

**Lietuvos valstybinis mokslo ir studijų fondas**  
**Vytauto Didžiojo universitetas**

Prašymo registracijos nr.: T-05267

Sutarties su fondu nr.: T-41/05

# **LIETUVIŲ KALBOS TEKSTO KIRČIAVIMAS INTERNETE**

**Baigiamoji ataskaita**

**Projekto vadovas dr. Gailius Raškinis**

**Kaunas**  
**2005**

## **Vykdytojų sąrašas:**

Dr. Gailius Raškinis (projekto vadovas)

Doc. dr. Asta Kazlauskienė

Dokt. Giedrius Norkevičius

Dr. Gailius Raškinis atsakingas už projekto veiklų koordinavimą, projekto įgyvendinimą ir atsiskaitymą Fondui už projekto rezultatus; taip pat už lietuvių kalbos gramatikose teikiamų kirčiavimo taisyklių struktūrinę analizę ir pritaikymą automatinio kirčiavimo metodams, vardažodžių automatinio kirčiavimo metodu konstravimą, kirčiavimo metodų kompiuterinę realizaciją ir baigiamosios ataskaitos Fondui rengimą.

Doc. dr. Asta Kazlauskienė atsakinga už papildomų duomenų ir kalbos resursų, reikalingų automatinio kirčiavimo metodams, surinkimą. Tai duomenys apie bendrinių lietuvių kalbos žodžių, dažnesnių vietovardžių ir asmenvardžių kirčiavimą (antraštinė žodžio forma, kalbos dalis, kirčiuotė, kirčio vieta ir priegaidė), priesaginių daiktavardžių ir būdvardžių kirčiavimo dėsnų kompiuterinį pateikimą. Doc. dr. Asta Kazlauskienė taip pat atsakinga už automatinio kirčiavimo metodų veikimo tikslumo įvertinimą, baigiamosios ataskaitos Fondui rengimą.

Dokt. Giedrius Norkevičius atsakingas už lietuvių kalbos gramatikose teikiamų kirčiavimo taisyklių struktūrinę analizę ir pritaikymą automatinio kirčiavimo metodams, kirčiavimo metodų kompiuterinę realizaciją, kirčiavimo technologijos pateikimą internete.

## Turinys

1	Įvadas	5
2	Analogiški darbai	5
3	Lietuvių kalbos vardažodžių kirčiavimas	6
3.1	Kalbos resursai, reikalingi automatizuotam vardažodžių kirčiavimui	6
3.1.1	Pirminiai duomenys	6
3.1.2	Priesaginių daiktavardžių ir būdvardžių kirčiavimo duomenys	8
3.2	Morfologinės analizės ir skiemonavimo procedūra	9
3.3	Kirčiavimo taisyklių struktūrinė analizė ir pritaikymas automatinio kirčiavimo metodams	10
4	Rezultatai	11
5	Automatinio vardažodžių kirčiavimo pateikimas internete	11
6	Išvados	12
7	Literatūra	13
	1 priedas. Detalus kirčiavimo tikslumo įvertinimas	14
	2 priedas. Parengtos publikacijos	13
	3 priedas. Projekto išlaidų sąmata	24
	4 priedas. Mokslinio – ekonominio produktyvumo suvestinė	25

## Santrauka

Lietuvių kalboje gali būti kirčiuotas ne tik bet kuris žodžio skiemuo, bet žodžio paradigmoje kirtis keičia vietą, kuri dažnai priklauso nuo priegaidės. Nors lietuvių kalbos kirčio ir priegaidės kitimo dėsningumai buvo išsamiai nagrinėti lietuvių kalbos mokslo tyrimo darbuose, tačiau įprastinėje gramatikoje pateikiamos kirčiavimo taisyklės negali būti tiesiogiai taikomos automatiniam kirčiavimui dėl dviejų priežasčių: a) kompiuteriui dar reikalingi papildomi duomenys, kuriuos apibendrintai galima vadinti žmogui būdingu „kalbos jausmu“, b) būtina žinoti kirčiavimo taisyklių taikymo tvarką. Taigi projekto tikslas ir yra išanalizuoti dabartinės lietuvių bendrinės kalbos norminio kirčiavimo taisyklių visumą, nustatyti šių taisyklių taikymo dėsningumus ir prioritetus, surinkti papildomas deklaratyvias ir procedūrinės žinias, reikalingas kirčiavimui automatizuoti, sukurti vardažodžių automatinio kirčiavimo metodus.

Darbo metu sukurti vardažodžių automatinio kirčiavimo metodai pateikti interneto serveryje ir yra prieinamos visiems vartotojams. Taip suteikiama galimybė radijo ir televizijos laidų vedėjams, diktoriaus, aktoriams, mokiniams ir studentams, užsieniečiams pasimokyti taisyklingos sakinės lietuvių kalbos.

## Summary

The accentuation of Lithuanian texts is complicated because of the word stress is not attached to any particular syllable. Word stress can move from one syllable to another within the same accentuation paradigm and often depends on word intonation. The accentuation of Lithuanian has been actively investigated by the Lithuanian language research community. However, the rules present in Lithuanian grammars cannot be directly applied to automatic accentuation by computer. Firstly, computer needs additional knowledge accounting for the “language sense” proper to humans. Secondly, computer needs additional knowledge, i.e. preference rules, to order the set of accentuation rules by the priorities of their applicability. The main objective of this project is to investigate the set of accentuation rules as they are defined by Lithuanian grammars, to disambiguate overlapping domains of their application, to collect an additional knowledge in both declarative and procedural form still necessary for the automation of the accentuation process, and, finally, to construct an algorithm capable of automatic accentuation of all parts-of-speech of Lithuanian.

The tools of automatic accentuation of Lithuanian constructed during this project are placed on the Internet, ensuring their accessibility to all Internet users. Potential users of these tools are radio and television anchormen, announcers, actors, schoolchildren and students, foreigners who learn conversational Lithuanian, and all those who study the culture of Lithuanian accentuation and language.

## 1 Įvadas

Lietuvių kalbos vardažodžių kirčiavimas išsamiai išanalizuotas, neblogai kodifikuotas ir aprašytas daugelio lietuvių kalbininkų, tam skirta ne viena įvairaus pobūdžio knyga ar vadovėlis. Vis dėlto kalbos technologijų darbuose pasigendama išsamaus struktūrinio lietuvių kalbos vardažodžių kirčiavimo modelio, be kurio neįmanomi kokybiški šnekos sintezės programiniai produktai, negalima parengti modernių kompiuterinių mokomųjų kirčiavimo programų, kalbos vartotojai neturi šiuolaikinių galimybių greitai pasitikrinti, ar taisyklingai kirčiuoja konkrečius žodžius. Šiame darbe nagrinėjame tik vardažodžių automatinį kirčiavimą, automatinio kirčiavimo struktūrinio modelio sudarymo bei lingvistinių taisyklių pritaikymo kompiuteriui problemas bei jų sprendimą.

## 2 Analogiški darbai

Dauguma kalbų turi fiksuotą kirtį, t.y. kirčio vietą galima nusakyti griežtomis taisyklėmis. Dažniausiai tai būna visai paprasti teiginiai, nurodą kirčio nutolimą nuo žodžio pradžios ar pabaigos. Pagal nuotolį galima išvelgti tris pagrindinius fiksuoto kirčio modelius:

- *Pastoviai kirčiuojamas pirmasis žodžio skiemuo.* Šią sistemą turi latvių, čekų, slovakų, islandų, estų, suomių, vengrų kalbos.
- *Pastoviai kirčiuojamas paskutinis skiemuo.* Šios rūšies kirčiavimas būdingas daugumai tiurkų kalbų, taip pat persų (ir tadžikų) kalbai. Panašiai kirčiuojama ir prancūzų kalboje, tik kirtį gauna ne žodžiai, o tam tikros reikšminės jų grupės.
- *Pastoviai kirčiuojamas priešpaskutinis skiemuo.* Priešpaskutinio skiemens kirtį turį, pavyzdžiui lenkų kalba.

Galimi ir sudėtingesni fiksuoto kirčio modeliai, kai kirčio vieta priklauso ne tik nuo žodžio ribų, bet ir nuo balsių bei skiemenų kiekybės. Pavyzdžiui, mongolų kalboje kirtį gauna pirmasis ilgasis žodžio skiemuo, o kai visi žodžio skiemenys trumpi, pirmasis skiemuo.

Fiksuotą kirtį turinčiose kalbose automatinis kirčiavimas nesukelia ypatingų problemų. Lietuvių kalba, kaip ir rusų, bulgarų, serbų, chroatų, italų, ispanų, anglų, turi laisvąją kirtį. Kai kuriose laisvąją kirtį turinčiose kalbose daugelis vienodas galūnes turinčių žodžių kirčiuojami vienodai, pavyzdžiui, taip yra italų kalboje. Kalbose, kuriose teksto transkribavimas yra sudėtingas, pavyzdžiui anglų, vokiečių kalbose teksto kirčiavimas ir transkribavimas atliekami kartu naudojant vieną kirčiuotų ir transkribuotų žodžių ar žodžių dalių žodyną. Lietuvių kalboje gali būti kirčiuotas ne tik bet kuris žodžio skiemuo, bet žodžio paradigmoje kirtis keičia vietą, kuri dažnai priklauso nuo priegaidės.

