

TALLINNA TEHNIKAÜLIKOOL  
Infotehnoloogia teaduskond

Maria Letta 164372IAPB

# **Android nutitelefonide mäng eesti keele vokaalide hääldustreeninguks**

Bakalaureusetöö

Juhendaja: Einar Meister  
PhD

Tallinn 2021

## **Autorideklaratsioon**

Kinnitan, et olen koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

Autor: Maria Letta

16.05.2021

## Annotatsioon

Käesoleva bakalaureusetöö eesmärk on Android operatsioonisüsteemile mängulise formaadiga nutitelefonirakenduse loomine, mis abistab kasutajat eesti keele vokaalide hääldustreeningul. Rakendus peab võimaldama kasutajal mänguvooru lõpus näha tagasisidet, mis peegeldab vokaalide hääldustulemusi ja annab vastavaid soovitusi häälduse parendamiseks.

Vokaalide kvaliteedi hindamisel tuvastatakse helisignaalist helirõhu taseme ja helikõrguse alusel kõne, seejärel teostatakse signaali formantanalüüs ning leitud formantsagedustest lähtudes tuvastatakse hääldatud vokaal ja hääldustäpsus.

Töö tulemusena valminud rakendus hindab vokaali kvaliteeti lähtudes kahest esimesest formandist, mis on otseselt seotud keele asendiga. Formantsagedusi võrreldakse eesti keele vokaalide referentsväärtustega, mis võimaldab rakendusel veenduda vokaali häälduse korrektsuses ning vajadusel kuvada kasutajale asjakohaseid soovitusi keele asendi muutmiseks. Vokaalide referentsväärtused valitakse vastavalt kasutaja soole ja vanusele.

Rakenduse valideerimiseks viidi läbi katse kuue isikuga, kellele anti võimalus rakendusega 15-minutilise katseperioodi vältel tutvuda. Küsimustiku vormis kogutud tagasisidest selgus, et vokaalide tuvastamine toimus kõigi katses osalenute arvates ja rakendusel nähti potentsiaali olla abiks hääldustreeningus.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 26 leheküljel, 5 peatükki, 20 joonist, 1 tabelit.

## **Abstract**

### **Android Smartphone Game for Estonian Vowel Pronunciation Training**

The aim of the thesis is to develop a smartphone game for the Android operating system that could guide a user in Estonian vowel pronunciation training. The application is required to provide the user with feedback regarding his/her pronunciation results after a game session has ended and, in addition, offer advice on changing his/her tongue position to improve his/her vowel pronunciation.

In determining the spoken vowel quality, speech detection depends on the sound pressure level and measured pitch. Subsequently, formants are derived from the confirmed speech signal by formant analysis. Finally, the spoken vowel, along with its quality, is derived from formant values.

The result of this thesis is an application that estimates the spoken vowel quality based on the first two formants, which are directly related to the position of the tongue. The first formant describes the tongue's height and the second formant reflects how forward its placement is. The formant frequencies are compared to Estonian vowels' reference values, which enables the application to confirm the vowel pronunciation's accuracy and, if necessary, provide the user with relevant information on how to improve his/her pronunciation by changing his/her tongue position. The vowel reference values are chosen according to the user's gender and age.

To verify the application, an experiment was carried out with six test subjects. The subjects were given a 15-minute testing period to try out the application. The feedback was collected in a questionnaire form. The results gathered from the questionnaire showed that vowel detection worked for all test subjects who saw potential in the application to be used for pronunciation training.

The thesis is in Estonian and contains 26 pages of text, 5 chapters, 20 figures, 1 table.

## Lühendite ja mõistete sõnastik

Diskreet	Signaali diskreetimisel saadud üksikväärtus
Formant	Kõnetrakti resonantssagedus, formantsagedus
F1	Esimene formant
F2	Teine formant
JSTK	<i>Java Speech Toolkit</i> teek
LPC	Lineaarprognoos
MARF	<i>Modular Audio Recognition Framework</i> raamistik
TarsosDSP	Digitaalse signaali töötlemise raamistik

## Sisukord

1 Sissejuhatus .....	9
1.1 Taust .....	9
1.2 Probleem .....	10
1.3 Eesmärk .....	11
1.4 Ülevaade tööst .....	11
2 Metoodika .....	12
2.1 Ülevaade objektist: vokaal .....	12
2.2 Ülevaade tööriistadest .....	14
2.2.1 TarsosDSP .....	15
2.2.2 Kõne tuvastamine .....	16
2.2.3 Formantanalüüs .....	17
2.3 Ülevaade protsessist .....	18
3 Tulemus .....	19
3.1 Mäng .....	19
3.1.1 Juhend .....	20
3.1.2 Kasutaja info sisestamise vaade .....	22
3.1.3 Uus mäng .....	23
3.1.4 Mängu tulemus .....	24
3.1.5 Tagasiside .....	25
3.2 Kõnetöötlus .....	26
4 Rakenduse valideerimine .....	29
4.1 Valideerimistulemuste analüüs .....	29
4.2 Järeldused .....	34
5 Kokkuvõte .....	35
Kasutatud kirjandus .....	36
Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks .....	38
Lisa 2 – Mängujuhendi leheküljed .....	39

## Jooniste loetelu

Joonis 1. Eesti keele rõhuliste vokaalide formantsagedused [5]. Näide meeste kõne põhjal. ....	13
Joonis 2. Keele asend vokaali /a/ hääldamisel (a) ja /i/ hääldamisel (b). ....	13
Joonis 3. Keele asend vokaali /e/ hääldamisel (a) ja /o/ hääldamisel (b). ....	14
Joonis 4. TarsosDSP audiotöötlemise protsessi kirjeldav skeem [6]. ....	15
Joonis 5. Lineaarprognoosi spekter vokaali /a/ hääldamisel. ....	17
Joonis 6. Lineaarprognoosi spekter vokaali /ö/ hääldamisel. ....	18
Joonis 7. Vaadetevahelise navigatsiooni ülevaade. ....	19
Joonis 8. Mängu avamenüü. ....	20
Joonis 9. Juhendi teine lehekülg „Kasside nõrkus“. ....	21
Joonis 10. Juhendi seitsmes lehekülg „F1 erinevuse näide“. Võrdluses on /i/ ja /a/. ....	22
Joonis 11. Kasutaja info sisestamise vaade. Sisestatud on kasutaja vanus ja sugu. ....	23
Joonis 12. Mängu vaade. Peategelasele läheneb /i/ vokaali nõrkusega vaenlane. ....	24
Joonis 13. Mängu tulemuse vaade võidu korral. ....	25
Joonis 14. Mängu tagasiside vaade vokaali /ö/ hääldusele. ....	25
Joonis 15. Mängujuhiste arusaadavus. ....	30
Joonis 16. Rakenduse vokaalivastuse toimivus. ....	31
Joonis 17. Tulemuste tagasiside mõistmine ja võime sellest õppida. ....	31
Joonis 18. Mängu graafiline kujundus. ....	32
Joonis 19. Rakenduse efektiivsus häälduse treenimisel. ....	32
Joonis 20. Huvi rakendust kasutada. ....	33

## **Tabelite loetelu**

Tabel 1. Testgrupi liikmete sooline ja vanuseline jaotuvus.....	30
---	----



# 1 Sissejuhatus

Järgnev peatükk keskendub käesoleva bakalaureusetöö eesmärgi püstitamisele. Selleks pööratakse tähelepanu hääldustreeningu taustale ja antakse ülevaade olemasolevatest hääldustreeningu rakendustest. Seejärel tuvastatakse probleem ja seatakse eesmärk probleemi lahendamiseks. Viimasena tutvustatakse töö sisu.

## 1.1 Taust

Igal vokaalil on selle hääldamiseks iseäralik keele asend. Seetõttu on kõnevilumuse saavutamisel oluline vokaalide suuline praktiseerimine. Järjepidev vokaalide hääldustreening arendab keele asendiga seotud lihasmälu ning täiustab selle kaudu vokaalide korrektset produtseerimist. See on tähtis keeleõppijate jaoks, kes peavad võõraste vokaalide hääldamiseks harjumatu keele asendeid harjutama. Samuti aitab hääldustreening kõnearengu probleemidega lapsi, kelle vokaalide hääldamise oskuses esineb olulisi puudujäike.

Tänapäeval on loodud mitmeid hääldustreeningu rakendusi, mis võimaldavad kasutajal iseseisvalt vokaalide hääldamist treenida ning saada vahetut tagasisidet oma sooritusele. Järgnevalt tutvustatakse mõnda sellist rakendust tuues näiteid vokaalide hääldustreeningu võimalustest eesti keeles.

Arvutiprogramm *EstonianVowelCAPT* [1] on eesti keele vokaalide hääldustreeningu programm. Programm pakub iga eesti keele vokaali hääldamiseks harjutusi, kus kasutaja saab lindistada oma vokaali hääldust ja näha ekraanil oma tulemuse täpsust. Tulemus, mis peegeldab kasutaja keele asendit, kuvatakse koordinaatteljestikus, kus y-teljele märgitakse keele kõrgus ja x-teljele keele ees-tagapoolsus. Lisaks vokaali hääldustulemusele on teljestikku kantud ka referentsala, mis iseloomustab korrektset keele asendit vastava vokaali hääldamisel. Kasutaja eesmärk on saavutada hääldustulemus, millele vastavad x- ja y-telje väärtused asuvad teljestiku referentsalas. Rakendus pakub kasutajale ka kuulamisharjutust eesti keele völdete eristuse õppimiseks.

Estoñol [2] on eesti keele vokaalide hääldustreeningu rakendus Android seadmetele. Rakenduse sihtgruppi kuuluvad eesti keelt õppida soovivad inimesed, kelle emakeeleks on hispaania keel. Tähelepanu pööratakse seitsme vokaalipaari /i-ü/, /u-ü/, /a-o/, /a-ä/, /e-ä/, /o-ö/, /o-õ/ taju ja häälduse arendamisele. Vokaalipaaride moodustamisel on võetud eesmärgiks luua võimalikult väljakutsuvad harjutused sihtgrupi jaoks. Kui kasutajal ebaõnnestub hääldusharjutuse vältel vähemalt 3 järjestikusel korral vokaali korrektne hääldamine, näeb kasutaja ekraanil häälduse parendamiseks nõuannet.

Hääldustreeningu rakendused on levinud abivahendid hääldustreeningus ja lisaks eesti keelele on sarnaseid rakendusi loodud ka teiste keelte õppimiseks. Veebipõhine hääldustreeningu platvorm *The Computer-Assisted Listening and Speaking Tutor*<sup>1</sup> (CALST) pakub hääldustreeninguks harjutusi nii inglise kui ka norra keelele. CALSTi harjutuste valik tugineb kasutaja emakeelele, sest selle kaudu on võimalik ennustada võimalikke raskuseid, mis õpitava keele häälduses esineda võivad.

