Eesti keele välted foneetika korpuse põhjal

Pärtel Lippus

Lühikokkuvõte

Siin peatükis kasutatakse eesti keele spontaanse kõne foneetilist korpust, et kirjeldada eesti keele vältesüsteemi. Korpus koosneb kõnesalvestustest (WAV-failid) ja mitmetasandilisest aegjoondusega annotatsioonist (TextGrid-failid). Juhtumiuuringus kasutatakse Praati skripte, et leida annotatsiooni põhjal üles uuritavad sõnad ning arvutada häälikute kestused. Lisaks kasutatakse helifaile, millest samuti Praati skriptide abil leitakse põhitooni ja vokaalide formantväärtused. Seejärel analüüsitakse saadud andmeid R-is. Kõnelejatevaheliste erinevuste vähendamiseks kasutatakse erinevaid normaliseerimismeetodeid. Tulemuste kirjeldamiseks, visualiseerimiseks ja väidete tõestamiseks kasutatakse lineaarseid segamudeleid.

1. Sissejuhatus

Käesolevas peatükis teeme läbi väikese välteuurimuse eesti keele spontaanse kõne foneetilise korpuse (Lippus jt 2023) andmetega. Uurimuse eeskujuks on Lippus jt (2013) sama korpuse põhjal läbi viidud uurimus. Eesti keele vältesüsteemi kirjeldamiseks analüüsime häälikute kestust, põhitooni ja vokaalide formante. Andmed kogume failidest Praati (Boersma & Weenink 2023) skriptide abil. Andmete analüüsiks kasutame R-i (R Core Team 2023).

Sissejuhatuseks teeme lühiülevaate eesti vältesüsteemist. Vältesüsteemi põhjalikuma käsitluse leiab näiteks Lippus jt (2013) sissejuhatusest või eesti keele häälduse tervikkäsitlusest (Asu, Lippus, Pajusalu, jt 2016)[1]. Vältesüsteem on kahtlemata eesti keele fonoloogilise süsteemi üks olulisemaid nähtusi, mida on ka kõige põhjalikumalt uuritud. Mis selle huvitavaks teeb, on esiteks see, et kolmese pikkusvastandusega keeli ei ole maailmas kuigi palju (eriti neid, millel oleks süsteemi kaasatud nii vokaalid kui konsonandid). Eesti keeles võib nii konsonant kui vokaal olla nii lühike, pikk kui ülipikk. Õigemini, nagu loodetavasti ka järgnev analüüs

[1] Lisaks võib terminoloogia kohta selgitusi otsida foneetika sõnastikust: https://sonaveeb.ee/ds/fon.

näitab, on eesti välde sõnatasandi (või täpsemini rõhulisest ja järgnevast rõhutust silbist koosneva kõnetakti) nähtus. **Kolme välte süsteemis** võivad väldet kanda kas rõhulise silbi vokaal (nimetame vokaalikeskseks malliks), rõhulise ja rõhutu silbi vaheline konsonant (konsonandikeskne mall) või mõlemad kombinatsioonis (vt näiteid tabelist 1). Esimeses vältes (Q1) on rõhuline vokaal ja järgnev konsonant alati lühikesed. Teises (Q2) ja kolmandas (Q3) vältes võib vokaalikeskses mallis rõhuline vokaal olla pikk või ülipikk monoftong või diftong, millele järgneb lühike konsonant. Konsonandikeskses mallis järgneb lühikesele vokaalile pikk või ülipikk konsonant või konsonantühend. Segamallis on teises ja kolmandas vältes rõhuline pikk vokaal või diftong, millele järgneb pikk konsonant või konsonantühend.

Tabel 1. Näited võimalikest fonoloogilistest struktuuridest kolmes vältes. Näitesõnad on tavaortograafia kõrval esitatud rahvusvahelises foneetilises transkriptsioonisüsteemis (IPA)

Mall		**Q1**		**Q2**		**Q3**	
Vokaal	monoftong	*sada*	sɑ.tɑ	*saada*	sɑ:.tɑ	*saada*	sɑ::.tɑ
	diftong			*paise*	pɑise	*paise*	pɑi:.se
Konsonant	geminaat			*kalla*	kɑl.lɑ	*kalla*	kɑl:.lɑ
	konsonantühend			*metsa*	met.sɑ	*metsa*	met:.sɑ
Mõlemad	monoftong + geminaat			*saate*	sɑ:t.te	*saate*	sɑ:t:.te
	diftong + geminaat			*võite*	vɤit.te	*võite*	vɤit:.te
	diftong + konsonantühend			*paista*	pɑis.tɑ	*paistma*	pɑis:t.mɑ

Teiseks teeb eesti keele välte huvitavaks see, et tegemist on kompleksse prosoodilise nähtusega, millel on seosed kõigi muude keele tasanditega, mis tähendab, et eesti keele kirjeldamisel peab vältega arvestama ka siis, kui vaatluse all on midagi muud kui välde. Näiteks eristab välde süstemaatiliselt käänet sõnatüüpides, mis on astmevahelduslikud (nt omastav *pluusi* – osastav *pluusi*) ja seetõttu võib see olla ka ainus tunnus, mis eristab täis- ja osasihitist: *Ma õmblesin pluusi* võib tähendada nii seda, et õmblesin parasjagu pluusi (Q3, osastav kääne, osasihitis), kui ka seda, et õmblesin pluusi valmis (Q2, omastav kääne, täissihitis). Teisalt võivad välte minimaalpaare või -kolmikuid moodustada täiesti erinevate tüvede eri sõnaliigi vormid, nt *kala* (nimisõna, 'vees elav kõigusoojane selgroogne') – *kalla* (nimisõna, taim ladinakeelse nimetusega *Zantedeschia*) – *kalla* (verbi *kallama* käskiva kõneviisi 2. isiku vorm).

Kuigi fonoloogiliselt võib väldet pidada rõhulise silbi omaduseks, kirjeldab eesti keele kolme välte süsteemi kõige paremini esimese ja teise silbi kestuste suhe. Fonoloogilise pikkuse kõige otsesem akustiline vaste on **kestus**. Kui võrrelda ainult rõhulise silbi (häälikute) kestust, siis esimene ja teine välde eristuvad selgelt: pikk häälik on umbes kaks korda pikem kui lühike. Teise ja kolmanda välte puhul on häälikutasandil erinevus aga väike: ülipikk häälik on ainult pisut pikem kui pikk. Samas kuigi rõhutus silbis fonoloogilised pikkusvastandused puuduvad, varieerub välteti ka rõhutu silbi kestus, mistõttu on nii Lehiste (1960) kui Liiv (1961) pakkunud välte kirjeldamiseks kõige sobivamaks mõõdikuks rõhulise ja rõhutu silbi kestuste suhet. Samuti on mõlemad kirjeldanud teist ja kolmandat väldet eristava lisatunnusena **põhitooni**, mille langus kolmandas vältes on varasem kui teises vältes. Hilisemad uurimused on näidanud, et põhitoon on ka vältetaju jaoks väga oluline tunnus, mis kolmandat väldet eristab.

Mõnevõrra varieerub hääliku pikkusest sõltuvalt ka **vokaalide kvaliteet**. Kui Lehiste (1960) seda väga selgelt ei leidnud ja Eek & Meister (1998) osutasid mõnele väikesele tendentsile, siis Lippus jt (2013) leidsid spontaanse kõne põhjal, et rõhulise silbi vokaalid olid teises ja kolmandas vältes selgemini hääldatud kui esimeses vältes ning kui rõhutu silbi vokaalid olid üldiselt redutseeritud, siis kolmanda välte rõhutu silbi vokaalid olid tugevasti redutseeritud.

Siin näidisuurimuses kontrollime, kas varasemates uurimustes leitud vältega seotud seaduspärasused kehtivad ka väikeses foneetika korpusest võetud valimis. Täpsemalt on eesmärk

1. kontrollida kestussuhteid:
   a) kas rõhulise vokaali kestused eristuvad kolmes vältes?
   b) kas rõhutu vokaali kestused eristuvad välteti?
   c) kas silbialguskonsonantide kestustes on välteti erinevusi?
2. kontrollida, kas põhitooni langus toimub kolmandas vältes varem kui esimeses ja teises vältes;
3. kontrollida, kas vokaalide kvaliteet sõltub vältest.

Neile küsimustele vastamiseks teeme läbi järgmised sammud:

1. Praati skriptiga otsime TextGrid-failidest üles huvipakkuvad sõnad ja kirjutame häälikute kestused tabelisse.
2. Avame tabeli programmis R, tutvume leitud materjaliga ja sorteerime sealt tingimustele mittevastava välja.
3. Analüüsime häälikute kestusi.
4. Leiame Praati skriptiga sõnade põhitoonikontuurid.
5. Normaliseerime R-is kõnelejatevahelist varieeruvust.
6. Analüüsime põhitooni liikumist seoses vältega.
7. Leiame Praati skriptiga vokaalide formantväärtused.
8. Normaliseerime kõnelejatevahelist varieerumist.

9. Normaliseerime vokaalikategooriate erinevused.
10. Analüüsime vokaalide kvaliteeti seoses vältega.

2. Materjal

Materjalina kasutame **eesti keele spontaanse kõne foneetilisest korpusest** (Lippus jt 2023)[2] pärit faile. Kuna tervikkorpusele on ligipääs piiratud, on selle näite jaoks valitud korpusest neli faili (kaks nais- ja kaks meeskeelejuhti), mida on võimalik avaandmetena kättesaadavaks teha. Need materjalid on kättesaadavad käesoleva õpiku repositooriumis[3].

Korpus sisaldab spontaanse kõne salvestusi, mis on märgendatud sõna, hääliku ja silbi tasandil, lisaks veel mitmeid erinevaid märgenduskihte. Spontaanne kõne on väga varieeruv ja kui me tahame uurida just nimelt seda, kuidas vaadeldavad akustilised tunnused on seotud meie uuritava tunnusega (vältega), siis peaksime võimalikult palju üritama kontrolli alla saada muid tunnuseid, mis samuti võivad akustiliste tunnuste varieerumist mõjutada. Niisiis, seame mõned tingimused sellele, milliste piirangutega materjali korpusest otsime:

- **kahesilbilised sõnad** – eesti keeles võib sõnas olla teoreetiliselt üks kuni kaksteist silpi (Asu, Lippus, Pajusalu, jt 2016: 154–156). Ühesilbilised sõnad kolme välte süsteemis ei osale. Välistame ka võimaliku sõnaisokroonia efekti – kui sõnas on rohkem silpe ja häälikuid, siis häälikud on lühema kestusega – ning piirdume ainult kahe silbi pikkuste sõnadega.
- **lahtised silbid** – teises ja kolmandas vältes võib rõhulise silbi riim koosneda pikast monoftongist või diftongist, lühikesest vokaalist ja konsonandist või ka pikast vokaalist ja konsonandist (vt tabel 1). Erinevatel kombinatsioonidel on ka mõningane mõju silbi kestusele üldiselt (Lippus & Šimko 2015). Et hoida ära võimalikku silbistruktuurist tingitud varieerumist, võtame uurimusse ainult vokaalikeskse malliga monoftongidega lahtiste silpidega sõnad (nt *sada – saada – saada*, *pole – poole – poole*).
- **välistame fraasilõpulised sõnad** – nagu paljudes keeltes, märgitakse prosoodilisi-süntaktilisi piire eesti keeles lõpupikenemisega, st fraasi lõpus häälikud pikenevad (Krull 1997). Seetõttu jätame välja sõnad, millele järgneb paus või mis on venitatud.
- **välistame sosina**, kuna tahame vaadata ka põhitooni liikumist ja sosinkõnes põhitoon puudub. Samuti võib sosin teha keerulisemaks vokaali formantide leidmise.

