

VILNIAUS UNIVERSITETAS
KAUNO HUMANITARINIS FAKULTETAS
INFORMATIKOS KATEDRA

Verslo informatikos studijų programa
Kodas 62409P104

TOMAS RASYMAS

MAGISTRO BAIGIAMASIS DARBAS

**BALSO KOMANDŲ ATPAŽINIMO METODAI, NAUDOJANT FONEMŲ
DISKRIMINAVIMĄ IR TRANSKRIPCIJŲ ADAPTAVIMĄ**

Kaunas, 2011

VILNIAUS UNIVERSITETAS
KAUNO HUMANITARINIS FAKULTETAS
INFORMATIKOS KATEDRA

Verslo informatikos studijų programa
Kodas 62409P104

TOMAS RASYMAS

MAGISTRO BAIGIAMASIS DARBAS

**BALSO KOMANDŲ ATPAŽINIMO METODAI, NAUDOJANT FONEMŲ
DISKRIMINAVIMĄ IR TRANSKRIPCIJŲ ADAPTAVIMĄ**

Leidžiama ginti _____

Magistrantas _____
(parašas)

Darbo vadovas _____
(parašas)

(darbo vadovo mokslo laipsnis, mokslo
pedagoginis vardas, vardas ir pavardė)

Darbo įteikimo data _____

Registracijos Nr. _____

Kaunas, 2011

TURINYS

PAVEIKSLŲ SĄRAŠAS	5
LENTELIŲ SĄRAŠAS	7
SANTRAUKA	8
ĮVADAS	9
1. KALBOS ATPAŽINIMO PROCESO APŽVALGA IR ANALIZĖ	11
1.1. Kalbos generavimas	11
1.2. Kalbos suvokimas	12
1.3. Kalbos atpažinimo sistemų architektūra	13
1.3.1 Pradinis kalbos signalo apdorojimas	13
1.3.2. Signalo požymių skaičiavimas	14
1.3.2.1. Tiesinis prognozavimas	14
1.3.2.2. Kepstrinė analizė	15
1.3.2.3. Spektrinė analizė	16
1.3.2.4. Vektorinis kvantavimas	17
1.3.4. Kalbos atpažinimo metodai	18
1.3.4.1. Akustiniai – fonetiniai metodai	18
1.3.4.2. Statistiniai kalbos atpažinimo metodai	19
1.3.4.3. Dirbtinio intelekto metodai	20
1.4. SAPI platformos apžvalga	21
2. ŠNEKOS IŠSKYRIMAS IŠ KALBOS SIGNALO	24
2.1. Pradinis kalbos signalo apdorojimas	24
2.2. Pradžios ir pabaigos nustatymas ilgame įrašė	25
2.3. Metodo tyrimo eksperimentas	26
3. LIETUVIŠKŲ KOMANDŲ ATPAŽINIMO, NAUDOJANTIS ANGLŲ KALBOS ATPAŽINIMO PRIEMONĖMIS, TYRIMAS	29
3.1. Programinės įrangos realizacija	31
3.2. Eksperimentinių duomenų sudarymas	31
3.3. Eksperimento rezultatų analizė	32
4. LIETUVIŠKŲ FONEMŲ PERRAŠYMO ANGLIŠKOMIS TAISYKLĖS	34
4.1. „ė“ ir „ch“ garsų aprašymo eksperimentas	35
4.2. „c“ ir „dz“ garsų aprašymo eksperimentas	38

4.3. „č“ ir „dž“ garsų eksperimentas.....	40
4.4. „ie“ ir „uo“ garsų eksperimentas	42
4.5. Mažos apimties žodyno, sudaryto pagal suformuluotas taisykles, atpažinimo eksperimentas	45
IŠVADOS	50
LITERATŪRA.....	52
PRIEDAI.....	54
1 Priedas. Magistro mokslinio darbo ruošimo forma	55
2 Priedas. Komandų išskyrimo metodo kokybės tikrinimo eksperimentų rezultatai	56
3 Priedas. Dalis gramatikos failo turinio.....	67
4 Priedas. Eksperimentinės programinės įrangos atpažinimo rezultatai.....	69

SANTRUMPŲ SĄRAŠAS

Hz – hercas, dažnio matavimo vienetas, reiškiantis (reiškinio, būsenos, padėties) pasikartojimų skaičių per sekundę;

kHz – kilohercas, dažnumo vienetas, lygus 1000 hercų;

SAPI – Microsoft produktas skirtas kalbos atpažinimui Speech Application Programming Interface;

GFT (angl. FFT) – greitoji Furjė transformacija;

DTF – diskrečioji Furjė transformacija;

Fonema – mažiausias kalbos garsinės sandaros vienetas;

TP – tiesinis prognozavimas.

PAVEIKSLŲ SĄRAŠAS

1 pav. Žmogaus balso trakto schema.....	11
2 pav. Žmogaus klausos sistema.....	12
3 pav. Supaprastinta kalbos atpažinimo sistemos architektūra.....	13
4 pav. Analoginis ir jį atitinkantis skaitmeninis signalai.....	14
5 pav. Kalbos tiesinis prognozavimo metodas.....	15
6 pav. Lygegretus filtrų bankas.....	16
7 pav. Persidengiančių juostinių filtrų banko schema.....	16
8 pav. Akustinių - fonetinių metodų blokinė schema.....	18
9 pav. Žodžio fonemų tinklelio pavyzdys.....	19
10 pav. Statistinių kalbos atpažinimo metodų blokinė schema.....	19
11 pav. Biologinio neurono pavyzdys.....	20
12 pav. Dirbtinio neurono pavyzdys.....	20
13 pav. Žinių integravimo kalbos atpažinimui schema.....	21
14 pav. SAPI naudojamo kalbos atpažinimo variklio veikimo schema.....	22
15 pav. Žodžio ribų nustatymo filtro taikymas.....	24
16 pav. Komandos ribų aptikimo algoritmas.....	25
17 pav. Žodžio "trys" spektograma su išskirtomis pradžios ir pabaigos ribomis.....	26
18 pav. Pateikiama žodžio "šeši"(viršuje) ir žodžio "du" (apačioje) spektogramos.....	27
19 pav. Pateikiama žodžio "septyni"(viršuje) ir žodžio "aštuoni" (apačioje) spektogramos.....	27
20 pav. Komandos viduryje pauzės identifikavimo algoritmas.....	28
21 pav. Programinės įrangos veikimo schema.....	31
22 pav. P. Kasparaičio pateikiamos perrašymo taisyklės.....	34
23 pav. Žodžių su „ė“ pasikartojimas procentais.....	36
24 pav. Žodžių su „ch“ pasikartojimas procentais.....	37
25 pav. Geriausių raidės „ė“ užrašymų reikšmės.....	37
26 pav. Geriausių raidės „ch“ užrašymų reikšmės.....	37
27 pav. Žodžių su „c“ pasikartojimas procentais.....	39
28 pav. Žodžių su „dz“ pasikartojimas procentais.....	39
29 pav. Geriausių raidės „c“ užrašymų procentas.....	40
30 pav. Geriausių raidės „dz“ užrašymo reikšmių procentas.....	40
31 pav. Žodžių su „č“ pasikartojimas procentais.....	41

32 pav. Žodžių su „dž“ pasikartojimas procentais.	41
33 pav. Geriausių raidės „č“ užrašymo reikšmių procentas.....	42
34 pav. Geriausių raidės „dž“ užrašymo reikšmių procentas.....	42
35 pav. Žodžių su „ie“ pasikartojimas procentais.	43
36 pav. Žodžių su „uo“ pasikartojimas procentais.	44
37 pav. Geriausių raidės „ie“ užrašymo reikšmių procentas.	44
38 pav. Geriausių raidės „uo“ užrašymo reikšmių procentas.	44
39 pav. Mažo žodyno atpažinimo tikslumas.	48

LENTELIŲ SĄRAŠAS

1 lentelė. SAPI atpažįstamų anglišku fonemų sąrašas.....	29
2 lentelė. Atpažinimo eksperimentiniai duomenys	31
3 lentelė. Geriausi komandų užrašymai.....	32
4 lentelė. Garsams „ė“ ir „ch“ tirti sudarytų žodžių aibė.....	35
5 lentelė. Garsams „c“ ir „dz“ tirti sudarytų žodžių aibė.....	38
6 lentelė. Garsams „č“ ir „dž“ tirti sudarytų žodžių aibė.....	41
7 lentelė. Garsams „ie“ ir „uo“ tirti sudarytų žodžių aibė	43
8 lentelė. Papildyta garsų perrašymo taisyklių lentelė	45
9 lentelė. Sudarytas mažos apimties žodynas	47

SANTRAUKA

RASYMAS, Tomas. (2011) *Voice commands recognition techniques using phonemic discrimination and pronunciation adaptation*. MBA Graduation Paper. Kaunas: Vilnaus Univesity, Kaunas Faculty of Humanities, Department of Informatics. 52 p.

SUMMARY

In the contemporary world of techniques, voice technologies, such as speech recognition, synthesis of speech signals and their combined versions, acquire more and more significance. These technologies are very popular for wide practical use: telecommunications, banking, army and so on.

The aim of the research paper is to analyze opportunity to adapt others languages speech recognition engine for Lithuanian language recognition using voice detection in speech signal and signal pre-emphasis methods. Microsoft SAPI English recognition engine was used for experiments in his paper.

P. Kasparaitis has offered how Lithuanian phonemes can be transcribed using English phonemes, but not all Lithuanian phonemes has been analyzed. So this paper also replenish P. Kasparaitis paper (P. Kasparaitis (2008). *Lithuanian Speech Recognition Using the English Recognizer.*) by offering new ways how some Lithuanian sounds can be transcribed by English phonemes.

The paper includes:

1. The analysis of general speech recognition methods;
2. General speech recognition methods advantages and disadvantages;
3. Overlook of Microsoft SAPI speech recognition engine;
4. Voice detection in continues speech signal method analysis and experiments;
5. Experiment of Lithuanian commands recognition using SAPI speech recognition engine;
6. Lithuanian sounds („ė“, „ch“, „c“, „č“, „dž“, „dz“, „ie“, „uo“) transcriptions experiments and generated rules.

IVADAS

Šiuolaikiniame technikos pasaulyje vis didesnę prasmę įgauna kalbos technologijos, tokios kaip kalbos atpažinimas, kalbos signalų sintezė, bei kombinuoti jų variantai. Pasaulyje visada buvo jaučiamas didelis poreikis kalbos atpažinimo sistemų, programų. Tokios sistemos labai populiarios dėl plataus savo pritaikymo spektro: karinė pramonė, įvairių prekių ar paslaugų užsakymų priėmimas, žmonių su įvairia negale darbo palengvinimas, telekomunikacijų, bankų sfera ir t.t.

Kalbos atpažinimas kelia skirtingus uždavinius kiekvienai kalbai, nes kiekviena kalba turi savo specifiką. Pastaruoju metu kitose šalyse yra sukurta daug įvairių atpažinimo sistemų, tačiau jos nėra naudojamos lietuvių kalbos atpažinimui. Todėl yra būtina eksperimentuoti bei kurti savai kalbai pritaikytas tokio tipo sistemas naudojant kitų kalbų atpažinimui pritaikytus produktus.

Darbo objektas:

Lietuviškų balso komandų atpažinimas naudojant kitų kalbų atpažinimo priemones.

Tyrimo tikslas:

Ištirti lietuviškų balso komandų atpažinimo galimybę, naudojant nelietuviškas atpažinimo priemones (Microsoft Speech Application Programming Interface).

Tyrimo uždaviniai:

1. Apžvelgti kalbos atpažinimo metodus;
2. Ištirti kalbos atpažinimo metodus, išskiriant jų privalumus ir trūkumus;
3. Apžvelgti kalbos atpažinimo paketą SAPI;
4. Atlikti komandų išskyrimo, iš kalbos signalo, metodo kiekybinę analizę;
5. Atlikti lietuviškų komandų atpažinimo eksperimentą, naudojantis SAPI varikliu;
6. Pateikti siūlymą kaip perrašyti lietuviškus garsus į angliškus, kad padidinti jų atpažinimą naudojantis SAPI varikliu.

Tyrimo problema:

Tyrimų, kuriuose būtų analizuojamas lietuviškos kalbos atpažinimas naudojant nelietuviškas atpažinimo priemones, trūkumas.

Mokslinis naujumas:

Šiame darbe buvo ištirta galimybė atpažinti lietuviško balso komandas naudojantis nelietuviškomis atpažinimo priemonėmis. Pasiūlytas metodas kaip atlikti lietuviškų balso komandų atpažinimą naudojantis SAPI paketu.

Naudoti metodai:

1. Rašant teorinę ir analitinę dalis remtasi moksline literatūra bei apibendrintais metodais;

2. Įgyvendinant eksperimentus – eksperimentiniu matavimu bei statistine analize.

Darbo struktūra ir apimtis:

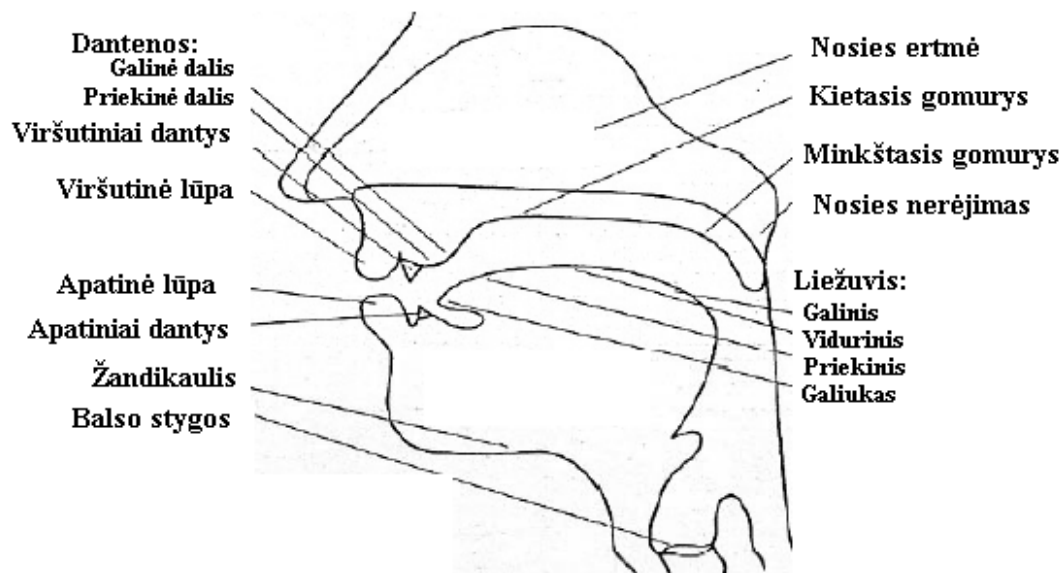
Darbą sudaro trys skyriai: teorinis – kuriame atliekama kalbos atpažinime naudojamų metodų apžvalga, analitinis – atliekamas tyrimas lietuviškų balsų komandų atpažinimui naudojantis SAPI paketu, tolimesniame skyriuje atliekami eksperimentai, norint įsitikinti balsų komandų atpažinimo galimybe.

Darbą sudaro 52 puslapių, 39 paveikslėliai, 9 lentelės ir 4 priedai.

1. KALBOS ATPAŽINIMO PROCESO APŽVALGA IR ANALIZĖ

1.1. Kalbos generavimas

Prieš kuriant kalbos atpažinimo sistemas, būtina suprasti žmogaus šnekos susidarymą. Šneka yra apibūdinama kaip oro molekulių susispaudimas laike, kurias generuoja žmogaus burna ir nosiaskylė (Huang, Acero, Hon, 2001, p. 24).



Šaltinis: sukurta autoriaus, pagal Xuedong Huang, Alex Acero, Hsiao-Wuen Hon (2001). Spoken language processing, p. 24.

1 pav. Žmogaus balso trakto schema.

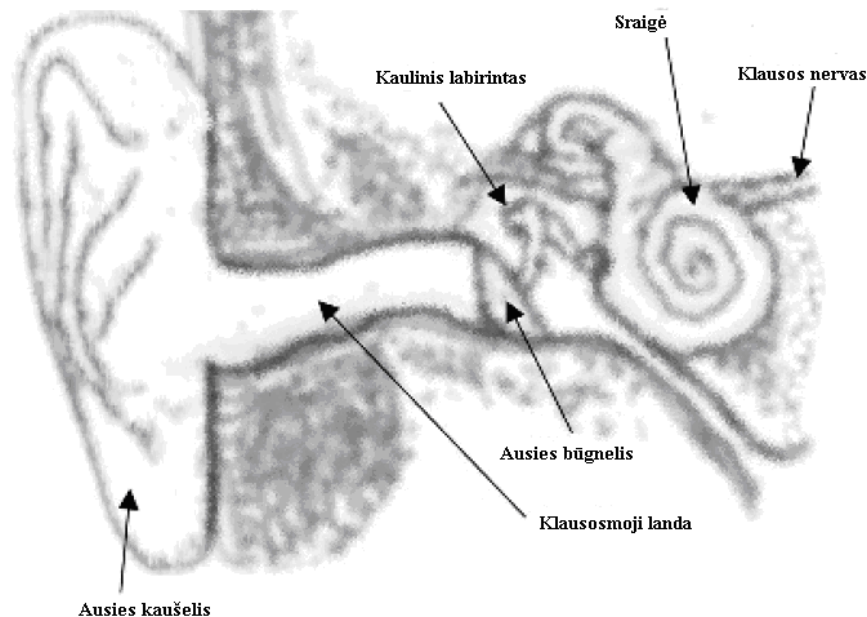
Taigi žmogaus balso traktas susideda iš (Huang, Acero, Hon, 2001, p. 25):

1. Plaučiai – oro šaltinis kalbėjimo metu;
2. Balso stygos – kai balso stygos yra įtemptos, sudaromi vokalizuoti garsai. Kai balso stygos yra laisvos, oro slėgis arba eina per susiaurėjimą balso trakte ir sudaro nevokalizuosius garsus, arba yra suslegiamas ir staigiai išleidžiamas sukuriant trumpalaikį garsą;
3. Minkštasis gomurys – veikia kaip vožtuvas, atsidaro sudarydamas oro koridorių per nosies ertmę;
4. Kietasis gomurys – ilgas, gana kietas paviršius, viršuje burnos, kuris kai liežuvis dedamas prieš jį generuoja priebalsinį skambėjimą;
5. Liežuvis – jo padėtis balso trakte nulemia skirtingus garsus;
6. Lūpos – gali būti suapvalintos, ištemptos, norint paveikti balsių skambėjimą arba visiškai uždarytos norint sustabdyti oro burnos sklidimą tam tikrais sąskambiais.

Apibendrinant kalbos susidarymo, visą procesą galima būtų aprašyti sekančiai: kalbos signalas yra sukuriamas pučiant iš plaučių oro srovę ir veikiant balso stygas. Oras judėdamas per balso stygas jas priverčia judėti ir sukelti tam tikrus slėgio svyravimus. Šie slėgio svyravimai pereina per keletą ertmių, kurios sustiprina arba susilpnina tam tikrų svyravimų dažnius. Pabaigoje iškvepiamas oras dar praeina per artikulacinius ir tada yra sukuriami periodiškai oro virpesiai (Huang, Acero, Hon, 2001, p. 25).

1.2. Kalbos suvokimas

Žmogui kalbant yra generuojami oro virpesiai, kuriuos priima žmogaus klausos sistema. Supaprastinta žmogaus klausos sistema pateikta apačioje.



Šaltinis: sukurta autoriaus, pagal Wai C Chu (2003). Speech Coding Algorithms Foundation and Evolution of Standardized Coders, p. 19.

2 pav. Žmogaus klausos sistema.

Ausies kaušelis yra paviršius supantis klausomąją landą į kurią yra nukreipiamos garso bangos. Garso bangos klausomąją landa keliauja iki ausies būgnelio – membrana, kuri elgiasi kaip signalo keitiklis. Garso bangos yra paverčiamos į mechaninius virpesius ir yra perduodamos į sraigę per kaulinį labirintą. Kaulinis labirintas pagerina garso sklaidimą. Garso bangos perėjusios kauliniu labirintu perduodamos į sraigėje esančią pagrindinę membraną. Ši membrana sukelia vidinės ausies virpėjimus, kurie per plaukuotąsias ląsteles transformuojami į nervinį impulsą. Gautą impulsą klausos nervai perduoda į smegenis, kur gauti impulsai yra analizuojami ir suvokiami kaip garsai (C Chu, 2003, p. 19).

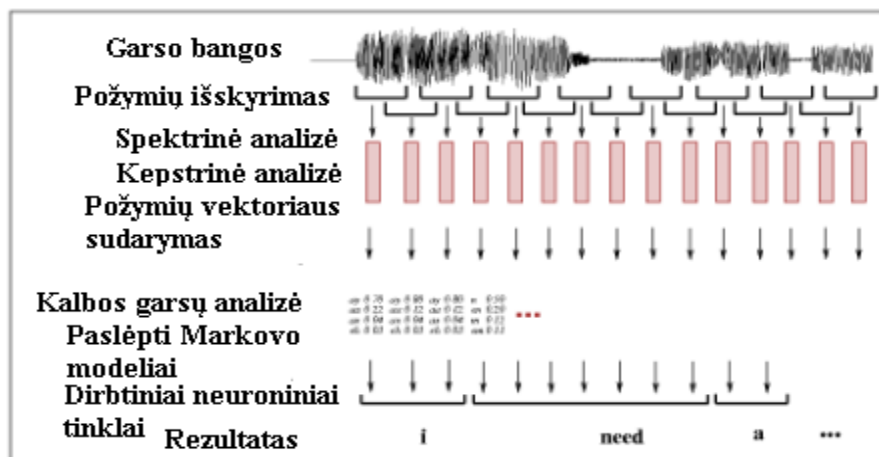
Žmogaus klausos sistema yra jautriausia garso dažnumui tarp 1 ir 4 kHz. Yra priimta, jog žemiau 20 Hz ir virš 20 kHz žmogaus klausos sistema visiškai nefunkcionuoja (C Chu, 2003, p. 19).

1.3. Kalbos atpažinimo sistemų architektūra

Kalbos atpažinimas – tai procesas, kai panaudojant įvairius matematinius skaičiavimus, leidžiama kompiuteriui užrašyti, atpažinti natūralią žmonių kalbą (Huang, Acero, Hon, 2001, p. 5).

Kalbos atpažinimo srityje mokslininkai dirba nuo 1940 – 1950 metų, tačiau rimtų produktų skirtų komerciniam naudojimui kol kas yra nedaug. Tai lemia didelis skirtingų kalbų skaičius pasaulyje, bei jų sudėtingumas (Jurafsky, Martin, 2000, p. 10).

Nors pasaulyje yra skirtingų kalbos atpažinimo sistemų, tačiau jų veikimo schemas yra labai panašios (Jurafsky, Martin, 2000, p. 235). Apačioje pateikiama supaprastinta kalbos atpažinimo sistemos architektūra.



Šaltinis: sukurta autoriaus, pagal Daniel Jurafsky, James H. Martin (2000). Speech and Language Processing, p. 239.

3 pav. Supaprastinta kalbos atpažinimo sistemos architektūra.

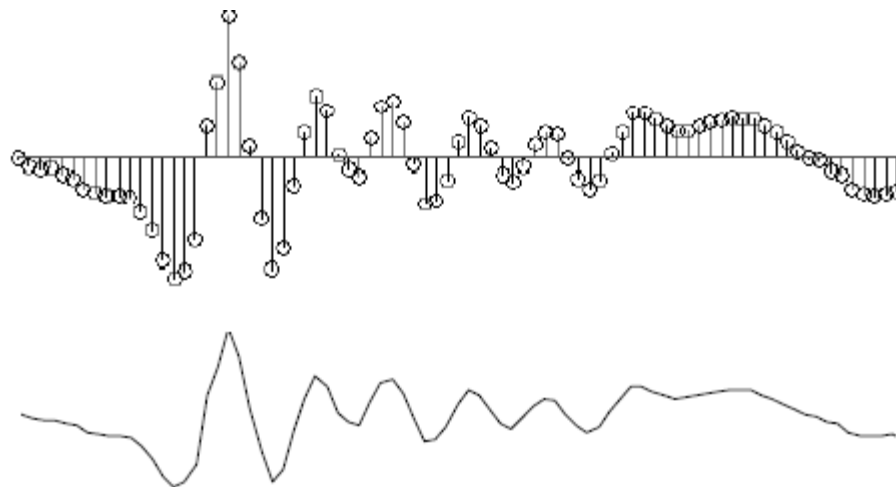
Analoginės garso bangos yra paverčiamos į skaitmeninį pavidalą, gauti duomenys suskirstomi į tam tikro ilgio langus, pjūvius (dažniausiai į 10, 15, 20 milisekundžių trukmės). Kiekvienam gautam langui atliekamas požymių išskyrimas, kuris pateikia informaciją apie tai kiek energijos yra signale skirtinguose dažniuose. Paskui naudojant tam tikrus metodus (pvz.: neuroniniai tinklai, paslėpti Markovo modeliai) atpažįstami individualūs kalbos garsai (pvz.: p arba b). Galiausiai gauti garsai yra lyginami su turimu žodžių tarimo žodynu ir ieškoma žodžių labiausiai atitinkančių turimus garsus (Jurafsky, Martin, 2000, p. 239).

Tai būtų apibendrintas kalbos atpažinimo sistemos veikimas. Tolimesniuose poskyriuose paanalizuosiu kiekvieną kalbos atpažinimo proceso dalį.

