



SUOMEN PITKIEN PAINOTTOMIEN VOKAALIEN HAVAITSEMINEN SPONTAANISSA JA LUETUSSA PUHEESSA

Joonas Vakkilainen, Tampereen yliopisto

Michael O'Dell, Tampereen yliopisto

Suomessa pitkien ja lyhyiden äänteiden välinen ero on fonologisesti distinktiivinen, eli se erottaa merkityksiä. Spontaanissa puheessa lyhyiden ja pitkien vokaalien väliset erot ovat pienempiä kuin lukupuhunnassa (de Silva ym., 2003). Suomessa vokaalin kesto on kuulijalle pääasiallinen vihje sen kvantiteetista, mutta painollisissa tavuissa on havaittu myös sävelkulkuun liittyviä eroja (Järvikivi ym., 2007). Tässä tutkimuksessa tarkastellaan painottomien jälkitavujen pitkien vokaalien havaitsemista spontaanissa puheessa. Tutkimuskysymyksenä on, vaikuttaako vokaalin fysikaalinen kesto spontaanissa puheessa pituuden havaitsemiseen eri tavalla kuin lukupuhunnassa. Tutkimuksessa tehtiin havaintokokeita, joissa koehenkilöt kuuluivat muokattuja havaintoärsyksiä. Ärsykkeet poimittiin spontaanista puheesta ja saman henkilön lukemista tämän spontaanissa puheessa esiintyneistä sanoista. Ärsykkeiden sisäiset kestoerot poistettiin muokkaamalla spontaani-lukupuhunta-ärsykepareista kestoiltaan identtiset. Jokainen alkuperäinen sana sisälsi pitkän jälkitavun vokaalin, jonka kestosta muokattiin neljä eri varianttia. Koehenkilöt vastasivat, kuuluivatko he lyhyt- vai pitkävokaalisen sanan. Tulosten perusteella spontaanissa puheessa sanan lopussa olevan vokaalin ei tarvitse olla yhtä pitkä kuin lukupuhunnassa tullakseen havaituksi pitkäksi. Spontaanin puheen ärsykkeet saivat keskimäärin enemmän pitkävokaalisia vastauksia kuin lukupuhuntaärsykkeet. Taustatekijöillä (F0, F1, F2) ei ollut selvää vaikutusta havaintoihin, joskaan niitä ei voida sulkea pois tämän tutkimuksen perusteella.

Avainsanat: havaitseminen, kesto, kvantiteetti, spontaani puhe, vokaali

1 JOHDANTO

Suomi on kvantiteettikieli, jossa merkitystä erottavat pitkät ja lyhyet äänneet esiintyvät vapaasti riippumatta tavun painosta ja melko vapaasti riippumatta vierekkäisten äänneiden pituuksista. Kvantiteetti tarkoittaa järjestelmää, jossa äänneen fysikaalista kestoa käytetään merkitystä erottavassa tehtävässä.

Äänneen fonologinen pituus on subjektiivinen kokemus siitä, kuinka paljon aikaa äänneen tuottamiseen on tarvittu. (Lehtonen, 1970.) Tämä tutkimus käsittelee Lehtosen (1970, s. 16) määrittelemää subjektiivista eli havaittavaa kvantiteettiä (*perceptual quantity*). Tutkimuksessamme keskitymme jälki-

Kirjoittajien yhteystiedot:

Joonas Vakkilainen

joonas.vakkilainen@tuni.fi

Michael O'Dell

michael.odell@tuni.fi

tavujen vokaalikvantiteetin havaitsemisen eroihin spontaanin ja luetun puheen välillä. Huomion kohteena on näin ollen kvantiteetin havaitseminen eli kuulijan kokemus fonologisesta pituudesta eikä pelkän foneettisen keston havaitseminen. Spontaanina puhetta koskevia tutkimuksia on tehty suomesta vain vähän (ks. de Silva, Iivonen, Bondarko & Pols, 2003; Anttila, 2009; Lennes, 2009).

Suomea voidaan kutsua autonomiseksi kvantiteettikieleksi, sillä kestojen foneeminen käyttö ei sisällä juurikaan laatueroja eikä rajoitu tiettyihin tavarakenteisiin (Lehtonen, 1970, s. 86). Fonologinen pituus perustuu suomessa äänteen foneettiseen keston. Pitkien ja lyhyiden vokaalien suhteen on yleensä havaittu olevan noin 1 : 2. Isei-Jaakkolan (2004, s. 43–44) tutkimuksessa /ɑ/:n ja /ɑɑ/:n suhde oli keskimäärin 1 : 2,3. Lyhyen vokaalin kesto oli keskimäärin 79,3 ms ja pitkän 180,5 ms. Lehtosen (1970, s. 89) mukaan pitkän ja lyhyen vokaalin suhde ensimmäisessä tavussa on keskimäärin 1 : 2,2 ja kestit 70 ms ja 153 ms, hitaassa tempossa jopa 1 : 3, mutta jälkitavujen suuren variaation vuoksi niistä ei ole hänen mukaansa mielekästä esittää keskiarvoa. Lehtosen (1970, s. 90) tutkimuksessa jälkitavuisissa pienin suhde (1 : 1,38) on tavarakenteiden CVCV ja CVCVV välillä ja suurin (1 : 3,19) tavarakenteiden CVVCCV ja CVVCCVV välillä. Typologisesti katsoen kvantiteettioppositio 1 : 2 on kielissä hyvin tyypillinen (Meister, Werner & Meister, 2011, s. 1363).

Tutkimuksissa yleensä esitetyt keistosuhteet perustuvat huolelliseen lukupuhuntaan. Spontaanista puheesta tehdyt mittaukset antavat toisenlaisia tuloksia. De Silva, Iivonen, Bondarko ja Pols (2003) vertailivat suomea, venäjää ja hollantia ja havaitsivat, että suomessa segmenttien kestit lyhenevät spontaanissa puheessa toisin kuin venäjässä ja hollannissa, joissa ne kasvavat. Myös pitkien ja lyhyiden vokaalien ero pieneni suomessa spontaanissa

puheessa: lukupuhunnassa pitkät vokaalit olivat 52 % pidempiä kuin lyhyet, spontaanissa puheessa vain 20 %. Lyhenemisen ja suhteen pienenemisen lisäksi Lennes (2009) havaitsi, että spontaanissa puheessa ja jälkitavuisissa vokaalien laadussa oli suurempaa variaatiota kuin lukupuhunnassa ja painollisissa tavuisissa. Vastaavasti Arai ja Warner (1999) huomasivat japanissa, että spontaanissa puheessa sanan kestoa ei voi tarkasti ennustaa morien lukumäärästä, sillä morien kestit vaihtelevat enemmän kuin lukupuhunnassa.

Myös prosodininen konteksti vaikuttaa kestitoihin. Pitkät vokaalit muuttuvat eniten. Lehtosen (1970, s. 146) tutkimuksessa lausepainollisessa asemassa lyhyet vokaalit olivat keskimäärin 15 % pidempiä kuin lausepainottomassa asemassa, kun pitkät vokaalit olivat 33 % pidempiä. Toisessa tavussa pitkät vokaalit pitenevät selvästi enemmän (50 %) kuin ensitavujen pitkät vokaalit (15 %).

Kvantiteetissa on kyse suhteellisuudesta. Kuulija ei perusta havaintoaan äänteen pituudesta sen absoluuttiseen fysikaaliseen keston vaan suhteuttaa keston muihin lähellä oleviin kestitoihin. Kvantiteettikielen puhujien tekemä luokittelu fonologisesti lyhyeen ja pitkään äänteeseen perustuu äänteen sijoittamiseen jompaankumpaan kategoriaan, mikä tarkoittaa, että puhuja saattaa verrata pituusluokkia niiden prototyyppeihin, jotka mahdollisesti vaativat kestojen vertailua laajempiin yksiköihin kuten tavuihin (Nenonen, 2001, s. 14–16). Kategorian rajalla on epävarma *havainnon raja-alue* (tai *kriittinen alue* tai *siirtymäalue*, vrt. Lehtonen, 1969, s. 369), jonka sisällä havainto pituudesta voi olla kumpi tahansa. Isei-Jaakkola (2010) vertaili suomen, japanin ja tšekin havainnon raja-alueita (*perceptual boundary range*) eri tavarakenteissa ja prosodisissa konteksteissa ja havaitsi, että suomen raja-alue oli suurin ja että prosodininen konteksti sekä vokaalienvälinen konsonantti vaikuttivat suomeen enemmän kuin kahteen

muuhun kieleen. Järvikiven, Aallon, Alangon ja Vainion (2007) tutkimuksessa painottoman jälkitavun lyheneminen kasvatti ensimmäisen tavun havaitsemista pitkäksi.

Lehtosen (1969, s. 369–370) mukaan kuulija havaitsee kestoeroja parhaiten kategorian rajalla, kun taas kategorian sisällä kestoja voidaan muuttaa enemmän kuulijan huomauttamatta. Kategorioiden olemassaolo näkyi myös siinä, että Ylisen, Shestakovan, Alun ja Huotisen (2005, s. 327–328) havaintokokeessa suomenkielisten puhujien reaktioaika kategorian rajalla olevissa ärsykkeissä oli pidempi kuin kategorian sisällä olevissa, ja vastaavasti venäjänkielisillä kuulijoilla tällaista eroa ei ollut.

Kvantiteetin suhteellisuus tulee esiin erityisesti spontaanissa puheessa. Spontaanissa puheessa kestojen variaatio on suurempaa kuin lukupuhunnassa. Erityisesti painottomat sanat redusoituvat ja menettävät kestoaan (Lennes, Alarotu & Vainio, 2001, s. 128). Spontaanin puheen havaitseminen on kontekstisidonnaisempaa kuin lukupuhunnan. Arai (1999) havaitsi, että spontaanissa japanissa koehenkilöiden kuulemien morien määrä väheni, kun sanoja kuuntelutettiin ilman kontekstia, mutta kuulijat kuulivat kaikki morat, kun lausekontekstia oli tarpeeksi.

