

VYTAUTO DIDŽIOJO UNIVERSITETAS  
MATEMATIKOS IR INFORMATIKOS INSTITUTAS

**Daiva Šveikauskienė**

**LIETUVIŲ KALBOS VIENTISINIŲ SAKINIŲ  
AUTOMATINĖ SINTAKSINĖ ANALIZĖ**

Daktaro disertacija

Fiziniai mokslai (P 000)

Informatika (09 P)

Dirbtinis intelektas (P 176)

Vilnius, 2009

Disertacija rengta 2000-2009 m. Matematikos ir informatikos institute  
Disertacija ginama eksternu

Mokslinis konsultantas:

prof. habil. dr. Laimutis Telksnys (Matematikos ir informatikos institutas, fiziniai mokslai,  
informatika – 09 P)

## **Padėka**

Nuoširdžiai dėkoju Vytautui Zinkevičiui už suteiktą galimybę naudotis jo sukurta lietuvių kalbos morfologinių modelių programine įranga.

Taip pat noriu labai padėkoti dr. Elenai Valiulytei už begalines konsultacijas lietuvių kalbos klausimais ir už lituanistinių klaidų ištaisymą visuose mano spausdintuose darbuose.

## Santrauka

Darbas priklauso dirbtinio intelekto sričiai, jame nagrinėjamas vienas žmogaus protinio darbo automatizavimo uždavinių – lietuvių kalbos automatinės sintaksinės analizės sukūrimas.

Dėl didelių skirtumų tarp lietuvių kalbos ir kitų indoeuropiečių kalbų negalima tiesiogiai pasinaudoti jau sukurta kitose šalyse automatinės sintaksinės analizės programine įranga ir būtina sukurti naują savitą metodą, kuris gerai atspindėtų specifinius lietuvių kalbos bruožus – didelį kaitomumą ir laisvą žodžių tvarką sakinyje.

Darbe apžvelgtos trijų kalbų – anglų, vokiečių ir rusų – sintaksinės analizės metodikos. Visos šios kalbos priklauso tai pačiai kalbų grupei (indoeuropiečių), kaip ir lietuvių kalba, ir skiriasi viena nuo kitos kaitomumo laipsniu bei žodžių tvarkos sakinyje laisvumu. Pagrindinis kriterijus, į kurį atsižvelgiama atliekant sintaksinę analizę anglų ir vokiečių kalbose, yra žodžio vieta sakinyje, nes šiose kalbose beveik tik nuo jos ir priklauso, kokią sintaksinę funkciją atlieka žodis. Lietuvių kalboje nesant griežtos, sugramatintos žodžių tvarkos didžiausias sintaksinės informacijos kiekis sukaupytas žodžių formose (jų galūnėse). Anglų bei vokiečių kalboms sukurtose sintaksinės analizės sistemose **nenumatytas sintaksinės informacijos paėmimas iš žodžių galūnių**. Rusų kalba artimesnė lietuvių kalbai kaitomumo požiūriu, bet rusų kalbos automatinės sintaksinės analizės sistemose nesinaudojama formaliu sintaksės aprašu, kaip anglų ir vokiečių kalbose ir analizę čia atlieka grupės algoritmu. Tačiau, jei ruošiant programinę įrangą, sprendžiamą uždavinį pavyksta aprašyti II tipo formalia gramatika (pagal Chomskio klasifikaciją), labai supaprastėja programavimas. Todėl šiame darbe buvo siekiama lietuvių kalbos sintaksę aprašyti nekontekstinės gramatikos taisyklėmis. Visoms sakinio dalims sudarytas aprašas BNF (Bekaus ir Nauro forma), nurodantis kokios žodžio morfologinės formos gali atlikti kiekvieną sintaksinę funkciją. Laisvai žodžių tvarkai lietuvių kalbos sakiniuose įvertinti BNF pateikta informacija apie tai, kokie žodžiai gali būti susiję tiesioginiu sintaksiniu ryšiu ir kokie kiti žodžiai gali būti tarp jų įsiterpę.

Nauja lietuvių kalbos automatinėje sintaksinėje analizėje yra tai, kad visos trys gramatikos sritys – morfologija, sintaksė ir semantika – sujungiamos į vieną visumą. Kitų kalbų sintaksinės analizės sistemose darbo rezultatų pagerinimui naudojama semantika (informacija apie žodžio reikšmę), morfologinių duomenų apie žodį šiam tikslui nenaudoja niekas.

Darbe pagrindžiama būtinybė lietuvių kalbos sintaksinei struktūrai naudoti priklausomybių grafą, nes kitoms kalboms naudojamas priklausomybių medis negali atspindėti visos sintaksinės informacijos, esančios lietuviškame sakinyje. Pateikiamas lietuvių kalbos sakinių apibendrinta sintaksinė struktūra, išreikšta grafu, kuris apima bet kokią vientisinį lietuvių kalbos sakinį. Kiekvienas konkretus sakinytis turi aktyvuoti tame grafe vieną kelią.

Aprašomas lietuvių kalbos sakinių sintaksinės analizės metodas, algoritmas bei programinė įranga. Programinės įrangos veikimas patikrintas dviem eksperimentais: pirmo etapo metu testui buvo naudojama 670 sakinių, paimtų iš įvairių rūšių tekstų (tekstyno ir spausdintos literatūros) Gauti tokie rezultatai: 629 sakiniai (93,88%) išnagrinėti teisingai. Antro etapo metu testas buvo atliekamas su 400 eksperto parinktų sakinių iš 8 skirtingų tekstyno dalių ir gautas tikslumas 92%. Pagrindinės klaidų priežastys – sutampančios žodžių formos ir semantinės informacijos trūkumas (automatinė semantinė analizė lietuvių kalbai kol kas dar nesukurta). Testui naudoti sakiniai ir analizės rezultatai pateikiami kompaktinėje plokštelėje, pridedamoje prie disertacijos.

## Turinys

<b>1</b>	<b>IŽANGA</b> .....	1
<b>2</b>	<b>ANALITINĖ APŽVALGA</b> .....	5
2.1	Įvadas.....	5
2.1.1	Kalbininkų darbai.....	5
2.1.2	Informatikų darbai.....	6
2.2	Automatinė sintaksinė analizė.....	8
2.2.1	Anglų kalbos sintaksinė analizė.....	10
2.2.2	Vokiečių kalbos sintaksinė analizė.....	20
2.2.3	Rusų kalbos sintaksinė analizė.....	25
2.3	Statistiniai metodai.....	29
2.4	Lietuvoje atlikti darbai.....	30
2.5	Skyriaus išvados.....	32
<b>3</b>	<b>KAI KURIOS LIETUVIŲ KALBOS YPATYBĖS</b> .....	33
<b>4</b>	<b>GRAFO PANAUDOJIMAS SAKINIO STRUKTŪRAI</b> .....	36
4.1	Tarininio pažyminio problema.....	36
4.2	Būtinumas panaudoti grafą.....	37
4.3	Apibendrinta lietuvių kalbos sakinio sintaksinė struktūra.....	39
4.3.1	Frazių metodo trūkumai.....	40
4.3.2	Priklausomybių gramatikos trūkumai.....	41
4.3.3	Apibendrintos sakinio struktūros formavimas.....	43
4.4	Skyriaus išvados.....	48
<b>5</b>	<b>SINTAKSĖS TAISYKLIŲ UŽRAŠYMAS BNF</b> .....	49
5.1	Grafo viršūnės.....	51
5.1.1	Veiksnių aprašas.....	52
5.1.2	Tarinio aprašas.....	54
5.1.3	Pažyminio aprašas.....	59
5.1.4	Papildinio aprašas.....	66
5.1.5	Aplinkybės aprašas.....	70
5.2	Grafo lankai.....	73
5.2.1	Laisvos žodžių tvarkos įvertinimas.....	77
5.2.2	Intarpo vaidmuo formalioje gramatikoje.....	78
5.2.3	Morfologinio daugiareikšmiškumo panaikinimas.....	80
5.3	Skyriaus išvados.....	83
<b>6</b>	<b>SINTAKSINĖS ANALIZĖS ALGORITMAS</b> .....	84
6.1	Sintaksinės analizės metodika.....	85
6.2	Algoritmo blokinė schema.....	86
6.3	Algoritmo iliustracija pavyzdžiu morfologiniam daugiareikšmiškumui panaikinti.....	86
6.4	Skyriaus išvados.....	90
<b>7</b>	<b>SINTAKSINĖS ANALIZĖS ALGORITMO PROGRAMINĖ ĮRANGA</b> .....	91
7.1	Programinės įrangos blokinė schema.....	92
7.2	Iteracinio metodo panaudojimas.....	99
7.3	Skyriaus išvados.....	101
<b>8</b>	<b>SINTAKSINĖS ANALIZĖS EKSPERIMENTAS</b> .....	102
8.1	Ekspерименто pirmasis etapas.....	102
8.1.1	Testų tipai.....	103
8.1.2	Ekspерименто pirmo etapo rezultatai.....	104
8.2	Ekspерименто antrasis etapas.....	108
8.2.1	Morfologinių duomenų patikslinimas.....	110
8.2.2	Gauti rezultatai.....	112
8.3	Skyriaus išvados.....	118
<b>9</b>	<b>SINTAKSINĖS ANALIZĖS TAIKYMAS</b> .....	119
9.1	Automatinis vertimas.....	119
9.2	Ištisinės šnekos atpažinimas.....	120
9.3	Klaidų suradimas tekstyne.....	122
	<b>BENDROSIOS IŠVADOS</b> .....	124
	<b>LITERATŪROS SĄRAŠAS</b> .....	125

## Paveikslėlių sąrašas

1 pav.	Automatinio vertimo etapai, naudojant keitimo metodiką .....	8
2 pav.	Žodžio <i>like</i> struktūrinės schemos pakeitimas žodžio <i>patikti</i> struktūrine schema .....	9
3 pav.	Gramatikos taisyklių rinkinys, su kuriuo galima išnagrinėti sakinį <i>John eats the cat</i> .....	10
4 pav.	Sakinio <i>John eats the cat</i> sintaksinė struktūra.....	10
5 pav.	Sakinio <i>John eats the cat</i> nagrinėjimo „iš viršaus į apačią“ žingsniai .....	11
6 pav.	Sakinio <i>John eats the cat</i> nagrinėjimo „iš apačios į viršų“ žingsniai .....	11
7 pav.	Formali gramatika, kuri nėra nepriekaištinga leistinių sakinių atžvilgiu .....	12
8 pav.	Gramatikos taisyklių rinkinys sakiniui <i>The blue can can hold the water</i> išnagrinėti.....	18
9 pav.	Sakinio <i>The blue can can hold the water</i> analizės diagrama .....	18
10 pav.	Gramatikos taisyklės, pagal kurias nagrinėjamas sakinytis <i>The green water evaporated</i> .....	19
11 pav.	Sakinio <i>The green water</i> sintaksinė struktūra .....	19
12 pav.	Sakinio <i>The green water evaporated</i> struktūrinė schema .....	20
13 pav.	Sakinio <i>John eats the cat</i> sintaksinė struktūra priklausomybių gramatikoje .....	20
14 pav.	Priklausomybių medžio viršūnės ryšiai su kitomis viršūnėmis .....	21
15 pav.	Sakinio <i>Der Robot hebt einen großen roten Block</i> sintaksiniai ryšiai DUG .....	21
16 pav.	Sakinio <i>Der Robot hebt einen großen roten Block</i> aprašas DUG .....	22
17 pav.	Veiksmažodžio <i>heben</i> sintaksinius ryšius atspindinčių šablonų rinkinys .....	23
18 pav.	Alternatyvus veiksnio šablonas .....	23
19 pav.	Žodžių <i>heben</i> ir <i>Robot</i> sintaksiniai rėmai.....	24
20 pav.	Pradinis sakinio sąrašas, gautas atliekant <i>Der Robot hebt einen großen roten Block</i> sintaksinę analizę.....	24
21 pav.	Taisyklė būdvardis-daiktavardis.....	26
22 pav.	Sakinio <i>В статье описаны методы анализа, его процедуры и полученные результаты</i> struktūra.....	27
23 pav.	Sakinio <i>Создаётся информационная система для анализа предложения и перевода текстов</i> struktūra .....	28
24 pav.	Žodyno įrašo pavyzdys .....	31
25 pav.	Semantinės informacijos kodavimo pavyzdys .....	31
26 pav.	Vokiško sakinio <i>Der Vater kam gestern verärgert</i> sintaksinė struktūra .....	38
27 pav.	Lietuviško sakinio <i>Tėvas vakar grįžo piktas</i> sintaksinė struktūra .....	38
28 pav.	Vokiško sakinio <i>Die Mutter aß die Mohrrüben roh</i> sintaksinė struktūra .....	38
29 pav.	Lietuviško sakinio <i>Motina morkas valgė žalias</i> sintaksinė struktūra .....	38
30 pav.	Sakinio <i>Mergina praėjo pro šalį su auskarais</i> struktūra, pavaizduota frazių metodu, traktuojant, kad mergina praėjo pro šalį, kuri buvo su auskarais, t.y. šalis buvo su auskarais.....	40
31 pav.	Sakinio <i>Mergina praėjo pro šalį su auskarais</i> struktūra, pavaizduota frazių metodu, traktuojant, kad mergina praėjo su auskarais, t.y. ji pro šalį ėjo su auskarais, taip kaip, pvz., važiuo su mašina .....	40
32 pav.	Sakinio su vienerūšiais tariniais schema .....	41
33 pav.	Apibendrinta lietuvių kalbos sakinio struktūrinė schema.....	44
34 pav.	Sakinio <i>Jis (buvo) geras mokytojas</i> sintaksinė struktūra .....	45
35 pav.	Sakinio <i>Ant aukšto statau kalno pasirodė stebuklingas žiburys</i> sintaksinė struktūra .....	45
36 pav.	Sakinio <i>Vakar aš mačiau savo motulę verkiančią</i> sintaksinė struktūra .....	46
37 pav.	Sakinio <i>Motina viską pastebi ir prašo negąsdinti vaiko ir neįvaryti jam baimės</i> sintaksinė struktūra .....	46
38 pav.	Sakinio <i>Įkaitęs oras ribuliuoja ir mirga virš laukų</i> sintaksinė struktūra .....	47
39 pav.	Formali gramatika reiškinių $a+b*c/a-d$ reikšmei apskaičiuoti.....	51
40 pav.	Medis, pagal kurį skaičiuojama reiškinių $a+b*c/a-d$ reikšmė.....	51
41 pav.	Gramatika, neturinti papildomų neterminalinių simbolių .....	52
42 pav.	Reiškinių $a+b*c/a-d$ medis, sudarytas pagal gramatiką, neturinčią papildomų neterminalinių simbolių.....	52
43 pav.	Sakinio <i>Aš girdėjau netoliese žmones kalbant</i> sintaksinė struktūra .....	67
44 pav.	Derinamojo pažyminio įsiterpimas į žodžių junginį <i>mano namas</i> .....	77
45 pav.	Nederinamojo pažyminio įsiterpimas į žodžių junginį <i>mano namas</i> .....	77
46 pav.	Naujų žodžių junginių susiformavimas .....	77
47 pav.	Intarpas, išreikštas derinamojo pažyminio gija .....	78
48 pav.	BNF taisyklių rinkinio pavyzdys .....	79
49 pav.	Sintaksiškai išnagrinėto sakinio pavyzdys .....	79
50 pav.	Sakinio <i>Tamsūs pušų sakai blizgėjo saulėje</i> sintaksinė struktūra.....	80
51 pav.	Sakinio <i>Per tyliai sakai tuos reikšmingus žodžius</i> sintaksinė struktūra.....	80
52 pav.	Sintaksiniai ryšiai tarp žodžių, pavaizduoti linijinėje sakinio struktūroje .....	80
53 pav.	Veiksnių suradimo kelias sakinyje <i>Tamsūs pušų sakai blizgėjo saulėje</i> .....	81
54 pav.	Sakinio <i>Per tyliai sakai tuos reikšmingus žodžius</i> sintaksinė struktūra.....	82
55 pav.	Apibendrintas sakinio vaizdas formalios gramatikos požiūriu .....	84
56 pav.	Algoritmo blokinė schema .....	86
57 pav.	Segmentavimo ir morfologinės analizės iliustracija pavyzdžiu .....	87
58 pav.	Morfologinių sakinių variantai .....	87
59 pav.	Sintaksinės analizės iliustracija pavyzdžiu.....	87

