaan pois saattaa olla tarpeellista hyvien tulosten saavuttamiseksi.

Tutkimuksemme tarkoituksena on vielä vertailla miten hyvin eri kohdistusjärjestelmät toimivat, kun useita prosodisia lausekkeita puhumattomine jaksoineen on samassa tiedostossa, ja toisaalta kun aineisto on valmiiksi paloitetu lauseisiin (tms. minimaalisiin jaksoihin). Jatkotutkimuksisamme tulemme vertailemaan kohdistusjärjestelmien tuloksia tilastollisesti.

Lähteet

- Boersma, P. & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Versio 6.0.17, haettu 21.04.2016 osoitteesta <http://www.praat.org/>
- Bigi, B. (2015-toappear). SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech the Phonetician - International Society of Phonetic Sciences, Volume 111-112, Pages 55-69.
- Brognaux, S., Roekhaut, S., Drugman, T., Beaufort, R. (2012) Train&Align: a new online tool for automatic phonetic alignment, Miami, Florida, *IEEE Workshop on Spoken Language Technology*,
- DiCanio, C., Nam, H., Whalen, D. H., Bunnell, H. T., Amith, J. D., García, R. C. (2013) Using automatic alignment to analyze endangered language data: testing the viability of untrained alignment. *J. Acoust Soc Am.* 134 (3): 2235-46.
- Freeware package 'Munich AUtomatic Segmentation' (MAUS). <ftp://ftp.bas.uni-muenchen.de/pub/BAS/SOFTW/MAUS>
- Goldman, J.-Ph. (2011). *EasyAlign: an automatic phonetic alignment tool under Praat* Proceedings of InterSpeech, September 2011, Firenze, Italy.
- Gorman, K., Howell, J. & Wagner, M. (2011): Prosodylab-Aligner: a tool for forced alignment of laboratory speech. *Canadian Acoustics*, 39, s. 192-193.
- Itä-Suomen yliopisto (2011–2014). SILK: Suomen itäpuolisten lähialueiden kielikorpukset. <http://www.uef.fi/silk>
- Kielipankki. <https://www.kielipankki.fi/>
- Kurimo, M. (2016). *AaltoASR: Aalto-yliopiston automaattinen puheentunnistin*. <https://github.com/aalto-speech/AaltoASR>

- Lubbers, M. & Torreira, F. (2013-2016). *Praatalign: an interactive Praat plug-in for performing phonetic forced alignment*, <https://github.com/dopefishh/praatalign>,
- van Niekerk, D. R. & Barnard, E. (2009). Phonetic alignment for speech synthesis in under-resourced languages. *INTERSPEECH 2009*, 10th Annual Conference of the International Speech Communication Association, Brighton, United Kingdom.
- Stan, A., Mamiya, Y., Yamagishi, J., Bell, P., Watts, O., Clark, R., King, S.(2016). "ALISA: An automatic lightly supervised speech segmentation and alignment tool", In *Computer Speech and Language*, vol. 35, pp. 116-133.
- Tavi, S. (2015). *Rajakarjalaismurteiden venäläiset lainasanat*. Pro gradu -tutkielma. Itä-Suomen yliopisto.
- Yuan, J. & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *Proceedings of Acoustics '08 (P2FA)*.

Puheenkierrätystutkimus suomessa

Joona Vakkilainen

Tampereen yliopisto

vakkilainen.joonas.j@student.uta.fi

Puheenkierrätys on puheen rytmin tutkimiseen kehitetty menetelmä. Vanhat menetelmät eivät ole tuottaneet näyttöä rytmitypologian puolesta, minkä vuoksi kielen rytmiä on pidetty epätodellisenä. Koskavapaa puhe on epäsäännöllistä ja rytmi puolestaan hierarkkista, spontaaniin puheeseen tai lukupuhuntaan perustuvat mittaukset eivät saa rytmin järjestymistä nousemaan esiin. Puheenkierrätys pakottaa puheen säännölliseen muottiin käyttäen ääniärsykeitä. Tällä tavalla mahdollistetaan rytmiin vaikuttavien elementtien tarkastelu. Tässä tutkimuksessa havainnoidaan suomen rytmiä. Koehenkilöt lukivat ääniärsyksen tahdissa kolmisanaa fraaseja, joiden keskimäistä sanaa varioitiin mora- ja tavuluvun puolesta. Tuloksissa morailuku vaikutti sanan kestoon, mutta fraasin kesto pysyi keskimäärin lähes vakiona. Tavuluku vaikutti sekä sanan että fraasin kestoon enemmän. Tavuluku vaikutti myös varioitavan sanan alkuajankohtaan enemmän. Saatujen tulosten perusteella sekä morat että tavut vaikuttavat suomen rytmiin.

1. Taustaa

Rytmitypologia jakaa kielet erilaisiin rytmityyppeihin sen mukaan, mitkä elementit niissä pyrkivät olemaan isokronisia, eli mitkä niistä esiintyvät säännöllisin välein. Perinteinen jako määrittelee kielet painoajoitteiseen ja tavuajoitteiseen tyyppiin. Painoajoitteisissa kielissä painojen intervallit pyrkivät olemaan isokronisia tavuluvusta riippumatta, ja tavuajoitteisissa kielissä tavujen kestot ovat säännöllisiä. Näiden lisäksi puhutaan myös mora-ajoitteisuudesta, jossa isokroninen yksikkö on tavun sijasta mora. (Tajima 1998: 3.) Vanhojen tyyppien rinnalle on otettu myös paino- ja tavuajoitteisuuden yhdistävä tyyppi sekä ajoittamaton tyyppi, jossa mikään yksikkö ei tavoittele isokroniaa (Mairano 2011: 29–30). Koska suomen rytmityypistä ei ole yksimielisyyttä (Dauer 1987: 447; Isei-Jaakkola 2004: 35), suomi ei luultavasti edusta mitään tyyppiä puhtaasti.

Perinteinen rytmitypologia perustuu kuitenkin hyvin pitkälti vaikutelmiin, eikä sen puolesta ole aiemmin saatu kokeellista näyttöä. Foneettiset mittaukset eivät ole tukeneet ajatusta isokroniasta, eivätkä havaitsemiskokeet ole tuottaneet yksimielisiä tuloksia kaikkien kielten rytmityypeistä (Dauer 1987). Rytmiä on jopa väitetty epätodelliseksi. Pamies Bertrán (1999) sai tavu- ja painoajoitteisina pidetyistä kielistä samankaltaisia mittaustuloksia, mikä hänen mukaansa kertoo

siitä, että kielitieteessä rytmien käsite on antiikin metriikasta periytynyt käsite, jolla ei ole tekemistä todellisen fonologian kanssa. Roachin (1982) mukaan kieli edustaa jotakin rytmityyppiä, jos se kuulostaa siltä. Dauer (1987) esittää, että rytmi perustuu erilaisiin osatekijöihin, jotka kertovat, miten kyseisen kielen rytmi suhteutuu painointervalleihin, ja näiden perusteella kieli asetetaan liukuvalle asteikolle painoajoitteinen vs. ei-painoajoitteinen.

Rytmitypologian kokeellisen tutkimuksen ongelmaksi voidaan katsoa menetelmien puute. Spontaani puhe sekä lukupuhunta ovat epäsäännöllisiä, ja siksi rytmien hierarkia ei pääse esiin niin, että sitä voitaisiin havainnoida. Rytmi perustuu hierarkiaan, jossa eri yksiköt asettuvat erilaisiin asemiin, mikä luo lopullisen rytmivaikutelman. Puheenkierrätys (*speech cycling*) pakottaa puheen säännölliseen muottiin, jolloin vapaaseen puheeseen liittyvät häiriötekijät eivät pääse vaikuttamaan rytmiin ja rytmien muodostavat temporaaliset tekijät nousevat tarkastelun mahdollistavaan asemaan. Puheenkierrätyksessä puhuja synkronoi puheensa ääniärsykkeisiin saamiensa ohjeiden mukaisesti. Ohjeilla voidaan muokata koeasetelmaa halutulla tavalla. Puheenkierrätystä on kritisoitu epäluonnollisuudesta ja keinotekoisuudesta. Kritiikkiin voidaan vastata, että ihmisellä on luonnollinen kyky tuottaa puhetta rytmisesti, mikä näkyy mm. runoudessa, mielenosoitusten iskulauseissa ja samanaikaisessa puhunnassa, ja että laboratorio-olosuhteet ovat aina jokseenkin keinotekoisia. Rytmi perustuu oskillaattoreiden säännöllisyyteen (esim. kävelyaskeleiden väli), ja puheessa nämä oskillaattorit vain ovat kognitiivisia. Puheenkierrätyksen tehtävänä on houkutella puheen kognitiiviset oskillaattorit esiin pakottamalla ne fyysiseen säännöllisyyteen. (Tajima 1998: 17; Cummins 2002: 123; Port, Tajima & Cummins 1998: 70.)

Puheenkierrätyksen tarkoitus on havainnoida rytmiä. Sillä ei pyritä todistamaan perinteisen rytmitypologian puolesta tai kumoamaan sitä. Puheenkierrätystutkimukset ovatkin tuottaneet perinteisiä rytmityyppejä sekä tukevaa että haastavaa näyttöä. Tajima (1998: 85–91) vertaili japania ja englantia varioimalla koefraasin tavujen määrää. Japanissa tavujen lisääminen aiheutti fraasin muiden painollisten tavujen siirtymistä enemmän kuin englannissa, mikä tukee tavu- ja mora-ajoitushypoteesia. Myös englannissa tavujen siirtymistä tapahtui hieman, mikä on vastoin perinteistä painoajoitushypoteesia. Tällä tavalla puheenkierrätys on onnistunut tuomaan rytmiä esiin konkreettisesti ja tarkastelemaan, kuinka todellisia rytmityypit ovat. Jo puheenkierrätyksen soveltuvuus tutkittavaan kieleen kertoo jotakin kielen rytmiominaisuuksista. Tämä näkyy Cumminsin (2002) tutkimuksessa, jossa puhujien tuli sijoittaa pää- ja sivupainot erilaisiin ääniärsykkeisiin. Englanninkieliset suoriutuivat tehtävästä helposti, mutta italian- ja

espanjakielisille tehtävä oli vaikea, mikä johtuu siitä, että italian ja espanjan painot eivät muodosta tahteja kuten englannissa.

2. Aineisto ja menetelmät

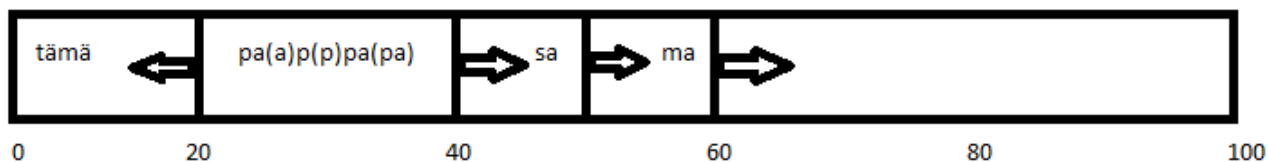
Omassa tutkimuksessani tutkin suomen rytmitystä. Tavoitteena on tarkastella, mitä rytmisyksiköitä puheenkierrätyksessä nousee esiin. Erityisenä mielenkiintona on moran ja tavu suhde: onko suomi rytmityypiltään enemmän tavu- vai mora-ajotteinen kieli. Koska suomessa on foneeminen kvantiteetti, ainakaan puhdas tavuajoitus ei ole oletettava tulos.

Aineisto perustuu yhdeksän suomenkielisen koehenkilön lukemiin koefraaseihin. Kokeessa käytettiin viittä erilaista fraasia, jotka koehenkilöt lukivat ääniärsykkeen tahdissa. Ääniärsykkeenä toimi 440 hertsin lyhyt piippaus, joka toistui 1,5 sekunnin välein. Koehenkilöitä ohjeistettiin kuuntelemaan ensin neljä piippausta rytmiin pääsemiseksi ja sitten toistamaan fraasi aloittaen sen aina piippauksen kohdalla. Jokaista fraasia kohden oli 11 toistoa. Koefraasi alkoi aina sanalla *tämä* ja päättyi sanaan *sama*. Keskellä olevaan koeseen varioitiin kvantiteettihahmoltaan ja tavuluvultaan. Viisi koeseen olivat *papa*, *paapa*, *pappa*, *paappa* ja *papapa*. Jos koehenkilö ei onnistunut tehtävässä heti, tehtävä uusittiin. Koeasetelma olikin koehenkilöille selvästi outo, ja erityisesti viides koeseinä osoittautui monelle ongelmalliseksi.

Analyyseissä mielenkiinnon kohteena ovat fraasin rajat eli fraasin loppuajankohta, fraasien välisen hiljaisuuden kesto, sekä fraasin osien kestot ja alkuajankohdat. Koska rytmisissä on kyse suhteellisista ominaisuuksista, fraasi asetetaan suhteelliselle asteikolle. Mittauksissa toistosyklin alku on sama kuin fraasin alku ja loppu sama kuin seuraavan syklin alku, joten koehenkilön tarkkuus ääniärsykkeen suhteen ei ole tutkimuksessa oleellista. Yhdestä tällaisesta fraasin ja fraasien välisen hiljaisuuden sisältävää jaksoa kutsutaan Tajiman (1998: 32–33) mukaisesti ulkoiseksi sykliksi (*external phase*) ja fraasia ilman hiljaisuutta sisäiseksi sykliksi (*internal phase*). Ulkoisessa syklissä tarkastellaan rytmisiä säännönmukaisuuksia, kuten fraasin kestoa, kun taas sisäinen sykli mahdollistaa fraasin sisäisten ajankohtien havainnoinnin.