1.3.1 Pradinis kalbos signalo apdorojimas

Prieš atliekant kalbos atpažinimo procesą reikia atlikti signalo diskretizavimą, tai yra analoginio signalo pavertimą į skaitmeninį. Jei apibrėšime analoginį signalą $x_a(t)$, kaip funkciją besikeičiančią

laiką t ir sudalintume signalą x į T dalių (T – signalo pavyzdžio ėmimo dažnumas (sampling period), $t = nT$), tai skaitmeninį signalą galėtume užrašyti sekančiai: $x[n] = x_a(nT)$ (Huang, Acero, Hon, 2001, p. 202). Analoginio ir skaitmeninio signalų grafinė iliustracija pateikta apačioje.



Šaltinis: Xuedong Huang, Alex Acero, Hsiao-Wuen Hon (2001). Spoken language processing, p. 203.

4 pav. Analoginis ir jį atitinkantis skaitmeninis signalai.

Kaip matyti iš viršuje pateikto paveikslėlio kiekvieną sekundę iš analoginio signalo daug kartų imami pavyzdiniai duomenys kurie yra registruojami. Kuo dažniau duomenys yra imami tuo labiau skaitmeninis signalas bus artimesnis analoginiam.

1.3.2. Signalų požymių skaičiavimas

Turint kalbos signalo skaitmeninę formą, galima išskirti naudingus požymius, taikytinus kalbos atpažinime. Pagrindiniai kalbos signalų požymių skaičiavimo metodai – tiesinis prognozavimas, spektrinė ir keprinė analizės bei vektorinis kvantavimas (Rabiner, Juang, 1993, p. 69).

Tolimesniuose poskyriuose analizuojami dažniausiai praktikoje taikomi požymių skaičiavimo metodai.

1.3.2.1. Tiesinis prognozavimas

Tiesinio prognozavimo (TP) (Linear Predictive Coding) teorija žinoma jau senai. Tai vienas pirmųjų metodų pradėtų naudoti kalbos signalų analizei (Rabiner, Juang, 1993, p. 100). Pagrindinė TP metodo idėja yra ta, jog turima kalbos signalo diskreta (speech sample) laiko momentu n $s(n)$, gali būti aproksimuotas kaip tiesinė kombinacija prieš tai buvusių p kalbos signalo reikšmių:

$$s(n) \approx a_1 s(n-1) + a_2 s(n-2) + \dots + a_p s(n-p), \quad (1)$$

kur koeficientai a_1, a_2, \dots, a_p yra laikomi pastoviais visame analizuojamame signalo lange. Įtraukią į formulę papildomą narį $G_u(n)$ gauname tokią lygybę:

$$s(n) = \sum_{i=1}^p a_i s(n-i) + Gu(n), \quad (2)$$

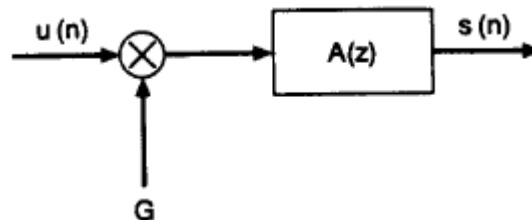
kur $u(n)$ yra normalizuotas žadinimas, o G yra žadinimo stiprėjimas. Jei iš priklausomybės (2) perrašytume z , tada gautume lygybę:

$$S(z) = \sum_{i=1}^p a_i z^{-i} S(z) + GU(z), \quad (3)$$

kuri yra pagrindinė perrašymo funkcijai:

$$H(z) = \frac{S(z)}{GU(z)} = \frac{1}{1 - \sum_{i=1}^p a_i z^{-i}} = \frac{1}{A(z)}. \quad (4)$$

Apacioje pateikiamas (4) išraiškos grafinė schema.



Šaltinis: sukurta autoriaus, pagal Lawrence Rabiner, Biing – Hwang Juang (1993). Fundamentals of speech recognition, p. 100.

5 pav. Kalbos tiesinis prognozavimo metodas.

Schemoje pateikta normalizuotas sužadinto šaltinis $u(n)$, suskaldytas pagal G stiprumą, veikiantis kaip įeiga $H(z) = \frac{1}{A(z)}$, kalbos signalo žadinimui $s(n)$ (Rabiner, Juang, 1993, p. 100).

Tiesinio prognozavimo metodas sueikvoja mažiau resursų lyginant su kitais metodais (Huang, Acero, Hon, 2001, p. 288).

1.3.2.2. Kepstrinė analizė

Spekto atvirkštinė logaritmo transformacija yra vadinama kepstru. Kepstrinė analizė taip pat yra labai naudinga konkreto konteksto stochastinių signalų parametrų gauti, ypač tada, kai spekto signalų stiprumas yra gaunamas iš autoregresinių modelių. Simbolinė kepstro išraiška yra tokia: $\text{Kepstras}(d) = \text{IFFT}(\log_{10}|\text{FFT}(s(n))|)$. Yra du kepstrų tipai: realieji (real cepstrum) ir kompleksiniai (complex cepstrum) (Huang, Acero, Hon, 2001, p. 305).

Realusis kepstras skaitmeniniam signalui $x[n]$ skaičiuojamas pagal formulę:

$$c[k] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln(|X(e^{j\Omega})|) e^{j\Omega k} d\Omega, \quad (5)$$

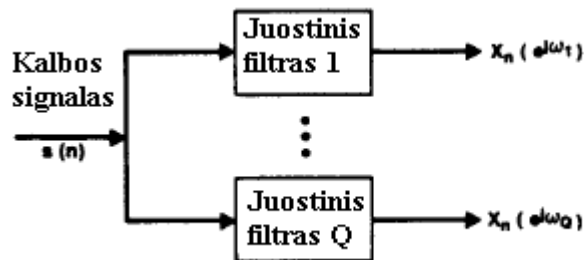
o kompleksiniam kepstrui skaičiuoti naudojamas sudėtingas logaritmas:

$$\hat{X}(e^{j\Omega}) = |X(e^{j\Omega})| + j\theta \Omega, \quad (6)$$

Signalas $x[n]$ yra tikras, taip pat tikrasis $c[n]$ ir sudėtingas $\hat{X}[n]$ kepsrai yra tikri signalai (Huang, Acero, Hon, 2001, p. 305).

1.3.2.3. Spektrinė analizė

Norint išgauti svarbius kalbos signalo parametrus įvairiose dažnių juostose yra plačiai naudojama spektrinės signalo analizė (Rabiner, Juang, 1993, p. 70). Apačioje pateikta filtrų bankų metodo veikimo schema.

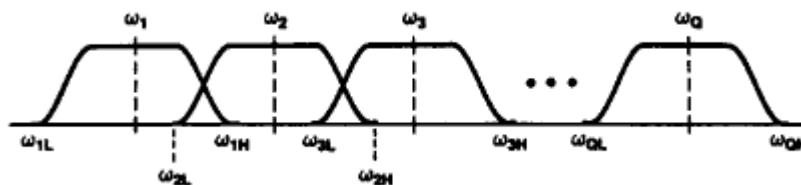


Šaltinis: sukurta autoriaus, pagal Lawrence Rabiner, Biing – Hwang Juang (1993). Fundamentals of speech recognition, p. 72.

6 pav. Lygegretus filtrų bankas.

Kalbos signalas $s(n)$ yra praleidžiamas per Q kiekį juostinių filtrų sudarytą juostų banką, kuris perdengia dominantį signalo dažnių diapazoną (pvz.: 100 – 3000 Hz telefoninio signalo kokybei, 100 – 8000 Hz plačiajuosčio radijo signalo kokybei) (Rabiner, Juang, 1993, p. 70).

Individualūs filtrai gali ir dažniausiai persidengia savo dažnių diapazonais (Rabiner, Juang, 1993, p. 72). Tai iliustruoja schema pateikta apačioje.



Šaltinis: Lawrence Rabiner, Biing – Hwang Juang (1993). Fundamentals of speech recognition, p. 72.

7 pav. Persidengiančių juostinių filtrų banko schema.

Galima matyti, jog juostinių filtrų bankų metode kiekvienas juostinis filtras apdoroja kalbos signalą individualiai, savarankiškai, norint pasiekti spektrinės analizės rezultatą $X(n)$. Praktikoje dažniausiai naudojami nevienodo pločio filtrų bankai, kas leidžia pagreitinti skaičiavimo procesą (Rabiner, Juang, 1993, p. 72).

Spektrinėje signalo analizėje taip pat naudojamas greitosios Furjė transformacijos metodas (GFT). GFT yra naudojama gauti energijos reikšmes laiko/dažnių tam tikrame taške (Rabiner, Juang,

1993, p. 72). Furjė algoritmų taikymas yra daug paprastesnis negu juostų bankų metodas, todėl praktikoje šis metodas yra dažniau naudojamas. Pagrindinė GFT idėja yra padalinti visą diskretinės Furjė transformacijos skaičiavimą į mažesnius skaičiavimus, tada atlikti šiuos mažesnius skaičiavimus ir tada sugeneruoti rezultata (Vary, Martin, 2006, p. 57). GFT koeficientų skaičiavimas yra pagrįstas diskrečios Furjė transformacijos lygybe:

$$X_{\mu} = \sum_{k=0}^{M-1} x(k) e^{-j \frac{2\pi \mu k}{M}}, \quad \mu = 0, 1, \dots, M-1. \quad (7)$$

Iš šios lygybės galima išsireikšti signalo $y(k)$ Furjė transformaciją trumpame laikotarpyje naudojant lango funkciją $v(k)$:

$$X_{\mu} = \frac{1}{M} \sum_{k=0}^{M-1} v(k) e^{-j \frac{2\pi \mu k}{M}}, \quad \mu = 0, 1, \dots, M-1. \quad (8)$$

Furjė transformacija leidžia atlikti signalo konvertacijas tiek iš laiko skalės į dažnių skalę ir atvirkščiai (Vary, Martin, 2006, p. 57).

1.3.2.4. Vektorinis kvantavimas

Vektorinis kvantavimas (VK) atlieka multidimencinį žymėjimą nuo šaltinio visumos iki atskirų dalių. VK Q pasižymi tuo, jog jis k – vektorius, x vektorinėje erdvėje R^k apjungia į baigtinį rinkinį $Y = \{y_i; i = 1, 2, \dots, N\}$, generuodamas klasterių seką pradedant nuo m ir kiekviename žingsnyje dalinant vieną arba daugiau pasirinktų klasterių į dvi dalis, taip didinant klasterių skaičių. Kiekvienas vektorius y_i vadinamas kodiniu vektoriumi:

$$Q = \{y_i\}_{i=1}^N, \quad (9)$$

$$x = [x_1, x_2, \dots, x_k]^T, \quad (10)$$

$$(y_1, y_2, \dots, y_k) \in Q, \quad (11)$$

$$y_i = [y_{i1}, y_{i2}, \dots, y_{iM}]^T, \quad i = 1, 2, \dots, N, \quad (12)$$

o visas tokių vektorių rinkinys vadinamas vektorine knyga:

$$Q(x) = \{y_i, i = 1, 2, \dots, N\} \quad (C \text{ Chu, 2003, p. 185}). \quad (13)$$

Vienas iš pagrindinių vektorių kvantavimo privalumų yra tas, jog šios technologijos taikymas leidžia maksimaliai sumažinti bitų skaičių reikalingą duomenims perduoti arba saugoti. Atliekant vektorinį kvantavimą tolydūs stebėjimo vektoriai išreiškiami kodinės knygos indeksais (Huang, Acero, Hon, 2001, p. 164).

1.3.4. Kalbos atpažinimo metodai

Metodus naudojamus kalbos atpažinime galima skirstyti į tris pagrindines grupes (Rabiner, Juang, 1993, p. 42):

1. Akustiniai – fonetiniai metodai;
2. Statistiniai kalbos atpažinimo metodai;
3. Dirbtinio intelekto metodai.

1.3.4.1. Akustiniai – fonetiniai metodai

Naudojant šiuos metodus kompiuteris nuosekliai dekoduoja įeinantį kalbos signalą pagal signalo akustinius požymius bei žinomais ryšiais tarp akustinių požymių ir fonetinių simbolių. Šie metodai buvo nuodugnai tyrinėjamas apie 40 metų. Tačiau dėl daugybės priežasčių, šis metodas nepasiekė tokios sėkmės praktiniuose pritaikymuose kaip alternatyvūs kalbos atpažinimo metodai. Akustiniai – fonetiniai metodai remiasi akustinės fonetikos teorija, jog šnekamojoje kalboje egzistuoja baigtiniai fonetiniai vienetai ir šie vienetai yra apibūdinami eilės savybių, kurios atsispindi kalbos signalo kitimas laike (Rabiner, Juang, 1993, p. 44).

Apačioje pateikiama akustinių – fonetinių metodų blokinė schema.



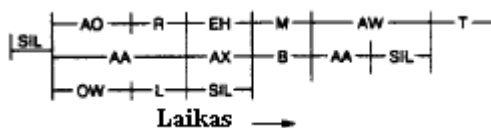
Šaltinis: sukurta autoriaus, pagal Lawrence Rabiner, Biing – Hwang Juang (1993). Fundamentals of speech recognition, p. 45.

8 pav. Akustinių - fonetinių metodų blokinė schema.

Kaip matyti iš paveikslėlio viršuje šiuose metoduose požymiai yra išskiriami naudojant filtrų bankus arba tiesinės prognozės metodus. O dažniausiai naudojami požymiai yra: formantės, pagrindinis tonas ir energija.

Pirmiausia kalbos signalas patenka į kalbos analizės sistemos modulį, kur iš laike kintančio kalbos signalo išgaunamos jo spektrinės charakteristikos. Vėliau gauti spektriniai matavimai paverčiami į požymių rinkinius. Paskui segmentavimo ir žymėjimo metodas sužymi regionus, kuriuose

gauti požymiai keičiasi labai mažai. Šio metodo rezultatas yra fonemų tinklelis (Rabiner, Juang, 1993, p. 43). Jo pavyzdys pateikiamas apačioje.



Šaltinis: sukurta autoriaus, pagal Lawrence Rabiner, Biing – Hwang Juang (1993). Fundamentals of speech recognition, p. 43.

9 pav. Žodžio fonemų tinklelio pavyzdys.

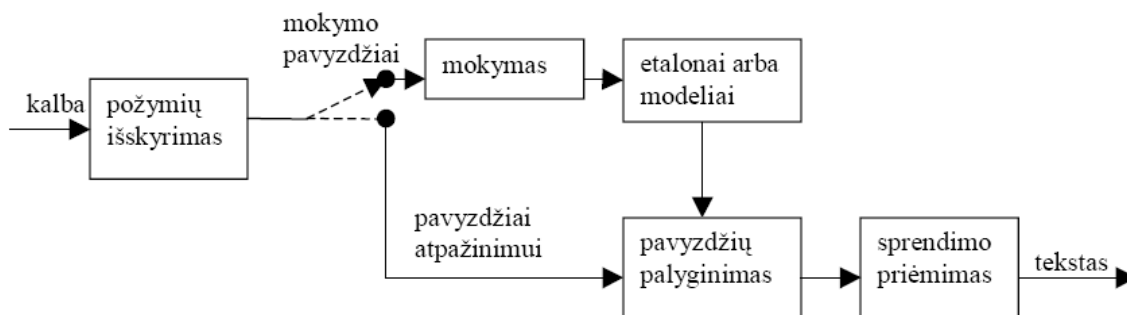
Paskui parenkamas žodis labiausiai atitinkantis turimų fonemų rinkinį ir galiausiai formuojama išeiga, kur pateikiamas atpažintas žodis arba žodžių seka (Rabiner, Juang, 1993, p. 43).

Akustinių – fonetinių metodų pagrindiniai trūkumai (Rabiner, Juang, 1993, p. 44):

1. Norint tinkamai naudotis metodu būtinos geros fonetikos teorijos žinios;
2. Kalbos signalo savybių išrinkimas ir signalo klasifikatorių formavimas nėra optimalūs.

1.3.4.2. Statistiniai kalbos atpažinimo metodai

Statistiniuose kalbos atpažinimo metoduose kalbos šablonų modeliai yra naudojami tiesiogiai (Rabiner, Juang, 1993, p. 51). Apačioje pateikiama statistinių kalbos atpažinimo metodų blokinė schema.



Šaltinis: sukurta autoriaus, pagal Lawrence Rabiner, Biing – Hwang Juang (1993). Fundamentals of speech recognition, p. 51.

10 pav. Statistinių kalbos atpažinimo metodų blokinė schema.

Požymių išskyrimo etape naudojami filtrų bankai, DFT arba TP. Paskui mokymo etape vykdomas mokymo procesas, kurio metu sukuriama etalonai. Visos žinios apie kalbą patenka apmokymo metu. Toliau vykdomas kalbos atpažinimas, pasakytų garsų lyginimas su turimais etalonais. Etalonai naudojami dinaminio laiko skalės kraipyme, o modeliai – paslėptuose Markovo modeliuose. Pavyzdžių palyginimo metu kalbos pavyzdys, kurį norim atpažinti, yra lyginamas su visais turimais pavyzdžiais ir parenkamas tas etalonas, kuris labiausiai atitinka įėjimo duomenis. Atlikus palyginimus pateikiamas atpažintas tekstas (Rabiner, Juang, 1993, p. 51).

Pagrindiniai statistinių kalbos metodų privalumai (Rabiner, Juang, 1993, p. 52):

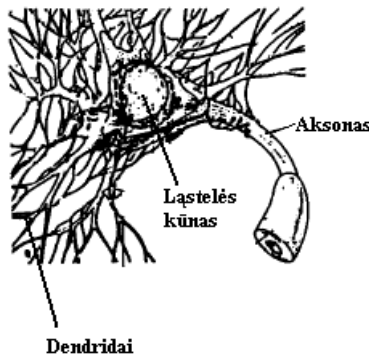
1. Tinka įvairiems garsams atpažinti;
2. Nereikalingos specifinės kalbos žinios;
3. Metodus galima taikyti norint atpažinti žodžių dalis, žodžius ar ištikus sakinius.

Pagrindiniai statistinių kalbos metodų trūkumai (Rabiner, Juang, 1993, p. 52):

1. Atpažinimo kokybė priklauso nuo mokymo duomenų skaičiaus;
2. Metodai yra jautrūs kalbos aplinkai.

1.3.4.3. Dirbtinio intelekto metodai

Dirbtinio intelekto metodai remiasi žmogaus organizme vykstančiais atpažinimo procesais ir bando analizuoti garso akustines savybes taip kaip tai darytų žmogaus smegenys. Šiuose metoduose bandoma kopijuoti biologinio neurono darbą (Rabiner, Juang, 1993, p. 54). Biologinio neurono pavyzdys pateikiamas apačioje.

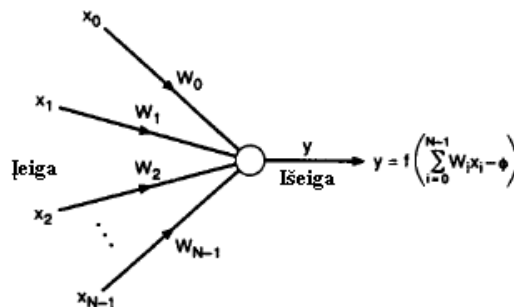


Šaltinis: sukurta autoriaus, pagal Lawrence Rabiner, Biing – Hwang Juang (1993). Fundamentals of speech recognition, p. 58.

11 pav. Biologinio neurono pavyzdys.

Į biologinį neuroną informacija patenka per dendridus. Kaip matoma iš paveikslėlio neuronas turi daug informacijos įėjimų, tačiau tik vieną išėjimą vadinamą aksonu.

Apačioje pateikiamas dirbtinio neurono pavyzdys.



Šaltinis: sukurta autoriaus, pagal Lawrence Rabiner, Biing – Hwang Juang (1993). Fundamentals of speech recognition, p. 57.

12 pav. Dirbtinio neurono pavyzdys.

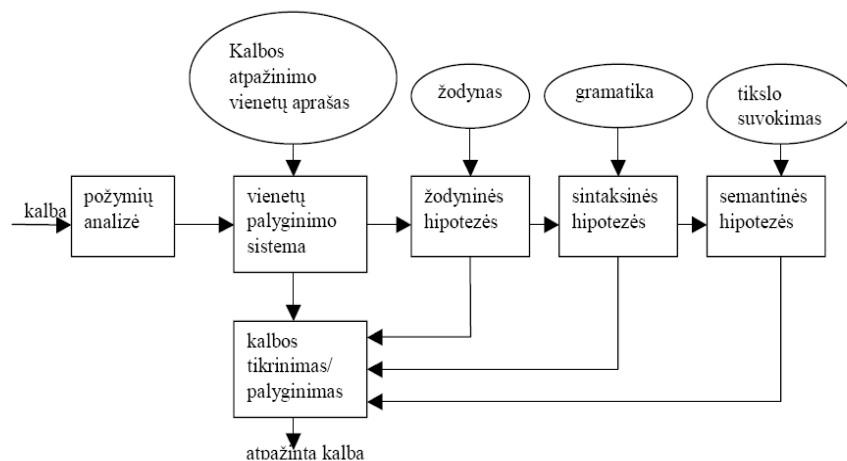
X_0, X_1, \dots, X_{N-1} yra neurono įėjimai, W_0, W_1, \dots, W_{N-1} – svoriniai koeficientai, išeiga Y yra apibrėžta kaip:

$$Y = f\left(\sum_{i=1}^N W_i X_i - \phi\right), \quad (14)$$

kur f yra svorinė funkcija, o ϕ – slenksčio reikšmė (Rabiner, Juang, 1993, p. 57).

Dirbtiniame neurone kiekvienas įėjimas turi savo svorinę reikšmę, o pats neuronas turi slenksčio reikšmę. Kai neuronas yra sužadinamas rezultatas gaunamas suskaičiuojant svorinę įėjimo signalų sumą ir iš jų atimant neurono slenksčio reikšmę. Pagal sužadinimo reikšmę taikant neurono perdavimo formulę yra apskaičiuojamas neurono išėjimas (Rabiner, Juang, 1993, p. 58).

Šiuose metoduose yra integruojama foneminė, žodyninė, semantinė ir pragmatinė informacija apie kalbą (Rabiner, Juang, 1993, p. 54). Tai iliustruoja schema esanti apačioje.



Šaltinis: sukurta autoriaus, pagal Lawrence Rabiner, Biing – Hwang Juang (1993). Fundamentals of speech recognition, p. 55.

13 pav. Žinių integravimo kalbos atpažinimui schema.

Šių metodų privalumai (Rabiner, Juang, 1993, p. 62):

1. Metodai atsparūs triukšmams;
2. Gali realizuoti didelius lygiagrečius skaičiavimus, kas paspartina atpažinimo procesą;
3. Neuroninių tinklų struktūrai būdinga prisitaikymo mokymosi koncepcija (adaptive learning), kas rodo, jog jų įėjimo reikšmių svorių reikšmės gali būti pertvarkomos realiame laike.

Šių metodų trūkumai (Rabiner, Juang, 1993, p. 62):

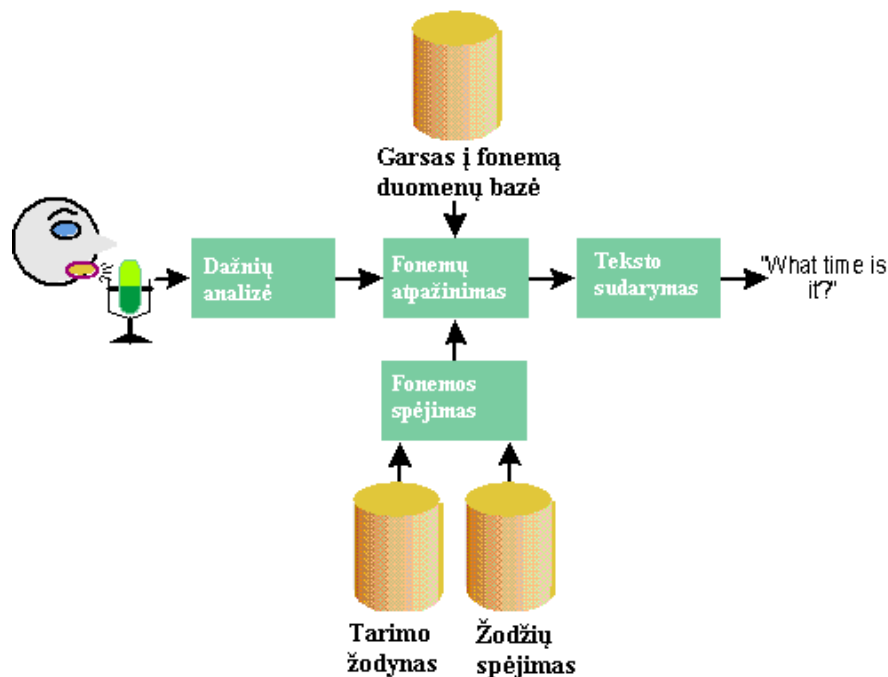
1. Sudėtingas bandymų kelias nustatant ir sudarant neuronų ryšių koeficientus.

1.4. SAPI platformos apžvalga

Speech Application Programming Interface (SAPI) yra Microsoft kompanijos sukurtas paketas leidžiantis palengvinti angliškos, ar kitų didžiųjų kalbų kalbos atpažinimo bei sintezavimo programų

kūrimą. Pati pirmoji SAPI versija buvo SAPI 1, kuri pristatyta 1995 metais. Šiuo metu yra dešimt šio paketo versijų (Rozak, 1996).

Apacioje pateiktas SAPI kalbos atpažinimo variklio veikimo schema.



Šaltinis: sukurta autoriaus, pagal Mike Rozak (1996). Talk to Your Computer and Have It Answer Back with Microsoft Speech API.

14 pav. SAPI naudojamo kalbos atpažinimo variklio veikimo schema.