[2] https://foneetikakorpus.ut.ee/

[3] https://osf.io/xqzsf/

3. Praat

Kestuse, põhitooni ja formantide andmete kogumiseks kasutame vabavaralist programmi Praat (Boersma & Weenink 2023)[4]. Praatis on vahendid akustiliseks analüüsiks graafilises kasutajaliideses, mille võimalused vaatame siin põgusalt üle. Lisaks graafilisele kasutajaliidesele on Praatis ka täismahus skriptikeel, mida kasutame käesoleva peatüki andmete kogumiseks. Selle näidisuurimuse skriptid on mõnevõrra kommenteeritud, aga väga põhjaliku ja algajale sobiliku Praati skriptikeele õpetuse leiab Praati kasutusjuhendist[5]. Samuti selgitab põhjalikumalt nii akustilise foneetika aluseid kui Praati kasutamist eesti keeles akustilise foneetika meetodite õpik (Lippus 2026).

Korpusest on meil vaja helifaile (WAV-vormingus) ja helifaili juurde kuuluvaid aegjoondusega TextGrid-faile. TextGrid[6] on Praati märgendusvorming. Nagu öeldud, leiab siin peatükis analüüsitavad failid õpiku repositooriumist. Faili avamiseks valime Praati objektiaknas *Open* menüüst käsu *Read from file*.... Sama käsuga saab avada nii helifaili kui TextGrid-faili. Nüüd peaks olema failid ilmunud *Objects* loendisse. Valime mõlemad objektid ja vajutame nuppu *View & Edit*, mille peale avaneb uus aken, kus on kolm jaotust: helilaine, spektrogramm ning TextGridi objekti märgenduskihid (vt joonis 1). Selles aknas on graafilise kasutajaliidese vahenditega võimalik helilaine põhjal kõne akustilisi tunnuseid mõõta. Pikemalt me sellest siiski siin peatükis ei räägi, vaid kasutame olemasolevaid skripte.

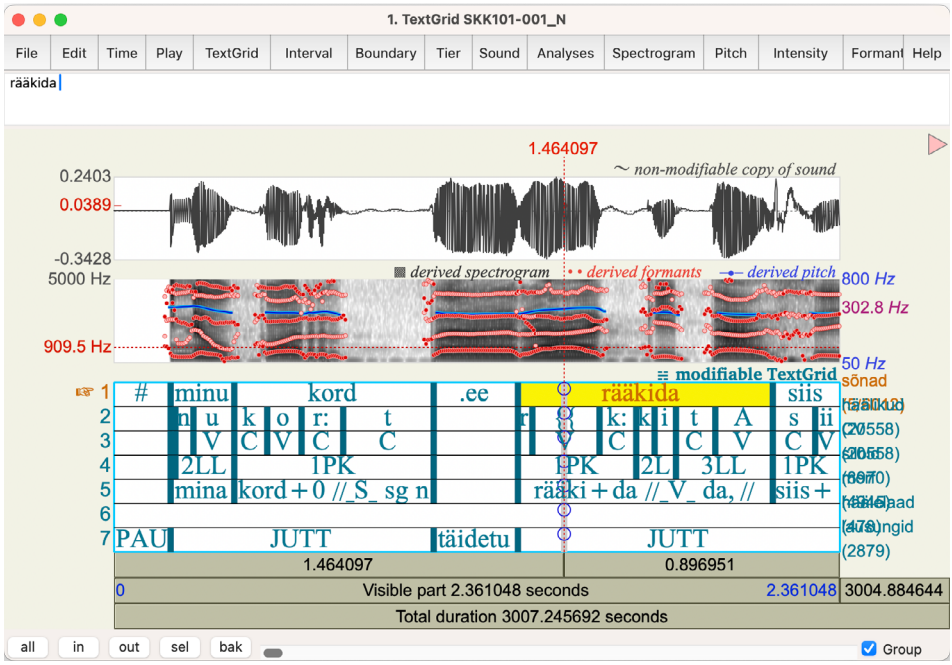
Kõigepealt kogume kestusandmed. Valime Praati objektiakna menüüst *Praat* käsu *Open Praat script*... ja avame faili *korp_opik_fonkorp_valteandmed.praat*. Skript avaneb uues skriptiaknas ja on oma olemuselt tavaline tekstifail. Enne skripti käivitamist tuleks skripti algusest üles otsida ja enda arvuti failisüsteemi arvestades ära muuta kaustade aadressid, kus asuvad korpuse failid (muutuja *kaust$*) ja kuhu kirjutatakse tulemused (*tulemusteKataloog$*). Tuleb meeles pidada, et kausta aadressi lõpus oleks kindlasti kaustaeristaja /. Seejärel valime menüüst *Run* käsu *Run*. Kui skript lõpetab, siis on tulemuste kausta tekkinud fail nimega *kestuste_tulemused.txt*.

Kuigi siin peatükis kasutame Praati, et TextGridi märgendusega helifailidest andmed kätte saada, siis on ka R-i ja Pythoni pakette, millega saab selle töö ära teha. TextGrid-failide lugemiseks ja muude Praatiga tekitatavaid objekte analüüsimiseks on R-is pakett `rPraat` (Bořil & Skarnitzl 2016). Selle paketiga töötades peab aga näiteks põhitooni- ja formantanalüüsid tegema siiski Praatis (ja kui me ei taha teha korraga tervest failist analüüse, vaid eri lõikudest erinevate sätetega, siis on see tülikas). Päris alternatiiviks Praatile on R-i pakett `phonTools` (Barreda 2023), mis võimaldab R-is teha põhitooni- ja formantanalüüsi ja palju muud (Praatis on siiski mõnevõrra laiemad võimalused täpsemalt seadeid määrata). On

[4] https://www.praat.org/

[5] https://www.fon.hum.uva.nl/praat/manual/Scripting.html

[6] https://www.fon.hum.uva.nl/praat/manual/TextGrid.html



Joonis 1. Praati toimetamisakna ekraanipilt

ka pakett `speakr` (Coretta 2024), millega saab R-is Praati skripte käivitada, aga Praatile suunatavad käsud peavad seal olema Praati skriptikeeles. Pythonis on Praati kasutamiseks teek `Parselmouth` (Jadoul, Thompson & de Boer 2018), mille kaudu saab kasutada suuremat osa Praati funktsioone Pythoni keeles.

4. Kestused

Alustame analüüsi häälikukestustest. Kuna vaatluse all on kahesilbilised lahtiste silpidega (ehk vokaalikeskse malliga) sõnad, siis on igas sõnas neli häälikut: alguskonsonant (C1), esimese silbi rõhuline vokaal (V1), teise silbi alguskonsonant (C2) ja teise silbi rõhutu vokaal (V2), mille kestusi analüüsi kaasame. R-i skript, millega käesolev analüüs on tehtud, on kättesaadav õpiku repositooriumis. Kõik kirjeldatud sammud on skriptis kommenteeritud, et analüüsi oleks võimalik ise läbi teha.

4.1. Andmete lugemine R-i ja täiendav filtreerimine

Esimese sammuna loeme Praati häälikukestuste leidmise skriptiga saadud andmed R-i ja teeme tabelis mõned kohendused:

- kodeerime kõneleja soo ja välte tunnustena, mille tasemetel on kindel järjekord. Sellega välistame analüüsis tasemete järjekorra varieerumise juhusliku esinemisjärjekorra tõttu erinevates andmestiku alamosades, mille tagajärjel võiks erinevatel joonistel olla kategooriate järjekord erinev;
- arvutame häälikute algus- ja lõpuaegade põhjal häälikukestused ja teisendame need sekunditest millisekunditeks, sest häälikukestusi on nii visuaalselt parem jälgida. Selle sammu oleks võinud teha ka Praati skriptis, aga kestuse arvutamine on väga lihtne tehe: lahutame lõpuajast algusaja. Igal juhul on mõistlik andmestikus talletada ka algus- ja lõpuajad, sest kui hiljem on analüüsi käigus tarvis midagi helifailist täpsustada või juurde mõõta, on aegade kaudu võimalik sõna failist üles otsida;
- otsime morfoloogilisest märgendusest üles sõnaliigi tähise ja paneme selle eraldi tulpa. Ka seda oleks võinud teha Praati skriptis, aga R-is on see natuke lihtsam.

Seejärel filtreerime veel andmestikust üht-teist välja, et vähendada häälikukestuste varieerumist, mis võiks olla seotud muude nähtustega kui välde. Jällegi, kõike seda saaks teha ka Praati skriptis, kuid R-iga on mõnevõrra lihtsam.

Kuna andmestikku jäi sisse ka häälikutasandil venitatuks märgitud sõnu, siis viskame nad nüüd välja. Venitus on häälikutasandil märgitud topeltkooloniga (nt sõna *kõne* oleks venitatud rõhutu silbi vokaaliga SAMPA transkriptsioonis transkribeeritud [k7ne::]). **Välja jäävad** kõik sõnad, kus mõni häälik on märgitud venitatuks.

Jätame välja sõnaliigi poolest grammatilised sõnad, mis võivad olla suure tõenäosusega redutseeritud. Andmestikku jätame alles adjektiivid (sõnaliigitähisega *A*, *C*, *U*), pärisnimed (*H*), hüüdsõnad (*I*), põhi- ja järgarvsõnad (*N*, *O*), nimisõnad (*S*) ja tegusõnad (*V*). Välja jäävad määrsõnad, sidesõnad, kaassõnad, asesõnad, lühendid jms. (Täpsemalt vaata morfoloogilise analüüsi kohta Vabamorfi kirjeldusest[7].) Tabelis 2 on loend sõnadest, mis selle tulemusena andmestikku jäid.

[7] https://www.filosoft.ee/html_morf_et/morfoutinfo.html#2

Tabel 2. Andmestikus esinevad sõnad

Välde	N	Sõnad
Q1	248	*kõne, ole, tohi, kadu, sõna, tere, kobe, mõju, näha, koha, pära, lumi, võru, kana, vahe, kogu, neli, keele, teha, mage, ole/.?, kahe, pidi, suhe, pähe, vali, müra, sõnu, nime, sada, saja, muna, kuju, sära, nõme, tuli, pere, mina, tore, koma, kena, tänu, huvi, sisu, rida, vaba, büroo, Mare/.isikunimi, Pire, vajab, kada, pole, viga, sobi, mari*
Q2	111	*keele, räägi, hääle, Mona, Lisa, tuua, näeme, joone, liigu, viisi, saame, kiire, sööme, roosa, kuule, tüübi, geeni, paari, Via, suuda, viie, viide, loome, teeme, luua, ruumi, poole, moora, Leelo*
Q3	50	*tuua, jooni, jääma, saada, poole, maali, haara, keeli, tooma, tuuma, soola, tooli, jääda, suuri, saama, luua, piisa, moodi, viia/.?, kuulu, viia*

Tabelist 2 torkavad silma mõned sõnad, mis on mitteootuspäraselt kategoriseeritud ja mille puhul peaks kaaluma, kas need tuleks veel andmestikust eemaldada. Esmavälteliste sõnade hulgas on sõna *büroo*, mis on võõrsõna rõhuga teisel silbil. See ei sobitu ülejäänud materjali hulka, kus on rõhk esisilbil. Mingil põhjusel on see TextGridil märgendatud nii, nagu teine silp oleks lühike. Võiks minna Praati, otsida sõna helifailist üles ja kontrollida, kuidas seda on hääldatud, aga selle asemel viskame käesolevas analüüsis selle sõna lihtsalt andmestikust välja.