Keston lisäksi kvantiteetti voi perustua myös äänteen sisäisiin ominaisuuksiin. Vaikka lyhyiden ja pitkien vokaalien väliset laadulliset erot on tiedetty ennestäänkin, aiemmin niillä ei ole katsottu olevan merkitystä kvantiteetin havaitsemisessa (Lehtonen, 1970, s. 21–22). Uudemmat tutkimukset ovat kuitenkin löytäneet laadullisten tekijöiden mahdollisen yhteyden havaintoon. Meisterin ym. (2011) mukaan vokaalin laadulliset ominaisuudet määrittävät kvantiteetin kategorian rajan, sillä heidän tutkimuksessaan eri vokaalit saivat eri rajan. O'Dell (2003, s. 57–79) tutki saman vokaalin eri pituuksien välisten laatuerojen vaikutusta. Hänen havaintokokeessaan lyhytvokaalisesta sanasta tehty kestosarja sai enem-

män lyhytvokaalisia vastauksia kuin samat kestoportaat sisältävä pitkävokaalisesta sanasta tehty kestosarja. Voi ajatella, että kvantiteetti perustuuakin kuulijan kannalta vaikutelmaan artikulaatioiden ajoituksesta (sisäinen ajoitus, *intrinsic timing*) eikä absoluuttisiin kestoihin (ulkoinen ajoitus, *extrinsic timing*). Tähän vaikutelmaan voivat potentiaalisesti liittyä kaikki puhujan ajoituksesta johtuvat seikat (O'Dell, 2003, s. 9–11, 111–112).

Myös prosodisilla tekijöillä, esimerkiksi sävelkululla, on havaittu olevan merkitystä (O'Dell, 2003; O'Dell, Nieminen & Vakkilainen, 2019; Vihanta, 1988). Järvikivi ym. (2007) huomasivat, että sävelkulku vaikuttaa painollisen vokaalin havaitsemiseen pitkäksi. Vainion, Järvikiven, Aallon ja Sunin (2010) tutkimus vahvisti, että kaksimoraisilla painollisilla tavuilla on usein laskeva sävelkulku, mikä tarkoittaa, että kuulija saattaa käyttää muitakin vihjeitä kuin kestoja.

Tutkimuksemme tavoitteena on vertailla kvantiteetin havaitsemista jälkitavuisissa spontaanin puheen ja lukupuhunnan välillä, sillä spontaanin puheen tutkimuksessa erityisesti havaitseminen on vielä vähän tutkittu alue (Warner, 2012, s. 633). Koska spontaanissa puheessa kestot varioivat enemmän kuin lukupuhunnassa, oletamme, että myös kvantiteetin havaitseminen eli kestojen fonologinen tulkinta spontaanissa puheessa eroaa kvantiteetin havaitsemisesta huolellisessa lukupuhunnassa. Kestojen variaatio on myös jälkitavuisissa suurta, eivätkä aiemmat tutkimukset ole raportoineet niissä prosodisia tekijöitä, jotka vaikuttaisivat havaintoon. Tutkimuksemme tarkastelee ainoastaan jälkitavun kestoja, ja siksi kontekstin vaikutus on neutraloitu. Hypoteesimme on, että jälkitavujen vokaalien ei tarvitse spontaanissa puheessa olla yhtä pitkiä kuin lukupuhunnassa tullakseen havaituiksi pitkiksi. Näin ollen pelkkä kesto ei voi toimia vihjeenä pituudesta, koska variaatio painaisi sen alleen.

2 AINEISTO JA MENETELMÄT

Tutkimuksen spontaanin puheen ärsykkeet saatiin puhuttamalla suomenkielistä hämäläismurteista miespuolista henkilöä. Tässä tutkimuksessa käytettiin yhden puhujan puhetta, jotta puhujan vaikutus koehenkilöiden vastauksiin voitiin poissulkea. Luonnollisen ja spontaanin puheen aikaan saamiseksi henkilö sai puhua vapaasti itseään kiinnostavista asioista, minkä lisäksi tehtiin karttatehtävä, jossa puhuja kuvaili reittiohjeet kartalle merkityn kahden paikan välille. Spontaani puhe täytyi valmiin puheaineiston sijaan tuottaa uutena, jotta oli mahdollista saada lukupuhuntaärsykkeet samalta puhujalta. Puhujaa ei ohjailtu käyttämään tiettyjä sanoja, jotta materiaali olisi mahdollisimman autenttista spontaania puhetta. Lukupuhuntaa varten spontaanin puheen aineistosta poimittiin yksittäisiä sanoja, jotka puhuja luki ääneen (vrt. Warner, 2012, s. 628). Havaintokoetta varten valittiin fonologisesti pitkän loppuvokaalin sisältäviä sanoja, joista ensimmäisessä havaintokokeessa käytettiin kolmea ja toisessa kahdeksaa. Spontaanista aineistosta poimittiin sanat, jotka olivat rytmillisesti tiiviisti fraasin osana (esim. ei taukoa molemmilla puolilla), jotta ne vastaisivat todellisen puheen redusoituneita muotoja ja olisivat kaukana sitaattimuodosta (Lennes ym., 2001, s. 128), ja joiden äänenlaatu oli riittävän hyvä. Lukupuhuntasanat olivat puolestaan yksittäin luettuina luonnollisesti painollisia, minkä ansiosta sanaparien välille saatiin mahdollisimman suuri ero. Lisäksi pitkävokaalisilla ja vastaavilla lyhytvokaalisilla muodoilla on eri merkitys puhekielessä. Valitut sanat olivat *aikuisikää*, *pelästyy* ja *riippuu* ensimmäisessä havaintokokeessa ja *isoo*, *jatkuu*, *kattoo*, *koiraa*, *nähtävää*, *saksaa*, *sanoo* ja *seuraa* toisessa kokeessa. Ne esiintyivät spontaanissa aineistossa seuraavissa lauseissa:

- *en oikeestaan ennen aikuisikää oo kauheesti ollu kissojen kanssa tekemisissä*
- *jos se jotenkin pelästyy erityisesti niin ei se heti ryntää siitä*
- *se riippuu tietysti sitten montako kertaa mää sen ruokin siitä että ketä on kotona*
- *kaks tai kolme isoo koiraa, ne voi olla vaikka kaukasianpaimenkoiria*
- *kävellään edelleen sitä katua joka siitä sillan jälkeen jatkuu*
- *täytyy kattoo semmonen reitti että tästä löytää oudompi perille*
- *mutta siellä on paljon nähtävää ja se on jotenkin se kaupungin ilmapiiri on miellyttävä*
- *siellä on ei saksaa puhuvan ennen käymättömän ihmisen helppo liikkua*
- *sen joen nimee mää en osaa sanoo mutta sitten on pieni kävelysilta*
- *kun seuraa sitä radan vartta*

Ärsykeiksi valittujen sanojen pitkä loppuvokaali oli odotuksenmukaisesti suhteellisesti lyhyempi spontaanissa puheessa (lukuunottamatta sanaa *seuraa*, jonka loppuvokaalin suhteellinen kesto oli lähes identtinen). Kaikki sanat olivat kokonaiskestoltaan pidempiä lukupuhunnassa (vrt. taulukko 1).

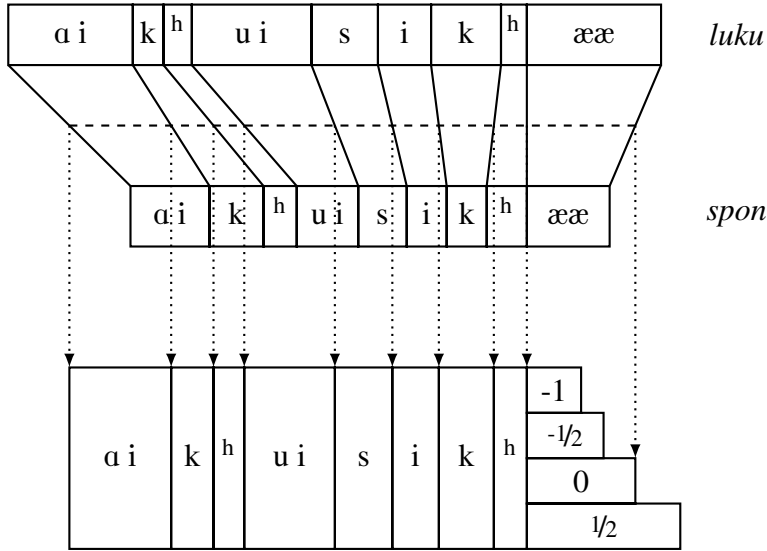
TAULUKKO 1. Alkuperäisten sanojen kokonaiskestot ja loppuvokaalin osuudet.