Kuvassa 1 nähdään ulkoisen syklin malli. Puhtaassa painoajoituksessa siirtymistä nuolten osoittamiin suuntiin koeseinän kasvaessa ei tapahtuisi vaan koeseinä säilyttäisi kestopaikkansa. Tavu- ja mora-ajoituksessa fraasin loppupuolen rajat siirtyisivät pois lisättyjen tavujen tai morien

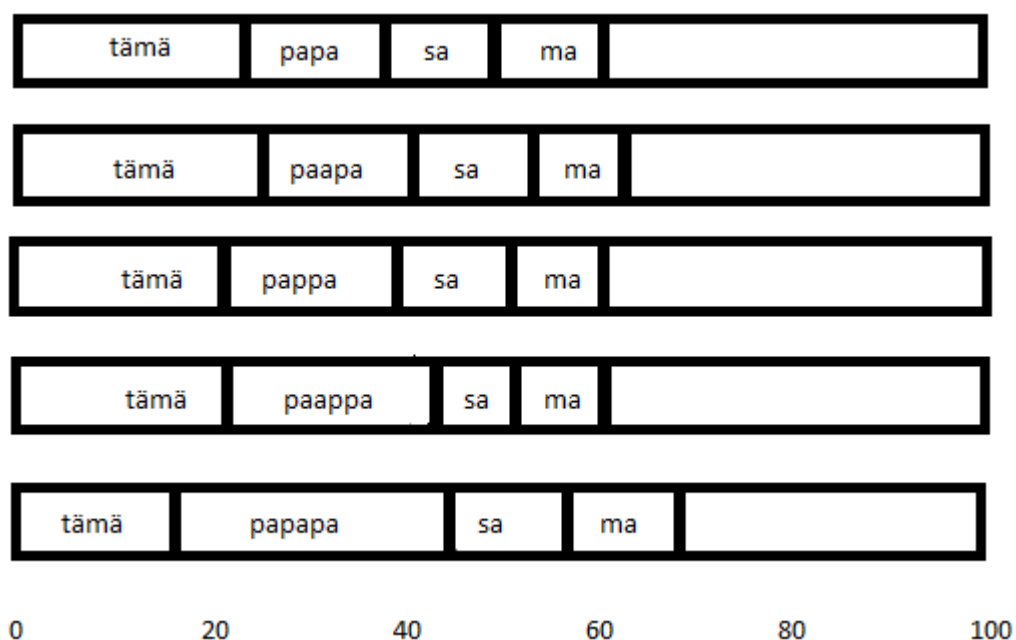
tieltä ja mahdollisesti fraasin alkupuolen sana saattaisi niin ikään tehdä tilaa kasvaneelle koesanalle. On oletettavaa, että suomessa rajojen siirtymistä tapahtuu.



Kuva 1. Ulkoinen sykli. Nuolet osoittavat fraasin osien rajojen odotettuja siirtymäsuuntia koesanana kasvatettaessa. Luvut esittävät suhteellista osuutta syklistä.

3. Tulokset

Kuvassa 2 nähdään kaikkien koehenkilöiden kaikkien toistojen keskiarvoihin perustuvat rajat. Koesanalla on vaikutus rajojen ajankohtiin. Kuvasta on myös helppo nähdä, että viimeinen fraasi erottuu muista selvästi.



Kuva 2. Fraasien rajat ulkoisessa syklistä. Kaikkien koehenkilöiden kaikkien toistojen keskiarvot suhteellisella asteikolla.

Kaksitavuisen koesanan sisältävät fraasit ovat keskimäärin lähes yhtä pitkiä, vaikka koesanan osuus muuttuukin. Kolmitavuisen koesanan sisältävä fraasi on kestoltaan suurempi. Koesanan kasvu saa fraasin viimeisen sanan siirtymään oikealle, mutta myös fraasin ensimmäinen sana antaa tilaa. Koesanan alkuraja siirtyy vasemmalle erityisesti *papapa*-sanankohdalla. Viides koesana alkaa keskimäärin ajankohdassa 18, kun taas muut koesanat ajankohdassa hieman yli 20. Koesana vaikuttaa myös viimeisten tavujen siirtymiseen. *Sa-* ja *ma-*tavut liikkuvat aika lailla samassa suhteessa, joten tämän perusteella on vaikea sanoa, onko tahdin alku kiinteämmin kiinni tietyssä ajankohdassa. Tahteihin perustuvassa ajoituksessa *sa*-tavun siirtyminen olisi vähäisempää kuin *ma*-tavun (Tajima 1998: 73–77), mutta tällaista tendenssiä ei näiden tulosten pohjalta nähdä.

Sisäisessä syklissä tulokset ovat melko odotettuja. Koesanan alku sijoittuu neljässä ensimmäisessä fraasissa melko tarkasti noin toisen kolmanneksen alkuun. Viides koesana alkaa hieman aiemmin. Fraasin kolmas sana alkaa noin kolmannen kolmanneksen alussa. Odotettu tulos on myös koesanan kesto. *Papa* vie fraasista 27 %, *papapa* 38 % ja kolme muuta 30–34 %.

Myös absoluuttiset kestot osoittavat joitakin odotettuja tuloksia. Koesanoista *papapa* on pisin, *paappa* on toiseksi pisin, ja *papa* lyhin. Sanan kesto kasvaa järjestyksessä. Samoin fraasin eli sisäisen syklin kestossa *papapa*-fraasi on selvästi pisin, ja muut sijoittuvat lähemmäs toisiaan, mutta kestojen suuruusjärjestys ei ole aivan sama kuin koesanalla. On myös odotettua, että viidennen fraasin tauko on lyhempi kuin muilla fraaseilla. Koesanan kasvu vie siis tilaa toistokertojen väliseltä tauolta.

Taulukko 1. Absoluuttisten kestojen keskiarvot millisekunneissa. Punainen merkitsee suurinta kestoa ja sininen pienintä.

	sana	fraasi	tauko	sykli
papa	252	921	586	1507
paapa	275	931	569	1494
pappa	287	907	597	1504
paappa	309	920	586	1506
papapa	378	994	504	1498

Ulkoisen syklin kesto ei luonnollisestikaan vaihtelee paljoa, koska se perustuu pakotettuun puherytmiin, mutta siinäkin on noin 10 millisekunnin suuruista huojuntaa. Mielenkiintoista on,

että ensimmäisen fraasin sykli on pisin mutta viidennen fraasin sykli kuuluu toiseksi lyhin. Viides koena mahdollisesti saa puhujat kiihottamaan, jotta toisto saadaan pysymään annetuissa rajoissa. Tätä saattoi havaita myös koetilanteessa, sillä osa koehenkilöistä selvästi nopeutti tuotostaan *papapa*-sanon kohdalla, ja yhdellä koehenkilöllä viides fraasi täytti lähes koko syklin.

4. Johtopäätökset

Voisiko saatujen tulosten perusteella sanoa jotakin suomen rytmistä? Fraasin rajat osoittavat siirtymistä niin paljon, että suomi ei kuulu ainakaan painoajoitteiseen tyyppiin. Kvantiteetti siirtää rajoja, joten moroluvulla on vaikutusta suomen rytmisä, mikä tarkoittaa, että suomi ei ole puhtaasti tavuajoitteinen kieli. Tavun lisääminen vaikuttaa enemmän kuin moron lisääminen, joten suomi ei ole puhtaasti myöskään mora-ajoitteinen.

Sekä moralla että tavulla on siis merkitystä suomen rytmisä. Fraasin kesto ja näin ollen syklin tauko ovat kutakuinkin keskimäärin yhtä pitkiä kaksitavuisen koenan sisältävissä sykleissä, mutta koenan kesto riippuu tästä huolimatta moroluvusta. Kolmitavuinen koena saa fraasin keston kasvamaan, ja myös itse koena on selvästi pitempi kuin kaksitavuiset koenat. Koena *paappa* voidaan katsoa nelimoraiseksikin, mutta vaikka se tulkittaisiin vain kolmimoraiseksi, mitä tämä tutkimus puoltaa, se ei silti saa aikaan yhtä suurta vaikutusta kuin niin ikään kolmimorainen *papapa*. Tavuluku on tässä määrävssä asemassa. Huomionarvoista on myös, että tavun lisääminen koenaan saa koenan laajenemaan erityisesti vasemmalle eli viemään tilaa fraasin ensimmäiseltä sanalta. Tämä voi tuki johtua koeasetelmasta, jossa puhuja on pakotettu mahduttamaan pitkän koenan annettuihin rajoihin. Kiihottamisesta kertoo sekin, että viidennen koenan loppuajankohta ei eroa muista koenoista niin paljon kuin sen alkuajankohta. Samanlaista vaikutusta ei kuitenkaan tapahdu yhtä paljon moron lisäämisen yhteydessä, joten tavulla voidaan katsoa olevan moraa suurempi merkitys. Tässä tutkimuksessa ei tarkkailtu tahteja, mutta niilläkin on mahdollisesti vaikutuksensa suomen rytmisä. Kaksitavuisen tahdin sisältämät fraasit ovat kestoiltaan samankaltaisia, mutta kolmannen tavun lisääminen tahtiin pidentää fraasin kestoa. Lisätavu ikään kuin häiritsee kaksitavuisuuteen perustuvaa järjestelmää, mikä näkyy fraasin keston lisäksi siinä, että kolmas tavu vaikuttaa erityisesti koenan eli tahdin alkuajankohtaan.

Suomessa on siis sekä mora- että tavuaajoituksen piirteitä. Morat kasvattavat sanan kestoa, mutta tavuluku kasvattaa sitä enemmän. Tavuluku vaikuttaa myös fraasin kestoon moraalukua enemmän. Tässä tutkimuksessa tehty puheenkierätykskoe antaa selvää näyttöä suomen rytmistä. Samalla se myös osoittaa, että kielen rytmityyppi ei ole määriteltävissä yhdellä piirteellä, vaan se koostuu useista tekijöistä.

5. Lähteet

Cummins, F. (2002). Speech rhythm and rhythmic taxonomy. *Proceedings of Speech Prosody 2002*, Aix en Provence, 121–126.

Dauer, R. (1987). Phonetic and phonological components of language rhythm. *Proceedings of the XIth international congress of phonetic sciences*, Tallinn, 447–450.

Isei-Jaakkola, T. (2004). *Lexical quantity in Japanese and Finnish*. Helsingin yliopiston fonetiikan laitoksen julkaisuja 48. Helsinki: Hakapaino oy.

Mairano, P. (2011). *Rhythm typology. Acoustic and perceptive studies* (Väitöskirja). Dipartimento di Scienze Letterarie e Filologiche. Università degli Studi di Torino.

Pamies Bertrán, A. (1999). Prosodic typology: On the dichotomy between stress-timed and syllable-timed languages. *Language Design*, 2, 103–130.

Port, R., Tajima, K., Cummins, F. (1998). Speech and rhythmic behavior. Teoksessa Geert J. P. Savelsburgh, Han van der Maas & Paul C. L. van Geert (toim.), *The non-linear analysis of developmental processes*, 53–78. Amsterdam: Elsevier.

Roach, P. (1982): On the distinction between 'stress-timed' and 'syllable-timed' languages. Teoksessa David Crystal (toim.), *Linguistic controversies*, 73–179. London: Edward Arnold (Publishers) Ltd.

Tajima, K. (1998). *Speech rhythm in English and Japanese: Experiments in speech cycling* (Väitöskirja). Department of Linguistics. Indiana University.

On the prosody of CV.CV words in the Eastern Finnish dialects

Riikka Ylitalo, University of Oulu

Abstract

Contrastively accented CV.CV words were produced by speakers of five different subgroups of the Eastern Finnish dialects. The analyses show that the half-lengthening took place in every variety studied, but there were some differences in the location of the F_0 peak between the varieties. When the C_2 was voiced, the peak was in Lappeenranta and Joensuu varieties located in C_2 and thus in the second syllable, in the other varieties the peak was in the first syllable. When the C_2 was voiceless, the peak was in all the varieties in the first syllable, but unfortunately, Lappeenranta had to be excluded from the analyses of the words with a voiceless C_2 .

1 Background and goals of the study

In many Finnish regional varieties, the so-called half-lengthening takes place in the CV.CV(X)-structured words: the second single vowel of those words has a much longer duration than the first single vowel of the same word. But there are also many regional varieties of Finnish, in which the half-lengthening does not occur. In those varieties, the V_2 of the CV.CV(X) words is about the same length as V_1 , or even shorter than V_1 . The half-lengthening does not occur roughly in the varieties based on Häme and Southern Ostrobothnian dialects, and takes place in at least most of the other areal varieties. [1, 2.]

There are also some differences between the Finnish regional varieties in the F_0 curves realising accent during the CV.CV(X) words. In Oulu and Rovaniemi varieties, which both have half-lengthening, the F_0 peak is located in the first syllable, as well as in all the varieties without half-lengthening studied so far. In Turku variety, which has half-lengthening, the F_0 peak is in the second syllable: in the immediate beginning of V_2 when C_2 is voiceless, and close to the midpoint of C_2 , when C_2 is voiced. The situation seems to be similar also in the Kuopio variety, which has half-lengthening: when C_2 is voiced, the F_0 peak is located roughly in the middle of the C_2 [3, 4, 5], but CV.CV(X) words with voiceless C_2 have not been investigated from Kuopio variety.

Kuopio belongs to the area where Eastern Finnish dialects are spoken. The main goal of this study is to find out where the F_0 peak during accented CV.CV words is located in some other varieties based on Eastern Finnish dialects, when C_2 is voiced, and when it is voiceless. At the same time, further information on F_0 curves during the CV.CV words and on the amount of the half-lengthening in the eastern dialects is achieved.

2 Methods

Kajaani, Joensuu, Jyväskylä, Mikkeli and Lappeenranta towns, with their surroundings belonging to the same dialectal subgroup as the central town, were chosen to the study. All these towns belong to the Eastern dialect area. Kajaani, Joensuu, Jyväskylä and Mikkeli belong to the area of the Savo subgroup, and Lappeenranta belongs to the area of the South-Eastern subgroup of the Eastern dialects. Kajaani, Joensuu, Jyväskylä and Mikkeli dialects further belong to different subgroups of the Savo dialects. [6.]

The informants of the study were all female, at the time of the recordings 20–27 years of age. They were born, or lived since early childhood, in the dialect area they represent with at most short stays elsewhere. At the time of the recordings they all were studying in Oulu, where they had lived a time varying from a few weeks to a few years. At the current stage of the study, 5 speakers from Kajaani region have been recorded, from Joensuu 4, from Jyväskylä 4, from Mikkeli 3, and from Lappeenranta 3 speakers. The final goal is to record an even amount of speakers from all the varieties, but unfortunately more suitable speakers have not been reached so far. The informants were recorded in 2014–16 in the recording studio of logopedics and phonetics at the University of Oulu. The informants' speech was recorded directly on a hard disc using high-quality equipment (44.1 kHz, 16 bits).