Samuti võiks välja jätta esimese välte alla sattunud sõna *keele*. Tõenäoliselt on tegemist (fookus)rõhulise sõnale järgneva deaktsentueeritud häälduse ja pikk vokaal on redutseerunud. Kuna esmavältelisi sõnu on andmestikus nagunii rohkem kui teisi, võime need ka rahumeeli välja jätta.

Veel on näha, et andmestikku on sattunud sõna *ole*, mis ortograafias ei alga konsonandiga, aga sõnaalguskonsonandiks on märgenduses /j/. Ilmselt on see tekkinud koartikulatoorselt järjendis *ei ole* sõnade piirile. Kuna ei ole ühtegi empiirilist uurimust, mis võrdleks selliseid koartikulatsioonist tekkinud konsonante muude konsonantidega ja ei ole päris selge, kas need kestuste poolest muudest konsonantidest ei eristu, siis jätame siin need ka igaks juhuks välja.

Teisevälteliste sõnade hulgas leidub mõningaid võõrnimesid, mis ortograafia järgi on lühikese vokaaliga (*Mona, Lisa*) või sobimatu struktuuriga (*Via*), aga kuna andmed valisime häälikutasandi märgenduse põhjal, siis need on õiges kohas teisevälteliste sõnade hulgas (SAMPA transkriptsioonis [moonA], [liisA], [viijA]) ja võivad jääda andmestikku alles.

4.2. Kestusandmete analüüs

Teeme nüüd väikese kokkuvõtte, kui palju andmeid õnnestus kätte saada. Kokku on andmestikus 345 vaatlust kahelt nais- ja kahelt meeskeelejuhilt. Teeme selle põhjal risttabeli kõneleja soo ja välte esinemisjuhtudest (tabel 3). Tegemist on seega suhteliselt väikese andmestikuga, aga sellele vaatamata on iga välte kohta üle 10 vaatluse. Erinevus andmete jaotuse osas välteti on ka mõnevõrra ootuspärane: esmavältelisi sõnu on oluliselt rohkem kui kõiki muid, sest andmete otsingul seatud piirang, et kaasata analüüsi ainult lahtised silbid, piirab teise ja kolmanda vältega sõnade kaasamist oluliselt rohkem.

Tabel 3. Analüüsitavate sõnade arv välteti mees- ja naiskeelejuhtidel

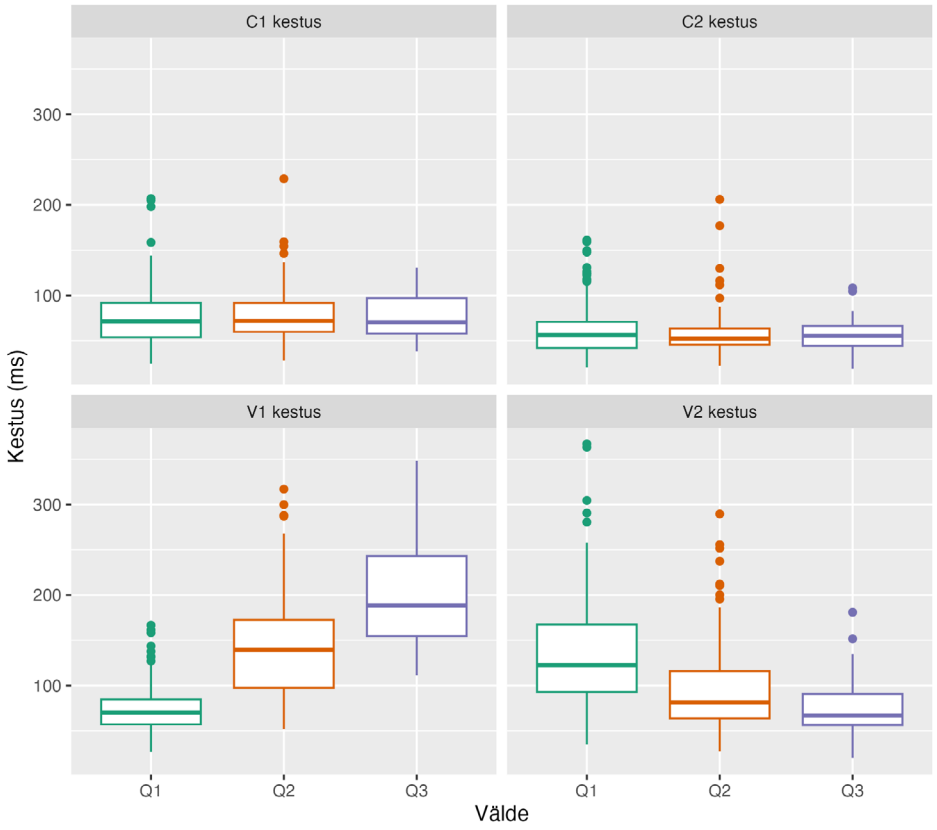
sugu	**Q1**	**Q2**	**Q3**
N	93	56	26
M	91	55	24

Et andmetest paremat ülevaadet saada, joonistame häälikute kestused välteti karpdiagrammina (joonis 2). Karpdiagramm näitab andmete jaotust: karbi keskel jämedam joon tähistab mediaanväärtust, karp kvartiilhaaret ja vurrud (kasti alumisest ja ülemisest küljest välja ulatuvad jooned) usaldusvahemikku; karpdiagrammi kohta leiab rohkem infot õpiku peatükist 6.1.5.2 „Karpdiagramm".

Lisaks joonisele teeme keskmiste ja standardhälvete tabeli (tabel 4). Nagu joonisel kast ja vurrud, annab standardhälve tabelis infot andmete jaotuse ja hajuvuse kohta. Kui võrdleme kolme välterühma keskmisi kestusi, siis standardhälbe piirid näitavad seda, kui selgelt rühmad eristuvad (vt ka õpiku peatükki 6.1.3 „Aritmeetiline keskmine ja standardhälve").

Tabel 4. Häälikukestuste keskmised (*kesk*) ja standardhälbed (*sd*) millisekundites

	C1		**V1**		**C2**		**V2**	
Välde	kesk	*sd*	kesk	*sd*	kesk	*sd*	kesk	*sd*
Q1	77	31	74	26	61	28	133	58
Q2	77	30	145	58	58	25	98	53
Q3	77	24	202	59	57	18	75	30



Joonis 2. Häälikute kestused välteti: C1 – sõna alguskonsonant, C2 – vokaalidevaheline konsonant, V1 – rõhulise silbi vokaal, V2 – rõhutu silbi vokaal

Tabelist 4 ja jooniselt 2 näeme, et konsonantide puhul on väldetevahelised erinevused väiksemad kui standardhälbed ühe välte piires, vokaalide puhul on väldetevaheline varieerumine suurem, aga standardhälbe piirides on siiski ka kattumisi. Järgnevalt analüüsime tabelis 4 ja joonisel 2 esitatud tulemusi põhjalikumalt häälikute kaupa ning kontrollime järeldava statistika abil, kas kestuse vältega seotud varieerumine on selle hääliku puhul statistiliselt oluline või mitte.

4.2.1. Rõhulise vokaali kestus

Välte kirjeldamisel huvitab meid enim rõhulise silbi vokaali kestus, sest see on vokaalikeskses mallis fonoloogiliselt välte kandja. Tabelist 4 ja jooniselt 2 võib näha, et esmavältelise vokaali kestus on keskmiselt 74 millisekundit, teises vältes vokaali kestus keskmiselt 145 millisekundit, mis tähendab, et teises vältes vokaal on keskmiselt umbes poole pikem kui esimeses vältes. Kolmandas vältes on vokaali

kestus keskmiselt 202 millisekundit ja see on teisest vältest üksjagu pikem. Kui aga arvesse võtta ka standardhälbeid, siis on välterühmade esimese vokaali kestuses mõningane kattuvus, mida näeme joonisel 2 sellest, et eri väldete kastid on osaliselt üksteisega kohakuti.

Et kontrollida, kas vokaali kestus eristab kolme väldet, testime tulemusi **lineaarse segamudeliga**, kasutades R-i paketti `lme4` (Bates jt 2015). Siin mudelis võtame uuritavaks tunnuseks hääliku kestuse ja seletavaks tunnuseks sõna välte. Peale selle lisame mudelisse ka **juhuslikud vabaliikmed** (ingl *random intercept*) keelejuhtidele ja sõnadele, kuna mõned keelejuhid võivad rääkida kiiremini kui teised ja mõningaid sõnu võidakse hääldada kiiremini kui teisi. Juhuslikke kaldeid, mis arvestaksid lisaks sellega, et seletava tunnuse ehk välte mõju võib samuti sõnuti erineda, siin mudelis lisada ei saa. Seda seetõttu, et enamik sõnavorme on alati samas vältes (nt sõna *keele* on alati Q2 või sõna *mage* on alati Q1) ja ainult väike hulk sõnu on leksikonis varieeruva vältega (vt nt Piits & Kalvik 2017). Samuti ei lisa me mudelisse juhuslikke kaldeid, mis võimaldaksid arvestada sellega, et välte mõju võib ka kõnelejati erineda (nt võivad mõne kõneleja häälikud olla esimeses vältes keskmisest lühemad, aga teises vältes keskmisest pikemad, teisel kõnelejal jällegi esimeses vältes keskmisest pikemad, aga teises vältes ainult natuke pikemad jne). Regressioonimudelite kohta saab rohkem lugeda õpiku ptk-st 6.2.2 „Mitmetunnuseline seoste analüüs: statistilised mudelid".

Kuna seletaval tunnusel *välde* on kolm taset, siis annab mudeli väljund meile kaks võrdlust baastaseme ja mõlema alternatiivse taseme vahel. Vaikimisi oleks meil baastase Q1 ja mudel esitaks võrdluse Q1 *vs.* Q2 ning Q1 *vs.* Q3. Alternatiivsete tasemete võrdlemiseks (Q2 ja Q3 vahelise erinevuse leidmiseks) peaksime tegema *post-hoc*-testi. Kuna me teame, et kõige suurem erinevus on Q1 ja Q3 vahel ja me tahame kontrollida, kas kummagi tasemega ka Q2 erinevus oluline on, võiksime määrata baastasemeks hoopis Q2. Nii saame hiljem *post-hoc*-testi vajadusest kõrvale põigata, sest lineaarse mudeli väljund annab võrdlused Q2 *vs.* Q1 ning Q2 *vs.* Q3.

Kuna uuritav tunnus on kestus ja kestusandmete jaotus on tihti paremale kaldu, siis **logaritmime** uuritava tunnuse väärtused, et paremini täita mudeli eeldusi normaaljaotuse osas (vt ka ptk 6.2 „Järeldav ehk inferentsiaalne statistika"), st paneme uuritavale tunnusele *V1_kest* ümber käsu *log()*. Pikemalt mudeli eelduste kontrollimisega siiski siin praegu ei tegele, jäägu see statistikaõpikutele.