Kai vartotojas kalba į mikrofoną, kompiuterio garso plokštė analoginį kalbos signalą paverčia į skaitmeninį. Skaitmeninis signalas toliau siunčiamas į dažnių analizės modulį, kuris gauna reikšmingą kalbos signalo informaciją. Paskui fonemų atpažinimo modulyje signalas išskaidomas į fonemas. Šis modulis ieško garsas į fonemą duomenų bazėje panašiausios fonemos į tą kurią išskyrė. Šioje duomenų bazėje kiekviena fonema turi šabloną, kuriame nurodoma kaip kiekviena fonema turi skambėti. Tačiau tinkamos fonemos ieškojimas šioje duomenų bazėje užtrunka ilgai, nes joje yra gausybė etalonų, tad norint paspartinti fonemų atpažinimą yra naudojamas fonemų spėjimo modulis. Žodžių spėjimo duomenų bazė naudojama sumažinti galimų tinkamų fonemų skaičiui, ji atmeta fonemas, kurios nesudarytų korektiškų žodžių. Kai fonemos yra atpažintos, jos yra siunčiamos į teksto sudarymo modulį, kuris sukonstruoja žodžius ir pateikia juos taikomajai programai (Rozak, 1996).

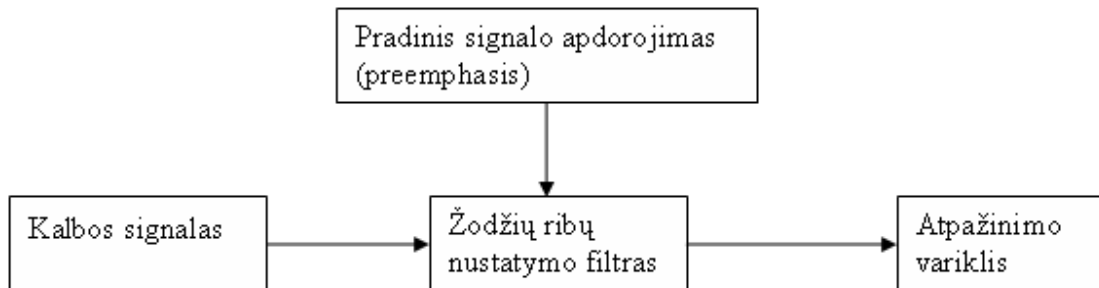
Kaip matoma iš SAPI kalbos atpažinimo variklio schemos, jį pritaikyti kitų kalbų atpažinimui yra įmanoma dviem būdais:

- 1) Perrašant SAPI naudojamų fonemų etalonus;
- 2) Perrašant lietuviškus garsus kitos kalbos garsais.

Pirmasis būdas yra sudėtingas, nes kompanija sukūrusi šią platformą nepateikia informacijos, kur ir kaip reiktų tą padaryti. Jei ir pateiktų, etalonų perrašymas būtų sunkus procesas, reikalaujantis daug laiko. Antras būdas yra daug efektyvesnis laiko ir kaštų sąnaudų atžvilgiu, tačiau tinkantis ne visoms kalboms.

2. ŠNEKOS IŠSKYRIMAS IŠ KALBOS SIGNALO

Automatinei kalbos atpažinimo sistemai paduoti visą įrašomą signalą yra neefektyvu nei kokybės, nei atpažinimo greičio atžvilgiu. Tam tikslui yra taikoma tarpinė grandis tarp signalo įėjimo iš mikrofono ir atpažinimo variklio, tai – komandų ribų nustatymo filtras (G.S. Ying, C.D. Mitchell, L.H. Jamieson, 2010).



Šaltinis: sukurta autoriaus.

15 pav. Žodžio ribų nustatymo filtro taikymas.

Analizuojamas bus ribų nustatymo filtras, pagrįstas šnekos energetinio mato charakteristikomis. Kalbos signalo energija yra dažniausiai taikoma signalų segmentavimui, kai reikia atskirti naudingą informaciją nuo pašalinio triukšmo, kadangi aplinkos triukšmo energija yra mažesnė negu kalbos signalo energija (Kiril Sakhnov, Boris Simak, 2009).

2.1. Pradinis kalbos signalo apdorojimas

Norint gauti tikslesnius komandų išskyrimo duomenis visų pirma yra taikomas pradinio kalbos signalo apdorojimo filtras.

Pradinis kalbos signalo apdorojimas (preemphasis) yra taikomas norint aplyginti kalbos signalo spektrą, kad signalo spektro energija būtų pasiskirsčiusi visame dažinių diapazone tolygiai. Kalbos signalo spektro išlyginimui yra naudojamas pirmos eilės ribotos impulsinės reakcijos filtras:

$$H(z) = -az^{-1} . \quad (15)$$

Laiko srityje ši funkcija yra išreiškiamą:

$$\tilde{S}(n) = s(n) - as(n-1) , \quad (16)$$

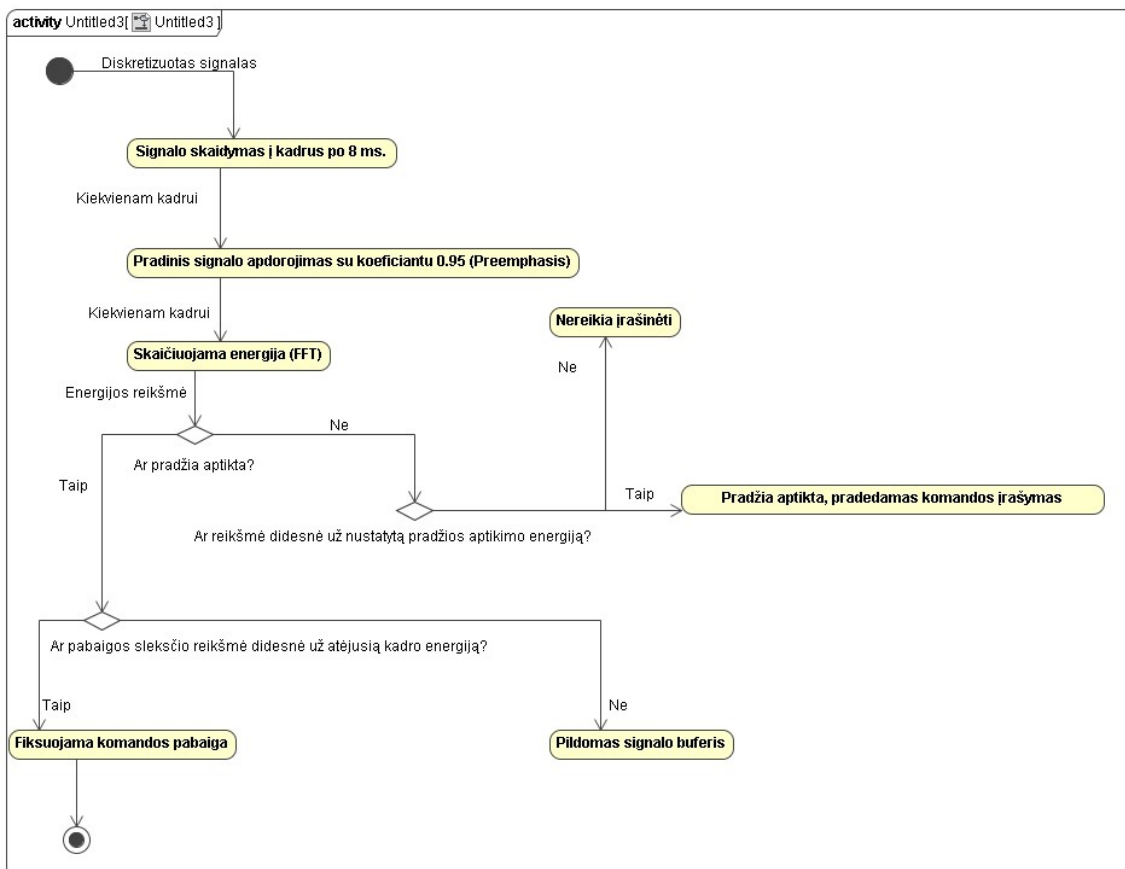
čia $\tilde{S}(n)$ – apdorojimo signalo išėjimo signalas, $s(n)$ – diskretizuoto kalbos signalo reikšmė, a – koeficientas, nurodantis spektro aplyginimo laipsnį, dažniausiai taikomas 0,95.

Pagrindinė priežastis kodėl yra naudojamas šis filtras: naudingą informaciją turinti kalbos signalo dalis turi neigiamą nuožulnumą ir yra susikoncentravusi žemų dažinių srityje ir tai sukelia didesnes baigtinio tikslumo paklaidas signalo apdorojime (Joseph Picone, 1993).

2.2. Pradžios ir pabaigos nustatymas ilgame įrašė

Naudojamo ribų nustatymo filtro pagrindinė funkcija yra aptikti sakomų komandų pradžią ir pabaigą. Aptikimas paremtas energijos stebėjimu atskiruose signalo pjūviuose.

Ateinantis signalas skaidomas į 8 ms. kadrus iš kurių skaičiuojamos energijos vertės naudojantis FFT algoritmu. Kiekvieno pjūvio gauta energija yra lyginama su nustatytais pradžios ir pabaigos energijomis, kurios identifikuoja komandos pradžią ir pabaigą. Kai pjūvio energija viršija slenksčio reikšmę identifikuojama kalbos pradžia, kai energija nukrinta žemiau pabaigos slenksčio ribos identifikuojama kalbos pabaiga. Visa tai galima pavaizduoti schema pateikta apačioje (Mark Marzinzik, Birger Kollmeier, 2002).

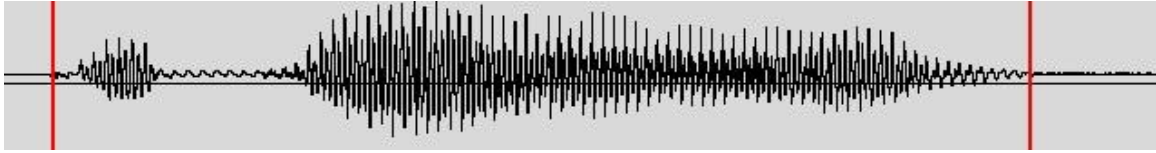


Šaltinis: sukurta autoriaus.

16 pav. Komandos ribų aptikimo algoritmas.

Pritaikius šį algoritmą vientisam kalbos signalui gaunamas tik kalbos signalas su informacija, kuri bus apdorota tolimesniuose etapuose.

Pateikiamas spektogramos vaizdas su nurodytomis komandos pradžia ir pabaiga:



Šaltinis: sukurta autoriaus.

17 pav. Žodžio "trys" spektrograma su išskirtomis pradžios ir pabaigos ribomis.

2.3. Metodo tyrimo eksperimentas

Eksperimentui įgyvendinti buvo sukurta speciali programinė įranga, kuri veikia pagal pateiktą algoritmą.

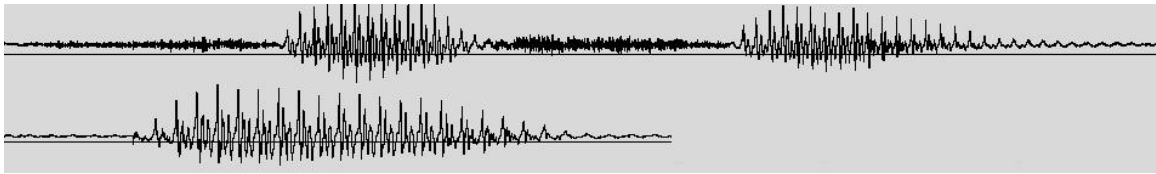
Pasirinkus 4 skirtingo amžiaus ir lyties asmenis, kad būtų modeliuojamos skirtingos tarties savybės, buvo atliekamas jų balso įrašymas triukšmingoje aplinkoje. Eksperimente triukšmai buvo suskirstyti į tris grupes: mažas triukšmas – nuo 10 iki 15 dB, vidutinis – nuo 15 iki 20 dB ir aukštas – nuo 20 iki 30 dB. Eksperimento dalyviai sakė komandas: nulis, vienas, du, trys, keturi, penki, šeši, septyni, aštuoni, devyni, dešimt. Sukurta programinė įranga, kuri turėjo apdoroti įeinantį ilgą, tačiau su pauzėmis, sakomų komandų signalą, išskirdavo tik komandas, be pašalinio triukšmo ir pauzių. Eksperimente buvo tikrinama metodo kokybė esant skirtingoms energijos slenksčių reikšmėms. Naudojami slenksčiai (pradžios – pabaigos): 25 – 15 dB; 30 – 25 dB; 35 – 25 dB ir 35 – 30 dB. Gauti eksperimento rezultatai yra pateikti antrame priede.

Apibendrinus rezultatus galima teigti, kad analizuojamo metodo naudojimas duoda geriausius rezultatus kai aplinkos triukšmas svyruoja diapazone 0 – 20 dB, o slenksčių reikšmės yra +/- 20 – 30 dB. Taip pat nustatyta, jog šio metodo naudojimas nėra labai efektyvus, jeigu triukšmas svyruoja labai dideliame diapazone. Esant didesniai triukšmui yra nepatogu įvedinėti komandas į sistemą, nes vartotojui reikia panaudoti daug energijos, norint viršyti aplinkos triukšmą.

Eksperimento rezultatai parodė, jog tiksliausia komandų pradžia aptinkama, kai ji prasideda sprogstamuoju priebalsiu. Sprogstamieji priebalsiai yra: „B“, „D“, „G“, „P“, „K“, „T“. Sprogstamieji priebalsiai yra gana sudėtingos artikuliacijos garsai, todėl jų savaiminės trukmės svyravimo intervalas yra nedidelis: 15 – 20 ms. (Asta Kazlauskienė, Gailius Raškinis, 2005). Tariant sprogstamąjį priebalsį burnoje susidaro tam tikra užtvara, kuri neleidžia garsui patekti į aplinką, o kai užtvara atidaroma yra iššaukiamas trumpas, tačiau turintis gan didelę energiją garsas, tad žodžių su šiais priebalsiais aptikimas yra gan tikslus.

Frikatyviniu priebalsiu prasidedančią komandą aptikti yra sunkiau, nes signalas laiko juostoje palaipsniui įgauna vis didesnę energiją ir yra užfiksuojamas tik tada kai pasiekia nustatytą ribą. Tad iš eksperimento rezultatų matyti, jog šių priebalsių aptikimas yra su didesne paklaida negu sprogstamųjų.

Palyginimui apačioje pateikiama komandų prasidedančių sprogstamuoju ir friktyviniu priebalsiu spektograma.

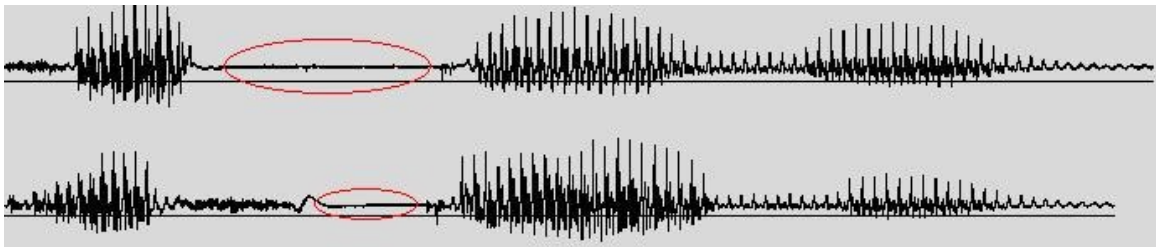


Šaltinis: sukurta autoriaus.

18 pav. Pateikiama žodžio "šeši"(viršuje) ir žodžio "du" (apačioje) spektogramos.

Kaip matyti iš paveikslėlio friktyvinių priebalsių aptikimas yra sudėtingesnis. Pabaigų nustatymas vykdomas gan efektyviai kai triukšmas neviršija nustatytos slenksčio reikšmės. Kadangi komandų pabaigoje dažniausiai pasitaiko friktyvinis priebalsis „S“ arba pusbalsis „I“, kurie laiko kitimo skalėje mažina savo energiją palaipsniui, tad maža dalis kalbos signalo pabaigos yra nukandama. Tačiau tai didelės įtakos atpažinimui neturi.

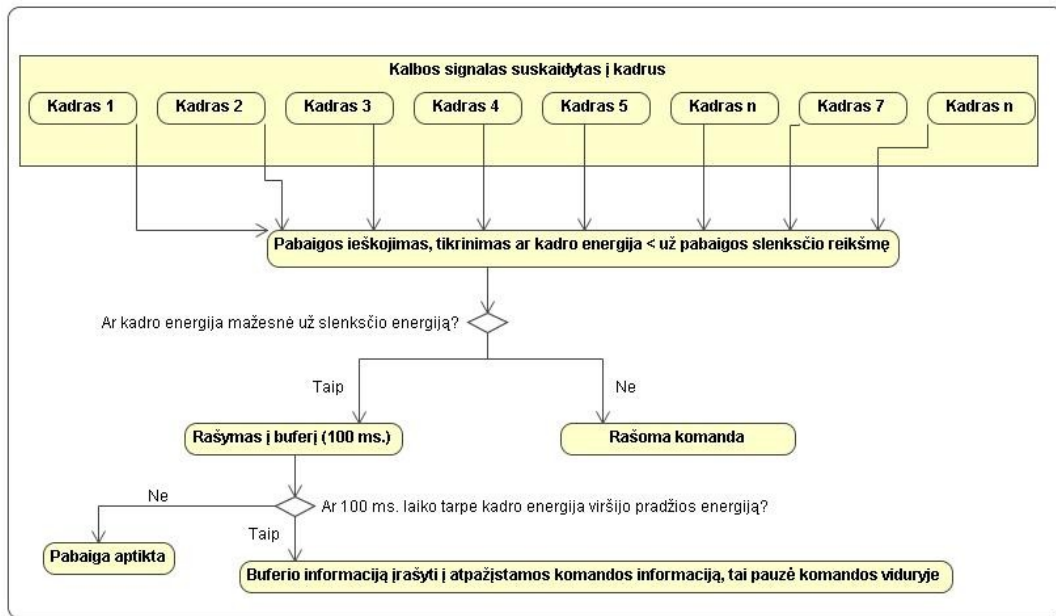
Taip pat pastebėta, jog metodas veikia neefektyviai kai viduryje komandos užfiksuojama pauzė: „sep_tyni“, „aš_tuoni“ ir pan. Tada kalbos energija nukrenta žemiau komandos pabaigos slenksčio ribos ir yra identifikuojama komandos pabaiga. Tai pavaizduota paveikslėlyje apačioje.



Šaltinis: sukurta autoriaus.

19 pav. Pateikiama žodžio "septyni"(viršuje) ir žodžio "aštuoni" (apačioje) spektogramos.

Sukurtoje eksperimentinėje programinėje įrangoje šis metodo trūkumas yra eliminuojamas naudojant papildomus buferius, kuriuose saugoma prieš tai buvusių kadro reikšmės. Algoritmas pateikiamas apačioje.



Šaltinis: sukurta autoriaus.

20 pav. Komandos viduryje pauzės identifikavimo algoritmas.

Paskaičiuota vidutinė metodo daroma paklaida: 10 – 50 ms. esant triukšmo lygiui 0 – 20 dB ir slenksčio reikšmėms 20 – 30 dB. Tai yra gan geras rezultatas ir metodo naudojimas, automatiniai atpažinimo sistemai, kuri veiks aplinkoje su panašiu triukšmu, yra tinkamas. Taip pat svarbu paminėti, jog metodo kokybė taip pat priklauso nuo asmens kalbėjimo ypatybių.

Papildant metodą automatiniais komandų išskyrimo slenksčių skaičiavimais, kai slenksčiai nustatomi pagal aplinkos ypatybes, metodą galima pritaikyti triukšmingoms aplinkoms (Kiril Sakhnov, Boris Simak, 2009).

3. LIETUVIŠKŲ KOMANDŲ ATPAŽINIMO, NAUDOJANTIS ANGLŲ KALBOS ATPAŽINIMO PRIEMONĖMIS, TYRIMAS

Šiam tyrimui atlikti buvo sukurta programinė įranga, kurios pagalba būtų galima ištirti SAPI atpažinimo variklio galimybes, atpažinti lietuviškos kalbos tam tikrus žodžius. Atpažįtamų žodžių bazė sudaryta iš šių žodžių: nulis, vienas, du, trys, keturi, penki, šeši, septyni, aštuoni, devyni. Tyrime dalyvavo 5 asmenys, skirtingo amžiaus, bei lyties, turintys skirtingas balso savybes. Kadangi šiuo atveju yra tiriama SAPI atpažinimo kokybė, aiškinamasi tinkamiausios garsų transkribavimo taisyklės, tad į atpažinimo variklį paduodamas ne kalbos signalas įvedamas iš mikrofono, o jau apdorotas, t.y. iš kalbos signalo išskirta komanda (WAV failas), nurodant jo vietą kietajame diske, taip yra išvengiama skirtingų įvedimo duomenų, susikoncentruojama ties tinkamiausių garsų užrašymo analize. Kiekviena kiekvieno tyrimo dalyvio komanda į atpažinimo variklį paduodama po 10 kartų. Priklausomai nuo analizuojamo žodžio parenkamas atitinkamas gramatikos failas, kuriame yra aprašyta to pačio žodžio keli variantai pvz.: vienas, vhienas ir t.t. Yra analizuojama ryšys tarp lietuviškų ir anglišių fonemų.

Kadangi ne visi lietuviški žodžiai turi panašius atitikmenis anglų kalboje, gramatikos faile žodžius galima aprašyti ir kaip anglišių fonemų rinkinius, kurie skambėtų panašiai kaip lietuviški (P. Kasparaitis, 2008). Jei žodis turi skambėjimo atitikmenį anglų kalboje jis gramatikos faile aprašomas taip:

```
<RULE NAME="Taip" ID="RID_Taip" TOPLEVEL="ACTIVE">
  <P>Type</P>
</RULE>
```

Kiti žodžiai, kurie neturi garsinių atitikimų anglų kalboje yra užrašomi kaip fonemų rinkiniai:

```
<RULE NAME="Ka" ID="RID_Ka" TOPLEVEL="ACTIVE">
  <P PRON="c aa">Ką?</P>
</RULE>
```

SAPI variklis atpažįsta 49 angliškas fonemas (Microsoft SAPI pagalbos dokumentas), jų sąrašas pateikiamas apačioje.

1 lentelė.

SAPI atpažįtamų anglišių fonemų sąrašas

Simbolis	Pavyzdys	Fonemos ID
-	syllable boundary (hyphen)	1
!	Sentence terminator (exclamation mark)	2
&	word boundary	3
,	Sentence terminator (comma)	4
.	Sentence terminator (period)	5
?	Sentence terminator (question mark)	6
_	Silence (underscore)	7
1	Primary stress	8
2	Secondary stress	9

1 lentelės tęsinys

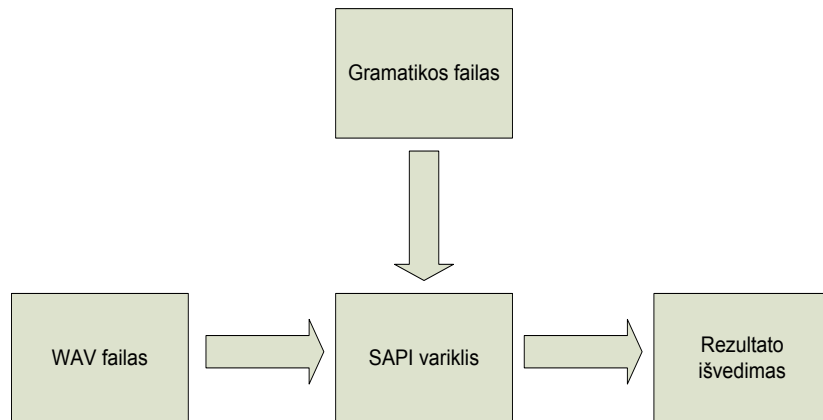
Simbolis	Pavyzdys	Fonemos ID
aa	f <u>a</u> ther	10
ae	c <u>a</u> t	11
ah	c <u>a</u> t	12
ao	d <u>o</u> g	13
aw	f <u>o</u> ul	14
ax	ag <u>o</u>	15
ay	b <u>i</u> te	16
b	<u>b</u> ig	17
ch	<u>ch</u> in	18
d	<u>d</u> ig	19
dh	<u>th</u> en	20
eh	pe <u>t</u>	21
er	f <u>u</u> r	22
ey	at <u>e</u>	23
f	<u>f</u> ork	24
g	<u>g</u> ut	25
h	<u>h</u> elp	26
ih	<u>fi</u> ll	27
iy	fe <u>el</u>	28
jh	jo <u>y</u>	29
k	<u>k</u> ut	30
l	<u>l</u> id	31
m	<u>m</u> at	32
n	<u>n</u> o	33
ng	si <u>ng</u>	34
ow	g <u>o</u>	35
oy	to <u>y</u>	36
p	<u>p</u> ut	37
r	<u>r</u> ed	38
s	<u>s</u> it	39
sh	<u>sh</u> e	40
t	<u>t</u> alk	41
th	<u>th</u> in	42
uh	bo <u>o</u> k	43
uw	to <u>o</u>	44
v	<u>v</u> at	45
w	<u>w</u> ith	46
y	<u>y</u> ard	47
z	<u>z</u> ap	48
zh	plea <u>su</u> re	49

Šaltinis: Microsoft SAPI pagalbos dokumentas.

Nors yra 58 lietuviškos fonemos (A. Girdenis, 1995), dauguma garsų galima sėkmingai aprašyti angliškomis fonemomis.