## 1.2 Probleem

Kõnevilumuse saavutamisel on oluline vokaalide suuline praktiseerimine, kuid eesti keele vokaalide hääldustreeningu rakendusi leidub väga vähe, millest tulenevalt puudub kasutajal valikuvõimalus enda jaoks sobivaima hääldustreeningu abivahendi leidmiseks. Kaks vastavat rakendust toodi esile eelnevas alapeatükis. Estoñol [2] rakenduse sihtgrupiks on eesti keele õppijad, kelle emakeeleks on hispaania keel. EstonianVowelCAPT [1] on seevastu suunatud laiemale kasutajaskonnale, kuhu kuuluvad eesti keelt võõrkeelena õppijad ja logopeedilist abi vajavad lapsed. Kuigi EstonianVowelCAPT arvutiprogrammi võimaliku sihtgrupina on arvestatud ka lapsed, tundub programmi kasutajaliides lastele iseseisvaks kasutamiseks liialt keeruline.

Eelnevast tulenevalt võiks lisanduda olemasolevatele eesti vokaalide hääldustreeningu rakendustele ka selline rakendus, mis võimaldaks vokaalide hääldustreenida mängulises formaadis. Kõnealune rakendus oleks eelkõige suunatud lastele, kuid sobiks kasutamiseks ka täiskasvanutele.

---

<sup>1</sup> <https://www.ntnu.edu/isl/calst>

### 1.3 Eesmärk

Käesoleva töö eesmärgiks on luua eesti keele vokaalide hääldustreeningut võimaldav nutitelefonis mäng Vokaalimäng, mille sihtgruppi kuuluvad logopeedilist abi vajavad lapsed ja eesti keelt võõrkeelena õppivad inimesed. Töö eesmärk on saavutatud, kui rakendus vastab järgnevatele kriteeriumitele:

- Võimaldab hääldustreeningut mängulises formaadis;
- Hõlmab harjutuses eesti keele 9 vokaali;
- Arvestab vokaali kvaliteedi hindamisel kasutaja soo ja vanusega;
- Kuvab kasutaja hääldustulemused mänguvooru lõpus;
- Jagab häälduse parendamiseks asjakohaseid soovitusi hääldustulemuste põhjal;
- Kasutaja ei taju mängus vokaali tuvastamisel viidet;
- Suunatud Android operatsioonisüsteemile;
- Arendatud Java programmeerimiskeeles.

### 1.4 Ülevaade tööst

Töö esimeses osas antakse ülevaade vokaali olemusest ja vokaali kvaliteedi hindamisest, lisaks tutvustatakse kõnetöötuse metoodikat ja kõnetöötuses kasutusel olevaid tööriistu. Teises osas keskendutakse töö tulemusele, kus tutvustatakse loodud mängu. Seal kirjeldatakse rakenduse kasutajaliidest, mängu loogikat ja implementeeritud kõnetöötuse protsessi. Töö viimases osas pööratakse tähelepanu töö valideerimismeetodi tutvustamisele, esitletakse kogutud valideerimistulemusi, neid analüüsitakse ja viimaks tehakse selle analüüsi põhjal ka järeldused.

## 2 Metoodika

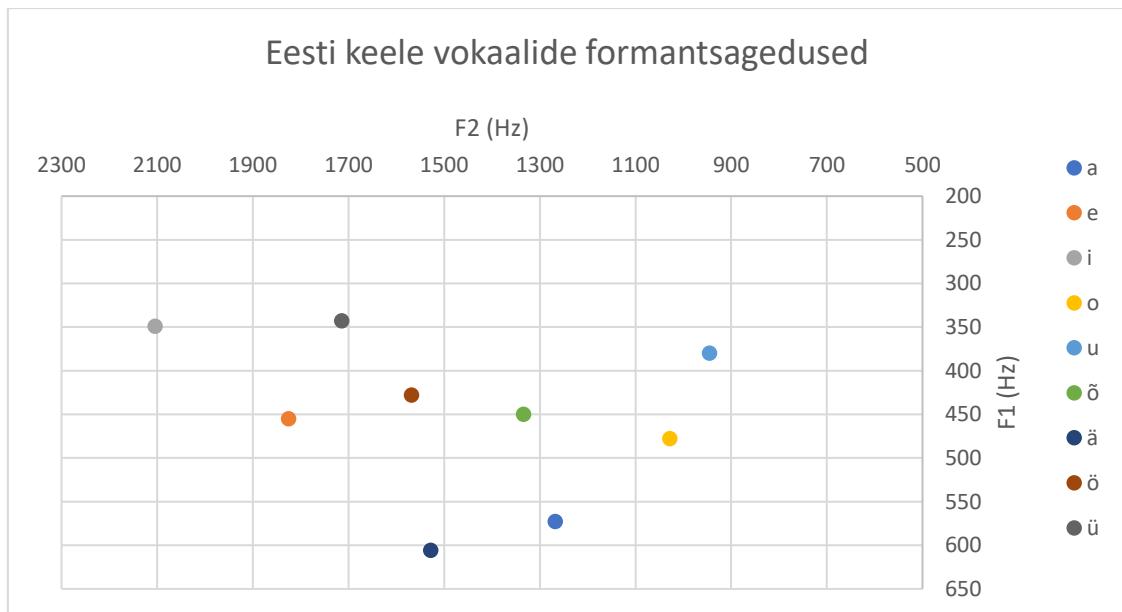
Järgnevas peatükis tutvustatakse kõigepealt vokaali olemust. Seejärel antakse ülevaade kõnetöötuse töövahenditest ja põhjendatakse arenduseks kasutatud raamistiku valikut. Viimasena selgitatakse töös rakendatud kõnetöötuse meetodeid ning kirjeldatakse töö protsessi eesmärgi saavutamisel.

### 2.1 Ülevaade objektist: vokaal

Vokaal on häälik, mille artikuleerimisel ei teki kõnetraktis sulgust ega ahtust ning õhuvool saab selle tulemusena takistuseta väljuda suust või suust ja ninast [3]. Vokaali ehk täishääliku hääldamisel on hääleallikaks häälekurrud, mis tekitavad võnkumisel heli. Hääli läbib filtri, milleks on kõnetrakt, mis muudab keele ja huulte liikumise tõttu pidevalt kuju. Kõnetrakt kujutab endast varieeruva kujuga resonantstorude kogumikku. Igal resonantstorul on lähtudes ruumalast ja kujust oma resonantssagedus, mida nimetatakse formandiks. Kõnetraktis võimenduvad resonantssageduste lähedased sagedused ning ülejäänud sumbuvad, mille tulemusena on igal vokaalil iseäralik spektrikuju [4].

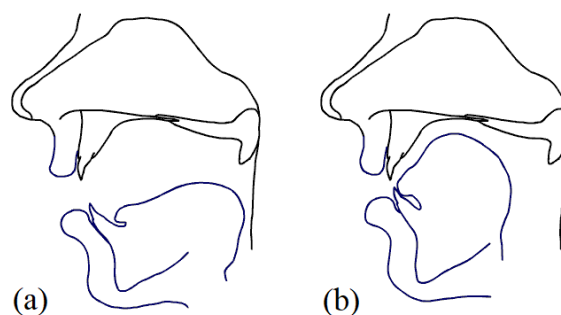
Vokaali kvaliteedi määramiseks piisab kahe esimese formandi (F1 ja F2) väärtusest, kuid täpsemaks kirjeldamiseks esitatakse kuni nelja formandi sagedused. Formandid on seotud vokaalide artikuloorsete põhitunnustega. F1 vastab keele kõrgusele, F2 keele ees-tagapoolsusele. Mida suurem on F1 väärtus, seda madalamal paikneb hääldamise ajal keel. Suurem F2 väärtus on iseloomulik eespoolsemale vokaalile. Vokaalide formantsagedused esitatakse tavapäraselt F1-F2 teljestikus [4].

Joonis 1 esitab näite eesti keele vokaalide formantsagedustest F1-F2 teljestikus. Näites peegeldavad andmed täiskasvanud meessoost isikute kõne keskmisi formantväärtusi rõhuliste vokaalide hääldamisel [5].



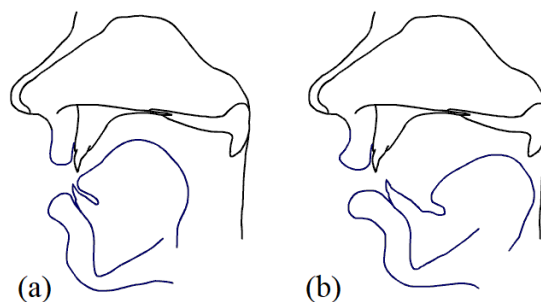
Joonis 1. Eesti keele rõhuliste vokaalide formantsagedused [5]. Näide meeste kõne põhjal.

Joonis 2 ja Joonis 3 piltide autor on Jürgen Lasn. Joonis 2 illustreerib keele asendit vokaalide /a/ ja /i/ hääldamisel. Uurides Joonis 1 andmeid on meeste kõnes vokaali /a/ tavapärase F1 väärtus ligikaudu 570 Hz, mis on vokaali /i/ korral umbes 350 Hz. Keele asend on /a/ hääldamisel (vt Joonis 2a) märkimisväärselt madalam kui /i/ artikuleerimisel (vt Joonis 2b), kuna /a/ F1 väärtus on oluliselt suurem kui /i/ F1 väärtus.



Joonis 2. Keele asend vokaali /a/ hääldamisel (a) ja /i/ hääldamisel (b).

Joonis 3 keskendub vokaalide /e/ ja /o/ hääldamise keele asendile, täpsemalt keele ees-tagapoolsusele. Võrreldes vokaalide F2 väärtusi, mis on välja toodud Joonis 1 teljestikul, ilmneb oluline suuruse erinevus: /e/ F2 väärtus on ligi 1820 Hz, seejuures /o/ F2 on umbes 480 Hz. Seetõttu on /e/ hääldamisel keele asend eespoolsem (vt Joonis 3a) kui /o/ puhul (vt Joonis 3b).



Joonis 3. Keele asend vokaali /e/ hääldamisel (a) ja /o/ hääldamisel (b).

## 2.2 Ülevaade tööriistadest

Android operatsioonisüsteemile suunatud nutitelefonide rakenduse loomisel kasutatakse Android Studio arenduskeskkonda. Arenduskeeleks on valitud Java, kuna töö autoril on vastava programmeerimiskeelega kõige pikemaajalisem kogemus.