<b>60 pav.</b>	Sintaksinių sakinių variantai.	88
<b>61 pav.</b>	Tiesioginių sintaksinių ryšių nustatymas.	89
<b>62 pav.</b>	Pradinio lango grafinis vaizdas.	91
<b>63 pav.</b>	Programinės įrangos blokinė schema.	94
<b>64 pav.</b>	Pranešimo kortelė, informuojanti kad sakinio struktūros sudaryti nepavyko.	96
<b>65 pav.</b>	Sakinio struktūros grafinio vaizdo kontūrai.	97
<b>66 pav.</b>	Sakinio struktūros grafinio vaizdo kontūrai antrame lange.	98
<b>67 pav.</b>	Išnagrinėto sakinio pavyzdys.	98
<b>68 pav.</b>	Sakinio struktūra, gauta naudojant pirmą morfologinės informacijos variantą.	99
<b>69 pav.</b>	Sakinio struktūra, atstačius teisingą veiksni.	99
<b>70 pav.</b>	Sakinio struktūra, nustačius veiksnio pažyminį.	100
<b>71 pav.</b>	Sakinio struktūra, suradus veiksnio pažyminio išplėtimą.	100
<b>72 pav.</b>	Galutinė teisingai išnagrinėto sakinio struktūra.	100
<b>73 pav.</b>	Programinei įrangai taikomų apribojimų sąrašas.	103
<b>74 pav.</b>	Tekstyno dalių sąrašas kiekvienai iš pirmų dviejų imčių.	103
<b>75 pav.</b>	Klaida, atsirandanti dėl semantinės informacijos trūkumo.	105
<b>76 pav.</b>	Klaida, atsirandanti dėl neoptimalaus morfologinių duomenų pateikimo.	105
<b>77 pav.</b>	Klaida, atsirandanti dėl sintaksinės analizės programos kaltės.	106
<b>78 pav.</b>	Sakinys, kuriame semantinis laiko požymis padeda atskirti papildinį nuo aplinkybės.	107
<b>79 pav.</b>	Sakinys, kuriame atsiranda klaida dėl morfologinių duomenų neoptimalaus pateikimo.	107
<b>80 pav.</b>	Tekstyno apdoravimo pavyzdys.	109
<b>81 pav.</b>	Žodžio <i>susitikimas</i> morfologiniai duomenys.	110
<b>82 pav.</b>	Žodžio <i>stovi</i> morfologiniai duomenys.	110
<b>83 pav.</b>	Žodžių, kuriems nepateikta išsami informacija pavyzdžiai.	111
<b>84 pav.</b>	Teisingai išnagrinėto sakinio pavyzdys (37 sakiny iš tekstyno Nr. 1107).	114
<b>85 pav.</b>	Teisingai išnagrinėto sakinio pavyzdys (6 sakiny iš tekstyno Nr. 10012).	114
<b>86 pav.</b>	Teisingai išnagrinėto sakinio pavyzdys (13 sakiny iš tekstyno Nr. 4032).	114
<b>87 pav.</b>	Teisingai išnagrinėto sakinio pavyzdys (22 sakiny iš tekstyno Nr. 10012).	115
<b>88 pav.</b>	Klaidingai išnagrinėtas sakiny (50 sakiny iš tekstyno Nr. 1107).	115
<b>89 pav.</b>	Klaidingai išnagrinėtas sakiny (29 sakiny iš tekstyno Nr. 1107).	115
<b>90 pav.</b>	Klaidingai išnagrinėtas sakiny (31 sakiny iš tekstyno Nr. 1107).	116
<b>91 pav.</b>	Klaidingai išnagrinėtas sakiny (23 sakiny iš tekstyno Nr. 2038).	116
<b>92 pav.</b>	Praleisto veiksnio atstatymas.	119
<b>93 pav.</b>	Ištisinės šnekamosios kalbos atpažinimo apibendrinta blokinė schema.	121
<b>94 pav.</b>	Sakinio su korektūros klaida analizė.	122
<b>95 pav.</b>	Sakinio su ištaisyta klaida struktūra.	123

## Lentelių sąrašas

<b>1 lentelė</b>	Sintaksinės informacijos pasiskirstymas įvairiose kalbose.	6
<b>2 lentelė</b>	Galimų sakinio struktūros variantų skaičiaus priklausomybė nuo pavartotų prielinksninių konstrukcijų kiekio.	49
<b>3 lentelė</b>	Linksniai, kurie traktuojami, kaip veiksmožodžių semantiniai požymiai.	55
<b>4 lentelė</b>	Linksniai su prielinksniais, kurie traktuojami, kaip veiksmožodžių semantiniai požymiai.	56
<b>5 lentelė</b>	Prielinksninės konstrukcijos, galinčios eiti aplinkybe.	71
<b>6 lentelė</b>	Masyvo elementuose (x-kryptimi) saugomos informacijos apie žodį pobūdis.	93
<b>7 lentelė</b>	Testų rezultatai.	108
<b>8 lentelė</b>	Vientisinių sakinių dalis rišliame lietuvių kalbos tekste.	109
<b>9 lentelė</b>	Automatinės sintaksinės analizės tikslumas.	113
<b>10 lentelė</b>	Klaidų tipai.	113

## Sutrupinimai

Sutrupinimas	Sutrupinimo iššifravimas	Vertimas į lietuvių kalbą
APLINKYB	Aplinkybė	
ART	Artikelis	
BENDR	Bendratis	
BENDRAT	Bendratis	
BEVG	bevardė giminė	
BNF	Bekaus ir Nauro forma	
BS	bazinė struktūra	
BŪDV	Būdvardis	
DAIKT	Daiktavardis	
DALYV	Dalyvis	
DERIN	Derinamasis	
DET	determinantas	
DF	daiktavardinė frazė	
DGS	daugiskaita	
DLL	dynamic link library	dinaminių saitų biblioteka
DUG	dependency unification grammar	unifikuota priklausomybių gramatika
GAL	galininkas	
Geta	Groupe d'Etudes pour la Traduction Automatique	automatinio vertimo tyrimų grupė
INAG	inagininkas	
IVARD	įvardis	
JUNGT	jungtis	
KILM	kilmininkas	
MODALV	modalinis veiksmažodis	
MOTG	moteriškoji giminė	
NAUD	naudininkas	
NEDERIN	nederinamasis	
PAGV	pagalbinis veiksmažodis	
PAPILD	papildinys	
PAŽYM	pažyminys	
PF	prielinksninė frazė	
PRIEV	prieveiksmis	
SKAITV	skaitvardis	
TARIN	tarinys	
VARD	vardininkas	
TDAIKT	tikrinis daiktavardis	
VEIKSM	veiksmažodis	
VEIKSN	veiksny	
VF	veiksmažodinė frazė	
VIET	vietininkas	
VYRG	vyriškoji giminė	
VNS	vienaskaita	



# 1 ĮŽANGA

## Tyrimų objektas

Lietuvių kalbos vientisinių sakinių sintaksinės analizės automatizavimas.

## Temos aktualumas

Daugeliui pasaulio kalbų jau yra sukurtos automatinės sintaksinės analizės sistemos. Lietuvių kalba tokios sistemos kol kas dar neturi ir šio darbo tikslas būtų užpildyti tą spragą.

Pagrindinė automatinės sintaksinės analizės pritaikymo sritis būtų automatinis vertimas. Kol kas patenkinamai dirbančios automatinio vertimo sistemos lietuvių kalbai dar neturime.

Automatinio vertimo sistemose keitimas tarp kalbų vyksta dviem lygmenimis: leksiniu ir sintaksiniu. Leksiniame lygmenyje vienos kalbos žodžiai keičiami kitos kalbos žodžiais, o sintaksiniame lygmenyje keičiamos sakinių sintaksinės struktūros. Skirtingose kalbose to paties sakinio struktūros kartais būna nevienodos, pvz., sakinyje *He likes this book* žodis *he* yra veiksnys, o žodis *book* – papildinys. Lietuviškame šio sakinio vertime *Jam patinka ši knyga* žodis *jam* jau atlieka papildinio funkciją, o žodis *knyga* yra veiksnys.

Tačiau, norint atlikti vienos kalbos sintaksinės struktūros pakeitimą kitos kalbos sintaksine struktūra, visų pirma reikia tą struktūrą turėti kompiuteryje. Tai reiškia, kad reikia turėti formalizuotą sintaksę, nes kompiuteris negali pats pasiskaityti lituanistų išleistų knygų šia tema ir iš ten pasiimti informacijos apie sakinio dalis. Tai jam turi būti pateikta formalizuotai. Taigi, toks ir būtų šio darbo tikslas: pasakyti kompiuteriui, jam suprantama kalba, kaip nustatomos sakinio dalys.

## Darbo tikslas

Sukurti sistemą, kuri galėtų atlikti lietuvių kalbos vientisinių sakinių automatinę sintaksinę analizę pasitelkiant šiam tikslui kompiuterių galimybes, t.y. paruošti metodiką bei jos programinę realizaciją vientisinio sakinio sintaksinei struktūrai sudaryti kompiuteriu.

## Uždaviniai

1. Išanalizuoti jau sukurtas pasaulyje sintaksinės analizės sistemas ir išsiaiškinti jų pritaikymo galimybes lietuvių kalbai.
2. Išryškinti specifines lietuvių kalbos savybes ir sukurti metodą lietuvių kalbos sakinių automatinei sintaksinei analizei atlikti.
3. Sukurti formalią gramatiką lietuvių kalbos sintaksei aprašyti, pateikiant ją BNF (Bekaus ir Nauro forma).
4. Sudaryti algoritmą bei parengti programinę įrangą, nustatančią žodžių sintaksines funkcijas lietuvių kalbos vientisiniame sakinyje.
5. Pasinaudojant informacija apie žodžių sintaksines funkcijas, pateikti išnagrinėto sakinio sintaksinės struktūros grafinį vaizdą kompiuterio ekrane.
6. Sukurtos sistemos darbą išbandyti su rinkiniu sakinių.

## Tyrimo metodai

Teoriniai tyrimai atlikti pasitelkiant kompiuterinės lingvistikos, automatinės sintaksinės analizės, lituanistikos, algoritmų teorijos, programavimo kalbų teorijos ir programavimo žinias bei metodus.

Eksperimentiniai tyrimai padaryti naudojantis specialiai šiam tikslui sukurta programine įranga, parašyta Visual Basic'6 programavimo kalba, Vytauto Zinkevičiaus sukurta lietuvių kalbos morfologinių modelių programine įranga bei lietuvių kalbos tekstynu.

## Mokslinis naujumas

Pirmą kartą sakinio sintaksinei struktūrai pavaizduoti naudojamas priklausomybių grafai. Kitų kalbų automatinio vertimo sistemose šios struktūros vaizduojamos priklausomybių medžiu. Darbe parodoma, kad, naudojant medį lietuvių kalbai, prarandama informacija, nes medis iš principo negali atspindėti visų sintaksinių ryšių, esančių lietuviškame sakinyje.

Sudaryta apibendrinta lietuvių kalbos sakinių sintaksinė struktūra, apimanti bet kokią vientisinį sakinį, t.y. apibrėžianti visus leistinus lietuvių kalbos vientisinio sakinio atvejus. Analogiškos struktūros kitų tautų kalboms literatūroje nepateikiamos.

Atsižvelgiant į didelį lietuvių kalbos kaitomumą, sistemos darbui pagerinti žodžių sintaksinės funkcijos diferencijuojamos pagal morfologines kategorijas. Toks atvejis nėra žinomas nė vienoje iš jau sukurtų automatinio vertimo sistemų. Net rusų kalbos sintaksinė analizė neskaido sakinio dalių pagal morfologinius požymius, nors rusų kalba savo kaitomumo laipsniu yra gana artima lietuvių kalbai. Visos iki šiol sukurtos sistemos sintaksinės analizės rezultatus bando pagerinti pasitelkdamos semantinę (žodžio reikšmės) informaciją.

Įvertinant kitą specifinį lietuvių kalbos bruožą – laisvą žodžių tvarką sakinyje, įvedamas formalus parametras GIIA, kuris fiksuoja ryšius tarp žodžių, nurodydamas jų tarpusavio padėtį. Gijos aprašas BNF susideda iš trijų dalių: pirmoje ir trečioje pozicijoje nurodomi žodžiai (tiksliau sakinio dalys), tarp kurių ieškomas sintaksinis ryšys, o viduryje parašomas neterminalinis simbolis INTARPAS, kuris atspindi informaciją apie tai, kas gali būti įsiterpę tarp Gijos žodžių. Tokiu principu žodžių tvarka neapdorojama nė vienoje iš jau sukurtų automatinio vertimo sistemų.

## Praktinė reikšmė ir darbo realizavimas praktikoje

Lietuvių kalbos sakinių automatinė sintaksinė analizė atveria galimybes sukurti automatinio vertimo sistemas tarp lietuvių kalbos ir anglų, vokiečių bei rusų kalbų.

Ištisinės šnekos atpažinimo sistemose taip pat naudojama sintaksinė informacija, todėl šiam tikslui taip pat galės būti panaudota automatinė sintaksinė lietuvių kalbos sakinių analizė.

Informacija apie žodžių sintaksines funkcijas sakinyje, pateikta kompiuteriniam apdorojimui tinkamu būdu, gali padėti panaikinti lietuvių kalbos morfologijos modelių programinės įrangos darbo rezultatų daugiareikšmiškumą. Pvz., pateikus trijų simbolių seką *mes* gaunami du variantai: 1) asmeninis įvardis, daugiskaitos vardininkas; 2) veiksmažodžio *mesti* būsimasis laikas, 3-as asmuo. Atlikus sintaksinę sakinio analizę ir nustatčius, kad nagrinėjamas žodis yra veiksnys, atkris antras variantas, nes veiksmažodžio forma (ne bendratis) negali būti veiksniu, o jei bus nustatyta, kad tai yra tarinys, bus atmetamas pirmas variantas, nes asmeninio įvardžio vardininkas niekada negali eiti vientisiniu tariniu. Taigi, po sintaksinės analizės bus paliekamas tik vienas galimas variantas.