Lineaarse segamudeli jaoks oleks nüüd tarvis aktiveerida paket `lme4` ning selle abipakett `lmerTest` (mis lisab mudeli väljundisse *p*-väärtused). Mudeli saab käsuga *lmer()* ning regressioonivalem on järgmine:

```
log (V1_kest) ~ välde + (1|fail) + (1|sõna)
```

Valemis on täpsustatud, et (logaritmitud) rõhulise silbi vokaali kestus sõltub vältest ning mudeli vabaliiget korrigeeritakse iga faili/kõneleja ning iga sõna suhtes.

Mudeli koguväljundit saame vaadata käsuga *summary()*, aga kuna kõige rohkem huvitab meid sealt välte **fikseeritud mõjude** osa, siis ruumi kokkuhoiu mõttes

esitame siin teksti sees väljundist ainult fikseeritud mõjude (ingl *fixed effects*) tabeli (tabel 5), mis antud mudeli puhul sisaldab ainult välte mõju kestusele.

Tabel 5. Rõhulise vokaali kestuste segamudeli fikseeritud mõjud

	Hinnang	Standard-viga	Vabadus-astmete arv	t-väärtus	Pr(>|t|)
(vabaliige)	4,961	0,143	3,510	34,661	<0,001
välde Q1	-0,599	0,054	103,373	-11,158	<0,001
välde Q3	0,299	0,062	168,127	4,784	<0,001

Tabeli 5 viimasest tulbast näeme, et kõik mõjud on mudeli seisukohast statistiliselt olulised ($p < 0{,}001$ tähendab, et on väiksem kui 0,1% tõenäosus, et näeksime enda andmetes väldete vahel sellist erinevust, kui tegelikult kehtiks nullhüpotees ja väldete vahel vokaalikestuses erinevusi ei oleks).

Mudeli väljundi esimeses tulbas on esitatud mudeli uuritava tunnuse (V1 kestus) hinnangulised väärtused. Esimesel real näeme mudeli **vabaliikme** hinnangulist väärtust, see on uuritava tunnuse väärtus seletava tunnuse baastaseme korral ehk V1 kestus siis, kui välde on Q2. Teistel ridadel näeme vastava taseme ja baastaseme erinevust: kui palju kestus keskmiselt muutub, kui Q2 asemel on Q1 või kui Q2 asemel on Q3. Aga kuna mudelisse sisestasime uuritava tunnuse V1 kestuse logaritmitud kujul, siis on mudeli hinnangulised väärtused samuti logaritmitud kestused. Selleks, et neid millisekunditesse tagasi saada, tuleks hinnangute väärtused kokku liita ning astendada[8]:

- Q2 (ehk vabaliikme) hinnang on 4,961 ehk $e^{4{,}961} = 142{,}8$ millisekundit;
- Q1 hinnang on 4,961 + -0,599 = 4,362 ehk $e^{4{,}362} = 78{,}5$ millisekundit;
- Q3 hinnang on 4,961 + 0,299 = 5,26 ehk $e^{5{,}26} = 192{,}5$ millisekundit.

Niisiis on mudeli hinnangul esimeses vältes rõhuline vokaal lühem kui teises vältes ja see erinevus on statistiliselt oluline ($\beta = -0{,}599$, $t = -11{,}158$, $p = <0{,}001$)[9]. Kolmandas vältes on rõhuline vokaal pikem kui teises vältes ja see erinevus on samuti statistiliselt oluline ($\beta = 0{,}299$, $t = 4{,}784$, $p = <0{,}001$).

[8] Kui logaritmimisel teisendatakse väärtused irratsionaalarvu *e* ehk Euleri arvu (mille väärtus on umbes 2,71828) astmeteks, siis logaritmitud väärtuse originaalühikutesse teisendamiseks tuleb astendada *e* logaritmitud väärtuse astmesse (nt 2,71828 astmes 4,961).

[9] Tunnuse olulisuse hinnang tuleb lineaarse segamudeli väljundist, siin tabeli 5 teiselt realt: Q2 (baastaseme) ja Q1 erinevus on -0,599, see on mudeli hinnang ja siin tähistatud kui β ning tõenäosus, et leiaksime valimis sellise erinevuse, kui tegelik erinevus tasemete vahel on 0, on väiksem kui 0,1% (ehk $p < 0{,}001$).

4.2.2. Rõhutu silbi vokaali kestus

Rõhutu silbi vokaali kestuste juures näeme tabelist 4 ja jooniselt 2 pööratud seost esimese vokaaliga: mida suurem välde, seda lühem vokaal, erinevused jäävad paarikolmekümne millisekundi piiresse. Standardhälve on kõige väiksem lühikesel kolmanda välte vokaalil. Teeme siin samasuguse segamudeli nagu rõhulise silbi vokaali kestusega, ent uuritav tunnus on sedapuhku rõhutu vokaali silbi kestus.

Tabel 6. Rõhutu vokaali kestuste segamudeli fikseeritud mõjud

	Hinnang	**Standard-viga**	**Vabadus-astmete arv**	**t-väärtus**	**Pr(>|t|)**
(vabaliige)	4,470	0,148	3,630	30,266	<0,001
välde Q1	0,338	0,062	63,194	5,488	<0,001
välde Q3	-0,244	0,076	110,726	-3,222	0,002

Mudeli väljundist tabelis 6 näeme, et esimeses vältes on rõhutu vokaal pikem kui teises vältes ja erinevus on oluline (β = 0,338, t = 5,488, p = <0,001). Kolmandas vältes on rõhutu vokaal lühem kui teises vältes ja see erinevus on samuti statistiliselt oluline (β = -0,244, t = -3,222, p = 0,002).

4.2.3. Konsonantide kestus

Varasemad uurimused (Lippus jt 2013) on näidanud, et ka silbialguskonsonantidel on välteti väikesed kestuserinevused. Tavaliselt jäävad need 10 millisekundi piiresse ja tavaliselt on konsonandid kõige lühemad teises vältes. Käesolevas andmestikus nii suuri erinevusi ei ole: tabelist 4 näeme, et sõna alguskonsonandi (C1) kestus on kõigis kolmes vältes keskmiselt sama väärtusega ja teise silbi alguskonsonandi (C2) keskmiste kestuste erinevused jäävad 5 ms piiresse.

Tabel 7. Sõna alguskonsonandi kestuste segamudeli fikseeritud mõjud

	Hinnang	**Standard-viga**	**Vabadus-astmete arv**	**t-väärtus**	**Pr(>|t|)**
(vabaliige)	4,210	0,105	5,141	40,110	<0,001
välde Q1	0,037	0,069	74,426	0,537	0,593
välde Q3	0,053	0,078	143,129	0,680	0,498

Tabel 8. Teise silbi alguskonsonandi kestuste segamudeli fikseeritud mõjud

	Hinnang	Standard-viga	Vabadus-astmete arv	t-väärtus	Pr(>|t|)
(vabaliige)	4,090	0,122	4,947	33,536	<0,001
välde Q1	0,016	0,077	108,527	0,203	0,839
välde Q3	-0,057	0,076	261,538	-0,749	0,454

Mudelid osutavad, et olulist vältega seotud varieerumist konsonantide kestustes ei ole: tabelites 7 ja 8 toodud mudelite väljundid näitavad, et ei esimese ega kolmanda välte konsonantide kestuse erinevus ei ole baastaseme ehk teise välte konsonandi väärtustest oluliselt erinev[10].

4.2.4. Kestussuhe

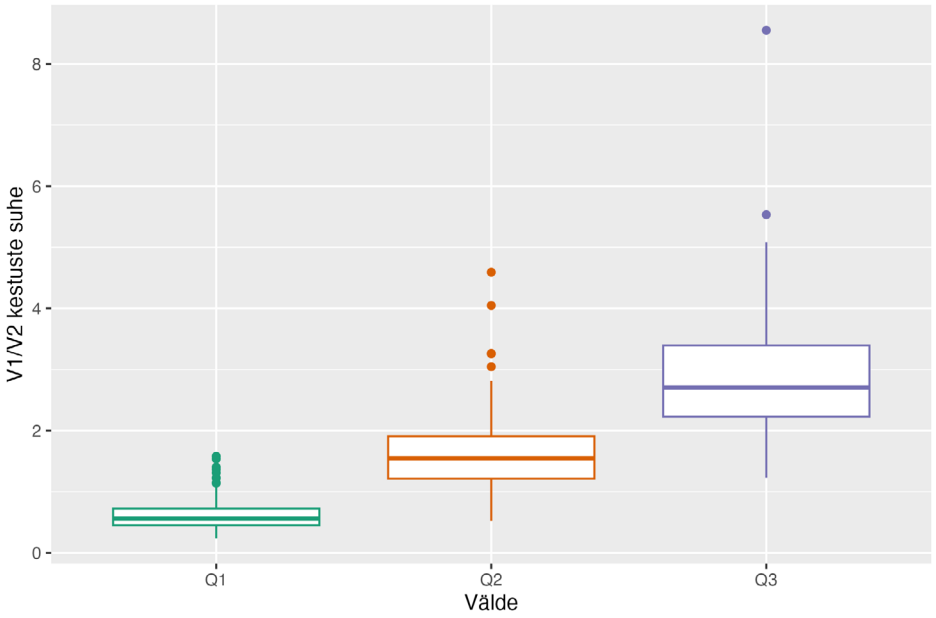
Kuna rõhulise vokaali puhul on väga selge erinevus esimese ja teise välte vahel, aga teise ja kolmanda välte erinevus on võrdlemisi väike, tasub vaadata V1/V2 kestuste suhet, mida kõige sagedamini on kasutatud välte eristuse mõõdikuna (Asu, Lippus, Pajusalu, jt 2016: 134–137). Kuigi sellest räägitakse tihti ka kui silbisuhtest, siis jäetakse enamasti arvutusest välja silpide alguskonsonandid ja arvestatakse ainult silbiriimi, mis meie andmestikus tähendab ainult vokaali (kinniste silpide puhul moodustaks riimi vokaal ja koodakonsonant või -konsonandid). Tavaliselt esitatakse kestussuhet nii, et rõhulise vokaali kestus jagatakse rõhutu vokaali kestusega (vt joonis 3).

Silbisuhete hindamiseks viidatakse kõige sagedamini Lehiste (1960) üldistatud sihtarvudele:

- esimeses vältes 2/3 (ehk 0,7),
- teises vältes 3/2 (ehk 1,5),
- kolmandas vältes 2/1 (ehk 2,0).

Lehiste esitas neid murdarvudena ning tegemist ei olnud täpsete mõõtmisandmetega, vaid mõõtmiste põhjal üldistatud kirjeldusega, mille lähedale võiksid jääda ka mõõtmisandmed. Kui võrrelda eri uurimuste tulemusi (vt nt Asu, Lippus, Pajusalu, jt 2016: 136, tabel 4.1), siis on neis üksjagu varieerumist, aga üldine ja püsiv on see, et esmavältelise sõna rõhuline silp on oluliselt lühem kui rõhutu silp (suhe <1),

[10] Tabelites 7 ja 8 näitab seda teise ja kolmanda rea viimases tulbas olev *p*-väärtus, mis on suurem kui 0,05, väljendades seda, et tõenäosus leida meie korpusvalimist tabeli hinnangu lahtrites esitatud imeväikesi, peaaegu nullilähedasi erinevusi, kui tegelikkuses kehtib nullhüpotees ja eesti keeles väldete vahel konsonandi kestustes väldetevahelisi erinevusi ei ole, on suurem kui 5%.Teisisõnu on suur tõenäosus, et näeme väga väikeseid erinevusi seetõttu, et tegelikkuses ongi erinevused väga väikesed ja välde uuritavate konsonantide kestust oluliselt ei mõjuta.



