



Íslenskir sérhljóðar

Tveir og þrír formendur

Jón Þór Hrannarsson og Viktor Díar Jónasson

Lokaverkefni til BS-gráðu
Sálfræðideild
Heilbrigðisvísindasvið



HÁSKÓLI ÍSLANDS

Íslenskir sérhljóðar
Tveir og þrír formendur

Jón Þór Hrannarsson og Viktor Díar Jónasson

Lokaverkefni til BS-gráðu í sálfræði

Leiðbeinandi: Jörgen L. Pind

Sálfræðideild

Heilbrigðisvísindasvið Háskóla Íslands

Júní 2012

Ritgerð þessi er lokaverkefni til BS gráðu í sálfræði og er óheimilt að afrita ritgerðina á nokkurn hátt nema með leyfi rétthafa.

© Jón Þór Hrannarsson og Viktor Díar Jónasson 2012

Prentun: Prentsmiðja Háskólaprent
Reykjavík, Ísland 2012

Efnisyfirlit

Útdráttur	5
Þakkarorð	6
Formendur sérhljóða	7
Íslensk sérhljóð	8
Rannsókn Ástu Svavarsdóttur og félaga á formendum íslenskra sérhljóða	9
Hljóðmyndunarkenningin	11
Haskins rannsóknarstofan	11
Klatt talgervillinn	14
Mel-kvarðinn	15
Hljóðmyndun sérhljóða	16
Staða tungu og vara í myndun sérhljóða	16
Algengi og mismunandi gerðir sérhljóða í heiminum	18
Tilgangur rannsóknarinnar	20
Aðferð	21
<i>Þátttakendur</i>	21
<i>Áreiti</i>	21
<i>Mælitæki</i>	22
<i>Framkvæmd</i>	23
<i>Tölfræðileg úrvinnsla</i>	23
Niðurstöður	26
Umræða	35
Heimildir	38
Myndaheimildir	39

Útdráttur

Formendur eru skilgreindir sem samhljómunarsvið munnhols. Þeir eru mótaðir af tungustöðu, opnustigi og kringingu. Hljóðmyndunarkenningin (*acoustic theory of speech production*) segir að tveir meginþættir séu sameiginlegir við myndun allra máhljóða. Sá fyrri er hljóðgjafi og sá seinni er samhljómur. Ein fyrsta talskynjunarrannsóknin var gerð á Haskins-rannsóknarstofunni þar sem rannsökuð voru áhrif fjölda formenda á gæði sérhljóða. Tilgangur þessarar rannsóknarinnar var tvíþættur. Í fyrsta lagi að kortleggja sérhljóðarými í íslensku með skynjunartilraun. Í öðru lagi að athuga mun á tveimur og þremur formendum í skynjun á íslenskum sérhljóðum. Tilgáta rannsakenda er sú að sérhljóð með þremur formendum verði greind skýrari heldur en sérhljóð með tveimur formendum. Þátttakendur voru 17 talsins, valdir eftir hentugleika. Áreitnir voru 126 sérhljóð með annaðhvort tveimur eða þremur formendum. Þau voru búinn til í Praat forritinu og tilraunin var einnig keyrð í því. Þátttakendur heyrðu hljóð og áttu að tilgreina hvaða sérhljóð þeir heyrðu og meta gæði hvers hljóðs. Niðurstöður rannsóknarinnar voru að ekki var marktækur munur á milli sérhljóða með tveimur eða þremur formendum. Það sem spilar hugsanlega inn í niðurstöðurnar er að þriðji formendinn hafði alltaf fast gildi, 3500 Hz. Í ritgerðinni er greint frá því með hvaða hætti hlustendur skynja íslenska sérhljóðarýmið.

Þakkarorð

Við viljum byrja á að þakka leiðbeinanda okkar Jörgen L. Pind fyrir alla hans aðstoð, leiðsögn og ábendingar. Einnig viljum við þakka samnemendum okkar Heiðari Hrafni Halldórssyni og Benóný Þór Björnssyni fyrir stuðning og ábendingar auk allra þeirra samverustunda sem við áttum saman við gerð rannsóknarinnar. Að lokum viljum við þakka foreldrum okkar fyrir þeirra stuðning í gegnum árin.

Formendur sérhljóða

Formandi er hugtak sem er mikið notað í hljóðeðlisfræði en skilgreiningin á hugtakinu er að formendur eru samhljómunarsvið munnhols (Ásta Svavarsdóttir, Halldór Ármann Sigurðsson, Sigurður Jónsson og Sigurður Konráðsson, 1982). Til þess að skoða þetta aðeins nánar er gott að líta á munnhol sem hljóðpípu. Þessi hljóðpípa myndar samhljómun við hljóðbylgjur á ákveðnum tíðnisviðum. Þessi tíðni er auðvitað mismunandi og ræðst hún af stærð og lögun pípunnar. Sá þáttur sem virðist þó skipta mestu máli er lengd pípunnar. Þegar sérhljóð eru mynduð eru ákveðnir þættir sem spila ákveðin hlutverk í myndun þeirra. Til dæmis eru raddböndin hljóðgjafi, munnholið hljóðpípa og loks eru formendur samhljómunarsvið þeirrar hljóðpípu. Með formendum er verið að skoða tíðni og lögun samhljómunarsviða munnholsins þegar ákveðið sérhljóð er myndað. Það er til ákveðið lögmál í þessum fræðum sem segir að því styttri sem hljóðpípan er, þeim mun hærra er tíðnisvið samhljómunar hennar. Þar sem munnhol kvenna er almennt styttra en munnhol karla, þá eru sérhljóðaformendur á hærra tíðnisviði í tali hjá konum en körlum (Ásta Svavarsdóttir o.fl., 1982).

Önnur skilgreining á hugtakinu formendur er að vísa til hljóðdeyfingar en ekki hljóðmögnunar eins og á við í samhljómunarsviði hljóðpípu. Það gerist stundum á ákveðnum tíðnisviðum að það verður hljóðdeyfing í hljóðpípunni. Í ákveðnum tilvikum gerist það þegar samsett hljóð fara gegnum hljóðpípunna að styrkur hljóða sem hafa sömu tíðni eykst en um leið minnkar styrkur annarra tóna hljóðsins. Í þessum tilvikum virka hljóðpípur sem hljóðsúr. Samsett hljóðbylgja er bylgja sem er samsett úr mörgum tónum með mismunandi tíðni, myndast við sveiflur í raddböndum. Samsetta hljóðbylgjan er í aðalatriðum eins í öllum sérhljóðum hjá einum ákveðnum einstaklingi en gjörbreytist við það að fara í gegnum hljóðsúu. Það sem hljóðsían (í þessu tilviki munnholið) gerir er að magna ákveðna tóna á tilteknu tíðnisviði og deyfa aðra. Þetta tíðnisvið sem hljóðsía magnar upp er það sama og formendur sérhljóðsins. Það er þess vegna sem meirihluti hljóðstyrks sérhljóða er á þessu

ákveðna tíðnibili (formendur sérhljóðsins). Því er hægt að fá flestar hljóðeðlislegar upplýsingar sérhljóðsins frá fyrstu tveimur eða þremur formendum þess. (Formendur eru tilgreindir með bókstafnum F. Fyrsti formandi er því F_1 , annar formandi F_2 , þriðji formandi F_3 , og svo framvegis). Sumir vilja meina það að fyrstu tveir formendurnir dugi til að fólk geti greint öll sérhljóð í ákveðnu tungumáli. Flestir eru þó sammála um að það þurfi að hafa fyrstu þrjá formendurnar þegar um er að ræða frammælt sérhljóð (Ásta Svavarsdóttir o.fl., 1982).

Hægt er að lýsa sérhljóði nákvæmlega með því að tilgreina tíðni formenda sérhljóðsins. Fólk með stórt höfuð hefur dýpri endurrómun (*resonance*) sem leiðir af sér formendur með tiltölulega lága tíðni. Að sama skapi hefur fólk með lítið höfuð lægri endurrómun og það leiðir af sér formendur með hærri tíðni. Besta leiðinn til að framsetja (*represent*) sérhljóð er að teikna myndrit þar sem formandi 1 (F_1) er á öðrum ásnunum og formandi 2 (F_2) er á hinum ásnunum (Ladefoged og Disner, 2012).

Íslensk sérhljóð

Í íslensku er máhljóðum skipt í tvennt, sérhljóð og samhljóð. Hljóðmyndunarlega er munurinn á þeim að sérhljóð eru mynduð þegar munnurinn er opinn sem gerir leiðina greiða fyrir loftið að fara úr lungunum og út um munninn. Þetta leiðir það af sér að það er líttill þrýstingur í munninum þegar sérhljóð eru mynduð. Á hinn bóginn eru samhljóð mynduð við meiri þrengsli og stundum er munnurinn alveg lokaður (Jörgen Pind, 1997).

Íslensk sérhljóð eru 13 talsins. Þeim er skipt í einhljóð og tvíhljóð. Einhljóðin eru átta og tvíhljóðin eru fimm. Það sem greinir tvíhljóð frá einhljóðum er að tvíhljóð byrjar sem eitt hljóð og breytist í annað hljóð. Sérhljóðin eru oft flokkuð eftir tungustöðu, opnustigi og kringingu. Með tungustöðu er átt við staðsetningu tungunnar inni í munninum þegar sérhljóðin eru mynduð. Tungustaðan skiptir sérhljóðum í frammælt hljóð og uppmælt hljóð. Frammælt hljóð er þegar tungunni er skotið fram í munninum og uppmælt hljóð er þegar

tungan er aftarlega í munninum. Opnustig segir til um staðsetningu tungunnar við góminn þegar sérhljóðið er myndað, þ.e.a.s. hversu nálægt gómnum tungan er við myndun sérhljóðsins. Kringing gefur til kynna að settur sé stútur á varirnar þegar sérhljóðið er myndað (Jörgen Pind, 1997).

Rannsókn Ástu Svavarsdóttur og félagar á formendum íslenskra sérhljóða

Ásta Svavarsdóttir, Halldór Ármann Sigurðsson, Sigurður Jónsson og Sigurður Konráðsson (1982) framkvæmdu rannsókn þar sem þau skoðuðu formendur íslenskra sérhljóða. Í íslensku eru sérhljóðarnir átta talsins en hvor um sig hefur langt og stutt afbrigði svo í raun eru 16 sérhljóð. Tilraunin var þannig framkvæmd að átta hljóðhafar lásu hvert sérhljóð upp þrisvar sinnum. Sérhljóðin voru ekki lesin upp ein og sér heldur voru þau sett inn í umhverfið *s_s(s)a*. Þannig urðu til nokkur orð sem ekki voru til í alvörunni en það skipti ekki öllu máli. Rannsakendur völdu þetta ákveðna umhverfi því samhljóðinn /s/ hafði víst einna minnst áhrif á formendur þessara sérhljóða. Í heildina fengu rannsakendur 384 mælingar (3 mælingar x 8 hljóðhafar x 16 hljóð).