SANA	kokonaiskesto			loppuvokaalin osuus		
	luku (ms)	spon (ms)	suhde	luku (%)	spon (%)	suhde
<i>aikuisikää</i>	863	634	1,36	20,6	17,3	1,19
<i>pelästyy*</i>	668	482	1,39	29,0	21,2	1,37
<i>riippuu</i>	615	315	1,95	31,6	24,0	1,32
<i>isoo</i>	571	343	1,66	60,9	43,4	1,40
<i>jatkuu</i>	621	539	1,15	40,7	35,1	1,16
<i>kattoo*</i>	723	325	2,22	37,3	20,3	1,84
<i>koiraa*</i>	657	411	1,60	40,2	29,9	1,34
<i>nähtävää</i>	704	439	1,60	28,4	19,6	1,45
<i>saksaa</i>	686	426	1,61	39,9	25,8	1,55
<i>sanoo</i>	486	351	1,38	39,5	21,9	1,80
<i>seuraa</i>	618	372	1,66	35,4	35,5	1,00

*mittaus alkaa eksploosiosta

Spontaani-lukupuhunta-sanaparit muokattiin vastaamaan toisiaan siten, että sanat jaettiin segmentteihin, joiden kestoista muokattiin sanaparin keskiarvo (ks. kuva 1). Muokkaus tehtiin Praatin *Overlap-add* manipulation -työkalun avulla (Boersma & Weenink, 2020). Tällä tavalla sanan sisäisten kestojen vaikutus neutralisoiitiin. Muokkaamalla sanaparin kumpaakin osapuolta vältettiin luonnollisuuden tuoma vinouma, joka olisi voinut vaikuttaa tuloksiin, mikäli vain osa ärsykkeistä olisi ollut muokattuja ja osa segmentteiltään alkuperäisiä. Segmenttien muokkauksen ansiosta myöskään sanan alkuperäiset sisäiset kestot eivät aiheuta vinoumaa loppuvokaalin havaitsemisessa. Sanan loppuvokaalista tehtiin neljä varianttia: kaksi lyhyempää, yksi alkuperäinen eli ensimmäisen muokkauksen mukainen, ja yksi pidempi. Ärsykkeiden vokaalien kestot erosivat toisistaan tasavälein logaritmisella asteikolla siten, että

lyhyimmän vokaalin kesto oli puolet alkuperäisen pitkän (keskiarvotetun) vokaalin kestoista. Verrattuna alkuperäisen pitkän vokaalin kestoön ärsykkeiden vokaalien kestot olivat $1/2$, $\sqrt{1/2}$, 1 , $\sqrt{2}$ (eli log₂-asteikolla -1 , $-1/2$, 0 , $1/2$; ks. kuva 1). Täten saman sanan spontaani-lukupuhunta-ärsykeparit olivat sekä kokonaiskestoltaan että segmenttensä kestoilta samanpituiset. Ärsykkeissä ei muokattu laatua tai prosodiaa. Näin ollen ärsykeitä oli yhtä sanaa kohden kahdeksan eli neljä eri kestovarianttia spontaanista puheesta ja lukupuhunnasta poimitusta sanasta ja täten ärsykkeen muuttujat olivat TYYPPI (*spon* = spontaani vs. *luku* = luettu), SANA (*aikuisikää*, *pelästyy*, *riippuu*, jne) ja loppuvokaalin KESTO (neljä astetta).



KUVA 1. Ärsykkeiden kestojen muokkaus (esimerkkinä aikuisikää).

Ensimmäiseen havaintokokeeseen osallistui yhteensä 11 koehenkilöä (3 naista, 7 miestä, 1 muu, iät 18–55 vuotta). Koehenkilöt olivat suomenkielisiä. Koehenkilöiden taustatiedoiksi kysyttiin lisäksi asuinpaikkakunnat murretaustan selvittämiseksi. Ensimmäisessä havaintokokeessa käytettiin sanoja *aikuisikää*, *pelästyy* ja *riippuu*. Jokainen yksittäinen ärsyke toistui kokeessa viisi kertaa, joten koehenkilöt kuulivat yhteensä 120 ärsykettä. Ärsykkeet soitettiin satunnaisjärjestyksessä, kuitenkin niin, että sama ärsyke ei koskaan esiintynyt kaksi kertaa peräkkäin, ja sama sana korkeintaan kolme kertaa peräkkäin. Koehenkilöt vastasivat lomakkeeseen alleviivaamalla, kuulivatko he *aikuisikä* vai *aikuisikää*, *pelästyy* vai *pelästyy*, *riippuu* vai *riippuu*. Vastaamatta ei saanut jättää. Ennen koetta koehenkilöt kuulivat esimerkkiärsykejonoa, joka ei vastannut kokeen ärsykkeiden järjestystä. Ärsykkeiden välinen viive oli verraten lyhyt, vain 800 ms, jotta leksikaalinen prosessointi jäisi mahdollisimman vähäiseksi. Vaikka tahti oli ripeä ja vaati keskittymistä, kukaan koehenkilöistä ei

valittanut tai maininnut tästä kokeen jälkeisissä kommentteissa. Koehenkilöt osallistuivat kokeeseen kahdessa ryhmässä. Vastausten analyysistä jätettiin pois koehenkilö, joka ei kyennyt vastaamaan suurimpaan osaan kohdistusta. Hän raportoi kokeen jälkeen suullisesti, että hän ei ollut kuullut ärsykkeissä kestoeroa, mikä vaikeutti kokeeseen vastaamista.

Toinen havaintokoe täydensi aineistoa, ja siinä käytettiin sanoja *isoo*, *jatkuu*, *kattoo*, *koiraa*, *nähtävää*, *saksaa*, *sanoo* ja *seuraa*. Toiseen kokeeseen osallistui 16 koehenkilöä (8 naista, 8 miestä, iät 20–62 vuotta). Ensimmäisen kokeen osallistujista neljä osallistui myös toiseen kokeeseen. Toinen havaintokoe toteutettiin Praatin *Experiment*-ohjelmalla, jossa koehenkilöt valitsivat ruudulla näkyvistä kahdesta vaihtoehdosta (esim. *koira* ja *koiraa*) sen, jonka he mielestään kuulivat. Ensimmäisestä kokeesta poiketen koehenkilöt saivat edetä omaan tahtiin. Vastauksen jälkeen koehenkilö kuuli palauteäänen, jonka jälkeen uusi ärsyke kuului vaihtelevan pituisen tauon jälkeen (n. 275–650 ms) siten, että saman sanan koh-

dalla tauko oli aina sama. Yhteensä yhdeksän vastausta hylättiin, koska koehenkilö reagoi vahingossa ennen ärsykkeen akustista alkua. Hyväksytyjen vastausten reaktioajan mediaani ärsykkeen lopusta laskien oli n. 796 ms.

2.1 Tilastollinen käsittely

Havaintojen analyysiin käytettiin bayesilaista hierarkkista logistista regressiota (Gelman ja Hill, 2007). Bayesilaisessa analyysissä ei lasketa datan todennäköisyyttä jonkin nollahypoteesin mukaan vaan mallin parametrien eli ennalta tuntemattomien suureiden posterioritodennäköisyyttä, joka tarkoittaa todennäköisyyttä sen jälkeen, kun data on otettu huomioon. Lopputulos näyttää parametrien todennäköisimmät arvot mutta myös jäljelle jäävän epävarmuuden (posteriorijakauman) tämän arvion ympärillä.

Laskennat suoritettiin R:ssä (versio 4.0.3, R Core Team, 2020) käyttämällä JAGS-ohjelmaa (versio 4.3.0, Plummer, 2017) ja *rjags*-pakettia (versio 4-10, Plummer, 2019). Tilastollisessa mallissa (ks. yhtälöt 1–9) pitkän vokaalin havaintotodennäköisyys $p_{ijk}(x)$ on KESTO x :n logistinen funktio, joka riippuu kahdesta parametrista, μ_{ijk} ja β_{ijk} (yhtälö 1). Tässä x on ärsykkeen vokaalin suhteellinen kesto \log_2 -asteikolla (ts. $x = -1, -1/2, 0, 1/2$), i kertoo TYYPPI-muuttujan tason ($i = \text{luku, spon}$), j kertoo SANA:n ($j = \text{aikuisikä, pelästyy, riippuu, jne}$) ja k koehenkilön ($k = 1, 2, \dots, 22$). Yksi näistä parametreista on rajaparametri μ_{ijk} , joka kertoo millä KESTO-arvolla x pitkän vokaalin havaintotodennäköisyys $p_{ijk}(x)$ saavuttaa 50 % (eli logit-asteikolla 0; $\text{logit}(1/2) = 0$). Toinen parametri on kulmakerroin β_{ijk} , joka kertoo, kuinka jyrkästi todennäköisyyskäyrä nousee, eli kuinka nopeasti havainto vaihtuu lyhyestä vokaalista pitkään. Tämän parametrin käänteisarvo ($1/\beta_{ijk}$) kertoo rajan leveyden.

Koko aineiston keskimääräisen rajapara-

metrin (μ_0) priorijakaumana käytettiin normaalijakaumaa, jonka keskiarvo on $-1/4$, joka vastaa ärsykesarjan keskipistettä, ja varianssi 10 (eli keskihajonta on $\sqrt{10} \approx 3,162$; näin suuri varianssi ei juurikaan rajoita mahdollisia arvoja) (ks. yhtälö 4). Keskimääräisen kulmaparametrin (β_0) priorina käytettiin normaalijakaumaa, jonka keskiarvo on 1 ja varianssi 10 (ks. yhtälö 7). Tämä priorijakauman keskiarvo vastaa rajan leveyttä, jossa pitkän vokaalin todennäköisyyden nousu 20 % \rightarrow 80 % kattaa noin kaksi ärsykesarjan askelta.

Kummankin parametrien annettiin vaihdella myös koehenkilön (KH), sanan¹ (SANA) ja tyyppin (TYYPPI) mukaan (yhtälöt 2 ja 3) siten, että jokainen efekti ja niiden interaktiot jakautuivat normaalijakauman mukaan, jonka keskiarvo on 0 (yhtälöt 5 ja 8; Näissä yhtälöissä X tarkoittaa TYYPPI, SANA, KH, TYYPPI \times SANA, TYYPPI \times KH, SANA \times KH TAI TYYPPI \times SANA \times KH, ja ”...” osoittaa sopivat alaindeksit). Yhden efektin parametrien keskiarvon on oltava 0, koska parametrit kertovat hierarkkisessa mallissa poikkeaman efektin yläluokan keskiarvosta. Näiden normaalijakaumien keskihajonnat toimivat hyperparametreina, jotka kertovat kuinka suuri merkitys kyseisellä efektillä on. Mitä suurempi hajonta on, sitä enemmän sen vaihtelut vaikuttavat. Kaikkien keskihajontaparametrien priorina käytettiin puoli-Cauchy-jakaumaa, jonka skaala on 1 (yhtälöt 6 ja 9). Tämä priorijakauma on tähän tutkimukseen sopiva, koska ryhmiä on verrattain vähän. Käytetty skaala valittiin, koska voidaan olettaa, että efektien kaikki keskihajonnat ovat paljon pienempiä kuin 4 (Gelman, 2006).