Joonis 3. Rõhulise ja rõhutu silbi vokaalide kestuste suhe

teisevältelise sõna rõhuline silp on pikem kui rõhutu silp (suhe on >1, aga <2) ning kolmandavältelise sõna rõhuline silp on väga palju pikem kui rõhutu silp (suhe >2). Siin uurimuses on keskmised silbisuhted sellised:

- esimeses vältes 0,6,
- teises vältes 1,7,
- kolmandas vältes 3.

Testime ka lineaarse segamudeliga, kas silbisuhe eristab välteid.

Tabel 9. Silbiriimide kestussuhete segamudeli fikseeritud mõjud

	Hinnang	**Standard-viga**	**Vabadus-astmete arv**	**t-väärtus**	**Pr(>|t|)**
(vabaliige)	0,471	0,082	5,853	5,757	0,001
välde Q1	-0,922	0,060	70,577	-15,311	<0,001
välde Q3	0,570	0,072	120,321	7,885	<0,001

Mudeli väljundist tabelis 9 näeme, et esimeses vältes on kestussuhe väiksem kui teises vältes ja erinevus on statistiliselt oluline (β = -0,922, t = -15,311, p = <0,001). Kolmandas vältes on kestussuhe suurem kui teises vältes ja see erinevus on samuti statistiliselt oluline (β = 0,570, t = 7,885, p = <0,001).

5. Põhitoonianalüüs

Kuigi välte primaarne tunnus on pikkus ja kestuste analüüs näitas, et rõhulise ja rõhutu silbi kestuste suhe eristab selgelt kõiki kolme väldet, siis varasemad uurimused (eriti tajukatsed) on näidanud, et põhitoonikontuur on kestussuhete kõrval oluline sekundaarne vältetunnus.

5.1. Põhitooni andmete kogumine Praatis

Põhitooni analüüsimiseks läheme uuesti Praati. Praati toimetamisaknas kuvatakse põhitoonikontuuri sinise joonena spektrogrammi peal ning selle skaala kuvatakse akna paremas servas (vt joonis 1). Põhitoonianalüüsi puhul on oluline määratleda võimalikult täpselt analüüsitav sageduste vahemik, et vältida oktavivigu. Põhitooni seadeid saab Praati toimetamisaknas määrata menüüst avanevas valikute aknas: *Pitch*: *Pitch settings…*.

Põhitooni analüüsimisel võib kergesti juhtuda, et ekslikult peab algoritm kahte täisvõnget üheks või vastupidi, mistõttu tekivad andmetesse nn **oktavivead**. See tähendab seda, et põhitooni väärtust näidatakse oktavi võrra kõrgema või madalamana, kui see tegelikult on. Veaohtu on võimalik vähendada, kui seada analüüsile täpsed piirid, mis vahemikus kõneleja põhitoon liigub ja kust algoritm seda otsima peab. Samas ei tohi piire liiga kitsalt seada, sest kui põhitooni tegelik väärtus jääb analüüsitavast vahemikust välja, siis surutakse põhitoonianalüüsile samuti peale oktaviviga, et näidata väärtust etteantud vahemikus.

Meeshääle ulatus on tavaliselt vahemikus 75–200 Hz, naishääle ulatus 150–300 Hz. Mida täpsemalt oskame konkreetse kõneleja hääleulatust määratleda, seda vähem vigu andmestikku satub. Analüüsivahemiku määramiseks peaksime tegema mõningaid vaatlusi, et teada saada, mis vahemikus kõneleja põhitoon varieerub. Et seda teha automaatselt ja parimal viisil, kasutame meetodit, mida on kirjeldanud De Looze & Hirst (2008): teeme esmalt põhitoonianalüüsi küllaltki suure analüüsivahemikuga, siis leiame sealt esimese ja kolmanda kvartiili ja seame põhitoonianalüüsi alumiseks piiriks 0,75-kordse 1. kvartiili ja ülemiseks 2-kordse 3. kvartiili.

Käivitame nüüd Praatis skripti *korp_opik_fonkorp_pohitoon.praat*. Ka siin skriptis tuleb konkreetse arvuti failisüsteemi järgi ära vahetada kaustade aadressid, kust skript faile otsib ja kuhu tulemused kirjutab. See skript eeldab, et oleme juba jooksutanud eelmist, kestuste leidmise skripti, sest siit on märgenduse järgi

sõnade otsimise osa välja jäetud ja skript võtab aluseks kestuste tabeli, kuhu põhitooni andmed juurde lisatakse. Seega peaks tulemuste kaustas juba olema fail *kestuste_tulemused.txt*. Skripti jooksutamise lõpuks on töökataloogi tekkinud uus fail *pohitooni_tulemused.txt* (mille leiad ka peatüki repositooriumist). Loeme selle faili R-i (arvestades, et Praati *--undefined--* väärtused võiks saada siin *NA*-ks) ning kohendame andmestikku samamoodi nagu kestuste puhul tegime: arvutame häälikute algus- ja lõpuaegadest kestused, eemaldame valimist valesti klassifitseeritud ja venitatud sõnad ning grammatilised sõnad.

5.2. Põhitooni väärtuste normaliseerimine

Kuna meil on andmeid erinevatelt kõnelejatelt, kelle häälekõrgus erineb, siis on vaja neid võrdlemiseks **normaliseerida**. Normaliseerimiseks ei ole kahjuks ühte head meetodit, sest selle käigus kaotame alati midagi ära algsest varieeruvusest ja tekitame andmetesse moonutusi.

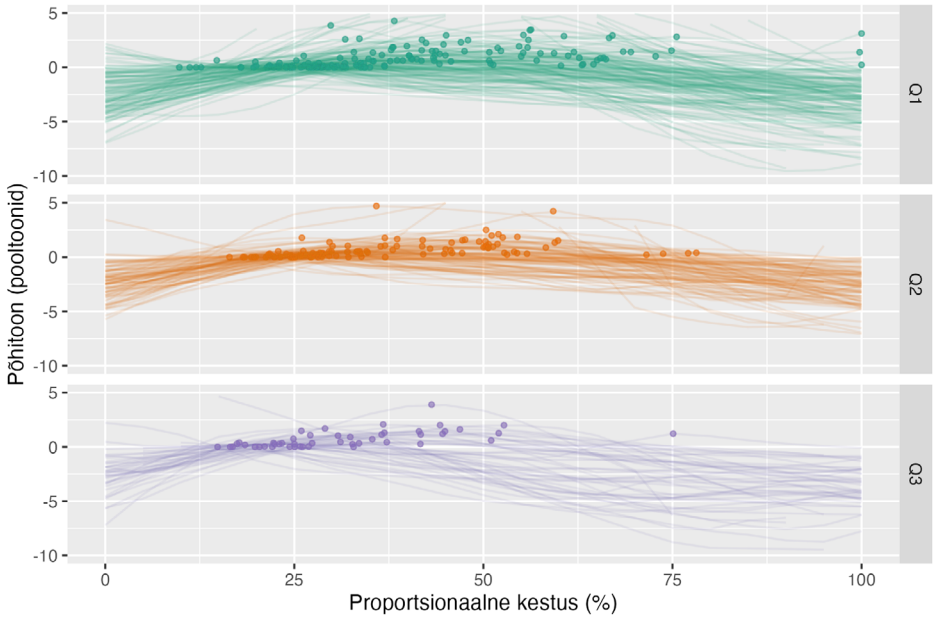
Hertsiskaala on inimese taju jaoks logaritmiline: näiteks muutust 50 hertsilt 100 hertsile tajume me sama suurena nagu muutust 200 hertsilt 400 hertsile, mistõttu kõrgemates sagedusvahemikes on ka varieerumine arvuliselt suurem. Selleks, et logaritmilisust eemaldada, on mõistlik teisendada andmed lineaarsele skaalale. Hääle põhitooni kirjeldades teisendatakse hertsides mõõdetud põhitoon tihti pooltoonideks. **Pooltooniskaala** on muusikas kasutatav intervallskaala, mis jagab oktavi 12 pooltooniks, võttes tihti nullpunktiks esimese oktavi la. Kui me tahame aga normaliseerida kõne põhitooni kõnelejate vahel, siis võime nullpunktiks võtta iga kõneleja keskmise häälekõrguse.

Siin uurimuses vaatame põhitooni liikumist üksikute sõnade piires, mistap lisaks kõnelejate individuaalsele häälekõrgusele oleks tarvis normaliseerida ka lauseintonatsioonist tingitud varieerumist – kui üldiselt algab lausung kõrgema põhitooniga ja intonatsioonifraasi jooksul põhitoon langeb (Asu, Lippus, Salveste, jt 2016), siis siin uurimuses meid ei huvita, kas sõna asus fraasi alguses, keskel või lõpus. Seetõttu teisendame iga sõna põhitooni pooltoonideks selle sõna alguspunkti suhtes. Või õigemini, kuna sõna päris algus võib olla helitu (algab näiteks klusiiliga), siis võtame normaliseerimise aluseks esimese vokaali alguse. Hertside pooltooniskaalale teisendamiseks kasutame funktsiooni *f2st()* paketist `hqmisc` (Quene 2022). Andmestikus on põhitoon mõõdetud vokaalide algusest ja lõpust ning lisaks ka kontuur kogu sõna jooksul (21-st võrdsete vahedega punktist). Pooltooniskaala teisenduseks teisendame iga sõna kontuuri V1 alguspunkti suhtes.

5.3. Põhitoonikontuuride visualiseerimine

Järgnevalt vaatame põhitoonikontuure joonistel, et nende kujust ülevaadet saada. Joonisel 4 on välja joonistatud kõikide sõnade individuaalsed põhitoonikõverad ning täpiga on tähistatud põhitooni maksimumpunkt. Kuigi kontuurid on

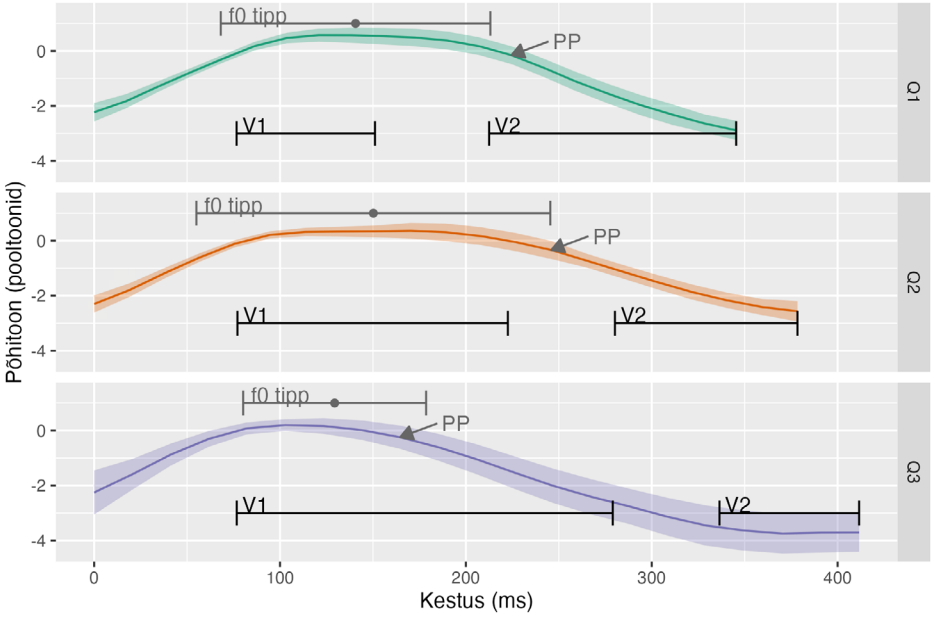
võrdlemisi sarnased kõigis väldetes, siis mõnevõrra joonistub juba välja, et kolmanda välte puhul on põhitooni tipp x-teljel natuke varasem kui esimeses ja teises vältes, kuigi varieerumine on kaunis suur. Joonisel on **kestustelg normaliseeritud** sõna kogukestuse suhtes nii, et kõik kontuurid on sama pikkusega. Seega kui võrrelda tipu asukohta absoluutteljel, võib erinevus tipu asukoha osas kolmandavältelistes sõnades ära kaduda.