3.1. Programinės įrangos realizacija

Eksperimentinės programinės įranga realizuota C# .NET programavimo kalba. Programinė įranga yra sukurta atskirai nuo pirmame eksperimente naudotos. Programinės įrangos veikimo schema pateikiama apačioje.



Šaltinis: sukurta autoriaus.

21 pav. Programinės įrangos veikimo schema.

Pirmiausia nuskaitytas jau paruoštas WAV failas su tam tikra ištarta komanda, nuskaičius informacija paduodama atpažinimo varikliui, kuris prieš vykdydamas atpažinimą užsikrauna gramatikos failą, kuriama yra aprašytos atpažinimo taisyklės. Turėdamas taisykles ir reikiamą atpažinti garsinę informaciją atpažinimo variklis atlieka palyginimus ir rezultatą išveda į ekraną.

3.2. Eksperimentinių duomenų sudarymas

Eksperimentiniai duomenys sudaromi pagal žodžių skambėjimo panašumą į angliškus žodžius, arba į atskiras fonemas. Sudarytų eksperimentinių duomenų aibės pateikiamos apačioje.

2 lentelė.

Atpažinimo eksperimentiniai duomenys

Nulis	Vienas	Du	Trys	Keturi	Penki	Šeši	Septyni	Aštuoni	Devyni
Nuhlis	Viehnas	Do	Trees	Kehtoori	P eh n ih	Sh eh sh ih	Septeeni	Ae sh th uh ao n ih	D eh v iy n ih
N uh l ih s	V ih ey n ah s	D uh	Treas	K eh t uh r ih	P eh n k ih	Sh ey sh ih	S eh p th ey n ih	Ae sh th uh ao ng ih	D eh v iy n iy
Noolis	Viehnahs	Doo	T r iy s	K eh t ow r ih	P ey n k ih	Sheshe	S eh p t ey n ih	Ah sh t uh n ih	Dh eh v iy n ih
Nolihs	V ey n ah s	D uw	Th r iy s	K ey th uh r ih	P ey n k iy	Sheshy	S eh p t iy n ih	Ah sh th uh ao n ih	Dh eh v iy n iy

2 lentelės tęsinys

Nulis	Vienas	Du	Trys	Keturi	Penki	Šeši	Septyni	Aštuoni	Devyni
N u w l i h s	V e y n g a h s	Dh uh	Thre e s	K e y t u h r i h	Pencki	Shishi	S e h p t h i y n i h	Ashtooni	Deveeni
	Veehnas	Dho	Threa s	Ketoori			S e h p t h i y n i	Ahshtuooni	D e h v i h n i h
		Dow					S e h p t e y n i		

Šaltinis: sukurta autoriaus.

Naudojantis sukurtais eksperimentiniais duomenimis yra sudaromas gramatikos failas, kuris bus naudojamas SAPI atpažįstant ištartas komandas. Dalis gramatikos taisyklių yra pateikiama trečiame priede.

Sukūrus gramatikos failą, atliekami eksperimentai, analizuojama geriausi komandų užrašymo būdai.

3.3. Eksperimento rezultatų analizė

Atlikto eksperimento rezultatai pateikiami 4 priede. Iš gautų rezultatų galima matyti, jog iš visų, tam tikros komandos, aprašymui naudotų išraiškų, dažniausiai panaudota būna kelios, priklausomai nuo žmogaus balso savybių.

Žemiau esančioje lentelėje yra pateikiama eksperimento rezultatai, kur nurodoma komanda ir ją labiausiai atitinkantis užrašymas SAPI gramatikos faile.

3 lentelė.

Geriausi komandų užrašymai

Komanda/ Užrašymas			
Nulis	N u h l i h s	Nolihs	
Vienas	Viehnas	Viehnahs	V e y n g a h s
Du	Dho	Dow	
Trys	T r i y s	Threes	
Keturi	K e h t u h r i h	K e y t h u h r i h	
Penki	P e y n k i h		
Šeši	S h e y s h i h		
Septyni	S e h p t i y n i h		
Aštuoni	A h s h t u h n i h	A h s h t h u h a o n i h	Ashtooni
Devyni	D e h v i y n i h	D h e h v i y n i h	

Šaltinis: sukurta autoriaus.

Sukurta eksperimentinė programinė įranga bus naudojama atlikti eksperimentus perrašymo taisyklių kūrimui. Atpažįstamas žodynas bus praplėstas ir stebimas programinės įrangos veikimas ir klaidų kaičius. Dėl praplėsto žodyno pailgės atpažinimo greitis, nes reiks apdoroti daugiau duomenų.

tad sprendimas bus priiminėjamas ilgiau. Taip pat dėl šių savybių, toks metodas negali būti taikomas ten kur reikalinga atpažinti didelę žodžių aibę. Mažam kiekiui žodžių atpažinti, šis metodas yra tinkamas ir veikia efektyviai.

4. LIETUVIŠKŲ FONEMŲ PERRAŠYMO ANGLIŠKOMIS TAISYKLĖS

Pagrindinė problema plėtojant lietuviškos kalbos atpažinimo sistemas, kurios paremtos kitų kalbų atpažinimo varikliais, yra ta, jog nėra sudarytos patikimos perrašymo taisyklių aibės, kuria būtų galima remtis kuriant sistemas. Taip pat yra tam tikrų lietuviškų fonemų, kurių atitikmenų angliškose fonemose nėra.

P. Kasparaitis savo darbe (P. Kasparaitis (2008). *Lithuanian Speech Recognition Using the English Recognizer.*) pateikia tam tikras perrašymo taisykles. Tačiau darbe minimi ir kai kurie atviri klausimai į kuriuos nėra pateikiami atsakymai:

- 1) Kokios angliškos fonemos geriausiai tinka aprašyti garsus „ė“ ir „ch“;
- 2) Kaip geriau aprašyti garsus „c“ ir „dz“, viena fonema ar keliomis;
- 3) Kaip geriau aprašyti garsus „č“ ir „dž“, viena fonema ar keliomis;
- 4) Kokios angliškos fonemos geriausiai tinka aprašyti garsus „ie“ ir „uo“.

Savo darbe P. Kasparaitis pateikia tokias fonemų perrašymo taisykles:

Lithuanian phonemes	Lithuanian letters	English phoneme	Example	Lithuanian phonemes	Lithuanian letters	English phoneme	Example
1 /a/	“a”	[ah]	cut	19 /ts/, /ts’/	“c”	–	–
2 /e/	“e”	[eh]	pet	20 /dz/, /dz’/	“dz”	–	–
3 /i/	“i”	[ih]	fill	21 /tʃ/, /tʃ’/	“č”	[ch] ³	chin
4 /o/, /o:/	“o”	[ao] ¹	dog	22 /dʒ/, /dʒ’/	“dž”	[jh] ⁴	joy
5 /u/	“u”	[uh]	book	23 /s/, /s’/	“s”	[s]	sit
6 /a:/	“a”, “ą”	[aa]	father	24 /z/, /z’/	“z”	[z]	zap
7 /e:/	“e”, “ę”	[ae]	cat	25 /ʃ/, /ʃ’/	“š”	[sh]	she
8 /i:/	“i”, “y”	[iy]	feel	26 /ʒ/, /ʒ’/	“ž”	[zh]	pleasure
9 /u:/	“u”, “ū”	[uw]	too	27 /x/, /x’/	“ch”	–	–
10 /ė:/	“ė”	–	–	28 /h/, /h’/	“h”	[h]	help
11 /ie/	“ie”	–	–	29 /f/, /f’/	“f”	[f]	fork
12 /uo/	“uo”	–	–	30 /j’/	“j”	[y]	yard
13 /p/, /p’/	“p”	[p] ²	put	31 /v/, /v’/	“v”	[v]	vat
14 /b/, /b’/	“b”	[b]	big	32 /l/, /l’/	“l”	[l]	lid
15 /t/, /t’/	“t”	[t] ²	talk	33 /m/, /m’/	“m”	[m]	mat
16 /d/, /d’/	“d”	[d]	dig	34 /n/, /n’/	“n”	[n]	no
17 /k/, /k’/	“k”	[k] ²	cut	35 /r/, /r’/	“r”	[r] ⁵	red
18 /g/, /g’/	“g”	[g]	gut				

¹ The same phoneme is used for short and long vowel.

² It is pronounced with aspiration in some cases.

³ Two consonants [t][sh] can be used instead of this one.

⁴ Two consonants [d][zh] can be used instead of this one.

⁵ It sounds very differently from the Lithuanian counterpart.

Šaltinis: Pijus Kasparaitis (2008). *Lithuanian Speech Recognition Using the English Recognizer*. Matematikos ir informatikos fakultetas, Vilniaus Universitetas, Vilnius, Vol. 19, No. 4, 505 – 516

22 pav. P. Kasparaičio pateikiamos perrašymo taisyklės.

Toliau darbe bus atliekamas eksperimentas norint išsiaiškinti anksčiau darbe paminėtas problemas. Eksperimentai bus atliekami naudojant eksperimentinę programinę įrangą, kurioje yra

taikomi prieš tai aptarti metodai, t.y. atliekamas pradinis kalbos signalo apdorojimas, tada signalas paduodamas šnekos išskyrimo algoritmui, o iš kalbos signalo išskyrus šneką, gautas pjūvis paduodamas į SAPI atpažinimo variklį.

4.1. „ė“ ir „ch“ garsų aprašymo eksperimentas

Šio eksperimento tikslas yra nustatyti, kaip geriau aprašyti garsus „ė“ ir „ch“. Eksperimentui atlikti buvo pasirinktos angliškos fonemos, kurios panašiausiai skamba į tiriamus garsus. Su pasirinktomis fonemomis buvo sudaryta tiriamų žodžių aibė. Sudarinėjant žodžių aibę buvo stengtasi, kad tiriamos fonemos būtų kuo labiau išsimėtę žodyje, t.y. būtų žodžio pradžioje, viduryje ir pabaigoje. Eksperimento dalyviai (kurių buvo dvidešimt) po 30 kartų sakė tiriamus žodžius. Analizuojami buvo žodžiai su „ė“: kėdė, saulė, ėdalas, ė, o su „ch“: choras, echidna, bronchitas, ch. Žodžių kombinacijos pateikiamos apačioje esančioje lentelėje.

4 lentelė.

Garsams „ė“ ir „ch“ tirti sudarytų žodžių aibė

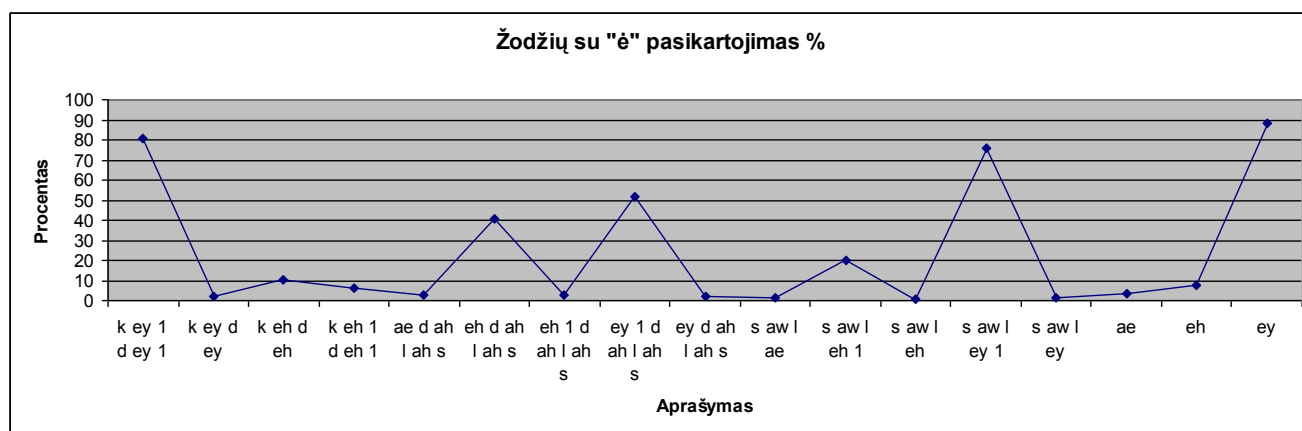
„ė“	„ch“
k ey 1 d ey 1	ch ow r ah s
k ey d ey	ch 1 ow r ah s
k eh d eh	h ow r ah s
k eh 1 d eh 1	h 1 ow r ah s
k ae d ae	ch h ow r ah s
k ae 1 d ae 1	ch h 1 ow r ah s
k ae aa d ae aa	eh ch ih d n ah
k ae aa 1 d ae aa 1	eh ch 1 ih d n ah
k ae ey d ae ey	eh h ih d n ah
k ae ey 1 d ae ey 1	eh h 1 ih d n ah
ae ey d ah l ah s	eh ch h ih d n ah
ae ey 1 d ah l ah s	eh ch h 1 ih d n ah
ae aa d ah l ah s	b r ao n ch ih t ah s
ae aa 1 d ah l ah s	b r ao n ch 1 ih t ah s
ae 1 d ah l ah s	b r ao n h ih t ah s
ae d ah l ah s	b r ao n h 1 ih t ah s
eh d ah l ah s	b r ao n ch h ih t ah s
eh 1 d ah l ah s	b r ao n ch h 1 ih t ah s
ey 1 d ah l ah s	ch
ey d ah l ah s	ch 1
s aw l ae ey 1	h
s aw l ae ey	h 1

4 lentelės tęsinys

„ė“	„ch“
s aw ae aa 1	ch h
s aw l ae aa	ch h 1
s aw l ae	
s aw l ae 1	
s aw l eh 1	
s aw l eh	
s aw l ey 1	
s aw l ey	
ae ey 1	
ae ey	
ae aa	
ae aa 1	
ae	
ae 1	
eh 1	
eh	
ey 1	
ey	

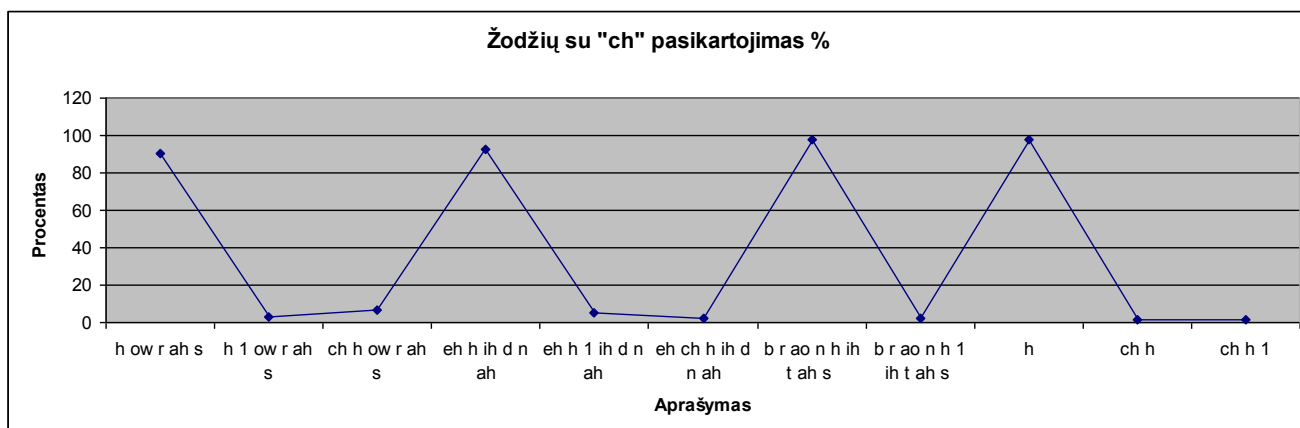
Šaltinis: sukurta autoriaus.

Iš tirtų žodžių užrašymų kombinacijų atrinktos bent kartą pasitaikę ir pateiktos žemiau esančiuose grafikuose.



Šaltinis: sukurta autoriaus.

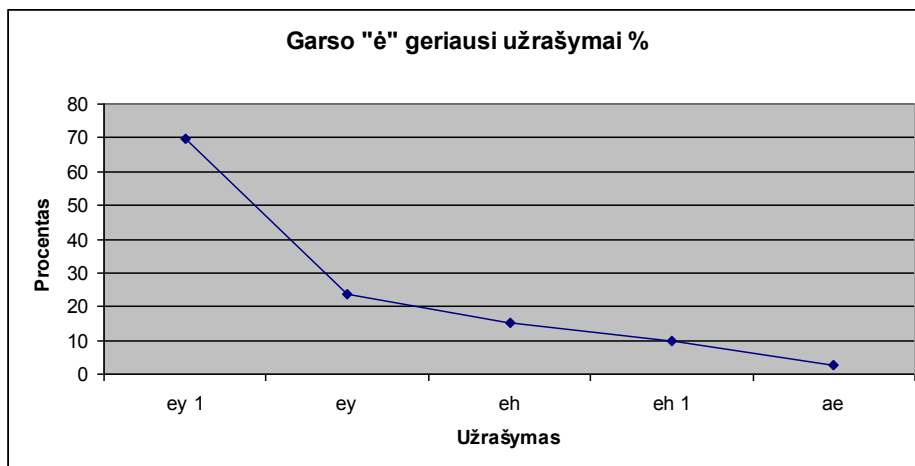
23 pav. Žodžių su „ė“ pasikartojimas procentais.



Šaltinis: sukurta autoriaus.

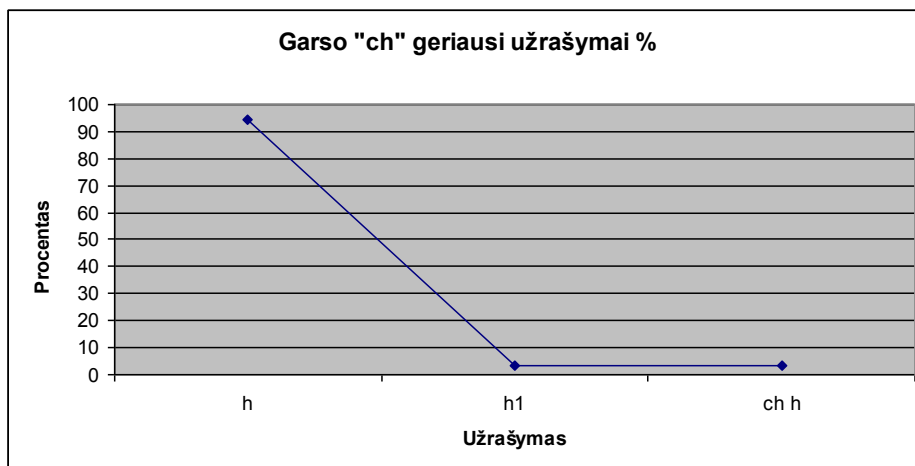
24 pav. Žodžių su „ch“ pasikartojimas procentais.

Gauti duomenys sutraukiami ir grafiškai pateikiama garsų „ė“ ir „ch“ geriausių užrašymų procentų vidurkiai.



Šaltinis: sukurta autoriaus.

25 pav. Geriausių garso „ė“ užrašymų reikšmės.



Šaltinis: sukurta autoriaus.

26 pav. Geriausių garso „ch“ užrašymų reikšmės.

Atlikus analizę galima daryti išvadą, kad garsą „ė“ galima užrašyti kaip „ey 1“ su patikimumu 70%, o garsą „ch“ kaip „h“ su patikimumu 95%. Garso „ė“ užrašymo procentas nėra aukštas, tad kuriant sistemas paremtas kitų kalbų atpažinimo varikliais, siūloma padubliuoti žodžius su „ė“ ir juose panaudoti fonemą „ey“, taip bus padidintas atpažinimo procentas iki apytiksliai 90 %.

Lyginant su P. Kasparaičio darbu (P. Kasparaitis (2008). *Lithuanian Speech Recognition Using the English Recognizer.*) eksperimente naudojama daugiau žmonių balsų, todėl gaunamas didesnis potencialus tikslumas. Taip pat eksperimente pavyko gauti geresnį garso „ch“ atpažinimo procentą.

4.2. „c“ ir „dz“ garsų aprašymo eksperimentas

Šis eksperimentas yra svarbus tuo, jog padės nustatyti, koks užrašymas yra efektyvesnis fonemų „c“ ir „dz“ užrašymui. Šias dvi fonemas galima aprašyti atskirais garsais: „c“, kaip [t] ir [s] ar tik [s] ir „dz“, kaip [d] ir [z] ar [z] arba [jh] ir [z].

Eksperimentas buvo atliktas su dvidešimties žmonių aibe, kurie po 30 kartų sakė, apačioje lentelėje, parašytus žodžius.

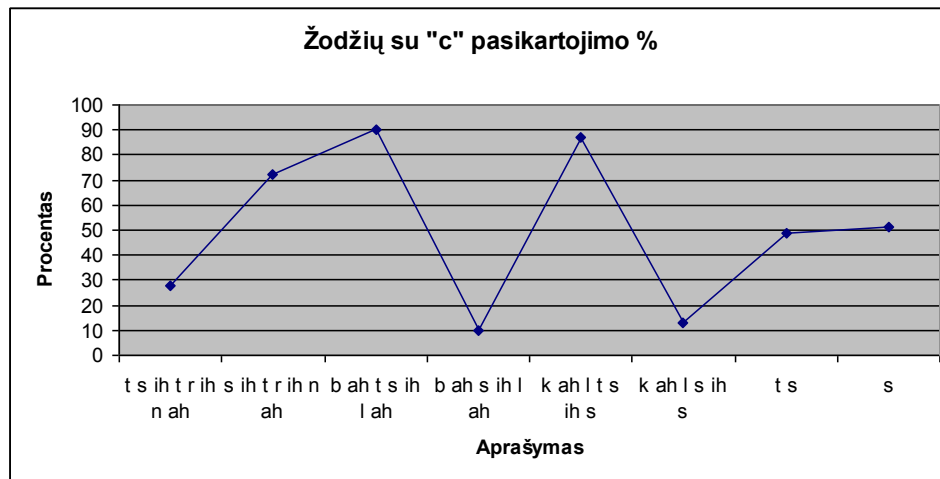
5 lentelė.

Garsams „c“ ir „dz“ tirti sudarytų žodžių aibė

„c“	„dz“
t s i h t r i h n a h	d z u w k a h s
s i h t r i h n a h	z u w k a h s
b a h t s i h l a h	j h z u w k a h s
b a h s i h l a h	p a h d z i h n s e y l t i h
k a h l t s i h s	p a h z i h n s e y l t i h
k a h l s i h s	p a h j h z i h n s e y l t i h
t s	d z
s	z
	j h z

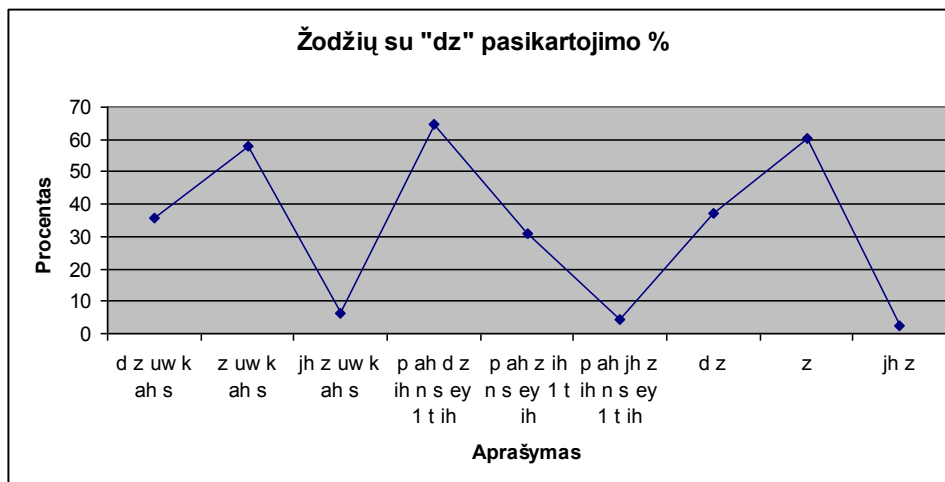
Šaltinis: sukurta autoriaus.

Atlikto eksperimento duomenys pateikti grafikuose apačioje.



Šaltinis: sukurta autoriaus.

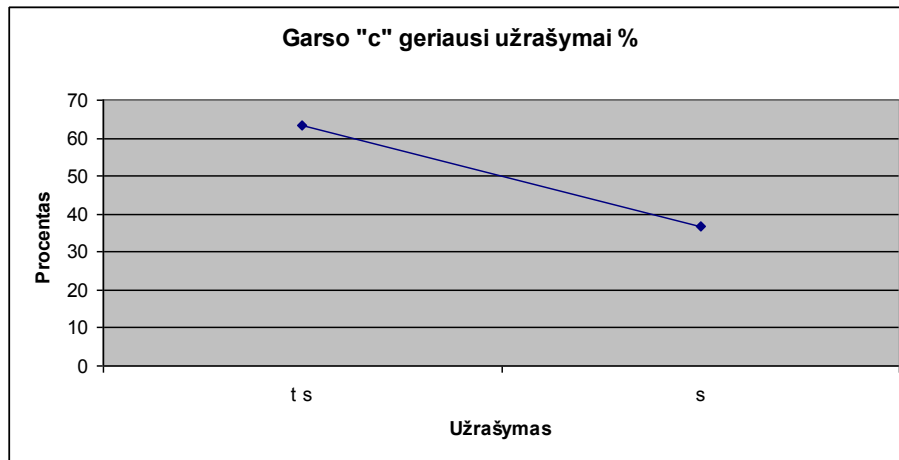
27 pav. Žodžių su „c“ pasikartojimas procentais.



Šaltinis: sukurta autoriaus.