Pasinaudojant sukurta programine įranga galima pagerinti lietuvių kalbos tikrintuvo darbą tais atvejais, kai neteisingai pavartojama šiaip jau leistina lietuvių kalboje žodžio forma. Pvz., jei žodis *būtų* parašomas su trumpąja *u*, dabartiniai tikrintuvai nerodo klaidos, nes yra toks lietuviškas žodis – žodžio *butas* daugiskaitos kilmininko (daug ko? – būtų) linksnis.

Žodis *butų* gali eiti tik pažyminiu ar papildiniu ir, jei sintaksinės analizės rezultatu duomenimis tai yra tarinys, fiksuojama klaida.

Dar viena sritis, kur gali būti pritaikoma automatinė sintaksinė analizė – tai klaidų suradimas tekstyne. Jei sintaksinės analizės metu negaunama sakinio struktūra, viena iš priežasčių gali būti korektūros klaidos sakinyje. Negavusi teigiamo analizės rezultato programinė įranga pateikia pranešimą „Sakinio struktūros nepavyko sudaryti“, ir tuo atkreipia dėmesį į tą sakinį, kuriame gali būti klaida.

## Darbo aprobavimas

Tarpiniai disertacijos rezultatai buvo pristatyti:

- Seminare Vytauto Didžiojo universiteto Kompiuterinės lingvistikos centre, ir kasmetiniuose Atpažinimo procesų skyriaus seminaruose Matematikos ir informatikos institute,
- Publikacijose: žurnaluose “Informatica”, “Information Technology and Control”

## Publikacijos

Pagrindiniai disertacijos darbo teiginiai ir rezultatai paskelbti 3 publikacijose, iš kurių 1 yra paskelbta ISI indeksuojamame žurnale “Informatica”.

1. Šveikauskienė, D. *Graph Representation of the Syntactic Structure of the Lithuanian Sentence*. Informatica. 2005, Vol. 16, Nr. 3, p. 407-418.
2. Šveikauskienė, D. *A System for Automatic Syntactic Analysis of Lithuanian Simple Sentences*. Information Technologies and Control. 2007 Vol. 36, Nr. 2, p.221-237.
3. Šveikauskienė, D. *Formal Description of the Syntax of the Lithuanian Language*. Information Technologies and Control. 2005 Vol. 34, Nr. 3, p.245-256.

## Struktūra ir apimtis

Disertacija susideda iš šių dalių:

- Aiškinamojo teksto, kurio apimtis – 124 puslapiai;
- Literatūros sąrašo, pateikiančio 63-jų cituotų šaltinių nuorodas;
- Kompaktinės plokštelės, kurioje patalpinti 5 priedai.

Disertacijoje pateikta 10 lentelių ir 95 paveikslėliai.

## Darbo turinys

Pratarmėje išdėstytas disertacijos tyrimo objektas, temos aktualumas, suformuluotas darbo tikslas bei uždaviniai, aprašyti tyrimo metodai, mokslinis naujumas, praktinė reikšmė, darbo aprobavimas, pateiktas darbo publikacijų sąrašas ir darbo struktūra bei turinys.

Įvade aprašomi kalbininkų ir informatikų darbai, kurie sudarė prielaidas tautų kalbų kompiuteriniam apdorojimui.

Analitinėje apžvalgoje nagrinėjamos trijų kalbų (anglų, vokiečių ir rusų) automatinės sintaksinės analizės sistemos. Parodoma, kodėl jų negalima tiesiogiai perimti lietuvių kalbai.

Trečiame skyriuje nurodomos kai kurios specifinės lietuvių kalbos ypatybės, skiriančios ją nuo anglų, vokiečių ir rusų kalbų.

Ketvirtame skyriuje pagrindžiama būtinybė lietuviško sakinio sintaksinei struktūrai naudoti priklausomybių grafą, o ne priklausomybių medį, kaip yra anglų, vokiečių bei rusų kalbų sintaksinės analizės sistemose, t.y. modifikuojamas priklausomybių metodas papildant jį dvigubais sintaksiniais ryšiais, kuriuos turi tarininis pažyminys lietuvių kalboje.

Penktame skyriuje pateikiamas lietuvių kalbos sintaksės taisyklių aprašas Bekaus ir Nauro forma (BNF), apimantis visas penkias sakinio dalis ir pagrindinius sintaksinius ryšius tarp žodžių sakinyje.

Šeštą skyrių apima sintaksinės analizės metodą ir algoritmą, kurio veikimas pailiustruojamas lietuvių kalbos vientisinio sakinio pavyzdžiu.

Septintame skyriuje aprašoma automatinės sintaksinės analizės algoritmą realizuojanti programinė įranga.

Aštuntame skyriuje aprašomas sukurtos programinės įrangos patikrinimas ir pateikiami eksperimento rezultatai.

Devintame skyriuje parodomas automatinės sintaksinės analizės taikymo galimybės.

Išvadose trumpai aprašyti apibendrinti disertacijos tyrimų rezultatai.

Prie disertacijos pridedamoje plokštelėje pateikiami priedai: testams naudoti sakiniai bei kompiuteriu sudarytos sintaksinės struktūros.

### **Gynimui pateikiama:**

- Lietuvių kalbos sintaksės taisyklių formalus aprašas BNF (Bekaus ir Nauro forma).
- Žodžių sintaksinių funkcijų nustatymo lietuviškame vientisiniame sakinyje metodas, įvertinantis specifinius lietuvių kalbos bruožus – didelį kaitomumą ir laisvą žodžių tvarką sakinyje.
- Programinė įranga lietuvių kalbos vientisinių sakinių automatinei sintaksinei analizei atlikti.

## 2 ANALITINĖ APŽVALGA

### 2.1 Įvadas

Atsiradus kompiuteriams labai greitai paaiškėjo, kad jie gali apdoroti ne tik skaičius, bet ir kitus simbolius, pvz., raides, taigi – jie gali apdoroti ir kalbas. Daugelis lingvistų pradėjo kurti formalius kalbų aprašus, kad ir tautų kalboms būtų galima panaudoti kompiuterių pajėgumus [Winograd, psl. 23].

Dirbtinio intelekto sferoje tautų kalbų apdorojimas tapo viena pagrindinių pritaikymo sričių, nes globalus viso dirbtinio intelekto uždavinys yra pasiekti, kad dirbtinai sukurtas daiktas išmanytų apie tikrovę [Chabris, psl. 71]). Nuo 1968 m. iki 1978 m. tautų kalbų gramatinė analizė buvo pagrindinė dirbtinio intelekto tyrinėtojų darbo tema [Nirenburg, psl. 26].

Tačiau tautų kalbos, lengvai pasiduodančios žmonėms, pasirodė sunkiai įkandamos kompiuteriams [Jensen, Heidorn, Richardson, psl. 2]. Todėl labai gerų rezultatų automatinio tautų kalbų apdorojimo srityje iki šiol dar nepasiekta.

#### 2.1.1 Kalbininkų darbai

Mes galime sudaryti bei pasakyti sakinius, kurių niekada anksčiau nesame ištarę ar girdėję, ir tuos sakinius supranta kiti žmonės, kuriems jie yra iš viso nauji, nežinomi, niekada anksčiau negirdėti. Tačiau mūsų laisvė sakyti nauja yra išprausta į tam tikrus rėmus. Visos kalbos turi sintaksę, kuri uždeda griežtas toje kalboje vartojamų modelių ribas [Winograd, psl. 35].

Lingvistikos teoretikai stengiasi struktūriškai aprašyti kalbas, apibūdinti bendrus jų sandaros principus, t.y. rasti tai, kuomet remiasi visų tautų kalbos. Jie mažai dėmesio skiria kokios nors vienos atskiros kalbos tyrinėjimui. Svarbiausias jų tikslas – sudaryti taisykles, kurios aprašo galimus atvejus kalbose, ir įvesti apribojimus, nurodančius, kokie atvejai yra neleistini. Lingvistų teorijos turi būti teisingos bendrai ir tikti įvairių tautų kalboms [Allen, psl. 1]. N. Chomskis rašė, kad pagrindinis lingvistikos tikslas yra sukurti universalią gramatiką, kuri būtų bendra visoms žmonių kalboms. Ir, žinoma, nėra tokios gramatikos, kuri būtų bendra bei pasitvirtintų visų pasaulio tautų kalbose [Winograd, psl. 174].

Kompiuterinės lingvistikos specialistai domisi, kaip žmonės vartoja bei suvokia kalbą. Jų tikslas – sukurti kalbos apdorojimo mechanizmo, esančio žmoguje, kompiuterinį analogą [Allen, psl. 2].

Psicholingvistus domina tiek kalbos struktūrų pavaizdavimo būdai, tiek kelias, kuriuo žmogus gali sudaryti tokią struktūrą iš konkretaus sakinio. Su žmonėmis atliekami eksperimentai – konkretūs matavimai, kaip žmonės vartoja ar suvokia kalbą, pvz., kiek laiko žmogui reikia perskaityti kiekvieną žodį sakinyje, kiek laiko jam reikia nuspręsti, ar tai teisingas (leistinas) žodis, ar ne; taip pat analizuojama, kokias klaidas daro žmonės atlikdami įvairias lingvistines užduotis [Allen, psl. 2]. Pavyzdžiui, eksperimento metu pateikus gimtosios kalbos tekstą su praleistomis raidėmis buvo tiriama, kokią dalį prasminės informacijos turi atskiros žodžio morfemos. Kai trūko galūnių, t.y. tų raidžių, kurios turi gramatinę, morfologinę informaciją, tekstas buvo atkuriamas 95-99% tikslumu. Kai trūko žodžio pradžios, t.y. leksinės informacijos, šis procentas žymiai sumažėjo. Pvz., uzbekai negalėjo atstatyti gimtosios kalbos teksto, kuriame buvo paliktos tik žodžių darybos morfemos – priešdėliai, priesagos, galūnės [Билан, Крисевич, psl. 81].

Tyrimai taip pat parodė, kad sintaksinės informacijos dalis, esanti tam tikroje žodžio vietoje (šaknyje, galūnėje), atskiroms kalboms yra labai nevienoda. 1 lentelėje (psl. 6) matyti, kad anglų kalboje žodžių galūnės teturi labai mažai sintaksinės informacijos, palyginus su prancūzų kalba [Билан, Крисевич, psl. 82].

**1 lentelė** Sintaksinės informacijos pasiskirstymas įvairiose kalbose.

Kalbos	Sintaksinės informacijos kiekis %, perduodamas:		
	žodžiu	žodžių junginiu	morfologiniais afiksais (priešdėliu, priesaga, galūne)
Anglų kalboje	65,0	81,0	3,5
Prancūzų kalboje	76,5	86,1	16,0

Psicholingvistinių eksperimentų duomenys kartais naudojami patvirtinant bei atmetant įvairias hipotezes apie kalbą, kurias iškelia teorinės ar kompiuterinės lingvistikos tyrinėtojai [Allen, psl. 2]. Tokios hipotezės pavyzdys galėtų būti Rusijos mokslininkų iškelta mintis, kad, remiantis tekste pavartotų sakinių sintaksinėmis struktūromis, galima nustatyti autorių pagal jo individualų sintaksinį braižą [Белецкая, psl. 37].

Įdomi kalbos suvokimo ypatybė yra ir tai, kad painūs, sudėtingi pasakymai (*The book the article the index listed mentioned quoted the Kama Sutra*) kartais gali būti labai sunkiai suprantami, o braižant pieštuku ant popieriaus sakinio struktūrą, galima daug lengviau atsekti, kas su kuo yra susiję [Winograd, psl. 145].

### 2.1.2 Informatikų darbai

Lingvistikos tyrinėtojams susidomėjus kalbų formalizavimu buvo bandoma sudaryti matematinę kalbos modelį. Taip gimė formalios kalbos bei formalios gramatikos idėja. Tik panaudota ji buvo programavime daug anksčiau, nei tautų kalbose, ir matyt todėl, kad formalų aprašą lengviau taikyti dirbtiniam daiktui (t.y. programavimo kalbai).

Formali kalba, kaip ir kiekviena kalba, turi savo abėcėlę, tik tai šios abėcėlės simbolius reikia suprasti abstrakčiau, nei mums gerai žinomas 32 lietuviškas raides. Tai gali būti raidės, skaitmenys, skyrybos ženklai; netgi keli spausdinti ženklai gali būti vienas tos abėcėlės simbolis.

Formalios kalbos abėcėlė susideda iš dviejų rūšių simbolių: terminalinių (galinių) ir neterminalinių (negalinių). Terminaliniai simboliai – tai kalbos, kurią reikia aprašyti (ar tai būtų kurios nors tautos kalba, ar programavimo kalba), vienetai. Jų pavyzdžiai tokie: programavimo kalboms – *if, while, do, a, +, =, 5, 7*; tautų kalboms – tai morfologinės kategorijos ar tiesiog sakinio žodžiai. Neterminaliniai simboliai nurodo pačios formalios kalbos sąvokas. Programavimo kalbose jie yra konstrukcijų pavadinimai (*kintamasis, prieskyros ženklas* ir t.t.), o tautų kalbose – sintaksinės kategorijos (*daiktavardinė frazė, sakiny*s ir kt.) [Dagienė, Grigas].

Terminalinių (galinių) simbolių seka vadinama eilute. Tautų kalbose sakiny s yra žodžių seka [Batori, Lenders, psl. 622]. Taigi, žodžiai yra terminaliniai simboliai, o sakiny s eilutė.

Formalią kalbą sudaro eilučių rinkinys, tačiau ne visų galimų, o tik tam tikrų (taip pat, kaip ir tautų kalbose: juk ne bet koks lietuviškų žodžių kratinys yra lietuvių kalbos sakiny s). Ar eilutė priklauso formaliai kalbai ar ne, nustatoma naudojantis formaliomis gramatikomis. (Lygiai taip pat, kaip ir tautų kalbų atveju: pvz., anglų kalbos gramatika teigia, kad sakiny s *Put some paper in the printer* yra galimas, o sakiny s *Printer some put the paper in* yra negalimas anglų kalboje [Arnold, Balkan, psl. 39].)

Kadangi formalios kalbos elementai yra eilutės, tai jų apdorojimui taikomos eilučių operacijos. Programuotojams gerai žinoma eilučių keitimo komanda. Operacija  $u \rightarrow v$  reiškia, kad simbolių seką  $u$ , rastą pradinėje eilutėje, reikia pakeisti seka  $v$ . Šitaip gaunama nauja eilutė. Pvz., jei turime eilutę  $aab$ , tai taikydami operaciją  $ab \rightarrow bac$  gausime dvi naujas eilutes:  $aab \rightarrow abac \rightarrow bacac$ , o taikant operaciją  $b \rightarrow bb$  iš pradinės eilutės galima gauti be galo daug naujų eilučių:  $aab \rightarrow aabb \rightarrow aabb \rightarrow aabbbb \dots$  Formalių gramatikų taisyklėms užrašyti ir buvo pasirinkta ši operacija. Formalią gramatiką sudaro keturios dalys:

- 1) neterminaliųjų simbolių baigtinė aibė  $N$ ;
- 2) terminalinių simbolių baigtinė aibė  $T$ ;
- 3) gramatikos taisyklių baigtinė aibė  $P$ ;
- 4) pradinis neterminalinis simbolis  $S$ , nuo kurio pradedamas eilučių generavimas, arba kuris turi būti gaunamas analizės metu.

