

**VILNIAUS UNIVERSITETAS
KAUNO HUMANITARINIS FAKULTETAS**

INFORMATIKOS KATEDRA

Verslo informatikos studijų programa
Kodas 62109P101

KĘSTUTIS TOMKEVIČIUS

MAGISTRO BAIGIAMASIS DARBAS

**BIOMETRIJOS TAIKYMO E-KOMERCIJOS
AUTENTIFIKAVIMO MODULIUOSE GALIMYBIŲ TYRIMAS**

Kaunas 2008

**VILNIAUS UNIVERSITETAS
KAUNO HUMANITARINIS FAKULTETAS**

INFORMATIKOS KATEDRA

KĘSTUTIS TOMKEVIČIUS

MAGISTRO BAIGIAMASIS DARBAS

**BIOMETRIJOS TAIKYMO E-KOMERCIJOS
AUTENTIFIKAVIMO MODULIUOSE GALIMYBIŲ TYRIMAS**

Leidžiama ginti _____

Magistrantas _____
(parašas)

Darbo vadovas _____
(parašas)

dr. doc. Vytautas Rudžionis
(darbo vadovo mokslo laipsnis, mokslo
pedagoginis vardas, vardas ir pavardė)

Darbo įteikimo data _____

Registracijos Nr. _____

Kaunas 2008

TURINYS

PAVEIKSLŲ SĄRAŠAS.....	4
LENTELIŲ SĄRAŠAS.....	5
SUTRUMPINIMŲ SĄRAŠAS	6
Santrauka.....	7
ĮVADAS.....	8
1. ANALIZĖS DALIS.....	11
1.1. Autentifikacijos priemonės	11
1.2. Slaptažodžiai	12
1.3. Slaptažodžių patikimumas: stiprūs ir silpni slaptažodžiai.....	13
1.4. Slaptažodžių saugumo tyrimai	14
1.4.1. Užsienyje.....	15
1.4.2. Lietuvoje.....	15
1.4.3. Autoriaus atliktas Lietuvos kompiuterių naudotojų slaptažodžių stiprumo galimybių ir elgsenos tyrimas I.....	17
1.4.4. Autoriaus atliktas Lietuvos kompiuterių naudotojų slaptažodžių stiprumo galimybių ir elgsenos tyrimas II	19
1.4.5. Slaptažodžių tyrimų išvados.....	21
1.5. Biometrinės autentifikavimo priemonės	22
1.6. Skirtingų autentifikavimo priemonių palyginimas.....	27
1.7. Diktoriaus atpažinimas	29
1.7.1. Tyrimai diktoriaus atpažinimo srityje	29
1.7.2. Diktoriaus atpažinimas kalbos atpažinimo srities kontekste.....	34
1.7.3. Algoritmai diktoriaus atpažinimo sistemose	35
1.7.4. Diktoriaus atpažinimo praktinio taikymo sritys	38
1.8. Analiztinės dalies išvados	38
2. Diktoriaus atpažinimas naudojant fonemų diskriminavimo metodą.....	40
2.1. Duomenų paruošimas	40
2.2. Diktoriaus klasifikacija pagal fonemas	45
3. EKSPERIMENTINIS SKYRIUS	47
3.1. Duomenys.....	47
3.2. Duomenų apdorojimas ir analizė.....	48
3.3. Rezultatų analizė ir apibendrinimas	50
IŠVADOS ir pasiūlymai.....	73
LITERATŪROS SĄRAŠAS.....	75
Mokslinė literatūra	75
Informacijos šaltiniai	76
PRIEDŲ SĄRAŠAS	78

PAVEIKSLŲ SĄRAŠAS

1 pav. Informacijos apsaugos priemonių klasifikacija	12
2 pav. Dažniausios slaptažodžių sudarymo technikos	16
3 pav. Dažniausios slaptažodžių sudarymo technikos (atsižvelgiant į klavišų patogumą)	16
4 pav. Dažniausios slaptažodžių sudarymo technikos (neatsižvelgiant į klavišų patogumą)	16
6 pav. Biometrijos sistemų veikimo principas ir sandara.....	22
7 pav. Žmogaus biometrinių savybių klasifikacija.....	24
8 pav. S. Parthasarathy ir A. E. Rosenberg pasiūlytas sistemos veikimo principas	31
9 pav. Verbalinės informacijos verifikavimo schema.....	33
10 pav. Principinė diktoriaus atpažinimo sistemos diktoriaus identifikavimo schema.....	34
11 pav. Principinė diktoriaus atpažinimo sistemos diktoriaus verifikavimo schema	34
12 pav. Požymių vektoriaus sudarymo schema	37
13 pav. Sistemos, naudojančios fonemų diskriminavimo metodą diktoriui atpažinti, schema.....	40
14 pav. LTDIGITS garsyno failinė struktūra	42
15 pav. *.PHN failo eilutės struktūra	43
16 pav. *.RBB failo eilutės struktūra	43
17 pav. Fonemos pradžios ir pabaigos nustatymo algoritmo blokinė schema	44
18 pav. Lietuvių kalbos fonemų LTDIGITS garsyne klasifikacija.....	47
19 pav. Papildyta LTDIGITS failinė struktūra.....	49
20 pav. Diktorių vyrų fonemų „a“, „b“, „d“, „e“, „r“ ir „t“ skaitinių įverčių histogramos.....	53
21 pav. Vyrų diktorių fonemų panašumo skaitinių įverčių minimalių, maksimalių ir skirtumų grafiškas atvaizdavimas.....	54
22 pav. Diktorių moterų fonemų „a“, „b“, „d“, „e“, „r“ ir „t“ skaitinių įverčių histogramos.....	55
23 pav. Moterų diktorių fonemų panašumo skaitinių įverčių minimalių, maksimalių ir skirtumų grafiškas atvaizdavimas.....	56
24 pav. Fonemų skaitinių įverčių skirtumai tarp lyčių.....	57
25 pav. 1 eilutėje dažniausiai pasitaikančios fonemos pagal panašumų įverčius	59
26 pav. 13 eilutėje dažniausiai pasitaikančios fonemos pagal panašumų įverčius	60
27 pav. 17 eilutėje dažniausiai pasitaikančios fonemos pagal panašumų įverčius	60
28 pav. 1 eilutėje dažniausiai pasitaikančios fonemos pagal panašumų įverčius	62
29 pav. 14 eilutėje dažniausiai pasitaikančios fonemos pagal panašumų įverčius	63
30 pav. 21 eilutėje dažniausiai pasitaikančios fonemos pagal panašumų įverčius	63
31 pav. Diktorių vyrų ir moterų fonemų pasikartojimo skirtumai 1 eilutėje	66
32 pav. Diktorių vyrų ir moterų fonemų pasikartojimo skirtumai 14 eilutėje	66
33 pav. Diktorių vyrų ir moterų fonemų pasikartojimo skirtumai 20 eilutėje	67
34 pav. Unikalių fonemų skaičiaus skirtumų vyrų ir moterų diktorių eilutėse diagrama	67
35 pav. Dominuojančių fonemų skaičiaus skirtumų vyrų ir moterų diktorių eilutėse diagrama	68
36 pav. Vyrų diktorių I+IV ir II+III kvartilų rezultatų skirtumų histograma	71
37 pav. Moterų diktorių I+IV ir II+III kvartilų rezultatų skirtumų histograma	72

LENTELIŲ SĄRAŠAS

1 lentelė. 2-ojo slaptažodžių tyrimo duomenys sugrupuoti pagal slaptažodžių tipus	19
2 lentelė. 2-ojo slaptažodžių tyrimo duomenys: slaptažodžių ilgiai	20
3 lentelė. „Įvairių biometrinių technologijų palyginimas pagal A. K. Jain“	24
4 lentelė. Biometrinių savybių palyginimas	26
5 lentelė. „Prieigos valdymo sprendimų palyginimas“	28
6 lentelė. Eksperimento metu naudotos fonemos ir jų žymėjimas	47
7 lentelė. Netinkamo formato pradinės tyrimo fonemos	50
8 lentelė. Neapdoroti vyrų diktorių fonemų apdorojimo rezultatai	51
9 lentelė. Neapdoroti moterų diktorių fonemų apdorojimo rezultatai	51
10 lentelė. Vyrų diktorių fonemų įverčių minimalios, maksimalios, skirtumo ir vidutinės reikšmės 52	
11 lentelė. Moterų diktorių fonemų įverčių min, max, skirtumo ir vidutinės reikšmės	54
12 lentelė. Fonemų įverčių skirtumai lyčių atžvilgiu	56
13 lentelė. Vyrų diktorių fonemų skaitiniai įverčiai išrikiuoti pagal tinkamumo eilę	58
14 lentelė. Moterų diktorių fonemų skaitiniai įverčiai išrikiuoti pagal tinkamumo eilę	58
15 lentelė. Dažniausiai diktorių vyrų atpažinimo eilės eilutėse pasikartojančios fonemos	60
16 lentelė. Skirtingų bei dominuojančių fonemų skaičius ir vidutinis fonemos pasikartojimas diktoriaus atpažinimo tinkamumo eilėje	61
17 lentelė. Dažniausios fonemos kiekvienoje eilutėje	63
18 lentelė. Skirtingų bei dominuojančių fonemų skaičius ir vidutinis fonemos pasikartojimas diktoriaus atpažinimo tinkamumo eilėje	64
19 lentelė. Fonemų charakteristikų diktorių vyrų ir moterų atpažinimo tinkamumo eilėje skirtumai 65	
20 lentelė. Dominuojančių fonemų skaičiaus įtaka fonemų pasikartojimui	68
21 lentelė. Diktorių duomenys ir dažniausiai eilutėse pasikartojančios fonemos	69
22 lentelė. Fonemų sutapimų su eilės dominuojančiomis fonemomis rezultatai	70
23 lentelė. Rezultatų apibendrinimo lentelė	70

SUTRUMPINIMŲ SĄRAŠAS

CHMM – Ciklinių paslėptųjų Markovo modelių metodas

DV – Diktoriaus verifikavimas

ERR – Klaidos atmetimo santykis

HMM – Paslėptųjų Markovo modelių metodas

LLR – Prisijungimo tikimybės koeficientas

LTRHMM – Iš kairės į dešinę paslėptųjų Markovo modelių metodas

MAR – Multivariacinis autoregresinis modelis

MFCC, MDKK – Melo dažnių kepstiniai koeficientai

PIN – Asmeninis identifikavimo numeris

LPC – Tiesinis prognozavimas

TDA – Tiesinio diskriminavimo analizė

VIV – verbalinės informacijos verifikavimas

SANTRAUKA

TOMKEVICIUS, Kestutis. (2008) Research of Applying Biometrics in E-commerce Authentication Modules. MBA Graduation Paper. Kaunas: Vilnius University, Kaunas Faculty of Humanities, Department of Informatics. 77 p.

SUMMARY

Contemporary society is highly dependent on machines and technology. As a result, the reassurance in information security becomes a global issue, when going to the internet world. Information security is closely associated with privacy protection and thus, is of high importance nowadays. There are several major tasks in the Master's thesis. First, this paper aims to investigate the world-wide standards of authentication systems and the devices used. Second, after the foregoing views on certain effects of the evaluation procedure and its possible drawbacks, it aims to explore, whether or not the speaker recognition based on the discrimination of phoneme is considered to be one of the most promising areas for the future research. In other words, the process of investigation seeks to find out whether it is possible to provide a thorough analysis of the speaker verification system and distinguish the significant phonemes in the subsets of different speaker phonemes that would allow forming the characteristic password phrases and if they would, how these phonemes were to be selected. In the feature-extraction stage of the speaker verification system, the Mel frequency cepstral coefficients (MFCC) give the bases for the research paper and various most suitable phonemes of the speaker are counted, while using the likelihood methods. The results of experiments indicate that variations in gender of speaker-based phoneme features as well as the proper recognition of the digital likelihood differ greatly. And the best way for the speaker recognition is the whole set of phonemes of the second and fourth quarter phonemes. The Master's thesis consists of 77 pages. This paper includes 23 tables and 37 pictures.

ĮVADAS

Šiandien žmogus, kadaise sakęs, kad gyvenimas persikels į kompiuterį ir internetą jau nebebūtų laikomas svajotoju ar ligoniu. Visuomenė vis labiau tampa priklausoma nuo mašinų. Informacijos saugumo užtikrinimas, persikraustant į elektroninę erdvę, vis aktualesnis. Informacijos srautai negrižtamai išaugo, padidėjo apdorojamos ir kaupiamos informacijos kiekis tiek privačioje, tiek verslo aplinkoje, o verslo žmonės turėtų būti itin susirūpinę savo informacijos saugumu. Informacija informacinėje visuomenėje tampa pinigais. Laikai pasikeitė – galingiausi, įtakingiausi ir turtingiausi yra tie, kurie valdo informaciją. Siekiant geresnius rezultatus duodančių sprendimų, verslas yra tiesiogiai susijęs su mokslu. Savo informacijos apsaugojimas nuo neleistino panaudojimo yra kiekvieno iš mūsų reikalas. Taigi baigiamajame darbe nagrinėsime informacijos apsaugos problemas.

Mokslo požiūriu tyrimai informacijos saugumo srityje galimai veda į naujų prekių ir paslaugų teikimą naudotojams.

Darbe remtasi įvairių mokslininkų, dirbančių biometrijos ir diktoriaus verifikavimo srityje, tyrimų rezultatais. Pasinaudota A. K. Jain, S. Furui, T. Matsui, R. J. Mammone, Xiaoyu Zhang ir kitų mokslininkų darbų rezultatais.

Darbo objektas – balso ir diktoriaus atpažinimo algoritmai ir metodai, fonemų ir diktorių ryšiai ir jų įvertinimas.

Darbo tema aktuali tuo, kad analizuojami metodai yra ne tik moksliniai, tačiau visi šie ir anksčiau atlikti tyrimai veda į praktinį pritaikymą ir galutiniam naudotojui skirtos priemonės sukūrimą. Autentifikavimo tyrimų rezultatai taip pat yra svarbūs vertinant socialinių tinklų rezultatus, kurie patys savaime gali būti atskira autentifikavimo priemonė.

Darbo tikslas – atlikti dabar naudojamų informacijos apsaugos priemonių vertinimą, apžvelgti perspektyviausius artimiausioje ateityje naudotinus būdus bei atlikus eksperimentinius tyrimus, įvertinti, ar diktoriaus diskriminavimas fonemų atžvilgiu gali būti taikomas diktoriaus atpažinimo uždaviniams spręsti.

Siekiant darbo tikslo, darbo eiga skaidyta į smulkesnius etapus. Darbo *uždaviniai* yra:

1. Atlikti tyrimus Lietuvos kompiuterių naudotojų dažniausiai naudojamų informacinių saugos priemonių srityje, siekiant išsiaiškinti pagrindinius elgsenos aspektus.
2. Apžvelgti kitus pažangiausius ir plačiausias galimybes praktiniame pritaikyme turinčius duomenų saugos metodus.
3. Išanalizavus paskutiniųjų metų mokslininkų veiklos tendencijas balso atpažinimo srityje, pasirinkta diktoriaus atpažinimo probleminė sritis.

4. Paruošus praktiniam tyrimui reikalingus duomenis, atlikti praktinį eksperimentą, kurio metu bus siekiama įvertinti, ar siūlomas sprendimas yra vertas tolesnių tyrimų ir praktinės realizacijos.
5. Iširti, ar skirtingiems diktoriams galima suformuoti juos identifikuojantį fonemų aibės poaibį, ir jei taip, kokios fonemos turėtų būti paimamos iš visos fonemų eilės.
6. Atlikti eksperimento rezultatų apibendrinimą, pateikti išvadas bei siūlymus.

Iškelta hipotezė: naudojant fonemų diskriminavimą iš visos kalboje naudojamos aibės galima atrinkti fonemų, geriausiai identifikuojantį atskirą diktorių, rinkinį. Remiantis eksperimentiniu tyrimu siekiama šią hipotezę patvirtinti arba paneigti.

Darbas sudaro trys pagrindinės dalys – teorinė (analizės) dalis, siūlomos sprendimo metodikos dalis bei eksperimento dalis. Kiekvienoje iš šių dalių pateikiama specifinė informacija apie darbo sritį ir darbo objektus bei atliekamus eksperimentinius tyrimus.

- *Teorinėje/Analizės dalyje* atlikta šiuo metu dažniausiai taikomų duomenų apsaugos priemonių apžvalga. Siekiant išsiaiškinti, kiek saugios šiuo metu labiausiai paplitę duomenų apsaugos priemonės, atlikti du slaptažodžių saugumo tyrimai. Taip pat šiame skyriuje aptariamos alternatyvios apsaugos priemonės, apžvelgiamos biometrijos taikymo galimybės, nustatoma, kad tolimesni tyrimai bus atliekami sprendžiant diktorius atpažinimo uždavinius. Apžvelgiami iki šiol atlikti tyrimai šioje srityje.
- *Siūloma sprendimo metodika* pateikiama antrojoje dalyje. Aptariama duomenų struktūra, reikalingi veiksmai duomenims apdoroti, aprašomas fonemų diskriminavimo metodas.
- *Eksperimentinėje dalyje* aprašomas atliktas tyrimas, analizuojami bei apibendrinami gauti rezultatai. Pateikiamas galutinis vertinimas, kad pasiūlytas metodas tinka diktoriams atpažinti.

Rašant baigiamąjį darbą, informacija buvo renkama iš įvairių šaltinių, naudotos įvairios priemonės informacijai surinkti: naudotasi internetiniais šaltiniais, mokslinių straipsnių duomenų bazėmis nustatyti, kokie naujausi tyrimai vykdomi, remtasi šioje srityje dirbančių Universiteto mokslininkų patirtimi ir atradimais, naudotos apklausos siekiant surinkti reikalingus duomenis.

Darbe naudoti *metodai*:

- Visuotinio pažinimo metodas (metodas naudotas darbo tikslų nustatymui, uždavinių formavimui, informacijos apie vertimų biurų veiklą rinkimui ir analizei; duomenims apibendrinti; išvadoms formuluoti).
- Bendrieji mokslinio tyrimo metodai:
 - pilnosios indukcijos metodas (naudotas darbo išvadoms formuluoti);

- dedukcijos metodas.
- Kokybiniai tyrimo metodai:
 - anketavimas, naudotas norint išsiaiškinti naudotojų elgseną susijusią su jų informacijos apsauga.
- Analizės duomenų analizės metodas.
- Apibendrinimo metodas.
- Abstrakcijos metodas – juo remiantis daromos kiekvienos darbo dalies ir galutinės viso darbo išvados.

Darbo metu pasiekti rezultatai įrodo, kad pasirinktas metodas yra tinkamas diktoriaus atpažinti. Įvertinus fonemų eilės pagal fonemos panašumo skaitinius įverčius ir dažniausiai toje eilėje pasikartojančių fonemų sutapimo/nesutapimo dažnius, daroma išvada, kad nors fonemos esančios pirmame imties kvartilyje, turi didžiausius panašumo įverčius, tačiau šios fonemos per dažnai sutampa su diktorių aibės atitinkamos eilutės fonemomis. Be to, pirmojo ir ketvirtojo kvartilio fonemos dominuoja visų eilėje esančių fonemų aibės atžvilgiu.

Darbo rezultatų praktinė reikšmė – apjungiant Vilniaus universiteto mokslininkų gautus rezultatus su šio darbo metu gautais rezultatais, plėtoti tolimesnius lietuvių kalbos fonemų atpažinimo tyrimus ir perkelti juos į praktiką, kuriant priemones natūraliam naudotojo bendravimui su kompiuteriu.

Darbo atlikimo metu pastebėta, jog turėtų duomenų nepakanka itin giliems tyrimams, kurių metu galima būtų pastebėti detalesnes tendencijas diktoriaus atpažinime diskriminuojant fonemas. Pagrindinė to priežastis – LTDIGITS duomenų bazės žodyno siaurumas. Siekiant rezultatus pritaikyti praktiškai, pradžioje dar turėtų būti atliktas garsyno bazės plėtimas kasdienėje kalboje vartojamais žodžiais, o tada atlikti papildomus rezultatų vertinimus.

Darbo apimtis – 77 puslapių be priedų, su priedais – 129 puslapiai. Be to, darbe yra 23 lentelės, 37 paveikslai bei 14 priedų. Prieduose yra 6 lentelės, 1 anketų maketas, 21 paveikslas, 4 programų išeities tekstai, 1 straipsnis.

Rezultatų aprobavimas – tarpiniai 2008 metais rezultatai buvo paskelbti studentų mokslinėje-praktinėje konferencijoje „Akademinių jaunimo siekiai: Ekonomikos, vadybos ir technologijų išvalgos“ straipsnyje „Diktoriaus atpažinimo sistema paremta fonetinių vienetų diskriminavimu“ („Akademinių Jaunimo siekiai: Ekonomikos, Vadybos ir Technologijų išvalgos“. 2008. ISSN 2029-0217. p. 168 - 172).

1. ANALIZĖS DALIS

1.1. Autentifikacijos priemonės

Informacijos apsaugos sritis neįsivaizduojama be termino „autorizacija“ ir „autentifikacija“.

Internete autentifikacija (iš graikiško *αυθεντικός*; tikras arba nesuklastotas, autentiškas, ir iš *authentēs*; autorius) apibrėžiama kaip kažko arba kažkieno patvirtinimo tikru veiksmas (Wikipedia, 2007a). Kompiuterių moksle autentifikavimas suprantamas kaip bandymas patvirtinti komunikacijos siuntėjo skaitmeninę tapatybę, pavyzdžiui siunčiant prašymą prisijungti prie sistemos. Internete autentifikacija yra būdas, skirtas užtikrinti, kad naudotojai yra tai, kas jie sakosi esą – naudotojas, kuris siekia atlikti tam tikras funkcijas sistemoje turi teises tiems veiksams atlikti (Wikipedia, 2007a). Autentifikacija, paprastai šnekant, suprantama kaip naudotojo veiksmų laisvės padidinimo kompiuterinėse ar interneto sistemose laidas ar priemonė.

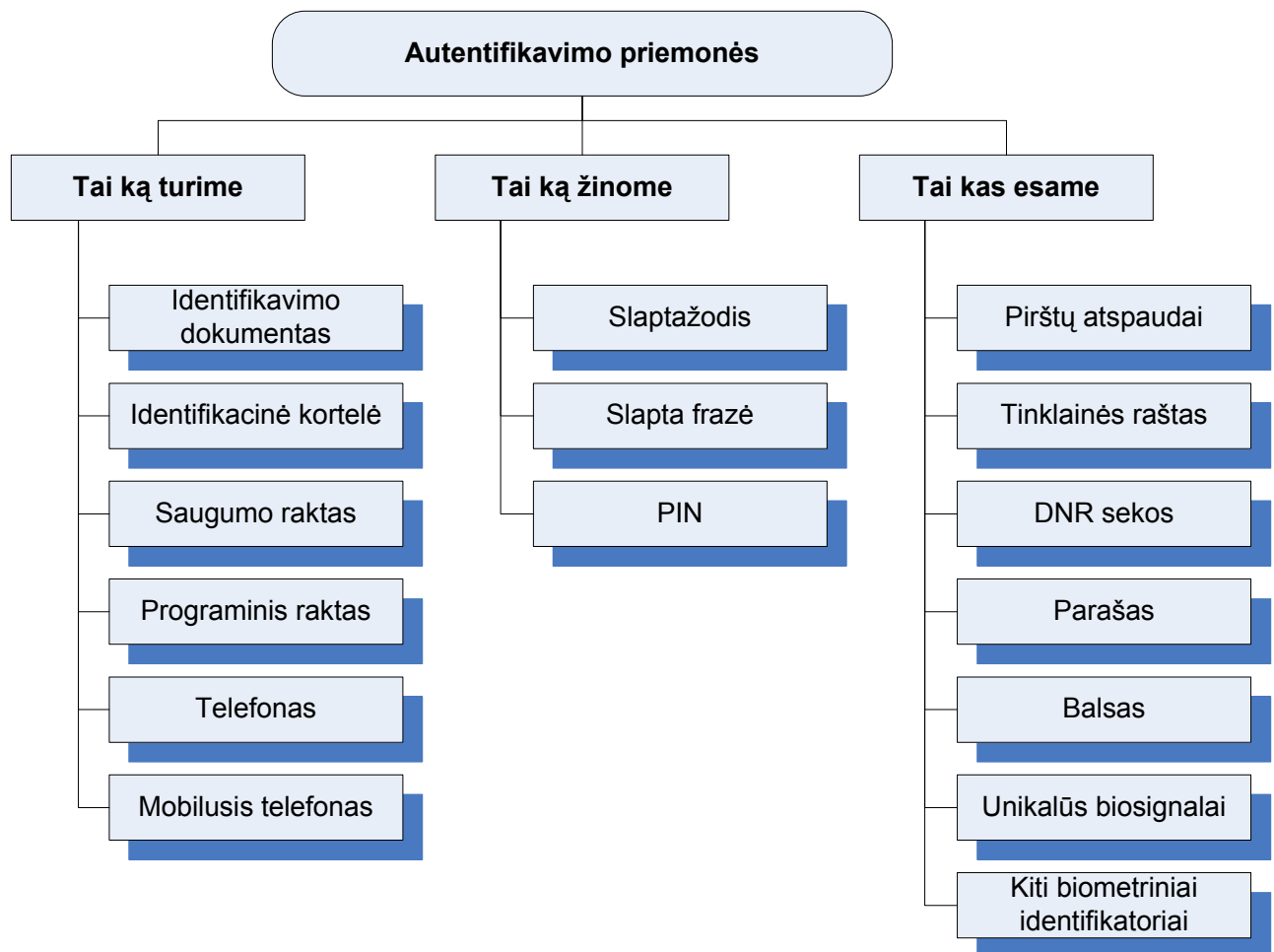
Autentifikavimas pagal naudojamas priemones, įprastai klasifikuojamas į tris grupes:

- Autentifikavimas pagal tai, ką turite;
- Autentifikavimas pagal tai, ką žinote;
- Autentifikavimas pagal tai, kas esate ar ką darote.

Pirmosios dvi grupės apibrėžia daiktus ar informaciją, kurią naudotojas turi pateikti sistemai, norėdamas identifikuoti save ir atlikti tam tikrus veiksmus ar prisijungti prie sistemos. Pirmajai grupei priskiriami asmens tapatybės kortelės (angl. *ID card*), saugumo raktai (angl. *security token*), programiniai raktai (angl. *software token*), telefonas arba mobilusis telefonas. Šios grupės elementai daugiausiai naudojami bankų, tinklų ir programinės įrangos autentifikacijai. Antroji grupė – informacija, kurią naudotojas žino: slaptažodis, slapta frazė, asmeninis identifikacinis numeris (angl. *personal identification number, PIN*). Šios grupės elementai yra plačiausiai paplitę ryšio priemonių, interneto svetainių ir aplikacijų autentifikacijoje, todėl darbo sritis yra susijusi būtent su slaptažodžiais.

Trečioji sritis apibrėžia naudotojo autentifikavimą naudojant žmogaus biometrinių savybių tikrinimą. Biometrinių savybių panaudojimas geriausiai plačiai visuomenei žinomas iš fantastikos ir mokslinės fantastikos filmų: rankų pirštų atspaudų atpažinimas, akies rainelės nuskaitymas, balso atpažinimas. Tačiau ši sritis apima ne tik jau paminėtąsias biometrikas, tačiau ir daugelį kitų, pvz. delnų venų raštas, ausis, eiseną ir t.t.

Šios grupės ir joms priklausančių priemonių klasifikavimas pateikiamas 1 paveiksle.



Šaltinis: sudaryta autoriaus.

1 pav. Informacijos apsaugos priemonių klasifikacija

Kaip jau minėta, viena iš plačiausiai šiuo metu paplitusių priemonių yra informacija, skirta autentifikacijos siekiančiam asmeniui identifikuoti. Kasdiniame gyvenime iš šios grupės priemonių naudojami slaptažodžiai, todėl tolesniuose skyreliuose pirmiausia plačiau analizuojami slaptažodžiai, jų stiprumas, naudotojų elgsena ir kiti saugumą lemiantys veiksniai.

1.2. Slaptažodžiai

Kompiuteriniame pasaulyje, norint pasinaudoti tam tikrais ištekliais arba juos apsaugoti, visur naudojami slaptažodžiai – vos įsijungus kompiuterį, norint prieiti prie jo išteklių, prisijungti prie interneto teikėjo serverio, prieiti prie savo el. pašto dėžutės, el. komercijos sistemos ir t.t. Slaptažodžiai reikalingi naudotojo identifikacijai ir apsaugai nuo kitų naudotojų, kurie galėtų pasinaudoti svetimais ištekliais. Tačiau net slaptažodžiai ne visada gelbėja, yra tam tikros situacijos, kad slaptažodžiai piktavališkai „nulaužiami“, arba naudotojai juos išduoda patys – viena iš pagrindinių problemų apsisaugoti nuo neautorizuoto informacijos peržiūrėjimo internete yra naudotojo nesugebėjimas tinkamai apsaugoti savo informacijos slaptažodžiu. Naudotojai dažniausiai naudoja netinkamus slaptažodžius savo interneto pašto dėžučių apsaugai, bankiniuose

tinklapiuose, kur yra reikalaujama naudoti ne tik banko suteiktą naudotojo identifikacijos numerį ir slaptažodžius iš slaptažodžių kortelės arba generatoriaus, bet ir savo pasirinktą slaptažodį, registruojantis į įvairius forumus, diskusijas ir t.t.

1.3. Slaptažodžių patikimumas: stiprūs ir silpni slaptažodžiai

Bet kokie slaptažodžiai, nepriklausomai nuo to, ar jie naudojami interneto svetainėse, intraneto aplikacijose, ar paprasčiausiai asmeninio kompiuterio programose, turinčiose galimybę naudotojams naudotis slaptažodžiais informacijai apsaugoti, pagal savo stiprumą ir galimybes jį „nulaužti“ (atspėti, perrinkti iš visų įmanomų slaptažodžių variantų ir naudojantis kitais būdais), yra skirstomi į silpnus ir stiprius slaptažodžius.

Silpnas slaptažodis yra toks slaptažodis, kurį nesunku atspėti tiek žmonėms, tiek kompiuteriams. Žmonės dažnai naudoja akivaizdžius slaptažodžius tokius kaip vaikų vardai, jų namų numeriai, arba gimtadienių datos, tam kad jų neužmirštų. Tačiau kuo paprastesnis slaptažodis, tuo jį yra lengviau nustatyti (TheFreeDictionary, 2006 a).

Stiprus slaptažodis yra toks slaptažodis, kurį yra sunku nustatyti tiek žmonėms, tiek kompiuteriams. Du dalykai padaro slaptažodį stipresniu: 1) didesnis kiekis raidžių ir 2) skaitmenų įmaišymas, mažosios ir didžiosios raidės bei specialiųjų simbolių buvimas (\$, #, etc.) (TheFreeDictionary, 2006 b).

Slaptažodžių stiprumas matuojamas taip vadinamuoju bito stiprumu, tuo pačiu metodu, kuris skirtas matuoti užkodavimo stiprumui. Bito stiprumas paprastai nurodo bendrą „permutations“ skaičių. Bito stiprumas paskaičiuojamas paimant visų vieno simbolio galimų variantų skaičių ir padauginant jį iš slaptažodžio ilgio (simbolių skaičiaus). Vieno simbolio variacijų skaičius yra išreiškiamas slaptažodžio politika (schema, kurios turi būti laikomasi sudarant slaptažodį) bei informacijos entropija (slaptažodžio atsitiktinumumu). (Wikipedia, 2007 b).

Slaptažodžių stiprumą kiekvienas naudotojas gali nesunkiai pasitikrinti neatliekant sudėtingų matematinių skaičiavimų. Internete gausu slaptažodžių stiprumo matuoklių, besiskiriančių savo algoritmais, stiprumo matavimo skalėmis, prieinamumu. Didžioji dalis jų yra nemokamos. Kai kurios svetainės, ypač pašto, bankinės sistemos, kai kurie turinio valdymų sistemų moduliai jau turi slaptažodžio stiprumo tikrinimo priemonę, taigi naudotojas, kurdamas slaptažodį iš karto gali pasitikrinti, ar pakankamai saugūs jo duomenys.

Kita vertus, ne vien slaptažodžio stiprumas lemia duomenų apsaugą. Dažnai saugumas yra aukojamas vardan patogumo – naudotojas, norėdamas prisiminti savo slaptažodį, dažniausiai jį sudaro taip, kad slaptažodis keltų jam asociacijas, kurios padėtų jam prisiminti slaptažodį. Dažnai tai yra įvairios datos, artimų žmonių vardai, kiti bendriniai ir tikriniai žodžiai. Visa tai, pažįstant asmenį arba jo aplinką, padidina galimybes atspėti naudotojo slaptažodį. O jei naudotojas turi ne

vieną vietą, kur naudojamas tas pats slaptažodis (šiuolaikinėje informacijos visuomenėje tai yra vis rečiau pasitaikantis atvejis)? Atspėjus slaptažodį vienoje vietoje, pažeidžiama tampa ir informacija kitose vietose, kur prieiga prie jos saugoma tuo pačiu slaptažodžiu.

Trečias atvejis kaip informacija apie asmens prisijungimus gali nesunkiai patekti į piktavališkas rankas yra pakankamas kiekis šnipinėjimo ir duomenų perėmimo programinės ir techninės įrangos. Naudotojo kompiuteryje gali būti įdiegta viena ar kelios programos, kurios fiksuoja visus naudotojo atliekamus veiksmus ir spaudomus klavišus. Paskui šią informaciją gali patikrinti programą instaliavęs žmogus. Dažnai net nereikia fiziškai prieiti prie kompiuterio – dažną šnipinėjimo programą galima taip sukonfigūruoti, kad ji ataskaitas apie dienos, valandos ar kito periodo atliktus veiksmus kompiuteryje siųstų į tam tikslui naudojamą pašto adresą. Tokios programos veikia visiškai pasislėpusios, dažnai naudojasi standartinių operacinės sistemos procesų pavadinimais, o suaktyvinti jų langą reikia žinoti trijų keturių klavišų kombinaciją. Mažiau patyrusiam kompiuterio naudotojui tai yra beveik neįmanoma ją aptikti. Atskiras duomenų perėmimo atvejis yra duomenų perėmimas, kuomet duomenys į internetą yra siunčiami nesaugiais kanalais. Šiuo atveju taip pat užtenka valandą pasirausti interneto lobynuose ir labiau patyręs kompiuterių naudotojas jau bus apsiginklavęs pakankamu kiekiu programinės įrangos, skirtos kaimyno ar kito žmogaus tinkle sekimui ir jo internetu siunčiamų informacijos paketų perėmimui. Net jei paketai ir siunčiami užšifruoti, kodą vis tiek galima dešifruoti vėliau, kuomet duomenų paketai jau saugiai guli piktus kėslus turinčio asmens kompiuteryje.

Naudotojas gali sukurti saugų slaptažodį, kurį sunku nuspėti, perrinkti ar kitais būdais „nulaužti“, tačiau vis dar nemenką dalį problemų sukelia tai, jog turėdamas kelis slaptažodžius, naudotojas pradeda juos maišyti, užmiršti, ir tuomet visus juos susirašo į užrašų knygutę, mobilųjį telefoną ar dar blogiau – tiesiog į tekstinę bylą kompiuterio darbastalyje. Tai slaptažodžio saugumą tik sumažina – užrašų knygutę gali pažiūrėti bet kas, kas žino kur ji guli, bylą kompiuteryje gali pažiūrėti bet kas, kas turi teisę dirbti kompiuteriu, telefoną gali pavogti, jį galima pamesti ir t.t.

Taigi, apibendrinant visus šiuos atvejus, galima teigti, kad kuo slaptažodis yra lengviau įsimenamas, statistikai jis yra lengviau pažeidžiamas, ir priešingai, kuo jis sudėtingesnis, sudarytas iš 9 – 12 įvairių simbolių kombinacijų, kurios tarpusavyje neturi jokios prasmės, tuo sunkiau jį įsiminti naudotojui. Kuo sunkiau jį įsiminti, tuo dažniau naudotojai į kur nors užsirašo. Taigi, nors slaptažodžiai yra dažniausiai kompiuterio bei duomenų saugumui internete naudojama priemonė, ji turi nemažai trūkumų.

1.4. Slaptažodžių saugumo tyrimai

Kadangi slaptažodžiai, kaip jau minėta, yra viena iš populiariausių identifikavimo ir autentifikavimo priemonių internete, šiai apsaugos priemonių grupei tyrinėtojai skiria nemažai

dėmesio. 2002 – 2004 metais JAV ir Lietuvoje buvo atlikta keletas su naudotojų slaptažodžių saugumu ir jų sudarymo manieromis tyrimų.

1.4.1. Užsienyje

Psichologijos tyrinėtojai tyrė (2004 m.) kaip žmonės renkasi slaptažodžius ir patvirtino saugumo ekspertų didžiausius nuogąstavimus.

Alan'as Brown'as iš Southern Methodist University universiteto (JAV) bei dar trys jo kolegijos apklausė 218 pirmakursių studentų. Dauguma apklausos dalyvių nurodė, kad stengiasi įkvėpti save kuomet jiems tenka kurti slaptažodį, 45% respondentų nurodė, kad naudoja savo vardą bent vienam slaptažodžiui.

Saugumo ekspertai pasisako prieš vieno slaptažodžio naudojimą keliems daiktams, tačiau toks slaptažodžių dauginimas buvo dažnas. Ekspertai taip pat išpėja neužsirašinėti slaptažodžių, tačiau daugiau nei pusė prisipažino tai darantys.

Tyrimas atskleidė, kad vidutiniškas studentas turi 8,18 slaptažodžių vartojimo vietų. Tam naudojama vidutiniškai 4,45 slaptažodžio, vidutiniškai vienam slaptažodžiui tenka 1,84 tinklapio arba programos.

Atliktų tyrimų specifinės išvados:

- Du trečdaliai slaptažodžių yra sukurti naudojant asmenines naudotojo charakteristikas, dauguma jų sukurti remiantis giminaičių, draugų ar mylimųjų informacija.
- Beveik visi respondentai panaudoja slaptažodžius iš naujo, dviejų trečdalių respondentų slaptažodžiai yra dublikatai.
- Slaptažodžius pamiršta kas trečias respondentas, o daugiau kaip pusė turi juos kur nors užsirašę.

Tyrėjai sako, kad atrado empirinį kai kurių blogų slaptažodžių kūrimo praktikos pavyzdžių, aptariamų literatūroje, patvirtinimą. Jie pateikia patarimus slaptažodžių kūrimui ir naudojimui. Vienas iš jų yra skirstyti slaptažodžius į tuos, kurie turi būti aukšto saugumo lygio ir tuos kurie ne, o atsižvelgiant į tai naudoti skirtingas slaptažodžių kūrimo technikas. Patariama, kad sugalvojus slaptažodį ir patikrinus jo stiprumą, jei tikrinimo rezultatai tenkina, patariama pabandyti slaptažodį prisiminti po dienos ir tik tuomet, jei jis gerai prisimenamas patariama naudoti savo duomenų apsaugai (UsabilityNews, 2004).





1.4.2. Lietuvoje

Lietuvoje atlikto tyrimo apie slaptažodžius „Iš ko paprastai sudarote savo slaptažodžius?“ rezultatai buvo pateikti interneto svetainėje Banga.lt. (2002 09 30 duomenys)

Iš ko paprastai sudarote savo slaptažodžius?

Atsako: visi

(iš viso atsakė 188)

gimimo datos, telefono		10%	(19)
savo vardo, pavardės		13%	(25)
artimųjų, pažįstamų, gyvūnėlių vardų		27%	(50)
beprasmių simbolių rinkinių		50%	(94)





Šaltinis: (Verslo banga, 2003).

2 pav. Dažniausios slaptažodžių sudarymo technikos

Iš ko paprastai sudarote savo slaptažodžius?

Atsako: atsižvelgiantys į patogų klavišų išsidėstymą respondentai

(iš viso atsakė 66)

gimimo datos, telefono		17%	(11)
savo vardo, pavardės		12%	(8)
artimųjų, pažįstamų, gyvūnėlių vardų		14%	(9)
beprasmių simbolių rinkinių		58%	(38)





Šaltinis: (Verslo banga, 2003).

3 pav. Dažniausios slaptažodžių sudarymo technikos (atsižvelgiant į klavišų patogumą)

Iš ko paprastai sudarote savo slaptažodžius?

Atsako: neatsižvelgiantys į patogų klavišų išsidėstymą respondentai

(iš viso atsakė 122)

gimimo datos, telefono		7%	(8)
savo vardo, pavardės		14%	(17)
artimųjų, pažįstamų, gyvūnėlių vardų		34%	(41)
beprasmių simbolių rinkinių		46%	(56)

Šaltinis: (Verslo banga, 2003).

4 pav. Dažniausios slaptažodžių sudarymo technikos (neatsižvelgiant į klavišų patogumą)

Kito tyrimo rezultatai pateikiami interneto svetainėje Ebiz.lt:

„IBM Lietuva“ užsakymu atliktas „Vilmorus“ tyrimas parodė, kad dauguma Lietuvos gyventojų gali prisiminti 2-3 slaptažodžius, saugančius jų mokėjimo korteles, mobiliuosius telefonus, kompiuterius, el. pašta ir kitas taikomas programas.

Populiariausiomis duomenų ir informacijos apsaugos priemonėmis – slaptažodžiais – naudojasi net 95 proc. lietuvių.

Tiek jų turi bent po vieną slaptažodį, naudojamą banko kortelei, elektroniniam paštui ar mobiliojo telefono SIM kortelei. Vis dėlto daugiausia gyventojų naudojami 2-3 ar 4-5 slaptažodžiais. Vienu slaptažodžiu apsiriboja beveik 18 proc. lietuvių, o daugiau kaip penkis turi apie 17 proc.

„Skirtingų slaptažodžių naudojimas jungimuisi prie skirtingų sistemų padidina jų saugumą, tačiau dažnai sukelia galvos skausmus naudotojams, kai reikia juos visus prisiminti.

Kol biometrinių duomenų ar lustinių kortelių naudojimas naudotojo identifikavimui dar nėra plačiai paplitęs, programų sistemų kūrimo kompanijos siūlo įvairius sistemų naudotojų darbą palengvinančius sprendimus,“ - sako Tautvydas Cibas, „IBM Lietuva“ programinės įrangos vadovas (Ebiz.lt, 2006).

Kaip rodo Lietuvos naudotojų apklausa, panaši dalis vyrų ir moterų naudojami atitinkamu slaptažodžių skaičiumi. Daugiausia slaptažodžių turi 30-39 metų daugiau uždirbantys vilniečiai su aukštuoju išsilavinimu. Dauguma pensininkų naudojami 2-3 slaptažodžiai.

Tiek pat slaptažodžių turi ir kaimo gyventojai. 2-3 slaptažodžius atmintinai prisimena 40 proc. lietuvių. Be užrašų knygutės prisiminti 4-5 slaptažodžius gali ketvirtadalis apklaustųjų. Daugiau kaip penkis – tik 11 proc.

Tik vieną slaptažodį įsimena 24 proc. gyventojų. Ir vyrai, ir moterys, visų amžiaus grupių, socialinių sluoksnių, profesijų, bet kurio išsilavinimo, gyvenantys ir mieste, ir kaime daugiausiai prisimena tą patį kiekį – du arba tris naudojamus slaptažodžius.

„IBM Lietuva“ užsakymu visuomenės nuomonės ir rinkos tyrimų centras „Vilmorus“ reprezentatyvią Lietuvos gyventojų apklausą atliko balandžio 7-9 d (Ebiz.lt, 2006).

1.4.3. Autoriaus atliktas Lietuvos kompiuterių naudotojų slaptažodžių stiprumo galimybių ir elgsenos tyrimas I

2006 – 2007 metais tyrimo autorius atliko apklausą, siekdamas išsiaiškinti, kokiais metodais remiasi Lietuvos kompiuterių naudotojai kurdami savo slaptažodžius, kaip su jais elgiasi.

Apklausa buvo vykdoma tiesioginio interviavimo metodu apklausiant įvairaus amžiaus žmones jų studijų ir darbo vietose, bei netiesioginio interviavimo metodą, kuomet anketa buvo siunčiama atsitiktinai pasirinktiems žmonėms, prašant jų atsakyti į anketos klausimus ir atsiųsti juos atgal tyrimo autoriui. Apklausa atlikta, naudojant specialiai šiam tyrimui sukurtą anketą. Tyrimo metu apklausta virš 150 respondentų. Tyrimo metu naudota anketa pateikiama 1 priede. Atliekant tyrimą taip pat buvo siekiama patikrinti, ar pasikeitė kompiuterių naudotojų elgsena lyginant su

2003 metų portale Banga.lt apklausoje dalyvavusių kompiuterių naudotojų pateiktais atsakymais apie savo slaptažodžių sudarymo principus ir įpročius.

Atlikus tyrimą paaiškėjo, kad mažiau nei 4% visų apklaustųjų nesinaudoja slaptažodžiais internete, neturi jais apsaugotos informacijos. Likusioji dalis tyrime dalyvavusių žmonių paminėjo, kad yra nuo 4 iki 9 vietų (70%) kur jie naudojami slaptažodžiais, o 30% respondentų atsakė, kad naudojami iki slaptažodžiais iki 15 svetainių. Absoliuti dauguma naudotojų (net 98%) slaptažodžiais naudojami el. pašto sistemose. Toliau pagal populiarumą seka el. bankininkystės sistemos, darbo paieškos portalai, pramogų ir bendravimo portalai, interneto diskusijos bei pažinčių portalai. Patys nepopuliariausi naudotojų tarpe liko el. komercijos sistemos, jose registruotų naudotojų yra apytikriai 4% iš visų apklaustų asmenų. Didžiajai tyrime dalyvavusiųjų daliai (92%) jų duomenų, apsaugotų slaptažodžiais, saugumas yra svarbus arba šiek tiek svarbus, tačiau yra ir tokių, kurie juo visiškai nesirūpina.

Nors 86% tyrime dalyvavusių respondentų teigė, kad jiems yra svarbu, kad jų slaptažodžiai būtų saugūs, tačiau didelė dalis jų naudoja nesaugius slaptažodžius, teigia, kad nėra nieko girdėję apie saugaus slaptažodžio sudarymo principus (nors norėtų apie tai sužinoti). Tik 15% visų respondentų galvoja, kad slaptažodis bus saugus, jei nebus naudojamas asmens vardas, antrosios pusės vardas, šeimos narių vardai, augintinio vardas, automobilio, muzikos grupės pavadinimai, savo ar artimųjų žmonių gimimo datos, taip pat bet kokie prasme turintys žodžiai ar vien tik skaičių kombinacijos. Dažniausiai paminėti atsakymai buvo „Naudoti bet kokius kitus prasme turinčius žodžius“, naudoti savo šeimos narių vardus, taip pat atsakymai susiję su datų naudojimu. Reikia pastebėti, kad didžioji dauguma naudotojų naudoja trumpus slaptažodžius – slaptažodžių ilgis varijuoją nuo 3 iki 8 simbolių. Ilgesnius nei 9 simboliai slaptažodžius naudoja tik 3 – 4% respondentų.

Slaptažodžius Lietuvos naudotojai pagrindė saugo savo galvoje (90%), likusieji savo slaptažodžius saugo užrašę knygutėse ar mobiliuosiuose telefonuose. Galima tik pasidžiaugti, kad nė vienas apklausos dalyvis nepaminėjo, jog savo slaptažodžius saugo po pelės kilimėliu (kita vertus, galbūt tai lemia faktas, jog pelių kilimėliai po truputį dingsta nuo kompiuterių naudotojų stalų). Žymiai prastesnė situacija yra su slaptažodžio keitimu ir slaptažodžių perdavimu kitiems asmenims. Tik mažiau nei 6% dalyvių pažymėjo, kad savo slaptažodžius keičia dažniau nei kartą per tris mėnesius. Labiausiai Lietuvos naudotojai įpratę savo slaptažodį keisti kartą per metus arba netgi rečiau nei kartą per metus. Savo kompiuterio naudojimo praktikoje savo slaptažodžio kitiems yra neišdavę tik ne daugiau kaip 15% kompiuterių naudotojų, o net 70 procentų asmenų, dirbančių su kompiuteriais slaptažodžius kitiems asmenims pasakę nuo dviejų iki penkių kartų. Dar apie 10% respondentų nurodė, kad savo slaptažodį peravę kitiems daugiau nei 5 kartus. El. pašto dėžutės patikrinimą kaip pagrindinę slaptažodžio išdavimo priežastį nurodė 65%, dar 20% nurodė, kad

slaptažodžio prašė sistemos administratorius, o apytikriai 8% respondentų nurodė, kad jų slaptažodžiai buvo gauti ir panaudoti prieš juos neteisėtai, be pačių naudotojų žinios.

Apibendrinus slaptažodžių kūrimo principus, saugojimo principus ir žinias apie slaptažodžių saugumą apimančių klausimų atsakymus, galima teigti kad saugius slaptažodžius naudoja ir tinkamai juos saugo ne daugiau nei 2% Lietuvos kompiuterių naudotojų.

1.4.4. Autoriaus atliktas Lietuvos kompiuterių naudotojų slaptažodžių stiprumo galimybių ir elgsenos tyrimas II

Siekiant patikrinti pirmojo tyrimo išvadas, buvo gautas nedidelis realių duomenų kiekis iš vieno lietuviško puslapio, turinčio savo bendruomenę.

Tyrimo duomenys: 151 žmonių vardai, prisijungimo vardai ir jų slaptažodžiai (informacija pateikiama 2 priede, slaptažodžiai esamoje sistemoje saugomi koduotai, tačiau pasinaudojus dekodavimo algoritmu, buvo iš koduoti ir parengti naudojimui).

Tyrimo metu buvo siekiama nustatyti, kokius slaptažodžius Interneto naudotojai naudoja savo prisijungimams prie Interneto sistemų.

Turimų slaptažodžių lentelę pagal tai, kokio tipo slaptažodžius naudotojai naudoja prisijungimui prie sistemos, galima suskirstyti į kelias grupes. Suskaičiavus kiekvienai grupei priskirtų slaptažodžių skaičių, rezultatai pateikiami 2 lentelėje.

1 lentelė. 2-ojo slaptažodžių tyrimo duomenys sugrupuoti pagal slaptažodžių tipus

Eil. nr.	Grupės pavadinimas	Grupei priklausančių slaptažodžių skaičius	Procentinis dažnumas
1.	Žodis	83	55,3%
2.	Skaičiai	15	10,0%
3.	Vardas	14	9,3%
4.	Žodis + skaičiai	11	7,3%
5.	Kitas vardas	8	4,7%
6.	Žodžių junginys	6	4,0%
7.	Datos	6	4,0%
8.	Pavardė	2	1,3%
9.	Miestas	2	1,3%
10.	Žodžių junginys + skaičiai	1	0,7%
11.	Raidės + skaičiai	1	0,7%
12.	Raidės + spec. simbolis	1	0,7%

Šaltinis: sudaryta autoriaus

Galima teigti, jog šio paprasto tyrimo duomenys atskleidžia, kad naudotojai renkasi paprastus, prasmę turinčius žodžius, vardus arba skaičius, kurie jiems kažką reiškia, jog galėtų nesunkiai atsiminti slaptažodį.

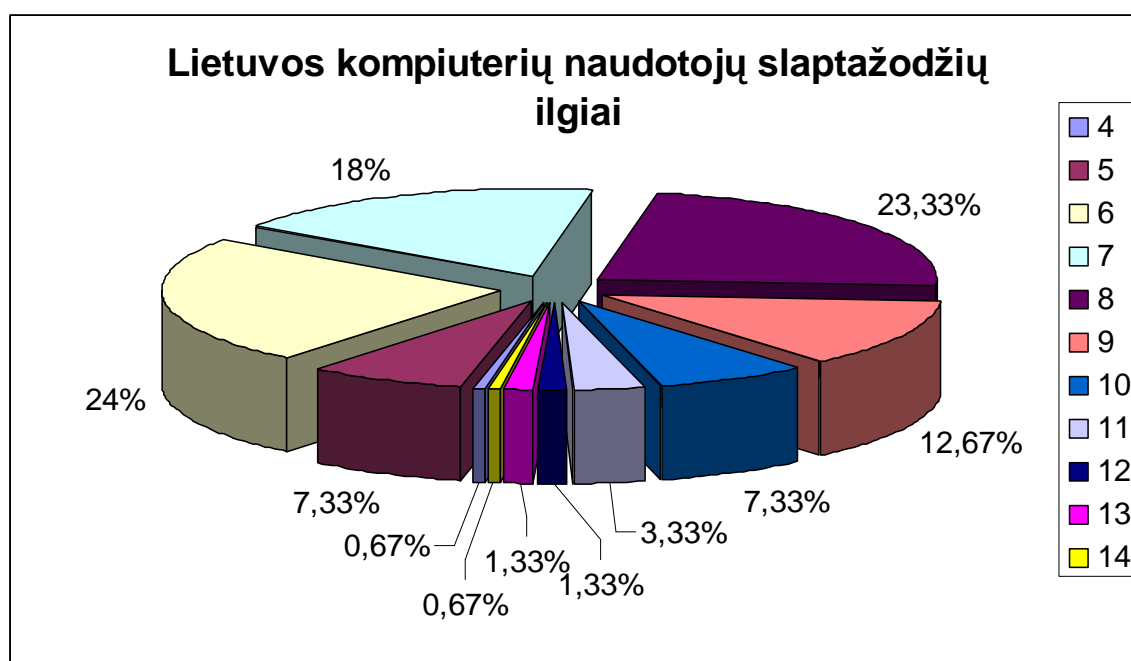
Toliau buvo suskaičiuoti visų slaptažodžių ilgiai, rezultatai sugrupuoti ir pateikiami apačioje esančioje 3 lentelėje.

