

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Jaanus Tärnpuu 204058IAPM

EESTI KEELE HÄÄLDUSTREENINGU RAKENDUS

Magistritöö

Juhendaja: Einar Meister
PhD

Tallinn 2025

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: Jaanus Tärnpuu

12.05.2025

Annotatsioon

Magistritöö raames töötati välja kaks iOS-platvormil toimivat rakendust – andmekogumisrakendus ja hääldustreenimiserakendus – koos neid toetavate taustateenustega. Rakenduste arendamisel on kasutatud kaasaegseid tehnoloogiaid nagu Swift, Python, Java, Spring Boot ja PostgreSQL. Lahendus tugineb mikroteenuste arhitektuurile, mis võimaldab paindlikku arendust ja süsteemi skaleeritavust.

Andmekogumisrakenduse abil koguti 1463 unikaalset helisalvestist, mida analüüsi ning mille põhjal koostati vokaalikaart hääldustreenimiserakenduse tarbeks.

Hääldustreenimiserakendus tutvustab kasutajale vokaalide hääldamise juhiseid ning pakub hääldustreenimise käigus kahte tüüpi visuaalset ja tekstilist tagasisidet: referentsväärtustel ning normaliseeritud väärtustel põhinevat.

Hääldustreenimiserakenduse töökindluse ja kasutajamugavuse paremaks mõistmiseks viidi läbi tagasiside küsitlus, kus osales 19 inimest. Tagasiside oli enamasti positiivne, küll aga toodi välja mõningad puudused, millega peaks arvestama rakenduse järgmistes versioonides.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 37 leheküljel, 6 peatükki, 21 joonist, 3 tabelit.

Abstract

Estonian Language Pronunciation Tutoring Application

As part of the master's thesis, two iOS-based applications were developed – a data collection app and a pronunciation training app – along with supporting backend services. The development of these applications utilized modern technologies such as Swift, Python, Java, Spring Boot, and PostgreSQL. The solution is based on a microservices architecture, which enables flexible development and system scalability.

Using the data collection app, a total of 1,463 unique audio recordings were gathered, analyzed, and used to construct a vowel chart for the pronunciation training application.

The pronunciation training application provides users with Estonian vowel pronunciation instructions and offers two types of visual and textual feedback during training: feedback based on reference values and feedback based on normalized values.

To better understand the reliability and usability of the pronunciation training app, a user feedback survey was conducted with 19 participants. The feedback was mostly positive, though several issues were identified that should be addressed in future versions of the application.

The thesis is written in Estonian and consists of 37 pages of text, 6 chapters, 21 figures, and 3 tables.

Lühendite ja mõistete sõnastik

L1	Emakeel
L2	Keel, mida õpitakse pärast emakeele omandamist
CALL	Arvutipõhine keeleõpe (inglise keeles - <i>Computer Assisted Language Learning</i>)
Formant	Suuõõne resonantssagedus
CAPT	Arvutipõhine hääldusõpe (inglise keeles - <i>Computer-Assisted Pronunciation Training</i>)
VAD	Kõnetegevuse detektor (inglise keeles - <i>Voice Activity Detector</i>)

Sisukord

1	Sissejuhatus	9
1.1	Taust	9
1.2	Probleem	10
1.3	Olemasolevad lahendused	11
1.3.1	SayEst	11
1.3.2	Vokaalimäng	11
1.3.3	Veebirakendus eesti keele häälduse treeninguks	12
1.3.4	Olemasolevate lahenduste analüüs	12
1.4	Töö eesmärk	12
2	Metoodika	14
2.1	Ülevaade planeeritavast rakendusest	14
2.1.1	Rakenduste arhitektuur	14
2.2	Tööriistad ja tehnoloogiad	15
2.2.1	PostgreSQL	16
2.2.2	Swift	17
2.2.3	Java	17
2.2.4	Praat	18
2.2.5	Python	18
2.3	Algoritmide arendus	19
3	Arendus	21
3.1	Algoritmide väljatöötamine	21
3.1.1	Vokaalide formantide leidmine	21
3.1.2	Tagasisidestamine	25
3.2	Rakendused	27
3.2.1	Vocal-recorder taustteenus	27
3.2.2	Vocal-processor taustteenus	28
3.2.3	Vocal-feedback taustteenus	28
3.2.4	Andmekogumisrakendus	29
3.2.5	Hääldustreenimisrakendus	29
3.2.6	Rakenduste evitamine	30
3.2.7	Rakenduste arhitektuuri ülevaade	30
4	Tulemused	32

4.1	Andmekogumirakendus	32
4.1.1	Vaade 1: Nõusolek osalemiseks ja privaatsuspoliitika tutvustamine	32
4.1.2	Vaade 2: Küsimustik	33
4.1.3	Vaade 3: Sissejuhatus	34
4.1.4	Vaade 4: Salvestamine	35
4.2	Häädustreenimisirakendus	36
4.2.1	Vaade 1: Nõusolek osalemiseks ja privaatsuspoliitika tutvustamine	36
4.2.2	Vaade 2: Küsimustik	37
4.2.3	Vaade 3: Ülevaade tagasiside meetoditest ja häädusõpetus	38
4.2.4	Vaade 4: Salvestamine	39
5	Tulemuste analüüs ja edasised arendused	41
5.1	Analüüs	41
5.1.1	Andmekogumirakendus	41
5.1.2	Häädustreenimisirakendus	41
5.2	Edasised arendused	45
6	Kokkuvõte	46
	Kasutatud kirjandus	47
	Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks	51
	Lisa 2 – Vokaalide referentsväärtused	52
	Lisa 3 – Rakenduse GitLab lingid	53

Jooniste loetelu

1	Ülevaade mikroteenuste arhitektuurist	15
2	Eesti keele vokaalide paiknemine F1 ja F2 teljestikus	20
3	Vokaalide formantide leidmise algoritmi 1 pseudokood	22
4	Vokaalide formantide leidmise algoritmi 2 pseudokood	23
5	Näitesalvestis sõnale "saak" avatuna Praat rakenduses	24
6	Rakenduste üldine arhitektuur, mis näitab rakenduste seoseid	31
7	Andmekogumisrakenduse vaade: Nõusolek osalemiseks ja privaatsuspoliitika tutvustamine	33
8	Andmekogumisrakenduse vaade: Küsimustik	34
9	Andmekogumisrakenduse vaade: Sissejuhatus	35
10	Andmekogumisrakenduse vaade: Salvestamise olekud 1, 2 ja 3 (vasakult paremale)	36
11	Hääldestreenimiskogumisrakenduse vaade: Nõusolek osalemiseks ja privaatsuspoliitika tutvustamine	37
12	Hääldestreenimiskogumisrakenduse vaade: Küsimustik	38
13	Hääldestreenimiskogumisrakenduse vaade: Ülevaade tagasiside meetoditest ja hääldestusjuhised	39
14	Hääldestreenimiskogumisrakenduse vaade: Salvestamine - kõik võimalikud olekud (ülemises reas olekud 1,2 ja 3, alumises 4, 5 ja 6)	40
15	Tagasiside andmisel osalenud kasutajate vanuseline jaotus	41
16	Tagasiside andmisel osalenud kasutajate sooline jaotus	42
17	Tagasiside andmisel osalenud kasutajate emakeel	42
18	Tagasiside andmisel osalenud kasutajate hinnang juhendite ja meetodide kirjelduste arusaadavusele	43
19	Tagasiside andmisel osalenud kasutajate hinnang rakenduse hääldestuse hindamise täpsusele	43
20	Tagasiside andmisel osalenud kasutajate hinnang rakenduse tagasiside meetodile 1	44
21	Tagasiside andmisel osalenud kasutajate hinnang rakenduse tagasiside meetodile 2	44

Tabelite loetelu

1	PostgreSQL ja selle konkurentide võrdlus	16
2	Eesti keele vokaalide formandid ning keele ja huulte asendid	20
3	Andmekogumisrakenduse abil kogutud salvestuste põhjal arvutatud referentsväärtused kasutades Praat skripti	52

1. Sissejuhatus

1.1 Taust

Eesti riigi keerulise ajaloo ja tormilise geopoliitilise olukorra tõttu 21. sajandil, elab Eestis palju inimesi, kelle emakeel ei ole eesti keel. Viimase rahva ja eluruumide loenduse kohaselt elas Eestis kokku 1 331 824 inimest, kelle seast emakeelena rääkis eesti keelt 895 493 [1]. Võrreldes kahe viimase aasta rahvastiku uuringutega on kasvanud Eestis mitte-eestlaste arv 237 497 pealt 247 375 -ni [2]. Samas on riigikeele oskusel oluline roll integratsioonipoliitikas.

Põhimõtteliselt on iga inimene võimeline lisaks emakeelele omandama mitmeid muid keeli, kuid kui võõrkeele õppimisega alustatakse alles täiskasvanuna, kaasneb sellega reeglina võõrkeele aktsent – keelele omasest tüüpilisest hääldusest hälbiv kõne. Selle peamiseks põhjuseks on imikueas omandatud emakeelsed hääldus- ja tajumallid [3, lk 12].

Lennebergi kriitilise perioodi hüpotees ütleb, et parim aeg keele õppimiseks on imikueast kuni puberteedieani. Lennebergi avastuste kohaselt on kriitiline periood seotud bioloogiliste ja neuroloogiliste muudatustega ajus. Varajases lapsepõlves on aju rohkem plastiline ning kohanemisvõimelisem, mis teeb uue keele õppimise ja omandamise lihtsamaks [4].

Traditsiooniliste keeleõppe meetodite kõrval kasutatakse üha enam arvutipõhist keeleõpet CALL. CALL-süsteemid pakuvad uudseid õppe- ja õpetamismeetodeid, mis motiveerivad õppijaid ning hõlbustavad keeleõpetajate tööd. Nii õppijad kui õpetajad on hinnanud neid vahendeid valdavalt kasulikuks [5, 6, 7].

Keele omandamise protsess jaguneb üldjoontes neljaks põhivaldkonnaks: lugemine, kirjutamine, kuulamine ja hääldamine. Esimest kolme on võimalik tõhusalt arendada iseseisva töö kaudu, kasutades veebipõhiseid õppematerjale ja -keskkondi. Häälduse kujundamine seevastu eeldab reeglina interaktiivset tagasisidet – selleks on vajalik keeleoskaja või sobiva tehnoloogilise lahenduse olemasolu, mis võimaldaks hinnata ning korrigeerida õppija hääldust vastavuses sihtkeele foneetiliste normidega.

1.2 Probleem

Iseseisev keeleõpe on viimastel aastatel muutunud üha populaarsemaks, mida toetavad nii paindlikud kaugõppevormid kui ka tehnoloogia areng. Selle tulemusel on loodud mitmeid keeleõppekeskkondi ja -rakendusi, mis aitavad õppijatel omandada uusi keeli oma tempos ja kohas. Joan Palmiter Bajorek analüüsis oma artiklis "L2 Pronunciation Tools: The Unrealized Potential of Prominent Computer-assisted Language Learning Software" mitmeid tuntumaid CALL platvorme ning jõudis järeldusele, et nende kõnetuvastuse ja häälduse tagasiside võimalused ei vastanud sageli ootustele. Uuringus käsitleti selliseid platvorme nagu Rosetta Stone, Duolingo, Babbel ja Mango Languages, keskendudes eelkõige häälduse õpetamisele ja õppija tagasiside kvaliteedile [8].

Kuigi keeleõpperakendusi ja -keskkondi on loodud arvukalt, keskendub enamik neist suure kasutajaskonnaga keelele nagu inglise, hispaania, saksa või prantsuse keel. Vähemlevinud keeled, sealhulgas eesti keel, on sageli esindatud piiratud kujul või puuduvad üldse.

Eesti keelele suunatud veebikeskkonnad ja rakendused on sageli piiratud nii funktsionaalsuse kui ka kättesaadavuse poolest. Eriti puudulikud on häälduse õpetamise harjutused, kuna see eeldab tehniliselt keerukamaid lahendusi. Näiteks Simply Learn Estonian [9] pakub õppimiseks erinevaid eestikeelseid fraase, mida saab kuulata, kuid puudub tagasiside-funktsionaalsus.

Arvutipõhine hääldusõpe on osutunud eriti tõhusaks raskesti eristatavate foneemikontrastide, näiteks vokaalide, tajumise ja hääldamise arendamisel. CAPT-süsteemide üks oluline funktsionaalsus on õppija hääldusele kohese ja arusaadava tagasiside andmine, mis võib olla: 1) visuaalne – õppija hääldatud vokaali asukohta visualiseerimine sihtkeele vokaaliruumis, 2) auditivne – õppija ja emakeelse kõneleja häälduse võrdlus, 3) numbriline – akustilisel sarnasusel põhinev skoor [10, 11, 12].

Hääldustreeningu meetodika sisaldab tavaliselt kuulamis- ja hääldusharjutusi. Kuulamisharjutustes kasutatakse identifitseerimis- ja diskrimineerimisülesandeid, mille abil arendatakse võõrkeele häälikute taju. Hääldusharjutustes peab õppija kordama ja jäljendama etteantud häälikut või sõna, misjärel visualiseeritakse nt vokaali asukoht vokaaliruumis ja antakse juhiseid, mis aitavad suunata õppija hääldust sihtkeelele vastavaks.