Sutrupintai gramatika  $G$  gali būti užrašoma:  $G = (N, T, P, S)$ .

Noamas Chomskis suskirstė formalias gramatikas į keturias klases, įveddamas tam tikrus apribojimus pačiai bendriausiai formalios gramatikos taisyklei, išreikštai  $u \rightarrow v$  ( $u$  vadinama kairiąja taisyklės puse,  $v$  – dešiniąja). Tokią gramatiką, kuriai netaikomi jokie apribojimai, jis pavadino nulinio tipo gramatika.

Pirmasis apribojimas buvo toks:  $uAv \rightarrow uvv$ . Šio tipo gramatikos vadinamos kontekstinėmis, nes eilutės pakeitimą  $A \rightarrow w$  galima atlikti tik tada, kai keičiama dalis  $A$  yra tam tikrame kontekste: jos kairėje pusėje yra eilutė  $u$ , o dešinėje – eilutė  $v$ . Pats kontekstas ir po taisyklės taikymo lieka tas pats, tik jis jau gaubia naują eilutę  $w$ .

Antras apribojimas:  $A \rightarrow w$ . Tokios gramatikos vadinamos nekontekstinėmis, kadangi neterminalinį simbolį  $A$  galima keisti dešiniąja taisyklės dalimi nepriklausomai nuo to, kokioje eilutės vietoje jis bebūtų, t.y. nepriklausomai nuo kaimyninių simbolių. Beveik visoms programavimo kalboms ir naudojama šio tipo gramatika.

Trečiasis apribojimas yra:  $A \rightarrow a|aB$ . Trečiojo tipo gramatikos vadinamos automatinėmis, nes tokio tipo kalbas galima apibrėžti kaip baigtinius automatus.

Tautų kalbose sunku rasti pavyzdžių, kad sintaksinės struktūros nebūtų galima gauti naudojantis nulinio ar pirmo tipo gramatikomis. Problema yra tai, kad norint sudaryti tokias struktūras, kokios mums reikalingos, kartais neužtenka paprastų gramatikų [Winograd, psl. 145]: automatinės, o kartais ir nekontekstinės.

Formaliųjų gramatikų taisyklėms užrašyti naudojama Bekaus ir Nauro forma (BNF). Tai formalizuotas gramatikos taisyklių užrašas, pavadintas jį pasiūliusių autorių vardu. BNF galima nusakyti taip:

- 1) vietoje rodyklės  $\rightarrow$  rašomas simbolis  $::=$ ,
- 2) neterminaliniai simboliai rašomi tarp kampinių skliaustų  $< >$ ,
- 3) vertikalus brūkšnylis  $|$  (kaip ir formaliųjų gramatikų taisyklėse) skiria alternatyvas, t.y. turi reikšmę „arba“,
- 4) laužtiniuose skliaustuose  $[ ]$  esanti simbolių seka gali būti praleista,
- 5) suskliaudus simbolių grupes į paprastuosius lenktinius skliaustus, galima alternatyviai nurodyti ne tik visą dešiniąją taisyklės pusę, bet ir jos dalį, pvz., skaičius, būtinai turintis ženklą, aprašomas taip:  $\langle \text{skaičius su ženklu} \rangle ::= (+ \mid -) \langle \text{skaičius} \rangle$ ,
- 6) simbolių seka, kuri gali būti pakartota, rašoma į riestinius skliaustus, pvz., vardų sąrašą galima aprašyti taip:  $\langle \text{vardų sąrašas} \rangle ::= \langle \text{vardas} \{, \text{vardas} \} \rangle$ .

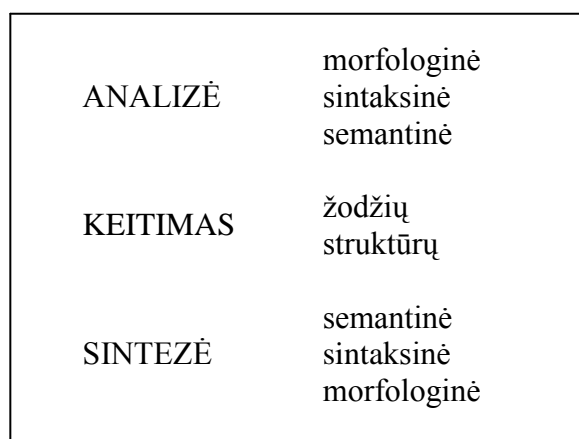
Lietuvių kalbos sintaksės taisyklės šiame darbe bus aprašomos BNF.

## 2.2 Automatinė sintaksinė analizė

Žodis „sintaksė“ kilęs iš graikų kalbos *syntaxis*, kurio pirminė reikšmė buvo – sąryšis, išsidėstymas, išsirikiavimas, sąveika. Graikai jį vartojo karo moksluose, kalbėdami apie kariuomenę; vėliau jis pateko į logiką, o po to – į gramatiką [Labutis, psl. 8].

Sakinio sintaksinė struktūra rodo, kaip žodžiai sakinyje yra susiję vienas su kitu [Allen, psl. 9]. Labiausiai paplitęs sakinio struktūros pavaizdavimo metodas yra grafai, tiksliau, medis [Allen, psl. 41]. Kalbininkų medžiai visada braižomi viršūne žemyn, t. y. jų šaknis yra viršuje, o lapai apačioje [Batori, Lenders, psl. 23]. Žinomi du iš principo skirtingi medžio sudarymo būdai: frazių metodas ir priklausomybių metodas. Frazių metodas atėjo iš Amerikos struktūralizmo (Bloomfield'o darbai) ir daugiausiai naudojamas anglų kalbos sintaksinei analizei, o priklausomybių gramatika kilo iš europietiškos tradicinės gramatikos (kalbininko Tesniere'o darbai) [Batori, Lenders, psl. 625] ir labiau taikoma kitoms Europos kalboms.

Pagrindinis automatinės sintaksinės analizės pritaikymas yra automatinis vertimas. Šiuo metu labiausiai paplitusiose automatinio vertimo sistemose naudojama keitimo metodika, kai vertimas skaidomas į tris fazes [Hutchins Sommers, psl. 4]. Detaliau automatinio vertimo sudedamosios dalys parodytos 1 pav.:



1 pav. Automatinio vertimo etapai, naudojant keitimo metodiką.

Analizės metu šaltinio tekstas performuojamas į tarpinę struktūrą. Tai atliekama trimis etapais:

- a) morfologinė analizė;
- b) sintaksinė analizė;
- c) semantinė analizė.

Morfologinės analizės etape kiekvienam teksto žodžiui surandama pradinė forma (bendratis, vardininko linksnis ir pan.) bei nusakoma morfologinė informacija apie sakinyje pavartotą žodžio formą (giminė, skaičius, linksnis, laikas, asmuo ir kt.). Lietuvių kalbos morfologinę analizę gali atlikti Vytauto Zinkevičiaus sukurta Lietuvių kalbos morfologinių modelių programinė įranga [Zinkevičius].

Sintaksinės analizės metu nustatomos žodžių sintaksinės funkcijos bei nurodomi ryšių tipai tarp jų. Lietuvių kalbai sintaksinės analizės sistema kol kas dar nėra sukurta, ir šio darbo tikslas būtų paruošti programinę įrangą kompiuteriniam lietuvių kalbos sintaksės apdorojimui. Iš jau atliktų šios srities darbų paminėtinas magistrinis darbas, kuriame nagrinėjamos galimybės pritaikyti DGS (Dependency Grammar with States – priklausomybių gramatikos su būsenomis) metodiką lietuvių kalbai. Pagal šią metodiką analizė atliekama dviem etapais: pradžioje sudaromas grafai, atspindintis visus galimus sintaksinius ryšius sakinyje, o vėliau iš jo formuojamas priklausomybių medis [Daudaravičius, psl. 55].



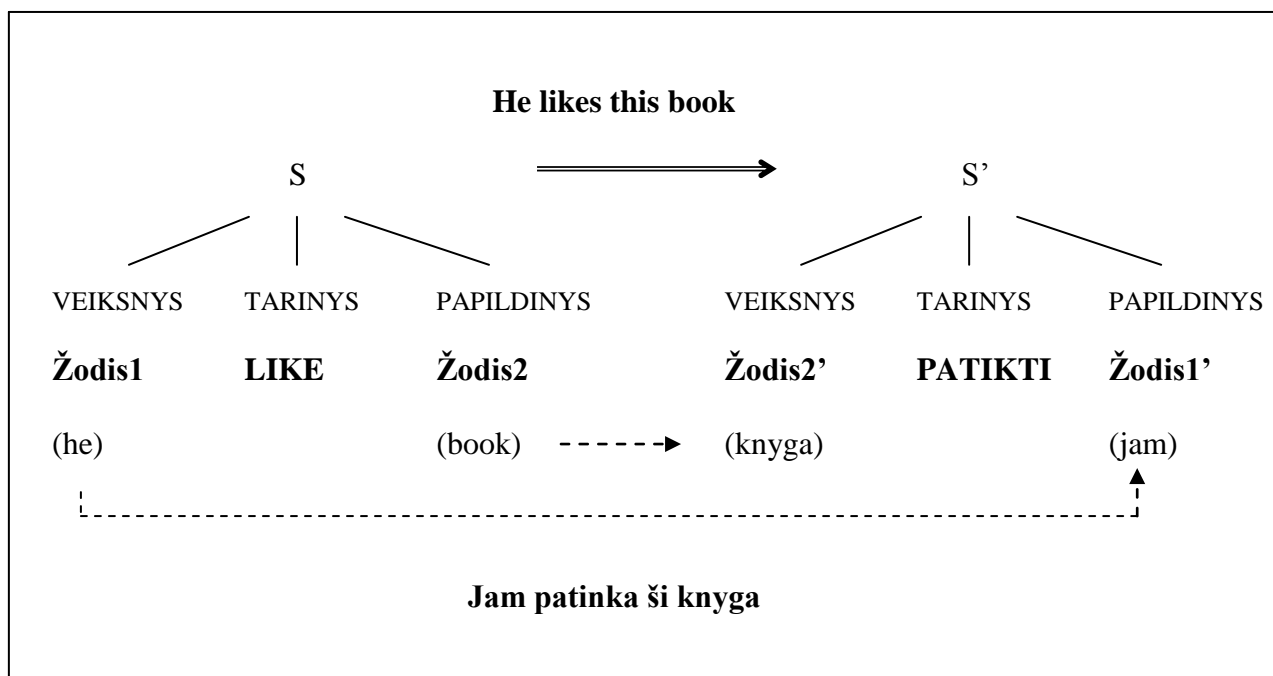
Semantinė analizė naudojama sintaksinės analizės rezultatams pagerinti. Jos metu žodžiams priskiriami tam tikri reikšmės požymiai. Kai kuriais atvejais, pasitelkiant tuos požymius, galima panaikinti sakino sintaksinės struktūros daugiareikšmiškumą. Be to, jie padeda tiksliau parinkti žodžius rezultato kalboje [Langer, psl. 8]. Sakykim, norint nustatyti tinkamiausią žodį iš kelių galimų, reikalingi, pvz., tokie požymiai:

- a) ar tai gyvas daiktas, ar ne, svarbu verčiant santykinius ir klausiamuosius įvardžius į rusų kalbą (*что, кто*), taigi, visi šaltinio kalbos daiktavardžiai turi būti skirstomi į dvi semantines grupes – gyvas ar negyvas daiktas [Марчук, psl. 96];
- b) ar tai skaičiuojamas daiktas, ar ne, aktualu verčiant žodį *daug* į anglų kalbą – *much time, many books* ir pan.

Lietuvių kalbos semantinės analizės sistema taip pat dar nėra sukurta. Kai kurie semantiniai požymiai šiame darbe bus naudojami kaip sintaksinės analizės sudedamoji dalis.

Antras automatinio vertimo etapas skirtas susieti abi sistemos kalbas. Jo metu šaltinio teksto tarpinė forma pakeičiama analogiška rezultato kalbos tarpine forma. Keitimas vyksta dviem lygmenimis – leksiniu ir struktūriniu:

1. Leksiniame lygmenyje vienos kalbos žodžiai keičiami kitos kalbos žodžiais.
2. Struktūriniame lygmenyje vienos kalbos sakinio sintaksinė struktūra keičiama kitos kalbos sintaksine struktūra. Labai dažnai jos nesutampa, todėl, neatlikus sintaksinės analizės, geros kokybės vertimo gauti negalima. Pvz., sakinyje *He likes this book* veiksmožodžio *like* struktūrinė schema, verčiant jį į lietuvių kalbą, būtų keičiama taip (2 pav.):



2 pav. Žodžio *like* struktūrinės schemos pakeitimas žodžio *patikti* struktūriniu schema.

Iš esmės visi sintaksiniai pakeitimai vertimo metu yra veiksmai, atliekami su medžiais [Гарина, psl. 43].

Sintezės paskirtis yra iš pakeistos tarpinės struktūros suformuoti išverstą tekstą. Tai yra atvirkščias teksto analizei procesas, ir turi jis tas pačias sudedamąsias dalis, tik išsidėsčiusias atvirkščia tvarka. Todėl sukurtomis morfologinės bei sintaksinės analizės programomis bus galima pasinaudoti ruošiant atitinkamas sintezės etapo sistemas.

Džordžauno universiteto mokslininkas L. Dostert buvo bene pirmasis (1954 m.), kuris sintaksinei analizei teikė tiek pat reikšmės kiek ir leksiniam (žodyno) apdorojimui [Hutchins, psl. 71]. O jau 1984 m. H. U. Block'as savo disertacijoje rašė: „Šaltinio kalbos sintaksinė analizė yra svarbiausia visos [vertimo] sistemos dalis” („Die syntaktische Analyse der Quellsprache ist der wichtigste Teil des gesamten Systems.”) [Block, psl. 24].

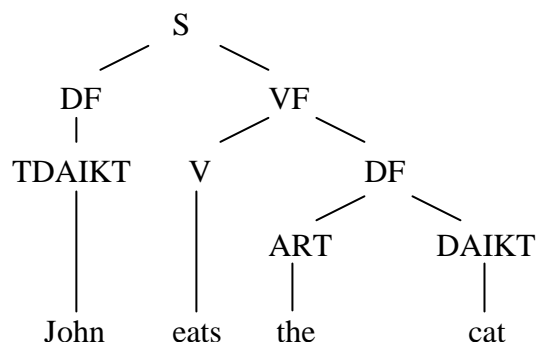
### 2.2.1 Anglų kalbos sintaksinė analizė

Anglų kalbos sakiny (S) paprastai susideda iš daiktavardinės frazės DF ir veiksmažodinės frazės VF. Veiksmažodinę frazę sudaro veiksmažodis V ir daiktavardinė frazė DF. Sudėtingesnės veiksmažodinės frazės apima veiksmažodį V, daiktavardinę frazę DF ir prielinksninę frazę PF.

