

**VILNIAUS UNIVERSITETAS  
KAUNO HUMANITARINIS FAKULTETAS**

**INFORMATIKOS KATEDRA**

Verslo informatikos studijų programa  
Kodas 62109P101

**ARVYDAS MINKŠTIMAS**

**MAGISTRO BAIGIAMASIS DARBAS**

**FONETINIO SPROGIMO KALBOS SIGNALE  
PAIEŠKA IR TYRIMAS**

Kaunas 2009

**VILNIAUS UNIVERSITETAS  
KAUNO HUMANITARINIS FAKULTETAS**

**INFORMATIKOS KATEDRA**

**ARVYDAS MINKŠTIMAS**

**MAGISTRO BAIGIAMASIS DARBAS**

**FONETINIO SPROGIMO KALBOS SIGNALE  
PAIEŠKA IR TYRIMAS**

Leidžiama ginti \_\_\_\_\_

Magistrantas \_\_\_\_\_  
(parašas)

Darbo vadovas \_\_\_\_\_  
(parašas)

**Dr. Kęstutis Driaunys** \_\_\_\_\_  
(darbo vadovo mokslo laipsnis, mokslo  
pedagoginis vardas, vardas ir pavardė)

Darbo įteikimo data \_\_\_\_\_

Registracijos Nr. \_\_\_\_\_

Kaunas 2009

# TURINYS

PAVEIKSLĖLIŲ SĄRAŠAS .....	4
LENTELIŲ SĄRAŠAS.....	5
PRIEDŲ SĄRAŠAS .....	6
SANTRAUKA .....	7
ĮVADAS.....	8
1 ANALIZĖS DALIS.....	11
1.1 Sprogstamieji priebalsiai .....	11
1.2 Atvaizdavimas spektrogramose.....	12
1.3 Sprogstamųjų priebalsių nustatymo metodai .....	13
1.3.1 Akustinėmis žiniomis grindžiamas modelis.....	13
1.3.2 V. Colotte ir Y. Laprie modelis.....	18
1.3.3 Hierarchinis fonemų klasifikavimo modelis su fonemų grupėmis.....	19
1.4 Analizės dalies išvados.....	22
2 SPROGSTAMŲJŲ PRIEBALSIŲ ATPAŽINIMO METODIKA .....	23
2.1 Kalbos signalo požymių skaičiavimas .....	23
2.1.1 Energijos maksimumų kitimų ir uždarumos skaičiavimas.....	24
2.2 Sprogstamųjų priebalsių atpažinimo siūlomi patobulinimai.....	25
2.2.1 Energijos minimumų ir maksimumų santykių skaičiavimais paremtas algoritmas .....	26
2.3 Sprogstamųjų priebalsių atpažinimo metodikos išvados .....	34
3 EKSPERIMENTINIS SKYRIUS .....	35
3.1 Duomenys.....	35
3.2 Eksperimentinės aplinkos formavimas.....	36
3.3 Algoritmo realizacija.....	36
3.4 Rezultatai.....	37
3.4.1 Eksperimentas kai santykiečiai skaičiuojami vienoje filtro juostoje .....	37
3.4.2 Eksperimentas kai santykiečiai skaičiuojami skirtingose filtrų juostose.....	39
3.5 Eksperimentinio skyriaus išvados .....	42
IŠVADOS.....	43
LITERATŪRA.....	44
PRIEDAI.....	47

## PAVEIKSLĖLIŲ SĄRAŠAS

1 pav.	Sprogstamieji priebalsiai. ....	11
2 pav.	Sprogstamojo garso susidarymas. ....	11
3 pav.	Spektrograma. Sprogstamųjų priebalsių „p“ ir „b“ sprogimai. ....	12
4 pav.	Priebalsių „K“ ir „G“ sprogimas. ....	13
5 pav.	Uždarumo sprogstamuosiuose priebalsiuose nustatymo algoritmas. ....	14
6 pav.	Artikuliacijos vietos nustatymo algoritmas. ....	16
7 pav.	Sprogstamųjų ir pučiamųjų priebalsių nustatymo algoritmas. ....	18
8 pav.	Hierarchinio fonemų klasifikavimo algoritmo scenarijus. ....	20
9 pav.	Sprogstamųjų priebalsių atpažinimo algoritmo schema. ....	23
10 pav.	Siūlomas sprogstamųjų priebalsių atpažinimo algoritmo patobulinimas. ....	25
11 pav.	Minimumų ir maksimumų santykių skaičiavimo algoritmas. ....	26
12 pav.	Energijos minimumų ir maksimumų santykių skaičiavimo principinė schema. ....	27
13 pav.	Žodžio „keturi“ vaizdavimas visuose filtrų juostose. ....	28
14 pav.	Fonemos „p“ palyginimas skirtingose vietose. ....	29
15 pav.	Sprogstamųjų priebalsių uždarumų prieš ir po sprogimo trukmės. ....	30
16 pav.	Akustinio sprogimo fonemoje „t“ (žodyje „septyni“) nustatymas. Nustatymui panaudota 3 filtro juosta (2500-3500Hz). ....	31
17 pav.	Sprogimo nustatymo rezultatai. ....	33
18 pav.	Fonemų atpažinimo rezultatai. ....	38
19 pav.	Fonetinio sprogimo atpažinimo daromos klaidos. ....	40

## LENTELIŲ SĄRAŠAS

1 lentelė	Rezultatai gauti naudojant uždarumo nustatymo algoritmą.....	15
2 lentelė	Rezultatai gauti naudojant artikuliacijos vietos nustatymo algoritmą.....	17
3 lentelė	Rezultatai nustatant uždarumą ir artikuliacijos vietą.....	17
4 lentelė	Sprogstamųjų priebalsių atpažinimo tikslumas.....	21
5 lentelė	Fonemų atpažinimo metodų suvestinė.....	21
6 lentelė	Fonemų energijos pauzių vidurkiai.....	26
7 lentelė	Minimumų ir maksimumų santykiai.....	32
8 lentelė	Gautų rezultatų palyginimas skaičiuojant papildomus požymius ir ne.....	37
9 lentelė	Atskirų fonemų gauti atpažinimo rezultatai.....	37
10 lentelė	Fonetinio sproginimo atpažinimo klaidos.....	39
11 lentelė	Fonetinio sproginimo atpažinimas santykius skaičiuojant skirtingose filtrų juostose.....	39
12 lentelė	Padarytos klaidos nustatant fonetinį sproginimą.....	40
13 lentelė	Atpažinimo rezultatai.....	41
14 lentelė	Klaidų palyginimas skaičiuojant ir neskaičiuojant papildomus požymius.....	41

## **PRIEDŲ SĄRAŠAS**

1 PRIEDAS.	Fonemos pradžios ir pabaigos ribų nustatymas.....	47
2 PRIEDAS.	Filtro juostos nustatymas pagal žodžio ir fonemos energijas.....	48
3 PRIEDAS.	Energijos maksimumų langų nustatymas. ....	49
4 PRIEDAS.	Energijos minimumų langų prieš maksimumus nustatymas. ....	50
5 PRIEDAS.	Energijos minimumų po maksimumų nustatymas. ....	51
6 PRIEDAS.	Minimumų ir maksimumų santykių įvertinimas. ....	52
7 PRIEDAS.	Fonetinio sproginimo nustatymo rezultatų apjungimas. ....	53

# Santrauka

MINKŠTIMAS, Arvydas. (2009) *Stop Consonant Burst Detection*. MBA Graduation Paper. Kaunas: Vilnius University, Kaunas Faculty of Humanities, Department of Informatics. 54 p.

## SUMMARY

Scientists for a time are interested in language recognition sphere. There are made a lot of researches on purpose to make various mechanisms understand human language and let easily communicate humans and mechanism. The problem is that the same rules can't be used for different languages recognition. It is because people have their speaking manners and languages differs one from other. For these reasons every country scientists searches best rules for their language recognition.

Probably most difficult to recognize stop consonants (Lithuanian language have 6 stop consonants: /b/, /d/, /g/, /p/, /k/, /t/) for their specific characteristics. Stop consonants are very short phonemes and this feature often hides burst. Usually stop consonants burst detection algorithms can recognize stop consonants with ~90% accuracy.

Have been set 4 goals for this work:

1. To analyze methods of stop consonants recognition;
2. To make algorithm for stop consonant recognition based on occlusive and maximum variation of energy detection and add new characteristics search and analysis;
3. To make new algorithm experimental research with LTDIGITS speech database;
4. To determinate whether new characteristics calculation leads to better results.

In this research algorithm was extended adding calculations of occlusive length and aspiration after plosion length, which helped to detect burst in phoneme better. Using these calculations 221 stop consonants more were recognized than without it (for experiments used more than 2500 phonemes, from which were 6454 stop consonants) and lets achieve 90.68% accuracy.

# IVADAS

## **Magistro baigiamojo darbo tema**

Fonetinio sprogimo kalbos signale nustatymas ir tyrimas.

## **Temos aktualumas**

Kalbos atpažinimas – mokslininkus seniai dominanti sritis. Atliekamais tyrimais siekiama priversti mašinas kuo tiksliau suprasti žmogaus tariamus garsus ir jų reikšmes, taip supaprastinant žmonių ir mašinų tarpusavio bendravimą. Dėl kalbų skirtumų ir žmonių kalbėjimo manierų tos pačios kalbos signalo atpažinimo taisyklės netinka visoms kalboms atpažinti. Dėl šių priežasčių, kiekvienos šalies mokslininkai ieško geriausių taisyklių rinkinių savo šalies kalbos signalo atpažinimui.

Bene sunkiausiai sekasi atpažinti sprogstamuosius priebalsius (lietuvių kalboje jų yra 6: /b/, /d/, /g/, /p/, /k/, /t/) dėl jų specifinių savybių. Sprogstamieji priebalsiai - tai labai trumpos, nuo diktoriaus kalbėjimo manierų, akcento, lyties, aplinkos triukšmo ir pan. priklausančios, fonemos. Atliekant signalo analizę dėl fonemos trumpumo dažnai sudėtinga pastebėti sprogstamųjų priebalsių požymius. Atlikti tyrimai ir sudaryti modeliai leidžia atpažinti sprogstamuosius priebalsius ~90 % tikslumu.

## **Problemos ištyrimo lygis**

Atliktos metodų analizės: Ahmed M. Abdelatty Ali ir k.t. uždarumos ir artikuliacijos vietos nustatymo; V. Colotte ir Y. Laprie fonemų atpažinimo pagal jų energijas; K. Driaunio uždarumos paieškos ir energijos maksimumų kitimų skaičiavimo. Visi nagrinėti modeliai skirsto fonemas į sprogstamųjų ir nesprogstamųjų fonemų grupes gana tiksliai, tačiau 100 % tikslumo nepasiekia nei vienas nagrinėtas modelis. Geriausius rezultatus leidžia pasiekti uždarumos paieškos bei energijos maksimumų kitimų skaičiavimais paremtas metodas.

## **Darbo objektas**

Sprogstamųjų priebalsių (b, p, k, g, d, t) atpažinimo algoritmas.

## **Baigiamojo darbo tikslas**

Patobulinti sprogstamųjų priebalsių atpažinimo algoritmą papildomai įvedant naujų požymių skaičiavimus.



## **Darbo uždaviniai**

1. išanalizuoti sprogstamųjų priebalsių atpažinimui naudojamus metodus bei įvertinti jų efektyvumą;
2. realizuoti algoritmą, naudojant uždarumos paieškos ir energijos maksimumų kitimų skaičiavimus bei papildant naujais sprogstamųjų priebalsių atpažinimo požymiais: fonemos energijos minimumų ir maksimumų santykių skaičiavimai;
3. atlikti sudaryto algoritmo eksperimentinius tyrimus panaudojant LTDIGITS garsyno įrašus;
4. nustatyti ar pasiteisino naujų požymių skaičiavimai remiantis gautais rezultatais.

## **Darbe naudojama metodika**

Išsikeltų uždavinių įgyvendinimui panaudoti metodai: mokslinės literatūros analizės, palyginimo, duomenų analizės, modeliavimo bei modelio testavimo ir analizės metodai.

## **Darbo rezultatų praktinė reikšmė**

Įtraukus papildomų požymių skaičiavimus gautas geresnis sprogstamųjų priebalsių, kurių trukmė iki 200 ms, atpažinimo rezultatas. Skaičiuojant fonemos energijos minimumų ir maksimumų santykius galima nustatyti daugiau trumpų sprogstamųjų priebalsių, tačiau trumpuose signalo atkarpose dažnai staigesnis energijos pokytis buvo fiksuojamas kaip fonetinis sproginimas todėl padaryta klaidų nustatant fonetinį sproginimą ten kur jo nebuvo. Taigi, eksperimentu metu skaičiuojant papildomus požymius buvo atpažinta 221 sprogstamąja fonema daugiau, nustatyta 106 klaidingais fonetiniiais sproginimais daugiau, o bendras fonemų skirstymo į sprogstamųjų priebalsių ir nesprogstamųjų fonemų grupes tikslumas pagerėjo nuo 90,23% iki 90,67%.

## **Rezultatų aprobavimas**

Darbo rezultatai paskelbti publikacijoje:

1. Minkštinas A. Fonetinio sproginimo kalbos signale paieška ir tyrimas. 14-osios tarpuniversitetinės magistrantų ir doktorantų mokslinės konferencijos „Informacinės technologijos“. Pranešimų medžiaga. Kaunas, 2009, p. 12-16.

## **Darbo struktūra ir apimtis**

Magistrinis darbas sudarytas iš įvado, teorinės/analitinės dalies, metodinės dalies, eksperimentinės dalies, išvadų, literatūros sąrašo ir 7 priedų. Darbo aprašas apima 43 puslapius. Darbe pateikta 19 paveikslėlių ir 14 lentelių.

## **Darbo struktūros aprašymas**

Įvade aptariamas temos aktualumas, problemos ištyrimo lygis, darbo objektas, baigiamojo darbo tikslas, uždaviniai bei struktūra.

Pirmojoje darbo dalyje analizuojamos fonetinio sprogimo susidarymo priežastys bei sprogstamųjų priebalsių požymiai. Taip pat pateikti trys sprogstamųjų priebalsių atpažinimo darbai, juose naudotos atpažinimo metodikos bei gauti rezultatai.

Antroje darbo dalyje aprašoma metodika, kuria remiantis parašytas algoritmas, ir siūlomi algoritmo patobulinimai.

Eksperimentinėje dalyje aprašomi du eksperimentiniai bandymai atlikti su „The MathWorks“ produktu MATLAB R2008a sukurtu sprogstamųjų priebalsių atpažinimo algoritmu. Skyriaus pabaigoje palyginami gauti rezultatai.

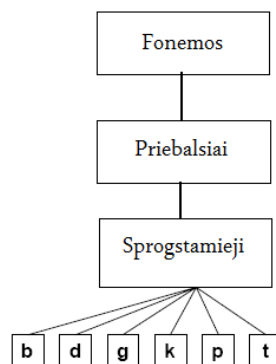
Darbo apibendrinimas pateiktas darbo išvadose.

# 1 ANALIZĖS DALIS

Šioje dalyje yra analizuojamos sprogstamųjų priebalsių fonetinės savybės. Taip pat apžvelgti trys eksperimentiniai sprogstamųjų priebalsių atpažinimo tyrimai.

## 1.1 Sprogstamieji priebalsiai

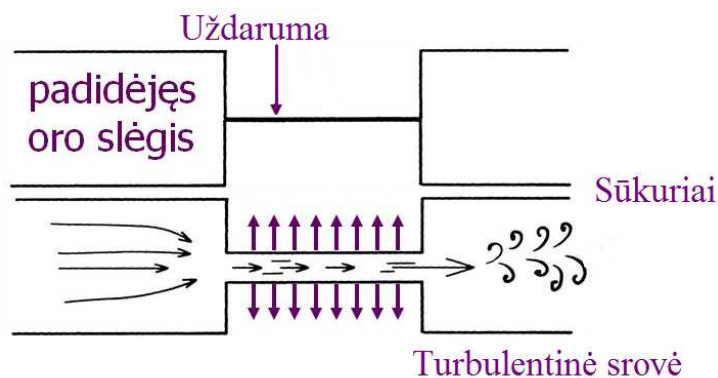
Fonemos – elementarūs šnekamosios kalbos lingvistiniai vienetai. Kalbos garsų sistemoje yra išskiriamos dvi pagrindinės garsų grupės – balsiai ir priebalsiai (K. Driaunys, 2006). Sprogstamieji priebalsiai tai tokie garsai, kurie susidaro stabdant oro srovę burnos ertmėje ir staiga ją išleidus gaunasi tarsi sprogimas (Ian Maddieson, 1984). Lietuvių kalboje yra šešios sprogstamųjų priebalsių grupės fonemos (*b, d, g, p, k, t*). Pagal akustines savybes priebalsiai yra skirstomi į skardžiuosius ir dusliuosius. Taip pat sprogstamieji priebalsiai skirstomi pagal artikuliacijos vietą į lūpinius, gomurinius ir dantinius (Ian Maddieson, 1984).



Šaltinis: sudaryta autoriaus.

1 pav. Sprogstamieji priebalsiai.

Visiems sprogstamiesiems priebalsiams būdinga tai, kad iš pradžių yra sudaroma aklina pertvara – uždaruma. (R. Ambrazevičius, 2008)



Šaltinis: sudaryta autoriaus remiantis Ambrazevičius R. *Sprogstamieji priebalsiai*.

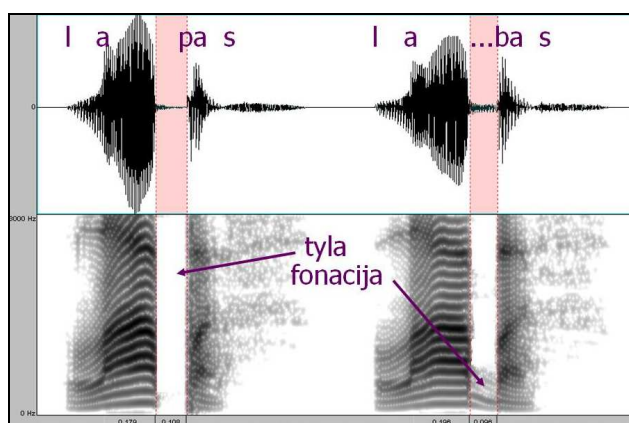
2 pav. Sprogstamojo garso susidarymas.

Kol yra susidarius uždaruma, oro srovė iš burnos ertmės negali išeiti, todėl netiriamas joks garsas. Oro srovė įveikdama uždarumą, jį staigiai atidaro todėl susidaro sproginimas (Ian Maddieson, 1984).

## 1.2 Atvaizdavimas spektrogramose

Spektrogramoje horizontalioji ašis yra laiko ašis, o vertikalioji – dažnis. Šešėlių intensyvumas nusako trečiąją dimensiją – amplitudę. Spektrograma – tai visa eilė spektrų, atvaizduotų žiūrint iš viršaus (J. Purlytė, 2003).

Sprogstamieji priebalsiai spektrogramose yra vaizduojami kaip smarkus energijos išsiveržimas po tylos (pauzės). Šių priebalsių skiriamasis bruožas yra beveik momentinis perėjimas nuo mažos arba neakustinės energijos prie trumpo aukštos energijos išsiveržimo plačioje dažnio juostoje (J. Purlytė, 2003).