2 lentelė. 2-ojo slaptažodžių tyrimo duomenys: slaptažodžių ilgiai

Slaptažodžio ilgis simboliais	Slaptažodžių skaičius	Procentinis slaptažodžių kiekis
4	1	0,67%
5	11	7,33%
6	36	24%
7	27	18%
8	35	23,33%
9	19	12,67%
10	11	7,33%
11	5	3,33%
12	2	1,33%
13	2	1,33%
14	1	0,67%

Šaltinis: sudaryta autoriaus.

Lentelės duomenys apipavidalinti grafiškai, jie pateikiami 5 paveiksle.



Šaltinis: sudaryta autoriaus.

5 pav. Lietuvos kompiuterių naudotojų sudaromų slaptažodžių ilgiai

Kaip matome iš lentelės, Lietuvos kompiuterių naudotojų tarpe vyrauja 6 – 8 simbolių ilgio slaptažodžiai, o slaptažodžiai iki 9 simbolių ilgio sudaro net 73,33% visų naudojamų slaptažodžių.

Taigi, antrasis tyrimas pavirtina pirmojo tyrimo rezultatus: žmonės kuria ir naudojami nesaugiais slaptažodžiais, kuriuos dažniausiai sudaro trumpi prasmingi žodžiai, skaičiai, asmens vardas arba žodis plius du skaičiai. Tokie slaptažodžiai yra itin nesaugūs, nekalbant jau apie naudojamo ryšio kanalo nesaugumą. Patikrinus slaptažodžius su slaptažodžių stiprumo matuokliais, nustatyta, kad tik 2% slaptažodžių nebuvo nesaugūs. Tai iš esmės patvirtina pirmojo tyrimo išvadas, jog tik 2% naudotojų kuria pakankamai saugius slaptažodžius ir moka juos saugoti. Taip pat

antrasis tyrimas patvirtino, kad slaptažodžiai nėra saugus būdas duomenims internete saugoti dėl pačių naudotojų elgesio su jais.

Tyrimo metu neatsižvelgta į tai, kaip naudotojas saugo savo slaptažodžius, ar tai yra jo vienintelis slaptažodis, ar jis turi ir daugiau slaptažodžių, ar šis slaptažodis naudojamas tik toje sistemoje, ar ir kitose, nežinomas naudotojų amžius (pagal svetainės turinį tikėtina, kad svetainės bendruomenė yra jaunimas nuo 18 iki 25 metų), nežinomas jų išsilavinimas ir kiti kriterijai, galintys įtakoti tyrimo išvadas.

1.4.5. Slaptažodžių tyrimų išvados

Taigi, atsižvelgiant į abu autoriaus atliktus tyrimus, galima daryti tokias darbo tiriamosios dalies išvadas:

1. Slaptažodžių naudojimas yra labiausiai paplitęs, tačiau nesaugus autentifikacijos būdas. Nesaugumo lygį dar labiau padidina pačių naudotojų elgesys kuriant ir saugant savo slaptažodžius.
2. Slaptažodžius lengva pamiršti, arba juos nesunku gauti ir panaudoti be naudotojo sutikimo ar žinios.
3. Nors dalis naudotojų ir žino, kaip reikia teisingai kurti slaptažodžius arba juos saugoti, šių taisyklių yra nesilaikoma: naudotojai vis tiek kuria slaptažodžius, kurie:
 - Yra lengvai įsimenami, nes
 - naudoja asmeninę naudotojo arba su naudotoju susijusią informaciją, todėl
 - yra nesunkiai įveikiami net ir menkai patyrusiam kompiuteriniam nusikaltėliui.
4. Didelė dalis Lietuvos kompiuterių naudotojų visiškai neskiria dėmesio savo duomenų saugumui.

Tiriant nebuvo atsižvelgta į ryšio kanalo saugumą. Tikėtina, kad didžioji dalis kompiuterių naudotojų visiškai nekreipia dėmesio į tai, kad duomenys iš jų kompiuterio į internetą keliauja neapsaugotu kanalu, todėl, pasinaudojant specialia programine įranga, jie yra nesunkiai perimami piktavalių.

Be to, kaip pateikia interneto svetainė Lockdown.co.uk, greičiausias pasaulio kompiuteris šiuo metu perrenka 76,1 milijardo slaptažodžių per sekundę (2007 01 22 dienos informacija) (Kompiuterija, 2007), (Lockdown.co.uk, 2007), o Blue Gene superkompiuteris 10 simbolių dydžio raktą nulaužtų per sekundės dalių (Saswato, 2007).

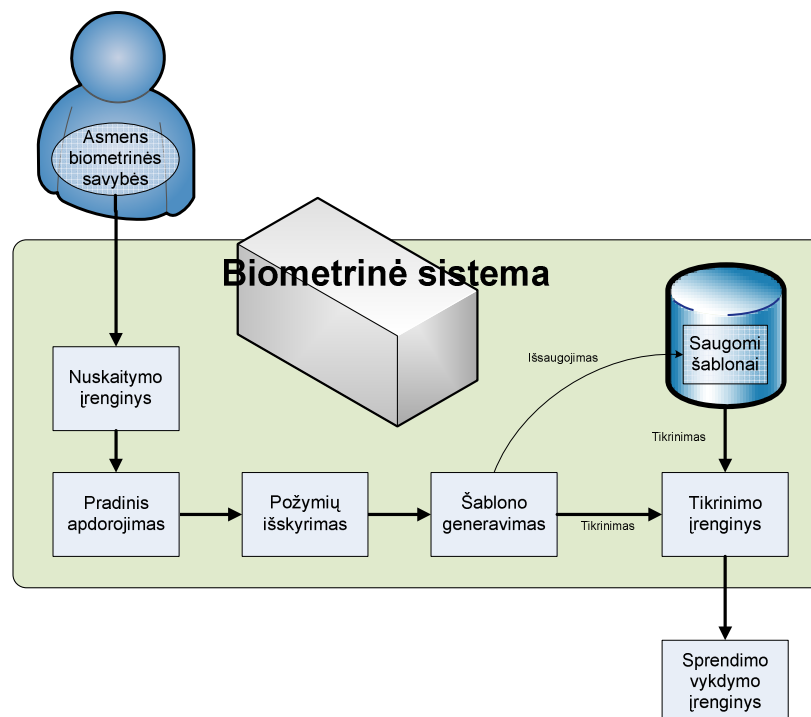
1.5. Biometrinės autentifikavimo priemonės

Atsižvelgiant į autoriaus atliktų tyrimų rezultatus ir išvadas, darosi akivaizdu, kad interneto pasaulyje duomenims apsaugoti neužtenka populiariausios šių dienų autentifikavimo priemonės, tačiau reikia ieškoti naujo mechanizmo duomenims saugoti. Viena sričių, kurioje jau dirba daugelio šalių mokslininkai ir kurioje vis dar galima pasiekti laukiamų gerų rezultatų yra biometrija.

Biometrija (angl. *Biometrics*) – biologinės informacijos matavimo ir analizavimo mokslas ir technologija. Informacinėse technologijose biometrija siejama su technologijomis, autentifikavimo tikslais skirtomis analizuoti žmogaus kūno charakteristikas, tokias kaip pirštų atspaudai, akies rainelė, balsas, veidas (Search Security, 2006).

Pastaruoju metu autentifikavimas naudojant biometrinių žmogaus savybių tikrinimą tampa vis labiau populiarus korporacinėse ir viešose saugumo sistemose, naudotojo elektroninėse ir prekybinėse (angl. *point of sale, POS*) aplikacijose. Atsižvelgiant į technologijų kaitos greitį ir poreikį, dabartinių priemonių nepakankamumą, biometrinis autentifikavimas turi geras perspektyvas ateityje.

Kai kurios biometrijos priemonės naudojamos jau daugiau kaip šimtą metų – pirštų atspaudai buvo pirmoji biometrinė priemonė, skirta nusikaltusių asmenų identifikavimui. Yra faktų, kad net jau pačioje pradžioje, kol pirštų atspaudai tarpusavyje buvo lyginami rankiniu būdu, pirštų atspaudų tyrimai yra padėję sugauti kelis iki tol ilgai laisvėje vaikščiojusius nusikaltėlius.



Šaltinis: sudaryta autoriaus pagal (Wikipedia, 2007 c).

6 pav. Biometrijos sistemų veikimo principas ir sandara

Kaip matome 6 paveiksle, biometrinių savybių atpažinimu grįstų priemonių sandara bendruoju atveju susideda iš:

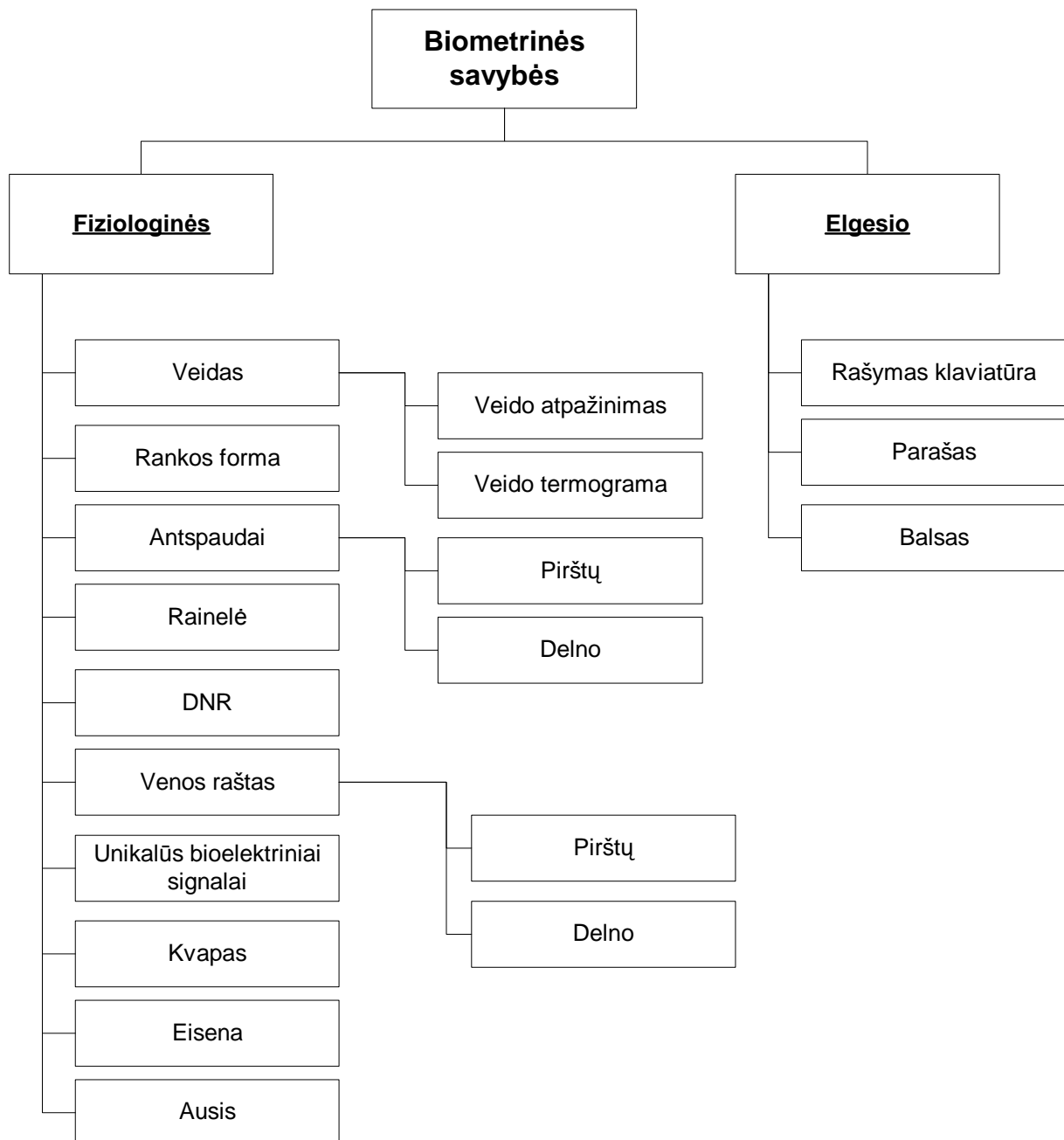
- Nuskaitymo ar nuskenavimo įrenginio;
- Informacijos konvertavimą iš nuskaitytos į skaitmeninę bei informacijos pirminį apdorojimą atliekančios programinės įrangos.
- Informacijos patikrinimą atliekančios programinės įrangos.
- Duomenų bazės, kurioje saugoma biometriniame atpažinime naudojama informacija.

Reikia pastebėti, kad norint apsisaugoti nuo identiteto vagystės, biometrinė informacija dažnai užkoduojama iš karto po to, kai ji buvo nuskaityta. Biometrinis verifikavimas veikia tokiu principu: siekiant konvertuoti biometrinius įvesties duomenis programinė įranga panaudojama duomenų palyginimo specifiniams taškams nustatyti. Duomenų bazėje saugomi palyginimo specifinių taškų informacija yra paverčiama į skaitmeninę išraišką ir tuomet duomenų bazėje saugotos biometrinių savybių palyginimo taškų reikšmės yra sulyginamos su tomis, kurias nuskaite įrenginys kai asmuo pateikė užklausimą tapatybei nustatyti. Autentiškumas yra patvirtinamas arba paneigiamas priklausomai nuo gaunamo sulyginimo atsakymo.

Visos biometrijos priemonės skirstomos į dvi pagrindines grupes:

- Fiziologinės biometrinės priemonės;
- Elgesio biometrinės priemonės.

Biologinių priemonių klasifikacija pateikiama 6 paveiksle.



Šaltinis: sudaryta autoriaus.

7 pav. Žmogaus biometrinių savybių klasifikacija

A. K. Jain 2004 metais savo darbuose “Biometric recognition: how do I know who you are?” bei “Biometric Recognition: Security and Privacy Concerns“ atliko biometrinių faktorių palyginimą, kuris pateikiamas 5 lentelėje (H=Aukštas, M=Vidutinis, L=Žemas).

3 lentelė. „Įvairių biometrinių technologijų palyginimas pagal A. K. Jain“

Biometrika	Universa-lumas	Unika-lumas	Pasto-vumas	Surenka-mumas	Atlikimas	Priima-mumas	Apeina-mumas
Veidas	H	L	M	H	L	H	H
Veido termograma	H	H	L	H	M	H	L
Pirštų ansapudai	M	H	H	M	H	M	M

Biometrika	Universa- lumas	Unika- lumas	Pasto- vumas	Surenka- mumas	Atlikimas	Priima- mumas	Apeina- mumas
Rankos geometrija	M	M	M	H	M	M	M
Rankos venos	M	M	M	M	M	M	L
Delno anspaudai	M	H	H	M	H	M	M
Rašymas klaviatūra	L	L	L	M	L	M	M
Parašas	L	L	L	H	L	H	H
Rainelė	H	H	H	M	H	L	L
Tinklainės skenavimas	H	H	M	L	H	L	L
Balsas	M	L	L	M	L	H	H
Kvapas	H	H	H	L	L	M	L
DNR	H	H	H	L	H	L	L
Eisena	M	L	L	H	L	H	M
Ausis	M	M	H	M	M	H	M

Šaltinis: sudaryta autoriaus pagal (Jain, 2004 a) ir (Jain, 2004 b).

Pavertus 3 lentelės duomenis skaitinėmis reikšmėmis ($H = 3$, $M = 2$, $L = 1$), atliekamas biometrinių savybių palyginimas pagal atskirus kriterijus. Palyginimo rezultatai pateikiami 4 lentelėje.

4 lentelė. Biometrinių savybių palyginimas

Eil. Nr.	Lyginimo kriterijus					
	Unikalumas	Surenkamumas	Apeinamumas	Vidurkis 1	Vidurkis 2	Vidurkis 3
1	Veido termograma	Rankos geometrija	Veido termograma	Veido termograma	Pirštų anspaudai	Pirštų anspaudai
2	Rainelė	Veidas	Rainelė	Pirštų anspaudai	Delno anspaudai	Delno anspaudai
3	Tinklainės skenavimas	Veido termograma	Rankos venos	Delno anspaudai	Veido termograma	Veido termograma
4	Kvapaspas	Eisena	Kvapaspas	Rankos geometrija	Rainelė	Veidas
5	DNR	Parašas	DNR	Veidas	Ausis	Rainelė
6	Pirštų anspaudai	Rainelė	Tinklainės skenavimas	Parašas	Veidas	Ausis
7	Delno anspaudai	Pirštų anspaudai	Rankos geometrija	Rainelė	DNR	Rankos geometrija
8	Rankos venos	Delno anspaudai	Eisena	Ausis	Rankos geometrija	DNR
9	Rankos geometrija	Ausis	Pirštų anspaudai	Eisena	Tinklainės skenavimas	Parašas
10	Ausis	Rankos venos	Delno anspaudai	Balsas	Kvapaspas	Tinklainės skenavimas
11	Eisena	Rašymas klaviatūra	Ausis	Tinklainės skenavimas	Rankos venos	Kvapaspas
12	Rašymas klaviatūra	Balsas	Rašymas klaviatūra	Kvapaspas	Eisena	Eisena
13	Veidas	Kvapaspas	Veidas	DNR	Parašas	Balsas
14	Parašas	DNR	Parašas	Rankos venos	Balsas	Rankos venos
15	Balsas	Tinklainės skenavimas	Balsas	Rašymas klaviatūra	Rašymas klaviatūra	Rašymas klaviatūra

Šaltinis: sudaryta autoriaus

Palyginimas pagal unikalumą, geriausiomis biometrinėmis savybėmis galima laikyti veido termograma, akies rainelę, pirštų anspaudus, delno anspaudus, tinklainės skenavimą, kvapą ir DNR, kuris apskritai kiekvieno žmogaus yra unikalus. Kadangi vertinti vien unikalumą bendru atveju neužtenka, nes skirtingos biometrinės savybės skiriasi ir savo surenkamumo galimybėmis, išrikiavus pagal surenkamumą, pastebėtume, jog geriausios biometrinės savybės yra veido termograma, rankos geometrija, eiseną, veidas ir parašas, parašas. Surenkant šias savybes reikia paprasčiausių technologinių sprendimų. Sunkiausiai surenkami tinklainės skenuoti pavyzdžiai, kvapo pavyzdžiai ir DNR – jiems reikia sudėtingų technologijų. Pagal apeinamumą geriausios technologijos yra veido termograma, rainelės skenavimas, rankos venų skenavimas, tinklainės skenavimas, kvapo atpažinimas ir DNR atpažinimas. Lyginant biometrines savybes pagal unikalumą, surenkamumą ir apeinamumą (skaičiuojant tik šių trijų savybių vidurkį, lentelėje – stulpelis „Vidurkis1“), gauname, jog veido termograma, pirštų anspaudai, delno anspaudai rankos geometrija yra geriausios savybės. Skaičiuojant visų savybių bendrą įvertinimą (savybių skaitinių reikšmių vidurkį, lentelėje stulpelis „Vidurkis2“), arba svorinį savybių įvertinimą (unikalumui, surenkamumui ir apeinamumui suteikiami svoriai „2“, kitoms savybėms – „1“) matome, kad daugiausiai perspektyvų turi pirštų anspaudai, delno anspaudai, veido termograma, veidas ir akies rainelė, nepaisant kai kurių iš šių savybių surinkimo sudėtingumo.

Apibendrinant biometrinių savybių pagal įvairius kriterijus palyginimą, galima teigti, jog prie geriausiai biometriniam atpažinimui tinkamų priemonių yra delno anspaudai, pirštų anspaudai, rainelė, bei veido termograma, o prie prasčiausiai – balsas, parašas, rankos venos ir rašymas klaviatūra (žr. 3 priedą.)

1.6. Skirtingų autentifikavimo priemonių palyginimas

Atsižvelgiant į prieigos valdymo sprendimų palyginimo lentelės duomenis (2006 m. duomenys) (SearchSecurity.com, 2007), dviejų faktorių autentifikacija galbūt būtų tinkamas sprendimas aukštesniam duomenų saugumo lygiui užtikrinti: naudojant dviejų faktorių, kurių vienas būtų biometrinių duomenų atpažinimas, rizika yra žema, vadinasi tai yra pakankamai saugu, be to, palyginus su tomis autentifikavimo priemonėmis, priskiriamoms grupei „tai ką turime“, tikimybė pamesti biometrinius duomenis yra artima nuliui, be to, biometriniai duomenys daugiau ar mažiau yra unikalūs.

5 lentelė. „Prieigos valdymo sprendimų palyginimas“

Prieigos valdymo sprendimas	Rizikos	Privalumai	Trūkumai
Naudotojo identifikaciniai numeriai ir slaptažodžiai	Jeigu neteisigai valdomis ar saugomi, naudotojo identifikaciniai numeriai ir slaptažodžiai gali būti nesunkiai pavogti ir taip užtikrintas priėjimas prie prie naudotojo tinklo, sistem ar duomenų. Rizikos lygis: AUKŠTAS	<ul style="list-style-type: none"> Nesunku įgyvendinti ir dažnai naudojama tinkle ir sistemų prieigai kontroliuoti. Su slaptažodžiais ir naudotojo identifikaciniais numeriais naudotojai susipažinę geriau nei su bet kuriomis kitomis autentifikavimo sistemomis ar priemonėmis. 	<ul style="list-style-type: none"> Slaptažodžiai gali būti atspėti, jeigu jie parenkami dažnai vartojamų žodžių, vardu ar kitos asmeninės informacijos pagrindu. Naudotojų ID numeriai ir slaptažodžiai nesunkiai gali būti pavogti naudojant laisvai Internetu platinamas išsilaužimo programas, naudojant Trojos arklio tipo programėles ar naudotojo veiksmus fiksuojančią programinę įrangą.
Biometrinės savybės	Pirštų atspaudų atveju yra tikimybė nukopijuoti naudotojo piršto atspaudus. Taip pat yra tikimybė, kad kad biometrinės priemonės bus nuskaitytos ir panaudotos iš skaitmeninės biometrijos savybių saugyklos. Rizikos lygis: ŽEMAS	<ul style="list-style-type: none"> Tai yra viena iš stipriausių prieigos kontrolės technologijų – yra beveik neįmanoma pavogti kieno nors ragenos raštą, veidą ar pirštų atspaudus. Geriausiai tinka kaip antrasis faktorius dviejų faktorių autentifikavimo sistemoje šalia naudojamo naudotojo ID/slaptažodžio ar sumaniųjų kortelių sistemos. Plačiausiai naudojamas fizinės prieigos prie sistemos užtikrinimui, tačiau auga panaudojimų skaičius, kur ši technologija naudojama kaip atskira sistema interneto prieigos užtikrinimui. 	<ul style="list-style-type: none"> Reikalauja didelių aparatūrinių priemonės šdiegimo kaštų. Ši technologija nėra 100% patikima – nemažas neteisingų atpažinimų kiekis.
Dviejų faktorių autentifikavimas	Egzistuoja nedidelė tikimybė, kad abu autentifikavimo metodai bus nulaužti iš karto vienas po kito. Rizikos lygis: ŽEMAS	<ul style="list-style-type: none"> Užtikrinamas papildomas apsaugos sluoksnis. Pavyzdžiui, naudojami naudotojo ID ir slaptažodis, antrasis faktorius – OTP raktas (OTP token). Jei vienas yra pralaužiamas, antrasis vis dar užtikrina duomenų apsaugą. 	<ul style="list-style-type: none"> Reikalauja papildomų programinės ir/ar aparatūrinių priemonės įdiegimo, siekiant priversti dvi skirtingas autentifikavimo sistemas dirbti kartu.

Šaltinis: sudaryta autoriaus pagal (SearchSecurity.com, 2007).

Šie apibendrinimai įrodo, kad tai, kas dabar yra plačiausiai naudojama savo duomenims užtikrinti, nėra vertinama kaip saugus būdas – naudojant slaptažodžius smarkiai rizikuojama informacijos saugumu. Siekiant rizikos minimizavimo, įprastines saugumo sistemas reikia keisti į biometrines savybes tikrinančias priemones arba dviejų faktorių autentifikavimo priemones, kuriose viena dedamoji galėtų būti tradicinė autentifikavimo priemonė (iš „tai ką turime“ ar „tai ką žinome“ autentifikavimo priemonių grupės), o antrąją dalimi gali būti biometrinių savybių autentifikavimo priemonė. Vertinant šiuos apibendrinimus, dar reikėtų atsižvelgti į tai, kad biometriniai duomenys, ypač tokie kaip pirštų antspaudai ar balsas, negali būti pirminiai autentifikatoriai (angl. *primary authenticators*). Biometrinės savybės nėra slaptos – žmonės

kasdien viešai demonstruoja savo pirštų antspaudus liesdami daiktus, balsą – bendraudami su kitais žmonėmis, veidą – rodydamiesi viešumoje, atlikdami kasdienę savo veiklą. Taip sukuriamos galimybės biometrinių duomenų vagystei. Galimos įvairios papildomos priemonės šiai problemai sumažinti, bet teoriškai problema yra neišsprendžiama. Kita vertus, nors žmogus negali pamiršti savo biometrinių duomenų (nes biometriniai duomenys priskiriami autentifikavimo faktorių grupei, „tai kas esu“, o ne „tai, ką žinau“), tačiau įvairūs organizmo ar fiziniai pokyčiai – traumas arba ligos – gali trumpam sutrikdyti biometrinėmis apsaugos priemonėmis saugomų duomenų pasiekiamumą: suskeldėję pirštai ar po laringito pasikeitęs balsas greičiausiai bus neatpažintas sistemos. Todėl atsižvelgiant į anksčiau paminėtus biometrinių savybių trūkumus, naujas sprendimas yra dviejų faktorių autentifikavimas.

1.7. Diktorius atpažinimas

Nepaisant to, kad vertinant kitų biometrinių savybių fone balso technologijos nėra laikomos pačia geriausia technologija atpažinimui, tačiau praktiniu požiūriu tokios sistemos realizavimas nereikalauja sudėtingų technologijų panaudojimo, be to, kalbos ir diktorius atpažinimas yra svarbus, siekiant užtikrinti natūralią žmogaus ir sumaniųjų sistemų komunikaciją.

1.7.1. Tyrimai diktorius atpažinimo srityje

Tyrimai diktorius atpažinimo srityje atliekami jau gana seniai. Vieni iš aktyviausių mokslininkų šioje srityje yra Sadaoki Furui (Japonija, Tokyo universitetas, NTT Human Laboratories) ir Tomoko Matsui. Sritis išsirutuliojo iš automatinės kalbos atpažinimo srities. 1990 metais S. Furui savo straipsniu „Nuo diktorius priklausomų savybių išskyrimas, atpažinimas ir apdorojimo technikos“ tarsi padėjo pirmąjį akmenį diktorius atpažinimo sričiai. 1993 – 1994 S. Furui ir T. Matsui paskelbė eilę straipsnių apie savo tyrimus diktorius atpažinimo srityje.

Matsui ir Furui 1993 metais tyrinėjo nustatė, kad taikant priklausomo nuo teksto diktorius atpažinimo metodą, kuriame slaptažodinius teksto raktus pateikia atpažinimo priemonės, yra reikalingi diktoriui būdingų fonemų modeliai. Jie reikalingi tam, kad būtų atpažįstamas teksto raktas (ženklas) ir pats diktorius, todėl tyrinėti du diktoriui būdingų fonemų nustatymo modeliai: fonemos prisitaikymo prie nepriklausomo nuo fonemos diktorius modelis ir diktorius prisitaikymą prie universalių fonemos modelių. Autoriai aptardami šiuos metodus taip pat įtraukia nepriklausomą nuo fonemos modelį, kad užpildytų žinių trūkumą apie diktorių. Eksperimentų rezultatai rodo, kad yra atmetama 98,5 procento neteisingai pasakyto teksto, tačiau diktorius atpažįstami 100% (Matsui, Furui, 1993).

1994 m. S. Furui 1993 m. straipsnio pagrindu straipsnyje „Diktorius atpažinimo technologijos apžvalga“ apžvelgė šios technologijos klasifikaciją, būtent išskyrė diktorius

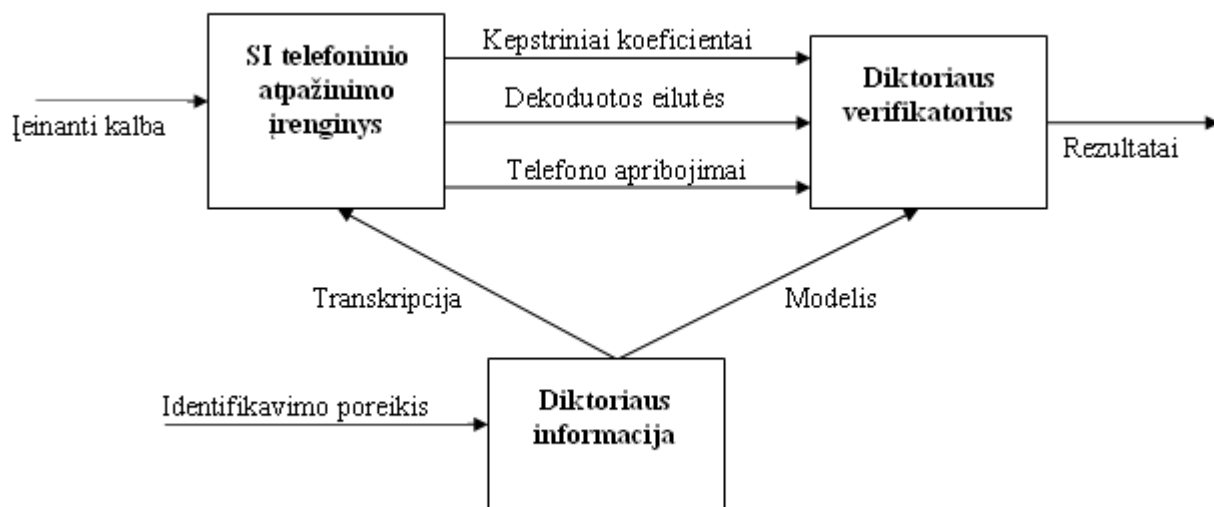
autentifikavimo srities sprendžiamus uždavinius į diktoriaus identifikavimą ir diktoriaus verifikavimą, aprašė skirtumus tarp jų. Taip pat šiame darbe jis nurodė, kad diktoriaus atpažinimas gali būti skirstomas į priklausomą nuo teksto ir nepriklausomą nuo teksto. Taip pat šiame straipsnyje aprašytos pagrindinės diktoriaus atpažinimo sistemų struktūros, toliau aptariamos darbo 1.7.2 skyriuje. Be viso to, Furui šiame straipsnyje apžvelgė įvairius metodus, taikomus diktoriaus atpažinimo srities uždaviniams spręsti. Pagrindiniai metodai, apžvelgti šioje apžvalgoje toliau analizuojami 1.7.3 skyriuje.

1994 m. C. Griffin kartu su Furui ir Matsui paskelbė straipsnį apie savo eksperimentą, kurio metu analizavo multivariacinį autoregresinį modelį (MAR). Autoriai tyrimo metu įvertina dvi atstumo matavimo priemones: prisijungimo tikimybės koeficientą (LLR) ir simetriško išdėstymo tikimybės santykį perdavimo greitį (SLR), kuris yra simetrinė tikimybės koeficiento matavimo priemonė. Eksperimentų rezultatai rodo, kad LLR dirba geriau, nei SLR matavimo priemonė, kai tyrime naudojama daugiau nei 5 sakiniai. Kuomet apmokymui naudojami 10 sakinių, identifikavimo ir verifikavimo santykis gaunamas beveik toks pat, kaip naudojant HMM paremtus metodus. Autorių prieita prie išvados, kad SLR geriau tinka naudoti matavimams esant trumpesniai vykdymo periodui arba mažesniai duomenų kiekiui (Griffin et al, 1994).

J.-L. Le Floch, C. Montaci ir M.-J. Caraty 1995 metais pristatė savo eksperimentinius tyrimus remiantis NTIMIT garsynu. Darbe palyginami trijų diktoriaus atpažinimo sistemų gauti rezultatai (GMM, AHSM, ARVM), kurie naudojo TIMIT ir NTIMIT garsynų duomenimis. Buvo pristatytos dvi naujos, sudėtingos sistemos, skirtos patobulinti rezultatus, kuomet tyrimai atliekami NTIMIT duomenų bazės pagrindu. Pirmoji sistema sukurta remiantis ARMAVector modeliu, o antra keletu AR-Vector modelių panaudojimu vienam diktoriui (Flock, 1995).

Priklausomo nuo teksto diktoriaus verifikavimo sistemos, naudojančios bendruosius frazinius slaptažodžius, tyrimus 1996 metais atliko S. Parthasarathy ir A. E. Rosenberg. Iškelta hipotezė, kad galima sumodeliuoti teksto slaptažodžių (raktinių žodžių) dikciją ir jų fonemos transkripciją. Tyrimo metu paaiškėjo, kad iškyla problemų dėl tinkamo vienetų pasirinkimo, norint sudaryti planuojamą diktoriaus modelį, ir dėl bazinių modelių pasirinkimo tikimybės koeficiento apskaičiavimui. Tyrimų metu jie sukuria diktoriaus verifikavimo sistemą, kuri naudojami anglų kalbos fraziniais slaptažodžiais. Sistemos pagrindu tapo anksčiau sukurtas skaitmenų eilučių slaptažodžių sistema. Buvo įrodyta, kad naudojantis žodžio ar frazės modeliais yra gaunams daug geresnis rezultatas, nei naudojantis telefoninio duomenų taikymo modeliais. Buvo įrodyta, kad daug pastangų reikalaujantis (angl. robust) rezultatų sumavimo metodas, kuris naudojami freimų išskyrimu remiantis fonemų energijagai būti naudingas. Tikslinių frazių modelio ir sparčiojo skaičiavimo metodų suliejimas sugebėjo pasiekti 33% mažesnę klaidų skaičių. Buvo pateikti

preliminarūs įrodymai norint įrodyti, kad yra naudinga įtraukti laikiną struktūrą į bazinius modelius. Sistemos veikimo principas pateikiamas 8 paveiksle (Parthasarathy, Rosenberg, 1996).



Šaltinis: (Parthasarathy, Rosenberg, 1996).

8 pav. S. Parthasarathy ir A. E. Rosenberg pasiūlytas sistemos veikimo principas

Tais pačiais metais šie du mokslininkai patvirtina, kad tikimybės koeficientas arba kohortos norminis sumavimas, yra efektyvios priemonės diktoriaus atpažinimo sistemų tobulinimui. Pradinių duomenų įvedimas laikomas viena svarbiausių problemų galinčių užkirsti kelią tam, kad būtų sukurta diktoriaus bazinė sistema ar kohortos modeliai, kurie pateikia labiausiai matomus norminius rezultatus. Dėl šios priežasties studijuojami keletas rūšių diktoriaus bazinių sistemų modeliai: individualūs diktoriaus modeliai, modeliai paremti daugelio diktorių vientisa dikcija, modeliai parinkti panašumo į tipinius modelius pagrindu, atsitiktinės diktorių atrankos būdu sudaryti modeliai, atsitiktinai parinktų diktorių modeliai, ir kitomis (skirtingomis) sąlygomis, nei tipiniai modeliai, kurti duomenų bazės modeliai. Eksperimentų rezultatai rodo, kad vientisi modeliai, paremti panašumu į nuorodos diktorių, veikia geriau, nei individualūs kohortos modeliai, kurie naudojami ta pačia diktorių grupe. Vientisi baziniai modeliai, kurie naudojami maža diktorių grupe, paremta panašumu, pasiekia geriausius rezultatus, bet ne tokius išskirtinius, kaip naudojantis atsitiktinio parinkimo metodu, kurio metu atrankama daugiau nei 40 subalansuotų pagal lytį diktorių su panašiomis rengimo sąlygomis, pritaikytomis etaloniniui diktoriui (Rosenberg, Parthasarathy, 1996).

Dar vienas tyrimas atlikas 1996 metais buvo Mammone, R. J. Xiaoyu Zhang ir R. P. Ramachandran buvo susijęs su energinga diktoriaus atpažinimo sistema paremta savybių atpažinimo metodu. Energinga diktoriaus atpažinimo sistema paremta savybių atpažinimo metodu. Autoriai pateikia dabartinio diktoriaus sistemų ir savo užfiksuotų problemų apžvalgą. Jie taip pat sutelkia dėmesį į pradinių bruožų surinkimo procesą, skirtą daug energingiems kalbos metodams ir jų tobulinimui. Pirmas žingsnis į savybių išskyrimą yra linijinė nuspėjamoji analizė (angl. Linear

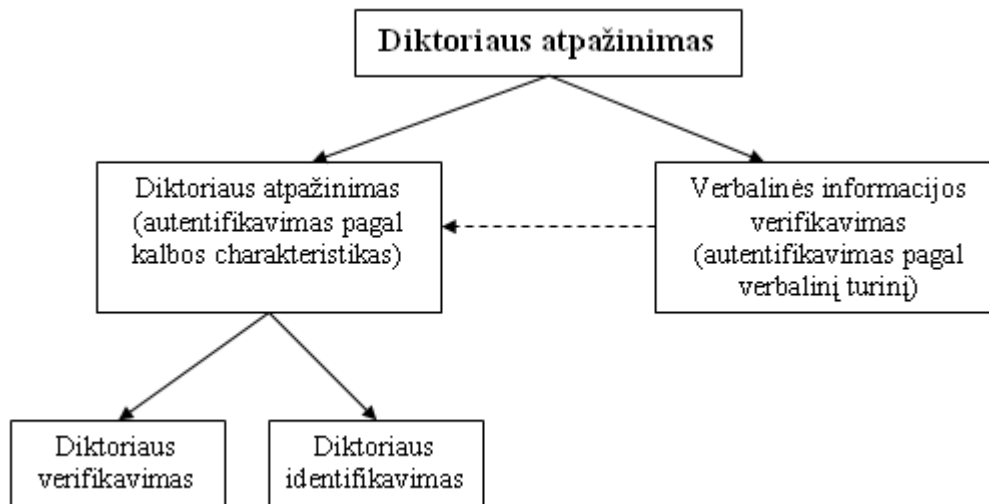
predictive analysis, LP). Tyrimų metu taip pat apžvelgiamos įvairaus sudėtingumo kepsrinės savybės, įgytos LP koeficientų nustatymo metu (Mammone et al, 1996).

1997 metais A. M. Ariyaeinia, P. Sivakumaran (University of Hertfordshire, Hatfield, UK) Šie autoriai pristato tyrimą apie įvairius galimus rezultatų norminimo metodus, skirtus diktorius verifikavimui. Nuodugni analizė apima įvairius metodus, skirtus sunorminti verifikavimo rezultatus, ir taip pat nagrinėja juos pagal identišką eksperimento sąlygas. Eksperimentai yra paremti Brento poaibio (telefono kokybės) kalbos duomenų baze, sudaryta iš pasikartojančio izoliuoto skaitmens posakių (nuo 0 iki 9), pasakytų anglakalbių diktorių. Aptikta, kad lyginant su kitais aptartais metodais, konkretus kohortinio norminimo formos metodas funkcionuoja geriausiai dėl verifikacijos tikslumo (Ariyaeinia, Sivakumaran 1997).

Suzuki Motoyuki, Abe Toshiaki ir kiti 1999 metais savo straipsnyje „Speaker Adaptation Using Phoneme-Dependent Tree-Structured Speaker Clustering“ aprašo dar vieną metodą, kuris naudoja diktorius klasterizavimą medžio struktūra, kaip vieną iš galimų diktorius pritaikymo technikų. Šiame metode, įvairių klasių modeliai, pradedant neapibrėžtais baigiant apibrėžtais diktorius, buvo kvalifikuojami į struktūras. Be to, buvo nustatyta kad yra galimas didelio greičio modelio paieškos atlikimas ir jis tinkamas diktorius adaptacijai. Taigi, yra nagrinėjama sudėtingesnė diktorius klasifikavimo technika, naudojanti priklausomo nuo fonemos diktorių klasifikavimą medžio struktūra. Fonemos atpažinimo eksperimentų metu, atpažinimo dažniai pralenkė medžio struktūras, būdingas fonemai. Be viso to, buvo pasiūlytas metodas, naudojantis panašią medžio fonemos struktūrą į tas, kurios nėra neįtraukiamos į adaptacines kalbas (Motoyuki ir kt., 1999).

Qi Li; Bing-Hwang Juang; Chin-Hui Lee 2000 metais pristatė tyrimus verbalinės informacijos verifikavimo srityje. Tradicinė diktorius autentifikacija fokusuoja dėmesį į diktorius verifikavimą (DV) ir diktorius atpažinimą. Pastarasis yra užbaigiamas suderinant diktorius balsą su jo ar jos registruotais kalbos modeliais. Šiame tyrime siūloma nauja technika, pavadinta verbalinės informacijos verifikavimu (VIV). Šio metodo esmė yra patikrinti kalbančiojo dikciją ir automatiškai palyginti su raktine (įprastai slapta) informacija registruoto diktorius profilyje bei nuspręsti, ar šį diktorių priimti ar atmesti. Naudojantis pateiktos tvarkos procedūra buvo įtraukti trys klausimai su atsakymais. Rezultatas be jokių klaidų buvo pasiektas naudojantis telefoninio diktorius autentifikavimo eksperimentu su 100 diktorių. Be to, VIV diktorius autentifikavimo sistema buvo sujungta su DV. Šioje sistemoje, vartotojas patvirtinamas VIV sistemos per pirmus keturis ar penkis priėjimus, dažniausiai skirtingose garso aplinkose. Šiuose panaudojimuose, vienas esminių klausimų siejamas su slaptažodinės frazės dikcija. VIV sistema savo ruožtu surenka ir patvirtina šias dikcijas naudojimui. Jo metu atlikimo duomenys naudojami diktorius modelio formavimui. Kai priklausomas nuo diktorius modelis yra sukuriamas, sistema pereina į SV. Šis

metodas vengia formalios registracijos precedūros nepatogumų, užtikrina DV apmokymo duomenų kokybę ir sumažina neatitikimus, kuriuos sukelia skirtingi balso duomenys, apmokymo ir testavimo metu. Eksperimentai parodė, kad pasiūlyta sistema patobulino diktoriaus verifikavimo procesų vykdymą daugiau 40% vienodų klaidų dažnyje, lyginant su tradicine diktoriaus verifikavimo sistema (Qi Li ir kt., 2000).



Šaltinis: (Qi Li ir kt., 2000).

9 pav. Verbalinės informacijos verifikavimo schema

Ismail Shahin 2004 – 2005 metais tyrinėjo ciklišką paslėptųjų Markovo modelių (CHMMs) metodą, kuris buvo naudojamas natūraliomis kalbėjimo sąlygomis, siekiant patobulinti priklausomo nuo teksto izoliuotus žodžius tariančio diktoriaus atpažinimo procesą. Rezultatai rodo, kad CHMMs apima diktoriaus autentifikavimo atlikimo procesą tokiomis sąlygomis, lyginant su iš kairės į dešinę paslėptųjų Markovo modelių metodais (LTRHMMs). Vidutiniškas diktoriaus autentifikavimo atlykimas, paremtas CHNNs, buvo gerokai patobulintas, lyginant su LTRHMMs metodu (Shahin, 2005).

Charl Johannes van Heerden ir Etienne Barnard tyrinėjo nuo konteksto priklausomas fonemas ir jų trukmę. 2007 pasiūlytas naujas metodas skirtas diktoriaus verifikavimui - priklausomo nuo teksto diktoriaus verifikacinė sistema, paremta paslėptųjų Markovo modelių metodais. Savybių rinkinys, paremtas laikinų nuo konteksto priklausomų fonemų trukme, yra naudojamas tam, kad būtų atpažįstamas tarp kitų diktorių. Atrasta, kad HMM paremta sistema pasiekė vienodą klaidų dažnį apie 0,68 procentų naudojantis tradicinėm (akustinėm) savybėm, o ERR pasiekė apie 0,32%, kai laiko savybės buvo sujungtos su akustinėmis. Šis testas parodė diktoriaus verifikavimo tikrąją laiko savybių vertę. Šios savybės taip pat galėtų būti naudojamos kitiems tikslams, tokiems, kaip atsakymų nustatymui, diktoriaus verifikavimo sistemų sudėtingumo patobulinimui ar diktoriaus variacijoms. Rezultatai patvirtina ankstesnius atradimus, pritaikytus nuo teksto priklausomam diktoriaus atpažinimui ir nuo teksto priklausomam diktoriaus

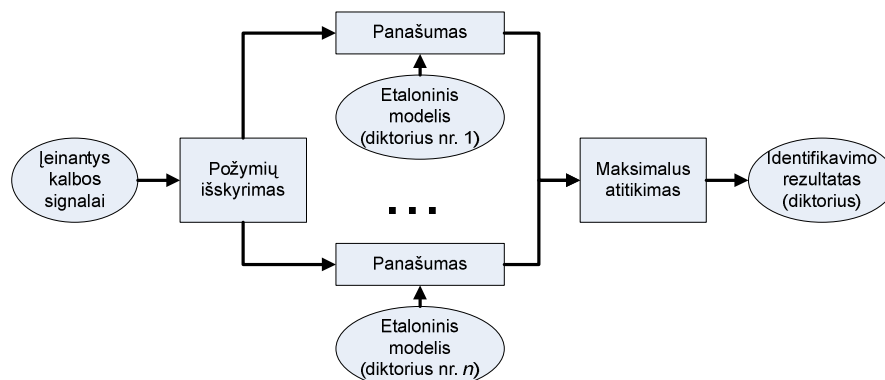
verifikavimo užduotims ir pateikia daugybę pasiūlymų dėl kitų patobulinimų (Heerden, Barnard, 2007).

Lietuvoje diktoriaus atpažinimo tyrimai neatliekami, Lietuvos mokslininkai (Žvynys, Rudžionis, Driaunys, Raškinis, Šilingas, Dereškevičiūtė, Filipovič ir kiti) kol kas plačiai dirba automatinio kalbos atpažinimo srityje, tačiau nebando atpažinti diktoriaus.

1.7.2. Diktoriaus atpažinimas kalbos atpažinimo srities kontekste

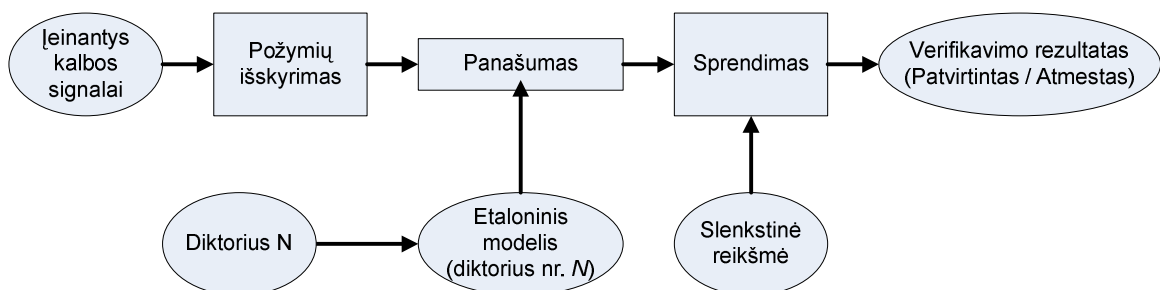
Kaip jau minėta, pagal atliekamas funkcijas diktoriaus atpažinimo sistemos klasifikuojamos į diktoriaus identifikavimo (angl. *Speaker identification*) ir diktoriaus verifikavimo (angl. *Speaker verification*) sistemas, o pagal tai, kokį tekstą turi sakyti diktorius, kad sistema jį atpažintų, skirstomos į nepriklausomas nuo teksto (angl. *Text independent*), priklausomas nuo teksto (Text dependent) ir pateikiamo teksto (angl. *Text-Prompted*) diktoriaus atpažinimo sistemas (Matsui, Furui, 1993).

10 paveiksle pateikiama principinė diktoriaus atpažinimo sistemos diktoriaus identifikavimo schema, o diktoriaus patvirtinimo schema – 11 paveiksle (Matsui, Furui, 1993).



Šaltinis: (Matsui, Furui, 1993).

10 pav. Principinė diktoriaus atpažinimo sistemos diktoriaus identifikavimo schema



Šaltinis: (Matsui, Furui, 1993).

11 pav. Principinė diktoriaus atpažinimo sistemos diktoriaus verifikavimo schema

Nepriklausomos nuo teksto sistemos yra lankstesnės, jos „nepirios“ prie teksto, kuri būtina turi pasakyti naudotojas. Tai nevaržo naudotojo pasirinkimo, jos yra saugesnės slaptažodžio

nuspėjamo požiūriu, nes naudotojas kaskart gali pasirinkti kitą frazę, tačiau tokios sistemos yra sunkiau realizuojamos. Tuo tarpu nuo teksto priklausomos diktorius atpažinimo sistemos naudojamos ten, kur reikia griežtos kontrolės naudotojo įvedamam tekstui, todėl frazę ar frazes, kurias gali ištarti diktorius, sistema saugo duomenų bazėje kaip išanksto įvestą šabloną, pagal kurį gali būti atliekamas lyginimas, o diktorius ištariama frazė veikia kaip slaptažodis.

Be to, pastebėta, kad labai trumpame periode (5 – 100 msek) balso charakteristikos yra labai panašios, tačiau situacija iš esmės pasikeičia ilgame laikotarpyje (1 – 5 sek), kuomet yra ištiriamas didesnis kiekis garsų.

1.7.3. Algoritmai diktorius atpažinimo sistemose

Šiuolaikinis diktorius, kaip ir kalbos, atpažinimas organizuojamas dviem etapais:

1. Požymių išskyrimas (angl. *Feature extraction*);
2. Požymių vektorių palyginimas (angl. *Feature matching*).

Požymių išskyrimo tikslas yra konvertuoti kalbos signalų įrašo diskretas į iš anksto pasirinkto tipo parametrinį atvaizdavimą, vadinamą savybių vektoriais, kurie naudojami tolimesnėje analizėje ir klasifikatorių apdorojime.

Atliekant paruošiamuosius veiksmus, pirmiausia įeinančiam kalbos signalui yra atliekamas pirminis apdorojimas: jis yra diskretizuojamas, suskaidomas į vienodo ilgio langus. Iš jų, naudojant įvairius požymių apskaičiavimo metodus, pateikiamus šiame straipsnyje, yra sudaromi požymių vektoriai, savyje saugantys informaciją, efektyviai aprašančią akustinį kalbos signalą (Viiki, 1999). Kitame diktorius atpažinimo etape naudojantis akustiniais bei kalbos modeliais yra atliekamas požymių vektorių statistinis palyginimas su duomenų bazėje esančių diktorių požymių vektoriais ir išrenkamas labiausiai atitinkantis diktorius.

Požymių išskyrimo algoritmai, naudojami diktoriaus atpažinimo sistemų, signalų požymių išskyrimo:

- Filtrų banku paremti kepstriniai koeficientai, pvz. Melo dažnių kepstriniai koeficientai (angl. *Mel Frequency Cepstral Coefficients, MFCC*) (Mermelstain, Davis, 1982);
- Tiesinis prognozavimas (angl. *Linear Predictive Coding, LPC*) (Vienas pirmųjų algoritmų, pradėtų naudoti kalbos signalų analizei) (Markel, Gray, 1976);
- Tiesinio prognozavimo kepstrinis koeficientas (angl. *Linear Predictive Cepstral Coefficients, LPCC*) (Makhoul, 1975);
- Percepcinis tiesinis prognozavimas (angl. *Perceptual Linear Prediction, PLP*) (Hermansky, 1991);
- Neuroninis prognozavimas (angl. *Neural Predictive Coding, NPC*);
- Spektrinė analizė (angl. *Spectral analysis*).

Šie metodai jau nusistovėję ir naudojami ganėtinai ilgai. Visiems šiems metodams papildomai galima paskaičiuoti šias charakteristikas (Young ir kt., 2001):

- Energiją;
- Delta koeficientus.
- Akceleracijos koeficientus.

Jos irgi naudojamos požymių vektoriaus sudaryme.

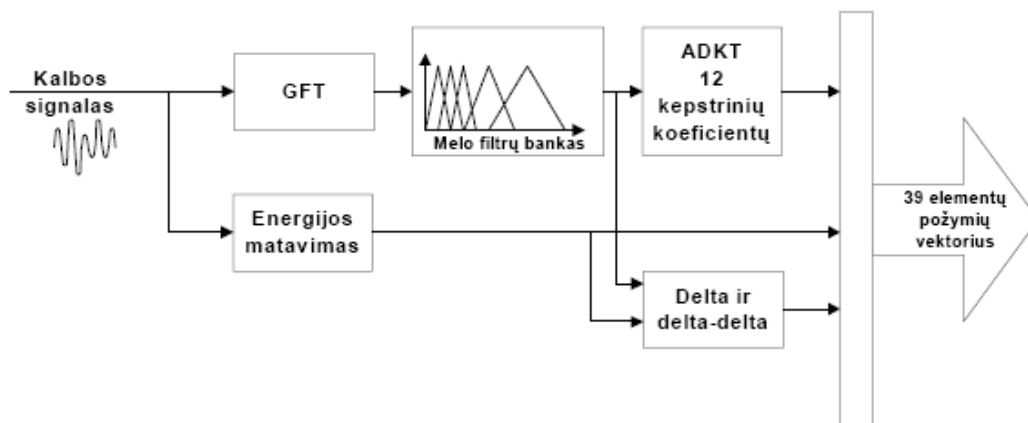
Atlikus kiekvieno lango analizę ir apskaičiavus kalbos signalo požymius iš jų galima sudaryti požymių vektorius, kuris bus naudojamas kaip automatinio diktoriaus kalbos atpažinimo sistemos įėjimas:

$$X_k = [c_k | E_k | \Delta E_k | \Delta \Delta_k | \Delta \Delta E_k], \quad (1)$$

kur X_k – k-tojo signalo požymių vektorius, c_k – MDKK koeficientas arba kiti kepstrinės, spektrinės ar tiesinės prognozės koeficientai, E_k – kalbos signalo energijos, ΔE_k – energijos delta koeficientas, $\Delta \Delta_k$ – MDKK akceleracijos koeficientas (dviguba delta, arba delta delta), $\Delta \Delta E_k$ – energijos akceleracijos koeficientas.