Kõne analüüs hõlmab helisignaalist kõne tuvastamist ja seejärel signaalil formantanalüüsi teostamist. Lähtudes rakenduse märgulisest formaadist on oluline, et kõne analüüs toimiks reaajas ja kasutaja ei tajuks mänguvoorude ajal vokaali tuvastamisel viidet. Seetõttu otsiti kõnetöötlemiseks raamistikku, mis võimaldaks heli analüüsida otse helisignaali diskreetimisel saadud üksikväärtuste ehk diskreetide põhjal. Lisaks reaajas opereerimise nõudele, on töövahendite oluliseks kriteeriumiks ka sobivus Androidi virtuaalmasinaga.

Rakenduse arendamiseks katsetati *Modular Audio Recognition Framework*<sup>1</sup> (MARF) raamistikku, mille võimaluste seast oleksid käesolevas töös kasulikud helikõrguse leidmine ja lineaarprognostise teostamine. MARF kujunes aga sobimatuks, kuna selle integreerimisel projekti selgus, et implementeeritud klassid sõltusid Java paketi `javax.sample`, mida Android rakendusliides ei toeta. Sobivaks osutus TarsosDSP [6] raamistik, mis võimaldab reaajas audio tunnusoonte eraldamist ja omab Android tuge. Antud raamistikus on implementeeritud muu hulgas vahendid audio helirõhu taseme ja helikõrguse leidmiseks ning mõõdetud audiosignaali diskreetimissageduse muutmiseks. Nimetatud võimalused on töö eesmärgi saavutamisel vajalikud. Formantanalüüsi tarbeks

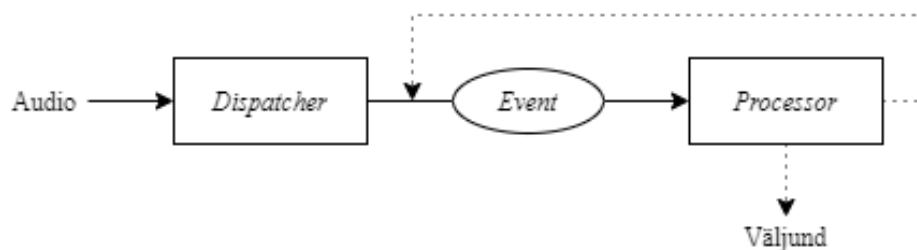
---

<sup>1</sup> <http://marf.sourceforge.net/docs/marf/0.2.0/report.pdf>

kaaluti Recognito<sup>1</sup> raamistikku ja *Java Speech Toolkit*<sup>2</sup> (JSTK) teeki. Mõlemad pakkusid lineaarproгноosi meetodi formantanalüüsiks, ent JSTK osutus valituks, kuna sellel teegil on implementeeritud lineaarproгноosi spektrist formantide tuletamine, mis Recognitol puudus. Lisaks on JSTK meetodite kasutamine näidetega illustreeritud, mis tegi sellega tutvumise kiireks ja mugavaks.

### 2.2.1 TarsosDSP

TarsosDSP [6] on audiotöötamise Java raamistik. Joonis 4 kirjeldab TarsosDSP sisemist protsessi heli töötlemisel. Katkendlikud nooled joonisel kujutavad valikulisi protsessivooge ja audio allikaks võib olla helifail või mikrofoni. Töö korraldaja *AudioDispatcher* objekt (joonisel *Dispatcher*) loeb sisse diskreete vastavalt etteantud analüüsiajaskaala suurusel. Diskreetide amplituud skaleeritakse vahemikule  $[-1, 1]$  vastavaks ja lisatakse sündmust kujutavale *AudioEvent* objektile (joonisel *Event*), mida hakatakse järgemööda edastama *AudioDispatcher*i külge registreerunud audiot töötlevatele *AudioProcessor* objektidele (joonisel *Processor*). *AudioProcessor*id võivad töötlemise tulemusena muuta *AudioEvent* sündmuse küljes olevat diskreetide massiivi või genereerida väljundi, millele reageerib töötleja külge registreeritud väljundi kuulaja. Järgmiste diskreetide lugemine on võimalik kui kõik registreeritud *AudioProcessor*id on lõpetanud töö viimase *AudioEvent* andmetega [6].



Joonis 4. TarsosDSP audiotöötamise protsessi kirjeldav skeem [6].

TarsosDSP tööprotsessi lihtsus on raamistiku eeliseks. TarsosDSP eesmärgiks oli luua praktiline raamistik, mis oleks ühtlasi keerukuse poolest hea sissejuhatus audiotöötlemisse sellega alustavatele tudengitele [6]. Raamistiku kasutamine on seeläbi mugav ja lihtne on

---

<sup>1</sup> <https://github.com/amaurycrickx/recognito>

<sup>2</sup> <https://github.com/sikoried/jstk>

kohandatud AudioProcessori lisamine, mis implementeeriks TarsosDSP võimaluste seast puuduvat analüüsi meetodi.

### 2.2.2 Kõne tuvastamine

Kõne olemasolu tuvastamiseks kasutatakse käesolevas töös helirõhu taset ja helikõrgust. Helirõhu taset jälgides on võimalik tuvastada ajahetk kui heli valjus ületab seatud lävendi, millest järeldub, et helisignaali võib sisaldada kõne. Seejärel saab helikõrguse põhjal hinnata, kas tuvastatud valjema heli puhul on potentsiaalselt tegemist inimese kõnega võrreldes sel hetkel leitud helikõrgust inimkõne tavapärase põhitooni sagedusvahemikuga. Mõlema heli tunnusjoone leidmisel kasutatakse TarsosDSP raamistikus olemasolevaid vahendeid.

**Helirõhu taseme** jälgimiseks pole vaja eraldi audio töötajat TarsosDSP töö korraldaja külge registreerida. Selle info on kättesaadav kõigile töötajatele saadetavalt sündmuse objektilt. Kõne helirõhu taseme (mõõtühik dB-SPL) arvutamisel kasutab TarsosDSP allolevat valemit, kus  $b$  tähistab diskreetide massiivi suurusega  $n$  [7]:

$$L_p = 20 \log_{10} \frac{\sqrt{\sum_{i=0}^n b^2[i]}}{n}$$

**Helikõrguse** mõõtmiseks on TarsosDSP töötajate seas implementeeritud PitchProcessor, millel saab määrata helikõrguse hindamisel kasutatava algoritmi implementeeritud meetodite seast. Saadaval on näiteks järgnevad analüüsimeetodid: YIN, *McLeod Pitch Method* (MPM) ja *Average Magnitude Difference Function* (AMDF). YIN meetodil on mitu implementatsiooni: ajavallal põhinev YIN ja sagedusvallal põhinev YIN FFT [7]. Vastavate meetodite efektiivsust võrreldes on kõige täpsemaks osutunud YIN ning YIN FFT meetodid [8].

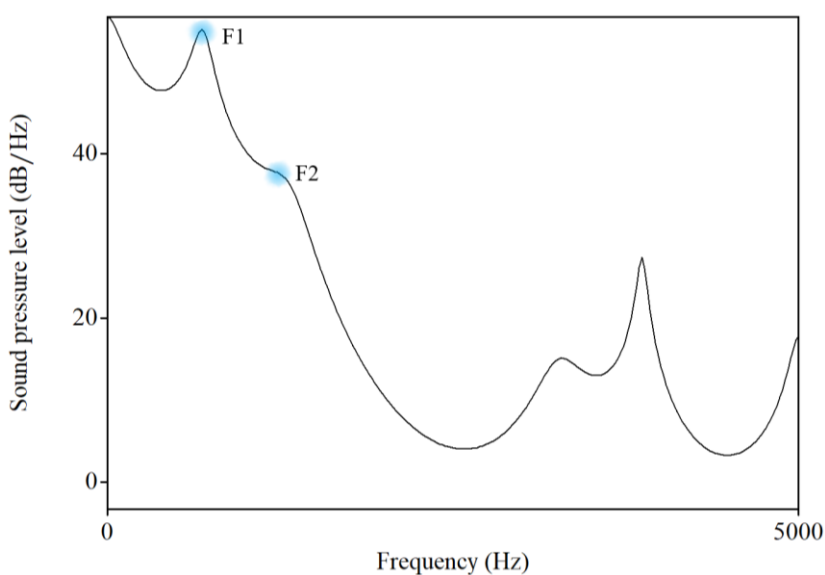
YIN meetod on autokorrelatsioonifunktsiooni edasiarendus, mille eesmärk oli vähendada autokorrelatsiooni meetodi vigu. YIN meetodi arendajate hinnangul teeb YIN kolm korda vähem vigu kui konkureerivad meetodid [9]. YIN FFT on YIN meetodiga sarnane, ent ligi kolm korda kiirem tänu Fourier' teisendusele diferentseeruva funktsiooni arvutamisel [8].



### 2.2.3 Formantanalüüs

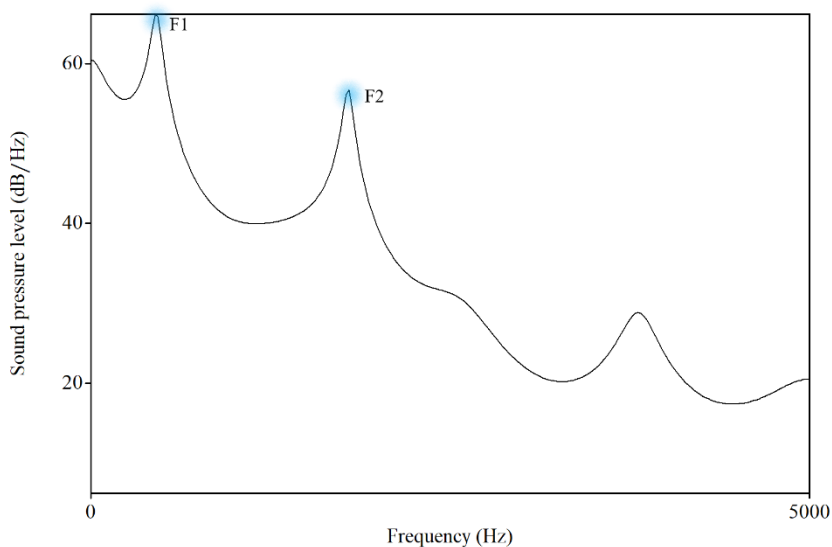
Formantanalüüsiga leitakse signaalist formantide väärtused. Peamine meetod, mida formantanalüüsil rakendatakse, on lineaarprognosi (LPC) meetod. Lineaarprognos on ajavallal põhinev meetod, mis modelleerib kõnesegmendi produtseerinud kõnetrakti kui filtri omadusi. Kõnesegmendist eraldatakse hääleallika signaal ja filtri kordajad, mille alusel oleks võimalik väljundsignaal taas konstrueerida. Filtri kordajate määramisel tuginetakse ideele, et signaali väärtust on võimalik tuletada sellele eelnenud väärtuste põhjal. Filtri kordajate abil saab tuletada LPC spektri [10].