The informants read the target words with their frame sentences from a computer screen. The target words are Finnish nouns with the structure CV.CV. The speakers from Kajaani, Mikkeli and Jyväskylä regions read 30 CV.CV words whose consonants were chosen from the group /l m n/, and 20 CV.CV words whose C_2 was chosen from the group /k p t s/. Also 3 speakers of the Joensuu variety and one speaker of the Lappeenranta variety read the 50 target words just described. Two speakers of the Lappeenranta variety and one speaker of the Joensuu variety read only the 30 CV.CV words with consonants from the group /l m n/. The reason for that was that their recordings were made in an earlier stage of the study, when the research plan did not have its final form. The 30 CV.CV words with the consonants from the group /l m n/, which all the

informants read, are the same than in Ylitalo's earlier studies on F_0 and segment durations in Kuopio, Rovaniemi, Turku, Helsinki and Seinäjoki varieties [3, 4, 5]. Also the frame sentences of the words are the same as in Ylitalo's earlier studies. 9 target words of the 20 CV.CV words whose C_2 was chosen from the group /k p t s/, and their frame sentences, are the same than Ylitalo's study concerning Oulu, Turku and Tampere varieties [3]. 11 additional target words whose consonants are from the group /k p t s/ were selected, and their frame sentences were composed first time for this study.

The target words were placed in the frame sentences to positions where they would probably be contrastively accented, as in Ylitalo's earlier studies [3, 4, 5]. An example of a sentence which the informants read, is *Sanoin että hyvä NIMI puuttuu, en sanonut että hyvä AJATUS puuttuu* 'I said that a good NAME is missing, I didn't say that a good THOUGHT is missing', where the target word is *nimi*. The informants were asked to emphasize the words written with capitals. All the sentences had the same grammatical structure.

The acoustic measurements were made using Praat [7]. The segmentation criteria were those used by Lehtonen [8] and Ylitalo [3]. Segment durations were measured using displays of the waveform and broadband spectrogram, as well as auditory criteria. Also the distance of the F_0 peak from the beginning of the word was measured. Fundamental frequency was measured 1) at the beginning and end of the voiced portion of the syllable preceding the target word, 2) at the beginning, middle and end of the voiced portion of the initial syllable of the target word, and half-way between the beginning and the middle as well as half-way between the middle and the end, 3) at the corresponding five voiced temporal locations of the second syllable of the target word, 4) at the beginning and end of the voiced portion of the syllable following the target word, and 5) at the F_0 peak during the target word.

All measured F_0 values were normalised before the statistical analyses: from each F_0 value measured, the F_0 value at the beginning of the voiced portion of the syllable preceding the target word was subtracted, and to the remainder the grand mean F_0 value at the beginning of the voiced portion of the syllable preceding the target word across all syllables preceding all the target words was added. Univariate ANOVAs were used to test the statistical effects of the grouping variable Variety on the means of each dependent variable, Variety acting as fixed factor. In addition, Tukey post hoc tests were used, with an alpha level $p < 0.05$.

3 Results

There were differences between varieties in the total word duration [$F(4,883) = 38.54, p < 0.001$]. The word duration was longest in Joensuu variety, in average 478 ms, and second longest in the Kajaani variety, 444 ms. The word duration was shorter in the Jyväskylä variety, 413 ms, and even shorter in the Lappeenranta variety, 388 ms. The word duration in the Mikkeli variety, 409 ms, differed from the durations in Joensuu and Kajaani, but not from the durations in Lappeenranta and Jyväskylä. All following analyses of the durations are therefore based on percentage values, which represent each durational variable's proportion of the total word duration.

The analyses are made separately of words, whose consonants are voiced, and separately of words, whose consonants are voiceless. One reason for dividing the materials in two this way is that the segments of the target words have different durations in these different word groups. A voiceless C_1 , on average 131 ms, is longer than a voiced C_1 , 101 ms [$F(1,867) = 163.05, p < 0.001$]. V_1 is shorter, 93 ms, between voiceless consonants than between voiced consonants, 105 ms [$F(1,873) = 74.90, p < 0.001$]. A voiceless C_2 , on average 117 ms, is longer than a voiced C_2 , 70 ms [$F(1,873) = 1319.18, p < 0.001$]. V_2 is shorter, 124 ms, after voiceless C_2 than after voiced C_2 , 135 ms [$F(1,873) = 27.23, p < 0.001$]. In the words that include only voiceless consonants, the consonants are thus relatively long and the vowels relatively short. On the contrary, in the words which include only voiced consonants, the consonants are relatively short, and the vowels relatively long. The total word duration is longer, 464 ms, when the word's consonants are voiceless, than when they are voiced, 412 ms.

Other reasons for analysing the words whose consonants are voiced separately from those whose consonants are voiceless, are that syllable-bound F_0 measuring points are located quite differently in the word depending on the consonants' ability to realise F_0 . And of course the F_0 peak can be located in C_2 only if the segment is voiced. One target word of the original materials, *näky* 'a sight' was excluded from the analyses above, because the other consonant in the word is voiced and the other is voiceless. The word *näky* was also excluded from all the following analyses.

3.1 The words with voiced consonants

3.1.1 Segment durations in the words with voiced consonants

Variety had an effect on the proportional duration of C_1 [$F(4,569) = 5.85, p < 0.001$]. The percentage duration of C_1 was greater in Mikkeli, 26.3% of the total word duration, than in Joensuu (23.6%), Kajaani (23.7%) and Lappeenranta (24.2%), which did not differ statistically from each other with respect to the duration of C_1 . Jyväskylä (24.9%) did not differ significantly from any other Variety. Variety also had an effect on the proportional duration of V_1 [$F(4,569) = 5.01, p < 0.01$]. The percentage duration of V_1 was greater in Lappeenranta (26.6%) than in Mikkeli (24.6%). Joensuu (25.5%) did not differ statistically from any other Variety. Kajaani (25.2%) differed only from Lappeenranta, and Jyväskylä (26.3%) differed only from Mikkeli.

Variety did not have an effect on the duration of C_2 . The proportional duration of V_2 was affected by Variety [$F(4,569) = 11.08, p < 0.001$]. The percentage duration was greater in Kajaani (34.5%) than in Jyväskylä (31.4%) and Mikkeli (31.5%), which did not differ from each other with respect to the duration. Joensuu (33.8%) differed only from Jyväskylä and Mikkeli, Lappeenranta (32.0%) only from Kajaani. The half-lengthening was stronger in Kajaani (V_2 duration 141.0% of the duration of V_1) and Joensuu (the corresponding ratio 135.2%) than in Jyväskylä and in Lappeenranta, which both had the ratio 122.8%. Mikkeli (132.2%) did not differ significantly from any other Variety. [$F(4,569) = 7.34, p < 0.001$.] The proportion of the first syllable of the total word duration was also counted. The percentage proportion of the first syllable was greater in Jyväskylä (51.1% of total word duration) than in Kajaani (48.9%) and Joensuu (49.1%), which did not differ from each other. Lappeenranta (50.8%) and Mikkeli (50.9%) differed only from Kajaani. [$F(4,569) = 5.61, p < 0.001$.]

To analyse the location of the highest point of the fundamental frequency, the measured distance from the beginning of the word to the F_0 peak was converted to a percentage proportion of the total word duration. Variety had an effect of this proportion [$F(4,539) = 9.61, p < 0.001$]. The distance from the beginning of the word to the F_0 peak was greater in Lappeenranta (54.3% of total word duration) than in the other varieties, which did not differ statistically from each other in this respect. The percentages are for Kajaani 46.8, Jyväskylä 47.1, Mikkeli 48.6 and Joensuu 50.1. When the F_0 peak location is compared to the location of the syllable border, it can be seen that the average location of the peak is in Lappeenranta and Joensuu in the second syllable. In the other varieties, the peak is located in average in the first syllable. In Joensuu, the peak is though very close to the syllable border.

3.1.2 Fundamental frequencies in the words with voiced consonants

In this article, I will discuss only fundamental frequencies during the target words, not during the syllables preceding and following the target words. It will still be remembered from above that all measured F_0 values were normalised before performing the analyses, with reference to the beginning of the voiced portion of the syllable preceding the target word.

Variety had no effect on the fundamental frequency at the F_0 peak. At the first measurement point of the first syllable, F_0 was lower in Lappeenranta than in Jyväskylä and Joensuu. Mikkeli and Kajaani did not differ from any other Variety. [$F(4,536) = 4.44$, $p < 0.01$.] At the second measurement point of the first syllable, F_0 was lower in Lappeenranta than in the other Varieties, which did not differ from each other [$F(4,537) = 5.17$, $p < 0.001$]. At the third measurement point of the first syllable, F_0 was lowest in Lappeenranta, higher in Kajaani and highest in Jyväskylä [$F(4,538) = 8.63$, $p < 0.001$]. Joensuu and Mikkeli did not differ significantly from Kajaani and Jyväskylä, but they differed from Lappeenranta. At the fourth measurement point of the first syllable, F_0 was lower in Lappeenranta than in Mikkeli and Jyväskylä [$F(4,538) = 6.02$, $p < 0.001$]. Kajaani and Joensuu did not differ from Lappeenranta and Mikkeli, but differed from Jyväskylä. At the fifth measuring point of the first syllable, Variety had no effect on F_0 .

At the first measurement point of the second syllable, Variety had no effect on F_0 . At the second measurement point of the second syllable, F_0 was higher in Joensuu than in Jyväskylä [$F(4,538) = 3.55$, $p < 0.01$]. The other Varieties did not differ from any other Variety. At the third measurement point of the second syllable, F_0 was lowest in Mikkeli and Jyväskylä, higher in Kajaani and highest in Joensuu. Lappeenranta did not differ from Kajaani and Joensuu, but differed from Mikkeli and Jyväskylä. [$F(4,532) = 20.83$, $p < 0.001$]. At the fourth measurement point of the second syllable, F_0 was lowest in Mikkeli, higher in Jyväskylä and highest in Kajaani, Lappeenranta and Joensuu [$F(4,494) = 23.62$, $p < 0.001$]. Finally, at the fifth measurement point of the second syllable, F_0 was lower in Mikkeli and Jyväskylä than in Kajaani, Joensuu and Lappeenranta [$F(4,493) = 17.12$, $p < 0.001$].

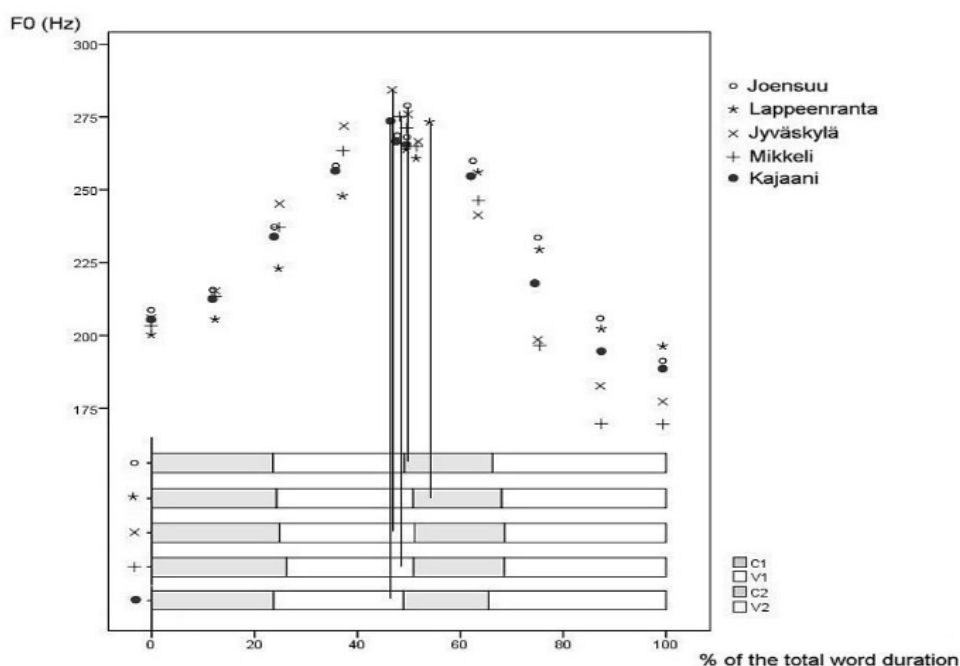


Figure 1: F_0 and segment durations in CV.CV words with voiced consonants in five Varieties.

The F_0 curves and segment durations of the words with voiced consonants can be seen in Figure 1. From the highest point of F_0 in each Variety, a line is drawn to the duration bar of the same Variety.

3.2 The words with voiceless consonants

3.2.1 Segment durations in the words with voiced consonants

The Lappeenranta Variety was excluded from the analyses of the words containing voiceless consonants, because unfortunately, only one speaker from Lappeenranta produced the target words with voiceless consonants (see above).

Variety had an effect on the proportional duration of C_1 . The percentage duration of C_1 was greater in Joensuu (30.2% of the total word duration) and in Mikkeli (29.3%) than in Kajaani (26.2%). Jyväskylä (28.2%) did not differ statistically from any other Variety. [$F(3,278) = 7.40$, $p < 0.001$]. Variety also had an effect on the proportional duration of V_1 [$F(3,278) = 8.94$, $p < 0.001$]. The percentage duration was greater in Kajaani (21.0%) and Joensuu (20.0%) than in Mikkeli (18.1%). Jyväskylä (19.5%) did not differ from any other Variety. The percentage duration of C_2 was significantly greater in Jyväskylä (27.3%) and Mikkeli (26.9%), which did not differ from each other, than in Kajaani (23.6%) and Joensuu (23.9%), which did not differ

from each other. [$F(3,278) = 30.68, p < 0.001$.] The proportional duration of V_2 was greater in Kajaani (29.1%) than in the other Varieties, which did not differ from each other [$F(3,278) = 14.35, p < 0.001$]. The percentage durations in the other Varieties were 25.0% in Jyväskylä, 25.7% in Mikkeli and 25.9% in Joensuu.