Efektiivse hääldusõppevahendi loomine on mitmeastmeline protsess: esmalt tuleb tuvastada õppija hääldus, seejärel seda akustiliselt analüüsida ning lõpuks esitada tagasiside õppijale arusaadava kujul. Selle protsessi muudab keerukaks asjaolu, et iga inimese häääl on unikaalne ning selle akustilisi omadusi mõjutavad mitmed tegurid, näiteks vanus, sugu

ja emakeele (L1) taust. Seetõttu on õppija-spetsiifilise tagasiside pakkumine tehniliselt keeruline ja sageli seda välditakse.

1.3 Olemasolevad lahendused

Ka varasemalt on üritatud eesti keele hääldustreeningu rakendusi luua. Alltoodud lõigus on mõningad näited välja töötatud lahendustest, ülevaade nende funktsionaalsusest ning ka lühike tehniline kirjeldus.

1.3.1 SayEst

SayEst on Tartu Ülikooli foneetika laboris 2022. aastal arendatud Android mobiilirakendus, mis on mõeldud eesti keele täishäälikute ja kaashäälikute õppimiseks ning nende häälduste harjutamiseks [13]. See rakendus pakub nelja liiki harjutusi:

- **Kuula ja korda (*exposure*)** – õppija kuulab eesti emakeelega kõneleja hääldatud sõna ja kordab seda, saades võimaluse võrrelda enda ja etteantud hääldust.
- **Tunne ära (*discrimination*)** – õppija kuuleb üht sõna ja valib kahe ekraanil kuvatud sõna vahel selle, mida kuulis; treenib kuulmiseristust.
- **Häälda (*pronunciation*)** – õppija hääldab etteantud sõna ning automaatne kõnetuvastus hindab selgust ja täpsust.
- **Kinnistamine (*mixed mode*)** – harjutused eelnevatest tüüpidest esinevad juhuslikus järjestuses, et kinnistada omandatud oskusi.

Sayest kasutab häälduse õigsuse hindamiseks Tallinna Tehnikaülikoolis loodud automaatset kõnetuvastust, mis põhineb *Kaldi toolkitil* [14].

1.3.2 Vokaalimäng

Vokaalimäng (*Vocal game*) on Maria Letta poolt arendatud Android-rakendus, mille peamine eesmärk on treenida eesti keele häälikute hääldust [15]. Mängus on mitu taset ning keeleõppija peab hääldama talle etteantud häälikuid ning kui hääldus on korrektne, viiakse mängija edasi järgmisele tasemele. Häälikute ebakorrektsel hääldusel korral antakse tekstilist tagasisidet, kuidas võiks hääldust parandada. Vokaalimäng kasutab häälikute leidmiseks ning formantide arvutamiseks Java teeket TarsosDSP [16] ning JSTK [17].

1.3.3 Veebirakendus eesti keele häälduse treeninguks

Veebirakendus eesti keele häälduse treeninguks on internetipõhine rakendus arendatud Erko Petersoni poolt [18]. Rakenduses on erinevad vaated õpilastele, õpetajatele kui ka rakenduse administraatorile. Keeleõpe rakenduses põhineb kahel harjutuse tüübil:

1. **Kuulamisharjutused**, kus õppija kuulab etteantud sõna ja peab valima, milline sõna talle esitati.
2. **Hääldusharjutused**, kus õppija kuulab etteantud sõna, kordab seda ning rakendus analüüsib tema hääldust ja annab tekstilist tagasisidet häälikute õigsuse kohta.

Rakendus kasutab vokaalide formantide tuvastamiseks Java teeki TarsosDSP ja JSTK. Aastal 2025 ei ole antud rakendus enam avalikult kättesaadav ning rakendusel oli mõningatel juhtudel probleeme vokaalide leidmisega heliklipist ning selle tulemusel oli ka tagasiside ebatäpne.

1.3.4 Olemasolevate lahenduste analüüs

Olemasolevad lahendused "Veebirakendus eesti keele häälduse treeninguks" ning "Vokaalimäng" kasutavad vokaalide leidmiseks ning formantide arvutamiseks TarsosDSP ning JSTK teeki. Mõlema rakenduse puhul on piirjuhtudel vokaalide tuvastamine heliklipist mitte korrektne ning seetõttu ei anta ka piisavalt korrektset tagasisidet häälduste parendamiseks. Mõlemas rakenduses antakse ainult tekstilist tagasisidet. 2025. aastal ei ole kumbki rakendus enam avalikult kättesaadav.

Rakendus SayEst kasutab vokaalide leidmiseks *Kaldi toolkitil*. Rakendus on kättesaadav Google Play Store'is ning on kasutatav ainult Android-seadmetega.

1.4 Töö eesmärk

Kuna puudub Apple'i ökosüsteemis töötav hääldustreenimisrakendus, siis on magistritöö põhiline **eesmärk välja töötada eesti keele hääldustreenimise iOS mobiilirakendus, mis annaks tagasisidet kasutaja hääldusele.**

Lahenduse välja töötamise käigus tuleb leida vastused ka uurimisküsimustele:

- Kuidas eristada kõnesignaalis erinevaid segmente (vokaale ja konsonante)?
- Kuidas hääldavad eesti vokaale L1 kõnelejad?

- Kuidas anda kasutajale visuaalset tagasisidet?
- Kuidas peaks kasutajat juhendama, et ta saaks oma hääldust parandada?

2. Metoodika

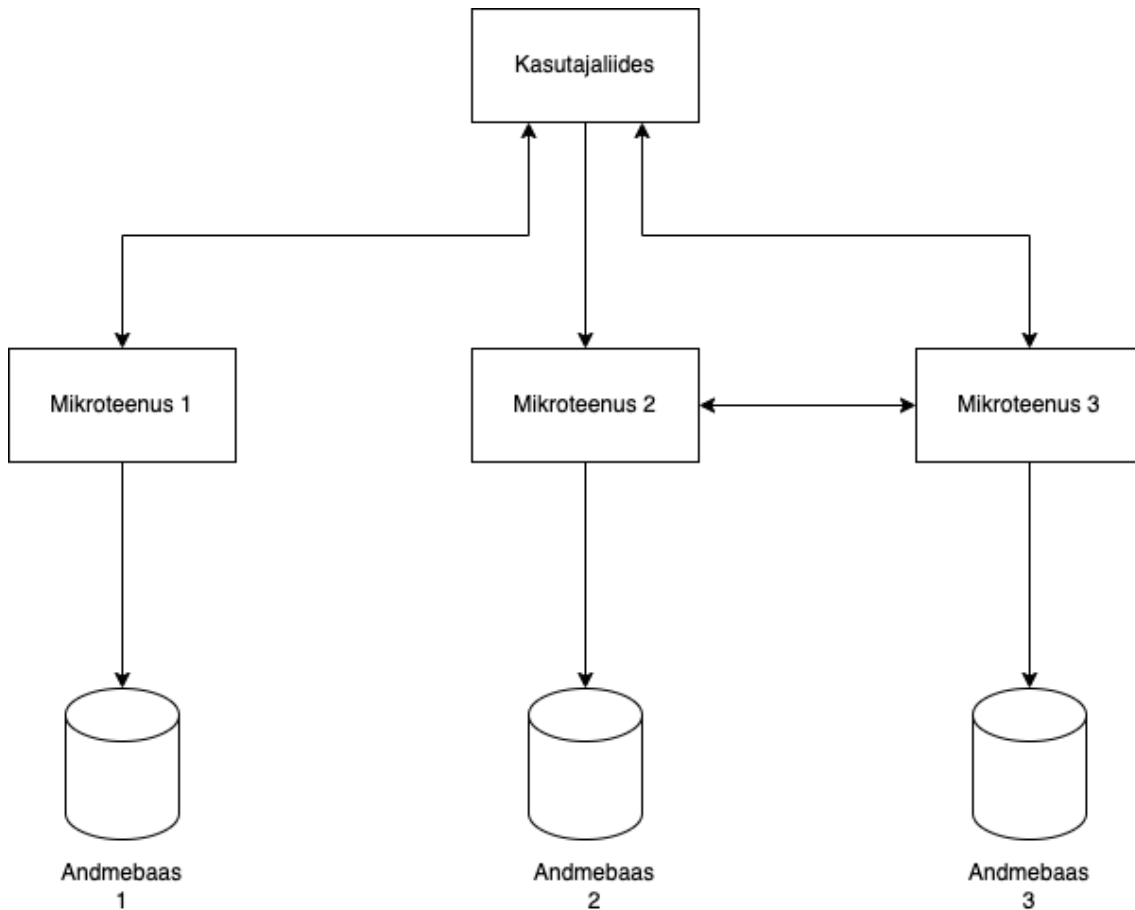
Järgnevas peatükis antakse ülevaade planeeritavast rakendusest ja arhitektuuri valikutest, loetletakse kasutatud tehnoloogiad ning kirjeldatakse vokaalide akustilisi omadusi, mis on aluseks algoritmide väljatöötamisele.

2.1 Ülevaade planeeritavast rakendusest

Antud magistritöö raames on planeeritud välja töötada kaks omavahel seotud iOS seadmetele mõeldud rakendust. Esimene rakendus on mõeldud eesti emakeelega inimeste häälde kogumiseks, et neist eraldada akustilised tunnused. Teine planeeritud rakendus on hääldestreenimise rakendus, mis on magistritöö põhiväljundiks. Teises rakenduses kasutatakse esimesest rakendusest kogutud andmeid häälde hindamiseks ning tagasiside andmiseks.

2.1.1 Rakenduste arhitektuur

Rakendused planeeritakse mikroteenuste arhitektuuri mudeli järgi. Mikroteenuste arhitektuur on kaasaegne tarkvaraarhitektuuri stiil, mis võimaldab süsteemi jagada väikesteks, iseseisvateks teenusteks. See lähenemine sobib hästi mobiilirakenduste arendamiseks, kus on vaja skaleeritavust, paindlikkust ja kiiret arendustsüklit [19]. Joonisel 1 on väljatoodud ülevaatlik pilt mikroteenuste arhitektuurist.



Joonis 1. Ülevaade mikroteenuste arhitektuurist

Kuigi rakenduste planeerimisel tuginetakse põhiliselt mikroteenuste arhitektuuri printsiipidele, siis tehakse üks mõndus andmebaasi osas - arendustegevuse lihtsustamiseks on kasutusel üks andmebaas, isoleeritus tagatakse tabelite abil.

Rakenduste arendamisel arvestatakse lisaks mikroteenuste arhitektuurile ka "õhuke klient (*thin client*)" lähenemisega ehk kogu ärilogika paigutatakse taustteenuste juurde. Selle lähenemise peamine eelis on see, et kui on vaja teha loogikas muudatusi, siis selleks ei pea tarnima uut versiooni kasutajate telefonisse, vaid saab muudatused teha märkamatuult taustteenustes. Negatiivne külg on see, et selle tulemusena ei ole võimalik mobiilirakendusi kasutada ilma internetiühenduseta ning sama kehtib ka taustteenuste töö häiringute korral.

2.2 Tööriistad ja tehnoloogiad

Rakenduste loomiseks kasutatakse mitmeid tänapäevaseid tarkvaraarendusvahendeid ja -tehnoloogiaid. Valitud tehnoloogiad lähtuvad nii töö autori kogemusest kui ka süsteemi arendamise ja hooldamise mugavusest.

Klientrakendused kirjutatakse Swift raamistikus. Taustteenused kasutavad andmebaasina PostgreSQL relatsioonilist andmebaasi. Taustrakendused saab jagada kaheks kategooriaks: 1) Java’l põhinev taustrakendus, mis kasutab Spring Boot raamistiku ning 2) Python’il põhinevad taustrakendused, mis kasutavad Flask’i. Helifailide töötlemiseks kasutatakse Praat’i Python’i teeki Parselmouth.

Rakenduste evitamiseks võetakse kasutusele konteinerlahendus Docker [20]. Rakendused serveeritakse Nginx *reverse-proxy* [21] abil avalikusele. Avalikud ühendused krüpteeritakse ning need kasutavad LetsEncrypt’i [22] poolt väljastatud sertifikaate.

2.2.1 PostgreSQL

PostgreSQL on avatud lähtekoodiga objekt-relatsiooniline andmebaasisüsteem, mis kasutab ja laiendab SQL-keelt ning sisaldab rohkelt võimalusi andmete turvaliseks talletamiseks ja suuremahuliste andmemahtude skaleerimiseks [23]. Järgnevas tabelis 1 on toodud välja PostgreSQL’i võrdlus põhiliste konkurentide Oracle DB [24], Microsoft SQL Serveri [25] ja MySQL’iga [26].

Tabel 1. PostgreSQL ja selle konkurentide võrdlus

Andmebaas	Litsents	Tugevused	Peamised kasutusvaldkonnad
PostgreSQL	Avalik	Standarditele vastav, rikkalik funktsionaalsus, laiendatav, ruumiandmete tugi	Veebi- ja ärirakendused, geoinfosüsteemid, akadeemilised projektid
Oracle DB	Kommertslitsents	Väga skaleeritav ja turvaline, sobib keerukatele süsteemidele	Suuretevõtted, finantsasutused, valitsused
Microsoft SQL Server	Kommertslitsents	Tugev Windows integratsioon	Microsofti ökosüsteemis rakendused, ärisüsteemid
MySQL	Avalik (GPL)	Kiirus ja lihtsus	Väiksed ja keskmise suurusega (veebi)rakendused

Peamisi andmebaasisüsteeme võrreldes oleks projekti oma omaduste poolest sobinud ka MySQL, kuid antud projekti raames osutus valituks PostgreSQL töö autori kogemuste tõttu antud süsteemi kasutamise ja ülesseadmise osas.

