

**KAUNO TECHNOLOGIJOS UNIVERSITETAS**  
**ELEKTROS IR ELEKTRONIKOS FAKULTETAS**

**Ignas Makauskas**

**SKAIČIŲ IR VARDŲ GARSYNŲ ATPAŽINIMO,  
NAUDOJANT ŽODŽIAIS GRĮSTUS PASLĖPTUOSIUS  
MARKOVO MODELIUS, TYRIMAS**

Baigiamasis magistro projektas

**Vadovas**

Doc. dr. Kastytis Ratkevičius

**KAUNAS, 2017**

**KAUNO TECHNOLOGIJOS UNIVERSITETAS**  
**ELEKTROS IR ELEKTRONIKOS FAKULTETAS**  
**AUTOMATIKOS KATEDRA**

**SKAIČIŲ IR VARDŲ GARSYNŲ ATPAŽINIMO, NAUDOJANT  
ŽODŽIAIS GRĮSTUS PASLĖPTUOSIUS MARKOVO  
MODELIUS, TYRIMAS**

Baigiamasis magistro projektas  
Valdymo technologijos (621H66001)

**Vadovas**

(parašas) Doc. dr. Kastytis Ratkevičius  
(data)

**Recenzentas**

(parašas) Doc. A. Knyš  
(data)

**Projektą atliko**

(parašas) Ignas Makauskas  
(data)



KAUNO TECHNOLOGIJOS UNIVERSITETAS

ELEKTROS IR ELEKTRONIKOS FAKULTETAS

Ignas Makauskas

(Studento vardas, pavardė)

Valdymo technologijos (621H66001)

(Studijų programos pavadinimas, kodas)

Baigiamojo projekto

„Skaičių ir vardų garsynų atpažinimo, naudojant paslėptuosius Markovo modelius, tyrimas“

**AKADEMINIO SAŽININGUMO DEKLARACIJA**

20 17 m. Gegužės 26 d.  
Kaunas

Patvirtinu, kad mano **Igno Makausko** baigiamasis projektas tema „Skaičių ir vardų atpažinimo, naudojant paslėptuosius Markovo modelius, tyrimas“ yra parašytas visiškai savarankiškai, o visi pateikti duomenys ar tyrimų rezultatai yra teisingi ir gauti sąžiningai. Šiame darbe nei viena dalis nėra plagijuota nuo jokių spausdintinių ar internetinių šaltinių, visos kitų šaltinių tiesioginės ir netiesioginės citatos nurodytos literatūros nuorodose. Įstatymų nenumatytų piniginių sumų už šį darbą niekam nesu mokėjęs.

Aš suprantu, kad išaiškėjus nesąžiningumo faktui, man bus taikomos nuobaudos, remiantis Kauno technologijos universitete galiojančia tvarka.

\_\_\_\_\_  
(vardą ir pavardę įrašyti ranka)

\_\_\_\_\_  
(parašas)

Makauskas, Ignas. Skaičių ir vardų garsynų atpažinimo, naudojant paslėptuosius Markovo modelius, tyrimas. Valdymo sistemų magistro baigiamasis projektas / vadovas doc. dr. Kastytis Ratkevičius; Kauno technologijos universitetas, Elektros ir elektronikos fakultetas, Automatikos katedra.

Mokslo kryptis ir sritis: Elektros ir elektronikos inžinerija, Technologiniai mokslai

Reikšminiai žodžiai: *Paslėptieji Markovo modeliai, HTK, WEKA, šnekos atpažinimas*

Kaunas, 2017. 50 p.

## SANTRAUKA

Baigiamajame darbe tiriamas lietuviškai ištartų skaičių ir vardų atpažinimas naudojant paslėptuosius Markovo modelius ir kitakalbį atpažintuvą. Atliekamas atpažintuvų sujungimo galimybių tyrimas.

Informacijos šaltinių analizės dalyje nagrinėjami šnekamosios kalbos atpažinimo metodai, paslėptieji Markovo modeliai, „HTK“ programinis paketas. Apžvelgiamos lietuvių kalbos atpažinimo sistemos.

Metodinėje dalyje aptariamas garsynų paruošimas tyrimams. Nusakoma žodžių ir vardų atpažinimo tyrimams naudojamo „HTK“ paketo ir atpažintuvų sujungimo tyrimams naudojamos programos „WEKA“ metodikos.

Tyrimu rezultatų dalyje pateikiami lietuviškai ištartų skaičių ir vardų atpažinimo rezultatai naudojant paslėptuosius Markovo metodus. Rezultatai gaunami keičiant būsenų ir Gauso mišinių skaičių Markovo modeliams. Pateikiami kitakalbio atpažintuvo tikslumo rezultatai. Atliekamas atpažintuvų sujungimo tyrimas naudojant skirtingus sprendimų priėmimo klasifikatorius. Gauti rezultatai išanalizuojami ir palyginami.

Makauskas, Ignas. Investigation of digits and names corpora recognition by words based hidden Markov models. Final project of control systems Master's degree / supervisor doc. dr. Kastytis Ratkevičius; Kaunas University of Technology, Faculty of Electrical and Electronics Engineering, department of Automation.

Research area and field: Electrical and Electronics Engineering, Technological Sciences

Key words: Hidden *Markov Models*, *HTK*, *WEKA*, *Speech Recognition*

Kaunas, 2017. 50 p.

## SUMMARY

The final project presents research of numbers and names, spoken in Lithuanian, recognition system using hidden Markov models and Spanish speech recognizer. Possibilities to combine two different speech recognizers are also presented.

In the source of information analytic part speech recognition methods, hidden Markov models and its toolkit are described. Lithuanian speech recognition systems are analyzed.

In the methodological part preparation of speech recording database is described. Hidden Markov model toolkit and "WEKA" programs are analyzed.

In research results part Lithuanian numbers and names recognition using hidden Markov models results are presented. They are obtained by changing number of states and Gaussian mixtures in hidden Markov models. Results of using Spanish speech identifier for the same words are also presented. Two different speech recognizers are combined into one system with decision making block for better recognizer selection. Different classifiers are used to control decision making block. The obtained results are compared and analyzed.

## TURINYS

ĮVADAS.....	7
1. Informacinių šaltinių analizė.....	9
1.1. Automatinis šnekamosios kalbos atpažinimas .....	9
1.2. Kalbos ir balso signalų apdorojimo istorija .....	10
1.3. Šnekamosios kalbos samprata .....	12
1.4. Šnekamosios kalbos atpažinimo metodai ir jų klasifikavimas.....	13
1.5. Paslėptieji Markovo modeliai ir jų taikymas šnekos atpažinimui .....	16
1.6. Žodžių modeliais grįstas šnekamosios kalbos atpažinimas.....	18
1.7. HTK programinė įranga.....	19
1.8. Garsynai.....	21
1.9. Lietuvių kalbos atpažintuvų apžvalga.....	22
2. Metodinė dalis .....	26
2.1. Vardų ir skaičių garsynų paruošimas atpažinimo tyrimams .....	26
2.2. Hidden Markov Model Toolkit taikymas.....	27
2.3. HTK panaudojimas šnekos atpažinimui taikant žodžių modelių metodą .....	28
2.4. Vardų ir skaičių atpažinimo tyrimo, naudojant žodžiais grįstus modelius, metodika.....	32
2.5. Vardų ir skaičių atpažinimo tyrimo, naudojant Ispanišką atpažintuvą, metodika.....	33
2.6. Dviejų skirtingų šnekos atpažintuvų sujungimas .....	34
3. Tyrimų rezultatų dalis.....	38
3.1. Skaičių ir vardų atpažinimo tyrimas naudojant žodžiais grįstus modelius.....	38
3.2. Skaičių ir vardų atpažinimo tyrimas naudojant ispanišką šnekos atpažintuvą .....	42
3.3. Atpažintuvų sujungimo galimybių tyrimas .....	42
4. Išvados ir rezultatai .....	48
5. Literatūros šaltiniai .....	49
6. Priedai .....	51

## ĮVADAS

Šnekamoji kalba žmonėms yra natūrali bendravimo ir informacijos pateikimo priemonė. Tačiau kompiuterinės sistemoms ne taip lengvai kaip žmogus supranta garsu pateikiamą informaciją. Vystantis technologijoms atsiranda poreikis valdyti šias kompiuterines sistemas mums natūraliausiu būdu – balsu. Todėl yra kuriamos automatinės šnekos atpažinimo (AŠA) sistemos, galinčios suprasti garsiniu signalu gautą informaciją.

Automatinio kalbos atpažinimo sistemos sparčiai tobulėja ir jas jau galime pastebėti kasdieniniame gyvenime. Kiekvienas naujas išmanusis telefonas turi AŠA funkcijas (*Google assistant, Siri, Bixby*). Taip pat kalbos atpažinimo taikymą kasdien galime pastebėti automobiliuose, navigacinėse sistemose, išmaniuosiuose namuose ar net laisvų rankų įrangoje pritaikyti neįgaliesiems. Kalbos atpažinimo sistemos yra naudojamos karo pramonėje, medicinoje, teismų procesų įrašymui, kalbų mokymosi kursuose, automatinėse subtitrų sudarymo sistemose ar net „NASA“ kosminiuose projektuose. Todėl galima teigti, kad tokių sistemų pritaikymo galimybės yra labai plačios. Didelės kompanijos kaip „*Google Inc*“ investuoja daugybę resursų tobulinant kalbos atpažinimo technologijas. Didinant apmokymo garsynų apimtį ši kompanija jau šiandien pasiekia labai didelius atpažinimo tikslumus net ir nepopuliarioms kalboms. Tačiau, net ir „*Google Inc*“ atpažintuvai nepasiekia tokio gero atpažinimo kaip įprasta žmogaus klausa, todėl šių sistemų tobulinimas yra būtinas.

Šiame darbe yra atliekamas lietuviškai tariamų skaičių ir vardų atpažinimo tyrimas naudojant paslėptuosius Markovo modelius ir ispanišką atpažintuvą. Ankstesni tyrimai buvo atliekami atskirai atpažįstant skaičius ir vardus. Šiuo atveju siekiamą sudaryti bendrą skaičių ir vardų atpažinimo sistemą. Darbe naudojamą garsyną sudaro 10 skirtingų skaičių ir 26 skirtingos vardų komandos (viso 16920 skirtingų balso įrašų). Garsyną sudarė 51 diktorius (21 – vardams, 30 – skaičiams).

**Darbo tikslas** – atlikti lietuviškai tariamų skaičių ir vardų atpažinimo tyrimus naudojant HTK programinį paketą ir kitakalbį atpažintuvą. Sudaryti hibridinę skaičių ir vardų atpažinimo sistemą.

**Darbo objektas** – lietuviškai tariamų skaičių ir vardų daugiadiktoriai garsynai.

**Darbo tyrimo metodai.** Atpažinimo tyrimai taikant paslėptuosius Markovo modelius atliekama HTK programiniu paketu. Rezultatai gaunami keičiant būsenų ir Gauso mišinių skaičių Markovo eilutėse. Kitakalbio atpažintuvo tyrimai atliekami „*Microsoft Speech Recognizer 8.0*“ programa, ją pritaikius lietuvių kalbos supratimui. Atpažintuvų sujungimo galimybių tyrimai atlikti „*WEKA*“ duomenų analizės programa. Požymių atrankai taikome skirtingus klasifikatorius ir tiriame jų efektyvumą.

**Darbo aktualumas.** Darbo rezultatus galima panaudoti sudarant žodinėmis komandomis tariamų ligų atpažintuvą. Gydytojai pildant formas nusako ne pilnus ligų pavadinimus, bet ligų kodus aprašytus TLK-10-AM. Panaudojus darbe sudarytus atpažintuvus galima nusakyti paciento vardą ir ligos kodą.

**Darbo struktūra.** Darbas susideda iš įvado, informacinių šaltinių analizės, metodinės dalies, tyrimų rezultatų dalies, išvadų ir rezultatų, literatūros šaltinių ir priedų.

Informacijos šaltinių analizės dalyje nagrinėjami šnekamosios kalbos atpažinimo metodai, paslėptieji Markovo modeliai, „HTK“ programinis paketas. Apžvelgiamos lietuvių kalbos atpažinimo sistemos. Metodinėje dalyje aptariamas garsynų paruošimas tyrimams. Nusakoma žodžių ir vardų atpažinimo tyrimams naudojamo „HTK“ paketo ir atpažintuvų sujungimo tyrimams naudojamos programos „WEKA“ metodikos. Nusakoma tyrimų eiga. Tyrimų rezultatų dalyje pateikiami ir analizuojami gauti rezultatai. Išvadose yra pateikti apibendrinti tyrimų rezultatai.



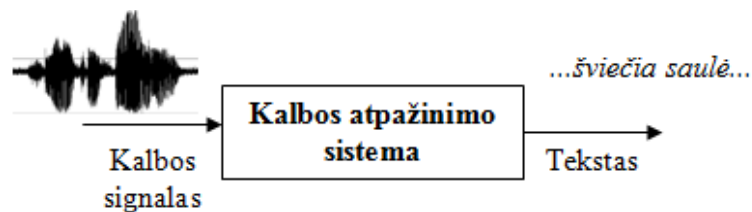
# INFORMACINIŲ ŠALTINIŲ ANALIZĖ

## 1.1. AUTOMATINIS ŠNEKAMOSIOS KALBOS ATPAŽINIMAS

Šnekamoji kalba – tai žmogaus plaučių, burnos ertmės ir balsu stygų sukuriamas garsas. Žmogui šnekamosios kalbos suvokimas atrodo lengvas ir natūralus reiškinys, tačiau kompiuteriui automatinis šnekos atpažinimas (AŠA) yra sudėtingas procesas. Pagrindinis automatinio šnekos atpažinimo sistemų kūrimo ir tobulinimo tikslas yra sukurti mašinas, kurios galėtų girdėti, suprasti, kalbėti ir veikti pagal balsu gautą informaciją. Kalbos atpažinimo technologijos gali padėti valdyti kompiuterius ir jų sistemas balsu. Toks valdymo metodas yra natūralesnis nei valdymas įprastinėmis valdymo priemonėmis, kaip pelė ar klaviatūra. [1]

Automatinis kalbos atpažinimas tai garsinio šnekos signalo vertimas į šį signalą atitinkantį tekstą ar kitokią balsu signalą atitinkančią formą. Automatinė šnekos atpažinimo sistema – tai aparatinė ir programinė įranga be pagalbos galinti suvokti ir atpažinti šnekamąją kalbą ir balsu komandas. Tokios sistemos yra kuriamos ir tobulinamos tam, kad kompiuteriai galėtų suprasti žmonių šneką, veikti pagal šnekamąsias komandas ar net kalbėti žmonių kalba. [2]

Kalbos atpažinimo procesas, pavaizduotas 1.1 paveiksle. Sistema gavusi garsinį kalbos signalą paverčia jį teksteine ar kompiuterinio kodo atitikmenimis.[1]



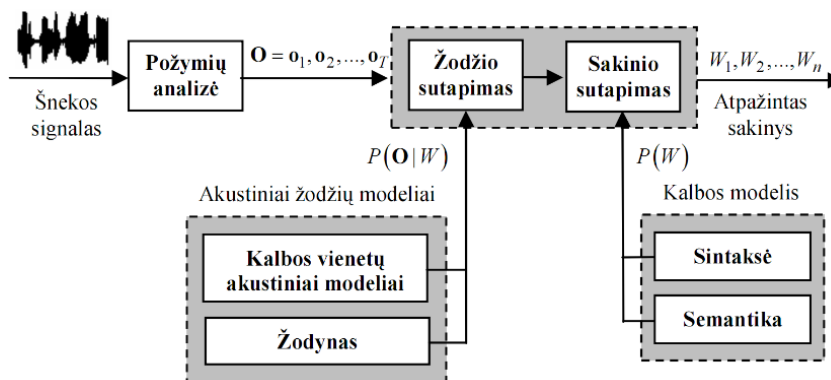
1.1 pav. Šnekamosios kalbos atpažinimo procesas [4]

Šnekamosios kalbos supratimas technologijoms yra komplikotas uždavinys, reikalaujantis perprasti sudėtingą fizikinį procesą, priklausantį nuo konkrečios kalbos, kalbinės aplinkos, diktoriaus artikuliacinių savybių ir kitų faktorių, lemiančių šnekos atpažinimo sudėtingumą.

Kalbos atpažinimo procesas yra komplikotas uždavinys, reikalaujantis suprasti sudėtingą fizikinį procesą. Šnekos atpažinimo kokybė priklauso nuo kalbėtojo kalbos, artikuliacinių savybių ar net aplinkos kurioje kalbą diktoriaus. Kompiuteris turi atsižvelgti į visus šiuos faktorius norint kokybiškai atpažinti šnekamąją kalbą.

Siekiant identifikuoti kas yra sakoma iš kalbos signalo turi būti išskiriami požymiai, būdingi tik analizuojamam signalo fragmento turiniui ir kuriuos matematinių bei statistinių algoritmų pagalba galima analizuoti informacinėse sistemose. [3]

Bendroji šnekamosios kalbos atpažinimo schema pateikiama 1.2 paveiksle.



**1.2 pav.** Šnekamosios kalbos atpažinimo sistemos schema [4]

Pagrindinėmis automatinės šnekamosios kalbos atpažinimo sistemos dalimis laikoma:

- požymių išskyrimas iš šnekos signalo;
- garsinę informaciją atitinkančių modelių (kalbos ir akustinių) formavimas;
- nežinomo ištariamo klasifikavimas vienam iš atitinkančių modelių — žodžių ir sakinių atpažinimas. [4]

Automatinės šnekamosios kalbos atpažinimo technologijos pagal jų sudėtingumą ir poreikius gali būti skirstomos į:

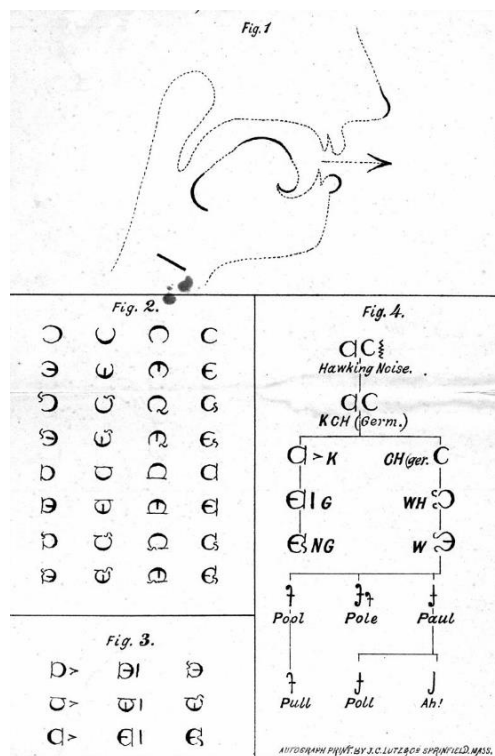
- kalbos sintezės – įgarsina įrašytą tekstą;
- gramatiškai apribotos šnekos – taikoma klientų aptarnavimo sferoje;
- kalbos atpažinimo arba šnekos identifikavimo – garsinis signalas verčiamas į tekstą;
- natūralios arba ištisinės šnekos – sudėtingiausios formos kalbos atpažinimo įranga – taikoma norint suprasti ištartas ilgas kalbos komandas. [5]

Technologinių įrenginių pažanga ir vis didesnis naudojimas reikalauja kokybiškų automatinių kalbos atpažinimo sistemų. Nors AŠA sistemos vis dar neatpažįsta kalbos taip gerai kaip žmogus, tačiau pastaraisiais dešimtmečiais technologijos paklausumas didėja. Didelį postūmį šioms technologijoms suteikė išmaniųjų telefonų programinės įrangos gamintojai norėdami užtikrinti lengvesni ir natūralesni įrenginio valdymą. Atsiranda vis daugiau šnekos atpažinimo sistemų kurios vis geriau sprendžia atpažinimo uždavinius.