Nors lietuvių kalbos kirčio ir priegaidės kitimo dėsningumai buvo išsamiai nagrinėti lietuvių kalbos mokslo tyrimo darbuose, tačiau įprastinėje gramatikoje teikiamos kirčiavimo taisyklės negali būti tiesiogiai taikomos automatiniam vardažodžių kirčiavimui dėl dviejų priežasčių: a) kompiuteriui dar reikalingi papildomi duomenys, kuriuos apibendrintai galima vadinti žmogui būdingu „kalbos jausmu“, b) būtina žinoti kirčiavimo taisyklių taikymo tvarką..

Yra bandymų spręsti ir lietuvių kalbos automatinio kirčiavimo problemą . P. Kasparaitis [1] gana plačiai išnaginėjęs daiktavardžių ir būdvardžių kirčiavimą, veiksmažodžiams skyrė, atrodo, mažiau dėmesio. Kasparaičio pateikta metodika pasižymi tuo, kad yra nagrinėjama ne morfologinė informacija [4], o žodžio sandaros ypatumai, todėl ir kirčiavimo taisyklės paremtos žodžio sandaros ypatumais, todėl jas sunku suvokti, patikrinti, ar tinkamai modifikuotos ir pritaikytos tradicinės kirčiavimo taisyklės (kurios remiasi morfologine informacija). Minėtina ir tai, kad P.Kasparaitis [1], testuodamas savo sukurtą kirčiavimo algoritmą, naudojo apie 900 žodžių, kurie nebuvo sugrupuoti pagal kalbos dalis. Turint galvoje tokią testavimo imtį ir neaiškią jos struktūrą, sunkoka nustatyti siūlomo algoritmo kirčiavimo tikslumą.

Taigi šio projekto tikslas ir yra išnagrinėti lietuvių bendrinės kalbos vardažodžių kirčiavimo sistemą, nustatyti kirčiavimo taisyklių taikymo prioritetus, surinkti papildomus duomenis, reikalingus kirčiavimui automatizuoti, sukurti vardažodžių automatinio kirčiavimo metodus. Be to, šio projekto tikslas buvo įdiegti sukurtus metodus internete bei Lietuvos ir pasaulio vartotojams suteikti galimybę naudotis pirmąja automatinio kirčiavimo programa . O ji labai reikalinga radijo ir televizijos laidų vedėjams, diktoriaus, aktoriams, mokiniams, studentams, užsieniečiams, visi, besimokantiems taisyklingo lietuvių kalbos kirčiavimo.

### **3 Lietuvių kalbos vardažodžių kirčiavimas**

#### **3.1 Kalbos resursai, reikalingi automatizuotam vardažodžių kirčiavimui**

##### *3.1.1 Pirminiai duomenys*

Daugelyje lietuvių kalbos gramatikų ir įvairių vadovėlių aprašytos kirčiavimo taisyklės skirtos žmonėms, t.y. jos suformuluotos tikintis, kad skaitytojas turi įgimtą kalbos jausmą, pavyzdžiui, vardažodžių kirčiuotė nustatoma pagal tai, kurioje žodžio dalyje – kamiene ar galūnėje – kirčiuojamos daugiskaitos naudininko ir galininko formos. Daroma prielaida, kad kirčiuojantysis (žmogus) pajėgia tai nustatyti. Kiek kitoks požiūris į kirčiavimą turėtų būti užsieniečiams skirtuose vadovėliuose, juo iš dalies galėtų pasinaudoti automatinio kirčiavimo programų kūrėjai. Deja, šiuose vadovėliuose sisteminiam kirčiavimui beveik neskiriama dėmesio. Manoma, kad taisyklingo kirčiavimo pagrindų reikia išmokti kartu su žodžiais, o vėliau, jau pasiekus tam tikrą lygį, sisteminio kirčiavimo galima mokytis ir iš gimtakalbiams skirtų leidinių.

Kad galima būtų kalbininkų žmonėms sudarytas vardažodžių kirčiavimo taisykles ir nustatytus dėsningumus panaudoti automatinio kirčiavimo metodams kurti, *reikalingi pirminiai duomenys* (žr. 1 lent.).

**1 lentelė. Lietuvių kalbos bendrinių žodžių kirčiavimo duomenys (ilustracija).**

Antraštinė žodžio forma	Kalbos dalis	Kirčiuotė	Kirčiuotas skiemuo nuo žodžio pradžios	Kirtis ir priegaidė
abonementas	dktv.	2	4	~
abonementinis	būdv.	1	4	~
abonentas	dktv.	2	3	~
abonentinis	būdv.	1	3	~
aborigenas	dktv.	2	1	`
...	...	...	...	...

Iš 1 lentelės matyti, kad reikia žinoti: kirčiuotę, vardininko kirčio vietą ir priegaidę, kalbos dalį (daiktavardis, būdvardis, skaitvardis, įvardis). Informacija apie kalbos dalį svarbi ne visais atvejais, tačiau mūsų surinktoje vardažodžių duomenų bazėje ji pateikiama. Iš viso šiame sąrašė yra beveik 25 tūkst. daiktavardžių, 7,5 tūkst. būdvardžių, 116 skaitvardžių ir 48 įvardžių. Reikia pasakyti, kad į šį sąrašą nebuvo įrašyti daiktavardžio priesagų *-imas*, *-ymas* vediniai. Tai labai darios priesagos (beveik 20 tūkst. žodžių), su kuriomis abstraktusis daiktavardis gali būti padaromas iš kiekvieno veiksmazodžio, jų kirčiavimas nesudėtingas ir gali būti aprašomas pagal elementarias taisykles.

Duomenys buvo renkami iš „Dabartinės lietuvių kalbos žodyno“. Šiame žodyne pateikiami tik bendriniai ir dažnesni tarptautiniai žodžiai. Todėl mūsų rengiamas sąrašas buvo papildytas dar dažnesniais asmenvardžiais ir vietovardžiais (žr. 2 ir 3 lent.), apie juos pateikiama ta pati informacija, kaip ir apie bendrinius lietuvių kalbos žodžius.

**2 lentelė. Lietuvių kalbos asmenvardžių kirčiavimo duomenys (ilustracija).**

Antraštinė žodžio forma	Kirčiuotė	Kirčiuotas skiemuo nuo žodžio pradžios	Kirtis ir priegaidė
Abraomas	2	3	~
Abrys	4	2	~
Achilas	2	2	`
Ada	2	2	`
Adas	2	1	~
Adauktas	1	2	'
Adelė	2	2	`
Adelija	1	2	`
Adelina	2	4	`
...	...	...	...

Sąrašas papildytas beveik 3,5 tūkst. vardų. Būtina paminėti, kad kiti asmenvardžiai (pavardės ir pravardės) į sąrašą neįtraukti, nes būdingesnių priesaginių pavardžių kirčiavimą daugeliu atvejų galima aprašyti elementariomis taisyklėmis, o

pavardes, kilusias iš bendrinio žodžio, galima kirčiuoti remiantis ankstesniais sąrašais. Pravardės yra šnekamosios kalbos leksiniai vienetai, todėl jų kirčiuoti kol kas neketinama. Kita vertus, nemaža jų dalis yra kilusi iš bendrinių žodžių ir juos galima kirčiuoti pagal bendrąjį sąrašą.

**3 lentelė. Dažnesnių lietuvių kalbos vietovardžių kirčiavimo duomenys (ilustracija).**

Antraštinė žodžio forma	Kirčiuotė	Kirčiuotas skiemuo nuo žodžio pradžios	Kirtis ir priegaidė
Ablinga	3b	3	`
Abromiškės	1	2	~
Acokavai	1	2	'
Adakavas	2	3	~
Adomėniškė	1	3	'
Adomynė	2	3	~
Adutiškis	1	2	`
...	...	...	...

Parengtas sąrašas papildytas beveik 3 tūkst. dažnesnių lietuvių kalbos vietovardžių (daugiausiai gyvenamųjų vietų vardai). Sąrašą būtina dar papildyti: reikia įtraukti dažniau vartojamų hidronimų, miestų dalių, istorinių vietų pavadinimų.