2.2.2 Swift

Swift on Apple'i loodud võimas ja intuitiivne programmeerimiskeel, mis on optimeeritud töötama iOS-i, macOS-i ning muudel Apple'i platvormidel. Swift on küll tihedalt seotud C programmeerimiskeelega, kuid on välja arendatud toimima iseseisvalt. See võimaldab sujuvat juurdepääsu olemasolevatele C ja Objective-C raamistikutele ning toetab vaikumisi mitme lõimega töötamist. Täieliku ja iseseisva keelena pakub Swift põhifunktsioone nagu vooahaldus, andmestruktuurid ja funktsioonid. Swift on avatud lähtekoodiga ning selle arengusse saavad panustada kõik huvilised. Antud projekti juures Swifti alternatiive ei kaalutud, kuna projekti alguses seati paika idee luua Apple'i rakendus ning Swift on Apple'i ametlik soovitus rakenduste loomiseks, kuna see töötab Apple'i ökosüsteemis kõige optimaalsemalt [27].

2.2.3 Java

Java on 1995. aastal SUN Microsystems organisatsiooni poolt välja arendatud programmeerimiskeel, mis on tõestanud end kui töökindel platvorm ning seetõttu on üks enimkasutatavaid keeli suurte süsteemide arendamisel [28]. Java põhilised omadused on [29]:

- **Platvormiülene:** Java programmid kompileeritakse baitkoodiks, mis töötab igas operatsioonisüsteemis, kus on olemas Java virtuaalmasin (JVM).
- **Objektorienteeritud:** Java toetab klassidel ja objektidel põhinevat lähenemist, mis soodustab korduvkasutatavat ja modulaarset koodi.
- **Turvaline:** Java-l on sisseehitatud turvamehhanismid (nt liivakastimudel, tüübitugevus), mis aitavad vältida pahatahtlikku käitumist.
- **Töökindel:** Java on loodud vigade vältimiseks juba kompileerimise ajal ning sisaldab automaatset mäluhaldust (Garbage Collection).
- **Mitme-lõime (*multithread*) tugi:** Java toetab mitme lõimega töötamist, mis võimaldab efektiivselt kasutada mitmetuumalisi protsessoreid.
- **Andmebaasi ühenduvus:** Java on sisseehitatud JDBC - *Java Database Connectivity*, mille abil saab lihtsasti ühendada kõikide populaarsemate andmebaasidega.

Java osutus antud projektis valituks autori kogemuste tõttu. Samuti on Javas implementeeritud tööriistad nagu Spring Boot ja Flyway, mis lihtsustavad oluliselt arendustegevust.

Spring boot

Spring Boot on Java-põhine arendusraamistik, mis lihtsustab rakenduste loomist, pakku- des vaikumisi seadistusi ja sisseehitatud komponente. Raamistiku eesmärk on vähendada

arendusprotsessi keerukust, eemaldades vajaduse käsitsi konfigureerida infrastruktuuri-komponente. Spring Boot võimaldab kiiresti käivitada iseseisvaid (*stand-alone*) rakendusi. Spring Boot peamised omadused on [30]:

- Võimaldab luua iseseisvaid (*stand-alone*) Springi rakendusi.
- Toetab veebiserverite nagu Tomcat, Jetty või Undertow sisseehitamist – ei ole vaja eraldi WAR-faile lisada ning konfigureerida.
- Pakub keskkonda Spring Initializer [31], mille abil saab luua kõikidest vajalikest sõltuvustest koosneva projekti läbi kasutajaliidese.
- Konfigureerib võimalusel automaatselt Springi ja kolmandate osapoolte teegid.
- Sisaldab toodangukeskkonna jaoks valmis funktsionaalsusi nagu monitooring, seisundi kontroll (*health checks*) ja väliste konfiguratsioonifailide tugi.
- Ei kasuta koodi genereerimist ning ei nõua XML-konfiguratsioone.

Flyway

Flyway on andmebaasi sisu versioneerimise tarkvara, mille abil saab luua andmebaasi tabeleid, seoseid ning käivitada teisi andmebaasi struktuuriga seotud käsklusi ja lisada vajadusel ka andmeid [32, 33]. Muudatused saab kirjutada näiteks SQL skriptina ning nende versioneerimise ajalugu hoitakse andmebaasis, et vältida sama skripti mitmekordset käivitamist. Antud tarkvara osutus valituks, et hoida arendaja arvutis ning serveris käitavat andmebaasi võrdväärsena.

2.2.4 Praat

Praat on tasuta ja avatud lähtekoodiga tarkvara, mida kasutatakse laialdaselt kõneanalüüsiks, sünteesiks ja töötlemiseks. Selle on loonud Paul Boersma ja David Weenink ning see pakub hulgaliselt funktsioone foneetiliseks uurimistöök, sealhulgas spektraalanalüüs, kõne põhitooni, intensiivsuse, formantide mõõtmine ja skriptimine [34].

2.2.5 Python

Python on 1990. aastate alguses Guido van Rossumi poolt Hollandi teadusasutuses Stichting Mathematisch Centrumis (CWI) välja arendatud, tuginedes keelele nimega ABC. Guido van Rossum oli pikka aega Python'i peaarhitekt, kuigi tänapäeval panustavad keelearendusse ka paljud teised, kuna tegemist on avatud lähtekoodiga platvormiga. Python'it levitatakse GPL-litsentsi alusel, mis võimaldab keelt vastavalt oma vajadustele modifitseerida. Python valiti välja Parselmouth-teegi tõttu, mis on Python'i teek Praat'i tarkvara jaoks [35, 36].

Parselmouth

Parselmouth on Python'i vabavaraline teek, mille abil saab käivitada Praat funktsioone otse Pythoni raamistikus. Võrreldes teiste sarnaste teekidega, käivitab Parselmouth otse Praat'i C ning C++ funktsioone, mille tulemuseks on käivitavate skriptide väljund alati sama, mis Praatis endas [37].

Flask

Flask on kergekaaluline ja paindlik avatud lähtekoodiga veebiraamistik Pythonis, mis võimaldab kiiresti luua veebirakendusi. Flask põhineb WSGI töötlusmudelil ja kasutab mallide genereerimiseks Jinja2 teeki. Raamistik pakub tuge näiteks URL-teekondadele, päringute töötlemisele, vormide haldamisele ja sessioonide käsitlemiseks [38].

2.3 Algoritmide arendus

Algoritmide arenduse saab oma loogika järgi jagada kahte gruppi. Esimene algoritm on seotud vokaalide tuvastamisega analüüsitava helilõigust. Teine algoritm on seotud tagasiside andmisega ning põhineb vokaalide hääldamise seostel keele- ning huulte asendiga.

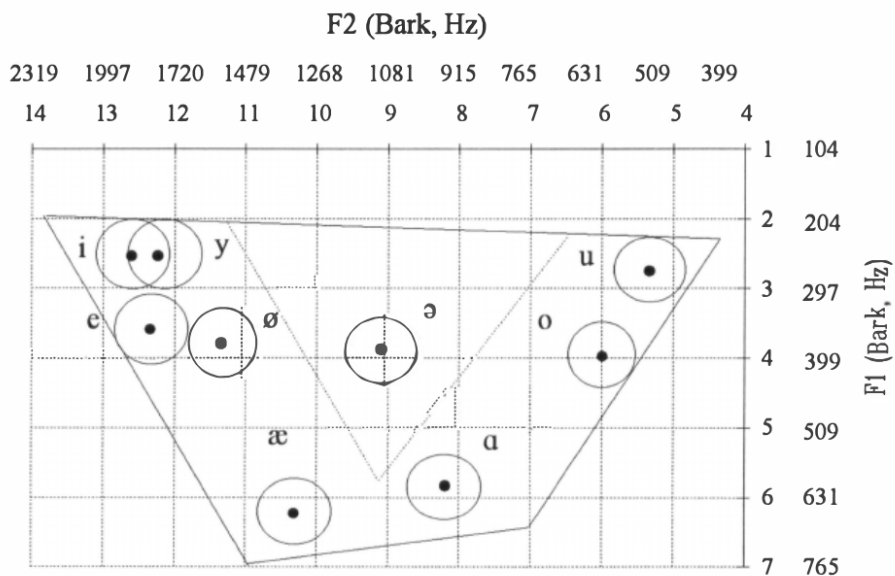
Vokaalide tuvastamine

Vokaalide tuvastamine põhineb foneetikauuringute tulemusel, et vokaalidel on oluliselt kõrgem intensiivsus kui konsonantidel ning olemas on põhitoon [39].

Tagasisidestamine

Tagasiside andmine põhineb vokaalil ning selle hääldamise seostel keele- ning huuleasendil. Vokaal on häälik, mille hääldamisel puudub suu keskel õhuvoolu takistav ahtus või sulg. Vokaali kvaliteeti kirjeldatakse kolme artikulaatorse mõõtmega: keele kõrgus, keele tagapoolsus ja huulte ümarus. Keele kõrgust kirjeldatakse kolme tunnusega: madal, keskkõrge ning kõrge. Keelekeha eest-taha liikumist kirjeldatakse kolme tunnusega: ees, kesk ning taga. Huulte ümardatust kirjeldatakse tunnusega ümar [40, lk 47].

Vokaalide artikulatsiooni akustilisteks korrelaatideks on formandid ehk kõnetrakti resonantssagedused. Need ilmnevad vokaali spektri tippudena ja neid tähistatakse F1, F2 ja F3. F1 on pöördvõrdelises suhtes keele kõrgusega, F2 korreleerub keele liikumisega ettepoole/-tahapoole, F3 on osalt seotud huulte ümardatusega ja kajastab kõnelejate individuaalseid erinevusi [41]. Vokaalide eristamiseks kasutatakse peamiselt formante F1 ja F2. Järgeval joonisel 2 on toodud eesti keele vokaalide paiknemine F1 ning F2 teljestikul [42, lk 65].



Joonis 2. Eesti keele vokaalide paiknemine F1 ja F2 teljestikus

Järgnevas tabelis 2 on välja toodud vokaalid, keeleasend ning formantide paiknemine F1 ja F2 teljestikus [40, lk 74].

Tabel 2. Eesti keele vokaalide formandid ning keele ja huulte asendid

Vokaal	F1	F2	Keele kõrgus	Keele ees-/tagaasend	Huulte asend
/i/	madal	kõrge	kõrge	ees	ümardamata
/ü/	madal	kõrge	kõrge	ees	ümardatud
/e/	keskmine	kõrge	kesk/kõrge	ees	ümardamata
/ö/	keskmine	kõrge	kesk	ees	ümardatud
/õ/	keskmine	madal	kesk	taga	ümardamata
/o/	keskmine	madal	kesk	taga	ümardatud
/u/	madal	madal	kõrge	taga	ümardatud
/a/	kõrge	madal	madal	kesk/taga	ümardamata

Tabelit analüüsid saame teha järgnevad tähelepanekud F1 ning F2 asukoha ja keele paiknemise kohta:

- **F1** korreleerub keele vertikaalse asendiga:
 - Madal F1 → keel on kõrgel.
 - Kõrge F1 → keel on madalal.
- **F2** peegeldab keele horisontaalset asendit (ees- või tagaasetus suus) ja huulte ümardatust:
 - Kõrge F2 → keel on ees.
 - Madal F2 → keel on taga ja/või huuled ümardatud.

3. Arendus

Magistritöö raames on arendatud välja kaks iOS rakendust ning neid toetavad kolm tausta-teenust. Lisaks rakendustele moodustab olulise osa tööst heliklippidest vokaalide tuvastamise ja tagasiside andmise algoritmide arendus.

3.1 Algoritmide väljatöötamine

Järgnevas lõigus on välja toodud algoritmi kirjeldus ning pseudokood nii vokaalide formantide leidmise kui ka tagasiside andmise kohta.

3.1.1 Vokaalide formantide leidmine

Varasemalt tehtud tööd eesti keele häälduste tagasiside andmise puhul on kasutusel olnud TarsosDSP ning JSTK Java teegid. Antud töö üks eesmärgi oli proovida kasutada alternatiivset lahendust. Vokaalide leidmiseks on kasutusele võetud Praat skriptid, mida käivitatakse läbi Parselmouth teegi otse Python'i koodis.