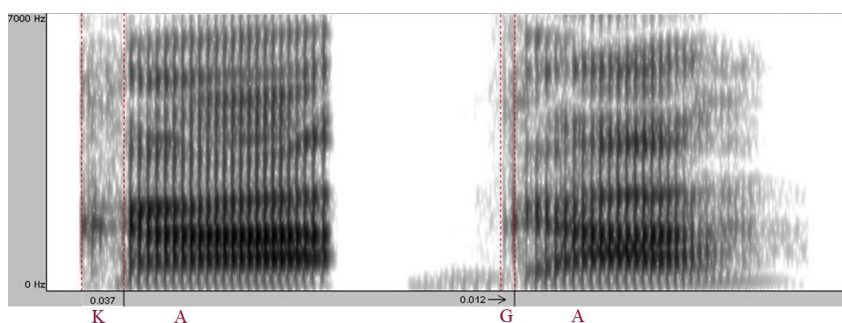


Šaltinis: Ambrazevičius R. *Sprogstamieji priebalsiai*.

### 3 pav. Spektrograma. Sprogstamųjų priebalsių „p“ ir „b“ sproginimai.

Spektrogramoje (žr. *Paveiksle 3*) pavaizduoti du žodžiai „lapas“ ir „labas“ bei juose įvykstantys sproginimai. Spektrogramoje matoma, jog prieš sprogstamąjį priebalsį atsiranda tylos spraga, kurios gale „sprogsta“ priebalsis. Tipiška tylos trukmė yra ~50 – 150 ms. (R. Ambrazevičius, 2008) Pastebėtina, kad prieš tylos tarpą esantis balsis yra ilgesnis.

4 *paveiksle* matyti, jog skardžiųjų sprogstamųjų priebalsių sproginimo laikas yra trumpesnis nei dusliųjų. Visiems sprogstamiesiems priebalsiams būdinga trumpa trukmė, jų sproginimas vyksta tik iki ~75 ms, kai tuo tarpu afrikatos (sutaptinis priebalsis, susidaręs iš sprogstamojo ir pučiamojo priebalsių (c, č, dz, dž)) trunka nuo 75 – 130 ms, o pučiamieji priebalsiai virš 130 ms (R. Ambrazevičius, 2008).



Šaltinis: Ambrazevičius R. Sprogstamieji priebalsiai.

#### 4 pav. Priebalsių „K“ ir „G“ sprogimas.

Sprogstamųjų priebalsių identifikavimui naudojami sprogimo parametrai (R. Ambrazevičius, 2008):

- Sprogimo spektro gaubtinė (nuolydis, maksimumai);
- Sprogimo spektro laikinis kitimas (iki įbalsio);
- Sprogimo amplitudė.

### 1.3 Sprogstamųjų priebalsių nustatymo metodai

#### 1.3.1 Akustinėmis žiniomis grindžiamas modelis

Ahmed M. Abdelatty Ali ir k.t. atliko eksperimentą panaudodami du naujus metodus, tai – uždarmo nustatymo ir artikuliacijos vietos nustatymo. Iš pradžių uždarmo algoritmu buvo nustatoma koks tai priebalsis (skardusis ar duslusis), o artikuliacijos metodo pagalba pagal išskirtus bruožus nustatomas koks tai priebalsis. Šie algoritmai rėmėsi trimis signalo savybėmis (Ahmed M. Abdelatty Ali ir k.t., 2001):

- a) formančių trajektorijos ir sprogimo spektras;
- b) sprogimo lygis;
- c) trukmė ir uždaruma.

Eksperimentas buvo atliktas su TIMIT garsynu, iš kurio pasirinkta 60 diktorių su 7 skirtingomis tarmėmis (Ahmed M. Abdelatty Ali ir k.t., 2001).

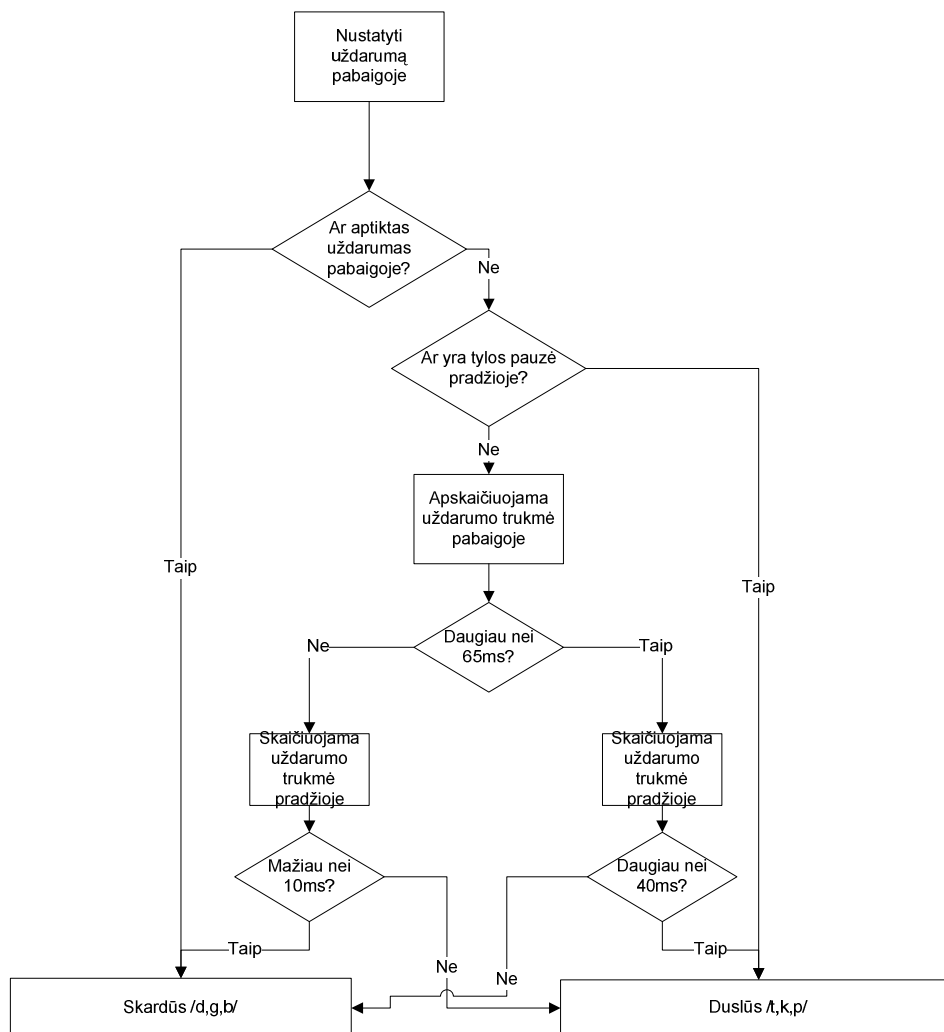
Nustatinėjant priebalsio uždarumą buvo naudoti trys signalo požymiai, pagal kuriuos garsai buvo skirstomi į skardžiuosius ir dusliuosius (Ahmed M. Abdelatty Ali ir k.t., 2001):

- a) uždarumas pabaigoje;
- b) uždarumas pradžioje;
- c) trukmės nustatymas.

Uždarumas pabaigoje buvo fiksuojamas paskaičiuojant koeficientą tarp žemo dažnio vidutinės energijos (iki 450Hz) trunkančios 20ms signalo pabaigoje ir maksimalaus koeficiento viso signalo

metu. Jei yra peržengiamas nustatytas slenkstis, tuomet fiksuojama uždaruma (Ahmed M. Abdelatty Ali ir k.t., 2001).

Kaip matoma *paveiksle 5* jei nustatomas uždarumas garso signalo pabaigoje, tuomet iš karto jis yra priskiriamas prie skardžiųjų, o jei randamas pradžioje - prie dusliųjų. Kitais atvejais yra skaičiuojamos uždarumo trukmės signalo pradžioje ir pabaigoje.



Šaltinis: sudaryta autoriaus remiantis Abdelatty Ali A.M., Van der Spiegel J., Muller P. (2001) Acoustic – Phonetic Features for the Automatic Classification of Stop Consonants.

### 5 pav. Uždarumo sprogstamuosiuose priebalsiuose nustatymo algoritmas.

Naudojant šį algoritmą buvo pasiektas 96% sprogstamųjų priebalsių atpažinimo tikslumas (žr. Lentelė 1). Nors eksperimento autorius pabrėžia, jog uždarumo pabaigoje nustatymas nėra esminė algoritmo dalis (Ahmed M. Abdelatty Ali ir k.t., 2001), tačiau nenaudojant šios algoritmo dalies priebalsių tipo nustatymo tikslumas nuo 96% sumažėja iki 90%.

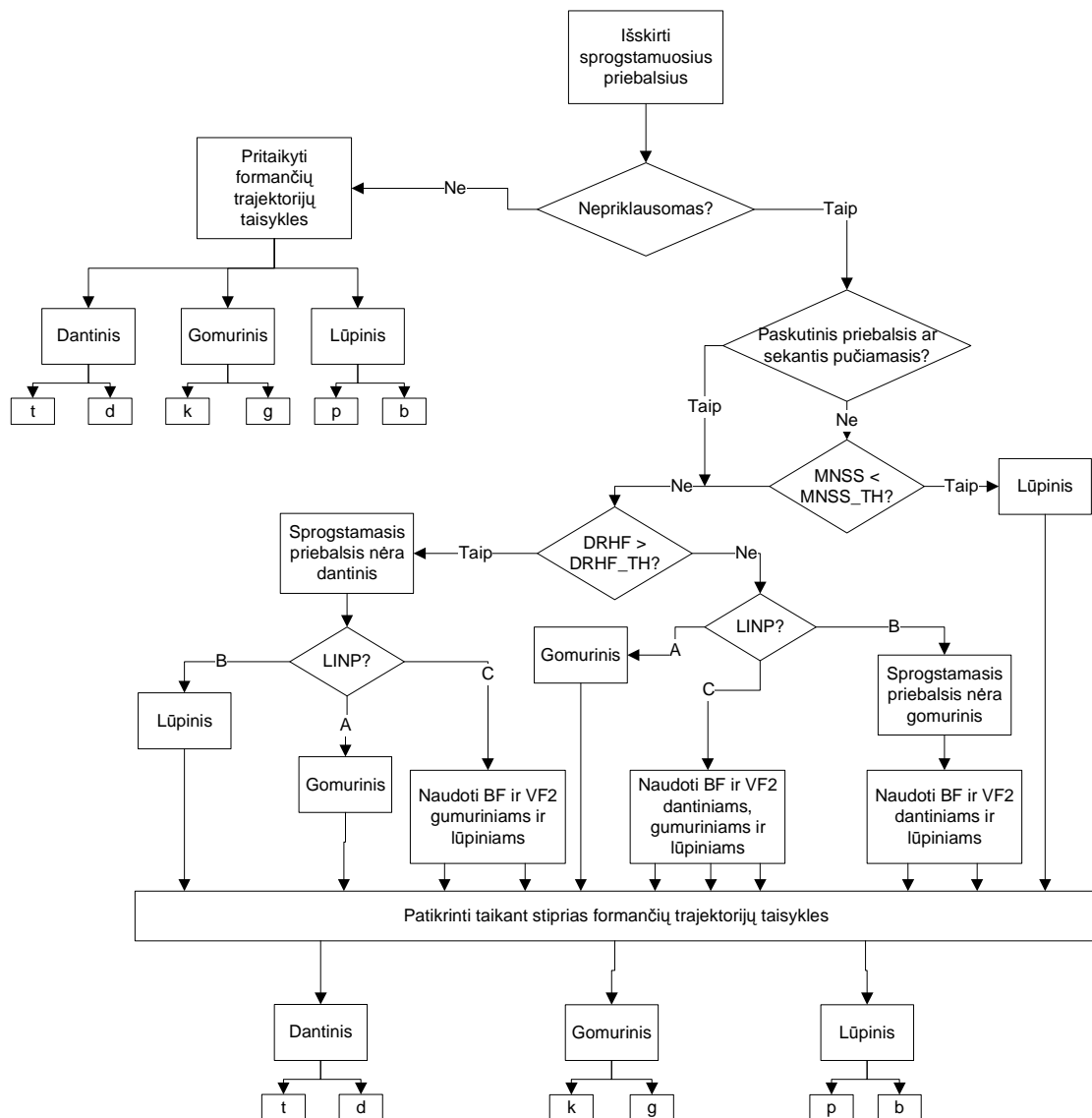
## Rezultatai gauti naudojant uždarumo nustatymo algoritmą.

	Nustatyti kaip skardieji	Nustatyti kaip duslieji
Skardieji	95%	5%
Duslieji	3%	97%

Šaltinis: sudaryta autoriaus remiantis Abdelatty Ali A.M., Van der Spiegel J., Muller P. (2001) Acoustic – Phonetic Features for the Automatic Classification of Stop Consonants.

Kitas pristatomas metodas yra artikuliacijos vietos nustatymas, kuris atpažįsta priebalsius (d, g, b, t, k, p) pagal požymius (Ahmed M. Abdelatty Ali ir k.t., 2001):

- a) dažnio sproginimas (BF) (*The burst frequency*);
- b) antra formantė (F2) sekančiame balsyje (VF2) (įbalsis) (*The second formant of the following vowel*);
- c) maksimaliai normalizuotas spektro nuolydis (MNSS) (*The maximum normalized spectral slope*);
- d) dažnio sproginimo maksimumai (DRHF ir LINP) (*The burst frequency prominence*);
- e) formančių trajektorijos sprogstamojo signalo pradžioje ir pabaigoje (*Formant transitions before and after the stop*);
- f) uždarumos nustatymas (pagal uždarumos nustatymo algoritmą) (*The voicing decision (using the previous section algorithm)*)



Šaltinis: sudaryta autoriaus remiantis Abdelatty Ali A.M., Van der Spiegel J., Muller P. (2001) Acoustic – Phonetic Features for the Automatic Classification of Stop Consonants.

### 6 pav. Artikuliacijos vietos nustatymo algoritmas.

Pagal priebalsių požymius buvo sudarytas algoritmas matomas 6 paveikslėlyje. Kai į algoritmą yra įvedamas sprogstamojo signalo požymių vektorius, visų pirma patikrinama ar sprogstamasis signalas nėra priklausomas nuo kitų garsų, jei sąlyga patenkinama tai sprogstamajam priebalsiui iš karto pritaikomos formančių trajektorijos ir atpažįstamas sprogstamasis signalas. Priešingu atveju signalo tipas nustatomas pagal išskirtus požymius (Ahmed M. Abdelatty Ali ir k.t., 2001).

Su šiuo algoritmu Ahmed M. Abdelatty Ali ir k.t. atliktame eksperimente gautas 90% bendras tikslumas (žr. Lentelė 2). Taip pat eksperimentas buvo atliktas iš algoritmo pašalinus formančių perėjimo taisykles tikslumas nuo 90% nukrito iki 86% (Ahmed M. Abdelatty Ali ir k.t., 2001).



2 lentelė

**Rezultatai gauti naudojant artikuliacijos vietos nustatymo algoritmą.**

	<b>Atpažinti kaip dantiniai (t,d)</b>	<b>Atpažinti kaip gomuriniai (k,g)</b>	<b>Atpažinti kaip lūpiniai (p,b)</b>
Dantiniai	91%	6%	3%
Gomuriniai	3%	88%	9%
Lūpiniai	6%	6%	86%

Šaltinis: sudaryta autoriaus remiantis Abdelatty Ali A.M., Van der Spiegel J., Muller P. (2001) Acoustic – Phonetic Features for the Automatic Classification of Stop Consonants.

Apjungus abu algoritmus gautas bendras 86% sprogstamųjų priebalsių atpažinimo tikslumas (žr. *Lentelė 3*). Šiam eksperimentui buvo panaudota 1200 sprogstamųjų priebalsių signalų (Ahmed M. Abdelatty Ali ir k.t., 2001).

3 lentelė

**Rezultatai nustatant uždarumą ir artikuliacijos vietą.**

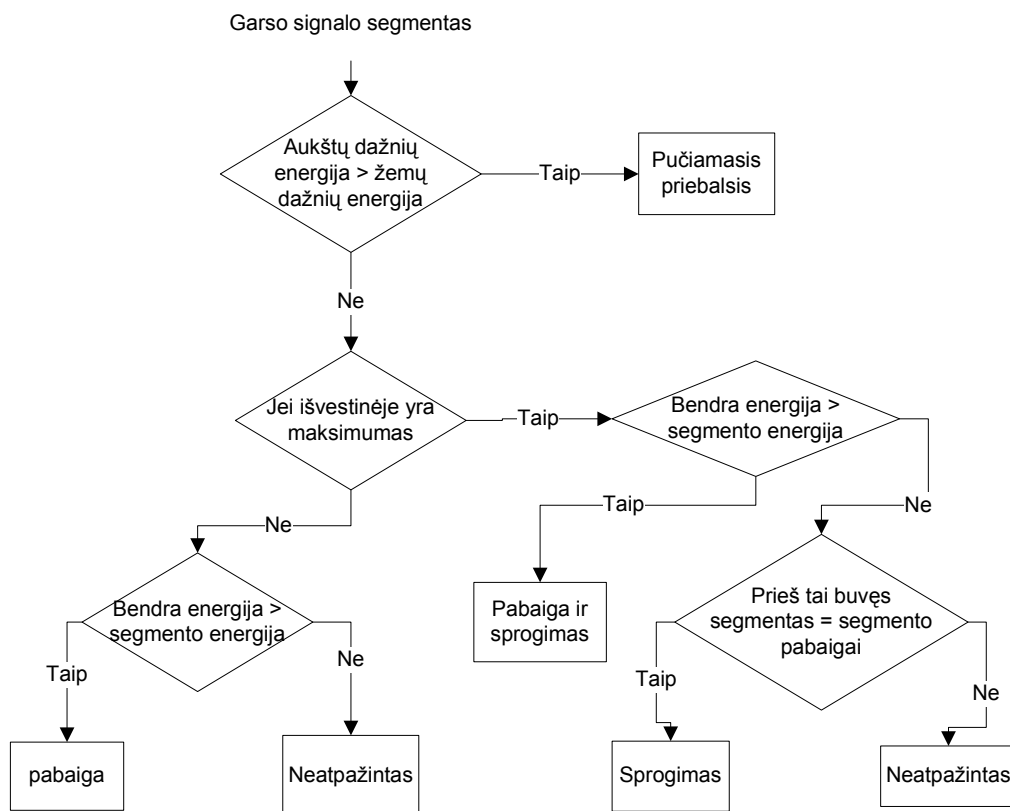
	<b>Atpažinta kaip /t/</b>	<b>Atpažinta kaip /d/</b>	<b>Atpažinta kaip /k/</b>	<b>Atpažinta kaip /g/</b>	<b>Atpažinta kaip /p/</b>	<b>Atpažinta kaip /b/</b>
/t/	87.5%	3.5%	5%	0.5%	3%	0.5%
/d/	3%	88%	0.5%	7%	0.5%	1%
/k/	2.5%	0.5%	87.5%	1%	8%	0.5%
/g/	2%	2%	10%	76%	0%	10%
/p/	7%	0%	7%	0%	83.5%	1%
/b/	0%	5%	0%	5%	2.5%	85.5%

Šaltinis: sudaryta autoriaus remiantis Abdelatty Ali A.M., Van der Spiegel J., Muller P. (2001) Acoustic – Phonetic Features for the Automatic Classification of Stop Consonants.