### 3.1.2 Priesaginių daiktavardžių ir būdvardžių kirčiavimo duomenys

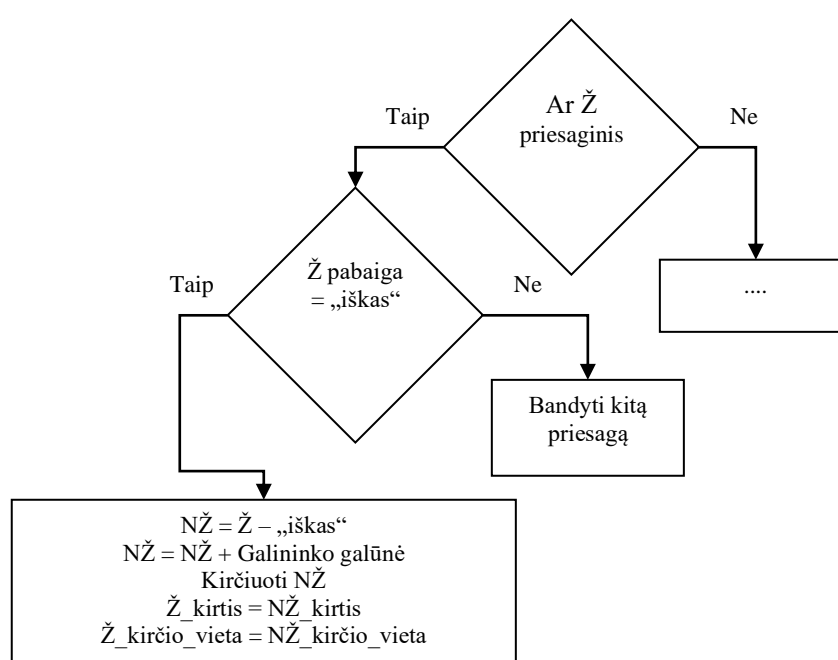
Atidžiai išanalizavus sudarytą pirminį daiktavardžių ir būdvardžių sąrašą, matyti, kad nemažai priesaginių vedinių kirčiuojami pagal nesunkiai suprogramuojamą kirčiavimo modelį. Taip būtų sumažintas pradinis kirčiavimo duomenų sąrašas. Be to, kai kurios priesagos labai produktyvios, su jomis nuolat pasidaroma naujų žodžių (pvz.: jau minėtos *-imas*, *-ymas*, *-inis*, *-elis*, *-ėlis* ir kt.), o jų visų sąrašą pateikti būtų neįmanoma.

Kirčiavimo požiūriu dažnesnės priesagos gali būti dvejopos: vienų akcentinis statusas lietuvių bendrinėje kalboje nusistovėjęs (t.y. jų vediniai visada vienodai kirčiuojami), kitų – ne (jų vediniai gali būti kirčiuojami skirtingai). Pirmajai grupei priklausančios priesagos taip pat nevienodos, jos skiriasi kirčio vieta: vienos visada kirtį turi priesagoje, kitos išlaiko pamatinio žodžio kirčio vietą ir priegaidę. Todėl priesagos kuriamų metodų požiūriu suskirstytos į dvi grupes. Pirmajai grupei – nusistovėjusio akcentinio statuso, kirčiuojamos priesagoje – priskirtos 92 priesagos (žr. 4 lent.), antrajai (nenusistovėjusio akcentinio statuso arba išlaikančios pamatinio žodžio kirtį žr. 1 pav.) – 13.



4 lentelė. Nusistovėjusio akcentinio statuso priesagų kirčiavimas (ilustracija).

Priesaga	Kalbos dalis	Kirčiuotė	Kirčiuotas skiemuo nuo priesagos pradžios	Kirtis ir priegaidė
-ukas	dktv.	2	1	`
-ukas	būdv.	2	1	`
-ukė	dktv.	2	1	`
-ukė	būdv.	2	1	`
-ingas	būdv.	1	1	`
-inga	būdv.	1	1	`
-ierius	dktv.	2	1	~
-ierė	dktv.	2	1	~
...	...	...	...	...



1 pav. Priesagos „iškas“ kirčiavimo schema. Ž – kirčiuojamas žodis, NŽ – naujas žodis, Ž\_kirtis – kirčiuojamo žodžio kirčio tipas, Ž\_kirčio\_vieta – kirčiuojamo žodžio kirčio vieta

### 3.2 Morfologinės analizės ir skiemenavimo procedūra

Automatizuojant vardažodžių kirčiavimą, neužtenka tik aprašyti kalbininkų nurodomus kirčiavimo tipus – kirčiuotes. Reikia papildomos žodžių

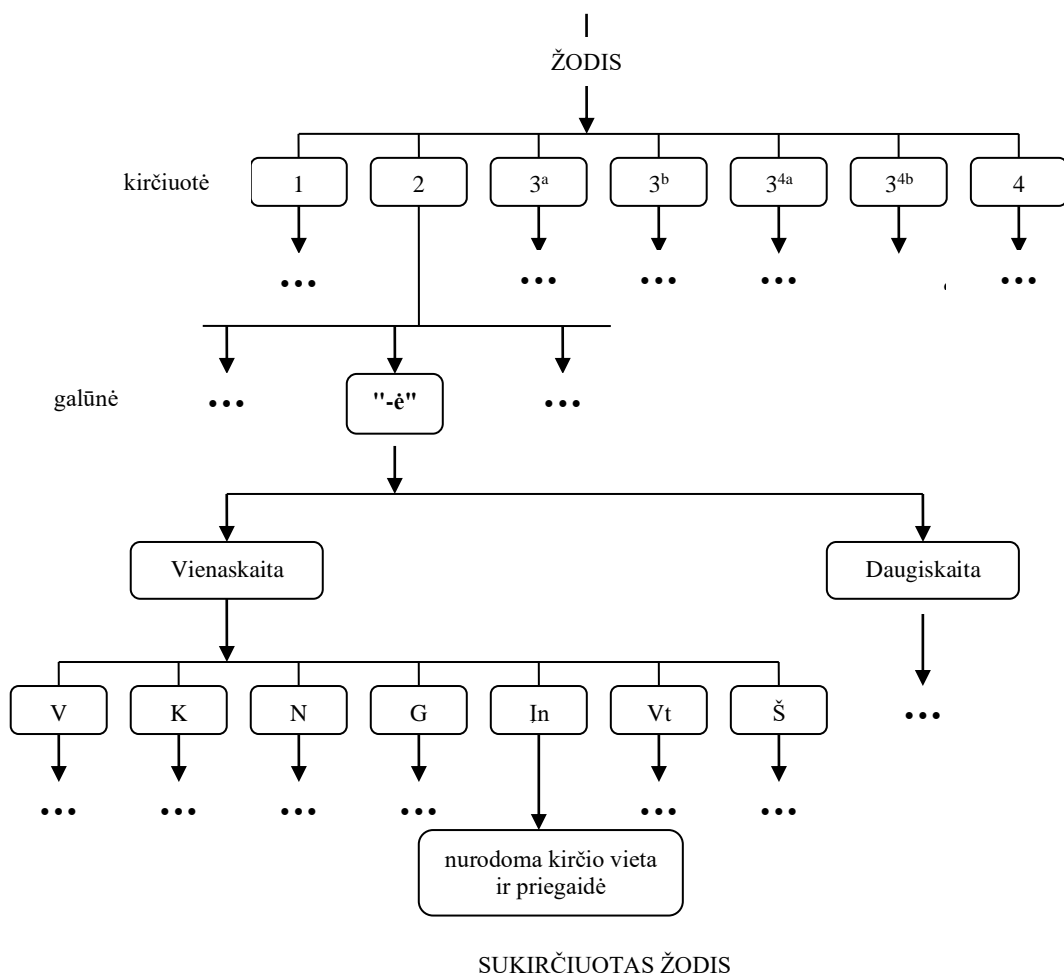
- morfologinės (kalbos dalis, giminė, skaičius linksnis, pagrindinė forma) ir
- sandaros (kiek skiemenų sudaro žodį) informacijos.

Šią informaciją gali pateikti morfologinis analizatorius ir žodžių skiemenavimo algoritmas. Tokių įrankių kūrimas nėra šio projekto tikslas, todėl šiuo atveju turėtume pasinaudoti jau esančiais.

### 3.3 Kirčiavimo taisyklių struktūrinė analizė ir pritaikymas automatinio kirčiavimo metodams

Morfologinės analizės informacija, skiemenavimo metu gauta informacija ir žodyne saugomi pirminiai vardažodžių kirčiavimo duomenys pateikiami kirčiavimo procedūrai, turinčiai medžio pavidalą.

Visuose kalbininkų darbuose paprastai nurodomos 4 kirčiuotės: 1, 2, 3 (ir jos variantai 3<sup>a</sup>, 3<sup>b</sup>, 3<sup>4a</sup>, 3<sup>4b</sup>), 4. Tiek visiškai užtenka žmogui, norinčiam taisyklingai sukirčiuoti. Tačiau tokios informacijos per mažai automatiniam kirčiavimo metodams. Kirčiavimo paradigmos skiriasi ne tik kitokia kirčio vieta, priegaidė, bet ir skirtingomis galūnėmis, kurias lemia žodžio linksniuotė. Tad visas skirtingas kirčiavimo paradigmas (t.y. 7) turėtume padauginti iš skirtingų galūnių skaičiaus (t.y. 9). Atidžiai išanalizavus gautų kirčiavimo paradigmu (63) panašumus ir skirtumus, paaiškėjo, kad vardažodžių kirčiavimui aprašyti reikia 27 modelių.



2 pav. Vardažodžių kirčiavimo medžio iliustracija.

2 pav. nurodytos kirčiavimo procedūros užtenka daiktavardžiams. Kai kurie būdvardžiai, skaitvardžiai, įvardžiai gali turėti įvardžiuotines formas, o ne *-inis, -ė* būdvardžiai kaitomi ir laipsniais. Įvardžiuotinės formos kirčiuojamos pagal 3 individualius algoritmus, aukštesniojo ir aukščiausiojo laipsnio būdvardžiai pagal 2 algoritmus.