Požymių vektoriaus elementų skaičius priklauso nuo naudojamo MDKK koeficientų skaičiaus, pavyzdžiui, jei naudojame 12 MDKK, tai požymių vektorius sudarys 39 elementai (Driaunys, 2004).

Grafiškai požymių vektoriaus schema pavaizduota 12 paveiksle.



Šaltinis: (Driaunys, 2004).

12 pav. Požymių vektoriaus sudarymo schema

Reikia paminėti, kad mokslininkai vis dar nerado universalus požymių rinkinio, kuris vienareikšmiškai identifikuotų nagrinėjamą kalbos signalo fragmentą. Visi išskiriami požymiai turi savų privalumų ir trūkumų. Kadangi šiuo metu geriausius rezultatus duoda MDKK metodas, jis pasirinktas tyrimo eksperimento metu atlikti požymių išskyrimą.

Požymių palyginimui (angl. *Feature matching*) naudojami įvairūs metodai. Nepriklausomos nuo teksto sistemos atpažinimą atlieka remiantis tokiais metodais:

- Stochastiniai modeliai:
 - Diskretiniai ir tolydiniai paslėptieji Markovo modeliai (Hidden Markov Models);
 - Gauso mišiniais paremti PMM (Gaussian Mixture models);
- Etalonų palyginimo modeliai:
 - Vektorių kvantavimas (Vector Quantization, VQ);
 - Dinaminis laiko transformavimas (Dynamic Time Wrapping, DTW);
- Dirbtiniai neuroniniai tinklai (Neural Networks).

Priklausomos nuo teksto sistemos diktorių atpažįsta pagrinde naudojant paslėptuosius Markovo modelius (Huang ir kt., 1990). Tačiau nepaisant dažno šio metodo naudojimo, šio metodo pagrindinis trūkumas yra tas, jog jis yra jautrus foniniam triukšmui, todėl dažnai šis metodas dažniau pateikia žemesnius teisingo atpažinimo rezultatus (FURUI 1994).

1.7.4. Diktoriaus atpažinimo praktinio taikymo sritys

Tyrimais diktoriaus atpažinime siekiama supaprastinti sąsajų su Diktoriaus atpažinimo sistemų praktinis pritaikymas:

- Numerio rinkimas balsu (angl. *Voice dialing*);
- Telefoninė bankininkystė (angl. *Banking by phone*) (McLaugling, Reynolds, 2002);
- Duomenų bazių prieigos paslaugos (angl. *Databases access services*);
- Konfidencialios informacijos apsauga (angl. *Security control to confidential information*);
- Nutolusių kompiuterių prieiga (angl. *Remote access of computers*);
- Namų automatika, protingieji namai (angl. *Smart houses*);
- Teismo ekspertizė (angl. *Forensics*) ir kt.

Eksperimentinės diktoriaus atpažinimo sistemos jau dabar naudojamos įvairiose veiklos srityse. Kaip teigiama „Speech technology magazine“, rašančiame apie signalų apdorojimo ir kalbos atpažinimo technologijų naujienas, eksponentiškai augant mobiliųjų telefonų skaičiui, JAV kompanija Planet Payment bando išnaudoti šią terpę verslo poreikiui ja pasiekti klientus. Todėl buvo sukurta mokėjimo naudojantis mobiliuoju ryšiu sprendimų linija BuyVoice, kurioje naudojama kompanijos PerSay balso autentifikavimo technologija. Kadangi kiekvieno žmogaus balsas, kaip ir jo pirštų anspaudai ar akies rainelė yra unikalūs, teigiama jog ši autentifikavimo technologija užtikrins pakankamai aukšto lygio paslaugų kokybę, o naudotojams paprastą būdą atlikti norimus veiksmus (Speechtechmag.org, 2007).

Pastaruoju metu kalbos ir diktoriaus atpažinimo sistemomis domisi ir Lietuvos specialiosios tarnybos.

1.8. Analiztinės dalies išvados

1. Atlikus dabartinių dažniausiai naudojamų autentifikavimo priemonių apžvalgą, pastebėta, kad šiomis dienomis dažniausiai naudojama informacijos apsaugos ir naudotojo autentifikavimo priemonė yra slaptažodis.
2. Slaptažodžiai nėra saugios priemonės – jie nesunkiai gali būti nulaužiami perrenkant visas simbolių kombinacijas ar tiesiog atspėjant juos remiantis naudotojo asmenine informacija.
3. Remiantis užsienyje ir Lietuvoje atliktų tyrimų rezultatus (kartu įvertinant ir autoriaus atliktų tyrimų rezultatus), galima teigti, kad slaptažodžiai nėra saugi priemonė ir dėl to, kad patys naudotojai netinkamai elgiasi su jais – kuria nesaugius slaptažodžius, naudoja asmeninius duomenis jiems sudaryti, nepakankamai juos

saugo, dažnai slaptažodžius pasako draugams. Lietuvos naudotojų sudaromi slaptažodžiai yra nepakankamai ilgi, kad juos būtų sudėtinga perrinti naudojant brutalios jėgos metodą.

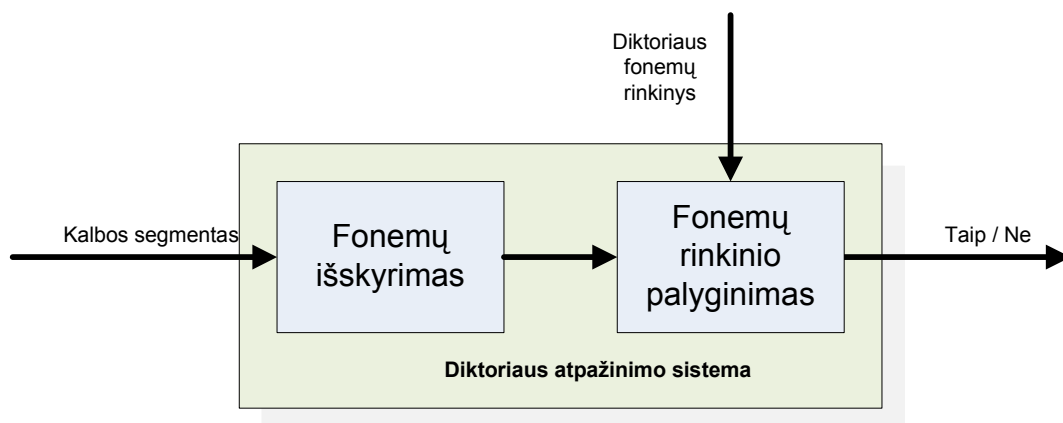
4. Palyginus slaptažodžius, biometrines atpažinimo sistemas ir dviejų faktorių atpažinimo sistemas nustatyta, kad dvi pastarosios yra geriausios pagal savo paskirtį iš visų šiuo metu naudojamų autentifikavimo sistemų.
5. Nepaisant ne itin gerų įvertinimų lyginant su kitų biometrijos technologijų naudojimu, tolimesniam darbui pasirinkta kalbos (balso) atpažinimo technologijų sritis. Šioje srityje su informacijos apsauga yra susiję diktorius atpažinimo tyrimai.
6. Atlikus pasaulyje ir Lietuvoje šioje srityje dirbančių mokslininkų apžvalgą, nustatyta, kad tyrimai šioje srityje atliekami nuo praėjusio amžiaus X dešimtmečio pradžios, yra daug mokslininkų dirbančių šioje srityje (vieni iš daugiausiai dirbančių šioje srityje yra Azijos mokslininkai), sukurta nemažai metodų diktorius atpažinimui.
7. Apžvelgti metodai, naudotini diktoriui atpažinti. Nustatyta, kad plačiausiai paplitęs paslėptųjų Markovo modelių naudojimas požymių palyginimo fazėje, o požymių išskyrimo fazėje geriausius rezultatus duoda Melo dažnių kepstriniai koeficientų metodas. Taip pat plačiai taikomas tiesinės diskriminantinės analizės metodas.

2. DIKTORIAUS ATPAŽINIMAS NAUDOJANT FONEMŲ DISKRIMINAVIMO METODĄ

Siekiant įvertinti, ar diktoriaus atpažinimo naudojant fonemų diskriminavimo metodą galima taikyti diktoriaus, šnekančio lietuvių kalba, atpažinimo uždaviniui spręsti, kuriamas teorinis metodas eksperimentui atlikti. Dviejų žingsnių metodika nėra taikoma, bet kituose tyrimuose skirtuose fonemų diskriminavimui buvo pastebėta, kad skirtingi garsai turi skirtingas potencialias diktoriaus diskriminavimo savybes, todėl reikėtų patikrinti, ar tokia metodika būtų perspektyvi.

Tikėtina, kad tokiu principu paremta sistema turėtų veikti greičiau, nes sistema netikrina visų diktoriaus išstartų fonemų, tačiau tik jam būdingas fonemas. Greitis yra itin svarbi charakteristika tose sistemose, kuriose yra daug naudotojų, o norima, kad sistema veiktų greitai ir patikimai.

13 paveiksle pateikiamas tokiu principu veikiančios sistemos prototipo schema.



Šaltinis: sudaryta autoriaus.

13 pav. Sistemos, naudojančios fonemų diskriminavimo metodą diktoriui atpažinti, schema

Sistemos prototipo veikimo logikos supaprastintas pavyzdys:

Diktorius: Aš esu diktorius Jonas.

Diktoriaus atpažinimo sistema: Ištark "du devyni praDėti".

Diktorius: du DevynI praDėti.

Diktoriaus atpažinimos sistema: Tu ne diktorius Jonas.

Skyriuje aprašomi žingsniai ir metodika, pagal kurią turėtų būti atliekamas praktinius rezultatus duodantis eksperimentas.

2.1. Duomenų paruošimas

Pirmas eksperimento etapas turėtų būti duomenų sutvarkymas. Pirminiai duomenys, naudojami eksperimente yra LTDIGITS garsyno 100 diktorių fonemos ir Melo dažnių kepstriniai

koeficientai (MDKK). Turimą garsyno aibę matematiškai galime pažymėti $W = w_1, w_2, w_3, \dots, w_m$, čia kiekvienas w_i žymi žodį iš diktoriaus atpažinimui naudojamą garsyno.

Fonemų tinkamumui įvertinti bus naudojamas tiesinės diskriminantinės analizės (TDA) metodas. Tai yra vienas iš paprasčiausių kalbos atpažinime naudojamų klasifikavimo metodų. Diskriminantinės analizės metodikos pagrindus faktiškai sukūrė R. Fišeris. Šis metodas remiasi tokiomis prielaidomis:

- 1) Grupių (arba fonemų) Ph skaičius yra baigtinė aibė;
- 2) Grupės yra nepriklausomos viena nuo kitos bei neturi bendrų objektų (fonema vienu metu negali priklausyti kelioms grupėms);
- 3) Kintamųjų kovariacinės matricos, naudojamos diskriminacinės funkcijos, yra vienodo dydžio.

Taigi fonemos tinkamumo lygio diktoriui atpažinti uždavinį galima suformuluoti kaip stebimo požymių vektoriaus priskyrimą vienai iš galimų klasių. Naudojant klasifikatoriaus diskriminantinę funkciją $g(x)$, fonemos x priskyrimas vienai iš dviejų klasių $Ph1$ ir $Ph2$ nusakomas tokia sistema:

$$\begin{cases} jei & g(x) \geq 0, & \text{tai } x \text{ priklauso klasei } Ph1 \\ jei & g(x) < 0, & \text{tai } x \text{ priklauso klasei } Ph2 \end{cases} \quad (2)$$

Bendruoju atveju tiesinė diskriminantinė funkcija išreiškiama

$$g(x) = \omega_T x + \omega_0, \quad (3)$$

kur $\omega = \omega_0, \dots, \omega_T$ – diskriminantinės funkcijos koeficientai.

Jei laikytumėmės prielaidos, kad klasių, kurioms bus priskiriami diktoriai, požymiai yra pasiskirstę pagal gauso normalųjį skirstinį, tiesinės diskriminantinės funkcijos formulę galima perrašyti tokia išraiška :

$$g(x) = (x - \mu)' \Sigma^{-1} (x - \mu) \quad (4)$$

Iš šios lygties seka, kad norint klasifikuoti diktorius, reikės apskaičiuoti fonemos požymių vektoriaus elementų vidurkius ir kovariacines matricas. Vidurkiai skaičiuojami pagal formulę

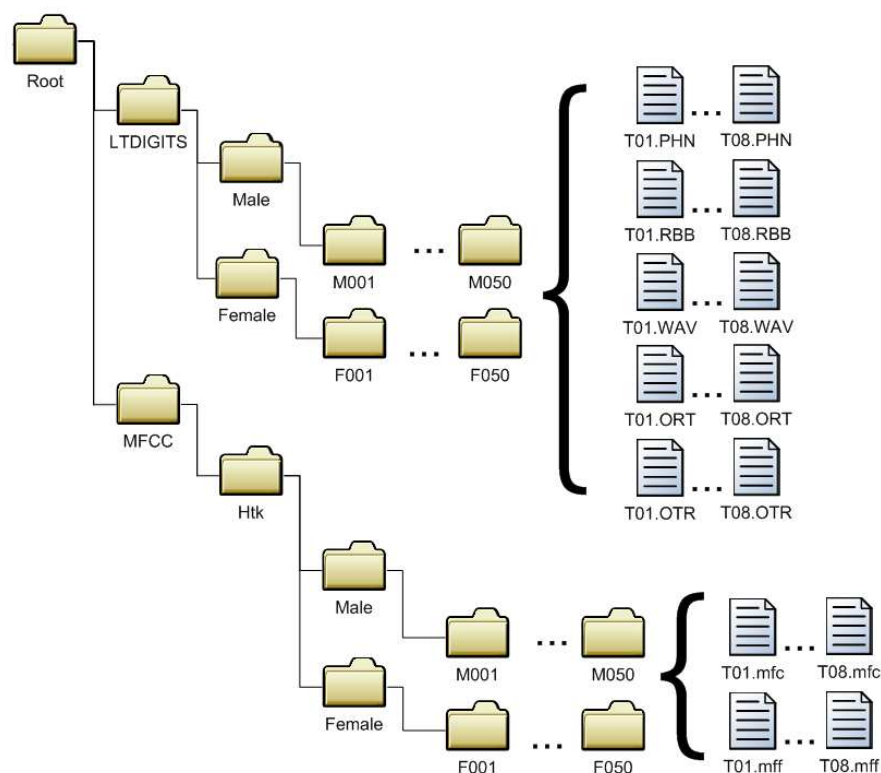
$$\mu = \frac{1}{N} \sum_{i=1}^N X_i \quad (5)$$

Kovariacinė matrica paskaičiuojama pagal formulę

$$\Sigma = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)(x_i - \mu)^T \quad (6)$$

Sudarant MDKK koeficientus naudotas 16 ms. lango ilgis, 6,25 ms žingsnis, jis buvo skaičiuojamas iš 20 filtrų Melo filtrų banke ir 12 keptrinių koeficientų. Papildomai buvo naudota energija, pirmosios ir antrosios išvestinės (delta ir delta delta). Požymių vektorius buvo sudaromas

remiantis tiesinio prognozavimo bendrąja metodika iš 39 elementų (Driaunys, 2004), aprašyta 1.7.3 skyrelyje. MFF failai, skirti darbui Matlab aplinkoje pavadinti *.mff, saugomi kataloge „Htk“. Failinė LTDIGITS garsyno struktūra, kuri bus naudojama eksperimente, pateikta paveiksle žemiau.



Šaltinis: sudaryta autoriaus

14 pav. LTDIGITS garsyno failinė struktūra

Apdorojant duomenis, visų pirma reikia patikrinti, ar visos fonemos yra tinkamos, ar nėra kokių nors neteisingų fonemų. Be to, dėl visa ko reikia suskaičiuoti visų diktorių visų frazių fonemų skaičių bei kiekvienos lyties diktorių išstartų fonemų skaičių atskirai.

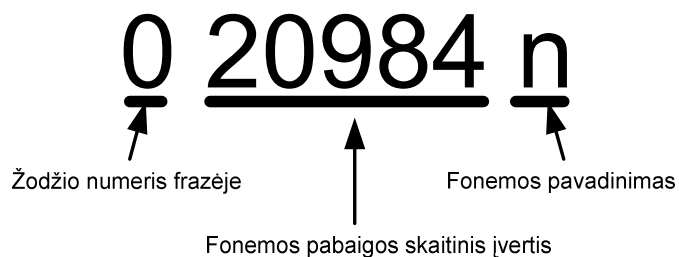
Tikrinimui skirtas fonemas reikia nuskaityti iš fonemų failo. Taip turėtų būti užtikrinamas programos funkcionalumas pasikeitus apdorojamų fonemų skaičiui.

Nuskaitomas diktorių failas, ir kiekviena apdorojama fonema turi būti patikrinta abiejų arba kurios nors iš lyčių, priklausomai nuo diktorių failo turinio, visų diktorių visose frazėse.

Atsižvelgiant į failinės sistemos specifiką, tiek diktorius, tiek frazės katalogo pavadinimas turi būti sudaromi dinamiškai. Todėl žinant diktorius numerį, frazės numerį, dinamiškai bus galima sukurti kelis iki ribų (failo plėtinys *.RBB) ir fonemų failų (*.PHN) failų. Iškart nustatomi keliai iki failų, kuriuose bus saugomi diktorius fonemos apdorojimo rezultatai: MFCC koeficientų vektorių masyvas (failo plėtinys *_mfcc_mas.vkt), ieškomos fonemos vektorių vidurkis (*_average.txt) ir kovariacinė matrica (*_covmat.txt).

Pradžią ir pabaigų nustatymas. Fonemos pabaigos ribos reikšmė bus nuskaitoma iš fonemų failo (*.PHN).

Fonemų failo eilutės sandara pavaizduota 15 paveiksle.



Šaltinis: sudaryta autoriaus.

15 pav. *.PHN failo eilutės struktūra

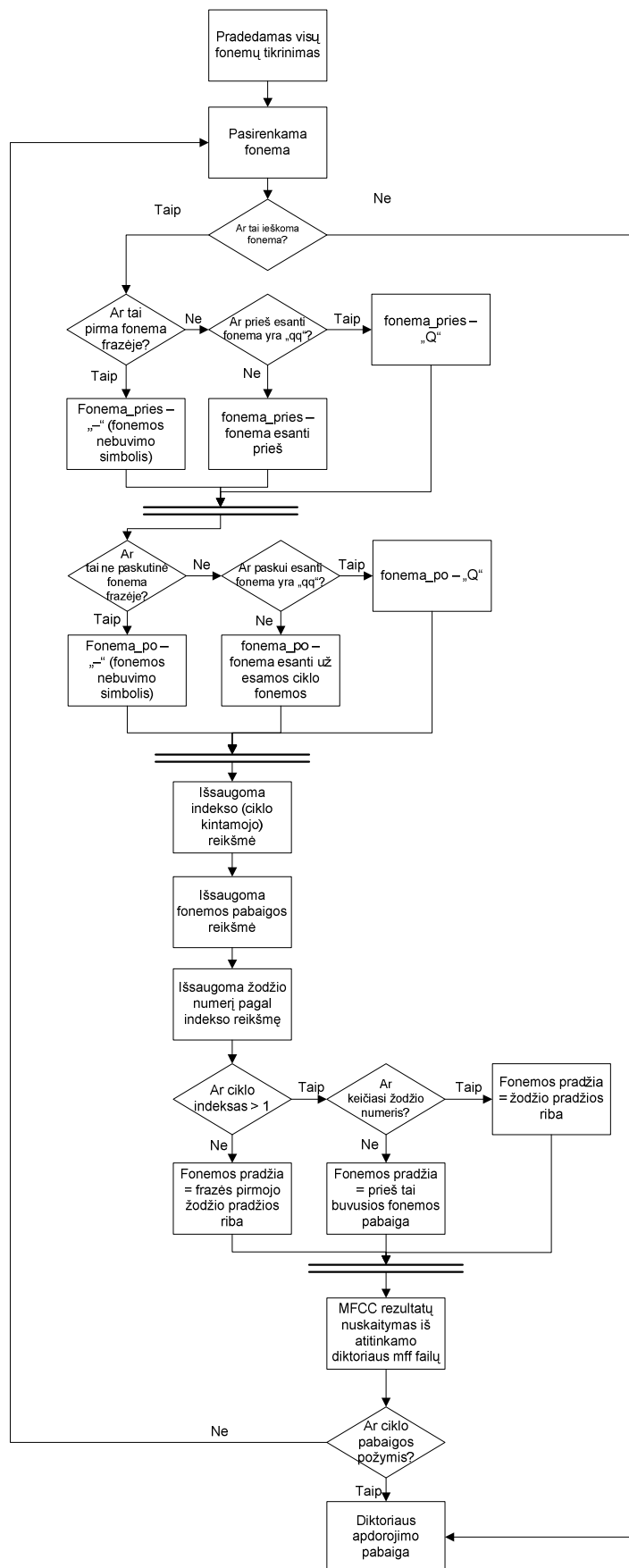
Fonemų pabaiga, priklausomai nuo aplinkybių, gali būti nuskaitoma iš fonemų arba žodžių ribų failo. Žodžių ribų failo eilutės sandara pavaizduota 16 paveiksle.



Šaltinis: sudaryta autoriaus.

16 pav. *.RBB failo eilutės struktūra

Toliau pateikta galima fonemos pradžios ir pabaigos nustatymo algoritmo blokinė schema.



Šaltinis: sudaryta autoriaus.

17 pav. Fonemos pradžios ir pabaigos nustatymo algoritmo blokinė schema

Pagal gautas fonemos pradžios ir pabaigos reikšmes, iš atitinkamo diktoriaus frazės Melo dažnių kepstrinių koeficientų failo iškerpamas reikalingas MDKK koeficientų skaičius. *.Mff faile fonemos Melo dažnio kepstrinio koeficiento pradžia randama fonemos pradžios skaičių padalinus iš 200, bei dauginant iš 4 ir iš koeficientų skaičiaus (šiuo atveju 39). Reikės apskaičiuoti MDKK baitų skaičių, kurį reikia iškirpti. Šiuo tikslu iš fonemos pabaigos reikšmės atimama fonemos pradžios reikšmė, skirtumas dalinamas iš 200 ir dauginama iš koeficientų skaičiaus.

Gautas fonemos kepstrinių koeficientų masyvas turės būti skaidomas į 39 koeficientų ilgio požymių vektorių masyvą $X = x_1, x_2, \dots, x_n$, kuris programiškai galėtų būti atvaizduotas kaip daugiamatis masyvas. Galimas sprendimas yra ir išsaugoti visus gautuosius rezultatus failinėje sistemoje, tam kad paskui, reikalui esant, galima būtų lengviau ieškoti klaidų. Tokiu atveju kiekvienos fonemos MDKK reikšmės turės būti grupuojamos po 39 elementus ir eilutėmis įrašinėjami į failus.

Apskaičiavus vienos fonemos visus įmanomus vektorius, bus suformuojamas diktoriaus fonemos savybių vektoriaus vidurkis, kuris kiekvienam diktoriui turėtų būti skirtingas. Apskaičiuotas vidurkis irgi turėtų būti saugomas failinėje sistemoje, siekiant išsaugoti jo duomenis tolesniuose tyrimuose.

Vienas iš paskutinių etapų duomenų ruošime prieš fonemų, geriausiai identifikuojančių diktorių, atrinkimą yra kovariacinės matricos diagonalės skaičiavimas. Diagonalė skaičiuojama todėl, kad likusieji kovariacinės matricos elementai neturės įtakos tolimesniuose skaičiavimuose. Apskaičiuota kovariacinės matricos diagonalė išsaugoma failinėje struktūroje.

Analogiški skaičiavimai atlikti su visomis fonemomis. Atlikant kiekvienos fonemos apskaičiavimus, reikėtų papildomai saugoti kiekvieno reikšmingesnio žingsnio rezultatus, taip pat informacinius pranešimus apie fonemos, diktoriaus, frazės apdoravimo pradžią ir pabaigą.

2.2. Diktoriaus klasifikacija pagal fonemas

Turint pirminius fonemų vektorių vidurkių ir kovariacinių matricų skaičiavimų reikšmes, turėtų būti atrenkamos tinkamiausios fonemos, geriausiai išskiriamos individualios savybės ir tinkamiausios fonemos. Tinkamumas suprantamas kaip fonemos sugebėjimas su mažiausia klaidos tikimybe iš fonemos išgauti maksimalias diktorių identifikuojančias savybes.

Tinkamų fonemų parinkimas reikalauja gilesnio tyrimo, todėl antruoju etapu sudaromos fonemų eilės – fonemos klasifikuojamos. Naudojamas standartinis fonemų atpažinimo modelis, kai klasifikatorius apmokomas naudojant paskaičiuotas kiekvieno diktoriaus kiekvienos fonemos požymių vektoriaus vidurkius bei kovariacines matricas, o diktoriaus klasifikavimo procese skaičiuojamas analizuojamos fonemos panašumo skaitiniai įverčiai, kad požymių vektorių ir

etaloninės reikšmės yra panašios arba nepanašios. Šis skirtumas tarp vektoriaus ir etalono nusakomas išraiška

$$k_i = (S - \mu_i)S_i^{-1}(S - \mu_i)' \quad (7)$$

Gautosios k_i reikšmės lyginamos tarpusavyje, kuo fonemos įverčio reikšmės didesnės, tuo geriau fonema turėtų būti identifikuoti diktorių.

Duomenų apdorojimui ir manipuliacijoms su MFCC reikšmėmis atlikti nuspręsta naudoti Matlab 6.1 programinį paketą.

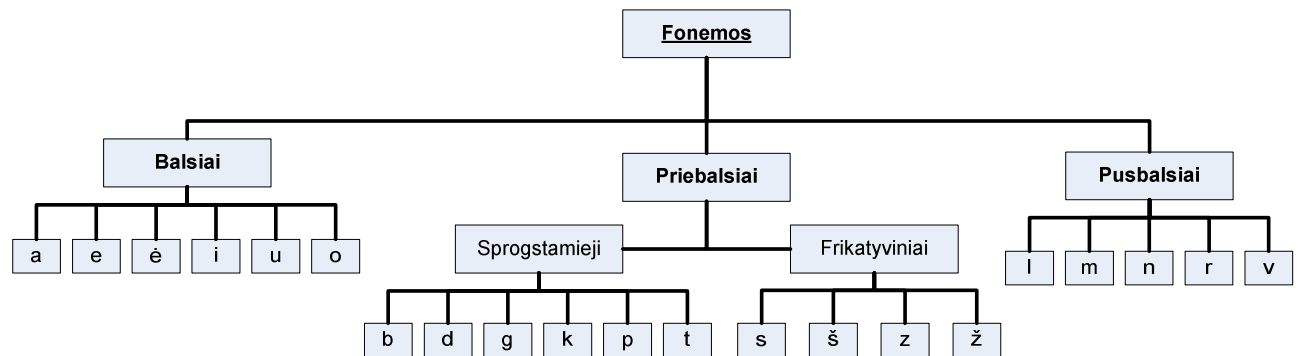
Paskutinis žingsnis – atlikti fonemų skaitinių įverčių grupavimus ir skaičiavimus įvairiais pjūviais, siekiant išsiaiškinti geriausią fonemų aibę ar jos ribas, kurios galėtų geriausiai reprezentuoti diktorių. Skaičiavimus atlikti naudojant MS Excel 2003 ir/ar Statistica 7.

3. EKSPERIMENTINIS SKYRIUS

3.1. Duomenys

Tyrimai atliekami naudojant lietuviško garsyno LTDIGITS duomenis, duomenų apdorojimui naudojami 100 abiejų lyčių diktorių kalbos duomenys (50 diktorių vyrų ir tiek pat diktorių moterų).

Kaip aprašyta metodinėje tyrimo dalyje, turimi duomenys yra suskirstyti į 50 diktorių vyrų katalogus ir 50 diktorių moterų katalogus, kiekvienas diktorius priklausomai nuo lyties turi savo identifikacinį kodą. Kodo struktūra yra LOXY, kur L – lyties identifikatorius (F arba M), X – skaičius nuo 0 iki 5, Y – skaičius nuo 0 iki 9 (diktoriaus identifikacinio kodo pavyzdžiai yra F032, M005). Kiekvienas diktorius yra įgarsinęs po 8 frazes, kiekvienoje iš jų sudaro 6 žodžiai (skaičių pavadinimai bei kiti žodžiai: vienas, du, trys, keturi, penki, šeši, septyni, aštuoni, devyni, nulis, taip, ne, pagalba, saugoti, start, stop, pirmyn, atgal, į pradžią, į pabaigą, sekantis, perduoti, pradėti, baigti, sustoti, pauzė, laukti, tęsti).



Šaltinis: sudaryta autoriaus pagal (Girdenis, 1995).

18 pav. Lietuvių kalbos fonemų LTDIGITS garsyne klasifikacija

Tyrimo metu naudotos LTDIGITS garsyno fonemos yra prieš tai susegmentuotos ankstesniuose Vilniaus universiteto Kauno Humanitarinio fakulteto mokslininkų darbuose – kaip aprašyta 2 skyriuje, žodžiai išskaidyti į atskiras fonemas priskiriant jiems pradžios ir pabaigos reikšmes, įrašant pastabas apie fonemą, išskiriant klaidas. Segmentavimo metu kai kurios fonemos žymimos kiek kitaip, nei fonema yra tariama, pradžioje tai gali kelti sumaištį. LTDIGITS garsyno fonemos ir jas atitinkantis žymėjimas pateikiamas 6 lentelėje.

6 lentelė. Eksperimento metu naudotos fonemos ir jų žymėjimas

Eil. nr	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Fonema	a	b	d	e	ė	g	i	k	l	m	n	o	p	r	s	š	t	u	v	z	ž
Žymėjimas	a	b	d	e	y	g	i	k	l	m	n	o	p	r	s	S	t	u	v	z	Z

Šaltinis: sudaryta autoriaus.

Kaip matome iš fonemų, naudojamų tyrime, lentelės, fonema „è“ tyrimo duomenyse žymima „y“, „š“ – simboliu „S“, o fonema „ž“ – simboliu „Z“. Reikia paminėti, kad tyrimo metu netyrinėta fonemų sprogstamųjų priebalsių „f“, „h“ bei pusbalsio „j“ įtaka diktoriaus atpažindinime. Be to, tyrime neatsižvelgiama į dvibalsių ir dvigarsių įtaką diktorių identifikuojančių fonemų rinkinio sudarymui.

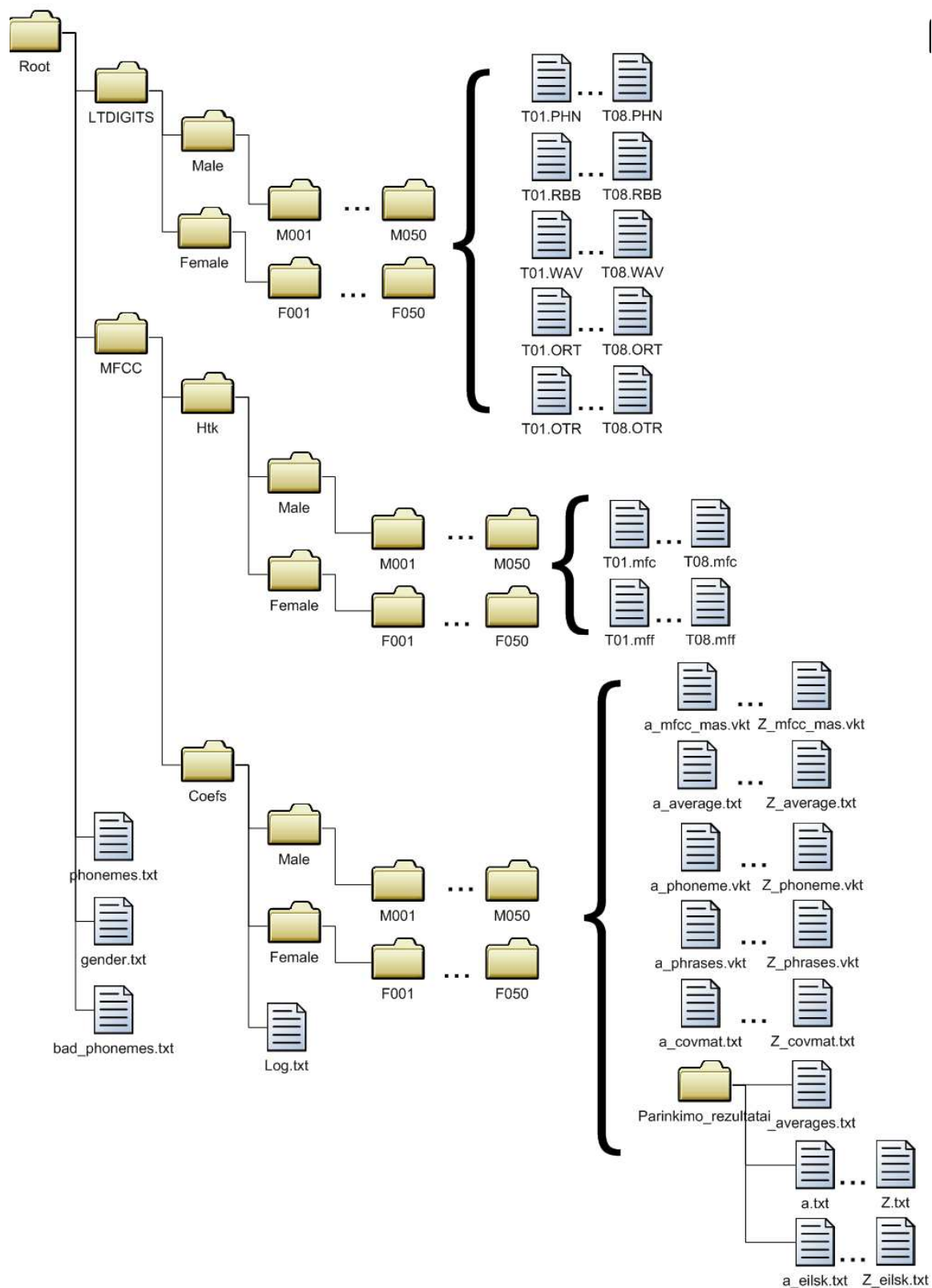
Kadangi atliekant eksperimentinių duomenų pasiruošimą, duomenų failų analizės metu buvo nustatyta, kad fonemų failuose yra ne tik reikšmines fonemas identifikuojančios reikšmės, o ir tarpus tarp fonemų arba kitas vietas, nenešančias reikšminės informacijos. Tokios vietos fonemų faile turi reikšmę „qq“ vietoje fonemos pavadinimo, todėl eksperimento metu pradinis algoritmas buvo patobulintas – skaičiuojant prieš ir už ieškomos fonemos esančias fonemas, jei viena iš ieškomų fonemų yra pažymėta simboliais „qq“, išsaugant informaciją apie rastą fonemą (fonemos pradžios ir pabaigos diskretų faile reikšmės) ir frazę, kurioje ji rasta (frazės numeris, fonema prieš ieškomą fonemą ir už ieškomos fonemos), simboliai „qq“ verčiami į simbolį „Q“, taip suvienodinant visų fonemų ir bei prieš ir už jų esančių fonemų ilgį (žr. 17 pav.).

3.2. Duomenų apdorojimas ir analizė

Duomenų apdorojimui remiantis 2 skyriuje nustatyta metodika ir apsiraušytais preliminariais algoritmais, buvo sukurtos penkios Matlab programos, skirtos:

- 1) Katalogų sistemos automatiniam sukūrimui;
- 2) LTDIGITS garsyno fonemų patikrai ir netinkamų tyrimui fonemų skaičiavimui;
- 3) Duomenų apdorojimui, savybių vektorių sudarymui, kovariacinės matricos sudarymui.
- 4) Panašumo įverčių tarp fonemų ir etalono pagal formulę nustatymui ir fonemų klasifikavimui.
- 5) Programa, skirta (1), (3) ir (4) programų automatiniam vykdymui ir bendro vykdymo laiko skaičiavimui.

Tyrimo metu buvo stengiasi išlaikyti esamos failinės struktūros logiką. Žemiau esančiame paveiksle pateiktas failinės sistemos schema atlikus tyrimą.



Šaltinis: sudaryta autoriaus.

19 pav. Papildyta LTDIGITS failinė struktūra

Kaip jau minėta, pradinę failinę struktūrą papildanti struktūra yra sukuriama automatiškai prieš atliekant fonemų apdorojimo ir klasifikavimo veiksmus. Programų išeities kodas pateikiamas 4 – 8 prieduose.

3.3. Rezultatų analizė ir apibendrinimas

Pirmiausia, remiantis turima informacija ir ankstesnėmis žiniomis apie LTDIGITS garsyno segmentavimą ir jo ypatumus, turimas garsynas patikrintas, kiek jame yra netinkamų tyrimui fonemų. Nepilnai išstartos ar dėl kitų priežasčių segmentavimo metu klaidos simbolių „x“ gavusios fonemos pateikiamos toliau einančioje 7 lentelėje.

7 lentelė. Netinkamo formato pradinės tyrimo fonemos

Eil. nr.	Bloga fonema	Diktorių vyrų	Diktorių moterų	Iš viso:
1	ix	0	1	1
2	xa	11	11	22
3	xb	1	2	3
4	xd	11	10	21
5	xe	12	14	26
6	xg	1	4	5
7	xi	26	21	47
8	xy	3	4	7
9	xk	2	2	4
10	xl	4	5	9
11	xm	3	1	4
12	xn	14	9	23
13	xo	9	12	21
14	xp	15	14	29
15	xr	6	12	18
16	xs	15	15	30
17	xt	22	20	42
18	xu	13	14	27
19	xv	3	5	8
20	xZ	5	6	11
	Iš viso	176	182	358

Šaltinis: sudaryta autoriaus.

Kaip matome, 100 diktorių frazėse visose rasta 358 fonemos, kurios segmentavimo etape buvo įvardintos kaip klaidingos. Visos klaidingos fonemos saugomos failinėje sistemoje (failas bad_phonemes.txt, žr. 19 pav.). Kiekviena šio failo eilutė saugo tokią informaciją:

- Diktoriaus numeris;
- Frazės numeris;
- Fonemos su klaidos ženklu pavadinimas;

Toliau atliekant eksperimentą duomenys buvo apdoroti Matlab aplinkoje parašytomis programomis, kurios, kaip minėta, pateikiamos 4 – 8 priede.

Pradiniai gautieji rezultatai – apskaičiuoti kiekvieno diktoriaus kiekvienos fonemos tinkamumo diktoriui klasifikuoti (atitikimo panašumo) skaitiniai įverčiai. 8 ir 9 lentelėje pateikiami vyrų ir moterų diktorių neapdorotų (originalių) skaitinių išraiškų palyginimas skirtingų diktorių tarpe.

8 lentelė. Neapdoroti vyrų diktorių fonemų apdorojimo rezultatai

Eil. Nr.	M002		M008		M013		M023		M028	
1	17.313593	a	13.196064	a	10.982662	a	20.377880	a	22.656697	a
2	15.063263	b	11.162712	b	29.112604	b	18.211842	b	10.395733	b
3	18.677880	d	10.183497	d	9.631016	d	13.809356	d	11.063216	d
4	18.097332	e	13.841417	e	11.260546	e	18.647865	e	23.489837	e
5	9.003720	y	9.152628	y	4.807655	y	14.704197	y	9.649149	y
6	15.170485	g	12.783992	g	18.157160	g	17.021678	g	12.891282	g
7	17.588577	i	16.078428	i	11.930601	i	19.685874	i	22.347844	i
8	22.919207	k	36.968343	k	22.591978	k	18.685666	k	16.969023	k
9	19.087194	l	17.154628	l	10.760599	l	17.835698	l	20.694142	l
10	8.450124	m	5.057821	m	7.002555	m	6.767811	m	8.035412	m
11	17.361121	n	14.558471	n	14.258607	n	20.092258	n	17.384320	n
12	22.526731	o	22.160686	o	14.459369	o	23.161461	o	24.052379	o
13	29.129596	p	33.341622	p	29.614835	p	31.885422	p	30.161233	p
14	23.296840	r	22.996706	r	24.843493	r	27.120281	r	25.465299	r
15	13.465025	s	10.866682	s	10.798155	s	12.422448	s	13.210884	s
16	13.465025	S	10.866682	S	10.798155	S	12.422448	S	13.210884	S
17	14.617467	t	19.463358	t	16.289939	t	14.977176	t	15.060893	t
18	20.354964	u	16.689994	u	12.357315	u	21.227704	u	21.864455	u
19	22.229784	v	17.026047	v	14.103374	v	17.898898	v	20.088122	v
20	17.486153	z	20.524746	z	18.758998	z	19.124275	z	28.118268	z
21	17.486153	Z	20.524746	Z	18.758998	Z	19.124275	Z	28.118268	Z

Šaltinis: sudaryta autoriaus.

9 lentelė. Neapdoroti moterų diktorių fonemų apdorojimo rezultatai

Eil. Nr.	F004		F011		F024		F026		F040	
1	12.018848	a	13.098696	a	16.268247	a	16.090484	a	17.731191	a
2	9.422276	b	5.829775	b	11.702812	b	11.573227	b	5.421475	b
3	8.943830	d	7.946329	d	9.105633	d	16.014940	d	10.267411	d
4	13.400383	e	12.826928	e	13.608070	e	14.916013	e	13.003989	e
5	7.238338	y	10.097757	y	9.998429	y	9.787212	y	8.740716	y
6	13.242391	g	10.410669	g	11.771410	g	10.844525	g	13.714889	g
7	14.210133	i	11.652537	i	13.349443	i	15.585751	i	16.878942	i
8	25.394368	k	26.075212	k	24.068317	k	19.801600	k	25.215669	k
9	9.495920	l	11.826616	l	14.684290	l	14.323501	l	17.178171	l
10	5.346261	m	5.470603	m	4.494494	m	4.728649	m	3.493473	m
11	10.296120	n	10.939723	n	11.547936	n	13.169216	n	14.231495	n
12	16.112202	o	12.871160	o	14.213994	o	21.330242	o	17.874956	o
13	29.844054	p	35.484559	p	43.175438	p	47.429545	p	33.842934	p
14	16.153211	r	24.297565	r	22.857767	r	22.787803	r	17.483406	r
15	9.709476	s	11.837233	s	11.504944	s	9.276783	s	15.941511	s
16	9.709476	S	11.837233	S	11.504944	S	9.276783	S	15.941511	S
17	15.097450	t	15.917415	t	16.294248	t	12.358423	t	17.766812	t
18	17.740130	u	11.323154	u	15.166925	u	18.930522	u	17.401215	u
19	12.948040	v	15.629004	v	13.339597	v	14.414935	v	13.478569	v
20	28.655805	z	11.810344	z	23.080237	z	10.486326	z	18.136206	z
21	28.655805	Z	11.810344	Z	23.080237	Z	10.486326	Z	18.136206	Z

Šaltinis: sudaryta autoriaus.

Originali, visus tyrimo metu tirtus diktorius ir jų fonemas apimanti lentelė, kurioje pateikiami diktorių vyrų fonemų įverčių duomenys pateikiama 9 priede, moterų – 10 priede.

Siekiant įvertinti kiekvienos fonemos tinkamumo atpažinti diktorių sklaidą, skaičiuotos skaitinių fonemų įverčių minimali, maksimali reikšmė, jų skirtumas, bei vidurkis. Gautieji vyrų diktorių rezultatai pateikiami 10 lentelėje.

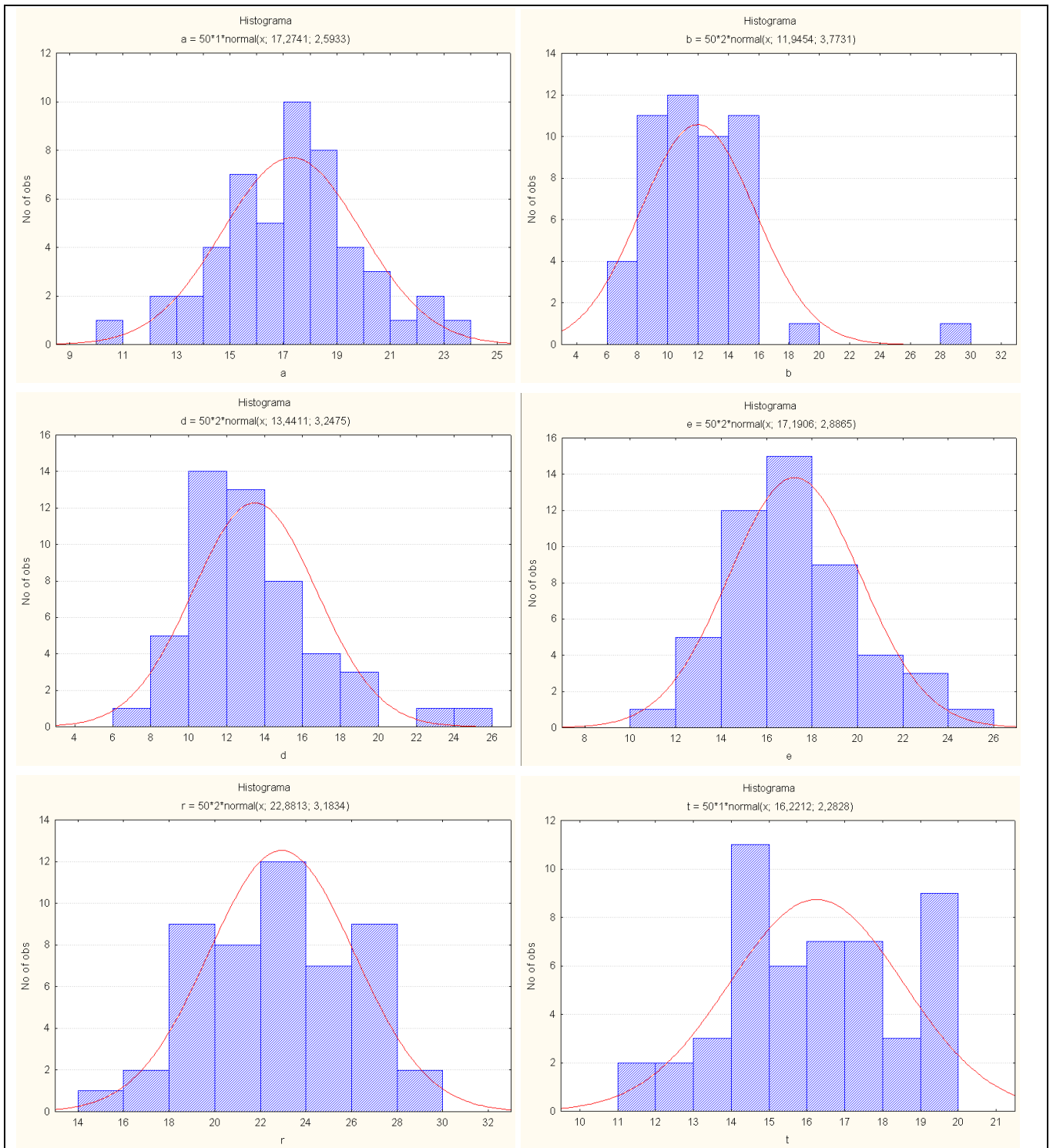
10 lentelė. Vyrų diktorių fonemų įverčių minimalios, maksimalios, skirtumo ir vidutinės reikšmės

Eil. nr	Fonema	Min	Max	Skirtumas (Max - Min)	Vidurkis	Mediana	Standartinis nuokrypis
1	a	10,9827	23,1897	12,2070	17,2741	17,3201	2,5933
2	b	6,8132	29,1126	22,2994	11,9454	11,1384	3,7731
3	d	7,1906	24,1225	16,9320	13,4411	12,8056	3,2475
4	e	11,2605	24,2002	12,9396	17,1906	16,5713	2,8865
5	y	4,8077	14,7042	9,8965	9,1509	8,7328	1,8957
6	g	8,3109	22,5426	14,2317	14,8569	14,5252	3,5857
7	i	11,0458	22,3478	11,3020	16,1046	16,0120	2,7989
8	k	9,9164	36,9683	27,0519	19,8979	19,7318	5,3097
9	l	10,7606	24,9562	14,1956	18,3087	18,3773	2,8081
10	m	3,7121	16,6863	12,9742	6,3899	5,8726	13,5562
11	n	11,8762	24,6486	12,7723	16,6100	16,0402	2,9706
12	o	14,4508	28,4490	13,9982	19,6085	19,6928	3,5452
13	p	19,2332	43,1170	23,8838	30,8769	31,1234	5,5261
14	r	15,7878	28,8500	13,0622	22,8813	23,0972	3,1834
15	s	8,6089	17,0640	8,4551	11,6442	11,3285	1,8405
16	S	8,6089	17,0640	8,4551	11,6442	11,3285	1,8405
17	t	11,4005	19,9716	8,5711	16,2212	16,1824	2,2828
18	u	12,3573	29,5771	17,2198	20,1318	19,8304	3,7456
19	v	9,1218	22,2298	13,1080	16,8248	17,0304	3,2340
20	z	12,1606	44,3760	32,2154	19,8150	18,4996	6,2999
21	Z	12,1606	44,3760	32,2154	19,8150	18,4996	6,2999

Šaltinis: sudaryta autoriaus.

Matome, kad skirtingas fonemas diktoriai vyrai taria skirtingai, nes gauti skirtingi fonemų skaitiniai įverčiai. Kai kurių fonemų, tokių kaip „s“ ir „š“, „z“ ir „ž“ visos charakteristikos nesiskiria, o kitos fonemos neturi vienodų įverčių. Skiriasi fonemų vidutinės reikšmės, mediana, standartiniai nuokrypiai.

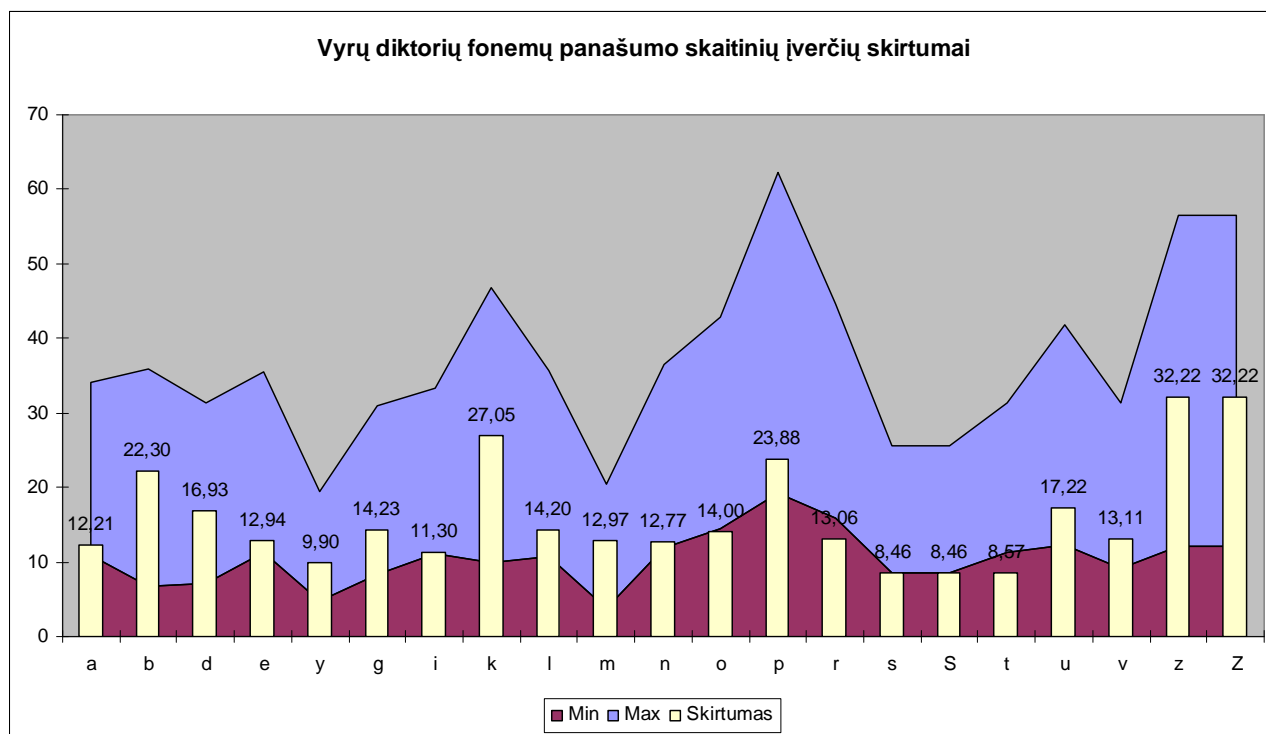
20 paveiksle pateikiamos atsitiktinai pasirinktų įvairių fonemų įverčių histogramos.