## 1.2. KALBOS IR BALSŲ SIGNALŲ APDOROJIMO ISTORIJA

Kalbos garsų analizės pradžia galima laikyti prieš beveik šimtmetį „Alexander Melville Bell“ sukurtą tariamų žodžių vizualinį atvaizdavimą, kuris leido aprašyti žymiai daugiau

informacijos nei bet kuriame tuometiniame žodyne naudojami kirčiavimo ženklai. Ši mokslininko sukurta sistema buvo pavadinta „matoma kalba“ (angl.: Visible Speech). Šios sistemos iliustracija matoma 1.3 paveiksle. [3]



1.3 pav. Matomos kalbos iliustracija [3]

Pirmuoju automatiniu kalbos atpažintuvu yra laikomas ATT&T's Bell Labs 1930 m. pasiūlytas elektroninės kalbos sintezatorius. Šis sintezatorius buvo naudojamas kariniams tikslams siekiant atpažinti kas yra sakoma. 1940 m. ATT&T's Bell Labs mokslininkai sukūrė kalbos signalo atvaizdavimo būdą, kuris ir šiandien plačiai naudojamas – tai spektrogramos. Tai vaizdinis kalbos signalo atvaizdavimo atvaizduojantis tris parametrus: dažnį, intensyvumą ir laiką. [3]

1990 metai buvo pristatytos pirmosios komercinės kalbos atpažinimo technologijos. Šiuo laikotarpiu kalbos atpažinimo žodžių bibliotekos jau buvo didesnės nei įprastas žmogaus žodynas. 1992 metais Karnegio Melono universitete buvo parodyta „Sphinx - II“ AŠA sistema. Ši sistema pirmoji nepriklausomai nuo kalbėtojo galėjo atpažinti nepertraukiamą kalbą.

Daugiausiai kalbos atpažinimo tyrimų atliekama anglų, rusų, ispanų ir kinų kalbomis. Mažiau paplitusioms kalboms kalbos atpažinimo tyrimų yra gerokai mažiau. Taip yra todėl, nes:

- pažangių kalbos atpažinimo technologijų rinka mažai paplitusioms kalboms yra nedidelė;

- mažiau paplitusios kalbos įprastai turi sudėtingesnę struktūrą;
- populiarios kalbos yra paprastesnės, nes laikui bėgant buvo supaprastintos dėl jų plataus paplitimo pasaulyje ir naudojimo. [6]

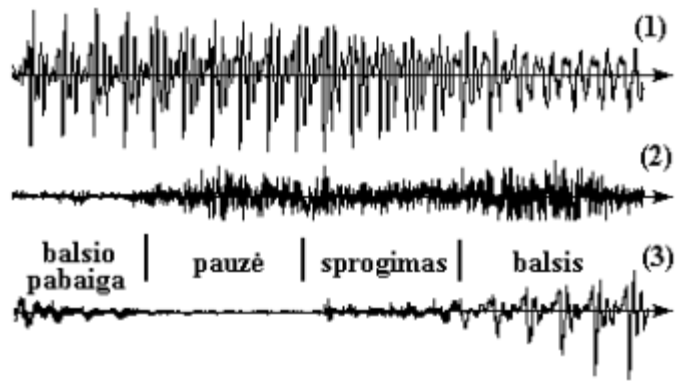
Pirmus kalbos signalų tyrimus Lietuvoje inicijavo P. Kemėšis bei L. Telksnys 20a. septintojo dešimtmečio viduryje. Palaipsniui susiformavo kelios grupės atliekančios mokslinius tiriamuosius darbus kalbos technologijų srityje: Kauno Technologijos universitetas (KTU) ir Vilniaus universitetas (VU), Vytauto Didžiojo universitetas (VDU), bei Matematikos informatikos institutas (MII). Vykdamas Lietuvių kalbos informacinėje visuomenėje 2000-2006 metų programą. 2000 metais pradėti sistemingi lietuvių kalbos automatinio atpažinimo tyrimai. Šiuose tyrimuose dalyvavo MII, Vytauto Didžiojo universiteto ir Lietuvos teismo ekspertizės centro mokslininkai. Tyrimų metu buvo dirbama ne tik kalbos, bet ir diktoriaus atpažinimo srityse. Šiuo metu atlikti diktoriaus atpažinimo tyrimai lietuvių kalba orientuoti į kalbančiojo išskyrimą naudojant Povilo Treigio ir Antano Lipeikos (2006) aprašomą vidutinio atstumo tarp tiriamojo ir lyginamųjų kalbėtojų nustatymą ir balso trakto charakteristikas. Vis dėlto Lietuvių kalbos apdorojimo tyrimai yra kuklesni ir mažiau išvystyti nei anglų, ar kitų plačiai tarptautiniame lygmenyje naudojamų kalbų. [3]

### 1.3. ŠNEKAMOSIOS KALBOS SAMPRATA

Šnekamoji kalba yra skirstoma į keturis skirtingus tipus:

- Izoliuoti pavieniai žodžiai (*isolated words*). Izoliuotų žodžių atpažinimo sistema atpažįsta vieną ištarimą, kuris yra apribotas ribojamą šnekos signalo charakteristikomis (žemo lygio triukšmais, tyla).
- Ištisinė šneka (*continuous speech*). Tokia šneka neturi lengvai suprantamų pauzių tarp žodžių.
- Sujungti žodžiai (*connected words*) – tai rišlios frazės. Naudojama kaip tarpinis variantas tarp ištisinės šnekos ir pavienių žodžių.
- Spontaniška šneka (*spontaneous speech*) – tai įprasta kasdienė šneka. Tokia šneka nėra sklandi, joje gali būti ilgų pauzių, nebaigtų tarti ar neteisingai ištarų žodžių, mikčiojimų ar neteisingos sintaksės sakinių. [1]

Kalbos signalai kinta laike. signalo kitimo būdas ir yra ta informacija, kurią žmonės naudoja komunikacijai. Lyginant su rašytine forma perduotu pranešimu, sakytiniame pranešime yra labai daug informacijos. [7]



1.4 pav. Kalbos signalų pavyzdžiai [7]

Šnekos analizės būdai naudojami projektuojant šnekos atpažinimo sistemas yra fonologija, kuri nagrinėja fonetinius kalbos elementus ir akustinė fonetika dažniausiai spektriniais metodais tiria patį garso signalą. Fonologija atsako į klausimą: kiek ir kokių fonetinių vienetų yra šnekamojoje kalboje, o akustinė fonetika pateikia šių vienetų charakteristikas spektriniais metodais. [gint]

Žodynas – tai AŠA sistemos naudojamų žodžių sąrašas su žodžių transkripcijomis ir jų akustiniais modeliais. Jis yra sudaromas iš apibrėžtų segmentinių kalbos vienetų aibių. Žodynas naudojamas tiek atpažinimo, tiek mokymo etapuose. Automatinių šnekos atpažinimo sistemų žodynai yra sudaromi pagal skaitmeninio formato kalbos transkripcijų žodynus. Naudojamo žodyno dydis ženkliai įtakoja AŠA sistemos tikslumą. Didelis žodynas leidžia turėti platų apmokymo masyvą taip didindamas tikslumą. [1]

#### 1.4. ŠNEKAMOSIOS KALBOS ATPAŽINIMO METODAI IR JŲ KLASIFIKAVIMAS

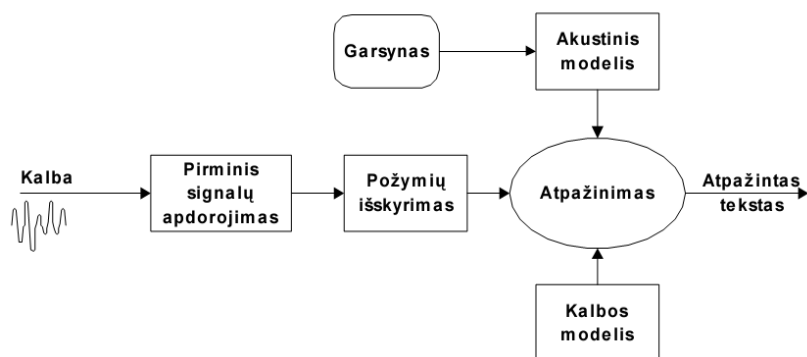
Automatinis šnekos atpažinimas yra sudėtingas procesas. Taip yra todėl nes garso signalas turi didelį kiekį informacijos. Pagrindinės priežastys įtakojančios atpažinimo sudėtingumą yra:

- Keletą kartų išstarto tam tikro garso akustinė realizacija labai skiriasi, net jei jį ištarė tas pats diktorius ir tame pačiame žodyje;
- Kalbėjimo greitis gali labai kisti, todėl skiriasi kelių to paties žodžių akustinių realizacijų ilgis. Kintant žodžių ilgiui atskirų garsų ilgis kinta netiesiškai;
- Garso akustinė realizacija priklauso nuo gretimų garsų, tai vadinama koartikuliacija;
- Kalbėjimo sraute nėra aiškių garsų ar žodžių ribų;
- Kiekvieno žmogaus tartis yra skirtinga, todėl reikalingas arba apmokymas konkrečiam diktoriui, arba sistema kūrimo metu turi būti apmokyta su kuo didesniu diktorių skaičiumi;

- Jei kuriama atpažinimo sistema remiasi žodžių atpažinimu, žodžių etalonų skaičius gali būti pernelyg didelis;
- Kalbėjimo sraute gali būti ir nekalbinių fragmentų (pvz., kosulys), kuriuos reikia atskirti ir pašalinti;
- Praktiniuose taikymuose papildomų problemų sukelia foninis triukšmas. [7]

Pagrindinis automatinės šnekos atpažinimo sistemos kokybės rodiklis yra šios sistemos atpažinimo tikslumas. Jis priklauso nuo sistemoje techniškai naudojamų algoritmų modelių ar taisyklių. Šie metodai įtakoja kaip sistema apdoroja akustinius modelius, priims sprendimus, išskirs požymius ar ieškos alternatyvų. 1.5 paveiklėse pateikta supaprastinta šnekos atpažinimo schema.

Atpažinimo tikslumą sąlygoja kalbų įvairovė, kalbėjimo aplinkos ir perdavimo kanalo neapibrėžtumai (triukšmas, iškraipymai) ir kitos kintančios aplinkybės apsunkinančios atpažinimo procesą. 1.3 paveiksle pateikiama supaprastinta automatinės šnekos atpažinimo schema.



**1.5 pav.** Automatinės šnekos atpažinimo schema [3]

AŠA sistema veikia gavusi garso signalą. Pirma įeinančiam kalbos signalui taikomas pirminis apdorojimas. Šio apdorojimo metu signalas diskretizuojamas ir skaidomas į vienodo ilgio intervalus. Po to pasitelkiant požymių apskaičiavimo algoritmus yra sudaromi požymių vektoriai. Galiausiai vyksta kalbos signalo atpažinimas palyginant jį su anksčiau sudarytais akustiniais ir kalbos modeliais. [3]

Pagrindiniai ir daugiausia naudojami šnekamosios kalbos atpažinimo metodai yra:

- Paslėptieji Markovo modeliai (PMM);
- Dirbtiniai neuroniniai tinklai (DNT);
- Jungtinis PMM ir DNT;
- Dinaminio programavimo.

Daugiausia kalbos atpažinimo sistemų yra pagrįsta paslėptųjų Markovo modelių teorija. Paslėptųjų Markovo modelių (angl. *Hidden Markov Models (HMM)*) metodas remiasi prielaida, kad kalbos signalas gali būti gerai apibūdintas kaip parametrinis atsitiktinis procesas ir stochastinio proceso parametrai gali būti tiksliai įvertinti. PMM šiuo metu yra pažangiausia technologija kalbos atpažinimui tačiau ji nėra ideali. Šių modelių atpažinimo ir apmokymo algoritmai turi esminių trūkumų. Jie kalbos signalą modeliuoja kaip tiesinę fonemų seką ir neatsižvelgia į šalutinius poveikius kaip tyla, triukšmai, mikčiojimai ar atodūsiai. Norint išvengti šių trūkumų naudojami kiti šnekamosios kalbos atpažinimo metodai. [5]

Dirbtinių neuroninių tinklų metodas kalbos atpažinimui taikomas kaip alternatyva Markovo modeliams. DNT (angl. *Artificial Neural Network (ANN)*) yra universalus atpažinimo uždaviniams naudojamas modeliavimo įrankis. Jis pasižymi didele duomenų atskiriamumo galia ir gali modeliuoti ypač sudėtingas sistemas. Dirbtinių neuroninių tinklų technologija sparčiai populiarėja, tačiau kalbos atpažinimui ji nėra ideali. DNT blogai modeliuoja ilgalaikes priklausomybes, kaip sakinyš ar frazė. Be to, neuroninių tinklų metodas reikalauja daug laiko procesui apmokyti. [6]

Jungtinis PMM ir DNT metodas yra taikomas kai norima išvengti pavienių šių sistemų trūkumų. Metodas yra pagrįsta paslėptųjų Markovo modelių ir dirbtinių neuroninių tinklų sujungimu į vientisą hibridinę sistemą. PMM yra naudojami šnekos signalų kitimo laike modeliavimui, o DNT – šnekos signalų akustiniam modeliavimui. Šis metodas gali pasiekti geresnius rezultatus, tačiau yra gerokai sudėtingesnis už pavienius metodus. [6]

Dinaminio laiko skalės iškraipymo (angl. *Dynamic Time Warping (DTW)*) metodas seniau buvo daugiausiai naudojamas šnekos atpažinimo būdas, tačiau jis buvo pakeistas sėkmingesniais aukščiau paminėtais metodais. Šis metodas yra pagrįsta algoritmu, kuris matuoja vektorių kintančių laike panašumus. Pagrindiniai trūkumai – didelės skaičiavimo apimtys ir sudėtingas apmokymo etalonų sudarymas. [8]

Automatiniam kalbos atpažinimui yra taikoma daugybe kitokių metodų nei aukščiau paminėti. Tačiau tie metodai nėra tokie populiarūs ir neužtikrina tokia gero atpažinimo tikslumo kaip PMM ar DNT metodai.

Šnekos atpažinimo sistemos yra skirstomos pagal naudojamus metodus ir algoritmus. Taip pat, klasifikuojama pagal diktorių skaičių, žodyno apimtį, kalbėjimo stilių.

**Pagal diktorių skaičių** yra skirstomos į nepriklausančias ir priklausančias nuo diktoriaus. Viendiktorinės sistemos gali atpažinti šneką tik to žmogaus pagal kurio balsą buvo sudarytas modelio apmokymas, o daugdiktorinės sistemos atpažįsta šneką nepriklausomai nuo diktoriaus.

**Pagal žodyno apimtį** AŠA sistemos skirstomos į siauros apimties žodyno (ji sudaro tik kelios dešimtis žodžių), vidutinės apimties (šimtai žodžių), didelės apimties (tūkstančiai žodžių) ir labai didelės apimties (dešimtys tūkstančių žodžių).

**Pagal kalbėjimo stilių** skirstomos pagal anksčiau minėtus šnekamosios kalbos tipus: izoliuotų žodžių, ištisinės šnekos, sujungtų žodžių, spontaniškos šnekos. [3]

## 1.5. PASLĖPTIEJI MARKOVO MODELIAI IR JŲ TAIKYMAS ŠNEKOS ATPAŽINIMUI

Paslėptieji Markovo modeliai (PMM) – tai yra populiarus statistinis Markovo modelis, kuriame sistema yra modeliuojama su paslėptais nariais, arba kitaip – būsenų kintamaisiais. Tai pats populiariausias šnekos atpažinimo modelis. Šio modelio teorija moksliniuose straipsniuose pasirodė šeštojo dešimtmečio gale, o panaudojimas šnekos atpažinimui įvyko po dešimties metų.

PMM modeliuoja atsitiktinius procesus. Atsitiktinis procesas juda būsenų seka, kiekvienoje būsenoje generuodamas kokį nors įvykį. Markovo modeliai yra skirstomi:

- Stebimas Markovo modelis. Viena būsena atitinka vieną įvykį, todėl stebint įvykių seką galima surasti kokia buvo būsenų seka.
- Paslėptasis Markovo modelis. Būsenų negalime stebėti, nes vienas įvykis yra tikimybinė būsenos funkcija. Viena būsena gali atitikti kelis įvykius. [1]

Šnekos atpažinimui naudojami deterministiniai ir stochastiniai signalų modeliai. Deterministiniai, tai tokie modeliai, kuriuose naudojamos jau žinomos signalo savybės norint nustatyti nežinomus signalo parametrus. Stochastiniai modeliai pasižymi tuo, kad signalų savybės yra aprašomos, kaip parametriniai atsitiktiniai procesai. Paslėptieji Markovo modeliai yra priskiriami prie stochastinių signalų modelių. [8]

Paslėptieji Markovo modeliai gerai tinka aprašyti laike kintančius procesus, jų savybių kitimą ir šio kitimo statines charakteristikas. Dėl šių priežasčių PMM puikiai tiko aprašyti garsinius šnekos signalus. Taikant PMM šnekamosios kalbos atpažinimui daroma prielaida, kad šneka gali būti pakankamai gerai aprašyta, kaip atsitiktinis parametrinis procesas ir šio stochastinio proceso parametrai gali būti pakankamai tiksliai nustatyti. [11]

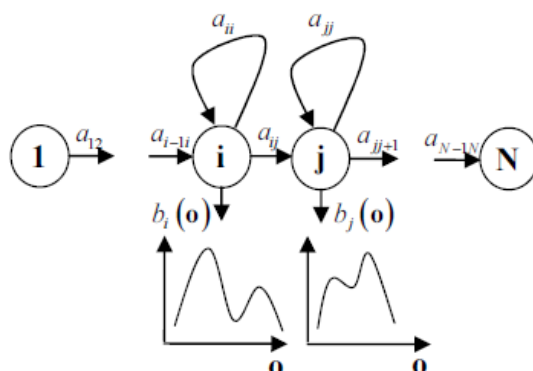
Paslėptieji Markovo modeliai yra aprašomi nusakant penkis dydžius:

- Būsenų skaičius,  $N$ ;
- Skirtingų stebėjimų būsenoje skaičių,  $D$ ;
- Tikimybių matrica nusakanti galimybę pereiti iš vienos būsenos į kitą,  $\mathbf{A}$ ;
- Stebėjimų tikimybinių skirstiniai,  $\mathbf{B}$ ;
- Praadinės buvimo būsenoje tikimybės.

Šie penki dydžiai ir nusako konkretų paslėptąjį Markovo modelį. [1]



Taigi PMM gali reprezentuoti šnekos signalą traktuojant jį kaip būsenų seką. Būsenos yra paslėptos ir jų negalima stebėti. Darant prielaidą, kad garsas gali būti skaidomas į tris dalis – pradžią, nekintančią dalį ir pabaigą, garsą įprasta aprašyti trimis būsenomis. Modelio pradžioje ir pabaigoje yra pridama po vieną papildomą neemituojančią (*non-emitting*) būseną, kurios sujungia kelis modelius. Struktūrinis PMM vaizdas pateiktas 1.6 paveiksle.



**1.6 pav.** Struktūrinis paslėptojo Markovo modelio vaizdas [1]

PMM tampa akustiniu modeliu tik po mokymo, kurio metu pagal stebėjimų duomenis įvertinami modelio parametrai, t. y. perėjimo iš būsenos į būseną tikimybių matrica ir stebėjimų tikimybės tankio funkcijos kiekvienai modelio būsenai. Ji yra išreiškiama stebėjimų arba požymių vektorių tikimybės tankio funkcija, nusakančia konkretaus požymio vektoriaus tikimybe.