Anglų kalbos sintaksinei analizei atlikti naudojamos formalios nekontekstinės gramatikos. Taisyklių, vadinamų pakeitimo taisyklėmis, rinkinys nurodo, kokia sakinio struktūra yra galima. Kiekviena taisyklė teigia, kad tam tikras simbolis medyje gali būti išskleistas kitų simbolių seka [Allen, psl. 41]. Labai paprasto gramatikos taisyklių rinkinio pavyzdys pateikiamas 3 pav. Šių taisyklių pagalba galima išnagrinėti tokį anglišką sakinį: *John eats the cat*. 4 pav. parodyta šio sakinio sintaksinė struktūra. Reikia pabrėžti, kad sintaksei visai nerūpi sakinio reikšmė. Ji turi išnagrinėti sakinį formaliai, visiškai nesigilindama į jo prasmę. Todėl net ir logiškai neteisingam sakiniui kompiuteris privalo sugebėti nustatyti sintaksinę struktūrą, jei tik formaliai (gramatiškai) šis sakiny yra taisyklingas. Turi būti galima automatiškai išvesti sakinį, pvz., *Asmuo, apsilankęs kavinėje, privalo ten išplauti indus*, nors vertimo sistema ir nesuprastų, kodėl tai turi būti daroma [Laffling, psl. 7].

S → DF VF	(1)	TDAIKT → John	(5)
VF → V DF	(2)	V → eats	(6)
DF → TDAIKT	(3)	ART → the	(7)
DF → ART DAIKT	(4)	DAIKT → cat	(8)

3 pav. Gramatikos taisyklių rinkinys, su kuriuo galima išnagrinėti sakinį *John eats the cat*.



4 pav. Sakinio *John eats the cat* sintaksinė struktūra.

Čia – S, DF, VF, V, TDAIKT, ART, DAIKT yra neterminaliniai simboliai, o John, eats, the, cat – terminaliniai (galiniai).

Yra dvi labiausiai paplitę sakinio nagrinėjimo technikos: „iš viršaus į apačią“ ir „iš apačios į viršų“. Nagrinėjimas „iš viršaus į apačią“ prasideda nuo simbolio S, pakeičiant jį į DF VF, kuriuos vėlgi galima keisti kitais simboliais, taikant kitas taisykles. Nagrinėjimo metu kairėje rodyklės pusėje esantys simboliai keičiami dešinė jos puse tol, kol nebelieka neterminalinių (negalinių) simbolių, t.y. kol gaunama eilutė, sudaryta vien iš terminalinių (galinių) simbolių.

Pasinaudojant tuo pačiu sakiniu *John eats the cat*, galima parodyti nagrinėjimo „iš viršaus į apačią“ žingsnius (5 pav.):

S	→ DF	VF	(1 taisyklė)
→	TDAIKT	VF	(3 taisyklė)
→	John	VF	(5 taisyklė)
→	John	V VF	(2 taisyklė)
→	John eats	DF	(6 taisyklė)
→	John eats	ART DAIKT	(4 taisyklė)
→	John eats the	DAIKT	(7 taisyklė)
→	John eats the	cat	(8 taisyklė)

5 pav. Sakinio *John eats the cat* nagrinėjimo „iš viršaus į apačią“ žingsniai.

Nagrinėjant sakinį „iš apačios į viršų“ (6 pav.), pradama nuo terminalinių simbolių eilutės, t.y. nuo konkretaus sakinio, ir žodžiai keičiami jų sintaksinėmis kategorijomis. Šiuo atveju dešinėje rodyklės pusėje esantys simboliai keičiami kairės pusės simboliais. Jei taikant taisykles pavyksta gauti simbolį S, reiškia sakinytis yra gramatiškai teisingas.

→	John	eats	the	cat	
→	TDAIKT	eats	the	cat	(5 taisyklė)
→	TDAIKT	V	the	cat	(6 taisyklė)
→	TDAIKT	V	ART	cat	(7 taisyklė)
→	TDAIKT	V	ART	DAIKT	(8 taisyklė)
→	DF	V	ART	DAIKT	(3 taisyklė)
→	DF	V		DF	(4 taisyklė)
→	DF			VF	(2 taisyklė)
→	S				(1 taisyklė)

6 pav. Sakinio *John eats the cat* nagrinėjimo „iš apačios į viršų“ žingsniai.

Su tokiu gramatikos taisyklių rinkiniu tegalima išnagrinėti šį vieną sakinį. Sudėtingesniems sakiniams nagrinėti reikia sudėtingesnių gramatikų. Pvz., 1978 m. Grenoblio universitete sukurta automatinio vertimo sistema ARIANE-78 sintaksinei analizei naudoja 176 taisykles [Guilbaud, psl. 297], o teoriniame darbe HARWARD SYNTACTIC ANALYSER anglų kalbai buvo sudaryta virš 3000 taisyklių [Winograd, psl. 360].

Formalios gramatikos gali būti labai įvairios. Vieną ir tą patį sakinį galima išnagrinėti panaudojant kelias skirtingas gramatikas, t.y. kelis skirtingus pakeitimo taisyklių rinkinius. Kuri gramatika geriausia, nulemia trys kriterijai [Arnold, Balkan, psl. 44]:

- 1) ar gramatika aprašo visus kalbos sakinius;
- 2) ar gramatika yra nepriekaištinga leistinių sakinių atžvilgiu, t.y. ar nelaiko teisingais gramatiškai netaisyklingų sakinių;
- 3) ar lengva ją suprasti bei taikyti automatiniam apdorojimui.

Pvz., 7 pav. pateikta gramatika nėra gera antro kriterijaus požiūriu. Ji laiko taisyklingais ir tokius sakinius: *The users cleans the printer*; *The users should cleans the printers*, *The user clean the printer* ir pan. [Arnold, Balkan, psl. 44].

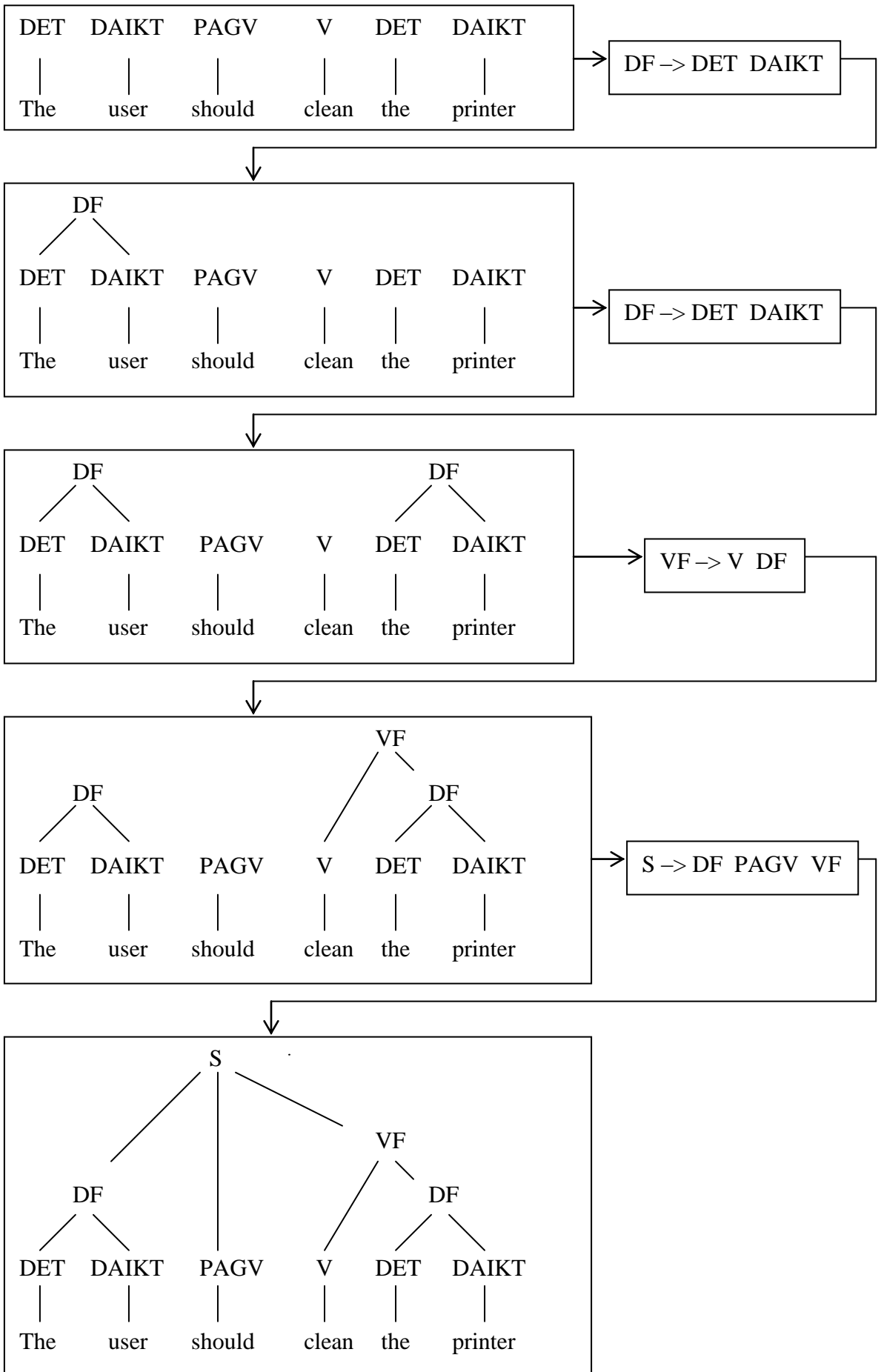
S → DF V	DAIKT → user
S → DF PAGV VF	DAIKT → users
DF → DET DAIKT	DAIKT → printer
DF → DET BŪDV DAIKT	DAIKT → printers
VF → V DF	V → clean
DET → the	V → cleans
DET → a	PAGV → should

**7 pav.** Formali gramatika, kuri nėra nepriekaištinga leistinių sakinių atžvilgiu.

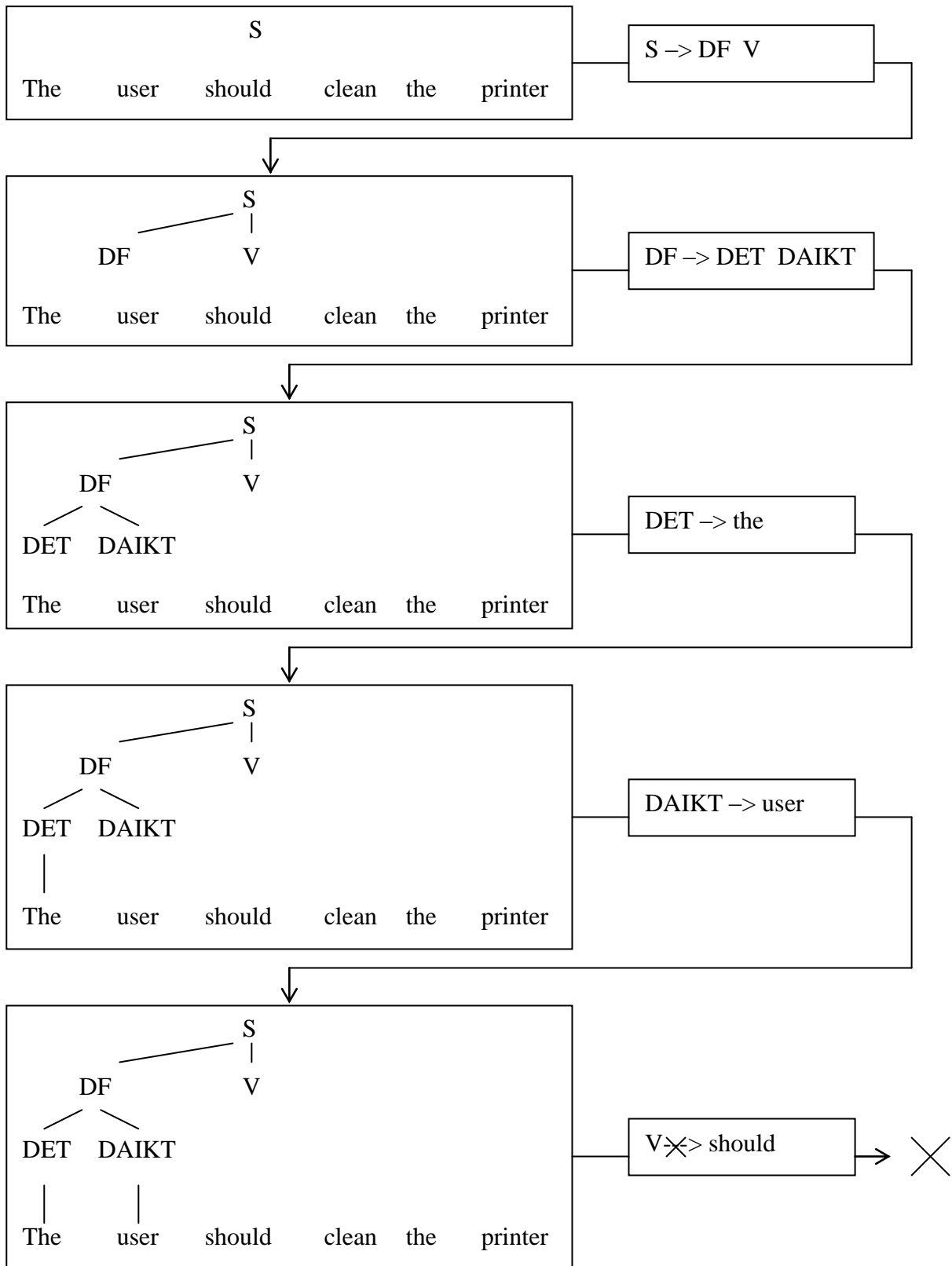
Tačiau sakinį *The user should clean the printer* ji gali išnagrinėti visai gerai. Todėl pasinaudojant šia gramatika bus parodyti nagrinėjimo žingsniai pačioje sakinio struktūroje. Tai turėtų aiškiau pademonstruoti kompiuterio darbą sintaksinės analizės metu.

Naudojant metodiką „iš apačios į viršų“ nagrinėjimas pradedamas nuo paties sakinio. Pirmas žingsnis – tai tarsi žodžių peržiūrėjimas anglų kalbos žodyne, t.y. nurodomos jų morfologinės kategorijos, iš kurių vėliau formuojami sintaksiniai vienetai [Arnold, Balkan, psl. 55].