Šaltinis: sudaryta autoriaus.

20 pav. Diktorių vyrų fonemų „a“, „b“, „d“, „e“, „r“ ir „t“ skaitinių įverčių histogramos

21 paveiksle pateikiamas grafinis 10 lentelėje pavaizduotų minimumo ir maksimumo bei jų skirtumo rezultatų vaizdas.



Šaltinis: sudaryta autoriaus.

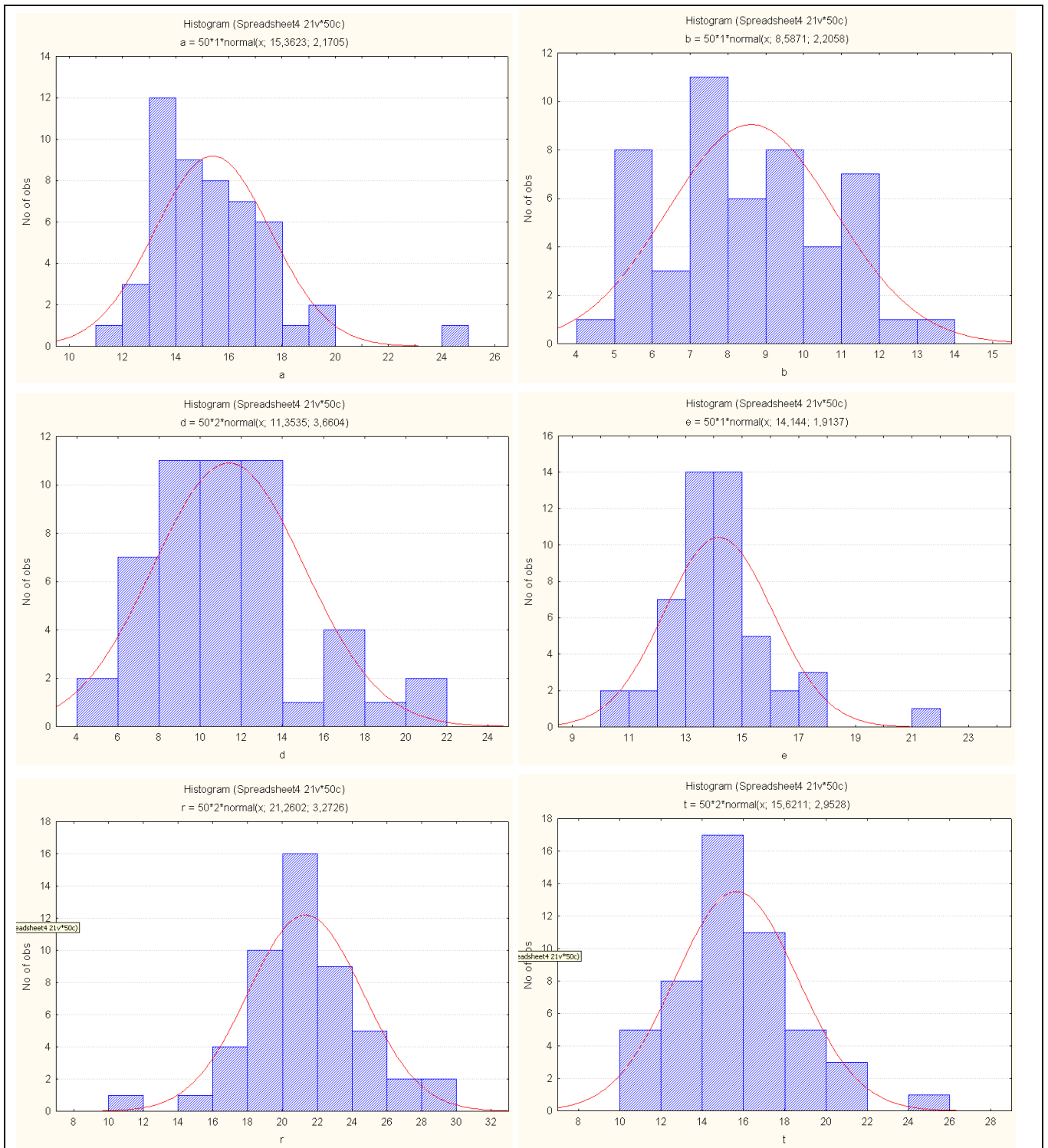
21 pav. Vyrų diktorių fonemų panašumo skaitinių įverčių minimalių, maksimalių ir skirtumų grafiškas atvaizdavimas

Analogiškai diktoriaus diktoriaus atliekami paskaičiavimai ir moterų diktorių fonemoms. Rezultatai pateikiami 11 lentelėje bei 22 ir 23 paveiksluose.

11 lentelė. Moterų diktorių fonemų įverčių min, max, skirtumo ir vidutinės reikšmės

Eil. nr	Fonema	Min	Max	Skirtumas	Vidurkis	Mediana	Standartinis nuokrypis
1	a	11,9654	24,0066	12,0412	15,3623	14,9929	2,1705
2	b	4,9518	13,5698	8,6180	8,5871	8,4372	2,2058
3	d	5,8422	21,3544	15,5122	11,3535	10,8044	3,6604
4	e	10,3063	21,9751	11,6688	14,1440	13,9229	1,9137
5	y	3,6788	25,8251	22,1463	8,7422	8,5284	3,0084
6	g	6,5277	24,6803	18,1526	11,4894	11,1329	2,7627
7	i	9,1830	27,2290	18,0459	13,6207	13,1129	2,8133
8	k	7,1045	34,0753	26,9708	20,8834	21,3550	5,6500
9	l	4,5885	24,0947	19,5062	14,5555	14,6589	3,2473
10	m	2,8729	10,7522	7,8793	5,2811	4,8024	1,4958
11	n	7,9856	25,8906	17,9049	12,4569	12,1350	3,0305
12	o	10,5332	27,6872	17,1540	17,2584	16,5924	3,6723
13	p	21,2893	52,9359	31,6466	34,3094	33,7029	6,1680
14	r	11,6302	29,5495	17,9193	21,2602	21,0284	3,2726
15	s	8,3166	17,2161	8,8996	10,6366	9,8364	1,8784
16	S	8,3166	17,2161	8,8996	10,6734	9,8364	1,9160
17	t	10,6106	25,5696	14,9589	15,6211	15,2462	2,9528
18	u	10,5067	23,8266	13,3198	16,6109	16,8834	3,0621
19	v	5,5306	21,2379	15,7073	13,6293	13,4492	3,5484
20	z	8,3822	28,6558	20,2736	15,0202	13,4250	4,8598
21	Z	8,3822	28,6558	20,2736	14,9112	13,4148	4,8480

Šaltinis: sudaryta autoriaus.



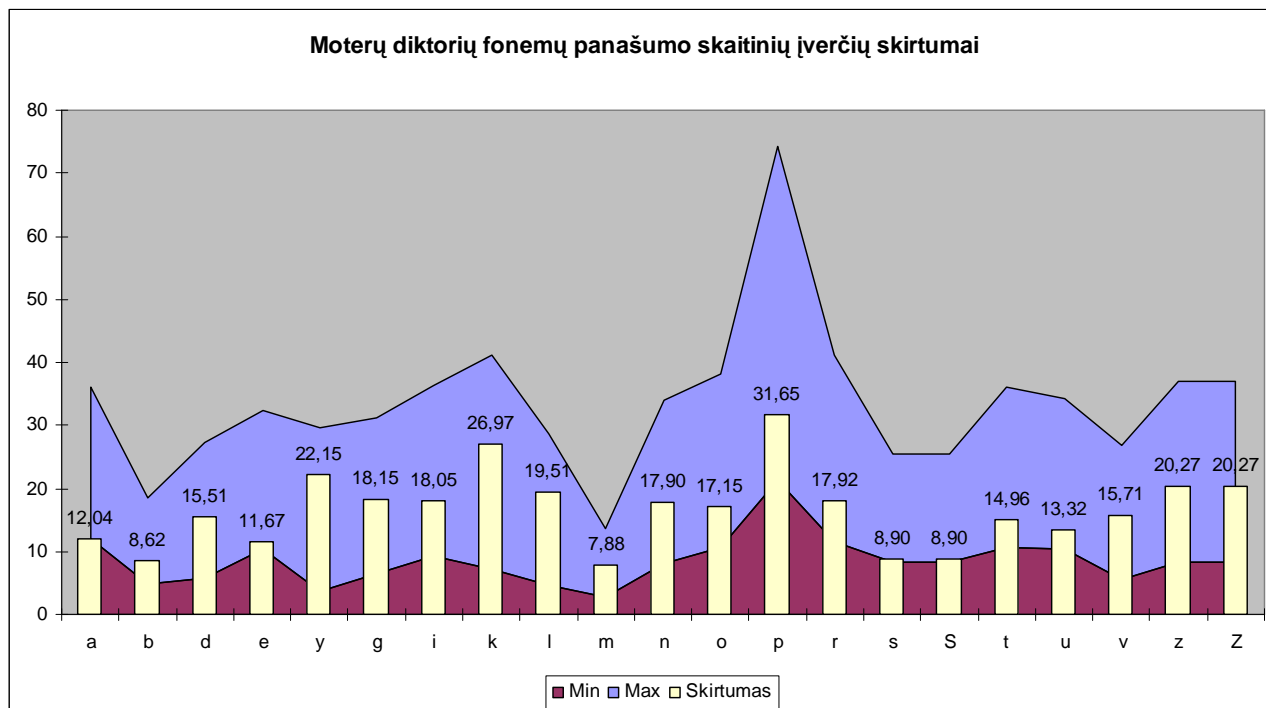
Šaltinis: sudaryta autoriaus.

22 pav. Diktorių moterų fonemų „a“, „b“, „d“, „e“, „r“ ir „t“ skaitinių įverčių histogramos

Palyginus šių kelių fonemų vyrų ir moterų histogramas matome, kad įvairių fonemų statistikos skiriasi. Fonemos „a“ vyrų diktorių skaitiniai įverčiai yra pasiskirstę beveik pagal normalųjį skirstinį, tuo tarpu moterų fiktorių fonema „a“ neturi tokio gero pasiskirstymo – svoris perkeltas į priekinę dalį, be to, yra viena reikšmė intervale 24 – 25, kuri nepatenka į normaliojo skirstinio pasiskirstymą. Tą patį galima pasakyti ir apie fonemos „e“ skaitinių įverčių skirtumus tarp skirtingų lyčių diktorių. Priešinga situacija yra su fonema „b“, kur vyrų diktorių svoris

persikėlęs į priekį, o galiniuose intervaluose yra vienas įvertis, nepatenkantis į normaliojo skirstinio ribas. Tuo tarpu moterų diktorių fonemos „b“ skaitiniai įverčiai yra arčiau normaliojo skirstinio, nors nėra visiškai taisyklingas normalusis skirstinys. Fonemos „r“ skaitiniai įverčiai moterų diktorių tarpe yra geriau pasiskirstę nei diktorių vyrų tarpe, tačiau vyrų diktorių tarpe skaitiniai įverčiai yra labiau išsisklaidę nei moterų diktorių.

Analogiškai diktoriams vyrams, braižomas 11 lentelės rezultatų grafikas.



Šaltinis: sudaryta autoriaus.

23 pav. Moterų diktorių fonemų panašumo skaitinių įverčių minimalių, maksimalių ir skirtumų grafiškas atvaizdavimas

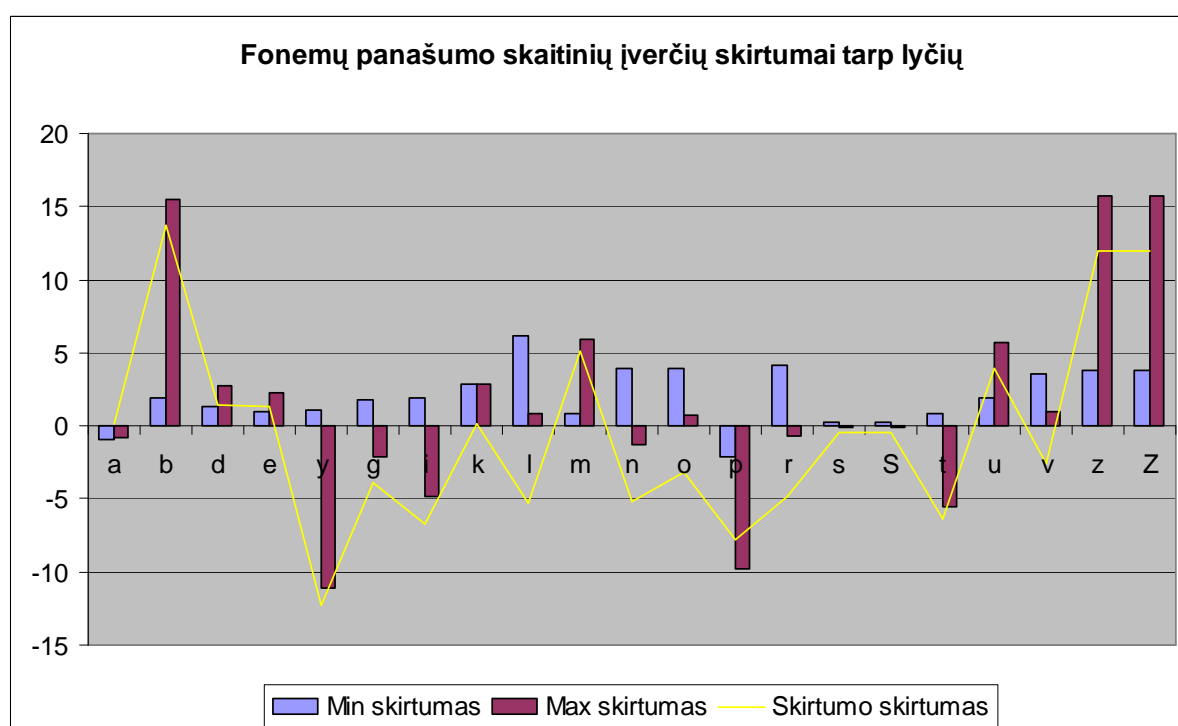
Apskaičiuojami atsikirų fonemų panašumų įverčių skirtumai tarp lyčių: minimalių reikšmių skirtumai, maksimalių reikšmių skirtumai ir skirtumų skirtumai. Skaičiavimų rezultatai pateikiami 12 lentelėje, grafiniai rezultatai – 24 paveiksle.

12 lentelė. Fonemų įverčių skirtumai lyčių atžvilgiu

Males - Females				
Eil. nr.	Fonema	Min skirtumas	Max skirtumas	Skirtumo skirtumas
1	a	-0,9827	-0,8169	0,1658
2	b	1,8614	15,5428	13,6814
3	d	1,3484	2,7681	1,4197
4	e	0,9543	2,2251	1,2708
5	y	1,1289	-11,1209	-12,2498
6	g	1,7832	-2,1376	-3,9209
7	i	1,8628	-4,8811	-6,7439
8	k	2,8119	2,8930	0,0811
9	l	6,1721	0,8615	-5,3107
10	m	0,8392	5,9341	5,0950

Eil. nr.	Fonema	Min skirtumas	Max skirtumas	Skirtumo skirtumas
11	n	3,8906	-1,2420	-5,1326
12	o	3,9176	0,7618	-3,1558
13	p	-2,0561	-9,8190	-7,7628
14	r	4,1577	-0,6995	-4,8571
15	s	0,2923	-0,1522	-0,4445
16	S	0,2923	-0,1522	-0,4445
17	t	0,7898	-5,5980	-6,3878
18	u	1,8506	5,7506	3,9000
19	v	3,5912	0,9919	-2,5994
20	z	3,7784	15,7202	11,9418
21	Z	3,7784	15,7202	11,9418

Šaltinis: sukurta autoriaus.



Šaltinis: sudaryta autoriaus.

24 pav. Fonemų skaitinių įverčių skirtumai tarp lyčių.

Taigi, apibendrinant pradinis rezultatus, galima teigti, kad stebimi akivaizdūs fonemų tarpusavio panašumų įverčių diktoriui atpažinti skirtumai, fonemų charakteristikos ir skaitiniai įverčiai skiriasi fonemų tarpe, lyginant fonemas pagal diktoriaus lytį taip pat gaunami ryškūs skirtumai – skirtingas fonemas įvairūs diktoriai taria nevienodai, o tai suteikia pagrindą pradinei hipotezei, kad diktorių galima atpažinti iš tam tikrų jo išstartų fonemų geriau nei iš likusiųjų.

Kadangi nustatyta, kad ne visos fonemos turi vienodą reikšmę identifikuojant diktorių pagal jas, analizuojant rezultatus toliau, bandyta nustatyti, kokios fonemos iš pagal reikšmingumą mažėjančiai išsidėsčiusių fonemų eilės būtų labiausiai tinkamos slaptažodinių frazių sudarymui.

Tuo tikslu atskirai vyrų diktorių fonemų įvertinimo rezultatai ir atskirai moterų buvo išrikiuoti mažėjimo tvarka pagal fonemų panašumo įvertį (galimą reikšmingumą diktoriaus atpažinimui). Dalis rezultatų, kuriuose akivaizdžiai matosi skirtingų fonemų reikšmingumo eilės vietų (eilučių) skirtumai, pateikiama žemiau esančiose lentelėse: 11 lentelėje pateikiami vyrų diktorių, o 12 lentelėje – moterų diktorių fonemų reikšmingumo eilės vietų (eilučių) skirtumai.

13 lentelė. Vyrų diktorių fonemų skaitiniai įverčiai išrikiuoti pagal tinkamumo eilę

Eil. nr.	M002		M008		M013		M023		M028	
1	29,1296	p	36,96834	k	29,61484	p	31,88542	p	30,16123	p
2	23,29684	r	33,34162	p	29,1126	b	27,12028	r	28,11827	z
3	22,91921	k	22,99671	r	24,84349	r	23,16146	o	28,11827	Z
4	22,52673	o	22,16069	o	22,59198	k	21,2277	u	25,4653	r
5	22,22978	v	20,52475	z	18,759	z	20,37788	a	24,05238	o
6	20,35496	u	20,52475	Z	18,759	Z	20,09226	n	23,48984	e
7	19,08719	l	19,46336	t	18,15716	g	19,68587	i	22,6567	a
8	18,67788	d	17,15463	l	16,28994	t	19,12428	z	22,34784	i
9	18,09733	e	17,02605	v	14,45937	o	19,12428	Z	21,86446	u
10	17,58858	i	16,68999	u	14,25861	n	18,68567	k	20,69414	l
11	17,48615	z	16,07843	i	14,10337	v	18,64787	e	20,08812	v
12	17,48615	Z	14,55847	n	12,35732	u	18,21184	b	17,38432	n
13	17,36112	n	13,84142	e	11,9306	i	17,8989	v	16,96902	k
14	17,31359	a	13,19606	a	11,26055	e	17,8357	l	15,06089	t
15	15,17049	g	12,78399	g	10,98266	a	17,02168	g	13,21088	s
16	15,06326	b	11,16271	b	10,79816	s	14,97718	t	13,21088	S
17	14,61747	t	10,86668	s	10,79816	S	14,7042	y	12,89128	g
18	13,46503	s	10,86668	S	10,7606	l	13,80936	d	11,06322	d
19	13,46503	S	10,1835	d	9,631016	d	12,42245	s	10,39573	b
20	9,00372	y	9,152628	y	7,002555	m	12,42245	S	9,649149	y
21	8,450124	m	5,057821	m	4,807655	y	6,767811	m	8,035412	m

Šaltinis: sudaryta autoriaus.

14 lentelė. Moterų diktorių fonemų skaitiniai įverčiai išrikiuoti pagal tinkamumo eilę

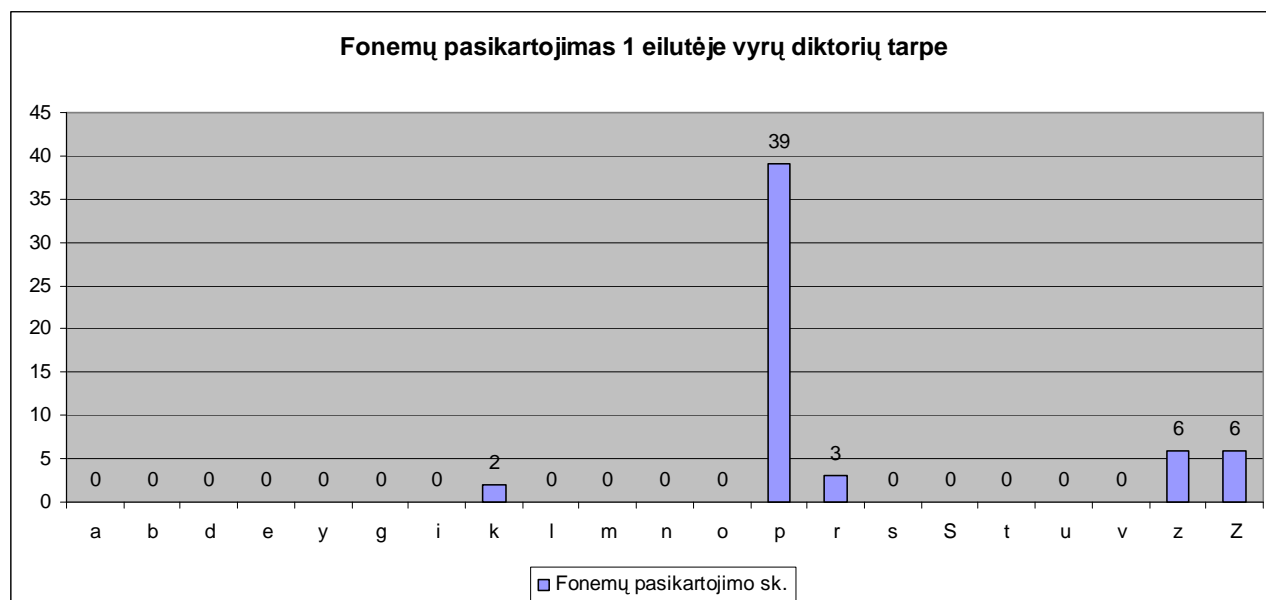
Eil. nr.	F004		F011		F024		F026		F040	
1	29,84405	p	35,48456	p	43,17544	p	47,42955	p	33,84293	p
2	28,65581	z	26,07521	k	24,06832	k	22,7878	r	25,21567	k
3	28,65581	Z	24,29757	r	23,08024	z	21,33024	o	18,13621	z
4	25,39437	k	15,91742	t	23,08024	Z	19,8016	k	18,13621	Z
5	17,74013	u	15,629	v	22,85777	r	18,93052	u	17,87496	o
6	16,15321	r	13,0987	a	16,29425	t	16,09048	a	17,76681	t
7	16,1122	o	12,87116	o	16,26825	a	16,01494	d	17,73119	a
8	15,09745	t	12,82693	e	15,16693	u	15,58575	i	17,48341	r
9	14,21013	i	11,83723	s	14,68429	l	14,91601	e	17,40122	u
10	13,40038	e	11,83723	S	14,21399	o	14,41494	v	17,17817	l
11	13,24239	g	11,82662	l	13,60807	e	14,3235	l	16,87894	i
12	12,94804	v	11,81034	z	13,34944	i	13,16922	n	15,94151	s
13	12,01885	a	11,81034	Z	13,3396	v	12,35842	t	15,94151	S
14	10,29612	n	11,65254	i	11,77141	g	11,57323	b	14,2315	n
15	9,709476	s	11,32315	u	11,70281	b	10,84453	g	13,71489	g
16	9,709476	S	10,93972	n	11,54794	n	10,48633	z	13,47857	v
17	9,49592	l	10,41067	g	11,50494	s	10,48633	Z	13,00399	e

Eil. nr.	F004		F011		F024		F026		F040	
18	9,422276	b	10,09776	y	11,50494	S	9,787212	y	10,26741	d
19	8,94383	d	7,946329	d	9,998429	y	9,276783	s	8,740716	y
20	7,238338	y	5,829775	b	9,105633	d	9,276783	S	5,421475	b
21	5,346261	m	5,470603	m	4,494494	m	4,728649	m	3,493473	m

Šaltinis: sudaryta autoriaus.

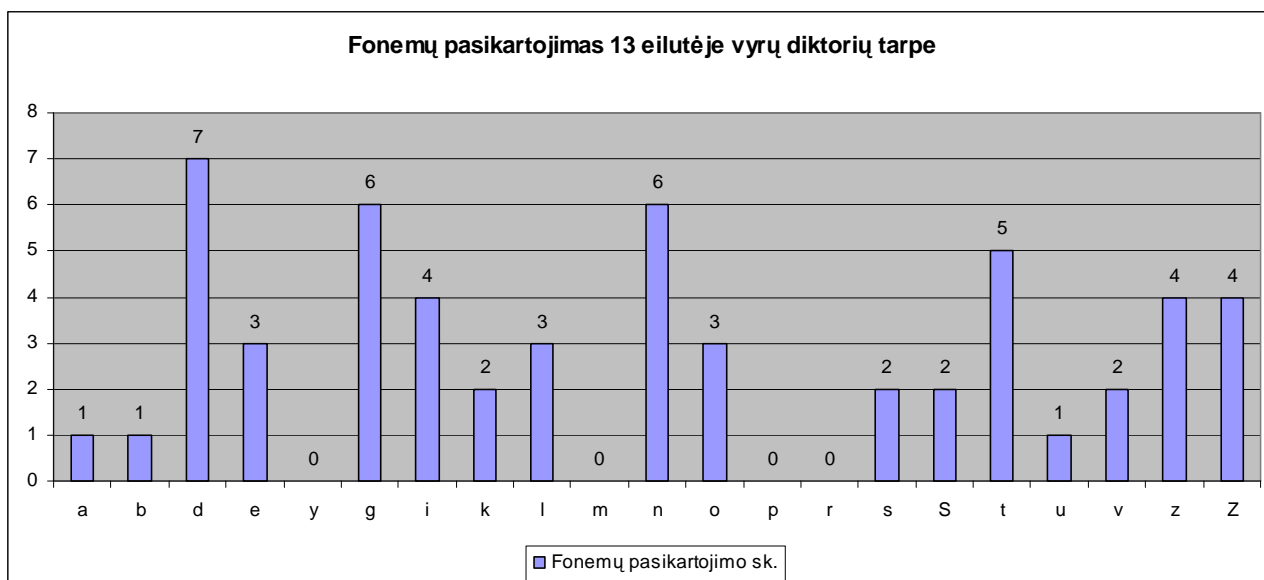
Visi išrikiuoti diktorių fonetinio panašumo skaitinių įverčių duomenys pateikiami 11 ir 12 prieduose.

Norint įvertinti, kurios fonemos geriausiai identifikuoja diktorių, papildomai nutarta įvertinti fonemų pasikartojimo dažnį. Skaičiuojamas kiekvienoje eilutėje esančių fonemų pasikartojimų skaičius, braižomos pasikartojimų skaičiaus diagramos. Diktorių vyrų pirmos atpažinimo dažnių eilės eilutėje dažniausiai pasikartojančios fonemos pavaizduotos 25 paveiksle, tryliktoje eilutėje – 26 paveiksle, o septynioliktoje eilutėje – 27 paveiksle.



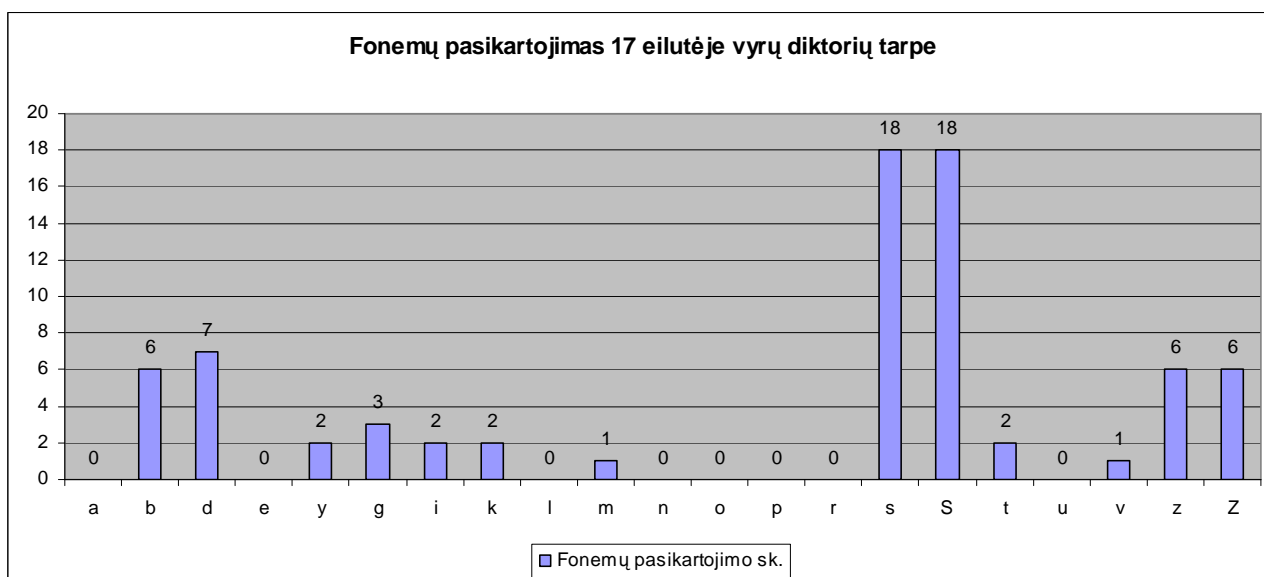
Šaltinis: sudaryta autoriaus.

25 pav. 1 eilutėje dažniausiai pasitaikančios fonemos pagal panašumų įverčius



Šaltinis: sudaryta autoriaus.

26 pav. 13 eilutėje dažniausiai pasitaikančios fonemos pagal panašumų įverčius



Šaltinis: sudaryta autoriaus.

27 pav. 17 eilutėje dažniausiai pasitaikančios fonemos pagal panašumų įverčius

Apibendrinus vyrų diktorių fonemų skaitines reikšmes ir išrikiavus fonemas pagal jų tinkamumą diktoriaus atpažinime, dažniausiai kiekvienoje eilutėje aptinkamos fonemos ir jų dažniai pateikti 15 lentelėje.

15 lentelė. Dažniausiai diktorių vyrų atpažinimo eilės eilutėse pasikartojančios fonemos

Eil. Nr.	Max pasikartojimas eilutėje	Fonema
1	39	p
2	14	r, z, Z
3	11	r
4	11	o
5	12	u
6	7	l

Eil. Nr.	Max pasikartojimas eilutėje	Fonema
7	8	a
8	10	z, Z
9	7	v, z, Z
10	9	a
11	8	n
12	7	n
13	7	d
14	6	a, v
15	9	s, S
16	14	s, S
17	18	s, S
18	26	s, S
19	21	s, S
20	24	y
21	41	m

Šaltinis: sudaryta autoriaus.

Paskaičiuojama, kiek yra skirtingų pasikartojančių fonemų kiekvienoje eilutėje, kiek iš jų – dominuojančių. Dominuojančia fonema laikyta fonema, kurios pasikartojimų skaičius yra didesnis nei pusė daugiausiai pasikartojančios ir mažiausiai pasikartojančios (tačiau bent kartą pasikartojančios) fonemos skirtumo. Skaičiavimų rezultatai pateikiami 16 lentelėje.

16 lentelė. Skirtingų bei dominuojančių fonemų skaičius ir vidutinis fonemos pasikartojimas diktoriaus atpažinimo tinkamumo eilėje

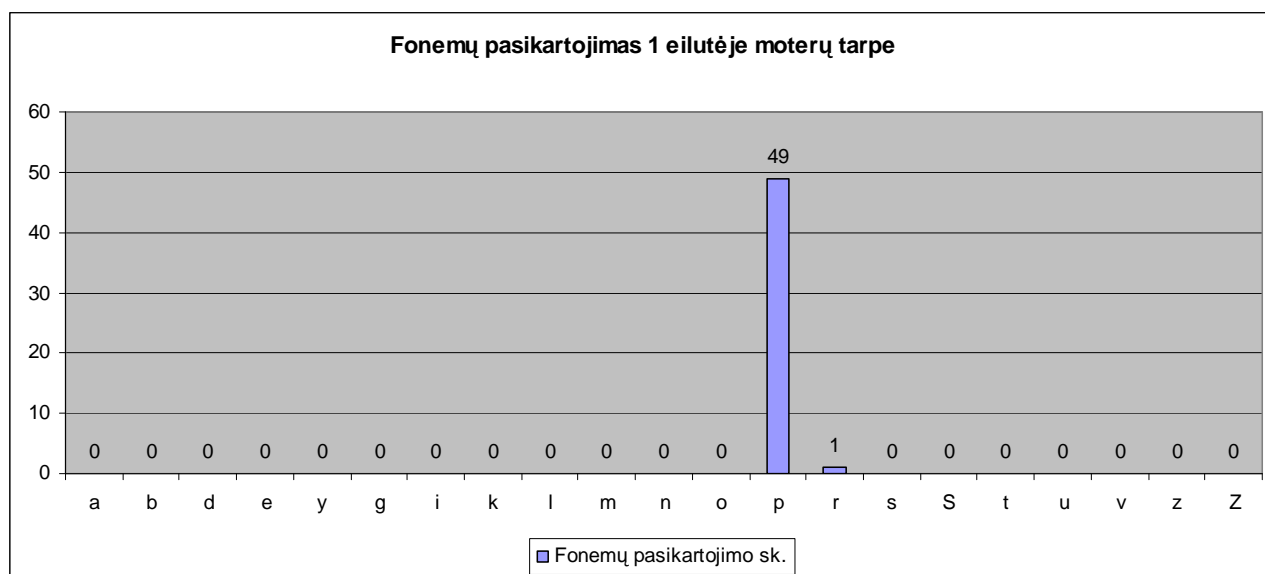
Eil. Nr.	Skirtingų fonemų skaičius	Dominuojančių skaičius	Vidutinis fonemos pasikartojimų
1	5	1	11,20
2	8	4	8,00
3	13	5	4,62
4	15	2	3,60
5	14	5	4,00
6	14	8	3,86
7	15	7	3,60
8	15	5	4,00
9	15	8	3,80
10	13	5	4,00
11	14	8	3,86
12	15	7	3,67
13	17	7	3,29
14	16	11	3,75
15	17	6	3,65
16	16	3	4,19
17	13	2	5,69
18	12	2	6,58
19	9	3	7,89
20	7	2	7,86
21	3	1	16,67

Šaltinis: sudaryta autoriaus.

Matome, kad daugiausiai dominuojančių fonemų yra atpažinimo tinkamumo eilės (fonemų panašumo skaitinių įverčių aibės) viduryje, o pirmame ir ketvirtame kvartiliuose ryškios po vieną dvi dominuojančios fonemos vyrų diktorių tarpe. Įvertinus skirtingų fonemų skaičių, matome tiesioginę priklausomybę tarp skirtingų fonemų skaičiaus eilutėje bei dominuojančių fonemų skaičiaus. Kita vertus, įvertinus antrojo ir trečiojo stulpelio reikšmes, galima pastebėti, jog reikšmės eilutėse yra paskirste pagal normalųjį skirstinį. Iš viso to seka, kad preliminariai galima daryti išvadą, jog į pirmąjį ir ketvirtąjį kvartilius patenkančios reikšmės netinka diktorių vyrų klasifikavimui ir atpažinimui pagal fonemas, nes:

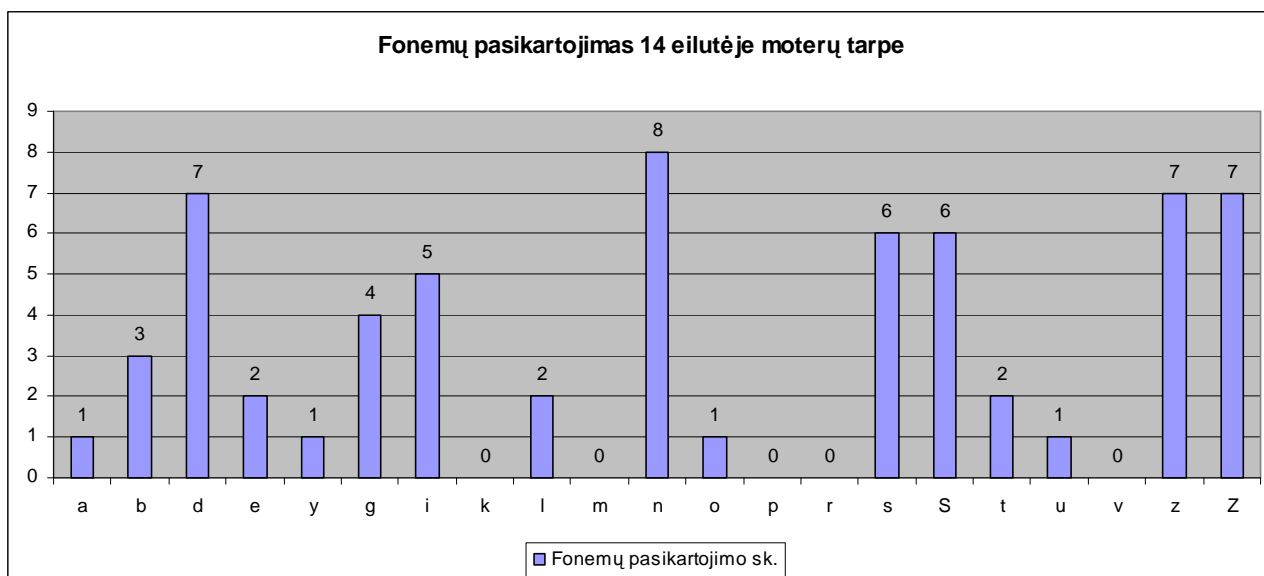
- Pirmojo kvartilio atžvilgiu – pasikartojančios fonemos per daug panašios ir per daug dominuojančios visoje tiriamų diktorių imtyje.
- Ketvirtojo kvartilio atžvilgiu – fonemos turi per daug silpnus skaitinius įverčius, kad tiktų diktoriaus atpažinimui, vėl pastebimas didelis vienos dviejų dominuojančių fonemų pasikartojimas.

Analogiškai paskaičiuojame fonemų pasikartojimus moterų diktorių tarpe. Skaičiavimų rezultatai, analogiškai kaip ir diktorių vyrų, pateikiami 28, 29 ir 30 paveiksluose bei 16 ir 17 lentelėse.



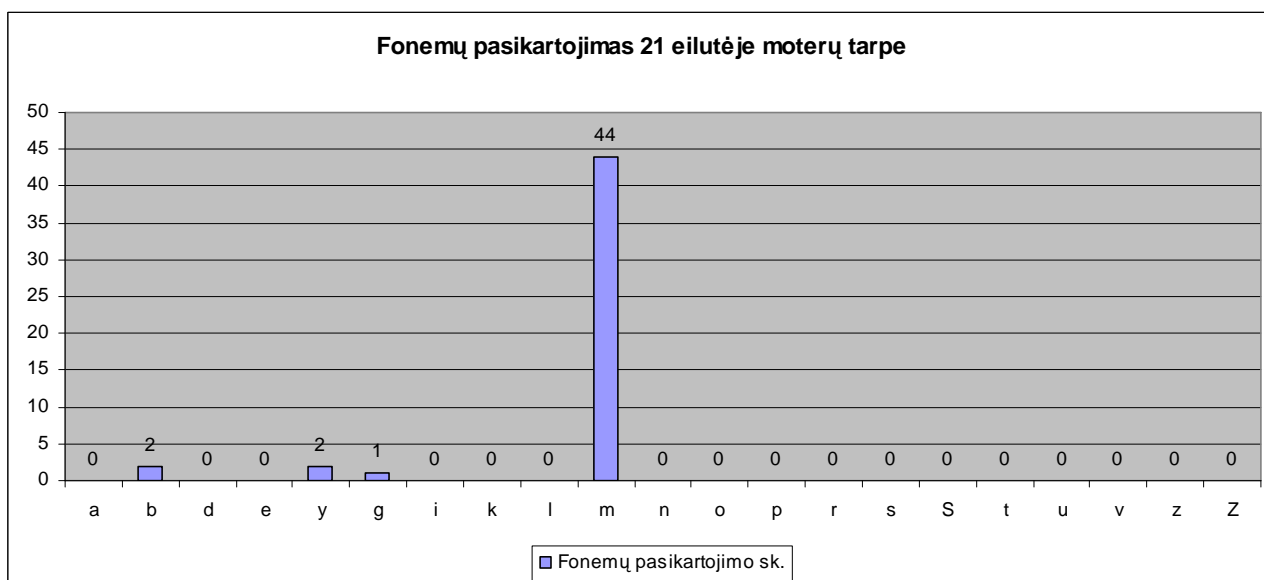
Šaltinis: sudaryta autoriaus.

28 pav. 1 eilutėje dažniausiai pasitaikančios fonemos pagal panašumų įverčius



Šaltinis: sudaryta autoriaus.

29 pav. 14 eilutėje dažniausiai pasitaikančios fonemos pagal panašumų įverčius



Šaltinis: sudaryta autoriaus.

30 pav. 21 eilutėje dažniausiai pasitaikančios fonemos pagal panašumų įverčius

Fonemos pasikartojimo kiekvienoje eilutėje dažnumas, dominuojančių fonemų skaičius, skirtingų pasikartojimų skaičius:

17 lentelė. Dažniausios fonemos kiekvienoje eilutėje

Eil. Nr.	Max pasikartojimas eilutėje	Fonema
1	49	p
2	19	k
3	15	r
4	9	u, z, Z,
5	12	u
6	9	t
7	9	l
8	8	e

Eil. Nr.	Max pasikartojimas eilutėje	Fonema
9	10	a
10	7	a, z, Z
11	8	e, i
12	12	i
13	11	z, Z
14	8	n
15	16	s, S
16	19	s, S
17	14	s, S
18	16	s, S
19	15	y
20	18	b, y
21	44	m

Šaltinis: sudaryta autoriaus.

18 lentelė. Skirtingų bei dominuojančių fonemų skaičius ir vidutinis fonemos pasikartojimas diktoriaus atpažinimo tinkamumo eilėje

Eil. Nr.	Skirtingų fonemų skaičius	Dominuojančių skaičius	Vidutinis fonemos pasikartojimų
1	2	1	25,0
2	10	2	5,4
3	10	4	5,9
4	11	7	5,4
5	11	2	5,0
6	13	4	4,1
7	13	5	4,2
8	14	8	4,0
9	16	4	3,8
10	16	8	3,8
11	16	7	3,4
12	15	5	4,1
13	14	6	4,9
14	16	8	3,9
15	16	3	4,4
16	15	2	4,9
17	14	3	4,9
18	14	3	4,9
19	12	4	5,0
20	8	2	6,4
21	4	1	12,3

Šaltinis: sudaryta autoriaus.

Kaip ir vyrų diktorių atveju, stebima, jog diktorių moterų tarpe gaunami panašūs rezultatai: pagal Gauso normalųjį skirstinį pasiskirstę skirtingų bei dominuojančių fonemų kiekiai visoje diktoriaus fonemų eilėje, vėl stebima tiesioginė priklausomybė tarp skirtingų fonemų skaičiaus eilėje, dominuojančių fonemų skaičiaus bei vidutinio fonemų pasikartojimo.

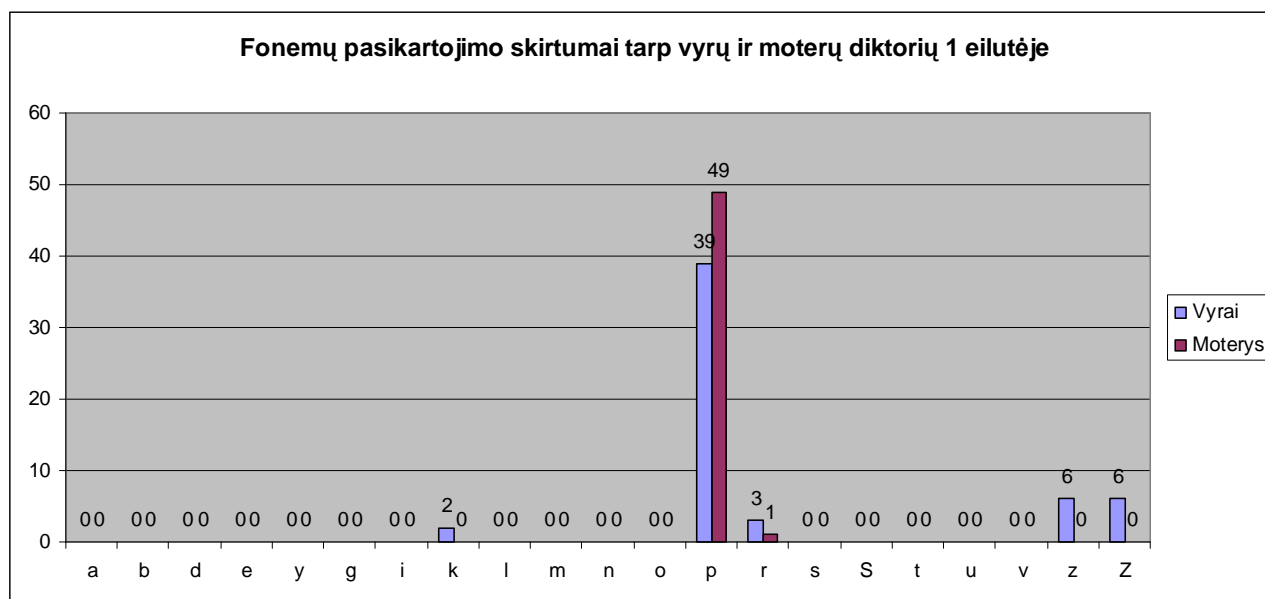
Sutraukus 16 – 18 lentelių duomenis į vieną lentelę, rezultatų tarp abiejų lyčių gaunama tokia duomenų lentelė:

19 lentelė. Fonemų charakteristikų diktorių vyrų ir moterų atpažinimo tinkamumo eilėje skirtumai

	Eil. Nr.	Max count	Phoneme	Fonemu sk.	Fonemos pasikartojimo vidurkis	Dominuojančių skaičius
Vyrai	1	39	p	5	11,20	1
Moterys	1	49	p	2	25,00	1
Vyrai	2	14	r, z, Z	8	8,00	4
Moterys	2	19	k	10	5,40	2
Vyrai	3	11	r	13	4,62	5
Moterys	3	15	r	10	5,90	4
Vyrai	4	11	o	15	3,60	2
Moterys	4	9	u, z, Z,	11	5,36	7
Vyrai	5	12	u	14	4,00	5
Moterys	5	12	u	11	5,00	2
Vyrai	6	7	l	14	3,86	8
Moterys	6	9	t	13	4,08	4
Vyrai	7	8	a	15	3,60	7
Moterys	7	9	l	13	4,15	5
Vyrai	8	10	z,Z	15	4,00	5
Moterys	8	8	e	14	4,00	8
Vyrai	9	7	v, z,Z	15	3,80	8
Moterys	9	10	a	16	3,75	4
Vyrai	10	9	a	13	4,00	5
Moterys	10	7	a, z, Z	16	3,75	8
Vyrai	11	8	n	14	3,86	8
Moterys	11	8	e, i	16	3,44	7
Vyrai	12	7	n	15	3,67	7
Moterys	12	12	i	15	4,13	5
Vyrai	13	7	d	17	3,29	7
Moterys	13	11	z, Z	14	4,86	6
Vyrai	14	6	a, v	16	3,75	11
Moterys	14	8	n	16	3,94	8
Vyrai	15	9	s, S	17	3,65	6
Moterys	15	16	s,S	16	4,38	3
Vyrai	16	14	s, S	16	4,19	3
Moterys	16	19	s,S	15	4,87	2
Vyrai	17	18	s, S	13	5,69	2
Moterys	17	14	s,S	14	4,93	3
Vyrai	18	26	s, S	12	6,58	2
Moterys	18	16	s,S	14	4,93	3
Vyrai	19	21	s, S	9	7,89	3
Moterys	19	15	y	12	5,00	4
Vyrai	20	24	y	7	7,86	2
Moterys	20	18	b, y	8	6,38	2
Vyrai	21	41	m	3	16,67	1
Moterys	21	44	m	4	12,25	1

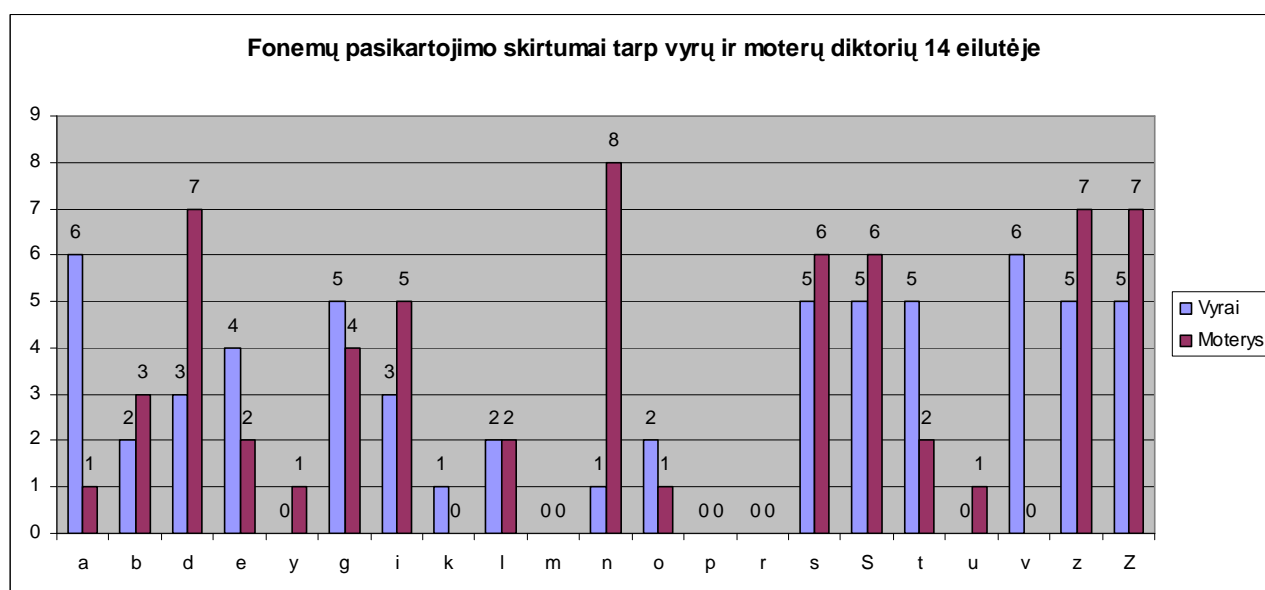
Šaltinis: sudaryta autoriaus.

Grafiniai atsitiktinai parinktų fonemų lentelėje esantys rezultatai pateikiami 31, 32 ir 33 paveiksluose.



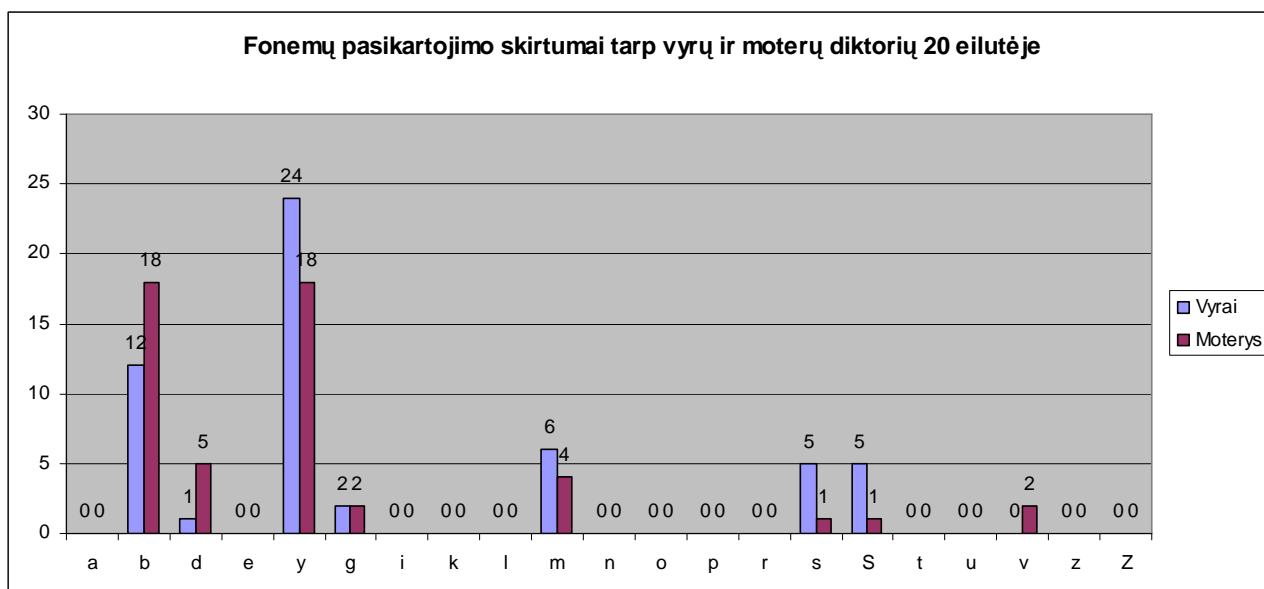
Šaltinis: sudaryta autoriaus.

31 pav. Diktorių vyrų ir moterų fonemų pasikartojimo skirtumai 1 eilutėje



Šaltinis: sudaryta autoriaus.