## 4 Rezultatai

Sukurtų vardažodžių kirčiavimo algoritmų tikslumas buvo tikrinamas remiantis dažniniu žodžių formų sąrašu, gautu iš 85 mln. žodžių lietuvių kalbos tekstyno [3]. Kiekvienai gramatinei kategorijai (pvz. daiktavardžio vns. v., daiktavardžio vns. k. ir t. t.) buvo išrinkta po 100 dažniausių žodžių iš minėto dažninio sąrašo, šie žodžiai buvo automatiškai sukirčiuoti ir pateikti vertinti ekspertui. Kai kurioms retoms gramatinėms kategorijoms dažniniame sąraše nepavyko rasti 100 žodžių. Tada buvo išrinkti visi dažniniame sąraše jai atstovaujantys žodžiai. Iš viso buvo palyginta 12870 automatiškai ir eksperto sukirčiuotų žodžių. Rezultatų suvestinė pateikta 5 lentelėje, o detalūs kiekvienos gramatinės kategorijos rezultatai – 1 priede.

5 lentelė. Vardažodžių automatinio kirčiavimo metodų tikslumas.

Kalbos dalis	Gramatinės kategorijos	Iš viso žodžių	Negali kirčiuoti, %	Klaidingai kirčiavo, %
Daiktavardis	65	5533	2,63	0,44
Būdvardis	73	4828	3,74	1,95
Skaitvardis	113	1410	1,91	0,36
Įvardis	124	1099	0,72	0,73
<b>Iš viso</b>	<b>375</b>	<b>12870</b>	<b>2,81</b>	<b>1,02</b>

Daiktavardis – dažniausia kalbos dalis – kirčiuojama beveik nepriekaištingai: realizuoti metodai klaidų beveik visiškai nedaro, o negali kirčiuoti 2,5 % daiktavardžių (tai rečiau vartojami tarptautiniai žodžiai). Būdvardžiai, įvardžiai, skaitvardžiai sudaro šiek tiek daugiau nei penktadalį tekste vartojamų žodžių. Iš šių kalbos dalių prasčiausiai kirčiuojami būdvardžiai: šiek tiek mažiau nei 2 % klaidų, dar beveik 4 % negali kirčiuoti (iš 4828 žodžių). Realizuoti metodai negali kirčiuoti apie 2 % skaitvardžių (tai dviejų žodžių, kurie nebuvo įtraukti į duomenų bazę, skirtingos formos), kitus sukirčiavo beveik be klaidų. Įvardžių kirčiavimo klaidos sudarė apie 1 % (tai visi vyr.gim.vns.n.įvardž. įvardžiai, turintys tvirtapradį a).

Būtina paminėti tai, kad pasiūlyta automatinio kirčiavimo metodika leidžia nesunkiai ištaisyti daugumą nustatytų klaidų. Tai galima padaryti, papildžius duomenų bazę retesniais žodžiais bei pakoregavus vieną kitą taisyklę ar viena kita taisykle papildžius dabartinį kirčiavimo taisyklių rinkinį. Šiuo požiūriu taisyklėmis paremta metodika, siūloma šiame darbe, pranašesnė už įprastus tikimybinus metodus.

## 5 Automatinio vardažodžių kirčiavimo pateikimas internete

Automatiniam vardažodžių kirčiavimui pateikti internete sukurta speciali interneto svetainė (žr. 3 pav.), kuri įdėta Vytauto Didžiojo universiteto Kompiuterinės lingvistikos centro serveryje adresu <http://donelaitis.vdu.lt/~astkar><sup>1</sup>.

<sup>1</sup> Tai laikinas svetainės adresas. Dėl lėšų trūkumo nebuvo galimybės įsigyti atskiro interneto serverio kirčiavimo paslaugoms teikti.

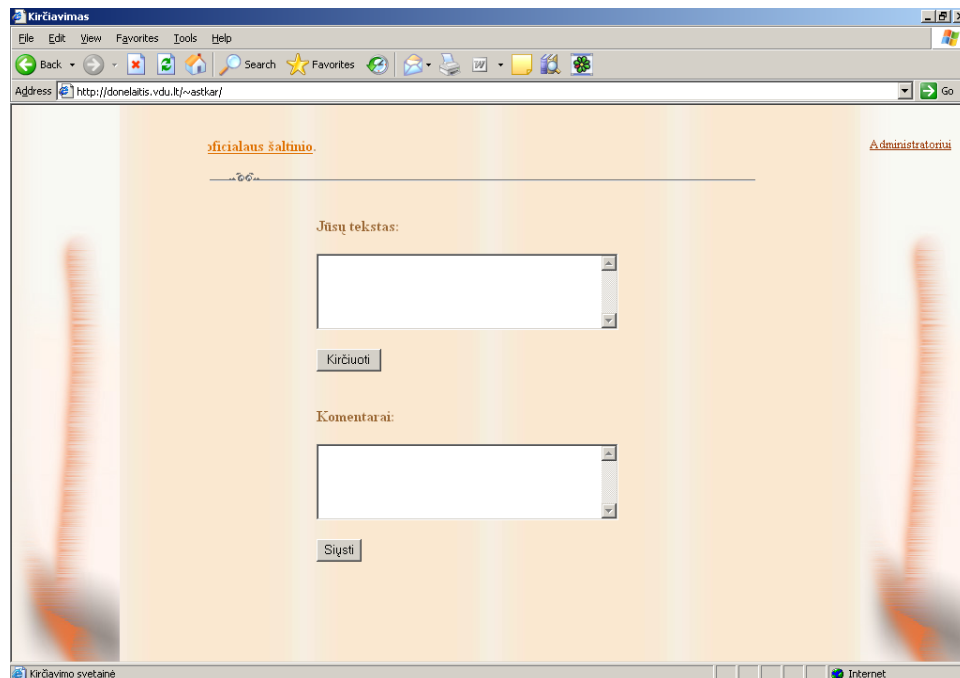
Svetainė yra viešai prieinama. Vartotojui reikia įsidiegti lietuvišką šriftą „Palemonas“ [5]. Svetainė naudotis labai paprasta:

1. Laukelyje „Jūsų tekstas“ reikia surinkti norimus kirčiuoti vardažodžius.
2. Paspausti mygtuką „Kirčiuoti“.
3. Vardažodžiai bus sukirčiuoti. Jeigu surinksite kitas kalbos dalis, jos liks nekirčiuotos.

Jeigu vartotjas turi pastabų apie kirčiavimą ar tiesiog komentarų, juos galima surašyti laukelyje „Komentari“ ir nusiųsti (paspausti mygtuką „Siųsti“) į tarnybinę stotį svetainės administratoriui .

Komentarų bei svetainės parametrų peržiūrai, koregavimui specialiai sukurta Administratoriaus skiltis.

3 pav. Automatinio vardažodžių kirčiavimo interneto svetainės vaizdas



## 6 Išvados

Darbe buvo ištirta dabartinės lietuvių bendrinės kalbos vardažodžių kirčiavimo taisyklių visuma, išryškintos persidengiančios šių taisyklių taikymo sritys ir nustatyti taisyklių taikymo prioritetai. Pasiūlytas bendrinės lietuvių kalbos vardažodžių kirčiavimo struktūrinis modelis bei pagrindinė jo dalis - kirčiavimo taisyklių medis – sudarytas morfologinės informacijos pagrindu. Modelis yra tinkamas automatinio kirčiavimo problemai spręsti ir yra tikslinga toliau tobulinti tiek šią konkrečią modelio realizaciją, tiek ir patį teorinį modelį.

Darbe buvo sukurta duomenų bazė, kurioje surinktos ir kompiuterine kalba užrašytos automatiniam vardažodžių kirčiavimui reikalingos papildomos žinios.

Darbe buvo sudarytas ir programiškai realizuotas daiktavardžių, būdvardžių, skaitvardžių ir įvardžių automatinis kirčiavimo algoritmas, ištirtas jo veikimo

tikslumas ir nustatyta, kad taisyklingai kirčiuojama apie 99 % lietuvių kalbos vardažodžių. Morfoliginis analizatorius gražina visus galimus morfologinių reikšmių variantus (t. y. visas homoforas), tačiau nenurodo, kurią pasirinkti. Šiuo metu daromas atsitiktinis pasirinkimas. Tad taisyklingai kirčiuoti lietuviškus tekstus bus galima tik išsprendus morfoliginio daugiareikšmiškumo problemą (o tai jau ne šio darbo tikslas). Darbo metu sukurti automatinio vardažodžių kirčiavimo metodaiskelbiami interneto serveryje ir yra prieinami visiems vartotojams. Jais galės naudotis radijo ir televizijos laidų vedėjai, diktoriai, aktoriai, mokiniai ir studentai, užsieniečiai, besimokantys taisyklingos sakytinės lietuvių kalbos, ir visi kiti, besimokantys kirčiavimo ir kalbos kultūros.

## 7 Literatūra

[1] P. Kasparaitis, 2001, Lietuvių kalbos kompiuterinės sintezė (daktaro disertacija, rankraštis, Vilniaus universitetas), Vilnius.

[2] P. Kniūkšta, A. Lyberis, 1989, Mokomasis lietuvių kalbos ir kirčiavimo žodynas, Šviesa.