Niðurstöður fyrir meðaltíðni formendana má sjá í töflu 1.

Tafla 1. Meðaltíðni íslenskra formenda í Hz. Löng sérhljóð eru táknuð með tvípunkti á eftir sérhljóðinu. Gögn frá Ástu Svavarsdóttur o.fl., 1982.

Sérhljóð	Konur			Karlar		
	F1	F2	F3	F1	F2	F3
í:	304	2613	3042	275	2229	2788
í	325	2450	2942	313	2125	2704
i:	521	2317	2908	363	2017	2654
i	546	2096	2825	417	1908	2588
e:	617	2075	2800	471	1813	2692
e	633	1825	2767	508	1588	2542
a:	783	1517	2663	696	1246	2429
a	729	1533	2675	663	1283	2479
ö:	575	1646	2675	479	1367	2350
ö	621	1546	2658	521	1325	2429
u:	533	1888	2642	392	1533	2208
u	550	1838	2658	417	1496	2325
o:	554	958	2618	463	854	2206
o	613	1229	2614	546	1071	2328
ú:	504	792	2588	354	738	2533
ú	538	942	2644	375	850	2367

Það sem er athyglisvert við þessar niðurstöður er að tíðni F_1 í stuttu hljóðunum er hærri en í löngu hljóðunum nema fyrir sérhljóðið /a/. Annað athyglisvert er að tíðni F_2 er alltaf hærri í stuttu sérhljóðunum en löngu fyrir F_2 í uppmæltum einhljóðum (/a/,/o/og/ú/). Einnig er meðaltíðni F_1 , F_2 og F_3 alltaf hærri hjá konum en körlum (Ásta Svavarsdóttir o.fl., 1982).

Hljóðmyndunarkenningin

Hljóðmyndunarkenningin (*acoustic theory of speech production*) segir að tveir meginþættir séu sameiginlegir við myndun allra málhljóða. Sá fyrri er hljóðgjafi, en hljóðgjafinn getur annaðhvort verið sveiflukenndar bylgjur eða suðhljóð. Sveiflukenndar bylgjur myndast þegar raddböndin sveiflast við myndun málhljóðsins en suðhljóð myndast þegar loft verður fyrir mótstöðu í þrengslum. Seinni þátturinn er samhljómur. Mótun málhljóða verður breytileg vegna ólíkrar legu talfæranna sem veldur samhljómun. Því má segja að talfærin virki sem einskonar síur á raddhljóðin (Jörgen Pind, 1997).

Haskins rannsóknarstofan

Í Haskins-rannsóknarstofunni í New Haven í Connecticut, Bandaríkjunum fóru fram fyrstu talskynjunarrannsóknirnar. Þetta byrjaði allt þegar menn á rannsóknarstofunni komu fram með „Pattern playback“ talgervilinn árið 1951. Hann var mjög frumstæður en hann las hljóðrófsrit sem máluð voru á glærar filmur og breytti þeim í einföld málhljóð. Einn helsti gallinn við þennan hljóðgervil var að öll málhljóð höfðu sama grunntóninn, 100 Hz, sem gerði það að verkum að málhljóðin urðu ekkert sérstaklega eðlileg. Þrátt fyrir þessa vankanta hentaði talgervillinn vel í það að athuga formendur og hvaða áhrif þeir hafa á talskynjun því mjög auðvelt var að teikna þá á filmurnar (Jörgen Pind, 1997).

Niðurstöður fyrstu rannsókna þeirra sýndu fram á að nóg var að hafa tvo formendur í málhljóði svo að hægt væri að greina öll algengustu sérhljóðin. Þeir breyttu svo tíðni formendanna en við það breyttist skynjun hlustenda á sérhljóðunum. Fyrir árið 1952 voru rannsakendur engan veginn sammála um hversu marga formendur þurfti til að greina öll sérhljóðin. Sumir sögðu að það þurfti þrjá formendur á meðan aðrir sögðu tvo og sumir sögðu að meira að segja einn formandi væri fullnægjandi. Á þessum tíma var búið að staðsetja tíðni allra formenda flestra sérhljóðana. M. Joos gaf út grein árið 1948 þar sem hann sagðist geta

greint alla sérhljóðana ágætlega með einungis tveimur formendum en hann þorði samt ekki að draga svo sterkar ályktanir svo sem að neita mikilvægi þriðja formenda þar sem rannsóknir hans voru einungis á byrjunarstigi (Delattre, Liberman, Cooper og Gerstman, 1952).

Delattre, Liberman, Cooper og Gerstman (1952) framkvæmdu rannsókn þar sem þeir notuðu „Pattern playback“ talgervilinn og reyndu að búa til öll 16 sérhljóð Alþjóðlega hljóðfræðisambandsins (*International Phonetic Association*). Þetta var gert með því að breyta handteiknuðum hljóðrófsmynstrum (*spectrographic patterns*) í hljóð. Tilgangur rannsókna þeirra var að sannreyna að einungis þurfti tvo formendur til að greina öll sérhljóðin að. Til að sanna það festu þeir fyrstu formendurnar á fjórum mismunandi stöðum þar sem /i/, /y/, /u/ og /u/ höfðu F_1 ávalt í 250 Hz en F_2 var á bilinu 480-3000 Hz. Næstu fjórir sérhljóðarnir /e/, /ø/, /ɤ/ og /o/ voru með F_1 fastan á 360 Hz en F_2 á bilinu 500-2520 Hz. Sérhljóðarnir /é/, /æ/ /ʌ/ og /ɔ/ höfðu F_1 í 510 Hz en F_2 á bilinu 840-2040 Hz. Síðustu fjórir sérhljóðarnir voru /æ/, /a/, /ɒ/ og /ɑ/ en þá var F_1 720 Hz og F_2 á bilinu 960-1800 Hz (Delattre o.fl., 1952).

Í næsta hluta tilraunarinnar festu þeir svo F_2 á meðan þeir breyttu F_1 og báru saman sérhljóðin sem höfðu verið með fastan F_1 við þau sérhljóð þar sem F_2 var fastur. Öll hljóðin sem þeir gerðu voru með grunntíðnina 120 Hz. Næst létu þeir þátttakendur segja til um hvaða sérhljóð þeir héldu að þeir væru að heyra. Niðurstöður þeirra leiddu í ljós að sérhljóðin sem voru utarlega á tíðniskalanum og voru þar af leiðandi ekki jafn umkringd öðrum sérhljóðum voru valinn rétt í fleiri tilvikum heldur en sérhljóð sem voru umkringd öðrum sérhljóðum (Delattre o.fl., 1952).

Í þriðja hluta tilraunar þeirra Delattre og félagar athuguðu þeir áhrif þess að lækka styrk formendanna á gæði sérhljóðanna. Þeir komust að því að þau sérhljóð sem voru með lengsta tíðnibilið á milli F_1 og F_2 /i, e, ε, y, ø og æ/ voru gjörsamlega ógreinanleg sem sérhljóð ef styrkur fyrsta formenda var minnkaður. Lækkun á hljóðstyrk F_2 olli ekki jafn miklum skaða á skýrleika sérhljóðsins eins og lækkun á F_1 . Allt aðrar niðurstöður fundust fyrir sérhljóð sem

voru ekki með langt bil á milli F_1 og F_2 /u, o, ɔ og a/. Hjá þeim minnkuðu ekki gæðin heldur breyttist hljóðið og fóru að hljóma eins og sérhljóðin við hliðina á til dæmis hljómaði /a/ eins og /æ/. Delattre og félagar reyndu að skýra þessar niðurstöður með meðaltali á milli formendanna. Þeir héldu að eyrað reiknaði meðaltal formenda svo ef styrkur F_1 væri lækkaður þá myndi vægi F_2 vera meira þegar eyrað reiknaði meðaltal formendanna (Delattre o.fl., 1952).

Ef þessi hugmynd þeirra er sönn þá myndi það þýða að aðeins þurfti einn formanda til að búa til sérhljóð þar sem eyrað reiknar bara meðaltal tveggja formenda og notar bara þetta eina meðaltal til að greina sérhljóð. Hugmynd þeirra gekk ekki upp þegar þeir bjuggu til tvo formendur með meðaltíðni sem átti að samsvara sérhljóða því hljóðið kom ávallt út eins og tvö mismunandi hljóð. Því næst bjuggu þeir til sérhljóð með einungis einum formanda og athuguðu á 3 punkta skala hversu auðþekkjanleg þau voru og niðurstöður þeirra voru að sumir sérhljóðar eins og /u/ og /o/ eru auðþekkjanlegir með einum formanda á meðan önnur eins og /i/ og /e/ eru gjörsamlega ógreinanleg. Svo virðist sem þeim mun hærri tíðni sem formandi ákveðins sérhljóðs hafði þeim mun erfiðara var að greina það. Niðurstöður þeirra voru skýrar, tveggja formenda sérhljóð er mun skýrara heldur en eins formanda sérhljóð (Delattre o.fl., 1952).