### 1.3.2 V. Colotte ir Y. Laprie modelis

Šio modelio veikimo principas kiek kitoks nei prieš tai aprašytas. V. Colotte ir Y. Laprie savo modelyje panaudojo TD-PSOLA metodą, kuris geba lengvai modifikuoti kalbos signalą ir F0 – pagrindinį toną. Šio metodo pagalba kalba yra padalinama į žingsnelius ir išskiriami kiekvieno žingsnelio požymiai. Toks kalbos apdorojimas vykdomas siekiant išryškinti signalo požymius (V. Colotte ir Y. Laprie, 2000).

Kalbos signalo svyravimams įvertinti naudojama spektro variacijos funkcija, kuri remiasi signalo energijos požymiu. Sprogstamųjų signalų energija buvo naudojama iš uždarumos laikotarpio. Buvo paimta signalo uždarumos vidurio energijos vidurkis ir lyginamas su viso segmento energijos vidurkiu (V. Colotte ir Y. Laprie, 2000). Apibendrintas signalo nustatymas matomas *paveiksle 7* pavaizduotame algoritme.



Šaltinis: sudaryta autoriaus remiantis Colotte V., Laprie Y. (2000) Automatic enhancement of speech intelligibility

#### 7 pav. Sprogstamųjų ir pučiamųjų priebalsių nustatymo algoritmas.

Eksperimentui atlikti buvo naudojamas TIMIT garsynas iš kurio parinkta 50 sakinių. 25 iš jų buvo modifikuoti TD-PSOLA metodu, o likę 25 palikti tokie kokie yra. Visi testavimo duomenys pasirinkti atsitiktine tvarka. Eksperimento autoriai išskyrė 4ias rezultatų būsenas (V. Colotte ir Y. Laprie, 2000):

1. 0 – nerastas atitikmuo;

2. 1 – atpažinta klaidingai;
3. 2 – atpažinta bent pusė fonemos teisingai;
4. 3 – fonema atpažinta.

Gauti eksperimento rezultatai parodė, kad modifikuojant kalbos signalo tempą neišdarakė signalo moduliacijos, tačiau artikuliacijos pasikeitimas vieną sprogstamąjį priebalsį atpažino kaip pučiamąjį priebalsį. Buvo pasiektas 86,8% sprogstamųjų priebalsių atpažinimo tikslumas (V. Colotte ir Y. Laprie, 2000).

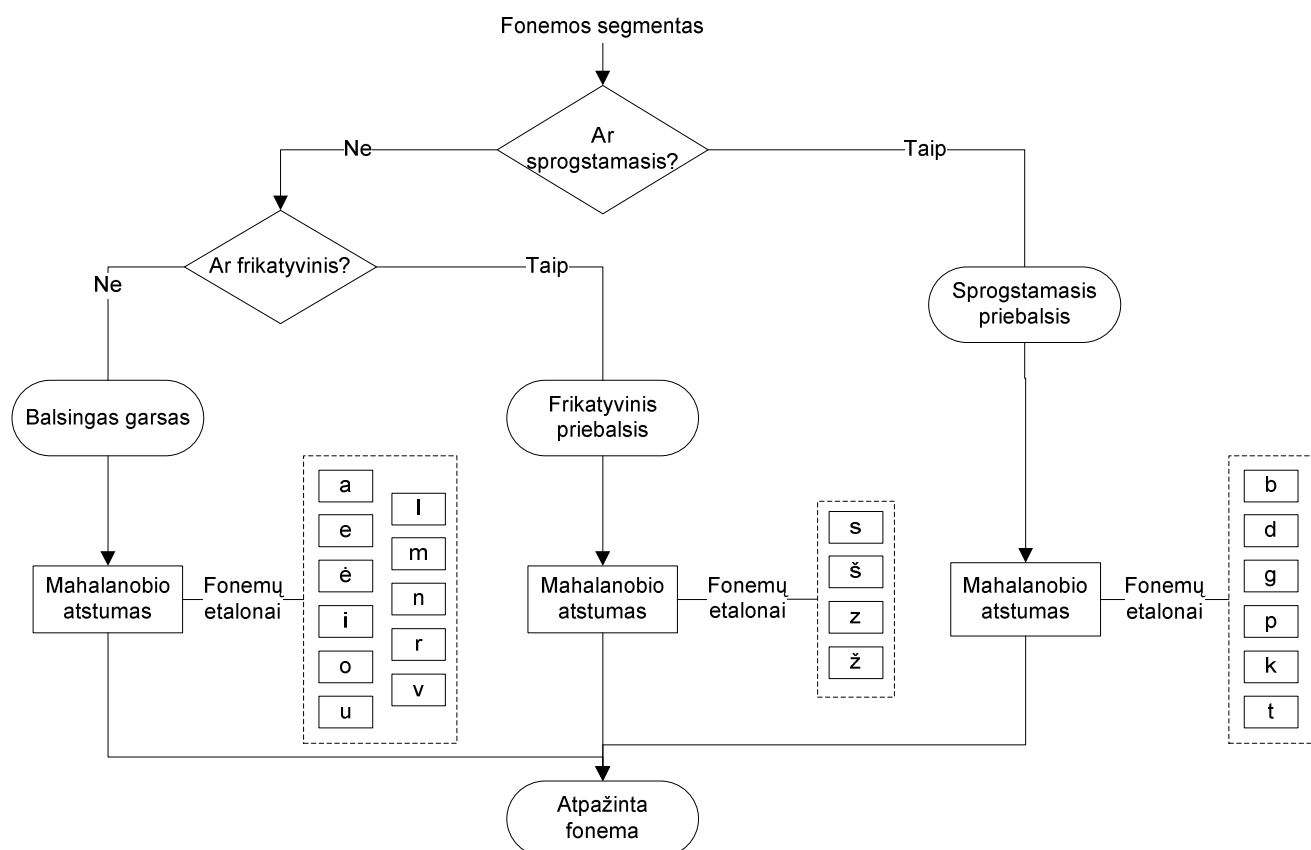
### **1.3.3 Hierarchinis fonemų klasifikavimo modelis su fonemų grupėmis**

Šiame modelyje naudojamas fonemų atpažinimo metodas analizuoja požymių vektorių, lygindamas jo požymius su etaloninėmis reikšmėmis t.y. tokiomis, kokios buvo sudarytos apmokymo metu. Etaloninės dalys yra trys, tai stacionarioji (vidurinioji fonemos dalis), ir dvi dinamiškos (kairės ir dešinės), kurios yra įtakojamos atitinkamai prieš ir po fonemos esančiomis fonemomis ar pauzėmis. Sprogstamųjų priebalsių nustatymui siūloma naudoti dešiniąją fonemos ribą, nes šio tipo garso pradžioje įvyksta pagrindinis įvykis – sproginimas (K. Driaunys, 2006).

K. Driaunys fonemų nustatymui naudoja hierarchinį metodą. Šio metodo esmė yra tai, kad pradžioje vyksta apmokymas, kurio metu sukuriama etalonų vienetai, kurie atpažinimo metu lyginami su atpažinimui pateiktu vektoriumi (K. Driaunys, 2006).

Šis modelis yra papildytas fonemų grupėmis (balsingų garsų, sprogstamųjų priebalsių ir friktyvinių priebalsių). Taigi, apmokymo etape modelis skirsto fonemas į grupes, kurios taip pat įgyja fonemų etalonines reikšmes. Vėliau atpažįstant vektorių jis yra lyginamas su grupės etaloninėmis reikšmėmis ir tik po to su fonemų reikšmėmis, taip išvengiant nereikalingo palyginimo su visomis fonemų etaloninėmis reikšmėmis (K. Driaunys, 2006).

*Paveiksle 8* yra pateiktas fonemų atpažinimo algoritmas. Kadangi visos fonemos suskirstytos į grupes, tai jau pirmame žingsnyje reikšmių vektorius yra palyginamas su sprogstamųjų priebalsių grupės etaloninėmis reikšmėmis. Jei sąlygos yra tenkinamos tuomet atpažinimas vykdomas šios grupės viduje, lyginant su kiekvienos fonemos etaloninėmis reikšmėmis taip atpažįstant fonemą. Jei sąlyga netenkinama, tuomet reikšmių vektorius yra lyginamas su kitų grupių etalonais.



Šaltinis: Driaunys K. Lietuvių šnekamosios kalbos segmentavimo ir fonetinio atpažinimo tyrimas naudojant LTDIGITS garsyno įrašus.

### 8 pav. Hierarchinio fonemų klasifikavimo algoritmo scenarijus.

Ekspertas buvo vykdomas su LTDIGITS garsyno duomenimis, buvo panaudota virš 25000 fonemų ištartų 100 skirtingų diktorių (50 vyrų, 50 moterų). Atlikto tyrimo su sprogstamaisiais priebalsiais metu gautas 93,9% tikslumas panaudojant fonemų grupes (K. Driaunys, 2006).

4 lentelė

## Sprogstamųjų priebalsių atpažinimo tikslumas.

	B	D	G	K	P	T
Fonemų skaičius	295	890	491	744	1447	2605
<b>Atpažinimas nenaudojant hierarchinės fonemų struktūros</b>						
Teisingai atpažino	27,1%	40,6%	59,9%	73,7%	84,3%	57,3%
Atpažino kaip balsį	6,1%	5,4%	3,9%	0,3%	1,5%	0,7%
Atpažino kaip pusbalsį	9,2%	16,4%	14,9%	2,8%	1,9%	2,7%
Atpažino kaip sprogstamąjį priebalsį	84,7%	78,1%	81,2%	96,5%	96,6%	96%
Atpažino kaip friktyvinį priebalsį	0%	0,1%	0%	0,4%	0%	0,6%
<b>Atpažinimas naudojant hierarchinę fonemų struktūrą</b>						
Teisingai atpažino	38,6%	71,9%	71,3%	77,8%	75,4%	79,9%
Atpažino kaip balsį	5,8%	1,8%	1,2%	0,1%	0,1%	0,1%
Atpažino kaip pusbalsį	18,3%	10,9%	11,8%	1,1%	3,2%	0,7%
Atpažino kaip sprogstamąjį priebalsį	75,9%	87,1%	86%	96,5%	95,4%	97,9%
Atpažino kaip friktyvinį priebalsį	0%	0,2%	1%	2,3%	1,3%	1,3%

Šaltinis: Driaunys K. Lietuvių šnekamosios kalbos segmentavimo ir fonetinio atpažinimo tyrimas naudojant LTDIGITS garsyno įrašus.

Kaip matome *lentelėje 4* sprogstamųjų priebalsių atpažinimo rezultatuose taikant hierarchinę fonemų struktūrą, rezultatai yra geresni ne tik fonemos priskyrimo grupei, bet ir atpažįstant pačią fonemą.

5 lentelė

## Fonemų atpažinimo metodų suvestinė.

Autorius	Garsynas	Metodas	Tikslumas	Analizuojamas garsų skaičius	Požymiai
Ahmed M. Abdelatty Ali ir k.t.	TIMIT	Uždarumos ir artikuliacijos vietos nustatymo	86%	1200	Naudojami uždarumos ir artikuliacijos vietos nustatymo metoduose.
V. Colotte ir Y. Laprie	TIMIT	Spekto variacijos funkcija	86,8%	215	Fonemos energija.
K.Driaunys	LTDIGITS	Hierarchinis fonemų klasifikavimas su fonemų grupėmis	93,9%	6472	Fonemos vidurinioji, dešinės ir kairės pusės dalys (atpažinimas vykdomas pagal fonemos pradžią (dešinėsios pusės dalį)).

Šaltinis: sudaryta autoriaus.

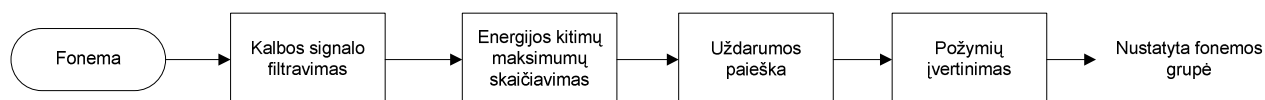
## **1.4 Analizės dalies išvados**

Apibendrinant aukščiau nagrinėtus sprogstamojo signalo atpažinimo metodus galima padaryti išvadas (žr. *Lentelė 5*):

- Apjungus uždarumos ir artikuliacijos vietos nustatymo metodus gautas 86% tikslumas. Eksperimentui atlikti buvo naudojama 1200 garsų iš TIMIT garsyno įrašytų 60 skirtingu diktorių.
- V. Colotte ir Y. Laprie atlikto eksperimento metu buvo pasiektas 86,8% tikslumas sprogstamųjų priebalsių atpažinimui. Šis eksperimentas buvo atliktas spektro variacijos metodu, jo testavimui buvo panaudota TIMIT garsyno duomenys.
- K. Driaunys atliko eksperimentą su LTDIGITS garsyno duomenimis, iš kurio buvo panaudota 6472 sprogstamųjų priebalsių garsų. Vykdamas fonemų atpažinimą buvo naudotas hierarchinis fonemų klasifikavimo metodas, kuris buvo papildytas fonemų grupėmis. Šio modelio atpažįstant sprogstamuosius garsus pasiektas tikslumas yra 93,9%.

## 2 SPROGSTAMŪJŲ PRIEBALSIŲ ATPAŽINIMO METODIKA

Sprogstamųjų priebalsių atpažinimo procesas – tai garso signalo požymių skaičiavimas pagal kuriuos nustatoma fonemos grupė – mūsų atveju sprogstamųjų arba nesprogstamųjų. Fonemos atpažinimo algoritmas naudoja jau segmentuotus garsus, t.y. kur yra pažymėtos žodžių bei fonemų pradžios ir pabaigos diskretos, kurių intervale skaičiuojami požymiai fonemos grupės nustatymui. Bendras algoritmo veikimo principas pavaizduotas *paveiksle 9*.



Šaltinis: sudaryta autoriaus remiantis K. Driaunio sprogstamųjų priebalsių atpažinimo algoritmu.

### 9 pav. Sprogstamųjų priebalsių atpažinimo algoritmo schema.

Kaip matoma algoritmo veikimo schemoje, iš pradžių pateiktas kalbos signalas (šiuo atveju vienos fonemos segmentas) yra filtruojamas, po to yra skaičiuojami energijos kitimai visuose filtrų juostose, vykdoma uždarumos paieška ir atlikus gautų rezultatų įvertinimą fonema pripažįstama kaip sprogstamasis priebalsis arba nesprogstamasis garsas.

#### 2.1 Kalbos signalo požymių skaičiavimas

Sprogstamųjų priebalsių atpažinimo algoritmas išskiria garsų požymius jau iš apdirbto kalbos signalo t.y. algoritmui turi būti pateiktas signalas, kuriame yra sužymėtos žodžių ir fonemų pradžios bei pabaigos ribos. Iš pradžių garso signalas yra diskretizuojamas, siekiant sumažinti kalbos signalo atvaizdavimui reikalingų duomenų kiekį, pagal iš anksto apibrėžtą leistiną maksimalią amplitudės reikšmę. Kalbos signalo amplitudės matavimai atliekami diskretizavimo dažniu (K. Driaunys 2006). Kai turime diskretizuotą garso signalą, tuomet galime atlikti tolesnius žingsnius su juo. Atliekamas pirminis filtravimas, kuriuo siekiama panaikinti diskretizavimo metu atsiradusius iškreipimus (K. Driaunys 2006). Gautasis signalas yra skaidomas į analizės langus, kurių naudojimas palengvina ir pagreitina darbą su analizuojamais duomenimis. Taikant įvairius metodus langų ilgis gali svyruoti nuo 10 ms iki 30 ms, o lango žingsnis nuo 5 ms iki 15 ms (lango žingsnis – tai lango postūmis, kuris skirtas efektyviau panaudoti dviejų gretimų langų informaciją) (K. Driaunys 2006). Siekiant gauti kuo tikslesnę informaciją iš analizės langų buvo pasirinktas lango ilgis 10 ms su 5 ms persidengimu. Kalbos signalo skaidymui į analizės langus naudojamas Hammingo (Hamming) langas. Toliau taip apdorotas signalas yra filtruojamas. Remiantis prielaida, jog sprogimo akustinis įvykis turėtų atsispindėti kaip energijos kitimo lokalus maksimumas įvairiose dažnių juostose, buvo realizuota 14 juostinių filtrų sistema. Filtru juostos realizuotos su 500 Hz pralaidumu, o bendras kalbos signalo spektro dengimas nuo 500 iki 7500 Hz. Filtru realizavimui panaudotas antros eilės Batervorto filtro prototipas (K. Driaunys 2006):

$$K(f) = \frac{1}{\sqrt{1 + \left(\frac{f}{f_0}\right)^{2L}}} \quad (1)$$

kur  $f_0$  - pjūvio dažnis,  $L$  - filtro eilė.

Atlikus filtravimą kiekvienam langui buvo skaičiuojamos energijos ir delta koeficientai.

Energijos skaičiavimas:

$$E = \log \sum_{i=1}^N y_i^2 \quad (2)$$

Delta koeficientų skaičiavimas:

$$\Delta_k(l) = c_k(l) - c_{k-1}(l) \quad (3)$$

kur  $k$  - lango eilės numeris.

Gauti koeficientai toliau naudojami kalbos signalo požymių sudarymui, pagal kuriuos galiausiai fonema priskiriama sprogstamųjų arba nesprogstamųjų grupei.

### 2.1.1 Energijos maksimumų kitimų ir uždarumos skaičiavimas

Atlikus tyrimą nustatyta, jog potenciali sprogimo vieta fiksuojama tada, kai daugiau nei pusės filtrų maksimumo pozicija sutampa. Sprogimo vertinimui yra įvedamas santykio vidurkis (SVMv) tarp visų filtrų ir nagrinėjamo filtro juostos energijos kitimo maksimumo santykio su sprogstamojo priebalsio kitimo maksimumu (K. Driaunys 2006). Maksimumo vietos nustatymui skaičiuojama moda, kuri įvertindama energijos kitimus visuose filtruose suranda kuriame analizės lange yra energijos maksimumas, bei suskaičiuoja patikimumą – kuris nusako tikimybę, kad būtent tame lange yra energijos maksimumas. Energijos kitimų maksimumų skaičiavimas pateiktas sekančiuose žingsniuose:

**1 žingsnis.** Suskaičiuojama moda (*Spro1*) ir jos patikimumas (*Proc*). Moda skaičiuojama, įvertinant energijos maksimumų kitimą kiekviename filtrų juostoje (K. Driaunys 2006).