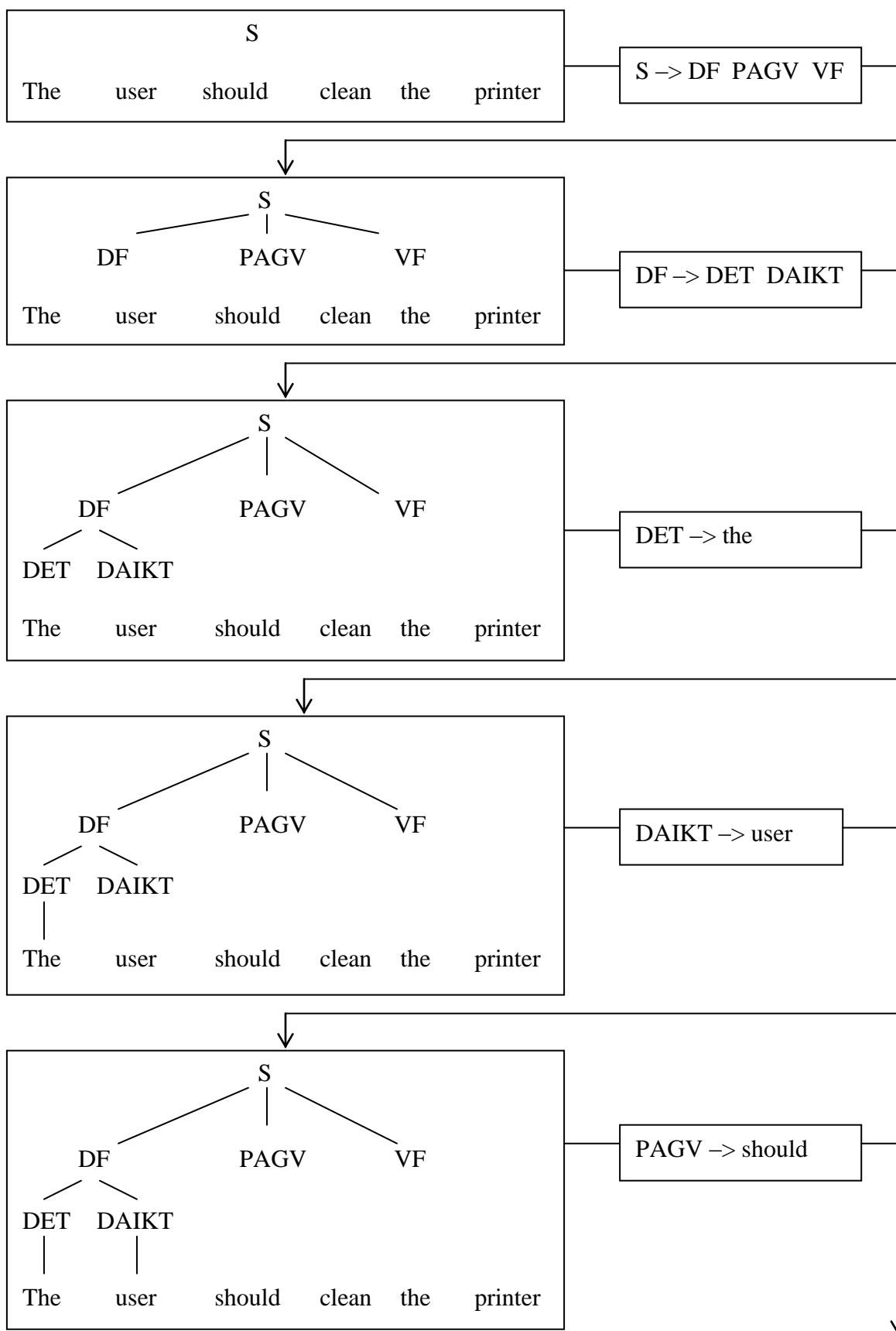
Būtent ši metodika „iš apačios į viršų“ ir bus naudojama lietuvių kalbos sakinių sintaksinei analizei atlikti, t.y. pradedama nuo terminalinių simbolių BNF apraše (arba nuo morfologinių kategorijų, rodančių pavartoto sakinyje žodžio formą). Toks analizės būdas pasirenkamas atsižvelgiant į tai, kad lietuvių kalba neturi griežtos, sugramatintos žodžių tvarkos [Labutis, psl. 15], ir didžiausia sintaksinės informacijos dalis yra sukaupta žodžių galūnėse, t.y. morfologinėje formoje.

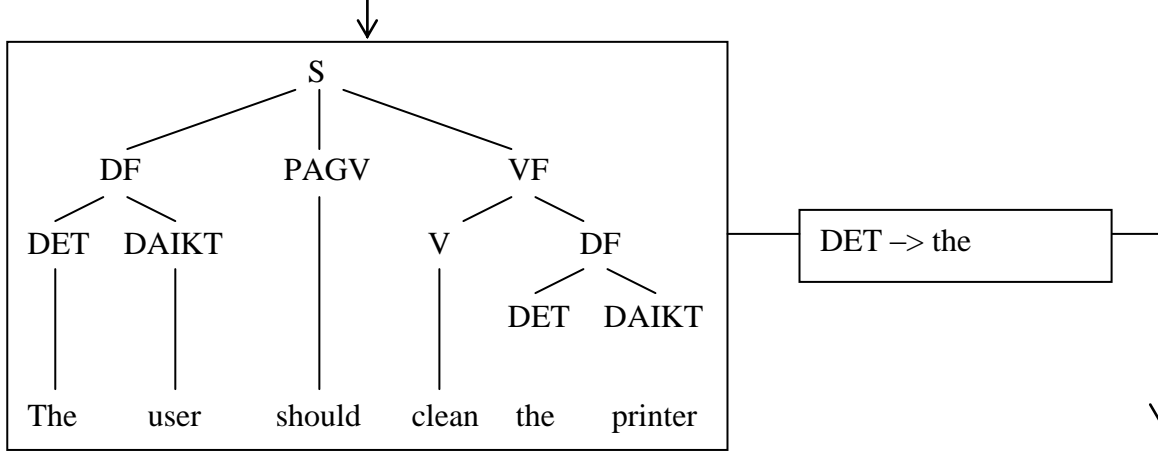
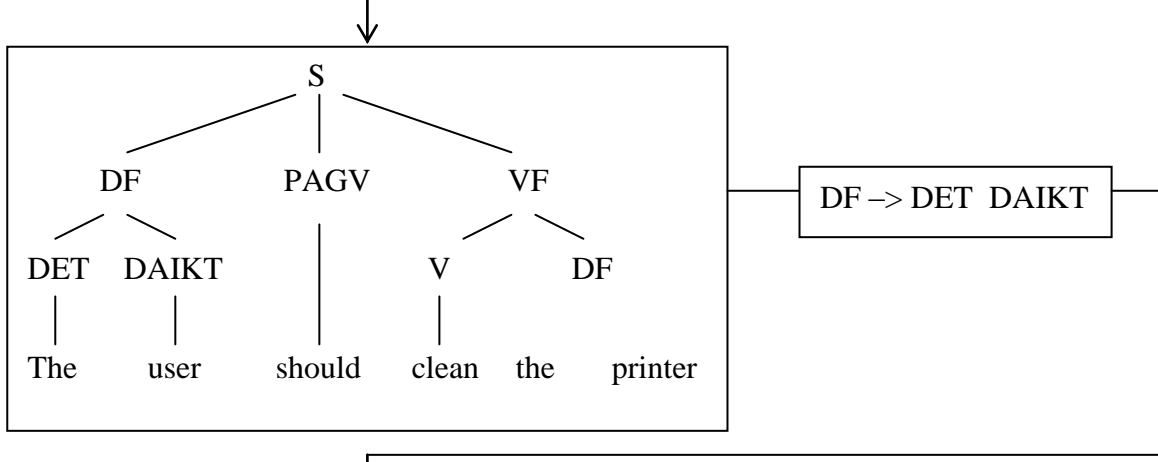
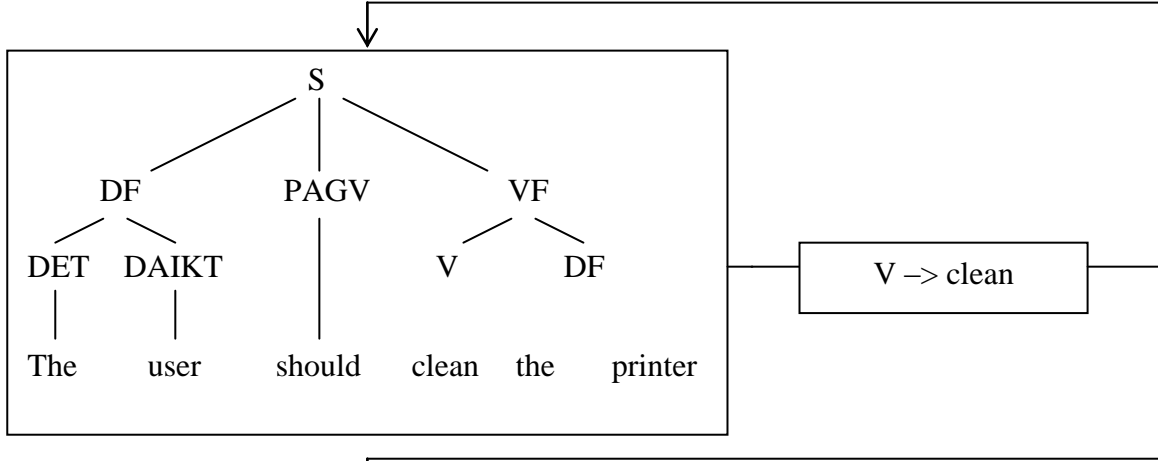
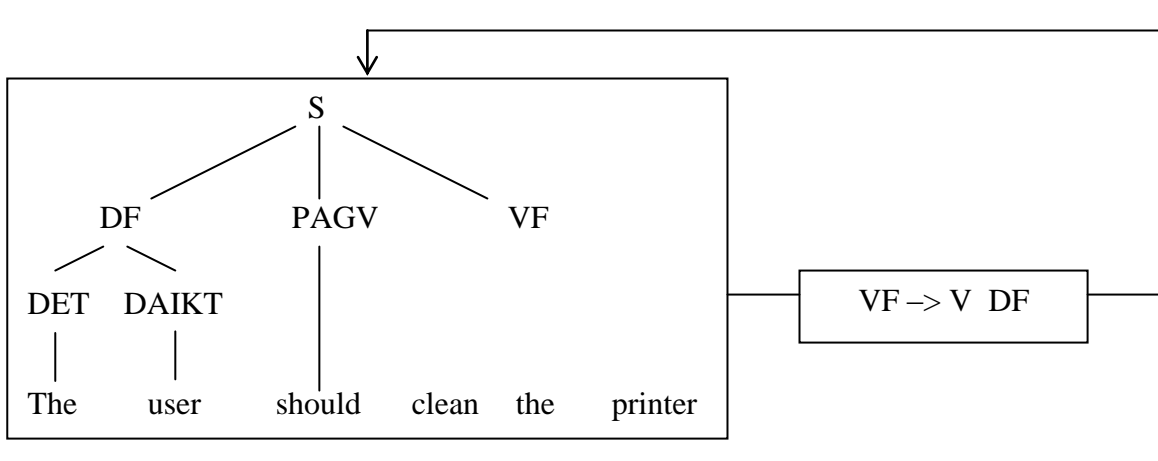


Nagrinėjant sakinį „iš viršaus į apačią“, pradedama nuo viršūnės S. Pasirenkama pirma alternatyva ir tikrinama, ar tikrai sakinytis prasideda daiktavardine fraze DF. Jei tai pasitvirtina, žiūrima, ar po jos eina veiksmažodis V. Kadangi taip nėra, grįžtama atgal

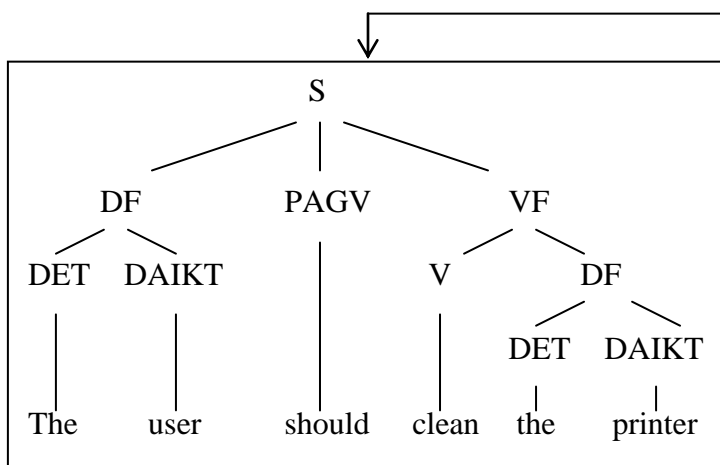
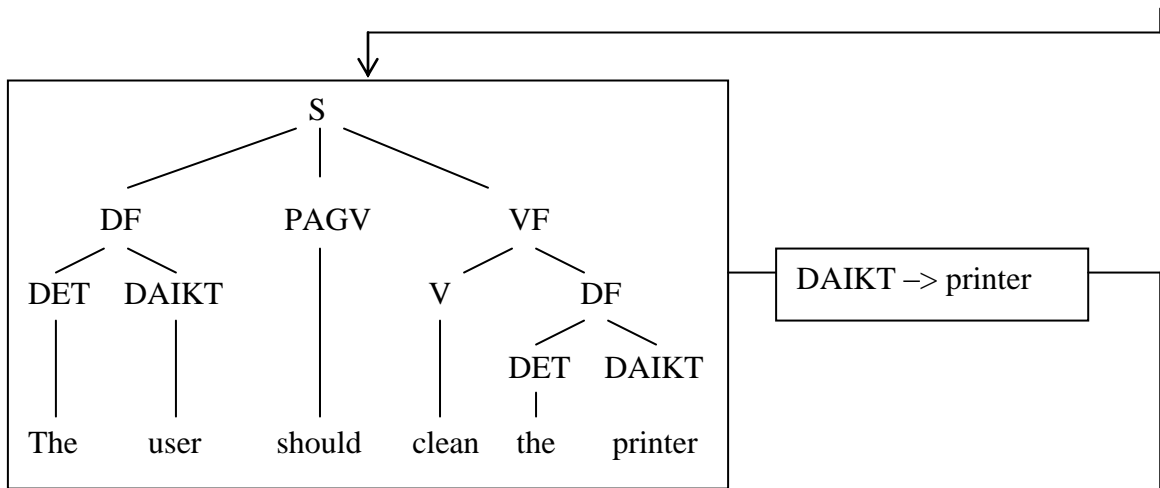


ir pasirenkama kita alternatyva: S → DF PAGV VF.









Dažnai galutinis rezultatas gaunamas tik po daugelio klaidingų žingsnių ir kiekvieną kartą iš naujo reikia nagrinėti tas sakinio vietas, kurios jau anksčiau buvo teisingai išnagrinėtos [Arnold, Balkan, psl. 55]. Pvz., sakinio pradžioje esanti daiktavardinė frazė DF antrą kartą vėl turi būti nustatoma kaip *the user*, nors pirmos alternatyvos metu ji jau buvo teisingai nustatyta. Tuo pasireiškia šio metodo („iš viršaus į apačią“) trūkumas. Nagrinėjant „iš apačios į viršų“ niekada nereikia antrą kartą ieškoti jau surastų frazių. Tačiau šio metodo silpnoji pusė yra ta, kad nagrinėjimo metu surandamos visos galimos frazės, taigi, ir tos, kurios niekada neatves prie teisingos sakinio struktūros.