32 pav. Diktorių vyrų ir moterų fonemų pasikartojimo skirtumai 14 eilutėje

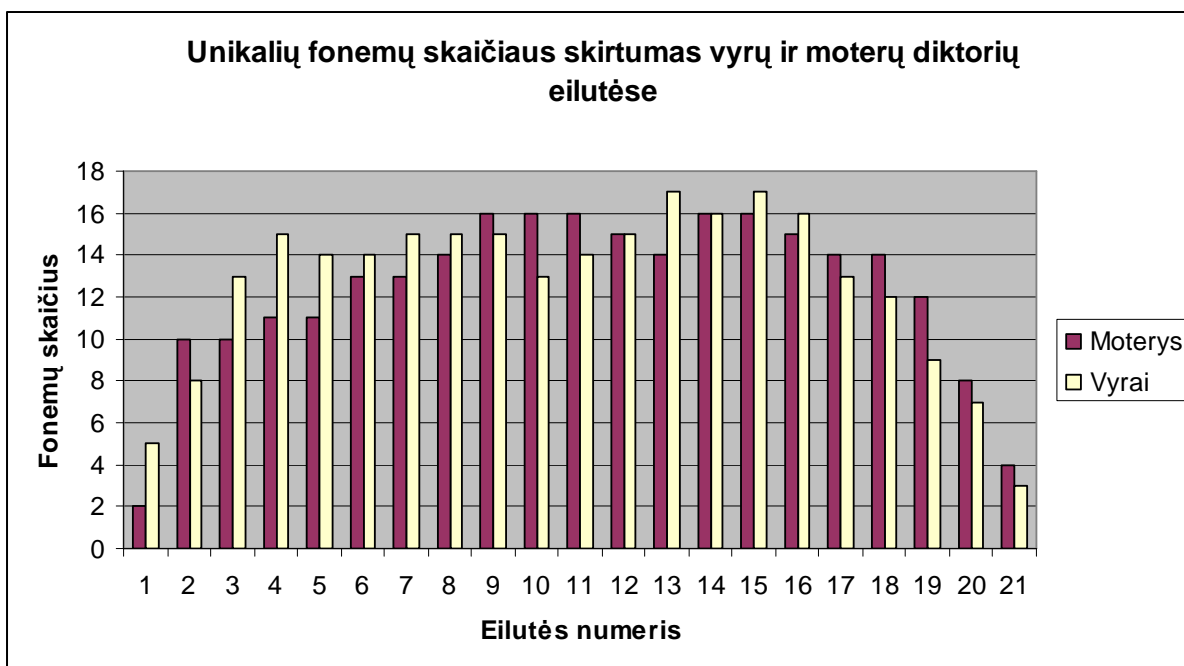


Šaltinis: sudaryta autoriaus.

33 pav. Diktorių vyrų ir moterų fonemų pasikartojimo skirtumai 20 eilutėje

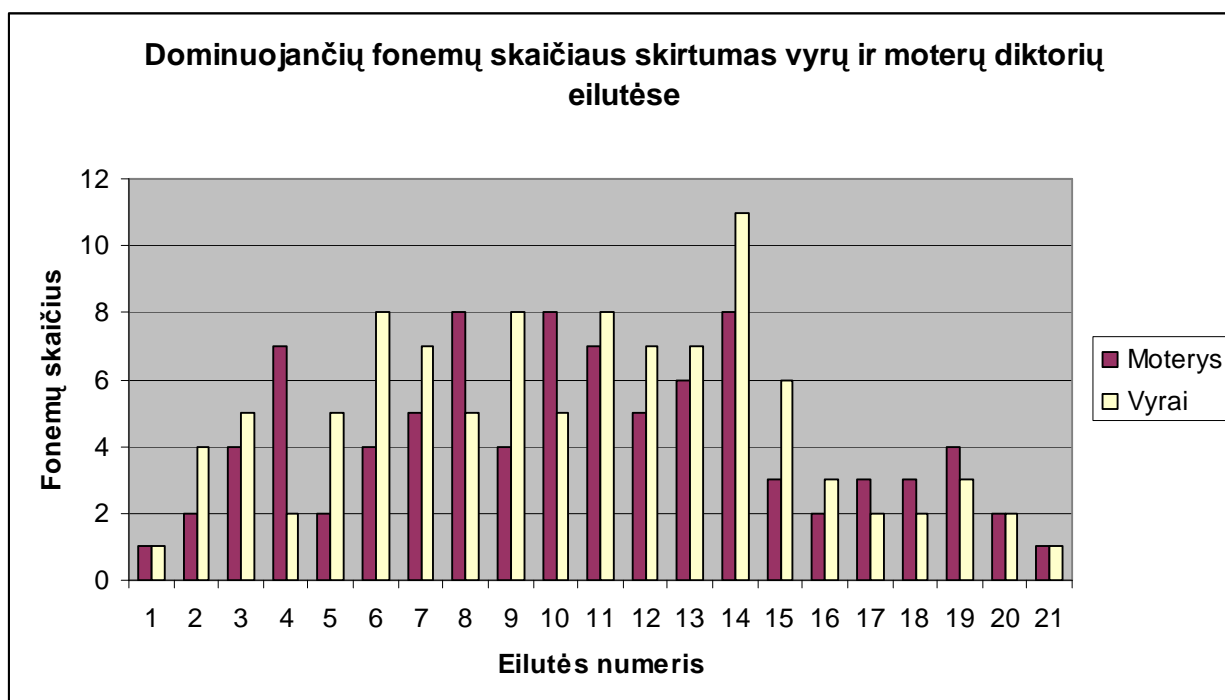
Kaip matome iš grafikų, skirtingose eilutėse vyrų ir moterų besikartojančios fonemos skiriasi, taip pat skiriasi pasikartojimų skaičiai, bei matomi fonemų pasikartojimų skirtumai tarp lyčių. Be to, galima pastebėti, kad kai kurių fonemų pasikartojimų skirtumai tarp diktorių lyčių yra gana reikšmingi (pvz. fonema „n“ 14 eilutėje). Daugiau šių rezultatų diagramų pavyzdžių pateikiama 13 priede.

Be to, kad geriau būtų matomi skirtumai, atskirai braižome unikalių fonemų skaičiaus skirtumų vyrų ir moterų diktorių eilutėse (34 pav.) bei dominuojančių fonemų skaičiaus skirtumų vyrų ir moterų diktorių eilutėse (35 pav.) diagramas.



Šaltinis: sudaryta autoriaus.

34 pav. Unikalių fonemų skaičiaus skirtumų vyrų ir moterų diktorių eilutėse diagrama



Šaltinis: sudaryta autoriaus.

35 pav. Dominuojančių fonemų skaičiaus skirtumų vyrų ir moterų diktorių eilutėse diagrama

Iš dviejų paskutiniųjų diagramų matome, kad tiek dominuojančių fonemų pasikartojimo skirtumai tarp diktorių vyrų ir diktorių moterų nėra labai ryškūs, yra beveik pasiskirstę pagal normalųjį skirstinį. Moterų diktorių dominuojančių fonemų skaičius daugiau nei dviem fonemomis lenkia vyrus 4, 8 ir 10 eilutėse, o vyrai diktoriai lenkia moteris dominuojančiomis fonemomis atitinkamai 5, 9, 14 ir 15 eilutėse. Toks dominuojančių fonemų pasiskirstymas irgi gali turėti įtakos diktoriaus atpažinimui pagal jam būdingų fonemų rinkinį arba nustatant diktoriui būdingų fonemų aibę.

Papildomai paskaičiuojama, kaip dominuojančių fonemų skaičius įtakoja fonemų pasikartojimą diktorių moterų ir diktorių vyrų aibėse, bei kaip kinta bendras vidurkis.

20 lentelė. Dominuojančių fonemų skaičiaus įtaka fonemų pasikartojimui

Fonemos pasikartojimo vidurkio priklausomybė nuo lyties ir fonemos pasikartojimų kiekio			
Dominuojančių skaičius	Moterys	Vyrai	Bendras vidurkis
1	18,625	13,933	16,279
2	5,410	5,933	5,672
3	4,744	6,038	5,262
4	4,682	8,000	5,345
5	4,144	4,154	4,150
6	4,857	3,647	4,252
7	4,401	3,520	3,872
8	3,896	3,838	3,867
11	-	3,750	3,750
Bendras vidurkis:	5,975	5,608	5,792

Šaltinis: sudaryta autoriaus.

Matome, kad bendru atveju kuo dažniau fonema pasikartoja eilutėje, tuo didesnis pasikartojimų vidurkis kiekvienoje diktorių populiacijoje ir bendrai. Galima daryti išvadą, kad jokios rezultatus keičiančios informacijos negavome.

Remiantis atlikta rezultatų duomenų analize, galime teigti, kad svarbu, kurioje diktoriaus atpažinimo fonemų eilės vietoje yra fonema ir kaip jos pasiskirsto tiriant skirtingus diktorius.

Galutis rezultatų įvertinimas atliekamas paskaičiavus kaip dažnai kiekvienoje eilutėje dažniausiai pasikartojanti fonema (ar kelios pasikartojančios fonemos, jei jų yra daugiau nei viena) sutampa su diktorių fonemomis atpažinimo eilėje. Tyrimo logikai iliustruoti naudojami diktorių M002, M008, M013, M023 ir M028 fonemų eilės bei dažniausiai pasikartojančios fonemos. Šie duomenys pateikti 21 lentelėje.

21 lentelė. Diktorių duomenys ir dažniausiai eilutėse pasikartojančios fonemos

Dažniausios fonemos			Eil. Nr	M002	M008	M013	M023	M028
p			1	p	k	p	p	p
r	z	Z	2	r	p	b	r	z
r			3	k	r	r	o	Z
o			4	o	o	k	u	r
u			5	v	z	z	a	o
l			6	u	Z	Z	n	e
a			7	l	t	g	i	a
z	Z		8	d	l	t	z	i
v	z	Z	9	e	v	o	Z	u
a			10	i	u	n	k	l
n			11	z	i	v	e	v
n			12	Z	n	u	b	n
d			13	n	e	i	v	k
a, v	v		14	a	a	e	l	t
s	S		15	g	g	a	g	s
s	S		16	b	b	s	t	S
s	S		17	t	s	S	y	g
s	S		18	s	S	l	d	d
s	S		19	S	d	d	s	b
y			20	y	y	m	S	y
m			21	m	m	y	m	m

Šaltinis: sudaryta autoriaus.

Šio etapo metu tikrinama, ar n-tosios eilutės dažniausiai pasikartojančios fonemos nesutampa su M-tojo diktoriaus toje pačioje eilutėje esančia fonema. Jei sutampa, į rezultatų lentelės atitinkamą vietą yra rašomas „1“, jei ne – „0“. Taip patikrinamos visos 21 fonetinės eilės eilutė. Visa fonetinė eilė yra sudalinama į ketvirčius (statistinius kvartilius, turinčius po 5 – 6 elementus). Tiriamas pasikartojimų pasiskirstymas kvartiliuose. Rezultatai pateikiami 22 lentelėje.

22 lentelė. Fonemų sutapimų su eilės dominuojančiomis fonemomis rezultatai

Eil. nr.		M002	M008	M013	M023	M028
I kvartilis	1	1	0	1	1	1
	2	1	0	0	1	1
	3	0	1	1	0	0
	4	1	1	0	0	0
	5	0	0	0	0	0
II kvartilis	6	0	0	0	0	0
	7	0	0	0	0	1
	8	0	0	0	1	0
	9	0	1	0	1	0
	10	0	0	0	0	0
III kvartilis	11	0	0	0	0	0
	12	0	1	0	0	1
	13	0	0	0	0	0
	14	0	0	0	0	0
	15	0	0	0	0	1
	16	0	0	1	0	1
IV kvartilis	17	0	1	1	0	0
	18	1	1	0	0	0
	19	1	0	0	1	0
	20	1	1	0	0	1
	21	1	1	0	1	1
Nesutampa iš viso		14	13	17	15	13
Nesutampa iš 1 kvartilio		2	3	3	3	3
Nesutampa iš 4 kvartilio		1	1	4	3	3
Nesutampa iš 1 ir 4 kvartilių		3	4	7	6	6
Nesutampa iš 2, 3 kvartilių		11	9	10	9	7
Nesutampa iš 1, 4 kvartilių (%)		21,43%	30,77%	41,18%	40,00%	46,15%
Nesutampa iš 2, 3 kvartilių (%)		78,57%	69,23%	58,82%	60,00%	53,85%
Skirtumai		57,14%	38,46%	17,65%	20,00%	7,69%

Šaltinis: sudaryta autoriaus.

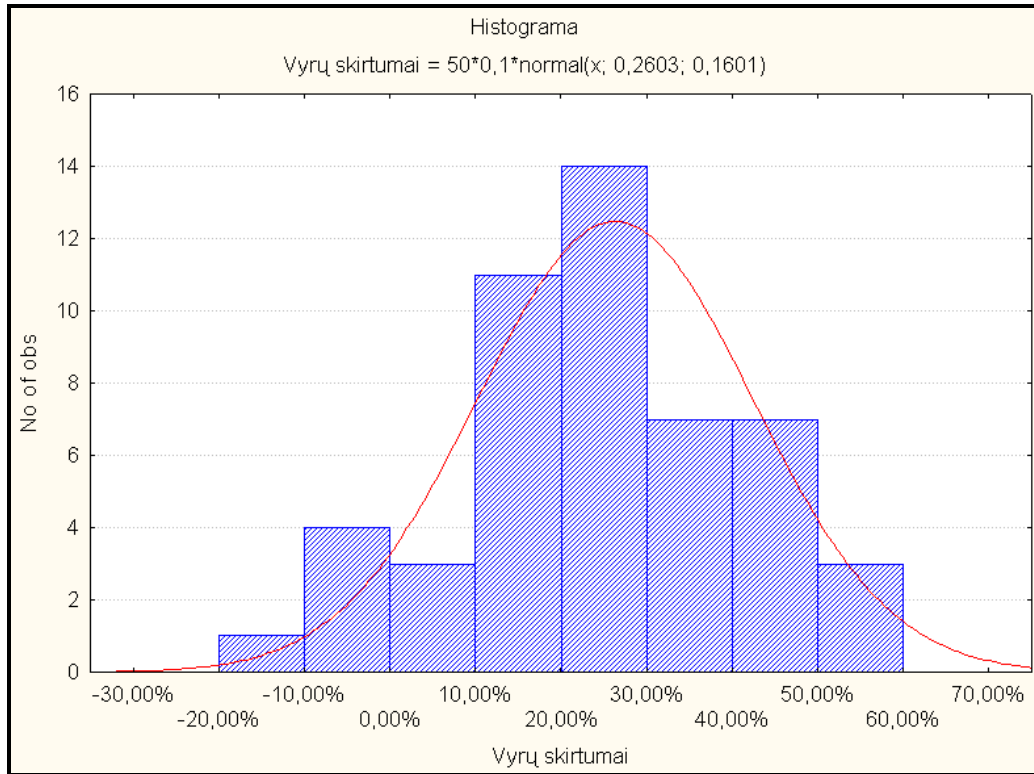
Analogiškai 22 lentelėje pateiktam vyrų diktorių fonemų atitikimo eilutėse įvertinimui, buvo skaičiuojami ir diktorių moterų fonemų atitikimo eilutėse įvertinimai. Įvertinus visų 50 diktorių vyrų ir visų 50 diktorių moterų fonemų pasikartojimų rezultatus statistinių tyrimu paketu Statistica, sudaryta rezultatus apibendrinanti lentelė (23 lentelė).

23 lentelė. Rezultatų apibendrinimo lentelė

	Vyrai	Moterys
Vidurkis	26,03%	23,37%
Mediana	26,79%	23,08%
Standartinis nuokrypis	16,01%	18,26%
Minimali reikšmė	-14,29%	-23,08%
Maksimali reikšmė	57,14%	66,67%
25%	0,166667	0,090909
75%	0,384615	0,333333

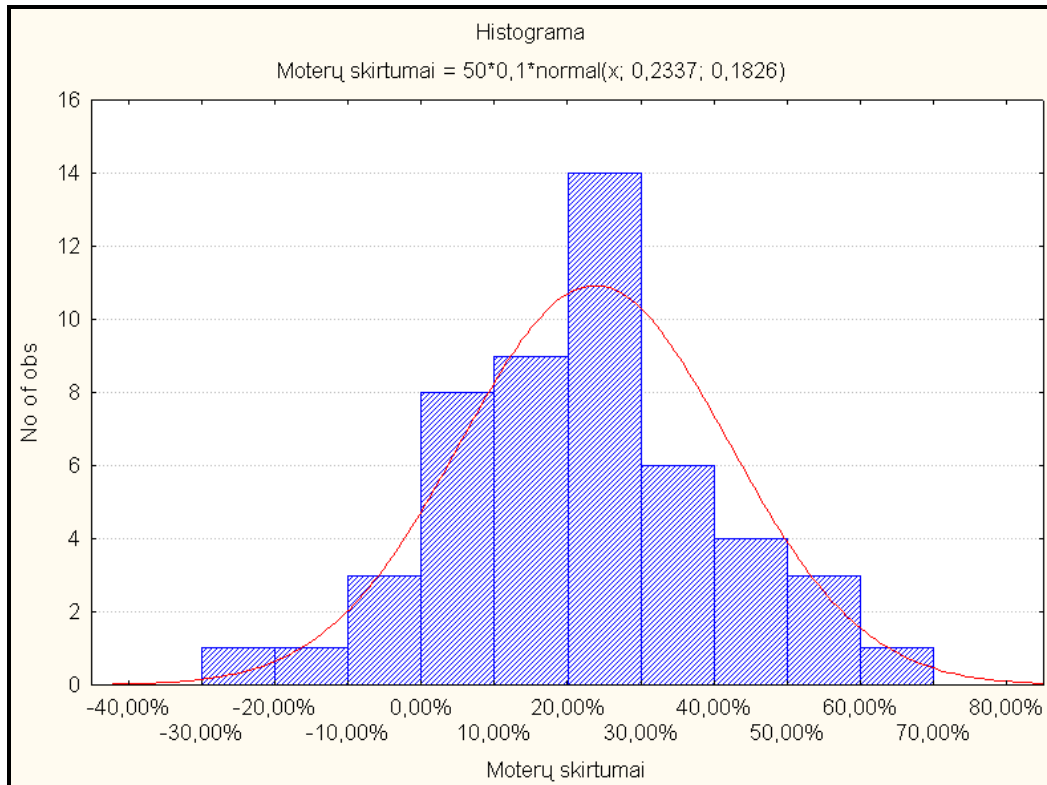
Šaltinis: sudaryta autoriaus.

Pateikiamos diktorių vyrų ir diktorių moterų nesutapimų skirtinguose kvartiliuose skirtumų pasiskirstymas imtyje.



Šaltinis: sudaryta autoriaus.

36 pav. Vyrų diktorių I+IV ir II+III kvartilių rezultatų skirtumų histograma



Šaltinis: sudaryta autoriaus.

37 pav. Moterų diktorių I+IV ir II+III kvartilių rezultatų skirtumų histograma

Kaip matome iš rezultatų, pasiūlyta sprendimo metodika, kuomet diktoriaus identifikavimui naudojami antrojo ir trečiojo ketvirčio fonemos bendru atveju turėtų pasiteisinti – atpažinimui naudojant fonemas iš šių ketvirčių, nors ir fonemų skaitiniai įverčiai neturi aukščiausių reikšmių, tačiau šiose imtyse esančios fonemos rečiausiai sutampa su visoje vienos lyties diktorių aibėje dažniausiai pasikartojančiomis fonemomis. Nepaisant to, kad moterų diktorių imtyje fonemos antrame ir trečiame kvartilyje sutapo su bendros aibės dažniausiai pasikartojančiomis fonemomis daugiau nei pirmame ir ketvirtame kvartilyje vieną kartą, o vyrų diktorių – net tris kartus (vienu atveju skirtumas yra 23%), apskritai taikant šį metodą gaunamas 23,37% nesutapimų skirtumų vidurkis moterų diktorių imtyje, ir 26,03% – vyrų diktorių imtyje.

Skaičiuojant įvairius imčių atitikimo santykio variantus buvo pastebėta, kad kuo platesnė fonemų imtis, tuo geresni rezultatai, tačiau pradėdant imtį nuo fonemų tinkamumo eilės vidurio, rezultatai vėl pradeda prastėti.

Atsižvelgiant į visus atliktus apskaičiavimus, fonemų panašumo skaitinių reikšmių transformavimų ir interpretavimų rezultatus galima teigti, kad nors diktoriaus atpažinimui fonemų diskriminavimo metodu gali būti parenkamos iš fonemų skaitinių įverčių eilės vidurio ir šitoks parinkimas turėtų duoti teigiamus rezultatus atpažįstant diktorių pagal slaptažodines frazes, kurių pagrindą sudaro parinktos fonemos.

IŠVADOS IR PASIŪLYMAI

1. Apibendrinus analitinės dalies tyrimus, galima teigti, jog „tai ką žinau“ grupės priemonės, o konkrečiai – slaptažodžiai ir naudotojo asmeniniai identifikavimo numeriai, yra šiandien populiariausia, tačiau praktiškai saugumo priemonėms keliamų kriterijų nebeatitinkanti informacijos apsaugos priemonė.
2. Slaptažodžių nesaugumas pasireiškia dėl keleto faktorių: tai yra priemonė, kurią naudotojas gali pamesti, pamiršti, slaptažodžiai gali būti pavogti arba atspėti pasinaudojus visų simbolių perrinkimo algoritmais arba asmenine slaptažodžio turėtojo informacija, nes dažnai pačių naudotojų neatsakingas slaptažodžių naudojimas pridaro daugiausiai rūpesčių.
3. Geresnis sprendimas – biometrijos priemonių naudojimas. Seniai žinoma, kad biometrinės žmogaus savybės yra unikalios arba smarkiai skiriasi. Įvertinus įvairius biometrinių savybių kriterijus, darbo tema susieta su balso technologijomis ir diktorius atpažinimu. Svarbų vaidmenį šio pasirinkimo procese suvaidino faktas, jog balso technologijos yra perspektyvios siekiant pagerinti natūralią komunikaciją su sumaniosiomis aplinkomis.
4. Apžvelgus pasaulyje šioje srityje atliktus tyrimus, pasiūlytas naujas metodas diktoriui atpažinti. Šio metodo pagalba siekiama įvertinti skirtingų fonemų savybes ir fonemų tinkamumo diktoriui atpažinti. Keliamą mintis, kad galima sudaryti tokį fonemų poaibį, kuris geriausiai tiktų diktoriui išskirti iš kitų diktorių tarpo.
5. Sudaryta metodika eksperimentui atlikti. Naudojant Melo dažnių kepstrinių koeficientų metodą bus atliekamas fonemų požymių išskyrimas ir požymių vektorių sudarymas, o fonemų panašumo ir tinkamumo diktoriui atpažinti skaitiniai įverčiai bus atliekami tiesinio diskriminavimo analizės metodo pagalba.
6. Atlikus eksperimentą ir vertinant jo rezultatus paaiškėjo, kad skirtingos fonemos turi skirtingus panašumo skaitinius įverčius. Taip pat tų pačių fonemų panašumai įverčiai skiriasi keičiantis diktoriui ir diktorius lyčiai. Išrikiavus šiuos įverčius buvo aptikta, kad skirtingų diktorių fonemų eilės yra skirtingos.
7. Įvertinus dažniausiai visos diktorių aibės fonemų eilės eilutėse pasikartojančių fonemų sutapimą su diktorių fonemomis tose eilutėse, gauti rezultatai, kuriuos galima interpretuoti taip: pasiūlyta sprendimo metodika, kuomet diktorius identifikavimui naudojami antrojo ir trečiojo ketvirčio fonemos bendru atveju turėtų pasiteisinti – atpažinimui naudojant fonemas iš šių ketvirčių, nors ir fonemų skaitiniai įverčiai neturi aukščiausių reikšmių, tačiau šiose imtyse esančios fonemos rečiausiai

sutampa su visoje vienos lyties diktorių aibėje dažniausiai pasikartojančiomis fonemomis.

8. Nepaisant to, kad moterų diktorių imtyje fonemos antrame ir trečiame kvartilyje sutapo su bendros aibės dažniausiai pasikartojančiomis fonemomis daugiau nei pirmame ir ketvirtame kvartilyje vieną kartą, o vyrų diktorių – net tris kartus (vienu atveju skirtumas yra 23%), apskritai taikant pasirinktą metodą gaunamas 23,37% nesutapimų skirtumų vidurkis moterų diktorių imtyje, ir 26,03% – vyrų diktorių imtyje.
9. Vertinant šiuos rezultatus galima teigti, kad nors diktorius atpažinimui fonemų diskriminavimo metodu gali būti parenkamos iš fonemų skaitinių įverčių eilės vidurio ir šitoks parinkimas turėtų duoti teigiamus rezultatus atpažįstant diktorių pagal slaptažodines frazes, kurių pagrindą sudaro parinktos fonemos.

Pasiūlymai

1. Atliekant praktinį eksperimentą pastebėta, kad turima lietuviško garsyno duomenų bazė nėra pakankamai didelė – naudojamas tik labai ribotas kiekis žodžių, kas lemia, kad fonemų skaitinių įverčių reikšmės, kurios gautos naudojant eksperimentinius duomenis gali nesutapti su rezultatais, kurie būtų gaunami naudojant daugiau kasdienės kalbos žodžių. Antra, negalima patikrinti, ar fonemų atpažinimui įtakos turi šalia fonemos esančios kitos fonemos – nors rengiant eksperimente naudotas programas buvo atsižvelgta į tokią galimybę, pastebėta, kad fonetinė įvairovė šalia fonemos nėra pakankamai didelė, kad galima būtų bandyti daryti kokius nors tyrinėjimus ir išvadas. Trečia, kadangi labai ribota fonemų aplinka, tęsiant tyrimus su tokios sudėties lietuviškojo garsyno duomenų baze, galimos problemos parenkant slaptažodžius iš geriausiai diktorius identifikuojančios fonemų aibės. Todėl atliekant tolimesnius tyrimus šioje srityje, rekomenduojama:
 - Atlikti tyrimą su didesniu LTDIGITS garsyno diktorių skaičiumi;
 - Plėsti LTDIGITS garsyno duomenų bazę įtraukiant į ją daugiau fonemų kombinacijų (difonų ir trifonų), atliekant jų segmentavimo ir MDKK koeficientų skaičiavimus.

LITERATŪROS SĄRAŠAS

Mokslinė literatūra

1. Ariyaeeinia, A. M., Sivakumaran P. (1997) Analysis and Comparison of Score Normalisation Methods for Text-Dependent Speaker Verification. 5th European Conference on Speech Communication and Technology [interaktyvus] . [žiūrėta 2007 m. lapkričio 29 d]. Prieiga per internetą: <http://www.isca-speech.org/archive/eurospeech_1997/e97_1379.html>
2. Driaunys K. (2004) Lietuvių kalbos signalų duomenų bazės LTDIGITS segmentavimas ir statistinis tyrimas. Informacijos mokslai. ISSN 1392-0561. Nr. 28. p.87-96.
3. Girdenis A. (1995) Teoriniai fonologijos pagrindai. Vilnius. Petro ofsetas.
4. Furui, S. Toward the Ultimate Synthesis/Recognition System. In Voice Communication Between Humans and Machines. Roe, D., Wilpon, J. (editors). National Academy Press, Washington D.C., 1994: 450-466.
5. Floch, Le J.-L., Montacié C., Caraty M.-J. (1995) Speaker recognition experiments on the NTIMIT database [interaktyvus]. [žiūrėta 2007 m. lapkričio 29 d]. Prieiga per internetą: <<http://citeseer.ist.psu.edu/cache/papers/cs/13149/http:zSzzSzwww-rtcd.lip6.frzSzPAROLEzSzarticleszSzeuro95jl.pdf/lefloch95speaker.pdf>>
6. Griffin, C. Matsui, T. Furui, S (1994). Distance measures for text-independent speaker recognition based onMAR model. Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE. pp. I/309-I/312
7. Heerden, C. J., Barnard, E. (2007). Durations of Context-Dependent Phonemes: A New Feature in Speaker Verification. Volume 4441/2007, p. 93-103
8. Hermansky, H., Morgan, N., Bayya, A., Kohn, P. (1991). RASTA-PLP Speech Analysis. ICSI Technical Report, 1991
9. Huang X.D., Ariki Y., Jack, M.A. (1990) Hidden Markov models for speech recognition, Edinburgh University press.
10. Young S., Evermann G., Kershaw D., Moore G., Odell J., Ollason D., Valtchev V., Woodland P. (2001) The HTK Book. Microsoft Corporation.
11. Jain, A. K., "Biometric recognition: how do I know who you are?", Signal Processing and Communications Applications Conference, 2004. Proceedings of the IEEE 12th: 3 – 5
12. Jain, A. K., A. Ross, S. Prabhakar, "An Introduction to Biometric Recognition", IEEE Trans. on Circuits and Systems for Video Technology, Vol. 14, No. 1, pp 4-19, January 2004
13. Makhoul, J. (1975). Linear prediction: A tutorial review. Proceedings of the IEEE, Vol. 63, No. 4., 1975, pp. 561 – 580.
14. Mammone, R. J., Zhang, X. and Ramachandran, R. P. (1996). "Robust speaker recognition," IEEE Signal Processing Magazine, 13, Sept., 58-71.
15. Markel D., Gray, Jr. A.H. (1976), Linear Prediction of Speech, Springer Verlag, Berlin.
16. Matsui T., Furui S. (1993). Concatenated phoneme models for text-variable speaker recognition. In ICASSP [ICA93], pp. 391–394.
17. McLaughlin J., Reynolds D.A. (2002). Speaker detection and tracking for telephone transactions. Acoustics, Speech, and Signal Processing, 2002. (ICASSP '02), pp. I-129 – I-132, vol.1.

18. Motoyuki, S, Toshiaki, A. (1999). Speaker Adaptation Using Phoneme-Dependent Tree-Structured Speaker Clustering. IEICE Transactions on Information and Systems. J82-D-2 , pp. 981-989
19. Qi Li Biing-Hwang Juang Chin-Hui Lee (2000). Automatic verbal information verification for user authentication. Speech and Audio Processing, IEEE Transactions, Vol. 8, Issue 5, pp. 585-596
20. Parthasarathy, S. Rosenberg, A.E (1996). General phrase speaker verification using sub-word backgroundmodels and likelihood-ratio scoring. Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference, p. 2403-2406 vol.4.
21. Rosenberg, A.E. Parthasarathy, S. (1996). Speaker background models for connected digit password speakerverification. Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference. pp. 81-84 vol. 1
22. Saswato D., (2007). „Qubits poised to reveal our best-kept secrets“. The New Scientist, Volume 195, Issue 2621, 15 September 2007, p. 30-31
23. Shahin (2005). Enhancing speaker authentication systems using circular hidden Markov models. Signal Processing and Its Applications, 2005. Proceedings of the Eighth International Symposium, p. 703-706, vol. 2.

Informacijos šaltiniai

24. Aclam ,P. J. (2003). MATLAB array manipulation tips and tricks [interaktyvus]. [žiūrėta 2008 03 15]. Prieiga per internetą: <<http://home.online.no/~pjacklam/matlab/doc/mtt/doc/mtt.pdf>>
25. Balakrishnama, S. Ganapathiraju, A. (1998) Linear discriminant analysis - a brief tutorial [interaktyvus]. [žiūrėta 2008 02 20] Prieiga per internetą: <http://lcv.stat.fsu.edu/research/geometrical_representations_of_faces/PAPERS/lda_theory.pdf>
26. Ebiz.lt. (2006) Slaptažodžiai - silpnoji lietuvių vieta (Straipsnis [interaktyvus]. [žiūrėta 2006 m. balandžio 16 d.]. Prieiga per internetą: <<http://www.vipt.lt/cms/app?service=external/ComNewsDetails&sp=2321&sp=SComNewsArchive>>
27. Kompiuterija. (2007) Visi slaptažodžiai įveikiami [interaktyvus]. [žiūrėta 2007 m. rugsėjo 20 d.]. Prieiga per internetą: <<http://kompiuterija.lrytas.lt/zurnale/idomu/2007-09-20/visi-slaptaazodziai-iveikiami/>>
28. Likit (2002). Balso technologijų pasiekimai pasaulyje. Balso technologijų taikymo lietuvių kalbai analizė ir perspektyvinių veiklos kryptių pagrindimas [interaktyvus]. [žiūrėta 2008 02 20]. Prieiga per internetą: <http://www.likit.lt/all/balso_tech/03_pasiekimai.htm>
29. Lockdown.co.uk - The Home Computer SecurityCentre. (2007) Password Recovery Speeds [interaktyvus]. [žiūrėta 2007 m. rugsėjo 20 d.]. Prieiga per internetą: <<http://www.lockdown.co.uk/?pg=combi&s=articles>>
30. Rudžionis, V. (1998) Priebalsių atpažinimas difonuose priebalsis – balsis. Informacijos mokslai [interaktyvus]. [žiūrėta 2008 02 20] Prieiga per internetą: <<http://www.leidykla.vu.lt/inetleid/inf-m-9/rudzionis.html>>
31. Scholarpedia (2008). Speaker recognition. S. Furui kuruojamas populiarinio mokslo puslapis [interaktyvus]. [žiūrėta 2008 02 20]. Prieiga per internetą: <http://www.scholarpedia.org/article/Speaker_recognition>
32. Search Security (2006). Biometrijos apibrėžimas [Interaktyvus]. [žiūrėta 2008 04 15] Prieiga per internetą < http://searchsecurity.techtarget.com/sDefinition/0..sid14_gci211666.00.html>

33. SearchSecurity.com. (2007) Cheat sheet: Access management solutions and their pros and cons [interaktyvus]. [žiūrėta 2008 m. sausio 5 d.]. Prieiga per internetą: <http://searchsecurity.techtarget.com/tip/0,289483,sid14_gci1157765,00.html>
34. Smith, S. (2006) Biometrics: 21st Century Security. Straipsnis [interaktyvus]. [žiūrėta 2008 02 20]. Prieiga per internetą: <http://www.infosecwriters.com/text_resources/pdf/Biometrics_SSmith.pdf>
35. Speechechmag.org (2007). Planet Payment Adds Voice Authentication [interaktyvus]. [žiūrėta 2008 m. sausio 5 d.]. Prieiga per internetą: <<http://www.speechechmag.com/Articles/News/Industry-News/Planet-Payment-Adds-Voice-Authentication-40412.aspx>>
36. TheFreeDictionary. (2006) Weak password, apibrėžimas [interaktyvus]. [žiūrėta 2006 m. balandžio 16 d.]. Prieiga per internetą: <<http://computing-dictionary.thefreedictionary.com/weak+password>>
37. TheFreeDictionary. (2006) Strong password, apibrėžimas [interaktyvus]. [žiūrėta 2006 m. balandžio 16 d.]. Prieiga per internetą: <<http://computing-dictionary.thefreedictionary.com/strong%20password>>
38. UsabilityNews (2004). People's Bad Password Habits exposed. Straipsnis [interaktyvus]. [žiūrėta 2006 m. balandžio 16 d.]. Prieiga per internetą: <<http://www.usabilitynews.com/news/article1926.asp>>
39. Verslo banga. (2003) Apklauskos rezultatai. Iš ko paprastai sudarote savo slaptažodžius?]. [žiūrėta 2006 m. balandžio 16 d.]. Prieiga per internetą: <<http://verslas.banga.lt/lt/poll.results/13>>
40. Wikipedia. (2007) Autentifikavimo apibrėžimas [interaktyvus]. [žiūrėta 2007 m. rugsėjo 29 d.]. Prieiga per internetą: < <http://en.wikipedia.org/wiki/Authentication> >
41. Wikipedia (2007). Stiprus slaptažodis [interaktyvus]. [žiūrėta 2006 m. balandžio 16 d.], <http://en.wikipedia.org/wiki/Password_strength>
42. Wikipedia (2007). Stiprus slaptažodis [interaktyvus]. [žiūrėta 2006 m. balandžio 16 d.], <<http://en.wikipedia.org/wiki/Biometrics>>

PRIEDŲ SĄRAŠAS

1	priedas. Anketa naudota slaptažodžių tyrime.....	I
2	Priedas. Antrojo slaptažodžių tyrimo duomenys.....	IV
3	Priedas. biometrinių savybių pagal įvairius kriterijus palyginimas	VIII
4	Priedas. Katalogų automatinio kūrimo programos išeities tekstas.....	IX
5	Priedas. Duomenų paruošimo programos išeities tekstas	X
6	Priedas. MFCC reikšmių iškirpimo iš *.Mfc failų funkcijos išeities kodas.....	XVII
7	Priedas. Fonemų klasifikavimo programos išeities tekstas.....	XVII
8	priedas. Programų paleidimo programos išeities kodas	XXIII
9	Priedas. eksperimento metu gauti Vyrų diktorių tariamų fonemų įverčiai	XXIV
10	Priedas. eksperimento metu gauti moterų diktorių tariamų fonemų įverčiai	XXVII
11	Priedas. Išrikiuoti vyrų diktorių fonemų panašumų ĮVERČIAI	XXXI
12	priedas. Išrikiuoti moterų diktorių fonemų panašumų ĮVERČIAI.....	XXXIV
13	Priedas. Fonemų pasikartojimo skirtumai vyrų ir moterų diktorių tarpe 1 – 21 eilutėse XXXVIII	
14	Priedas. Straipsnis „Diktoriaus atpažinimo sistema paremta fonetinių vienetų diskriminavimu“ XLIV	

1 PRIEDAS. ANKETA NAUDOTA SLAPTAŽODŽIŲ TYRIME

Esu Vilniaus universiteto magistrantas Kęstutis Tomkevičius (kestutis.tomkevicius@vukhf.lt) ir atlieku Lietuvos kompiuterių naudotojų elgsenos su slaptažodžiais tyrimą. Siekiant palyginti pastaraisiais metais JAV atliktų apklausų gautus duomenis apie JAV interneto naudotojų elgseną su situacija Lietuvoje, man reikalinga Jūsų pagalba.

Pagrindiniai šio tyrimo tikslai – išsiaiškinti naudotojų supratimą apie slaptažodžių kūrimo, keitimo, saugojimo principus, bei kaip jie susiję su individo charakteristikomis.

Prašau Jūsų atidžiai perskaityti žemiau pateiktus klausimus ir atsakyti į juos, pasirenkant Jums tinkamiausią atsakymą (jei nurodyta kitaip – kelis arba parašyti savo).

Anketa yra anoniminė, tyrimo rezultatai bus pateikti tik apibendrinta forma, pavieniai duomenys nebus publikuojami. Anketavimo metu gauti rezultatai bus panaudoti baigiamajam darbui.

Prašau užpildyti šią anketą

1. Ar naudojate interneto svetainėmis, kuriose reikalinga autentifikacija slaptažodžiu?
 Taip. Ne.
Jei pirmas atsakytas neigiamai, pereikite prie 14 klausimo.
2. Keliose skirtingose vietose (svetainėse, programose) naudojate slaptažodžiais?
 1 – 3;
 4 – 6;
 7 – 9;
 10 – 15;
 15 ir daugiau.
3. Kokio tipo svetainėse, reikalaujančiose prisijungimo slaptažodžiu, lankotės? (**Galimi keli variantai**).
 El. bankininkystės sistemose (pvz. hansa.net, seb.lt, Snoro bankas internetu ir pan.).
 Elektroninio pašto svetainės (pvz. Zebra (Takas), One, Delfi, Omni, Yahoo, Hotmail ir t.t.).
 El. parduotuvėse, aukcionuose (pvz. Pirkis, Amazon.com, E-bay, bid.lt ir pan.).
 Pramogų portaluose (pvz. One, Biteplius, pupauoga.lt ir pan.).
 Darbo paieškos portaluose (pvz. CV.lt, CV-online.lt, cvbankas.lt ir pan.).
 Pažinčių portaluose (pvz. draugas.lt, ieškok.lt, pazintys.lt, pazintys.biz ir pan.).
 Interneto diskusijose (pvz. diskusijos.lt, supermama.lt ir pan.)
 Kita. (parašyti) _____
4. Ar Jums svarbus Jūsų duomenų, saugomų slaptažodžiais, saugumas internete?
 Taip, labai svarbus.
 Šiek tiek svarbus.
 Ne, tuo visiškai nesirūpinu.
5. Kiek skirtingų slaptažodžių naudojate?
 1 – 3;
 4 – 6;
 7 – 9;
 10 – 15;
 15 ir daugiau.
6. Koks vidutinis Jūsų slaptažodžio ilgis?
 3 – 6 simboliai;
 7 – 8 simboliai;
 9 – 12 simbolių;
 Daugiau nei 12 simbolių.

7. Trumpiausio Jūsų slaptažodžio ilgis?
- 1 – 2 simboliai
 - 3 – 5 simboliai.
 - 6 – 7 simboliai.
 - 8 – 9 simboliai.
 - 10 ir daugiau simbolių.
8. Kokius metodus manote esamus tinkamais savo slaptažodžiams kurti: (**Galimi keli variantai.**)
- Naudoti savo vardą.
 - Naudoti savo antros pusės vardą.
 - Naudoti savo šeimos narių vardus.
 - Naudoti savo augintinio vardą.
 - Naudoti automobilio pavadinimą.
 - Naudoti muzikos grupės pavadinimus, atlikėjo vardus arba pavardes, muzikinių kūrinių pavadinimus.
 - Naudoti bet kokius kitus prasmę turinčius žodžius.
 - Naudoti savo arba artimų žmonių gimimo datas.
 - Naudoti kitas įsimintinas sau datas.
 - Naudoti vien tik kitų skaičių kombinacijas.
 - Naudoti programų sugeneruotus slaptažodžius.
 - Naudoti kitus principus. (parašyti)
-

9. Kaip saugote savo slaptažodžius?
- Turiu užsirašęs ant stalo po pelės kilimėliu.
 - Turiu užsirašęs slaptažodžių knygutėje.
 - Naudojuosi specialiomis slaptažodžių saugojimo programomis.
 - Slaptažodžiai saugiai paslėpti mano galvoje.
 - Naudoju kitą slaptažodžių saugojimo būdą: (parašyti)
-

10. Kaip dažnai keičiate savo slaptažodžius:
- Dažniau nei kartą per mėnesį.
 - Kartą per mėnesį.
 - Kartą per 3 mėnesius.
 - Kartą per pusę metų.
 - Kartą per metus.
 - Rečiau nei kartą per metus. (Parašyti, kaip retai _____)
 - Nekeičiu, visą gyvenimą naudoju tą patį slaptažodį.
11. Ar esate savo slaptažodžius dėl kokių nors priežasčių pasakę kitiems?
- Taip, daugiau nei 5 kartus.
 - Taip, 2 – 5 kartus.
 - Taip, vieną vienintelį kartą.
 - Ne, niekada.

Jei į prieš tai atsakyta klausimą atsakyta teigiamai, papildomai atsakykite į 12 – 13 klausimus:

12. Kiek žmonių žino bent vieną jūsų slaptažodį?
- 1 – 2;
 - 3 – 4;
 - 5 – 7;
 - 8 ir daugiau;
13. Kokios buvo slaptažodžio išdavimo priežastys?
- Slaptažodžio prašė sistemos administratorius.

- Slaptažodį perdaviau draugui, kad patikrintų mano paštą.
 - Slaptažodį pasakiau, nes bijojau pamiršti.
 - Slaptažodis buvo gautas nesankcionuotai.
 - Kitos priežastys (parašyti)
-

14. Ar esate girdėję apie stipraus slaptažodžio sudarymo principus?

- Girdėjau, stengiuosi jais vadovautis.
- Girdėjau, bet nesivadovauju.
- Negirdėjau, bet norėčiau sužinoti.
- Negirdėjau ir manęs jie nedomina.

15. Ar esate girdėję apie slaptažodžio saugojimo principus?

- Girdėjau, stengiuosi jais vadovautis.
- Girdėjau, bet nesivadovauju.
- Negirdėjau, bet norėčiau sužinoti.
- Negirdėjau ir manęs jie nedomina.

16. Jūsų lytis?

- Vyras.
- Moteris.

17. Jūsų amžius?

- 5 – 14.
- 15 – 24.
- 25 – 35.
- 36 – 50.
- 50 ir vyresni.

18. Jūsų išsilavinimas?

- Pagrindinis.
- Vidurinis.
- Aukštesnysis (proftechninis).
- Nebaigtas universitetinis.
- Aukštasis neuniversitetinis.
- Aukštasis universitetinis.

19. Jūsų darbo kompiuteriu patirtis (metais):

- 1 – 2.
- 3 – 5.
- 5 – 10.
- 10 ir daugiau.

20. Jūsų kompiuterinio raštingumo lygis.

- Pradedantysis naudotojas.
- Pažengęs naudotojas.
- Kompiuterių profesionalas.

21. Jūsų gyvenamoji vieta:

- Vilnius.
- Kaunas, Klaipėda, Šiauliai, Panevėžys, Alytus.
- Kiti miestai.
- Kaimas.

22. Jūsų mėnesio pajamų dydis:

- iki 600 Lt.
- 600 – 800 Lt.
- 801 – 1499 Lt.
- 1500 – 2500 Lt.
- Virš 2500 Lt.

Ačiū už sugaištą laiką, Jūs labai padėjote!

2 PRIEDAS. ANTROJO SLAPTAŽODŽIŲ TYRIMO DUOMENYS

Eil. Nr	Prisijungimo vardas	Asmens vardas	Slaptažodis
1	aiste501	aiste501	akelina
2	Andorada	Andorada	viesulas
3	Ipsum	Ipsum	pergrazus
4	Ruta.S	ozhkute	rutukas
5	smike	smike	draugas
6	laurrutte	laurrutte	litvinskaite
7	NeivardintasN	Neivardintas N	46422
8	Kestutis	vice-prezidentas	apsaukmane
9	Eglis	Klecckis	ogiss
10	Rasuole	Raudonplauke	pasimatymas
11	mokslininke	LaimaJan	1234
12	ingas	inga salyte	klaudziukas
13	geete	Geete	GEELI666
14	spanguole	spange	mantulis
15	sabris	sabris	sabris20
16	bambatrynis	bambatrynis	840305
17	alekas	alekas	tankistas
18	axvnka	meška	gabija
19	kshisia	kshisia	19841124
20	neziniuke	neziniuke	neziniuke
21	JolitaR	Jolita Rimšaitė	ciburaska
22	ginteg	gintukas	sopranos
23	kazkas	kazkas	belekas123
24	goda	godzila	digas
25	ibzuvele	Zuvelia	indrule
26	Simulcika	Simulcika	zinovas
27	nikitux	nikitux	talita
28	meskiukas	meskiukas	meskiukas
29	Vitkus	Arnas Vitkus	vitkus
30	nso04070	vilma	elvis
31	musiadayy	kolega	2003091812
32	inga	ingule	545160
33	jewl-e-ah	jewl-e-ah	matehuala
34	Armas	Arturas VDU	elektrenai
35	elektrikas	elektrikas	mentas
36	MeOnly	MeOnly	495658
37	Joli	jolli	studentas
38	creative	creative	bambalis
39	crazy_beautiful	crazy_beautiful	inulka
40	Mize	bombonke	angeliukas
41	sifus	svedas	svedas
42	pequena	pequena	boreanaz
43	Minosa	Minosa	abitura
44	Spinduliux	Spinduliux	mazamete20

Eil. Nr	Prisijungimo vardas	Asmens vardas	Slaptažodis
45	soliaris	soliaris	amzinybe
46	kristina	sauluze	angel
47	Daliux	Studencioke	vytukas
48	daivukas	dadada	daiva
49	gelvis	gelvis	kristina
50	pvspvs	pvspvs	radiohead
51	Inga.	Inga.	BORUZYTTE
52	zibuoklyte	violeta	zibuokle
53	Destytojas	BIFIDO	987654
54	MudzhyJ	MudzhyJ	aviena
55	cesnjuli	Julius	kaimas
56	BlackNo1	BlackNo1	itasitas
57	vikle	vikle	z-chebra
58	Charisma	Charisma	escada
59	Duokvarda	Prisukama citrina	outsidedoor
60	Phosphor	Phosphor	maqmba
61	karmelita	karmelita	jurgyte
62	Balamutuke	Ieviukas	liutukas
63	labas9	labas9	telefonas
64	zedas	zedas	adazas
65	LaurynasS	Lauris	runce
66	teletabiz	teletabiz	666666
67	dz0_	Jurgis Pralgauskis	123456
68	Laucius	Laucevicius	gaidys
69	povilas666	Evil Inside	alus666
70	mazhoji	Lina B	pagrandukas
71	linasj	linas	linas111
72	lauris	lauris	braskute
73	Anglas	Anglas	justdoit
74	Paloma	Paloma	valda19
75	leily	leily	pavasaris
76	Skruzdelyte	Skruzdelyte	19860310
77	murma	bambeklè	DIEGA
78	exhumacija	mergaichiuke	coliuke
79	mickus	k.parazitas	intuito
80	KRIU KRIU	KRIUKRIU	PUPCIUNE
81	adriana	adriana	zalgiris
82	eskime	eskime	ragana
83	rametvarijon	Raminta	pranas
84	jura	juoda	senagim111
85	Mergiukshte	Mergiukshte	ingute
86	gugne	gugne	gugne
87	EdHeaD	Eckagalvis	ghitzerai
88	Egipta	Egipta	andrius
89	asarele	angeles	isdykele

Eil. Nr	Prisijungimo vardas	Asmens vardas	Slaptažodis
90	Ciucela	Ciucela	tarka
91	Nenute	Nenute	uzaugaukaime16
92	spurgencija	spurgencija	murgita
93	Aura	Auryte	vasarele
94	Antonio	Banderas	20040501
95	bagyra7	bagyra7	pantera
96	vaida	student	darbas
97	studente	studente	711015
98	wishtajte	laimes kudikis	lakstingala
99	asta	zuika	suniuukas
100	ezis	ezis2	motorola
101	kayra	renaciuke	renatukas
102	birbizas	birbizass	987654
103	jonas	[DiZ]	jonas
104	Giedra	Giedra	dizainas
105	skorpione	skorpione	adele13
106	kitokia152	kitokia152	19870928
107	enigma	enigma	kliukva
108	MaterialG	MaterialG	lifeisgood
109	Kalade	Kalade	pasaulis
110	tinuska	smile:)	parole
111	Forget^It	Forget^It	722415
112	smiltis	smiltis	karamia
113	daktariukste	daktariukste	gimnaziste
114	RedHairMafia	RedHairMafia	trylika
115	waidele	Vaida V.	drugys
116	desert_fox	desert_fox	zuikis
117	Jecky	Jecky	jeckute
118	Abredukas	Abredukas	adamkus
119	zuikiuka	Katja	berlynas
120	gija	gijele	naktiniszygis
121	Dainora	Dainora	vasara
122	studente_karolina	studente_karolina	zuikis
123	liepa19	liepa19	saule
124	Initta	Initta	ramune
125	Fantasy	Fantasy	fantasyone
126	joosteete	me...	alinabea
127	julkam	Liulia	katinas
128	mambo	mambo	broliukas
129	cassawa	KTU-valdo	hazapijus
130	RooteLa	RooteLa	rootela
131	lawriz	lawriz	judesys
132	SKAMPAI	Victoria	SKAMPAI
133	stift	stiftas	987654
134	bebras	bebras	bebras

Eil. Nr	Prisijungimo vardas	Asmens vardas	Slaptažodis
135	swee	sweee	pomidoras
136	Ievute	Ievute	assirdele
137	Drasius	Drasius (VDU)	vidmantas123
138	asketas	asketas	kamuoliukas02
139	birbis	birbis	gaidys
140	mistica	Agne_	862107004
141	Martute	kaunas	tobiukas
142	saulyte	studente:)	sviecia
143	v0dka	Dainiokas	katinas
144	zandas	pupele	pupsikas
145	mistica_	niekas	862107004
146	tapke	tapke	darius16
147	Ieviukas	Ieviukas	baubukas
148	jusste21	dzesta	781521
149	DeimaV	Vileikiene	19850926
150	angeliukste	angeliukste	merkurijus