**2 žingsnis.** Skaičiuojamas kiekvieno filtro energijos deltos vidurkio ir kitimo maksimumo santykis (SVM) taip įvertinama sprogimo kokybė (K. Driaunys 2006):

$$\begin{cases} kit_z = \frac{1}{Spro1} \sum_{i=1}^{Spro1} \Delta Ef_{z;i} \\ SVM_z = \frac{kit_z}{\Delta Ef_{z;Spro1}} \end{cases} \quad (4)$$

kur  $z=1 \dots Z$



**3 žingsnis.** Skaičiuojamas gautojo santykio vidurkis, kuriam esant didesniai už vieną, reiškia, kad kitimas fiksuojamas ganėtinai triukšmingoje aplinkoje, kas nebūdinga sprogstamajam priebalsiui (K. Driaunys 2006):

$$SVM_{vid} = \frac{1}{Z} \sum_{i=1}^Z SVM_i \quad (5)$$

Tačiau vien energijos maksimumų kitimų įvertinimas nėra pakankamas sprogstamųjų priebalsių atpažinimui, siekiant didelio tikslumo, todėl įvesta uždarumos paieška. Uždaruma – tai energijos nebuvimas pasireiškiantis įvairiose filtrų juostose. Uždarumai nustatyti buvo naudojami filtruoti signalai atmetant pagrindinio tono diapazoną (iki 500 Hz), taip atskiriant triukšmą nuo uždarumos (K. Driaunys 2006). Taip pat buvo įvestas uždarumos slenkstis, kuris buvo lyginamas su fonemos fragmentais, tikrinant ar jie žemesni už slenkstinę uždarumos reikšmę. Kiekvienam fonemos langui paskaičiuojama jo energija ir lyginama su slenkstine uždarumos reikšme (K. Driaunys 2006):

$$E_k = \log\left(\frac{1}{n} \sum_{i=1}^n y_i\right) \quad (6)$$

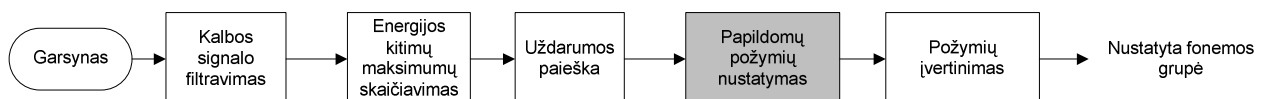
kur  $k=1 \dots \text{fonemos\_langu\_skaičius}$

Jei fonemoje randama langų, kurių energija mažesnė už slenkstinę uždarumos reikšmę, tuomet skaičiuojama kiek tokių langų fonemoje yra.

Pagal gautus požymius fonema yra priskiriama sprogstamųjų arba nesprogstamųjų grupėms.

## 2.2 Sprogstamųjų priebalsių atpažinimo siūlomi patobulinimai

Taikant 2.1 skyriuje aprašytą algoritmą gaunami geri sprogstamųjų priebalsių atpažinimo rezultatai, tačiau slenkstinės reikšmės netikslus parinkimas lemia nesprogstamųjų fonemų priskyrimą sprogstamųjų grupei, o sprogstamųjų – nesprogstamųjų grupei. Siūlomas papildomų požymių skaičiavimas pavaizduotas *paveiksle 10*.



Šaltinis: sudaryta autoriaus remiantis K. Driaunio sprogstamųjų priebalsių atpažinimo algoritmu.

### 10 pav. Siūlomas sprogstamųjų priebalsių atpažinimo algoritmo patobulinimas.

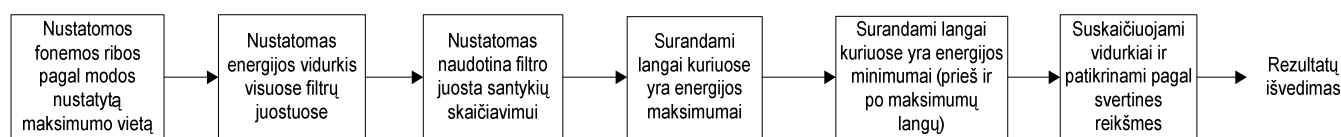
Remiantis Ahmed M. Abdelatty Ali ir k.t. (Ahmed M. Abdelatty Ali ir k.t., 2001) sudarytu metodu, kuriuo matuojant pradžios uždarumos ir aspiracijos trukmes, sprogstamieji priebalsiai skirstomi į skardžiuosius ir dusliuosius, pastebėta, jog sprogstamosiose fonemose išsiskiria uždaruma

ir aspiracija kaip energijos minimumas arba net jos nebuvimas, o sproginimas kaip didelis ir staigus pokytis po minimumų. Šis energijos šuolis keliasdešimt ar net kelis šimtus kartų turi daugiau energijos nei prieš ar po esančiuose atkarpose. Remiantis šiais pastebėjimais algoritmas tobulinamas sekančiais žingsniais (*paveiksle 11*):

- nustatomos fonemos ribos (jei reikia ribos padidinamos tiek į kairę, tiek į dešinę pusę);
- nustatoma filtro juosta, kurioje sproginimas yra ryškiausias;
- surandami maksimumų ir minimumų analizės langai;
- suskaičiuojami santykiai tarp minimumų prieš maksimumus ir maksimumų bei minimumų po maksimumų ir maksimumų santykiai;
- gauti santykiai įvertinami;
- rezultatų apjungimas.

## 2.2.1 Energijos minimumų ir maksimumų santykių skaičiavimais paremtas algoritmas

Algoritmo veikimo schema pateikta *paveiksle 11*. Šioje schemoje pateiktas bendras algoritmo veikimo principas.



Šaltinis: sudaryta autoriaus.

### 11 pav. Minimumų ir maksimumų santykių skaičiavimo algoritmas.

Papildomų požymių skaičiavimai atliekami nuosekliai, kaip parodyta *paveiksle 11*. Šie skaičiavimai atliekami tuo atveju, jei fonemos pauzės energija yra mažesnė nei 0,002. Žemiau pateiktoje lentelėje matyti fonemų pauzių energijų vidurkiai (vidurkiai suskaičiuoti atmetant 15% mažiausių ir didžiausių reikšmių). Kaip galima pastebėti, sprogstamųjų priebalsių vidurkiai gerokai mažesni nei kitų fonemų. Išanalizavus duomenis nustatyta, kad yra 85 sprogstamieji priebalsiai, kurių pauzės energija viršija 0,002. Taigi, siekiant išvengti kuo daugiau klaidingų nustatymų 85 sprogstamosios fonemos nepateks į minimumų ir maksimumų santykių skaičiavimus.

6 lentelė

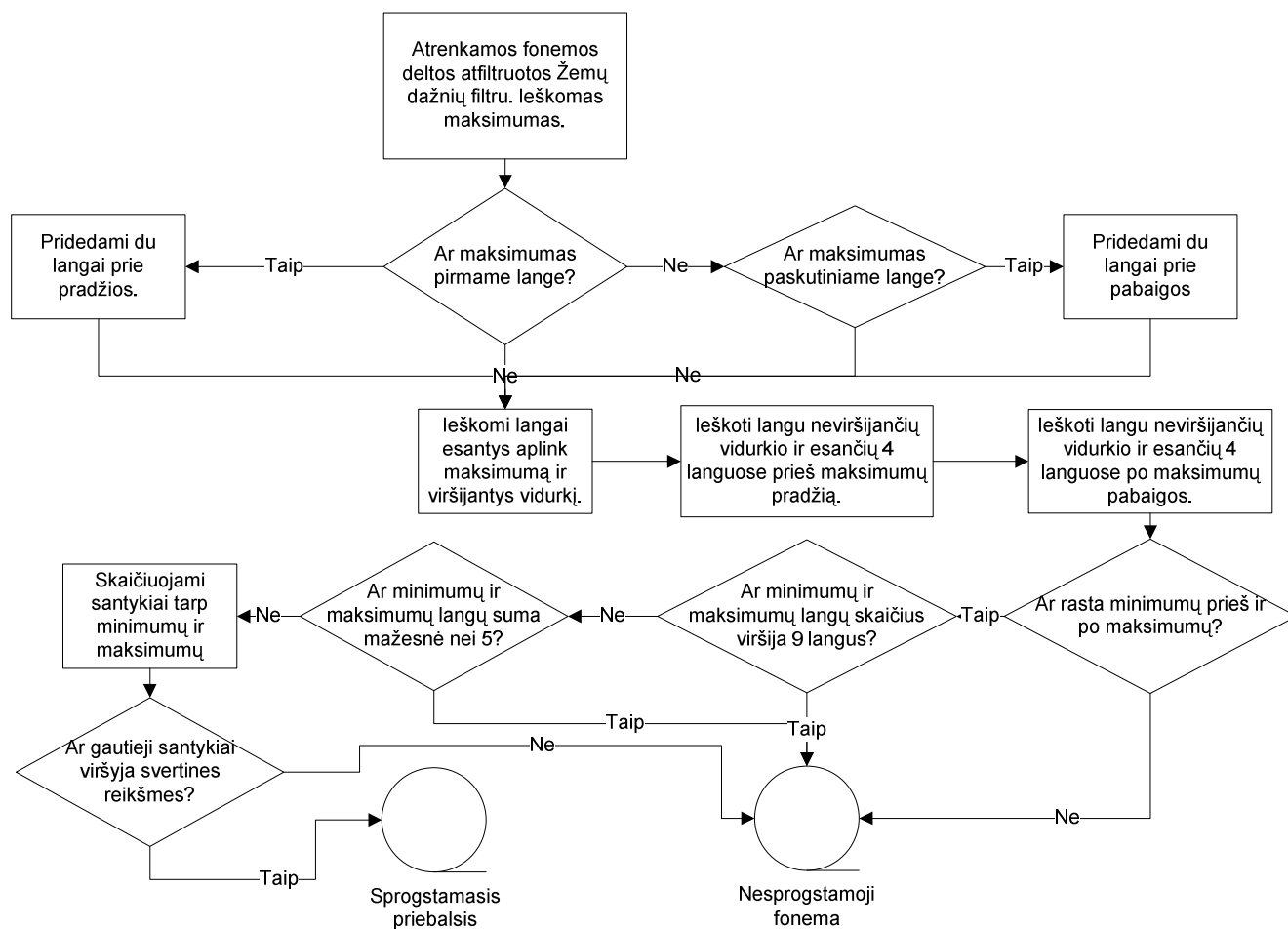
#### Fonemų energijos pauzių vidurkiai.

Sprogstamieji priebalsiai													
b	d	G	k	p	t	Viso							
0,000447	0,000338	0,00032	0,000119	0,000804	0,000083	0,0000351							
Nesprogstamosios fonemos													
a	e	i	y	l	M	n	o	r	s	u	v	z	Viso
0,004	0,003	0,002	0,003	0,001	0,002	0,001	0,003	0,001	0,0006	0,003	0,0006	0,0008	0,00214

Šaltinis: sudaryta autoriaus.

*Paveiksle 12* pateikta fonemos energijos minimumų ir maksimumų skaičiavimo principinė schema. Kaip matoma *paveiksle* iš pradžių išskiriami minimumų ir maksimumų langai, tačiau jei

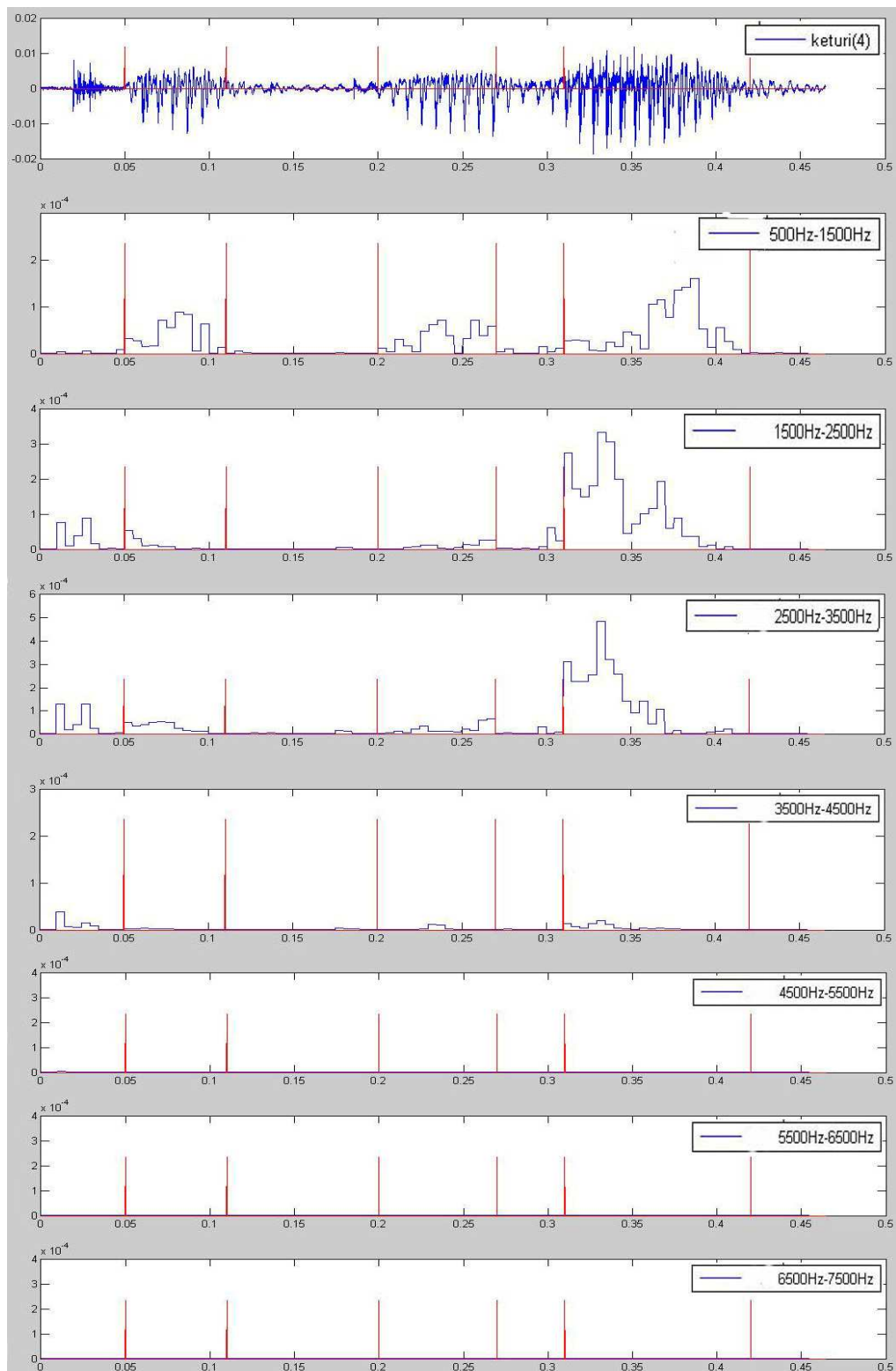
analizės langų gaunama per mažai ar per daug tuomet algoritmas baigia darbą nesuskaičiavęs ir neįvertinęs santykių.



Šaltinis: sudaryta autoriaus.

## 12 pav. Energijos minimumų ir maksimumų santykių skaičiavimo principinė schema.

Iš pradžių patikslinamos fonemos pradžios ir pabaigos ribos siekiant kuo tiksliau nustatyti minimumų ir maksimumų pradžias ir pabaigas (žr Priedas 1). Atlikus eksperimentinių duomenų analizę pastebėta, jog dažnai fonemų pradžios ir pabaigos pažymėtos neviseškai korektiškai, pvz.: pabaigos požymis yra padėtas arčiau pradžios, todėl galima nepastebėti energijos sumažėjimo po sprogo arba pradžia pažymėta toliau nei turėtų būti, todėl fonema prasideda sprogu. Fonemų ribos yra padidinamos jei fonemą sudaro mažiau nei 7 langai, tai gali reikšti, jog segmentavimo metu fonema buvo sutrumpinta. Didinant fonemos analizės langų skaičių yra atsižvelgiama ar fonema yra žodžio pradžioje ar pabaigoje, tokiais atvejais į vieną pusę didinimas neįmanomas. Ribos praplečiamos iš gretimos fonemos pradžios ar pabaigos pridedant du analizės langus. Šiame žingsnyje remiamasi moda, kuri nurodo kuriame analizės lange, visuose 14 filtrų juostose, energijos maksimumas pasitaikė dažniausiai (Driaunys K., 2006). Jei moda nurodo pirmą analizės langą tuomet pradžia dar padidinama dviem langais, jei – paskutinį tuomet pabaiga padidinama dviem analizės langais.



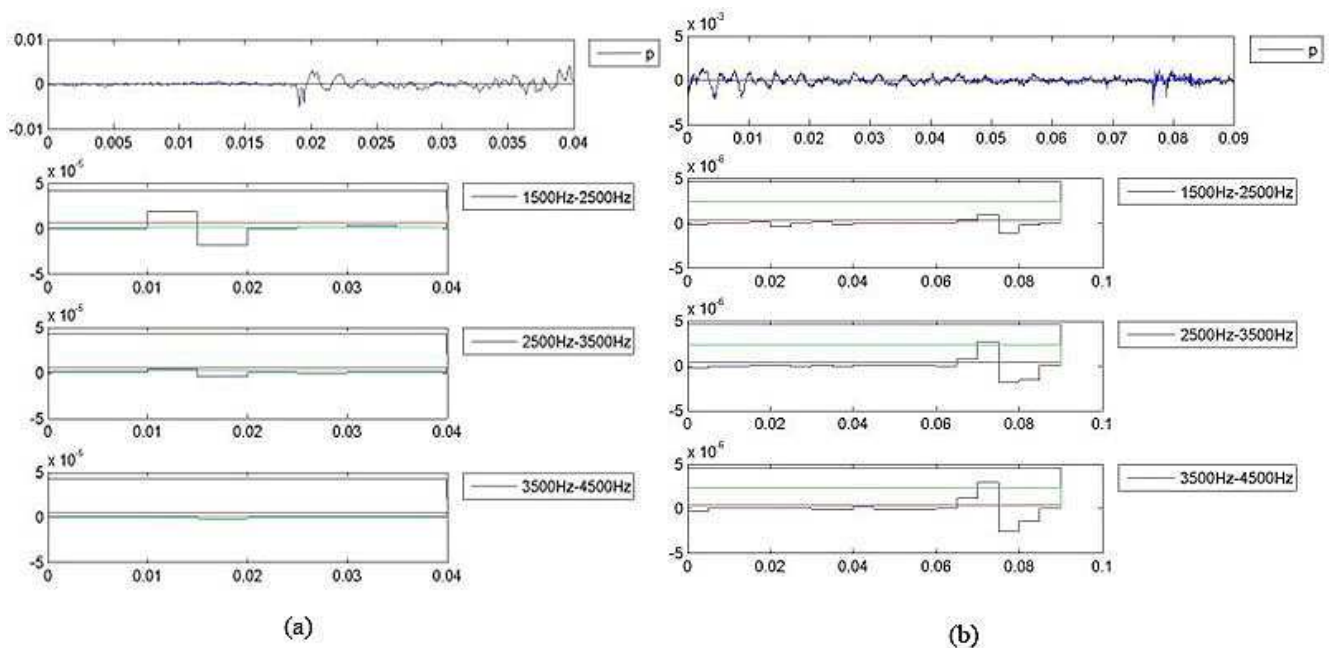
Šaltinis: sudaryta autoriaus.