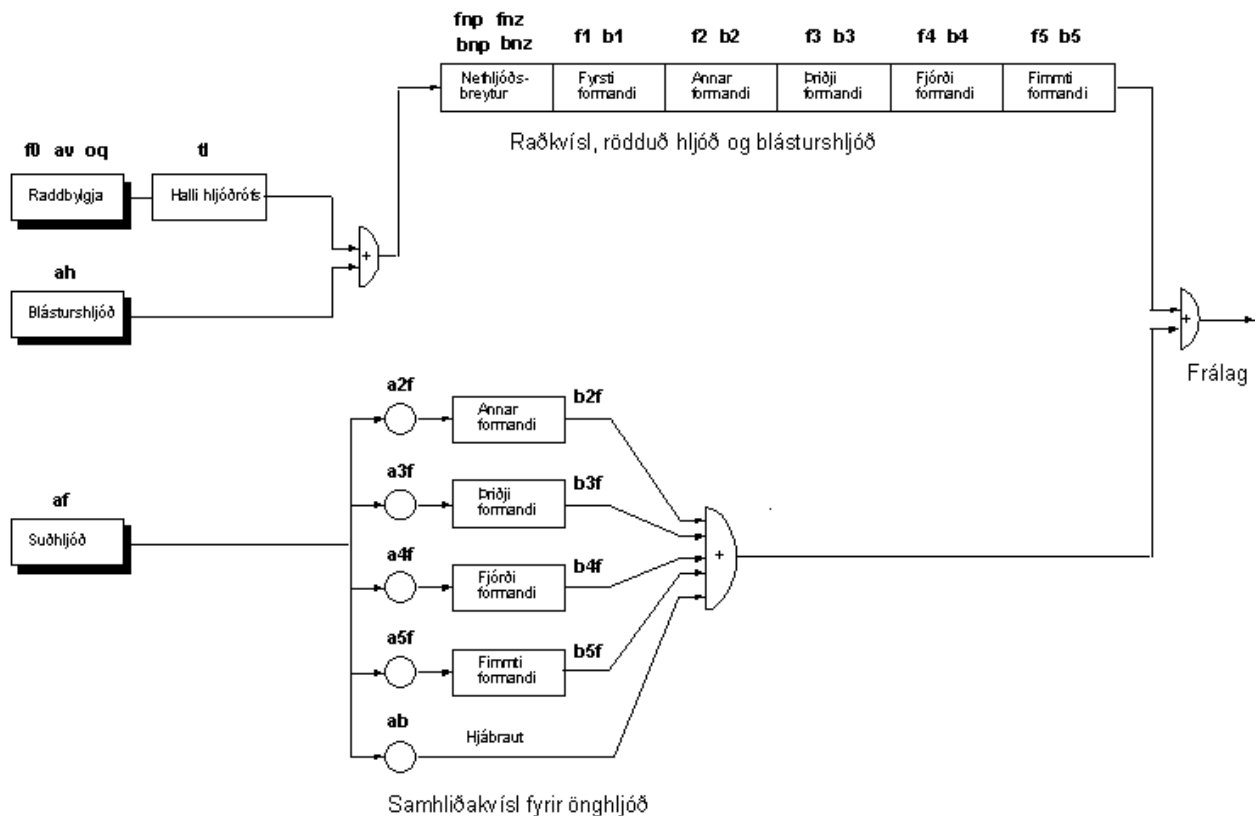
Að lokum báru þeir saman töluð sérhljóð með þremur formendum við gervisérhljóð með tveimur formendum. Niðurstöður þeirra leiddu í ljós að gervisérhljóð voru ávallt með hærri formendur og þá sérstaklega F_2 . Stærsti munurinn fannst fyrir /i/ þar sem F_2 í töluðu máli er 2500 Hz á meðan F_2 fyrir gervi sérhljóðið /i/ hafði tíðni 2900 Hz. Delattre og félagar töldu vera tvær líklegar skýringar fyrir þessu: (1) F_3 hefur mikil áhrif á gæði sérhljóðsins eða (2) að eyrað tekur meðaltal af einungis tveimur formendum. Þar sem F_2 og F_3 eru nálægt hver öðrum þá notuðu Delattre og félagar meðaltalið á milli F_2 og F_3 og notuðu það sem F_2 . Þar sem tíðni F_3 breytist afskaplega lítið á milli sérhljóða þá hefur það freistað margra

vísindamanna til að telja F_3 ómerkilegan en niðurstöður rannsókna Delattre og féлага sýna að þvert á móti þá hefur F_3 mikil áhrif (Delattre o.fl., 1952).

Klatt talgervillinn

Seinna meir breyttust talskynjunartilraunirnar með tilkomu tölvunnar. Við það urðu talgervlar mun betri og þróaðari. Sá talgervill sem náði mestum vinsældum var Klatt-talgervillinn en hann var hannaður af Dennis H. Klatt við MIT háskólann árið 1980. Það sem gerði hann svo vinsælan var hversu fullkominn hann var og að hver sem er gat sett hann upp í sinni tölvu, því Klatt birti forritið opinberlega 1980 (Jörgen Pind, 1997). Margir talgervlar höfðu verið hannaðir í gegnum tíðina, ýmist sem vélbúnaðar (*hardware*) eða sem hugbúnaðar (*software*), og margir hverjir misgóðir. Þeir bestu höfðu það þó sameiginlegt að formendurnir voru annaðhvort búnir til í raðkvísla- (*cascade*) eða í samhliðakvísla (*parallel*) talgervlum. Klatt ákvað að taka tvennt það besta úr fyrri talgervlum og setja saman. Svo í Klatt hugbúnaðar talgervlinum er hægt að velja bæði raðkvísla- og samhliðakvíslastillingu í myndun formenda. Munurinn á raðkvísla- og samhliðakvíslastillingu er sá að í raðkvísl er sveifluvidd formendatoppa sérhljóðanna sjálfgefin en í samhliðakvísl þarf að stilla sveifluvidd sérhvers formendatopps. En það sem samhliðakvíslin hefur fram yfir raðkvíslina er að hann getur gert önghljóð og lokhljóð en það getur raðkvíslin ekki (Klatt, 1980). Næsta áratug vann Klatt í að endurbæta talgervilinn þar sem hann bætti við möguleika á ólíkum röddum svo sem blásinni rödd (*breathy*), brakrödd (*creaky*) og strekktri rödd (*tense*). Einnig bætti hann við kvenröddum og gerði karlraddirnar raunverulegri. Þessi endurbætti talgervill nefnist KLSYN88 (Klatt og Klatt, 1990).

Talgervill Klatts er talsvert flókinn með hátt í 50 breytur en megingerð hans sést á mynd 1 hér að neðan.



Mynd 1. uppbygging Klatt- talgervilsins (með raðkvísl). Mynd fengin úr glærum Jörgens L. Pind, 2000.

Á myndinni hér að ofan sjást þrjár hljóðgjafar þar sem efsti gefur raddað hljóð, sá í miðjunni gefur blásturshljóð og neðsti gefur suðhljóð. Þessi þrjú hljóð geta því búið til rödduð hljóð eins og til dæmis sérhljóð, blásturshljóð eins og /h/ og einnig suðhljóð eins og /s/ og /f/. Aðrar breytur á myndinni eru f_1 , f_2 og f_3 sem standa fyrir tíðni fyrstu þriggja formendanna og b_1 , b_2 og b_3 standa fyrir bandbreidd fyrstu þriggja formendanna (Jörgen L. Pind, 2000).

Mel-kvarðinn

Mel-kvarði (mel scale) er kvarði fyrir tónhæð (*pitch*) þar sem jöfn bil á kvarðanum eiga að gefa jafnan mun á tónhæð. Viðmiðunarpunkturinn milli mel-kvarðans og venjulegrar tíðnimælinga er greindur með því að jafna (*equating*) 1000 Hz tón, 40 dB fyrir ofan þröskuld hlustandans, með tónhæð upp á 1000 mel. Undir 500 Hz eru hertz og mel kvarðinn

samsvarandi en eftir því sem tíðni mæld í Hz verður hærri því meira verður bilið sem hlustandinn greinir til að framleiða jafna stighækkun (*increment*) tónhæðar (Truax, 1999). Formúla til að breyta tíðni (f) í Hz í tónhæð (m) í melum er $m=2595 \log_{10}(1+f/700)$ (O'Shaughnessy, 1987).

Hljóðmyndun sérhljóða

Sérhljóðar geta verið búnir til í hvaða tónhæð (*pitch*) sem er innan tónsviðs (*range*) raddar talandans. Þó svo að tónhæð raddar breytist þá hafa sérhljóð ennþá að mestu sömu eiginleika, hvort sem talandinn öskrar eða talar rólega. Mismunandi sérhljóð eru eins og ólík hljóðfæri. Það er hægt að spila „tónleikatónhæð“ A (*concert pitch A*) á píanó, klarínnett eða fiðlu. Í hvert skipti verður sami tónn vegna stigs endurtekningar hljóðbylgjunar sem heildar, þ.e.a.s. grundvallartíðnin er sú sama. Eiginleikarnir verða öðruvísi vegna smávægilegra breytileika innan hvernar endurtekningar hljóðbylgjunnar. Þetta leiðir það af sér að yfirtónarnir (*overtones*) verða öðruvísi og því hljóðblærinn (*timbre*). Þegar við hlustum á sérhljóð eða tónlistarnótu þá getum við greint hvaða sérhljóð það er eða hvaða hljóðfæri skapaði út frá yfirtónunum sem koma fyrir. Blað (*reed*) klarínettsins eða raddböndin geta sveiflast (*vibrate*) 100 sinnum á sekúndu, en hljóðið sem er framleitt af munnni klarínettsins eða varanna hefur ólíka einkennishópa af yfirtónum (Ladefoged og Disner, 2012).

Staða tungu og vara í myndun sérhljóða

Öll sérhljóð eru mismunandi vegna stöðu tungu og vara. Þetta er vegna þess að tungan og varirnar þrengja fyrir farveg loftsins. Þegar þrengt er fyrir farveg hljóðbylgju þá hækkar tíðni hennar og þannig geta tungan og varirnar mótað hljóðbylgjuna í mismunandi sérhljóð. Tungan getur þrengt að farvegi hljóðbylgjunnar snemma eða seint á ferð hennar í gegnum munninn.

Þess vegna er greint á milli framsérhljóða og baksérhljóða. Undir framsérhljóða floknum eru þá /i/, /I/, /ε/ og /æ/ vegna þess að tungan er svo framarlega í myndun þeirra. En í baksérhljóða floknum eru /a/, /ʊ/ og /u/ því í myndun þeirra er tungan aftarlega. Þessir tveir flokkar eru svo flokkaðir frekar niður eftir hæð tungunnar þar sem /i/ er hár fremri sérhljóði; /I/ mið-hár fremri sérhljóði; /ε/ mið-lár fremri sérhljóði; og /æ/ lár fremri sérhljóði. Baksérhljóðarnir eru einnig flokkaðir í /a/ lágur baksérhljóði; /ʊ/ mið baksérhljóði; og /u/ hár baksérhljóði (Ladefoged og Disner, 2012).