¹ Koetyypin mahdollinen vaikutus (ensimmäinen ja toinen koe toteutettiin hiukan eri tavalla) sisältyy SANA-efektiin. Kokeen ja sanan vaikutuksia ei ole mahdollista erottaa, koska kahdessa kokeessa oli eri sanoja.

$$\text{logit}(p_{ijk}(x)) = \beta_{ijk} \cdot (x - \mu_{ijk}) \quad (1)$$

$$\begin{aligned} \mu_{ijk} = & \mu_0 + \mu_i^{\text{TYYPPI}} + \mu_j^{\text{SANA}} + \mu_k^{\text{KH}} \\ & + \mu_{i,j}^{\text{TYYPPI} \times \text{SANA}} + \mu_{i,k}^{\text{TYYPPI} \times \text{KH}} + \mu_{j,k}^{\text{SANA} \times \text{KH}} \\ & + \mu_{i,j,k}^{\text{TYYPPI} \times \text{SANA} \times \text{KH}} \end{aligned} \quad (2)$$

$$\begin{aligned} \ln(\beta_{ijk}) = & \beta_0 + \beta_i^{\text{TYYPPI}} + \beta_j^{\text{SANA}} + \beta_k^{\text{KH}} \\ & + \beta_{ij}^{\text{TYYPPI} \times \text{SANA}} + \beta_{ik}^{\text{TYYPPI} \times \text{KH}} + \beta_{jk}^{\text{SANA} \times \text{KH}} \\ & + \beta_{ijk}^{\text{TYYPPI} \times \text{SANA} \times \text{KH}} \end{aligned} \quad (3)$$

$$\mu_0 \sim \text{Norm}(-1/4, \sqrt{10}) \quad (4)$$

$$\mu_{\dots}^X \sim \text{Norm}(0, \sigma_{\mu}^X) \quad (5)$$

$$\sigma_{\mu}^X \sim \text{puoli-Cauchy}(1) \quad (6)$$

$$\beta_0 \sim \text{Norm}(1, \sqrt{10}) \quad (7)$$

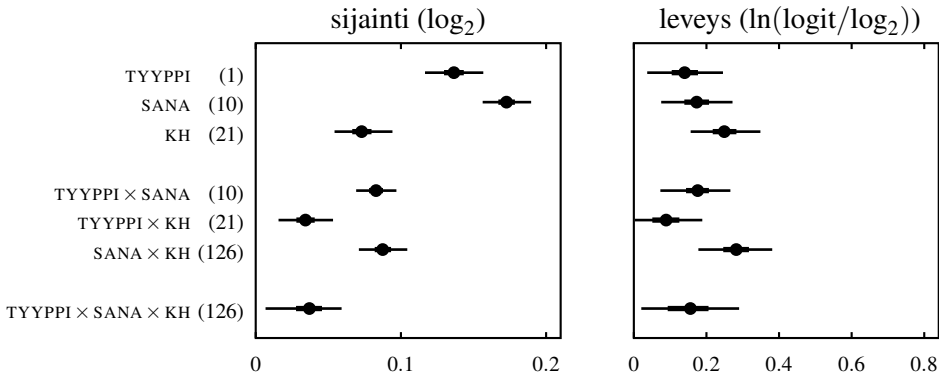
$$\beta_{\dots}^X \sim \text{Norm}(0, \sigma_{\beta}^X) \quad (8)$$

$$\sigma_{\beta}^X \sim \text{puoli-Cauchy}(1) \quad (9)$$

3 TULOKSET

Eri efektien arvioidut keskihajontaparametrit esitetään visuaalisesti kuvassa 2, jossa parametrien posteriorijakaumat on pelkistetty ns. toukkajanoiksi: ohut viiva näyttää posteriorijakauman 95% CI (*Credible Interval*, bayesilainen luottamusväli), paksumpi viiva näyttää 50% CI, ja pallo osoittaa jakauman mediaania. Janan pituus siis osoittaa kuinka paljon epävarmuutta jää sen jälkeen, kun data

on otettu huomioon. Efektin keskihajonta kiteyttää efektin vaikutusta. Efektin parametrien keskihajonta voi olla nolla ainoastaan, jos ne ovat kaikki keskenään samat, joten mitä selvemmin posteriorijakauma (ja pelkistävä jana) on erillään nolasta, sitä varmemmin efektillä on vaikutusta. Janan sijainti kertoo toisaalta myös vaikutuksen arvioidusta suuruudesta: parametrien isompi keskihajonta tarkoittaa, että efektillä on suurempi vaikutus.



KUVA 2. Rajan sijaintiin ja leveyteen vaikuttavien efektien arvioidut keskihajonnat (posteriorijakaumien mediaani, 50% ja 95% CI, suluissa efektin vapausasteet).

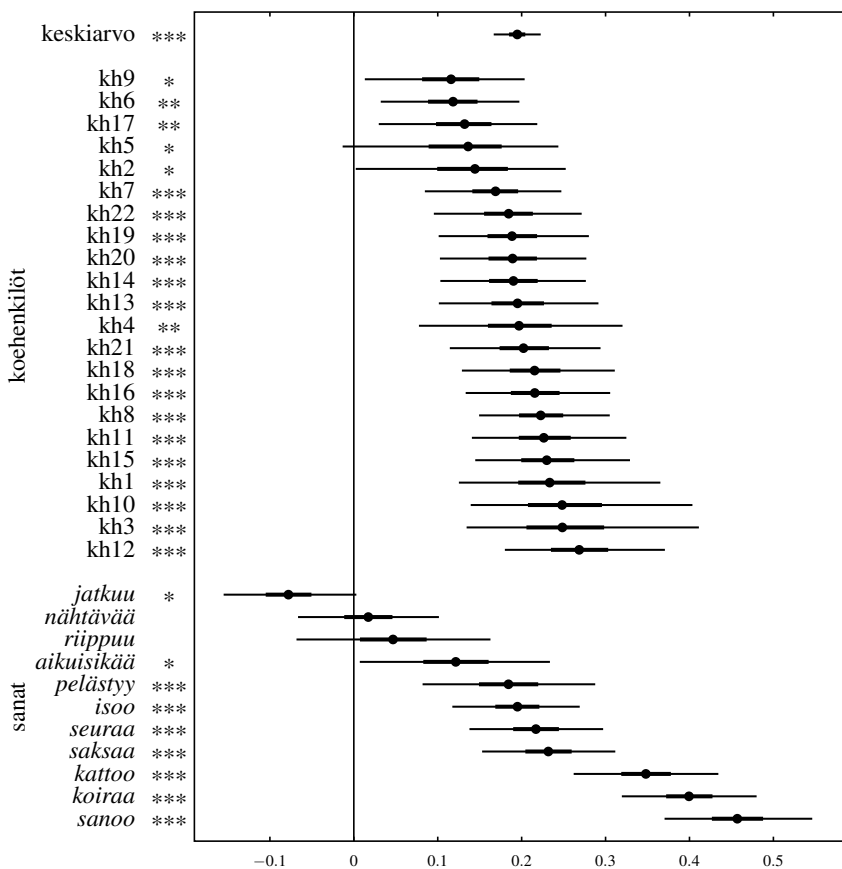
3.1 Rajan sijainti

Kuvasta 2 nähdään, että kaikilla pääefekteillä (TYYPPI, SANA JA KH) on selvä vaikutus rajan sijaintiin, koska keskihajontaparametri on selvästi > 0. Toisin sanoen eri koehenkilöillä (KH) oli keskenään eri keskiwertorajoja (ts. katsomatta SANA- ja TYYPPI-vaikutuksia), samoin eri sanoilla (SANA). Tämän tutkimuksen kannalta on huomionarvoista, että myös *spon-* ja *luku-*sarjoilla (TYYPPI) oli selvästi eri raja keskimäärin (katsomatta SANA- ja KH-vaikutuksia). TYYPPI-efektin keskihajonta ei sinänsä kerro *miten päin* rajat eroavat, mutta raja on keskimäärin lyhyemmällä kestolla *spon-*sarjassa, kuten alempana käy ilmi.

Rajan sijainnin efekteistä interaktioilla TYYPPI × KH JA TYYPPI × SANA × KH on verrat-

tain vähän vaikutusta, mikä helpottaa tulosten tulkintaa. On mahdollista päätellä, että esimerkiksi TYYPPI-efektin vaikutus on lähes sama eri koehenkilöillä, tai kääntäen, koehenkilöiden väliset erot vaikuttivat lähes samalla tavalla *spon-* ja *luku-*sarjoissa.

Sen sijaan interaktioilla TYYPPI × SANA ja SANA × KH oli selvä vaikutus, joka on samaa suuruusluokkaa kuin KH:n pääefektillä. Tämä tarkoittaa, että TYYPPI vaikutti eri sanoissa eri tavalla, ja toisaalta SANA vaikutti eri koehenkilöillä eri tavalla. Näistä erityisesti ensimmäinen seikka on tärkeä tämän tutkimuksen kannalta, ja siksi TYYPPI-efektin vaikutusta tarkastellaan paitsi aineistossa keskimäärin myös yksittäisten sanojen tasolla.



KUVA 3. Rajan sijainnin erotus (*luku-spon*) koehenkilöittäin ja sanoittain; posteriorijakaumien mediaani, 50% ja 95% CI (posterioritodennäköisyys, että erotus on saman suuntainen kuin mediaani: * $p > 0,95$, ** $p > 0,99$, *** $p > 0,999$).

Kuvassa 3 näkyvät *luku*-sarjan ja *spon*-sarjan rajan erotuksen (toukkajanoiksi pelkistetyt) posteriorijakaumat sekä keskimäärin että koehenkilöiden ja sanojen mukaan eriteltyinä. Sekä koehenkilöt että sanat esitetään erotuksen mediaanin mukaisessa järjestyksessä. Kuten kuvasta voi nähdä, *spon*-sarjan keskimääräinen raja sijaitsee *lyhyemmällä kestolla* kuin *luku*-sarjan raja (erotus on positiivinen) sekä koko aineistossa (eli ilman SANA- ja KH-Vaiikutuksia) että kaikilla koehenkilöillä. Vaikka koehenkilöiden välillä on selviä eroja, mitään selvää korrelaatiota ei löytenyt koehenkilöiden taustatietojen kanssa (murretausta, ikä, sukupuoli).