### 13 pav. Žodžio „keturi“ vaizdavimas visuose filtrų juostose.

Sekančiu žingsniu yra suskaičiuojamas visų fonemos analizės langų (pagal pakoreguotas ribas prieš tai buvusiame žingsnyje) energijų vidurkis visuose filtrų juostose, kuris bus naudojamas kaip svertinė reikšmė nustatant fonemos minimumus ir maksimumus. Nors energijos minimumai ir maksimumai bus skaičiuojami vienoje filtro juostoje, tačiau vidurkis skaičiuojamas visuose filtrų

juostose siekiant kuo tiksliau išskirti maksimumus ir minimumus. *Paveiksle 13* matyti, jog fonetiniai sproginiai (sprogstamieji priebalsiai „k“ ir „t“) geriausiai išryškėja 2, 3, 4 ir kartais 5 filtrų juostose, todėl filtras bus parenkamas būtent vienas iš šių 4.

Filtro juosta, kurioje bus ieškomi minimumai ir maksimumai nustatoma remiantis žodžio ir fonemos pauzių energijomis (žr. Priedas 2). Kiekvienai fonemai yra nustatinėjama filtro juosta, nes skirtingose žodžio vietose ar tariamas skirtingų diktorių akustinis sproginas išryškėja skirtingose filtrų juostuose. Kaip matoma *paveiksle 14* sprogstamojo priebalsio „p“ sproginas a) atveju ryškiausias - antroje



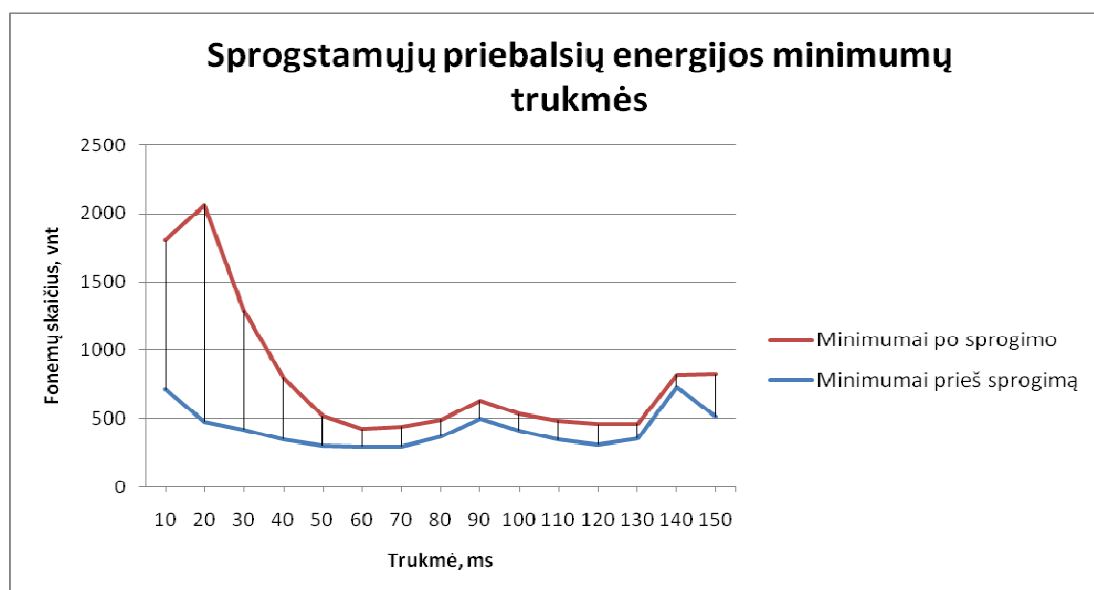
Šaltinis: sudaryta autoriaus.

#### 14 pav. Fonemos „p“ palyginimas skirtingose vietose.

filtrų juostoje (1500Hz – 2500Hz), o tuo tarpu b) atveju akustinis sproginas stipriausiai jaučiamas 4-oje filtrų juostoje (3500Hz – 4500Hz). Filtrų juosta yra nustatoma remiantis fonemos ir žodžio pauzių energijomis. Bandymų metu pastebėta, jog sproginas geriau išryškėja aukštesniuose filtrų juostose, jei žodžio pauzės energija yra mažesnė, ir atvirkščiai, jei žodžio pauzės energija didesnė tuomet sproginas ryškesnis žemesnėse filtrų juostose. *Paveiksle 16* pavaizduotas fonemos „t“ akustinis sproginas 2, 3 ir 4 filtrų juostose, o raudona linija parodo fonemos analizės langų energijų vidurkio lygį. Iš paveikslo galima spręsti, jog filtro juosta parinkta teisingai, nes jei būtų panaudota 2 filtro juosta tuomet pirmo ir paskutinio maksimumų langų energijos būtų žemesnės nei vidurkis ir jie būtų priskirti prie minimumų langų. Tuo tarpu jei būtų santykių skaičiavimui parinkta 4 filtro juosta tuomet energijos sumažėjimas būtų priskirtas prie maksimumų langų, tačiau sproginas būtų nustatytas pagal gautą santykį tarp minimumų (prieš sproginą) ir maksimumų langų.

Ketvirtoje algoritmo dalyje yra surandami fonemos langai viršijantys fonemos analizės langų energijų vidurkį (žr. Priedas 3). Šie langai yra nustatomi pasitikint moda t.y. yra nustatomi visi prieš ir po modos esantys analizės langai, kurių energijos yra didesnės nei vidurkis. Kaip matoma *paveiksle 16* fonemos pradžioje jokios energijos nėra (uždaruma), o po to įvyksta sproginimas ir galiausiai energija sumažėja iki minimumo, tačiau jos šiek tiek išlieka (aspiracija). Moda nustato, jog daugiausiai energijos visuose filtrų juostose yra antrajame lange, todėl ir maksimumų langai yra ieškomi prieš ir po modos nustatyto lango. Remiantis R. Ambrazevičiaus paskaitų medžiaga, kurioje teigiama, jog sproginimas trunka ~30 – 40 ms (Ambrazevičius R. Sprogstamieji priebalsiai, 2008), todėl patikrinama sproginimo trukmė ir jei identifikuotas sproginimas daugiau nei 5 (pridedamas papildomą langą, nes atliekant bandymus kartais sproginimas buvo identifikuojamas ilgesnis) languose, tuomet fonema nustatoma kaip nesprogstamoji. Tuo pačiu yra nustatoma sproginimo pradžia ir pabaiga.

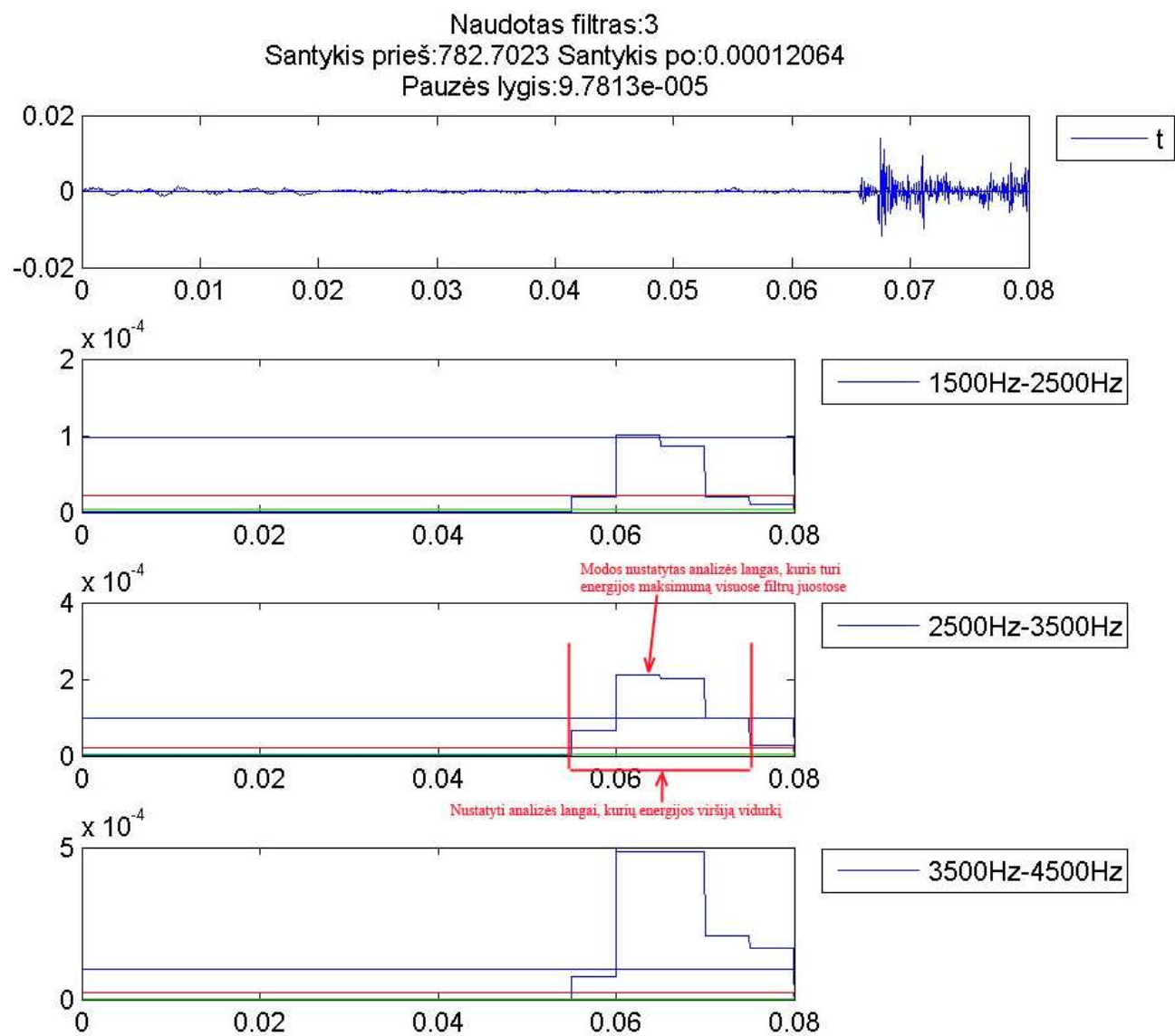
Kai jau turime nustatytus maksimumų langus, penktame algoritmo žingsnyje pradedame analizės langų prieš maksimumus su mažesne energija nei visų langų energijų vidurkis (žr. Priedas 4). Paieška pradedama nuo pirmojo fonemos lango ir vis pereinama prie sekančio. Jei bent vieno lango energija didesnė nei vidurkis, paieška tęsiama nuo sekančio lango iki pirmo nustatyto maksimumų analizės lango. R. Ambrazevičiaus paskaitų medžiagoje teigiama, jog uždaruma prieš sproginimą paprastai trunka ~50 – 150 ms (Ambrazevičius R. Sprogstamieji priebalsiai, 2008), tačiau algoritmas į tai neatsižvelgia. Po sproginimo įvykio paprastai seka energijos sumažėjimas, kurio trukmės vidurkis yra ~31 ms (Ringen C., Helgason P., 2004). Taip pat *paveiksle 15* matoma, jog po sproginimo energijos sumažėjimas dažniausiai jaučiamas nuo 10 ms iki 30 ms, todėl aspiracijos ieškosime 4iuose languose po sproginimo identifikuotos pabaigos.



Šaltinis: sudaryta autoriaus.

**15 pav. Sprogstamųjų priebalsių uždarumų prieš ir po sproginimo trukmės.**

Jei minimumų langų nerandama tuomet nustatoma jog sproginimas neaptiktas ir algoritmas baigia darbą.



Šaltinis: sudaryta autoriaus.

**16 pav. Akustinio sproginimo fonemoje „t“ (žodyje „septyni“) nustatymas. Nustatymui panaudota 3 filtro juosta (2500-3500Hz).**

Sekančiu žingsniu yra suskaičiuojami santykiai tarp minimumų ir maksimumų. Santykių skaičiavimas parodytas 7-oje ir 8-oje formulėse. 7-oje formulėje suskaičiuojami santykiai tarp maksimumų ir minimumų prieš maksimumus langų, o 8-ąją formule skaičiuojami santykiai tarp maksimumų ir minimumų po sprogimo. Formulės iš esmės yra vienodos, tik 7-ojoje – naudojami minimumų langai prieš maksimumus, o 8-oje – minimumų langai po maksimumų.

$$SAN_{pr} = \frac{\left( \sum_{i=k}^m MAX_i \right) / m}{\left( \sum_{j=l}^n MINPR_j \right) / n} \quad (7)$$

kur  $SAN_{pr}$  – santykis prieš maksimumus,  $m$  – maksimumų langų skaičius,  $k$  – pirmas maksimumų langas,  $n$  – minimumų prieš maksimumus langų skaičius,  $l$  – pirmas minimumų langas,  $MAX$  – maksimumų analizės langas,  $MINPR$  – minimumų prieš maksimumus analizės langas.

$$SAN_{po} = \frac{\left( \sum_{i=k}^m MAX_i \right) / m}{\left( \sum_{j=l}^n MINPO_j \right) / n} \quad (8)$$

kur  $SAN_{po}$  – santykis po maksimumo,  $MINPO$  – minimumų po maksimumų analizės langas.

Toliau gautieji santykiai yra lyginami su svertinėmis reikšmėmis, pagal kurias nustatoma ar santykiai pakankamai dideli, jog tai būtų galima užskaityti kaip akustinį sprogimą (žr. Priedas 6). Svertinės reikšmės yra nustatytos nustačius varijuojančius santykius tarp minimumų ir maksimumų langų. *Lentelėje 7* pateiktas sprogstamųjų priebalsių ir kitų fonemų santykių variavimas, atmetant 15% mažiausių ir didžiausių reikšmių.

7 lentelė

**Minimumų ir maksimumų santykiai**

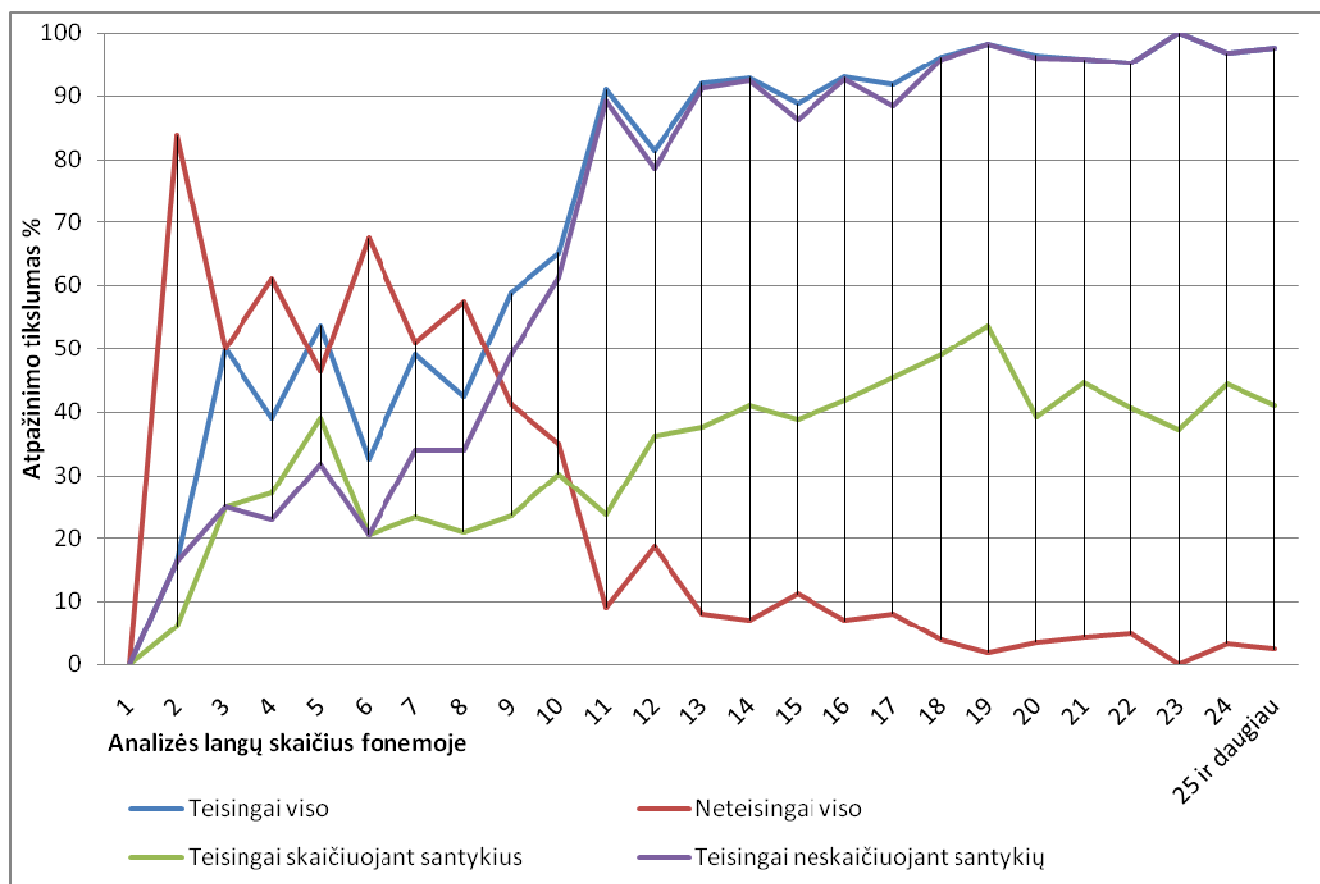
Fonemos	Santykis tarp minimumo prieš maksimumo ir maksimumo (atmetant po 15% mažiausių ir didžiausių reikšmių)		Santykis tarp minimumo po maksimumo ir maksimumo (atmetant po 15% mažiausių ir didžiausių reikšmių)	
	Nuo	Iki	Nuo	Iki
Sprogstamieji priebalsiai	0.00012	277.60712	0.00001	28.30124
Kitos fonemos	0.00001	20.48267	1.80263	15.80744

Šaltinis: sudaryta autoriaus.

Šiuo žingsniu minimumų ir maksimumų santykių skaičiavimas baigiamas, toliau gautas rezultatas apjungiamas su rezultatu gautu skaičiuojant maksimumų kitimus bei ieškant uždarumos (žr. Priedas 7). Apjungiant gautus rezultatus su energijos maksimumų kitimų ir uždarumos skaičiavimų rezultatais, didžiausias dėmesys kreipiamas į tai, jog naudojant maksimumų ir minimumų skaičiavimus



galima geriau nustatyti sprogimą trumpesnėse fonemose, todėl jei fonemą sudaro daugiau nei 20 analizės langų pasikliaujama energijos maksimumų kitimų ir uždarumos skaičiavimų rezultatais. *Paveiksle 17* yra pavaizduoti akustinio sprogo nustatymo rezultatai (procentaliai).



Šaltinis: sudaryta autoriaus.

### 17 pav. Sprogimo nustatymo rezultatai.

Žalia linija rodo sprogo nustatymo rezultatą skaičiuojant tik minimumų ir maksimumų langų santykius; violetinė – sprogo nustatymo rezultatą skaičiuojant tik energijos maksimumų kitimus ir uždarumas; mėlyna – teisingai atpažintų sprogo bendras rezultatas; raudona – neteisingai atpažintus sprogo. Paveiksle galime pastebėti, jog minimumų ir maksimumų santykių skaičiavimai geriau nustato sprogimą trumpesnėse fonemose nei ilgesnėse.