Variety had no effect on the amount of the half-lengthening. The ratio V_2/V_1 is 132.5% in Joensuu, 134.2% in Jyväskylä, 143.2% in Kajaani and 146.8% in Mikkeli. Variety had an effect on the first syllable proportional duration of the total word duration [$F(3,278) = 4.66, p < 0.01$]. The proportional duration of the first syllable was greater in Joensuu (50.2% of the total word duration) than in the other Varieties, which did not differ from each other. The proportional duration of the first syllable was 47.2% in Kajaani, 47.4% in Mikkeli and 47.7% in Jyväskylä. The distance between the F_0 peak location and the beginning of the word, expressed as a proportion of the total word duration, was greater in Joensuu (50.1%) than in Jyväskylä (39.4%). The other Varieties did not differ significantly from Joensuu, from Jyväskylä or from each other. [$F(3,263) = 6.58, p < 0.001$.] The proportional distance was 44.7% in Mikkeli and 44.6% in Kajaani. Thus, the highest point of F_0 was on average located in the first syllable in every Variety, although in Joensuu the average peak was very close to the syllable border. It may be worth noting that the first syllable had a greater proportional duration in Joensuu than in the other Varieties (see above). If the first syllable had even a little shorter duration, the average F_0 peak in Joensuu would inevitably have been located in the second syllable, as in the South-Western Turku variety [3].

3.2.2 Fundamental frequencies in the words with voiceless consonants

Variety had no effect on the fundamental frequency at the F_0 peak. At the first [$F(3,266) = 4.87, p < 0.01$], second [$F(3,266) = 4.53, p < 0.01$] and third [$F(3,265) = 3.38, p < 0.05$] measurement points of the first syllable, F_0 was higher in Jyväskylä than in Kajaani, and Joensuu and Mikkeli did not differ statistically from the other Varieties. In the fourth and fifth measurement points of the first syllable, Variety had no effect on F_0 . At the first [$F(3,258) = 19.84, p < 0.001$], second [$F(3,256) = 29.50, p < 0.001$], third [$F(3,243) = 34.41, p < 0.001$], fourth [$F(3,239) = 33.09, p < 0.001$] and fifth [$F(3,241) = 24.31, p < 0.001$] measurement points of the second syllable, F_0 was higher in Joensuu and Kajaani than in Jyväskylä and Mikkeli. Joensuu and Kajaani did not differ statistically from each other, nor did Jyväskylä and Mikkeli.

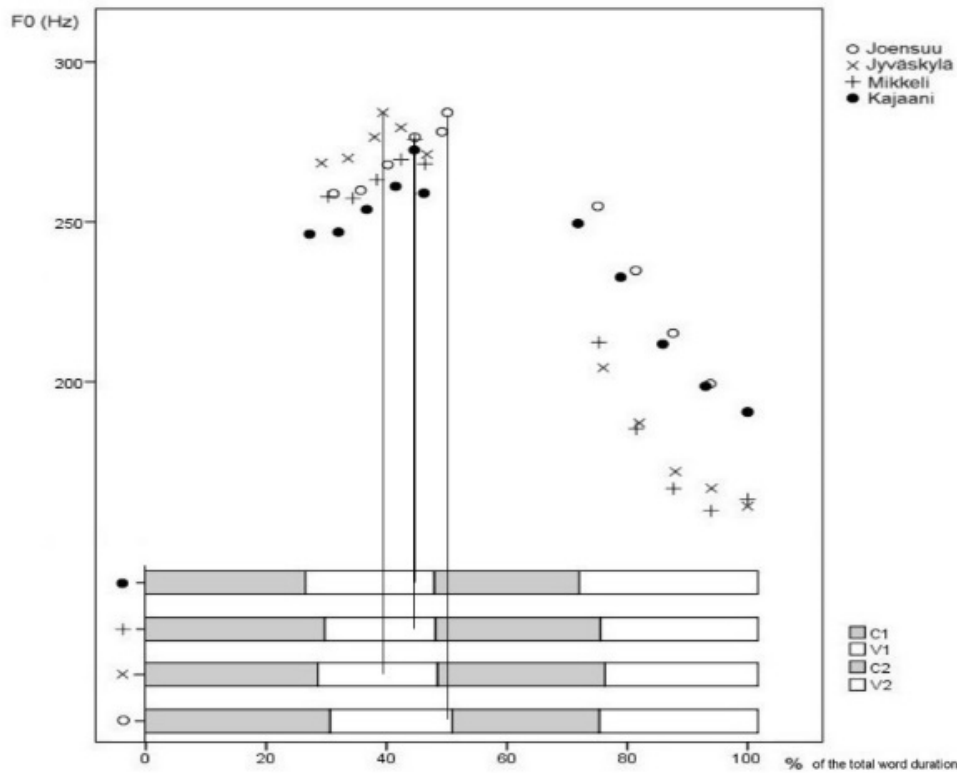


Figure 2: F_0 and segment durations in CV.CV words with voiceless consonants in four Varieties.

The F_0 curves and segment durations of the words with voiceless consonants can be seen in Figure 2. From the highest point of F_0 in each Variety, a line is drawn to the duration bar of the same Variety. Note that the lines for Mikkeli and Kajaani are very close to each other.

4 Summary and discussion

There were some differences in the segment durations between the varieties. Both in the words with voiced consonants and the words with voiceless consonants, C_1 was longer in Mikkeli than in Kajaani, and V_1 was longer in Kajaani than in Mikkeli. In both target word groups V_2 was longer in Kajaani than in Mikkeli and Jyväskylä. The half-lengthening took place in all the varieties, also in Lappeenranta, which has been in some earlier studies classified as an intermediate dialect between those with half-lengthening and those with no half-lengthening. In the materials of this study, the duration of V_2 was 122.8% of the duration of V_1 in Lappeenranta, when C_2 was voiced. That was a smaller amount of half-lengthening than in Kajaani and Joensuu in the same words, but exactly the same percent of half-lengthening that was measured also in Jyväskylä.

Variety had no effect on the Hz value at the F₀ peak, but there were some differences in the location of the F₀ peak between the varieties. When the C₂ was voiced, the peak was in Lappeenranta and Joensuu located in C₂, in the other varieties the peak was in V₁. When the C₂ was voiceless, the peak was in all the varieties in V₁, but unfortunately, those words were not analysed from Lappeenranta variety. The biggest shortcoming of this study is that CV.CV words with voiceless consonants were not recorded from Lappeenranta variety enough for statistical analysis. Additional recordings of that kind of words are needed in the future.

References

- [1] K. Wiik and I. Lehist, "Vowel Quantity in Finnish Disyllabic Words," P. Ravila (ed.), *Congressus Secundus Internationalis Fenno-Ugristarum. Helsingiae habitus 23.–28. VIII. 1965. Pars I. Acta Linguistica*, pp. 569–574. Helsinki: Suomalais-Ugrilainen Seura, 1968.
- [2] K. Wiik, "Suomen murteiden vokaalien kestoista," O. Aaltonen ja T. Hulkko (eds.), *Fonetiikan päivät – Turku 1985. XIII Fonetiikan päivillä Turun yliopistossa pidetyt esitelmät*, pp. 253–317. Turku: Turku University, 1985.
- [3] R. Ylitalo, *The realisation of prominence in three varieties of standard spoken Finnish*. PhD thesis. Acta Universitatis Ouluensis B 88. <http://herkules.oulu.fi/isbn9789514291142>. Oulu: Oulu University, 2009.
- [4] R. Ylitalo, "Aksentoitujen CV.CV-rakenteisten sanojen F₀ ja segmenttikestot Etelä-Pohjanmaalla ja Pohjois-Savossa," M. Lehtinen and U.K. Laine (eds.), *XXIX Fonetiikan Päivät, Espoo 20.-21.3.2015. Julkaisut – Papers*, pp. 12–15. Aalto University publication series Tiede + teknologia 7/2015. http://fp2015.aalto.fi/Fonetiikan_Paivat-2015_Aalto-yliopisto.pdf.
- [5] R. Ylitalo, "Puolipituussanojen alueellisesta prosodisesta vaihtelusta," to appear in *Virittäjä*.
- [6] L. Kettunen, *Suomen murteet III A. Murrekartasto*. <http://kettunen.fnhost.org/>. Helsinki: Suomalaisen Kirjallisuuden Seura, 1940.
- [7] P. Boersma and D. Weenink, Praat: doing phonetics by computer [computer program]. Version 6.0.20, retrieved 3 September 2016 from <http://www.praat.org/>.
- [8] J. Lehtonen, *Aspects of Quantity in Standard Finnish*. Studia philologica Jyväskyläensia VI. Jyväskylä: Jyväskylä University, 1970.

Passiivisen kuuntelun vaikutus alakouluikäisten lasten vieraan kielen havaitsemiseen

Posteri

Katja Immonen^{1,2}, Maija S. Peltola^{1,2}

¹ *Fonetiikka, Turun yliopisto*

² *Learning, Age and Bilingualism -laboratorio (LAB-lab)*

Vieraan kielen oppiminen asettaa oppijalle usein erilaisia haasteita. Kohdattavat oppimisvaikeudet voivat olla hyvin eriasteisia riippuen siitä, miten eri- tai samankaltainen oppijan äidinkieli on kohdekieleen verrattuna (esim. Speech Learning Model, Perceptual Assimilation Model). Lapset ovat kuitenkin kielenoppijoina yleensä aikuisia menestyksekkäämpiä aivojensa plastisuuden ansiosta. Tässä tutkimuksessa selvitettiin intensiivisen kuuntelun vaikutuksia 7-12-vuotiaiden lasten vieraan kielen äännekontrastin havaitsemiseen. Tutkimus on osa väitöskirjaprojektia, jossa selvitetään erilaisista kielitaustoista tulevien lasten vieraan kielen oppimiseen vaikuttavia tekijöitä.

Tutkimuksessa tarkasteltiin suomenkielisillä yksikielisillä lapsilla, kuinka intensiivinen passiivinen kuuntelu vaikuttaa vieraan kielen äänten havaitsemiseen tietoisella ja esitietoisella tasolla. Koeasetelmassa pyrittiin simuloimaan luonnollisessa kieliympäristössä tapahtuvaa äänteiden omaksumista ja muistijälkien muodostumista. Kokeen kuunteluosioissa lapsille soitettiin syntetisoituja vieraan kielen vokaalikategorian /ʌ/ eri variantteja niin, että he kuulivat /ʌ/ :n prototyyppistä edustajaa prosentuaalisesti eniten ja kategorian muita variantteja vähenevässä määrin suhteessa prototyyppiin. Lapset osallistuivat tutkimukseen kahtena peräkkäisenä päivänä. Molempiin päiviin sisältyi noin tunnin mittainen passiivinen kuunteluosio. Passiivisen altistuksen vaikutuksia mitattiin diskriminaatiokokeilla ja aivojen herätevastemittauksilla ennen ja jälkeen kuuntelun.

Alustavien tulosten valossa näyttää siltä, että tutkimuksen behavioraalinen osuus oli tehtävänä nuorimmille (7-8-vuotiaille) koehenkilöille liian haastava. Vanhempien (9-12-vuotiaiden) lasten tulokset antavat viitteitä jonkinasteisesta positiivisesta kehityksestä reaktioajoissa ja diskriminaatiotarkkuudessa. Herätevastemittausten osalta tulokset ovat vielä analyysivaiheessa. Kerron tutkimuksen tuloksista ja niiden merkityksestä tarkemmin posteriesityksessäni. Tulosten toivotaan antavan uutta ja arvokasta tietoa siitä, kuinka alakouluikäiset lapset oppivat havaitsemaan teoreettisesti mahdollisimman vaikean vieraan kielen äännekontrastin sekä kuinka uudet äänteiden muistijäljet muodostuvat lasten aivoihin.

Asiasanat: Vieraan kielen oppiminen, havaitseminen, lapset

Vaikuttaako kieliharrastus vieraan äänteen havaitsemiseen ikääntyneillä?

Katri Jähi¹, Paavo Alku², Maija S. Peltola¹

¹Fonetiikka ja Learning, Age and Bilingualism -laboratorio, Turun yliopisto

²Signaalinkäsittelyn ja akustiikan laitos, Aalto-yliopisto

Tämän tutkimuksen tarkoituksena oli selvittää, miten kieliharrastus vaikuttaa vieraan äänteen havaitsemiseen ikääntyneillä.

Koehenkilöt olivat 62–73 -vuotiaita eläkeläisiä ja heidät oli jaettu kahteen ryhmään. Kieliryhmän koehenkilöt rekrytoitiin jatkotason kielikurssilta ja toisessa ryhmässä olleilla koehenkilöillä oli muita kuin kielten oppimiseen liittyviä harrastuksia (esimerkiksi liikunta ja lukeminen). Kummankaan ryhmän koehenkilöt eivät osanneet ruotsin kieltä juuri perusteita enempää, eivätkä he käyttäneet säännöllisesti ruotsia. He eivät olleet myöskään asuneet muissa Pohjoismaissa kuin Suomessa. Koehenkilöt osallistuivat kaksipäiväiseen tutkimukseen, joka sisälsi EEG-rekisteröinnin (Mismatch Negativity, MMN), diskriminaatio- ja reaktioaikakokeen, identifikaatio- ja hyvyysarviointikokeen, tuottokokeen sekä kaksi harjoitusosiota. MMN-mittaus ja diskriminaatiokoe tehtiin kahdella eri ärsykeparilla, joista toinen oli harjoittelussa käytetty ärsykepari ja toisen parin ärsykkeet olivat lähellä toisiaan, mutta kuuluivat kuitenkin eri äännekategorioihin. Jälkimmäisessä parissa akustinen ero oli huomattavasti pienempi, joten eroa oli vaikeampi havaita ja tuottaa. Koeärsykeiksi valittiin semisynteettiset epäsanat /tʉ:ti/ ja /ty:ti/, joista kohdesanan ensimmäinen äänne /tʉ/ on suomenkieliselle vieras, mutta samankaltainen akustisesti lähellä olevien suomen /y/ ja /u/ äänteiden kanssa. Vieraan kielen oppimisen teorioiden valossa tämän kaltainen tilanne aiheuttaa oppimisvaikeuksia (SLM, Flege 1987 ja PAM, Best & Strange 1992).