Joonis 4. Üksikute sõnade põhitoonikontuurid. Põhitooni väärtused on teisendatud pooltooniskaalale V1 alguse suhtes; kestusteljel normaliseeritud aeg sõna kogukestuse suhtes; täpiga on tähistatud kontuuri maksimumpunkt

Joonisel 5 on põhitoonikontuurid skaleeritud **absoluutajateljele** ning selleks, et vähendada müra üksikute kontuuride varieerumisest (üksikute sõnade jooned joonisel 4), on joonisel kontuurid keskmistatud ja esitatud usaldusvahemiku piiridega (umbes sama, mis vurrud karpdiagrammil). Et kontuuri sõna struktuuriga paremini seostada, on kontuuride all näidatud vokaalide paiknemine. Samuti on joonisel osutatud põhitooni tipu asukoht standardhälbe piiridega. Nagu ka joonilt 4 nägime, on põhitooni tipu asukoht sõnati väga varieeruv. Lippus jt (2013) kirjeldavad seda kui kõrget platood, mis esimeses ja teises vältes püsib pea kogu esimese silbi jooksul ning kolmandas vältes esimese silbi esimeses pooles. Mõõtes tippu kui põhitooni absoluutset maksimumi võibki see sattuda võrdlemisi juhuslikku punkti kõrge platoo jooksul. Seetõttu oleks mõttekam tipu asemel kirjeldada

pöördepunkti, kust kõrge põhitoon pöördub langusele. Joonisel 5 on pöördepunkt (*PP*) tähistatud noolega ning see on siin defineeritud kui punkt, kus põhitoon on maksimumväärtusest langenud 0,5 pooltooni võrra. Nagu näeme, on esimeses ja teises vältes pöördepunktid võrdlemisi kohakuti, paiknedes teise silbi alguses, kolmandas vältes on pöördepunkt märksa varasem, asudes esimese silbi keskel. Tajukatsed on näidanud, et kolmanda välte puhul on oluline tunnus see, et põhitoon esimese silbi jooksul märgatavalt langeb (Lippus, Pajusalu & Allik 2011).

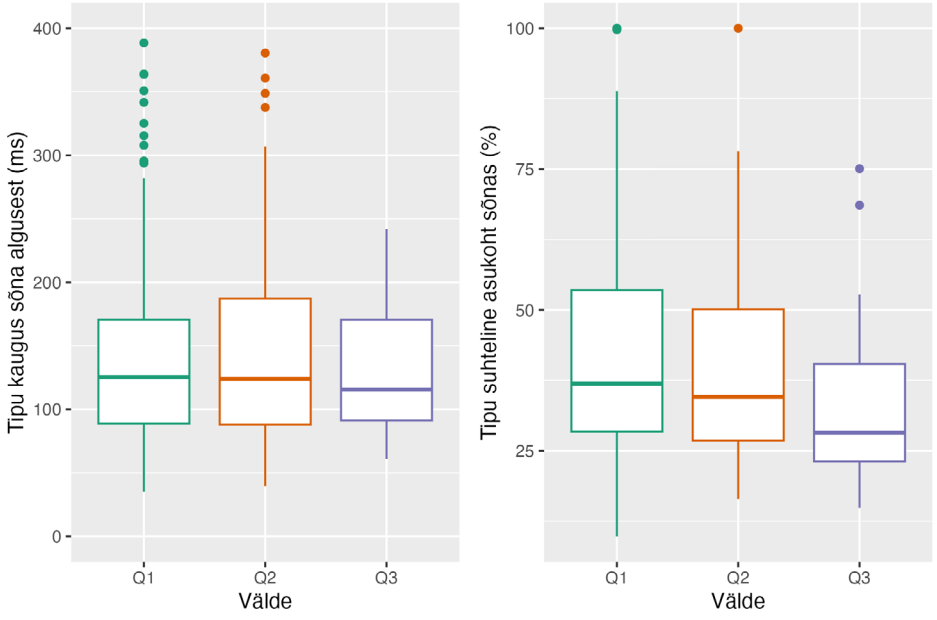


Joonis 5. Keskmised põhitoonikontuurid. Kontuuri kohal on näidatud põhitooni tipu asukoht standardhälbe piiridega ja languse algust tähistav pöördepunkt (*PP*); kontuuri all on näidatud esimese ja teise vokaali paiknemine

5.4. Põhitooni tipu kaugus välteti

Jooniselt 5 nägime, et pöördepunkt võiks välte kirjeldamiseks olla parem mõõdik kui põhitooni absoluutne tipp, aga kuna tippu on lihtsam defineerida, siis testime järgnevalt, kas tipu asukoht on ka statistiliselt oluline välte eristaja. Joonise 5 (ja ka varasemate uurimuste põhjal) võib väita, et sõna piires on kolmandas vältes põhitooni tipp varasem kui esimeses ja teises vältes. Testime põhitooni tipu suhtelist asukohta sõna piires jällegi lineaarse segamudeliga. Teeme kaks analüüsi: esmalt vaatame tipu absoluutset kaugust sõna algusest (joonisel 6 vasakul) ning seejärel

teisendame tipu kauguse sõna algusest tipu suhteliseks asukohaks esimeses silbis (joonisel 6 paremal).



Joonis 6. Põhitooni tipu kaugus sõna algusest (vasakul) ja tipu proportsionaalne asukoht sõnas (paremal)

Tabelist 10 ja joonise 6 parempoolselt paneelilt näeme, et millisekundites mõõdetud tipu absoluutkaugus sõna algusest ei ole ei esimese ja teise välte (β = -0,015, t = -0,230, p = 0,819) ega ka teise ja kolmanda välte vahel (β = 0,087, t = -0,946, p = 0,347) oluliselt erinev – keskmiselt jääb tipp esimeses vältes 140 ms, teises vältes 150 ms ja kolmandas 130 ms kaugusele, aga varieeruvus on väga suur.

Tabel 10. Põhitooni tipu absoluutkaugust kirjeldava segamudeli fikseeritud mõjud

	Hinnang	**Standard-viga**	**Vabadus-astmete arv**	**t-väärtus**	**Pr(>|t|)**
(vabaliige)	-2,025	0,118	4,215	-17,153	<0,001
välde Q1	-0,015	0,067	34,674	-0,230	0,819
välde Q3	-0,082	0,087	77,604	-0,946	0,347

Kui aga vaadata tipu suhtelist asukohta sõnas (joonis 6 paremal, tabel 11), siis näeme, et see ei ole oluliselt erinev esimese ja teise välte vahel (β = 0,048, t = 0,966, p = 0,379), aga kolmandas vältes on tipp natuke varasem kui teises vältes (β = -0,179, t = -2,562, p = 0,019).

Tabel 11. Põhitooni tipu suhtelist asukohta kirjeldava segamudeli fikseeritud mõjud

	Hinnang	**Standard-viga**	**Vabadus-astmete arv**	**t-väärtus**	**Pr(>|t|)**
(vabaliige)	3,583	0,042	4,353	86,190	<0,001
välde Q1	0,048	0,050	4,876	0,966	0,379
välde Q3	-0,179	0,070	19,710	-2,562	0,019

6. Vokaalikvaliteet

Kui varasemad eesti vokaalisüsteemi kirjeldused on leidnud, et vokaalikvaliteet vältega seoses ei varieeru (Lehiste 1960; Liiv 1962) või siis varieerub võrdlemisi vähe (Eek & Meister 1998), siis Lippus jt (2013) tulemused näitasid, et esiteks on rõhulised vokaalid teises ja kolmandas vältes siiski oluliselt perifeersemad kui esimeses vältes, teiseks on rõhutud vokaalid üldiselt redutseeritumad kui rõhulises silbis, ning kolmandaks on rõhutud vokaalid kolmandas vältes oluliselt redutseeritumad kui esimeses ja teises vältes. Vokaalikvaliteeti saab iseloomustada nende formantide järgi. Selleks viime läbi formantanalüüsi.

6.1. Formantanalüüs

Praati toimetamisaknas kuvatakse formante punaste täpikestena spektrogrammi peale ja skaala kuvatakse vasakul servas (vt joonis 1). Formantanalüüsi tulemused ja täpsus sõltuvad sellest, mis sagedusvahemikust mitut formanti otsime. See peaks olema kooskõlas kõneleja kõnetrakti pikkusega. Formantanalüüsi seadeid saab sättida menüüs *Formants → Formant settings...*. Kõige üldisemalt võiks seada formantanalüüsi laeks (ingl *formant ceiling*) meeskõnelejatel 5000 Hz ja naiskõnelejatel 5500 Hz. Kõige rohkem vigu teeb automaatne formantanalüüs tagavokaalidega /u/ ja /o/, mille keele kõrgust väljendav sagedus F1 ja keele tagapoolsust väljendav F2 on väga lähestikku. Nende vokaalide puhul aitaks automaatanalüüsi tulemusi parandada, kui kohandada seadeid igale kõnelejale ja igale vokaalikategooriale (vt nt Escudero jt 2009), kuid piirdume siin praegu ainult meeste ja naiste eristamisega.

Klassikaliselt kasutatakse vokaalikvaliteedi kirjeldamisel ***sihtväärtuse*** mõistet: igal häälikul on mingid artikulatoorsed sihid, mille poole liigutakse, ja mingi aja

jooksul hääliku algusest saavutatakse siht, mida natuke aega hoitakse, et siis edasi liikuda järgmise hääliku suunas. Võiks arvata, et häälikupiiridel on siirded ühelt häälikult teisele ning keskosas on just sellele häälikule omane stabiilne periood. Tegelikult erilist stabiilsust ei ole, pidev liikumine toimub kogu aeg. Aga kui me soovime monoftonge kirjeldada vokaaliklassiti ja mitte arvestada väga palju seda, mis konsonantide kontekstis nad on, siis oleks mõistlik mõõta formante vokaali keskosast ja piisab ühest väärtusest ühe vokaali kohta. Kui me aga tahame kirjeldada diftonge või ka just nimelt konsonandi konteksti, siis oleks mõistlik mõõta formantide kontuuri dünaamiliselt terve vokaali ulatuses.

Teeme siin lihtsamal viisil: mõõdame ühe punkti väärtused vokaali keskelt. Kui Praati toimetamisaknas vertikaalne kursoriosa panna vokaali keskele ja horisontaalne kursoriosa asetada kohakuti formantidega, võib formandi väärtust lugeda vertikaalskaalal spektrogrammist vasakul. Siin analüüsis aga mõõdame formantväärtusi skriptiga vokaali keskpunktist, mille arvutame välja TextGrid-failis märgitud häälikupiiride põhjal. Kuna formantanalüüs teeb siiski küllaltki palju vigu, siis võtame formantväärtused mitte vokaali keskpunkti ühest ajahetkest, vaid vokaali keskosa keskmised väärtused. Nii anname üksikutele hälbivatele analüüsipunktidele väiksema kaalu.