3 PRIEDAS. BIOMETRINIŲ SAVYBIŲ PAGAL ĮVAIRIUS KRITERIJUS PALYGINIMAS

Eil. Nr.	Lyginimo kriterijus																					
	Unikalumas	Surenkamumas	Apeinamumas	Vidurkis 1	Vidurkis 2	Vidurkis 3	Ausis	Balsas	Delno anspaudai	DNR	Eisena	Kvapas	Parašas	Pirštų anspaudai	Rainelė	Rankos geometrija	Rankos venos	Rašymas klaviatūra	Tinklainės skenavimas	Veidas	Veido termograma	
1	Veido termograma	Rankos geometrija	Veido termograma	Veido termograma	Pirštų anspaudai	Pirštų anspaudai	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	3
2	Rainelė	Veidas	Rainelė	Pirštų anspaudai	Delno anspaudai	Delno anspaudai	0	0	2	0	0	0	0	1	2	0	0	0	0	1	0	
3	Tinklainės skenavimas	Veido termograma	Rankos venos	Delno anspaudai	Veido termograma	Veido termograma	0	0	1	0	0	0	0	0	0	1	0	1	0	0	3	
4	Kvapas	Eisena	Kvapas	Rankos geometrija	Rainelė	Veidas	0	0	0	0	1	2	0	0	1	1	0	0	0	1	0	
5	DNR	Parašas	DNR	Veidas	Ausis	Rainelė	1	0	0	2	0	0	1	0	1	0	0	0	0	1	0	
6	Pirštų anspaudai	Rainelė	Tinklainės skenavimas	Parašas	Veidas	Ausis	1	0	0	0	0	0	1	1	1	0	0	0	1	1	0	
7	Delno anspaudai	Pirštų anspaudai	Rankos geometrija	Rainelė	DNR	Rankos geometrija	0	0	1	1	0	0	0	1	1	2	0	0	0	0	0	
8	Rankos venos	Delno anspaudai	Eisena	Ausis	Rankos geometrija	DNR	1	0	1	1	0	0	0	0	0	1	1	0	0	0	0	
9	Rankos geometrija	Ausis	Pirštų anspaudai	Eisena	Tinklainės skenavimas	Parašas	1	0	0	0	1	0	1	1	0	1	0	0	1	0	0	
10	Ausis	Rankos venos	Delno anspaudai	Balsas	Kvapas	Tinklainės skenavimas	1	1	1	0	0	1	0	0	0	1	0	1	0	0	0	
11	Eisena	Rašymas klaviatūra	Ausis	Tinklainės skenavimas	Rankos venos	Kvapas	1	0	0	0	1	1	0	0	0	0	1	1	1	0	0	
12	Rašymas klaviatūra	Balsas	Rašymas klaviatūra	Kvapas	Eisena	Eisena	0	1	0	0	2	1	0	0	0	0	0	2	0	0	0	
13	Veidas	Kvapas	Veidas	DNR	Parašas	Balsas	0	1	0	1	0	1	1	0	0	0	0	0	0	2	0	
14	Parašas	DNR	Parašas	Rankos venos	Balsas	Rankos venos	0	1	0	1	0	0	2	0	0	0	2	0	0	0	0	
15	Balsas	Tinklainės skenavimas	Balsas	Rašymas klaviatūra	Rašymas klaviatūra	Rašymas klaviatūra	0	2	0	0	0	0	0	0	0	0	0	3	1	0	0	

4 PRIEDAS. KATALOGŲ AUTOMATINIO KŪRIMO PROGRAMOS IŠEITIES TEKSTAS

```
% ----- Magistro darbo reikalingu direktoriju sukurimo skriptas -----

%close all; clear all; clc;
disp ('*****');

% ***** Pradiniai kintamieji *****

global_root = 'C:\Documents and Settings\Vycius\Desktop\Signalai'; % kelias iki duomenu folderio kompiuteryje
%ltdig_root = [global_root '\Ltdigits']; % kelias iki Ltdigis folderio
mff_root = [global_root '\Mfcc\Htk']; % kelias iki mff failu folderio
coefs_root = [global_root '\Mfcc\Coefs']; % kelias iki mano apskaiciuotu koeficientu folderio

dikt_nr = ""; % diktoriaus numeris
f_nr = -1; % fonema
DSK = 50; % DSK = max diktoriu skaicius
fpradzia = -1; % fonemos pradzia
fpabaiga = -1; % fonemos pabaiga
iesk_fonema = ""; % fonema kurios ieskome
pozymiu_sk = 39; % fonemos pozymiu skaicius
lytis = 'FM'; %Lytis
sex_root = "";

% ***** lyties nuskaitymas *****

gender_file = [global_root '\gender.txt'];
phonemes_file = [global_root '\phonemes.txt'];
log_file = [coefs_root '\phrases_phonemes_log.txt'];

% ***** Apdorojam visas lytis *****
fid = fopen(gender_file);
while 1
    tline = fgetl(fid);
    if ~ischar(tline), break, end
    if tline == 'M'
        lytis = tline;
        sex_root = 'Male';
    else
        if tline == 'F'
            lytis = tline;
            sex_root = 'Female';
        else
            end
    end
end
disp(lytis);
disp(sex_root);

% ***** fonemu failo nustatymas *****
[phonemes] = textread(phonemes_file, '%c');
phlen = length(phonemes);

% ***** diktoriaus nr nustatymas *****
for d = 1:DSK
    %disp(d);
    if d <= 9
        dikt_nr = [lytis '00'];
    else
        dikt_nr = [lytis '0'];
    end
end
```

```

end

dikt_nr = [dikt_nr " int2str(d)];

kelias_iki = [coefs_root '\ sex_root '\];

mkdir (kelias_iki, dikt_nr);
disp(['Diktoriaus ' dikt_nr ' katalogas sukurtas (' kelias_iki " dikt_nr ').']);

naujas_kelias = [kelias_iki '\ dikt_nr '\];

mkdir (naujas_kelias, 'Parinkimo_Rezultatai');

disp(['Katalogas ' naujas_kelias 'Parinkimo_Rezultatai sukurtas.'])

end % diktoriaus nr for'o end'as

end %Lyties while'o end'as

fclose(fid);
disp ('----- Scripto pabaiga-----');

```

5 PRIEDAS. DUOMENŲ PARUOŠIMO PROGRAMOS IŠEITIES TEKSTAS

```

% ----- Magistro darbo duomenu paruosimo skriptas -----

%close all; clear all;
disp ('*****');

% ***** Pradiniai kintamieji *****

global_root = 'C:\Documents and Settings\Vycius\Desktop\Signalai'; % kelias iki duomenu folderio kompiuteryje
ltdig_root = [global_root '\Ltdigits']; % kelias iki Ltdigis folderio
mff_root = [global_root '\Mfcc\Htk']; % kelias iki mff failu folderio
coefs_root = [global_root '\Mfcc\Coefs']; % kelias iki mano apskaiciuotu koeficientu folderio

dikt_nr = ""; % diktoriaus numeris
f_nr = -1; % fonema
DSK = 50; % DSK = max diktoriu skaicius
FSK = 8; % FSK = max fraziu skaicius, max sk - 8
fpradzia = -1; % fonemos pradzia
fpabaiga = -1; % fonemos pabaiga
iesk_fonema = ""; % fonema kurios ieskome
pozymiu_sk = 39; % fonemos pozymiu skaicius
lytis = 'FM'; % Lytis
sex_root = "";
irasyta = 0;

% ***** lyties nuskaitymas *****

gender_file = [global_root 'gender.txt'];
phonemes_file = [global_root 'phonemes.txt'];
log_file = [coefs_root 'log.txt'];
%countf_failas = [coefs_root 'count_phonemes.txt'];

% ***** Apdorojam visas lytis *****
fid = fopen(gender_file);
while 1
    tline = fgetl(fid);

```



```

if ~ischar(tline), break, end
if tline == 'M'
    lytis = tline;
    sex_root = 'Male';
else
    if tline == 'F'
        lytis = tline;
        sex_root = 'Female';
    else
        end
end
disp(lytis);
disp(sex_root);

countf_failas = [coefs_root '\' sex_root '\count_phonemes.txt'];

% ***** fonemu failo nustatymas *****
[phonemes] = textread(phonemes_file, '%c');
phlen = length(phonemes);

% ***** Apdorojam visas fonemas *****
for ph = 1:phlen
    iesk_fonema = phonemes(ph)

    fonsk = 0;
    %fonsk_m = 0;
    %fonsk_f = 0;

    log = fopen(log_file, 'a');
    fprintf(log, '%s', iesk_fonema);
    fprintf(log, '\n');
    fclose(log);

% ***** diktoriaus nr nustatymas *****
for d = 1:DSK
    %disp(d);
    if d <= 9
        dikt_nr = [lytis '00'];
    else
        dikt_nr = [lytis '0'];
    end

    dikt_nr = [dikt_nr " " int2str(d)]

    log = fopen(log_file, 'a');
    fprintf(log, '%s', dikt_nr);
    fprintf(log, '\n');
    fclose(log);

    %dir_root = [coefs_root '\' sex_root];

    %status = exist (dikt_nr, dir_root)

    %mkdir (dir_root, dikt_nr);

    %sukuriamas kelias iki failo fonemos ribu issagojimui
    frb_failas_irasymui = [coefs_root '\' sex_root '\' dikt_nr '\' iesk_fonema '_'_phoneme.vkt'];

    %sukuriamas kelias iki failo frazes numerio ir aplink esanciu fonemu issagojimui
    frbph_failas_irasymui = [coefs_root '\' sex_root '\' dikt_nr '\' iesk_fonema '_'_phrases.vkt'];

```

```

%sukuriamas kelias iki failo fonemos pozymių vidurkių išsagojimui
rez_failas_irasymui = [coefs_root '\ sex_root '\ dikt_nr '\ iesk_fonema '_mfcc_mas.vkt'];

%sukuriamas kelias iki failo fonemos pozymių vidurkių išsagojimui
vid_failas_irasymui = [coefs_root '\ sex_root '\ dikt_nr '\ iesk_fonema '_average.txt'];

%sukuriamas kelias iki failo fonemos pozymių vidurkių išsagojimui
cm_failas_irasymui = [coefs_root '\ sex_root '\ dikt_nr '\ iesk_fonema '_covmat.txt'];

for f=1:FSK
    if f <= 9
        fr_nr = ['T0' int2str(f)];
    else
        fr_nr = ['T' int2str(f)];
    end
    %disp(fr_nr);

    % sukuriami nuskaitomu duomenų failai
    final_root = [ltdig_root '\ sex_root '\ dikt_nr];
    fonemu_failas = [final_root '\ fr_nr '.PHN'];
    ribu_failas = [final_root '\ fr_nr '.RBB'];
    mff_failas = [mff_root '\ sex_root '\ dikt_nr '\ fr_nr '.mff'];

    % ***** fonemos failo atidarymas ir nuskaitymas *****

    [zodis_mas, fpabaiga_mas, fonema_mas] = textread(fonemu_failas, '%d %d %s');
    eil_sk = length (fonema_mas);
    z_eil_sk = length (zodis_mas);
    % ***** fonemos failo atidarymo ir nuskaitymo pabaiga *****

    % pakeiciam visas fonemas x* i *
    for i=1:eil_sk
        %fonema_mas(i)
        %length(char(fonema_mas(i)))

        if length(char(fonema_mas(i)))>1
            disp ([iesk_fonema ', ' dikt_nr ', ' fr_nr ', ' char(fonema_mas(i))]);
            tolog = [iesk_fonema ', ' dikt_nr ', ' fr_nr ', ' char(fonema_mas(i))];
            log = fopen(log_file, 'a');
            fprintf(log, '%s', tolog);
            fprintf(log, '\n');
            fclose(log);

            if strcmp (fonema_mas(i), 'qq') == 0

                fonema_mas(i) = cellstr(substr(char(fonema_mas(i)), 1));
                disp ('Pakeista');
                disp ([iesk_fonema ', ' dikt_nr ', ' fr_nr ', ' char(fonema_mas(i))]);

                tolog = ['Pakeista'];
                log = fopen(log_file, 'a');
                fprintf(log, '%s', tolog);
                fprintf(log, '\n');
                fclose(log);
                tolog = [iesk_fonema ', ' dikt_nr ', ' fr_nr ', ' char(fonema_mas(i))];
                log = fopen(log_file, 'a');
                fprintf(log, '%s', tolog);
                fprintf(log, '\n');
                fclose(log);

            end
        end
    end
end

```

```

end %fonemos pakeitimo is x* i * pabaiga

for i=1:eil_sk

    %disp(['Rasta frazeje ' fr_nr ' vietoje ' i]);

    %tikriname ar nuskaityta eilute yra lygi ieskomai fonemai
    if strcmp (fonema_mas(i), iesk_fonema) == 1

        fonsk = fonsk + 1;

        % issirenkame fonema esancia pries ir po tikrinamos fonemos
        if i ~= 1
            if strcmp(fonema_mas(i-1), 'qq') == 1
                fonema_pries = 'Q';
            else
                fonema_pries = fonema_mas(i-1);
            end

        else
            fonema_pries = '-';
        end

        if i < z_eil_sk
            if strcmp(fonema_mas(i+1), 'qq') == 1
                fonema_po = 'Q';
            else
                fonema_po = fonema_mas(i+1);
            end
        else
            fonema_po = '-';
        end

        % issaugom indeksu su reikiama fonema reiksme
        indeksas = i;

        %issaugom pabaigos reiksme
        fpabaiga = fpabaiga_mas(indeksas);

        %issaugom zodzio numeri
        zodis = zodis_mas(indeksas) + 1;

        if i > 1

            if zodis_mas(i) ~= zodis_mas(i-1)    %if i == 1 | lygu_prev == 1

                %disp ('Atidarom nauja faila');
                [rbpr, rbpab] = textread (ribu_failas, '%d %d');

                %issaugom pradzios reiksme
                fpradzia = rbpr(zodis);
            else
                %issaugom pradzios reiksme
                fpradzia = fpabaiga_mas(indeksas-1);
            end
        else
            [rbpr, rbpab] = textread (ribu_failas, '%d %d');

            %issaugom pradzios reiksme
            fpradzia = rbpr(1);
        end
    end
end

```

```

disp(['Fonemos pradzia ' int2str(fpradzia) ', fonemos pabaiga ' int2str(fpabaiga) ', frazes nr. ' fr_nr]);
disp(['Fonema pries ' char(fonema_pries) ', fonema uz ' char(fonema_po)']);

tolog = ['Fonemos pradzia ' int2str(fpradzia) ', fonemos pabaiga ' int2str(fpabaiga) ', frazes nr. ' fr_nr ' . Fonema pries '
char(fonema_pries) ', fonema uz ' char(fonema_po)];
log = fopen(log_file, 'a');
fprintf(log, '%s', tolog);
fprintf(log, '\n');
fclose(log);

if (fpabaiga - fpradzia) < 256
    tolog = ['!!! Galima tuscium rezultatu aibe!!! Fonemos pradzia ' int2str(fpradzia) ' ir fonemos pabaiga '
int2str(fpabaiga) ' frazeje ' fr_nr ' skiriasi per maziau nei 256.'];
    log = fopen(log_file, 'a');
    fprintf(log, '%s', tolog);
    fprintf(log, '\n');
    fclose(log);

else

    % ***** Issaugom fonemos ribu duomenis faile *****
    fwr = fopen (frb_failas_irasymui, 'a');
    fprintf (fwr, '%d', fpradzia);
    fprintf (fwr, '\t');
    fprintf (fwr, '%d', fpabaiga);
    fprintf (fwr, '\n');
    fclose (fwr);

    fwr = fopen (frbph_failas_irasymui, 'a');
    fprintf (fwr, '%s', fr_nr);
    fprintf (fwr, '\t');
    fprintf (fwr, '%c', char(fonema_pries));
    fprintf (fwr, '\t');
    fprintf (fwr, '%c', char(fonema_po));
    fprintf (fwr, '\n');
    fclose (fwr);

    % ***** kreipinys i K. Driaunio f-ja *****
    rezultatai = Fonema_is_Mfcc (mff_failas, fpradzia, fpabaiga, pozymiu_sk);
    % ***** kreipinio i K. Driaunio f-ja pabaiga *****

    if (isempty(rezultatai) == 0) % lyginame ar rezultatu masyvas netuscias

        % grazinto reiksmiu masyvo ilgis
        rezlen = length(rezultatai);

        eilciuskaicius = rezlen / 39;

        fwr = fopen (rez_failas_irasymui, 'a');

        % irasom rezultatus i faila
        skait = 0; eil_sk = 0;
        for k = 1:rezlen
            skait = skait + 1;
            fprintf (fwr, '%f', rezultatai(k));
            if skait <= 38
                fprintf (fwr, '\t');
            else
                fprintf (fwr, '\n');
                skait = 0;
                eil_sk = eil_sk + 1;
            end
        end % rezultatu irasyimo for'o end'as

```



```

% ***** apskaiciuojam kiekvieno stulpelio vidurki ir rasom i faila *****
vid = [mean(r1) mean(r2) mean(r3) mean(r4) mean(r5) mean(r6) mean(r7) mean(r8) mean(r9) mean(r10) mean(r11)
mean(r12) mean(r13) mean(r14) mean(r15) mean(r16) mean(r17) mean(r18) mean(r19) mean(r20) mean(r21) mean(r22) mean(r23)
mean(r24) mean(r25) mean(r26) mean(r27) mean(r28) mean(r29) mean(r30) mean(r31) mean(r32) mean(r33) mean(r34) mean(r35)
mean(r36) mean(r37) mean(r38) mean(r39)];

fvr = fopen (vid_failas_irasymui, 'w');
for c = 1:39
    fprintf(fvr, '%f', vid(c));
    if c <=38
        fprintf(fvr, '\t');
    end
end
fclose (fvr);
% ***** apskaiciavimo ir rasymo i faila pabaiga *****

% ***** skaiciuojam covariacine matrica ir rasom i faila *****

CMA = load(rez_failas_irasymui); % Covariance Matrix Array

[N, M] = size(CMA); % N - kintamuju skaicius, M - eiluciu skaicius

CM = diag(diag(cov(CMA))); % Covariance Matrix % CM = diag(cov(CMA));

fvr = fopen (cm_failas_irasymui, 'w');
for cmi = 1:39
    for cmj = 1:39
        fprintf(fvr, '%f', CM(cmi,cmj));
        if cmj <= 38
            fprintf(fvr, '\t');
        else
            fprintf(fvr, '\n');
        end
    end
end
fclose (fvr);

else

    disp(['Fonema ' iesk_fonema ' nebuvo rasta diktoriaus ' dikt_nr ' frazese']);

end % tikrinimo ar yra fonemos failas end'as

% ***** diktoriaus tikrinimo pabaiga *****
disp(['Diktorius ' dikt_nr ' apdorotas sekmingai.']);

tolog = ['Diktorius ' dikt_nr ' apdorotas sekmingai.'];
log = fopen(log_file, 'a');
fprintf(log, '%s', tolog);
fprintf(log, '\n');
fclose(log);

end % diktoriaus for'o end'as

fwr = fopen (countf_failas, 'a');
fprintf (fwr, '%d', fonsk);
fprintf (fwr, '\t');
fprintf (fwr, '%s', iesk_fonema);
fprintf (fwr, '\n');
fclose (fwr);

```

```

tolog = ['Fonema ' iesk_fonema ' rasta ' fonsk ' kartus.'];
log = fopen(log_file, 'a');
fprintf(log, '%s', tolog);
fprintf(log, '\n');
fclose(log);

% ***** fonemos tikrinimo pabaiga *****
disp(['Fonema ' iesk_fonema ' apdorota sekmingai.']);
disp(' ');

tolog = ['Fonema ' iesk_fonema ' apdorota sekmingai.'];
log = fopen(log_file, 'a');
fprintf(log, '%s', tolog);
fprintf(log, '\n');
fclose(log);

end % fonemos for'o end'as

end %Lyties while'o end'as

fclose(fid);
disp ('----- Scripto pabaiga-----');
log = fopen(log_file, 'a');
fprintf(log, '----- Scripto pabaiga-----');
fprintf(log, '\n');
fclose(log);

```

6 PRIEDAS. MFCC REIKŠMIŲ IŠKIRPIMO IŠ *.MFC FAILŲ FUNKCIJOS IŠEITIES KODAS

```

function [mas_mfcc] = Fonema_is_Mfcc (path4, pr, pb, kofsk);
%Funkcija ispjauia is failu Konkrecios fonemos Mfcc parametrus irgrazina i pagrindine funkcija
%IEJIMO KINTAMIEJI
%path4 - kelias iki mfcc parametru failo
%pr - fonemos pradzios riba
%pb - fonemos pabaigos riba
%ISEJIMO KINTAMIEJI
%mas_mfcc - konkretios fonemos MFCC prametru masyvas
% path4='C:\Darbai\Mfcc\Htk\Female\F001\t01.mff';
% pr=100;
% pb=300;
% kofsk=39;

ii=((floor(pr/200))*4*kofsk); %baitu skaicius fonemos pradzios is mfc failo
%Pradzios riba dalinam is 2 paversdami diskretos nr kuris div is 100 ir dauginam is 4 ir 12
%*****
%Skaiciuojam kiek reiksmiu reikia nuskaityti
dydis=(floor(((pb-pr)/2)/100))*kofsk;
f1=fopen(path4, 'rb');
fseek(f1, ii, -1);
mas_mfcc=fread(f1, dydis, 'float32');
fclose(f1);

```

7 PRIEDAS. FONEMŲ KLASIFIKAVIMO PROGRAMOS IŠEITIES TEKSTAS

```

% ----- Magistro darbo fonemu rezultatu apskaičiavimo ir parinkimo skriptas -----

%close all; clear all;
disp ('*****');

% ***** Pradiniai kintamieji *****

global_root = 'C:\Documents and Settings\Vycius\Desktop\Signalai'; % kelias iki duomenų folderio kompiuteryje
%ltdig_root = [global_root '\Ltdigits']; % kelias iki Ltdigits folderio
mff_root = [global_root '\Mfcc\Htk']; % kelias iki mff failų folderio
coefs_root = [global_root '\Mfcc\Coefs']; % kelias iki mano apskaičiuotų koeficientų folderio

dikt_nr = ""; % diktoriaus numeris
f_nr = -1; % fonema
DSK = 50; % DSK = max diktorių skaičius
fpradzia = -1; % fonemos pradžia
fpabaiga = -1; % fonemos pabaiga
iesk_fonema = ""; % fonema kurios ieskome
pozymiu_sk = 39; % fonemos pozymių skaičius
lytis = 'FM'; %Lytis
sex_root = "";
irasyta = 0;

% ***** lyties nuskaitymas *****

gender_file = [global_root '\gender.txt'];
phonemes_file = [global_root '\phonemes.txt'];
log_file = [coefs_root '\phrases_phonemes_log.txt'];

% ***** Apdorojam visas lytis *****
fid = fopen(gender_file);
while 1
    tline = fgetl(fid);
    if ~ischar(tline), break, end
    if tline == 'M'
        lytis = tline;
        sex_root = 'Male';
    else
        if tline == 'F'
            lytis = tline;
            sex_root = 'Female';
        else
            end
    end
    disp(lytis);
    disp(sex_root);

% ***** fonemu failo nustatymas *****
[phonemes] = textread(phonemes_file, '%c');
phlen = length(phonemes);

% ***** Apdorojam visas fonemas *****
for ph = 1:phlen
    iesk_fonema = phonemes(ph)

    log = fopen(log_file, 'a');
    fprintf(log, '%s', iesk_fonema);
    fprintf(log, '\n');
    fclose(log);

% ***** diktoriaus nr nustatymas *****
for d = 1:DSK

```



```

%disp(d);
if d <= 9
    dikt_nr = [lytis '00'];
else
    dikt_nr = [lytis '0'];
end

dikt_nr = [dikt_nr " int2str(d)]

log = fopen(log_file, 'a');
fprintf(log, '%s', dikt_nr);
fprintf(log, '\n');
fclose(log);

% sukuriamas kelias iki failo , kuriame saugomos fonemos ribos
rb_failas = [coefs_root '\ sex_root '\ dikt_nr '\ iesk_fonema '_phoneme.vkt'];

% sukuriamas kelias iki failo frazes numerio ir aplink esanciu fonemu issagojimui
phr_failas = [coefs_root '\ sex_root '\ dikt_nr '\ iesk_fonema '_phrases.vkt'];

% rez_failas = [coefs_root '\ sex_root '\ dikt_nr '\ iesk_fonema '_mfcc_mas.vkt'];

% sukuriamas kelias iki failo fonemos covariacines matricos issagojimui
vid_failas = [coefs_root '\ sex_root '\ dikt_nr '\ iesk_fonema '_average.txt'];

% sukuriamas kelias iki failo fonemos pozymių vidurkiu issagojimui
cm_failas = [coefs_root '\ sex_root '\ dikt_nr '\ iesk_fonema '_covmat.txt'];

% sukuriamas kelias iki failo fonemos pozymių vidurkiu issagojimui
rez_failas_irasymui = [coefs_root '\ sex_root '\ dikt_nr '\Parinkimo_Rezultatai\ iesk_fonema '.txt'];

% sukuriamas kelias iki failo fonemos pozymių eiluciu skaiciaus issagojimui
eilsk_failas_irasymui = [coefs_root '\ sex_root '\ dikt_nr '\Parinkimo_Rezultatai\ iesk_fonema '_eilsk.txt'];

% sukuriamas kelias iki failo fonemos pozymių eiluciu skaiciaus issagojimui
all_eilsk_failas = [coefs_root '\ sex_root '\eilsk.txt'];

% sukuriamas failas dideliu eiluciu skaiciu fonemoms
big_phonemes = [coefs_root '\ sex_root '\big_phonemes.txt'];

% sukuriamas kelias iki failo fonemos pozymių vidurkiu issagojimui
vid_failas_irasymui = [coefs_root '\ sex_root '\ dikt_nr '\Parinkimo_Rezultatai\ averages.txt'];

%atliekam palyginima ar yra irasyti failai *_average.txt ir *_cov_mat.txt
%ivedame palyginimo logika

if (exist(vid_failas) == 2) & (exist(cm_failas) == 2)

    % uzsikraunam i masyva AVG faila *_average.txt
    AVG = load(vid_failas);

    % uzsikraunam i masyva CM faila *_cov_mat.txt
    CM = load(cm_failas);

    % uzsikrauname i masyva RB faila *_phoneme.vkt
    %PHR = load(phr_failas)
    [PHR, fonema_pries_mas, fonema_po_mas] = textread(phr_failas, '%s %c %c');

    % uzsikrauname i masyva RB faila *_phoneme.vkt
    RB = load(rb_failas);

    [RBx, RBy] = size(RB);

```

```

%RBx = 1; % pradzioje naudojam tik vienos is visu visose frazese esamu x fonemos mfcc rezultatus

koefisum = 0;

%pereiname per visus galimus vienos fonemos atvejus, paskaiciuojame rezultatus
for rbi = 1:RBx
    fpradzia = RB(rbi, 1);
    fpabaiga = RB(rbi, 2);
    fr_nr = char(PHR(rbi));
    fonema_pries = fonema_pries_mas(rbi);
    fonema_po = fonema_po_mas(rbi);

    mff_failas = [mff_root '\ sex_root '\ dikt_nr '\ fr_nr '.mff'];

    rezultatai = Fonema_is_Mfcc (mff_failas, fpradzia, fpabaiga, pozymiu_sk);

    if (isempty(rezultatai) == 0)

        [rezx, rezy] = size(rezultatai);

        eil_sk = rezx/39

        fwr = fopen (eilsk_failas_irasymui, 'a');
        fprintf (fwr, '%d', eil_sk);
        fprintf (fwr, '\n');
        fclose (fwr);

        fwr = fopen (all_eilsk_failas, 'a');
        fprintf (fwr, '%d', eil_sk);
        fprintf (fwr, '\t');
        fprintf (fwr, '%s', iesk_fonema);
        fprintf (fwr, '\t');
        fprintf (fwr, '%s', dikt_nr);
        fprintf (fwr, '\n');
        fclose (fwr);

        % sudeliojam rezultatu masyva i atitnkamo formato dvimati masyva
        skait = 0; cx = 1; cy = 0; rez_mas = [];
        for k = 1:rezx
            cy = cy + 1;

            if cy <= 38
                rez_mas (cx, cy) = rezultatai(k);
            else
                rez_mas (cx, cy) = rezultatai(k);
                cx = cx + 1;
                cy = 0;
            end
        end

        index_sum = 0; index_eil = 0;
        %disp ('-->');

        if eil_sk > 60
            fwr = fopen (big_phonemes, 'a');
            fprintf (fwr, '%d', eil_sk);
            fprintf (fwr, '\t');
            fprintf(fwr, '%s', iesk_fonema);
            fprintf(fwr, '\t');
            fprintf(fwr, '%s', dikt_nr);
            fprintf(fwr, '\t');
            fprintf(fwr, '%s', fr_nr);
        end
    end
end

```

```

fprintf(fwr, '\n');
fclose (fwr);
else

% prisiskiriame kurios eilutes reiksmes istatinesime i Mahalanobio formule
if mod(eil_sk, 2) == 0

    if ((eil_sk/2) - 1) > 2
        xud = floor((eil_sk/2 - 1)/2);
    else
        xud = 0;
    end

    cpr = (eil_sk / 2) - xud;
    cpb = (eil_sk / 2) + 1 + xud;

    for i=cpr:cpb
        %rez_mas(i,1)
        index_sum = index_sum + rez_mas(i,:);
        index_eil = index_eil + 1;
    end
    index = index_sum/index_eil;

else
    vid_eil = floor(eil_sk / 2) + 1;

    if (vid_eil - 1) > 1
        xud = floor((vid_eil - 1)/2);
    else
        xud = 0;
    end

    cpr = vid_eil - xud;
    cpb = vid_eil + xud;

    for i=cpr:cpb
        %rez_mas(i,1)
        index_sum = index_sum + rez_mas(i,:);
        index_eil = index_eil + 1;
    end
    index = index_sum/index_eil;

end % rezultatu eiluciu vidurkiu skaiciavimo pabaiga

% apskaiciuojam koeficientus pagal Mahalanobio formule
koefi = (index - AVG) * inv(CM) * transpose((index - AVG));

fwr = fopen (rez_failas_irasymui, 'a');
fprintf (fwr, '%f', koefi);
fprintf (fwr, '\t');
fprintf (fwr, '%c', char(fonema_pries));
fprintf (fwr, '\t');
fprintf (fwr, '%c', char(fonema_po));
fprintf (fwr, '\t');
fprintf (fwr, '%s', fr_nr);
fprintf (fwr, '\n');
fclose (fwr);

koefisum = koefisum + koefi;

end % rezultatu 50 eiluciu tikrinimo pabaiga

```

```

else

    disp(['Fonemos ' iesk_fonema ' (diktorius ' dikt_nr ') rezultatai grazino tuscia aibe']);

    tolog = ['Fonemos ' iesk_fonema ' (diktorius ' dikt_nr ') rezultatai grazino tuscia aibe'];
    log = fopen(log_file, 'a');
    fprintf(log, '%s', tolog);
    fprintf(log, '\n');
    fclose(log);

    [rezx, rezy] = size(rezultatai);

    eil_sk = rezx/39

    %fwr = fopen (eilsk_failas_irasymui, 'a');
    %fprintf (fwr, '%d', eil_sk);
    %fprintf (fwr, '\n');
    %fclose (fwr);

    fwr = fopen (all_eilsk_failas, 'a');
    fprintf (fwr, '%d', eil_sk);
    fprintf (fwr, '\t');
    fprintf (fwr, '%s', iesk_fonema);
    fprintf (fwr, '\t');
    fprintf (fwr, '%s', dikt_nr);
    fprintf (fwr, '\n');
    fclose (fwr);

end

end % fonemos ribu tikrinimo for'o end'as

fwr = fopen (vid_failas_irasymui, 'a');
fprintf (fwr, '%f', koefisum/RBx);
fprintf (fwr, '\t');
fprintf (fwr, '%c', iesk_fonema);
fprintf (fwr, '\n');
fclose (fwr);

else

    disp(['Fonema ' iesk_fonema ' nebuvo rasta diktorius ' dikt_nr ' frazese']);

end

%MX = importdata(rez_failas_irasymui)
%MXDLM = dlmread(rez_failas_irasymui)
%MX2 = char(MX(1)), '\t';

% ***** diktorius tikrinimo pabaiga *****
disp(['Diktorius ' dikt_nr ' apdorotas sekmingai.']);

tolog = ['Diktorius ' dikt_nr ' apdorotas sekmingai.'];
log = fopen(log_file, 'a');
fprintf(log, '%s', tolog);
fprintf(log, '\n');
fclose(log);

end % diktorius for'o end'as

% ***** fonemos tikrinimo pabaiga *****
disp(['Fonema ' iesk_fonema ' apdorota sekmingai.']);
disp(' ');

```

```

tolog = ['Fonema ' iesk_fonema ' apdorota sekmingai.'];
log = fopen(log_file, 'a');
fprintf(log, '%s', tolog);
fprintf(log, '\n');
fclose(log);

end % fonemos for'o end'as

end %Lyties while'o end'as

fclose(fid);
disp ('----- Scripto pabaiga-----');

log = fopen(log_file, 'a');
fprintf(log, '----- Scripto pabaiga-----');
fprintf(log, '\n');
fclose(log);

```

8 PRIEDAS. PROGRAMŲ PALEIDIMO PROGRAMOS IŠEITIS KODAS

```

% ----- Magistro darbo paleidziamasis skriptas -----
close all; clear all; clc;

time0 = clock;

run folders;
run scriptas;
run fonemu_parinkimas;

time1 = clock;

timediff = etime(time1, time0)

```

9 PRIEDAS. EKSPERIMENTO METU GAUTI VYRŲ DIKTORIŲ TARIAMŲ FONEMŲ ĮVERČIAI

M001		M002		M003		M004		M005		M006		M007		M008		M009		M010	
18.576046	a	17.313593	a	17.512604	a	16.298790	a	17.053561	a	14.327118	a	16.453822	a	13.196064	a	18.395470	a	13.833794	a
15.284322	b	15.063263	b	12.141357	b	13.573124	b	13.953602	b	15.824133	b	15.612563	b	11.162712	b	12.820380	b	13.458719	b
9.965335	d	18.677880	d	12.323193	d	11.708168	d	13.497702	d	15.031324	d	14.945641	d	10.183497	d	12.514328	d	13.364883	d
15.239715	e	18.097332	e	15.877563	e	14.784000	e	16.237568	e	12.836847	e	17.659870	e	13.841417	e	15.991822	e	15.356433	e
8.006380	y	9.003720	y	7.930548	y	9.258505	y	6.762499	y	8.061472	y	8.470299	y	9.152628	y	7.455095	y	10.197641	y
12.283097	g	15.170485	g	17.382967	g	15.190743	g	14.587428	g	15.296938	g	11.309449	g	12.783992	g	10.934238	g	13.024025	g
12.550835	i	17.588577	i	15.027232	i	14.349045	i	15.797002	i	11.045798	i	12.900299	i	16.078428	i	15.291704	i	15.032867	i
27.457188	k	22.919207	k	19.553092	k	12.537820	k	13.045429	k	16.063113	k	18.229864	k	36.968343	k	27.197238	k	27.300872	k
19.217597	l	19.087194	l	18.558061	l	19.890219	l	19.752492	l	20.443317	l	19.714906	l	17.154628	l	16.716737	l	15.168916	l
6.422180	m	8.450124	m	4.548716	m	5.382635	m	7.653300	m	4.927584	m	5.431525	m	5.057821	m	NaN	m	7.671376	m
16.080642	n	17.361121	n	20.543777	n	17.845459	n	16.416127	n	17.899725	n	15.999658	n	14.558471	n	15.499666	n	19.972705	n
17.770682	o	22.526731	o	15.769642	o	19.727586	o	23.892774	o	16.477094	o	21.693117	o	22.160686	o	20.490812	o	15.965920	o
29.027354	p	29.129596	p	31.478000	p	36.591851	p	32.523155	p	39.141671	p	41.733366	p	33.341622	p	35.758295	p	40.413740	p
22.205027	r	23.296840	r	26.050345	r	25.458975	r	27.858051	r	23.197682	r	23.824485	r	22.996706	r	28.813802	r	23.458677	r
11.710570	s	13.465025	s	12.207455	s	10.924197	s	14.296372	s	10.677768	s	12.118873	s	10.866682	s	11.241383	s	9.673952	s
11.710570	S	13.465025	S	12.207455	S	10.924197	S	14.296372	S	10.677768	S	12.118873	S	10.866682	S	11.241383	S	9.673952	S
16.247346	t	14.617467	t	17.052015	t	19.512390	t	19.971586	t	13.811343	t	19.209989	t	19.463358	t	18.327886	t	17.408453	t
20.126492	u	20.354964	u	23.395581	u	25.367635	u	26.107437	u	29.577142	u	29.388592	u	16.689994	u	16.099569	u	20.027152	u
18.226924	v	22.229784	v	18.788443	v	15.508064	v	18.463495	v	17.407113	v	16.636299	v	17.026047	v	18.317584	v	21.191444	v
16.120698	z	17.486153	z	26.981953	z	25.612703	z	44.376012	z	18.240104	z	14.200745	z	20.524746	z	14.621219	z	16.223577	z
16.120698	Z	17.486153	Z	26.981953	Z	25.612703	Z	44.376012	Z	18.240104	Z	14.200745	Z	20.524746	Z	14.621219	Z	16.223577	Z

M011		M012		M013		M014		M015		M016		M017		M018		M019		M020	
15.993032	a	15.873953	a	10.982662	a	14.554288	a	14.928924	a	15.993909	a	15.345174	a	16.690699	a	12.295950	a	12.903256	a
12.527449	b	10.773665	b	29.112604	b	14.221604	b	8.923956	b	6.844104	b	10.705628	b	14.252671	b	11.078285	b	9.685281	b
18.456316	d	14.282767	d	9.631016	d	13.150366	d	9.558196	d	12.035308	d	17.753986	d	13.288120	d	15.508795	d	11.216891	d
15.837064	e	17.017698	e	11.260546	e	16.475891	e	16.572937	e	13.849930	e	16.122710	e	15.036917	e	12.743329	e	14.272855	e
8.820717	y	7.182014	y	4.807655	y	8.014157	y	10.162789	y	7.726440	y	6.707181	y	8.537543	y	6.509890	y	7.986341	y
17.205942	g	10.907901	g	18.157160	g	22.542609	g	10.929173	g	10.507979	g	14.074357	g	18.681575	g	11.422526	g	14.463007	g
14.471892	i	12.853672	i	11.930601	i	13.028838	i	16.171104	i	14.029360	i	13.736030	i	12.356365	i	11.062390	i	14.363944	i
15.589338	k	26.885387	k	22.591978	k	18.748603	k	19.929983	k	14.883321	k	12.045302	k	26.514530	k	9.916405	k	17.416672	k
21.225958	l	17.409151	l	10.760599	l	15.644104	l	17.778281	l	20.742623	l	15.589464	l	18.181920	l	14.361288	l	17.486283	l
8.882432	m	5.117733	m	7.002555	m	4.586301	m	3.960789	m	4.164572	m	6.598764	m	6.570825	m	5.910620	m	4.850871	m

M011		M012		M013		M014		M015		M016		M017		M018		M019		M020	
20.331214	n	16.371951	n	14.258607	n	14.338417	n	14.668900	n	13.438558	n	11.876224	n	13.510646	n	12.714058	n	13.220901	n
17.700476	o	17.022202	o	14.459369	o	15.980076	o	15.757036	o	17.001996	o	21.442344	o	15.659652	o	19.657986	o	16.158051	o
27.501350	p	33.430454	p	29.614835	p	35.211376	p	19.233182	p	21.173327	p	28.423604	p	43.116959	p	31.002709	p	19.553438	p
26.510694	r	25.004158	r	24.843493	r	19.885747	r	17.633486	r	15.787843	r	18.040700	r	19.878752	r	16.804796	r	19.681911	r
12.010980	s	8.968040	s	10.798155	s	9.642140	s	9.226119	s	8.608873	s	13.801177	s	12.096850	s	9.412951	s	13.290349	s
12.010980	S	8.968040	S	10.798155	S	9.642140	S	9.226119	S	8.608873	S	13.801177	S	12.096850	S	9.412951	S	13.290349	S
15.951139	t	18.720992	t	16.289939	t	13.537034	t	13.836020	t	11.400459	t	14.077399	t	14.198744	t	11.637714	t	12.749588	t
18.700027	u	15.134832	u	12.357315	u	16.810660	u	17.931166	u	15.276882	u	18.613277	u	18.453805	u	19.049862	u	20.219412	u
20.954098	v	17.996190	v	14.103374	v	19.287168	v	13.166349	v	13.450921	v	14.789737	v	15.546464	v	14.695959	v	15.643545	v
21.796178	z	14.606178	z	18.758998	z	27.196406	z	23.462873	z	26.351633	z	31.969092	z	12.160617	z	13.036248	z	21.963457	z
21.796178	Z	14.606178	Z	18.758998	Z	27.196406	Z	23.462873	Z	26.351633	Z	31.969092	Z	12.160617	Z	13.036248	Z	21.963457	Z

M021		M022		M023		M024		M025		M026		M027		M028		M029		M030	
19.015141	a	18.237598	a	20.377880	a	17.026745	a	16.044168	a	14.544602	a	19.192644	a	22.656697	a	22.385467	a	21.122518	a
15.254017	b	11.624723	b	18.211842	b	8.152959	b	14.795465	b	10.696972	b	9.224857	b	10.395733	b	6.813230	b	7.673647	b
14.205823	d	22.993748	d	13.809356	d	11.310431	d	16.024997	d	10.536123	d	13.307638	d	11.063216	d	12.667665	d	16.593475	d
16.372214	e	20.654094	e	18.647865	e	18.137871	e	17.420334	e	17.178034	e	22.857772	e	23.489837	e	19.740228	e	23.374336	e
9.951772	y	11.713007	y	14.704197	y	12.220009	y	8.159728	y	8.229355	y	12.262671	y	9.649149	y	8.169855	y	8.079097	y
22.210327	g	18.588116	g	17.021678	g	13.155140	g	17.724087	g	12.761636	g	15.451861	g	12.891282	g	10.242424	g	11.176543	g
17.766589	i	20.197900	i	19.685874	i	17.596346	i	14.252592	i	15.459968	i	18.667450	i	22.347844	i	16.152463	i	18.096503	i
23.350799	k	26.085385	k	18.685666	k	17.408076	k	20.569983	k	14.681355	k	23.484183	k	16.969023	k	19.910424	k	20.648981	k
24.956164	l	18.298545	l	17.835698	l	22.615991	l	16.152608	l	21.087977	l	21.633946	l	20.694142	l	20.035400	l	23.478367	l
16.686328	m	7.146236	m	6.767811	m	5.722508	m	5.470494	m	3.841922	m	7.033816	m	8.035412	m	7.073214	m	5.664084	m
24.648556	n	19.108892	n	20.092258	n	19.057266	n	17.693789	n	15.568214	n	17.613923	n	17.384320	n	13.707772	n	14.609118	n
19.120538	o	15.431278	o	23.161461	o	25.703203	o	21.136567	o	17.439109	o	16.201261	o	24.052379	o	27.372261	o	28.448981	o
30.234356	p	27.929588	p	31.885422	p	30.599764	p	38.631484	p	35.176265	p	32.034759	p	30.161233	p	30.972467	p	35.100972	p
25.732325	r	21.493814	r	27.120281	r	20.379438	r	22.256319	r	23.781810	r	26.760083	r	25.465299	r	21.285112	r	23.394159	r
10.689963	s	14.484023	s	12.422448	s	11.179519	s	10.456686	s	11.415652	s	12.438842	s	13.210884	s	11.238190	s	12.056226	s
10.689963	S	14.484023	S	12.422448	S	11.179519	S	10.456686	S	11.415652	S	12.438842	S	13.210884	S	11.238190	S	12.056226	S
19.154660	t	15.192729	t	14.977176	t	14.282800	t	19.689495	t	15.739957	t	17.772759	t	15.060893	t	16.384290	t	15.859003	t
19.102814	u	19.575387	u	21.227704	u	25.030444	u	22.421729	u	16.770118	u	16.870433	u	21.864455	u	16.041858	u	25.345689	u
18.239386	v	19.213761	v	17.898898	v	13.330838	v	22.198316	v	18.413552	v	15.593635	v	20.088122	v	20.918967	v	19.147333	v
19.687051	z	19.501124	z	19.124275	z	13.027286	z	13.216224	z	15.713149	z	21.862281	z	28.118268	z	29.807579	z	16.720867	z
19.687051	Z	19.501124	Z	19.124275	Z	13.027286	Z	13.216224	Z	15.713149	Z	21.862281	Z	28.118268	Z	29.807579	Z	16.720867	Z

M031		M032		M033		M034		M035		M036		M037		M038		M039		M040	
16.647285	a	17.788614	a	18.009916	a	18.613443	a	17.758581	a	20.274400	a	19.722154	a	18.656296	a	18.874226	a	19.608572	a
10.286962	b	8.073740	b	9.482074	b	12.086522	b	13.078201	b	9.053669	b	11.017133	b	13.853001	b	15.808524	b	10.056094	b
7.190567	d	9.846248	d	11.472934	d	24.122524	d	11.772996	d	11.312373	d	11.788322	d	18.136506	d	16.095130	d	14.330069	d
18.441972	e	18.375882	e	19.386716	e	16.719548	e	17.379511	e	19.396300	e	16.569690	e	21.732895	e	24.200160	e	18.351780	e
8.236024	y	9.883476	y	13.758486	y	9.388971	y	10.387214	y	11.394072	y	8.425935	y	11.211954	y	9.773505	y	10.191041	y
14.236479	g	11.884511	g	15.754358	g	17.577153	g	20.331067	g	15.117287	g	13.607450	g	20.460729	g	19.403188	g	17.207834	g
17.071277	i	15.863168	i	20.242854	i	16.835709	i	15.992772	i	17.587786	i	14.447095	i	17.190661	i	21.577130	i	17.614250	i
15.012977	k	20.061554	k	15.981264	k	27.092111	k	16.858945	k	17.155778	k	12.479298	k	22.601106	k	20.613863	k	21.241023	k
16.168998	l	15.351564	l	14.968553	l	16.753583	l	18.481919	l	12.921490	l	18.455969	l	20.384069	l	18.740074	l	17.084064	l
3.712084	m	7.187362	m	6.064510	m	5.611182	m	4.277247	m	5.079887	m	5.870958	m	4.981749	m	5.874182	m	8.469539	m
14.293235	n	18.382980	n	13.242257	n	20.876411	n	15.529638	n	16.627017	n	22.082796	n	14.207959	n	22.515108	n	17.937766	n
21.589994	o	18.948577	o	25.161413	o	26.197692	o	19.968449	o	21.027492	o	19.864365	o	23.023952	o	21.835472	o	19.783144	o
33.367491	p	38.611471	p	28.442757	p	24.063477	p	20.899941	p	29.539305	p	23.078118	p	26.924118	p	31.244175	p	24.960899	p
26.308201	r	21.408947	r	19.525517	r	26.514842	r	21.981626	r	19.546881	r	22.437505	r	22.391960	r	28.850026	r	18.865576	r
12.163318	s	10.452509	s	15.103216	s	13.881469	s	13.684895	s	9.614610	s	11.628416	s	9.478729	s	17.063964	s	12.914435	s
12.163318	S	10.452509	S	15.103216	S	13.881469	S	13.684895	S	9.614610	S	11.628416	S	9.478729	S	17.063964	S	12.914435	S
15.787635	t	16.675845	t	19.750147	t	16.636331	t	17.052511	t	17.553817	t	14.368942	t	14.595544	t	17.908524	t	12.853152	t
18.593538	u	16.032732	u	23.828415	u	22.393912	u	19.223432	u	17.436006	u	20.617642	u	22.653243	u	26.670714	u	21.559368	u
12.833061	v	10.547564	v	17.034669	v	15.670782	v	15.749832	v	9.894515	v	13.592972	v	16.623729	v	15.630423	v	18.956331	v
22.825426	z	15.554279	z	25.898784	z	13.799721	z	15.782411	z	23.843402	z	17.323688	z	20.756937	z	24.610085	z	12.287846	z
22.825426	Z	15.554279	Z	25.898784	Z	13.799721	Z	15.782411	Z	23.843402	Z	17.323688	Z	20.756937	Z	24.610085	Z	12.287846	Z

M041		M042		M043		M044		M045		M046		M047		M048		M049		M050	
15.954019	a	18.553266	a	17.583552	a	23.189695	a	15.453113	a	17.165797	a	17.903162	a	20.490405	a	17.326603	a	15.006337	a
15.129515	b	8.085226	b	12.952751	b	11.114001	b	8.188414	b	11.014822	b	14.804528	b	8.175846	b	8.190188	b	6.956460	b
13.585924	d	9.435367	d	14.592992	d	12.880981	d	12.730208	d	11.986974	d	15.740957	d	11.746969	d	11.982079	d	11.698783	d
14.443769	e	17.957410	e	15.518173	e	20.946965	e	16.014771	e	15.165097	e	13.926315	e	20.747753	e	16.500866	e	14.768552	e
8.018868	y	8.409460	y	8.644812	y	9.413464	y	9.966110	y	6.924354	y	7.128971	y	10.615895	y	9.750366	y	12.131039	y
11.056509	g	10.571317	g	19.412639	g	8.564517	g	8.310915	g	15.043252	g	17.931755	g	14.206591	g	21.457632	g	12.631619	g
13.748482	i	18.113984	i	16.501812	i	22.117411	i	13.330093	i	14.717260	i	17.561824	i	21.654723	i	17.140922	i	16.031135	i
21.986035	k	17.106906	k	17.547669	k	18.412453	k	11.751749	k	20.019435	k	14.981025	k	25.318779	k	24.775273	k	26.308362	k
13.214964	l	18.038482	l	20.449621	l	21.598675	l	14.707223	l	17.323050	l	17.640725	l	20.377323	l	21.750703	l	19.352400	l
4.745658	m	5.269356	m	4.957362	m	5.699851	m	9.671949	m	4.678219	m	11.223813	m	8.474458	m	6.519883	m	8.104093	m
15.802246	n	14.163132	n	12.322618	n	16.600811	n	14.372667	n	15.898424	n	17.291051	n	22.920047	n	13.458907	n	15.583955	n

M041		M042		M043		M044		M045		M046		M047		M048		M049		M050	
19.802059	o	16.665943	o	22.214042	o	14.971028	o	15.618009	o	20.205636	o	14.450788	o	18.383818	o	16.446375	o	18.886122	o
33.059113	p	24.800541	p	29.483947	p	25.049827	p	31.793469	p	32.420307	p	24.357472	p	33.696796	p	32.397013	p	35.528873	p
23.599993	r	26.360992	r	20.845681	r	21.942279	r	21.893365	r	19.084124	r	24.777729	r	24.278158	r	18.523632	r	26.028546	r
10.730848	s	9.516824	s	14.181307	s	13.765706	s	9.000143	s	10.160871	s	11.165083	s	12.839924	s	9.773551	s	10.458936	s
10.730848	S	9.516824	S	14.181307	S	13.765706	S	9.000143	S	10.160871	S	11.165083	S	12.839924	S	9.773551	S	10.458936	S
16.517411	t	14.899478	t	16.117473	t	14.411105	t	19.752921	t	14.116861	t	14.283602	t	18.900100	t	19.551940	t	17.192144	t
20.507338	u	19.996360	u	19.664461	u	21.446600	u	18.469248	u	15.800527	u	14.670275	u	24.416341	u	21.102915	u	17.573099	u
14.834303	v	9.121818	v	12.882460	v	20.413873	v	11.440585	v	16.140242	v	19.315617	v	19.777044	v	20.804911	v	21.502143	v
13.579088	z	19.049852	z	12.665245	z	25.937663	z	14.715933	z	16.755690	z	13.747600	z	25.608986	z	15.794743	z	17.746541	z
13.579088	Z	19.049852	Z	12.665245	Z	25.937663	Z	14.715933	Z	16.755690	Z	13.747600	Z	25.608986	Z	15.794743	Z	17.746541	Z

10 PRIEDAS. EKSPERIMENTO METU GAUTI MOTERŲ DIKTORIŲ TARIAMŲ FONEMŲ ĮVERČIAI

F001		F002		F003		F004		F005		F006		F007		F008		F009		F010	
14.761123	a	16.783375	a	14.318348	a	12.018848	a	16.230449	a	17.277418	a	16.670912	a	15.136385	a	19.195387	a	13.582768	a
7.662653	b	9.788700	b	8.812167	b	9.422276	b	5.895426	b	10.638745	b	11.500507	b	6.366044	b	7.879231	b	7.815959	b
16.799063	d	21.354421	d	12.409323	d	8.943830	d	7.997440	d	10.046071	d	20.285821	d	9.278836	d	9.662867	d	6.113419	d
15.171266	e	14.009808	e	17.111982	e	13.400383	e	14.524320	e	17.835858	e	17.161795	e	14.723216	e	14.130780	e	14.820110	e
8.856630	y	12.023363	y	8.811379	y	7.238338	y	8.353773	y	9.132852	y	11.646324	y	8.681918	y	9.748707	y	7.935068	y
11.248570	g	13.401835	g	24.680255	g	13.242391	g	8.872653	g	11.019577	g	10.666869	g	11.007373	g	6.527698	g	12.057650	g
14.676777	i	17.639361	i	16.799775	i	14.210133	i	13.600780	i	15.244354	i	18.565390	i	13.170886	i	12.618789	i	11.810148	i
18.796139	k	16.651326	k	12.427202	k	25.394368	k	29.161400	k	20.278248	k	19.094877	k	17.110629	k	7.104478	k	20.721556	k
14.251715	l	14.717371	l	11.731246	l	9.495920	l	12.986784	l	12.644690	l	21.421097	l	17.535365	l	13.767973	l	14.953933	l
5.552073	m	4.428965	m	4.511932	m	5.346261	m	7.632954	m	5.652188	m	3.881047	m	7.071971	m	4.594611	m	4.090811	m
13.790524	n	11.152218	n	15.072710	n	10.296120	n	12.201963	n	13.297265	n	18.610521	n	11.475882	n	10.811222	n	11.502601	n
13.400047	o	14.412142	o	18.578800	o	16.112202	o	16.528455	o	18.041241	o	18.447120	o	21.340048	o	18.234709	o	21.164295	o
24.753828	p	31.462794	p	32.051134	p	29.844054	p	42.901841	p	39.080116	p	29.221083	p	26.536694	p	28.835152	p	29.890489	p
21.185955	r	21.572793	r	19.728729	r	16.153211	r	22.054735	r	22.320008	r	21.202993	r	24.872637	r	16.299769	r	21.559874	r
13.666632	s	9.855284	s	14.467697	s	9.709476	s	9.314396	s	9.817607	s	17.216114	s	12.401699	s	10.923582	s	9.201043	s
13.666632	S	9.855284	S	14.467697	S	9.709476	S	9.314396	S	9.817607	S	17.216114	S	12.401699	S	10.923582	S	9.201043	S
11.244298	t	11.503835	t	14.382763	t	15.097450	t	19.494220	t	11.474633	t	15.259759	t	15.232622	t	17.037646	t	11.528773	t
15.274567	u	17.499273	u	23.826559	u	17.740130	u	19.909158	u	15.926916	u	16.859402	u	17.417111	u	16.835125	u	16.369331	u
14.513698	v	10.621441	v	20.084839	v	12.948040	v	7.282774	v	19.854864	v	18.686825	v	10.744004	v	8.814397	v	10.861323	v
20.109308	z	26.338677	z	12.929346	z	28.655805	z	11.346248	z	19.862047	z	23.199376	z	19.069965	z	10.395314	z	13.037571	z