Paslėptieji Markovo modeliai akustiniu modeliu tampa tik po mokymo proceso. Mokymo metu stebėjimo duomenys yra įvertinami modelio parametrais (**A,B**). Šie parametrai nusako perėjimo iš būsenos į būseną tikimybių matrica ir stebėjimų tikimybės tankio funkcijas kiekvienai modelio būsenai. Tikimybiniai skirstiniai gali būti įvairių formų:

- Tolydinis skirstinys (*continuous mixture density*);
- Diskretus skirstinys (*discrete density*);
- Pusiau-tolydaus (*semi-continuous density*). [10]

Kuriant atpažintuvą pirma yra nustatomos PMM aibės formos ir struktūros. Paskui yra prognozuojami PMM parametrai, remiantis apmokymu ir modeliuojamais duomenimis. Šnekos signalas įprastai yra koduojamas parametrinėmis formomis. PMM naudojant atpažinimui yra daroma prielaida, kad ištarimą atitinkanti vektorių seka yra lyginama apmokymų vektorių sekoms. Panašiausia seka remiantis geriausio sprendimo metodu identifikuoja ištarimą. [9]

Šnekos atpažinimo problema, kuriai spręsti taikomi PMM, yra keleto rišliai ištartų žodžių atpažinimas. Šiuo atveju atpažinimas yra pagrįstas pilnai ištartų žodžių modeliais. Atpažinimo problema (kai yra sudaromi pilnai ištartų žodžių modeliai) – rasti optimaliausią žodžių modelių seką, kuri būtų panašiausia (labiausiai atitiktų) į nežinomų žodžių sekas. [8]

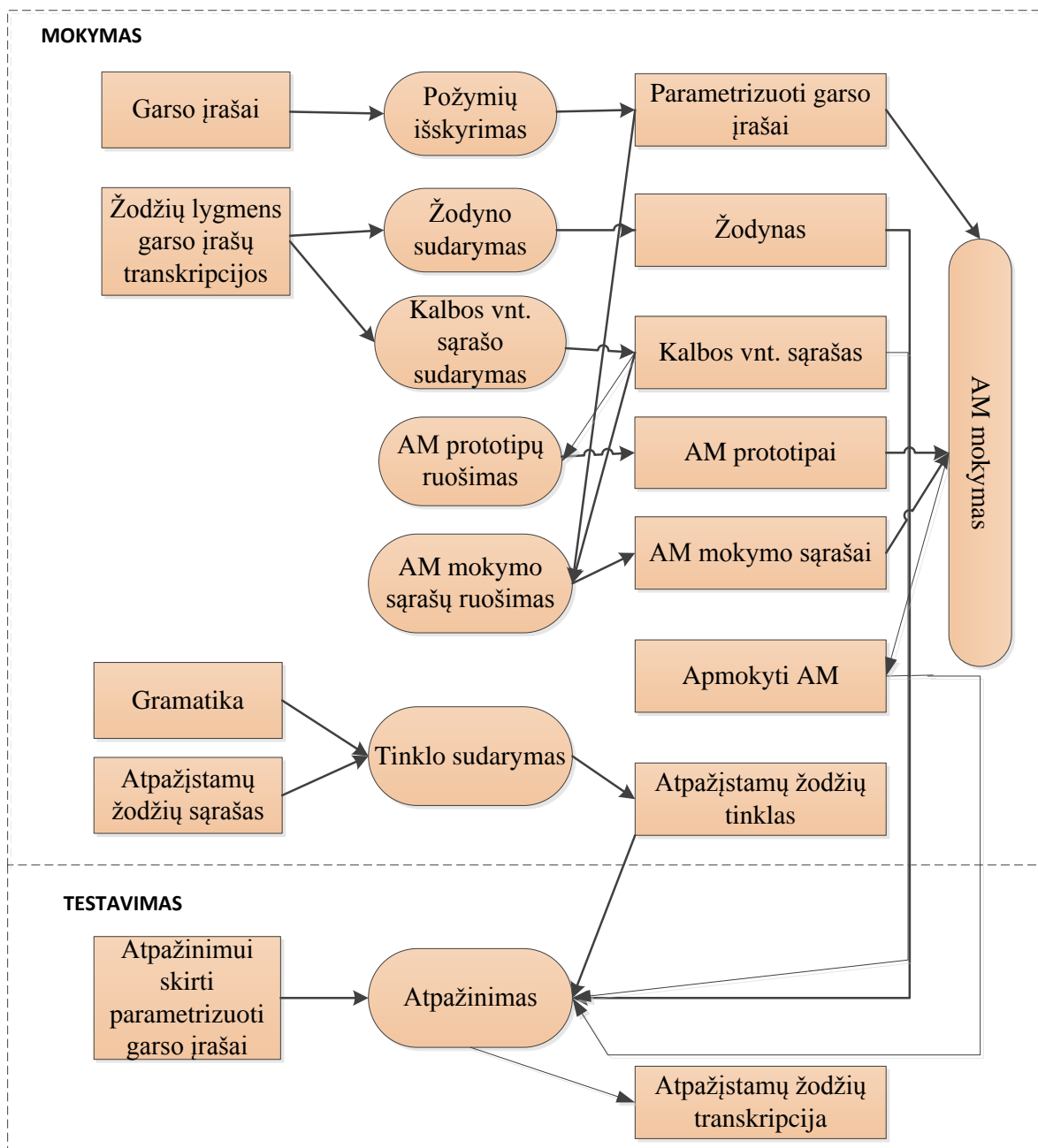
## 1.6. ŽODŽIŲ MODELIAIS GRĮSTAS ŠNEKAMOSIOS KALBOS ATPAŽINIMAS

Toks kalbos atpažinimas yra skirtas atpažinti izoliuotus žodžius. Šis atpažinimo metodas yra paprastesnis nei nepertraukiamas ar natūralios šnekos atpažinimas. Izoliuotų žodžių atpažinimas yra vykdomas dvejais etapais. Pirma yra naudojama spektrinė informacija šnekos pavyzdžiams palyginti. Paskui naudojant statistinius modelius turima spektrinė informacija yra apibendrinama.

Pirmosios izoliuotų pavienių žodžių atpažinimo sistemos buvo kuriamos pavyzdžiais grįsto šnekos atpažinimo principu. Tokios atpažinimo sistemos tipinis modelis yra sudarytas iš požymių išskyrimo, pavyzdžių palyginimo ir sprendimo priėmimo blokų. Toks atpažinimo modelis yra patogus dėl gero praktinio veikimo, nesudėtingo įgyvendinimo ir gero pavyzdžių palyginimo algoritmo. Jei atpažinimo sistema yra priklausoma nuo kalbėtojo, užtenka vieno ar dviejų kalbėtojo ištarimų spektrinės informacijos. Jeigu automatinė atpažinimo sistema nepriklauso nuo kalbėtojo yra atliekamas vektorinis kvantavimas (kalbėtojų ištarimų spektrinės informacijos pavyzdžiai suskirstomi i mažesnę pavyzdžių skaičių). Tokios AŠA atpažinimo tikslumas tiesiogiai priklauso nuo kalbėtojo, žodyno sudėtingumo. [1]

Automatinė izoliuotų žodžių atpažinimo schema taikant paslėptuosius Markovo modelius pateikta 1.7 paveiksle. Schemoje ovalai vaizduoja procesus, o stačiakampiai – procesams pateikiamus ar po procesų gaunamus objektus. Pagal šią schemą pradedant modeliuoti pavienių žodžių atpažinimo sistemą, „požymiu išskyrimo“ bloke yra išskiriami kiekvieno garso įrašo parametrai. „Žodyno sudarymo“ ir „Kalbos vienetų sudarymo“ blokuose iš žodžių lygmens garso įrašų transkripcijų yra sudaromi žodynai ir kalbos vienetų sąrašas. Turint kalbos vienetų sąrašą yra kuriami akustinių modelių (AM) prototipai ir paruošiami AM mokymo sąrašai. Pastaraisiais etapais gauti rezultatai kartu su parametrizuotais garso įrašais yra pradiniai duomenys AM mokymui, tai atliekama „AM mokymas“ bloke. Turint apmokytus akustinius modelius galime atlikti sistemos testavimą, jis vyksta testavimo bloke „atpažinimas“. Čia atpažinti skirti parametrizuoti garso įrašai yra verčiami į žodžių transkripcijas. Šiam procesui yra naudojami mokymo metu gauti žodynai, kalbos vienetai, akustiniai modeliai ir žodžių tinklai. [8]

Visi aukščiau aprašyti procesai yra vieni nuo kitų priklausomi, todėl juos sukeisti vietomis negalime. Visi žodžių modeliais grįsti atpažinimo procesai, išskyrus garso įrašų segmentavimą ir garsynų sudarymą yra atliekami paslėptų Markovo modelių programine įranga.



1.7 pav. Izoliuotų žodžių automatinės atpažinimo sistemos, mokymo ir testavimo etapų detalizavimo schema [8]

### 1.7. HTK PROGRAMINĖ ĮRANGA

Paslėptųjų Markovo modelių programinė įranga - HTK (angl. *Hidden Markov Model Toolkit*) – programinių priemonių rinkinys skirtas sudaryti ir valdyti PMM. Ši programinė įranga sukurtas Kembridžo universitete, kompiuterinio intelekto laboratorijoje. HTK yra daugiausiai

naudojama kalbos atpažinimo tyrimams. Tačiau šią programinę įrangą galime pritaikyti ir kitose srityse, pavyzdžiui, simbolių atpažinimo, kalbos sintezės ar DNR sekų tyrimuose.

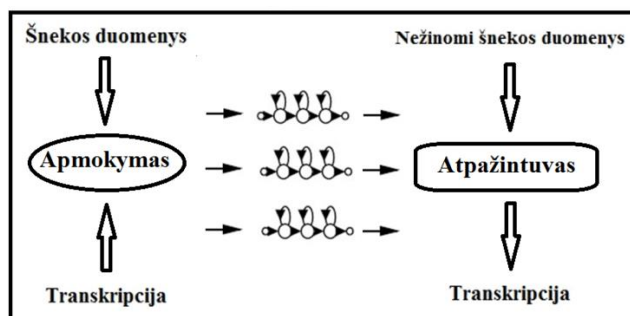
HTK programinį paketą sudaro programiniai įrankiai parašyti C programine kalba ir bibliotekų modelių rinkiniai. Šiame programiniame rinkinyje galima rasti priemonių, skirtų duomenims apdoroti, mokytį ir testuoti akustinius ir šnekos modelius, taip pat priemones akustiniams modeliams adaptuoti konkrečiam kalbėtojui bei rezultatų analizės priemones. [11]

„Hidden Markov Model Toolkit“ programinė įranga gali modeliuoti ištisines kalbos atpažinimo sistemas ar izoliuotus žodžius. HTK galima pasirinkti įvairius paslėptųjų Markovo modelių parametrus, topologijas ar signalą atitinkančius požymius. HTK priemonėms valdyti naudojama tekstinė sąsaja. Šnekos atpažinimo rezultatai, gaunami HTK programinėmis priemonėmis, būna pateikti tik skaitine forma. Todėl atpažinimo klaidų analizė yra pakankamai sudėtinga, nėra paprasta išsiaiškinti klaidų priežastis. [12]

HTK programinė įranga turi du pagrindinius etapus skirtus apdoroti šnekos signalą:

- HTK programiniame pakete šnekamosios kalbos atpažinimui apmokymo įrankiai yra naudojami siekiant apytiksliai apskaičiuoti paslėptųjų Markovo modelių parametrus, naudojant apmokymui pateiktus komandų įrašus ir parinktas komandų transkripcijas;
- Nežinomi komandų įrašai yra transkribuojami naudojant HTK atpažinimo įrankius. [11]

1.8 paveiksle pavaizduota schema, apibūdina šiuos du apdorojimo etapus.



1.8 pav. HTK šnekos apdorojimo schema [8]

Naudojant HTK programinį paketą šnekai atpažinti, kiekviename etape reikia panaudoti po atskirą programą iš HTK programinių įrankių paketo. Visos programos yra valdomos iš komandinės eilutės per papildomus failus ir eilutės komandas. [11]

Toliau pateikiami keturi pagrindiniai darbo su HTK programiniu paketu etapai žodžių atpažinimo modeliavimui:

1. **Žodyno sudarymas.** Žodyno sudarymas yra reikalingas tam, kad projektuojama sistema žinotų, kokios žodžių (akustinių vienetų) kombinacijos yra naudojamos atpažinimui;

2. **Požymių apskaičiavimas.** Atpažinimo sistemose naudojami iš balso komandų (žodžių) įrašų apskaičiuoti požymiai. Požymiams apskaičiuoti naudojama programa „Hcopy“, kuri naudodama įrašų failus apskaičiuoja požymius;

3. **Modelių apmokymas.** Apmokymas vyksta naudojant „HRest“ įrankį, kuris naudoja Baum-Welch algoritmą apmokymo procesui vykdyti. Svarbus apmokymo etapas – modelių failų kūrimas, kuriuose nurodomi pagrindiniai PMM modelio parametrai. Modelių failai – tai specialaus formato failai, kuriuose yra surašomi atitinkamo modelio parametrai: perėjimų tikimybės ir išėjimų tikimybės. HTK operuoja tik tolydiniais PMM modeliais, todėl išėjimų tikimybės modeliuojamos Gauso skirstiniais. Vadinasi išėjimų tikimybės aprašomos vidurkiais ir standartiniais nuokrypiais;

4. **Atpažinimas.** Atpažinimo testavimas vykdomas naudojant „HVite“ įrankį. Po apmokymo etapo, naudojant testavimui skirtus įrašus (naudojami kiti įrašai nei apmokymui) galima atlikti atpažinimo testavimą ir patikrinti atpažinimo kokybę su esamais modeliais. Atpažinimo kokybei gerinti iš naujo atliekami visi etapai nuo pirmojo, keičiant transkripcijas ir modelių parametrus. [2]

## 1.8. GARSYNAI

Kalbos inženerijos projektų sėkmė labiausiai priklauso nuo to, kokius duomenų išteklius turime. Šnekamosios kalbos atpažinimo tyrimams atlikti reikia turėti kalbos garsyna. Garsynas tai surinkti ir struktūrizuoti garso įrašų failai ir tekstinės transkripcijos (atitinkančios garso įrašus). Jie gali būti skirtingų kalbų ir nusakyti žodžius, fonemas, skiemenis.

Garsynų kūrimas ir anotavimas yra daug darbo sąnaudų reikalaujantys uždaviniai. Norint dirbti su jais reikia kvalifikacinių įgūdžių ir kruopštumo. Sudaromo garsyno pobūdis ir panaudojimas labai priklauso nuo to kur ir kokiems tikslams jis bus naudojamas. Garsynai gali skirtis trukme, diktorių skaičiumi, kalba, garso įrašų turiniu, transkribavimu ar naudojamu žodymo dydžiu. Labai svarbu užtikrinti nepriekaištingą visų kalbos įrašų kokybę, nes visi girdimi pašaliniai trukdžiai daro labai didelę įtaką. [20]

Šnekos atpažinimo sistemos tampa vis paklausesnės, todėl ir jų sudarymui reikalingų garsynu paklausa sparčiai didėja. Automobilių pramonė, mobiliųjų ryšių įmonės ar net tokios kompanijos kaip „google“ ar „amazon“ investuoja dideles pinigų sumas garsynų kūrimams. Kuriami įvairių kalbų ir dydžių sisteminiai, gerai anotuoti garsynai moksliniams tyrimams.

TIMIT (*Texas Instruments/Massachusetts Institute of Technology*) šnekos duomenų bazė yra prototipas pagal kurį yra sudaromi nauji garsynai. TIMIT laikomas klasikiniu garsyno pavyzdžiu, kurio esmė – užkoduota pagrindinė garsyno organizacinė informacija direktorių struktūroje, o tekstiniuose failuose ir lentelėse yra pateikiama informacija apie šnekos signalus, diktorius ir jų anotacijas. [21]

Garsynai gali būti skirstomi pagal turinį: nespontaniškos šnekos ir spontaniškos šnekos. Pagal paskirtį garsynai yra skirstomi dviem skirtingoms sistemoms:

- Atpažįstančioms balsą nepriklausomai nuo diktorius;
- Atpažįstančioms balsą priklausomai nuo diktorius (turi būti apmokytos prie diktorius).

Garsynų anotavimas ir segmentavimas gali būti atliekamas:

- Neuroniniais tinklais;
- Paslėptųjų markovo modelių;
- Dinaminio laiko skalės kraimymo modelių.

Garsynų įvairovė yra labai didelė ir jų nuolat daugėja. Kiekvienas garsynas yra labai skirtingas ir turi jam būdingas charakteristikas. Šnekos atpažinimo sistemų kūrimui pasirinkti garsynai turi užtikrinti:

- Duomenų korektriškumą ir integralumą;
- Manipuliavimą duotais duomenimis;
- Duomenų paieškos sistemas;
- Lengvą vartotojo sąsają.

## 1.9. LIETUVIŲ KALBOS ATPAŽINTUVŲ APŽVALGA

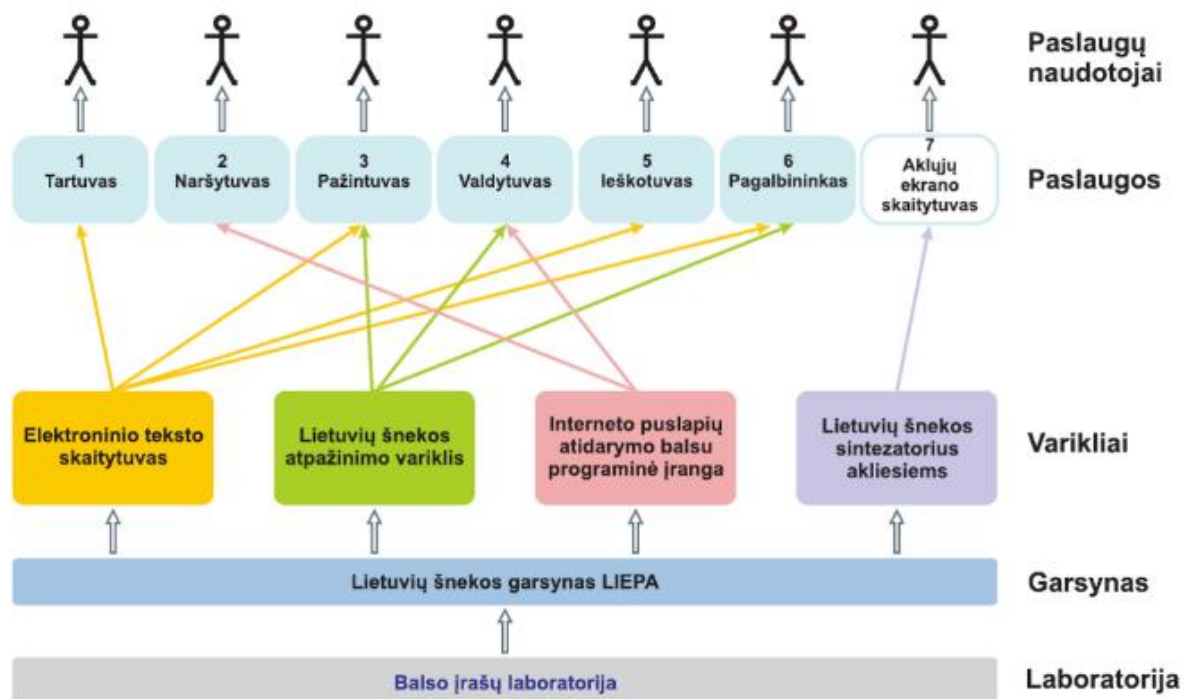
### Projektas „Liepa“

Pasaulyje atsiranda vis daugiau technologijų ir įtaisų galinčių suprasti kalbą. Tačiau tokie įtaisai nėra pritaikyti lietuvių kalbai ir jos tiesiog nesupranta. Šį trūkumą šalina darbai, atlikti vykdančią projektą LIEPA. Projekto metu yra sukurtos lietuvių kalba valdomos paslaugos, kurios padeda moksleiviams, neįgaliesiems ar suaugusiems žmonėms savo veikloje naudoti tokią šneką suprantančius ir atsakančius kompiuterius. Projekto „Liepa“ paslaugos vartotojams:

- 1) tartuvas (leidžia išgirsti lietuvių kalbos naujažodžių tarimą);
- 2) naršytuvas (leidžia naršyti internete naudojant kompiuteriui nurodomas komandas balsu);
- 3) pažintuvas (padeda moksleiviams mokytis ląstelės sandarą naudojant garsines komandas);
- 4) valdytuvas (leidžia valgyti kompiuterį balsu);
- 5) ieškotuvai (balsu galima ieškoti informacijos apie UNESCO kultūros paveldą pasaulyje);
- 6) pagalbininkas (padeda mokytis socialinių gebėjimų negalios turintiems žmonėms);
- 7) lietuvių šnekos sintezatorius akliesiems (kompiuterio ekrane rodomos informacijos skaitymas).