Tämän posterin yhteydessä esittelen koesarjan tuloksista MMN-mittausten ja diskriminaatiokokeen tulokset. Alustavien tulosten perusteella näyttäisi siltä, että Kieliryhmään kuuluvien koehenkilöiden reaktioajat nopeutuvat jopa vaikeamman ärsykeparin kohdalla, vaikkakaan behavioraalisessa erotteluherkkyydessä ei näytä olevan eroja ryhmien välillä. MMN-mittauksessa eri ärsykeparit saavat aikaan erilaiset vasteet, mutta ryhmien välillä ei näyttäisi olevan suuria eroja. Lopulliset tulokset esittelen XXX Fonetiikan päivillä Oulussa.

Asiasanat: vieraan kielen oppiminen, ikääntyminen, harjoittelu

Title: Statistical learning in the perception of sentence prominence

Authors: Sofoklis Kakouros & Okko Räsänen

Affiliation: Department of Signal Processing and Acoustics, Aalto University

Abstract:

The study of probabilistic effects in language has played a central role in many models of language production and perception. However, most models focus on the probabilistic effects observed at the level of linguistic units, such as words and syllables. In the context of prominence, the majority of the existing work has focused on analyzing the acoustic prosodic correlates such as energy, F0, duration, and spectral tilt and how they are manifested in prominent words and syllables. In this work we investigated a hypothesis that predictability of the acoustic prosodic features in speech and the perception of sentence prominence are mutually connected so that less frequent prosodic outcomes would be perceived as more prominent independently of their actual form or magnitude. We tested this hypothesis by conducting a listening test where subjects were first exposed to a stream of Finnish utterances with a fixed ratio of sentence-final words having either rising or falling F0 trajectories. We then asked them to grade their perceived level of prominence in a set of novel sentences. The results show that subjects were more likely to select words with low-probability trajectories as prominent independently of the direction of the F0 change. This result provides first behavioral evidence towards the predictability hypothesis in prominence perception and suggests that prominence perception mechanisms quickly adapt to the distributional properties of prosodic patterns encountered in speech.

Keywords: statistical learning, sentence prominence, prosody

Suullisen kielitaidon arvioinnin foneettisesta perustasta

Heini Kallio

Svenska Folkskolans vänner -säätiön rahoittama DigiTala -hanke käynnistyi 1.3.2015 tavoitteenaan kehittää suullisen kielitaidon arviointia luotettavammaksi ja käytännöllisemmäksi. Hankkeessa kehitetään tietokoneavusteista suullista kielikoetta suomalaista ylioppilastutkintoa varten. Monitieteisessä hankkeessa on mukana Helsingin yliopiston opettajankoulutuslaitos ja fonetiikan oppiaine sekä Aalto-yliopiston signaalinkäsittelyn ja akustiikan laitos. Yhteistyökumppaneita ovat Ylioppilastutkintolautakunnan lisäksi Jyväskylän yliopiston kielten laitos sekä soveltavan kielentutkimuksen keskus.

DigiTala-hankkeen yhteydessä aloittamani väitöstutkimuksen tavoite on parantaa suullisen kielitaidon arviointimenetelmien objektiivisuutta ja luotettavuutta tarkentamalla arviointikriteerejä foneettisin menetelmin. Saatavat tulokset voivat lisätä kieltenopettajien ja -opiskelijoiden foneettista tietoisuutta ja auttaa kehittämään uusia kieltenopettamisen menetelmiä ja materiaaleja. Tutkimukseni lähtökohtana on puheen ymmärrettävyys ja sen yhteys sujuvuuden käsitteeseen. Suullisen kielitaidon opetuksessa on perinteisesti keskitytty vieraan kielen äänteiden opetteluun, mutta sekä sujuvuutta että ymmärrettävyyttä mittaavissa tutkimuksissa on todettu, että puheen prosodia on yksittäisiä äänteitä tärkeämpää. Tässä tutkimuksessa selvitetään, mitkä puheen prosodiset ominaisuudet ovat oleellisia suullisen kielitaidon arvioinnissa. Tutkimus koostuu menetelmällisesti kolmesta vaiheesta: 1) suullisen kielikokeen kehittämisestä ja puheaineiston keruusta, 2) puheaineiston käsittelystä, analysoinnista ja ymmärrettävyyden mittaamisesta sekä 3) kohdekielten asiantuntijoiden sujuvuusarvioiden keräämisestä.

Esittelen hankkeessa kehitettyä suullisen kielitaidon arviointijärjestelmän prototyyppiä ja sen suunnitteluprosessia sekä ensimmäisiin ihmisarvioihin perustuvien pilottikokeiden tuloksia.

Asiasanat: suullisen kielitaidon arviointi, ymmärrettävyys, prosodia

Pronunciation of word-initial *h* in Estonian

Mari-Liis Kalvik; Liisi Piits

Institute of the Estonian Language

The presentation will report the results of the reading experiment aimed at establishing whether and on what conditions word-initial *h* is pronounced in Estonian. The pronunciation of word-initial *h* has become optional, having ceased to be an obligatory requirement. The liberal attitude is based on the general tendency that even in the early 20th century the word-initial *h* was consistently pronounced only in a few dialect areas. Modern pronunciation variation, however, and its causal factors still need exploring because of the current needs of the development of Estonian speech synthesis.

The study is based on the recordings of read sentences that contain 42 target words with word-initial *h*. 30 words of them were chosen according to their frequency and 12 words with the word-initial *h* occurring in semantically distinctive positions. Study revealed that 95% of occasions the word-initial *h* was pronounced. The main reason of the pronunciation of *h* is supposed to be the visual factor in reading experiment as well as the formality of reading situation. Still, there were differences in the duration and intensity of the examined *h*-s. The aim was to investigate:

- a) how the frequency of the examined words influences the duration of the word-initial *h*;
- b) whether *h* is pronounced more accurately if presented in semantically distinctive position;
- c) is there any correlation between the intensity of word-initial *h* and the speaker's dialectal background.

Keywords: Estonian, word-initial *h*, reading experiment

Havaintoja suomenkielisten oppijoiden ruotsin kielen ääntämisestä

Kautonen Maria & Heinonen Henna

Jyväskylän yliopisto

Aiemmat suomenkielisten puhujien ruotsin ääntämistä käsittelevät tutkimukset ovat harvassa. Suuri osa tutkimuksista käsittelee Ruotsissa asuvien suomenkielisten ruotsinruotsin ääntämistä. Esimerkiksi Bannertin (1980, 2004) tutkimus maahanmuuttajien ääntämisestä on paljastanut useita haasteita, joita suomenkielinen oppija kohtaa ruotsinruotsin segmenttien ja prosodian kohdalla, mm. sibilantit, klusiilit, konsonanttiklusterit, sanapaino ja lauserytmi. Pääasiassa formaalin opetuksen kautta kieltä oppineiden tutkimus on huomattavasti vähäisempää. Suomessa asuvista kielenoppijoista tehdyissä tutkimuksissa on selvitetty mm. yliopisto-opiskelijoiden ruotsinruotsin sana-aksenttien ääntämistä (esim. Kuronen 2015), suomenruotsin klusiileita (Nygård 1977) sekä abiturienttien ääntämistä suullisen kielitaidon testissä (Hildén 2000).

Tässä esitelmässä tarkastelemme Svenska Litteratursällskapet i Finland r.f.:n rahoittamassa FOKUS-projektissa tehtyä ajankohtaista tutkimusta suomenkielisten oppijoiden ruotsin ääntämisestä. Esittelemme tuloksia audiitiivisesta analyysistä 9.-luokkalaisten suomenruotsin ääntämisestä, alustavia tuloksia lukiolaisten ruotsin kielen prosodiasta audiitiivisen ja akustisen analyysin pohjalta sekä tuloksia akustisista mittauksista aikuisten kielenoppijoiden suomenruotsin intonaatiosta. Tutkimuksissa on analysoitu sekä kielenoppijoiden tuottamaa vapaata puhetta (9.-luokkalaiset ja aikuiset) sekä lukupuhuntaa (lukiolaiset).

Tulokset antavat tärkeää tietoa eritasoisista oppijoista, ääntämisen kehittymisestä sekä haasteista ääntämisen oppimisessa. Tätä tietoa voidaan hyödyntää entistä tehokkaamman ääntämisen opettamisen kehittämisessä. Tutkimuksella peruskoululaisten ja lukiolaisten ruotsin kielen ääntämisestä saadaan tärkeää tietoa ääntämistaitojen kehityksestä peruskoulun päättövaiheen sekä lukion päättövaiheen välillä. Suullisen kielitaidon mittaamisen tullessa osaksi kielten ylioppilaskirjoituksia keväällä 2019 tutkimustieto eritasoisten opiskelijoiden vieraiden kielten ääntämistaidoista on tarpeellista ja ajankohtaista.

Asiasanat: ruotsi, ääntäminen, oppiminen

Lähteet:

Bannert, R. (1980). Svårigheter med svenskans uttal: inventering och prioritering. Praktisk lingvistik 5. Lund University, Department of Linguistics.

Bannert, R. (2004). På väg mot svenskt uttal. Lund: Studentlitteratur.

Hildén, R. (2000). Att tala bra, bättre och bäst. Helsingin yliopiston opettajankoulutuslaitos.

Kuronen, M. (2015). Tonaccenterna i avancerade finska inlärares svenska – en akustisk studie. Nordand 2015, nr. 1.

Nygård, A. (1977). Svenska klusiler i finsk mun. Teoksessa: Palmberg, R. & Ringbom, H. (toim), Föredrag vid konferensen om kontrastiv lingvistik och felanalys: Stockholm & Åbo 7-8 februari 1977. (s.71-84). Åbo: Åbo akademi.

Abstrakti

/s/-KOARTIKULAATION KEHITYSPIIRTEISTÄ SUOMENKIELISILLÄ LAPSILLA SPEKTRIMOMENTTIEN VALOSSA

Suomen konsonanttijärjestelmän ainoa koronaalinen frikatiivi on /s/ (S). On ilmeistä, että tämän sibilanttifrikatiivin foneettisissa manifestaatioissa sallitaan suhteellisen paljon vapausasteita. Eräs potentiaalinen S:n laadullisia variaatiota aiheuttava tekijä on sen vokaalikonteksti: viereisten vokaalien artikulaatiopiirteillä, mm. pyöreä vs. ei-pyöreä, tiedetään aikuisten puheessa olevan vaikutuksia S:n foneettisiin piirteisiin. Akustisia tai muita tutkimuksia tästä suomen kielen erästä koartikulaatioilmiöstä on kuitenkin varsin niukalti niin aikuisten kuin lastenkin puheesta.

Tässä äidinkielen äänteellistä kehitystä tarkastelevassa akustisessa, spektrimomentteihin perustuvassa tutkimuksessa, lasten S-tuotosten akustisia piirteitä neljässä symmetrisessä vokaalikonteksteissa (/iSi, uSu, ySy, aSa/) verrattiin vastaviin aikuisten puhujien tuotoksiin. Tutkimuksen koehenkilöinä oli 10–12 lasta kolmesta ikäryhmästä (6v, 9v ja 11v) ja 10 aikuista puhujaa. Koehenkilöiden S-tuotokset ylipäästösuodatettiin (700Hz) ja niiden spektrimomentit M1 (kohinaspektrin painopistetaajuus) ja M2 (kohinaspektrin keskihajonta) laskettiin 11 aikakohdalta tasavälein (0, 10, 20, 30 40 50 60 70 80 90 100%) kustakin S-segmentistä. Akustinen analyysi ja momenttilaskenta tehtiin TF32-sovelluksella (Milenkovic, Madison, Wisc.).

Tulokset paljastivat, että a) vokaalikontekstin vaikutus S-segmenttien M1- ja M2-arvoihin oli systemaattisinta aikavälillä 30–80% segmentin kestosta, ts. S:n *onset*- ja *offset*-vaiheista ei löytynyt selvää vokaalikontekstivaikutusta, b) että vokaalikontekstin vaikutuksessa etenkin S-segmenttien M1-piirteisiin oli systemaattisia, tilastollisesti merkitseviä eroja lasten ja aikuisten verrokkiryhmän välillä ja c) että näitä piirre-eroja oli myös eri-ikäisten lasten tuotoksissa. Hieman yllättäen nuorimman lapsiryhmän S-tuotosten M1-piirteissä oli eniten aikuispuhujien tuotosten kaltaisia koartikulaatiopiirteitä.

Tämän tutkimuksen tuloksia voidaan pitää vain alustavina, sillä koehenkilöiden määrä eri ikäryhmissä on varsin pieni. Alustavat tulokset S-koartikulaation akustisista ilmentymistä näyttäisivät tukevan käsitystä, että S-artikulaation omaksuminen lapsilla kehittyy asteittain. Tulosten perusteella voidaan tietyin varauksin myös kysyä, kehittyvätkö ja ilmenevätkö aikuismaiset foneettiset S-piirteet suomenkielisillä lapsilla ehkä oletettua myöhemmin.

keywords: koartikulaatio, äidinkielen kehitys, frikatiivien akustiikka, spektrin momentit

Suomen kielen alveolaarisen frikatiivin soinnillisuudesta

Paula Laine
Michael O'Dell
Tampereen yliopisto

Tämän tutkimuksen tarkoituksena on selvittää äänneympäristön vaikutusta alveolaarisen frikatiivin esiintymiseen soinnillisena suomen kielessä. Tarkoituksena on myös selvittää miten frikatiivien soinnillisuutta voitaisiin mitata. Aineistona käytän aiemmin nauhoitettua äänitettä, josta segmentoin *s*-äänteiden esiintymät erilaisissa ympäristöissä.

Suomen kielen frikatiivien soinnillisuutta ei ole juurikaan tutkittu. Sen sijaan jonkin verran alveolaaristen frikatiivien soinnillisuutta on tutkittu espanjan kielessä, jossa – kuten suomessakaan – äänten soinnillisuus ei ole distinktiivinen piirre (ks. esim. Hualde & Prieto 2014: 109). Yleensä on katsottu suomen kielessä esiintyvän ainoastaan soinniton alveolaarinen frikatiivi (VISK § 3). Aineiston perusteella näyttää kuitenkin selvältä, että myös soinnillista varianttia tästä äänteestä käytetään.