### ***2.3 Sprogstamųjų priebalsių atpažinimo metodikos išvados***

- Sprogstamųjų priebalsių atpažinimo algoritmas bus sudarytas remiantis energijos maksimumų kitimų skaičiavimu ir uždarumos nustatymu.
- Algoritmas patobulintas įvedant fonemos energijos minimumų ir maksimumų santykių skaičiavimus.
- Algoritmo patobulinimai paremti energijos sumažėjimo prieš sproginimą, sproginimo ir energijos sumažėjimo po sproginimo langų paieška skirtingose filtrų juostose, kurios parenkamos atsižvelgiant į žodžio ir fonemos pauzių energijas.
- Minimumų ir maksimumų santykių skaičiavimai geriau nustato akustinį sproginimą trumpose fonemose nei ilgose, todėl sproginimo nustatymui ilgose fonemose naudojamas tik energijos maksimumų kitimų skaičiavimų ir uždarumos paieškos metodas.

### 3 EKSPERIMENTINIS SKYRIUS

Šioje dalyje yra aprašomi duomenys, kurie buvo naudojami eksperimentams atlikti, algoritmo realizacija bei pateikiami eksperimentų metu gauti rezultatai. Algoritmo kurimo metu buvo atliekami eksperimentai, o gaunami rezultatai buvo analizuojami ir atžvelgiant į juos algoritmas buvo papildomas vienomis ar kitomis funkcijomis. Šioje dalyje pateikiami rezultatai iliustruojantys pradinio bei galutinio algoritmo veikimą.

#### 3.1 Duomenys

Kalbos signalų tyrimams reikalingi duomenys ir dažnai didelis jų kiekis. Kuo duomenys įvairesni, tuo tyrimai labiau atitinka realybę. Atliekant eksperimentus su garsais, geriausiai, kai jie yra tariami skirtingų diktorių iš kuo įvairesniu vietovių, nes skiriasi tarmės ir pan. Tuo tikslu yra kuriami garsynai – surinktų kalbos signalų duomenų bazės. Garsynai yra sudaromi priklausomai nuo uždavinių, kuriems spręsti jie renkami. (K. Driaunys 2006).

Pasaulyje sudarytų garsynų yra labai daug (pvz.: TIMIT, BUPT, SPEECHDAT ir pan.), kadangi tyrėjai dažnai susidaro savo garsynus skirtus jų iškeltiems uždaviniams spręsti. Taip pat yra organizacijos koordinuojančios garsynų sudarymo klausimus.

Nuo 1999 metų pradėta sudarinėti pirmoji sisteminė Lietuviška kalbos signalų duomenų bazė – LTDIGITS. Kuriant garsyną, buvo numatyti uždaviniai, kuriuos jis turėjo padėti spręsti (Rudžionis 1999):

1. Kurti daugiadiktoreis lietuviškų skaičių pavadinimų sekų atpažinimo sistemas;
2. Kurti daugiadiktoreis kai kurių lietuviškų kompiuterio valdymo komandų atpažinimo sistemas;
3. Spręsti atskirų fonemų diskriminavimo uždavinius, kurie aktualūs ne vien tik lietuvių kalbos požiūriu.

Taigi, atsižvelgiant į iškeltus uždavinius, kuriems spręsti skirtas LTDIGITS garsynas, jis yra sudarytas iš lietuviškų skaitmenų ir lietuviškų valdymo komandų pavadinimų pvz.: garsyną sudaro skaitmenų frazės: vienas, trys, nulis, keturi, devyni ir t.t. ir valdymo komandos: pradėti, baigti, sustoti, laukti, tęsti, saugoti, taip, ne ir pan. Garsyne jau yra sukaupta virš 350 įrašų, kurie priklauso skirtingiems diktoriais: 220 – moterų ir 130 – vyrų

LTDIGITS garsynas nėra idealiai sudarytas ir viena iš pagrindinių to priežasčių yra diktorių neprofesionalumas. Diktoriais buvo pasirinkti neprofesionalai, todėl skaitydami tekstą darė tarties klaidas, pasimesdavo ir pan. (K. Driaunys 2006).

Ekperimentui buvo panaudota 100 diktorių (50 vyrų ir 50 moterų) įrašų. Buvo naudojamos aštuonios kiekvieno diktoriaus frazės, kuriose yra ištarti skaitmenų pavadinimai ir valdymo komandos. Visuose eksperimentuose yra išanalizuota daugiau nei 25000 fonemų, iš kurių 6454 sprogstamasis priebalsis.

### ***3.2 Eksperimentinės aplinkos formavimas***

Fonetinio sprogimo paieškai buvo panaudoti energijos pokyčio ir uždarumos paieškos bei energijos minimumų ir maksimumų santykių skaičiavimai. Algoritmas sudarytas remiantis teorine medžiaga bei bandymu metu gautais statistiniais duomenimis. Eksperimentams buvo naudojami specialiai paruošti duomenys t.y. buvo nurodytos žodžių bei fonemų pradžios ir pabaigos ribos. Kiekviena frazė buvo atskiruose garso failuose (\*.wav formate). Algoritmas buvo paruoštas taip, kad pagal nurodytas žodžio bei to žodžio fonemų pradžios bei pabaigos ribas galėtų iš garso failų nuskaityti žodžius bei fonemas, kurių požymiai po to buvo analizuojami.

### ***3.3 Algoritmo realizacija***

Ekperimentams atlikti buvo sukurtas įrankis, kurio dėka galima vykdyti sprogstamųjų priebalsių atpažinimą tiek naudojant energijos minimumų ir maksimumų skaičiavimus tiek be jų, analizuoti gautus rezultatus, keisti eksperimentų sąlygas.

Įrankis buvo sukurtas panaudojant „The MathWorks“ produktą MATLAB R2008a, kurio dėka atliekami reikalingų požymių skaičiavimai. Šis produktas pasirinktas dėl didelio operacijų atlikimo greičio, daugybės funkcijų, kurių nereikia papildomai kurti, ir galimybėmis analizuoti gautus rezultatus įvairiais aspektais.

### 3.4 Rezultatai

Fonetinio sproginimo nustatymo eksperimentai buvo atliekami algoritmo kūrimo eigoje. Kai algoritmas galėjo atlikti energijos maksimumų kitimų skaičiavimus ir uždarumos paieška buvo atlikti bandymai, kurie buvo panaudoti kaip atsvaros taškas tolimesniems bandymams. Sekantys algoritmo testavimai papildžius papildomų požymių skaičiavimais, tačiau nepilnai apjungus visų skaičiavimų gaunamus rezultatus ir tinkamai neišanalizavus bandymų rezultatų, parodė kiek prastesnius rezultatus. Galutinis algoritmo variantas pateisino papildomų požymių išskyrimą ir skaičiavimų rezultatų apjungimą. Algoritmo testavimai buvo atliekami su LTDIGITS garsyno įrašais.

#### 3.4.1 Eksperimentas kai santykiai skaičiuojami vienoje filtro juostoje

Šio eksperimento metu buvo atliekami bandymai naudojant tik vieną filtro juostą santykių skaičiavimams ir gauti įvairiapusiški rezultatai. Skaičiuojant minimumų ir maksimumų santykius naudojant 3 filtro juostą (2500-3500Hz), kaip matoma *lentelėje 8* sprogstamųjų priebalsių buvo atpažinta ~3,95% daugiau t.y. 255 fonemomis daugiau.

8 lentelė

Gautų rezultatų palyginimas skaičiuojant papildomus požymius ir ne.

	Netaikant algoritmo	Taikant algoritmą
Atpažino sprogstamąjį priebalsį teisingai	4927	5182
Atpažino nesprogstamąjį garsą kaip sprogstamąjį priebalsį	963	1439
Sprogstamųjų priebalsių atpažinimo tikslumas, %	76,34	80,29

Šaltinis: sudaryta autoriaus.

Tačiau nors ir pagerėjo akustinio sproginimo nustatymo tikslumas, tačiau nauji skaičiavimai dažnai nustatydavo sproginimą ten kur jo ir nebūdavo, todėl bendras atpažinimo tikslumas suprastėjo nuo 90,23% iki 89,82%. Taigi algoritmas pagerino sprogstamųjų priebalsių atpažinimo rodiklį, tačiau bendrą sprogstamųjų priebalsių išskyrimo iš garsų aibės tikslumą suprastino, nes interpretuodamas nesprogstamuosius garsus kaip sprogstamuosius padidina klaidų skaičių. *Lentelėje 9* pateikti atskirų fonemų atpažinimo metu gauti rezultatai.

9 lentelė

Atskirų fonemų gauti atpažinimo rezultatai.

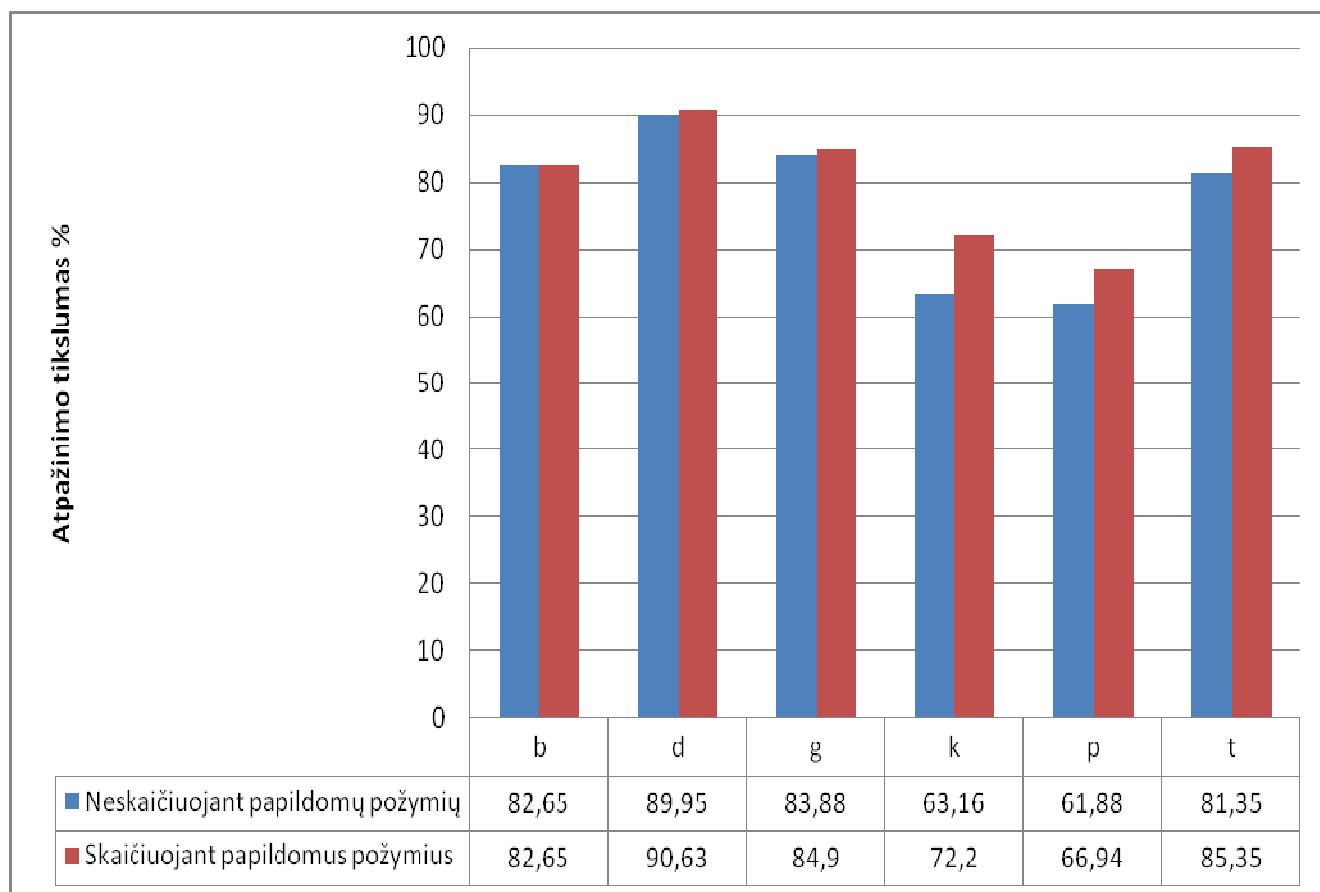
	Netaikant algoritmo						Taikant algoritmą					
	b	d	g	k	p	t	b	d	g	k	p	t
Atpažino kaip sprogstamąjį priebalsį	243	797	411	468	893	2115	243	803	416	535	966	2219
Atpažino kaip nesprogstamąjį garsą	51	89	79	273	550	485	51	83	74	206	477	381

Šaltinis: sudaryta autoriaus.

Gautus rezultatus palyginus su *lentelėje 7* pateiktais duomenimis galima teigti, jog atpažinimo tikslumas pagerintas tiems sprogstamiesiems priebalsiams, kurių energijos maksimumų ir minimumų skirtumai yra didesni. Tik skardžiojo priebalsio *b* atpažinimas liko toks pat, taigi skardžiųjų priebalsių

atpažinimą algoritmas pagerino nestipriai, su tuo susidurta atpažįstant fonemas energijos pokyčio skaičiavimais ir uždarmo paieška, kuomet atpažįstant dusliuosius sprogstamuosius priebalsius buvo pasiekti geresni rezultatai nei skardžiuosius (Driaunys K. (2006)).

Kaip matoma *paveikslėlyje 18* pokytis atpažįstant fonemas nežymus, tačiau sprogstamųjų priebalsių atpažįstama daugiau skaičiuojant papildomus požymius (raudonas stulpelis), nei jų neskaičiuojant (mėlynas stulpelis).



Šaltinis: sudaryta autoriaus.

### 18 pav. Fonemų atpažinimo rezultatai.

Iš *paveiksle 18* matomų atpažinimo rezultatų galime pastebėti, jog minimumų ir maksimumų skaičiavimai fonetinio sprogimo atpažinimą pagerino ~4%. Atpažinta daugiau tų sprogstamųjų priebalsių, kurių nustatymo tikslumas buvo mažesnis, o kitų (šiuo atveju skardžių) praktiškai nepakito.

*Lentelėje 10* pateiktos atpažinimo klaidos aptiktos eksperimento metu. Algoritmas nustatė daugiausiai fonemų kaip sprogstamuosius priebalsius „s“ ir „i“, tačiau daugiausiai neteisingai identifikuotų yra iš pusbalsių klasės („n“, „v“, „r“, „l“, „m“). Taip pat ypač daug duslių sprogstamųjų priebalsių buvo atpažinta neteisingai.

## Fonetinio sprogimo atpažinimo klaidos

Klaidingai identifikuotas fonetinis sproginimas														
s	i	n	v	r	e	a	u	l	s	o	z	y	m	ž
296	281	272	159	102	78	76	62	49	19	18	14	5	5	3
Neidentifikuotas sproginimas														
p		t		k		d		g		b				
477		381		206		83		74		51				

Šaltinis: sudaryta autoriaus.

Šio eksperimento metu rezultatai buvo gauti prastesni nei neskaičiuojant papildomų požymių, tačiau didesnis skaičius atpažintų sprogstamųjų priebalsių rodė tai, jog yra padaryta klaidų apjungiant atskirų skaičiavimų gaunamus rezultatus bei galbūt galima patikrinti energijų minimumų ir maksimumų skaičiavimus ir įvertinimus.

### 3.4.2 Eksperimentas kai santykiai skaičiuojami skirtingose filtrų juostose

Kadangi algoritmas darė pakankamai daug klaidų ir siekiant jas sumažinti buvo pradėta analizuoti kokiais atvejais vienoje ar kitoje filtro juostoje sprogstamojo priebalsio energijos minimumų ir maksimumų langai geriausiai išryškėja, o kitų fonemų išsiskiria mažiausiai. Taip pat pastebėta, jog daug klaidų daroma netinkamai apjungus minimumų ir maksimumų santykių bei energijos maksimumų kitimų ir uždaromos paieškos atpažinimo rezultatus. Algoritmas buvo papildytas filtro juostos pasirinkimu ir rezultatų apjungimu (žr. 2.2.1 skyriuje).

Įvedus filtro juostos pasirinkimą algoritme buvo atpažinta daugiau sprogstamųjų priebalsių nei prieš tai, o pataisytas rezultatų apjungimas sumažino klaidų skaičių. Fonemų skirstymo į sprogstamųjų priebalsių ir nesprogstamųjų fonemų grupes tikslumas pagerėjo iki 90.68 %. Detalesni bandymo rezultatai pateikti *lentelėje 11*.

11 lentelė

## Fonetinio sprogimo atpažinimas santykius skaičiuojant skirtingose filtrų juostose.

	Nekeičiant filtro juostos	Keičiant filtro juostą
<b>Atpažino sprogstamąjį priebalsį teisingai</b>	5182	5148
<b>Atpažino nesprogstamąjį garsą kaip sprogstamąjį priebalsį</b>	1439	1069
<b>Sprogstamųjų priebalsių atpažinimo tikslumas, %</b>	80,29	79,78

Šaltinis: sudaryta autoriaus.

Taigi kelių patobulinimų įvedimas parodė geresnius rezultatus ir pagerino prieš tai buvusio bandymo rezultatus ~1%. Kaip galima matyti *lentelėje 11* sprogstamųjų priebalsių nustatymas, lyginant su prieš tai vykdyto eksperimento rezultatais, suprastėjo ir buvo nustatyta 34 sproginiais mažiau, tačiau algoritmas padarė mažiau neteisingų fonetinių sproginų identifikavimų, todėl tai pakėlė bendrą rezultatą. Žemiau esančioje lentelėje pateiktos atpažinimo padarytos klaidos. Kaip galima pastebėti, sproginimas buvo neteisingai nustatytas daugiausiai pusbalsiuose.