F001		F002		F003		F004		F005		F006		F007		F008		F009		F010	
20.109308	Z	26.338677	Z	12.929346	Z	28.655805	Z	11.346248	Z	19.862047	Z	23.199376	Z	19.069965	Z	10.395314	Z	13.037571	Z

F011		F012		F013		F014		F015		F016		F017		F018		F019		F020	
13.098696	a	14.874856	a	18.073191	a	17.401876	a	15.062943	a	16.449498	a	19.115633	a	16.566462	a	15.418060	a	24.006616	a
5.829775	b	9.405937	b	7.855479	b	7.809061	b	5.338360	b	9.175650	b	10.452485	b	9.855854	b	8.737424	b	5.416558	b
7.946329	d	10.890032	d	10.031182	d	11.161703	d	11.229759	d	12.905232	d	8.110455	d	12.714158	d	8.153423	d	13.718227	d
12.826928	e	10.836855	e	16.304877	e	15.964228	e	14.293510	e	14.210652	e	15.795144	e	11.968254	e	14.156701	e	16.538033	e
10.097757	y	3.678796	y	5.405415	y	8.793339	y	8.416694	y	10.610242	y	11.010250	y	10.496848	y	8.918593	y	25.825140	g
10.410669	g	10.991503	g	10.914892	g	6.746072	g	10.135017	g	9.927421	g	11.246217	g	13.552433	g	12.751619	g	17.481696	i
11.652537	i	9.183030	i	12.663494	i	14.491504	i	9.889768	i	12.624721	i	17.304968	i	14.004133	i	12.805211	i	27.228952	k
26.075212	k	25.473503	k	22.782786	k	25.814255	k	22.661339	k	24.087562	k	22.563536	k	11.261335	k	8.383932	k	24.528898	l
11.826616	l	14.358153	l	17.159470	l	16.715806	l	15.353430	l	15.884601	l	14.674692	l	12.073690	l	14.738513	l	4.588453	m
5.470603	m	4.841058	m	2.872924	m	10.208111	m	4.543079	m	7.129193	m	5.321507	m	4.354019	m	3.642191	m	10.752193	n
10.939723	n	9.208628	n	9.352123	n	9.748991	n	10.772116	n	13.171103	n	14.448839	n	12.522636	n	13.680083	n	25.890571	o
12.871160	o	10.533194	o	12.702211	o	19.672482	o	20.676278	o	14.542379	o	13.778589	o	15.061131	o	16.656260	o	27.687230	p
35.484559	p	38.593910	p	40.181387	p	28.930789	p	33.349819	p	33.688696	p	33.108826	p	31.742496	p	31.760290	p	21.289322	r
24.297565	r	23.788374	r	21.450126	r	25.408553	r	24.785185	r	20.658676	r	18.980805	r	22.496564	r	19.916937	r	11.630184	s
11.837233	s	9.044116	s	9.759874	s	9.754576	s	10.497906	s	9.540281	s	10.806296	s	12.789122	s	11.072828	s	11.630184	S
11.837233	S	9.044116	S	9.759874	S	9.754576	S	10.497906	S	9.540281	S	10.806296	S	12.789122	S	11.072828	S	13.473900	t
15.917415	t	15.052835	t	14.365799	t	17.573277	t	15.667546	t	17.693151	t	14.735359	t	13.338936	t	20.295067	t	19.236760	u
11.323154	u	11.336175	u	18.356253	u	20.651474	u	12.310588	u	20.239280	u	18.290918	u	17.358223	u	19.468048	u	10.506724	v
15.629004	v	10.660730	v	9.116696	v	13.419920	v	14.834235	v	15.393307	v	14.900804	v	13.838221	v	14.134611	v	20.359322	z
11.810344	z	13.905407	z	13.238016	z	14.960103	z	21.082610	z	15.289120	z	11.319536	z	13.435139	z	11.712854	z	20.359322	Z
11.810344	Z	13.905407	Z	13.238016	Z	14.960103	Z	21.082610	Z	15.289120	Z	11.319536	Z	13.435139	Z	11.712854	Z		

F021		F022		F023		F024		F025		F026		F027		F028		F029		F030	
15.590147	a	13.304735	a	14.765464	a	16.268247	a	12.969738	a	16.090484	a	13.869207	a	13.876230	a	13.951191	a	17.743506	a
11.835675	b	5.392116	b	10.678661	b	11.702812	b	5.926934	b	11.573227	b	11.941958	b	11.048361	b	8.286073	b	6.406365	b
16.395999	d	12.328464	d	11.715795	d	9.105633	d	8.725844	d	16.014940	d	13.205678	d	13.846697	d	9.186412	d	6.465904	d
13.802088	e	13.658436	e	14.045422	e	13.608070	e	13.418116	e	14.916013	e	13.742881	e	13.183946	e	14.018550	e	13.167022	e
8.936643	y	7.601666	y	7.223862	y	9.998429	y	8.030251	y	9.787212	y	6.024421	y	7.294059	y	7.284415	y	5.734566	y
10.654011	g	8.605028	g	12.034764	g	11.771410	g	12.508127	g	10.844525	g	14.928063	g	8.123441	g	9.762526	g	9.093265	g
11.572008	i	11.575256	i	12.629774	i	13.349443	i	12.061403	i	15.585751	i	12.430452	i	12.830349	i	10.642665	i	12.966800	i
33.319621	k	14.390006	k	30.656865	k	24.068317	k	17.171520	k	19.801600	k	21.169428	k	16.902047	k	26.022994	k	21.882421	k
14.643173	l	11.103701	l	16.533340	l	14.684290	l	15.066453	l	14.323501	l	16.198847	l	11.422008	l	14.188059	l	16.883014	l

F021		F022		F023		F024		F025		F026		F027		F028		F029		F030	
3.905685	m	4.781950	m	6.556449	m	4.494494	m	3.804074	m	4.728649	m	4.822858	m	5.926770	m	4.416614	m	4.220299	m
9.615078	n	7.985626	n	14.354306	n	11.547936	n	9.594289	n	13.169216	n	12.759718	n	8.730822	n	13.931013	n	10.256599	n
17.639818	o	15.917687	o	11.431664	o	14.213994	o	12.722866	o	21.330242	o	15.149797	o	15.996817	o	14.144295	o	16.353592	o
52.935930	p	31.980756	p	46.121514	p	43.175438	p	36.503890	p	47.429545	p	39.458642	p	33.207120	p	36.923782	p	40.702458	p
21.332651	r	20.168969	r	26.850939	r	22.857767	r	26.124397	r	22.787803	r	20.843130	r	18.422135	r	19.890076	r	17.688703	r
10.180504	s	8.787356	s	11.461006	s	11.504944	s	8.826548	s	9.276783	s	8.351484	s	9.789269	s	9.189573	s	9.649607	s
10.180504	S	8.787356	S	11.461006	S	11.504944	S	8.826548	S	9.276783	S	8.351484	S	9.789269	S	9.189573	S	9.649607	S
13.017465	t	13.262227	t	21.432488	t	16.294248	t	10.610644	t	12.358423	t	17.309171	t	16.141098	t	18.655266	t	14.744764	t
20.278820	u	12.958678	u	16.907460	u	15.166925	u	14.270897	u	18.930522	u	13.674808	u	14.933779	u	14.000934	u	19.447414	u
14.728304	v	11.474380	v	12.483224	v	13.339597	v	8.560687	v	14.414935	v	17.488787	v	11.865594	v	11.721701	v	17.716598	v
17.648241	z	16.974162	z	15.030128	z	23.080237	z	13.847822	z	10.486326	z	13.414774	z	17.460381	z	10.690535	z	8.382227	z
17.648241	Z	16.974162	Z	15.030128	Z	23.080237	Z	13.847822	Z	10.486326	Z	13.414774	Z	17.460381	Z	10.690535	Z	8.382227	Z

F031		F032		F033		F034		F035		F036		F037		F038		F039		F040	
15.250232	a	14.536689	a	15.856392	a	15.273831	a	14.613118	a	13.901137	a	17.602546	a	14.922876	a	14.168355	a	17.731191	a
7.130530	b	9.277153	b	13.569833	b	9.732187	b	8.721326	b	7.578990	b	11.585018	b	7.048153	b	5.924542	b	5.421475	b
11.135199	d	13.341740	d	14.467894	d	10.718744	d	9.116321	d	5.842183	d	19.097047	d	12.368998	d	17.697473	d	10.267411	d
12.680096	e	13.618126	e	21.975062	e	13.835930	e	14.926751	e	12.427560	e	13.191203	e	14.857218	e	11.894335	e	13.003989	e
7.698892	y	7.822631	y	9.725302	y	8.910883	y	5.374511	y	7.592287	y	8.564750	y	8.960341	y	8.672258	y	8.740716	y
12.824175	g	11.776138	g	12.383046	g	10.768585	g	11.798409	g	13.485902	g	12.000936	g	11.561876	g	11.270345	g	13.714889	g
12.874575	i	13.184675	i	16.531209	i	13.235058	i	13.742633	i	13.575929	i	14.898436	i	13.054953	i	13.705972	i	16.878942	i
21.540473	k	19.605793	k	18.192716	k	23.036172	k	23.886879	k	21.802128	k	15.793670	k	15.774891	k	20.267639	k	25.215669	k
13.491682	l	14.958207	l	24.094682	l	16.528251	l	12.088644	l	16.990036	l	21.992846	l	13.992284	l	12.465199	l	17.178171	l
6.899133	m	5.665630	m	4.609661	m	5.947280	m	6.601171	m	5.508544	m	6.275132	m	5.061820	m	4.595106	m	3.493473	m
12.320153	n	13.245306	n	21.420149	n	11.758094	n	9.501608	n	10.556679	n	13.467889	n	11.250361	n	11.363399	n	14.231495	n
22.347735	o	16.969177	o	23.558551	o	26.786107	o	17.340268	o	22.322650	o	15.770614	o	16.821225	o	21.417085	o	17.874956	o
36.552744	p	25.973440	p	33.819656	p	37.770151	p	32.638213	p	35.548567	p	30.055454	p	36.162921	p	24.477174	p	33.842934	p
19.099574	r	20.738902	r	28.667803	r	21.313008	r	15.649707	r	18.133307	r	20.562597	r	24.268443	r	23.700257	r	17.483406	r
10.784493	s	9.650420	s	11.799304	s	12.355213	s	10.750087	s	9.472494	s	12.489676	s	12.673562	s	8.822499	s	15.941511	s
10.784493	S	9.650420	S	11.799304	S	12.355213	S	10.750087	S	9.472494	S	12.489676	S	12.673562	S	8.822499	S	15.941511	S
14.035593	t	16.189039	t	17.160657	t	17.576182	t	12.339609	t	14.267128	t	20.438427	t	18.393464	t	18.320735	t	17.766812	t
18.652864	u	17.705933	u	22.103528	u	21.356579	u	18.460298	u	13.357533	u	17.271448	u	14.052346	u	16.285726	u	17.401215	u
11.844008	v	11.542678	v	21.237923	v	14.652812	v	14.148189	v	5.530577	v	9.968330	v	12.892561	v	16.094987	v	13.478569	v
13.225017	z	14.899253	z	21.375018	z	20.467139	z	10.732530	z	11.991570	z	14.923077	z	10.885653	z	8.414934	z	18.136206	z
13.225017	Z	14.899253	Z	21.375018	Z	20.467139	Z	10.732530	Z	11.991570	Z	14.923077	Z	10.885653	Z	8.414934	Z	18.136206	Z

F041		F042		F043		F044		F045		F046		F047		F048		F049		F050	
13.492425	a	13.327024	a	15.245957	a	13.871027	a	14.831682	a	13.235422	a	11.965397	a	13.579245	a	12.426483	a	17.814021	a
6.757023	b	7.127461	b	10.757139	b	7.509944	b	9.181764	b	7.418264	b	8.025045	b	12.596364	b	4.951814	b	8.588416	b
7.957637	d	5.924184	d	8.422087	d	7.869770	d	13.973156	d	9.686838	d	10.669152	d	13.916519	d	6.932028	d	11.525790	d
12.512865	e	12.495920	e	10.306291	e	12.080779	e	15.416344	e	13.152704	e	14.183682	e	15.588993	e	12.662485	e	13.043206	e
10.346568	y	6.898942	y	6.634114	y	7.791243	y	8.103708	y	8.491994	y	7.222700	y	6.703756	y	6.914649	y	12.344903	y
9.313230	g	10.082478	g	10.021927	g	14.724021	g	10.934686	g	11.743327	g	7.766362	g	10.572089	g	10.946736	g	11.571350	g
14.346679	i	11.592526	i	11.590283	i	12.216253	i	15.459116	i	11.923774	i	10.780305	i	11.965923	i	10.163342	i	13.487710	i
26.456518	k	34.075300	k	13.185450	k	23.566357	k	17.744417	k	18.988404	k	14.729321	k	25.662960	k	16.797960	k	22.090622	k
13.787068	l	9.273495	l	16.581827	l	10.399469	l	18.355286	l	13.912791	l	10.601952	l	15.339240	l	13.752433	l	16.066397	l
4.547226	m	4.845568	m	4.332260	m	3.595760	m	6.452136	m	4.655450	m	4.737510	m	5.607662	m	4.673070	m	6.394572	m
12.084000	n	12.186041	n	12.354739	n	11.891553	n	14.894383	n	11.288787	n	11.155823	n	12.359928	n	12.634072	n	12.991461	n
14.994168	o	14.764117	o	14.075469	o	14.621516	o	21.470762	o	17.926941	o	15.342634	o	17.029718	o	15.788344	o	20.354770	o
33.717006	p	40.080990	p	37.175892	p	25.382502	p	32.703838	p	34.495957	p	35.979423	p	26.250855	p	38.061689	p	33.632703	p
20.261523	r	23.510735	r	19.823262	r	18.255030	r	19.941952	r	20.870821	r	20.005083	r	29.549495	r	20.350477	r	23.492812	r
8.316562	s	9.378208	s	9.012503	s	9.414615	s	10.796832	s	9.643820	s	11.395109	s	9.434249	s	8.421745	s	10.951969	s
8.316562	S	9.378208	S	9.012503	S	9.414615	S	10.796832	S	9.643820	S	11.395109	S	9.434249	S	8.421745	S	10.951969	S
14.520838	t	25.569571	t	16.203736	t	13.095107	t	15.467430	t	15.530592	t	13.270244	t	12.212844	t	14.506500	t	14.234886	t
13.412238	u	14.945695	u	14.153161	u	14.474985	u	18.091526	u	19.068354	u	11.114542	u	14.518694	u	14.907165	u	21.142970	u
12.339025	v	12.132618	v	11.541205	v	14.555594	v	16.975088	v	11.958462	v	18.115713	v	19.884270	v	8.197300	v	13.879909	v
9.696302	z	10.559596	z	23.779622	z	15.086871	z	12.054202	z	8.973443	z	12.767395	z	10.531680	z	10.583359	z	12.856641	z
9.696302	Z	10.559596	Z	23.779622	Z	15.086871	Z	12.054202	Z	8.973443	Z	12.767395	Z	10.531680	Z	10.583359	Z	12.856641	Z

11 PRIEDAS. IŠRIKIUOTI VYRŲ DIKTORIŲ FONEMŲ PANAŠUMŲ ĮVERČIAI

M001		M002		M003		M004		M005		M006		M007		M008		M009		M010	
29,02735	p	29,1296	p	31,478	p	36,59185	p	44,37601	z	39,14167	p	41,73337	p	36,96834	k	NaN	m	40,41374	p
27,45719	k	23,29684	r	26,98195	z	25,6127	z	44,37601	Z	29,57714	u	29,38859	u	33,34162	p	35,7583	p	27,30087	k
22,20503	r	22,91921	k	26,98195	Z	25,6127	Z	32,52316	p	23,19768	r	23,82449	r	22,99671	r	28,8138	r	23,45868	r
20,12649	u	22,52673	o	26,05035	r	25,45898	r	27,85805	r	20,44332	l	21,69312	o	22,16069	o	27,19724	k	21,19144	v
19,2176	l	22,22978	v	23,39558	u	25,36764	u	26,10744	u	18,2401	z	19,71491	l	20,52475	z	20,49081	o	20,02715	u
18,57605	a	20,35496	u	20,54378	n	19,89022	l	23,89277	o	18,2401	Z	19,20999	t	20,52475	Z	18,39547	a	19,97271	n
18,22692	v	19,08719	l	19,55309	k	19,72759	o	19,97159	t	17,89973	n	18,22986	k	19,46336	t	18,32789	t	17,40845	t
17,77068	o	18,67788	d	18,78844	v	19,51239	t	19,75249	l	17,40711	v	17,65987	e	17,15463	l	18,31758	v	16,22358	z
16,24735	t	18,09733	e	18,55806	l	17,84546	n	18,4635	v	16,47709	o	16,6363	v	17,02605	v	16,71674	l	16,22358	Z
16,1207	z	17,58858	i	17,5126	a	16,29879	a	17,05356	a	16,06311	k	16,45382	a	16,68999	u	16,09957	u	15,96592	o
16,1207	Z	17,48615	z	17,38297	g	15,50806	v	16,41613	n	15,82413	b	15,99966	n	16,07843	i	15,99182	e	15,35643	e
16,08064	n	17,48615	Z	17,05202	t	15,19074	g	16,23757	e	15,29694	g	15,61256	b	14,55847	n	15,49967	n	15,16892	l
15,28432	b	17,36112	n	15,87756	e	14,784	e	15,797	i	15,03132	d	14,94564	d	13,84142	e	15,2917	i	15,03287	i
15,23972	e	17,31359	a	15,76964	o	14,34905	i	14,58743	g	14,32712	a	14,20075	z	13,19606	a	14,62122	z	13,83379	a
12,55084	i	15,17049	g	15,02723	i	13,57312	b	14,29637	s	13,81134	t	14,20075	Z	12,78399	g	14,62122	Z	13,45872	b
12,2831	g	15,06326	b	12,32319	d	12,53782	k	14,29637	S	12,83685	e	12,9003	i	11,16271	b	12,82038	b	13,36488	d
11,71057	s	14,61747	t	12,20746	s	11,70817	d	13,9536	b	11,0458	i	12,11887	s	10,86668	s	12,51433	d	13,02403	g
11,71057	S	13,46503	s	12,20746	S	10,9242	s	13,4977	d	10,67777	s	12,11887	S	10,86668	S	11,24138	s	10,19764	y
9,965335	d	13,46503	S	12,14136	b	10,9242	S	13,04543	k	10,67777	S	11,30945	g	10,1835	d	11,24138	S	9,673952	s
8,00638	y	9,00372	y	7,930548	y	9,258505	y	7,6533	m	8,061472	y	8,470299	y	9,152628	y	10,93424	g	9,673952	S
6,42218	m	8,450124	m	4,548716	m	5,382635	m	6,762499	y	4,927584	m	5,431525	m	5,057821	m	7,455095	y	7,671376	m

M011		M012		M013		M014		M015		M016		M017		M018		M019		M020	
27,50135	p	33,43045	p	29,61484	p	35,21138	p	23,46287	z	26,35163	z	31,96909	z	43,11696	p	31,00271	p	21,96346	z
26,51069	r	26,88539	k	29,1126	b	27,19641	z	23,46287	Z	26,35163	Z	31,96909	Z	26,51453	k	19,65799	o	21,96346	Z
21,79618	z	25,00416	r	24,84349	r	27,19641	Z	19,92998	k	21,17333	p	28,4236	p	19,87875	r	19,04986	u	20,21941	u
21,79618	Z	18,72099	t	22,59198	k	22,54261	g	19,23318	p	20,74262	l	21,44234	o	18,68158	g	16,8048	r	19,68191	r
21,22596	l	17,99619	v	18,759	z	19,88575	r	17,93117	u	17,002	o	18,61328	u	18,45381	u	15,5088	d	19,55344	p
20,9541	v	17,40915	l	18,759	Z	19,28717	v	17,77828	l	15,99391	a	18,0407	r	18,18192	l	14,69596	v	17,48628	l
20,33121	n	17,0222	o	18,15716	g	18,7486	k	17,63349	r	15,78784	r	17,75399	d	16,6907	a	14,36129	l	17,41667	k
18,70003	u	17,0177	e	16,28994	t	16,81066	u	16,57294	e	15,27688	u	16,12271	e	15,65965	o	13,03625	z	16,15805	o
18,45632	d	16,37195	n	14,45937	o	16,47589	e	16,1711	i	14,88332	k	15,58946	l	15,54646	v	13,03625	Z	15,64355	v

M011		M012		M013		M014		M015		M016		M017		M018		M019		M020	
17,70048	o	15,87395	a	14,25861	n	15,98008	o	15,75704	o	14,02936	i	15,34517	a	15,03692	e	12,74333	e	14,46301	g
17,20594	g	15,13483	u	14,10337	v	15,6441	l	14,92892	a	13,84993	e	14,78974	v	14,25267	b	12,71406	n	14,36394	i
15,99303	a	14,60618	z	12,35732	u	14,55429	a	14,6689	n	13,45092	v	14,0774	t	14,19874	t	12,29595	a	14,27286	e
15,95114	t	14,60618	Z	11,9306	i	14,33842	n	13,83602	t	13,43856	n	14,07436	g	13,51065	n	11,63771	t	13,29035	s
15,83706	e	14,28277	d	11,26055	e	14,2216	b	13,16635	v	12,03531	d	13,80118	s	13,28812	d	11,42253	g	13,29035	S
15,58934	k	12,85367	i	10,98266	a	13,53703	t	10,92917	g	11,40046	t	13,80118	S	12,35637	i	11,07829	b	13,2209	n
14,47189	i	10,9079	g	10,79816	s	13,15037	d	10,16279	y	10,50798	g	13,73603	i	12,16062	z	11,06239	i	12,90326	a
12,52745	b	10,77367	b	10,79816	S	13,02884	i	9,558196	d	8,608873	s	12,0453	k	12,16062	Z	9,916405	k	12,74959	t
12,01098	s	8,96804	s	10,7606	l	9,64214	s	9,226119	s	8,608873	S	11,87622	n	12,09685	s	9,412951	s	11,21689	d
12,01098	S	8,96804	S	9,631016	d	9,64214	S	9,226119	S	7,72644	y	10,70563	b	12,09685	S	9,412951	S	9,685281	b
8,882432	m	7,182014	y	7,002555	m	8,014157	y	8,923956	b	6,844104	b	6,707181	y	8,537543	y	6,50989	y	7,986341	y
8,820717	y	5,117733	m	4,807655	y	4,586301	m	3,960789	m	4,164572	m	6,598764	m	6,570825	m	5,91062	m	4,850871	m

M021		M022		M023		M024		M025		M026		M027		M028		M029		M030	
30,23436	p	27,92959	p	31,88542	p	30,59976	p	38,63148	p	35,17627	p	32,03476	p	30,16123	p	30,97247	p	35,10097	p
25,73233	r	26,08539	k	27,12028	r	25,7032	o	22,42173	u	23,78181	r	26,76008	r	28,11827	z	29,80758	z	28,44898	o
24,95616	l	22,99375	d	23,16146	o	25,03044	u	22,25632	r	21,08798	l	23,48418	k	28,11827	Z	29,80758	Z	25,34569	u
24,64856	n	21,49381	r	21,2277	u	22,61599	l	22,19832	v	18,41355	v	22,85777	e	25,4653	r	27,37226	o	23,47837	l
23,3508	k	20,65409	e	20,37788	a	20,37944	r	21,13657	o	17,43911	o	21,86228	z	24,05238	o	22,38547	a	23,39416	r
22,21033	g	20,1979	i	20,09226	n	19,05727	n	20,56998	k	17,17803	e	21,86228	Z	23,48984	e	21,28511	r	23,37434	e
19,68705	z	19,57539	u	19,68587	i	18,13787	e	19,6895	t	16,77012	u	21,63395	l	22,6567	a	20,91897	v	21,12252	a
19,68705	Z	19,50112	z	19,12428	z	17,59635	i	17,72409	g	15,73996	t	19,19264	a	22,34784	i	20,0354	l	20,64898	k
19,15466	t	19,50112	Z	19,12428	Z	17,40808	k	17,69379	n	15,71315	z	18,66745	i	21,86446	u	19,91042	k	19,14733	v
19,12054	o	19,21376	v	18,68567	k	17,02675	a	17,42033	e	15,71315	Z	17,77276	t	20,69414	l	19,74023	e	18,0965	i
19,10281	u	19,10889	n	18,64787	e	14,2828	t	16,15261	l	15,56821	n	17,61392	n	20,08812	v	16,38429	t	16,72087	z
19,01514	a	18,58812	g	18,21184	b	13,33084	v	16,04417	a	15,45997	i	16,87043	u	17,38432	n	16,15246	i	16,72087	Z
18,23939	v	18,29855	l	17,8989	v	13,15514	g	16,025	d	14,68136	k	16,20126	o	16,96902	k	16,04186	u	16,59348	d
17,76659	i	18,2376	a	17,8357	l	13,02729	z	14,79547	b	14,5446	a	15,59364	v	15,06089	t	13,70777	n	15,859	t
16,68633	m	15,43128	o	17,02168	g	13,02729	Z	14,25259	i	12,76164	g	15,45186	g	13,21088	s	12,66767	d	14,60912	n
16,37221	e	15,19273	t	14,97718	t	12,22001	y	13,21622	z	11,41565	s	13,30764	d	13,21088	S	11,23819	s	12,05623	s
15,25402	b	14,48402	s	14,7042	y	11,31043	d	13,21622	Z	11,41565	S	12,43884	s	12,89128	g	11,23819	S	12,05623	S
14,20582	d	14,48402	S	13,80936	d	11,17952	s	10,45669	s	10,69697	b	12,43884	S	11,06322	d	10,24242	g	11,17654	g
10,68996	s	11,71301	y	12,42245	s	11,17952	S	10,45669	S	10,53612	d	12,26267	y	10,39573	b	8,169855	y	8,079097	y
10,68996	S	11,62472	b	12,42245	S	8,152959	b	8,159728	y	8,229355	y	9,224857	b	9,649149	y	7,073214	m	7,673647	b
9,951772	y	7,146236	m	6,767811	m	5,722508	m	5,470494	m	3,841922	m	7,033816	m	8,035412	m	6,81323	b	5,664084	m

M031		M032		M033		M034		M035		M036		M037		M038		M039		M040	
33,36749	p	38,61147	p	28,44276	p	27,09211	k	21,98163	r	29,53931	p	23,07812	p	26,92412	p	31,24418	p	24,9609	p
26,3082	r	21,40895	r	25,89878	z	26,51484	r	20,89994	p	23,8434	z	22,43751	r	23,02395	o	28,85003	r	21,55937	u
22,82543	z	20,06155	k	25,89878	Z	26,19769	o	20,33107	g	23,8434	Z	22,0828	n	22,65324	u	26,67071	u	21,24102	k
22,82543	Z	18,94858	o	25,16141	o	24,12252	d	19,96845	o	21,02749	o	20,61764	u	22,60111	k	24,61009	z	19,78314	o
21,58999	o	18,38298	n	23,82842	u	24,06348	p	19,22343	u	20,2744	a	19,86437	o	22,39196	r	24,61009	Z	19,60857	a
18,59354	u	18,37588	e	20,24285	i	22,39391	u	18,48192	l	19,54688	r	19,72215	a	21,7329	e	24,20016	e	18,95633	v
18,44197	e	17,78861	a	19,75015	t	20,87641	n	17,75858	a	19,3963	e	18,45597	l	20,75694	z	22,51511	n	18,86558	r
17,07128	i	16,67585	t	19,52552	r	18,61344	a	17,37951	e	17,58779	i	17,32369	z	20,75694	Z	21,83547	o	18,35178	e
16,64729	a	16,03273	u	19,38672	e	17,57715	g	17,05251	t	17,55382	t	17,32369	Z	20,46073	g	21,57713	i	17,93777	n
16,169	l	15,86317	i	18,00992	a	16,83571	i	16,85895	k	17,43601	u	16,56969	e	20,38407	l	20,61386	k	17,61425	i
15,78764	t	15,55428	z	17,03467	v	16,75358	l	15,99277	i	17,15578	k	14,4471	i	18,6563	a	19,40319	g	17,20783	g
15,01298	k	15,55428	Z	15,98126	k	16,71955	e	15,78241	z	16,62702	n	14,36894	t	18,13651	d	18,87423	a	17,08406	l
14,29324	n	15,35156	l	15,75436	g	16,63633	t	15,78241	Z	15,11729	g	13,60745	g	17,19066	i	18,74007	l	14,33007	d
14,23648	g	11,88451	g	15,10322	s	15,67078	v	15,74983	v	12,92149	l	13,59297	v	16,62373	v	17,90852	t	12,91444	s
12,83306	v	10,54756	v	15,10322	S	13,88147	s	15,52964	n	11,39407	y	12,4793	k	14,59554	t	17,06396	s	12,91444	S
12,16332	s	10,45251	s	14,96855	l	13,88147	S	13,6849	s	11,31237	d	11,78832	d	14,20796	n	17,06396	S	12,85315	t
12,16332	S	10,45251	S	13,75849	y	13,79972	z	13,6849	S	9,894515	v	11,62842	s	13,853	b	16,09513	d	12,28785	z
10,28696	b	9,883476	y	13,24226	n	13,79972	Z	13,0782	b	9,61461	s	11,62842	S	11,21195	y	15,80852	b	12,28785	Z
8,236024	y	9,846248	d	11,47293	d	12,08652	b	11,773	d	9,61461	S	11,01713	b	9,478729	s	15,63042	v	10,19104	y
7,190567	d	8,07374	b	9,482074	b	9,388971	y	10,38721	y	9,053669	b	8,425935	y	9,478729	S	9,773505	y	10,05609	b
3,712084	m	7,187362	m	6,06451	m	5,611182	m	4,277247	m	5,079887	m	5,870958	m	4,981749	m	5,874182	m	8,469539	m

M041		M042		M043		M044		M045		M046		M047		M048		M049		M050	
33,05911	p	26,36099	r	29,48395	p	25,93766	z	31,79347	p	32,42031	p	24,77773	r	33,6968	p	32,39701	p	35,52887	p
23,59999	r	24,80054	p	22,21404	o	25,93766	Z	21,89337	r	20,20564	o	24,35747	p	25,60899	z	24,77527	k	26,30836	k
21,98604	k	19,99636	u	20,84568	r	25,04983	p	19,75292	t	20,01944	k	19,31562	v	25,60899	Z	21,7507	l	26,02855	r
20,50734	u	19,04985	z	20,44962	l	23,1897	a	18,46925	u	19,08412	r	17,93176	g	25,31878	k	21,45763	g	21,50214	v
19,80206	o	19,04985	Z	19,66446	u	22,11741	i	16,01477	e	17,32305	l	17,90316	a	24,41634	u	21,10292	u	19,3524	l
16,51741	t	18,55327	a	19,41264	g	21,94228	r	15,61801	o	17,1658	a	17,64073	l	24,27816	r	20,80491	v	18,88612	o
15,95402	a	18,11398	i	17,58355	a	21,59868	l	15,45311	a	16,75569	z	17,56182	i	22,92005	n	19,55194	t	17,74654	z
15,80225	n	18,03848	l	17,54767	k	21,4466	u	14,71593	z	16,75569	Z	17,29105	n	21,65472	i	18,52363	r	17,74654	Z
15,12952	b	17,95741	e	16,50181	i	20,94697	e	14,71593	Z	16,14024	v	15,74096	d	20,74775	e	17,3266	a	17,5731	u
14,8343	v	17,10691	k	16,11747	t	20,41387	v	14,70722	l	15,89842	n	14,98103	k	20,49041	a	17,14092	i	17,19214	t
14,44377	e	16,66594	o	15,51817	e	18,41245	k	14,37267	n	15,80053	u	14,80453	b	20,37732	l	16,50087	e	16,03114	i

M041		M042		M043		M044		M045		M046		M047		M048		M049		M050	
13,74848	i	14,89948	t	14,59299	d	16,60081	n	13,33009	i	15,1651	e	14,67028	u	19,77704	v	16,44638	o	15,58396	n
13,58592	d	14,16313	n	14,18131	s	14,97103	o	12,73021	d	15,04325	g	14,45079	o	18,9001	t	15,79474	z	15,00634	a
13,57909	z	10,57132	g	14,18131	S	14,41111	t	11,75175	k	14,71726	i	14,2836	t	18,38382	o	15,79474	Z	14,76855	e
13,57909	Z	9,516824	s	12,95275	b	13,76571	s	11,44059	v	14,11686	t	13,92632	e	14,20659	g	13,45891	n	12,63162	g
13,21496	l	9,516824	S	12,88246	v	13,76571	S	9,96611	y	11,98697	d	13,7476	z	12,83992	s	11,98208	d	12,13104	y
11,05651	g	9,435367	d	12,66525	z	12,88098	d	9,671949	m	11,01482	b	13,7476	Z	12,83992	S	9,773551	s	11,69878	d
10,73085	s	9,121818	v	12,66525	Z	11,114	b	9,000143	s	10,16087	s	11,22381	m	11,74697	d	9,773551	S	10,45894	s
10,73085	S	8,40946	y	12,32262	n	9,413464	y	9,000143	S	10,16087	S	11,16508	s	10,6159	y	9,750366	y	10,45894	S
8,018868	y	8,085226	b	8,644812	y	8,564517	g	8,310915	g	6,924354	y	11,16508	S	8,474458	m	8,190188	b	8,104093	m
4,745658	m	5,269356	m	4,957362	m	5,699851	m	8,188414	b	4,678219	m	7,128971	y	8,175846	b	6,519883	m	6,95646	b

12 PRIEDAS. IŠRIKIUOTI MOTERŲ DIKTORIŲ FONEMŲ PANAŠUMŲ ĮVERČIAI

F001		F002		F003		F004		F005		F006		F007		F008		F009		F010	
24,75383	p	31,46279	p	32,05113	p	29,84405	p	42,90184	p	39,08012	p	29,22108	p	26,53669	p	28,83515	p	29,89049	p
21,18596	r	26,33868	z	24,68026	g	28,65581	z	29,1614	k	22,32001	r	23,19938	z	24,87264	r	19,19539	a	21,55987	r
20,10931	z	26,33868	Z	23,82656	u	28,65581	Z	22,05474	r	20,27825	k	23,19938	Z	21,34005	o	18,23471	o	21,1643	o
20,10931	Z	21,57279	r	20,08484	v	25,39437	k	19,90916	u	19,86205	z	21,4211	l	19,06997	z	17,03765	t	20,72156	k
18,79614	k	21,35442	d	19,72873	r	17,74013	u	19,49422	t	19,86205	Z	21,20299	r	19,06997	Z	16,83513	u	16,36933	u
16,79906	d	17,63936	i	18,5788	o	16,15321	r	16,52846	o	19,85486	v	20,28582	d	17,53537	l	16,29977	r	14,95393	l
15,27457	u	17,49927	u	17,11198	e	16,1122	o	16,23045	a	18,04124	o	19,09488	k	17,41711	u	14,13078	e	14,82011	e
15,17127	e	16,78338	a	16,79978	i	15,09745	t	14,52432	e	17,83586	e	18,68683	v	17,11063	k	13,76797	l	13,58277	a
14,76112	a	16,65133	k	15,07271	n	14,21013	i	13,60078	i	17,27742	a	18,61052	n	15,23262	t	12,61879	i	13,03757	z
14,67678	i	14,71737	l	14,4677	s	13,40038	e	12,98678	l	15,92692	u	18,56539	i	15,13639	a	10,92358	s	13,03757	Z
14,5137	v	14,41214	o	14,4677	S	13,24239	g	12,20196	n	15,24435	i	18,44712	o	14,72322	e	10,92358	S	12,05765	g
14,25172	l	14,00981	e	14,38276	t	12,94804	v	11,34625	z	13,29727	n	17,21611	s	13,17089	i	10,81122	n	11,81015	i
13,79052	n	13,40184	g	14,31835	a	12,01885	a	11,34625	Z	12,64469	l	17,21611	S	12,4017	s	10,39531	z	11,52877	t
13,66663	s	12,02336	y	12,92935	z	10,29612	n	9,314396	s	11,47463	t	17,1618	e	12,4017	S	10,39531	Z	11,5026	n
13,66663	S	11,50384	t	12,92935	Z	9,709476	s	9,314396	S	11,01958	g	16,8594	u	11,47588	n	9,748707	y	10,86132	v
13,40005	o	11,15222	n	12,4272	k	9,709476	S	8,872653	g	10,63875	b	16,67091	a	11,00737	g	9,662867	d	9,201043	s
11,24857	g	10,62144	v	12,40932	d	9,49592	l	8,353773	y	10,04607	d	15,25976	t	10,744	v	8,814397	v	9,201043	S
11,2443	t	9,855284	s	11,73125	l	9,422276	b	7,99744	d	9,817607	s	11,64632	y	9,278836	d	7,879231	b	7,935068	y
8,85663	y	9,855284	S	8,812167	b	8,94383	d	7,632954	m	9,817607	S	11,50051	b	8,681918	y	7,104478	k	7,815959	b
7,662653	b	9,7887	b	8,811379	y	7,238338	y	7,282774	v	9,132852	y	10,66687	g	7,071971	m	6,527698	g	6,113419	d

F001		F002		F003		F004		F005		F006		F007		F008		F009		F010	
5,552073	m	4,428965	m	4,511932	m	5,346261	m	5,895426	b	5,652188	m	3,881047	m	6,366044	b	4,594611	m	4,090811	m

F011		F012		F013		F014		F015		F016		F017		F018		F019		F020	
35,48456	p	38,59391	p	40,18139	p	28,93079	p	33,34982	p	33,6887	p	33,10883	p	31,7425	p	31,76029	p	27,68723	p
26,07521	k	25,4735	k	22,78279	k	25,81426	k	24,78519	r	24,08756	k	22,56354	k	22,49656	r	20,29507	t	27,22895	k
24,29757	r	23,78837	r	21,45013	r	25,40855	r	22,66134	k	20,65868	r	19,11563	a	17,35822	u	19,91694	r	25,89057	o
15,91742	t	15,05284	t	18,35625	u	20,65147	u	21,08261	z	20,23928	u	18,98081	r	16,56646	a	19,46805	u	25,82514	g
15,629	v	14,87486	a	18,07319	a	19,67248	o	21,08261	Z	17,69315	t	18,29092	u	15,06113	o	16,65626	o	24,5289	l
13,0987	a	14,35815	l	17,15947	l	17,57328	t	20,67628	o	16,4495	a	17,30497	i	14,00413	i	15,41806	a	24,00662	a
12,87116	o	13,90541	z	16,30488	e	17,40188	a	15,66755	t	15,8846	l	15,79514	e	13,83822	v	14,73851	l	21,28932	r
12,82693	e	13,90541	Z	14,3658	t	16,71581	l	15,35343	l	15,39331	v	14,9008	v	13,55243	g	14,1567	e	20,35932	z
11,83723	s	11,33618	u	13,23802	z	15,96423	e	15,06294	a	15,28912	z	14,73536	t	13,43514	z	14,13461	v	20,35932	Z
11,83723	S	10,9915	g	13,23802	Z	14,9601	z	14,83424	v	15,28912	Z	14,67469	l	13,43514	Z	13,68008	n	19,23676	u
11,82662	l	10,89003	d	12,70221	o	14,9601	Z	14,29351	e	14,54238	o	14,44884	n	13,33894	t	12,80521	i	17,4817	i
11,81034	z	10,83686	e	12,66349	i	14,4915	i	12,31059	u	14,21065	e	13,77859	o	12,78912	s	12,75162	g	16,53803	e
11,81034	Z	10,66073	v	10,91489	g	13,41992	v	11,22976	d	13,1711	n	11,31954	z	12,78912	S	11,71285	z	13,71823	d
11,65254	i	10,53319	o	10,03118	d	11,1617	d	10,77212	n	12,90523	d	11,31954	Z	12,71416	d	11,71285	Z	13,4739	t
11,32315	u	9,405937	b	9,759874	s	10,20811	m	10,49791	s	12,62472	i	11,24622	g	12,52264	n	11,07283	s	11,63018	s
10,93972	n	9,208628	n	9,759874	S	9,754576	s	10,49791	S	10,61024	y	11,01025	y	12,07369	l	11,07283	S	11,63018	S
10,41067	g	9,18303	i	9,352123	n	9,754576	S	10,13502	g	9,927421	g	10,8063	s	11,96825	e	8,918593	y	10,75219	n
10,09776	y	9,044116	s	9,116696	v	9,748991	n	9,889768	i	9,540281	s	10,8063	S	11,26134	k	8,737424	b	10,50672	v
7,946329	d	9,044116	S	7,855479	b	8,793339	y	8,416694	y	9,540281	S	10,45249	b	10,49685	y	8,383932	k	5,416558	b
5,829775	b	4,841058	m	5,405415	y	7,809061	b	5,33836	b	9,17565	b	8,110455	d	9,855854	b	8,153423	d	4,588453	m
5,470603	m	3,678796	y	2,872924	m	6,746072	g	4,543079	m	7,129193	m	5,321507	m	4,354019	m	3,642191	m		

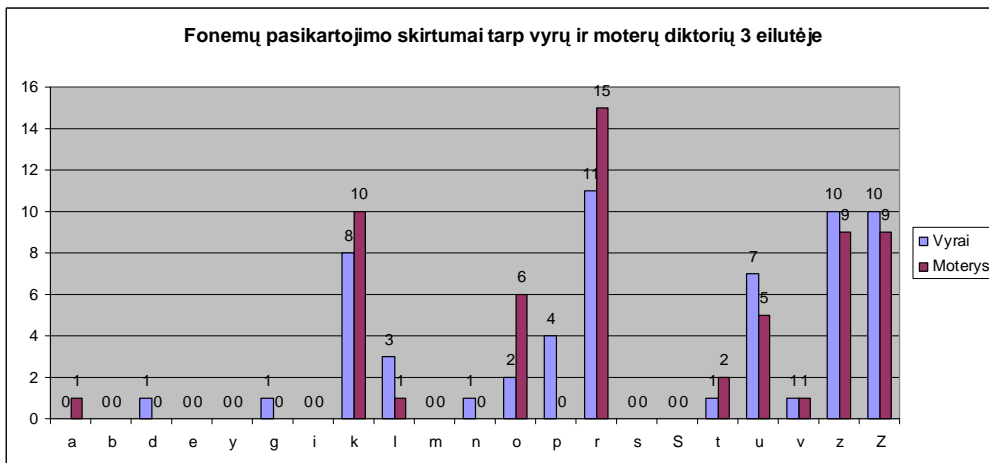
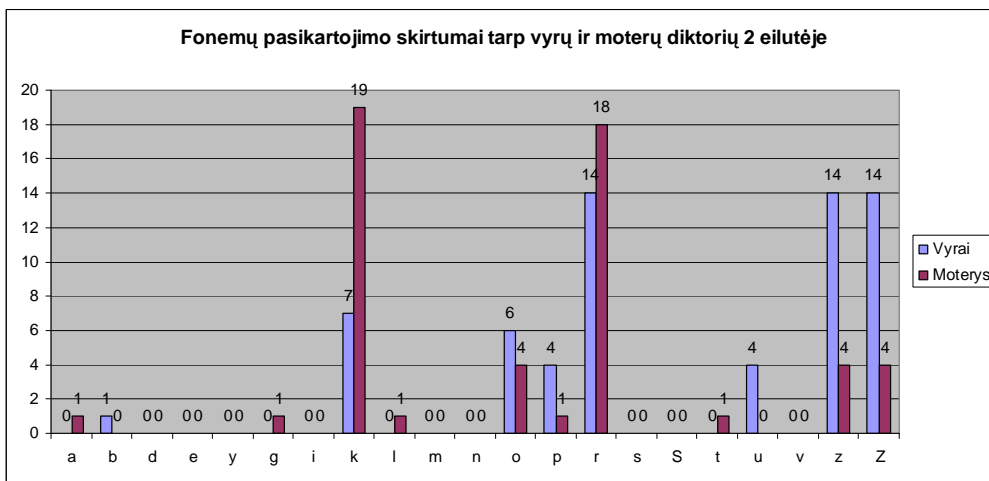
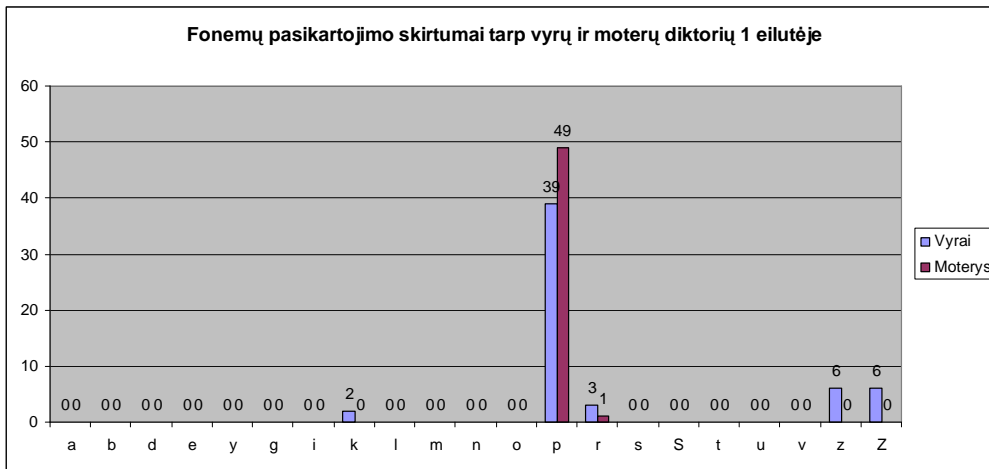
F021		F022		F023		F024		F025		F026		F027		F028		F029		F030	
52,93593	p	31,98076	p	46,12151	p	43,17544	p	36,50389	p	47,42955	p	39,45864	p	33,20712	p	36,92378	p	40,70246	p
33,31962	k	20,16897	r	30,65687	k	24,06832	k	26,1244	r	22,7878	r	21,16943	k	18,42214	r	26,02299	k	21,88242	k
21,33265	r	16,97416	z	26,85094	r	23,08024	z	17,17152	k	21,33024	o	20,84313	r	17,46038	z	19,89008	r	19,44741	u
20,27882	u	16,97416	Z	21,43249	t	23,08024	Z	15,06645	l	19,8016	k	17,48879	v	17,46038	Z	18,65527	t	17,74351	a
17,64824	z	15,91769	o	16,90746	u	22,85777	r	14,2709	u	18,93052	u	17,30917	t	16,90205	k	14,18806	l	17,7166	v
17,64824	Z	14,39001	k	16,53334	l	16,29425	t	13,84782	z	16,09048	a	16,19885	l	16,1411	t	14,1443	o	17,6887	r
17,63982	o	13,65844	e	15,03013	z	16,26825	a	13,84782	Z	16,01494	d	15,1498	o	15,99682	o	14,01855	e	16,88301	l
16,396	d	13,30474	a	15,03013	Z	15,16693	u	13,41812	e	15,58575	i	14,92806	g	14,93378	u	14,00093	u	16,35359	o
15,59015	a	13,26223	t	14,76546	a	14,68429	l	12,96974	a	14,91601	e	13,86921	a	13,87623	a	13,95119	a	14,74476	t

F021		F022		F023		F024		F025		F026		F027		F028		F029		F030	
14,7283	v	12,95868	u	14,35431	n	14,21399	o	12,72287	o	14,41494	v	13,74288	e	13,8467	d	13,93101	n	13,16702	e
14,64317	l	12,32846	d	14,04542	e	13,60807	e	12,50813	g	14,3235	l	13,67481	u	13,18395	e	11,7217	v	12,9668	i
13,80209	e	11,57526	i	12,62977	i	13,34944	i	12,0614	i	13,16922	n	13,41477	z	12,83035	i	10,69054	z	10,2566	n
13,01747	t	11,47438	v	12,48322	v	13,3396	v	10,61064	t	12,35842	t	13,41477	Z	11,86559	v	10,69054	Z	9,649607	s
11,83568	b	11,1037	l	12,03476	g	11,77141	g	9,594289	n	11,57323	b	13,20568	d	11,42201	l	10,64267	i	9,649607	S
11,57201	i	8,787356	s	11,7158	d	11,70281	b	8,826548	s	10,84453	g	12,75972	n	11,04836	b	9,762526	g	9,093265	g
10,65401	g	8,787356	S	11,46101	s	11,54794	n	8,826548	S	10,48633	z	12,43045	i	9,789269	s	9,189573	s	8,382227	z
10,1805	s	8,605028	g	11,46101	S	11,50494	s	8,725844	d	10,48633	Z	11,94196	b	9,789269	S	9,189573	S	8,382227	Z
10,1805	S	7,985626	n	11,43166	o	11,50494	S	8,560687	v	9,787212	y	8,351484	s	8,730822	n	9,186412	d	6,465904	d
9,615078	n	7,601666	y	10,67866	b	9,998429	y	8,030251	y	9,276783	s	8,351484	S	8,123441	g	8,286073	b	6,406365	b
8,936643	y	5,392116	b	7,223862	y	9,105633	d	5,926934	b	9,276783	S	6,024421	y	7,294059	y	7,284415	y	5,734566	y
3,905685	m	4,78195	m	6,556449	m	4,494494	m	3,804074	m	4,728649	m	4,822858	m	5,92677	m	4,416614	m	4,220299	m

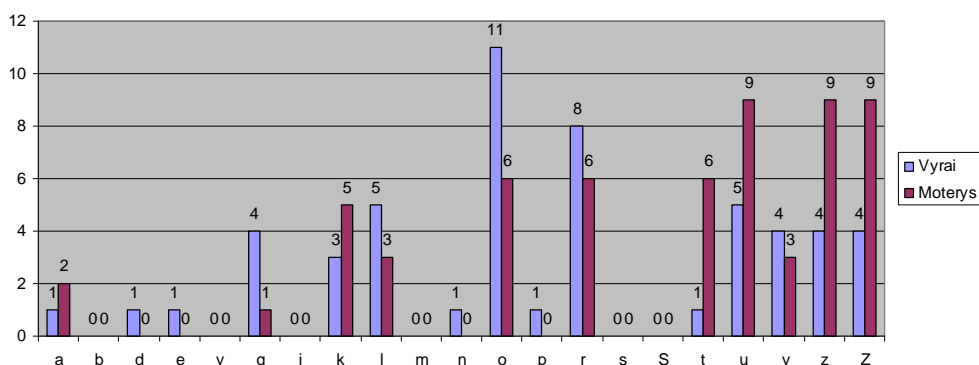
F031		F032		F033		F034		F035		F036		F037		F038		F039		F040	
36,55274	p	25,97344	p	33,81966	p	37,77015	p	32,63821	p	35,54857	p	30,05545	p	36,16292	p	24,47717	p	33,84293	p
22,34774	o	20,7389	r	28,6678	r	26,78611	o	23,88688	k	22,32265	o	21,99285	l	24,26844	r	23,70026	r	25,21567	k
21,54047	k	19,60579	k	24,09468	l	23,03617	k	18,4603	u	21,80213	k	20,5626	r	18,39346	t	21,41709	o	18,13621	z
19,09957	r	17,70593	u	23,55855	o	21,35658	u	17,34027	o	18,13331	r	20,43843	t	16,82123	o	20,26764	k	18,13621	Z
18,65286	u	16,96918	o	22,10353	u	21,31301	r	15,64971	r	16,99004	l	19,09705	d	15,77489	k	18,32074	t	17,87496	o
15,25023	a	16,18904	t	21,97506	e	20,46714	z	14,92675	e	14,26713	t	17,60255	a	14,92288	a	17,69747	d	17,76681	t
14,03559	t	14,95821	l	21,42015	n	20,46714	Z	14,61312	a	13,90114	a	17,27145	u	14,85722	e	16,28573	u	17,73119	a
13,49168	l	14,89925	z	21,37502	z	17,57618	t	14,14819	v	13,57593	i	15,79367	k	14,05235	u	16,09499	v	17,48341	r
13,22502	z	14,89925	Z	21,37502	Z	16,52825	l	13,74263	i	13,4859	g	15,77061	o	13,99228	l	14,16836	a	17,40122	u
13,22502	Z	14,53669	a	21,23792	v	15,27383	a	12,33961	t	13,35753	u	14,92308	z	13,05495	i	13,70597	i	17,17817	l
12,87458	i	13,61813	e	18,19272	k	14,65281	v	12,08864	l	12,42756	e	14,92308	Z	12,89256	v	12,4652	l	16,87894	i
12,82418	g	13,34174	d	17,16066	t	13,83593	e	11,79841	g	11,99157	z	14,89844	i	12,67356	s	11,89434	e	15,94151	s
12,6801	e	13,24531	n	16,53121	i	13,23506	i	10,75009	s	11,99157	Z	13,46789	n	12,67356	S	11,3634	n	15,94151	S
12,32015	n	13,18468	i	15,85639	a	12,35521	s	10,75009	S	10,55668	n	13,1912	e	12,369	d	11,27035	g	14,2315	n
11,84401	v	11,77614	g	14,46789	d	12,35521	S	10,73253	z	9,472494	s	12,48968	s	11,56188	g	8,822499	s	13,71489	g
11,1352	d	11,54268	v	13,56983	b	11,75809	n	10,73253	Z	9,472494	S	12,48968	S	11,25036	n	8,822499	S	13,47857	v
10,78449	s	9,65042	s	12,38305	g	10,76859	g	9,501608	n	7,592287	y	12,00094	g	10,88565	z	8,672258	y	13,00399	e
10,78449	S	9,65042	S	11,7993	s	10,71874	d	9,116321	d	7,57899	b	11,58502	b	10,88565	Z	8,414934	z	10,26741	d
7,698892	y	9,277153	b	11,7993	S	9,732187	b	8,721326	b	5,842183	d	9,96833	v	8,960341	y	8,414934	Z	8,740716	y
7,13053	b	7,822631	y	9,725302	y	8,910883	y	6,601171	m	5,530577	v	8,56475	y	7,048153	b	5,924542	b	5,421475	b
6,899133	m	5,66563	m	4,609661	m	5,94728	m	5,374511	y	5,508544	m	6,275132	m	5,06182	m	4,595106	m	3,493473	m