Algoritmi esimeses iteratsioonis kasutatakse teadmist, et vokaalide esinemise ajal on heli intensiivsus kõrge ning olemas on põhitoon. Vokaali algushetke määramiseks kontrollitakse iga ajahetke intensiivsuse taset ja põhitooni olemasolu ning väärtust. Algushetkeks märgitakse esimene hetk, mil intensiivsus ületab eelmääratud läviväärtuse, põhitoon on olemas ja väiksem kui 250 Hz ning esimene formant (F1) on väiksem kui 1000 Hz. Lõpphetkeks loetakse ajahetk, mil intensiivsus langeb alla läviväärtuse või põhitoon kaob. Seejärel leitakse leitud lõigus F1 ja F2 keskmised väärtused. Antud algoritmi pseudokood on esitatud joonisel 3.

```

1 Funktsioon extract_vowel_formants_v1(heli, maks_formant):
2   Resample'i heli sagedusele 10 kHz ja skaleeri intensiivsus 70 dB-ni;
3   Arvuta heli põhjal järgmised omadused;
4     Intensiivsus (helitugevus ajas);
5     Hääletoon ehk pitch;
6     Formandid (vokaalide resonantssagedused);
7   Sea intensiivsuse läviväärtus: maksimaalne intensiivsus - 10 dB;
8   vokaali_algus = tühi, vokaali_lõpp = tühi;
9   aeg = intensiivsuse algusaeg;
10  while aeg väiksem or võrdne intensiivsuse lõppajaga do
11    Leia selle ajahetke intensiivsus, hääletoon ja F1 väärtus;
12    if intensiivsus suurem kui läviväärtus and
13    hääletoon kehtiv and väiksem kui 250 Hz and
14    F1 väiksem kui 1000 Hz then
15      | vokaali_algus = aeg;
16    end
17    if vokaali_algus on määratud and
18    (hääletoon puudub or intensiivsus väiksem kui lävi) then
19      | vokaali_lõpp = aeg;
20      | katkesta tsükkel;
21    end
22    aeg = aeg + aja_samm;
23  end
24  if vokaali_algus and vokaali_lõpp on määratud then
25    | Arvuta F1 ja F2 keskmised väärtused nende aegade vahel;
26    | Tagasta F1 ja F2;
27  end
28  else
29    | Tagasta None;
30  end

```

Joonis 3. Vokaalide formantide leidmise algoritmi 1 pseudokood

Algoritmi edasiarendamiseks võeti kasutusele Praat'i enda funktsioon *Sound: To TextGrid (silences)* [43] heliliste lõikude leidmiseks – täishäälikud on helilised ning kaashäälikud mitte või on nende intensiivsus oluliselt väiksem kui vokaalidel. Antud funktsioon kasutab heliliste lõikude leidmiseks samuti intensiivsust ja selle muutumist ajas, kuid lisaks rakendab erinevaid filtreid. Kui heliline lõik on leitud, siis otsitakse antud lõigus maksimaalne intensiivsuse väärtus ning leitakse selle ümbrusest (± 20 ms) F1 ja F2 keskmised väärtused. Antud algoritmi pseudokood on leitav joonisel 4.

```

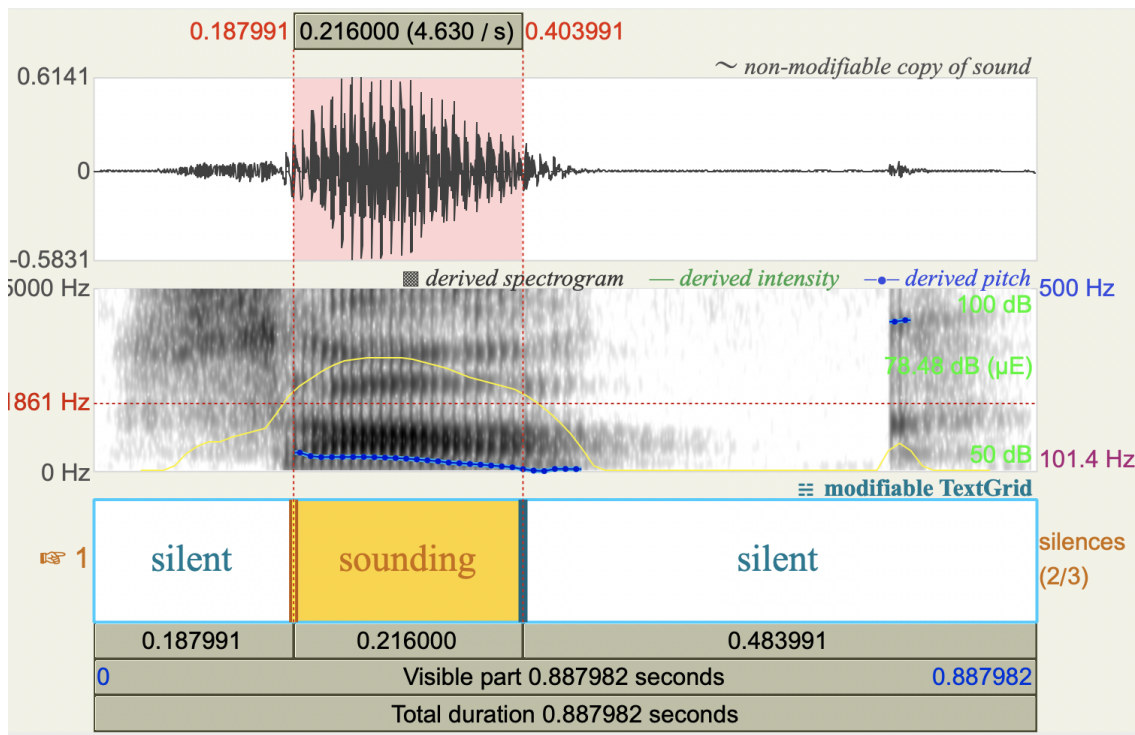
1 Funksioon extract_vowel_formants_v2(heli, maks_formant):
2   Resample'i heli sagedusele 10 kHz ja skaleeri intensiivsus 70 dB-ni;
3   Arvuta helist formandid;
4   Leia helist praat textgrid fn abil helilised ja helitud lõigud ("silent",
   "sounding");
5   Intervallide arv = textgrid intervallide koguarv;
6   for iga intervall (1 kuni intervallide arv) do
7     intervalli silt = tekstvõre silt selles intervallis;
8     if silt on "sounding" then
9       algusaeg = intervalli algus;
10      lõppaeg = intervalli lõpp;
11      Arvuta intensiivsus heli põhjal;
12      Leia maksimaalse intensiivsuse ajahetk vahemikus
        algus-lõpp;
13      vokaali_algus = maksimaalne intensiivsus antud lõigus 20
        ms;
14      vokaali_lõpp = maksimaalne intensiivsus antud lõigus + 20
        ms);
15      Arvuta F1 ja F2 keskmised väärtused vokaali_algus kuni
        vokaali_lõpp;
16      Tagasta F1 ja F2;
17   end

```

Joonis 4. Vokaalide formantide leidmise algoritmi 2 pseudokood

Algoritmide valideerimine

Väljatöötatud eesti keele vokaalide formantide leidmise algoritme saab valideerida kõnetöötuse rakenduse Praat abil. Järgneval joonisel 5 on välja toodud töö autori poolt salvestatud sõna "saak" avatuna Praat rakenduses.



Joonis 5. Näitesalvestis sõnale "saak" avatuna Praat rakenduses

Eeltoodud joonisel 5 on rakendatud algoritmis 2 kasutusel olevat funktsiooni *Sound: To TextGrid (silences)* samade parameetritega, mis algoritmis. Joonisel on välja kuvatud ka intensiivsus kollase joonena ning ka põhitoon sinise joonena.

Esimese algoritmi abil leitakse analüüsitav lõik 0.192 s - 0.408 s. Seejärel rakendatakse formantide keskmise leidmist antud lõigule ning tulemuseks saadakse $F1 = 694.70$ Hz ja $F2 = 1032.62$ Hz.

Teise algoritmi abil leitakse vokaali lõik 0.188s - 0.404s. Misjärel rakendatakse maksimaalse intensiivsuse leidmist antud lõigul. Selle tulemusel saadakse analüüsitavaks lõiguks 0.255s - 0.295s. Seejärel rakendatakse formantide keskmise leidmist antud lõigule ning tulemuseks saadakse $F1 = 712.31$ Hz ja $F2 = 1022.06$ Hz.

Eeltoodud näite puhul töötasid algoritmid võrdväärselt – kui esimeses algoritmis oleks rakendatud maksimaalse intensiivsuse ümber formantide otsimist, siis oleksid tulemused samad. Kuid autori silmis on sisseehitatud funktsionaalsuse kasutamine vokaali tuvastamiseks töökindlam piirjuhtude korral ning helilise lõigu maksimaalse intensiivsuse ümber formantide otsimine peaks töökindlust veelgi tõstma. Eelnimetatud põhjustel kasutatakse valminud taustateenustes algoritmi 2.

3.1.2 Tagasisidestamine

Tagasiside saamiseks on kasutajal võimalik valida kahe meetodi vahel. Esimene meetod põhineb kasutaja vokaali formantide võrdlusel referentsväärtustega. Teine meetod põhineb kasutaja vokaali formantide võrdlusel sama sihtgrupi (vanus, sugu) teiste kõnelejate normaliseeritud väärtustega.

Formantide referentsväärtusel põhinev tagasisidestamine

Esimene tagasiside meetod toimib leitud kasutaja hääldatud sõnast leitud vokaali formantide võrdlemisega eelarvutatud sama sõna vokaali referentsväärtustega. Tagasiside ei arvesta kasutaja vanuse ega sooga. Referentsväärtused ja tolerantsid on leitavad lisas 2 ning need on arvutatud andmekogumisrakendusest kogutud sisendi põhjal, kasutades selle tarbeks loodud Praat skripti.

Tagasiside andmiseks kasutatakse nii visuaalset kuva kui ka esitatakse vajadusel juhendid häälduse parendamiseks.

Visuaalse tagasiside andmiseks kuvatakse kasutajale ellips referentsväärtuste abil, kus Y-telg joonistatakse vokaali F2 referentsväärtuse ning \pm F2 tolerantsi järgi. X-telg joonistatakse vastavalt F1 referentsväärtuse ning \pm F1 tolerantsi järgi. Joonisele märgitakse ka kasutaja poolt hääldatud vokaali F1 ning F2 väärtused punktina.

Tekstiliste juhendite andmiseks kontrollitakse esmalt, kas kasutaja poolt hääldatud vokaal asub eelnevalt välja arvutatud ellipsi sees. Selleks arvutatakse kasutaja poolt hääldatud formantide (F1 ja F2) väärtuse erinevused vokaalide referentsväärtusest (ellipsi keskpunkt) ning skaleeritakse need vastava tolerantsiga. Seejärel kasutatakse normaliseeritud erinevusi Eukleidese kauguse valemis:

$$kaugus = \sqrt{\left(\frac{F1_{kaugus} - F1_{referents}}{F1_{tolerants}}\right)^2 + \left(\frac{F2_{kaugus} - F2_{referents}}{F2_{tolerants}}\right)^2} \quad (3.1)$$

Kui kaugus jääb alla 1, siis kasutaja tulemus on ellipsi sees ning hääldus on perfektne. Kui kaugus on üle 4, siis järelikult kasutaja hääldus erines referentsist liiga palju ning tagasisidestamine lõpetatakse. Kui kaugus jääb vahemiku 1 - 4 vahele, antakse tagasisidet järgneva loogika abil:

Kui F1 on madalam kui referentsväärtus, tähendab see, et kasutaja keel on liiga kõrgel \rightarrow

"Proovi keelt veidi madalamale lasta."

Kui F1 on kõrgem kui referentsväärtus, siis keel on liiga madalal → "Proovi keelt veidi tõsta."

Kui F2 on madalam kui referentsväärtus, siis keel on liiga taga → "Proovi keelt veidi ettepoole liigutada."

Kui F2 on kõrgem kui referentsväärtus, siis keel on liiga ees → "Proovi keelt veidi tahapoole tõmmata."

Tagasisidestamisel kombineeritakse F1 tagasisidet keele kõrguse kohta ning F2 tagasisidet keele kauguse kohta.

Formantide normaliseeritud väärtustel põhinev tagasisidestamine

Teine tagasiside meetod toimib kasutaja vokaali formantide võrdlemisega normaliseeritud teiste samasse soo- ja vanusekategoriasse kuuluvate väärtustega.

Vanusegrupid jaotuvad järgnevalt:

- **Lapsed ja noored:** alla 15-aastased.
- **Noored täiskasvanud:** 15–25-aastased.
- **Täiskasvanud:** 26–60-aastased.
- **Eakad:** 61-aastased ja vanemad.

Kasutaja sisendist leitud vokaali formandid ja teiste samasse gruppi kuuluvate salvestatud formandid normaliseeritakse z-skoori meetodi abil:

$$F_{1norm} = \frac{F1 - \mu_{F1}}{\sigma_{F1}}, \quad F_{2norm} = \frac{F2 - \mu_{F2}}{\sigma_{F2}} \quad (3.2)$$

ning tulemused kantakse joonisele. Lisaks joonistatakse ellipsid eeldefineeritud väärtustega X=1, Y=1 ning X=2, Y=2.

Tagasiside andmiseks kasutatakse Eukleidese kauguse valemit ning kui kasutaja on esimese ellipsi sees (X=1, Y=1), siis loetakse antud sisend ideaalseks ning tagasisidestamine lõppeb. Kui kasutaja hääldus jääb teise ellipsi sisse (X=2, Y=2), siis antakse kasutajale tagasisidet keele kõrguse ja kauguse kohta, kasutades samasugust loogikat nagu meetodis 1 välja

toodud. Kui kasutaja sisend ei jää kummagi ellipsi sisse, siis tagasisidestamine lõpetatakse ning kasutajal palutakse uuesti proovida.

3.2 Rakendused

3.2.1 Vocal-recorder taustteenus

Vocal-recorder taustteenus on Java programmeerimiskeeles kirjutatud rakendus, mis kasutab Spring Boot'i ja Flyway teeke. Rakenduse põhiroll on hallata salvestusi ja pakkuda salvestustega seotud funktsionaalsust. Rakendus pakub kolme otspunkti:

- POST -> /api/v1/profile - võtab sisendiks metaandmed (vanus, sugu) ning loob andmebaasi kirje selle kohta ning tagastab loodud profiili id.
- GET -> /api/v1/all-tasks - tagastab nimekirja sõnadest, mille salvestusi on tarvis koguda.
- POST -> /api/v1/task - võtab sisendiks loodud profiili id, salvestatud sõna ning salvestuse helifaili wav formaadis. Rakendus salvestab salvestuse kohta kirja andmebaasi ning salvestuse helifaili etteantud asukohale serveri kettal.
- GET -> /api/v1/sample-recording - tagastab sisendis saadud sõnale näitesalvestuse helifaili.