Selvää on myös, että soinnillisuudessa ei ole kyse joko–tai-ilmiöstä, vaan kyseessä on jatkumo täysin soinnillisen ja soinnittoman äänten välillä. Soinnin voimakkuus vaihtelee frikatiiveissa sekä siten, että täysin soinnillisten ja täysin soinnittomien äänteiden välille jää suuri joukko tapauksia, joissa sointi ei ole täysin selvää. Toisaalta soinnin määrä myös vaihtelee sen mukaan, esiintyykö sointia koko äänten ajan, vai vain osassa äännettä.

Luultavasti myös muut tekijät vaikuttavat sointiin äänneympäristön lisäksi. Kyse saattaa olla myös puhujakohtaisesta ilmiöstä siten, että toiset puhujat käyttävät soinnillista äännettä joskus ja toiset eivät koskaan. Puhujakohtaisen variaation tutkimiseen tarvittaisiin kuitenkin perehtymistä suurempaan eri puhujilta koottuun aineistoon, eikä siihen tämän tutkimuksen puitteissa valitettavasti ole mahdollisuuksia.

Lähteet

HUALDE, JOSÉ IGNACIO – PRIETO, PILAR 2014: Lenition on intervocalic alveolar fricatives in Catalan and Spanish. – *Phonetica* 71 s.109-127.
VISK = HAKULINEN, AULI – VILKUNA, MARIA – KORHONEN, RIITTA – KOIVISTO, VESA – HEINONEN, TARJA RIITTA – ALHO, IRJA 2004: *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura. Verkkoversio, viitattu 27.1.2016. Saatavissa: <http://scripta.kotus.fi/visk>

Unto K. Laine

Signaalinkäsittelyn ja akustiikan laitos, Aalto-yliopisto

Puhesignaalin käsittelyä aika-alueessa permutaatioita soveltaen

Avainsanat: puheanalyysi, puheenmallinnus, aika-alue, ordinal patterns, permutaatiot

Signaalien analyysissä vallitseva paradigma on, että kaikki keskeiset operaatiot tehdään taajuusalueessa. Puheen osalta Mel-spektri ja -kepstri ovat tästä ehdottomasti tärkeimmät esimerkit. Vain jotkut harvat menetelmät, jotka perustuvat signaalin ajallisen aaltomuodon suoraan prosessointiin, ovat menestyneet. Autokorrelaatioiden laskenta on yksi niistä. Myös lineaarista ennustamista (lineaariprediktio) voidaan pitää aika-alueen menetelmänä, joskin siihen liittyy hyvin läheisesti tulosten taajuustulkinta, mallien tuottaessa myös spektrikuvauksen. Myös joitain perustaajuuden ilmaisuun käytettyjä metodeja on kehitetty toimiviksi aika-alueessa, kuten AMDF (*Average Magnitude Difference Function*).

TKK:n Akustiikan laboratoriossa (nykyisin Aalto-yliopiston Signaalinkäsittelyn ja Akustiikan laitos) permutaatioihin liittyvä tutkimusta on tehty jo yli vuosikymmen. Permutaatiomuunnos toimii lyhyen signaalin osan, tyypillisesti 5–7 näytettä, aaltomuodon kvantisoijana ja koodaajana. Jatko-operaatioina voidaan laatia tilastoja näiden koodisanojen esiintymisistä, erityisesti siitä, miten ne liittyvät toisiinsa. Näin laajemmassa aikaikkunassa oleva aaltomuoto saadaan kuvattua tilastollisesti.

Puheen heikosti ohjatun segmentoinnin lisäksi (Fonetiikan päivät 2015) permutaatioihin perustuvia menetelmiä on alustavasti tutkittu ja sovellettu perustaajuuden ja formanttitaajuuksien analyysiin, yhden ääninäytteen opetusdatan käyttöön sanojen luokituksessa (*one-shot learning*), vokaalien mallinnukseen ja luokitukseen, sekä moneen muuhun käyttökohteeseen. Saadut tulokset viittaavat siihen, että menetelmä tuottaa taajuusalueen menetelmiin nähden keskimäärin yhtäläisiä tuloksia. Tämä on ymmärrettävää, koska taajuusmuunnokset eivät voi tuota mitään lisäinformaatiota alkuperäiseen aika-aluesignaaliin nähden. Toisaalta permutaatiopohjaiset aika-aluekuvaukset nostavat paremmin esiin tietynlaisia signaalirakenteita, ja täydentävät siten olemassa olevia aika- ja taajuusalueen analyysitekniikoita.

Tässä puheenvuorossa on tarkoitus antaa metodin matemaattisista perusteista tiivis periaatekuvaus, sekä valaista esimerkein mitä näillä uusilla metodeilla on mahdollista saavuttaa puheetutkimuksessa. Tulokset ovat olleet lupaavia, mutta yhä nämä uudet metodit etsivät niiden optimaalista soveltamistapaa myös puheen maailmassa.

**Helsingin yliopiston fonetiikan opiskelijoiden tieteenalakäsityksen
muodostuminen**

Mona Lehtinen & Heini Kallio

Helsingin yliopisto, fonetiikka

mona.lehtinen@helsinki.fi, heini.h.kallio@helsinki.fi

Fonetiikan historia Suomessa ulottuu 1800-luvulle; Knut Hugo Pipping perusti ensimmäisen fonetiikan laboratorion Helsinkiin tultuaan nimitetyksi fonetiikan opettajaksi 1891. Tuolloinen laboratorio oli osa fysiikan laitosta. Historiansa aikana fonetiikan oppiaine on muiden tavoin käynyt läpi useita muutoksia, ja nyt olemme jälleen uudessa tilanteessa: Iso pyörä -koulutus uudistuksessa muuttuu tutkintorakenne, ja lisäksi Helsingin yliopiston fonetiikan oppiaine siirtyy käyttäytymistieteellisestä tiedekunnasta takaisin humanistiseen. Taitekohdat ovat osuvia ydinainesanalyysin ja historiakatsauksen paikkoja. Tutkimalla aikaisempia tieteenalakuvauksia sekä haastatteleamalla fonetiikan opiskelijoita pyrimme antamaan hyödyllisiä näkökulmia uudistusten keskelle ja helpottamaan aineen profiloitumista uudessa ympäristössä.

Haastattelimme Helsingin yliopiston fonetiikan pääaineopiskelijoita syksyllä 2015 ja keväällä 2016 (N=10). Kysymysten avulla selvitimme mm. opiskelijoiden tieteenalakäsitystä sekä opiskeluorientaatiota. Tässä työssä keskitymme erityisesti fonetiikan opiskelijoiden tieteenalakäsityksen muotoutumiseen opintojen aikana. Haastateltavia oli kaikista opintojen vaiheista ensimmäiseltä vuosikurssilta maisterivaiheeseen. Vastaukset olivat vaihtelevia ja niissä korostui fonetiikan monitieteinen luonne; sidoksia nähtiin niin luonnontieteisiin, käyttäytymistieteisiin kuin kliinisiin ja humanistisiin aloihin. Vastauksista voidaan nähdä tieteenalakäsityksen muuttuvan sekä rakentuvan opintojen edetessä. Lisäksi opiskeluorientaatiolla ja tieteenalakäsityksellä vaikuttaa olevan yhteys.

Asiasanat: fonetiikka, tieteenalakäsitys, opiskelu

The effect of the artificial palate on the acoustics of speech

Anton Malmi (University of Tartu)

Einar Meister (Tallinn University of Technology, Institute of Cybernetics at TUT)

Pärtel Lippus (University of Tartu)

Keywords: acoustic phonetics, articulatory phonetics, EPG

This paper studies the effect of the artificial palate of the electropalatography system (EPG3) on the quality and quantity measurements of the palatalized and unpalatalized intervocalic /l/. The test subject was asked to read isolated test words in three quantity degrees, with and without EPG3 pseudo-palate. The lateral was in the context of the vowels /a, e, i, u/. Only the acoustic recordings were analyzed for this study. Preliminary results show that the average F₂ frequency of the lateral was approximately 100-200 Hz higher and the duration showed the tendency to be longer when the pseudo-palate was used. The quality of the vowels preceding the lateral was also affected: the F₂ frequency of /a, e, u/ was approximately 30-100 Hz higher, but the F₂ frequency of /i/ was approximately 100-200 Hz lower when produced with the palate. Further analysis of formant frequencies and durations of the vowels and the lateral confirms that the artificial palate does affect the quality and quantity of speech segments suggesting that the acoustic data collected in the EPG experiments are not directly comparable with the data from the acoustic studies.

Suspect speech database for automatic speaker comparison

Authors: Tuija Niemi, Juha Koponen

Forensic Laboratory, Finland

Keywords: Speaker comparison, Forensics

Abstract

Analysis of speech is used in forensics as a biometric identifier of a person. Traditionally the speech comparison is done by an expert and an official statement is given. Novel signal analysis technology provides also technical tools for using a semi-automatic system that allows comparing unknown speech to the known (suspects) speech samples in a database. Especially in telephone interception cases there might be several speakers in each case not known to the investigators.

During years 2009-2013 a system consisting of a speech database and an automatic speaker comparison program was constructed to serve whole Police organization as well as the Border Guard and the Customs in Finland. A speech sample is taken from a suspect as part of the registration process. In the future, every unknown speech sample necessary to be studied will be compared to the speech samples in the database, from which the system will give a list of potential speaker candidates. However, there are number of legal constraints which have to be taken into account when selecting the unknown speech samples.

This method does not necessarily require an official request from the police investigator nor an official expert opinion from the laboratory. This is a new way of working in Forensic Laboratory which has until now given only official expert opinions to the Police and other public authorities in Finland. If an expert opinion is needed for the court, it is still possible to do a more precise research of speech samples.

Konsonanttiyhtymässä olevan lateraalin akustinen variaatio suomessa

Michael O'Dell & Tommi Nieminen
Tampereen yliopisto & Itä-Suomen yliopisto

Suomen kielessä lateraali // on ainakin vokaalien välissä usein hyvin lyhyt ja ballistinen. Aikaisemmin (Nieminen & O'Dell 2010) olemme esittäneet arvion, että sama kenties pätee myös konsonanttiyhtymän alussa, mikä selittäisi epenteettisen vokaalin eli švaan taipumusta näissä konteksteissa: Artikulatorisen fonologian mallin mukaisesti sen helposti aiheuttaisi variaatio lateraalin ja seuraavan konsonantin eleen suhteellisessa ajoituksessa. Yritämme luonnehtia konsonanttiyhtymän //:n akustista variaatiota juoksevassa spontaanissa puheessa sekä //:n dynamiikan kannalta että ajallisessa suhteessa seuraavaan konsonanttiin. Vertaamme tuloksia myös aineistossamme olevaan tietoon puhujien alueellisesta taustasta.

asiasanat: lateraali, švaa, suomen kieli

Nieminen, T. & O'Dell, M. 2010. Epenthetic Vowels in Finnish: An Articulatory Phonology Approach. Esitelmä 26.8.2010 Joensuussa (24th Scandinavian Conference of Linguistics)

Identifying the First Structure to Move in Ultrasound Videos

Pertti Palo, Sonja Schaeffler and James M. Scobbie

Clinical Audiology, Speech and Language (CASL) Research Centre,
Queen Margaret University

We study the effect that phonetic onset of an utterance has on acoustic and articulatory reaction times (RTs). An acoustic study by Rastle et al. (2005) shows that the place and manner of the first consonant or consonant cluster in a target affects acoustic RT. An articulatory study by Kawamoto et al. (2008) shows that the same effect is not present in articulatory reaction time of the lips.

In this study we seek to identify the articulatory structures that move first in each context. We do this by developing an automated analysis method for ultrasound tongue videos. Ultrasound data in its raw form consists of scanlines (Figure 1). The analysis is based on calculating a Euclidean distance measure for each scanline of the raw ultrasound data. Similar methods have been employed as more holistic measures by e.g. McMillan and Corley (2010), Raeesy et al. (2011), and Drake et al. (2013). The location of the first movement is identified by first finding the movement onset for each individual scanline and then taking the minimum of the corresponding time stamps thus identifying the time and position of the first movement. The results from analysing delayed naming data from a group of participants is the subject of another submission to Fonetiikan Päivät 2016 by the same authors.

Keywords: Articulatory RT, automatic analysis

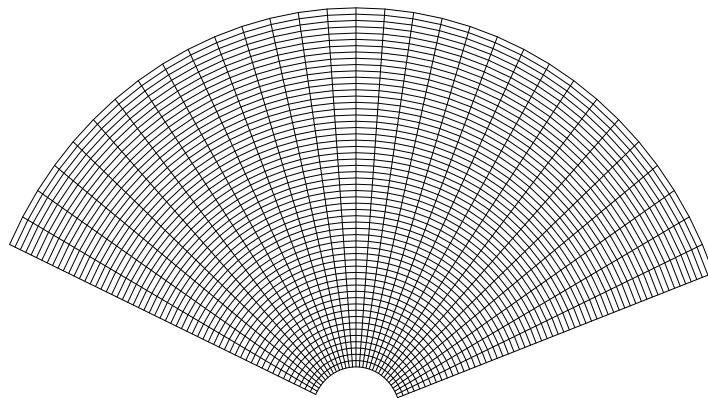


Figure 1: Ultrasound fan consisting of 31 radial scanlines and 64 pixels along each scanline.

References

- Drake, E., Schaeffler, S., and Corley, M. (2013). Articulatory evidence for the involvement of the speech production system in the generation of predictions during comprehension. In *Architectures and Mechanisms for Language Processing (AMLaP)*, Marseille.
- Kawamoto, A. H., Liu, Q., Mura, K., and Sanchez, A. (2008). Articulatory preparation in the delayed naming task. *Journal of Memory and Language*, 58(2):347 – 365.
- McMillan, C. T. and Corley, M. (2010). Cascading influences on the production of speech: Evidence from articulation. *Cognition*, 117(3):243 – 260.