Kuten yllä todettiin, TYYPPI \times SANA-interaktio oli verrattain suuri, mikä näkyy kuvassa 3: kahden sarjan rajan erotus vaihtelee sanan mukaan paljon. Enemmistöllä erotus on melko varmasti positiivinen, mutta kahdella sanalla (*nähtävää* ja *riippuu*) ero on niin pieni, että sarjojen keskinäinen järjestys on epävarma, ja yhdellä sanalla (*jatkuu*) järjestys on jopa todennäköisesti päinvastainen. Pohdintaluvussa palataan tämän variaation mahdollisten syiden pohdiskeluun.

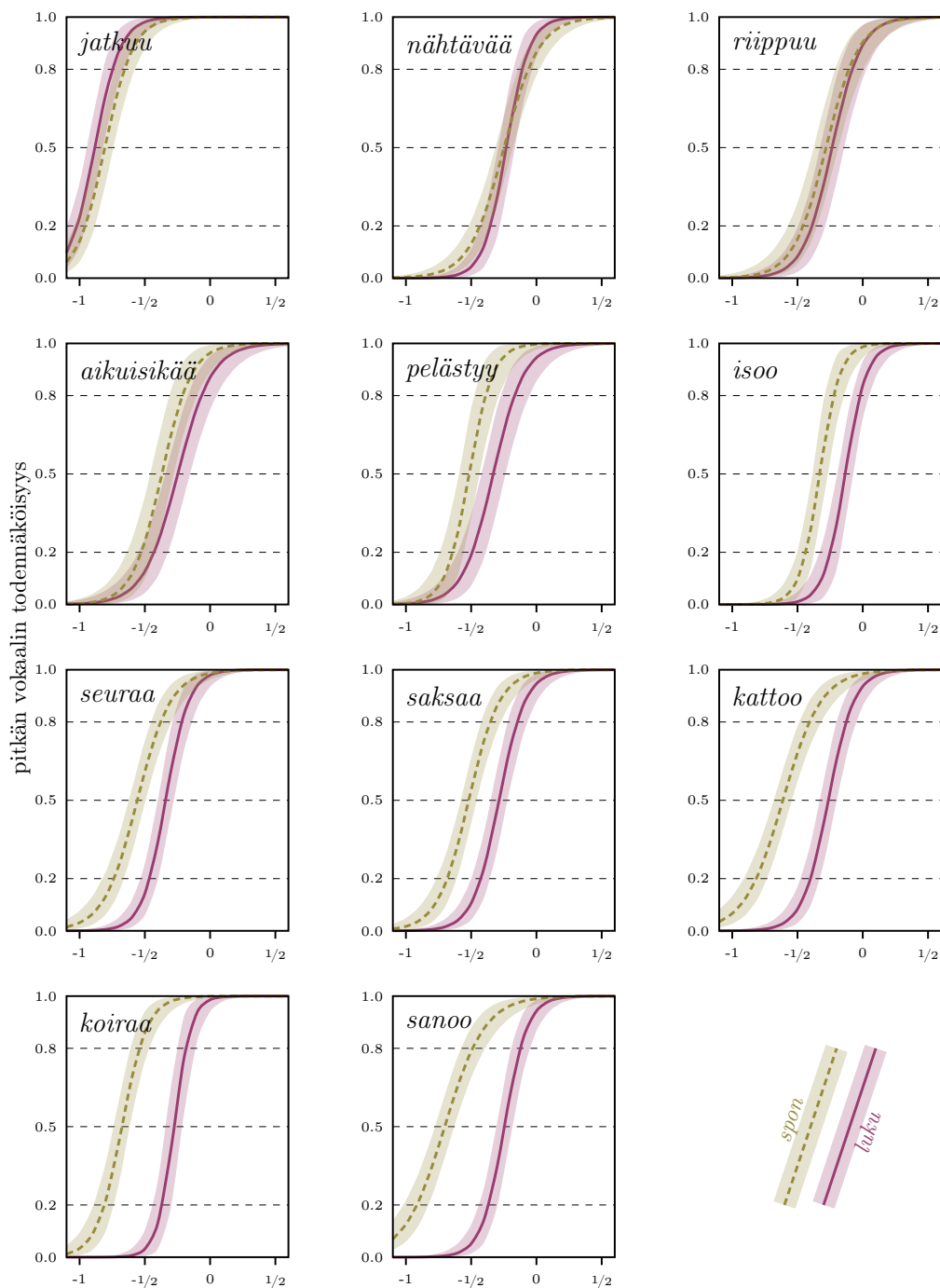
3.2 Rajan leveys

Kulmakerroin (β_{ijk} yhtälössä 1) kertoo kuinka nopeasti pitkän vokaalin havainnon todennäköisyys kasvaa, kun siirrytään ärsykesarjassa oikealle eli kohti pidempiä vokaaleja. Saman asian voi tulkita myös lyhyen ja pitkän vokaalin välisenä rajan leveytenä ($1/\beta_{ijk}$): kun kulmakerroin suurenee, havainnon todennäköisyys kasvaa nopeammin, ja rajan leveys on pienempi.

Leveysvaikutukset olivat kaiken kaikkiaan melko pieniä. Kuvasta 2 nähdään, että KH:n pääefektillä on selvä vaikutus rajan leveyteen. Toisilla koehenkilöillä oli näin ollen kapeampi raja kuin toisilla. Muut pääefektit eivät vaikuttaneet yhtä varmasti. Esimerkiksi TYYPPI:n pääefektillä oli vain heikko vaikutus, mutta on todennäköisempää, että keskimäärin *luku*-sarjalla oli kapeampi raja kuin *spon*-sarjalla.

Interaktioista selvin vaikutus on interaktiolla SANA \times KH. Yksittäisillä koehenkilöillä oli nähtävissä leveyseroja eri sanojen suhteen, mutta koehenkilöt erosivat keskenään siinä, minkä sanan kohdalla raja oli kapeampi tai leveämpi. TYYPPI \times SANA -interaktion vaikutus oli melko pieni; posterioritodennäköisyys sille, että *luku*-sarjan raja oli kapeampi, oli useammalla sanalla $> 0,5$, mutta joillakin sanoilla $< 0,5$.

Yhteenvetona esitetään koko aineiston estimoidut, posteriorijakauman mukaiset identifikaatio- eli todennäköisyyskäyrät kuvassa 4. Kuvassa esitetään käyrät sanoittain, ja mukana ovat 95 % luottamusvälit kuvaamassa jäljelle jäävää epävarmuutta, kun data on otettu huomioon.



KUVA 4. Posteriorin mukaiset keskivertoidentifikaatiokäyrät (mediaani ja 95% CI).

4 POHDINTA

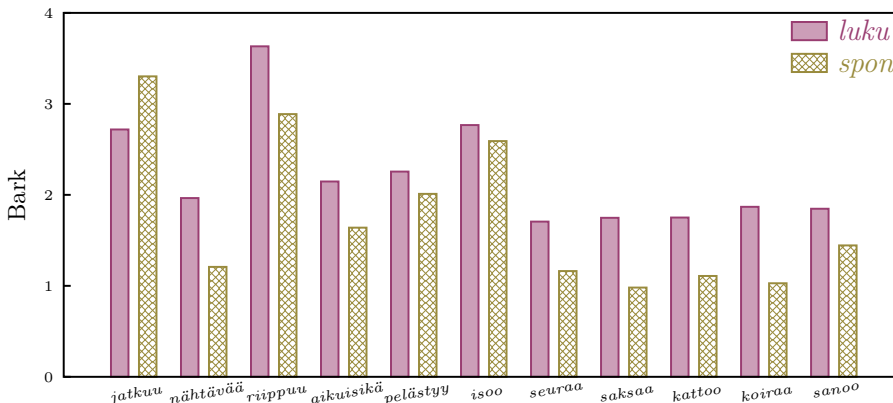
Hypoteesi siitä, että spontaanissa puheessa jälkitavun pitkän vokaalin ei tarvitse olla yhtä pitkä kuin huolellisessa lukupuhunnassa tullakseen havaituksi pitkäksi, sai vahvistusta tuloksista. Koska ärsykkeiden segmenttien kestoerot neutraloitiin, kvantiteetin havaitseminen ei täten perustu vain yksittäisen äänteen fyysiseen keston ja segmenttien kestojen suhteisiin, jotka on esitetty aiemmassa kirjallisuudessa (Lehtonen, 1970; Meister ym., 2011).

Havaintojen erilaisuus spontaanin ja lukupuhunnan välillä sopii osaltaan yhteen sen kanssa, että aiemmissa tutkimuksissa prosodinen konteksti on vaikuttanut suomessa kvantiteetin havaitsemiseen enemmän kuin muissa tutkituissa kielissä (Isei-Jaakkola, 2004; 2010). Spontaani puhe on prosodisesti erilaista kuin lukupuhunta, joten on perusteltua olettaa, että spontaanisuus on yksi prosodinen konteksti, joka lisää kvantiteetin havaitsemisen variaatiota.