Projekto „Liepa“ infrastruktūrinės paslaugos yra elektroninio teksto skaitytuvas, lietuvių šnekos atpažinimo variklis, interneto puslapių atidarymo balsu programinė įranga ir 113 valandų anotuotas lietuvių kalbos garsynas. Lietuvių šneka valdomų paslaugų architektūros schema pavaizduota 1.9 paveiksle.

## Lietuvių šneka valdomos paslaugos Architektūra



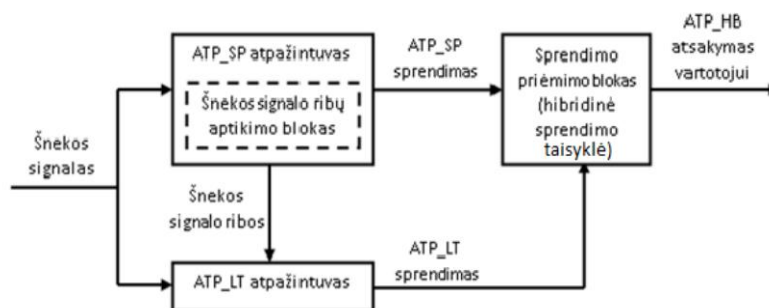
1.9 pav. Lietuvių šnekos valdomų paslaugų architektūra [15]

Garsynas „LIEPA“ sudarytas šnekos atpažinimo ir sintezavimo tikslais. Parengta 100 val. ir 13 val. trukmės garsyno dalys. Garso įrašų formatas – „wav“, diskretizavimo dažnis - 22 kHz, kvantavimas - 16 bitų. 100 val. trukmės garsyno dalį sudaro garso įrašai ir transkribuotų garsyno žodžių žodynas. Diktorių skaičius – 376 (248 mot. ir 128 vyr., 83 proc. diktorių amžius 15-22 m.). Garso įrašai: žodžių, frazių, sakinių ir jų dalių. 13 val. trukmės garsyno dalį sudaro garso įrašai ir juos atitinkančios anotacijos. Diktorių skaičius – 4 (2 mot. ir 2 vyr.). Garso įrašai: 5000 sakinių. [15]

### Projektas „Hibridinė atpažinimo technologija balsu sąsajai“

Šiame 2012 – 2013 metais vykdytame projekte buvo norima ištirti ar įmanoma tuo pačiu metu hibridiškai naudoti du skirtingus atpažintuvus, panaudojant kiekvieno iš jų privalumus ir tokiu būdu optimizuoti bendrą sistemos veikimo tikslumą.

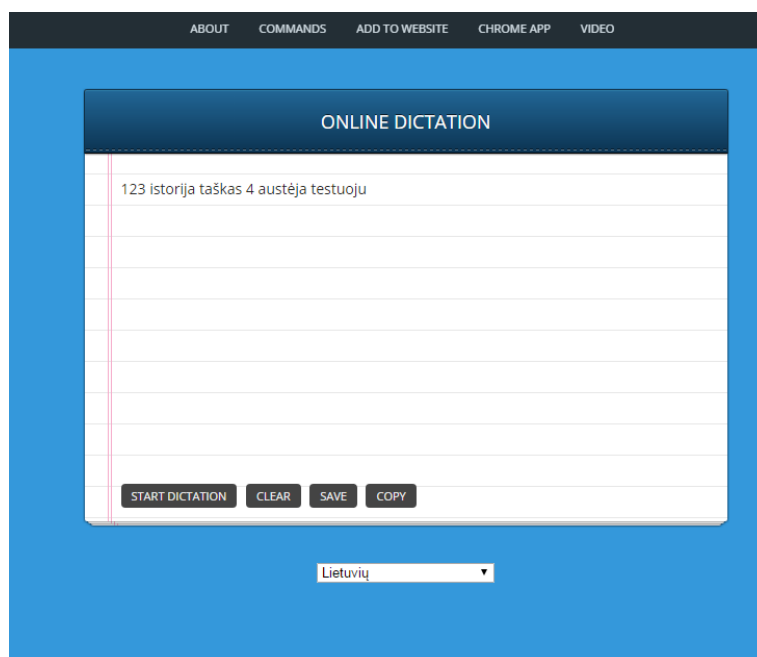
Projekto tikslas buvo sukurti kalbos komandų atpažinimo sistemą, kuri būtų orientuota į gydytojus, farmacininkus, siekiant informacijos paiešką farmacijos duomenų bazėje padaryti patogesne ir tuo pačiu greitesne, neprarandant kontakto su pacientu. [13]



1.10 pav. Hibridinio atpažintuvo struktūra [16]

### Teksto diktavimas naršyklėje (*dictation.io*)

Apsilankius internetiniu adresu „<https://dictation.io/>“, galima rašyti rašinius, laiškus ir kitus tekstus savo naršyklėje visiškai neprisiliečiant prie klaviatūros ar pelės. Ši balso atpažinimo sistema gali atpažinti įvairios kalbos žodžius, taip pat ir lietuviškus. Užtenka tiesiog turėti mikrofoną prijungtą prie kompiuterio, pasirinkti norimą kalbą ir paspausti pradėti diktavimą („*start dictation*“) mygtuką. Išstarti žodžiai yra matomi kompiuterio ekrane. Baigę diktavimą spaudžiame nustoti klausytis („*stop listening*“) mygtuką. Paspaudę išsaugoti („*save*“) galime išsaugoti visą sudiktuotą tekstą. „*Dictation.io*“ tinklalapio langą ir jame sudiktuotus žodžius matome 1.11 paveiksle.



1.11 pav. Dictation.io tinklalapio langas



Naudotis šiuo atpažintuvu labai paprasta. Tereikia nueiti į nurodytą puslapį naudojant „Google“ naršyklę. Atpažintuvas yra paremtas „Google Now“ virtualiu asistentu.

### **„Google Now“ virtualus asistentas**

„Google Now“ tai virtualus asistentas sukurtas bendrovės „Google“ ir naudojamas „Google Chrome“ naršyklėje tiek personaliniame kompiuteryje, tiek mobiliuosiuose telefonuose su „Android“ ar „iOS“ operacinėmis sistemomis.

Kompanijos inžinieriai 2014 metų vasarą pasitelkę į pagalbą kelias dešimtis asmenų, įrašinėjo lietuviškus žodžius ir frazes. Nors virtualus asistentas pradėtas naudoti jau 2012 metais, tik nuo 2015 metų užklausas pradėjo suprasti lietuvių kalba. Vykdamas paiešką balsu, asistentas puikiai susidoroja su lietuviškais vietovardžiais ir gatvių pavadinimais. Pavyzdžiui, paklausus, „kaip nuvažiuoti į Studentų gatvę“, gauname nuorodą į atitinkamą vietą žemėlapyje.

Kalbos atpažinimas nėra iki galo užbaigtas ir vis dar testuojamas, kadangi „Google Now“ mokymas lietuviškai yra gan ilgai procesas. Nors lietuvių kalbos atpažinimą paieškos programose ir „Google Translate“ gali išbandyti visi vartotojai, tačiau „iOS“ operacinių sistemų naudotojams vis dar nėra galimybės valdyti balsu kalbant lietuviškai.

## METODINĖ DALIS

### 1.10. VARDŲ IR SKAIČIŲ GARSYNŲ PARUOŠIMAS ATPAŽINIMO TYRIMAMS

Norint atlikti skaičių ir vardų atpažinimo tyrimus visų pirma reikia turėti norimų žodžių ištarimo garsynus. Skaičių ir vardų garsynas yra sudaromas naudojant daug skirtingų ištarimų ir diktorių. Garsyno balso komandos buvo įrašytos panaudojus MS DOS operacinėje sistemoje programa „inp\_sr16.exe“.

Garsynas buvo sudarytas iš 26 skirtingų vardų balso komandų ir 10 skirtingų skaičių balso komandų. Vardų balso komandas įrašinėjo 21 skirtingas diktorius, o skaičių komandas sudarė 30 skirtingų diktorių. Kiekviena balso komanda buvo diktuojama po 20 kartų. Garsynas buvo sudarytas tylioje aplinkoje ištariant reikiamas komandas. Nekokybiškai įrašytos komandos arba komandos kuriose girdimi triukšmai buvo pakartotinai ištariamoms ir pakeičiamoms geriau girdimu variantu. Pilną tiriamą garsyną sudarė šešiolika tūkstančių devyni šimtai dvidešimt (16920) skirtingų balso įrašų. Šeši tūkstančiai (6000) skaičių balso įrašų ir dešimt tūkstančių devyni šimtai dvidešimt (10920) vardų balso įrašų. Garsyną sudarantys balso įrašai pateikti priede 6.1 ir 6.2 lentelėse.

Kiekvienas balso įrašas buvo specifiškai pavadintas nurodant tarėjo lytį, vardo ir pavardės pirmąsias raides, tariamą komandą ir ištarimo numerį. Balso įrašo pavadinimas yra paaiškintas žemiau esančioje lentelėje.

#### 1.1 lentelė. Balso įrašo pavadinimo paaiškinimas

Ženklo numeris	Ženklo reikšmė	MKASRATC12 balso įrašo pavadinimo paaiškinimas
1 raidė	Nurodo asmens lytį, M – vyras, F – moteris	M - nurodo vyrišką lytį
2-4 raidės	Tarėjo pirmosios vardo raidės	KAS – nurodo „Kastytis“ pirmąsias 3 vardo raides
5-7 raidės	Tarėjo pirmosios pavardės raidės	RAT – nurodo „Ratkevičius“ pirmąsias 3 pavardės raides
8 ženklas	Nurodo ištartą žodį: 0-9 nurodo ištartą skaičių, a-z nurodo ištartą vardo pirmąją raidę	C – nurodo ištartą vardą „Cecilija“
9-10 skaičiai	Nurodo nuo 0 iki 19 žodžio ištarimo	12 – dvyliktą kartą ištartas žodis

	numerį	„Cecilija“
--	--------	------------

Darbe naudojamas garsynas buvo suskirstytas į 7 lygias dalis (sluoksnius), kurios pagal diktorius yra pateiktos priede, 6.3 lentelėje. Taip buvo padaryta todėl, nes atliekant tyrimus 6 sluoksniai yra naudojami apmokymui, o likęs sluoksnis skirtas testavimui. Tokiu metodu galima atlikti testavimus su septyniais skirtingais sluoksniais testavimui, toks tyrimas vadinamas kryžminių. Darbe testavimui naudojami 5 skirtingi tarėjai vardams ir 3 skirtingi tarėjai skaičiams, tai sudaro 15,13 % pilno garsyno masyvo.



**1.12 pav.** Garsyno suskirstymas mokymui ir testavimui

### **1.11.HIDDEN MARKOV MODEL TOOLKIT TAIKYMAS**

Šiame skyriuje yra aprašoma darbo eiga, skaičių ir vardu pavadinimų atpažinimo tikslumo tyrimams atlikti, naudojant HTK (*Hidden Markov Model Toolkit*) programinį paketą. HTK programinio paketo pritaikymas šnekamosios kalbos atpažinimui ir pagrindiniai bruožai yra aptarti teorinėje darbo dalyje, tačiau taip pat svarbu ir praktinis šio įrankio pritaikymas. Taigi šiame skyriuje aptariamas praktinis HTK programinio paketo pritaikymas, įsisavinimas ir reikalingų duomenų paruošimas šnekamosios kalbos atpažinimui.

HTK žodžių atpažinimas grįstas paslėptaisiais Markovo modeliais (PMM), taigi sistema pirmiausiai turi būti apmokyta, o tik paskui naudojama šnekai atpažinti. PMM apmokymo metu visiems žodyno žodžiams pagal parinktas žodžių transkripcijas sudaromi tikimybiniai akustiniai modeliai ir apskaičiuojamas kiekvieno akustinio modelio tikėtumo įvertis. Sistema iš žodyno pateikia kaip atpažintą tą žodį, kurį reprezentuoja modelis, kurio tikėtumo įvertis yra didžiausias. Šnekamosios kalbos atpažinimo sistemos, kurioje kalbos atpažinimui taikomas

PMM metodas, apmokymui rekomenduojama naudoti kitų diktorių garsyno įrašus nei atpažinimo kokybės testavimui.

HTK programiniam paketui reikalingų duomenų redagavimui, katalogavimui, perkėlimui, grupavimui ir kopijavimui buvo naudojamos papildomos programos - „Total Commander“ ir „Far Manager 3“, „Notepad++“, Jas naudojant lengviau sutvarkyti duomenis, padidėja darbo su duomenimis efektyvumas.

HTK programinio paketo programoms aktyvinti naudojami „batch“ (trumpinys „\*.bat“) failai. Šių failų pagrindinė paskirtis yra automatizuoti tam tikrą procesą. Failų viduje parašoma tam tikra veiksmų seka, pvz. „*HCOPY -C CONFIG -S failai.scp*“, kuri vykdoma aktyvavus „batch“ failą. HTK paketo programos yra naudojamos nuosekliai, veiksmai atliekami vienas po kito.

## 1.12.HTK PANAUDOJIMAS ŠNEKOS ATPAŽINIMUI TAIKANT ŽODŽIŲ MODELIŲ METODĄ

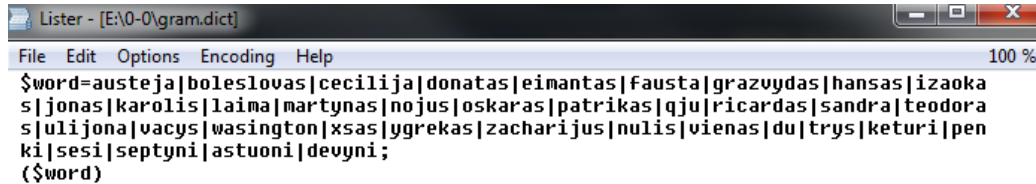
Šnekamosios kalbos atpažinimui taikyti HTK programinių įrankių paketą galima modeliuojant balso komandų atpažinimą žodžių modelių metodu. Šiame skyriuje aprašomas HTK programinių įrankių naudojimas balso komandų atpažinimo modeliavimui taikant žodžių modelių metodą, kurio taikymą sudaro 7 nuoseklūs etapai. Pirmi šeši etapai apima komandų akustinių modelių apmokymą, o septintasis etapas apima modelių testavimą. Trumpa visų etapų apžvalga pateikiama 2.2 lentelėje.

**1.2 lentelė.** Žodžių modelių metodo taikymas.

		<b>Pavadinimas</b>	<b>Programa</b>	<b>Reikalingi failai</b>	
<b>ETAPAS</b>	<b>1</b>	Gramatikos failo kūrimas	-	Gram.dict	<b>MOKYMAS</b>
	<b>2</b>	Reikalingų failų paruošimas	<i>Hparse.exe</i>	Gram.dict, wordnet.txt	
	<b>3</b>	Požymių kūrimas	<i>Hcopy.exe</i>	Config.txt, failai.scp	
	<b>4</b>	Modelių parengimas	-	Trans_eimantas.scp, hmm_eimantas, ...	
	<b>5</b>	Modelių apmokymas	<i>HRest.exe</i>	Trans_eimantas.scp, hmm_eimantas, ...	
	<b>6</b>	Duomenų paruošimas testavimui	<i>HCOPY.exe</i>	Config.txt, test.scp	<b>TES</b>

	7	Modelių testavimas	<i>Hvite.exe</i>	Hmmlist, test_mfc.scp, wordnet.txt, dict.txt
--	---	--------------------	------------------	--

**Pirmas etapas** – „*gram.dict*“ gramatikos failo kūrimas. Šiame faile eilės tvarka yra surašomos balso komandos, kurias bandoma atpažinti. Sudarytas failas pateiktas 2.2 paveiksle.



```

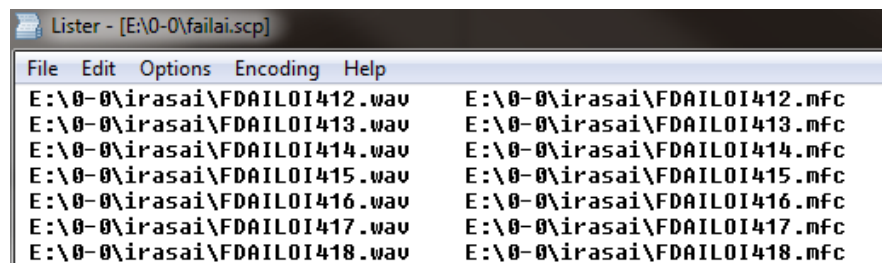
Lister - [E:\0-0\gram.dict]
File Edit Options Encoding Help 100 %
$word=austeja|boleslovas|cecilija|donatas|eimantas|fausta|grazvydas|hansas|izaokas|jonas|karolis|laima|martynas|nojus|oskaras|patrikas|qju|ricardas|sandra|teodora|s|ulijona|vacys|wasington|xsas|ygrekas|zacharijus|nulis|vienas|du|trys|keturi|penki|sesi|septyni|astuoni|devyni;
($word)

```

**1.13 pav.** Darbe naudojamas „*gram.dict*“ failas

**Antras etapas** – Žodžių modelių sudarymui reikalingų failų paruošimas. Naudojant HTK programinio paketo programą „*Hparse.exe*“ ir jau sudarytą „*gram.dict*“ failą yra automatiškai sukuriamas failas „*wordnet.txt*“

**Trečias etapas** – požymių kūrimas. Pirma yra sukuriamas failas „*failai.scp*“. Šiame faile nurodoma, kur yra darbe naudojamu skaičių ir vardų garso failai. Taip pat nurodoma, kur bus sukurti ir talpinami požymių failai kiekvienam balso įrašui. Darbe naudojamas skaičių ir vardų failo dalis „*failai.scp*“ pavaizduota 2.3 paveiksle.



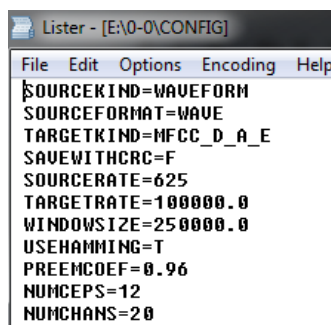
```

Lister - [E:\0-0\failai.scp]
File Edit Options Encoding Help
E:\0-0\irasai\FDAILOI412.wav E:\0-0\irasai\FDAILOI412.mfc
E:\0-0\irasai\FDAILOI413.wav E:\0-0\irasai\FDAILOI413.mfc
E:\0-0\irasai\FDAILOI414.wav E:\0-0\irasai\FDAILOI414.mfc
E:\0-0\irasai\FDAILOI415.wav E:\0-0\irasai\FDAILOI415.mfc
E:\0-0\irasai\FDAILOI416.wav E:\0-0\irasai\FDAILOI416.mfc
E:\0-0\irasai\FDAILOI417.wav E:\0-0\irasai\FDAILOI417.mfc
E:\0-0\irasai\FDAILOI418.wav E:\0-0\irasai\FDAILOI418.mfc

```

**1.14 pav.** Darbe naudojamas „*failai.scp*“ failas

Taip pat yra sukuriamas failas „*CONFIG*“, kuriame nurodoma požymių skaičiavimo būdus. Toks failas yra pateiktas 2.4 paveiksle.



```

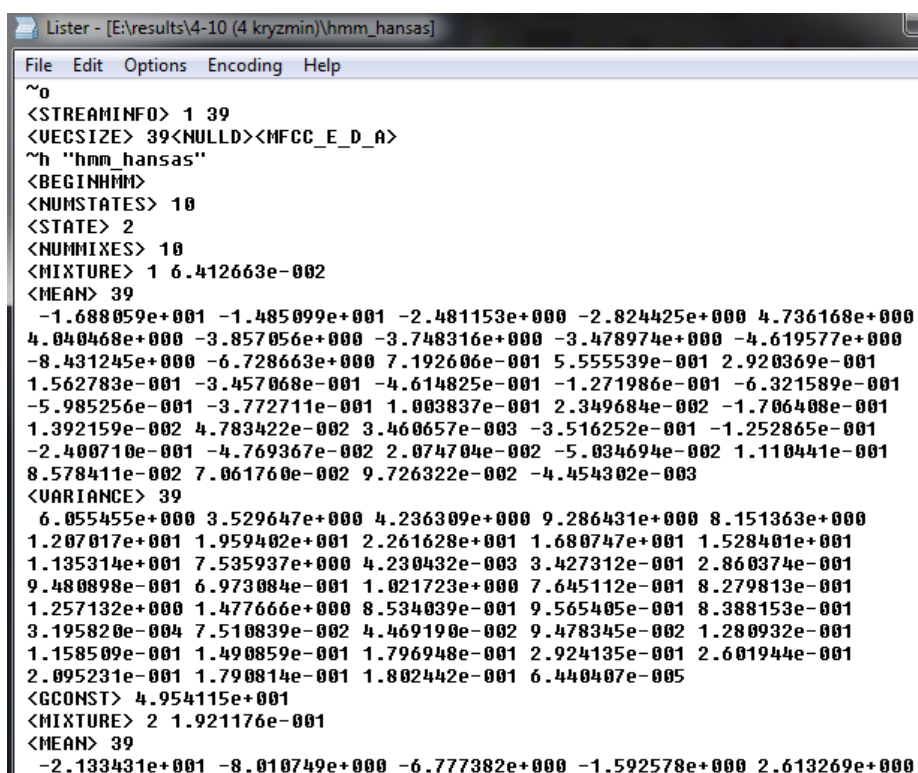
Lister - [E:\0-0\CONFIG]
File Edit Options Encoding Help
SOURCEKIND=WAVEFORM
SOURCEFORMAT=WAVE
TARGETKIND=MFC_D_A_E
SAVEWITHCRC=F
SOURCERATE=625
TARGETRATE=100000.0
WINDOWSIZE=250000.0
USEHAMMING=T
PREEMCOEF=0.96
NUMCEPS=12
NUMCHANS=20

```

### 1.15 pav. Darbe naudojamas „CONFIG“ failas

Turint failus „*failai.scp*“ ir „*CONFIG*“ naudojama HTK programinio paketo programa „*Hcopy.exe*“ ir taip sukuriama požymių failai „\*.mfc“ formatu. Atpažinimo sistema atpažinimo tyrimams naudoja požymių failus, o ne pačius garso failus.