Jooksutame Praatis skripti *korp_opik_fonkorp_formandid_vokaali_keskelt.praat* ja loeme tulemuseks saadud faili *formandid_tulemused.txt* R-i.

Koondame vokaali märgendeid nii, et märgend oleks ainult ühe tähemärgiga (kaotame ära pikkuse ja lisakvaliteedimärgid) ning teisendame SAMPA transkriptsiooni IPA sümboliteks, et sümbolid näeks kenamad välja. Tabel 12 annab ülevaate sellest, kui palju erinevaid vokaale andmestikus on.

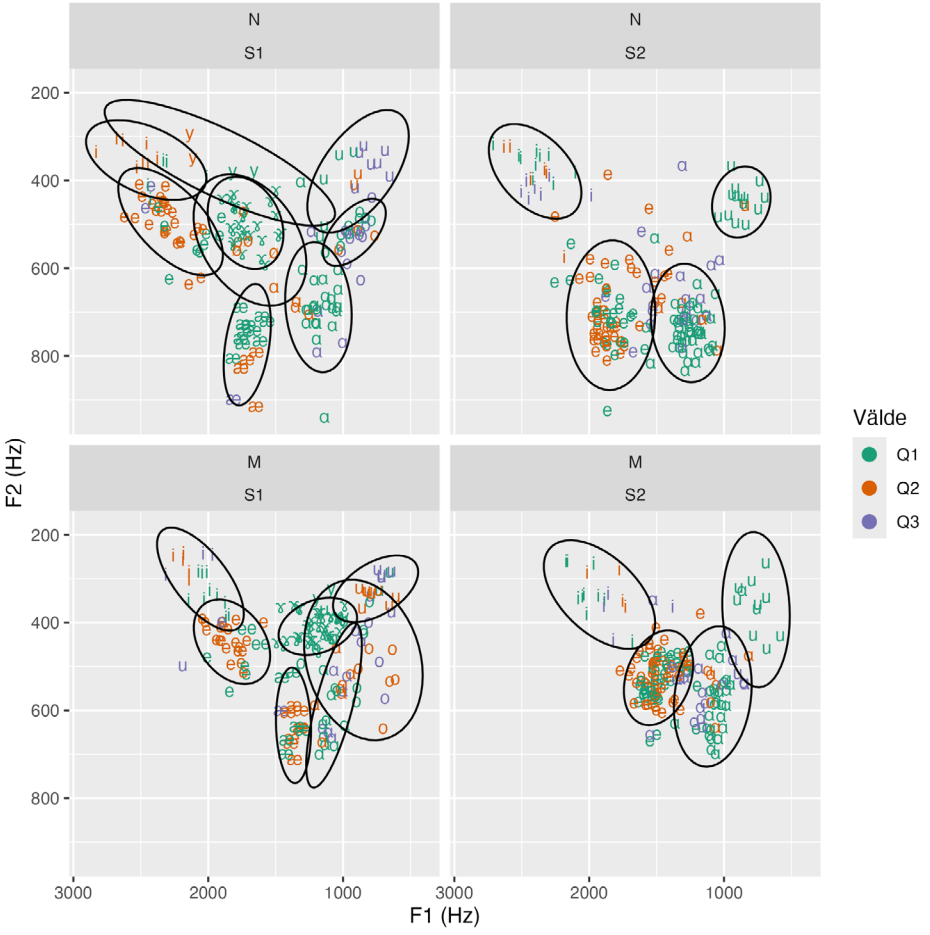
Tabel 12. Vokaalide arv

Välde	Silp	ɑ	e	i	u	o	y	æ	ɤ	ø
Q1	S1	33	17	15	8	20	5	17	69	
Q1	S2	71	68	20	24			1		
Q2	S1	11	48	13	10	10	2	13		4
Q2	S2	10	89	10	1	1				
Q3	S1	12	3	3	15	15		2		
Q3	S2	34	4	11	1					

Kuna meil on võrdlemisi väike andmestik, siis paratamatult ei esine siin kõiki vokaale kõigis vältusastmetes. Vähemsagedastest vokaalidest on andmestikus olemas /y/ ainult esma- ja teisevälteliste, /ɤ/ ainult esmavälteliste ja /ø/ ainult teisevälteliste sõnade rõhulistes silpides.

Järgsilbis on /æ/ märgitud ühes sõnas (*pähe*), aga kuna see on fonoloogiliselt /e/, siis võime selle ümber kategoriseerida.

Joonisel 7 on nüüd kujutatud vokaalid formantruumis hertsiskaalal, meeste ja naiste andmed eraldi.



Joonis 7. Mees- ja naiskõnelejate rõhuliste (S1) ja rõhutute (S2) silpide vokaalid F1-F2 formantruumis

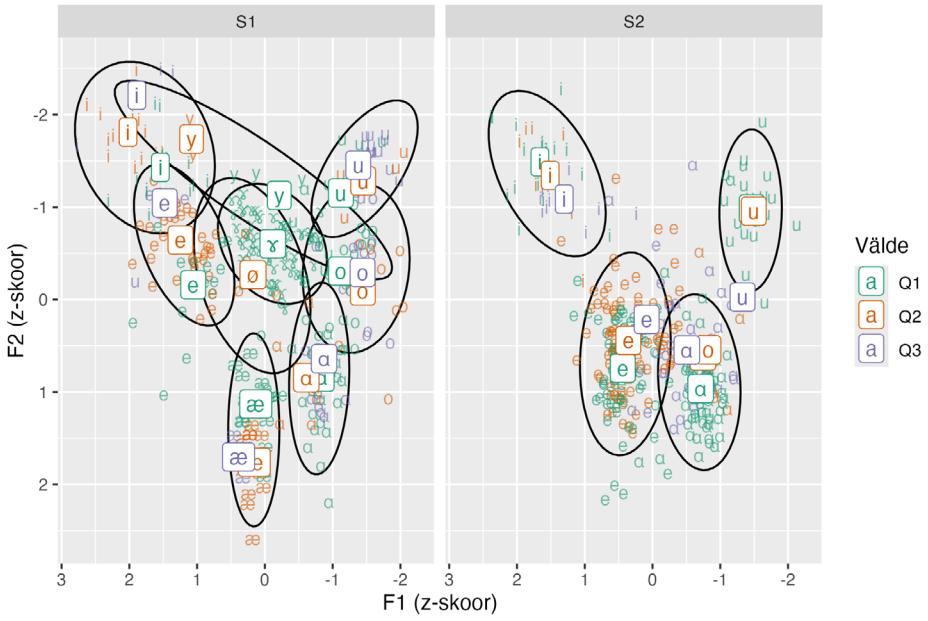
6.2. Formantväärtuste normaliseerimine

Kui põhitooniandmeid teisendatakse enamasti pooltooniskaalale, siis formantväärtuste esitamiseks on peale hertsiskaala ka teisi skaalasid, nt barkide, *mel*'ide või

ERB skaala, mis on inimtaju jaoks lineaarsed. Andmete teisendamine hertsidest barkideks muudab jaotust normaaljaotusele lähemaks, kuid ei kaota ära kõnelejatevahelisi kõnetrakti suurusest tulenevaid erinevusi.

Nagu põhitoon, varieeruvad ka formantväärtused kõnelejati. Suur erinevus on meeste ja naiste vahel, mis ei ole tingitud mitte erinevustest keelekasutuses, vaid kõnetrakti mõõtmetest. Kuna naistel on kõnetrakt veidi lühem kui meestel, siis on nende formantväärtused natuke suuremad. Aga ka samast soost kõnelejate vahel on varieerumist. Kui me tahame vaadelda vokaalide paiknemist kõneleja füsioloogilistest eripäradest sõltumata ja kujutada vokaaliruumi kõigi jaoks ühel joonisel, siis peaksime **formantandmeid normaliseerima**.

Üks normaliseerimisviis, kuidas vähendada kõnelejatevahelisi erinevusi nii, et jääks alles vokaalide omavahelised suhted vokaaliruumi sees, oleks teisendada väärtused **z-skoorideks**. Selle käigus teisendatakse kõneleja formantväärtused tema vokaaliruumi standardhälbe ühikuteks (vt allolevat valemit, kus x_i on formantväärtus, $\bar{x}$ on kõneleja formantväärtuste keskmine ning σ on standardhälve). Seda meetodit tuntakse foneetikas ka Lobanovi normaliseerimise all (Boris Lobanovi (1971) järgi, kes seda esimesena vokaalide kirjeldamisel kasutas, kuigi meetod on signaalitöötluses väga laialt kasutusel). Sellel normaliseerimisel on ainult üks häda: ei ole enam sisukat skaalat. Z-skoor kirjeldab ühe andmepunkti kaugust keskväärtusest rühma standardhälvetes. Seega on z-skooride ühikuks üks standardhälve ning väärtused võivad varieeruda umbes vahemikus -3 (standardhälvet) ja 3 (standardhälvet).



Joonis 8. Normaliseeritud vokaalide formantväärtused F1-F2 formantruumis

$$z = \frac{(x_i - \bar{x})}{\sigma}$$

Joonis 8 kujutab vokaalide normaliseeritud formantväärtusi formantruumis ilma sugu eristamata. *Z*-skoorideks teisendatud formantväärtuste puhul saab paremini võrrelda erinevate kõnelejate vokaalide paiknemist üksteise suhtes.

6.3. Vokaalide redutseerumine ja välde

Kuidas aga mõõta vokaalikvaliteedi seost vältega? Vokaalikvaliteet on kirjeldatud mitme formandi väärtusega, mis on seotud vokaalikategooriaga ja redutseerumine mõjutab eri vokaalikategooriate puhul väärtuste muutumist eri suundades: kui madal tagavokaal /ɑ/ redutseerub, siis tema F1 kahaneb ja F2 kasvab, aga kui kõrge eesvokaal /i/ redutseerub, siis tema F1 kasvab ja F2 kahaneb. Lahendus oleks üritada **normaliseerida vokaali kvaliteeti** vokaalikategooria suhtes.

Vokaalide redutseerumist üldiselt iseloomustab liikumine vokaaliruumi keskpunkti suunas. Et saada kahemõõtmelises (F1-F2) ruumis kirjeldatud andmed ühemõõtmeliseks, võiks mõõta iga vokaali kaugust vokaaliruumi keskpunktist. Selleks arvutame eukleidilise kauguse. **Eukleidiline kaugus** on otsekaugus kahe punkti vahel mitmemõõtmelises ruumis ja selle arvutamiseks kasutame Pythagorase teoreemi. Allolevat valemit vaadeldes võib kujutleda, et kui ja on vastavalt kõneleja ühe vokaali F1 ja F2 väärtus, siis ja on selle kõneleja formantruumi keskpunkti F1 ja F2 väärtused.

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2}$$

Eukleidilise kauguse arvutamiseks on parem pöörduda tagasi mõõdetud väärtuste juurde ja mitte kasutada z-skoorideks teisendatud väärtusi, sest algandmetel on ühikud, mille põhjal saab hinnata muutuse suurust. Logaritmilisuse eemaldamiseks teisendame andmed hertsiskaalalt **barki skaalale**, mis arvestab paremini tajuga. Hertside barkideks teisendamiseks leiab käsu *f2bark()* paketist `hqmisc` (Quene 2022).