- Raeesy, Z., Baghai-Ravary, L., and Coleman, J. (2011). Parametrising degree of articulator movement from dynamic MRI data. In *12th Interspeech*, pages 2853 – 2856.
- Rastle, K., Harrington, J. M., Croot, K. P., and Coltheart, M. (2005). Characterizing the motor execution stage of speech production: Consonantal effects on delayed naming latency and onset duration. *Journal of Experimental Psychology: Human Perception and Performance*, 31(5):1083 – 1095.

Patterns of Articulatory Activation in Delayed Naming

Pertti Palo, Sonja Schaeffler and James M. Scobbie

Clinical Audiology, Speech and Language (CASL) Research Centre,
Queen Margaret University

We study the effect that phonetic onset of an utterance has on acoustic and articulatory reaction times (RTs). An acoustic study by Rastle et al. (2005) shows that the place and manner of the first consonant or consonant cluster in a target affects acoustic RT. An articulatory study by Kawamoto et al. (2008) shows that the same effect is not present in articulatory reaction time of the lips.

In this study we answer the question: Can we predict the location of the first tongue movement with the phonological identity of the utterance onset? Knowing the phonological features of the initial consonantal segment one might think that it is easy to predict that certain structures are going to move first. However, from a coarticulatory point of view, we know that what happens first might actually be related to the production of the second, third or even fifth segment.

To answer our research question we analyse tongue ultrasound data from a delayed naming experiment using 58 lexical English words. The words all have a high lexical frequency and are of the types /VC/, /CVC/, /CCVC/, and /CCCVC/. The consonants and vowels are varied systematically in the corpus. At the time of writing data has been successfully recorded for three participants. Data from these and possibly further participants will be presented at the conference. The analysis technique is the subject of another submission to Fonetiikan Päivät 2016 by the same authors.

Keywords: Articulatory RT, first movement

References

- Kawamoto, A. H., Liu, Q., Mura, K., and Sanchez, A. (2008). Articulatory preparation in the delayed naming task. *Journal of Memory and Language*, 58(2):347 – 365.
- Rastle, K., Harrington, J. M., Croot, K. P., and Coltheart, M. (2005). Characterizing the motor execution stage of speech production: Consonantal effects on delayed naming latency and onset duration. *Journal of Experimental Psychology: Human Perception and Performance*, 31(5):1083 – 1095.

Esikielellisen tavun tutkimus sonoriteettiin perustuvana rytmisenä yksikkönä

Okko Räsänen

Signaalinkäsittelyn ja akustiikan laitos, Aalto-yliopisto

Tavu on yleisesti tunnettu puhutun ja kirjoitetun kielen yksikkö ja sillä on myös keskeinen rooli useissa lapsen varhaisen kielenoppimisen teorioissa ja tutkimuksissa. Huolimatta tavun yleismaailmallisuudesta ja usein selkeästä havaittavuudesta, aiempi tutkimus ei ole kyennyt yksiselitteisesti määrittelemään tavun akustista tai foneettista luonnetta. Lisäksi on epäselvää miten esimerkiksi fonologisiin esitysmuotoihin perustuvat tavun määritelmät vastaavat esikielellisen lapsen puheenhavaitsemisen esitysmuotoja, eli tilanteessa jossa lapselle ei ole vielä kehittynyt kykyä tunnistaa kielen äänteitä tai sanoja. Ymmärtääksemme paremmin tällaisen esikielellisen tavun luonnetta, tässä tutkimuksessa pyrittiin selvittämään minkälaisia tavun kaltaisia yksiköitä puhesignaalista voidaan havaita käyttämällä pelkästään puheen sonoriteettiverhokäyrän rytmisiä rakenteita, ts. informaatiota jonka tiedetään olevan myös kielen äännejärjestelmää ja sanoja vasta opettelevien lasten käytössä. Tutkimuksessa on aluksi määritetty tavu sonoriteettiverhokäyrän kahden minimin välisenä äänekäänä yksikkönä, eli rakenteena joka toteuttaa akustisessa mielessä tavurakenteen sonoriteettihierarkiaperiaatteen (eng. sonority hierarchy principle). Tämän jälkeen kehitettiin kuulojärjestelmää mukaileva laskennallinen malli tällaisten tavunkaltaisten yksiköiden segmentointiin jatkuvasta puhesignaalista. Mallin tuottaman tavusegmentoinnin tulosta vertailtiin tutkitun puhemateriaalin foneettiseen annotaatioon englannin-, eestin- ja suomenkielillä. Tulokset osoittavat että yleismaailmallinen rytmisen segmentointi toimii samankaltaisesti kaikilla tutkituilla kielillä, ja että sonoriteettiverhokäyrän rytmiset vihjeet ovat potentiaalisesti erittäin hyödyllisiä tavu- ja sanarajojen löytämisessä kielenoppimisen alkuvaiheessa. Toisaalta tulokset osoittavat että tutkimuksessa määritetty akustinen tavunkaltainen yksikkö ei vastaa täsmällisesti fonologista tavutasoa, ja siksi oletusta vauvojen kyvystä jäsentää puhetta täsmällisiksi sanarajoja mukailevaksi tavujonoiksi tulisi jatkossa lähestyä varovaisen kriittisesti.

Asiasanat: puheen havaitseminen; rytmi; tavu

A study of metrics for models of phone-like units obtained by a time domain process called rank predictor

Shreyas Seshadri and Unto K. Laine

Department of Signal Processing and Acoustics, Aalto University

Keywords: speech analysis, speech segmentation, sub-word models, permutations

Abstract

As a part of our studies on new time-domain methods based on permutation mathematics we have developed an unsupervised system for automatic segmentation of speech utterances to phone-like units. The system utilizes, the so-called *rank predictor*. Here an N sample (typically 20-30) long frame is considered, where the permutation code of the first five signal samples, p , and the rank of the last sample in the whole N -frame, r , are considered. When the N -frame is moved over a signal segment $\{p, r\}$ pairs of permutation are collected at each position of the indices N -frame to form a rank predictor matrix (RPM). The RPM is used to model the actual signal segment. The RPMs reveal the structural similarity between different speech segments.

The segmentation process utilizes a multiphase process and the information collected in the RPMs to give estimates for the segment boundaries. Initial results show that these are typically located at the phone boundaries (manuscript submitted for publication).

In this study we continue this line of research by analyzing Finnish sentences (ACORNS Caregiver) including the same keyword in each sentence, but at different locations. The data is initially segmented to phone-like units using the above-mentioned process and modeled with the RPMs. These segment models should reveal the locations of similar segments at least in the simple case of a single talker. We test this hypothesis using different similarity measures for the RPMs and also model the co-occurrence of segments to learn word models.

Bayesian non-parametric clustering of speech into phone-like categories

Shreyas Seshadri, Ulpu Remes, and Okko Räsänen

Department of Signal Processing and Acoustics, Aalto University, Finland

Keywords – Speech modeling, Bayesian clustering, unsupervised learning.

Abstract

Modeling and segmenting of sub-word components (such as phone-like units) from continuous speech is an important problem in speech research. The most successful methods in the field are supervised algorithms that train specific models for these units based on annotated data. However, there are many languages in the world for which these annotations are not readily available. Development of unsupervised techniques, called clustering, for this task could therefore prove very useful. Studying the clustering problem can also potentially contribute to language acquisition research by providing an understanding of how much can be learnt purely from speech data without any other sources of supporting information.

The models of speech typically used in unsupervised speech systems are Gaussian mixture models (GMMs) and Hidden Markov models (HMMs). The problem with these models is that they require the user to define a specific number of clusters/states to find in the data, which, in general, is not known.

In contrast to these standard approaches, the present paper investigates the clustering problem using a class of unsupervised algorithms called Bayesian non-parametric methods. These methods allow us to cluster the data without explicitly specifying the number of clusters in advance. Instead, they find the number of clusters required based on the statistical properties of the data.

The study reports results from two methods of Bayesian non-parametric clustering namely the Dirichlet process Gaussian Mixture models (DPGMMs) and Hierarchical Dirichlet Process Hidden Markov models (HDPHMMs) on speech data. These results are then compared to traditional supervised clustering methods.

Foneettisen harjoittelun vaikutus alakouluikäisten lasten vieraan kielen tuottamiseen ja havaitsemiseen

Taimi, L.^{1,2}, Alku, P.³, Peltola, M.S.^{1,2}

¹Fonetiikka ja Learning, Age and Bilingualism -laboratorio, Turun yliopisto

²Signaalinkäsittelyn ja akustiikan laitos, Aalto-yliopisto

Posteri

Tämän tutkimuksen tarkoituksena oli selvittää miten alakouluikäiset suomenkieliset lapset oppivat havaitsemaan ja tuottamaan heille vieraan äänteen, joka ei erota merkitystä suomen kielessä. Tutkimukseen osallistui 12 äidinkieleltään suomenkielistä lasta (8 tyttöä, keski-ikä 8;11). Ärsykkeinä tutkimuksessa oli semi-syntetisoidut epäsanat /tɘ:ti/ ja /ty:ti/, joista /tɘ:ti/ toimi kohdesanana. Kukin lapsi osallistui kahtena peräkkäisenä päivänä äänteiden erottelu- ja tuottokokeeseen, sekä aivojen sähköistä toimintaa mittaavaan EEG-mittaukseen. Lisäksi koehenkilöt harjoittelivat sanoja kuuntele-ja-toista -menetelmän avulla yhteensä 240 toiston verran. EEG-mittauksessa kiinnitettiin erityisesti huomiota poikkeavuusnegatiivisuusvasteeseen (Mismatch Negativity, MMN), joka syntyy silloin jos tutkimuksessa käytetyt sanat pystytään erottamaan toisistaan ja kohdesanalle on syntynyt aivoihin muistijälki.

Tutkimuksessa käytetty kohdesana valittiin niin, että kielen oppimisen teorioiden (esim. Speech Learning Model, SLM, Flege 1987 ja Perceptual Assimilation Model, PAM, Best & Strange 1992) valossa se aiheuttaisi oppimisvaikeuksia oppijoille. Kohdesanan /tɘ:ti/ ensimmäinen vokaali ei ole fonologinen suomen kielessä, mutta sijaitsee akustisesti lähellä suomen kielen vokaaleja /y/ ja /u/, joten sen havaitseminen ja tuottaminen omaksi kategoriaksi tuottaa vaikeuksia suomenkielisille puhujille.

Tutkimuksen päätuloksena oli, että harjoittelun myötä lasten MMN-vaste aikaistui ja kasvoi tilastollisesti merkitsevästi. Myös tuottokokeen tulokset muuttuivat kohdesanan osalta mallin suuntaan. Erottelukokeessa sen sijaan ei huomattu tilastollisesti merkitseviä muutoksia. Esitietoinen prosessointi siis voimistui ja nopeutui harjoittelun myötä ja tuottaminen muuttui. Tulokset viittaavat mahdollisesti siihen, että esitietoinen havaitseminen edeltää tietoisien erottelun oppimista lapsilla ja psykofysiologisten ja tuottomuutosten saavuttamiseksi lyhytkin foneettinen harjoittelu on riittävä.

Asiasanat: vieraan kielen oppiminen, foneettinen harjoittelu, lapset

The Prosody of Emergency Calls

Lauri Tavi
University of Eastern Finland

The purpose of this study is to examine emergency calls using methods from experimental and forensic phonetics. The dissertation study will consist of three different parts: One investigation aims to find prosodic differences between authentic and hoax emergency calls, especially focusing on callers' voice features of psychological stress. The second investigation will deal with the prosodic features of modal sentence types such as interrogative and directive clauses in real emergency calls. The third research topic is the interaction between caller and duty officer from the prosodic point of view; how does the voice of the duty officer affect the voice quality of the caller?

Furthermore, the study aims to refine speaker classification and recognition methods. The study is currently in the phase of research material collection. The research material will consist of authentic recordings of emergency calls. From the forensic point of view, it is important to investigate telephone speech data from real life situations. After collection of the data, it will be acoustically analyzed with Praat software and differences in acoustical measurements will be determined statistically. The study will investigate prosodic and acoustic features such as F0, jitter, shimmer, speech tempo, formants, spectral balance and intensity.

Key words

Prosodic features, emergency calls

APUA NIMIKOINTIIN AUTOMAATTISESTA KOHDISTAMISESTA?

Lauri Tavi, Stefan Werner

Itä-Suomen yliopisto

Koneen säätöön *SILK*-hankkeen (Suomen itäpuolisten lähialueiden kielikorpukset) yhteydessä kokeiltiin erilaisia automaattisia sana-/foneemi-/foonikohdistusjärjestelmiä ja niiden mukautuvuutta uusiin murteisiin. Esitelmässämme kerromme käyttö- ja adaptointikokeuksistamme seuraavien järjestelmien kanssa: *AaltoASR*, *SPPAS*, *Prosodylab-Aligner* ja *Praatalign*.

*Prosodylab-Aligner*in ja *AaltoASR*:n käytön ja mukautuvuuden tutkimisessa aineistona käytettiin tunti äänenlaadultaan kohtalaista entisen Raja-Karjalan Ilomantsin murretta. *Prosodylab-Aligner*iin loimme kyseiselle murrevariantille akustisen mallin, joka toimii kohtalaisesti. Kohdistaminen on varsinkin sanatasolla tarkkaa, mutta tuloksissa esiintyy tasaisesti myös virheitä. Erityisesti äännitteen sisältämät luonnolliset häiriöäänäiset kuten yskäisyt ja kolinat aiheuttavat kohdistukseen epätarkkuutta. *Prosodylab-Aligner* ei myöskään hyväksy foneemien käytössä UTF-formaattia, mikä olisi ollut aineiston kannalta käytännöllisempää.

AaltoASR on ainakin toistaiseksi tarkoitettu yleiskieliselle suomelle, joten tietyn murrevariantin käyttö aineistona oli puhtaasti kokeellista. *AaltoASR* ei myöskään tarjoa loppukäyttäjälle mahdollisuutta lisätä kohdistimeen uutta kielimuotoa. Kohdistustarkkuus osoittautui hyvin vaihtelevaksi, ja jotkin yleiskielisetkin sanat tulivat tavatuksi äänteittäin. Toisaalta osa aineistosta kohdistui tarkasti murteellisesta muodosta huolimatta. Kuitenkin samoin kuin *Prosodylab-Aligner*issa myös *AaltoASR*:ssä häiriöäänäiset aiheuttivat virheitä.