Järgnevad joonised kujutavad LPC spektreid, mille maksimumid kirjeldavad formante. Joonised on loodud Praat foneetikaprogrammi [11] abiga. F1 ja F2 asukohad on markeritud sinise värviga. Joonis 5 peegeldab vokaali /a/ hääldamisel tekkinud LPC spektrit.



Joonis 5. Lineaarprognosi spekter vokaali /a/ hääldamisel.

Joonis 6, millel kujutatud spekter kirjeldab hääldatud vokaali /ö/, on toodud võrdluseks /a/ vokaali spektrile. /ö/ ja /a/ spektrite uurimisel saab F1 ja F2 väärtuste põhjal järeldada, et /ö/ hääldamisel on keele asend kõrgem ja eespoolsem.



Joonis 6. Linearprognosi spekter vokaali /ö/ hääldamisel.

### 2.3 Ülevaade protsessist

Tööd alustatakse kõnetötluse teooria ning olemasolevate hääldustreeningu rakendustega tutvumisega. Seejärel pööratakse tähelepanu kõnetötluse raamistike/teekide otsimisele ja katsetamisele, et leida sobivad abivahendid rakenduse kõnetötluse osa implementeerimiseks.

Rakenduse loomine algab kõnetötluse arendamisest, mis hõlmab helisignaalist kõne tuvastust ja formantanalüüsi abil vokaali kvaliteedi määramist. Paralleelselt arendustööga alustatakse mängu animatsioonikaadrite ja taustpiltide visandamist ning joonistamist rakenduse mängugraafika kunstniku Marilyn Letta abiga. Kõnetötluse järel keskendutakse mängu realiseerimisele ning seejärel ülejäänud vaadetele, nende hulgas tagasiside kuvamisele.

Arendustöö lõpus testitakse rakendust 6 isiku peal. Igale testgrupi liikmele antakse võimalus rakendusega iseseisvalt tutvuda 15-minutilise katseperioodi vältel. Seejärel kogutakse küsimustikuga rakenduse kasutamise kogemuse kohta tagasisidet, et tagasiside tulemusi analüüsida ja teha järeldusi rakenduse toimivuse kohta.

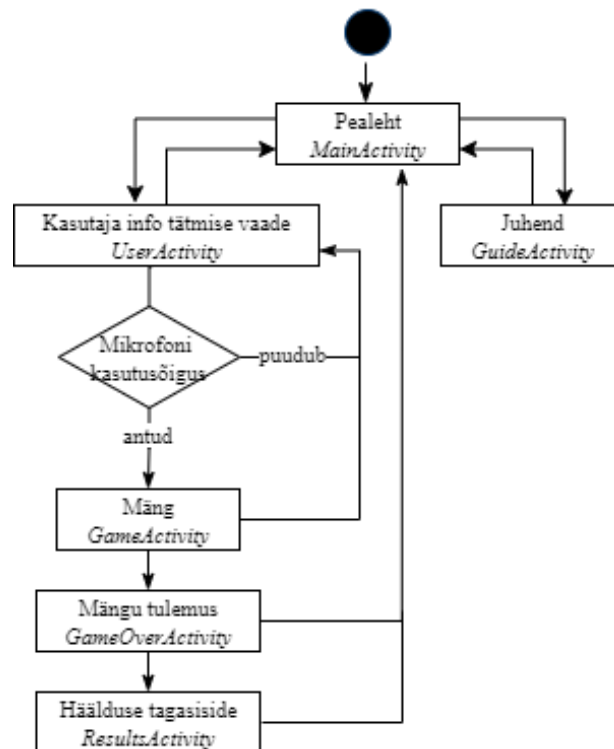
### 3 Tulemus

Käesolevas peatükis kirjeldatakse töö arenduskäiku, seejuures langetatud otsuseid ja antakse ülevaade töö tulemusest. Peatükis kuvatud joonised on loodud ülesvõtetena Android Studio Android virtuaalseadme Pixel 3a nutitelefoni, mille Android versioon on 11.0.

Loodud rakendus toetab Androidi versioone alates versioonist 8.0, kuna arendustöö käigus kasutati Android rakendusliidese versiooni 26. Seadmele, mille Android operatsioonisüsteemi versioon on mainitust madalam, ei ole rakendus kättesaadav. Rakendus eeldab ka mikrofoni olemasolu mänguvoorus kõnetötluse võimaldamiseks.

#### 3.1 Mäng

Kuna rakenduse sihtgruppi kuuluvad lapsed, on rakenduse ülesehitus ja kujundus võimalikult minimalistlik, intuiitivne ja visuaalne. Joonis 7 esitab seosed vaadete vahel.



Joonis 7. Vaadetevahelise navigatsiooni ülevaade.

Mängu graafilises disainis on kasutatud animeeritud tegelasi. Peategelasel corgil on tavaoleku ja haukumise animatsioonid. Vaenlastel ehk kassidel on animeeritud tegevusteks kõndimine ja põgenemisel tolmupilve haihtumine. Animatsioonikaadrite ja mängus kasutatud taustpiltide autor on Marilyn Letta. Ühe animatsioonikaadri kestusajaks on sätestatud 42 ms, millest tuleneb animatsioonide liikumise kiirus 24 kaadrit sekundis. Kõik rakenduse vaated on rakenduse manifestis deklareeritud rõhtvorminguga ja telefoni pööramine vormingut ei muuda

Mängu käivitamisel alustab kasutaja pealehel (vt Joonis 8), kus tal on võimalik alustada uut mängu ja tutvuda mängujuhendiga.



Joonis 8. Mängu avamenüü.

### 3.1.1 Juhend

Mängujuhend koosneb 8 leheküljest. Juhendis tutvustatakse kasutajale mängu eesmärki, milleks on peategelase aitamine lähenevate vaenlaste minema peletamisel. Seejärel kirjeldatakse mängu võitmis- ja kaotamistingimusi, tuuakse visuaalne näide vaenlasest ning tema nõrkusest (vt Joonis 9) ja selgitatakse, et vaenlase võitmiseks peab hääldama tema kohal paiknevat vokaali.



Joonis 9. Juhendi teine lehekülg „Kasside nõrkus“.

Samuti sisaldab mängujuhend visuaalset näidist corgi haugatuse animatsioonist, mis on üheks vokaali tuvastamise indikaatoriks. Kasutajat julgustatakse vokaali hääldamist kordama, kui vokaali ei suudetud tuvastada.

Mängujuhendi keskel annab töö autor kasutajale soovitus, mida järgida kui vokaali tuvastamine sageli ebaõnnestub. Kirjeldatakse ka mikrofoni kasutust mängus ning kinnitatakse, et mänguvooru jooksul analüüsitud heli ei salvestata ega edastata mängu arendajale.

Viimased juhendi leheküljed keskenduvad peamiselt mänguvooru tagasiside formaadi selgitamisele. Varasema uurimistöö [1] tagasisides selgus, et kasutajad pidasid hääldamise õppimisel abistavaks animatsioone, mis peegeldasid keele asendi muutumist vokaalide hääldamisel. Sellest lähtuvalt on ka käesoleva töö mängujuhendisse lisatud animatsioonid (vt Joonis 10), millega visualiseeritakse kasutajale, mida tähendavad kõrge, madal, ees- ja tagapoolne hääldus keele asendi kontekstis. Kasutatud animatsioonide autor on Jürgen Lasn.



Joonis 10. Juhendi seitsmes lehekülg „F1 erinevuse näide“. Võrdluses on /i/ ja /a/.

Ülejäänud mängujuhendi leheküljed, mida antud peatüki jooniste hulgas ei esitletud, leiab lisade hulgast (vt Lisa 2).

Vaates kasutatakse fragmenti. Fragment esindab vaate taaskasutatavat osa ning võimaldab kuvada kõiki mängujuhendi lehekülgi ühes vaates, muutes vaid fragmendi instantsi. Fragmendi kasutamisega välditi duplitseeritud koodi. Nooltega pöörab kasutaja juhendi lehti. Esimesel leheküljel navigeerib tagasiviiv nool kasutaja avamenüü juurde.

### 3.1.2 Kasutaja info sisestamise vaade

Enne uue mängu alustamist peab kasutaja sisestama oma vanuse ja soo (vt Joonis 11), kuna formantsagedused sõltuvad kõnetrakti anatoomilisest suurusest [5]. Seeläbi on rakenduse vokaalitu vastus täpsem, kuna formantide referentsväärtused saab valida kasutaja soo- ja vanusepõhised.



Joonis 11. Kasutaja info sisestamise vaade. Sisestatud on kasutaja vanus ja sugu.

Nupu edasiliikumiseks on aktiivne vaid juhul kui kasutaja on sisestanud oma vanuse ja soo. Nupu staatust peegeldab värvitooni erinevus – aktiivne nupp on helesinine, mitteaktiivne nupp on hall. Edasiliikumisel siseneb kasutaja mängu, tagasiliikumisel suundutakse tagasi avamenüü juurde.

### 3.1.3 Uus mäng

Alates Android versioonist 6.0, millele vastav rakendusliidese versioon on 23, kehtib nõue taodelda kasutajalt ohtlikuks kategoriseeritud õigused käitusajal [12]. Mikrofone kasutamiseks vajaminev õigus RECORD\_AUDIO on kategoriseeritud ohtlikuks [13]. Sellest lähtudes küsitakse kasutajalt mänguvooru alguses mikrofone kasutamiseks luba juhul, kui vastavat õigust pole veel rakendusele antud. Loa andmisest keeldumisel pole mängu alustamine võimalik ja kasutaja suunatakse tagasi eelnenud vaatesse.

Mikrofone kasutusõiguse olemasolul algab uus interaktiivne mäng (vt Joonis 12), kus peategelast ründavad ühekaupa vaenlased, kelle peletamiseks peab kasutaja hääldama piisava korrektsusega vaenlase nõrkuseks olevat vokaali. Mängus esinevad kõik eesti keele 9 vokaali juhuslikus järjekorras 2 korda. Hääldades õiget vokaali enne kui vaenlane jõuab peategelaseni, haihtub ekraanil olev vaenlane ning lähenema hakkab järgmine. Ekraani keskel paiknev loendur aitab kasutajal järge pidada, mitmes vokaal on parajasti ekraanil.