Vocal-recorder taustteenuse põhifunktsionaalsus:

- Pakkuda teenust, mille abil luuakse profiil ning salvestada see andmebaasi.
- Pakkuda teenust, mis annab eeldefineeritud sõnad andmebaasist, mille hääldusi soovitakse koguda.
- Pakkuda teenust, mille abil saab näitesalvestused siduda profiiliga ning salvestused serverile salvestada.
- Pakkuda teenus, mis tagastab võimaluse korral eeldefineeritud näite helifaili antud sõna kohta.
- Hallata andmebaasi skeeme ning tabeleid Flyway abil.

Rakendus on konteinerdatud Docker-lahenduse abil ning turvalisuse tagamiseks on tundlikud otspunktid kaitstud API token'iga.

3.2.2 Vocal-processor taustteenus

Vocal-processor taustteenus on Python programmeerimiskeeles kirjutatud rakendus, mis kasutab Parselmouth teeki. Rakenduse peamine eesmärk on leida Vocal-recorder poolt kogutud helifailidest üles olulised akustilised tunnused (formandid F1 ja F2) ning salvestada need uude tabelisse. Samuti vastutab see rakendus andmete kvaliteedi eest ehk salvestused, millest leitud formantväärtused on liiga suure erinevusega, märgitakse pehme kustutamise (*soft-delete*) abil mitteaktiivseteks ehk edasisel töötlemisel neid väärtusi ei kasutata. Rakendus käivitub ajastatud protsessina kord tunnis ning seejärel töötleb kõik uued salvestused läbi.

Vocal-processor taustteenuse põhifunktsionaalsus:

- Töödelda läbi salvestute tabelis olevad uued kirjed.
- Leida salvestatud helifailidest formantide väärtused (F1 ja F2), siduda need profiiliga ning luua saadud info põhjal uus kirje töödeldud salvestuste tabelisse edasiseks protsessimiseks.
- Kontrollida töödeldud salvestustete tabelit ning märkida liiga suurte kõikumistega väärtused mitteaktiivseteks.

3.2.3 Vocal-feedback taustteenus

Vocal-feedback taustteenus on Python programmeerimiskeeles kirjutatud rakendus, mis kasutab Flaski ja Parselmouthi teeki. Rakenduse põhiroll on anda tagasisidet kasutaja hääldestele Vocal-processor abil kogutud andmete ja eeldefineeritud formantväärtuste põhjal. Rakendus annab tagasisidet vastavalt kasutaja valitud meetoditele: 1) formantide normaliseeritud väärtustel põhinev tagasisidestamine või 2) formantide referentsväärtusel põhinev tagasisidestamine. Lisaks esitab graafiliselt kasutaja andmeid sessiooni kohta.

Rakendus pakub kahte otspunkti:

- POST -> /api-p/v1/statistics-table - võtab sisendiks kasutaja andmed sessiooni kohta ning tagastab selle põhjal graafilise tabeli kasutaja sessiooni kohta,
- POST ->/api-p/v1/upload-recording - võtab sisendiks helisalvestuse ning tagasiside meetodi valiku ja tagastab kasutajale visuaalse ja tekstilise tagasiside antud salvestusele.

Vocal-feedback taustteenuse põhifunktsionaalsus:

- Anda kasutajale tagasisidet salvestatud heliklipile valitud meetodi põhjal.
- Anda kasutajale visuaalne ülevaade tema sessiooni kohta.

Rakendus on konteinerdatud Docker lahenduse abil ning turvalisuse tagamiseks on tundlikud otspunktid kaitstud API tokeniga.

3.2.4 Andmekogumisrakendus

Andmekogumisrakenduseks on Swift programmeerimiskeeles kirjutatud rakendus ning see on suunatud iPhone kasutajatele. Andmekogumisrakenduse eesmärk on koguda kasutajatelt kõneandmeid edasisteks akustilisteks analüüsideks. Rakenduse põhifunktsionaalsus hõlmab endas vaadet, mille kaudu kasutaja sisestab metaandmed (vanus, sugu) ning kasutajale esitatakse hääldamiseks ning häälduse salvestamiseks etteantud eestikeelsed sõnad. Rakenduse taustateenusena töötab Vocal-recorder.

Andmekogumisrakenduse põhifunktsionaalsus:

- Juhendada kasutajat rakendust kasutama.
- Koguda kasutaja metaandmed (vanus, sugu).
- Küsida taustteenuselt sõnad, mille hääldusi on tarvis koguda ning need kasutajale välja kuvada.
- Salvestada kasutaja poolt hääldatud sõnad ning saata need koos metaandmetega taustteenusele edasiseks töötluks.

3.2.5 Hääldustreenimisrakendus

Hääldustreenimisrakenduse eesmärk on juhendada kasutajaid eestikeelsete vokaalide hääldamises ning anda tagasisidet kasutaja vokaalide hääldustele etteantud sõnades. Hääldustreenimisrakendus on Swift programmeerimiskeeles kirjutatud ning suunatud iPhone kasutajatele. Rakenduses kogutakse metaandmed (vanus, sugu) ning tutvustatakse, kuidas eesti keele vokaale tuleks hääldada. Samuti antakse ülevaade võimalikest tagasiside meetoditest. Seejärel on kasutajal võimalik hääldada etteantud sõnu ja kuulata näitehääldusi ning saada tagasisidet vastavalt enda valitud tagasiside meetodile. Lisaks saab kasutaja vaadata sessiooni andmeid.

Hääldustreenimisrakenduse põhifunktsionaalsus:

- Juhendada kasutajat rakendust kasutama.

- Koguda kasutaja metaandmed (vanus, sugu).
- Anda ülevaade tagasiside meetodidest ning vokaalidest ja nende akustilistest tunnustest.
- Kuvada kasutajale sõnu Vocal-recorder taustteenusest.
- Esitada kasutajale vajadusel sõnade näitehääldusi Vocal-recorder taustteenuse abil.
- Koguda kasutaja hääldusi etteantud sõnadele. Häälduste kogumiseks kasutatakse VAD abi ehk rakendus tuvastab ise häälduse lõpu. Peale etteantud sõna häälduse kogumist saadetakse sõna Vocal-feedback taustteenusele ning selle tulemusel saab kasutaja tagasisidet oma häälduse kohta vastavalt eelnevalt valitud tagasiside meetodile.
- Kuvada kasutajale sessiooni kokkuvõtte Vocal-feedback taustteenuse abil kasutaja häälduste kohta kasutades eelnevalt valitud tagasiside meetodit.

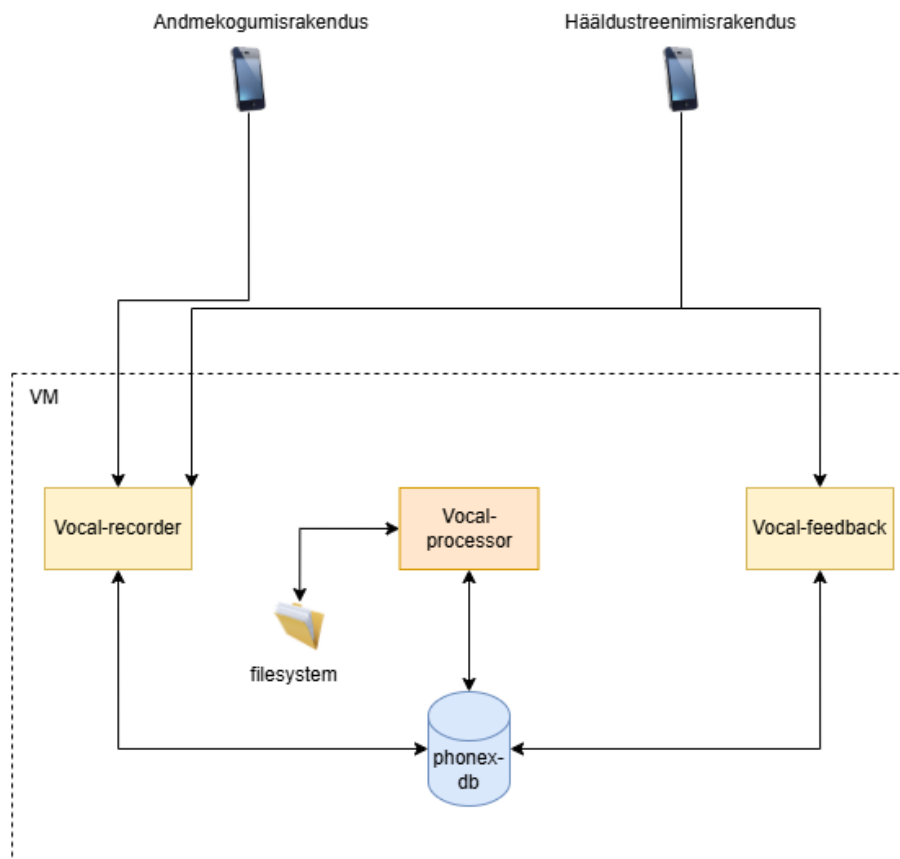
3.2.6 Rakenduste evitamine

Kasutajaliidese rakendused ehk hääldustreenimiserakendus ja andmekogumiserakendus on evitatud kasutades Apple ametlikku keskkonda App Store Connect [44]. Rakendused on tasuta kättesaadavad kõikides regioonides ning neid saab alla laadida Apple Store'ist [45] kaudu.

Taustteenused on evitatud Docker'i abil Taltech poolt pakutud virtuaalmasinas. Avaliku veebi ja sisemiste teenuste vahel töötab Nginx *reverse-proxy*. Lisaks lisab *reverse-proxy* Letsencrypt'i abil päringutele külge kehtiva TLS-sertifikaadi, et pakkuda krüpteeritud (*HTTPS*) ühendusi.

3.2.7 Rakenduste arhitektuuri ülevaade

Rakendused on arendatud mikroteenustena ehk iga komponent on vastutav mingi kindla piiratud funktsionaalsuse eest. Järgneval joonisel 6 on välja toodud üldine (*high-level*) arhitektuurijoonis rakenduste töö kohta.



Joonis 6. Rakenduste üldine arhitektuur, mis näitab rakenduste seoseid

Joonisel 6 on näha, et andmekogumisrakendus kasutab enda tööks Vocal-recorderi taustteenuse pakutud API meetodeid. Häälustreenimisrakendus kasutab nii Vocal-recorderi kui ka Vocal-feedback taustteenuste API meetodeid. Salvestused seotakse andmebaasi Vocal-processor'i töö abil, mis suhtleb andmebaasiga ja loeb faile süsteemikettalt.

4. Tulemused

Järgnevas peatükis tutvustatakse valminud rakendusi ning nende vaateid.

4.1 Andmekogumisrakendus

Andmekogumisrakendus on iOS rakendus, mis on kättesaadav Apple Store'is nimega "Phonex vocal recorder". Andmekogumisrakenduse peamine eesmärk on koguda näitehääldusi eelmääratud sõnadele erinevate kasutajate poolt. Rakendus koosneb neljast vaatest: 1) "Nõusolek osalemiseks ja privaatsuspoliitika tutvustamine", 2) "Küsimustik", 3) "Sissejuhatus" ja 4) "Salvestamine"

4.1.1 Vaade 1: Nõusolek osalemiseks ja privaatsuspoliitika tutvustamine

Rakenduse esimeses vaates tutvustatakse kasutajale tingimusi ja privaatsuspoliitikat ning nõutakse kasutajapoolset kinnitust neile, et liikuda edasi. Järgneval joonisel 7 on toodud välja eelkirjeldatud vaade.

15:51

Consent to Participate and Privacy Policy

The goal of the project "Estonian Language Pronunciation Tutoring Application for Estonian Language Learners" is to develop an application that helps users learn and improve the pronunciation of Estonian vowels. To achieve this, we collect recordings of Estonian words and analyze them to extract acoustic data essential for the application's development.

By participating, you agree that:

- Your recordings will be used anonymously for acoustic analysis.
- You will provide your age and sex, which will be anonymized and cannot be traced back to you.

[Click here to read the Terms of Use](#)

I agree to the Terms of Use and Privacy Policy



Accept

Joonis 7. Andmekogumiskrakenduse vaade: Nõusolek osalemiseks ja privaatsuspoliitika tutvustamine

4.1.2 Vaade 2: Küsimustik

Rakenduse teises vaates kogutakse kasutaja kohta metaandmed – kasutaja sugu ning vanus. Järgneval joonisel 8 on toodud välja eelkirjeldatud vaade.

The screenshot shows a mobile application interface for a questionnaire. At the top, the time is 15:51, and there are icons for signal strength, Wi-Fi, and battery. Below the status bar is a blue back arrow and the text "Back". The main heading is "Questionnaire" in a large, bold, black font. Underneath, the instruction "Select your gender" is followed by two radio button options: "MALE" (which is selected) and "FEMALE". Below this is the instruction "Enter your age" followed by a text input field containing the number "25". At the bottom of the form is a large blue button with the text "Submit".