[3] R. Marcinkevičienė, 2000. Tekstynų lingvistika (teorija ir praktika) Darbai ir dienos, 24, Vytauto Didžiojo universitetas, p. 7-63.

[4] V. Zinkevičius, 2000, Lemuoklis – morfologinei analizei. – Darbai ir dienos, T. 24, 245–274.

[5] Valstybinė lietuvių kalbos komisija, šriftas „Palemonas“  
<http://www.vlkk.lt/palemonas.html>

### Parengtos publikacijos

Darbo pagrindu parengta ir tarptautinėje recenzuojamoje konferencijoje užsienyje išspausdinta publikacija (žr. 2 priedą):

Norkevičius G., Raškinis G., Kazlauskienė A., Knowledge-based grapheme-to-phoneme conversion of Lithuanian words // In Proceedings of the 10th International Conference on Speech and Computer - Specom. Patras, Greece, 2005, p. 235-238.

Projekto vadovas

dr. Gailius Raškinis

## 1 PRIEDAS

### DETALUS KIEKVIENOS GRAMATINĖS KATEGORIJOS KIRČIAVIMO TIKSLUMO ĮVERTINIMAS

#### Daiktavardžiai:

Gramatinė kategorija	Iš viso žodžių	Klaidų kiekis	Negali kirčiuoti	Klaidos, %	Atsisakė, %
dktv.	100	0	65	0,0	65,0
dktv. bendr.gim. dgsk. G.	81	0	1	0,0	1,2
dktv. bendr.gim. dgsk. In.	55	0	0	0,0	0,0
dktv. bendr.gim. dgsk. K.	100	1	2	1,0	2,0
dktv. bendr.gim. dgsk. N.	40	0	0	0,0	0,0
dktv. bendr.gim. dgsk. S.	19	0	0	0,0	0,0
dktv. bendr.gim. dgsk. V.	100	0	3	0,0	3,0
dktv. bendr.gim. dgsk. Vt.	12	0	0	0,0	0,0
dktv. bendr.gim. vnsk. G.	96	1	3	1,1	3,1
dktv. bendr.gim. vnsk. In.	100	0	7	0,0	7,0
dktv. bendr.gim. vnsk. K.	100	0	4	0,0	4,0
dktv. bendr.gim. vnsk. N.	61	0	2	0,0	3,3
dktv. bendr.gim. vnsk. S.	1	0	0	0,0	0,0
dktv. bendr.gim. vnsk. V.	100	0	6	0,0	6,0
dktv. bendr.gim. vnsk. Vt.	19	1	0	5,3	0,0
dktv. mot.gim. dgsk. G.	100	0	2	0,0	2,0
dktv. mot.gim. dgsk. In.	100	0	2	0,0	2,0
dktv. mot.gim. dgsk. K.	100	3	3	3,1	3,0
dktv. mot.gim. dgsk. N.	100	2	1	2,0	1,0
dktv. mot.gim. dgsk. S.	100	1	2	1,0	2,0
dktv. mot.gim. dgsk. V.	100	0	0	0,0	0,0
dktv. mot.gim. dgsk. Vt.	100	0	1	0,0	1,0
dktv. mot.gim. vnsk. G.	100	0	3	0,0	3,0
dktv. mot.gim. vnsk. In.	100	0	2	0,0	2,0
dktv. mot.gim. vnsk. K.	100	1	3	1,0	3,0
dktv. mot.gim. vnsk. N.	100	0	3	0,0	3,0
dktv. mot.gim. vnsk. S.	55	0	2	0,0	3,6
dktv. mot.gim. vnsk. V.	100	0	4	0,0	4,0
dktv. mot.gim. vnsk. Vt.	100	1	3	1,0	3,0
dktv. vyr.gim. dgsk. G.	100	2	1	2,0	1,0
dktv. vyr.gim. dgsk. G. nesngr.	100	0	0	0,0	0,0
dktv. vyr.gim. dgsk. In.	100	0	0	0,0	0,0
dktv. vyr.gim. dgsk. In. nesngr.	100	0	0	0,0	0,0
dktv. vyr.gim. dgsk. K.	100	1	1	1,0	1,0
dktv. vyr.gim. dgsk. K. nesngr.	100	0	0	0,0	0,0
dktv. vyr.gim. dgsk. K. sngr.	10	0	0	0,0	0,0
dktv. vyr.gim. dgsk. N.	100	0	1	0,0	1,0
dktv. vyr.gim. dgsk. N. nesngr.	100	0	0	0,0	0,0
dktv. vyr.gim. dgsk. S.	100	0	3	0,0	3,0
dktv. vyr.gim. dgsk. S. nesngr.	100	0	0	0,0	0,0
dktv. vyr.gim. dgsk. S. sngr.	8	0	0	0,0	0,0
dktv. vyr.gim. dgsk. V.	100	0	3	0,0	3,0
dktv. vyr.gim. dgsk. V. nesngr.	100	0	0	0,0	0,0

dktv. vyr.gim. dgsk. V. sngr.	8	0	0	0,0	0,0
dktv. vyr.gim. dgsk. Vt.	100	2	1	2,0	1,0
dktv. vyr.gim. dgsk. Vt. nesngr.	89	0	0	0,0	0,0
dktv. vyr.gim. vnsk. G.	100	0	2	0,0	2,0
dktv. vyr.gim. vnsk. G. nesngr.	100	0	0	0,0	0,0
dktv. vyr.gim. vnsk. G. sngr.	100	0	0	0,0	0,0
dktv. vyr.gim. vnsk. In.	100	1	0	1,0	0,0
dktv. vyr.gim. vnsk. In. nesngr.	100	0	0	0,0	0,0
dktv. vyr.gim. vnsk. In. sngr.	84	0	0	0,0	0,0
dktv. vyr.gim. vnsk. K.	100	0	2	0,0	2,0
dktv. vyr.gim. vnsk. K. nesngr.	100	0	0	0,0	0,0
dktv. vyr.gim. vnsk. K. sngr.	100	0	0	0,0	0,0
dktv. vyr.gim. vnsk. N.	100	1	0	1,0	0,0
dktv. vyr.gim. vnsk. N. nesngr.	100	0	0	0,0	0,0
dktv. vyr.gim. vnsk. N. sngr.	81	0	0	0,0	0,0
dktv. vyr.gim. vnsk. S.	100	1	4	1,0	4,0
dktv. vyr.gim. vnsk. V.	100	0	1	0,0	1,0
dktv. vyr.gim. vnsk. V. nesngr.	100	0	0	0,0	0,0
dktv. vyr.gim. vnsk. V. sngr.	100	3	0	3,0	0,0
dktv. vyr.gim. vnsk. Vt.	100	2	3	2,1	3,0
dktv. vyr.gim. vnsk. Vt. nesngr.	100	0	0	0,0	0,0
dktv. vyr.gim. vnsk. Vt. sngr.	14	0	0	0,0	0,0
Iš viso	5533	24	146	0,4	2,6

## Būdvardžiai

Gramatinė kategorija	Iš viso žodžių	Klaidų kiekis	Negali kirčiuoti	Klaidos, %	Nekirčiuoja, %
bdvr.	32	6	5	22,2	15,6
bdvr. bevrd.gim. neivardz.	100	1	0	1,0	0,0
bdvr. mot.gim. dgsk. G.	1	0	1	0,0	100,0
bdvr. mot.gim. dgsk. G. ivardz.	100	0	0	0,0	0,0
bdvr. mot.gim. dgsk. G. neivardz.	100	1	2	1,0	2,0
bdvr. mot.gim. dgsk. In.	1	0	1	0,0	100,0
bdvr. mot.gim. dgsk. In. ivardz.	100	0	1	0,0	1,0
bdvr. mot.gim. dgsk. In. neivardz.	100	1	1	1,0	1,0
bdvr. mot.gim. dgsk. K.	1	0	1	0,0	100,0
bdvr. mot.gim. dgsk. K. ivardz.	100	0	2	0,0	2,0
bdvr. mot.gim. dgsk. K. neivardz.	100	1	2	1,0	2,0
bdvr. mot.gim. dgsk. N.	1	0	1	0,0	100,0
bdvr. mot.gim. dgsk. N. ivardz.	100	1	2	1,0	2,0
bdvr. mot.gim. dgsk. N. neivardz.	100	1	4	1,0	4,0
bdvr. mot.gim. dgsk. V.	1	0	1	0,0	100,0
bdvr. mot.gim. dgsk. V. ivardz.	100	0	0	0,0	0,0
bdvr. mot.gim. dgsk. V. neivardz.	100	2	0	2,0	0,0
bdvr. mot.gim. dgsk. Vt. ivardz.	79	0	1	0,0	1,3
bdvr. mot.gim. dgsk. Vt. neivardz.	100	1	1	1,0	1,0
bdvr. mot.gim. vnsk. G.	1	0	1	0,0	100,0
bdvr. mot.gim. vnsk. G. ivardz.	100	0	1	0,0	1,0
bdvr. mot.gim. vnsk. G. neivardz.	100	3	0	3,0	0,0
bdvr. mot.gim. vnsk. In.	1	0	1	0,0	100,0
bdvr. mot.gim. vnsk. In. ivardz.	100	0	0	0,0	0,0
bdvr. mot.gim. vnsk. In. neivardz.	100	8	0	8,0	0,0