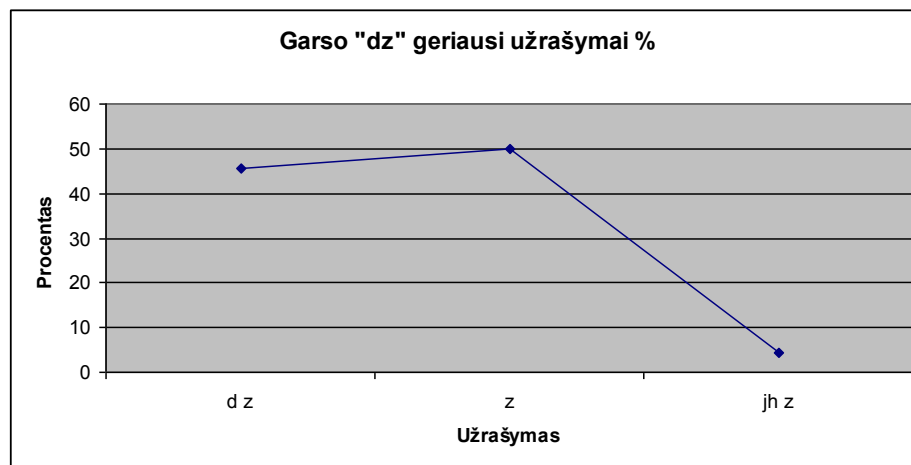
28 pav. Žodžių su „dz“ pasikartojimas procentais.

Kaip matome iš grafikų, kai garsas „c“ pasitaiko žodžio pradžioje jo užrašymas kaip „s“ yra geresnis, atitinkamai ir garso „dz“ geresnis užrašymas yra „z“. Tai paaiškinama tuo, jog pradedant tarti žodį pirmasis garsas kuris yra trumpas ir sprogstamas iš užrašymų „t s“ ir „d z“ neturi didelės įtakos atpažinimo varikliui. Tuo tarpu kai garsai yra žodžių viduryje geresnis užrašymas garsui „c“ tampa „t s“, o „dz“ – „d z“. Žodžio viduryje nuskambėjęs pirmasis garsas labiau įtakoja atpažinimo variklį. Apačioje pateikiama bendros garsų užrašymo diagramos.



Šaltinis: sukurta autoriaus.

29 pav. Geriausių garso „c“ užrašymų procentas.



Šaltinis: sukurta autoriaus.

30 pav. Geriausių garso „dz“ užrašymo reikšmių procentas.

Apžvelgus gautus duomenis galima daryti išvadą, jog atpažinimo variklis tiksliau atpažįsta užrašymą „t s“ nei „s“, tačiau jei garsas „c“ yra žodžio pradžioje siūloma naudoti užrašymą „s“. Taip pat tiksliau yra atpažįstamas užrašymas „z“ nei „d z“, tačiau jei garsas „dz“ pasitaiko žodžio viduryje siūloma naudoti „d z“ užrašymą.

4.3. „č“ ir „dž“ garsų eksperimentas

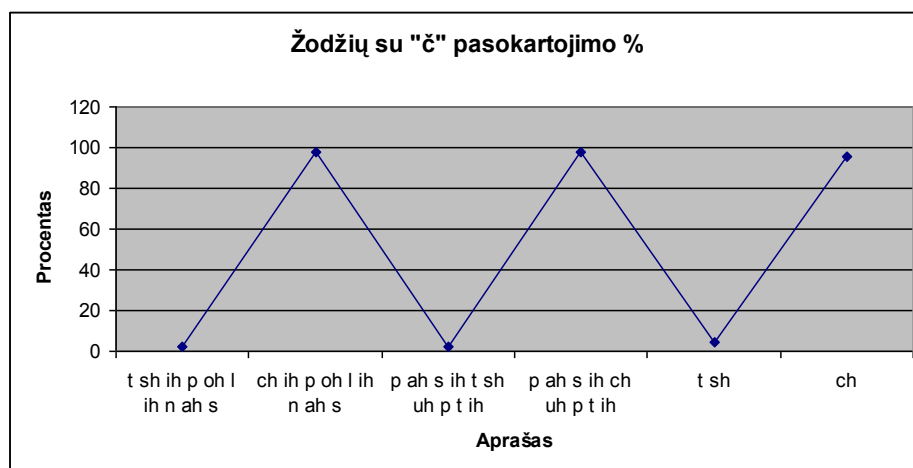
Šio eksperimento tikslas yra išsiaiškinti koks angliškas užrašymas geriau tinka „č“ ir „dž“ garsams aprašyti. P. Kasparaitis savo darbe (P. Kasparaitis (2008). *Lithuanian Speech Recognition Using the English Recognizer.*) teigia, jog „č“ garsą galima užrašyti kaip „t sh“ arba „ch“, o garsą „dž“ kaip „d zh“ arba „jh“. Šio eksperimento tikslas yra išsiaiškinti kuris variantas yra geresnis. Eksperimentas buvo atliktas su dvidešimties žmonių aibe, kurie po 30 kartų sakė, apačioje lentelėje, parašytus žodžius. Sudaryta eksperimento žodžių aibė pateikta apačioje esančioje lentelėje.

Garsams „č“ ir „dž“ tirti sudarytų žodžių aibė

„č“	„dž“
t sh ih p oh l ih n ah s	jh ih n ah s
ch ih p oh l ih n ah s	d zh ih n ah s
p ah s ih t sh uh p t ih	s k uh r jh uh s
p ah s ih ch uh p t ih	s k uh r d zh uh s
t sh	jh
ch	d zh

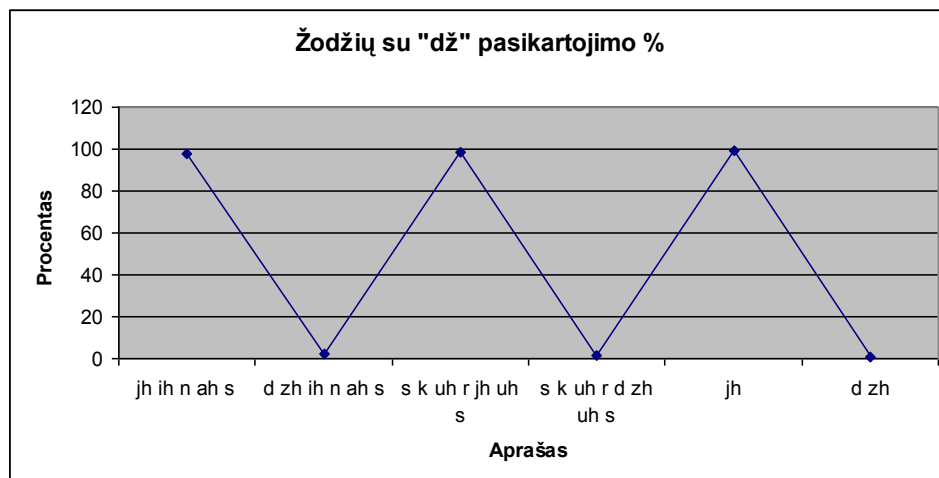
Šaltinis: sukurta autoriaus.

Atlikto eksperimento rezultatas pateikiamas apačioje esančiuose paveikslėliuose.



Šaltinis: sukurta autoriaus.

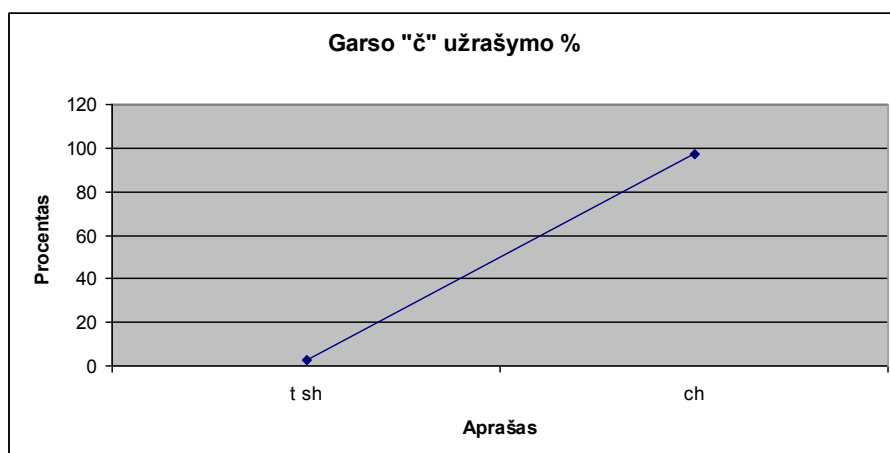
31 pav. Žodžių su „č“ pasikartojimas procentais.



Šaltinis: sukurta autoriaus.

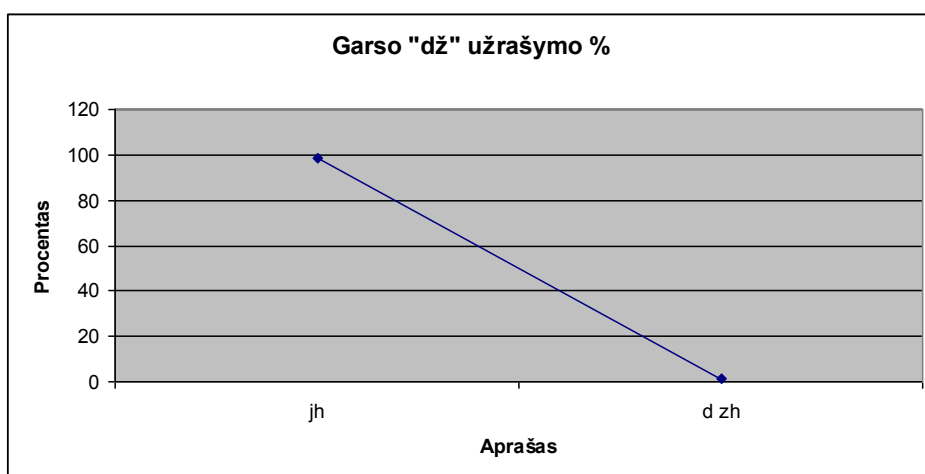
32 pav. Žodžių su „dž“ pasikartojimas procentais.

Kaip galima pamatyti iš atlikto eksperimento rezultatų, „č“ užrašymas „t sh“ ir „dž“ užrašymas kaip „d zh“ yra beveik neveiksmingi. Apačioje pateikiama užrašymų procentinis įvertinimas.



Šaltinis: sukurta autoriaus.

33 pav. Geriausių garso „č“ užrašymo reikšmių procentas.



Šaltinis: sukurta autoriaus.

34 pav. Geriausių garso „dž“ užrašymo reikšmių procentas.

Pagal eksperimento rezultatus galima teigti, „č“ užrašymas kaip „ch“, o „dž“ kaip „jh“ yra daug efektingesni. Vartojant užrašymus „t sh“ ir „d zh“ rizikuojama sumažinti atpažinimo variklio gerai atpažįstamų žodžių procentą.

Gauti eksperimento rezultatai pranoksta P. Kasparaičio (P. Kasparaitis (2008). *Lithuanian Speech Recognition Using the English Recognizer.*) atliktų eksperimentų rezultatus.

4.4. „ie“ ir „uo“ garsų eksperimentas

Šio eksperimento tikslas yra išsiaiškinti koks angliškas užrašymas geriausia tinka „ie“ ir „uo“ garsams aprašyti. P. Kasparaitis savo darbe (P. Kasparaitis (2008). *Lithuanian Speech Recognition Using the English Recognizer.*) gavo rezultatus, jog garsą „ie“ geriausia užrašyti kaip [iy ax], o garsą „uo“ kaip [uw ax]. Eksperimentas atliekamas, norint išsiaiškinti, ar šie užrašymai yra patys efektyviausi. Šio eksperimento tikslas yra išsiaiškinti kuris variantas yra geresnis. Eksperimentas buvo

atliktas su dvidešimties žmonių aibe, kurie po 30 kartų sakė, apačioje lentelėje, parašytus žodžius. Sudaryta eksperimento žodžių aibė pateikta apačioje esančioje lentelėje.

7 lentelė.

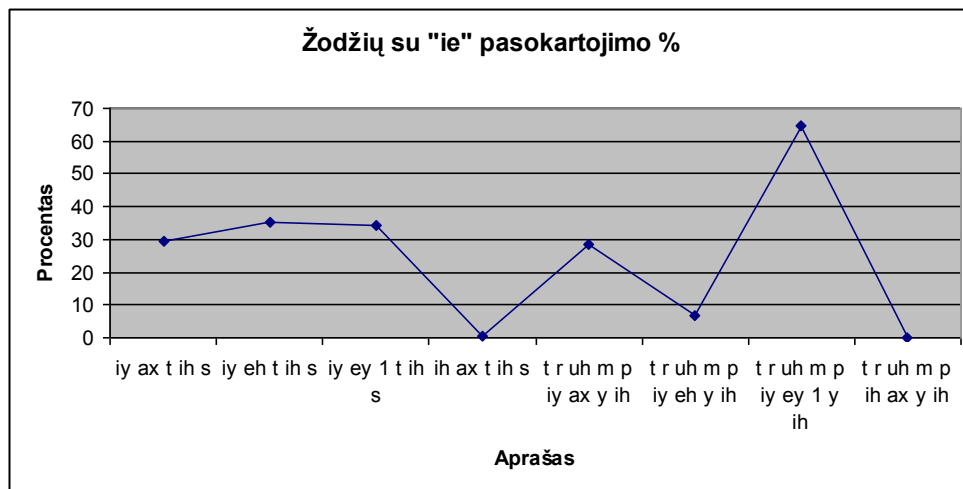
Garsams „ie“ ir „uo“ tirti sudarytų žodžių aibė

„ie“	„uo“
iy ax t ih s	w ao l ah
iy eh t ih s	uw ax l ah
iy ey l t ih s	w aw l ah
ih ax t ih s	uw eh l ah
t r uh m p iy ax y ih	m iy l w ao t ih
t r uh m p iy eh y ih	m iy l uw ax t ih
t r uh m p iy ey l y ih	m iy l w aw t ih
t r uh m p ih ax y ih	m iy l uw eh t ih
	n w ao
	n uw ax
	n w aw
	n uw eh

Šaltinis: sukurta autoriaus.

Sudarius gramatinius failus, su sukurtais aprašymai, juos paduodame į eksperimentinę programinę įrangą ir stebime atliekamo eksperimento rezultatus.

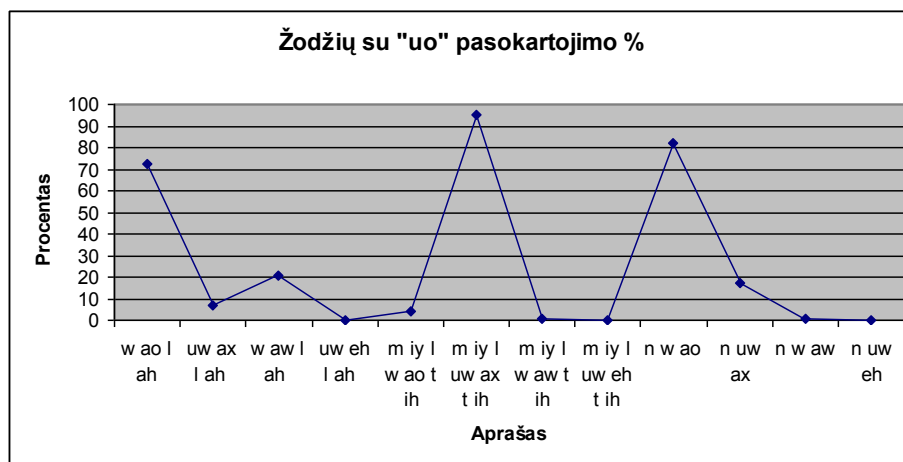
Atlikto eksperimento rezultatai pateikiami apačioje.



Šaltinis: sukurta autoriaus.

35 pav. Žodžių su „ie“ pasikartojimas procentais.

Iš grafiko matyti, jog žodžio „ietis“ aprašymuose nėra aiškaus užrašymo, visi užrašymai pasikartoja maždaug vienodai. Tuo tarpu žodžio „trumpieji“ užrašymuose matomas išsiskiriantis užrašymas.

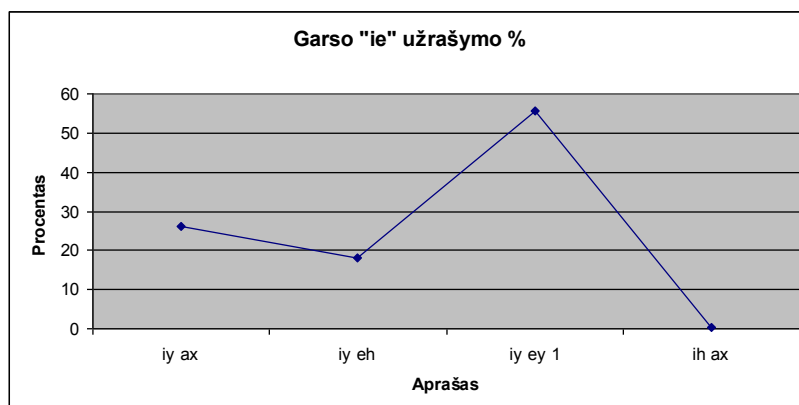


Šaltinis: sukurta autoriaus.

36 pav. Žodžių su „uo“ pasikartojimas procentais.

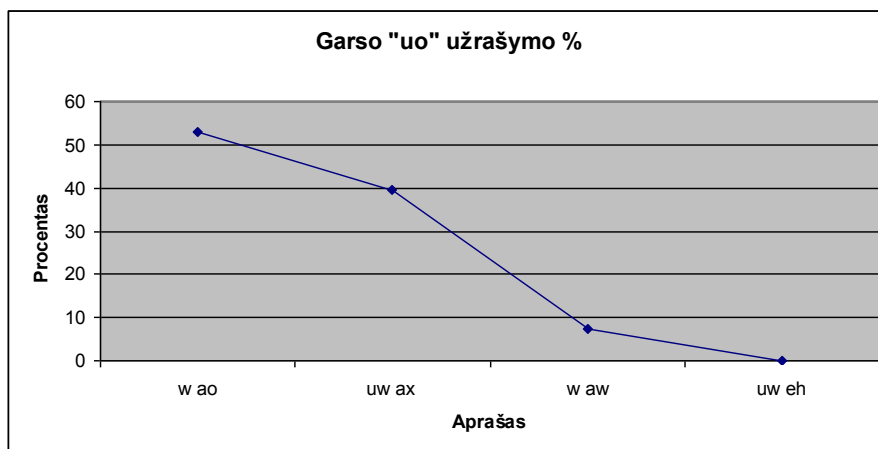
Iš šio grafiko taip pat matyti, jog yra keli išsiskiriantys žodžių užrašymai.

Išskyrus tik tiriamus užrašymus gaunamos tokios pasikartojimų procentalios reikšmės (pateikta apačioje esančiuose diagramose).



Šaltinis: sukurta autoriaus.

37 pav. Geriausių garso „ie“ užrašymo reikšmių procentas.



Šaltinis: sukurta autoriaus.

38 pav. Geriausių garso „uo“ užrašymo reikšmių procentas.

Analizuojant garso „ie“ užrašymo eksperimento rezultatus galima pamatyti, jog jei garsas yra žodžio pradžioje, atpažinimas tirtų užrašymų yra gan vienodas, tačiau jei šis garsas yra žodžio viduryje matome, jog užrašymas [iy ey 1] parodo labai aukštus rezultatus, beveik 65%. Lyginant su P. Kasparaičio darbu (P. Kasparaitis (2008). *Lithuanian Speech Recognition Using the English Recognizer.*) gautas rezultatas yra gan vienodas, tačiau šiame eksperimente geriausiai atpažįstamas užrašymas yra P. Kasparaičio netirtas užrašymas [iy ey 1].

Garso „uo“ užrašymo eksperimento rezultatai yra gan nustebino, nes labai akivaizdžiai matosi, jog garsą „uo“, jei jis yra žodžio pradžioje arba pabaigoje geriausiai jį galima aprašyti kaip [w ao], tačiau jei garsas yra viduryje žodžio labiausiai tinkamas užrašymas yra [uw ax]. Bendri užrašymų rezultatai neišskiria kažkurio užrašymo kaip geriausio [w ao] – 53%, [uw ax] – 40%. Kaip matome užrašymų atpažinimo procentas atskirai yra mažas, tačiau naudojant vieną užrašymą kai garsas yra žodžio pradžioje arba pabaigoje, kitą – kai viduryje, galima labai padidinti atpažinimo tikslumą. Gauti rezultatai šiek tiek nusileidžia P. Kasparaičio darbe (P. Kasparaitis (2008). *Lithuanian Speech Recognition Using the English Recognizer.*) gautiems rezultatams, tačiau šio eksperimento dėka, nustatyta, kuris užrašymas, kurioje žodžio vietoje yra efektyvesnis.

4.5. Mažos apimties žodyno, sudaryto pagal suformuluotas taisykles, atpažinimo eksperimentas

Pagal atliktus eksperimentų rezultatus, galima papildyti P. Kasparaičio pateiktą garsų perrašymo taisyklių lentelę, kuri atrodytų taip:

8 lentelė.

Papildyta garsų perrašymo taisyklių lentelė

	Lietuviška fonema	Lietuviška raidė	Angliška fonema	Pvz.		Lietuviška fonema	Lietuviška raidė	Angliška fonema	Pvz.
1	/a/	„a“	[ah]	Cut	19	/t/, /t' /	„č“	[ch] efektyviau negu [t sh]	Chin
2	/e/	„e“	[eh]	Pet	20	/d3/, /d3' /	„dž“	[jh] efektyviau negu [d zh]	Joy
3	/i/	„i“	[ih]	Fill	21	/s/, /s' /	„s“	[s]	Sit
4	/o/, /o:/	„o“	[ao]	Dog	22	/z/, /z' /	„z“	[z]	Zap
5	/u/	„u“	[uh]	Book	23	/ʃ/, /ʃ' /	„š“	[sh]	She
6	/a:/	„a“, „ą“	[aa]	Father	24	/3/, /3' /	„ž“	[zh]	Peasure

8 lentelės tęsinys

	Lietuviška fonema	Lietuviška raidė	Angliška fonema	Pvz.		Lietuviška fonema	Lietuviška raidė	Angliška fonema	Pvz.
7	/e:/	„e“, „ė“	[ae]	Cat	25	/x/, /x’/	„ch“	[h] 95%	Chorus
8	/i:/	„i“, „y“	[iy]	Feel	26	/h/, /h’/	„h“	[h]	Help
9	/u:/	„u“, „ū“	[uw]	Too	27	/f/, /f’/	„f“	[f]	Fork
10	/ė:/	„ė“	[ey 1] 70%, [ey] 25%		28	/j’/	„j“	[y]	Yard
11	/p/, /p’/	„p“	[p]	Put	29	/v/, /v’/	„v“	[v]	Vat
12	/b/, /b’/	„b“	[b]	Big	30	/l/, /l’/	„l“	[l]	Lid
13	/t/, /t’/	„t“	[t]	Talk	31	/m/, /m’/	„m“	[m]	Mat
14	/d/, /d’/	„d“	[d]	Dig	32	/n/, /n’/	„n“	[n]	No
15	/k/, /k’/	„k“	[k]	Cut	33	/r/, /r’/	„r“	[r]	Red
16	/g/, /g’/	„g“	[g]	Gut	34	/ie/	„ie“	Pradžioje [iy ey 1] arba [iy ax], pabaigoje tik [iy ey 1]	
17	/ts/, /ts’/	„c“	Pradžioje [s], viduryje [t s]	Civilian	35	/uo/	„uo“	Pradžioje ir pabaigoje [w ao], viduryje [uw ax]	
18	/dz/, /dz’/	„dz“	Žodžio pradžioje [z], viduryje [d z]						

Šaltinis: sukurta autoriaus.

Pagal šias taisykles sudarytas eksperimento žodynas pateiktas apačioje.

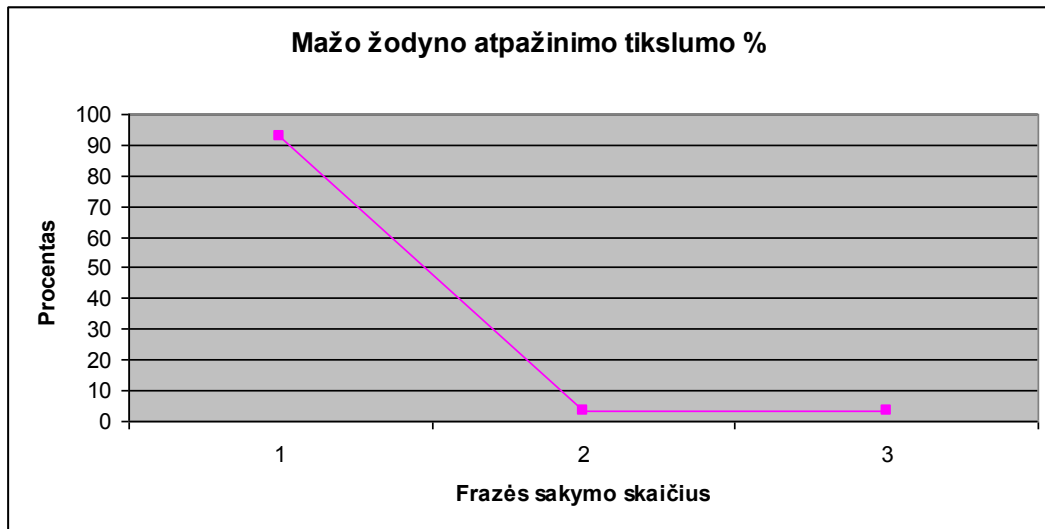
9 lentelė.