Ši blogybė ypač išryškėja anglų kalboje, kur labai aštri žodžių daugiareikšmiškumo problema (sutampančios daiktavardžių ir veiksmažodžių formos). Pvz., sakinyje *The blue can can hold the water*, galimi tokie žodžių junginiai: *can can* – PAGV V, *can hold* – taip pat PAGV V, *The blue* – ART DAIKT, *The blue can* – ART DAIKT PAGV ir *The blue can* – ART BŪDV DAIKT ir t.t. Sudarius šiam sakiniui gramatikos taisyklių rinkinį (8 pav., psl. 18), kompiuterio darbą analizės metu galima stebėti diagramoje (9 pav., psl. 18) [Allen, psl. 66].

S → DF VF	(1 taisyklė)	ART → the	(8 taisyklė)
S → DF PAGV V	(2 taisyklė)	BŪDV → blue	(9 taisyklė)
S → DF V	(3 taisyklė)	DAIKT → blue	(10 taisyklė)
DF → ART DAIKT	(4 taisyklė)	DAIKT → can	(11 taisyklė)
DF → ART BŪDV DAIKT	(5 taisyklė)	DAIKT → hold	(12 taisyklė)
VF → PAGV V DF	(6 taisyklė)	DAIKT → water	(13 taisyklė)
VF → V DF	(7 taisyklė)	PAGV → can	(14 taisyklė)
		V → can	(15 taisyklė)
		V → hold	(16 taisyklė)
		V → water	(17 taisyklė)

8 pav. Gramatikos taisyklių rinkinys sakiniui *The blue can can hold the water* išnagrinėti.

<b>S4 (taisyklė 1)</b>						
S3 (taisyklė 2)						
S2 (taisyklė 3)						
S1 (taisyklė 3)			<b>VF2 (taisyklė 6)</b>			
<b>DF2 (taisyklė 5)</b>			VF1 (taisyklė 7)			
DF1 (taisyklė 4)		<b>DAIKT2</b>	DAIKT3	<b>DF3 (taisyklė 4)</b>		
	DAIKT1	V1	V2	<b>V3</b>		V4
<b>ART1</b>	<b>BŪDV1</b>	PAGV1	<b>PAGV2</b>	DAIKT4	<b>ART2</b>	<b>DAIKT5</b>
<b>the</b>	<b>blue</b>	<b>can</b>	<b>can</b>	<b>hold</b>	<b>the</b>	<b>water</b>

9 pav. Sakinio *The blue can can hold the water* analizės diagrama.

Diagramos apatinėje eilutėje pateiktas pats sakinys. Virš kiekvieno sakinio žodžio nurodytos visos galimos jo morfologinės kategorijos. Kaip matyti iš diagramos, kiekvienas savarankiškas žodis (ne artikkelis) turi jų po dvi ar tris. Tos alternatyvos, kurios pasitvirtina šiame sakinyje, pažymėtos geltona spalva. Balti langeliai rodo atvejus, kurie šiam sakiniui netinka.

*Lietuvių kalbai žodžių daugiareikšmiškumo problema neturėtų būti labai aktuali, nes sutampančios veiksmažodžių ir daiktavardžių formos yra retas atvejis. Pvz., žodis „laiką“ gali būti ir daiktavardis (neramūs laikai), ir veiksmažodis (tu laikai), tačiau tokių žodžių pas mus yra mažai. Taigi, lietuvių kalbos sintaksinei analizei bus naudojamas metodas „iš apačios į viršų“. Tuo labiau, kad, esant laisvai žodžių tvarkai sakinyje, pradėti jo analizę „iš viršaus“ iš viso neįmanoma, nes kiekvienas žodis gali būti bet kurioje sakinio vietoje. Įvertinant dar ir tai, kad dauguma žodžių lietuvių kalboje turi virš 20 formų, sakinio sudaryto iš 10 ar 20 žodžių galimų išsidėstymo variantų kiekis bus  $10^{20}$  ar  $20^{20}$ , o tai jau yra nepolinominio sudėtingumo uždavinys, t.y. realiam laikui jis neišsprendžiamas [Cormen, psl. 917].*

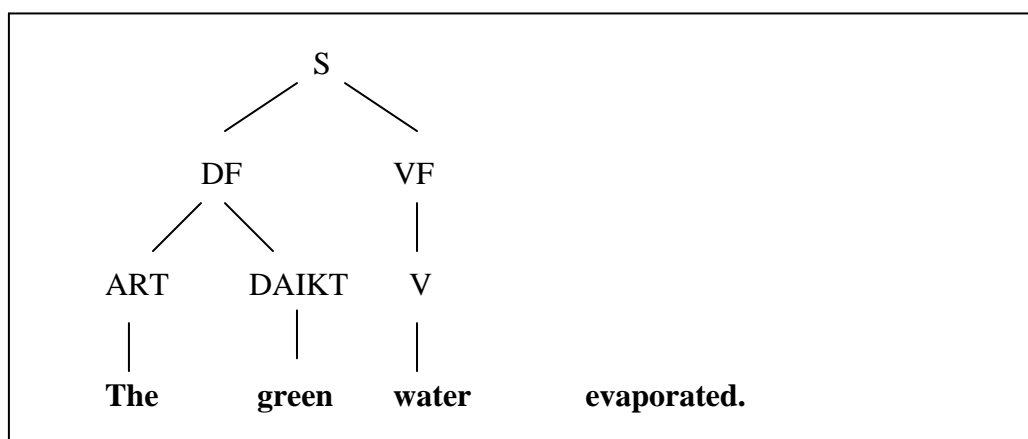
Bendrą angliško sakinio nagrinėjimo strategiją galima pailiustruoti pavyzdžiu *The green water evaporated*, kur žodis *green* gali būti būdvardis arba daiktavardis, o žodis *water* – daiktavardis arba veiksmažodis ir gramatikos taisyklės yra tokios [Allen, psl. 67] (10 pav.):

S → DF VF	(1 taisyklė)	ART → the	(7 taisyklė)
DF → ART DAIKT	(2 taisyklė)	DAIKT → green	(8 taisyklė)
DF → ART BŪDV DAIKT	(3 taisyklė)	DAIKT → water	(9 taisyklė)
VF → V	(4 taisyklė)	BŪDV → green	(10 taisyklė)
VF → PAGV V	(5 taisyklė)	V → water	(11 taisyklė)
VF → V DF	(6 taisyklė)	V → evaporated	(12 taisyklė)
		PAGV → can	(13 taisyklė)

10 pav. Gramatikos taisyklės, pagal kurias nagrinėjamas sakinys *The green water evaporated*.

Kiekvienas analizės žingsnis atliekamas įvertinant jau turimą informaciją. Apdorojimas prasideda visada nuo kairiausio simbolio. Jei tai angliško sakinio žodis, t.y. terminalinis (galinis) simbolis, pereinama prie sekančio simbolio į dešinę, kol sutinkamas neterminalinis (negalinis). Pirmas sutiktas neterminalinis simbolis keičiamas pagal taisyklę dešinėje rodyklės pusėje esančia seka.

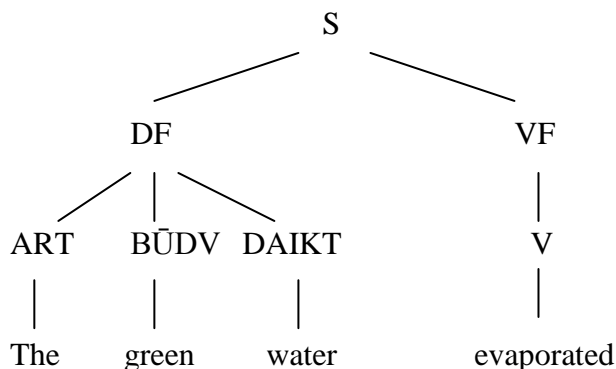
Pradžioje sakinys S išskleidžiamas į daiktavardinę ir veiksmažodinę frazes DF ir VF (pagal 1 taisyklę) ir tikrinama, ar jos pasitvirtina sakinyje: pirmiausia apdorojamas kairysis simbolis DF ir žiūrima pagal 2 taisyklę (DF → ART DAIKT), ar sakinys tikrai prasideda artikeliu ir ar po jo eina daiktavardis. Tokia frazė yra galima: ART → the (remiantis 7 taisykle) ir DAIKT → green (8 taisyklė), todėl ji įsimenama ir toliau apdorojamas dar nesukonkretintas simbolis VF. Tikrinama pirmoji jo alternatyva VF → V. Ji pasitvirtina, t.y. trečias žodis sakinyje gali būti veiksmažodis: V → water (11 taisyklė). Taigi, formuojamas sakinys (11 pav.):



11 pav. Sakinio *The green water* sintaksinė struktūra.

Tačiau ne visi sakinio žodžiai dar aptarti, todėl šis variantas atmetamas ir grįžtama prie kitos VF alternatyvos; pasirenkama: VF → PAGV V. Kadangi žodis *water* negali būti PAGV, grįžtama prie kito DF atvejo: DF → ART BŪDV DAIKT. Taip būti gali: ART → the (7 taisyklė), BŪDV → green (10 taisyklė) ir DAIKT → water (9 taisyklė); taigi, vėl žiūrimas pirmas VF variantas: VF → V (4 taisyklė). Jis pasitvirtina: V → evaporated (12

taisyklė) ir šiuo atveju jau nebelieka neapartų žodžių. Tokia sakinio struktūra ir pateikiama kaip analizės rezultatas (12 pav.).

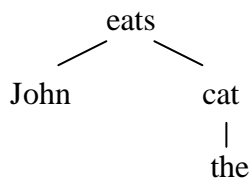


12 pav. Sakinio *The green water evaporated* struktūrinė schema.

*Iš esmės angliško sakinio nagrinėjimas – tai suradimas ribos, kur baigiasi veiksnys (daiktavardis bei jį išplečiantys žodžiai) ir prasideda tarinys. Jie anglų kalboje yra išsidėstę griežtai vienas po kito. Lietuvių kalboje, turinčioje laisvą žodžių tvarką, šituo pasinaudoti visiškai negalima.*

## 2.2.2 Vokiečių kalbos sintaksinė analizė

Kita gramatikos teorija – priklausomybių gramatika – sako, kad sintaksė ne tiek turi grupuoti žodžius į frazes, kiek nustatyti tiesioginius ryšius tarp pačių žodžių [Kay, Gawron, Norvig, psl. 53]. Šioje gramatikoje sakinio sintaksinė struktūra, pavaizduota medžiu, skiriasi nuo frazių metodo. Priklausomybių gramatikos medžio šaknyje stovi kaitomas veiksmožodis, o žemyn išdėstomi jį išplečiantys žodžiai. Tas pats sakinyje *John eats the cat* (žr. psl. 10, 4 pav.) priklausomybių gramatikoje atrodytų taip (13 pav.):

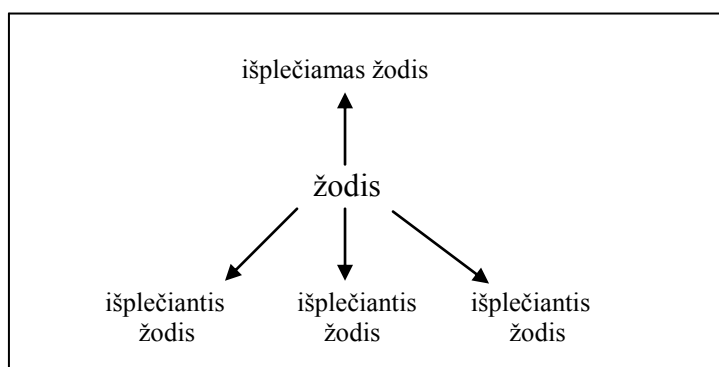


13 pav. Sakinio *John eats the cat* sintaksinė struktūra priklausomybių gramatikoje.

Ši struktūra yra nesusijusi su žodžių tvarka sakinyje ir todėl labiau paplito kalbose, turinčiose laisvą žodžių tvarką. Ja patogiau pasinaudoti automatinio vertimo sistemose, kada galima neatsižvelgti į šaltinio kalbos žodžių išsidėstymą, ir, remiantis vien žodžių priklausomumu, formuoti išverstą sakinį su rezultato kalbai būdinga žodžių tvarka [Kay, Gawron, Norvig, psl. 55).

Du pagrindiniai sintaksinės struktūros pavaizdavimo būdai skiriasi tuo, kad priklausomybių gramatikos medyje kiekvienoje viršūnėje yra sakinio žodis, ir žemiau jo pateikiami jį išplečiantys (patikslinantys, papildantys) žodžiai, o frazių gramatikoje tik medžio lapuose yra žodžiai, virš kurių išsidėstę jų morfologinės kategorijos (daiktavardis, veiksmožodis, būvardis ir t.t.); visose kitose viršūnėse yra sintaksinės kategorijos (daiktavardinė frazė, sakinyje ir pan.) [Winograd, psl. 75].

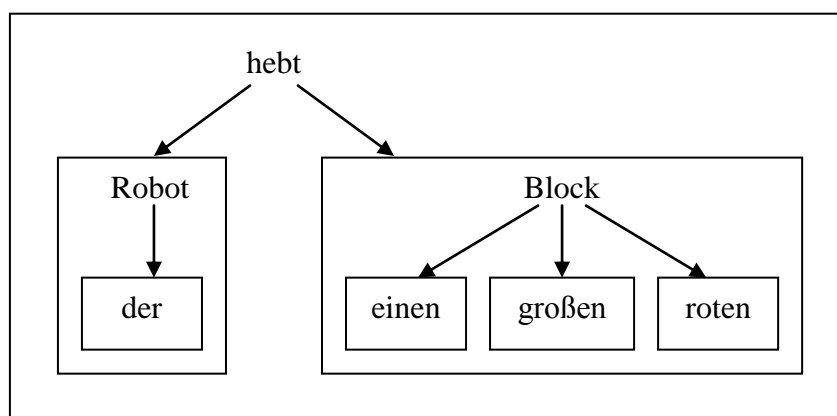
Apibendrintai priklausomybių medžio viršūnę galima pavaizduoti kaip parodyta 14 pav. Kiekvienoje viršūnėje yra žodis, kuris gali turėti vieną ar kelis jį išplečiančius žodžius. Tačiau pats jis tegali išplėsti tik vieną virš jo esantį žodį.



14 pav. Priklausomybių medžio viršūnės ryšiai su kitomis viršūnėmis.

Sintaksinės analizės užduotis surasti kiekvienam sakinio žodžiui visus išplečiančius žodžius ir nustatyti, kurį žodį jis pats išplečia.

Priklausomybių medžiui apdoroti vokiečių mokslininkas P. Hellwig pasiūlė formalų aprašą DUG (Dependency Unification Grammar). Šis aprašas remiasi idėja, kad sakinyje žodį išplečia ne kuris nors vienas atskiras žodis, o visa žodžių grupė, esanti medyje žemiau jo. Tą grupę sudaro vienas tiesioginiu sintaksiniu ryšiu susietas žodis su visais jį patį išplečiančiais žodžiais. Sakinį, tokiu būdu, galima išsivaizduoti kaip sudarytą iš dėžučių, ir kiekvieną žodį išplečia visas dėžučių blokas. Rodyklės čia rodo sintaksinį ryšį ne tarp dviejų žodžių, bet tarp žodžio ir jį išplečiančios grupės (dėžučių bloko). 15 pav. pateikta sakinio *Der Robot hebt einen großen roten Block* sintaksinių ryšių schema, remiantis DUG aprašu [Hellwig, 2003, psl. 13].