**Ketvirtas etapas** – modelių failų parengimas. Pirma sukuriama „*trans*“ failai kiekvienam žodžiui. Šiuose failuose nurodomi atskirų žodžių požymių failai ir jų direktorijos. Paskui Kiekvienam norimam atpažinti žodžiui yra kuriamas modelių failas „*hmm*“. Kiekviename modelių faile yra nurodoma kiek Gauso mišinių ir būsenų bus naudojama modeliuojant žodį. Aprašomos perėjimų tikimybių reikšmės, dispersijų ir vidurkių vektoriai ir kiti parametrai. 2.5 paveiksle yra pateiktas „*hmm\_hansas*“ nepilnas modelio failas.



```
Lister - [E:\results\4-10 (4 kryzmin)\hmm_hansas]
File Edit Options Encoding Help
~o
<STREAMINFO> 1 39
<VECSIZE> 39<NULLD><MFCC_E_D_A>
~h "hmm_hansas"
<BEGINHMM>
<NUMSTATES> 10
<STATE> 2
<NUMMIXES> 10
<MIXTURE> 1 6.412663e-002
<MEAN> 39
-1.688059e+001 -1.485099e+001 -2.481153e+000 -2.824425e+000 4.736168e+000
4.040468e+000 -3.857056e+000 -3.748316e+000 -3.478974e+000 -4.619577e+000
-8.431245e+000 -6.728663e+000 7.192606e-001 5.55539e-001 2.920369e-001
1.562783e-001 -3.457068e-001 -4.614825e-001 -1.271986e-001 -6.321589e-001
-5.985256e-001 -3.772711e-001 1.003837e-001 2.349684e-002 -1.706408e-001
1.392159e-002 4.783422e-002 3.460657e-003 -3.516252e-001 -1.252865e-001
-2.400710e-001 -4.769367e-002 2.074704e-002 -5.034694e-002 1.110441e-001
8.578411e-002 7.061760e-002 9.726322e-002 -4.454302e-003
<VARIANCE> 39
6.055455e+000 3.529647e+000 4.236309e+000 9.286431e+000 8.151363e+000
1.207017e+001 1.959402e+001 2.261628e+001 1.680747e+001 1.528401e+001
1.135314e+001 7.535937e+000 4.230432e-003 3.427312e-001 2.860374e-001
9.480898e-001 6.973084e-001 1.021723e+000 7.645112e-001 8.279813e-001
1.257132e+000 1.477666e+000 8.534039e-001 9.565405e-001 8.388153e-001
3.195820e-004 7.510839e-002 4.469190e-002 9.478345e-002 1.280932e-001
1.158509e-001 1.490859e-001 1.796948e-001 2.924135e-001 2.601944e-001
2.095231e-001 1.790814e-001 1.802442e-001 6.440407e-005
<GCNST> 4.954115e+001
<MIXTURE> 2 1.921176e-001
<MEAN> 39
-2.133431e+001 -8.010749e+000 -6.777382e+000 -1.592578e+000 2.613269e+000
```

### 1.16 pav. Darbe naudojamas „hmm\_hansas“ modelio failas

Pirmosiose eilutėse yra pateikiama bendra informacija apie modelių failą:

- „<STREAMINFO>“ nurodo naudojamus 39 požymių rinkinius išdėstyti pirmoje eilutėje;
- „<VECSIZE>“ nurodo vektorių dydį;
- „<NUMSTATES>“ – naudojamų būsenų skaičius;
- „<STATE>“ – modelio pradžios būseną;
- „<NUMMIXES>“ – naudojamų Gauso mišinių skaičius;
- „<MIXTURE>“ – pirmasis Gauso mišinys ir jo vertė;

- „<MEAN>“ – požymių vidurkiai. Jie mokymo metu perskaičiuojami iš naudo;
- „<VARIANCE>“ – požymių dispersijos. Jos mokymo metu perskaičiuojamos.

**Penktas etapas** – modelių apmokymas. Sudarius „hmm“ ir „trans“ failus kiekvienam norimam atpažinti žodžiui yra naudojama HTK programa „HRest.exe“. Ši programa išskiria pagalbinį apmokymo failą „kk“ kiekvienam žodžiui. Šis failas nurodo apmokymo iteracijų skaičių. Vardo Patrikas ištarimo pagalbinis apmokymo failas „kk16“ pateiktas 2.6 paveiksle.

```

Lister - [E:\0-3\kk16]
File Edit Options Encoding Help
Reestimating HMM hmm_patrikas . . .
States : 2 3 4 5 6 7 8 9 10 11 (width)
Mixes s1: 10 10 10 10 10 10 10 10 10 10 ( 39 )
Num Using: 0 0 0 0 0 0 0 0 0 0
Parm Kind: MFCC_E_D_A
Number of owners = 1
SegLab : NONE
MaxIter : 20
Epsilon : 0.000100
Updating : Transitions Means Variances

- system is PLAIN
360 Examples loaded, Max length = 160, Min length = 96
Ave LogProb at iter 1 = -6090.82847 using 360 examples
Ave LogProb at iter 2 = -6057.47986 using 360 examples change = 33.34875
Ave LogProb at iter 3 = -6056.24028 using 360 examples change = 1.23970
Ave LogProb at iter 4 = -6055.72014 using 360 examples change = 0.52010
Ave LogProb at iter 5 = -6055.37639 using 360 examples change = 0.34383
Ave LogProb at iter 6 = -6055.01319 using 360 examples change = 0.36327
Ave LogProb at iter 7 = -6054.77361 using 360 examples change = 0.23957
Ave LogProb at iter 8 = -6054.61389 using 360 examples change = 0.15955
Ave LogProb at iter 9 = -6054.50347 using 360 examples change = 0.11030
Ave LogProb at iter 10 = -6054.40208 using 360 examples change = 0.10133
Ave LogProb at iter 11 = -6054.29792 using 360 examples change = 0.10394
Ave LogProb at iter 12 = -6054.19028 using 360 examples change = 0.10757
Ave LogProb at iter 13 = -6054.08264 using 360 examples change = 0.10779
Ave LogProb at iter 14 = -6053.97431 using 360 examples change = 0.10821
Ave LogProb at iter 15 = -6053.89514 using 360 examples change = 0.07898
Ave LogProb at iter 16 = -6053.85139 using 360 examples change = 0.04363
Ave LogProb at iter 17 = -6053.82847 using 360 examples change = 0.02309
Ave LogProb at iter 18 = -6053.80625 using 360 examples change = 0.02236
Ave LogProb at iter 19 = -6053.78611 using 360 examples change = 0.02004
Ave LogProb at iter 20 = -6053.77222 using 360 examples change = 0.01391
Estimation aborted at iteration 20

```

**1.17 pav.** Vardo Patrikas „kk16“ pagalbinis apmokymo failas

**Šeštasis etapas** – duomenų paruošimas testavimui. Šiame etape kuriami „\*.mfc“ požymių failai kiekvienam testavime naudojamam garsui. Tai atliekama taip pat kaip ir 3 etape buvo kuriami požymių failai apmokymui. Sukuriamas „dict“ failas kuris nurodo kokius žodžius reikia atpažinti.

**Septintasis etapas** – modelių testavimas. Atlikus visus paminėtus veiksmus ir pasiruošus reikiamus failus galime atlikti modelių testavimą. Testavimui atlikti naudojama HTK programa „Hvite“. Ji automatiškai pateikia atpažinimo tikslumo rezultatus. Rezultatuose nurodomi diktoriai, jų tariami žodžiai, atpažinti žodžiai ir atpažinimo tikimybės. Vieno žodžio rezultato pavyzdys pateiktas 2.7 paveiksle.

```

"E:/0-3/pozymiai2/MDAIGUS203.rec"
0 5900000 du -2798.197266
...

```

**1.18 pav.** Vieno žodžio atpažinimo rezultatas

Pagal 2.1 lentelę galime nuspręsti, kad „MDAIGUS207“ failas nusako žodį „du“. Toks žodis ir yra atpažintas pavyzdyje.

Norint apibendrinti rezultatus sudarome „*testref*“ failą, kur nurodoma kiekvieno testavime naudojamo žodžio tikroji reikšmė. Panaudojame HTK „*Hresults*“ programą, kuri sulygina tikrąsias žodžių reikšmes su atpažintomis ir taip gauname rezultatus.

### **1.13. VARDŲ IR SKAIČIŲ ATPAŽINIMO TYRIMO, NAUDOJANT ŽODŽIAIS GRĮSTUS MODELIUS, METODIKA**

Pagal šiame skyriuje aprašyto HTK programinių įrankių paketo taikymą balso komandų atpažinimui tirti, naudojant žodžių modelių metodą, metodiką buvo sumodeliuoti skaičių ir vardų garsyno komandų modeliai, kurių atpažinimo tikslumo rezultatai, naudojant skirtingų modelių būsenų ir Gauso mišinių skaičių, aprašomi 2.3 skyriuje.

Norint gauti kuo tikslesnius vardų komandų modelius, reikia iširti komandų modelių atpažinimo tikslumą su įvairiomis parametrų vertėmis, redaguojamomis HTK programiniame pakete.

Kiekviena balso komanda buvo modeliuojama viena ištisine Markovo grandine, sukuriant komandoms atitinkamus PMM, kuriuose pagrindiniai redaguojami parametrai, kurie gali lemti balso komandų atpažinimo tikslumą yra būsenų ir Gauso mišinių skaičius grandinėje.

#### **Papildomų būsenų įterpimas į žodžių modelius**

Atpažinimo tikslumui tirti su skirtingu būsenų skaičiumi, būsenų skaičius buvo parenkamas apytiksliai lygus balso komandą sudarančių fonetinių elementų skaičiui intervale, imtinai nuo 3 iki 10. Skaičiaus komandos „*du*“ modeliui buvo naudotos tik 2 būsenos, komandų „*Boleslovas*“ ir „*Zacharijus*“ modeliams buvo naudota 10 būsenų ir kitų balso komandų modeliams atitinkamai parenkamas būsenų skaičius. Toliau būsenų skaičius buvo parenkamas prie pirminio būsenų skaičiaus palapsniui pridedant 2, 4, 6, 10 papildomų būsenų.

#### **Papildomų Gauso mišinių įterpimas į žodžių modelius**

Toliau tęsiant žodžių atpažinimo tikslumo su skirtingu būsenų ir Gauso mišinių skaičiumi tyrimus, testuojant tą patį žodžių garsyną, buvo įterpiami Gauso mišiniai ir tiriamas komandų atpažinimo tikslumas naudojant modelius su dviem papildomomis būsenomis ir skirtingu Gauso mišinių skaičiumi – 2, 3, 6, ir 10 Gauso mišinių, ir naudojant modelius su keturiomis papildomomis būsenomis ir 2, 3, 6, 10 ir Gauso mišinių.

Atlikto vardų komandų atpažinimo tikslumo tyrimo, naudojant žodžių modelių metodą, rezultatai įvertinami 3.1 skyriuje.

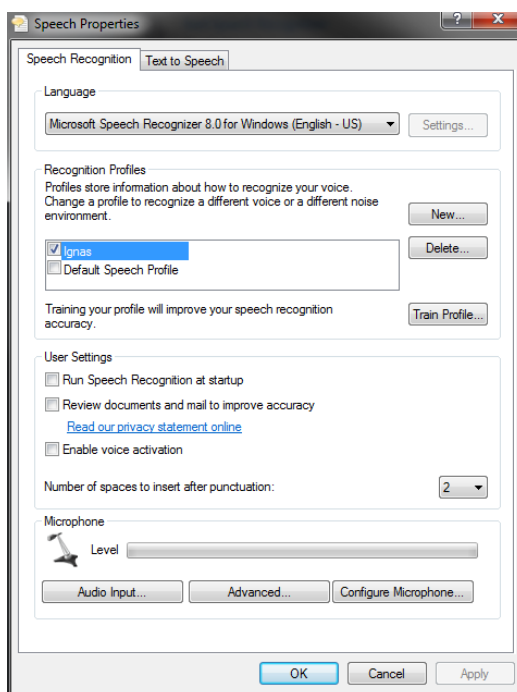


## 1.14. VARDŲ IR SKAIČIŲ ATPAŽINIMO TYRIMO, NAUDOJANT IŠPANIŠKĄ ATPAŽINTUVĄ, METODIKĄ

Ši darbo dalis yra atliekama „Microsoft Speech Recognizer“ programinius paketu. Šis programinis paketas yra įdiegtas „Microsoft Windows“ operacinėje sistemoje ir jis yra skirtas kalbos atpažinimui. „Microsoft Speech Recognizer“ gali atpažinti tik populiariausias kalbas (anglų, prancūzų, ispanų ir kita), lietuvių kalbos šis atpažintuvas neatpažįsta. Tačiau galime pritaikyti kitakalbį atpažintuvą suprasti specifinius norimus žodžius. Tai yra atliekama parenkant tinkamas fonetinių vienetų sekas (parenkami kitakalbio akustiniai modeliai atitinkantis lietuviškus fonetinius vienetus).

Darbui buvo parinktas „Microsoft Speech Recognizer 8.0 for Windows (Spanish-Spain)“ automatinis ispanų kalbos atpažintuvas. Šio atpažintuvo fonetika labiausiai atitiko lietuvišką ir juo gauti rezultatai yra geresni už kitus „Microsoft Windows“ aplinkoje naudojamus atpažintuvus. Ši AŠA buvo naudojama kaip savaime apsimokanti sistema su visais naudojamai diktorių įrašais.

„Microsoft Speech Recognizer“ programinio paketo langas pavaizduotas 2.2 paveiksle.



1.19 pav. „Microsoft Speech Recognizer“ kalbos atpažintuvo langas

„Language“ skiltyje galime pasirinkti norimą atpažinimo kalbą. „Recognition profiles“ – parenkame atpažinimo profilį. „Train profile“ – apmokomas profilis. Profilis yra sukuriamas vienam kalbėtojui tariant frazes. Kalbos atpažinimo sistemai sukūrus profilį buvo atliktas

daugiadiktorinis skaičių ir vardų atpažinimo tyrimas. Tyrimo metu buvo pateikti kiekvieno naudojamo garso failo atpažinimo rezultatai.

### 1.15. DVIEJŲ SKIRTINGŲ ŠNEKOS ATPAŽINTUVŲ SUJUNGIMAS

Šio tyrimo metu yra sudaroma atpažinimo sistema, kuri gali parinkti geriau veikiančią atpažintuvą iš darbe tirtų lietuviško (naudojančio paslėptuosius Markovo modelius) ir ispaniško atpažintuvų. Šio tyrimo metu yra naudojama duomenų analizės sistema „WEKA“.

Skirtingi atpažintuvai skirtingai atpažįsta tariamus žodžius ir turi skirtingas atpažinimo tikimybes ir požymius. Pagal abiejų atpažintuvų pateiktus atsakymus (sprendimus), įrašai buvo suskirstyti į poaibius. Įrašų suskirstymo pavyzdys pateiktas 2.9 paveiksle..

Diktorius	Failas	Teisingas sprendimas	LT sprendimas 1	LT logprob	SP sprendimas	SP prop	gender	class	TRUE LT	TRUE SP	lt_a	lt_b	lt_c	lt_ė
FGINTRA	FGINTRA005	Boleslovas	Boleslovas	-9809,379883	Gražvydas	465	F	LT	1	0	10,0	0,0	0,0	0,0
FGINTRA	FGINTRA04	Fausta	Fausta	-6205,787598	Oskaras	593	F	LT	1	0	33,3	0,0	0,0	0,0
FGINTRA	FGINTRA09	Fausta	Fausta	-6257,910645	Oskaras	636	F	LT	1	0	33,3	0,0	0,0	0,0
FGINTRA	FGINTRA11	Fausta	Fausta	-6510,578613	Oskaras	596	F	LT	1	0	33,3	0,0	0,0	0,0
FGINTRA	FGINTRA13	Fausta	Fausta	-6162,312988	Oskaras	637	F	LT	1	0	33,3	0,0	0,0	0,0
FGINTRA	FGINTRA19	Fausta	Fausta	-6216,69873	Oskaras	629	F	LT	1	0	33,3	0,0	0,0	0,0
FGINTRA	FGINTRA00	Ygrekas	Ygrekas	-7433,833984	Hansas	454	F	LT	1	0	14,3	0,0	0,0	0,0
FGINTRA	FGINTRA06	Ygrekas	Ygrekas	-6790,686035	Eimantas	516	F	LT	1	0	14,3	0,0	0,0	0,0
FGINTRA	FGINTRA08	Ygrekas	Ygrekas	-7159,348145	Oskaras	469	F	LT	1	0	14,3	0,0	0,0	0,0
FGINTRA	FGINTRA02	Karolis	Karolis	-6845,609375	Zacharijus	580	F	LT	1	0	14,3	0,0	0,0	0,0
FGINTRA	FGINTRA11	Karolis	Karolis	-6429,882324	Teodoras	557	F	LT	1	0	14,3	0,0	0,0	0,0
FGINTRA	FGINTRA14	Martynas	Martynas	-8669,970703	Gražvydas	565	F	LT	1	0	25,0	0,0	0,0	0,0
FGINTRA	FGINTRA07	Nojus	Nojus	-6784,269043	Zacharijus	503	F	LT	1	0	0,0	0,0	0,0	0,0
FGINTRA	FGINTRA11	Patrikas	Patrikas	-7521,931641	Ygrekas	617	F	LT	1	0	25,0	0,0	0,0	0,0
FGINTRA	FGINTRA00	Kju	Kju	-5326,892578	Jonas	548	F	LT	1	0	0,0	0,0	0,0	0,0
FGINTRA	FGINTRA01	Kju	Kju	-6625,562012	Izaokas	540	F	LT	1	0	0,0	0,0	0,0	0,0
FGINTRA	FGINTRA03	Kju	Kju	-4388,038086	Ulijona	454	F	LT	1	0	0,0	0,0	0,0	0,0
FGINTRA	FGINTRA06	Kju	Kju	-5648,612793	Ulijona	591	F	LT	1	0	0,0	0,0	0,0	0,0
FGINTRA	FGINTRA08	Kju	Kju	-4232,661621	Ulijona	649	F	LT	1	0	0,0	0,0	0,0	0,0
FGINTRA	FGINTRA09	Kju	Kju	-4132,177246	Ulijona	543	F	LT	1	0	0,0	0,0	0,0	0,0
FGINTRA	FGINTRA12	Kju	Kju	-6903,913574	Ulijona	642	F	LT	1	0	0,0	0,0	0,0	0,0
FGINTRA	FGINTRA14	Kju	Kju	-4067,349365	Ulijona	647	F	LT	1	0	0,0	0,0	0,0	0,0
FGINTRA	FGINTRA16	Kju	Kju	-3766,903809	Ulijona	510	F	LT	1	0	0,0	0,0	0,0	0,0

1.20 pav. Skirtingų atpažintuvų balso įrašų suskirstymas

Ši į poaibius suskirstyta lentelė leido sudaryti turimų dviejų atpažintuvų papildomumą lentelę.