Koska ärsykkeillä ei ollut laajempaa kontekstia, havainnon täytyy perustua sanan sisäisiin ominaisuuksiin. O'Dellin (2003) tutkimuksessa pitkistä vokaalista tehdyt ärsykkeet tuottivat enemmän pitkävokaalisia havaintoja kuin vastaavan kestoiset lyhyestä vokaalista tehdyt ärsykkeet. Samaan tapaan

tutkimuksessamme spontaanista puheesta tehdyt ärsykkeet saivat enemmän pitkävokaalisia havaintoja kuin lukupuhunnasta tehdyt ärsykkeet. Näin ollen spontaanin puheen ja lukupuhunnan välillä on oltava jokin muu kuin keston perustuva ero samaan tapaan kuin lyhyiden ja pitkien vokaalien välillä. Tulosten perusteella voi spekuloida, että kuulija havaitsee tämän eron, minkä vuoksi hän vaatii vokaalille erilaisen keston eri tyyppin sanoille, eli toisin sanoen hän havaitsee kvantiteetin eri tavalla puhetyypin mukaan. Koska ärsykkeiden segmenttien kestoerot hämärrettiin, käsitys spontaanisuudesta ja täten vokaalin vaatimasta keston ei voi perustua yksinomaan segmenttien kestojen suhteisiin vaan sen täytyy liittyä muihin tekijöihin.

Tämän tutkimuksen tarkoituksena ei ollut tutkia, miten kuulija erottaa spontaanin puheen luetusta puheesta tai miten nämä puhetyypit eroavat toisistaan, vaan selvittää, onko kuulijan havainnoissa spontaanin ja lukupuhunnan välillä ero pitkän ja lyhyen vokaalin välisen rajan sijainnissa, ja tulokset paljastivat, että rajan sijainnissa on ero tyyppien välillä. Pohdimme kuitenkin mahdollisia tekijöitä tämän eron taustalla, vaikka tämän tutkimuksen perusteella ei voida sanoa, mikä eron aiheuttaa.

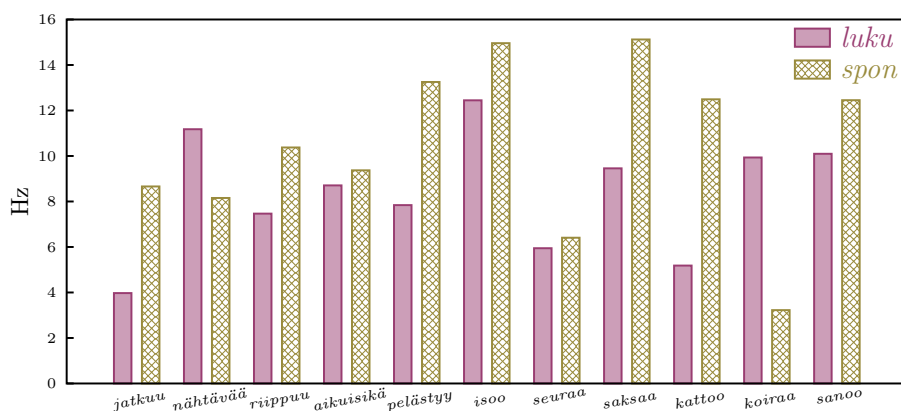


KUVA 5. Ärsykkeiden (sarjan 3. ärsyke) vokaalien keskimääräinen etäisyys vokaaliavaruuden (F1 × F2) keskuksesta Bark-asteikolla.

Spontaanisuuden ja lukupuhunnan erojen voi olettaa näkyvän vokaalien laadussa. Huolellisessa lukupuhunnassa vokaalien voisi olettaa ääntyvän keskimäärin perifeerisempinä (eli kauempana vokaaliavaruuden keskuksesta) kuin nopeassa spontaanissa puheessa. Perifeerisyyden arvioimiseksi kokeessa käytettyjen ärsykkeiden kaksi alinta formanttia (F1 ja F2) mitattiin Praatilla. Kunkin ärsykkeen keskimääräinen etäisyys Bark-asteikolla (Bark-asteikon käytöstä ks. esim. Iivonen, 1994) formanttikentän keskipisteestä eli koko aineiston keskiarvosta toimi karkeana perifeerisyyden mittarina. Kuvassa 5 esitetään tulokset ärsykeittäin *luku-* ja *spont-*sarjoissa. Kaikissa sanoissa *luku-*versio on näin lasketuna perifeerisempi kuin *spont-*versio lukuun ottamatta *jatkuu*-sanaa, jossa spontaani versio on perifeerisempi (kuva 5). *Jatkuu*-sana oli myös havainnoissa poikkeava, mutta perifeerisysero ei kuitenkaan selitä muita ärsykkeiden välisiä eroja havainnoissa, esimerkiksi *nähtävää* sai pienemmän eron kuin *isoo* (vrt. kuvat 3 & 4), vaikka *nähtävää*-sanalla on suu-rempi perifeerisysero kuin *isoo*-sanalla.

Toinen vaikuttava tekijä on mahdollisesti F0. Lennes, Aalto ja Palo (2009, s. 153) havaitsivat, että suomessa ”lukupuheen huippu [ts. F0-jakauman moodi] on keskimäärin kor-

keampi kuin spontaanin puheen” ja toisaalta ”luetussa puheessa käytetään korkeita taa-juuksia vähemmän kuin spontaanissa puheessa.” Monissa kielissä on todettu, että F0:n liikkeet voivat voimakkaastikin vaikuttaa puheen havaittuun tempoon, mikä puolestaan voi vaikuttaa kvantiteetin havaitsemiseen (suomessa esim. Aulanko, 1985, s. 45–46; O’Dell ym., 2019; Vihanta, 1988). Esimerkiksi dynaamisen sävelkulun on havaittu kasvattavan jälkitavujen vokaalin havaitsemista pitkäksi sekundaarisena vihjeenä japanissa (Kinoshita, Behne & Arai, 2002; Takiguchi, Takeyasu & Giriko, 2010). Ärsykeistä kaikilla paitsi *riippuu*-sanan *spont-*ärsykkeellä oli laskeva F0. Vainio ym. (2010) huomasi laskevan sävelkulun suomen pitkissä painollisissa vokaaleissa. Lisäksi Vihanta (1988) esitti, että äänenkorkeuseron suuruus ensimmäisen ja toisen tavun välillä voi toimia vihjeenä kvantiteetista. Tämä saattaa tarkoittaa, että spontaani *riippuu*-sarja ei kuulostanut yhtä spontaanilta tai yhtä nopealta kuin muut spontaanisarjan ärsykkeet, mikä saattaisi osittain selittää, miksi *riippuu*-ärsykkeiden havainnoissa ero kahden sarjan välillä jäi pieneksi. Tämän todeksi osoittaminen vaatisi kuitenkin erillisen tutkimuksen.



KUVA 6. Ärsykkeiden (sarjan 3. ärsyke) F0:n keskihajonta.

Ärsykkeiden äänenkorkeuden yleistä dynaamisuutta mitattiin laskemalla F0:n keskihajonta. Hypoteesin mukaan äänenkorkeus on dynaamisempi spontaanissa puheessa kuin luetussa puheessa. Käyttämässämme ärsykkeissä F0:n keskihajonta on spontaanissa versiossa suurempi kaikissa paitsi *nähtävää*- ja *koiraa*-ärsykkeissä (kuva 6). Se, että nähtävää-ärsykkeissä luku-versio osoittautui dynaamisemmaksi voisi potentiaalisesti selittää, miksi myös *nähtävää*-ärsykkeiden havainnoissa ero *luku*- ja *spon*-sarjojen välillä jäi pieneksi. Dynaamisuusero ei voi kuitenkaan yksin selittää kaikkia tapauksia. Esimerkiksi ärsykkeillä *kattoo*, *koiraa* ja *seuraa* on erilaiset F0:n keskihajonnat (*kattoo*- ja *koiraa*-ärsykkeillä erot keskihajonnassa ovat suuret mutta eri päin, ja *seuraa*-ärsykkeillä ero on pieni), mutta niillä kaikilla on selvä ero havainnoissa *spon*- ja *luku*-versioiden välillä. *Nähtävää*-ärsykkeillä havaintoerot ovat pienemmät kuin *koiraa*-ärsykkeillä, vaikka niillä molemmilla on käänteiset F0:n keskihajontaerot muihin ärsykkeisiin nähden.

Yllä mainittujen seikkojen lisäksi on katsottava myös alkuperäisten sanojen loppuvokaalien kestoja. Koska tutkimuksemme perusteella kuulija antaa todennäköisimmin eri vastauksen *spon*- ja *luku*-ärsykkeille, on syytä katsoa, onko alkuperäisen sanan ja sen loppuvokaalin kestolla yhteyttä havaintoihin. Havaintokokeessa käytettyjen ärsykkeiden segmenttien kestojen neutralisoinnin vuoksi *spon*- ja *luku*-sarjan ärsykkeiden välillä ei ollut kestoeroja, joten kuulija ei voinut perustaa vastaustaan kestoeroihin, mutta koska *spon*- ja *luku*-sarjojen ärsykkeet saivat eri vastauksia, on mahdollista, että alkuperäinen kesto vaikuttaisi kesto vaikutelmaan myös ärsykkeiden manipulaation jälkeen.

Kuten yllä on mainittu, *jatkuu*-ärsykkeet saivat muihin ärsykkeisiin nähden poikkeavat vastaukset, sillä sen identifikaatiokäyrissä *spon*-sarjan vastaukset sijoittuvat pidemmäl-

le kestolle kuin *luku*-sarjan vastaukset (kuva 4). Tämän perusteella voisi olettaa, että alkuperäisen *jatkuu*-sanan loppuvokaalien kestojen suhteet poikkeaisivat muista. Taulukosta 1 nähdään, että alkuperäisen sanan loppuvokaalien ero *spon*- ja *luku*-versioiden välillä on vain 1,16, joka on pienempi kuin useimmilla muilla sanoilla, mutta toisaalta alkuperäisellä *seuraa*-sanalla eroa ei juurikaan ole, mutta sen identifikaatiokäyrät kuitenkin erottuvat toisistaan selvästi. Alkuperäisellä *jatkuu*-sanalla on sen sijaan kaikista sanoista pienin ero *spon*- ja *luku*-versioiden kokonaiskestojen välillä (1,15, kun muilla sanoilla se on vähintään 1,36, enimmillään 2,22). F0:n keskihajonnassa tämä ei näy (kuva 6), mutta formanttien perifeerisyysero on käänteinen (kuva 5). Sanalla *kattoo* on suurin ero sekä kokonaiskestossa että loppuvokaalin osuudessa, ja ero sen identifikaatiokäyrissä on kolmanneksi suurin. Sitä suurempia ovat *koiraa* ja *sanoo*, joista *sanoo*-sanalla on niin ikään suuri loppuvokaalin ero (1,80) mutta ei yhtä suuri kokonaiskeston ero (1,38), kun taas *koiraa*-sanalla on suurempi kokonaiskeston ero (1,60) ja pienempi loppuvokaalin ero (1,34). Sanoilla *isoo* ja *nähtävää* on samankaltaiset suhdet (eritoten *isoo* kokonaiskesto 1,66 ja loppuvokaali 1,40, *nähtävää* 1,60 ja 1,45), mutta *nähtävää*-sanan identifikaatiokäyrien erot eivät ole suuret toisin kuin *isoo*-sanalla, joka sijoittuu identifikaatioerojen puoliväliin. Tämän perusteella alkuperäisten sanojen kestoista ei voi tehdä suoria johtopäätöksiä tuloksiin.