Joonis 12. Mängu vaade. Peategelasele läheneb /i/ vokaali nõrkusega vaenlane.

Mänguvoor lõpeb, kui kasutaja peletab kõik 18 vaenlast hääldades korrektselt nende nõrkuseks olevad vokaalid. Vooru katkestab ka vaenlase jõudmine peategelaseni, mis juhtub, kui kasutaja ei jõua 10 sekundi vältel vastavale vokaalile korrektset hääldust produtseerida. Samuti on võimalik kasutajal mängust ise lahkuda, vajutades selleks ekraani paremas ülanurgas paiknevale mängu sulgemise nupule, mille tulemusena jõutakse tagasi kasutaja info sisestamise vaatesse.

### 3.1.4 Mängu tulemus

Mängusessiooni lõppedes kuvatakse kasutajale tulemus (vt Joonis 13). Juhul, kui mängija peletas kõik peategelasele lähenenud 18 vaenlast, kuvatakse tulemuseks „Võit!“. Kui aga mõni vaenlastest jõudis peategelaseni, kuvatakse tulemusena „Kaotus!“. Peategelase emotsiooni peegeldav illustratsioon muutub vastavalt tulemusele.





Joonis 13. Mängu tulemuse vaade võidu korral.

Tulemuse lehelt on võimalik edasi liikuda avalehele või tutvuda mänguvooru tagasisidega. Tagasiside lehele suundumise nupp on nähtav vaid juhul, kui kasutaja hääldas mänguvoorus vähemalt ühe tuvastatud vokaali, mille kohta tagasisidet anda.

### 3.1.5 Tagasiside

Tagasiside lehel (vt Joonis 14) on iga hääldatud vokaali kohta esitatud eraldi leht. Vokaali lehel on esitatud kasutaja hääldatud formantide sagedused koos referentsvahemikega. Kui rakendus tuvastas ja registreeris hääldatud vokaali mitmel korral, on lehele koondatud tulemuste veergu formantide aritmeetiline keskmine.



Joonis 14. Mängu tagasiside vaade vokaali /ö/ hääldusele.

Olenevalt hääldustulemusest antakse kasutajale tagasiside koos soovitusena, kuidas hääldust parandada. Kui F1 hääldustulemus on referentsvahemikust väljas, kuvatakse

nõuanne muuta keele kõrgust vastavalt madalamaks või kõrgemaks. F2 mittedobivusele soovitatakse keele asendit korrigeerida ees- või tagapoolsemaks. Eksides hääldamisel mõlema formandiga näeb kasutaja soovitus muuta keele asendit mõlema muutuja suhtes. Formantide referentsvahemikele vastavaks osutumisel kiidetakse kasutajat korrektse häälduse eest.

Sarnaselt mängujuhendi vaatele on ka tagasiside vaates kasutusel fragment. Fragmenti eesmärk on vältida koodikordusi. Noolega tagasi navigeerimisel naaseb kasutaja avamenüü juurde.

### **3.2 Kõnetöötlus**

Kõnetöötlusel kasutatakse raamistikku TarsosDSP [6]. TarsosDSP tööprotsess algatatakse pealõimest eraldi lõimes. Protsessi korraldaja loomisel konfigureeritakse sätted helisignaali mõõtmiseks. Sätestatakse diskreetimissagedus 44 100 Hz ja analüüsiakna suurus 2048 diskreeti. See tähendab, et ühe sekundi jooksul mõõdetakse 44 100 korda signaali väärtust ning analüüs teostatakse iga 2048 diskreeti järel, mille tulemusena on analüüsiakna pikkuseks ligi 46 ms. Diskreetimissagedus valiti Android dokumentatsiooni info põhjal, mille kohaselt on 44,1 kHz hetkel ainus diskreetimissagedus, mis on garanteeritud töötama igal Android seadmel [14]. Analüüsiakna suuruse valimisel konfigureeritakse väärtus, mille sätestamist TarsosDSP vastava diskreetimissageduse puhul toetab, sooviga ühtlasi saavutada võimalikult lühike analüüsiakna pikkus, mis tagaks ka lühikese ajaintervalli TarsosDSP sündmuste kajastamise vahel.

TarsosDSP töö korraldaja AudioDispatcher külge registreeritakse kõne tuvastamiseks raamistiku PitchProcessor objekt, millele on määratud helikõrguse hindamise meetodiks YIN FFT. Töös lisatakse implementatsioon PitchDetectionHandler liidesele, mis registreeritakse PitchProcesori väljundi kuulajaks. Selle kuulaja roll on väljundi põhjal järeldada, kas mõõdetud analüüsiaken sisaldab kõne. Järelduse tegemisel uuritakse esmalt kajastatud sündmuse infost helirõhu taset. Valjuse alusel peetakse mõõtetulemusi kõneks, kui helirõhu tase ületab TarsosDSP poolt vaikumisi sätestatud vaikuslävendi. Seejärel kontrollitakse tuletatud helikõrguse sobivust inimkõne põhitooni tavapärase vahemikuga, milleks on määratud 100 Hz kuni 300 Hz. Mõlema tingimuse paikapidavusel teavitab väljundi kuulaja teisi töötlejaid kõne tuvastamisest. Inimkõne põhitooni vahemiku

tuletamisel lähtuti andmetest, et meeshääle tüüpiline kõne põhitoon on vahemikus 85 Hz kuni 180 Hz ja naishäälele omane põhitoon 165 Hz kuni 255 Hz [15]. Arvestati ka laste kõne põhitooniga. 9-aastastel poistel on kõne keskmine põhitoon 235 Hz, mis langeb märkimisväärselt häälemurde perioodil keskmiselt 108 Hz võrra. 9-aastaste tüdrukute kõne keskmine põhitoon on 245 Hz [16].

Järgmise sammu teostamiseks on AudioDispatcheri töötlejaks registreeritud RateTransposer. RateTransposer vastutab signaali ümberdiskreetimise eest. Teisendamise vajadus tuleneb soovist vähendada formantanalüüsi etapi keerukust. Formantanalüüsiks kasutatakse lineaarprognoosi, mille filtrikordajate arv määrab formantanalüüsi keerukuse. Filtrikordajate arvu sätestamisel kehtib soovituslik reegel lisada 2 kordajat iga diskreetimissageduse kHz kohta ning lisaks veel 2 kordajat spektri kõrgemate sageduste mõjuga arvestamiseks [17]. Määratud diskreetimissageduse 44,1 kHz korral oleks filtrikordajate arvuks 46. Arvestades asjaoluga, et vokaali kvaliteedi hindamisel kasutatakse vaid kahe esimese formandi väärtusi, mis ei ületa sagedust 3 kHz, piisab kõne analüüsiks ka diskreetimissagedusest 8 kHz, mille korral on Nyquisti teoreemi järgi kõige suurem mõõdetud signaali sagedus 4 kHz. Diskreetimissagedus 8 kHz vähendab filtri järku arvuni 10 ja seeläbi ka arvutuskooormust.

Seejärel toimub formantanalüüs. Kuna TarsosDSP ei pakkunud töövahendeid formantsageduste leidmiseks, implementeeriti töö käigus TarsosDSP AudioProcessor liidesele vastav töötleja, mis kasutab formantanalüüsiks JSTK<sup>1</sup> teegi lineaarprognoosi. Lisatud lineaarprognoos kasutab signaali silumisel Hamming aknafunktsiooni, filtrikordajate arvutusmeetodina autokorrelatsiooni ja formantide tuletamisel LPC spektrit. Kuigi käesolevas töös hinnatakse vokaali kvaliteeti vaid 2 esimese formandi põhjal, otsustati neid lineaarprognoosist tuletada varuga 3, arvestades võimalusega, et tulemustes võib esineda ka müra.

Iga analüüsiakna töötlus, mida teostavad AudioDispatcheri külge registreeritud töötlejad, peab saama teostatud analüüsiaknaga kestuselt võrdsel ajavahemikul, mis on 46 ms. Vastasel juhul tekib infokadu kahe analüüsiakna mõõtmistulemuste vahel, kuna AudioDispatcher ei loe järgmist analüüsiakent uute diskreetidega enne, kui viimati loetud

---

<sup>1</sup> <https://github.com/sikoried/jstk>

analüüsiakna töötlemine on lõpetatud. Kuna formantanalüüs on kõige mahukam etapp kõnetöötamise protsessis, delegeeritakse formantanalüüsi tööjupid AudioDispatcheri tööloimest eraldiseisvatele lõimedele, mis talletavad saadud tulemusi vahemällu, ja ei peata AudioDispatcherit tööd.

Kui PitchProcessor väljundi kuulaja tuvastab kõne lõppemise, teostatakse formantanalüüsiks implementeeritud AudioProcessoris vahemällu talletatud formantide põhjal vokaali tuvastamine. Esmalt filtreeritakse välja formandid, mis esindavad müra. Müra filtreerimine sõltub kasutaja soost. Kahe esimese formandi põhisel analüüsil peetakse naissoo esindajatel müraks formante, mille väärtused on alla 300 Hz või üle 3000 Hz, ning meeste puhul vastavalt formantsagedusi, mis on madalamad kui 250 Hz või kõrgemad kui 2500 Hz. Allesjäänud formantväärtustest eristatakse F1 formandid ja F2 formandid ning arvutatakse mõlema grupi aritmeetilised keskmised. Vokaali hääldust sisaldanud analüüsiakende keskmisi formantväärtusi võrreldakse parajasti mängus ekraanil oleva vokaali F1 ja F2 referentsväärtustega, et jõuda selgusele, kas häälduse sooritus oli piisavalt edukas vokaali tuvastamiseks. Lähtudes rakenduse eesmärgist võimaldada kasutajal trennida eesti keele vokaalide hääldust, ei saa eeldada kasutajalt referentsväärtustele  $\pm 1$  standardhälve [5] vastavaid hääldustulemusi. Lubatud eksimuse määra on suurendatud korrutades standardhälbeid eksperimentaalselt valitud kordajatega. F1 puhul, mis on väiksema varieeruvusega kui F2, on lubatud 2,0 kordajaga standardhälve, F2 korral standardhälve kordajaga 2,5.