Sudarytas mažos apimties žodynas

Žodis	Užrašymas
Vėtra	v ey l t r ah
Čipolinas	ch ih p oh l ih n ah s
Džinas	jh ih n ah s
Kalcis	k aa l t s ih s
Skurdžius	s k uh r jh uh s
Tįsti	t iy s t ih
Chaosas	h ah ao s ah s
Choras	h ow r ah s
Vėlu	v ey l l uh
Džipas	jh ih p ah s
Echidna	eh h ih d n ah
Pasičiupti	p ah s ih ch ih uh p t ih
Bronchitas	b r ao n h ih t ah s
Mėnulis	m ey l n uh l ih s
Ciklas	s ih k l ah s
Karosas	k ah r ao s ah s
Vizitas	v ih z ih t ah s
Uola	w ao l ah
Centas	s eh n t ah s
Rašyti	r ah sh iy t ih
Pravardė	p r ah v ah r d ey l
Dužti	d uh zh t ih
Ryškus	r iy sh k uh s
Civilizacija	s ih v ih l ih z ah t s ih y ah
Čežėti	ch eh zh ey l t ih
Porcija	p oh r t s ih y ah
Supančioti	s uh p ah n ch ih ao t ih
Visada	v ih s ah d ah
Linksnis	l ih n k s n ih s
Trumpieji	t r uh m p iy ey l y ih
Ironiškas	ih r ao n ih sh k ah s
Eržilas	eh r zh ih l ah s
Dziudo	z ih uh d oh

Šaltinis: sukurta autoriaus.

Dvidešimt eksperimento dalyvių atsitiktinai sakė 10 žodžių iš sudaryto žodyno ir buvo stebimas atpažinimo tikslumas. Gauti rezultatai pateikti apačioje esančiame grafike.



Šaltinis: sukurta autoriaus.

39 pav. Mažo žodyno atpažinimo tikslumas.

Kaip matyti iš grafiko, atpažinimo tikslumas yra labai aukštas – 93%. Tai rodo kad perrašymo taisyklės yra gan efektingos. Taip pat, jog metodas, kuriame kalbos signalui pirmiausia atliekamas pirminis apdorojimas, tada ieškoma ir išskiriama šneka iš vientiso kalbos signalo ir tik tada gautas pjūvis paduodamas į atpažinimo variklį yra efektingas ir mažiau naudojantis atpažinimo variklio resursų, nes atpažinimo variklis ne nuolatos analizuoja kalbos signalą, tačiau jau gauna paruoštą atpažinimui. Tačiau šis metodas yra šiek tiek rizikingas, nes priklauso nuo aplinkos garso lygio. Jei staiga pakiltų garso lygis aplinkoje, kurioje naudojama sistema, paremta šiuo metodu, atpažinimas suprastėtų, nes būtų klaidingai interpretuojamos šnekos pradžios ir pabaigos. Šią problemą galima išspręsti naudojant dinamišką aplinkos garso lygio tikrinimą ir pagal tai reguliuojant tam tikras programinės įrangos savybes, tačiau šiame darbo šio būdo nepritaikiau.

Šis metodas atpažinimo sistemoms, kurios naudoja mažą ir vidutinio dydžio žodyną ir veiks mažai kintančio pašalinio triukšmo sąlygomis, gali būti lengvai pritaikomas, be didesnių kaštų skirtų sudėtingiems programavimo darbams, bei palaikymui.

Atlikti eksperimentai papildė P. Kasparaičio ir R. Maskeliūno atliktus eksperimentus. Kadangi mano eksperimentuose naudojama didesnė balsų aibė negu P. Kasparaičio, gaunamas didesnis potencialus tikslumas, bei vietomis pavyksta pagerinti jo gautus rezultatus. R. Maskeliūnas savo darbe (P. Maskeliūnas, 2009, *LITHUANIAN VOICE COMMANDS RECOGNITION BASED ON THE MULTIPLE TRANSCRIPTIONS*) generuoja daug skirtingų transkripcijų ir bando rasti optimaliausias. Tačiau šiam metodui reikia didelės aibės mokymo duomenų. Mano atliktuose eksperimentuose

naudojamasi tik fonetinėmis taisyklėmis, todėl mano metodas reikalauja mažesnių resursų ir yra pranašesnis mažesniuose ir mažiau resursų reikalaujančiuose taikymuose. Tai yra geras kompromisas tarp brangaus ir daug išteklių reikalaujančio kūrimo ir prasto atpažinimo.

IŠVADOS

1. Kalbos atpažinimo metodai skirstomi į tris grupes: akustiniai – fonetiniai, statistiniai ir dirbtinio intelekto. Galima teigti, kad kiekvienas metodas leidžia pasiekti geriausius rezultatus, tačiau su tam tikrais trūkumais, kurie būdingi kiekvienam iš jų.
2. Praktikoje plačiausiai naudojami statistiniai metodai, nes šiuo metu demonstruoja geriausius rezultatus. Dirbtinio intelekto metodai praktikoje taikomi mažiau negu statistiniai, nes yra labai sudėtingas bandymų kelias norint nustatyti ir sudaryti neuronų ryšių koeficientus. Akustiniai – fonetiniai metodai praktikoje taikomi mažiausiai, nes signalo savybių išrinkimas ir signalo klasifikatorių formavimas nėra optimalūs, bei reikalingos fonetikos teorijos žinios.
3. Nors SAPI pritaikytas anglų kalbos atpažinimui iš jo veikimo schemos matosi, kad jį pritaikyti kitų kalbų atpažinimui galima keičiant naudojamus garsų etalonus, arba perrašant tam tikros kalbos garsus angliškais.
4. Svarbiausias eksperimento rezultatas – įvertintos kiekybinės šnekamosios šnekos signalo pradžios ir pabaigos nustatymo, naudojant energetinį matą, charakteristikos. Nustatyta, jog metodas yra neefektyvus kai triukšmas kinta dideliame diapazone ir persidengia su nustatytomis slenksčio reikšmėmis. Metodui lengviau nustatyti komandų pradžias kai komanda prasideda sprogstamuoju priebalsiu. Kadangi pabaigoje dažniausiai pasitaiko friktyvinis priebalsis ‚S‘ arba pusbalsis ‚I‘, kurie laiko kitimo skalėje mažina savo energiją palaipsniui, tad maža dalis kalbos signalo pabaigos yra prarandama.
5. Atlikus eksperimentą nustatyta, jog SAPI atpažinimo varikliui padavus gramatikos failą, kuriame atpažįstami garsai užrašyti skirtingais formatais (pvz.: ‚T r iy s‘ ir ‚Threes‘), yra padidinama atpažinimo kokybė ir praplečiamas tinkamai atpažįstamų balso savybių spektras. Nustatyta, jog toks metodas gali būti taikomas tik turint mažą arba vidutinio dydžio žodyną, nes kitaip atpažinimas užtruks ilgiau ir nebus efektyvus, nes visų lietuviškų garsų angliškais išreikšti negalima.
6. Atlikus eksperimentus patikrinta ir pasiūlyta kaip geriau perrašyti sudėtingus lietuviškus garsus angliškais: ‚ė‘ perrašyti kaip [ey 1] su 75% tikslumu (norint padidinti tikslumą siūloma sudvigubinti žodį ir vietoj [ey 1] panaudoti [ey], kuris davė 25% tikslumą), ‚ch‘ kaip [h] su 95% patikimumu, ‚c‘ kaip [t s], jei garsas žodžio pradžioje geresnis tikslumas gaunamas kai užrašoma kaip [s], ‚dz‘ kaip [d z], jei garsas yra žodžio pradžioje geresnis tikslumas užrašant kaip [z], ‚č‘ efektyvesnis užrašymas yra ‚ch‘, o ‚dž‘ – ‚jh‘, ‚ie‘ kaip [iy ey 1], šis užrašymas efektyviausias kai garsas yra žodžio viduryje, ‚uo‘ kaip [w ao], jei yra garsas žodžio pradžioje

arba pabaigoje, [uw ax], jei garsas yra žodžio viduryje. Sukūrus iš pasiūlytų taisyklių mažos apimties žodyną, gaunamas 93% atpažinimo tikslumas, kas parodo naudojamo metodo efektyvumą naudojant mažos ar vidutinės apimties žodyną. Kuriant atpažinimo sistemas, kurios naudosis tokio dydžio žodynais siūloma naudoti tokį atpažinimo metodą, dėl mažų sistemos kūrimo kaštų, bei mažo laiko reikalingo sistemai realizuoti. Eksperimentų rezultatai vietomis pranoksta P. Kasparaičio atliktų eksperimentų rezultatus, taip pat gaunamas didesnis potencialus tikslumas, nes naudojama didesnė balsų aibė. Lyginant su R. Maskeliūno atliktais eksperimentais galima teigti, jog mano metodas yra geras kompromisas tarp brangaus ir daug išteklių reikalaujančio kūrimo ir prasto atpažinimo.

LITERATŪRA

1. Girdenis Aleksas (1995). *Teoriniai lietuvių kalbos fonologijos pagrindai*. Vilniaus Universitetas, Vilnius. ISBN 9785420015018.
2. Huang Xuedong, Acero Alex, Hon Hsiao-Wuen (2001). *Spoken language processing*. USA, Prentice-Hall Inc. ISBN 0-13-022616-5.
3. Jurafsky Daniel, James H. Martin (2000). *Speech and Language Processing*. New Jersey 07632, Prentice – Hall Inc.
4. Joseph Picone (1993). *Signal Modeling Techniques In Speech Recognition*. Systems and Informatikon Sciences Laboratory Tsukuba Research and Development Center Tsukuba, Japan.
5. Kasparaitis Pijus (2008). *Lithuanian Speech Recognition Using the English Recognizer*. Matematikos ir informatikos fakultetas, Vilniaus Universitetas, Vilnius, Vol. 19, No. 4, 505 – 516.
6. Kazlauskienė Asta, Raškinis Gailius (2005). *Lietuvių kalbos sprogstamųjų priebalsių kiekybės*. Vytauto Didžiojo universitetas, Kompiuterinės lingvistikos centras. ISSN 1648-2824.
7. Marzinzik Mark, Kollmeier Birger (2002). *Speech Pause Detection for Noise Spectrum Estimation by Tracking Power Envelope Dynamics*. IEEE transactions on speech and audio processing, Vol. 10, NO. 2.
8. Maskeliūnas Rytis (2009). *Daktaro disertacija. LITHUANIAN VOICE COMMANDS RECOGNITION BASED ON THE MULTIPLE TRANSCRIPTIONS*. Kauno Technologijos Universitetas, Informatikos inžinerijos katedra, Kaunas.
9. Maskeliūnas R., Rudžionis A., Rudžionis V., Ratkevičius K. (2009). *Investigation of Foreign Languages Models for Lithuanian Speech Recognition*. Elektronika ir elektrotechnika. ISSN 1392-1215, No. 3(91). Prieiga per internetą: http://www.ktu.lt/lt/mokslas/zurnalai/elektros_z/z91/03__ISSN_1392-1215_Investigation%20of%20Foreign%20Languages%20Models%20for%20Lithuanian%20Speech%20Recognition.pdf.
10. Microsoft Speech SDK 5.1. pagalbos dokumentas. Prieiga per internetą: <http://www.microsoft.com/downloads/en/details.aspx?FamilyID=5e86ec97-40a7-453f-b0ee-6583171b4530&displaylang=en>.
11. Rabiner Lawrence, Binn – Hwang Juang (1993). *Fundamentals of speech recognition*. USA, Prentice – Hall Inc. ISBN 0-13-285826-6.

12. Rozak Mike (1996). *Talk to Your Computer and Have It Answer Back with Microsoft Speech API* [interaktyvus]. [žiūrėta 2010.03.15]. Microsoft system journal. Prieiga per internetą: <<http://www.microsoft.com/msj/archive/s233.aspx>>.
13. Sakhnov Kiril, Simak Boris (2009). *Dynamical Energy – Based Speech/Silence Detector for Speech Enhancement Applications*. Proceedings of the World Congress on Engineering Vol 1, London, U.K.
14. Vary Peter, Rainer Martin (2006). *Digital Speech Transmission Enhancement, Coding and Error Concealment*. England, John Wiley & Sons Ltd. ISBN-13 978-0-471-56018-9.
15. Wai C Chu (2003). *Speech Coding Algorithms Foundation and Evolution of Standardized Coders*. USA, John Wiley & Sons Inc. ISBN 0-471-37312-5.
16. Ying G.S., Mitchell C.D., Jamieson L.H. (2010). *Endpoint Detection of Isolated Utterances Based On A Modified Teager Energy Measurement*. Purdue University, West Lafayette, ISSN 47907-1285.

PRIEDAI

1 PRIEDAS Magistro mokslinio darbo ruošimo forma.....	50
2 PRIEDAS Komandų išskyrimo metodo kokybės tikrinimo eksperimentų rezultatai.....	51
3 PRIEDAS Dalis gramatikos failo turinio.....	62
4 PRIEDAS Eksperimentinės programinės įrangos atpažinimo rezultatai.....	64

1 Priedas. Magistro mokslinio darbo ruošimo forma

VILNIAUS UNIVERSITETAS
KAUNO HUMANITARINIS FAKULTETAS
INFORMATIKOS KATEDRA

VERSLO INFORMATIKOS MAGISTRANTŪROS PROGRAMOS MOKLSO TIRIAMOJO DARBO
PLANAS

Magistrantas Rasymas Tomas Tel. +37062453637
(pavardė, vardas)

Magistrantūros trukmė nuo 2009 m. iki 2011 m.

TEMA: Balso komandų atpažinimo metodai, naudojant fonemų diskriminavimą ir transkripcijų adaptavimą.

Vadovas: Vytautas Evaldas Rudžionis, Dr., doc., Vilniaus universitetas, +37061005649.
(vardas, pavardė, mokslinis laipsnis, pedagoginis vardas, darbovietė, telefonas)

Darbo anotacija:

Tikslas: Ištirti lietuviškų balso komandų atpažinimo galimybę, naudojant nelietuviškas atpažinimo priemones (Microsoft Speech Application Programming Interface).

Uždaviniai: Apžvelgti kalbos atpažinimo metodus; Ištirti kalbos atpažinimo metodus, išskiriant jų privalumus ir trūkumus; Apžvelgti kalbos atpažinimo paketą SAPI; Paruošti eksperimentinius duomenis; Atlikti balso komandų atpažinimo eksperimentus; Apibendrinti rezultatus.

Metodai, kuriuos ketinama ištirti panaudoti darbe: Fonemų diskriminavimo; Transkripcijų adaptavimo.

Laukiami rezultatai: Programinė įranga, gebanti atpažinti lietuviškas balso komandas; Tirtų metodų kokybės vertinimas.

Mokslo – tiriamojo darbo planas

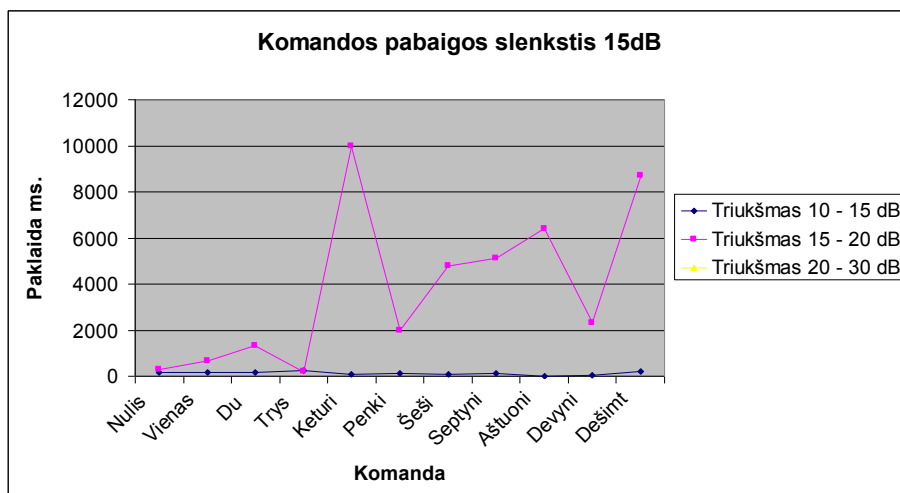
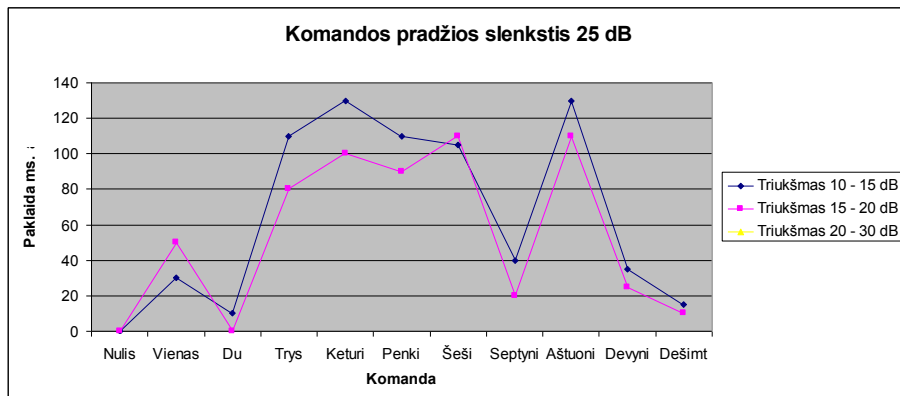
Semestras	(data)	Užduotys
S1	2009.09.01 – 2010.01.20	Literatūros apžvalga.
S2	2010.02.04 – 2010.06.01	Tiriamų metodų analizė.
S3	2010.09.01 – 2011.01.20	Programinio produkto realizacija.
S4	2011.02.04 – 2011.06.01	Eksperimentų atlikimas.

Magistrantas _____ Vadovas _____
(parašas) (parašas)

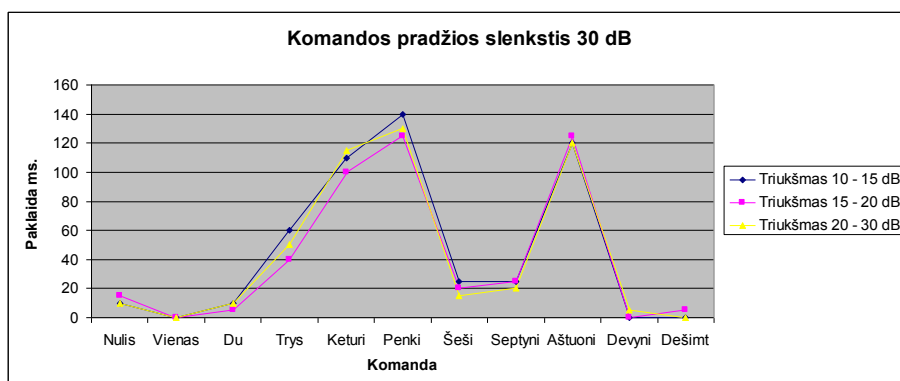
2 Priedas. Komandų išskyrimo metodo kokybės tikrinimo eksperimentų rezultatai

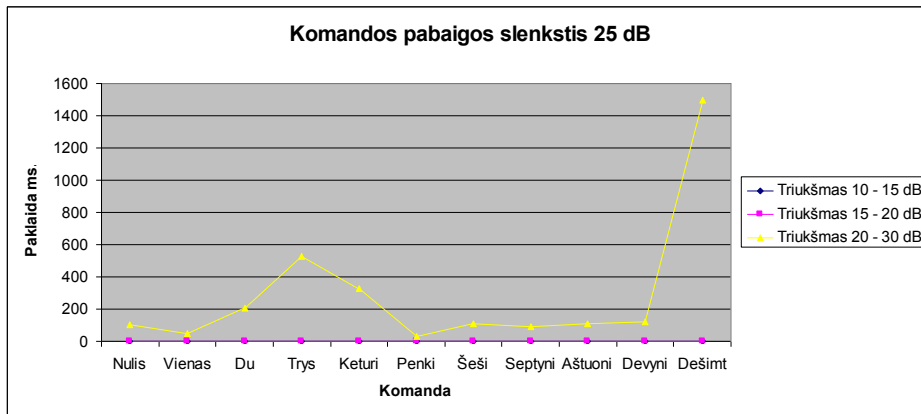
Pirmas asmuo.

Atpažinimo slenksčio reikšmė 25 – 15 dB.

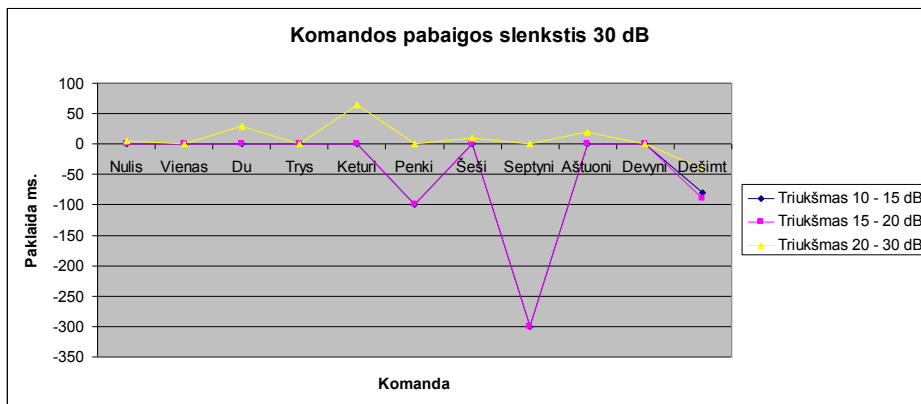
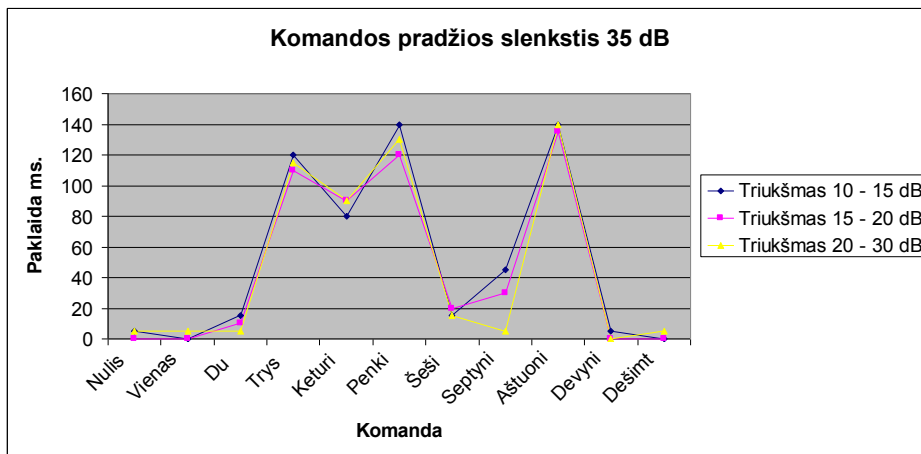


Atpažinimo slenksčio reikšmė 30 – 25 dB.

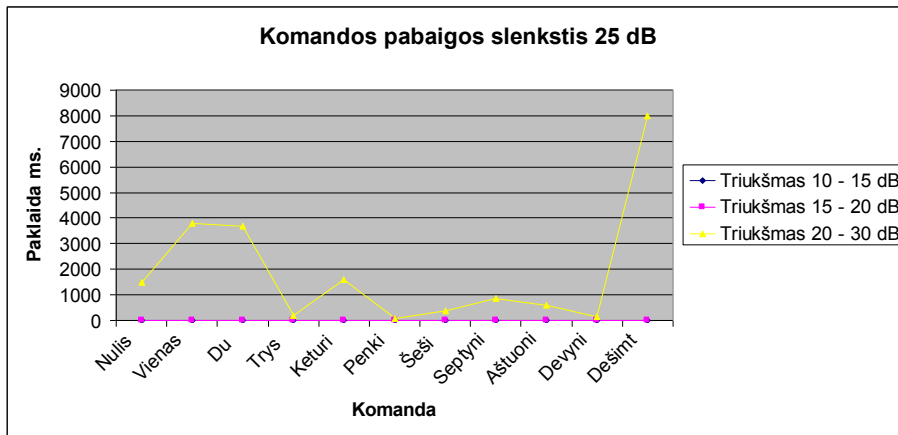
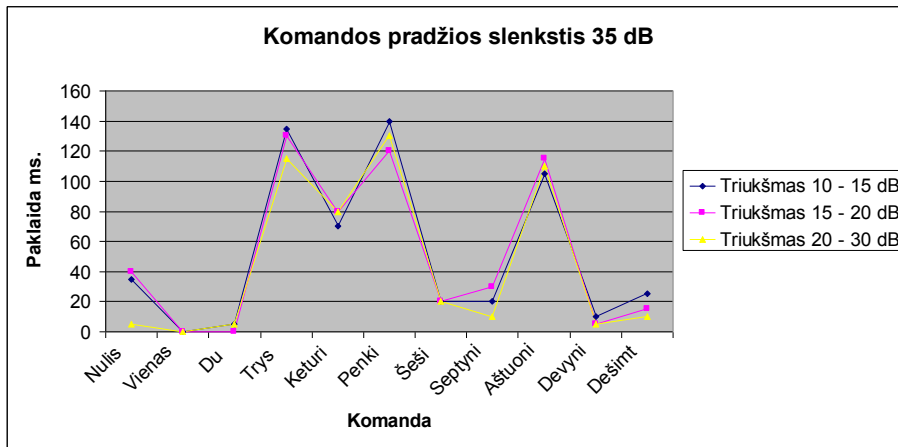




Atpažinimo slenkščio reikšmė 35 – 30 dB.