Yksittäisten sanojen lisäksi *spon*-sarjan sanojen alkuperäisellä kontekstilla voi olla vaikutusta. Sana *jatkuu* oli käytetyistä sanoista ainoa, joka oli fraasin lopussa, mikä saattaa selittää sen erilaisuuden muihin verrattuna, vaikka myös siinä *spon*-version loppuvokaali oli suhteellisesti lyhempi kuin *luku*-versiossa. Kontekstista irrrottamisen vuoksi osa lyhyt-pitkä-sanapareista ei ollut kokeen asetelmas- sa samassa merkityksessä kuin alkuperäisessä

kontekstissa. Tällainen oli esimerkiksi *seuraa*, joka alkuperäisessä spontaanissa puheessa oli verbi, mutta joka havaintokokeessa *seura-*vaihtoehdon vieressä mielletään helposti substantiiviksi. Koska kyse on kvantiteetin havaitsemisesta yksittäisen sanan kontekstissa, emme katso tällä olevan merkitystä tutkimuksen kannalta.

Loppuvokaalin kestoon sen sijaan sanan konteksti vaikuttaa. Alkuperäisessä spontaanissa puheessa *sanoo* ja *kattoo* esiintyvät infinitiiveinä, joissa on jäännöslopuke, minkä vuoksi seuraavan sanan alkukonsonantti piteenee (rajageminaatio), mikä puolestaan sulkee tavun. Umpitavun vokaali on usein lyhempi kuin vastaavan avotavun vokaali, joskaan yhteys ei ole aivan suoraviivainen (Lehtonen, 1970, s. 81–84). Tavun sulkeva konsonantti on leikattu ärsykkeistä pois, mutta sen jättämät formanttisiirtymät saattavat olla havaittavia ja siksi vetävät havaintoa lyhyen vokaalin suuntaan *spoon-*ärsykkeiden osalta. Tällä seikalla ei luultavasti ole vaikutusta itse tutkimuksen kannalta, sillä jokaisella spontaanista puheesta muodostetulla ärsykkeellä oli alkuperäisessä kontekstissaan omanlaisensa äänneympäristö, eikä tutkimuskysymyksenä ole, miksi vokaalien suhde on erilainen spontaanin ja lukupuhunnan välillä, vaan havaitseeko kuulija vokaalin kestoeron näiden välillä eri tavalla, eli onko hänen kokemuksensa kvantiteetista niiden välillä eri.

Tulosten voi nähdä toisaalta sopivan myös Lindblomin (1990) jaotteluun hypo- ja hyperpuheeseen, jossa spontaani puhe vastaisi systeemikeskeistä hypopuhetta ja lukupuhunta tuotoskeskeistä (*output-oriented*) hyperpuhetta. Spontaanissa puheessa puhuja näin ollen käyttäisi pienempää ponnistusta ja pienempiä fysikaalisia eroja kuin lukupuhunnassa, ja kuulija kykenisi tunnistamaan tämän eron, mikä puolestaan vaikuttaisi hänen arvioonsa vokaalin kvantiteetista. Tässä tutkimuksessa tällainen ero onkin todennä-

köinen, sillä lukupuhuntaärsykkeet olivat yksittäin luettuja sanoja, jolloin puhujan artikulaatiossa on todennäköisesti tällöin ollut suurempia liikkeitä, mikä näkyi formanttien suuremmassa periferisyydessä (kuva 5), joka puolestaan voi kertoa kuulijalle, minkä tyyppisestä puheesta on kyse.

Tutkimuksemme sopii myös hyvin yhteen sen kanssa, että kvantiteetti on kuulijan käsitys kestosta (Lehtonen, 1970, s. 19). Fyysisen keston ei tarvitse olla aina samanlainen kaikissa konteksteissa tai puhetilanteissa, jotta se silti tuottaa fonologisen pituuden. Se myös antaa näyttöä sisäisen ajoituksen puolesta, sillä spontaanin ja lukupuhunnan välillä samat kestot tuottivat eri havaintoja, joskaan se ei myöskään kumoa ulkoista ajoitusta, sillä vokaalin lyhimät ja pisimmät kestot tuottivat lähes yksinomaan lyhyitä tai pitkiä havaintoja. Erot syntyivät niin sanotulla epävarmalla alueella, jossa spontaanin puheen ärsykkeet saivat enemmän pitkiä havaintoja kuin lukupuhunnan ärsykkeet.

5 LOPUKSI

Koska spontaani puhe on lukupuhuntaa vaihtelevampaa, on selvää, että yksittäisissä esiintymissä on paljon variaatiota. Siksi voidaankin pitää yllättävänä tutkimuksemme tulosta, jossa näinkin pienessä aineistossa sanat saivat selvästi eri rajan sijainnin tyyppin mukaan. Jatkossa olisi kuitenkin hyvä saada enemmän dataa: useita puhujia, lisää sanoja ja jopa useita esiintymiä samoista sananmuodoista.

Muita kiinnostavia jatkotutkimusaiheita voisivat olla suomen oppijoiden ja muunkielisten havaintojen tutkiminen sekä spontaanin puheen ja lukupuhunnan pitkien vokaalien akustinen analyysi korpusaineistosta. Lisäksi olisi kiinnostavaa tietää, miten kuulijoiden reaktioaika vaihtelee eri kestojen ja spontaanin ja lukupuhunnan välillä samaan tapaan kuin Ylisen ym. (2005) tutkimuksessa,

jossa vertailtiin suomen- ja venäjänkielisten reaktioaikoja kvantiteetin havaitsemisessa.

Tulokset osoittavat, että spontaani puhe on otettava havaitsemista koskevissa tutkimuksissa huomioon. Kvantiteettijärjestelmän toimimisen ymmärtämiseksi on välttämätöntä, että havainnot perustuvat myös spontaaniin puheeseen. Koska kvantiteetin havaitseminen spontaanissa puheessa osoittautui erilaiseksi kuin lukupuhunnan havaitseminen, on spontaanin puheen tutkimuksella varmasti annettavaa myös fonologisille teorioille, kuten Warner (2012) esittää. Automaattisen puheentunnistuksen kehittämisessä olisi eduksi, mikäli kvantiteetti muodostettaisiin siinä spontaanin puheen mukaisesti.

LÄHTEET

- Anttila, H. (2009). Interrogative intonation in spontaneous Finnish. Teoksessa V. de Silva ja R. Ullakonoja (toim.), *Phonetics of Russian and Finnish* (s. 167–176). Frankfurt am Main: Peter Lang.
- Arai, T. (1999). A case study of spontaneous speech in Japanese. Teoksessa J. Ohala, Y. Hasegawa, M. Ohala, D. Granville & A. Bailey (toim.), *Proceedings of the 14th International Congress of Phonetic Sciences*, (s. 615–618). Berkeley: University of California.
- Arai, T. & Warner, N. (1999). Word level timing in spontaneous Japanese speech. Teoksessa J. Ohala, Y. Hasegawa, M. Ohala, D. Granville & A. Bailey (toim.), *Proceedings of the 14th International Congress of Phonetic Sciences*, (s. 1055–1058). Berkeley: University of California.
- Aulanko, R. (1985). Microprosodic features in speech: Experiments on Finnish. Teoksessa *XIII Fonetikan päivät. Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 26*, (s. 33–54). Turun yliopisto.
- Boersma, P. & Weenink, D. (2020). Praat: Doing phonetics by computer (Version 6.1.09) [tietokoneohjelma]. Haettu 10.2.2020 osoitteesta <http://www.praat.org/>.
- de Silva, V., Iivonen, A., Bondarko, L. V. & Pols, L. C. (2003). Common and language dependent phonetic differences between read and spontaneous speech in Russian, Finnish and Dutch. Teoksessa M. J. Solé, D. Recasens & J. Romero (toim.), *Proceedings of the 15th International Congress of Phonetic Sciences*, (s. 2977–2980). Universitat Autònoma de Barcelona.
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis 1.3*, 515–533.
- Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
- Iivonen, A. (1994). A Psychoacoustical explanation for the number of major IPA vowels. *Journal of the International Phonetic Association*, 24, 73–90.
- Isei-Jaakkola, T. (2004). *Lexical quantity in Japanese and Finnish*. Helsingin yliopiston fonetiikan laitoksen julkaisuja 48. Helsinki: Hakapaino Oy.
- Isei-Jaakkola, T. (2010). Durational variability of vowel quantity boundary for Japanese, Finnish and Czech Speakers in perception. Teoksessa *Speech Prosody 2010 – Fifth International Conference* (paper 192, s. 1–4). Chicago.