Joonis 8. Andmekogumisrakenduse vaade: Küsimustik

4.1.3 Vaade 3: Sissejuhatus

Rakenduse kolmandas vaates tutvustatakse kasutajale rakenduse eesmärki. Järgneval joonisel 9 on toodud välja eelkirjeldatud vaade.



Introduction

This app is used to record Estonian words. You will be guided through the process of recording words to help build a dataset for linguistic purposes.

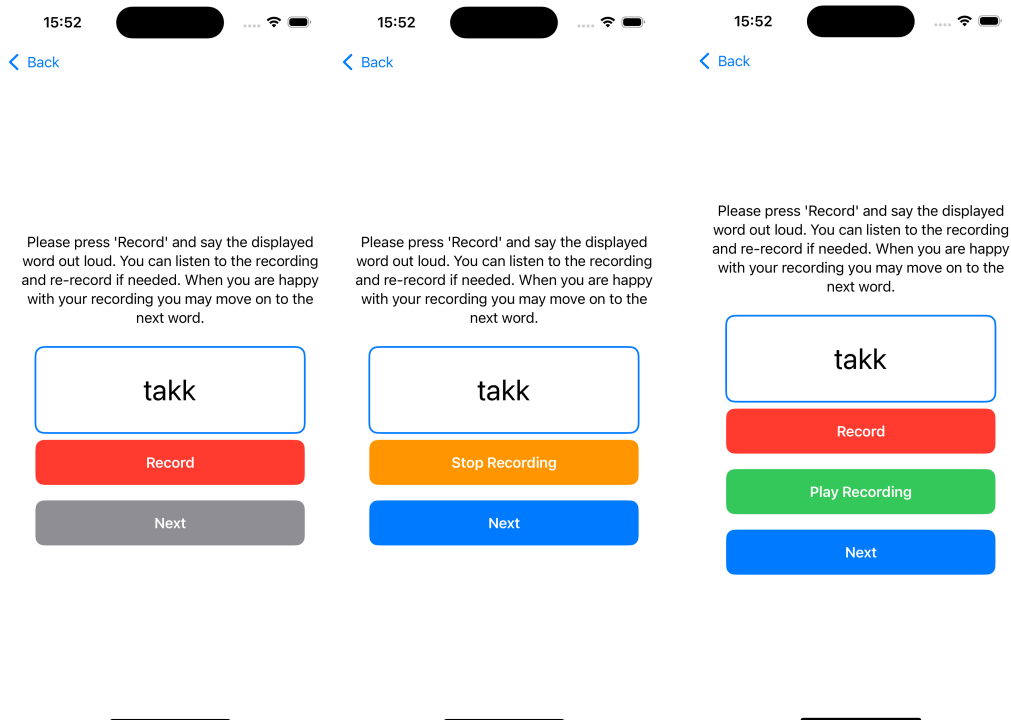


Joonis 9. Andmekogumisrakenduse vaade: Sissejuhatus

4.1.4 Vaade 4: Salvestamine

Rakenduse neljandas vaates juhendatakse kasutajat salvestama ekraanil kuvatud sõna. Neljas vaade koosneb mitmest erinevast olekust, sõltuvalt sellest, kas kasutaja on juba salvestanud sõna, salvestamine käib või salvestamist pole alustatud. Järgneval joonisel 10 on toodud võimalikud olekud:

- Esimene olek - kasutaja pole veel sõna salvestanud.
- Teine olek - salvestamine on töös.
- Kolmas olek - salvestamine on olnud edukas ning saab edasi liikuda järgneva sõna juurde.



Joonis 10. Andmekogumisrakenduse vaade: Salvestamise olekud 1, 2 ja 3 (vasakult paremale)

4.2 Hääldestreenimisrakendus

Hääldestreenimisrakendus on iOS rakendus, mis on kättesaadav Apple Store'is nimega "Phonex vocal learner". Hääldestreenimisrakenduse peamine eesmärk on anda kasutaja hääldestele visuaalset ja tekstilist tagasisidet. Hääldestreenimisrakendus pakub tagasisides-tamiseks kahte erinevat meetodit: 1) formantide referentsväärtusel põhinev tagasisides-tamine ja 2) formantide normaliseeritud väärtustel põhinev tagasisides-tamine. Kasutaja saab oma sessiooni kohta vaadata statistikat. Rakendus koosneb neljast vaatest: 1) "Nõusolek osalemiseks ja privaatsuspoliitika tutvustamine", 2) "Küsimustik", 3) "Ülevaade tagasiside meetoditest ja hääldeõpetus" ja 4) "Salvestamine".

4.2.1 Vaade 1: Nõusolek osalemiseks ja privaatsuspoliitika tutvustamine

Rakenduse esimeses vaates tutvustatakse kasutajale tingimusi ja privaatsuspoliitikat ning nõutakse kasutajapoolset kinnitust neile, et liikuda edasi. Järgneval joonisel 11 on toodud välja eelkirjeldatud vaade.

16:40

Terms of Use and Privacy Policy

The goal of the "Phonex-vocal-learner" is to help users learn and improve the pronunciation of Estonian vowels. This application uses pre-processed acoustic data collected from the "phonex-vocal-recorder" app to provide personalized feedback on your pronunciation.

By using this application, you agree that:

- Your pronunciation will be compared against anonymous, pre-processed acoustic data.
- Your age and gender will be used solely to match you with the appropriate acoustic benchmarks, ensuring tailored feedback.
- No new recordings or personal data will be collected or stored by this app.

[Click here to read the Terms of Use](#)

I agree to the Terms of Use and Privacy Policy



Accept

Joonis 11. Häädustreenimisrakenduse vaade: Nõusolek osalemiseks ja privaatsuspoliitika tutvustamine

4.2.2 Vaade 2: Küsimustik

Rakenduse teises vaates küsitakse kasutaja metaandmed - kasutaja sugu ning vanus. Järgneval joonisel 12 on toodud välja eelkirjeldatud vaade.

16:40



Questionnaire

Select your gender

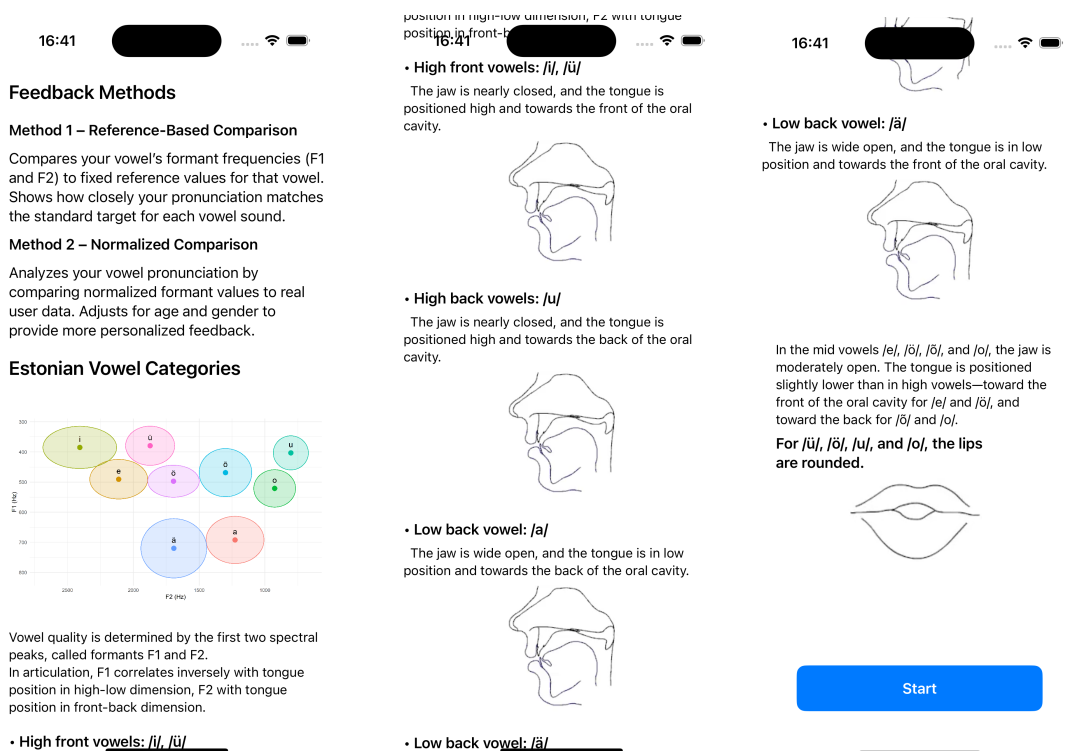
MALE FEMALE

Enter your age

Joonis 12. Häälustreenimisrakenduse vaade: Küsimustik

4.2.3 Vaade 3: Ülevaade tagasiside meetoditest ja häälusõpetus

Rakenduse kolmandas vaates tutvustatakse kasutajale võimalikke tagasiside meetodeid ja nende toimimist ning antakse ülevaade vokaalide häälendamise põhitõdedest koos visuaalsete näidetega. Järgneval joonisel 13 on toodud välja eelkirjeldatud vaade.



Joonis 13. Häälustreenimisrakenduse vaade: Ülevaade tagasiside meetoditest ja hääldejuhised

4.2.4 Vaade 4: Salvestamine

Neljas vaade koosneb mitmest erinevast olekust, sõltuvalt kasutaja tegevusest salvestamisrežiimis:

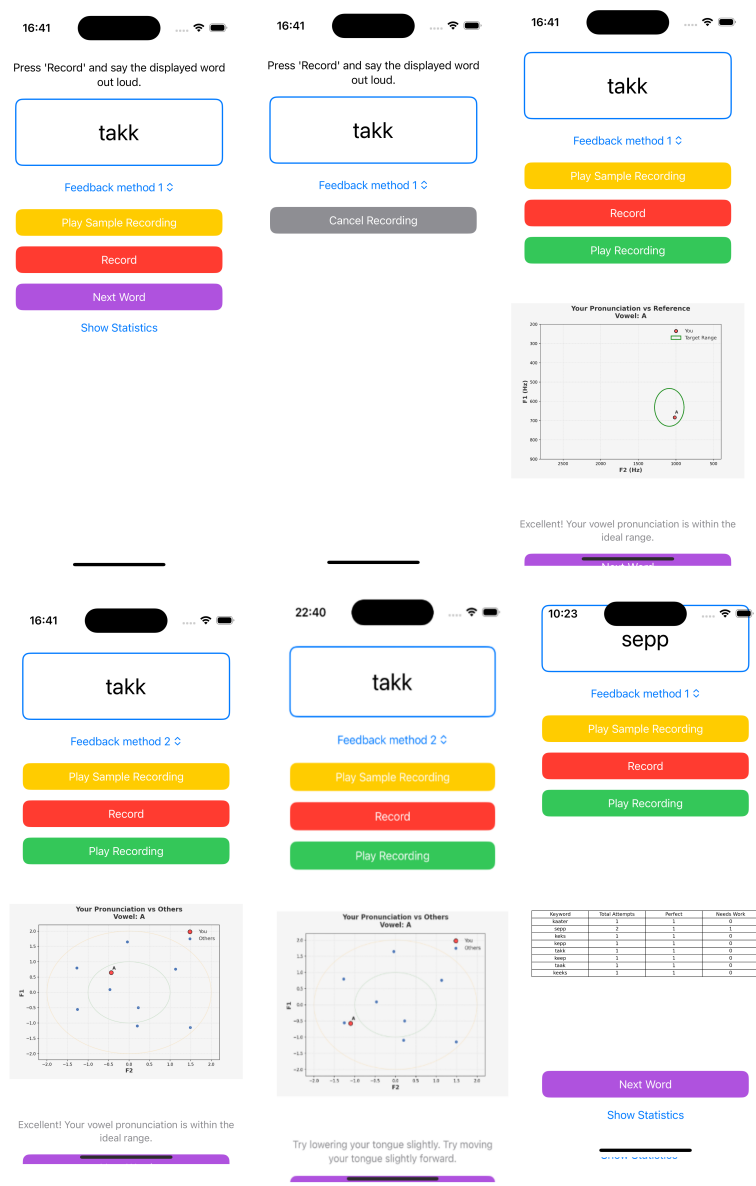
- kui salvestamist ei ole veel alustatud,
- kui salvestamine on käimas,
- kui sõna on juba salvestatud.

Lisaks mõjutab salvestamise väljundit valitud tagasiside meetod, mis omakorda võib muuta vaate sisu ja käitumist. Neljandas vaates on integreeritud ka statistika olek, kus antakse ülevaade kasutaja sessioonist. Järgneval joonisel 14 on toodud võimalikud olekud:

- Esimene olek - kasutaja pole veel sõna salvestanud.
- Teine olek - salvestamine on töös. Võrreldes Andmekogumisrakendusega on antud vaate puhul kasutusel VAD abil salvestuse lõpetamine ehk kasutaja ise enam ei pea käsitsi "Stop recording" vajutama.
- Kolmas olek - salvestamine oli edukas ja perfektne ning kasutaja saab tagasiside meetodi 1 abil. Kasutaja vokaali asukoht vokaalikaardil kuvatakse punase täpina ja

vokaalile vastav referentspiirkond roheline ellipsina.

- Neljas olek - salvestamine oli edukas ja perfektne ning kasutaja saab tagasiside meetodi 2 abil. Kasutaja vokaali asukoht vokaalikaardil kuvatakse punase täpina ning teiste kasutajate hääldused sinise täpina. Vokaalikaardil on kuvatud roheline ellipsina oodatud hääldus ning kollase ellipsiga piirkond, kus antakse juhised häälduse parendamiseks.
- Viies olek - salvestamine oli edukas, kuid mitte perfektne; kasutaja saab tagasiside meetodi 2 abil ja talle kuvatakse suunised häälduse parendamiseks (sarnast tagasiside kuvatakse ka meetodi 1 puhul kui hääldus ei ole perfektne).
- Kuues olek - sessiooni statistika.