bdvr. mot.gim. vnsk. K.	1	0	1	0,0	100,0
bdvr. mot.gim. vnsk. K. ivardz.	100	6	0	6,0	0,0
bdvr. mot.gim. vnsk. N.	1	0	1	0,0	100,0
bdvr. mot.gim. vnsk. N. ivardz.	100	0	2	0,0	2,0
bdvr. mot.gim. vnsk. N. neivardz.	100	2	2	2,0	2,0
bdvr. mot.gim. vnsk. V.	1	0	1	0,0	100,0
bdvr. mot.gim. vnsk. V. ivardz.	100	0	1	0,0	1,0
bdvr. mot.gim. vnsk. V. neivardz.	100	12	0	12,0	0,0
bdvr. mot.gim. vnsk. Vt.	1	0	1	0,0	100,0
bdvr. mot.gim. vnsk. Vt. ivardz.	100	0	1	0,0	1,0
bdvr. mot.gim. vnsk. Vt. neivardz.	100	1	2	1,0	2,0
bdvr. vyr.gim. dgsk. G.	1	0	1	0,0	100,0
bdvr. vyr.gim. dgsk. G. ivardz.	100	0	3	0,0	3,0
bdvr. vyr.gim. dgsk. G. neivardz.	100	11	3	11,3	3,0
bdvr. vyr.gim. dgsk. In.	1	0	1	0,0	100,0
bdvr. vyr.gim. dgsk. In. ivardz.	100	0	4	0,0	4,0
bdvr. vyr.gim. dgsk. In. neivardz.	100	7	1	7,1	1,0
bdvr. vyr.gim. dgsk. K.	1	0	1	0,0	100,0
bdvr. vyr.gim. dgsk. K. ivardz.	100	0	2	0,0	2,0
bdvr. vyr.gim. dgsk. N.	1	0	1	0,0	100,0
bdvr. vyr.gim. dgsk. N. ivardz.	100	0	3	0,0	3,0
bdvr. vyr.gim. dgsk. N. neivardz.	100	2	5	2,1	5,0
bdvr. vyr.gim. dgsk. V.	1	0	1	0,0	100,0
bdvr. vyr.gim. dgsk. V. ivardz.	100	0	3	0,0	3,0
bdvr. vyr.gim. dgsk. V. neivardz.	100	1	6	1,1	6,0
bdvr. vyr.gim. dgsk. Vt. ivardz.	95	0	1	0,0	1,1
bdvr. vyr.gim. dgsk. Vt. neivardz.	100	0	2	0,0	2,0
bdvr. vyr.gim. vnsk. G.	1	0	1	0,0	100,0
bdvr. vyr.gim. vnsk. G. ivardz.	100	0	2	0,0	2,0
bdvr. vyr.gim. vnsk. G. neivardz.	100	1	6	1,1	6,0
bdvr. vyr.gim. vnsk. In.	1	0	1	0,0	100,0
bdvr. vyr.gim. vnsk. In. ivardz.	100	1	2	1,0	2,0
bdvr. vyr.gim. vnsk. In. neivardz.	100	1	0	1,0	0,0
bdvr. vyr.gim. vnsk. K.	1	0	1	0,0	100,0
bdvr. vyr.gim. vnsk. K. ivardz.	100	3	2	3,1	2,0
bdvr. vyr.gim. vnsk. K. neivardz.	100	3	6	3,2	6,0
bdvr. vyr.gim. vnsk. N.	1	0	1	0,0	100,0
bdvr. vyr.gim. vnsk. N. ivardz.	100	0	68	0,0	68,0
bdvr. vyr.gim. vnsk. N. neivardz.	100	0	1	0,0	1,0
bdvr. vyr.gim. vnsk. S. neivardz.	100	6	2	6,1	2,0
bdvr. vyr.gim. vnsk. V.	1	0	1	0,0	100,0
bdvr. vyr.gim. vnsk. V. ivardz.	100	1	6	1,1	6,0
bdvr. vyr.gim. vnsk. V. neivardz.	100	5	1	5,1	1,0
bdvr. vyr.gim. vnsk. Vt.	1	0	1	0,0	100,0
bdvr. vyr.gim. vnsk. Vt. ivardz.	100	0	0	0,0	0,0
bdvr. vyr.gim. vnsk. Vt. neivardz.	100	2	0	2,0	0,0
Viso	4828	91	181	2,0	3,7

### Ivardžiai

Gramatinė kategorija	Iš viso žodžių	Klaidų kiekis	Negali kirčiuoti	Klaidos, %	Nekirčiavo, %
ivrd.	7	0	0	0,0	0,0



ivrd. beverd.gim.	4	0	0	0,0	0,0
ivrd. beverd.gim. neivardz.	8	0	0	0,0	0,0
ivrd. dgsk. G.	3	0	0	0,0	0,0
ivrd. dgsk. In.	3	0	0	0,0	0,0
ivrd. dgsk. K.	3	0	0	0,0	0,0
ivrd. dgsk. N.	3	0	0	0,0	0,0
ivrd. dgsk. V.	3	0	0	0,0	0,0
ivrd. dgsk. Vt.	2	0	0	0,0	0,0
ivrd. G.	12	0	0	0,0	0,0
ivrd. In.	9	0	0	0,0	0,0
ivrd. K.	14	0	0	0,0	0,0
ivrd. mot.gim. dgsk. G.	21	0	0	0,0	0,0
ivrd. mot.gim. dgsk. G. ivardz.	4	0	0	0,0	0,0
ivrd. mot.gim. dgsk. G. neivardz.	14	3	0	21,4	0,0
ivrd. mot.gim. dgsk. In.	16	0	0	0,0	0,0
ivrd. mot.gim. dgsk. In. ivardz.	3	0	0	0,0	0,0
ivrd. mot.gim. dgsk. In. neivardz.	8	0	0	0,0	0,0
ivrd. mot.gim. dgsk. K.	22	0	0	0,0	0,0
ivrd. mot.gim. dgsk. K. ivardz.	7	0	0	0,0	0,0
ivrd. mot.gim. dgsk. K. neivardz.	11	0	0	0,0	0,0
ivrd. mot.gim. dgsk. N.	15	0	0	0,0	0,0
ivrd. mot.gim. dgsk. N. ivardz.	3	0	0	0,0	0,0
ivrd. mot.gim. dgsk. N. neivardz.	7	0	0	0,0	0,0
ivrd. mot.gim. dgsk. V.	27	0	0	0,0	0,0
ivrd. mot.gim. dgsk. V. ivardz.	8	0	0	0,0	0,0
ivrd. mot.gim. dgsk. V. neivardz.	13	1	0	7,7	0,0
ivrd. mot.gim. dgsk. Vt.	14	0	0	0,0	0,0
ivrd. mot.gim. dgsk. Vt. ivardz.	1	0	0	0,0	0,0
ivrd. mot.gim. dgsk. Vt. neivardz.	8	0	0	0,0	0,0
ivrd. mot.gim. dvisk. G.	2	0	0	0,0	0,0
ivrd. mot.gim. dvisk. G. neivardz.	4	0	0	0,0	0,0
ivrd. mot.gim. dvisk. In.	2	0	0	0,0	0,0
ivrd. mot.gim. dvisk. K.	2	0	0	0,0	0,0
ivrd. mot.gim. dvisk. K. neivardz.	3	0	0	0,0	0,0
ivrd. mot.gim. dvisk. N.	2	0	0	0,0	0,0
ivrd. mot.gim. dvisk. V.	2	0	0	0,0	0,0
ivrd. mot.gim. dvisk. V. neivardz.	4	0	0	0,0	0,0
ivrd. mot.gim. G.	2	0	0	0,0	0,0
ivrd. mot.gim. In.	2	0	0	0,0	0,0
ivrd. mot.gim. K.	2	0	0	0,0	0,0
ivrd. mot.gim. N.	1	0	0	0,0	0,0
ivrd. mot.gim. V.	2	0	0	0,0	0,0
ivrd. mot.gim. vnsk. G.	23	0	0	0,0	0,0
ivrd. mot.gim. vnsk. G. ivardz.	8	0	0	0,0	0,0
ivrd. mot.gim. vnsk. G. neivardz.	14	0	0	0,0	0,0
ivrd. mot.gim. vnsk. In.	22	0	0	0,0	0,0
ivrd. mot.gim. vnsk. In. ivardz.	9	0	0	0,0	0,0
ivrd. mot.gim. vnsk. In. neivardz.	12	0	0	0,0	0,0
ivrd. mot.gim. vnsk. K.	27	0	0	0,0	0,0
ivrd. mot.gim. vnsk. K. ivardz.	8	0	0	0,0	0,0
ivrd. mot.gim. vnsk. K. neivardz.	13	0	0	0,0	0,0
ivrd. mot.gim. vnsk. N.	22	0	0	0,0	0,0