1.3 lentelė. LT ir SP atpažintuvų papildomumas

Poaibis	Aprašymas	Failų (frazių) kiekis
T=T	Atpažintuvų siūlomi sprendimai sutampa ir yra teisingi	15652
F=F	Atpažintuvų siūlomi sprendimai sutampa ir yra neteisingi	4
T-	Atpažintuvas LT siūlo teisingą sprendimą, atpažintuvas SP sprendimo nesiūlo	155
F-	Atpažintuvas LT siūlo neteisingą sprendimą, atpažintuvas SP sprendimo nesiūlo	4
-T	Atpažintuvas SP siūlo teisingą sprendimą, atpažintuvas LT sprendimo nesiūlo	0
-F	Atpažintuvas SP siūlo neteisingą sprendimą, atpažintuvas LT sprendimo nesiūlo	0
--	Abu atpažintuvai nesiūlo sprendimo	0
TF	Atpažintuvų siūlomi sprendimai nesutampa, LT sprendimas teisingas	1701

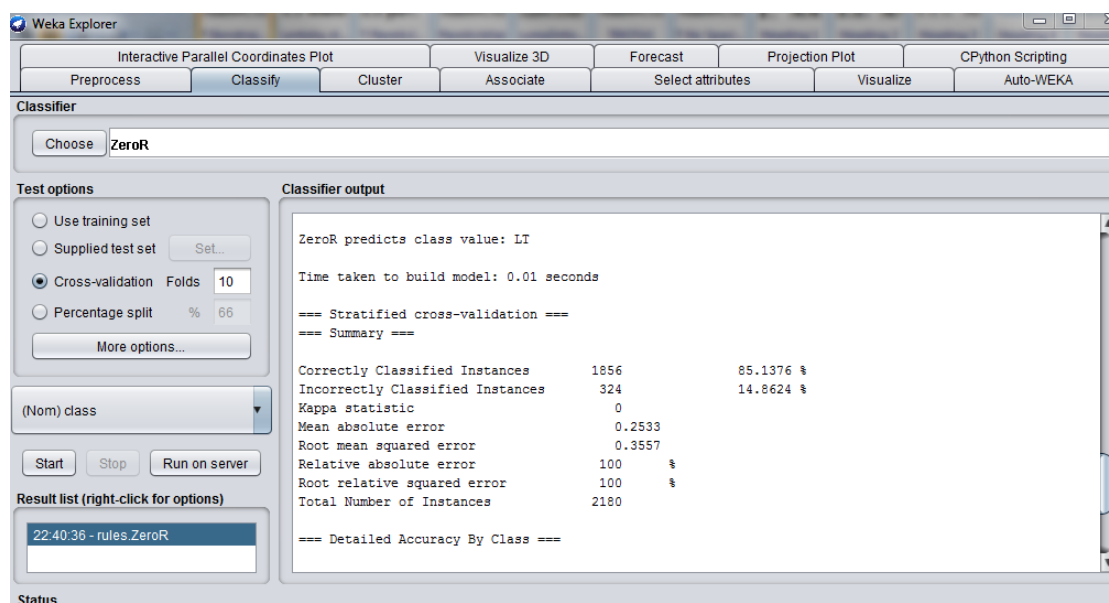


2.10 paveiksle matomų požymių paaiškinimai yra pateikti 2.4 lentelėje.

**1.4 lentelė.** Hibridinio sprendimų priėmimo bloko požymiai

Požymiai	Paaiškinimas
<i>LT_prob</i>	Paslėptaisiais Markovo modeliais sudaro atpažintuvo sprendimų patikimumų įvertįs. Matuojama vidutine logaritmine tikybine
<i>SP_prob</i>	Kitakalbio ispaniško atpažintuvo pateikto sprendimo patikimumo įvertįs
<i>gender</i>	Nusako diktoriaus lyti (M – vyras, F - moteris)
<i>lt_a</i>	Nusako kokį procentą, naudojant LT atpažintuvą, sudaro raidė „a“ atpažintam žodžiui. Pavyzdžiui, žodžiui „vienas“ sudaro 16,7% (1 raidė iš 6)
...	
<i>lt_zh</i>	Nusako kokį procentą, naudojant LT atpažintuvą, sudaro raidė „ž“ atpažintam žodžiui. Pavyzdžiui, žodžiui „Gražvydas“ sudaro 11,1% (1 raidė iš 9)
<i>sp_a</i>	Nusako kokį procentą, naudojant SP atpažintuvą, sudaro raidė „a“ atpažintam žodžiui.
...	
<i>sp_zh</i>	Nusako kokį procentą, naudojant SP atpažintuvą, sudaro raidė „ž“ atpažintam žodžiui.
<i>class</i>	Nusako atpažintuvą (LT arba SP), kuris gerai atpažino žodį

Tyrimai pirma atliekami „WEKA“ paketo įrankiu „Explorer“. Naudojant šį įrankį parenkame „\*.arff“ failą su visais aprašytais žodžiais, parenkame klasifikatorių ir norimus parametrus ir atliekame tyrimus. Tokio tyrimo fragmentas pateiktas 2.11 paveiksle.



**1.22 pav.** Duomenų klasifikavimo tyrimas „Explorer“ įrankiu

Tokiu būdu buvo atliekamas 10-kartų kryžminis patikrinimas. Paruošiamas vienas visų diktorių požymių failas, WEKA paketas pagal nutylėjimą atsitiktiniu būdu skirsto duomenis: 90% apmokymui, 10% testavimui, 10 kartų atlieka klasifikavimą, po to rezultatus vidurkina ir išveda į ekraną. Toks klasifikavimas leidžia prognozuoti klasifikavimo tikslumą (tuo pačiu ir hibridinio atpažintuvo tikslumą) žinomam kalbėtojui (vienam iš garsyno diktorių).

Taip pat galime panaudoti „Auto-WEKA“ įrankį, kuris pagal nustatyta laiko tarpa suranda geriausią klasifikatorių ir jo parametrus.

Norint atlikti tyrimus, kai yra nežinomas kalbėtojas, sudaromi „\*.arff“ failai tik testuojamiems ir apmokomiems sluoksniams. Paskui panaudojame „WEKA“ programinį įrankį „SimpleCLI“. Komandinėje eilutėje nurodomas norimas duomenų klasifikatorius, apmokymo duomenų direktorija ir testavimo duomenų direktorija. Paruošiami 2\*n failų, kur: apmokymui imama n-1 diktorių požymiai, o testavimui – 1 (n-tojo) diktoriaus požymiai. Klasifikavimas atliekamas n kartų per komandinę eilutę paduodant apmokymui n-1 diktorių požymių failą, o testavimui n-tojo diktoriaus požymių failą. Tai kartojama n kartų keičiant n-tąjį diktorių. Po to rezultatai rankiniu būdu vidurkinami. Tokio klasifikavimo rezultatai leidžia prognozuoti klasifikavimo tikslumą (tuo pačiu ir hibridinio atpažintuvo tikslumą). Tokio tyrimo fragmentas pateiktas 2.12 paveiksle.

```
> java weka.classifiers.trees.HoeffdingTree -t E:\2_etapas\F1_train.arff -T E:\2_etapas\F1_test.arff

LI (1512,000) NB1 NB adaptive1

Time taken to build model: 0.57 seconds
Time taken to test model on training data: 0.07 seconds

=== Error on training data ===

Correctly Classified Instances      1511           82.7945 %
Incorrectly Classified Instances    314           17.2055 %
Kappa statistic                    0
Mean absolute error                 0.2851
Root mean squared error             0.3774
Relative absolute error             100           %
Root relative squared error         100           %
Total Number of Instances          1825
```

**1.23 pav.** Duomenų klasifikavimo tyrimas „SimpleCLI“ įrankiu

## TYRIMŲ REZULTATŲ DALIS

### 1.16. SKAIČIŲ IR VARDŲ ATPAŽINIMO TYRIMAS NAUDOJANT ŽODŽIAIS GRĮSTUS MODELIUS

Pagal 4 skyriuje aprašyto HTK programinių įrankių paketo taikymo balso komandų atpažinimui tirti, naudojant žodžių modelių metodą, metodiką, buvo sumodeliuoti vardų garsyno komandų paslėptieji Markovo modeliai, kurių atpažinimo tikslumo rezultatai, naudojant skirtingų modelių būsenų ir Gauso mišinių skaičių, aprašomi šiame skyriuje. Visi rezultatai pateikti prieduose, 6.4 lentelėje.

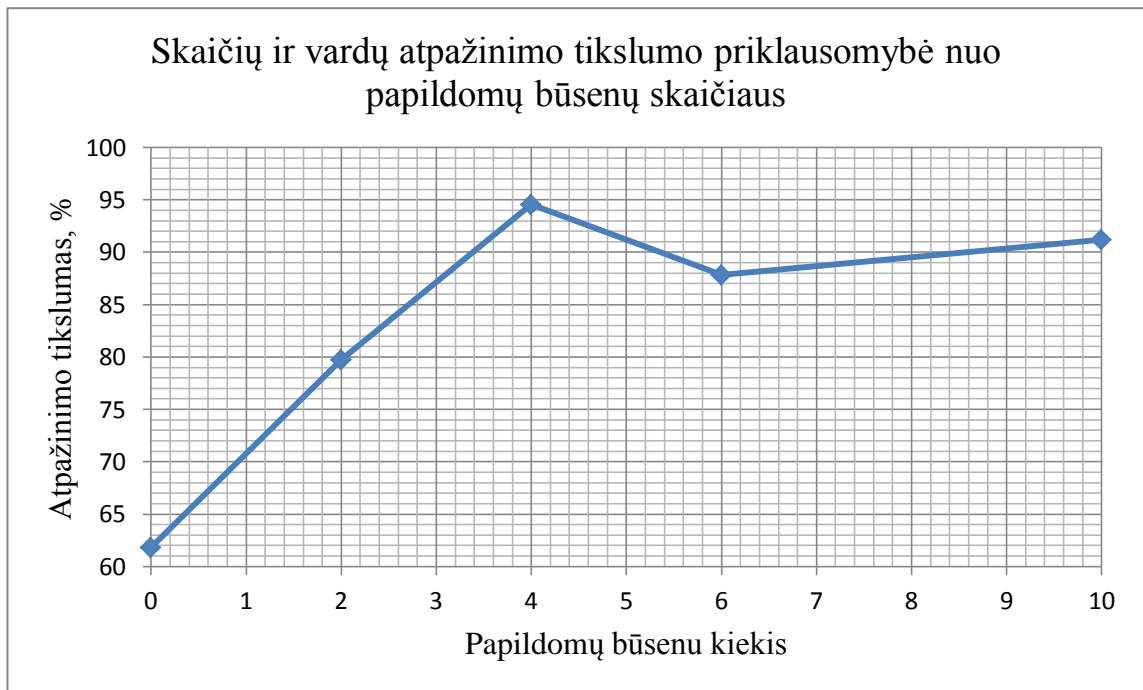
#### Papildomų būsenų įterpimas į žodžių modelius

Atpažinimo tikslumo tyrimo, išsamiai aprašyto 4 skyriuje naudojant modelius su skirtingais būsenų skaičiais, rezultatai procentais pateikti 3.1 lentelėje.

**1.5 lentelė.** Atpažinimo tikslumo tyrimų rezultatai naudojant modelius su skirtingais būsenų skaičiais.

Būsenų skaičius	0	+2	+4	+6	+10
Atpažinimo tikslumas, %	61,84	79,73	94,57	87,85	91,17

Didinant būsenų skaičių atpažinimas gerėja iki tam tikros ribos. Būsenų didinimas skaičių ir vardų garsyno atveju nebuvo naudingas, nes tyrimui naudojami žodžiai nėra labai ilgi, todėl didėjantis būsenų skaičius neilgiems žodžiams yra per didelis. Ši tendencija buvo pastebėta atlikus tyrimus komandų modeliuose didinat būsenų skaičių. Atpažinimo tikslumo priklausomybė procentais nuo būsenų skaičiaus pateikta 3.1 paveiksle.



**1.24 pav.** Skaičių ir vardų atpažinimo tikslumo priklausomybė nuo būsenų skaičiaus

Iš 3.1 lentelėje pateiktos žodžių atpažinimo tikslumo priklausomybės nuo būsenų skaičiaus . Galime daryti išvadą, kad skaičių ir vardų garsyno komandų modeliai su keturiomis papildomomis būsenomis atpažįstami didžiausiu tikslumu (94,57 %) lyginant su atpažinimo tikslumu kai modeliuose buvo naudojamos dvi, šešios, dešimt būsenų ar visai nenaudojamos papildomos būsenos. Prasčiausias atpažinimo rezultatas buvo gautas, kai modeliams nebuvo pridėdama papildomų būsenų (61,84 %). Pridėjus ir keturias ir šešias būsenas komandų modeliuose, atpažinimo tikslumas, nepaisant nuoseklaus mažėjimo didinat būsenų skaičių, vis tiek gaunamas aukštesnis nei, kad atpažinimo tikslumas visai nepridėjus papildomų būsenų komandų modeliams. Dėl to galima daryti prielaidą, kad modeliams pridėti papildomų būsenų yra naudinga, nes didėja atpažinimo tikslumas., Tačiau reikia surasti tinkamą papildomų būsenų skaičių, nes panaudojus per didelį papildomų būsenų kiekį modelyje, galima pakenkti atpažinimo tikslumo kokybei.

### **Papildomų Gauso mišinių įterpimas į žodžių modelius**

Tyrimo, išsamiai aprašyto 5 skyriuje, gautieji komandų atpažinimo tikslumo rezultatai procentais, naudojant modelius su keturiomis papildomomis būsenomis ir skirtingu Gauso mišinių skaičiumi, pateikiami 3.2 lentelėje. Pasirinkta atlikti tyrimus su keturiomis papildomomis būsenomis, nes prieš tai atlikto tyrimo metu su šiuo būsenų skaičiumi buvo gauti geriausi rezultatai. Pilnesni rezultatai pateikti priede 6.5 lentelėje.

**1.6 lentelė.** Atpažinimo tikslumo tyrimo rezultatai naudojant modelius su keturiomis papildomomis būsenomis ir skirtingu Gauso mišinių skaičiumi.

<b>+4 papildomos būsenos</b>					
<b>Gauso mišinių skaičius</b>	0	2	4	6	10
<b>Atpažinimo tikslumas, %</b>	94,57	93,01	97,7	95,98	99,49

Gauso mišinių įterpimas į žodžių komandų modelius akivaizdžiai pagerino atpažinimo tikslumo rezultatus. Iš 3.2 lentelėje pateiktos skaičių ir vardų atpažinimo tikslumo priklausomybės nuo Gauso mišinių skaičiaus matyti, kad vardų garsyno komandų modeliai su 10 Gauso mišinių atpažįstami didžiausiu tikslumu (99,49 %). Prasčiausias atpažinimo rezultatas buvo gautas, kai modeliams buvo įterpti 2 Gauso mišiniai (93,01 %). Tyrimo priklausomybė nuo Gauso mišinių skaičiaus pateikta ašinėje diagramoje, pavaizduotoje 3.2 paveiksle.



**1.25 pav.** Vardų ir skaičių atpažinimo tikslumo priklausomybė nuo Gauso mišinių skaičiaus, +6 būsenos



## Kryžminis atpažinimo tyrimas

Atliekamas kryžminis žodžių ir vardų atpažinimo tyrimas. Atpažįstant naudojami skirtingi garsyno failai apmokymui ir atpažinimui. Šiam tyrimui garsynas buvo suskirstytas į 7 dalis (sluoksnius) su skirtingais kalbėtojais atpažinimams. Tyrimas atliekamas naudojant keturias papildomas būsenas ir dešimt Gauso mišinių, nes prieš tai atliktais tyrimais tokiu būdu buvo gautas geriausias rezultatas (99,49 %). Daugiau informacijos apie garsyno suskirstymą pateikta priede. Kryžminio tyrimo metu gauti rezultatai pateikti 3.3 lentelėje.

**1.7 lentelė.** Kryžminio atpažinimo tikslumo tyrimas naudojant keturias papildomas būsenas ir 10 Gauso mišinių.

+4 papildomos būsenos ir 10 Gauso mišinių							
Kryžminio atpažinimo sluoksnis	1	2	3	4	5	6	7
Atpažinimo tikslumas, %	99,49	97,85	98,91	95	98,91	96,17	97,58

Tyrimo metu gauti rezultatai yra kiek blogesni už pradinį atpažinimo rezultatą (99,49 %). Vis dėlto, atpažįstama pakankamai gerai. Bendras žodžių ir vardų atpažinimo tikslumas naudojant keturias papildomas būsenas ir dešimt Gauso mišinių yra 97.7 %. Kryžminio atpažinimo tikslumo priklausomybė nuo testavimui naudojamam sluoksniui pateikta 3.3 paveiksle esančioje stulpelinėje diagramoje.



**1.26 pav.** Kryžminio atpažinimo tikslumo priklausomybė nuo testavimui naudojamo sluoksnio

## 1.17. SKAIČIŲ IR VARDŲ ATPAŽINIMO TYRIMAS NAUDOJANT ISPANIŠKĄ ŠNEKOS ATPAŽINTUVĄ

Tyrimo metodika naudojant ispanų kalbos atpažintuvą aprašyta 2.5 skyriuje. Atpažinimo tyrimas atliekamas panaudojus visą garsyną apmokymui ir testavimui. Gautų atpažinimo rezultatų pavyzdys pateiktas 3.4 paveiksle.

0	10094	527	nuhlejs	0nulis
1	10185	563	nojus nojus	
2	10276	639	nojus nojus	
3	10367	601	nuhlejs	0nulis
4	10462	567	nuhlejs 0nulis	
5	10556	599	nojus nojus	
6	10651	549	ehymahntahs	einantas
7	10746	573	nojus nojus	
8	10840	568	einantas	einantas
9	10935	601	nuhlejs 0nulis	
10	11037	571	nuhlejs 0nulis	
11	11132	581	trlys 3trys	
12	11230	573	trlys 3trys	
13	11329	581	trlys 3trys	
14	11427	580	nuhlejs 0nulis	
15	11526	584	nuhlejs	0nulis
16	11624	611	nojus nojus	
17	11726	575	nojus nojus	

**1.27 pav.** Ispanišku kalbos atpažintuvu gautų rezultatų pavyzdys

Ispaniško atpažintuvo rezultatai nurodo koks žodis yra atpažįstamas, atpažinimo tikimybę ir atpažinimui naudojamą failą. Naudojant šį atpažintuvą bendras žodžių atpažinimo tikslumas yra 89,15 %. Šis atpažinimo tikslumas yra blogesnis nei atpažinimo tikslumas naudojant žodžiais grįstus paslėptuosius Markovo modelius.

## 1.18. ATPAŽINTUVŲ SUJUNGIMO GALIMYBIŲ TYRIMAS

2.3 papildomųjų lentelė aiškiai parodo, kad lietuviško atpažintuvo vidutinis tikslumas yra 97,7% (T=T, T-, TF poaibiai), o ispaniško - 89,15% (T=T, FT poaibiai). Nors ispaniško atpažintuvo atpažinimo tikslumas yra mažesnis, tačiau šis atpažinimo metodas tam tikrus balso įrašus atpažįsta geriau nei lietuviškas atpažintuvas. Todėl galima daryti prielaidą, kad apjungus abu atpažintuvus galima pasiekti didesnę atpažinimo tikslumą nei naudojant tik lietuvišką atpažintuvą.