- Järvikivi, J., Aalto, D., Aulanko, R. & Vainio, M. (2007). Perception of vowel length: Tonality cues categorization even in a quantity language. Teoksessa J. Trouvain & W. J. Barry (toim.), *Proceedings of the 16th International Congress of Phonetic Sciences*, (s. 693–696). Universtiät des Saarbrückens.
- Kinoshita, K., Behne, D. M. & Arai, T. (2002). Duration and F0 as perceptual cues to Japanese vowel quantity. Teoksessa Z. Tan & P. Dalsgaard (toim.), *Proceedings of the Seventh International Conference on Spoken Language Processing*, (s. 757–760). Denver.
- Lehtonen, J. (1969). Huomioita kvantiteettien foneemirajoista ja subjektiivisista kestohavainnoista. *Virittäjä*, 73, 363–363.
- Lehtonen, J. (1970). *Aspects of quantity in standard Finnish*. Studia Philologica Jyväskyläläisiä VI. Jyväskylän yliopisto.
- Lennes, M. (2009). Segmental features in spontaneous and read-aloud Finnish. Teoksessa V. de Silva & R. Ullakonoja (toim.), *Phonetics of Russian and Finnish: General description of phonetic systems, experimental studies on spontaneous and read-aloud speech*, (s. 145–166). Frankfurt am Main: Peter Lang.
- Lennes, M., Aalto, D. & Palo, P. (2009). Puheen perustaajuuksajakaumat: Alustavia tuloksia. Teoksessa M. O'Dell & T. Nieminen (toim.), *Fonetiikan päivät 2008: XXV Fonetiikan päivillä Tampereen yliopistossa 11.–12.1.2008 pidetyt esitelmät*. Tampere Studies in Language, Translation and Culture, Series B 3. Tampere: Tampere University Press, s. 147–155.
- Lennes, M., Alarotu, N. & Vainio, M. (2001). Is the phonetic quality of unaccented words unpredictable? An example from spontaneous Finnish. *Journal of the International Phonetic Association*, 31, 127–138.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. Teoksessa W. J. Hardcastle & A. Marchal (toim.), *Speech Production and Speech Modelling* (s. 403–439). Springer, Dordrecht.
- Meister, E., Werner, S. & Meister, L. (2011). Short vs. long category perception affected by vowel quality. W. S. Lee & E. Zee (toim.), *Proceedings of the 17th International Congress of Phonetic Sciences* (s. 1362–1365). City University of Hong Kong.
- Neenonen, S. (2001). Venäläiset suomenoppijat suomen kielen pitkien painottomien vokaalien havaitsoijoina. M. Charles & P. Hiidenmaa (toim.), *Tietotyön yhteiskunta – kielen valtakunta: AFinLAn vuosikirja 2001* (s. 11–31). Jyväskylä: Jyväskylän soveltavan kielitieteen yhdistys AFinLA.
- O'Dell, M. (2003). *Intrinsic timing and quantity in Finnish*. Acta Universitatis Tamperensis 979. Tampere University Press.
- O'Dell, M., Nieminen, T. & Vakkilainen, J. (2019). F0:n, kvantiteetin ja puhutempon keskinäisistä vaikutussuhteista. *Esitelmä Fonetiikan päivillä 22.–23.3.2019*. Helsinki.
- Plummer, M. (2017). JAGS Version 4.3.0 user manual. <http://mcmc-jags.sourceforge.net>
- Plummer, M. (2019). rjags: Bayesian Graphical Models using MCMC. R package version 4–10. <https://CRAN.R-project.org/package=rjags>.
- R Core Team, (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org>
- Takiguchi, I., Takeyasu, H. & Giriko, M. (2010). Effects of a dynamic F0 on the perceived vowel duration in Japanese. *Speech Prosody 2010 – Fifth International Conference*, (paper 944, s. 1–4).
- Vainio, M., Järvikivi, J., Aalto, D. & Suni, A. (2010). Phonetic tone signals phonological quantity and word structure. *The Journal of the Acoustical Society of America*, 128, 1313–1321.
- Vihanta, V. (1988). F0:n osuudesta suomen kvantiteettioppositiossa. Teoksessa M. Karjalainen & U. K. Laine (toim.), *Fonetiikan päivät — Espoo 1988 / Papers from the 15th Meeting of Finnish Phoneticians* (s. 13–35). Helsinki University of Technology.
- Warner, N. (2012). Methods for studying spontaneous speech. Teoksessa A. C. Cohn, C. Fougeron & M. K. Huffman (toim.), *The Oxford handbook of laboratory phonology* (s. 621–633). Oxford: Oxford University Press.
- Ylinen, S., Shestakova, A., Alku, P. & Huotilainen, M. (2005). The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. *Language and Speech*, 48, 313–338.

PERCEPTION OF FINNISH LONG UNSTRESSED VOWELS IN SPONTANEOUS AND READ ALOUD SPEECH

Joonas Vakkilainen, Tampere University

Michael O'Dell, Tampere University

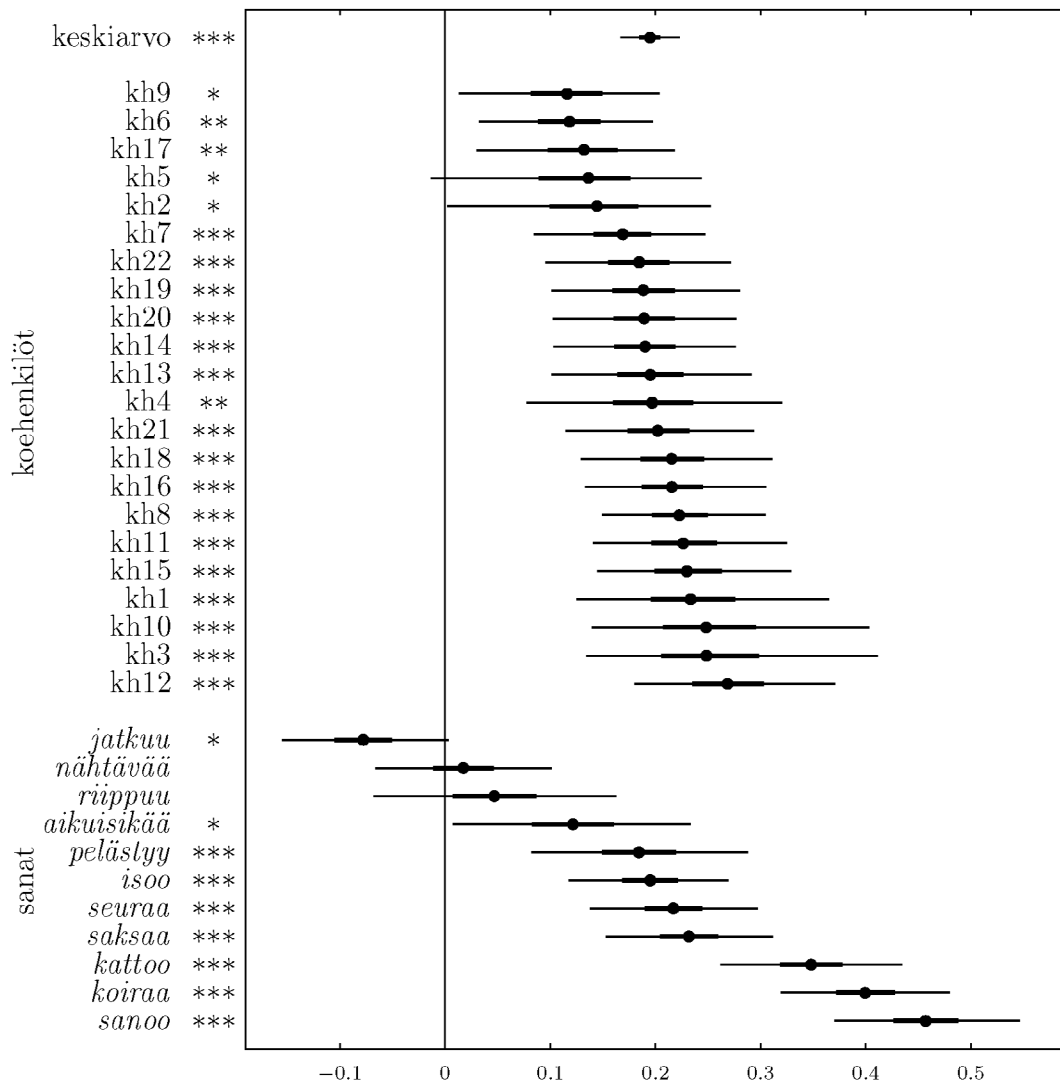
The difference between long and short vowels and consonants in Finnish is phonologically distinctive, i.e. it distinguishes meaning. In spontaneous speech, the difference between short and long vowels is smaller than in read aloud speech (de Silva et al., 2003). In Finnish, the duration of the vowel is the primary cue of quantity for the listener, but tonal differences have also been found in stressed syllables (Järvikivi et al., 2007). The present study deals with the perception of long vowels in unstressed word final syllables. The research question is whether the vowel's physical duration influences perception differently in spontaneous speech compared to read aloud speech. Perception tests were carried out in which the test subjects heard modified stimuli. The stimuli were words taken from spontaneous speech and from word lists read aloud by the same person. Internal duration differences in the stimuli were neutralized. Each original word had a long vowel in the last syllable, whose duration was modified into four different variants. Test subjects responded whether they heard a word containing a short or a long vowel. The results reveal that in spontaneous speech a word final vowel does not need to be as long as in read aloud speech to be perceived as long. On average the spontaneous speech stimuli evoked more long vowel responses than the read stimuli. Other factors (such as F0, F1, F2) did not show a clear correlation with responses, although such effects cannot be ruled out based on this study.

Keywords: duration, perception, quantity, spontaneous speech, vowel

ERRATUM

Joonas Vakkilainen ja Michael O'Dell

Puhe ja kieli -lehden numerossa 4/2021 artikkelissamme *Suomen pitkien painottomien vokaalien havaitseminen spontaanissa ja luetussa puheessa* oli kuvasta 3 leikkautunut ylä- ja alaosa pois. Kuva on alla kokonaisuudessaan.



Kuvassa 1 oli lisäksi ensimmäinen IPA-symboli korvautunut väärällä merkillä. Oikea symboli on α .