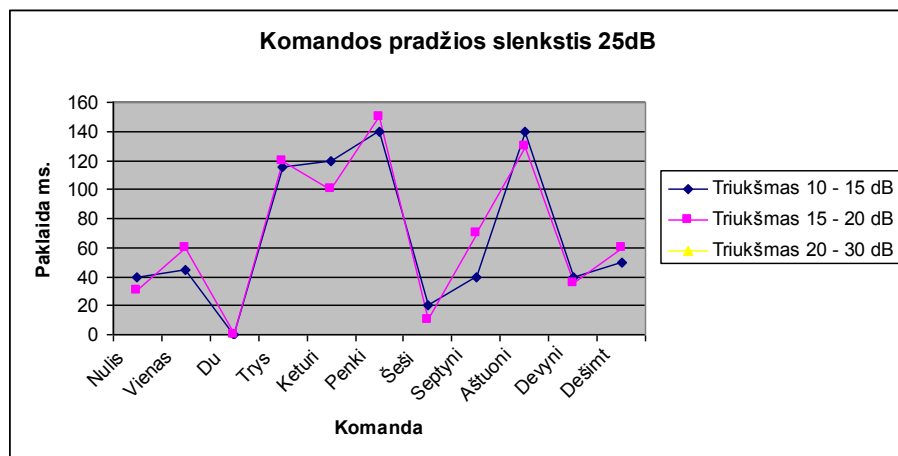


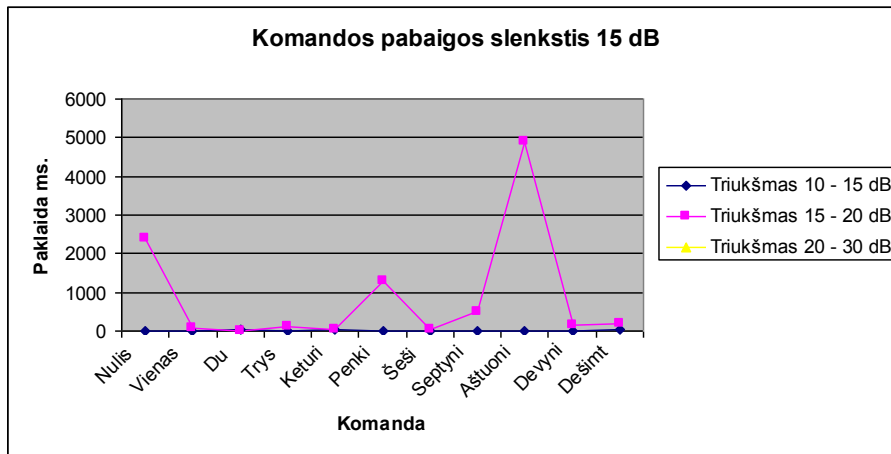
Atpažinimo slenkščio reikšmė 35 – 25 dB.



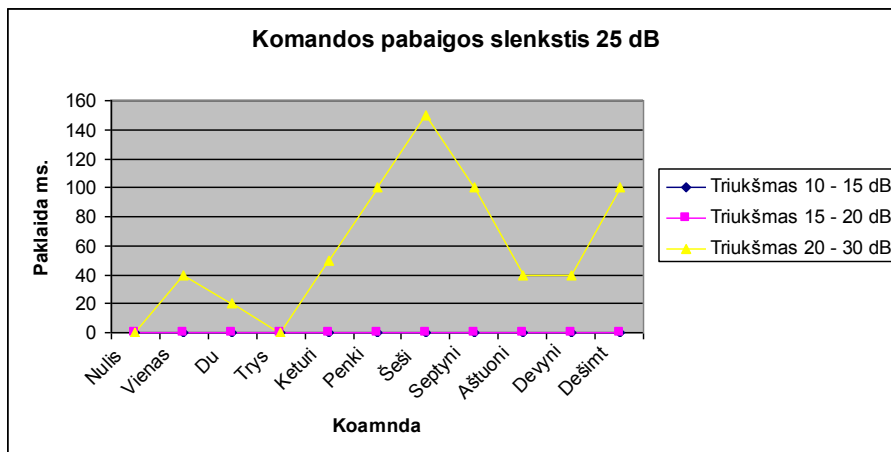
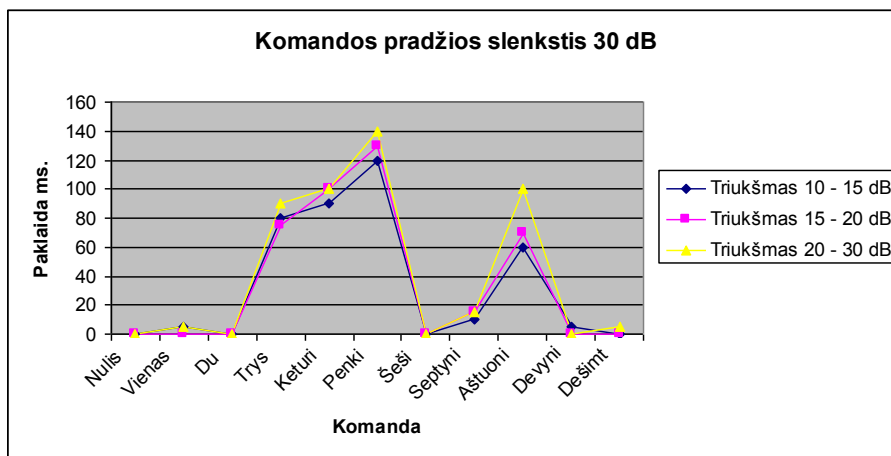
Antras asmuo.

Atpažinimo slenkščio reikšmė 25 – 15 dB.

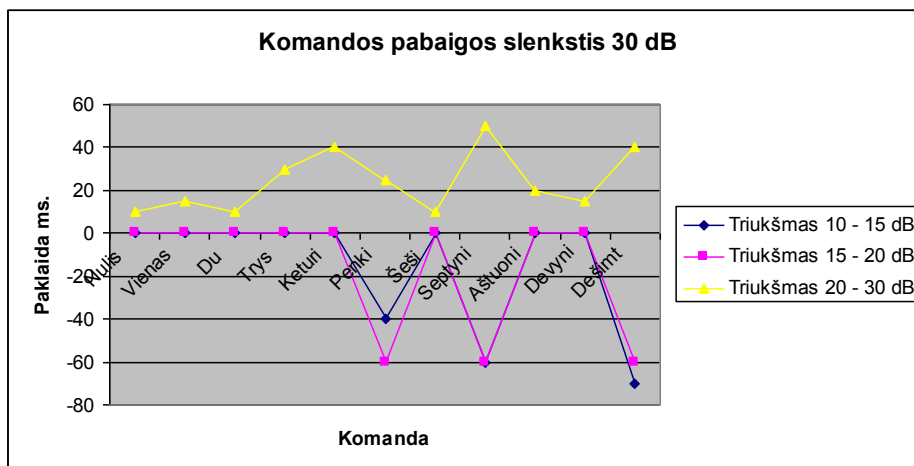
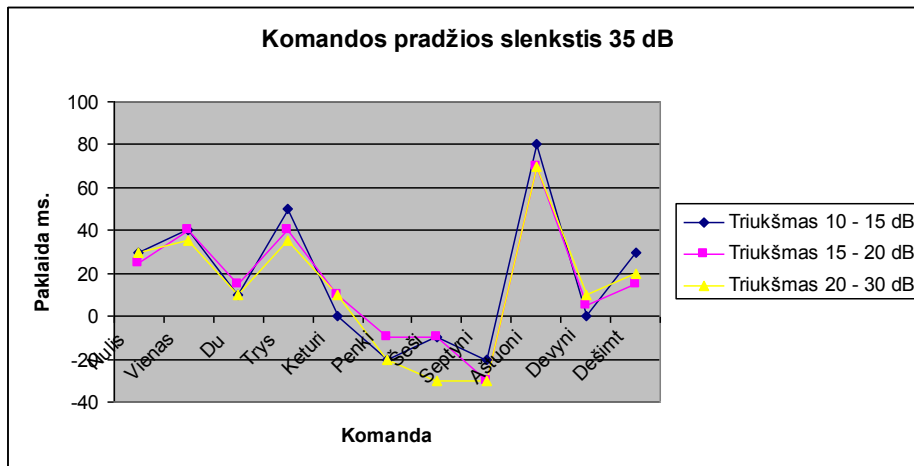




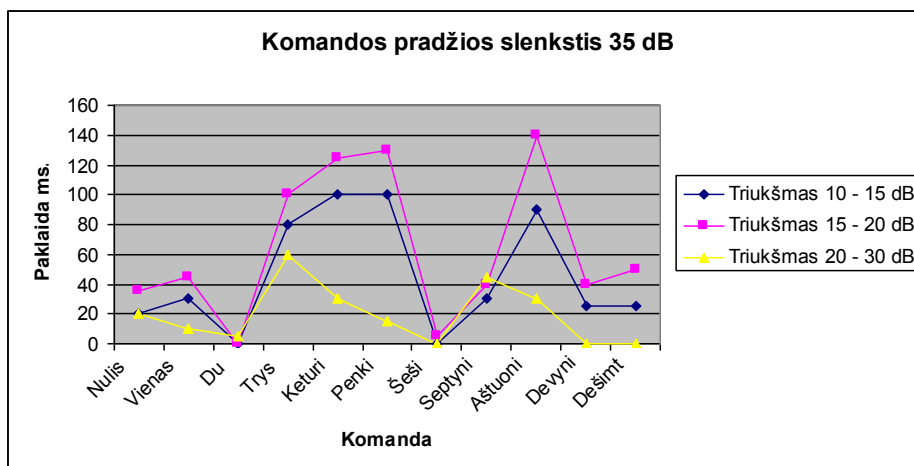
Atpažinimo slenksčio reikšmė 30 – 25 dB.

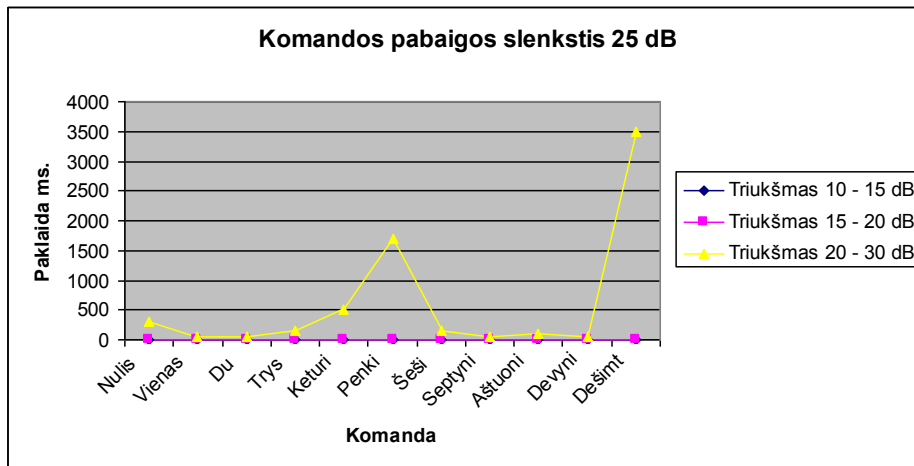


Atpažinimo slenksčio reikšmė 35 – 30 dB.



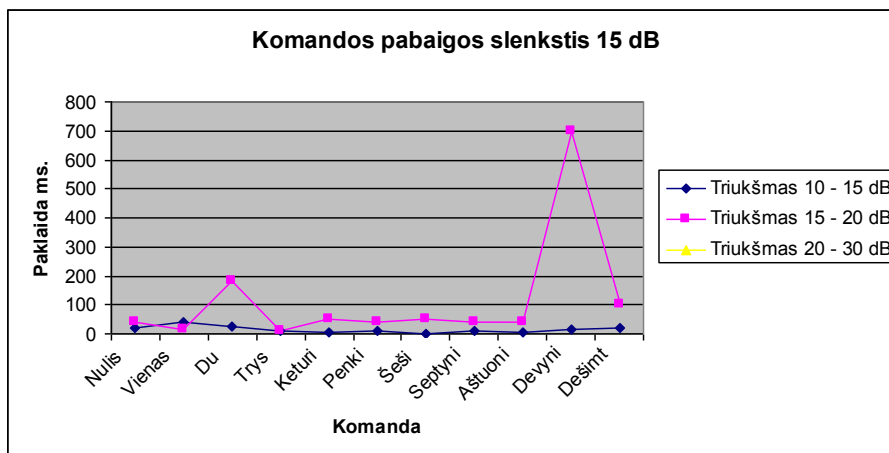
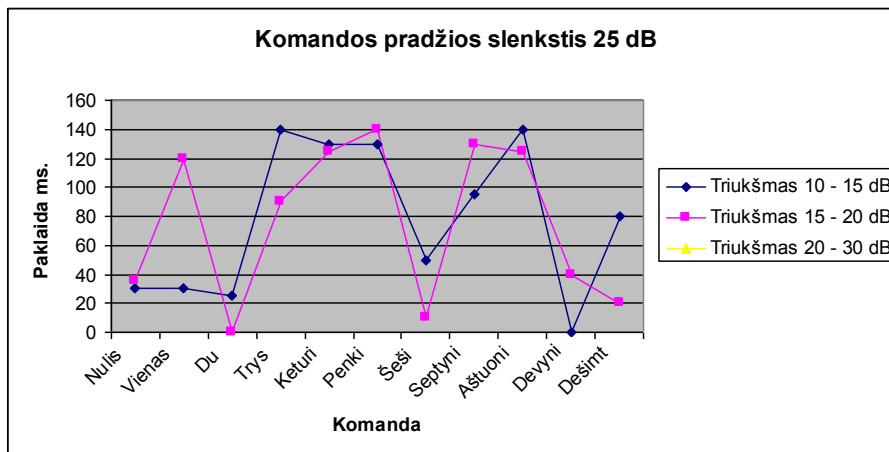
Atpažinimo slenkščio reikšmė 35 – 25 dB.



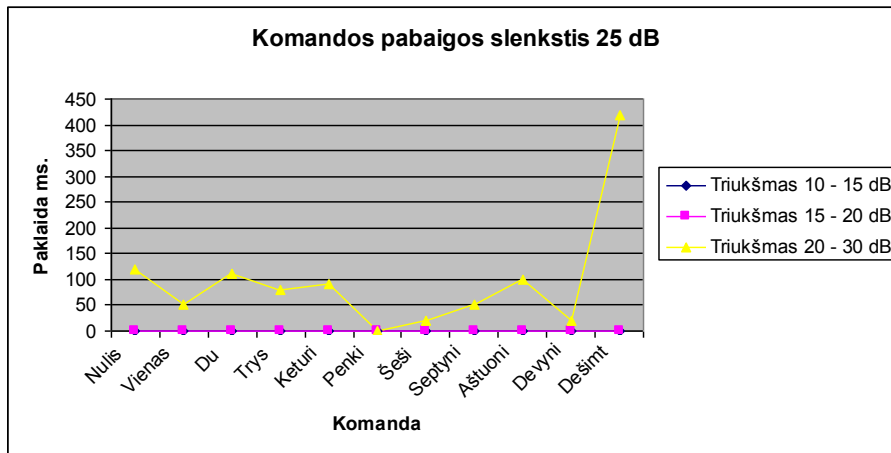
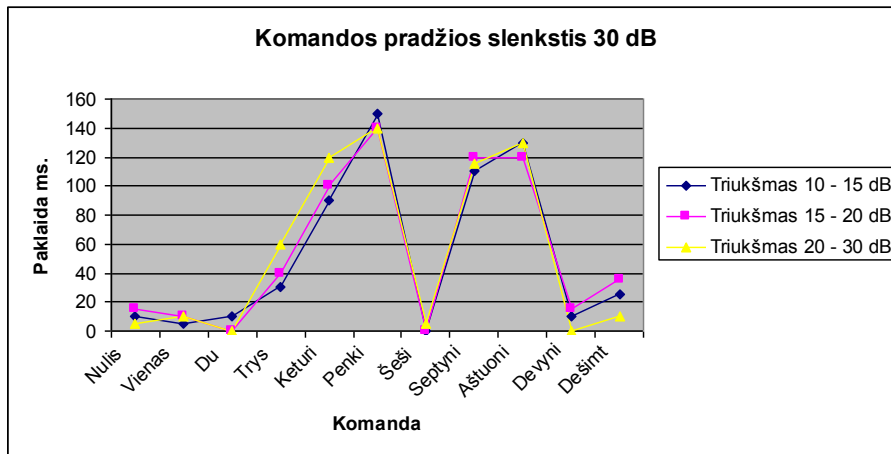


Trečias asmuo.

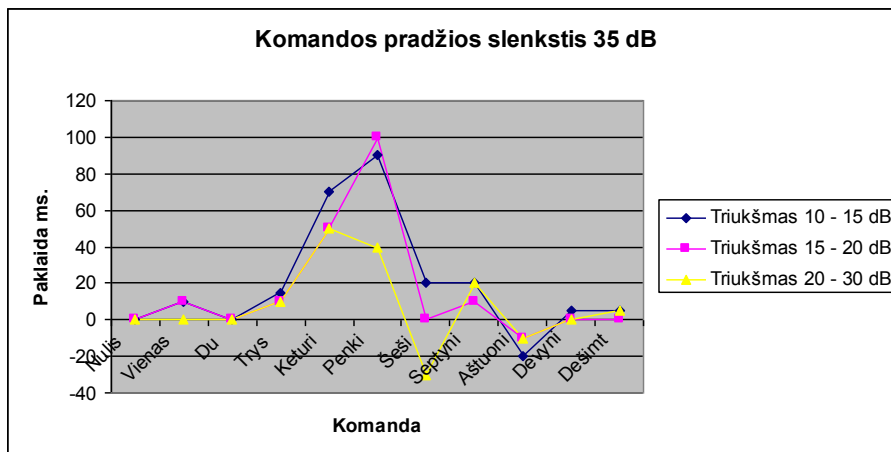
Atpažinimo slenksčio reikšmė 25 – 15 dB.

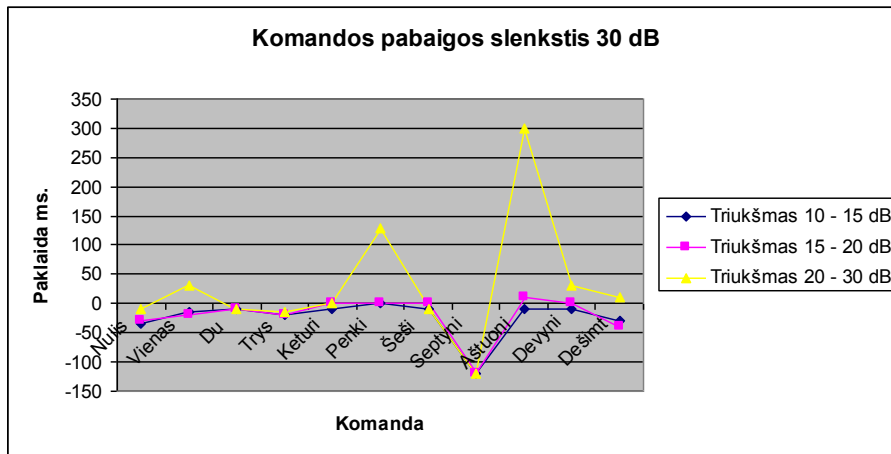


Atpažinimo slenksčio reikšmė 30 – 25 dB.

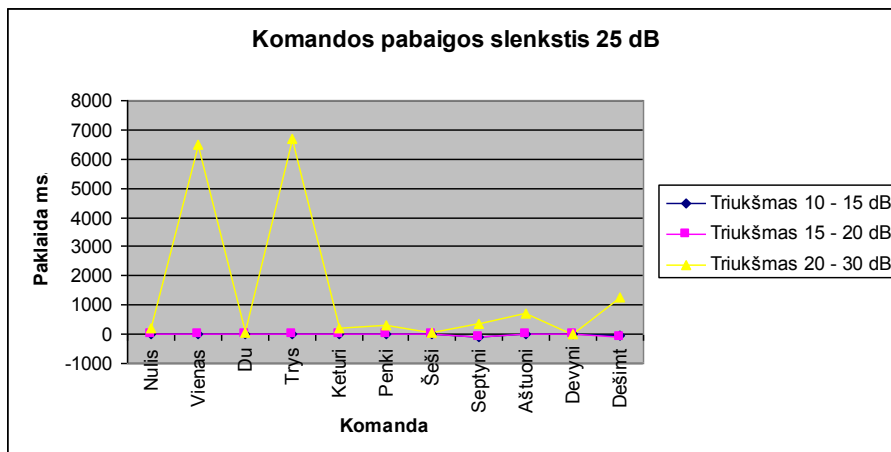
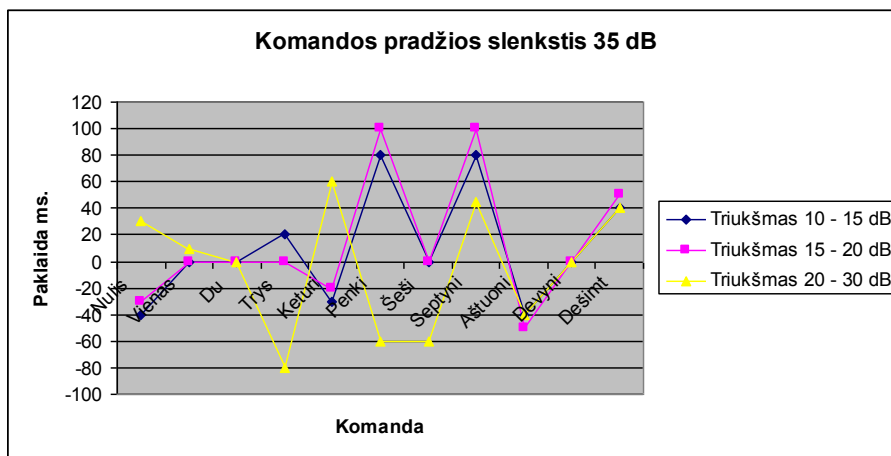


Atpažinimo slenkščio reikšmė 35 – 30 dB.



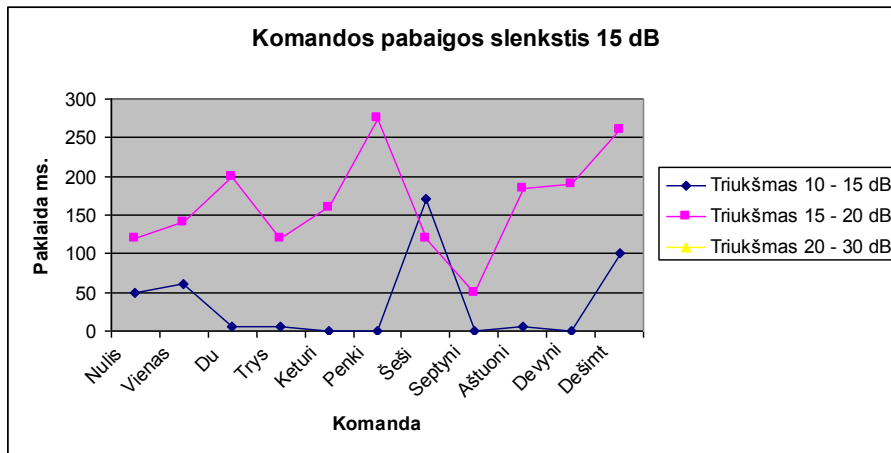
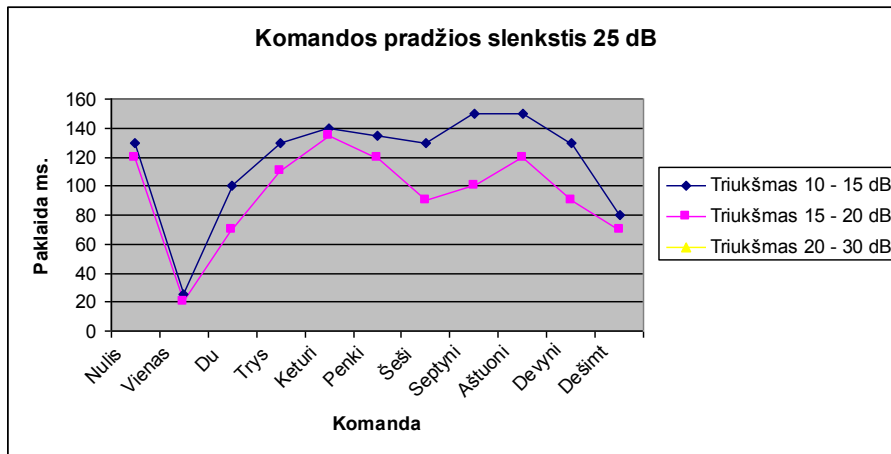


Atpažinimo slenkščio reikšmė 35 – 25 dB.