12 lentelė

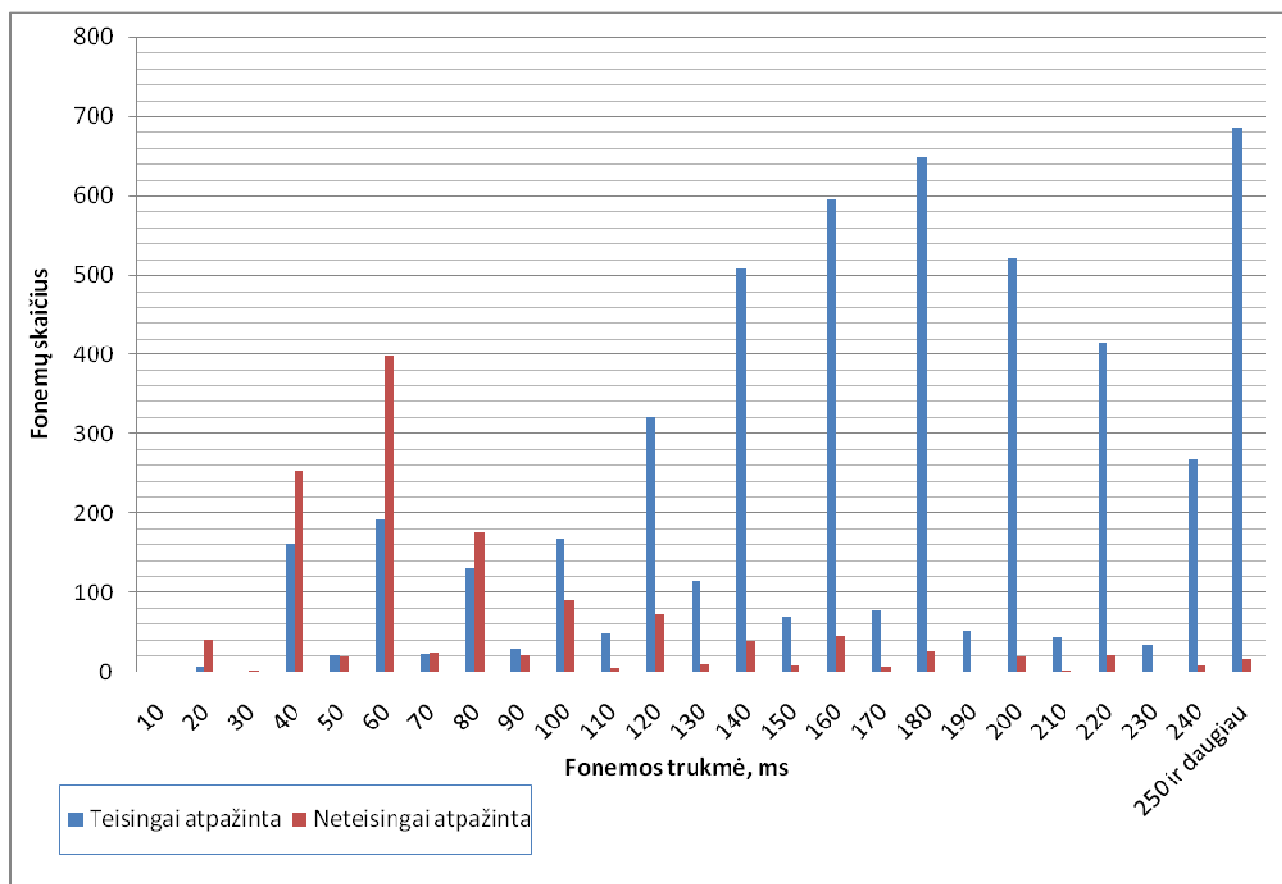
## Padarytos klaidos nustatant fonetinį sproginimą.

	Balsiai	Pusbalsiai	Frikatyviniai	Sprogstamieji
Atpažino kaip sprogstamąjį priebalsį	391	538	140	5148
Atpažino kaip nesprogstamąjį priebalsį	11191	3867	2901	1306
Padaryta klaidų, %	3.38	12.21	4.6	20.24

Šaltinis: sudaryta autoriaus.

Analizuojant klaidas pastebėta, jog sunkiausiai fonetinis sproginimas yra nustatomas trumpose fonemose iki 100 ms. Žemiau pateiktoje diagramoje yra pavaizduota kiek teisingai ir klaidingai nustatyta sprogstamųjų priebalsių priklausomai nuo jų ilgio.

Kaip matyti *paveiksle 18* prasčiausiai atpažįstamos trumpos fonemos, nes trumpas fonemas yra sunku teisingai segmentuoti, dėl segmentavimo metu padarytų klaidų sunku nustatyti sproginimo vietą bei uždarumą. Taigi, sprogstamieji priebalsiai trumpesni nei 100 ms yra silpnoji algoritmo vieta, tačiau trumpas fonemas ypatingai sudėtinga nustatyti, nes gali padidėti klaidų skaičius nustatant atsiktinius energijos šuolius kaip fonetinius sproginimus.



Šaltinis: sudaryta autoriaus.

19 pav. Fonetinio sproginimo atpažinimo daromos klaidos.



Sprogimui identifikuoti trumpose fonemose reikia taikyti kiek kitokias taisykles nei fonemoms ilgesnėms nei 100 ms. Kaip matoma *lentelėje 13* blogiausiai, kaip ir prieš tai atlikto bandymo metu, atpažįstami duslieji sprogstamieji priebalsiai, todėl algoritmo taisykles galima tobulinti nustatant dusliuosius priebalsius.

13 lentelė

**Atpažinimo rezultatai.**

	Testuota fonemų	Teisingai atpažinta	Neatpažinta	Skaičiuojant santykius atpažinta fonemų	Neskaičiuojant santykių atpažinta fonemų
<b>p</b>	1443	970	473	432	893
<b>k</b>	741	526	215	291	468
<b>t</b>	2600	2197	403	1249	2115
<b>b</b>	294	243	51	27	243
<b>d</b>	886	798	88	286	797
<b>g</b>	490	414	76	103	411
<b>Viso:</b>	6454	5148	1306	2388	4927

Šaltinis: sudaryta autoriaus.

Įvedus papildomų požymių skaičiavimus, būtent santykius tarp energijos minimumų ir maksimumų langų buvo aptikta 221 fonetiniu sprogimu daugiau nei vien ieškant uždardomos ir skaičiuojant energijos maksimumų kitimus. Sprogstamųjų priebalsių atpažinimo tikslumas buvo pagerintas nuo 76,34% iki 79,78%, o bendras fonemų skirstymas į sprogstamųjų ir nesprogstamųjų grupes padidėjo nuo 90,23% iki 90,68%. Žemiau pateikiamas padarytų klaidų palyginimas neskaičiuojant papildomų požymių ir juos skaičiuojant.

14 lentelė

**Klaidų palyginimas skaičiuojant ir neskaičiuojant papildomus požymius.**

Nesprogstamieji priebalsiai	Klaidingai identifikuotas fonetinis sprogimas															Viso
	n	i	v	s	r	a	l	e	u	o	z	m	y	š	ž	
Neskaičiuojant papildomų požymių	249	198	148	117	65	54	40	36	24	16	8	4	4	0	0	963
Skaičiuojant papildomus požymius	252	223	152	129	87	62	42	52	34	16	9	5	4	1	1	1069
Sprogstamieji priebalsiai	Neidentifikuotas sprogimas						Viso									
	p	t	k	d	g	b										
Neskaičiuojant papildomų požymių	550	485	273	89	79	51	1527									
Skaičiuojant papildomus požymius	473	403	215	88	76	51	1306									

Šaltinis: sudaryta autoriaus.

### ***3.5 Eksperimentinio skyriaus išvados***

- MATLAB R2008a įrankiu realizuotas sprogstamųjų priebalsių atpažinimo algoritmas, kurio dėka vykdomi energijos maksimumų kitimų skaičiavimai, uždarumos paieška ir energijos minimumų ir maksimumų santykių skaičiavimai.
- Atlikus eksperimentus nustatyta, jog pritaikius energijos minimumų ir maksimumų santykių skaičiavimo algoritmą bendras fonemų atpažinimo tikslumas suprastėjo nuo 90.23% iki 89.82%, tačiau sprogstamųjų priebalsių grupės nustatymas fonemoms pagerėjo nuo 76,34% iki 80,29%, buvo atpažinta 255 sprogstamaisiais priebalsiais daugiau.
- Į algoritmą papildomai įvedus filtro juostos keitimą minimumų ir maksimumų skaičiavimams priklausomai nuo žodžio ir fonemos pauzės energijos atpažinimo tikslumas pagerėjo iki 90,68%.

## IŠVADOS

1. Atliktus sprogstamųjų priebalsių atpažinimo darbus nustatyta, jog geriausius rezultatus parodo hierarchinis fonemų klasifikavimo metodas papildytas fonemų grupėmis. Jis leidžia pasiekti 93,9% atpažinimo tikslumą.
2. Eksperimentiniams tyrimams buvo paruoštas sprogstamųjų priebalsių atpažinimo algoritmas, paremtas uždarumos paieškos ir energijos maksimumų kitimų skaičiavimais. Algoritmas buvo patobulintas įvedus papildomus minimumų ir maksimumų langų santykių skaičiavimus.
3. Pasiūlytasis algoritmas eksperimentų metu pagerino sprogstamųjų priebalsių atpažinimo tikslumą, nuo 76,34% iki 79,78%, tačiau dėl algoritmo daromų klaidų bendras fonemų skirstymo į sprogstamųjų ir nesprogstamųjų grupes pagerėjo tik nuo 90,23% iki 90,68%.
4. Remiantis gautais rezultatais galima teigti, jog papildomai skaičiuojant energijos minimumų ir maksimumų analizės langų santykius ir parinkus teisingas svertines reikšmes galima pagerinti sprogstamųjų priebalsių atpažinimą. Šiame darbe pateiktame eksperimente buvo atpažinta 221 sprogstamuoju priebalsiu daugiau, nei netaikant papildomų skaičiavimų.

## LITERATŪRA

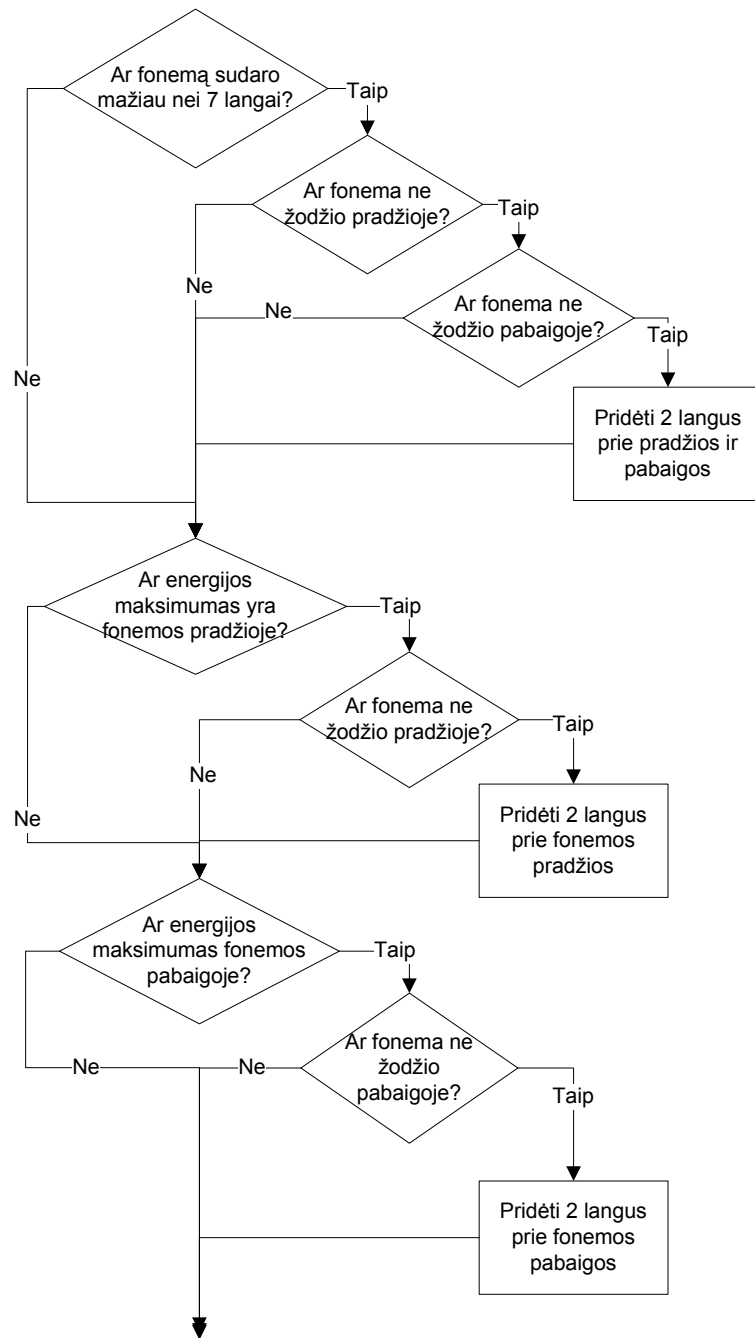
1. **Abdelatty Ali A.M., Van der Spiegel J., Muller P.** (2001) *Acoustic – Phonetic Features for the Automatic Classification of Stop Consonants*. IEEE Transactions on Speech and Audio Processing. Vol. 9. P.833-741.
2. **Ahmed M. Abdelatty Ali** (2000). *Acoustic-Phonetic Feature-Based Signal Processing for Automatic Speech Recognition: Brief Results*. Dept. of Electrical Engineering, University of Pennsylvania.
3. **Aissiou M., Guerti M.** (2007). *Genetic Algorithms Application for the Automatic Recognition of the Arabic Stop Sounds*. Journal of Applied Sciences Research 3(5): 358-366, 2007.
4. **Ambrazevičius R.** *Sprogstamieji priebalsiai*. [interaktyvus]. Paskaitos medžiaga. [žiūrėta 2008 m. sausio 8 d.]. Prieiga per internetą:  
<[http://www.rytisambrazevicius.hmf.ktu.lt/kalbos\\_akustika.html](http://www.rytisambrazevicius.hmf.ktu.lt/kalbos_akustika.html) >
5. **Colotte V., Laprie Y.** (2000) *Automatic enhancement of speech intelligibility*. In IEEE International Conference on Acoustics, Speech, and Signal Processing – ICASSP'2000, Istanbul, Turkey.
6. **Daunys G., Balbonas D.** *Garsų klasifikavimas panaudojant sprendimų medžius*. [interaktyvus]. Informacinės technologijos kalbų inžinerijoje. Šiaulių universitetas. Šiauliai. [žiūrėta 2008 m. sausio 9 d.]. Prieiga per internetą:  
<[http://internet.ktu.lt/lt/mokslas/konf05/konf\\_02/IT2005/Sekc05.pdf](http://internet.ktu.lt/lt/mokslas/konf05/konf_02/IT2005/Sekc05.pdf) >
7. **Don X. Sun, Deng L.** *Statistical Models for Speech Recognition*. [interaktyvus]. Tyrimų rezultatai. [žiūrėta 2008 m. sausio 8 d.]. Prieiga per internetą:  
<[http://cm.bell-labs.com/cm/ms/departments/sia/dxsun/speech/index.html#arti\\_feature](http://cm.bell-labs.com/cm/ms/departments/sia/dxsun/speech/index.html#arti_feature) >
8. **Driaunys K.** *Lietuvių šnekamosios kalbos segmentavimo ir fonetinio atpažinimo tyrimas naudojant LTDIGITS garsyno įrašus*. (2006). Daktaro disertacija. Fiziniai mokslai, Informatika. Vilnius. p. 13-19; 71-73;94-95.

9. **Guoning Hu, DeLianf Wang.** *Separation of Stop Consonants.* (2003) [interaktyvus]. In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03). [žiūrėta 2008 m. sausio 9 d.] Prieiga per internetą:  
<<http://www.cse.ohio-state.edu/~dwang/papers/Hu-Wang.icassp03.pdf>>
10. **Ghavami Modarresi Golnaz** (2002). *The Effects of Syllable Boundary, Stop Consonant Closure Duration, and VOT on VCV Coarticulation.* Doktoro disertacija. Teksaso Universitetas.
11. **Francis L. A., Ciocca V., Jojo Man Ching Yu** (2002) *Accuracy and variability of acoustic measures of voicing onset.* Journal Acoustic Society of America, 113(2), 2003.
12. **James K. Baker, Laurence Gillick.** (1989) *Method for speech recognition.* [interaktyvus]. Jungtinėse Amerikos valstijose išduotas patentas nr. 4805219. [žiūrėta 2008. Sausio 8 d.]. Prieiga per internetą:  
<<http://www.patentstorm.us/patents/4805219.html>>
13. **Jinjin Ye** (2004). *Speech recognition using time domain features from phase space reconstructions.* [interaktyvus]. Magistrinis darbas. Marquette University. Milwaukee, Wisconsin. USA. [žiūrėta 2008 m. sausio 8 d.]. Prieiga per internetą:  
<[http://speechlab.eece.mu.edu/johnson/papers/Ye\\_thesis.pdf](http://speechlab.eece.mu.edu/johnson/papers/Ye_thesis.pdf)>
14. **Kazlauskienė A., Raškinis G.** *Lietuvių kalbos sprogstamųjų priebalsių kiekybė.* (2006). Informacinės technologijos. Kalbų studijos. ISSN 1648-2824. 2006. 8 Nr.
15. **Kiefte M., Kluender K. R.** (2005). *Pattern Playback revisited: unvoiced stop consonant perception.* Journal Acoustical Society of America, 118 (4): p. 2599-2606.
16. *Kompiuterinės lingvistikos centras.*[interaktyvus] VDU (Vytauto Didžiojo Universitetas). [žiūrėta 2008 m. sausio 9 d.]. Prieiga per internetą:  
<<http://donelaitis.vdu.lt/main.php?id=5&nr=2>>
17. **Ladefoged P., Maddieson I.** (1984). *The sounds of the world's languages.* Oxford, OX, UK ; Cambridge, Mass., USA : Blackwell Publishers, 1996.

18. **Minkštimas A.** (2009) *Fonetinio sprogimo kalbos signale paieška ir tyrimas*. 14-osios tarpuniversitetinės magistrantų ir doktorantų mokslinės konferencijos „Informacinės technologijos“. Pranešimų medžiaga. Kaunas, 2009, p. 12-16.
19. *Moksliniai ir taikomieji balso technologijų darbai Lietuvoje*. [interaktyvus]. Mokslinių tiriamųjų darbų medžiaga. Kauno technologijos universitetas. [žiūrėta 2008 m. sausio 9 d.]. Prieiga per internetą:  
<[http://www.likit.lt/all/balso\\_tech/04\\_darbai.htm#\\_Toc533251325](http://www.likit.lt/all/balso_tech/04_darbai.htm#_Toc533251325) >
20. **Mundie C.** (2001). *Speech Transcript*. [interaktyvus]. Prepared Text of Remarks for Commercial Software Model. [žiūrėta 2008 m. sausio 8 d.]. Prieiga per internetą:  
<<http://www.microsoft.com/presspass/exec/craig/05-03sharesource.mspx> >
21. **Purlytė J.** *Kalbos technologijų mokymo Internetė tyrimai*. (2003) [interaktyvus] Magistrinis darbas. [žiūrėta 2008 m. sausio 8 d.]. Prieiga per internetą:  
<<http://kalba.mch.mii.lt/technologijos.htm> >
22. **Ringen C., Helgason P.** (2004) *Distinctive [voice] does not Imply Regressive Assimilation: Evidence from Swedish*. IJES, vol. 4 (2), 2004, p. 53-71.
23. **Truckenbrodt H.** (2005). *The motor theory of perception*. [interaktyvus]. Paskaitų medžiaga. [žiūrėta 2008.06.10]. Prieiga per internetą:  
<[http://www2.sfs.uni-tuebingen.de:9010/~hubert/Phon\\_05/lecture/perc2.pdf](http://www2.sfs.uni-tuebingen.de:9010/~hubert/Phon_05/lecture/perc2.pdf)>
24. **Uemura D., Kitazawa S.** (1992) *Distinction between Vowels and Unvoiced Stops using Features Observed in Speech Waveform*. [interaktyvus]. Garsų studijos. [žiūrėta 2008.06.01]. Prieiga per internetą:  
<[http://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/52467/1/soa026\\_049.pdf](http://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/52467/1/soa026_049.pdf)>
25. **Wen Yun Su V.** (2008). *The Gender Variable in Astralian English Stop Consonant Production*. Daktaro disertacija. Melburno Universitetas.