Nüüd arvutame iga vokaali kauguse selle kõneleja kõigi vokaalide keskmisest. See ei pruugi küll antud andmestiku puhul väga head tulemust anda, sest eesti keele vokaaliruum on ise pisut ebasümmeetriline ja meil on tasakaalustamata hulk vokaale igalt kõnelejalt, mistõttu neist keskmine võib olla mõnevõrra juhuslik. Lippus jt (2013) artiklis valiti iga kõneleja vokaaliruumi keskpunkti leidmiseks ainult nn nurgavokaalid /ɑ i u/. Vokaalide kaugus formantruumi keskpunktist välteti on näidatud joonisel 9. Jällegi testime tulemusi lineaarse segamudeliga (mille väljund rõhulise silbi vokaalidega on tabelis 13, rõhutu silbi vokaalidega tabelis 14).

Tabel 13. Rõhulise vokaali kvaliteedi segamudeli fikseeritud mõjud

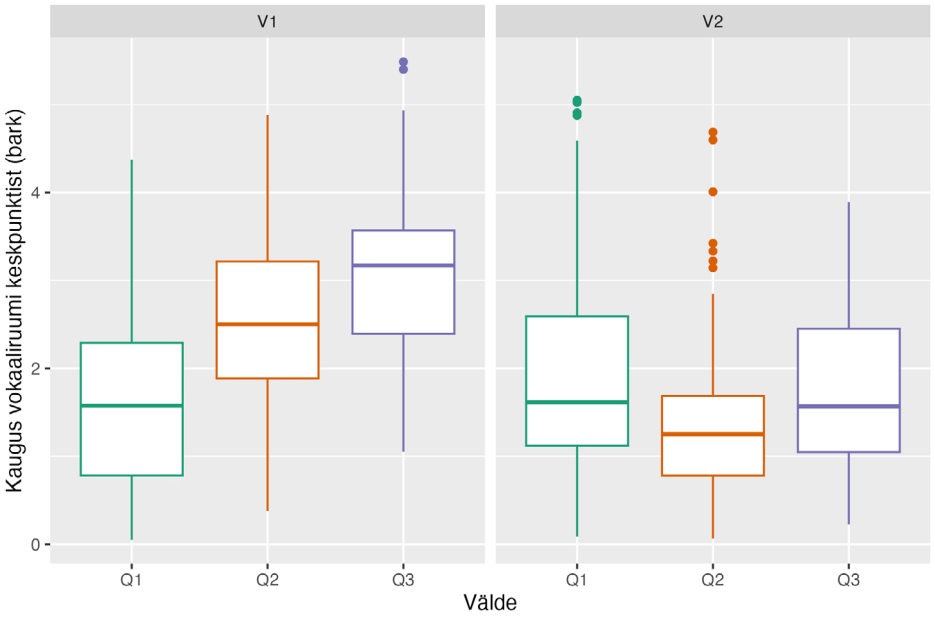
	Hinnang	Standard-viga	Vabadus-astmete arv	t-väärtus	Pr(>|t|)
(vabaliige)	2,943	0,182	45,331	16,211	<0,001
välde Q1	-0,851	0,214	112,315	-3,984	<0,001
välde Q3	0,101	0,185	322,385	0,547	0,584

Rõhulises silbis on esmavälteliste sõnade vokaalid tsentraalsemad kui teises vältes (β = -0,851, t = -3,984, p = <0,001). Erinevus teise ja kolmanda välte vahel ei ole statistiliselt oluline (β = 0,101, t = 0,547, p = 0,584).

Tabel 14. Rõhutu vokaali kvaliteedi segamudeli fikseeritud mõjud

Hinnang	Standard-viga	Vabadus-astmete arv	t-väärtus	Pr(>|t|)	
(vabaliige)	1,750	0,214	24,160	8,160	<0,001
välde Q1	0,452	0,233	112,043	1,939	0,055
välde Q3	-0,156	0,197	327,138	-0,790	0,430

Rõhutu silbi kvaliteedis ei ole siin andmestikus väldete vahel olulisi erinevusi (vt tabel 14).



Joonis 9. Rõhulise ja rõhutu silbi vokaalide kaugus formantruumi keskpunktist barkides väldete kaupa

Kokkuvõte

Siin peatükis analüüsisime kahesilbiliste sõnade häälikukestusi, põhitooni ja vokaalikvaliteeti, et kirjeldada eesti keele välteid. Andmed kogusime Praati skriptidega eesti keele spontaanse kõne foneetilisest korpusest pärit failidest. Nagu paljud varasemad uurimused on näidanud, eristab eesti keele kolme väldet kõige paremini rõhulise ja rõhutu silbi (vokaalide või silbiriimide) kestuste suhe, mis võtab kokku ühe mõõdikuna rõhulise silbi pikenemise ja rõhutu silbi lühenemise, mis vältega kaasneb.

Lisaks kestustele eristab välteid põhitooni langus, mis on absoluutajas mõõdetuna teisevältelistes sõnades veidi hilisem kui esma- ja kolmandavältelistes sõnades. Kui aga vaadata põhitooni joondumist suhtelisel ajateljel sõna piires, siis eristub esimesest ja teisest kolmas välde, kus langus toimub varem. Põhitooni tipp on esimeses vältes tihti teise silbi alguses, teises vältes esimese silbi lõpus ja kolmandas vältes esimese silbi esimeses pooles.

Vokaali kvaliteedi varieerumise seotust vältega nägime siin uurimuses ainult rõhulise silbi vokaali puhul ja tulemused olid kooskõlas varasemate tulemustega: esimeses vältes oli see teise vältega võrreldes oluliselt tsentraliseeritum, teise ja kolmanda välte vahel erinevust ei olnud. Rõhutus silbis välteerinevust ei õnnestunud

välja tuua. Peab siiski arvestama, et siin kasutatud valim on väike – ainult nelja kõneleja materjal –, mistõttu võivad tulemust rohkem mõjutada üksiku kõneleja individuaalsed eripärad, kõnesituatsioon, sõna esinemiskontekst jm faktorid, mida siin uurimuses ei kontrollitud.

Näidisuurimuse valmimist on toetanud projekt EKKD-TA6 „Liigutustest kõneni: suulise eesti keele multimodaalne analüüs" (2024–2027).

Kirjandus

Asu, Eva Liina, Pärtel Lippus, Karl Pajusalu & Pire Teras. 2016. *Eesti keele hääldus* (Eesti keele varamu 2). Tartu: Tartu Ülikooli Kirjastus. http://hdl.handle.net/10062/57960.

Asu, Eva Liina, Pärtel Lippus, Nele Salveste & Heete Sahkai. 2016. F0 declination in spontaneous Estonian: Implications for pitch-related preplanning in speech production. *Speech Prosody 2016*, 1139–1142. Boston, MA: Boston University. https://doi.org/10.21437/SpeechProsody.2016-234.

Barreda, Santiago. 2023. phonTools: Functions for phonetics in R.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Boersma, Paul & David Weenink. 2023. Praat: Doing phonetics by computer. https://www.praat.org.

Bořil, Tomáš & Radek Skarnitzl. 2016. Tools rPraat and mPraat. Petr Sojka, Aleš Horák, Ivan Kopeček & Karel Pala (toim), *Text, Speech, and Dialogue: Proceedings of the 19th International Conference, TSD 2016, Brno, Czech Republic, September 12-16, 2016*, 367–374. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-45510-5_42.

Coretta, Stefano. 2024. *speakr: A wrapper for the phonetic software Praat*. Manual. https://CRAN.R-project.org/package=speakr.

De Looze, Céline & Daniel Hirst. 2008. Detecting changes in key and range for the automatic modelling and coding of intonation. *Proceedings of the 4th International Conference on Speech Prosody*, 135–138. Campinas, Brazil. https://doi.org/10.21437/SpeechProsody.2008-32.

Eek, Arvo & Einar Meister. 1998. Quality of standard Estonian vowels in stressed and unstressed syllables of the feet in three distinctive quantity degrees. *Linguistica Uralica* 34(3). 226–233. https://doi.org/10.3176/lu.1998.3.11.

Escudero, Paola, Paul Boersma, Andréia Schurt Rauber & Ricardo A. H. Bion. 2009. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *The Journal of the Acoustical Society of America* 126(3). 1379–1393. https://doi.org/10.1121/1.3180321.

Jadoul, Yannick, Bill Thompson & Bart de Boer. 2018. Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics* 71. 1–15. https://doi.org/10.1016/j.wocn.2018.07.001.

Krull, Diana. 1997. Prepausal lengthening in Estonian: Evidence from conversational speech. Ilse Lehiste & Jaan Ross (toim), *Estonian Prosody: Papers from a Symposium*, 136–148. Tallinn: Institute of Estonian Language.

Lehiste, Ilse. 1960. Segmental and syllabic quantity in Estonian. Thomas A. Sebeok (toim), *American Studies in Uralic Linguistics* (Uralic and Altaic Series), 21–82. Bloomington: Indiana University Publications.

Liiv, Georg. 1961. Eesti keele kolme vältusastme vokaalide kestus ja meloodiatüübid. *Keel ja Kirjandus* 7–8. 412–424, 480–490.

Liiv, Georg. 1962. On the quantity and quality of Estonian vowels of three phonological degrees of length. Antti Sovijärvi & Pentti Aalto (toim), *Proceedings of the Fourth International Congress of Phonetic Sciences*, 682–687. The Hague: Mouton & Co.

Lippus, Pärtel. 2026. *Akustilised meetodid foneetikas. Kõneanalüüs programmiga Praat*. Tartu: Tartu Ülikooli Kirjastus. https://kodu.ut.ee/~partel/akustilised-meetodid-foneetikas/.

Lippus, Pärtel, Kätlin Aare, Anton Malmi, Tuuli Tuisk & Pire Teras. 2023. Phonetic corpus of Estonian spontaneous speech v1.3. Institute of Estonian and General Linguistics, University of Tartu. https://doi.org/10.23673/re-438.

Lippus, Pärtel, Eva Liina Asu, Pire Teras & Tuuli Tuisk. 2013. Quantity-related variation of duration, pitch and vowel quality in spontaneous Estonian. *Journal of Phonetics* 41(1). 17–28. https://doi.org/10.1016/j.wocn.2012.09.005.

Lippus, Pärtel, Karl Pajusalu & Jüri Allik. 2011. The role of pitch cue in the perception of the Estonian long quantity. Sónia Frota, Gorka Elordieta & Pilar Prieto (toim), *Prosodic Categories: Production, Perception and Comprehension*, 231–242. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-0137-3_10.

Lippus, Pärtel & Juraj Šimko. 2015. Segmental context effects on temporal realization of Estonian quantity. Maria Wolters, Judy Livingstone, Bernie Beattie, Rachel Smith, Mike MacMahon, Jane Stuart-Smith & Jim Scobbie (toim), *Proceedings of the 18th International Congress of Phonetic Sciences*, 1–5. Glasgow: University of Glasgow.

Lobanov, Boris M. 1971. Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America* 49(2B). 606–608. https://doi.org/10.1121/1.1912396.

Piits, Liisi & Mari-Liis Kalvik. 2017. Varieeruva vältega sõnade häälduskuuringud kõnesünteesi teenistuses. *Eesti Rakenduslingvistika Ühingu aastaraamat. Estonian Papers in Applied Linguistics* 13. 123–140. https://doi.org/10.5128/ERYa13.08.

Quene, Hugo. 2022. *hqmisc: Miscellaneous convenience functions and dataset*. Manual. https://CRAN.R-project.org/package=hqmisc.

R Core Team. 2023. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.