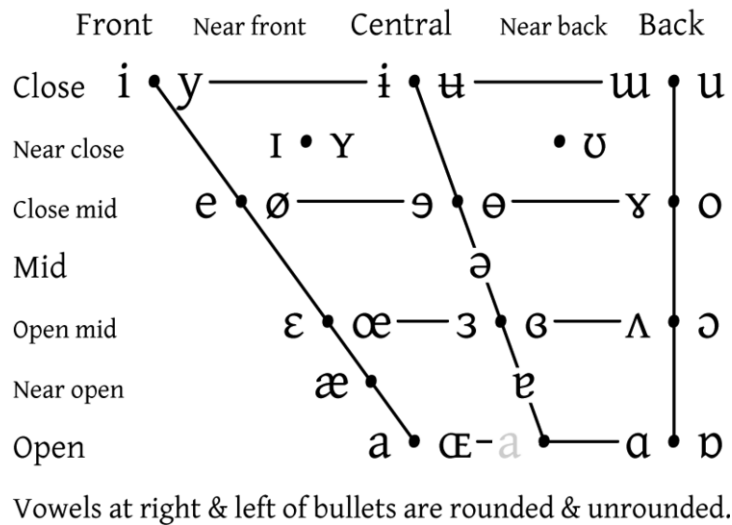
Tungan er ekkert nema vöðvi sem er fastur við kjálkana og höfuðkúpuna. Aðalvöðvin sem stýrir tungunni er höku- og tunguvöðvinn (*genioglossus*) en hann togar í rætur tungunnar í átt að kjálkabeininu (*mandible*). Þegar þessi vöðvi togar í tunguna þá fer fremri hlutinn upp í átt að harða gómnum (*hard palate*). Annar vöðvi sem er fastur við höfuðkúpuna rétt fyrir neðan eyrað sem kallast stíls- og tunguvöðvu (*styloglossus*) togar tunguna til baka og niður. Einnig er hægt að stýra hæð tungunnar með jaxla- og málbeinsvöðva (*mylohydoideus*) sem er undir tungunni á milli kjálkabeinanna. Ef þessi vöðvi dregst saman þá hækkar tungan. Vöðvarnir sem hafa mest áhrif á tungubroddinn eru efri langvöðvarnir (*superior longitudinalis*) en þeir eru beint undir hvorri hlið tungunnar. Þegar þeir dragast saman þá dregst efri partur tungunnar saman og tungubroddurinn krullast upp á við. Neðri langvöðvarnir (*inferior longitudinalis*) vöðvarnir halda áfram niður meðfram tungunni og þegar þeir dragast saman ýta þeir tungunni upp á við. Allir þessir vöðvar vinna saman við að hreyfa tunguna en það skiptir minna máli hvaða vöðvi gerði hvað. Það sem skiptir meira máli er staða tungunnar þegar hljóðbylgjan kemur í gegnum munninn. Aðal vöðvinn sem stýrir vörunum er hringvöðvi munnsins (*orbicularis oris*) en einnig er kjálkinn töluvert notaðir til að móta varirnar (Ladefoged og Disner, 2012).

Algengi og mismunandi gerðir sérhljóða í heiminum

Það er erfitt að svara þeirri spurningu hversu margir sérhljóðar eru til í öllum tungumálum heimsins vegna þess að mörk þeirra eru oft óljós. Það stafar af því að margir þættir koma að gerð sérhljóða svo sem tunga og varir. Rannsóknir Diane Kewley-Port hafa sýnt fram á að hlustendur geta greint að sérhljóð með einungis 12 Hz mun og að hlustendur geta greint 40 mismunandi hæðir tungunnar sem þýðir að það eru 40 mismunandi stig fyrsta formanda sem við getum greint. Einnig gátu hlustendur greint að með 5 til 30 gráðu nákvæmni hversu aftarlega tungan var sem segir til um tíðni annars formanda en nákvæmnin fór eftir því hversu hátt tungan var uppi. En í ónákvæmu hversdagstali þyrfti ábyggilega þrefalt meiri mun til að greina orð að heldur en 12 Hz og 5-30 gráður (Ladefoged og Disner, 2012).

Mikill munur er á fjölda sérhljóða milli tungumála. Sum tungumál nota einungis þrjú sérhljóða eins og flest tungumálin í Ástralíu, á meðan önnur nota allt að 22 sérhljóða eins og í sumum þýskum mállýskum en þar eru 12 löng sérhljóð og 10 stutt sérhljóð. Í þeim tungumálum sem eru með fæst sérhljóð þá eru sérhljóðin oftast /i/, /a/ og /u/. Þessi þrjú sérhljóð eru oftast valinn því þau eru lengst frá hvort öðru í sérhljóðarúminu (Ladefoged og Disner, 2012). (sjá mynd 2).

VOWELS



Mynd 2. Sérhljóðarúmið þar sem sérhljóðar sjást eftir því hvort þeir eru framarlega, miðlega eða aftarlega á tungunni og hvort þeir eru hátt, miðlægt eða neðarlega á tungunni. Myndin fengin af Wikipedia.org.

Svo virðist sem tungumál hafa tilhneigingu til að hafa oddatölufjölda sérhljóða. Í kringum 20 prósent allra tungumála í heimi eru með fimm sérhljóð. En þessi tilhneiging til að hafa sérhljóð í oddatölu gæti stafað af því að sérhljóðarúmið er þríkanta. Vegna flókinna samspila tungu og vara þá er ekki jöfn dreifing sérhljóða í sérhljóðarúminu. Það eru fleiri sérhljóð vinstra megin eins og /i/ og /e/ en hægra megin svo sem /u/ og /o/. Áhugavert er að vinsældir sérhljóða í tungumálum með fimm sérhljóð eru í réttri röð við stafrófið /a, e, i, o og u/ þar sem /a/ er algengast og /u/ er óalgengast (Ladefoged og Disner, 2012).

Í ensku eru frammælt sérhljóð búin til með ókringdum vörum á meðan aftari sérhljóðin eru búin til með kringdum vörum og kringingin eykst eftir því sem tungan er ofar í munninum. Þetta á við um flest tungumál heimsins. Eftir því sem varirnar kringjast meira þá lækkar tíðni allra þriggja formenda sérhljóðsins. En áhrifin eru mismikil eftir því hvaða sérhljóð er um að ræða. Formendur lækka mismikið eftir sérhljóðum við að kringja varirnar. Til dæmis hefur kringingin ekki mikil áhrif á F_1 og F_2 sérhljóðanna /i/ og /y/, en það sem

skilur þessi sérhljóð að er að kringingin lækkar tíðni F_3 mun meira hjá /y/ heldur en /i/ (Ladefoged og Disner, 2012).

Tilgangur rannsóknarinnar

Vísindamenn hafa ekki verið sammála um það hversu marga formendur það þarf til þess að greina sérhljóða. Sumir telja að tveir formendur séu nóg en aðrir telja að það þurfi þrjá formendur. Tilgangur rannsóknarinnar var tvíþættur. Í fyrsta lagi að kortleggja sérhljóðarými í íslensku. Í öðru lagi að athuga mun á tveimur og þremur formendum í skynjun á íslenskum sérhljóðum. Tilgáta rannsakenda er sú að sérhljóð með þremur formendum verði greind skýrari heldur en sérhljóð með tveimur formendum.

Aðferð

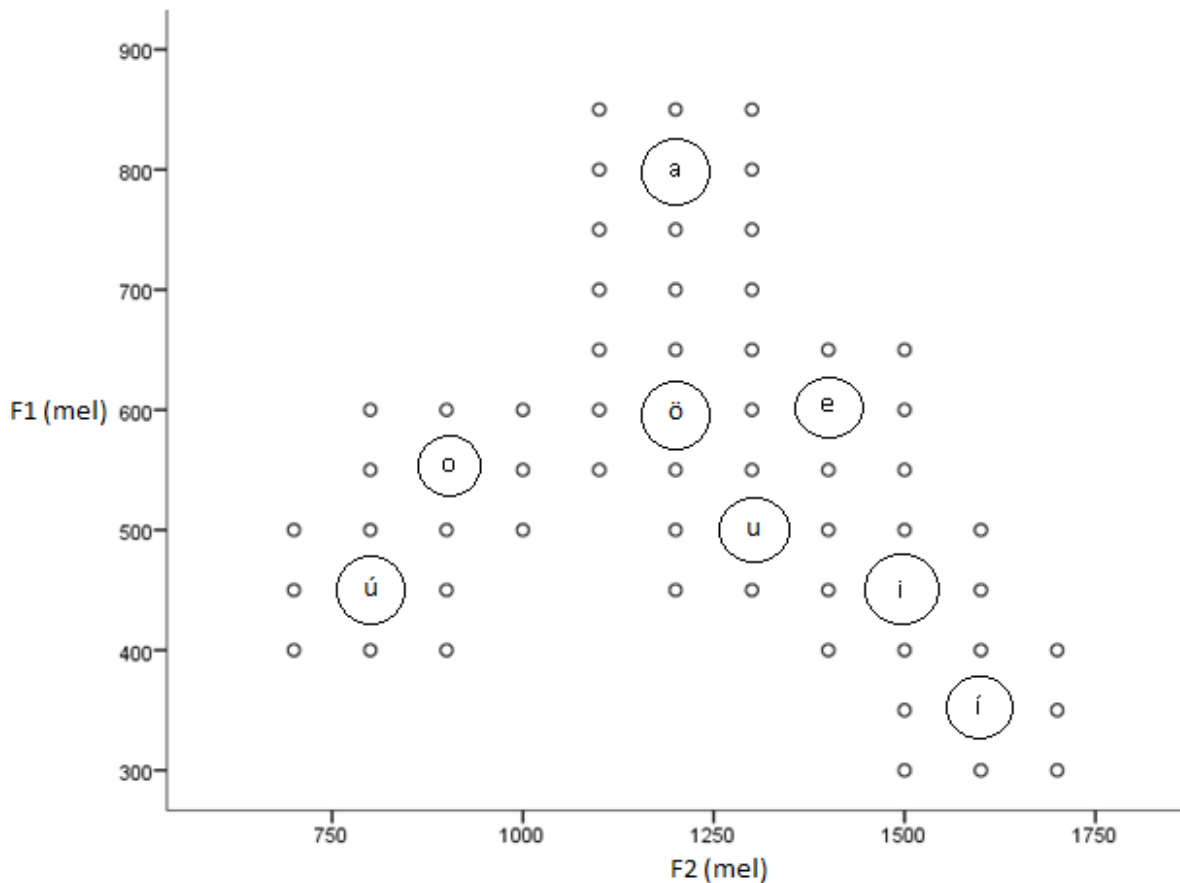
Þátttakendur

Í þessari rannsókn voru fengnir 17 þátttakendur. Þeir voru valdir af rannsakendum með hentugleikaúrtaki. Aldursspönn þátttakandana var frá 20-28 og meðaltalið 21,83 með staðalfrávikinu 2,5. Af þessum 17 þátttakendum voru 11 karlar ($M_{KK} = 21,8$, $SD_{KK} = 2,31$, $spönn_{KK}=20-26$) og 6 konur ($M_{KVK} = 22,2$, $STD_{KVK} = 3,32$, $Spönn_{KVK} = 20-28$).