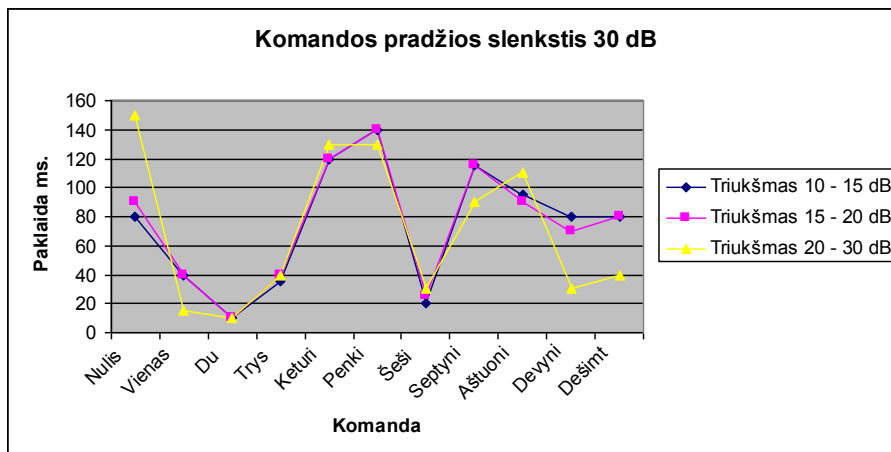


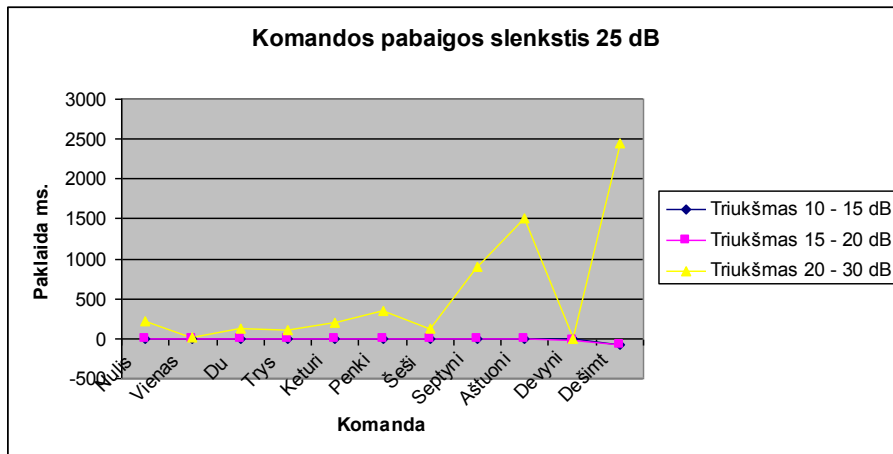
Ketvirtas asmuo.

Atpažinimo slenkščio reikšmė 25 – 15 dB.

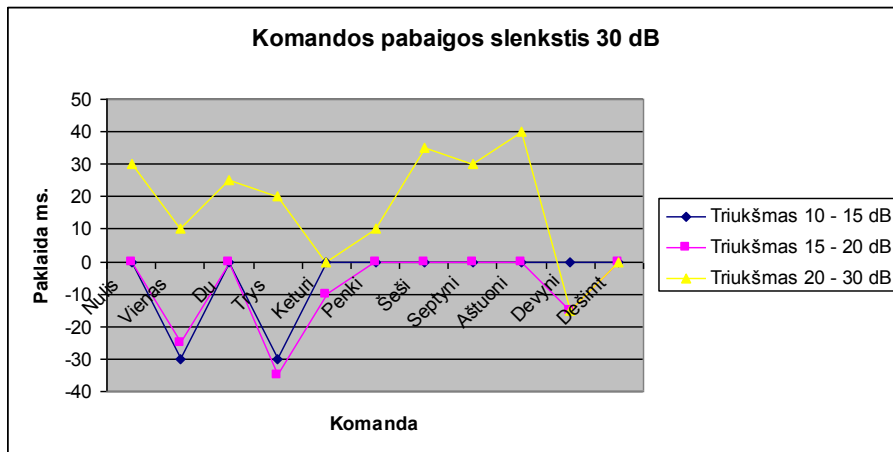
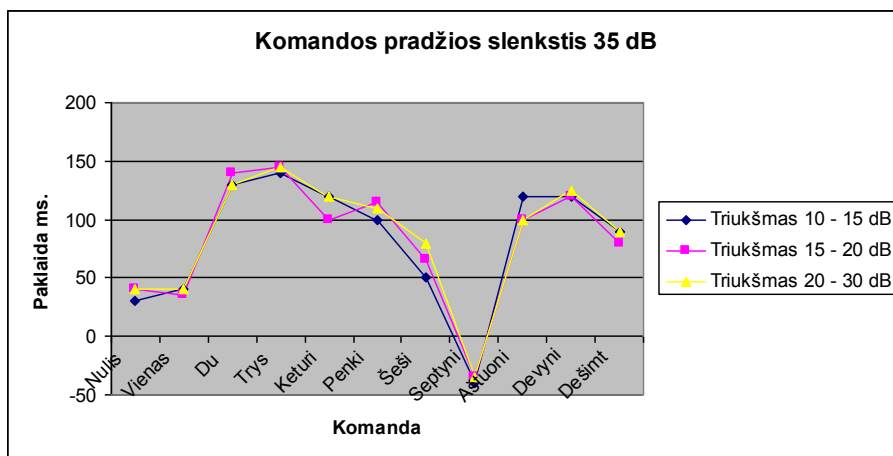


Atpažinimo slenkščio reikšmė 30 – 25 dB.

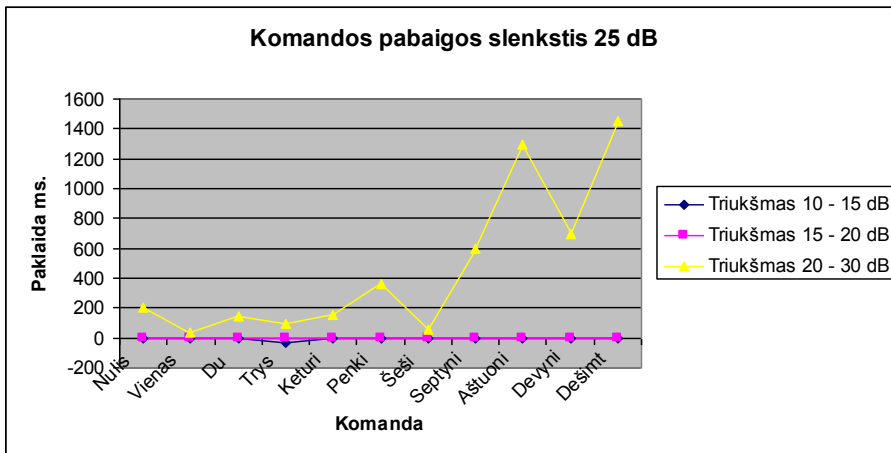
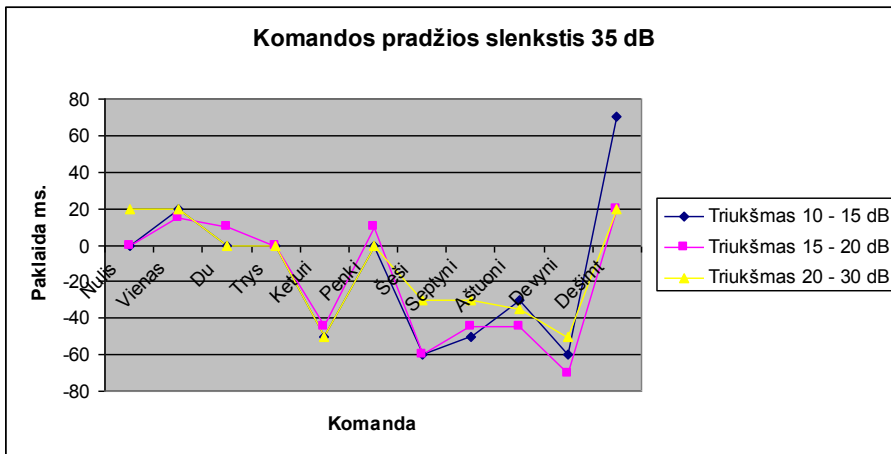




Atpažinimo slenksčio reikšmė 35 – 30 dB.



Atpažinimo slenksčio reikšmė 35 – 25 dB.



3 Priedas. Dalis gramatikos failo turinio.

```
<GRAMMAR LANGID="409">
  <DEFINE>
    <ID NAME="RID_Nulis_0" VAL="0"></ID>
    <ID NAME="RID_Nulis_1" VAL="1"></ID>
    <ID NAME="RID_Nulis_2" VAL="2"></ID>
    <ID NAME="RID_Nulis_3" VAL="3"></ID>
    <ID NAME="RID_Nulis_4" VAL="4"></ID>
    <ID NAME="RID_Vienas_0" VAL="5"></ID>
    <ID NAME="RID_Vienas_1" VAL="6"></ID>
    <ID NAME="RID_Vienas_2" VAL="7"></ID>
    <ID NAME="RID_Vienas_3" VAL="8"></ID>
    <ID NAME="RID_Vienas_4" VAL="9"></ID>
    <ID NAME="RID_Vienas_5" VAL="10"></ID>
    <ID NAME="RID_Du_0" VAL="11"></ID>
    <ID NAME="RID_Du_1" VAL="12"></ID>
    <ID NAME="RID_Du_2" VAL="13"></ID>
    <ID NAME="RID_Du_3" VAL="14"></ID>
    <ID NAME="RID_Du_4" VAL="15"></ID>
    <ID NAME="RID_Du_5" VAL="16"></ID>
    <ID NAME="RID_Du_6" VAL="17"></ID>
    <ID NAME="RID_Trys_0" VAL="18"></ID>
    <ID NAME="RID_Trys_1" VAL="19"></ID>
    <ID NAME="RID_Trys_2" VAL="20"></ID>
    <ID NAME="RID_Trys_3" VAL="21"></ID>
    <ID NAME="RID_Trys_4" VAL="22"></ID>
    <ID NAME="RID_Trys_5" VAL="23"></ID>
    <ID NAME="RID_Keturi_0" VAL="24"></ID>
    <ID NAME="RID_Keturi_1" VAL="25"></ID>
    <ID NAME="RID_Keturi_2" VAL="26"></ID>
    <ID NAME="RID_Keturi_3" VAL="27"></ID>
    <ID NAME="RID_Keturi_4" VAL="28"></ID>
    <ID NAME="RID_Keturi_5" VAL="29"></ID>
    <ID NAME="RID_Penki_0" VAL="30"></ID>
    <ID NAME="RID_Penki_1" VAL="31"></ID>
    <ID NAME="RID_Penki_2" VAL="32"></ID>
    <ID NAME="RID_Penki_3" VAL="33"></ID>
    <ID NAME="RID_Penki_4" VAL="34"></ID>
    <ID NAME="RID_Sesi_0" VAL="35"></ID>
    <ID NAME="RID_Sesi_1" VAL="36"></ID>
    <ID NAME="RID_Sesi_2" VAL="37"></ID>
    <ID NAME="RID_Sesi_3" VAL="38"></ID>
    <ID NAME="RID_Sesi_4" VAL="39"></ID>
    <ID NAME="RID_Septyni_0" VAL="40"></ID>
    <ID NAME="RID_Septyni_1" VAL="41"></ID>
    <ID NAME="RID_Septyni_2" VAL="42"></ID>
    <ID NAME="RID_Septyni_3" VAL="43"></ID>
    <ID NAME="RID_Septyni_4" VAL="44"></ID>
    <ID NAME="RID_Septyni_5" VAL="45"></ID>
    <ID NAME="RID_Septyni_6" VAL="46"></ID>
    <ID NAME="RID_Astuoni_0" VAL="47"></ID>
    <ID NAME="RID_Astuoni_1" VAL="48"></ID>
    <ID NAME="RID_Astuoni_2" VAL="49"></ID>
    <ID NAME="RID_Astuoni_3" VAL="50"></ID>
    <ID NAME="RID_Astuoni_4" VAL="51"></ID>
    <ID NAME="RID_Astuoni_5" VAL="52"></ID>
    <ID NAME="RID_Devyni_0" VAL="53"></ID>
    <ID NAME="RID_Devyni_1" VAL="54"></ID>
    <ID NAME="RID_Devyni_2" VAL="55"></ID>
    <ID NAME="RID_Devyni_3" VAL="56"></ID>
    <ID NAME="RID_Devyni_4" VAL="57"></ID>
    <ID NAME="RID_Devyni_5" VAL="58"></ID>
  </DEFINE>
  <RULE NAME="Nuhlis" ID="RID_Nulis_0" TOPLEVEL="ACTIVE">
    <P>Nuhlis</P>
  </RULE>
  <RULE NAME="N uh l ih s" ID="RID_Nulis_1" TOPLEVEL="ACTIVE">
    <P PRON="N uh l ih s">N uh l ih s</P>
  </RULE>
  <RULE NAME="Noolis" ID="RID_Nulis_2" TOPLEVEL="ACTIVE">
```

```

        <P>Noolis</P>
    </RULE>
    <RULE NAME="Nolihs" ID="RID_Nulis_3" TOPLEVEL="ACTIVE">
        <P>Nolihs</P>
    </RULE>
    <RULE NAME="N uw l ih s" ID="RID_Nulis_4" TOPLEVEL="ACTIVE">
        <P PRON="N uw l ih s">N uw l ih s</P>
    </RULE>

    <RULE NAME="Viehnas" ID="RID_Vienas_0" TOPLEVEL="ACTIVE">
        <P>Viehnas</P>
    </RULE>
    <RULE NAME="V ih ey n ah s" ID="RID_Vienas_1" TOPLEVEL="ACTIVE">
        <P PRON="V ih ey n ah s">V ih ey n ah s</P>
    </RULE>
    <RULE NAME="Viehnahs" ID="RID_Vienas_2" TOPLEVEL="ACTIVE">
        <P>Viehnahs</P>

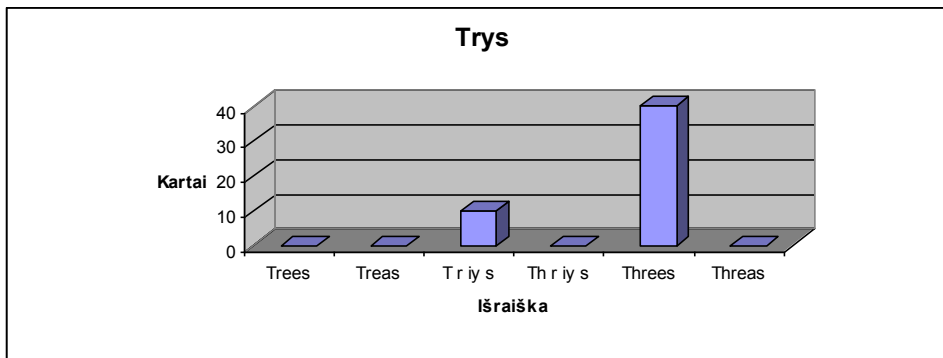
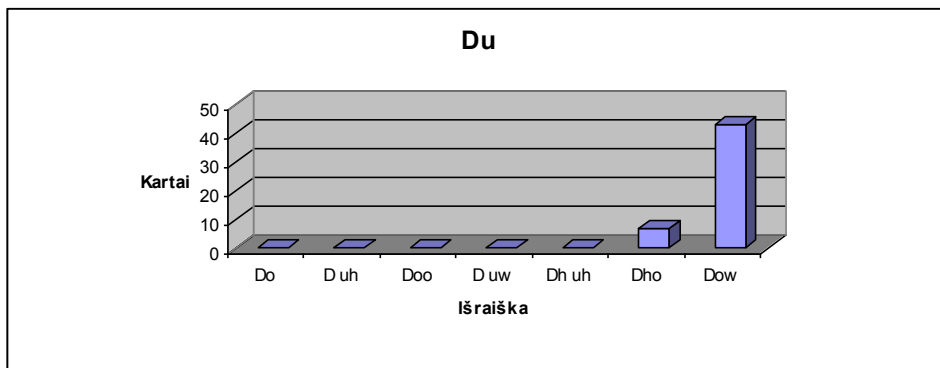
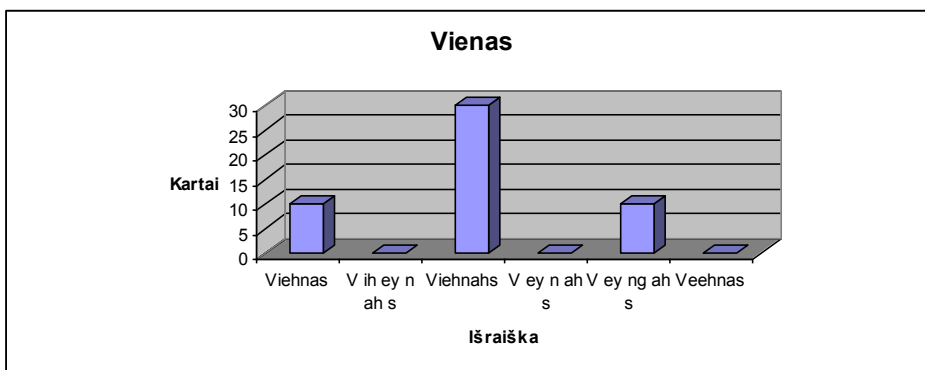
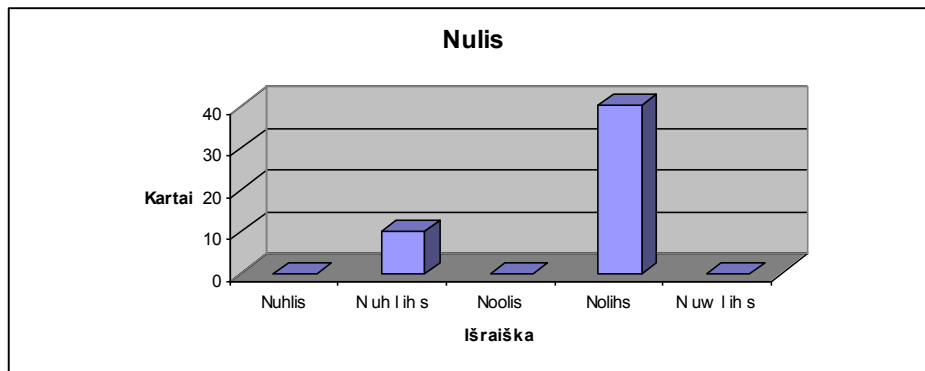
    <RULE NAME="K ey t uh r ih" ID="RID_Keturi_4" TOPLEVEL="ACTIVE">
        <P PRON="K ey t uh r ih">K ey t uh r ih</P>
    </RULE>
    <RULE NAME="Ketoori" ID="RID_Keturi_5" TOPLEVEL="ACTIVE">
        <P>Ketoori</P>
    </RULE>

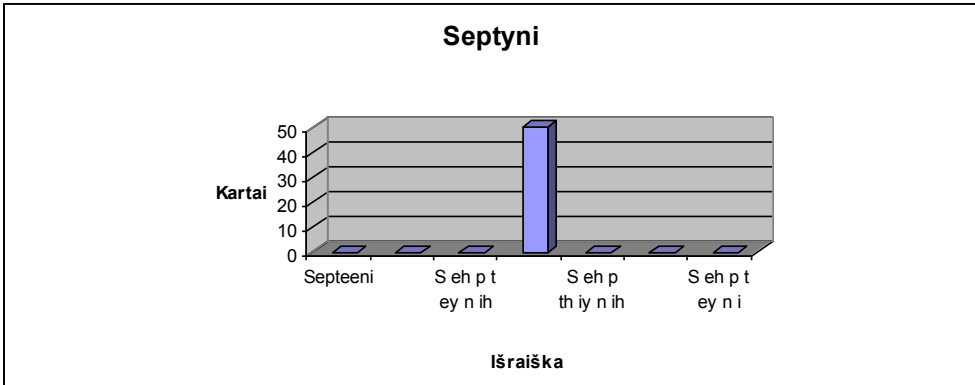
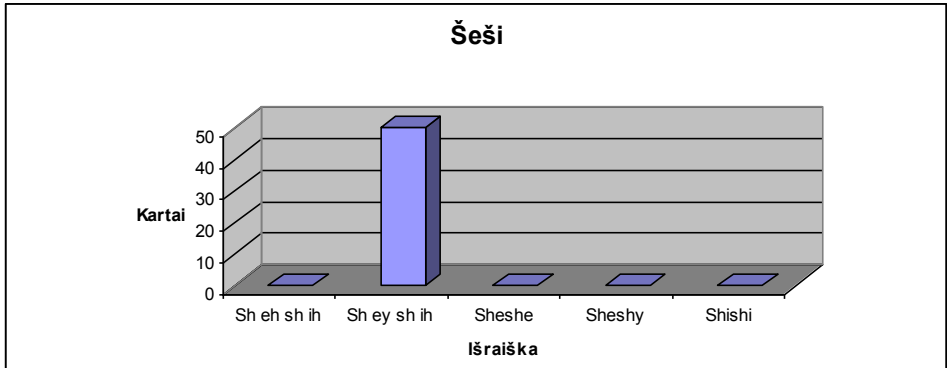
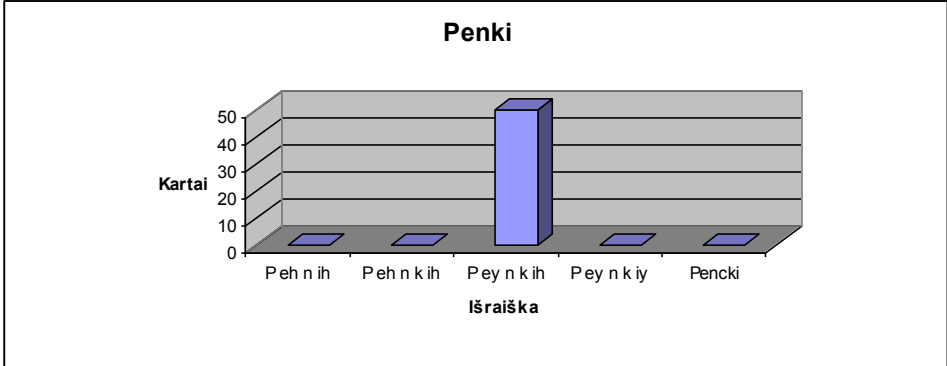
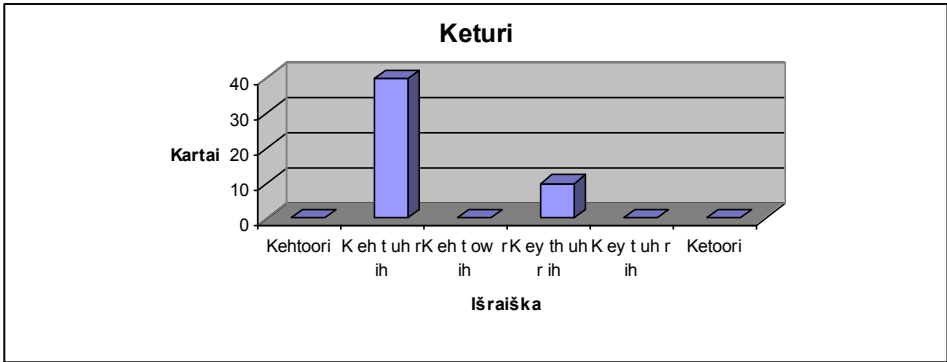
    <RULE NAME="P eh n ih" ID="RID_Penki_0" TOPLEVEL="ACTIVE">
        <P PRON="P eh n ih">P eh n ih</P>
    </RULE>
    <RULE NAME="P eh n k ih" ID="RID_Penki_1" TOPLEVEL="ACTIVE">
        <P PRON="P eh n k ih">P eh n k ih</P>
    </RULE>
    <RULE NAME="Septeeni" ID="RID_Septyni_0" TOPLEVEL="ACTIVE">
        <P>Septeeni</P>
    </RULE>
    <RULE NAME="S eh p th ey n ih" ID="RID_Septyni_1" TOPLEVEL="ACTIVE">
        <P PRON="S eh p th ey n ih">S eh p th ey n ih</P>
    </RULE>
    <RULE NAME="S eh p t ey n ih" ID="RID_Septyni_2" TOPLEVEL="ACTIVE">
        <P PRON="S eh p t ey n ih">S eh p t ey n ih</P>
    </RULE>
    <RULE NAME="S eh p t iy n ih" ID="RID_Septyni_3" TOPLEVEL="ACTIVE">
        <P PRON="S eh p t iy n ih">S eh p t iy n ih</P>
    </RULE>
    <RULE NAME="S eh p th iy n ih" ID="RID_Septyni_4" TOPLEVEL="ACTIVE">
        <P PRON="S eh p th iy n ih">S eh p th iy n ih</P>
    </RULE>
    <RULE NAME="S eh p th iy n i" ID="RID_Septyni_5" TOPLEVEL="ACTIVE">
        <P PRON="S eh p th iy n i">S eh p th iy n i</P>
    </RULE>
    <RULE NAME="S eh p t ey n i" ID="RID_Septyni_6" TOPLEVEL="ACTIVE">
        <P PRON="S eh p t ey n i">S eh p t ey n i</P>
    </RULE>

    <RULE NAME="Ae sh th uh ao n ih" ID="RID_Astuoni_0" TOPLEVEL="ACTIVE">
        <P PRON="Ae sh th uh ao n ih">Ae sh th uh ao n ih</P>
    </RULE>
    <RULE NAME="Ae sh th uh ao ng ih" ID="RID_Astuoni_1" TOPLEVEL="ACTIVE">
        <P PRON="Ae sh th uh ao ng ih">Ae sh th uh ao ng ih</P>
    </RULE>
    <RULE NAME="Ah sh t uh n ih" ID="RID_Astuoni_2" TOPLEVEL="ACTIVE">
        <P PRON="Ah sh t uh n ih">Ah sh t uh n ih</P>
    </RULE>
    .
    .
    .
    .
    </GRAMMAR>

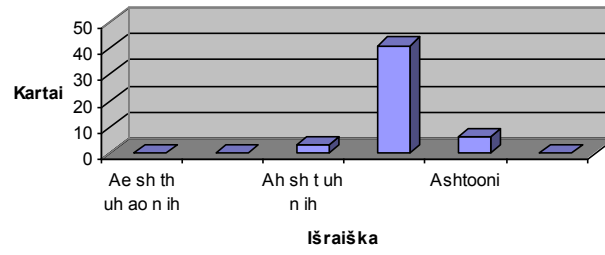
```

4 Priedas. Eksperimentinės programinės įrangos atpažinimo rezultatai.





Aštuoni



Devyni