Joonis 14. Hääldustreenimiskrakenduse vaade: Salvestamine - kõik võimalikud olekud (ülemises reas olekud 1,2 ja 3, alumises 4, 5 ja 6)

5. Tulemuste analüüs ja edasised arendused

Järgnevas peatükis antakse ülevaade rakenduste kasutusstatistikast, analüüsitakse kasutajakogemust ning tuuakse välja võimalikud edasised arendused.

5.1 Analüüs

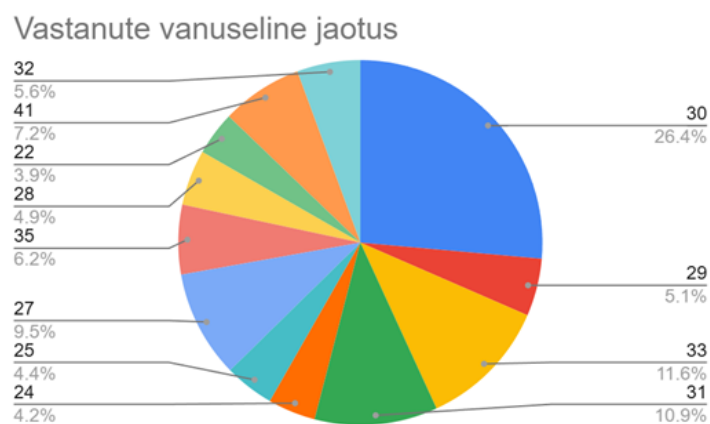
Rakenduste tulemuste paremaks mõistmiseks analüüsitakse rakenduse kasutusstatistikat ning Google Forms'i põhjal koostatud küsimustiku tulemusi.

5.1.1 Andmekogumisrakendus

Andmekogumisrakenduse abil koguti 1463 unikaalset salvestust. Andmekogumisrakenduses oli kokku 44 sessiooni, neist 22 sessiooni jooksul salvestati vähemalt 10 erinevat helifaili.

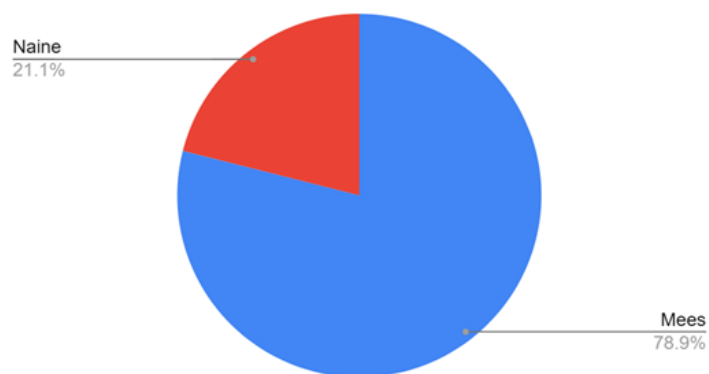
5.1.2 Häädustreenimiskrakendus

Häädustreenimiskrakenduse käitumise paremaks mõistmiseks viidi läbi Google Forms'il põhinev küsitlus. Küsitluse eesmärk oli saada põhjalikum ülevaade häädustreenimiskrakenduse kasutajakogemusest nii eesti keelt emakeelena kõnelevate (L1) kui ka eesti keelt teise keelena õppivate (L2) kasutajate seas. Küsitluses osales kokku 19 inimest, nende sooline jaotus on esitatud joonisel 16 ja vanuseline jaotus joonisel 15.



Joonis 15. Tagasiside andmisel osalenud kasutajate vanuseline jaotus

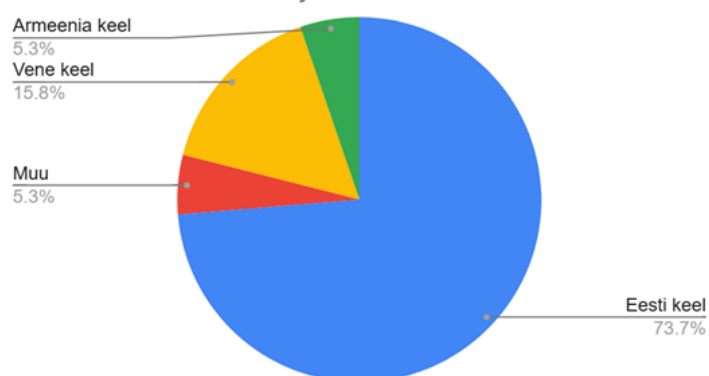
Vastanute sooline jaotus



Joonis 16. Tagasiside andmisel osalenud kasutajate sooline jaotus

Joonisel 17 on esitatud kasutajate keeleline jaotus: 73,7% vastanutest rääkis emakeelena eesti keelt, 15,8% vene keelt, 5,3% armeenia keelt ning 5,3% mõnda muud keelt.

Vastanute emakeele jaotus



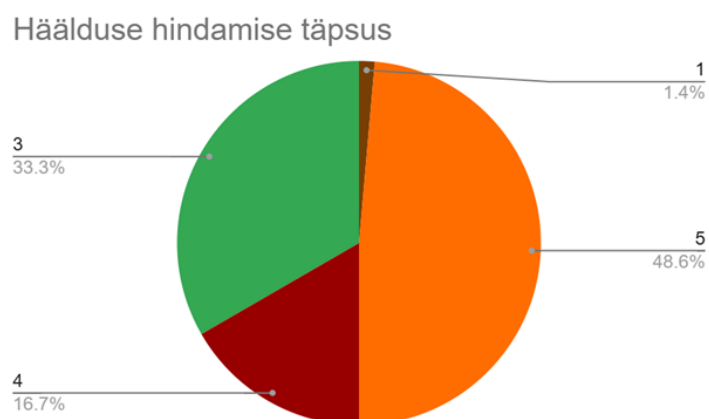
Joonis 17. Tagasiside andmisel osalenud kasutajate emakeel

Joonisel 18 on esitatud tagasiside andmisel osalenud kasutajate hinnang juhendite ja meetodite kirjelduste arusaadavusele: 56,3% andis hindeks 5, 30,0% hindeks 4, 11,3% hindeks 3 ning 2,5% hindeks 2.



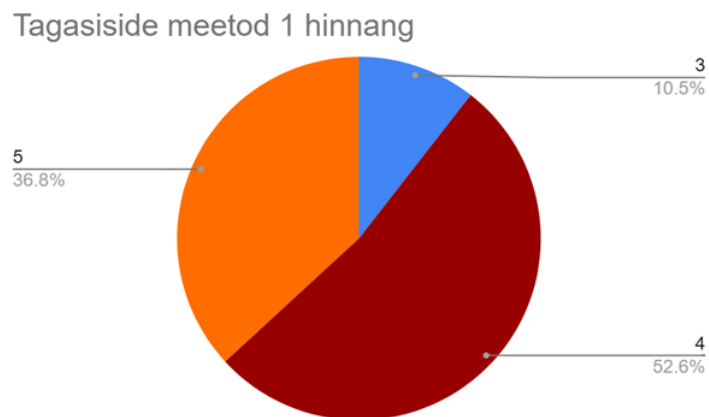
Joonis 18. Tagasiside andmisel osalenud kasutajate hinnang juhendite ja meetodide kirjelduste arusaadavusele

Joonisel 19 on esitatud tagasiside andmisel osalenud kasutajate hinnang rakenduse häälde hindamise täpsusele: 48,6% andis hindeks 5, 16,7% hindeks 4, 33,3% hindeks 3 ning 1,4% hindeks 1.



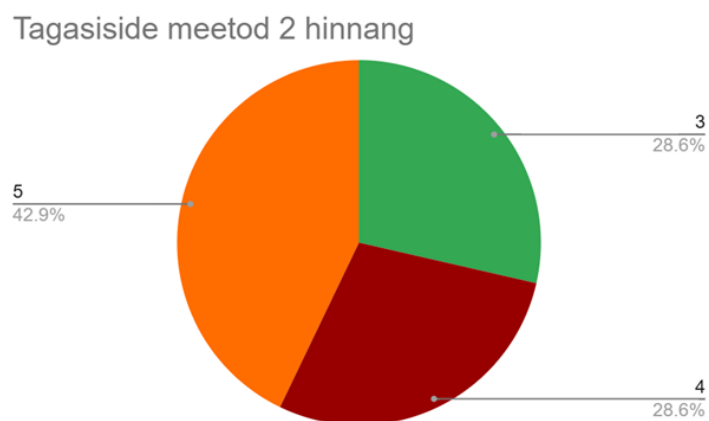
Joonis 19. Tagasiside andmisel osalenud kasutajate hinnang rakenduse häälde hindamise täpsusele

Joonisel 20 on esitatud tagasiside andmisel osalenud kasutajate hinnang rakenduse tagasiside meetodile 1: 36,8% andis hindeks 5, 52,6% hindeks 4 ning 10,5% hindeks 3.



Joonis 20. Tagasiside andmisel osalenud kasutajate hinnang rakenduse tagasiside meetodile 1

Joonisel 21 on esitatud tagasiside andmisel osalenud kasutajate hinnang rakenduse tagasiside meetodile 2: 42,9% andis hindeks 5, 28,6% hindeks 4 ning 28,6% hindeks 3.



Joonis 21. Tagasiside andmisel osalenud kasutajate hinnang rakenduse tagasiside meetodile 2

Tagasiside põhjal võib järeldada, et kasutajad hindasid rakenduse erinevat funktsionaalsust enamjaolt positiivselt. Kõige kõrgem hinnang anti juhendite ja meetodite arusaadavusele, kus 86,3% vastanutest andis hindeks 4 või 5. Samuti hinnati kõrgelt rakenduse häälduse hindamise täpsust ning erinevaid tagasiside meetodeid, kus mõlemas kategoorias moodustasid kõrged hinnangud üle 85% vastustest. Kuigi esines ka madalamaid hindeid, olid need vähemuses.

Tuleb siiski arvestada, et küsitlus põhineb kasutajate subjektiivsel hinnangul enda häälduse põhjal. Täpsemate järelduste tegemiseks ning arusaamiseks, miks kasutajad teatud

viisil hindasid, tuleks läbi viia uuring kontrollitud laboritingimustes ning kaasata ka fonetikaekspert, et tuvastada kasutaja tegelik hääldus ning hinnata rakenduse tagasiside vastavust.

5.2 Edasised arendused

Hääldustreenimiskrakenduses on implementeeritud häälduse kordamisharjutused, mis toetavad õppijate suulise keeleoskuse arengut. Siiski puuduvad rakenduses diskrimineerimisharjutused, mis on arvutipõhise keeleõppe süsteemide oluline komponent. Fonoloogiline diskrimineerimisvõime on keeleõppe varajases etapis keskse tähtsusega, kuna see võimaldab õppijal eristada ja tuvastada sarnaseid häälikuid või sõnavorme.

Seetõttu peaks lisama rakenduse edasistesse versioonidesse fonoloogilise diskrimineerimise harjutused, näiteks kontrastiivsete sõnade eristamise ülesanded ning valikvastustega testid.

Lisaks, võttes arvesse tagasisideküsitluse tulemusi, on ilmne, et kordamisharjutuste täpsuses esineb veel puudujääke. Täpsuse tõstmiseks tuleks laboritingimustes analüüsida, millistes olukordades probleemid tekivad, ning täiustada formantide leidmise algoritmi kogutud andmete põhjal. Üheks võimalikuks lahenduseks on närvivõrkudel põhineva formantide leidmise kasutuselevõtt. Vastav lähenemine on implementeeritud Python-teegis *Deep formants* [46], mille rakendamine võiks tõsta mitte-laboritingimustes salvestatud helifailide töötlemise täpsust.

Tagasisides toodi esile veel mitmeid kasutajaliidese ja funktsionaalsuse puudujääke. Näiteks puuduvad hääldusnäited naishäältega. Samuti sooviti võimalust valida, missugust vokaali soovitakse treenida, et keskenduda konkreetsematele raskuskohtadele. Treeningrežiimis toodi välja ka vajadus tagasinupu järele, mis suunaks kasutaja tagasi õpetuse juurde. Rakenduste edasistes versioonides tuleks arvestada nende märkustega.

6. Kokkuvõte

Magistritöö raames arendati välja kaks iOS-rakendust: andmekogumisrakendus Phonex Vocal Recorder ja hääldustreenimiserakendus Phonex Vocal Learner. Esimese rakenduse eesmärk oli koguda emakeelsete kõnelejate näitesalvestusi ettemääratud sõnadele. Kokku koguti 1463 unikaalset salvestust. Salvestusi kasutati, et koostada vokaalikaart ning eraldada referentsväärtused vokaalide hindamiseks.

Teine rakendus ehk töö lõpptulemusena valminud hääldustreenimiserakendus kasutab andmekogumisrakenduse abil salvestustest leitud formantide väärtusi ning annab kasutaja hääldusele nii visuaalset kui ka tekstilist tagasisidet. Tagasiside osas on võimalik kasutajal valida kahe meetodi vahel: 1) referentsväärtusel põhinev tagasiside ning 2) normaliseeritud väärtustel põhinev tagasiside.