Áreiti

Til að byrja með höfðu þátttakendur uppi á formendatíðni íslenskra sérhljóða sem Ásta Svavarsdóttir, Halldór Ármann Sigurðsson, Sigurður Jónsson og Sigurður Konráðsson höfðu fundið árið 1973. Þeir voru notaðir sem grunnur til að búa til 63 hljóð í kringum þá eða níu í kringum hvert sérhljóð. Hljóðin voru þannig gerð að F_1 var frá 400 melum að 800 melum í 50 mela skrefum en F_2 frá 800 melum að 1600 melum í 100 mela skrefum. Til að breyta Hz í mel var notað formúluna $m=2595 \log_{10}(1+f/700)$. Hljóðin voru annaðhvort með tveimur formendum eða þremur formendum svo alls voru hljóðin 126 talsins (sjá mynd 3). Ef formendur voru þrír var F_3 alltaf 3500 Hz.

Seinna kom í ljós að gerðar höfðu verið tvær innsláttarvillur þegar tilraunin hafði verið forrituð. Fyrri villan fól í sér að slegið var inn 778 Hz þegar það átt að vera 788 Hz. Hin villan fól það í sér að 1158 Hz kom óvart helmingi oftar fyrir en skyldi sem þýddi að 1330 Hz og 1519 Hz komu helmingi sjaldnar fyrir en þau áttu að gera.



Mynd 3. Sérhljóðin í melum eftir F1 og F2.

Mælitæki

Í þessari rannsókn var forritið Praat (hollenska orðið yfir að tala) notað en það forrit er hægt að sækja á netinu og er ætlað til þess að greina og vinna með talmál. Paul Boersma og David Weenink bjuggu það til í háskólanum í Amsterdam og hefur forritið verið í stöðugri þróun síðan þá. Forritið býður upp á að búa til hljóð með Klatt-talgervlinum með fyrirfram ákveðnum formendum sem gerir þetta forrit tilvalið fyrir þessa rannsókn. Annar kostur við forritið er að það býður einnig uppá að búa til hlustunartilraunir og keyra þær í forritinu.

Framkvæmd

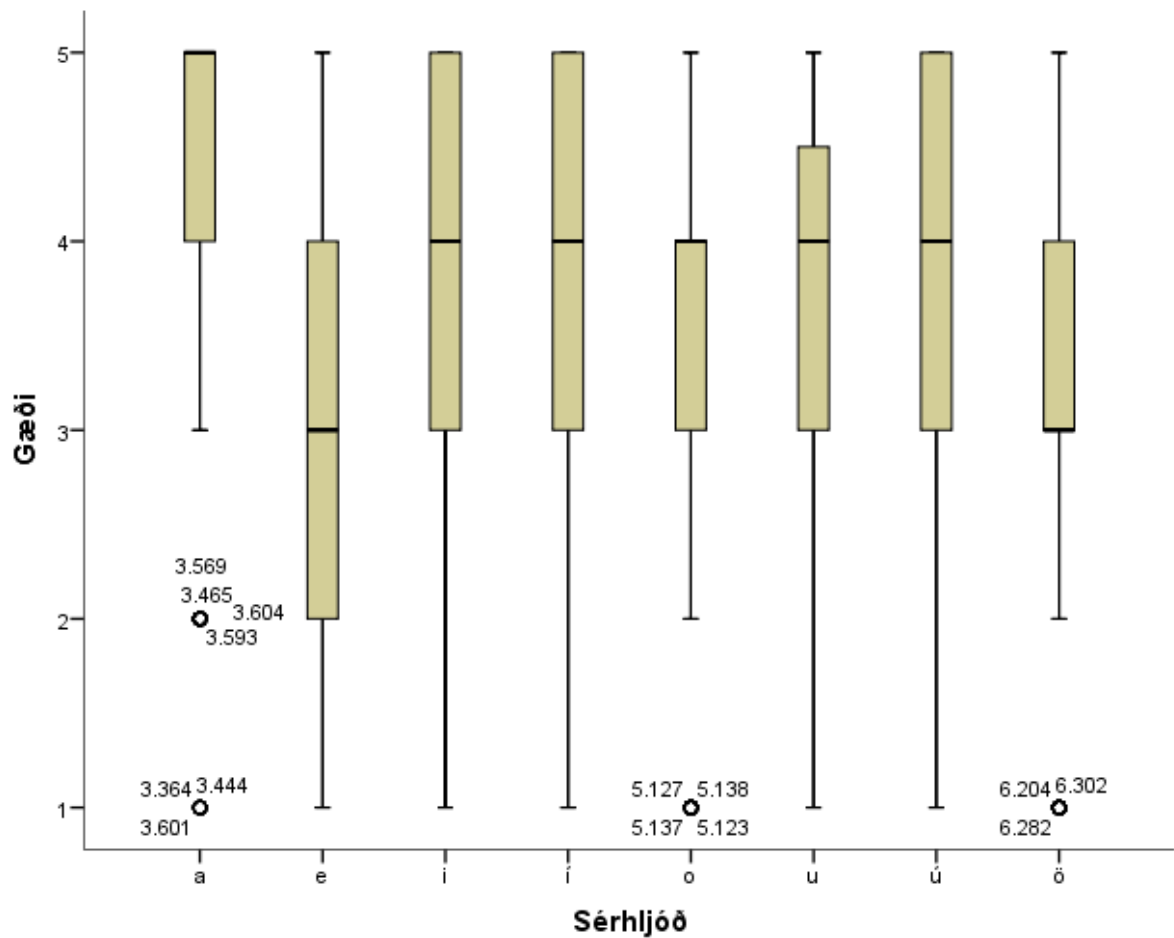
Þegar öll hljóðin höfðu verið búin til í Praat forritinu þá var tilraunin sett upp í Praat þar sem hljóðin heyrðust og þátttakendur áttu að segja hvaða hljóð heyrðist og velja gæði þeirra á fimm punkta Likert skala frá einum og uppi fimm. Hvert hljóð var spilað þrisvar sinnum í tilviljunarkenndri röð. Þátttakendur voru beðnir um að koma í sálfræðitilraunastofu við Háskóla Íslands. Inni í tilraunastofunni settust þátttakendur inn í hljóðklefa sem innihélt tölvu og heyrnartól. Þátttakendur voru spurður hvort þeir væru með einhverjar heyrnarskemmdir til að koma í veg fyrir skekkjur í niðurstöðum. Fyrst var keyrð prufutilraun sem innihélt 20 tölvugerð sérhljóð til að venja þátttakendur að tölvugerðu hljóðunum áður en þeir voru látnir taka alvöru tilraunina. Hljóðin voru spiluð á þægilegum styrk. Þátttakendum var síðan gert ljóst að þeir mættu hætta þátttöku hvenær sem er án afleiðinga. Þeir sem voru í sálfræðinámi gátu fengið rannsóknareiningu fyrir þátttöku sína.

Tölfræðileg úrvinnsla

Notast var við tölfræði forritið SPSS 20.0 fyrir Windows til þess að vinna úr gögnunum. En áður en hægt var að flytja gögnin frá Praat og yfir í SPSS þurfti að sameina þau í eitt Microsoft Excel skjal sem var síðan hægt að opna í SPSS.

Fyrst voru gögnin skoðuð myndrænt með dreifiriti, leifarriti, stöplariti og normalriti til að meta hvort forsendur fyrir dreifigreiningu væri til staðar. Forsendur dreifigreiningar eru annars vegar að dreifing villunnar sé einsleit (*homoscedasticity*) þ.e.a.s. að staðalfrávik hópanna sem bera á saman séu eins og hins vegar að leifin normaldreifist eða að það séu engir frávillingar.

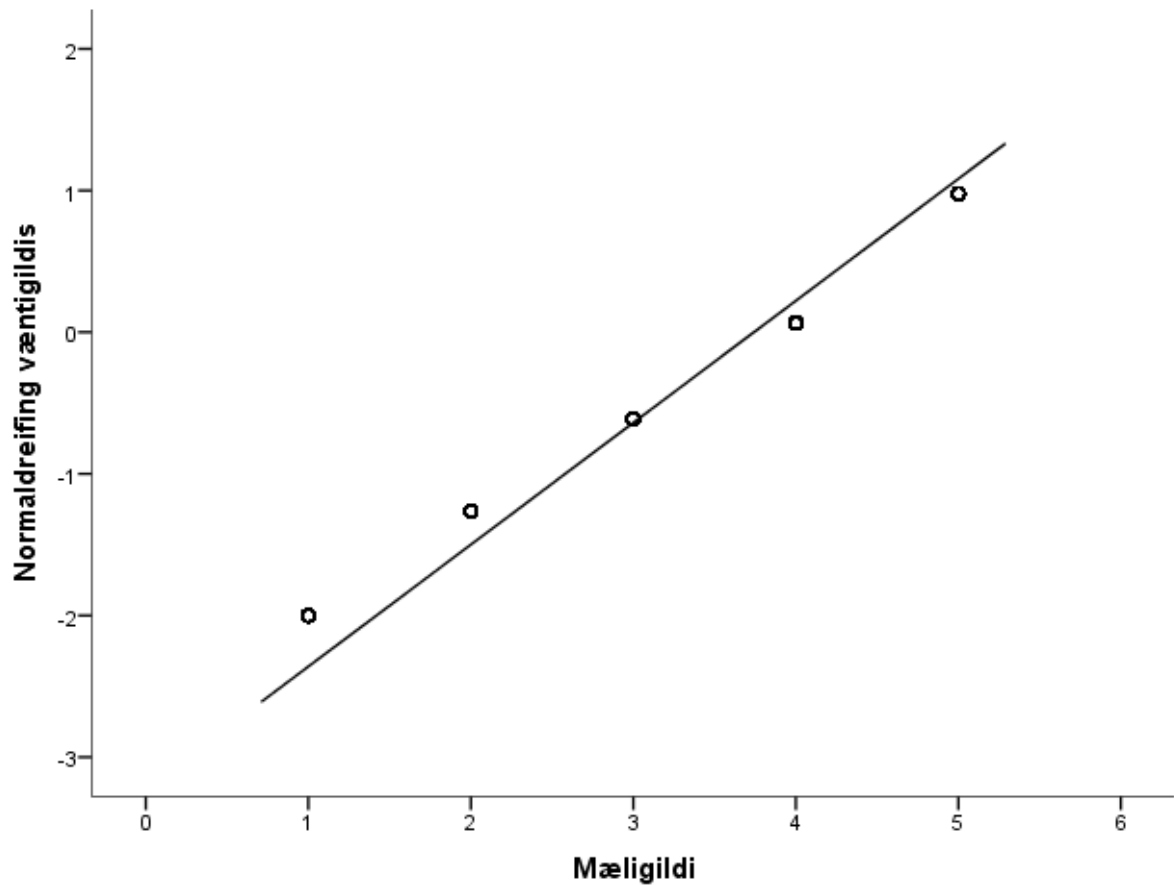
Til að athuga hvort frávillingar væru í gögnum voru kassarit skoðuð eins og sést á mynd 4.