ivrd. mot.gim. vnsk. N. ivardz.	5	0	2	0,0	40,0
ivrd. mot.gim. vnsk. N. neivardz.	11	1	0	9,1	0,0
ivrd. mot.gim. vnsk. V.	25	0	0	0,0	0,0
ivrd. mot.gim. vnsk. V. ivardz.	8	0	0	0,0	0,0
ivrd. mot.gim. vnsk. V. neivardz.	12	0	0	0,0	0,0
ivrd. mot.gim. vnsk. Vt.	18	0	0	0,0	0,0
ivrd. mot.gim. vnsk. Vt. ivardz.	3	0	0	0,0	0,0
ivrd. mot.gim. vnsk. Vt. neivardz.	9	0	0	0,0	0,0
ivrd. mot.gim. Vt.	2	0	0	0,0	0,0
ivrd. N.	11	1	0	9,1	0,0
ivrd. V.	12	0	0	0,0	0,0
ivrd. vnsk. G.	3	0	0	0,0	0,0
ivrd. vnsk. In.	3	0	0	0,0	0,0
ivrd. vnsk. K.	3	0	0	0,0	0,0
ivrd. vnsk. N.	3	0	0	0,0	0,0
ivrd. vnsk. V.	3	0	0	0,0	0,0
ivrd. vnsk. Vt.	2	0	0	0,0	0,0
ivrd. Vt.	8	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. G.	22	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. G. ivardz.	6	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. G. neivardz.	10	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. In.	21	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. In. ivardz.	4	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. In. neivardz.	10	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. K.	22	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. K. ivardz.	7	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. K. neivardz.	11	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. N.	21	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. N. ivardz.	4	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. N. neivardz.	7	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. V.	25	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. V. ivardz.	7	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. V. neivardz.	12	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. Vt.	14	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. Vt. ivardz.	1	0	0	0,0	0,0
ivrd. vyr.gim. dgsk. Vt. neivardz.	8	0	0	0,0	0,0
ivrd. vyr.gim. dvisk. G.	2	0	0	0,0	0,0
ivrd. vyr.gim. dvisk. G. neivardz.	4	0	0	0,0	0,0
ivrd. vyr.gim. dvisk. In.	2	0	0	0,0	0,0
ivrd. vyr.gim. dvisk. In. neivardz.	2	0	0	0,0	0,0
ivrd. vyr.gim. dvisk. K.	2	0	0	0,0	0,0
ivrd. vyr.gim. dvisk. K. neivardz.	3	0	0	0,0	0,0
ivrd. vyr.gim. dvisk. N.	2	0	0	0,0	0,0
ivrd. vyr.gim. dvisk. N. neivardz.	2	0	0	0,0	0,0
ivrd. vyr.gim. dvisk. V.	2	0	0	0,0	0,0
ivrd. vyr.gim. dvisk. V. neivardz.	8	0	0	0,0	0,0
ivrd. vyr.gim. G.	2	0	0	0,0	0,0
ivrd. vyr.gim. In.	2	0	0	0,0	0,0
ivrd. vyr.gim. K.	2	0	0	0,0	0,0
ivrd. vyr.gim. N.	2	0	0	0,0	0,0
ivrd. vyr.gim. V.	2	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. G.	23	0	0	0,0	0,0

ivrd. vyr.gim. vnsk. G. ivardz.	10	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. G. neivardz.	14	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. In.	20	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. In. ivardz.	6	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. In. neivardz.	11	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. K.	23	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. K. ivardz.	9	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. K. neivardz.	13	1	0	7,7	0,0
ivrd. vyr.gim. vnsk. N.	20	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. N. ivardz.	6	0	6	0,0	100,0
ivrd. vyr.gim. vnsk. N. neivardz.	9	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. S.	3	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. V.	24	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. V. ivardz.	11	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. V. neivardz.	14	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. Vt.	14	0	0	0,0	0,0
ivrd. vyr.gim. vnsk. Vt. ivardz.	5	1	0	20,0	0,0
ivrd. vyr.gim. vnsk. Vt. neivardz.	10	0	0	0,0	0,0
ivrd. vyr.gim. Vt.	1	0	0	0,0	0,0
Viso	1099	8	8	0,7	0,7

### Skaitvardžiai

Gramatinė kategorija	Iš viso žodžių	Klaidų kiekis	Negali kirčiuoti	Klaidos, %	Nekirčiavo, %
sktv. beverd.gim. kelintin. neivardz.	24	1	0	4,2	0,0
sktv. beverd.gim. kiekin.	1	0	0	0,0	0,0
sktv. kiekin.	9	0	0	0,0	0,0
sktv. kiekin. G.	9	0	0	0,0	0,0
sktv. kiekin. In.	9	0	0	0,0	0,0
sktv. kiekin. K.	9	0	0	0,0	0,0
sktv. kiekin. N.	9	0	0	0,0	0,0
sktv. kiekin. V.	9	0	0	0,0	0,0
sktv. kiekin. Vt.	9	0	0	0,0	0,0
sktv. kuopin. G.	8	0	0	0,0	0,0
sktv. kuopin. In.	8	0	0	0,0	0,0
sktv. kuopin. K.	8	0	0	0,0	0,0
sktv. kuopin. N.	7	0	0	0,0	0,0
sktv. kuopin. V.	8	0	0	0,0	0,0
sktv. kuopin. Vt.	7	0	0	0,0	0,0
sktv. mot.gim. daugin. G.	10	0	1	0,0	10,0
sktv. mot.gim. daugin. In.	6	0	1	0,0	16,7
sktv. mot.gim. daugin. K.	10	0	0	0,0	0,0
sktv. mot.gim. daugin. N.	7	0	1	0,0	14,3
sktv. mot.gim. daugin. V.	10	0	1	0,0	10,0
sktv. mot.gim. daugin. Vt.	10	0	0	0,0	0,0
sktv. mot.gim. kelintin. dgsk. G. ivardz.	17	0	0	0,0	0,0
sktv. mot.gim. kelintin. dgsk. G. neivardz.	26	0	0	0,0	0,0
sktv. mot.gim. kelintin. dgsk. In. ivardz.	8	0	0	0,0	0,0
sktv. mot.gim. kelintin. dgsk. In. neivardz.	3	0	0	0,0	0,0

sktv. mot.gim. kelintin. dgsk. K. ivardz.	26	0	0	0,0	0,0
sktv. mot.gim. kelintin. dgsk. K. neivardz.	19	0	0	0,0	0,0
sktv. mot.gim. kelintin. dgsk. N. ivardz.	11	0	0	0,0	0,0
sktv. mot.gim. kelintin. dgsk. N. neivardz.	7	0	0	0,0	0,0
sktv. mot.gim. kelintin. dgsk. V. ivardz.	25	0	1	0,0	4,0
sktv. mot.gim. kelintin. dgsk. V. neivardz.	22	0	0	0,0	0,0
sktv. mot.gim. kelintin. dgsk. Vt. ivardz.	8	0	0	0,0	0,0
sktv. mot.gim. kelintin. dgsk. Vt. neivardz.	12	0	0	0,0	0,0
sktv. mot.gim. kelintin. vnsk. G. ivardz.	25	0	0	0,0	0,0
sktv. mot.gim. kelintin. vnsk. G. neivardz.	25	0	0	0,0	0,0
sktv. mot.gim. kelintin. vnsk. In. ivardz.	18	0	0	0,0	0,0
sktv. mot.gim. kelintin. vnsk. In. neivardz.	24	0	0	0,0	0,0
sktv. mot.gim. kelintin. vnsk. K. ivardz.	25	0	1	0,0	4,0
sktv. mot.gim. kelintin. vnsk. K. neivardz.	22	0	0	0,0	0,0
sktv. mot.gim. kelintin. vnsk. N. ivardz.	11	0	0	0,0	0,0
sktv. mot.gim. kelintin. vnsk. N. neivardz.	14	0	0	0,0	0,0
sktv. mot.gim. kelintin. vnsk. V. ivardz.	27	0	0	0,0	0,0
sktv. mot.gim. kelintin. vnsk. V. neivardz.	24	0	0	0,0	0,0
sktv. mot.gim. kelintin. vnsk. Vt. ivardz.	16	0	0	0,0	0,0
sktv. mot.gim. kelintin. vnsk. Vt. neivardz.	19	0	0	0,0	0,0
sktv. mot.gim. kiek. dgsk. G.	3	0	0	0,0	0,0
sktv. mot.gim. kiek. dgsk. In.	2	0	0	0,0	0,0
sktv. mot.gim. kiek. dgsk. K.	10	0	0	0,0	0,0
sktv. mot.gim. kiek. dgsk. N.	2	0	0	0,0	0,0
sktv. mot.gim. kiek. dgsk. V.	2	0	0	0,0	0,0
sktv. mot.gim. kiek. dgsk. Vt.	2	0	0	0,0	0,0
sktv. mot.gim. kiek. G.	10	0	1	0,0	10,0
sktv. mot.gim. kiek. In.	9	1	0	11,1	0,0
sktv. mot.gim. kiek. K.	9	0	0	0,0	0,0
sktv. mot.gim. kiek. N.	9	1	0	11,1	0,0
sktv. mot.gim. kiek. V.	10	0	1	0,0	10,0
sktv. mot.gim. kiek. vnsk. G.	7	0	0	0,0	0,0
sktv. mot.gim. kiek. vnsk. In.	2	0	0	0,0	0,0
sktv. mot.gim. kiek. vnsk. K.	17	0	0	0,0	0,0
sktv. mot.gim. kiek. vnsk. N.	10	0	0	0,0	0,0
sktv. mot.gim. kiek. vnsk. S.	1	0	0	0,0	0,0
sktv. mot.gim. kiek. vnsk. V.	3	0	0	0,0	0,0