## 4 Rakenduse valideerimine

Järgnevas peatükis kirjeldatakse rakenduse testimist ja tagasiside kogumise meetodikat rakenduse valideerimiseks. Rakenduse testimiseks värvati uuringusse 6 isikut, kellele anti võimalus rakendusega iseseisvalt tutvuda ja selle põhifunktsioone 15-minutilise kohtumise jooksul isiklikult katsetada. Käesoleva töö autor viibis testitavatega samas ruumis, et jälgida testitavate üldist meeleolu katse jooksul, rakenduses orienteerumist ja tehnilist sujuvust. Katseperioodi lõpposas vastasid kõik testgruppi kuulunud isikud küsimustele, mis puudutasid äsja omandatud esimest kogemust rakenduse ja selle erinevate aspektidega. Küsimustik koosnes 6 küsimusest, kus oli vaja hinnata:

- Mängujuhiste arusaadavust;
- Rakenduse vokaalituvaatuse toimivust;
- Tulemuste tagasiside mõistmist ja võimet sellest õppida;
- Mängu graafilist kujundust;
- Rakenduse efektiivsust häälde treenimisel;
- Huvi rakendust kasutada.

Tagasisidestamiseks kasutati skaalapõhiseid vastuseid, kus testgrupi liikmed hindasid vahetult pärast rakendusega tutvumist erinevaid rakendust puudutavaid asjaolusid. Skaalal esinesid väärtused 1-5, kus hinne '1' - väga halb; '2' - halb; '3' - keskmine; '4' - hea; '5' - väga hea.

Küsitluse viimase osana avanes igal testgrupi liikmel võimalus vabas vormis rakendust täiendavalt kommenteerida ja esitada ettepanekuid selle kasutajasõbralikkuse parendamiseks.

### 4.1 Valideerimistulemuste analüüs

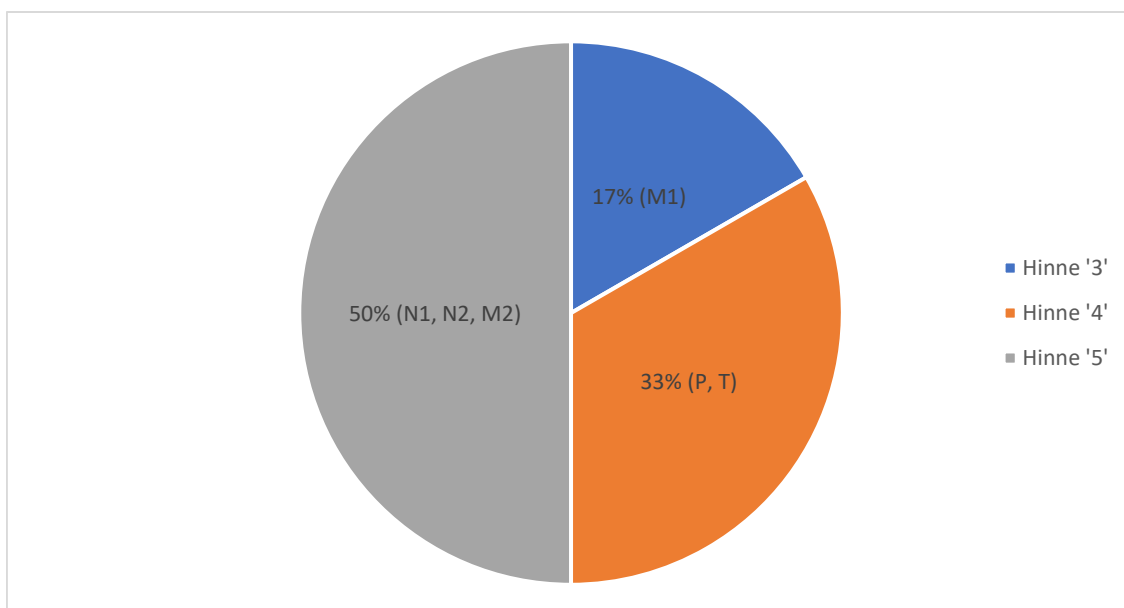
Rakenduse testimises osales kokku 6 inimest, kelle seas oli võrdselt nii mees- kui ka naissoost eesti keelt emakeelena rääkivaid isikuid (vt Tabel 1). Testgrupi kõige nooremad liikmed olid 13-aastased ja kõige vanem liige 32-aastane, teised liikmed olid vanuses 24, 26 ja 29 eluaastat. Testgrupi liikmete soolise jaotuvuse ja nende vanuste suhe oli eri

sugupoolte vahel ühtlane, sest osalenute seas oli 2 täiskasvanud meest (M1 ja M2), 2 täiskasvanud naist (N1 ja N2) ja 2 last, kellest üks oli tüdruk (T) ja teine poiss (P). Osalenud meessoost inimeste keskmine vanus oli 24,6 eluaastat ning naistel vastavalt 21 eluaastat.

Tabel 1. Testgrupi liikmete sooline ja vanuseline jaotuvus.

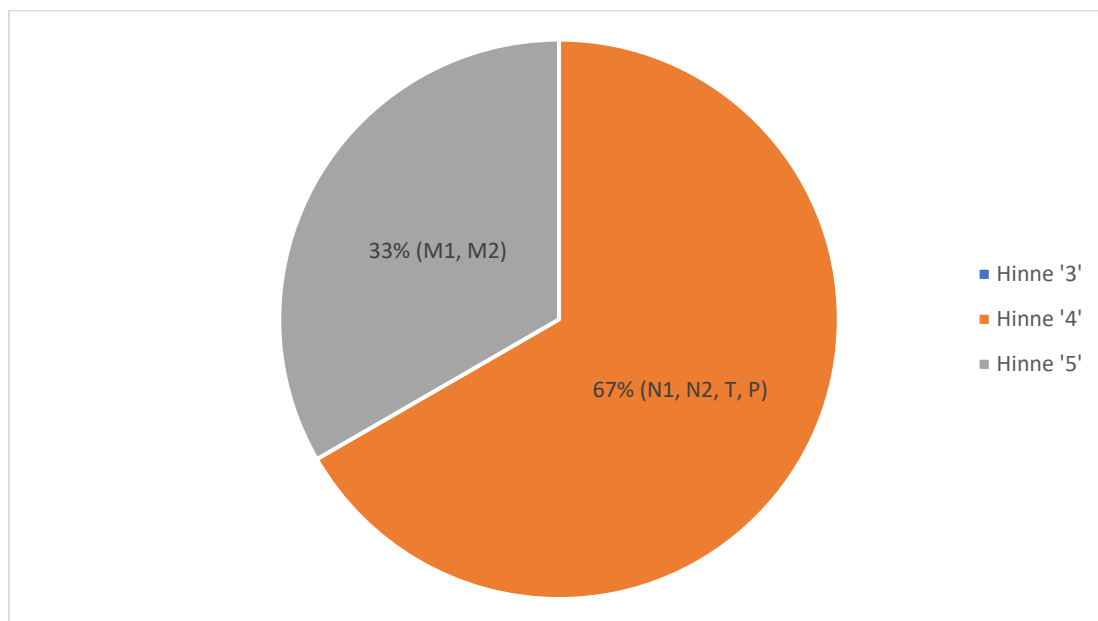
Testisik	N1	N2	M1	M2	T	P
Sugu	Naine	Naine	Mees	Mees	Naine	Mees
Vanus (aastat)	24	26	32	29	13	13

Esimesena hindasid testgrupi liikmed mängujuhiste arusaadavust (vt Joonis 15). 3 inimest andsid mängujuhiste hindeks '5', 2 inimest vastavalt '4' ja 1 inimene '3'. Suurem osa täiskasvanutest hindas mängujuhiseid skaala kõrgeima hindega, laste jaoks olid mängujuhised väärt ühe punkti võrra madalamat hinnet. Kõige madalamalt hindas mängule koostatud juhiseid testgrupi kõige vanem liige. Keskmiselt hindas testgrupp mängujuhiste arusaadavust hindega '4,3'.



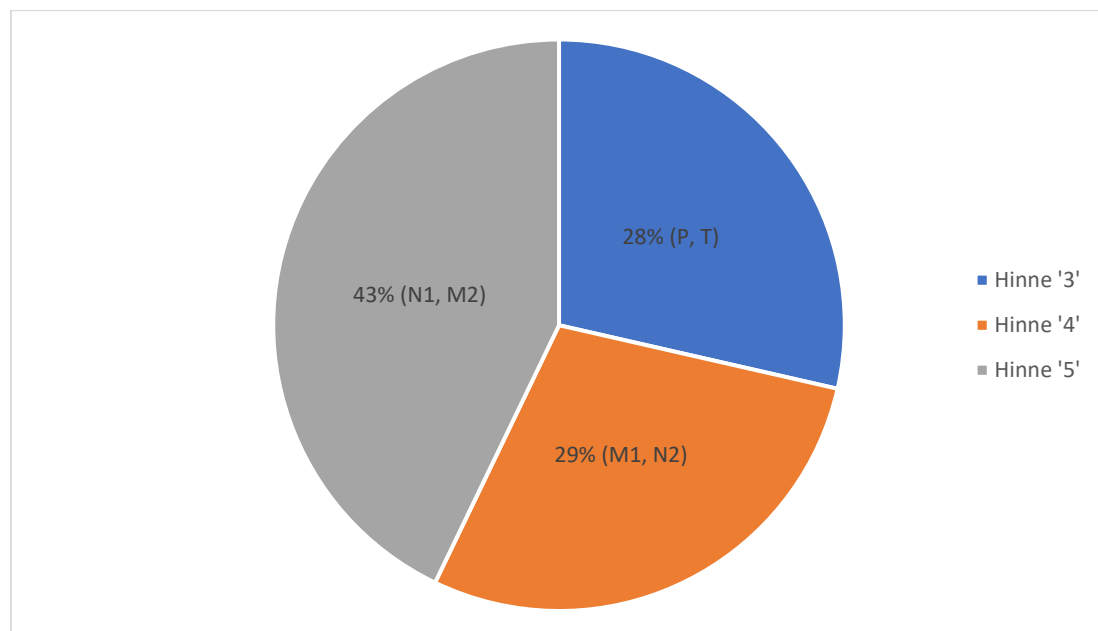
Joonis 15. Mängujuhiste arusaadavus.

Rakenduse vokaalituvastuse toimivusele andis 4 inimest hindeks '4' ja 2 inimest hindeks '5' (vt Joonis 16). Neist kahest madalama hinde olid andnud kõik osalenud naissoost isikud ja testgrupi ainus poisslaps. 2 meessoost isiku hinnangul väärts vokaalituvastus kõrgeimat hinnet. Rakenduse vokaalituvastus sai testgrupilt keskmiseks hindeks '4,3'.



Joonis 16. Rakenduse vokaalivastuse toimivus.