Mynd 4. Dreifing breytunnar gæði eftir sérhljóðum.

Á myndinni sjást 14 svör fyrir utan fjórðungsspönnina sem gætu verið frávillingar en þar sem dreifigreining er nokkuð traust gagnvart frávillingum þá hefur það ekki afgerandi áhrif á niðurstöður.

Til að skoða aðra forsendu normaldreifingar hvort að leifin normaldreifist er leifarrit skoðað eins og á mynd 5.



Mynd 5. Dreifing leifar um normallínu.

Þar sést að leifin normaldreifist vel um normallínu svo önnur forsenda dreifigreiningar um normaldreifingu leifar er heldur ekki brotinn.

1. Fyrsta forsenda fyrir dreifigreiningu er ekki brotinn þar sem staðalfrávik fyrir báða hópana eru nánast þau sömu (1,18 og 1,14).
2. Önnur forsenda um normaldreifingu leifar er heldur ekki brotinn þar sem leifin normaldreifist ágætlega og engir alvarlegir frávillingar voru til staðar.

Þar sem báðar forsendur dreifigreiningar eru til staðar þá var haldið áfram og dreifigreining framkvæmd.

Niðurstöður

Áberandi margir sögðust heyra sérhljóðið /a/ frekar en önnur sérhljóð þar sem /a/ var valið 1089 sinnum eða í 16,9% tilvika (sjá töflu 2). /e/ var valið næst sjaldnast í 584 skipti eða 9,1% og /i/ var valið sjaldnast í 514 skipti eða um 8%. /í/ var valið 734 sinnum sem jafngildir 11,4% en /o/ var valið nokkuð oftar eða 887 sinnum eða 13,8%. /u/ var valið örlítið oftar en /o/ eða um 896 sinnum sem nemur að 13,9%. Færri völdu /ú/ en /o/ og /u/ eða um 819 eða 12,7%. Næstvinsælasta hljóðið sem var valið í 14,1% tilvika var /ö/ en það var valið 903 sinnum (sjá töflu 2).

Aðeins fleiri völdu sérhljóðið /a/ með þremur formendum heldur en með tveimur formendum eða 0,5% oftar (sjá töflu 2). Sérhljóðið e var hins vegar 0,6% vinsælla þegar sérhljóðin voru með tveimur formendum heldur en þegar þau voru með þremur formendum. Það var ekki mikill munur á sérhljóðinu /i/ eða einungis 0,2% fleiri völdu það þegar sérhljóðin voru með tveimur formendum. Sérhljóðið /í/ var valið 28 sinnum oftar þegar sérhljóðið var með þrjá formendur heldur en þegar það var með tvo formendur sem jafngildir 0,9% oftar. /o/ var valið 0,6% oftar í þriggja formenda sérhljóði heldur en í tveggja en /u/ var nánast valið jafn oft í báðum tilvikum hvort sem það voru tveir eða þrír formendur. Ekki var mikill munur heldur á sérhljóðinu /ú/ en 0,3% völdu það oftar í sérhljóðum með tvo formendum heldur en í sérhljóðum með þremur formendum. Að lokum var stærsti munurinn á sérhljóðinu /ö/ þar sem 1,1% fleiri völdu það þegar sérhljóðin voru með tvo formendur heldur en þegar þau voru með þrjá (sjá töflu 2). Svo ekki mikill munur var á því hversu oft hvert sérhljóð var valið hvort sem sérhljóðin voru með tvo eða þrjá formendur.

Tafla 2. Tíðni og prósentu valdra sérhljóða eftir tveimur og þremur formendum

Sérhljóð	2 Formendur		3 Formendur		Heild	
	Tíðni	Prósent(%)	Tíðni	Prósent(%)	Tíðni	Prósent (%)
a	536	16,7	553	17,2	1089	16,9
e	302	9,4	282	8,8	584	9,1
i	259	8,1	255	7,9	514	8,0
í	353	11,0	381	11,9	734	11,4
o	433	13,5	454	14,1	887	13,8
u	446	13,9	450	14,0	896	13,9
ú	415	12,9	404	12,6	819	12,7
ö	469	14,6	434	13,5	903	14,1
Samtals	3213	100,0	3213	100,0	6426	100,0

Gæði sérhljóðanna var í hærri kantinum þar sem fjórir og fimm var valið töluvert oftur heldur en eitt og tvö á gæða skalanum (sjá töflu 3). Sérhljóðin þóttu í hinum bestu gæðum í 32,8% tilvika eða 2107 sinnum af 6426. Næstbestu gæðin voru valinn í 29,3% tilvika eða 1880 sinnum en sámileg gæði voru valinn aðeins sjaldnar eða 1406 sinnum eða 21,9% tilvika. Næstverstu gæðin voru valið 742 sinnum eða 11,5% á meðan verstu gæðin voru einungis valinn í 4,5% tilvika eða 291 sinni (sjá töflu 3).

Tafla 3. Tíðni og prósentu gæða sérhljóða.

Gæði	Tíðni	Prósentu (%)	Uppsafnað (%)
1(mjög vont)	291	4,5	4,5
2(vont)	742	11,5	16,1
3(ágætt)	1406	21,9	38,0
4(gott)	1880	29,3	67,2
5(mjög gott)	2107	32,8	100,0
Samtals	6426	100,0	

Ef gæðin eru skoðuð eftir fjölda formenda þá eru niðurstöðurnar mjög svipaðar (sjá töflu 4). Þegar formendur voru þrír þá völdu 4% verstu gæðin en 5,1% þegar formendurnir voru tveir. Næst verstu gæðin voru valin oftast þegar formendur voru tveir heldur en þegar formendur voru þrír eða 0,7% oftast. Ágæt gæði voru valinn 0,9% oftast þegar sérhljóðið var með þremur formendur heldur en tveimur formendur. Næstbestu gæðin voru valin 0,1% oftast þegar formendur voru tveir heldur heldur en þrír og bestu gæðin voru valinn 1,2% oftast þegar það voru þrír formendur (sjá töflu 4).

Tafla 4. Tíðni og prósentu gæða sérhljóða eftir tveim eða þrem formendum.

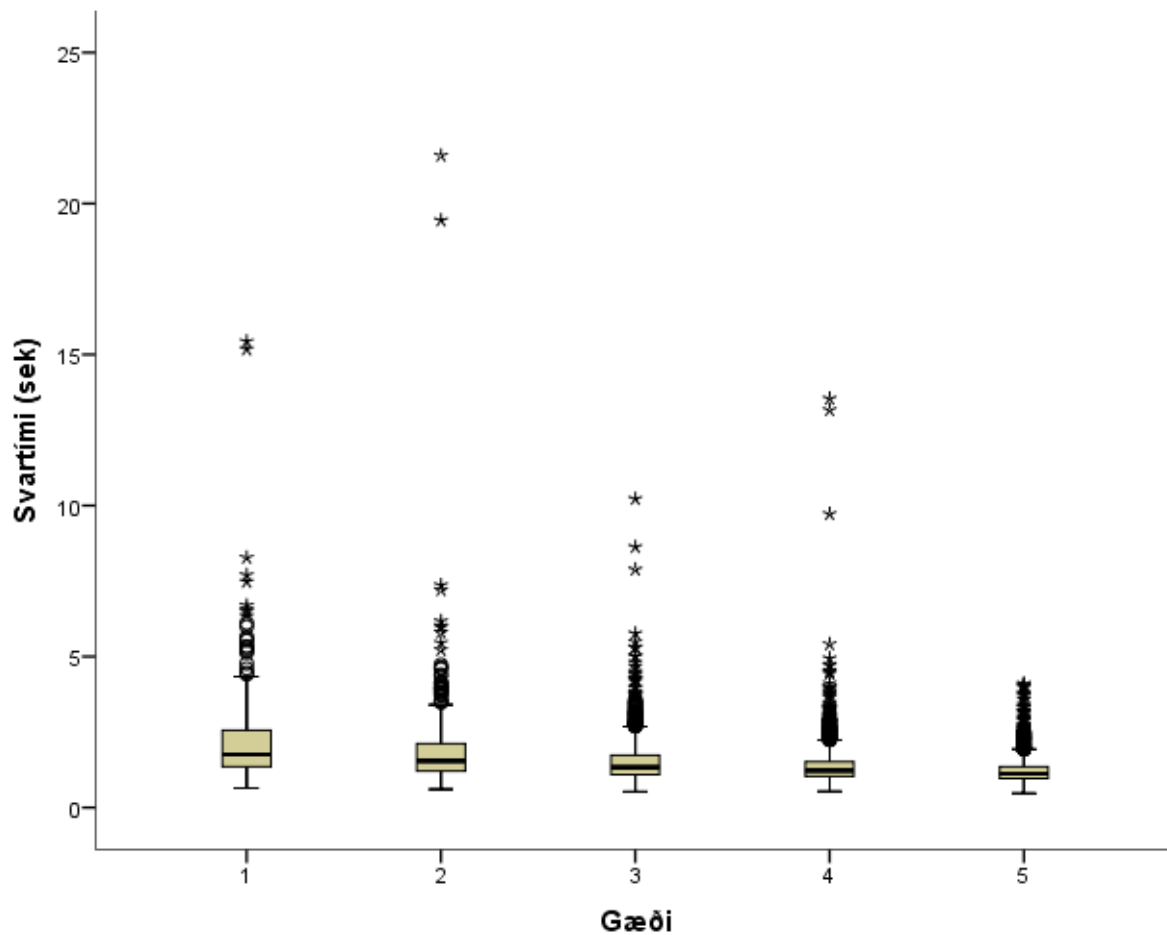
Gæði	2 Formendur		3 Formendur	
	Tíðni	Prósenta(%)	Tíðni	Prósenta(%)
1	164	5,1	127	4,0
2	383	11,9	359	11,2
3	689	21,4	717	22,3
4	943	29,3	937	29,2
5	1034	32,2	1073	33,4
Samtals	3213	100,0	3213	100,0

Meðalsvartími þátttakanda í tilrauninni var 1,45 sekúnda (sjá töflu 5). Stysti svartíminn var 47 sekúndubrot en lengsti svartími náði 21,59 sekúndum. Staðalfrávik svartímanna var 0,85 sekúnda (sjá töflu 5).

Tafla 5. Lýsandi tölfræði svartíma.