sktv. mot.gim. kiekis. vnsk. Vt.	9	0	0	0,0	0,0
sktv. mot.gim. kiekis. Vt.	9	0	0	0,0	0,0
sktv. vyr.gim. daugin. G.	10	0	1	0,0	10,0
sktv. vyr.gim. daugin. In.	10	0	1	0,0	10,0
sktv. vyr.gim. daugin. K.	10	0	0	0,0	0,0
sktv. vyr.gim. daugin. N.	10	0	1	0,0	10,0
sktv. vyr.gim. daugin. V.	10	0	1	0,0	10,0
sktv. vyr.gim. daugin. Vt.	4	0	0	0,0	0,0
sktv. vyr.gim. kelintin. dgsk. G. ivardz.	26	0	0	0,0	0,0
sktv. vyr.gim. kelintin. dgsk. G. neivardz.	21	0	0	0,0	0,0
sktv. vyr.gim. kelintin. dgsk. In. ivardz.	26	0	0	0,0	0,0
sktv. vyr.gim. kelintin. dgsk. In. neivardz.	17	0	0	0,0	0,0
sktv. vyr.gim. kelintin. dgsk. K. ivardz.	26	0	0	0,0	0,0
sktv. vyr.gim. kelintin. dgsk. K. neivardz.	19	0	0	0,0	0,0
sktv. vyr.gim. kelintin. dgsk. N. ivardz.	5	0	0	0,0	0,0
sktv. vyr.gim. kelintin. dgsk. N. neivardz.	4	0	0	0,0	0,0
sktv. vyr.gim. kelintin. dgsk. V. ivardz.	16	0	0	0,0	0,0
sktv. vyr.gim. kelintin. dgsk. V. neivardz.	23	0	0	0,0	0,0
sktv. vyr.gim. kelintin. dgsk. Vt. ivardz.	8	0	0	0,0	0,0
sktv. vyr.gim. kelintin. dgsk. Vt. neivardz.	5	0	0	0,0	0,0
sktv. vyr.gim. kelintin. vnsk. G. ivardz.	29	0	0	0,0	0,0
sktv. vyr.gim. kelintin. vnsk. G. neivardz.	25	0	0	0,0	0,0
sktv. vyr.gim. kelintin. vnsk. In. ivardz.	21	0	0	0,0	0,0
sktv. vyr.gim. kelintin. vnsk. In. neivardz.	18	0	0	0,0	0,0
sktv. vyr.gim. kelintin. vnsk. K. ivardz.	28	0	0	0,0	0,0
sktv. vyr.gim. kelintin. vnsk. K. neivardz.	24	0	0	0,0	0,0
sktv. vyr.gim. kelintin. vnsk. N. ivardz.	15	0	12	0,0	80,0
sktv. vyr.gim. kelintin. vnsk. N. neivardz.	15	0	0	0,0	0,0
sktv. vyr.gim. kelintin. vnsk. S. neivardz.	1	0	0	0,0	0,0
sktv. vyr.gim. kelintin. vnsk. V. ivardz.	30	0	1	0,0	3,3
sktv. vyr.gim. kelintin. vnsk. V. neivardz.	26	0	0	0,0	0,0
sktv. vyr.gim. kelintin. vnsk. Vt. ivardz.	21	0	0	0,0	0,0
sktv. vyr.gim. kelintin. vnsk. Vt. neivardz.	20	0	0	0,0	0,0

sktv. vyr.gim. kiekis. dgsk. G.	4	0	0	0,0	0,0
sktv. vyr.gim. kiekis. dgsk. In.	3	0	0	0,0	0,0
sktv. vyr.gim. kiekis. dgsk. K.	12	0	0	0,0	0,0
sktv. vyr.gim. kiekis. dgsk. N.	3	0	0	0,0	0,0
sktv. vyr.gim. kiekis. dgsk. V.	4	0	0	0,0	0,0
sktv. vyr.gim. kiekis. dgsk. Vt.	3	0	0	0,0	0,0
sktv. vyr.gim. kiekis. G.	16	0	1	0,0	6,3
sktv. vyr.gim. kiekis. In.	13	1	0	7,7	0,0
sktv. vyr.gim. kiekis. K.	15	0	0	0,0	0,0
sktv. vyr.gim. kiekis. N.	14	1	0	7,1	0,0
sktv. vyr.gim. kiekis. V.	16	0	1	0,0	6,3
sktv. vyr.gim. kiekis. vnsk. G.	3	0	0	0,0	0,0
sktv. vyr.gim. kiekis. vnsk. In.	3	0	0	0,0	0,0
sktv. vyr.gim. kiekis. vnsk. K.	11	0	0	0,0	0,0
sktv. vyr.gim. kiekis. vnsk. N.	3	0	0	0,0	0,0
sktv. vyr.gim. kiekis. vnsk. V.	3	0	0	0,0	0,0
sktv. vyr.gim. kiekis. vnsk. Vt.	3	0	0	0,0	0,0
sktv. vyr.gim. kiekis. Vt.	11	0	0	0,0	0,0
Iš viso	1410	5	27	0,4	1,9

## 2 PRIEDAS

### **PARENGTOS PUBLIKACIJOS**

Darbo pagrindu parengta ir tarptautinėje recenzuojamoje konferencijoje užsienyje išspausdinta publikacija:

Norkevičius G., Raškinis G., Kazlauskienė A., Knowledge-based grapheme-to-phoneme conversion of Lithuanian words // In Proceedings of the 10th International Conference on Speech and Computer - Specom. Patras, Greece, 2005, p. 235-238.

LIETUVOS VALSTYBINIS MOKSLO IR STUDIJŲ FONDAS  
PROJEKTO IŠLAIDŲ SĄMATA

Projekto pavadinimas **Lietuvių kalbos teksto kirčiavimas internete**

Atsakingoji mokslo ir studijų institucija **Vytauto Didžiojo universitetas**

Projekto trukmė 2005 m 05 mėn. – 2005 m 12 mėn.

Eil. Nr.	Išlaidų pavadinimas	2005 m.		Visam projektui	
		Lt	%	Lt	%
1.	Darbo užmokestis	6867,75	68,7	6867,75	68,7
2.	Socialinio draudimo įmokos	2129,01	21,3	2129,01	21,3
3.	Išlaidos prekėms įsigyti	998,8	9,9	998,8	9,9
4.	Komandiruočių išlaidos				
5.	Kitos išlaidos (tarp jų autorinis atlyginimas)				
6.	Ilgalaikio materialiojo turto įsigijimo išlaidos (laboratorinė įranga ir kt.)				
7.	Nematerialiojo turto įsigijimo išlaidos (programinė įranga ir kt.)				
8.	Išlaidos paslaugoms				
	<b>Iš viso:</b>	9995,56	99,9	9995,56	99,9

**Įsigytos prekės:**

1. Kompiuterio komponentas: DDR 512Mb PC3200 400MHz PQI – 1 vnt
2. Kompiuterio komponentas: HDD WD 1600JD SerialATA 160GB 7200rpm 8Mb – 2 vnt.



**LIETUVOS VALSTYBINIS MOKSLO IR STUDIJŲ FONDAS  
PROJEKTO MOKSLINIO EKONOMINIO PRODUKTYVUMO SUVESTINĖ**

Projekto pavadinimas **Lietuvių kalbos teksto kirčiavimas internete**  
 Projekto vadovas **Dr. Gailius Raškinis**  
 Atsakingoji mokslo ir studijų institucija **Vytauto Didžiojo universitetas**  
 Registracijos Nr. **T-05267**  
 Sutarties Nr. **T-41/05**  
 Sutarties terminai 2005 m. 05 mėn. 19 d. – 2005 m. 12 mėn. 10 d.  
 Mokslo kryptis **Informatika**  
 Mokslo sritis **Fiziniai mokslai**

Eil. Nr.	Produktyvumo kriterijai	Rezultatai (skaičius)
1	Naujos technologijos (įdiegtos)	1
2	Nauji gaminiai (aprobuoti)	
3	Įregistruota patentų:	
	3.1. Lietuvoje	
	3.2. Užsienyje	
4	Paskelbta mokslinių straipsnių:	1
	4.1. Recenzuojamuose Lietuvos leidiniuose	
	4.2. Užsienio mokslo leidiniuose, įtrauktuose į Mokslinės informacijos instituto pagrindinių žurnalų sąrašą	
	4.3. Kituose mokslo leidiniuose	1
	4.4. Įteikta straipsnių	
5	5.1. Doktorantai – projekto dalyviai	1
	5.2. Apginta disertacijų	
6	Mokslinės monografijos ir kiti mokslo veikalai:	
	6.1. Išleista	
	6.2. Įteikta spaudai	
7	Vadovėliai:	
	7.1. Išleista	
	7.2. Įteikta spaudai	
8	Mokslo renginiai	
9	Naujų darbo vietų skaičius	
10	Apyvartos padidėjimas (%)	

Užpildymo data 2005-12-10