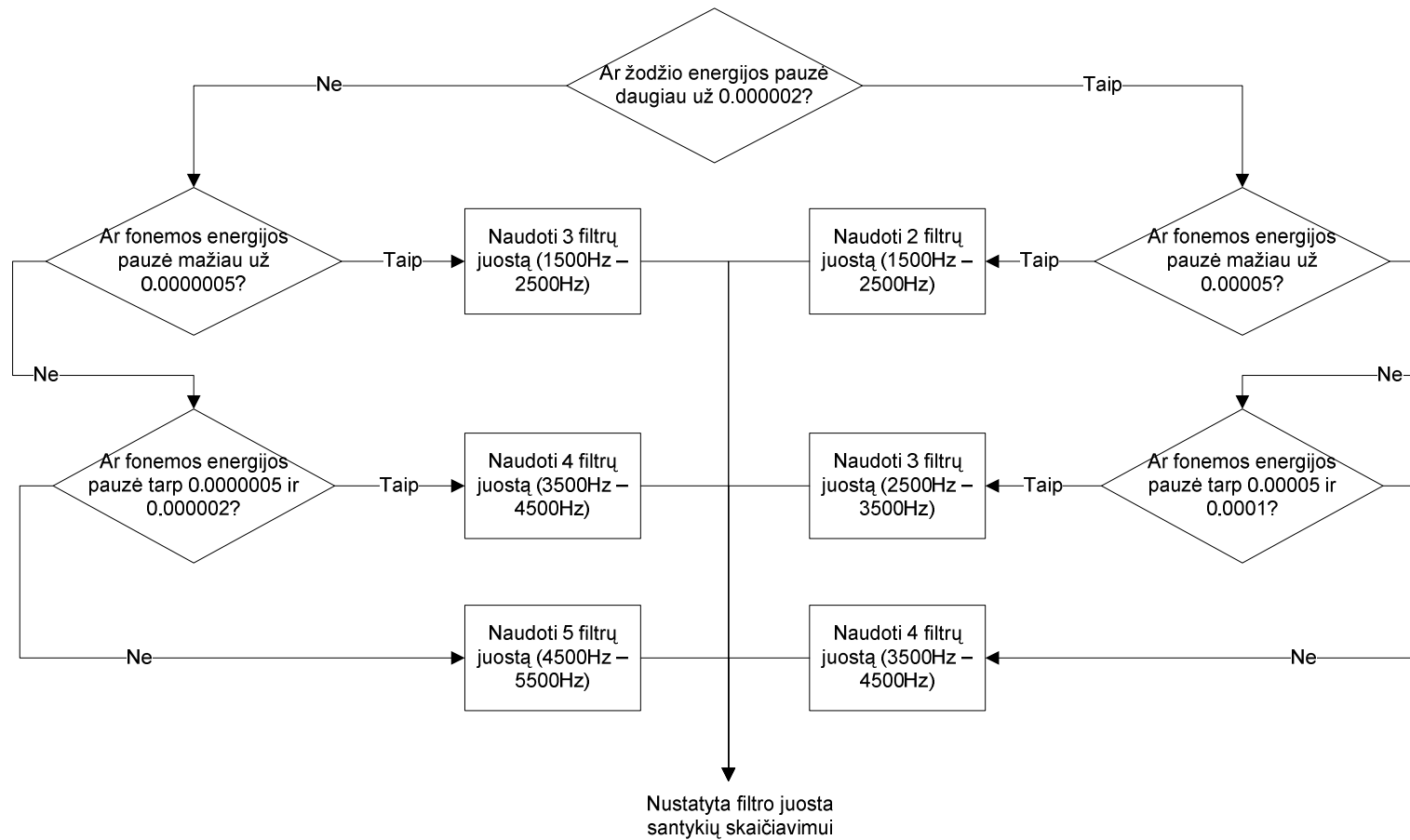
# PRIEDAI

## 1 PRIEDAS. Fonemos pradžios ir pabaigos ribų nustatymas.



Nustatytos fonemos ribos

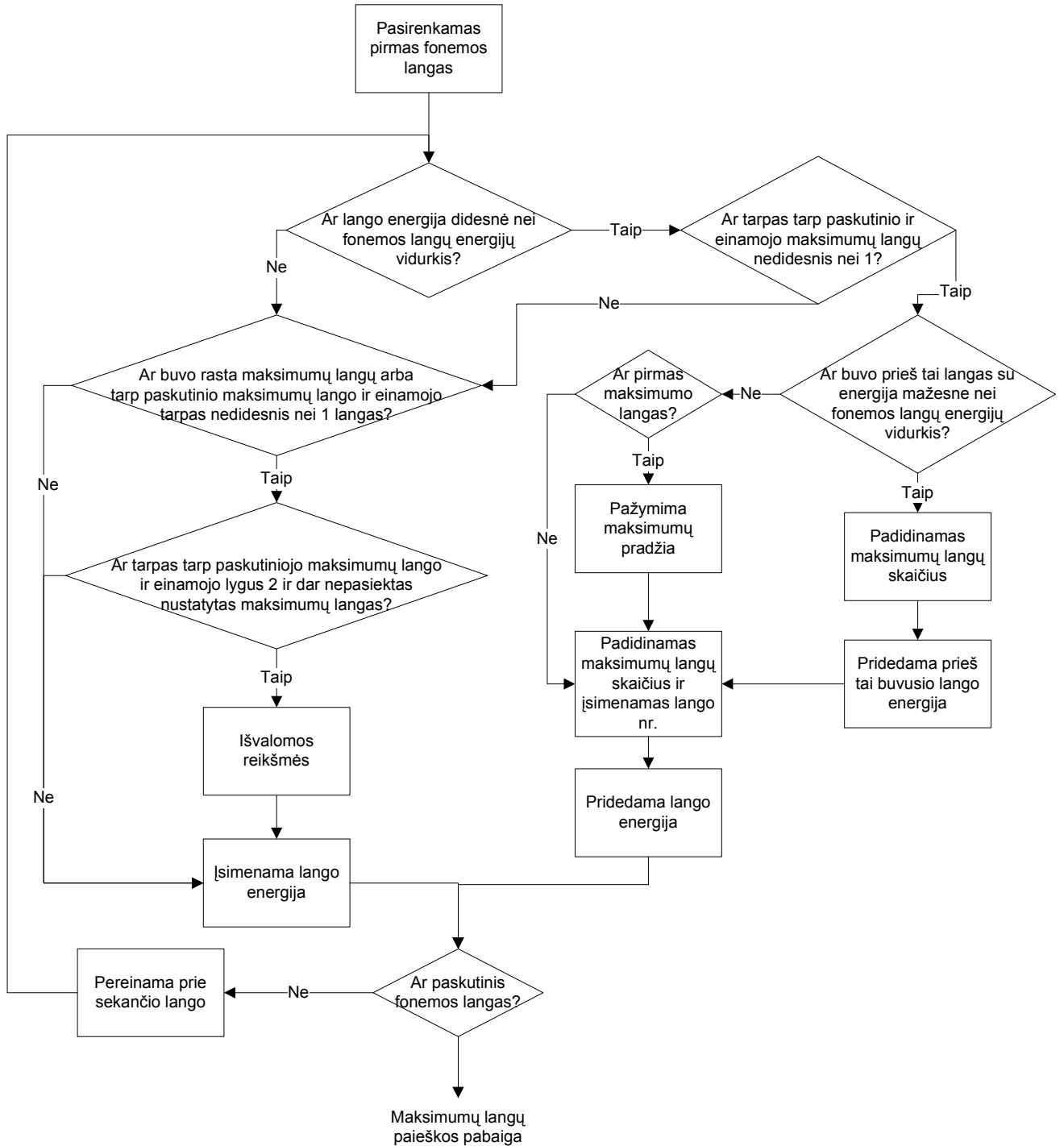
## 2 PRIEDAS. Filtro juostos nustatymas pagal žodžio ir fonemos energijas.



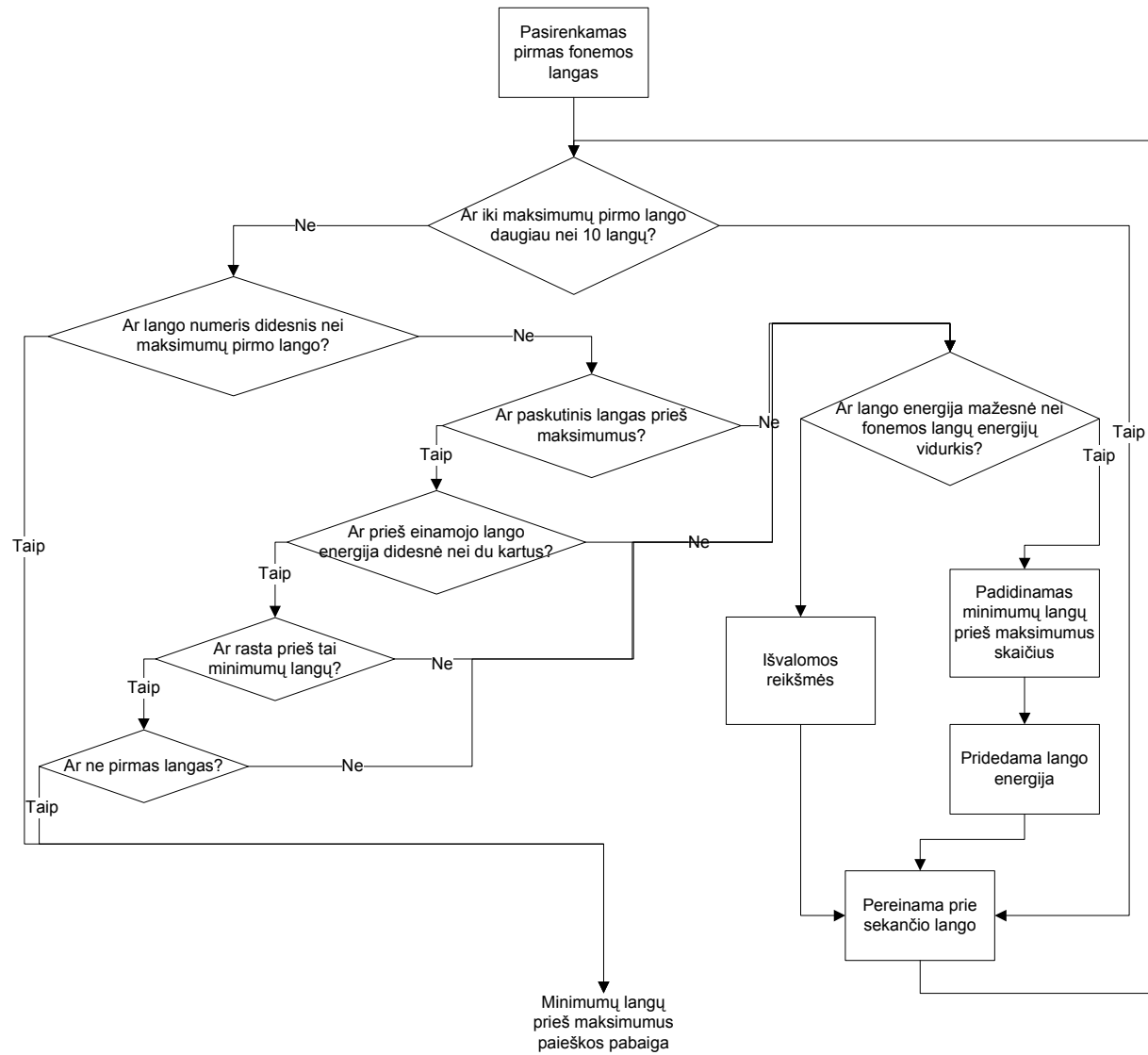


### 3 PRIEDAS.

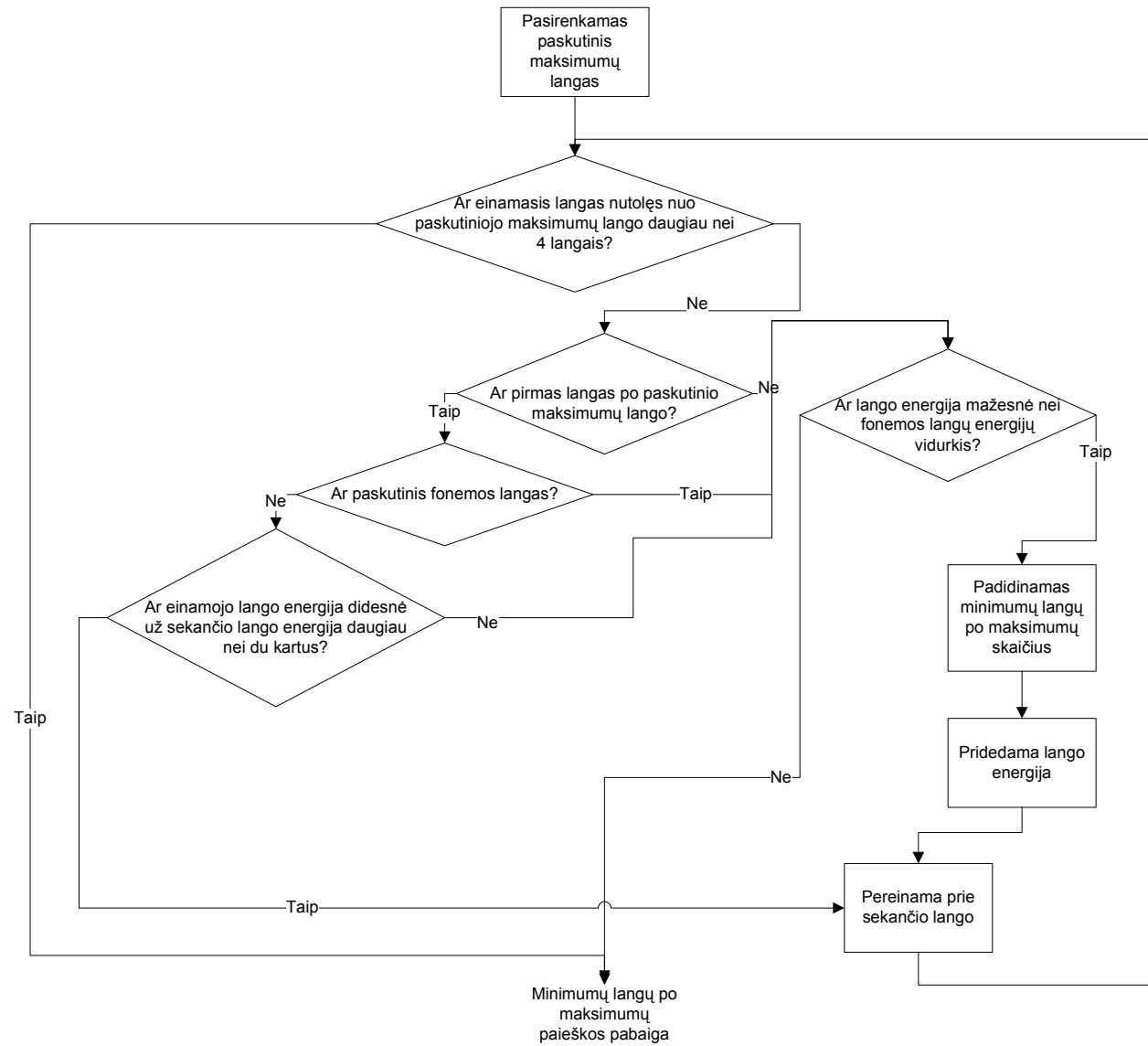
### Energijos maksimumų langų nustatymas.



#### 4 PRIEDAS. Energijos minimumų langų prieš maksimumus nustatymas.

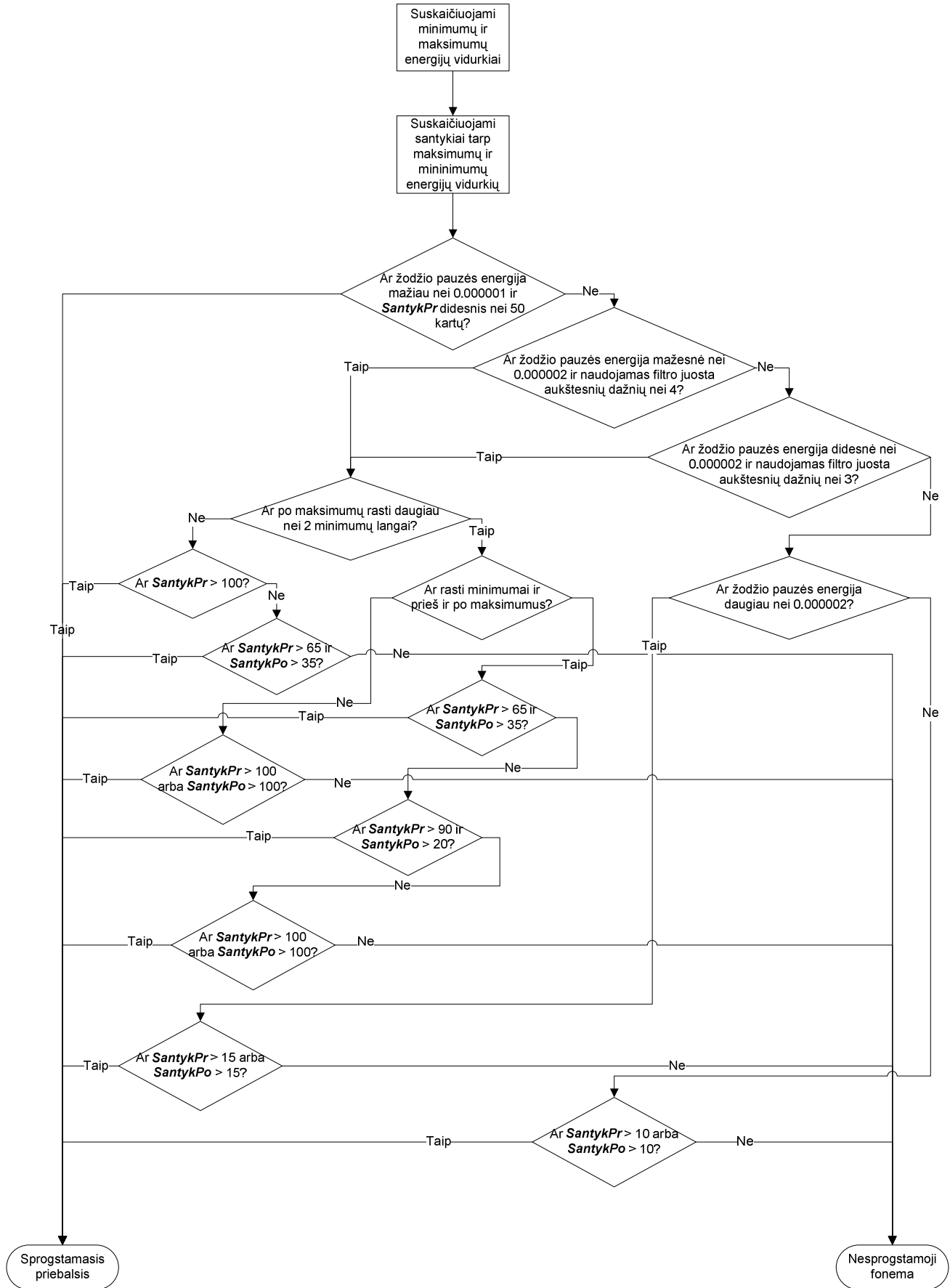


## 5 PRIEDAS. Energijos minimumų po maksimumų nustatymas.



## 6 PRIEDAS.

## Minimumų ir maksimumų santykių įvertinimas.



## 7 PRIEDAS. Fonetinio sproginio nustatymo rezultatų apjungimas.