	Lægsta gildi	Hæsta gildi	Meðaltal	Staðalfrávik
Svartími	0,47	21,59	1,45	0,85

Svartíma má sjá myndrænt eftir gæðum á mynd 6. Þar má sjá skýra hneigð að svartímar styttest með hæra gæðamati, til dæmis fór enginn fram yfir fimm sekúndur þegar gæði sérhljóðanna voru mjög góð (5).



Mynd 6. Kassarit fyrir svartími eftir gæðum

Sérhljóðið /a/ þótti mun skýrara heldur en öll hin sérhljóðin með gæði að meðaltali 4,28 af 5 mögulegum og minnsta staðalfrávikinu (sjá töflu 6). Næst skýrasta sérhljóðið var /í/ með meðaltalið 3,99 og þar rétt á eftir var /ú/ með 3,91. Sérhljóðið /i/ þótti nokkuð óskýrara heldur en /ú/ með 3,69. Sérhljóðið /u/ þótti örlítið óskýrara heldur en /i/ með meðaltalið 3,65. Þriðja óskýrasta hljóðið var /o/ með meðaltalið 3,51 en þar á eftir var /ö/ með 3,41 sem gerir það næst óskýrasta hljóðið. Sérhljóðið /e/ fékk lægstu einkunn í gæðum með meðaltalið 3,24 sem er rúmlega heilum minna heldur en sérhljóðið A (sjá töflu 6).

Tafla 6. Meðaltal og staðalfrávik breytunnar gæði sérhljóða.

Sérhljóð	Meðaltal	Lægst gildi	Hæsta gildi	Staðalfrávik
a	4,28	1	5	0,97
e	3,24	1	5	1,14
i	3,69	1	5	1,13
í	3,99	1	5	1,10
o	3,51	1	5	1,17
u	3,65	1	5	1,08
ú	3,91	1	5	1,19
ö	3,41	1	5	1,19

Sérhljóð með þremur formendum þótti skýrari heldur en sérhljóð með tveimur formendum (sjá töflu 7). Skýrleika tveggja formenda sérhljóða var með meðaltalið 3,72 á meðan meðaltal fyrir sérhljóð með þrjú formendur var 3,77. Þessi munur uppá 0,05 er ekki mikill en er þó í samræmi við tilgátu rannsóknarinnar.

Tafla 7. Meðaltal breytanna tveggja og þriggja formenda.

Formendur	Meðaltal	Staðalfrávik
2	3,72	1,18
3	3,77	1,14

Til að sjá hvort þessi munur á milli meðaltala hópanna tveggja væri marktækur eða einungis tilviljunarkenndur þá var framkvæmd einsleit dreifigreining (*e: one way ANOVA*). Samkvæmt F-prófi dreifigreiningarinnar $F(1, 6426) = 3,33$, $P = 0,068$ miðað við alpha 0,05 er munurinn

ekki marktækur. Aðaltilgátan að þriggja formenda íslenskt sérhljóð séu skýrari heldur en tveggja formenda sérhljóð er því ekki studd.

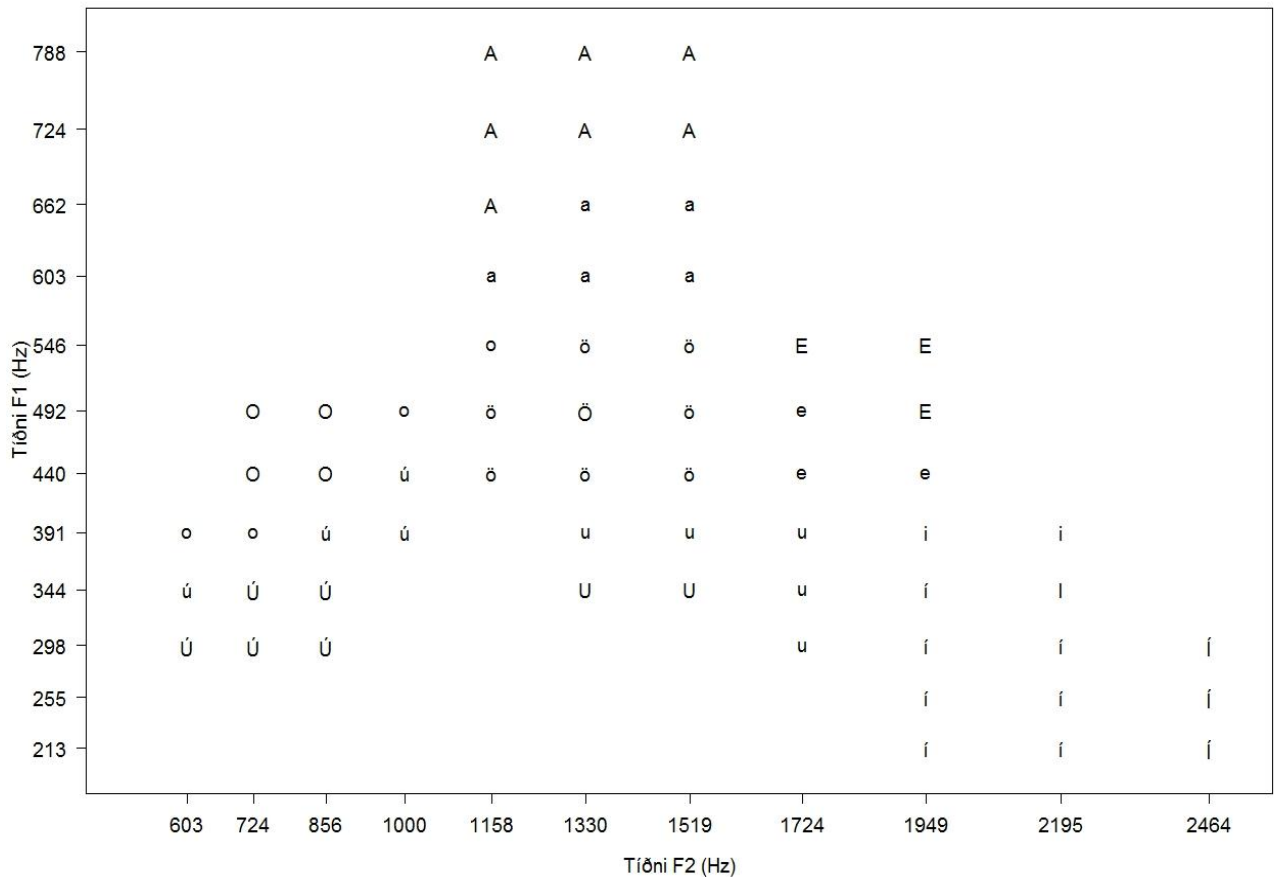
Þegar dreifgreining er framkvæmd fyrir hvert sérhljóð fyrir sig kemur í ljós að þrír formendur hafa meiri áhrif á ákveðin sérhljóð heldur en önnur þótt að engin þeirra séu marktæk (sjá töflu 8). Þrír formendur hafa ekki til dæmis jafn mikil áhrif á sérhljóðið /a/ og /e/ eins og það hafði á /i/ og /ö/. Þrír formendur höfðu ekki mikil áhrif á sérhljóðið /í/ né /o/ og ekki heldur /u/ og /ú/.

Tafla 8. Dreifgreining fyrir mun á tveimur eða þremur formendum fyrir hvert sérhljóð.

Sérhljóð	F	P-gildi
a	0,226	0,634
e	0,079	0,778
i	2,693	0,101
í	0,088	0,767
o	0,251	0,616
u	0,104	0,747
ú	0,032	0,859
ö	1,907	0,168

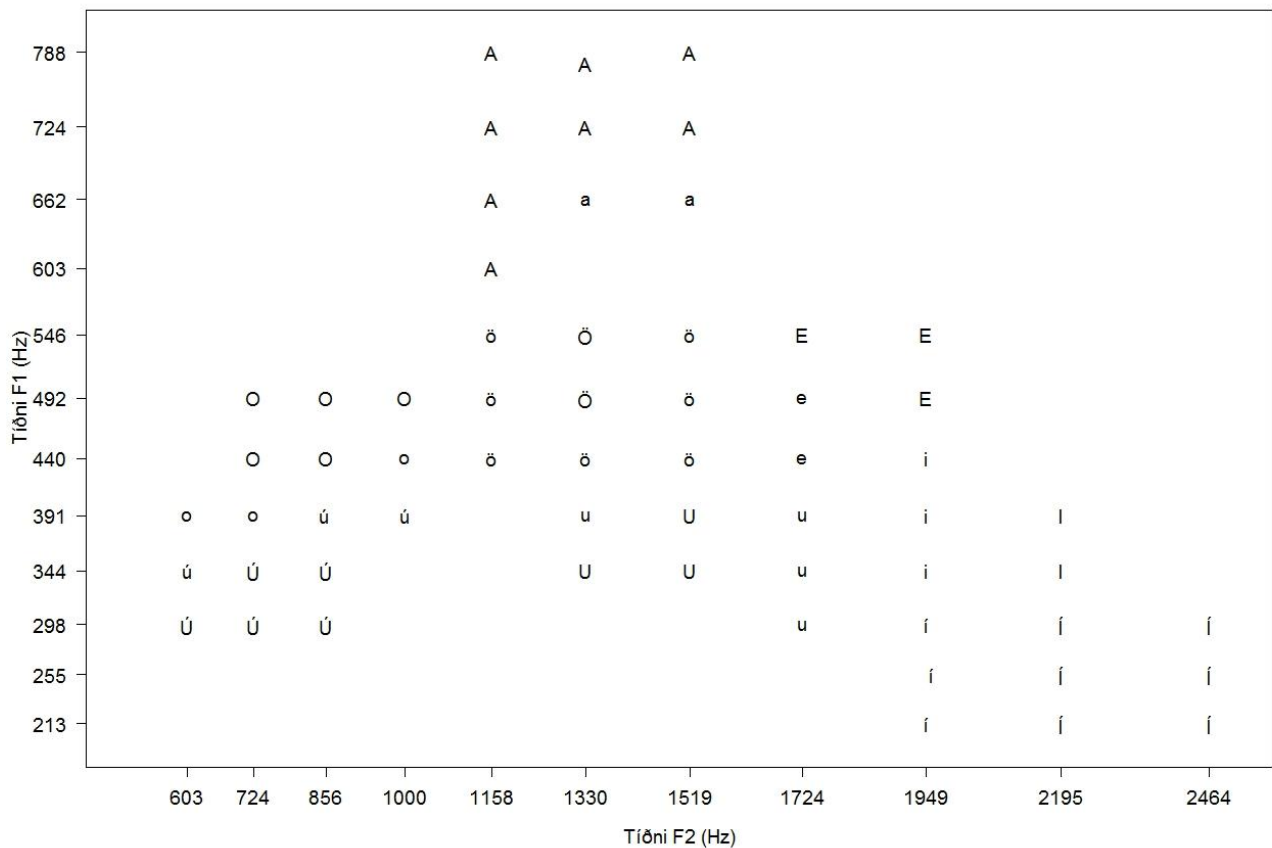
Á mynd 7 sést hvernig áreitni voru flokkuð fyrir tvo formendur. Áreitni sem voru flokkuð með samræmi yfir 80% voru táknuð með hástöfum en önnur með lægra en 80% voru táknuð með lágstöfum. Á myndinni sést vel hvernig sérhljóðarnir raðast eftir tíðni fyrsta og annars formenda. Sérhljóðið /a/ sést einna best efst á myndinni en það sérhljóð er töluvert aðskilið frá hinum sérhljóðunum og það getur verið hugsanleg ástæða fyrir því að það sé valið oftast eða tólf sinnum. En sérhljóðið /i/ kemur sjaldnast fyrir eða þrisvar sinnum, en það gæti verið

vegna þess að það er mitt á milli þriggja annarra sérhljóða /e/, /i/ og /u/. Á myndinni sést vel að flestir hástafir eru á jaðri myndarinnar nema /ö/ sem er akkúrat í miðjunni (sjá mynd7).