F041		F042		F043		F044		F045		F046		F047		F048		F049		F050	
33,71701	p	40,08099	p	37,17589	p	25,3825	p	32,70384	p	34,49596	p	35,97942	p	29,5495	r	38,06169	p	33,6327	p
26,45652	k	34,0753	k	23,77962	z	23,56636	k	21,47076	o	20,87082	r	20,00508	r	26,25086	p	20,35048	r	23,49281	r
20,26152	r	25,56957	t	23,77962	Z	18,25503	r	19,94195	r	19,06835	u	18,11571	v	25,66296	k	16,79796	k	22,09062	k
14,99417	o	23,51074	r	19,82326	r	15,08687	z	18,35529	l	18,9884	k	15,34263	o	19,88427	v	15,78834	o	21,14297	u
14,52084	t	14,9457	u	16,58183	l	15,08687	Z	18,09153	u	17,92694	o	14,72932	k	17,02972	o	14,90717	u	20,35477	o
14,34668	i	14,76412	o	16,20374	t	14,72402	g	17,74442	k	15,53059	t	14,18368	e	15,58899	e	14,5065	t	17,81402	a
13,78707	l	13,32702	a	15,24596	a	14,62152	o	16,97509	v	13,91279	l	13,27024	t	15,33924	l	13,75243	l	16,0664	l
13,49243	a	12,49592	e	14,15316	u	14,55559	v	15,46743	t	13,23542	a	12,7674	z	14,51869	u	12,66249	e	14,23489	t
13,41224	u	12,18604	n	14,07547	o	14,47499	u	15,45912	i	13,1527	e	12,7674	Z	13,91652	d	12,63407	n	13,87991	v
12,51287	e	12,13262	v	13,18545	k	13,87103	a	15,41634	e	11,95846	v	11,9654	a	13,57925	a	12,42648	a	13,48771	i
12,33903	v	11,59253	i	12,35474	n	13,09511	t	14,89438	n	11,92377	i	11,39511	s	12,59636	b	10,94674	g	13,04321	e
12,084	n	10,5596	z	11,59028	i	12,21625	i	14,83168	a	11,74333	g	11,39511	S	12,35993	n	10,58336	z	12,99146	n
10,34657	y	10,5596	Z	11,54121	v	12,08078	e	13,97316	d	11,28879	n	11,15582	n	12,21284	t	10,58336	Z	12,85664	z
9,696302	z	10,08248	g	10,75714	b	11,89155	n	12,0542	z	9,686838	d	11,11454	u	11,96592	i	10,16334	i	12,85664	Z
9,696302	Z	9,378208	s	10,30629	e	10,39947	l	12,0542	Z	9,64382	s	10,78031	i	10,57209	g	8,421745	s	12,3449	y
9,31323	g	9,378208	S	10,02193	g	9,414615	s	10,93469	g	9,64382	S	10,66915	d	10,53168	z	8,421745	S	11,57135	g
8,316562	s	9,273495	l	9,012503	s	9,414615	S	10,79683	s	8,973443	z	10,60195	l	10,53168	Z	8,1973	v	11,52579	d
8,316562	S	7,127461	b	9,012503	S	7,86977	d	10,79683	S	8,973443	Z	8,025045	b	9,434249	s	6,932028	d	10,95197	s
7,957637	d	6,898942	y	8,422087	d	7,791243	y	9,181764	b	8,491994	y	7,766362	g	9,434249	S	6,914649	y	10,95197	S
6,757023	b	5,924184	d	6,634114	y	7,509944	b	8,103708	y	7,418264	b	7,2227	y	6,703756	y	4,951814	b	8,588416	b
4,547226	m	4,845568	m	4,33226	m	3,59576	m	6,452136	m	4,65545	m	4,73751	m	5,607662	m	4,67307	m	6,394572	m

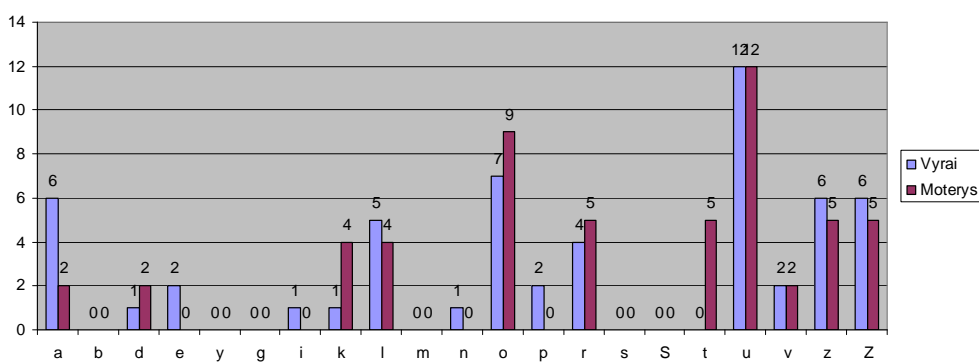
13 PRIEDAS. FONEMŲ PASIKARTOJIMO SKIRTUMAI VYRŲ IR MOTERŲ DIKTORIŲ TARPE 1 – 21 EILUTĖSE



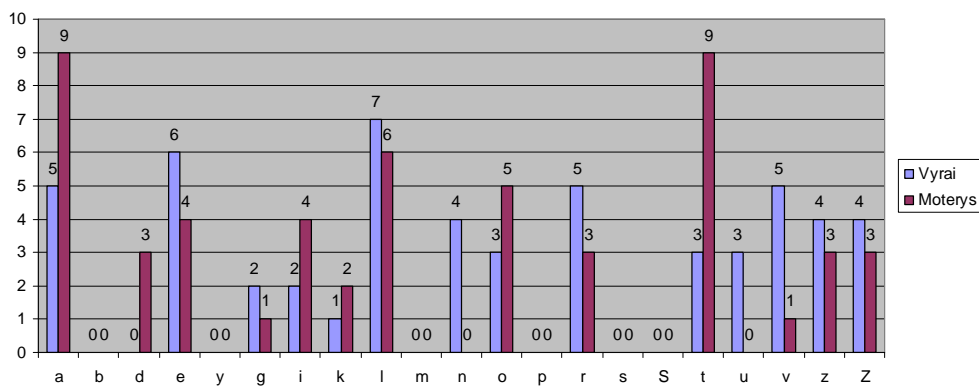
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 4 eilutėje



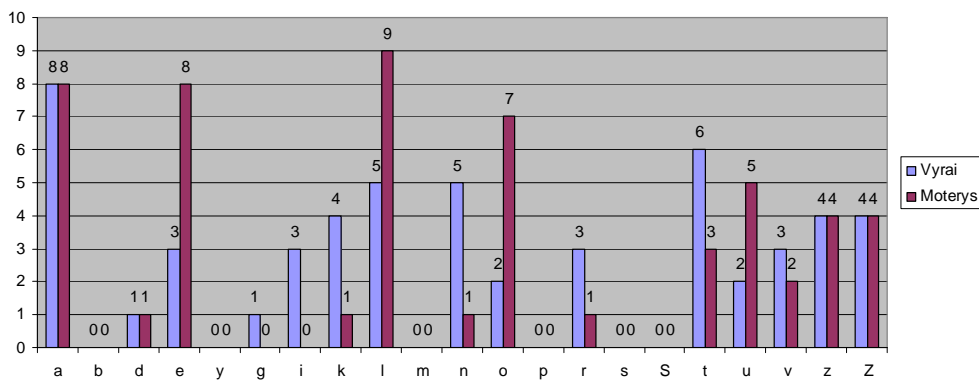
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 5 eilutėje



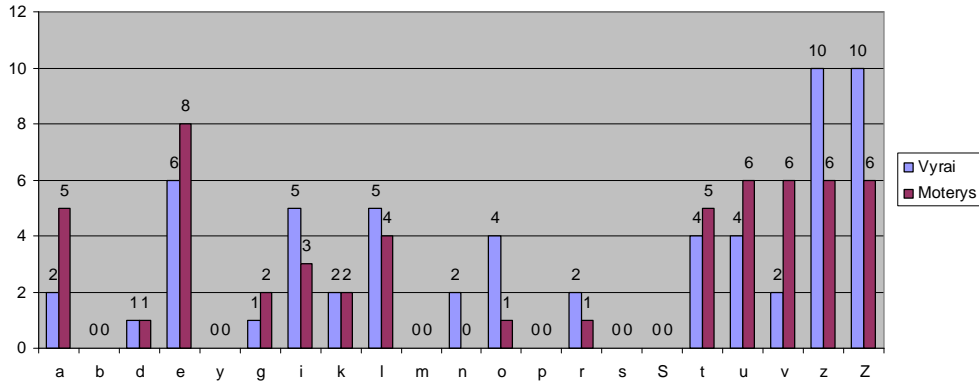
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 6 eilutėje



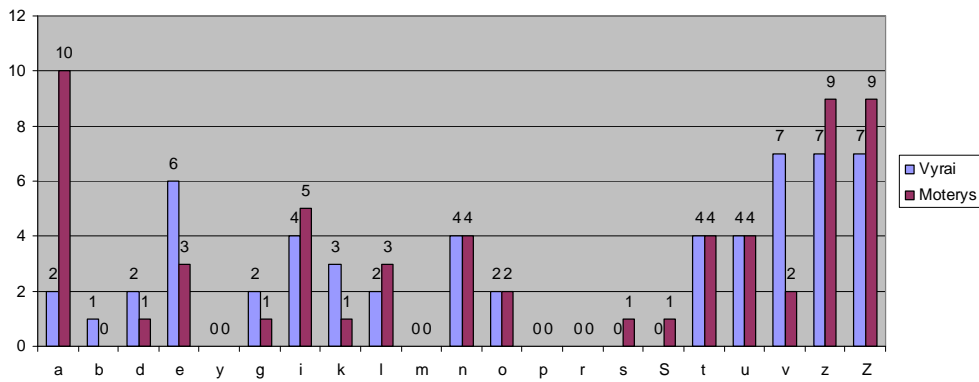
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 7 eilutėje



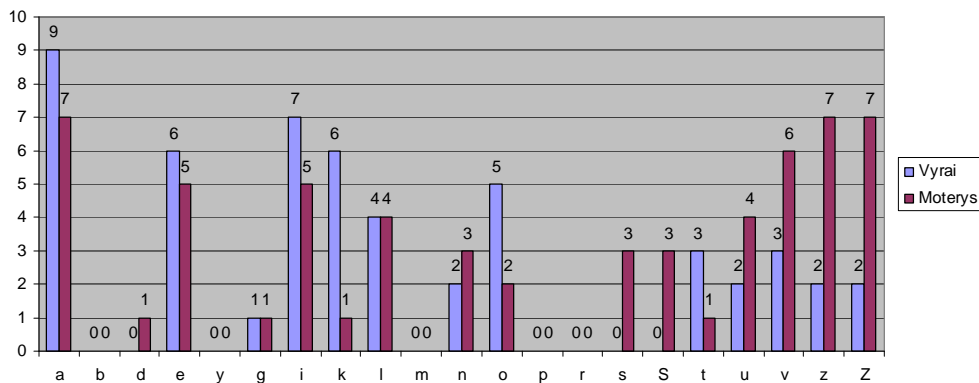
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 8 eilutėje



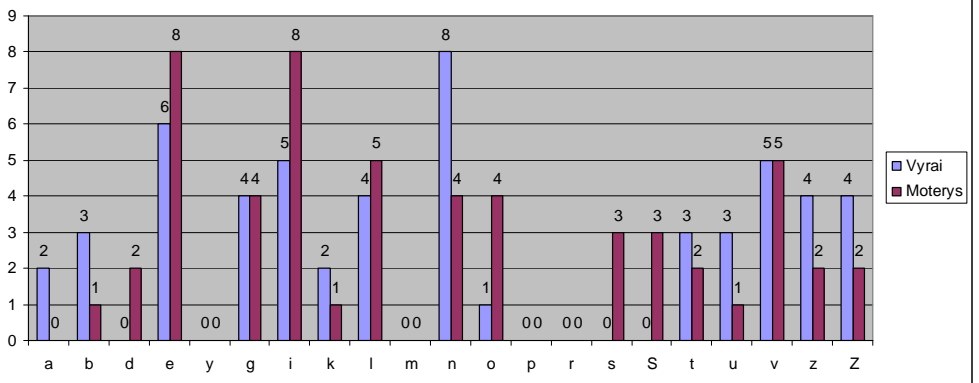
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 9 eilutėje



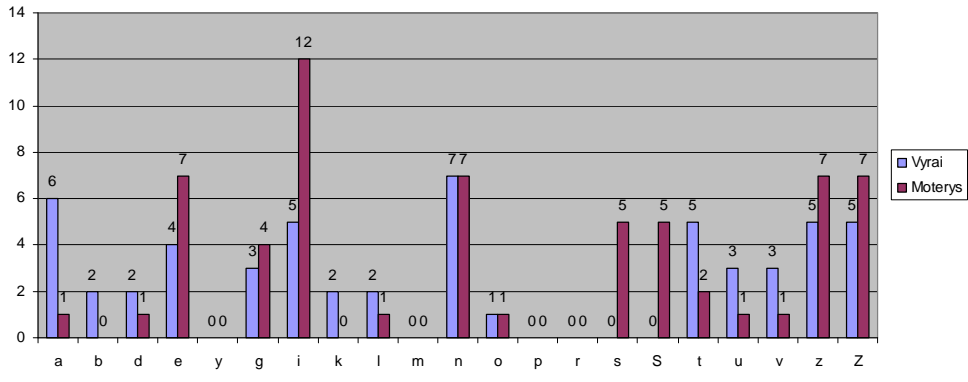
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 10 eilutėje



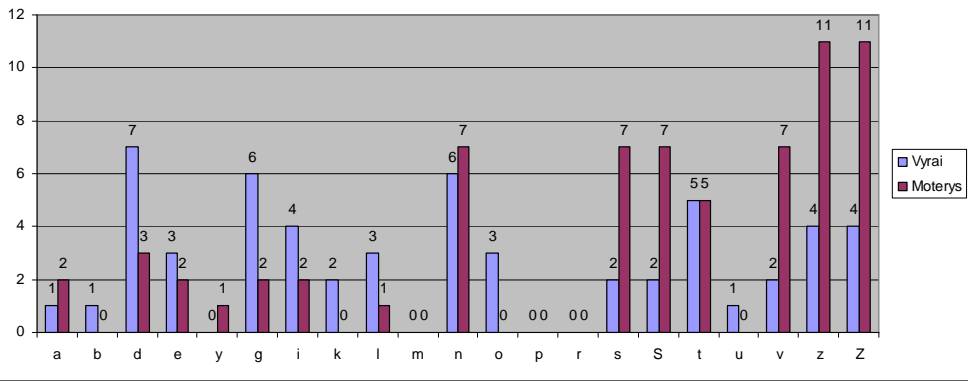
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 11 eilutėje



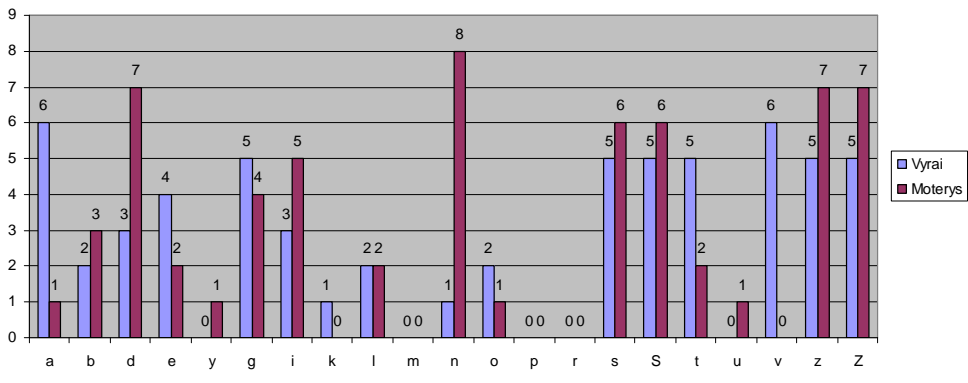
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 12 eilutėje



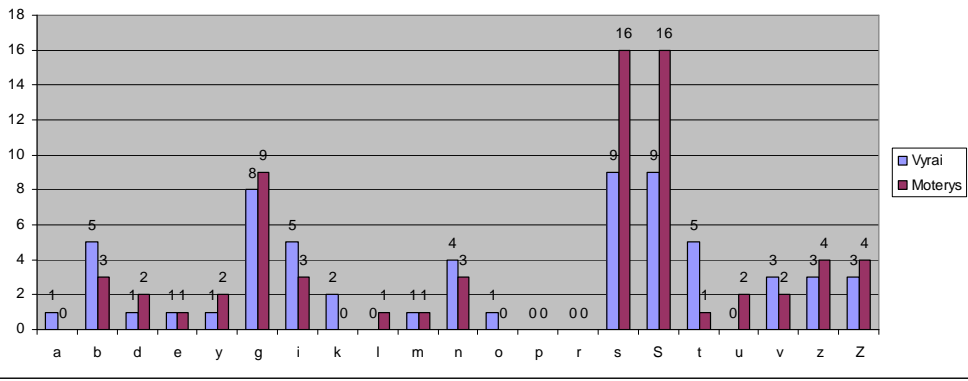
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 13 eilutėje



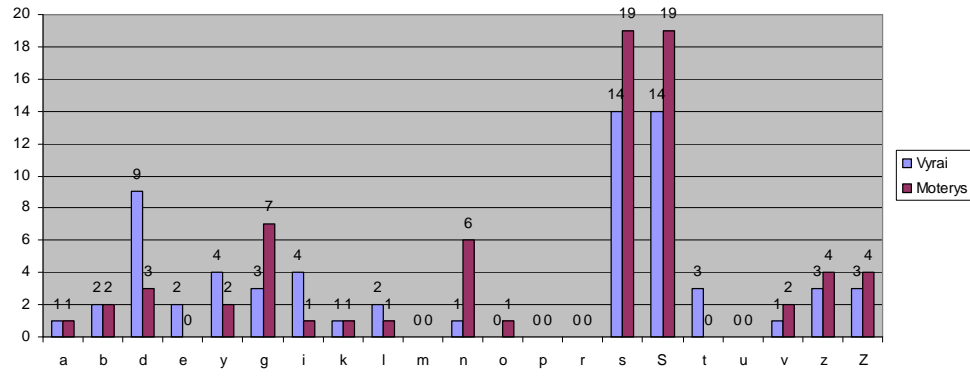
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 14 eilutėje



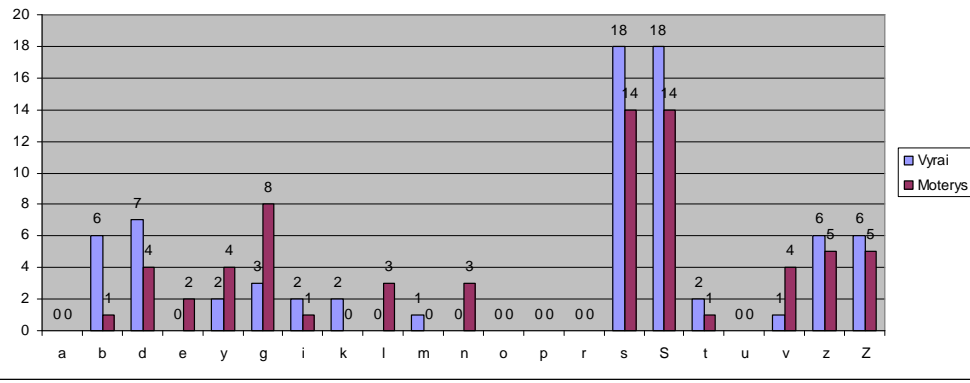
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 15 eilutėje



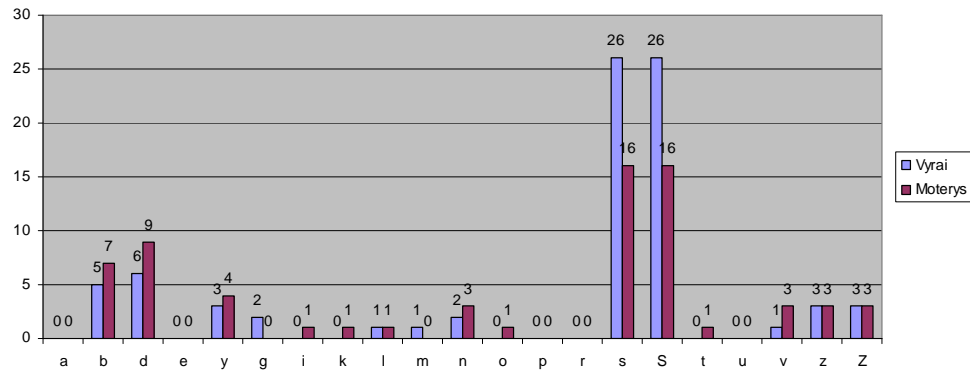
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 16 eilutėje



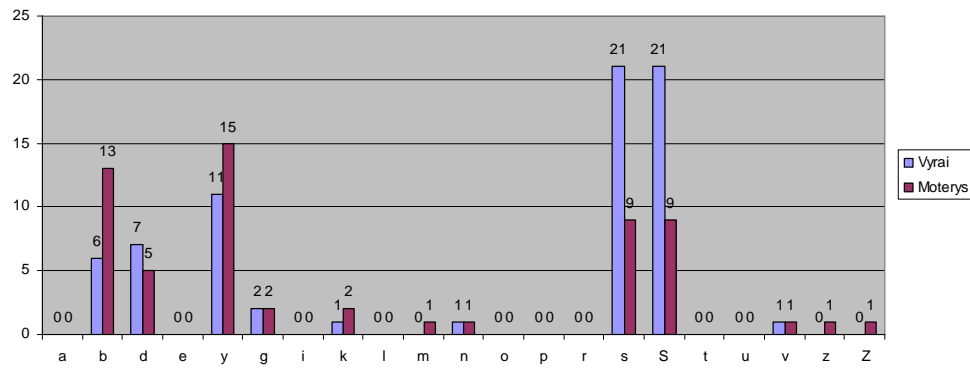
Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 17 eilutėje

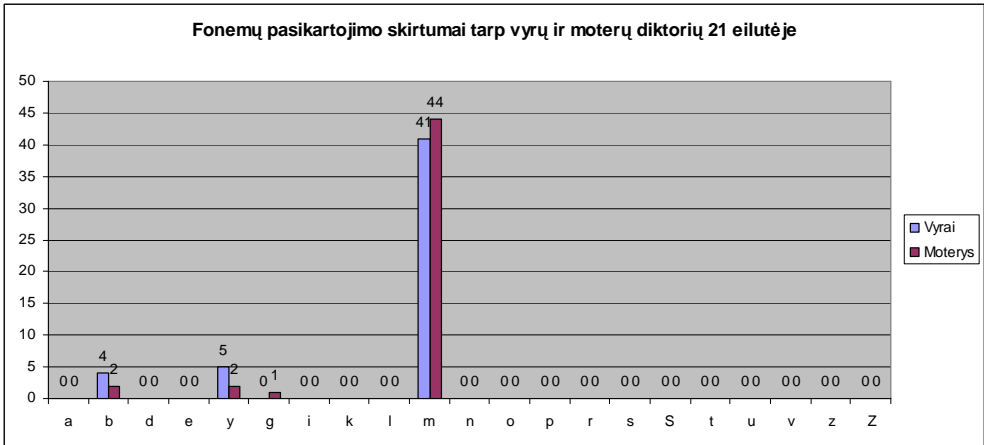
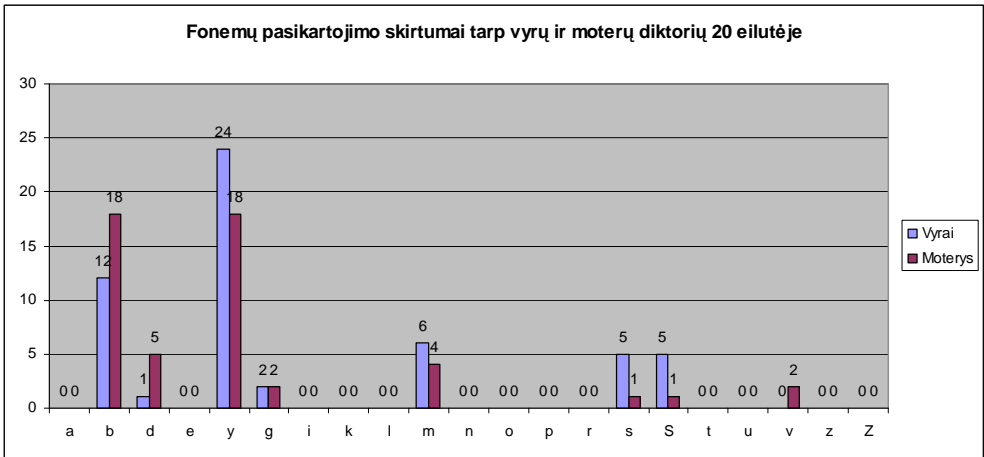


Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 18 eilutėje



Fonemų pasikartojimo skirtumai tarp vyrų ir moterų diktorių 19 eilutėje





14 PRIEDAS. STRAIPSNIS „DIKTORIAUS ATPAŽINIMO SISTEMA PAREMTA FONETINIŲ VIENETŲ DISKRIMINAVIMU“¹

Kęstutis Tomkevičius

Vilniaus Universiteto Kauno humanitarinis fakultetas, Muitinės 8, Kaunas

Doc. dr. Vytautas Rudžionis

Anotacija

PAGRINDINIAI ŽODŽIAI: diktoriaus atpažinimas, sistema, signalų apdorojimas, fonema.

Straipsnyje pristatomos diktoriaus atpažinimo sistemos, jų algoritmai bei taikymo galimybės. Taip pat pristatomos fonetinių vienetų atpažinimo ir jų naudojimo diktoriaus atpažinimui galimybės, pateikiama fonetinių vienetų savybių panaudojimu paremta diktoriaus atpažinimo sistema.

Abstract

KEYWORDS: speaker verification, system, signal processing, phoneme.

This article presents the speaker verification systems, their algorithms and the most reliable methods on which these systems are based. The article also reveals most common aspects of the implementation of these systems. The possibility for the identification and usage of the phonetic units' verification systems is introduced. The speaker verification system based on the phonetic units' recognition which is being developed is presented.

Įvadas

Šiandieniniame informacijos greittelyje tradicinių sprendimų informacijai valdyti, saugoti ir panaudoti nebepakanka. Norint palengvinti informacijos įsisavinimą ir apdorojimą, siekiama gerinti naudotojo komunikaciją su sistemomis. Vienas iš būdų – sistemų valdymas kalba. Todėl šiame kontekste svarbų vaidmenį vaidina kalbos signalų apdorojimo uždaviniai.

Šnekamosios kalbos signalų apdorojimas, pagal tai, kas yra tiriamasis objektas ir kokie laukiami tyrimų rezultatai, yra skirstoma į tris atskiras tyrimų sritis:

- Kalbos atpažinimas (Speech recognition).
- Šnekamosios kalbos atpažinimas (Language recognition).
- Diktoriaus atpažinimas (Speaker recognition).

Kalbos atpažinimo srityje dirbantys mokslininkai kuria sistemas, kurios nepriklausomai nuo kalbėjusiojo asmens, lyties, amžiaus, fiziologinių ar mechaninių žmogaus kalbos savybių, padėtų atpažinti kalbančiojo pasakytą tekstą. Antroje srityje dirbančių mokslininkų tikslas – sukurti kuo tikslesnius algoritmus, padedančius atskirti, kuria kalba buvo išreikšta gautoji informacija. Diktoriaus atpažinimo tikslas yra nustatyti kalbantį asmenį (diktoriaus autentifikacija) ir leisti jam atlikti tam tikrus veiksmus sistemoje (autorizacija).

Užsienyje tyrimai kalbos atpažinimo srityje atliekami jau daugiau kaip 40 metų. Lietuvoje pirmosios kalbos signalų tyrimų laboratorijos pradėtos kurti apie 1970-uosius metus, o 1999-aisiais metais pradėtas kurti pirmasis lietuviškas garsynas LTDIGITS.

Diktoriaus atpažinimas – tai automatinis kalbėtojo atpažinimo procesas, paremtas asmenine informacija, kurią turi apdorojamas signalas.

Diktoriaus atpažinimo uždaviniai glaudžiai siejasi su biometrinių sistemų sritimi, nes balsas yra viena iš keliolikos biometrinių charakteristikų, pagal kurias identifikuojami žmonės. Balsas ir kalba, kaip biometrinė priemonė, yra skirtingai vertinama – mokslininkai nesutaria dėl balso, kaip biometrinės priemonės patikimumo. Viena mokslininkų pusė teigia, kad balsas ir kalba neleidžia vienareikšmiškai identifikuoti asmens, kaip tai leidžia padaryti pirštų atspaudai, akies rainelė ar delno venų atspaudai. Kiti tuo tarpu teigia, kad balsas, balso tembras, bei kitos charakteristikos, kiekvieno žmogaus kalbą padaro unikalią. Sutariama dėl vieno – dirbant šioje srityje yra labai lengva surinkti duomenis tyrimams, be to, palyginus nesunkiai galima atlikti ir identifikavimą.

Kalbos atpažinimo sistemose labai didelę įtaką turi foninis triukšmas. Tradiciškai tokios sistemos turi funkcionuoti triukšmingoje aplinkoje (gatvės triukšmas, darbo aplinkos triukšmas, kolegų triukšmas ir pan.), o sistemos mokomos naudojant ramioje aplinkoje padarytus įrašus. Todėl gaunamas neatitikimas tarp mokymo ir sistemos funkcionavimo sąlygų. Antra vertus, mokyti sistemos triukšmingoje aplinkoje

¹ “Akademinio Jaunimo siekiai: Ekonomikos, Vadybos ir Technologijų įžvalgos”. 2008. ISSN 2029-0217. p. 168 – 172.

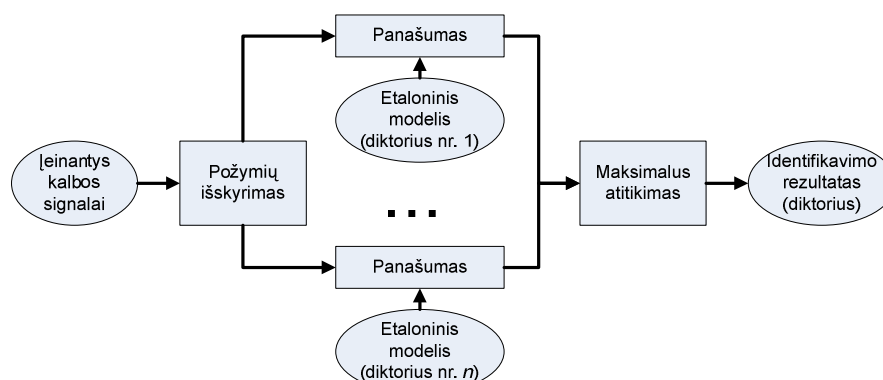
dažniausiai negalima, nes nėra žinoma, kokioje aplinkoje tai sistemai reikės funkcionuoti. Todėl reikia panaudoti specialius triukšmo įtaką mažinančius algoritmus (spektrinį įrašo kokybės pagerinimą, kepstro vidurkio atėmimą ir pan.). Tačiau praktiškai visada sistemos efektyvumas triukšmingoje aplinkoje sumažėja – skiriasi tik efektyvumo sumažėjimo laipsnis.

Diktorius atpažinimo sistemos

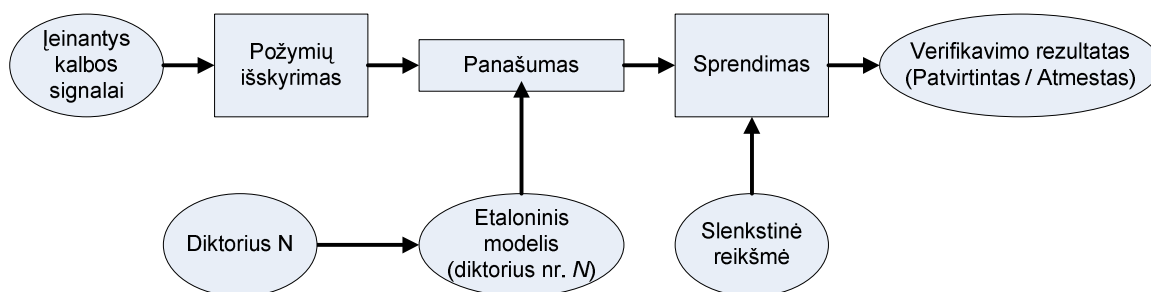
Pagal atliekamas funkcijas diktorius atpažinimo sistemos klasifikuojamos į diktorius identifikavimo (Speaker identification) ir diktorius verifikavimo (Speaker verification) sistemas, o pagal tai, kokį tekstą turi sakyti diktorius, kad sistema jį atpažintų, skirstomos į nepriklausomas nuo teksto (Text independent), priklausomas nuo teksto (Text dependent) ir pateikiamo teksto (Text-Prompted) diktorius atpažinimo sistemas (Matsui, Furui, 1993).

Pastebėta, kad labai trumpame periode (5 – 100 msec) balso charakteristikos yra labai panašios, tačiau situacija iš esmės pasikeičia ilgame laikotarpyje (1 – 5 sek), kuomet yra ištariamas didesnis kiekis garsų.

Pirmame paveiksle pateikiama principinė diktorius atpažinimo sistemos diktorius identifikavimo schema, o diktorius patvirtinimo schema – antrame paveiksle (Matsui, Furui, 1993).



38 pav. Principinė diktorius atpažinimo sistemos diktorius identifikavimo schema



39 pav. Principinė diktorius atpažinimo sistemos diktorius verifikavimo schema

Nepriklausomos nuo teksto sistemos yra lankstesnės, jos „neprištos“ prie teksto, kurį būtinai turi pasakyti naudotojas. Tai nevaržo naudotojo pasirinkimo, jos yra saugesnės slaptažodžio nuspėjamumo požiūriu, nes naudotojas kaskart gali pasirinkti kitą frazę, tačiau tokios sistemos yra sunkiau realizuojamos. Tuo tarpu nuo teksto priklausomos diktorius atpažinimo sistemos naudojamos ten, kur reikia griežtos kontrolės naudotojo įvedamam tekstui, todėl frazę ar frazes, kurias gali ištartti diktorius, sistema saugo duomenų bazėje kaip iš anksto įvestą šabloną, pagal kurį gali būti atliekamas lyginimas, o diktorius ištariama frazė veikia kaip slaptažodis.

Taikymai

Diktorius atpažinimo sistemų praktinis pritaikymas:

- Numerio rinkimas balsu (Voice dialing);
- Telefoninė bankininkystė (Banking by phone) (McLaughling, Reynolds, 2002);
- Duomenų bazių prieigos paslaugos (Databases access services);
- Konfidencialios informacijos apsauga (Security control to confidential information);

- Nutolusių kompiuterių prieiga (Remote access of computers);
- Namų automatika, protingieji namai (Smart houses);
- Teismo ekspertizė (Forensics) ir kt.

Eksperimentinės diktoriaus atpažinimo sistemos jau dabar naudojamos įvairiose veiklos srityse. Kaip teigiama „Speech technology magazine“, rašančiame apie signalų apdorojimo ir kalbos atpažinimo technologijų naujienas, eksponentiškai augant mobiliųjų telefonų skaičiui, JAV kompanija Planet Payment bando išnaudoti šią terpę verslo poreikiui ja pasiekti klientus. Todėl buvo sukurta mokėjimo naudojantis mobiliuoju ryšiu sprendimų linija BuyVoice, kurioje naudojama kompanijos PerSay balso autentifikavimo technologija. Kadangi kiekvieno žmogaus balsas, kaip ir jo pirštų anspaudai ar akies rainelė yra unikalūs, teigiama jog ši autentifikavimo technologija užtikrins pakankamai aukšto lygio paslaugų kokybę, o naudotojams paprastą būdą atlikti norimus veiksmus (Speechtechmag.org, 2007).

Algoritmai

Šiuolaikinis diktoriaus, kaip ir kalbos, atpažinimas organizuojamas dviem etapais:

- Požymių išskyrimas (Feature extraction);
- Požymių atitikimas (Feature matching).

Požymių išskyrimo tikslas yra konvertuoti kalbos signalų įrašo diskretas į iš anksto pasirinkto tipo parametrinį atvaizdavimą, vadinamą savybių vektoriais, kurie naudojami tolimesnėje analizėje ir klasifikatorių apdorojime.

Atliekant paruošiamuosius veiksmus, pirmiausia įeinančiam kalbos signalui yra atliekamas pirminis apdorojimas: jis yra diskretizuojamas, suskaidomas į vienodo ilgio langus. Iš jų, naudojant įvairius požymių apskaičiavimo metodus, pateikiamus šiame straipsnyje, yra sudaromi požymių vektoriai, savyje saugantys informaciją, efektyviai aprašančią akustinį kalbos signalą (Viiki, 1999). Kitame diktoriaus atpažinimo etape naudojantis akustiniais bei kalbos modeliais yra atliekamas požymių vektorių statistinis palyginimas su duomenų bazėje esančių diktorių požymių vektoriais ir išrenkamas labiausiai atitinkantis diktorius.

Požymių išskyrimo algoritmai, naudojami diktoriaus atpažinimo sistemų, signalų požymių išskyrimo:

- Filtrų banku paremti kepstriniai koeficientai, pvz. Melo dažnių kepstriniai koeficientai (Mel Frequency Cepstral Coefficients, MFCC) (Mermelstain, Davis, 1982);
- Tiesinis prognozavimas (Linear Predictive Coding, LPC) (Vienas pirmųjų algoritmų, pradėtų naudoti kalbos signalų analizei) (Markel, Gray, 1976);
- Tiesinio prognozavimo kepstrinis koeficientas (Linear Predictive Cepstral Coefficients, LPCC) (Makhoul, 1975);
- Perceptinis tiesinis prognozavimas (Perceptual Linear Prediction, PLP) (Hermansky, 1991);
- Neuroninis prognozavimas (Neural Predictive Coding, NPC);
- Spektrinė analizė (Spectral analysis).

Šie metodai jau nusistovėję ir naudojami ganėtinai ilgai. Visiems šiems metodams papildomai galima paskaičiuoti šias charakteristikas (Young ir kt., 2001):

- Energiją;
- Delta koeficientus.
- Akceleracijos koeficientus.
- Jos irgi naudojamos požymių vektorių sudaryme.

Atlikus kiekvieno lango analizę ir apskaičiavus kalbos signalo požymius iš jų galima sudaryti požymių vektorių, kuris bus naudojamas kaip automatinio diktoriaus kalbos atpažinimo sistemos įėjimas:

$$X_k = [c_k | E_k | \Delta E_k | \Delta \Delta_k | \Delta \Delta E_k],$$

kur X_k – k-tojo signalo požymių vektorius, c_k – MDKK koeficientas arba kiti kepstrinės, spektrinės ar tiesinės prognozės koeficientai, E_k – kalbos signalo energijos, ΔE_k – energijos delta koeficientas, $\Delta \Delta_k$ – MDKK akceleracijos koeficientas (dviguba delta), $\Delta \Delta E_k$ – energijos akceleracijos koeficientas.

Reikia paminėti, kad mokslininkai vis dar nerado universalus požymių rinkinio, kuris vienareikšmiškai identifikuotų nagrinėjamą kalbos signalo fragmentą. Visi išskiriami požymiai turi savų privalumų ir trūkumų.

Požymių atitikimui (Feature matching) naudojami įvairūs metodai. Nepriklausomos nuo teksto sistemos atpažinimą atlieka remiantis tokiais metodais:

- Stochastiniai modeliai:

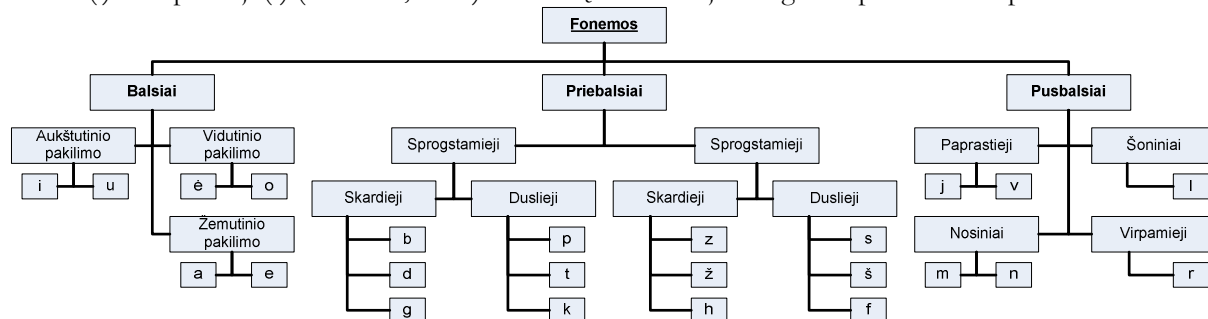
- Diskretiniai ir tolydiniai paslėptieji Markovo modeliai (Hidden Markov Models);
- Gauso mišiniais paremti PMM (Gaussian Mixture models);
- Etalonų palyginimo modeliai:
 - Vektorių kvantavimas (Vector Quantization, VQ);
 - Dinaminis laiko transformavimas (Dynamic Time Wrapping, DTW);
- Dirbtiniai neuroniniai tinklai (Neural Networks).

Priklausomos nuo teksto sistemos diktorių atpažįsta pagrinde naudojant paslėptuosius Markovo modelius (Huang ir kt., 1990).

Fonetiniai vienetai diktoriaus atpažinime

Fonemomis yra vadinami šnekamosios kalbos elementarūs lingvistiniai vienetai. Kiekvieną šnekamosios kalbos sistemos garsą (fonemą) galima apibūdinti artikuliaciniu, akustiniu ir funkciniu aspektais. Fonema įvertinta akustinėje aplinkoje vadinama alofonu, o sistemos įvertinimas – koartikuliacijos efektu.

Lietuvių kalbos fonemos pagal tarimo būdą, tradiciškai skirstomos į balsius, priebalsius ir pusbalsius. Balsiai savo ruožtu gali būti skirstomi į aukšutinio (i, u), vidutinio (ė, o) ir žemutinio pakilimo (a, e). Priebalsiai skirstomi į sprogstamuosius (pvz. b, p) ir pučiamuosius (z, š). Pusbalsiai būna paprastieji (j, v), nosiniai (m, n), šoniniai (l) ir virpamieji (r) (Girdenis, 1995). Fonemų klasifikacijos diagrama pateikiama 3 paveiksle.



40 pav. Fonemų klasifikacijos diagrama

Skirtingoms klasifikacijos šakoms priklausančios fonemos turi skirtingas pagrindines charakteristikas, todėl tiriama, kaip skirtingų kategorijų fonemos tinka diktoriaus identifikavimui. Be to, skirtingoms kategorijoms priklausančių fonemų grupei atpažinti atpažinti gali būti taikomi skirtingi metodai (pvz. skardžiuosius priebalsius tiksliausiai modeliuoja tiesinio prognozavimo metodas), tad atliekant diktoriaus atpažinimą, reikia įvertinti ir šią galimybę.

Kuriama sistema

Kuriamoje sistemoje siekiama įvertinti fonetinį frazės turinį, nes dauguma šiuolaikinių sistemų neatlieka autorizavimo frazės turinio fonetinio įvertinimo. Sistema kuriama lietuviško garsyno LTDIGITS pagrindu, duomenų apdorojimui naudojami 100 abiejų lyčių diktorių kalbos duomenys. Sistema remiasi ne visos frazės atpažinimu, tačiau fonemų, kurios labiau tinka duomenų užkodavimui ir atpažinimui, naudojimu.

Turimi duomenys yra suskirstyti į 50 diktorių vyrų katalogus ir 50 diktorių moterų katalogus, kiekvienas diktorius priklausomai nuo lyties turi savo identifikacinį kodą L0XY, kur L – lyties identifikatorius (F arba M), X – skaičius nuo 0 iki 5, Y – skaičius nuo 0 iki 9 (diktoriaus identifikacinio kodo pavyzdžiai yra F032, M005). Kiekvienas diktorius yra įgarsinęs po 8 frazes, kurias sudaro skaičių pavadinimai (pvz., „vienas“, „du“, „devyni“ ir kt.) bei kiti žodžiai (pvz. „taip“, „ne“, „start“, „stop“ ir kt.). Kiekvienas diktorius bent kartą yra ištaręs visus lietuvių kalbos garsus.

Sistemos kūrimas susideda iš dviejų etapų – fonemų tyrimo ir autorizavimo frazės parinkimo algoritmo sudarymo. Pirmajame etape atliekamas garsyno duomenų apdorojimas, kurio metu išskirti kalbos signalo požymiai, sudaryti kiekvienos fonemos požymių vektoriai. Remiantis turimais LTDIGITS garsyno duomenimis, pirmojo etapo metu naudoti 50 vyrų diktorių visos frazės ir 50 moterų diktorių visos frazės. Kadangi pastaruoju metu Melo dažnių kepstrinių koeficientų skaičiavimai rodo geriausius rezultatus sprendžiant kalbos atpažinimo problemas, keliama hipotezė, kad MDKK turėtų tikti ir diktoriaus atpažinimo sistemose, todėl tyrimui naudojami duomenys frazių ir fonemų ribų duomenys kartu panaudojant kalbos atpažinimo uždavinius sprendžiančių VU KHF mokslininkų atliktų Melo dažnių kepstrinių koeficientų skaičiavimų rezultatus (skaičiuojant MDKK koeficientus naudotas 16 ms. lango ilgis, 6,25 ms žingsnis, 20 filtrų Melo filtrų banke ir 12 kepstrinių koeficientų. Požymių vektorius buvo sudaromas remiantis tiesinio

prognozavimo bendraja metodika, papildomai šalia 12 MDKK koeficientų naudojant energiją bei pirmąsias ir antrąsias jų išvestines) (Driaunys, 2004). Atliekant duomenų analizę, fonemai skaičiuotas kiekvieno iš 100 diktorių tos fonemos, besikartojančios visose diktoriaus frazėse, požymių vektoriai, kuriuos sudaro 39 elementai. Vėliau, suskaičiavus visus vyrų diktorių vienos fonemos požymių vektorius, skaičiuotas vyrų diktorių požymių vektorių vidurkio vektorius, o suskaičiavus moterų – moterų vektorių vidurkio vektorius. Be to, paskaičiuotas visų 100 diktorių vienos fonemos vektorių vidurkio vektorius. Analogiški skaičiavimai atlikti su visomis fonemomis.

Turint pirminius fonemų vektorių vidurkių paskaičiavimus, išskiriamos individualios savybės ir tinkamiausios fonemos. Tinkamumas suprantamas kaip fonemos sugebėjimas su mažiausia klaidos tikimybe iš fonemos išgauti maksimalias diktorių identifikuojančias savybes. Preliminarūs atrinktų fonemų rezultatai kol kas neskelbiami, nes atliekamas šių fonemų efektyvumo patikrinimas ir įvertinimas.

Antroje fazėje sudaromos autorizavimo frazės, kurių pagrindu bus naudojamos atrinktos tinkamiausios fonemos. Numatoma atlikti frazių tinkamumo analizę ir įvertinimą. Šiai dienai projektuota, kad sistema bus priklausoma nuo teksto arba pateikiamo teksto tipo, tačiau tęsiant mokslinius tyrimus, didžiausias perspektyvas praktiniame panaudojime turėtų lietuviškai šnekančio diktoriaus nepriklausoma nuo teksto diktoriaus atpažinimo sistema. Siekiant užtikrinti, kad sistema veiktų teisingai, reiktų papildomai atlikti sistemos modulio, atpažįstančio šnekamąją kalbą, projektavimą. Modulis turėtų neleisti suveikti nepriklausomai nuo teksto sistemai, kuomet tekstas tariamas užsienio, o ypač artimiausia lietuvių kalbai latvių kalba. Tokios sistemos taikymu jau dabar domisi Lietuvos respublikos teisėsaugos institucijos ir specialiosios tarnybos, sistema galėtų būti taikoma kariuomenėje ar tiesiog kuriant protingąją namų aplinką.

Išvados

1. Straipsnyje analizuojamos diktoriaus atpažinimo sistemos, apžvelgiami dažniausi ir geriausi rezultatus duodantys metodai. Kuriant sistemą požymių išskyrimas atliekamas naudojant Melo dažnių kepstrinius koeficientus, o požymių palyginimas – paslėptųjų Markovo modelių metodo pagalba.
2. Straipsnyje atskleidžiama bendroji fonemų – pagrindinių signalų apdorojimo vienetų – klasifikacija. Fonemos yra skirstomos į balsius, priebalsius ir pusbalsius, kurie yra dar labiau smulkinami 2 – 3 lygių gylio kategorijas.
3. Kuriamą sistema, skirta atpažinti lietuviškai kalbantį diktorių. LTDIGITS garsyno duomenys yra apdorojami, išskirti fonemų požymiai, atlikti fonemų požymių vektorių analizės darbai. Eksperimentais pasirinktos fonemos bus naudojamos slaptąžodinių frazių sudaryme, tačiau kol kas atliekamas jų tinkamumo tikrinimas.

Literatūros šaltiniai

1. Driaunys K. (2004) Lietuvių kalbos signalų duomenų bazės LTDIGITS segmentavimas ir statistinis tyrimas. Informacijos mokslai. ISSN 1392-0561. Nr. 28. p.87-96.
2. Girdenis A. (1995) Teoriniai fonologijos pagrindai. Vilnius. Petro ofsetas.
3. Hermansky, H., Morgan, N., Bayya, A., Kohn, P. (1991). RASTA-PLP Speech Analysis. ICSI Technical Report, 1991
4. Huang X.D., Ariki Y., Jack, M.A. (1990) Hidden Markov models for speech recognition, Edinburgh University press.
5. Young S., Evermann G., Kershaw D., Moore G., Odell J., Ollason D., Valtchev V., Woodland P. (2001) The HTK Book. Microsoft Corporation
6. Makhoul, J. (1975). Linear prediction: A tutorial review. Proceedings of the IEEE, Vol. 63, No. 4., 1975, pp. 561 – 580.
7. Markel D., Gray, Jr. A.H. (1976), Linear Prediction of Speech, Springer Verlag, Berlin.
8. Matsui T., Furui S. (1993). Concatenated phoneme models for text-variable speaker recognition. In ICASSP [ICA93], pp. 391–394.
9. McLaughlin J., Reynolds D.A. (2002). Speaker detection and tracking for telephone transactions. Acoustics, Speech, and Signal Processing, 2002. (ICASSP '02), pp. I-129 – I-132, vol.1.
10. Speechechmag.org (2007). Planet Payment Adds Voice Authentication [interaktyvus]. [žiūrėta 2008 m. sausio 5 d.]. Prieiga per internetą: <<http://www.speechechmag.com/Articles/News/Industry-News/Planet-Payment-Adds-Voice-Authentication-40412.aspx>>
11. Viiki O. (1999) Adaptive Methods for Robust Speech Recognition. PhD thesis. Tampere University of Technology.

SPEAKER RECOGNITION SYSTEM BASED ON PHONEMES DISCRIMINATION

S u m m a r y

The quantity of information nowadays became extremely large so the common solutions are not efficient any more when it is speaking about information management. The new approach is to handle information using person's voice. Speech signal processing is split in to speech recognition, language recognition and speaker recognition. Speaker recognition systems (SRS) on the other hand are categorized into speaker identification and speaker verification systems by what they are designed to do. The text said by speaker identifies SRS can be Text independent, Text dependent and Text-Prompted. In practice SRS are used for voice dialing, banking by phone, databases access services, smart houses, etc. The most important methods applied in speech and speaker recognition systems are presented. Some of them, like Mel Frequency Cepstral Coefficients (MFCC) and Hidden Markov Models (HMM), are used in the experimental part of the research. The main material for these methods and systems that is why Lithuanian phonemes and its classification is described. In the experimental part 90 speaker phrases where used for counting the feature vectors and the average vectors of feature vectors of each phonemes. According to the results the most valuable features and phonemes where picked to create experimental password phrases. These phrases will be used in further experiments of text independent or test prompted speaker recognition system which is being developed.