Alustavien tulosten mukaan automaattisten kohdistimien tarkkuus on vielä suhteellisen kaukana manuaalisesta annotoinnista silloin, kun uuden murteen aineiston äänenlaatu ei ole laadukasta ja aineistona oleva puhe on luonnollista – eli puhekielistä ja erilaisia häiriöäänäisiä sisältävää. *Prosodylab-Aligner*in sekä *AaltoASR*:n tulokset uuden murteen kohdistamisessa osoittautuivat toisaalta sen verran tarkaksi ja manuaalista annotointia selvästi nopeammaksi, että automaattista nimikointia on jo nyt mahdollista käyttää ainakin alustavana nimikointina.

SPPAS on periaatteessa loppukäyttäjän muokattavissa, mutta jo oletuskielten segmentointi on melko epätarkkaa. Myös *Praatalign* sallii uusien kielten lisäämistä, mutta tältä osin työemme on vielä kesken.

Asiasanat: *murre, automaattinen kohdistaminen, puheentunnistus*

Keywords: *dialect, automatic alignment, ASR*

The durational structure of Estonian CVCV-words

Pire Teras, Eva Liina Asu, Pärtel Lippus, Karl Pajusalu

This paper investigates variation in the realisation of Estonian CVCV-words. The short unstressed vowel in the second syllable (V2) of such words is longer than the short stressed vowel in the first syllable (V1). Various studies have addressed duration ratios of CVCV-words in Estonian as well as in related Finnic languages (Finnish, Livonian, Soikkola Ingrian), but very few have treated this issue from the point of view of regional variation or accentuation conditions.

Earlier research reveals considerable differences in the duration ratios of CVCV-words in Standard Estonian. In read speech, the V1/V2 ratio varies between 0.52 and 0.83 (e.g. Lehiste 1968; Krull 1992), and in spontaneous speech between 0.64 and 0.82 (e.g. Krull 1998; Lippus et al. 2013). It has been demonstrated that the ratio is smaller in prepausal position than elsewhere (Krull 1997). The V1/V2 ratio can also be influenced by accentuation in that it is smaller in accented than in deaccented words (Lippus et al. 2013).

A study investigating materials other than Standard Estonian (Sepp 1980) has shown that the V1/V2 ratio is the smallest (0.55) in Võru South Estonian and the biggest in North-East Coastal dialects (0.8).

In our paper we analyse duration ratios of CVCV-words in read speech, where test-words occur both in accented and deaccented phrase-medial position. The aim of this research is to study in closer detail the role of speakers' regional background and accentuation in the variation of V1/V2 in the Estonian language area. The results are compared to work done on other Finnic languages.

Keywords: vowel length, Estonian, quantity

Kazoo-pillin vaikutukset suomenkielisten oppijoiden ruotsin puherytmin oppimiseen

Elina Tergujeff, Mikko Kuronen & Maria Kautonen

Jyväskylän yliopisto

Tässä tutkimuksessa selvitettiin kazoo-pillillä harjoittelun vaikutuksia suomenkielisten oppijoiden ruotsin puherytmin oppimiseen. Kazoo-pilli on yksinkertainen kalvosoitin, jota on suositeltu prosodian opetukseen ja harjoitteluun, koska se tarjoaa oppijalle segmenteistä riisutun ääntämismallin kiinnittäen huomion prosodiaan (Gilbert 1978). Pillin vaikutuksen tutkimiseksi järjestettiin opetuskokeilu, jossa pitkälle edistyneet suomenkieliset ruotsinopiskelijat osallistuivat ruotsin ääntämistä käsittelevälle kurssille. Kurssi koostui kahdestatoista 90 minuutin kontaktiopetustapaamisesta, jotka sisälsivät monipuolista ääntämisen opetusta (teoriaa ja käytännön harjoituksia). Opiskelijat opiskelivat kahdessa ryhmässä, joista ensimmäisen opiskelijat (n=12) muodostivat kontrolliryhmän. Toisen ryhmän opiskelijat (n=17) muodostivat interventior ryhmän, jonka opetus poikkesi kontrolliryhmän opetuksesta vain siinä, että he käyttivät kazoo-pillejä prosodiaa käsittelevissä käytännön harjoituksissa.

Kaikki opiskelijat testattiin ennen ja jälkeen opetuksen käyttäen sekä lukupuhuntaa että vapaan tuottamisen tehtäviä. Suoritusten analysoinnissa keskityttiin erityisesti rytmiin (esim. tavun kesto ja lausepaino) ja intonaatioon (intonaatiojakson kesto). Kyseiset parametrit eroavat suomessa (Iivonen 2009) ja ruotsissa (Vihanta ym. 1990) ja ovat siksi haastavia mutta tärkeitä oppimistavoitteita suomenkielisille ruotsinoppijoille. Oppijoiden kehittymisen arvioimiseksi heidän suorituksiaan ennen ja jälkeen kurssin verrattiin syntyperäisiltä ruotsinpuhujilta kerättyyn identtiseen materiaaliin. Tutkimuksen tulokset kertovat, oliko kazoo-pillillä harjoittelulla vaikutusta suomenkielisten oppijoiden ruotsin puherytmin oppimiseen, jolloin saadaan myös yleisempää tietoa kazoo-pillin hyödyllisyydestä toisen/vieraan kielen prosodian oppimisessa.

Asiasanat: puherytmi, intonaatio, opetuskokeilu

Lähteet:

Gilbert, J. 1978. GADGETS: Some non-verbal tools for teaching pronunciation. *TESL Reporter*, Winter 1978.

Iivonen, A. 2009. Finnish Sentence Accent and Intonation. Teoksessa: de Silva, V. & Ullakonoja, R. (toim.), *Phonetics of Russian and Finnish*. Frankfurt am Main: Peter Lang.

Vihanta, V.V. & Leinonen, K. & Pitkänen, A.J. 1990. On Rhythmic Features in Finland-Swedish and Sweden-Swedish. Teoksessa: K. Wiik & I. Raimo (toim.) *Nordic Prosody V. Papers from a Symposium*. Turku: University of Turku. 325–350.

Puheenkierrätystutkimus suomessa

Joonas Vakkilainen

Suomen kielen rytmityypistä ei ole yksimielisyyttä, ja kaikkia tyyppejä onkin ehdotettu. Rytmityypeistä yleisestikään ei ole aiemmin ollut selvää näyttöä, sillä eri rytmityyppejä edustavissa kielissä tehdyt foneettiset mittaukset eivät ole paljastaneet eroja kielten välillä. Aiemmat huonot tulokset ovat johtuneet menetelmien puutteesta. Uudempi menetelmä puheenkierrätys (speech cycling) pakottaa puheen säännölliseen rytmiin, mikä mahdollistaa rytmin hierarkian tarkastelun. Puheenkierrätyksessä puhuja synkronoi puheensa ääniärsykyksiin. Puheenkierrätystä ei ole aiemmin sovellettu suomeen. Tämän tutkimuksen tarkoituksena on kokeilla, soveltuuko puheenkierrätys suomeen ja millaista näyttöä se antaa suomen rytmityypin puolesta. Aineisto koostuu fraaseista, joiden keskimmäistä sanaa varioidaan tavu- ja moraaluvun puolesta. Tämän tarkoituksena on analysoida, miten fraasin osien sijainti ja kestot muuttuvat. Mittaukset osoittavat fraasien pysyvän kestoiltaan melko stabiileina, vaikka varioitavan sanan kesto kasvaisikin, mikä kertonee jonkinasteisesta tahtiajoitteisuudesta. Tämä tutkimus on osa tekeillä olevaa pro gradu -tutkielmaa, joka käsittelee suomen typologista rytmityyppiä käyttäen puheenkierrätystä menetelmänään.

asiasanat: rytmitypologia, puheenkierrätys (speech cycling), suomen kieli

Kissankello vai kissan kello – Suomen yhdyssanojen prosodisista ominaisuuksista

Päivi Virkkunen & Martti Vainio

Tavoitteet: Tutkimuksen tavoitteena oli selvittää suomen kielen yhdyssanojen tuottamista. Suomen kielen prosodiaa ja kontrastiivisen lausepainon vaikutusta puheentuottoon on tutkittu runsaasti, mutta yhdyssanojen tuottamista ei ole tutkittu tämänkaltaisilla koeasetelmilla. Tässä työssä tutkittiin, oliko puheen prosodisissa ominaisuuksissa eroja yhdyssanojen ja vertailuaineistona olleiden sanaliittojen välillä. Lisäksi tarkasteltiin lausepainon aiheuttaman prosodisen fokuksen vaikutusta yhdyssanojen tuottamiseen.

Menetelmät: Tutkimus oli kvantitatiivinen ja se koostui puheentuottokokeesta ja tilastollisesta tarkasteluosuudesta. Puheentuottokokeessa samanasaiset yhdyssanat ja sanaliitot (esimerkiksi kissankello ja kissan kello) toistuivat kolmessa eri prosodisessa fokus konditiossa. Nämä olivat painottoman lause, kontrastiivinen lausepainotus ensimmäisellä kohdesanalla ja kontrastiivinen lausepainotus toisella kohdesanalla. 28 koehenkilöä (14 naista ja 14 miestä) puhui yhteensä 1680 lausetta, jotka segmentoitiin ja annotoitiin. Sanoista mitattuja perustaajuuden, intensiteetin ja keston arvoja verrattiin sanatyypin välillä. Sanatyypin ja fokuksen vaikutusta akustisiin mittaustuloksiin selvitettiin tilastollisesti toistettujen mittausten varianssianalyysillä.

Tulokset ja johtopäätökset: Tutkimuksessa saatiin uutta tietoa siitä, millaisia strategioita puhujat käyttävät tuottaessaan yhdyssanoja eri painotusasemissa. Tulokset osoittivat, että yhdyssanat ja sanaliitot tuotettiin eri tavoilla laajassa fokus konditiossa. Sanatyypillä oli tilastollisesti merkitsevä vaikutus sekä perustaajuuden ja intensiteetin että tavujen keston tuottoon. Tuloksista nähtiin, että puhujat käsittelivät yhdyssanaa yhtenä ja sanaliittoa kahtena erillisenä sanana. Kontrastiivinen lausepainotus loi akustisiin arvoihin voimakkaita muutoksia, jotka peittivät alleen sanatyypin tuottamat erot.

Kontekstilla on suuri vaikutus siihen, miten kuulija ymmärtää tuotetut sanat, mutta käytössä on myös akustisia havaintovihjeitä sanatyypin erottelemiseksi toisistaan. Tutkimustulos luo vankan pohjan jatkotutkimukselle. Alustavan analyysin perusteella tulokset viittaavat merkittäviin eroihin sukupuolten välillä. Tutkimusta on helppo laajentaa myös esimerkiksi suomen eri murteisiin tai S2-puheeseen.

asiasanat: prosodia, puheentuotto, yhdyssanat

Puolipituussanojen prosodiikka suomen itämurteissa

Riikka Ylitalo

Suomen (C)V.CV(X)-rakenteisiin sanoihin liittyy prosodisia erityispiirteitä. Syynä tähän voi olla se, että näiden sanojen ensitavu on lyhyt, jolloin painotuksen kannalta tärkeä toinen mora sijaitsee vasta sanan toisessa tavussa. Tunnetuin (C)V.CV(X)-sanoihin liittyvistä erityispiirteistä on ns. puolipidennys: monissa suomen kielimuodoissa lyhyen ensitavun jälkeisen toisen tavun yksinäisvokaali on sanan ensitavun vokaalia huomattavasti pidempikestoinen. Puolipidennyksettömissä suomen varieteeteissa lyhyen ensitavun jälkeinen toisen tavun yksinäisvokaali on lyhempi tai suunnilleen yhtä pitkä kuin V₁. Kutsun sanoja, joiden rakenne mahdollistaa puolipidennyksen toteutumisen niissä, puolipituussanoiksi.

Lausepainollisiin puolipituussanoihin liittyy myös perustaajuuden eroja suomen varieteettien välillä. Lausepainoa toteuttavan perustaajuuden nousu–lasku-kuvion korkeimman kohdan on todettu sijaitsevan lounaismurteiden puolipituussanoissa vasta toisessa tavussa. Tilanne on samanlainen Kuopion seudulla Pohjois-Savon murrealueella, ainakin, jos sanan C₂ on soinnillinen. Yleensä perustaajuuden huippu sijoittuu suomessa sanan ensimmäiseen tavuun.

Tässä tutkimuksessa tarkastellaan puolipituussanojen prosodiikkaa seuraavista itämurteiden alueeseen kuuluvista kaupungeista ja niiden ympäristökunnista kotoisin olevien henkilöiden puheessa: Kajaani, Joensuu, Mikkeli, Jyväskylä ja Lappeenranta.

Kultakin informantilta äänitettävä materiaali koostuu yhteensä 50:stä kontrastiivisesti lausepainolliseen asemaan sijoitetusta CV.CV-rakenteisesta kohdesanasta kehysvirkkeineen. 30 kohdesanan C₂ on soinnillinen, 20 kohdesanassa C₂ on soinniton. Informantit lukevat kohdesanat kehysvirkkeineen ääneen tietokoneen ruudulta äänitysstudioissa. Heidän puhetuotoksensa tallennetaan suoraan tietokoneen kovalevylle (44,1 kHz, 16-bittiä). Tallennetut kohdesanat analysoidaan akustisesti äänenkäsittelyohjelman avulla. Äännesegmenttien kestot ja perustaajuushuipun etäisyys sanan alusta mitataan millisekunnin kymmenesosan tarkkuudella sekä perustaajuusarvot hertsin tarkkuudella viidestä kohtaa kohdesanojen ensitavua, viidestä kohtaa niiden toista tavua sekä perustaajuuden huippukohdasta. Mittaustulokset analysoidaan tilastollisesti, jolloin selviää, miten

voimakkaana puolipidennysilmiö toteutuu eri itämurteissa, ja millainen on tarkasteltavien varieteettien lausepainoa toteuttavan perustaajuuskäyrän muoto.

ISBN 978-952-62-1479-5

Oulun yliopisto

HuTK, Fonetikka