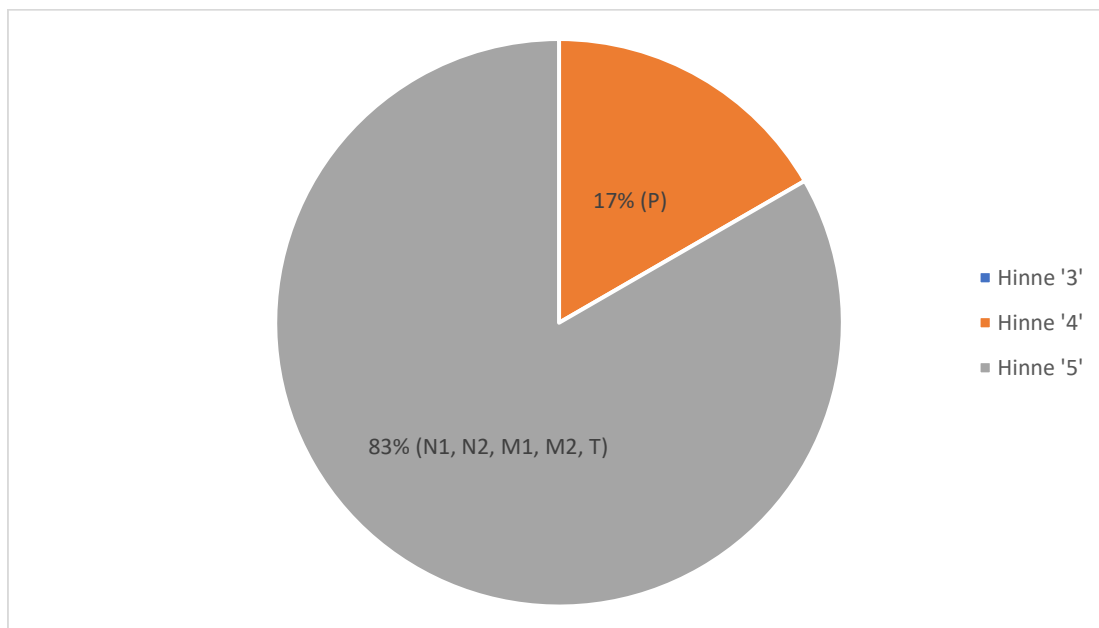
Mänguvooru lõppedes ilmuva tagasiside kujutusviisile ja selle mõistmisele andsid testgrupi liikmed kõige suurema variatsiooniga hindeid (vt Joonis 17). Kui täiskasvanud mehed ja naised hindasid tagasiside leheküljel kuvatut võrdsel määral hinnetega '4' ja '5', siis lapsed andsid samale vaatele ja oma isiklikule analüüsivõimele hindeks '3'. Tagasiside mõistmine ja sellest õppimine oli testgrupi hinnangul keskmiselt väärt hinnet '4'.



Joonis 17. Tulemuste tagasiside mõistmine ja võime sellest õppida.

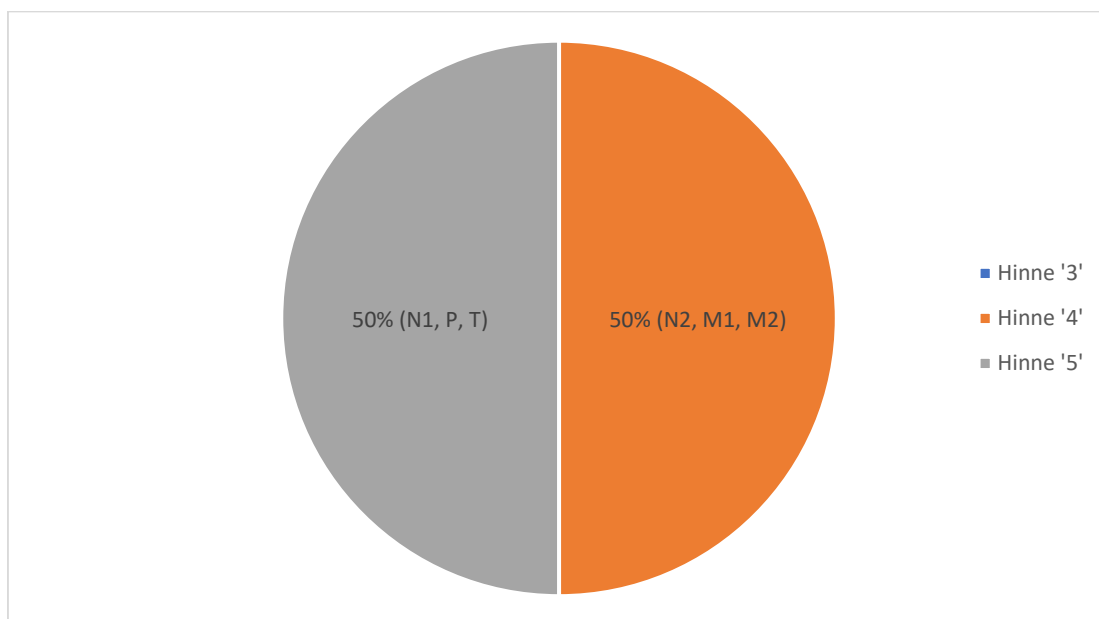
Rakenduse graafiline kujundus pälvis testgrupi liikmetelt kõige kõrgemad hinned (vt Joonis 18). Kõik täiskasvanud ja testgrupi ainus alaealine tüdruk hindasid mängu

graafilisi elemente hindega '5'. Testgrupi ainus poisslaps andis oma hindeks '4'. Keskmiselt hindas testgrupp rakenduse graafilist kujundust hindega '4,8'.



Joonis 18. Mängu graafiline kujundus.

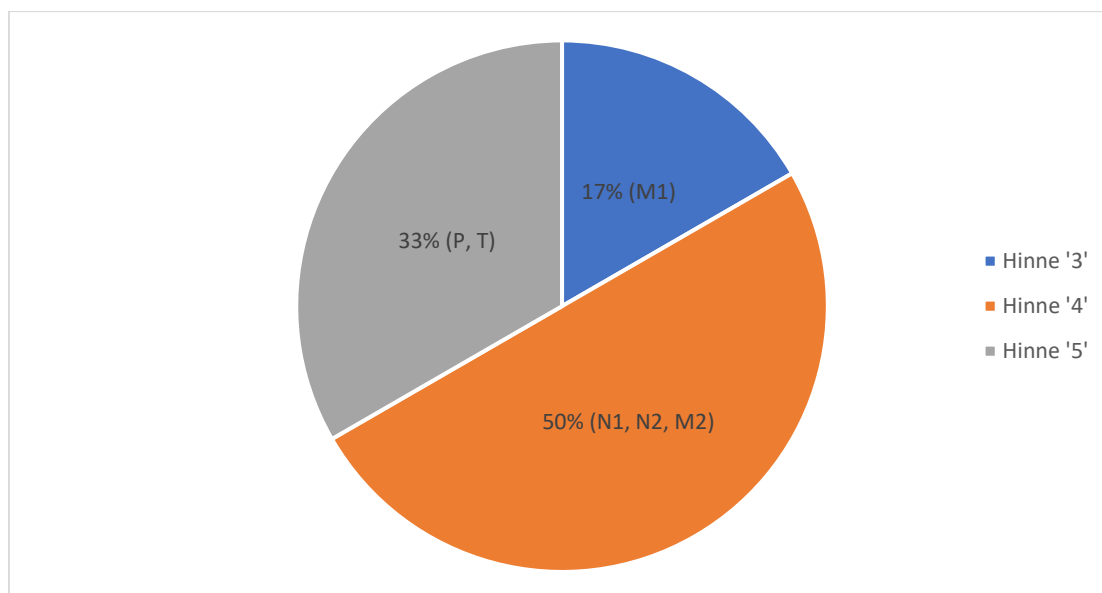
Testgrupi liikmetel paluti oma kogemuse põhjal hinnata ka rakenduse potentsiaali vokaalide häälduse treenimiseks (vt Joonis 19). Testgrupi liikmete hoiak oli positiivne ja tulemused järgmised: 3 inimest hindas rakenduse efektiivsust hindega '5' ja 3 inimest hindega '4', kusjuures lapsed olid andnud neist kahest kõrgeima hinde. Keskmiseks hindeks sai rakenduse efektiivsus '4,5'.



Joonis 19. Rakenduse efektiivsus häälduse treenimisel.



Huvi rakenduse kasutamise vastu hindasid täiskasvanud peamiselt hindegaga '4' ja lapsed hindegaga '5' (vt Joonis 20). Ühe täiskasvanu jaoks köitis mäng teda hinde '3' vääriliselt, kuid ükski testimises osalenud isik ei katkestanud mängimist huvi kaotamise tõttu. Testgrupi hinded kujundasid mänguhuvi tekke keskmiseks hindeks '4,2'.



Joonis 20. Huvi rakendust kasutada.

Viimaseks anti igale testgrupi liikmele võimalus mängu vabas vormis kommenteerida ja esitada ettepanekuid millegi muutmiseks vastavalt oma eelistustele. Testgrupi liikmete kommentaarid olid peamiselt seotud kasutajaliidese parendamisega, aga esines ka muid soovitusi. Ühe osalenud isiku hinnangul võiks mängujuhiste mõistmist lihtsustada juhiste ilmumise ja ekraanile mängu tegevusega jooksvalt, kui kasutaja esimest korda mänguvaatega tutvub. Teine osaleja oli välja toonud, et mängujuhiste lugemisel võiks olla võimalus kiiresti menüüvaatesse tagasi pääseda selle asemel, et sinna tulnud teed mööda tagasi minna. Veel öeldi, et mängujuhistes kasutatud anatoomilised animatsioonid kõnetrakti ja keele liikumise kohta võiksid olla ühesemalt arusaadavamad. Lisaks sellele võiks olla võimalus mänguvooru tulemuste tagasiside lehel mängujuhiseid uuesti näha, et kasutaja saaks vajadusel meelde tuletada endale, mida tähendab termin „formant“ ning kuidas on F1 ja F2 seotud keele asendiga. Testimises osalenud lapsed avaldasid soovi mängu peategelast metsas liigutada või tuua muid variatsioone mängu tegevusse, et mängu üldmulje oleks dünaamilisem. Paljud täiskasvanud mainisid peale rakenduse kasutamist, et ilmselt oleks kasulik testida rakendust selle otsese sihtgrupi peal. Testgrupi üldine tagasiside rakenduses kogetule oli positiivne, kõige enam kiideti mängu üldist ideed ja graafilisi lahendusi.

## 4.2 Järeldused

Mängujuhiste arusaadavuse hindamisel selgus, et laste jaoks on juhistest aru saamine keerulisem kui täiskasvanute jaoks ehkki kõik testgrupi liikmed said juhistest siiski piisavalt hästi aru, et mängimist alustada ja mängu põhifunktsioone rakendada.