Mynd 7. Flokkun áreita fyrir tvo formendur (hástafir þýða að samræmi var yfir 80 % á flokkun áreitanna).

Á mynd 8 hér fyrir neðan sést að hástafirnir koma einnig hér aðalega fyrir í jaðri myndarinnar eins og á mynd 7 fyrir tvo formendur. Á myndinni sést einnig að sérhljóðin eru 61 en ekki 63 þar sem það vantar tvö sérhljóð vegna innsláttarvillu í upphafi tilraunarinnar og eitt /a/ er í 778 Hz en ekki 788 Hz eins og það á að vera. Þar sést að sérhljóðið /a/ kemur oftast fyrir eða tíu sinnum og sérhljóðin /i/ og /e/ koma sjaldnast fyrir eða fimm sinnum.



Mynd 8. Flokkun áreita fyrir þrjá formendur (hástafir þýða að samræmi var yfir 80 % á flokkun áreitanna).

Þegar mynd 7 fyrir tvo formendur og mynd 8 fyrir þrjá formendur eru bornar saman kemur í ljós að það eru 26 hástafir þegar sérhljóðin eru með tvo formendur og 34 þegar sérhljóð eru með 3 formendur svo það eru átta fleiri hástafir á mynd 8 heldur en á mynd 7. Það þýðir að meira samræmi er í flokkun áreitanna þegar sérhljóðin eru með þremur formendum heldur en þegar sérhljóðin eru með tveimur formendum. Þegar myndir 7 og 8 eru bornar saman við mynd 3 sést greinilega að niðurstöður þessarar rannsóknar eru í samræmi við niðurstöður Ástu Svavarsdóttur og félaga um sérhljóðarými í íslensku.

Umraeda

Vísindamenn víðs vegar um heiminn hafa ekki verið sammála um hversu marga formendur þurfi til að greina að sérhljóð. Sumir ganga svo langt að segja að einn formandi nægi til að greina að öll sérhljóð þótt fáir taka undir það en flestir vilja meina að tveir formendur nægja til að greina að flest sérhljóð en þó þurfi þrjá formendur til að greina að örfáa sérhljóða. Ásta Svavarsdóttir vildi meina að frammælt sérhljóð svo sem /i/ og /e/ þyrftu þrjá formendur til að greina þau en fyrir önnur sérhljóð nægðu tvo formendur. Niðurstöður þessarar rannsóknar styðja að einungis þurfi tvo formendur til að greina að öll stutt íslensk sérhljóð og eru því í samræmi við niðurstöður M. Joos en eru ekki í samræmi við niðurstöður Ástu Svavarsdóttir og félagar og Delattre og félagar. Niðurstöður þessarar rannsóknar eru hins vegar tölfræðilega takmarkaðar vegna þess að þriðji formandi allra sérhljóðanna var með fast gildi svo niðurstöðum þessarar rannsóknar skal taka með fyrirvara.

Myndir 7 og 8 sýndu að 34 sérhljóð voru valinn með yfir 80% samræmi þegar sérhljóðin voru með þrem formendum en 26 sérhljóð voru valinn með yfir 80% samræmi þegar sérhljóðin voru með tveimur formendum. Þetta gæti verið vísbending um að þriggja formenda sérhljóð er talið skýrara heldur en tveggja formenda sérhljóð.

Tilgáta rannsakanda stóðst samt aðeins að hluta. Það var ekki marktækur munur á gæðum sérhljóða eftir fjölda formenda þeirra. Munurinn var eiginlega enginn. Þessi niðurstaða kom rannsakendum á óvart. Það sem spilar hugsanlega inn í niðurstöðurnar er að þriðji formendinn hafði alltaf fast gildi, 3500 Hz. Í rannsókn Ástu Svavarsdóttur og félagar fundu þau fyrstu þrjá formendurnar fyrir alla íslensku sérhljóðanna. Hefði verið notast við gildin fyrir þriðja formendann í öllum sérhljóðunum þá hefðu niðurstöðurnar hugsanlega orðið öðruvísi. Annar hugsanlegur möguleiki fyrir þessum niðurstöðum er að rannsókn Ástu Svavarsdóttir var birt fyrir 30 árum. Mælitækin sem þau notuðu eru ekki næstum því eins góð og þau mælitæki sem eru í boði í dag. Þess vegna gæti verið að sú meðaltíðni sem þau fengu fyrir formendur

íslenskra sérhljóða séu ekki alveg nákvæm. Það þyrfti að endurtaka þessa tilraun með nýjustu forritum og þá væri hægt að fá nákvæmari niðurstöður.

Sérhljóðinn /a/ var talinn vera skýrastur af þátttakendum. Meðtalið fyrir hann var 4,28 en hann var eini sérhljóðinn sem var með meðaltal hærra en 4. Þetta kom rannsakendum ekkert á óvart. Þegar þeir prufukeyrðu tilraunina á sjálfum sér voru þeir báðir sammála um það að /a/ var lang skýrasti sérhljóðin. Þetta kom líka í ljós þegar rannsakendum spurðu þátttakendur eftir að þeir hefðu tekið þátt í tilrauninni. Flestir ef ekki allir þátttakendur sögðu að /a/ hafi verið skýrasta sérhljóðið. Skýringin fyrir því gæti verið að hann er einhljóði og er nokkuð langt frá öðrum sérhljóðum í Hz tíðni og því er ekki verið að rugla honum saman við önnur hljóð. Að sama skapi var sérhljóðið /e/ talið vera óskýrast. Það var með meðaltal upp á 3,24 og næst á eftir kom /ö/ með meðaltal upp á 3,41. Hugsanleg ástæða fyrir því að þessir sérhljóðar séu taldir vera óskýrastir er að þeir eru nálægt hvor öðrum í Hz tíðni, /e/ hefur 508 Hz í F_1 og 1588 í F_2 og /ö/ hefur 521 Hz í F_1 og 1325 Hz í F_2 fyrir stutt hljóð í karlmansrödd. Þetta gæti útskýrt það í sumum tilvikum að erfitt var að greina á milli þessara sérhljóða og það gæti útskýrt af hverju þeir tveir eru taldir vera óskýrustu sérhljóðarnir.

Vandkvæði rannsóknar voru innsláttarvillur í forritun rannsóknarinnar þar sem eitt hljóð var birt helmingi oftar en það átti að vera birt sem leiddi til þess að tvö hljóð voru birt helmingi sjaldnar. Einnig var eitt sérhljóð birt í helming tilvika með 10 Hz lægri fyrsta formenda heldur en það átti að vera. Þessi mistök gætu skekkt niðurstöður eitthvað örlítið en hafa þó varla nein afgerandi áhrif. Annar vankantur rannsóknarinnar gæti verið að þegar tíðni formendanna var færður frá Hz og yfir í mel þá var tíðni fyrsta formenda námunduð að 50 mel og tíðni annars formenda námunduð að 100 mel. Þetta gæti einnig valdið örlitlum skekkjum í niðurstöðum rannsóknarinnar.

Þessi rannsókn innihélt einungis stutt afbrigði íslenskra sérhljóða en áhugavert væri að rannsaka hvort þrjár formendur hefðu meiri áhrif á löng afbrigði íslenskra sérhljóða heldur en

Þeir höfðu á stutt afbrigði íslenskra sérhljóða. Þessi rannsókn var líka einungis með karlmannsrödd þar sem grunntóninn var 120 Hz en framtíðarrannsóknir gætu athugað hvort þriðji formendi myndi hafa meiri áhrif fyrir kvenmannsraddir heldur en þriðji formendi hafði fyrir karlmannsraddir á íslensk sérhljóð.

Heimildir

- Ásta Svavarsdóttir, Halldór Ármann Sigurðsson, Sigurður Jónsson og Sigurður Konráðsson. (1982). Formendur íslenskra einhljóða: meðaltíðni og tíðnidreifing. *Íslenskt mál og almenn málfræði*, 4, 63-85.
- Delattre, P., Liberman, A. M., Cooper, F. S. og Gerstman, L. J. (1952). An experimental study of the acoustic determinants of vowel color; observation on one- and two formants vowel synthesized from spectorgraphy patterns. *Word*, 8, 195-210.
- Jörgen Pind. (1997). *Sálfræði ritmáls og talmáls*. Reykjavík: Háskólaútgáfan.
- Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, 67, 971–995.
- Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, 87, 820–857.
- Ladefoged, P. og Disner, S. F. (2012). *Vowels and consonants* (3. útgáfa). Chichester: Wiley-Blackwell.
- O'Shaughnessy, D. (1987). *Speech communication: Human and machine*. Michigan: Institute of Electrical and Electronic Engineers.
- Paul Boersma & David Weenink (2009):Praat: doing phonetics by computer (Version 5.1.05) [Computer program]. Sótt: 5.4.2012, frá <http://www.praat.org/>
- Truax, B. (ritstjóri). (1999). *Handbook for acoustic ecology*. (2. útgáfa). Sótt: 3.4.2012, frá <http://www.sfu.ca/sonic-studio/handbook/>

Myndaheimildir

Jörgen L. Pind. (2000). *Talgervill Klatts (drög)*. Bls: 4.

Vowel space á bls 16 í ritgerðinni er Sótt 23.4.2012 af:

http://upload.wikimedia.org/wikipedia/en/5/5a/IPA_vowel_chart_2005.png