Rakenduste loomisel kasutati kaasaegseid tehnoloogiaid, sealhulgas Swift, Python, Java, Spring Boot ja PostgreSQL. Lahendus põhineb mikroteenuste arhitektuuril, mis võimaldab paindlikku arendust ja süsteemi skaleeritavust.

Hääldustreenimiserakenduse toimimise ja kasutajakogemuse paremaks mõistmiseks viidi läbi tagasisideküsitlus, kus osales 19 inimest. Kasutajakogemus oli valdavalt positiivne, kuid toodi välja ka mõningad puudused ning võimalikud edasised arendussuunad. Rakenduse edasistes versioonides peaks analüüsima, mis olukordades ei tööta vokaali tuvastamine piisavalt hästi ning vastavalt analüüsi tulemustele algoritmi parandused sisse viima. Samuti viidati, et rakendusel puudub võimalus navigeerida edasi/tagasi ning ei saa valida häälikut, mida konkreetselt soovitakse treenida.

Kuigi väljatöötatud hääldustreenimiserakendusel ilmnes mõningaid puudusi, võib positiivse tagasiside põhjal järeldada, et töö eesmärgid on täidetud ning kõik magistritöö alguses püstitatud uurimisküsimused on käsitletud.

Töö tulemusena valmis toimiv rakendus, mis toetab eesti keele häälduse omandamist individuaalselt ning millel on potentsiaali ka laiemaks rakendamiseks haridusvaldkonnas.

Kasutatud kirjandus

- [1] Statistikaamet. *RV068: Rahvastik, 1. jaanuar | Aasta, elukoht ning kodakondsus 2023, 2024*. 2025. URL: https://andmed.stat.ee/et/stat/rahvastik__rahvastikunaitajad-ja-koosseis__rahvaarv-ja-rahvastiku-koosseis/RV068 (vaadatud 25.02.2025).
- [2] Statistikaamet. *RLV431: Rahvastik emakeele ja soo järgi (1979, 1989, 2000, 2011, 2021)*. 2025. URL: https://andmed.stat.ee/et/stat/rahvastik__rahvastikunaitajad-ja-koosseis__rahvaarv-ja-rahvastiku-koosseis/RLV431 (vaadatud 08.04.2025).
- [3] Lya Meister. „Eesti vokaali- ja kestuskategooriad vene emakeelega keelejuhtide tajus ja häälduses: Eksperimentaalfoneetiline uurimus“. Doktoritöö. Tartu Ülikool, 2011.
- [4] Eric H. Lenneberg. *Biological Foundations of Language*. New York: John Wiley & Sons, 1967.
- [5] John Gillespie. „CALL research: Where are we now?“ *ReCALL* 32.2 (mai 2020), lk. 127–144. DOI: 10.1017/S0958344020000051.
- [6] Javed Iqbal Mirani *et al.* „A Review of Computer-Assisted Language Learning (CALL): Development, Challenges, and Future Impact“. *Education and Linguistics Research* 5.1 (2019), lk. 37–45. ISSN: 2377-1356. DOI: 10.5296/elr.v5i1.14515. URL: <https://www.macrothink.org/journal/index.php/elr/article/view/14515> (vaadatud 12.05.2025).
- [7] Jin Ha Woo ja Heeyoul Choi. *Systematic Review for AI-based Language Learning Tools*. arXiv preprint arXiv:2111.04455. 2021. DOI: 10.48550/arXiv.2111.04455.
- [8] Joan Palmiter Bajorek. „L2 Pronunciation Tools: The Unrealized Potential of Prominent Computer-assisted Language Learning Software“. *Issues and Trends in Educational Technology* 5.1 (2017), lk. 1–12. URL: <https://journals.librarypublishing.arizona.edu/itlt/article/id/1497/>.
- [9] Simya Solutions Ltd. *Simply Learn Estonian*. 2025. URL: <https://simplylearnapp.com/estonian> (vaadatud 09.04.2025).
- [10] A. Neri, C. Cucchiari ja H. Strik. „The Effectiveness of Computer-Based Speech Corrective Feedback for Improving Segmental Quality in L2 Dutch“. *ReCALL* 20.2 (2008), lk. 225–243. DOI: 10.1017/S0958344008000724.

- [11] A. Neri *et al.* „The Pedagogy-Technology Interface in Computer Assisted Pronunciation Training“. *Computer Assisted Language Learning* 15.5 (2002), lk. 441–467. DOI: 10.1076/call.15.5.441.13473.
- [12] M. Amrate ja P.-H. Tsai. „Computer-Assisted Pronunciation Training: A Systematic Review“. *ReCALL* 37.1 (2025), lk. 22–42. DOI: 10.1017/S0958344024000181.
- [13] Anton Malmi, Katrin Leppik ja Pärtel Lippus. „SayEst – mobiilirakendus eesti keele häälduse harjutamiseks“. *Oma Keel* 45.Sügis (2022). URL: <https://www.emakeeleselts.ee/valjaanne/oma-keel-45-sugis-2022/> (vaadatud 08.05.2025).
- [14] Tanel Alumäe. „Full-duplex speech-to-text system for Estonian“. Teoses: *Human Language Technologies–The Baltic Perspective*. IOS Press, 2014.
- [15] Maria Letta. *Android nutitelefonil märg eesti keele vokaalide hääldustreeninguks*. 2021.
- [16] Joren Six. *TarsosDSP: A Real-Time Audio Processing Framework in Java*. <https://github.com/JorenSix/TarsosDSP>. 2013. (Vaadatud 08.05.2025).
- [17] Korbinian Riedhammer. *JSTK: Java Speech ToolKit*. <https://github.com/sikoried/jstk>. 2012. (Vaadatud 08.05.2025).
- [18] Erko Peterson. *Veebirakendus Eesti keele häälduse treeninguks*. 2023.
- [19] Red Hat. *What Are Microservices?* Red Hat. 2023. URL: <https://www.redhat.com/en/topics/microservices/what-are-microservices> (vaadatud 15.04.2025).
- [20] Docker Inc. *Docker Overview*. Docker. 2024. URL: <https://docs.docker.com/get-started/docker-overview/> (vaadatud 09.04.2025).
- [21] Inc. NGINX. *NGINX - High Performance Load Balancer, Web Server, Reverse Proxy*. NGINX.org. 2024. URL: <https://nginx.org/> (vaadatud 14.04.2025).
- [22] Internet Security Research Group. *About Let's Encrypt*. Let's Encrypt. 2021. URL: <https://letsencrypt.org/about/> (vaadatud 14.04.2025).
- [23] PostgreSQL Global Development Group. *About PostgreSQL*. PostgreSQL. 2024. URL: <https://www.postgresql.org/about/> (vaadatud 11.04.2025).
- [24] Oracle Corporation. *Oracle Database Features*. Oracle. 2024. URL: <https://www.oracle.com/database/technologies/> (vaadatud 11.04.2025).
- [25] Microsoft. *SQL Server Overview*. Microsoft. 2024. URL: <https://learn.microsoft.com/en-us/sql/sql-server/> (vaadatud 11.04.2025).
- [26] Oracle Corporation. *MySQL – The World's Most Popular Open Source Database*. Oracle. 2024. URL: <https://www.mysql.com/> (vaadatud 11.04.2025).

- [27] Apple. *Swift Programming Language*. 2025. URL: <https://github.com/swiftlang/swift> (vaadatud 11.04.2025).
- [28] Oracle Corporation. *What is Java?* Oracle. 2024. URL: https://www.java.com/en/download/help/whatis_java.html (vaadatud 11.04.2025).
- [29] Oracle Corporation. *Java Overview – Oracle Database 23c*. Oracle. 2024. URL: <https://docs.oracle.com/en/database/oracle/oracle-database/23/jjdev/Java-overview.html> (vaadatud 11.04.2025).
- [30] Inc. VMware. *Spring Boot – Features*. Spring.io. 2024. URL: <https://spring.io/projects/spring-boot#overview> (vaadatud 11.04.2025).
- [31] Inc. VMware. *Spring Initializr - Create a Spring Boot application*. Spring.io. 2024. URL: <https://start.spring.io/> (vaadatud 11.04.2025).
- [32] Redgate Software. *Redgate Flyway Community Edition*. Redgate. 2024. URL: <https://www.red-gate.com/products/flyway/community/> (vaadatud 14.04.2025).
- [33] Redgate Software. *Migrations-based approach - Redgate Flyway Documentation*. Redgate. 2025. URL: <https://documentation.red-gate.com/flyway/deploying-database-changes-using-flyway/deployment-approaches-with-flyway/migrations-based-approach> (vaadatud 14.04.2025).
- [34] Paul Boersma ja David Weenink. *Praat: doing phonetics by computer [Computer program]. Version 6.4.31*, <http://www.praat.org>. 2025. (Vaadatud 08.05.2025).
- [35] Guido van Rossum. *The Python Documentation*. Python Software Foundation. 2024. URL: <https://docs.python.org/3/> (vaadatud 14.04.2025).
- [36] Python Software Foundation. *Python – GitHub repository*. GitHub. 2024. URL: <https://github.com/python> (vaadatud 14.04.2025).
- [37] Yannick Jadoul. *Parselmouth – Praat in Python, the Pythonic way*. GitHub. 2024. URL: <https://github.com/YannickJadoul/Parselmouth> (vaadatud 14.04.2025).
- [38] Pallets Projects. *Flask Documentation (Stable Version)*. Pallets Projects. 2024. URL: <https://flask.palletsprojects.com/en/stable/> (vaadatud 14.04.2025).
- [39] Keith Johnson. *Acoustic and Auditory Phonetics*. 3. väljaanne. Malden, MA: Wiley-Blackwell, 2011.
- [40] Arvo Eek. *Eesti keele foneetika. I*. Toim. Pire Teras. Tallinn: Tallinna Tehnikaülikooli Kirjastus, 2008.

- [41] Eva Liina Asu *et al.* *Eesti keele hääldus*. Tartu: Tartu Ülikooli Kirjastus, 2016.
- [42] Einar Meister ja Arvo Eek. „Acoustics and Perception of Estonian Vowel Types“. Teoses: *PERILUS XVIII: Experiments in Speech Processes*. Toim. Mats Dufberg ja Olle Engstrand. Published. Stockholm: Stockholm University, 1994, lk. 55–90.
- [43] Paul Boersma ja David Weenink. *Praat: Sound To TextGrid (silences)*. 2025. URL: https://www.fon.hum.uva.nl/praat/manual/Sound__To_TextGrid__silences____.html (vaadatud 09.04.2025).
- [44] Apple Inc. *App Store Connect*. 2025. URL: <https://developer.apple.com/app-store-connect/> (vaadatud 16.04.2025).
- [45] Apple Inc. *Making the Most of the App Store*. 2025. URL: <https://developer.apple.com/app-store/> (vaadatud 16.04.2025).
- [46] Shua Dissen ja Joseph Keshet. *DeepFormants*. 2025. URL: <https://github.com/MLSpeech/DeepFormants> (vaadatud 02.05.2025).

Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks¹

Mina, Jaanus Tärnpuu

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose “Eesti keele hääldustreeningu rakendus”, 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.

12.05.2025

¹Lihtlitsents 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 - Vokaalide referentsväärtused

Gender	Vowel	F1 Ref	F1 Tol	F2 Ref	F2 Tol	F1 Bark	F1 Bark Tol	F2 Bark	F2 Bark Tol
FEMALE	a	742.6	55.07	1343.8	211.96	6.8	0.39	10.3	0.98
FEMALE	e	526.7	48.24	2284.0	256.37	5.2	0.40	13.7	0.82
FEMALE	i	425.7	70.78	2576.1	346.34	4.3	0.54	14.5	0.98
FEMALE	o	537.0	53.63	923.1	159.54	5.3	0.45	8.0	0.96
FEMALE	u	415.8	55.05	772.8	115.92	4.2	0.50	7.0	0.78
FEMALE	õ	486.1	52.65	1482.9	247.62	4.8	0.46	10.9	1.08
FEMALE	ä	769.1	85.01	1797.2	238.04	7.0	0.58	12.1	0.90
FEMALE	ö	510.1	34.85	1806.6	136.52	5.0	0.29	12.2	0.52
FEMALE	ü	402.7	62.82	1951.0	152.97	4.1	0.58	12.7	0.52
MALE	a	631.6	56.04	1087.7	130.35	6.0	0.43	9.0	0.71
MALE	e	445.0	70.59	1893.8	136.99	4.5	0.61	12.5	0.49
MALE	i	326.0	47.18	2161.1	150.54	3.4	0.45	13.4	0.45
MALE	o	495.7	66.63	930.8	159.25	4.9	0.56	8.0	0.95
MALE	u	383.1	53.86	850.3	145.56	3.9	0.49	7.5	0.93
MALE	õ	440.0	62.71	1208.8	187.27	4.4	0.55	9.75	0.87
MALE	ä	632.8	46.80	1506.6	138.50	6.0	0.36	11.0	0.61
MALE	ö	474.4	70.16	1500.3	112.17	4.7	0.60	11.0	0.49
MALE	ü	336.7	44.65	1729.3	154.95	3.5	0.41	11.9	0.60

Tabel 3. Andmekogumisrakenduse abil kogutud salvestuste põhjal arvatud referentsväärtused kasutades Praat skripti

Lisa 3 – Rakenduse GitLab lingid

Kõik rakendused on leitavad järgnevast repositooriumist: <https://gitlab.cs.taltech.ee/jaturn/phonex-audio>