### Klasifikatorių tyrimas su žinomu kalbėtoju

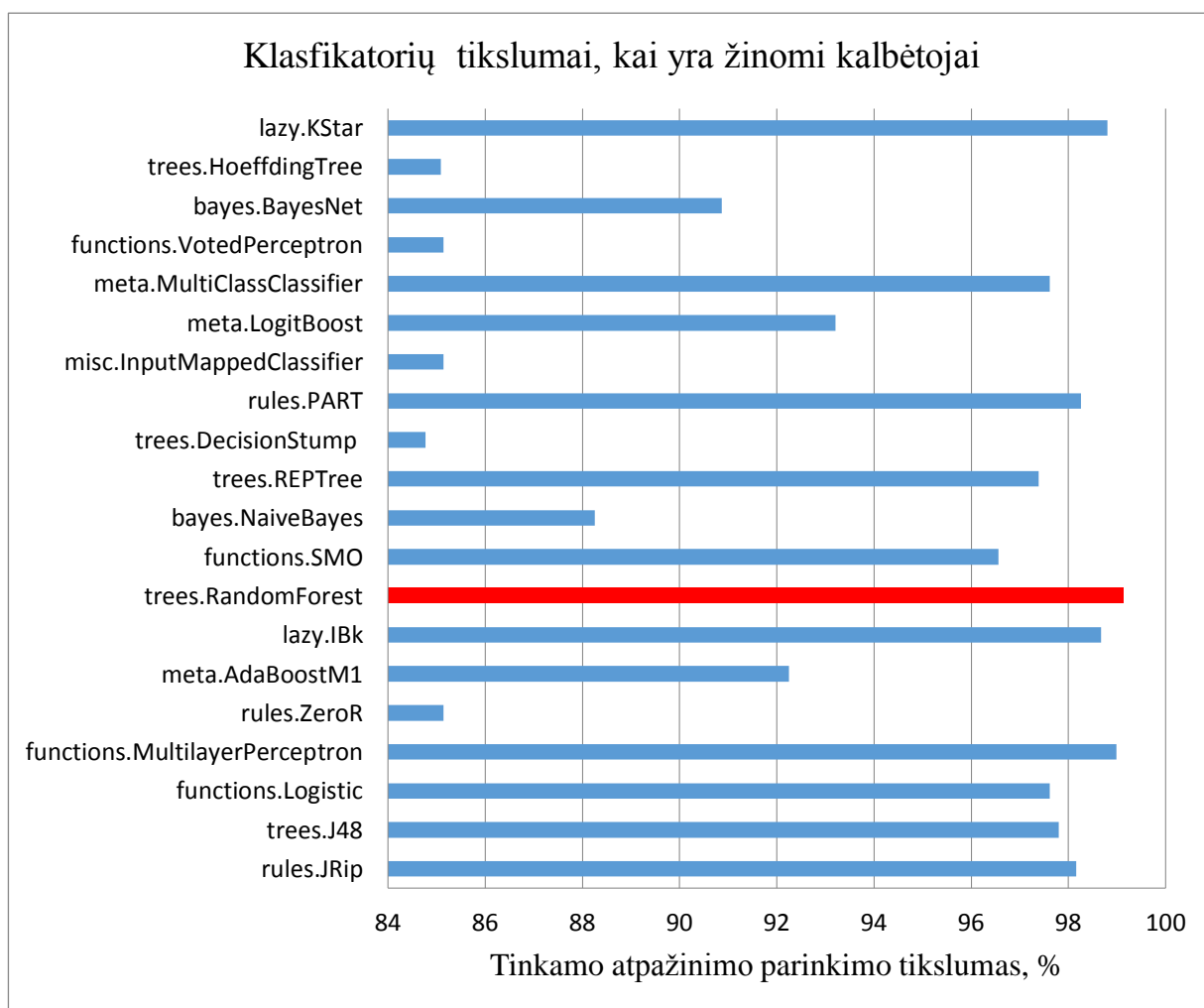
Atliekamas tyrimas su visais 2180 žodžiais, kuriems reikia parinkti tinkamą atpažintuvą. Tyrimas atliekamas apmokymui ir testavimui naudojant tą patį garsyno masyvą kuris buvo suskirstytas „WEKA“ programiniu įrankiu. Sprendimo priėmimo taisyklė buvo mokoma ir

testuojama 10-kartų kryžminio patikrinimo būdu. Buvo analizuojamas 20 skirtingų klasifikatorių efektyvumas. Tyrimo rezultatai pateikti 3.4 lentelėje.

**1.8 lentelė.** Klasifikatorių efektyvumo rezultatai kai yra žinomas kalbėtojas

Klasifikatorius	Atpažintuvo parinkimo tikslumas, %	Modelio sudarymo laikas, s
<i>rules.JRip</i>	98,1651	0,94
<i>trees.J48</i>	97,7982	0,45
<i>functions.Logistic</i>	97,6147	0,65
<i>functions.MultilayerPerceptron</i>	98,9908	39,54
<i>rules.ZeroR</i>	85,1376	0
<i>meta.AdaBoostM1</i>	92,25	0,49
<i>lazy.IBk</i>	98,6697	0
<i>trees.RandomForest</i>	99,1284	0,82
<i>functions.SMO</i>	96,5596	0,65
<i>bayes.NaiveBayes</i>	88,2569	0,08
<i>trees.REPTree</i>	97,3853	0,22
<i>trees.DecisionStump</i>	84,7704	0,08
<i>rules.PART</i>	98,2569	0,52
<i>misc.InputMappedClassifier</i>	85,1375	0
<i>meta.LogitBoost</i>	93,211	0,28
<i>meta.MultiClassClassifier</i>	97,6147	0,67
<i>functions.VotedPerceptron</i>	85,1376	0,11
<i>bayes.BayesNet</i>	90,8716	0,2
<i>trees.HoeffdingTree</i>	85,0917	0,29
<i>lazy.KStar</i>	98,8073	0

Iš 3.4 lentelės galime pastebėti, kad geriausias tikslumas buvo pasiektas naudojant atsitiktinio miško (angliškai – „*Random Forest*“) klasifikatorių. Kiti klasifikatoriai nebuvo tokie tikslūs. „*functions.MultilayerPerceptron*“ klasifikatorius pasiekė didelį tikslumą, tačiau šio klasifikatoriaus modelio sudarymo laikas yra labai ilgas (39,54 s). Nors „*rules.ZeroR*“, „*rules.ZeroR*“, „*lazy.IBk*“ ir „*misc.InputMappedClassifier*“ klasifikatorių modelio sudarymo laikas yra nulis sekundžių, tačiau jiems testuoti sugaištama laiko. „*Random Forest*“ modelio sudarymo laikas yra pakankamai trumpas norint šį klasifikatorių naudoti hibridiniam atpažintuvui. Grafiškai rezultatai pateikti 3.5 paveiksle esančioje stulpelinėje diagramoje.



**1.28 pav.** Klasifikatorių tikslumai, kai yra žinomas kalbėtojas

#### Klasifikatorių tyrimas su nežinomu kalbėtoju

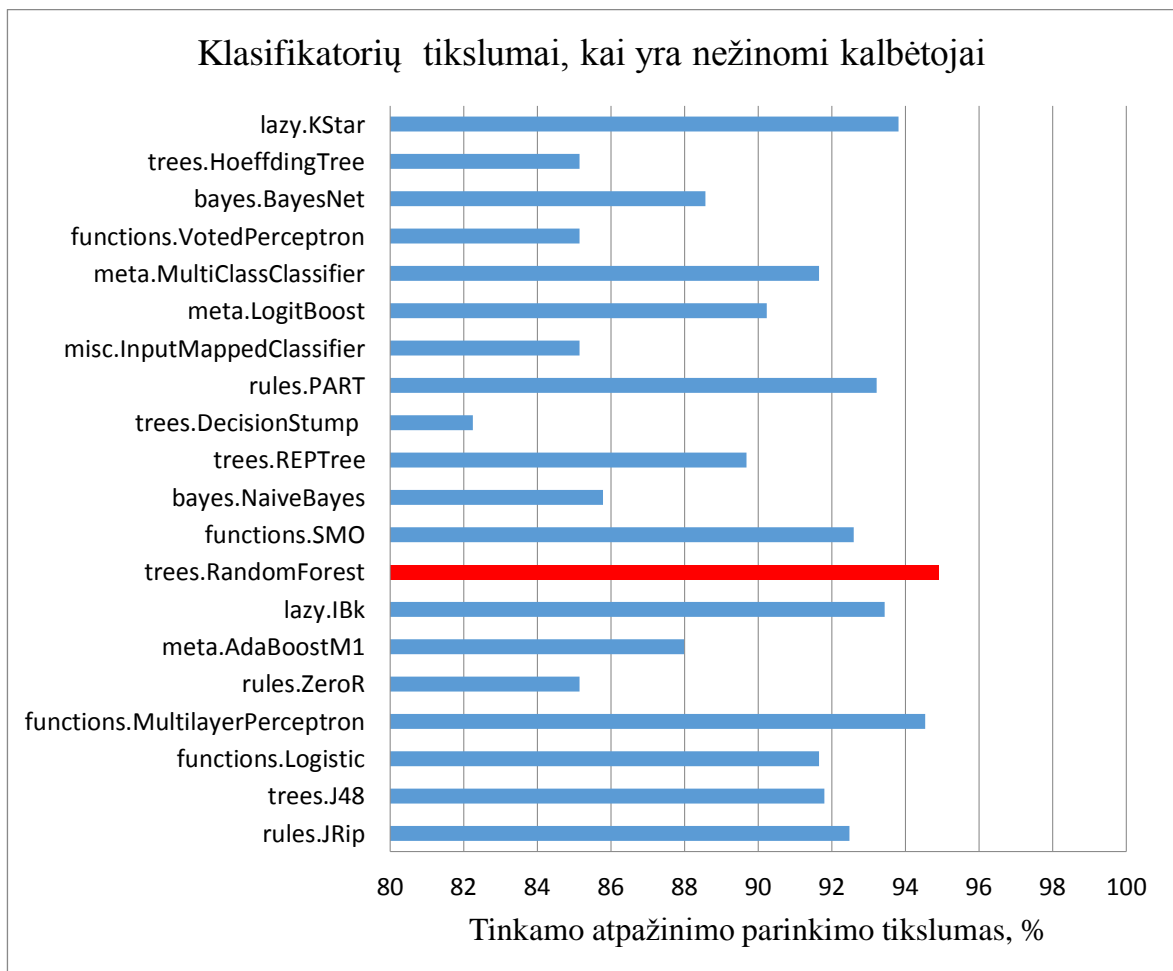
Šio tyrimo metu garsynas suskirstytas į septynias dalis (sluoksnius) su skirtingais diktoriais atpažinimui. Kiekvienas sluoksnis yra tiriamas su anksčiau minėtais 20 klasifikatorių. Kryžminiai klasifikatorių efektyvumų tyrimų rezultatai su nežinomu kalbėtoju yra pateikti 3.5 lentelėje. Pilni kiekvieno sluoksnio rezultatai pateikti prieduose 6.6 lentelėje.

**1.9 lentelė.** Klasifikatorių efektyvumo rezultatai kai yra nežinomi kalbėtojai

Klasifikatorius	Atpažintuvo parinkimo tikslumas %	Modelio sudarymo ir testavimo laikas, s
<i>rules.JRip</i>	92,48	0,48
<i>trees.J48</i>	91,79	0,23
<i>functions.Logistic</i>	91,65	0,43
<i>functions.MultilayerPerceptron</i>	94,54	41,2
<i>rules.ZeroR</i>	85,14	0,01
<i>meta.AdaBoostM1</i>	87,98	0,27
<i>lazy.IBk</i>	93,44	0,46

<i>trees.RandomForest</i>	94,9	0,58
<i>functions.SMO</i>	92,6	0,35
<i>bayes.NaiveBayes</i>	85,78	0,14
<i>trees.REPTree</i>	89,68	0,07
<i>trees.DecisionStump</i>	82,25	0,03
<i>rules.PART</i>	93,21	0,19
<i>misc.InputMappedClassifier</i>	85,14	0,02
<i>meta.LogitBoost</i>	90,23	0,15
<i>meta.MultiClassClassifier</i>	91,65	0,15
<i>functions.VotedPerceptron</i>	85,14	0,13
<i>bayes.BayesNet</i>	88,56	0,15
<i>trees.HoeffdingTree</i>	85,14	0,12
<i>lazy.KStar</i>	93,81	53,8

Šie tyrimo rezultatai parodo, kad geriausias klasifikatorius pasiekiantis didžiausia atpažintuvo parinkimo tikslumą yra atsitiktinio miško („*Random Forest*“). Pasiektas tikslumas yra 94,9 %, o modelio sudarymo ir testavimo laikas yra pakankamai trumpas 0,58 sekundės. Grafiškai rezultatai pateikti 3.6 paveiksle esančioje stulpelinėje diagramoje.



**1.29 pav.** Klasifikatorių tikslumai, kai apmokymui ir testavimui naudojamas kryžminiu būdu suskirstyti garsynai

### „Auto-WEKA“ klasifikatoriaus parinkimo tyrimas

Šio tyrimo metu atliekamas klasifikatoriaus panaudojamas „Auto-WEKA“ įrankį, kuris pagal nustatyta laiko tarpa suranda geriausią klasifikatorių ir jo parametrus pilnam garsynui. Pasirinkus atpažinimo laiką 15 minučių kaip geriausias klasifikatorius yra išskiriamas „functions.SMO“, jo tikslumas 99,9 %. Padidinus laiko tarpą iki 100 minučių geriausiu klasifikatoriumi parenkamas „Random Forest“, jo tikslumas 100 %.

### „Random Forest“ klasifikatoriaus panaudojimo tyrimas

Atlikus atpažintuvų sujungimo tyrimus galime daryti išvadą, kad „Random Forest“ klasifikatorius geriausiai parenka reikiamą atpažintuvą. Šio tyrimo tikslas yra apskaičiuoti kokią atpažinimo tikslumą gali pasiekti sistema su dvejais atpažintuvais ir „Random Forest“ klasifikatoriumi. Tam naudojami kryžminiai atsitiktinio miško klasifikatoriaus parinkimo tyrimo rezultatai, jie pateikti 3.6 lentelėje.

**1.10 lentelė.** Kryžminiai „Random Forest“ klasifikatoriaus atpažintuvų parinkimo rezultatai kai yra naudojami nežinomi kalbėtojai

Poaibis	Failų (frazių) kiekis atskiriems sluoksniams						
	1	2	3	4	5	6	7
T=T	2202	2234	2245	2148	2329	2219	2275
F=F	0	0	0	0	0	2	2
T-	19	6	32	22	4	55	17
F-	0	1	0	2	1	0	0
TF	326	265	255	262	199	188	206
FT	10	46	19	114	9	68	58
FF	3	8	9	12	18	28	2
Iš viso	2560	2560	2560	2560	2560	2560	2560
Teisingo atpažintuvo parinkimo tikimybė	0,99	0,93	0,98	0,81	1	0,98	0,99

Turint šiuos duomenis galime apskaičiuoti koku tikslumu veiks hibridinis atpažintuvas kai yra žinomas ir nežinomas kalbėtojas klasifikatorių parinkimui. Tokio atpažintuvo atpažinimo tikslumai yra apskaičiuoti pagal žemiau pateiktą formulę ir parodyti 3.7 lentelėje

$$\frac{((TF + FT + T -) * P_{klasf}) + T = T}{mas} * 100$$

Čia,

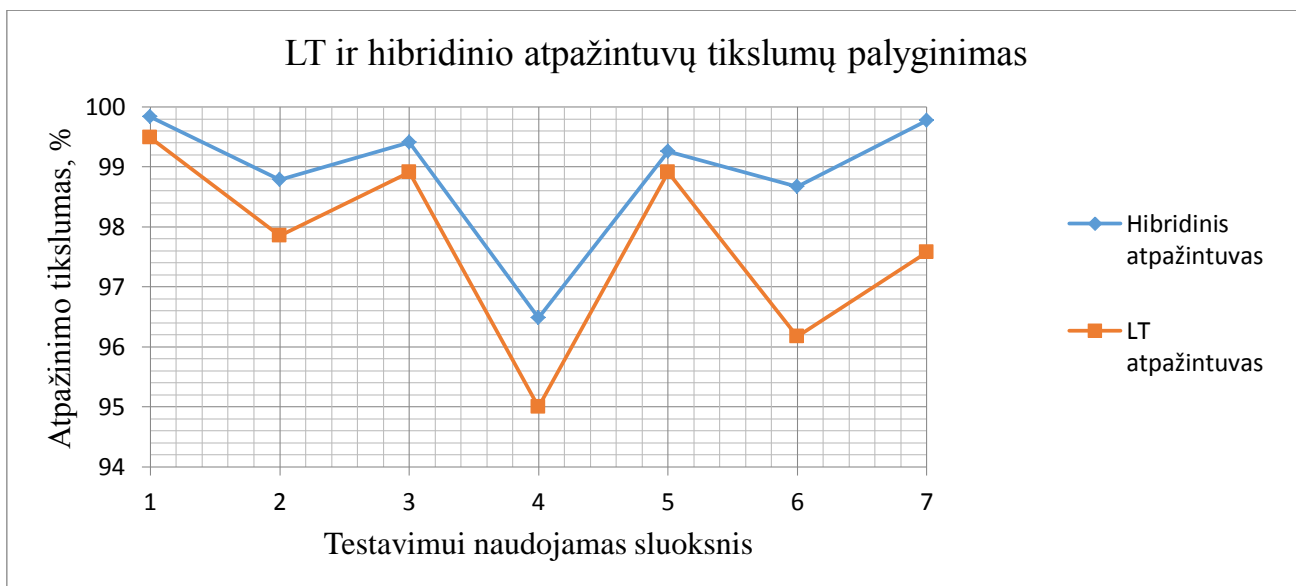
$P_{klasf}$  – klasifikatoriaus tikimybė teisingai parinkti atpažintuvą;

mas – iš viso atpažinti naudojamų frazių kiekis

**1.11 lentelė.** Hibridinio atpažintuvo atpažinimo tikslumai ir jų palyginimai kai testavimui yra naudojami nežinomi kalbėtojai

Sluoksnis	1	2	3	4	5	6	7
Hibridinio atpažintuvo atpažinimo tikslumas, %	99,84	98,79	99,41	96,48	99,26	98,67	99,77
Lietuviško atpažintuvo atpažinimo tikslumas, %	99,49	97,85	98,91	95	98,91	96,17	97,58
Pagerėjimas, %	0,35	0,94	0,50	1,48	0,35	2,50	2,19

Vidutinis hibridinio atpažintuvo skaičių ir vardų atpažinimo tikslumas yra 98.89%, o vidutinis lietuviško atpažintuvo naudojančio paslėptuosius Markovo modelius atpažinimo tikslumas yra 97,7%. Naudojant hibridinį atpažintuvą atpažinimo tikimybė vidutiniškai padidėjo 1,19%. Atpažintuvų tikslumų palyginimą galime matyti 3.7 paveiksle.