Vokaalituvastuse hindamisel ilmnas, et tuvastusmeetod toimus veidi paremini meessoost isikute katsetustel, kuid olulisi soolisi erinevusi ei esinenud. Kuigi testgrupi liikmete emakeeleks oli eesti keel, ei võitnud kõik testisikud mängu. See viitab võimalikule probleemile formantide tuvastamisel, mida tuleks põhjalikumalt uurida lindistades mängu vältel testisiku hääldusi ja analüüsides lindistusi Praat foneetikaprogrammiga [11]. Praati tuletatud formantide võrdlus rakenduse formantanalüüsi tulemustega aitaks hinnata rakenduse vokaalituvastust detailsemalt. Võidutingimust, milleks on kõigi 18 vokaali hääldamine tuvastataval tasemel, tuleks leevendada lubades mõne vokaali ebaõnnestumist, et rakenduse kasutamine mõjuks korrektse häälduse õpingutel julgustavamalt.

Olulisemalt paistis silma testgrupi hinnang kolmandale küsimusele, kus testimises osalenud lapsed hindasid tagasiside lehekülje arusaadavust ja nendest tulemustest õppimise võimet skaala keskmise hindega '3'. Sellest järeldub, et tulemuste tagasiside esitamise viis võiks olla selgem ja lihtsam. Kuna rakenduse sihtgruppi esindav kasutaja on tõenäoliselt rakenduseni jõudnud logopeedi või võõrkeeleõpetaja soovitusel, on ta ka juba teadlik vokaali kvaliteedi ja keele asendi vahelisest seosest. Selle tõttu võib eeldada, et rakenduse sihtgrupi jaoks võib tagasiside formaadi mõistmine tõenäoliselt olla lihtsam kui testgrupi jaoks.

Rakenduse efektiivsust hinnates olid lapsed kindlad, et rakenduse järjekindel kasutamine võib inimese vokaalide hääldust positiivselt mõjutada ning lapsed olid ka kõige rohkem huvitatud mängu edaspidisest kasutamisest. Sellest järeldub, et mängu üldine kontseptsioon on laste jaoks piisavalt atraktiivne ja väljakutsuv, et neis mänguhuvi tekitada. Täiskasvanute veidi vähesem huvi mängu vastu tulenes tõenäoliselt asjaolust, et nad ei kuulunud rakenduse otsesesse sihtgruppi, mille tõttu oli mängu üldine tase nende jaoks veidi liiga väljakutsetevaene ning mängu lihtsast kontseptsioonist ainuüksi ei piisanud, et maksimaalset huvi tekitada. Järgmist testgruppi koostades oleks seega oluline silmas pidada nende suuremat vastavust rakenduse sihtgrupiga.

## 5 Kokkuvõte

Käesoleva töö eesmärgiks oli luua Android nutitelefonile mängulises formaadis rakendus eesti keele vokaalide häälduse treenimiseks. Häälduse kvaliteedi hindamisel lähtuti kasutaja kõne kahest esimesest formantsagedusest, mis on otseses seoses keele asendiga.

Töö tulemusena valmis esialgne versioon mängust, mis võimaldab kasutajal vokaalide hääldust treenida ja saada ka tagasisidet iga hääldatud vokaali sooritusele. Referentsväärtused, millega hääldustulemusi võrreldakse, on kasutaja soo- ja vanusepõhised. Korrektseks peetakse hääldust, mille formantsagedused mahuvad eesti vokaalidele omaste referentsvahemike sisse. Kui vokaali esimene formant ei vasta referentsvahemikule, antakse kasutajale soovitus muuta hääldamisel keele kõrgust vastavalt madalamaks või kõrgemaks. Teise formantväärtuse mittedobivusel on nõuanne korrigeerida keele asendit ees- või tagapoolsemaks. Vajadusel soovitab rakendus muuta keele asendit mõlema muutuja suhtes.

Loodud rakenduse valideerimiseks viidi läbi katse 6 isiku peal, kellele anti võimalus iseseisvalt rakendusega 15 minutit tutvuda. Katseperioodi järel vastasid testgrupi liikmed küsimustikule, millega anti hinnang rakenduse vokaalitivastuse oskusele, mängujuhiste ja tagasiside formaadi arusaadavusele, mängu graafilisele kujundusele, efektiivsusele hääldustreeningul ning huvile rakendust ise kasutada. Tulemustest selgus, et kõige tugevamaks rakenduse aspektiks peeti selle graafilist kujundust ning kõige nõrgemaks mängu tulemuste tagasiside mõistmist. Vokaalitivastus toimus mõlema sugupoole esindajate häälduse puhul, kuigi testgrupi meessoost isikud hindasid vokaalitivastust veidi kvaliteetsemaks kui naissoost isikud.

Rakenduse edasiarenduse korral tuleks tähelepanu pöörata tagasiside formaadi lihtsustamisele, et nooremal sihtgrupil oleks kergem oma hääldustulemusi tõlgendada. Täiskasvanud mängijate tagasisidest selgus, et mäng oleks võinud olla väljakutsavam. Edasiarendusena saaks lisada valikusse ka keerulisema mänguvooru, kus võitmiseks ei piisa vaid korrektsest vokaali hääldusest, vaid oluline on tabada ka juhuslikult nõutud hääle kõrgust.

## Kasutatud kirjandus

- [1] M. Väljaots, “Eesti keele häädustreeningu rakendus,” [Magistritöö], Infotehnoloogia teaduskond, TalTech, Tallinn, Eesti, 2018. [Online]. Loetud aadressil: <http://digikogu.taltech.ee/et/item/6e11a437-4d5d-4eb7-965a-609dac7ed687> Kasutatud: 09.05.2021.
- [2] K. Leppik ja C. Tejedor-García, “Estoñol, a computer-assisted pronunciation training tool for Spanish L1 speakers to improve the pronunciation and perception of Estonian vowels,” *Eesti ja soome-ugri keeleteaduste ajakiri*, vol. 10, no. 1, pp. 89-104, 2019, doi: 10.12697/jeful.2019.10.1.05.
- [3] A. Eek, *Eesti keele foneetika I*. Tallinn: Tallinna Tehnikaülikooli Kirjastus, 2008.
- [4] E. L. Asu, P. Lippus, K. Pajusalu, P. Teras, *Eesti keele häädus*. Tartu Ülikooli Kirjastus, 2016.
- [5] E. Meister ja L. Meister, “Eesti laste kõne II. Vokaalide akustiline analüüs,” *Keel ja Kirjandus*, vol. 62, no. 4, pp. 282–295, 2019. [Online]. Loetud aadressil: <https://dea.digar.ee/cgi-bin/dea?a=d&d=AKkeeljagirjandus201904.2.8> Kasutatud: 09.05.2021.
- [6] J. Six, O. Cornelis, M. Leman, “TarsosDSP, a Real-Time Audio Processing Framework in Java,” in *Proceedings of the 53rd AES Conference (AES 53rd)*, London, 2014, pp. 1-7. Loetud aadressil: [https://0110.be/files/attachments/411/aes53\\_tarsos\\_dsp.pdf](https://0110.be/files/attachments/411/aes53_tarsos_dsp.pdf) Kasutatud: 10.05.2021.
- [7] *Digital Sound Processing and Java, Documentation for the TarsosDSP Audio Processing Library*, Institute for Psychoacoustics and Electronic Music (IPEM), Gent, Belgia, 2015. [Online]. Loetud aadressil: <https://0110.be/releases/TarsosDSP/TarsosDSP-2.3/TarsosDSP-2.3-Manual.pdf> Kasutatud: 11.05.2021.
- [8] M. Jaczyńska, P. Bobiński, A. Pietrzak, “Music recognition algorithms using queries by example,” in *Proc. Joint Conference - Acoustics 2018*, Ustka, Poola, pp. 108-111. doi: 10.1109/ACOUSTICS.2018.8502429.
- [9] A. de Cheveigné ja H. Kawahara, “YIN, a fundamental frequency estimator for speech and music,” *The Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1-14, 2002, doi: 10.1121/1.1458024.
- [10] J. Harrington, S. Cassidy, *Techniques in Speech Acoustics*. Springer, Dordrecht, 1999.
- [11] P. Boersma, D. Weenink, *Praat: doing phonetics by computer*. [Online]. Loetud aadressil: <http://www.fon.hum.uva.nl/praat/> Kasutatud: 18.05.2021.
- [12] Android Developers, *Request app permissions*. [Online]. Loetud aadressil: <https://developer.android.com/training/permissions/requesting> Kasutatud: 14.05.2021.
- [13] Android Developers, *Manifest.permission*. [Online]. Loetud aadressil: [https://developer.android.com/reference/android/Manifest.permission#RECORD\\_AUDIO](https://developer.android.com/reference/android/Manifest.permission#RECORD_AUDIO) Kasutatud: 14.05.2021.
- [14] Android Developers, *AudioRecord*. [Online]. Loetud aadressil: [https://developer.android.com/reference/android/media/AudioRecord.html#AudioRecord\(int,%20int,%20int,%20int\)](https://developer.android.com/reference/android/media/AudioRecord.html#AudioRecord(int,%20int,%20int,%20int)) Kasutatud: 14.05.2021.

- [15] I. R. Titze, *Principles of Voice Production*. Prentice Hall, 1994.
- [16] E. Meister ja L. Meister, "Eesti laste kõne I. Põhitooni akustiline analüüs," *Keel ja Kirjandus*, vol. 60, no. 7, pp. 518–533, 2017. [Online]. Loetud aadressil: <https://dea.digar.ee/cgi-bin/dea?a=d&d=AKkeeljakirjandus201707.2.6> Kasutatud: 23.05.2021.
- [17] P. Ladefoged, *Elements of Acoustic Phonetics*, 2<sup>nd</sup> ed. University of Chicago Press, 2017. [E-book]. Loetud aadressil: <https://books.google.ee/books?id=4iCMDwAAQBAJ&dq> Kasutatud: 14.05.2021.

## **Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks<sup>1</sup>**

Mina, Maria Letta

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Android nutitelefoni mäng eesti keele vokaalide hääldustreeninguks“, mille juhendaja on Einar Meister
  - 1.1. reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
  - 1.2. üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. Olen teadlik, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autorile.
3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

16.05.2021

---

<sup>1</sup> Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil. Kui lõputöö on loonud kaks või enam isikut oma ühise loomingu tegevusega ning lõputöö kaas- või ühisautor(id) ei ole andnud lõputööd kaitsvale üliõpilasele kindlaksmääratud tähtajaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. ja 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.

## Lisa 2 – Mängujuhendi leheküljed