**1.30 pav.** Hibridinio ir LT atpažintuvų palyginimas,

## IŠVADOS IR REZULTATAI

1. Atlikus skaičių ir vardu komandų atpažinimo tikslumo tyrimus, taikant paslėptuosius Markovo modelius ir keičiant būsenų skaičių pastebėta, kad aukščiausias tikslumas buvo pasiektas naudojant komandų modelius su keturiomis (4) papildomomis būsenomis (94,6%).
2. Atlikus skaičių ir vardu komandų atpažinimo tikslumo tyrimus, taikant paslėptuosius Markovo modelius ir keičiant Gauso mišinių skaičių pastebėta, kad aukščiausias tikslumas buvo pasiektas naudojant komandų modelius su dešimt Gauso (10) mišinių (99,5%).
3. Atlikus skaičių ir vardu komandų kryžminius atpažinimo tikslumo tyrimus, taikant paslėptuosius Markovo modelius ir naudojant 4 papildomas būsenas ir 10 Gauso mišinių, nustatyta, kad vidutinis atpažinimo tikslumas yra 97,7%.
4. Palyginus atpažinimų tikslumų rezultatus gautus su ispanišku atpažintuvu (89,2%) ir kryžminius atpažinimo tikslumų rezultatus gautus su „HTK“ programiniu paketu (97,7%), yra pastebėta, kad sistemos naudojančios paslėptuosius Markovo modelius atpažinimo tikslumas yra didesnis.
5. Atlikus klasifikatorių efektyvumo tyrimus žinomam ir nežinomam kalbėtojiui yra pastebėta, kad tiksliausiai atpažintuvą parenka atsitiktinio miško („*Random Forest*“) klasifikatorius.
6. Atlikus hibridinio atpažintuvo tyrimą su „*Random Forest*“ klasifikatoriumi sprendimų priėmimui, nežinomam kalbėtojiui, pastebėta, kad šis atpažintuvas prognozuoja didesnę vidutinį atpažinimo tikslumą (98,9%), nei pavieniai atpažintuvai.
7. Atlikus hibridinio atpažintuvo tyrimą su „*Random Forest*“ klasifikatoriumi sprendimų priėmimui, žinomam kalbėtojiui, pastebėta, kad šis atpažintuvas prognozuoja didžiausią atpažinimo tikslumą (99,4%).



## LITERATŪROS ŠALTINIAI

- 1. Laurinčiukaitė, S. (2008).** Lietuvių šnekos atpažinimo akustinis modeliavimas: daktaro disertacija: technologijos mokslai, informatikos inžinerija 07T / Vilniaus Gedimino technikos universitetas, Vilnius: Technika. 134 p. Interaktyvus [žiūrėta 2016-04-27]. Prieiga per internetą: [www.mii.lt/files/mii\\_dis\\_08\\_laurinciukaite.pdf](http://www.mii.lt/files/mii_dis_08_laurinciukaite.pdf)
- 2. Jurafsky, D.; Martin, J. H. (2000).** Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition. Prentice Hall, Upper Saddle River, New Jersey 07458. Inc, 934 p.
- 3. Driaunys K. (2006).** Lietuvių šnekamosios kalbos segmentavimo ir fonetinio atpažinimo tyrimas naudojant LTDIGITS garsyno įrašus: daktaro disertacija: fiziniai mokslai, informatika / Vilniaus universitetas. Vilnius. 171 p.
- 4. Vaičiukynas, E. (2009).** Balso dialogų technologijos (kompiuteriniai dialogai balsu): Informacinių technologijų metodai: Kauno technologijos universitetas (T120D004). Interaktyvus [žiūrėta 2016-04-29]. Prieiga per internetą: [http://www.elen.ktu.lt/studentai/lib/exe/fetch.php?media=balso\\_dialogu\\_tech.pdf](http://www.elen.ktu.lt/studentai/lib/exe/fetch.php?media=balso_dialogu_tech.pdf)
- 5. Pranevičius H., Raudys Š., Rudžionis A., Rudžionis V., Ratkevičius K., Sakalauskaitė J., Makackas D. (2008).** Agentinių sistemų modeliai: mokomoji knyga. Vilnius: Mokslo aidai. 255 p. ISBN 978-9955-591-55-9.
- 6. Filipovič M. (2005).** Atskirai tariamų lietuvių šnekos žodžių atpažinimo, grindžiamo dirbtiniais neuroniniais tinklais ir paslėptais Markovo modeliais, tyrimai: daktaro disertacijos santrauka: fiziniai mokslai, informatika / Vytauto Didžiojo universitetas. Kaunas. 138 p.
- 7. Kasparaitis, P. (2005).** Kompiuterinės lingvistikos paskaitų konspektai. Interaktyvus [žiūrėta 2016-05-12]. Prieiga per internetą: <http://www.mif.vu.lt/~pijus/CL/cl.htm>
- 8. Paškauskaitė G. (2015).** Balsu išstartų kodų atpažinimo sistemos sukūrimas ir tyrimas: Magistro baigiamasis; Kauno technologijos universitetas, Elektros, Kaunas. 70 p.
- 9. Maskeliūnas R. (2009).** Lietuviškų balsu komandų atpažinimas daugybinių transkripcijų pagrindu: daktaro disertacija: technologijos mokslai, informatikos inžinerija – 07T / Kauno technologijos universitetas. Kaunas. 159 p.
- 10. Gales M. (2008).** „The Application of hidden Markov Models in Speech Recognition“ Cambridge University Engineering Department, Cambridge. 110p.
- 11. Young S. (2002).** „The HTK Book (for HTK Version 3.2)“ Cambridge University Engineering Department, Cambridge. 335p.
- 12. Ž. Ringelienė, M. Filipovič (2011).** Žodžių atpažinimo, grįsto paslėptaisiais Markovo modeliais, vizualizavimo ir analizės programinė įranga. 10 p.
- 13. G. Bartišiūtė, G. Paškauskaitė (2015).** „Šnekos atpažintuvų sujungimo galimybių tyrimas“, E2TA-2015 konferencija, Kaunas, Lietuva. ISBN 978-609-02-1133-5. 20-23p.

- 14. Rudžionis A., Ratkevičius K., Rudžionis V., Kasparaitis P., Šalna B. (2001).** Balso technologijų taikymo lietuvių kalbai analizė ir perspektyvinių veiklos kryptių pagrindimas: ataskaita / Kauno technologijos universitetas, Vilniaus universitetas, Teismo ekspertizės centras. Kaunas, Vilnius. 84 p., [žiūrėta 2016-09-26]. Prieiga per internetą: [http://www.likit.lt/frames/balso\\_tech/balsotech\\_st.htm](http://www.likit.lt/frames/balso_tech/balsotech_st.htm).
- 15.** <https://www.xn--ratija-ckb.lt/liepa> [žiūrėta 2016-10-10]
- 16. Rudžionis, V., Ratkevičius, K., Rudžionis, A., Raškinis, G., Maskeliūnas, R. (2013).** “Hibridinė atpažinimo technologija balso sąsajai (INFOBALSAS)”. Aukštųjų technologijų plėtros 2011–2013 metų programa. Mokslo, inovacijų ir technologijų agentūra. Baigiamoji ataskaita. Kaunas, Lietuva, 80 p.
- 17. Fosberg, M. (2003).** Why is speech recognition difficult? Chalmers university, Departement of Computer science. Interaktyvus [žiūrėta 2015-03-17] Prieiga per internetą: [http://www.speech.kth.se/~rolf/gslt\\_papers/MarkusForsberg.pdf](http://www.speech.kth.se/~rolf/gslt_papers/MarkusForsberg.pdf)
- 18. Bahdanau, Dzmitry (2016).** End-to-End Attention-based Large Vocabulary Speech Recognition.
- 19. Ratkevičius, K. (2013).** Hibridinė atpažinimo technologija balso sąsajai (Infobalsas): ataskaita. Kauno technologijos universitetas, UAB „Softdent“. Kaunas. 50 p.
- 20. Raškinis, A., Kazlauskienė, A. Raškinis, G. (2003).** VDU bendrinės lietuvių šnekos universalus anotuotas garsynas. Vytauto Didžiojo universitetas.
- 21. Balvočius, B., Telksnys, L. (2003).** Garsynų duomenų modeliai ir programinės įrangos architektūros. Vytauto Didžiojo universitetas, Matematikos informatikos institutas.

## PRIEDAI

1.12 lentelė. Darbe atpažįstamų skaičių garsynas

Skaičius	Tarimas
0	Nulis
1	Vienas
2	Du
3	Trys
4	Keturi
5	Penki
6	Šeši
7	Septyni
8	Aštuoni
9	Devyni

1.13 lentelė. Darbe atpažįstamų vardų garsynas

Raidė	Tarimas	Raidė	Tarimas
A	Austėja	N	Nojus
B	Boleslovas	O	Oskaras
C	Cecilija	P	Patrikas
D	Donatas	R	Ričardas
E	Eimantas	Q	Kju
F	Fausta	S	Sandra
G	Gražvydas	T	Teodoras
H	Hansas	U	Ulijona
I	Izaokas	V	Vacys
J	Jonas	W	Vašington
K	Karolis	X	Iksas
L	Laima	Y	Ygrekas
M	Martynas	Z	Zacharijus

1.14 lentelė. Garsyno suskirstymas pagal diktorius į septynias dalis (sluoksnius)

Pirmas	Antras	Trečias	Ketvirtas	Penktas	Šeštas	Septintas
FIEVVIS	FAGNGRA	FDAILOI	FVANPEC	FLAUZET	MKAZANU	FUGNBUC
FJUSKIN	FAGNVIN	FGINGED	FVILVAI	FRAISAV	MMODSLE	FAISIZI
FUGNBUC	FAISIZI	FIEVJUR	FVIONAB	FRUTNAN	MRIMAPA	FIEVJUR
FUGNNOV	FAISZYM	FIEVSAB	MDAIGUS	FSIMMEI	MVYGVAI	FVIONAB
FZIVSTA	FAUSNEM	FKAMMOS	MEDGVO L	FVAIVAI	MLINJUR	MRIMAPA
FGINTRA	FANGRUM	FMLRAU	MJUOCES	FIVEVAL	MROKKUO	MZYGSVE
FGRETUB	FGINPAS	FSEVBUT	MKASRAT	FJULBAL	MSARNEM	MLAUBAR
FINDBEN	FLAUKLU	MANDMAR	FEGLZAJ	MJURBIZ	FGINBAR	MDARJEG

1.15 lentelė. Skaičių ir vardų atpažinimo tikslumo tyrimų rezultatai naudojant paslėptuosius Markovo modelius ir keičiant būsenų skaičių

Atpažinimo tikslumo priklausomybė nuo būsenų skaičius, %					
Žodis	Būsenų skaičius atitinka raidžių sk.				
	+0 būsenos	+2 būsenos	+4 būsenos	+6 būsenos	+10 būsenų
Austeja	100	86,7	96,7	95	91,7
Boleslovas	100	100	95	96,7	96,7
Cecilija	85	96,7	95	98,3	96,7
Donatas	58,3	98,3	100	15	100
Eimantas	83,3	100	98,3	100	45
Fausta	58,3	70	100	100	100
Grazvydas	100	83,3	100	96,7	96,7
Hansas	26,7	96,7	98,3	100	91,7
Izaokas	100	100	100	0	100
Jonas	21,7	48,3	71,7	100	95
Karolis	75	95	95	3,3	95
Laima	25	65	90	98,3	88,3
Martynas	56,7	100	100	100	100
Nojus	16,7	58,3	96,7	100	48,3
Oskaras	75	100	98,3	20	95
Patrikas	53,3	100	100	100	93,3
Qju	0	1,7	95	96,7	100
Ricardas	36,7	98,3	93,3	95	86,7
Sandra	16,7	98,3	98,3	96,7	98,3
Teodoras	98,3	98,3	100	100	96,7
Ulijona	100	98,3	100	91,7	96,7
Vacys	3,3	50	86,7	95	81,7
Wasington	100	100	100	100	100
Xsas	0	1,7	55	96,7	68,3
Ygrekas	73,3	95	100	16,7	70

Zacharijus	98,3	98,3	95	98,3	61,7
Nulis	45	67	100	100	92
Vienas	100	88	100	100	100
Du	0	7	61	97	99
Trys	0	51	96	100	89
Keturi	99	98	99	98	99
Penki	56	80	99	100	98
Šeši	77	81	96	99	97
Septyni	100	100	100	100	100
Aštuoni	100	96	98	98	95
Devyni	69	90	97	97	89
<b>Vidurkis, %</b>	<b>61,84</b>	<b>79,73</b>	<b>94,57</b>	<b>87,85</b>	<b>91,17</b>

1.16 lentelė. Skaičių ir vardų atpažinimo tikslumo tyrimų rezultatai naudojant paslėptuosius Markovo modelius ir Gauso mišinių skaičių

Atpažinimo tikslumo priklausomybė nuo Gauso mišinių skaičius, %					
Skaičius	4 papildomos būsenos				
	2 mišiniai	3 mišiniai	6 mišiniai	10 mišinių	16 mišinių
Austeja	98,3	95	69,7	100	100
Boleslovas	98,3	98,3	98,3	100	100
Cecilija	98,3	100	100	100	100
Donatas	91,7	98,3	96,7	100	100
Eimantas	96,7	90	100	100	100
Fausta	100	100	100	100	100
Grazvydas	100	98,3	100	100	100
Hansas	93,3	100	96,7	100	100
Izaokas	98,3	95	95	100	100
Jonas	56,7	83,3	100	100	100
Karolis	96,7	98,3	98,3	100	100
Laima	88,3	100	100	100	100
Martynas	100	100	100	100	100
Nojus	85	100	100	100	100
Oskaras	96,7	100	76,7	100	96,7
Patrikas	100	95	86,7	100	100
Qju	96,7	100	100	100	100
Ricardas	100	93,3	96,7	100	100
Sandra	80	100	100	100	100
Teodoras	100	86,7	96,7	100	100
Ulijonas	90	100	100	100	100
Vacys	43,3	100	95	96,7	96,7
Wasington	100	100	100	100	100
Xsas	85	100	83,3	100	100

Ygrekas	91,7	90	71,7	100	100
Zacharijus	85	100	78,3	100	100
Nulis	99	100	99	100	100
Vienas	91	97	98	97	99
Du	98	98	98	98	98
Trys	95	99	96	99	100
Keturi	95	97	96	97	98
Penki	96	100	100	100	100
Šeši	100	100	100	100	97
Septyni	100	100	100	100	100
Aštuoni	98	100	100	100	100
Devyni	87	97	90	98	98
Vidurkis, %	93,01	97,7	95,98	99,49	99,38

1.17 lentelė. Kryžminiai klasifikatorių parinkimo tyrimų rezultatai

Adaboost					
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s	
1	97,7465	355	0,04	0,36	
2	89,9054	317	0,02	0,21	
3	94,4444	306	0,02	0,27	
4	78,392	398	0,01	0,22	
5	97,6415	212	0,03	0,17	
6	73,6334	311	0,01	0,23	
7	88,6121	281	0,01	0,19	
Jrip					
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s	
1	99,1549	355	0,01	0,84	
2	92,429	317	0,03	0,46	
3	97,3856	306	0,02	0,34	
4	79,8995	398	0,01	0,26	
5	100	212	0,02	0,45	
6	87,1383	311	0,02	0,39	
7	96,7972	281	0,01	0,34	
logistic					
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s	
1	98,5915	355	0,04	0,62	
2	89,5899	317	0,01	0,45	
3	93,4641	306	0,02	0,44	
4	80,402	398	0,03	0,2	
5	99,0566	212	0,02	0,43	

6	92,283	311	0,04	0,28
7	92,8826	281	0,02	0,36
<b>naïve Bayes</b>				
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	87,3239	355	0,1	0,07
2	83,9117	317	0,24	0,03
3	89,5425	306	0,08	0,01
4	76,6332	398	0,07	0,01
5	87,7358	212	0,12	0,02
6	91,3183	311	0,09	0,01
7	87,1886	281	0,07	0,01
<b>SMO</b>				
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	98,0282	355	0,04	0,67
2	92,429	317	0,02	0,27
3	95,7516	306	0,01	0,34
4	76,6332	398	0,01	0,17
5	97,6415	212	0,01	0,27
6	95,8199	311	0,01	0,28
7	97,7357	281	0,01	0,27
<b>RepTree</b>				
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	94,0845	355	0,05	0,04
2	87,6972	317	0,02	0,05
3	97,0588	306	0,02	0,04
4	76,8844	398	0,02	0,04
5	98,5849	212	0,01	0,05
6	83,6013	311	0,01	0,04
7	96,4413	281	0,02	0,04
<b>rulesPART</b>				
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	98,5915	355	0,01	0,18
2	94,3218	317	0,01	0,17
3	96,7320	306	0,01	0,17
4	81,1558	398	0,01	0,15
5	99,5283	212	0,03	0,22
6	91,9614	311	0,01	0,19
7	95,0178	281	0,01	0,18
<b>logitBoost</b>				
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	96,338	355	0,02	0,08
2	85,489	317	0,02	0,09
3	93,4641	306	0,01	0,34

4	79,397	398	0,01	0,09
5	98,1132	212	0,01	0,1
6	95,1768	311	0,02	0,09
7	88,2562	281	0,01	0,09
<b>VotedPerceptron</b>				
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	97,1831	355	0,09	0,06
2	85,489	317	0,09	0,04
3	93,7908	306	0,1	0,04
4	71,3568	398	0,07	0,03
5	95,7547	212	0,11	0,04
6	78,135	311	0,09	0,03
7	79,3594	281	0,09	0,04
<b>trees.HoeffdingTree</b>				
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	97,1831	355	0,01	0,09
2	85,489	317	0,02	0,1
3	93,7908	306	0,01	0,12
4	71,3568	398	0,02	0,12
5	95,7547	212	0,04	0,09
6	78,135	311	0,01	0,1
7	79,3594	281	0,01	0,12
<b>lazyKSTAR</b>				
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	96,0563	355	0,02	50,77
2	92,7445	317	0,02	53,1
3	97,3856	306	0,02	53,87
4	81,407	398	0,01	48,25
5	100	212	0,03	63,28
6	97,4277	311	0,01	53,59
7	97,153	281	0,01	54,26
<b>BayesNET</b>				
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	89,8592	355	0,03	0,16
2	87,6972	317	0,03	0,07
3	88,5621	306	0,03	0,1
4	78,8945	398	0,02	0,15
5	92,9245	212	0,03	0,14
6	92,6045	311	0,03	0,11
7	93,9502	281	0,03	0,13
<b>MultiClassClassifier</b>				
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	98,5915	355	0,03	0,16



2	89,5899	317	0,03	0,07
3	93,4641	306	0,03	0,1
4	80,402	398	0,02	0,15
5	99,0566	212	0,03	0,14
6	92,283	311	0,03	0,11
7	92,8826	281	0,03	0,13

#### ImputMappetClassifier

Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	97,1831	355	0,02	0
2	85,489	317	0,03	0
3	93,7908	306	0,03	0
4	71,3568	398	0,01	0
5	95,7547	212	0,03	0
6	78,135	311	0,02	0
7	79,3594	281	0,02	0

#### DecisionStump

Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	88,7324	355	0,02	0,02
2	85,489	317	0,02	0,02
3	86,2745	306	0,01	0,02
4	71,3568	398	0,01	0,01
5	91,0377	212	0,01	0,01
6	78,135	311	0,01	0,01
7	79,3594	281	0,02	0,01

#### zeroR

Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	97,1831	355	0,01	0
2	85,489	317	0,01	0
3	93,7908	306	0,01	0
4	71,3598	398	0,01	0
5	95,7547	212	0,02	0
6	78,135	311	0,01	0
7	79,3594	281	0,01	0

#### Radom Forest

Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	99,7183	355	0,13	0,78
2	93,0599	317	0,07	0,54
3	98,0392	306	0,07	0,41
4	80,9048	398	0,06	0,37
5	100	212	0,07	0,45
6	98,7138	311	0,06	0,45
7	99,2883	281	0,07	0,48

#### multiLayer

Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	99,4366	355	0,05	37,22
2	94,0063	317	0,04	37,05
3	99,0196	306	0,06	43,77
4	80,6533	398	0,05	43,65
5	100	212	0,04	43,26
6	97,4277	311	0,04	41,95
7	96,4413	281	0,04	45,23
<b>lazyIBK</b>				
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	96,338	355	0,53	0,04
2	92,429	317	0,48	0,01
3	96,7320	306	0,44	0,01
4	79,397	398	0,39	0,02
5	99,5283	212	0,45	0,01
6	98,0707	311	0,39	0,01
7	97,5089	281	0,44	0,02
<b>J48</b>				
Sluoksniai	Tikslumas, %	Masyvas	Testavimo laikas, s	modelio sudarymo laikas, s
1	96,6197	355	0,03	0,42
2	89,9054	317	0,03	0,26
3	97,0588	306	0,02	0,16
4	85,9296	398	0,01	0,14
5	97,1698	212	0,01	0,16
6	82,3151	311	0,01	0,14
7	96,7972	281	0,01	0,12