15 pav. Sakinio *Der Robot hebt einen großen roten Block* sintaksiniai ryšiai DUG.

Įvesti į kompiuterį grafinį vaizdą yra nepatogu, todėl šiam tikslui naudojamas sąrašas. Kiekvieną sąrašo elementą sudaro sakinio žodis su nurodyta gramatine informacija apie jį. Sąrašas prasideda nuo pagrindinio elemento t.y. tarinio, nes priklausomybių gramatika jį laiko pagrindiniu žodžiu sakinyje. Kiekvienas elementas yra apskliaudžiamas kartu su jį išplečiančiais žodžiais, kurių kiekvienas taip pat yra apskliaudžiamas. Gauta skliaustų išraiška atitinka medį: sąrašo elementai atitinka medžio viršūnes, o skliaustų hierarchija – lankus tarp viršūnių. 16 pav. (psl. 22) pateikiamas sakinio *Der Robot hebt einen großen roten Block* formalus DUG aprašas.

(eilutė[.] funkcija[ne žodis] kategorija[ženklas]

(<eilutė[**hebt**] funkcija[*tarinys*] žodis[*heben*] kategorija[*veiksmazodis*]  
laikas[*esamasis*] rūšis[*veikiamoji*] asmuo[3]

(<eilutė[**Robot**] funkcija[*veiksny*] žodis[*Robot*] kategorija[*daiktavardis*]  
skaičius[*vns.*] giminė[*vyr.*] linksnis[*vard.*]

(<eilutė[**Der**] funkcija[*artikelis*] žodis[*der*] kategorija[*tarnybinis žodis*]  
skaičius[*vns.*] giminė[*vyr.*] linksnis[*vard.*]))

(>eilutė[**Block**] funkcija[*papildinys*] žodis[*Block*] kategorija[*daiktavardis*]  
skaičius[*vns.*] giminė[*vyr.*] linksnis[*gal.*]

(<eilutė[**einen**] funkcija[*artikelis*] žodis[*ein*] kategorija [ *tarnybinis žodis*]  
skaičius[*vns.*] giminė[*vyr.*] linksnis[*gal.*])

(<eilutė[**großen**] funkcija[*pažyminys*] žodis[*groß*] kategorija[*būdvardis*]  
skaičius[*vns.*] giminė[*vyr.*] linksnis[*gal.*])

(<eilutė[**roten**] funkcija[*pažyminys*] žodis[*rot*] kategorija[*būdvardis*]  
skaičius[*vns.*] giminė[*vyr.*] linksnis[*gal.*]))))

**16 pav.** Sakinio *Der Robot hebt einen großen roten Block* aprašas DUG.

Pasukus šį aprašą 90° kampu pagal laikrodžio rodyklę galima nesunkiai atkurti medį, brėžiant lankus tarp kiekvieno sąrašo elemento ir tiesiogiai su juo susijusių žodžių [Hellwig, 2003, psl. 4].

Reikėtų atkreipti dėmesį į tai, kad DUG aprašas yra jau kontekstinės gramatikos atvejis [Hellwig, 2003, psl. 3], nes jame nurodoma, kurioje vietoje yra išplečiantis žodis – prieš išplečiamą žodį ar po jo. Paprastame priklausomybių medyje nėra informacijos apie tai, kokioje vietoje yra žodžiai vienas kito atžvilgiu. DUG apraše išplečiantis žodis pažymimas ženkleliu “<”, jei jis yra į kairę, ir ženkleliu “>”, jei jis yra į dešinę nuo išplečiamo žodžio. Taigi, ženklo smaigalys yra nukreiptas į išplečiantį žodį. Jei jų toje pačioje pusėje yra daugiau negu vienas, sąrašė jie išdėstomi tokia pat tvarka, kaip ir pavartoti sakinyje [Hellwig, 2003, psl. 3].

*Negalima pasakyti, kad toks aprašas labai gerai atspindėtų lietuvių kalbos savybes. Daugumai žodžių lietuvių kalboje ženklo smaigalio kryptis negalėtų palengvinti sintakšinės analizės, nes lietuviškame sakinyje išplečiantis žodis iš esmės gali būti bet kurioje pusėje, t.y. ir prieš išplečiamą žodį ir po jo.*

Pagrindinis DUG tikslas – sudaryti formalią sakinio struktūrą. Norint tai pasiekti, reikia surasti visus ryšius tarp žodžių tame sakinyje, t.y. nustatyti, kas išplečia kiekvieną žodį ir kurį žodį išplečia pats aprašomasis žodis.

Tam tikslui sudaromos dviejų rūšių duomenų bazės: pirmoje – sukaupiama abstrakti informacija apie visus galimus toje kalboje sintaksinius ryšius, o antra bazė skirta konkrečių žodžių aprašymui. Joje nurodomos žodžio junglumo galimybės, t.y. su kokio tipo kitais žodžiais jis gali sudaryti junginius. Informacija abiejose duomenų bazėse taip pat pateikiama sąrašais.

Sintaksiniai ryšiai traktuojami kaip modeliai, kurie gali būti užpildyti konkrečiais sakinio žodžiais. Kiekvienas modelis atitinka tik vieną lanką priklausomybių medyje ir aprašo vieną sintaksinių ryšių tipą, pvz., ryšys tarp veiksnio ir tarinio, ryšys tarp tarinio ir papildinio ir pan. Jei žodis gali būti išplėstas daugiau nei vienu žodžiu, jis turi būti aprašytas keliais modeliais: po vieną kiekvienam išplečiančiam žodžiui. Pvz., veiksmažodžiui *heben* reikalingi trys modeliai: vienas jų turi aprašyti šio žodžio ryšį su veiksmiu, antras – su papildiniu, trečias – su aplinkybe. Visi jie parodyti 17 pav.

(šablonas[*veiksnys*] + kategorija[*veiksmažodis*] asmuo[3]  
(dėžutė[*būtina*] funkcija[*veiksnys*] kategorija[*daiktavardis*] linksnis[*vard*]))

(šablonas[*papildinys*] + kategorija[*veiksmažodis*] asmuo[3]  
(dėžutė[*būtina*] funkcija[*papildinys*] kategorija[*daiktavardis*] linksnis[*gal*]))

(šablonas[*aplinkybė*] + kategorija[*veiksmažodis*] asmuo[3]  
(dėžutė[*nebūtina*] funkcija[*aplinkybė*] kategorija[*prieveiksmis*]))

**17 pav.** Veiksmažodžio *heben* sintaksinius ryšius atspindinčių šablonų rinkinys.

Po modelio pavadinimo, atitinkančio kurią nors išplečiančią sakinio dalį (veiksnys, papildinys ir pan.) nurodomas tiesioginiu sintaksiniu ryšiu susietas žodis, kurį jis išplečia (t.y. priklausomybių medyje tiesiogiai virš jo esantis žodis). Antroje eilutėje aprašomas pats modelio pavadinime esantis žodis, nurodant jo sintaksinę funkciją ir morfologines kategorijas. Pagal analogiją su sintaksinės struktūros vaizdavimo principu DUG, jo pozicija įvardijama kaip „dėžutė“ ir pateikiama jos charakteristika reikalingumo požiūriu, pvz., nurodoma, ar veiksmažodis būtinai reikalauja tokio papildymo, ar šis papildymas yra fakultatyvus. Iš aprašo (pav.17) matyti, kad žodžiui *heben* yra būtini sintaksiniai ryšiai su dviem sakinio dalimis – veiksmiu ir papildiniu (svarbu, kas kelia ir ką kelia), jei bent vieno iš jų nebus, sakinyš atrodys neužbaigtas; o aplinkybės gali ir nebūti – net jei ir nepasakysime kaip kelia (sunkiai, lėtai ir pan.), sakinyš bus visai geras ir neatrodys, kad jam ko nors trūksta.

Jei ta pati sakinio dalis (pvz., veiksnys) gali būti išreikšta keliomis kalbos dalimis (pvz., daiktavardžiu ir įvardžiu), kiekvienai iš jų sudaromas atskiras modelis. Alternatyvus veiksnio modelis pateikiamas 18 pav.

(šablonas[*veiksnys*] + kategorija[*veiksmažodis*] asmuo[3]  
(dėžutė[*būtina*] funkcija[*veiksnys*] kategorija[*įvardis*] linksnis[*vard*]))

**18 pav.** Alternatyvus veiksnio šablonas.

Kitoje duomenų bazėje kiekvienam konkrečiam žodžiui sukaupiamas modelių rinkinys, atitinkantis visus galimus jo sintaksinius ryšius. Tas rinkinys vadinamas žodžio sintaksiniu rėmu. Sudarant jį, tikrinami visi modeliai, t.y. visi galimi toje kalboje sintaksiniai ryšiai, ir atrenkami tik tie, kuriais gali būti susijęs aprašomasis žodis. Veiksmažodžio *heben* ir daiktavardžio *Robot* sintaksiniai rėmai pateikiami 19 pav.

(žodis[ <i>heben</i> ] kategorija[ <i>veiksmažodis</i> ] (išplėtimas[ <i>veiksnys</i> ]) (išplėtimas[ <i>papildinys</i> ]) (išplėtimas[ <i>aplinkybė</i> ]))  (žodis[ <i>Robot</i> ] kategorija[ <i>daiktavardis</i> ] (išplėtimas[ <i>pažyminys</i> ]))
--

19 pav. Žodžių *heben* ir *Robot* sintaksiniai rėmai.

Atliekant sakinio *Der Robot hebt einen großen roten Block* sintaksinę analizę, iš sąrašų, pateiktų 17 pav. (psl. 23) ir 19 pav., gaunamas pradinis sakinio sąrašas, kuris parodytas 20 pav.:

(eilutė[ <i>hebt</i> ] žodis[ <i>heben</i> ] kategorija[ <i>veiksmažodis</i> ] asmuo[3] (dėžutė[ <i>būtina</i> ] funkcija[ <i>veiksnys</i> ] kategorija[ <i>daiktavardis</i> ] linksnis[ <i>vard.</i> ]) (dėžutė[ <i>būtina</i> ] funkcija[ <i>papildinys</i> ] kategorija[ <i>daiktavardis</i> ] linksnis[ <i>gal.</i> ]))
---

20 pav. Pradinis sakinio sąrašas, gautas atliekant *Der Robot hebt einen großen roten Block* sintaksinę analizę.

Analogišku būdu sujungus artikelius ir būdvardžius su daiktavardžiais, kuriuos jie išplečia, gaunamos frazės *der Robot* ir *einen großen roten Block*. Jos turi tuos pačius požymius, kaip ir dėžutės, aprašytos 17 pav. (psl. 23). Taigi, į dėžučių vietą įrašomos jas atitinkančios frazės ir gaunamas viso sakinio struktūros aprašas, parodytas 16 pav. (psl. 22).

*Vokiečių kalbos sintaksinė analizė remiasi tomis pačiomis kalbos savybėmis kaip ir anglų: griežtai nustatyta tarinio vieta sakinyje, prieš jį eina veiksnys, o po jo papildiniai bei aplinkybės. Ir anglų, ir vokiečių kalbose tarinys būtinai turi stovėti antroje vietoje, taigi, jį gana nesunkų surasti kompiuteriu. Todėl sakinio nagrinėjimas ir pradedamas nuo tarinio. Modeliuose pagal jį orientuojamasi į kitų žodžių vietą. Aplinkybė turi būti dešinėje tarinio pusėje (t.y. po jo), veiksnys dažniausiai būna kairėje (prieš tarinį), tik inversijos atveju jis gali susikęsti vietomis su papildiniu, tačiau tarinio vieta lieka nepajudinama. Kadangi lietuvių kalboje nėra griežtai fiksuotos tarinio (o taip pat ir kitų sakinio dalių) vietos sakinyje, todėl tiesiogiai perėmus šią metodiką, analizė nebūtų labai efektyvi.*



### 2.2.3 Rusų kalbos sintaksinė analizė

Pirmosios rusų kalbos sintaksinės analizės sistemos buvo kuriamos naudojantis frazių gramatika. Rusų mokslininkai rėmėsi tuo, kad automatinį tekstų vertimą galima atlikti frazių pagrindu [Истомин, Петровская, Печников, psl. 81], o svarbiausia sintaksinės analizės pritaikymo sritis tada ir buvo automatinis vertimas.

Rusų kalbos sintaksinei analizei naudojamos bazinės struktūros (BS). Bazinė struktūra – tai gramatiškai apiformintas šaltinio kalbos vienetas, kuris turi atitikmenį rezultato kalboje. Ne visos bazinės struktūros yra lygiavertės. Kai kurios BS yra elementarios ir jų pagalba formuojamos sudėtingesnės. Taip pat nelygiaverčiai yra ir ryšiai, tiek tarp pačių struktūrų, tiek jų viduje. Tai ir nulėmė BS hierarchinį pasiskirstymą: išskiriami penki lygiai [Билан, psl. 176]:

- 1) žodžių formos,
- 2) žodžių grupės,
- 3) žodžių junginiai,
- 4) sakinio dalys,
- 5) struktūriniai sakinių tipai.

Žodžių formų (1) lygmenyje rusų kalbai nustatytos septynios vardažodinės formos (daiktavardis, asmeninis įvardis, kiekinis skaitvardis ir kt.) ir penkios veiksmažodinės formos:

- 1a) asmenuojamoji veiksmažodžio forma,
- 1b) trumpoji būdvardžio forma,
- 1c) žodžiai *можно, нужно, надо*,
- 1d) žodis *хотел бы*,
- 1e) žodžiai *нет, нельзя*

Formų lygmens bazinės struktūros yra elementarios. Jų pagalba formuojamos BS kituose lygmenyse. Iš žodžių formų sudaromos grupių lygmens (2) bazinės struktūros: vardažodinės grupės, veiksmažodinės grupės ir prielinksninės-verdažodinės grupės.

Kiekvienai aukštesnio lygmens bazinei struktūrai formuoti surašomos taisyklės, koku būdu ji yra gaunama iš žemesnio lygmens bazinių struktūrų. Vardažodinė grupė gali būti sudaryta trimis būdais:

- 2a) viena vardažodinė forma (*вокзал, он, десять, никто*),
- 2b) vardažodinė forma su prieš ją einančiu pažyminiu, kuris gali būti išreikštas savybiniu bei parodomuoju įvardžiu, būdvardžiu, kelintiniu skaitvardžiu, prieveiksniu (*ваш багаж, этот музей, ближайшая остановка, пятый этаж, мало книг*),
- 2c) vardažodinė forma, prieš kurią eina keli pažyminiai (*ближайшая заправочная станция, очень мало книг, этот красивый парк*).

Veiksmažodinių grupių galimi du atvejai:

- 2d) viena veiksmažodinė forma (*живет, рад, был*),
- 2e) veiksmažodinė forma, po kurios eina bendratis ar kelios bendratys (*буду разговаривать, хочу начать заниматься*).

Prielinksninę grupę (2f) sudaro prielinksnis ir po jo esanti vardažodinė grupė (*в городе, на новой улице, к ближайшей заправочной станции*).

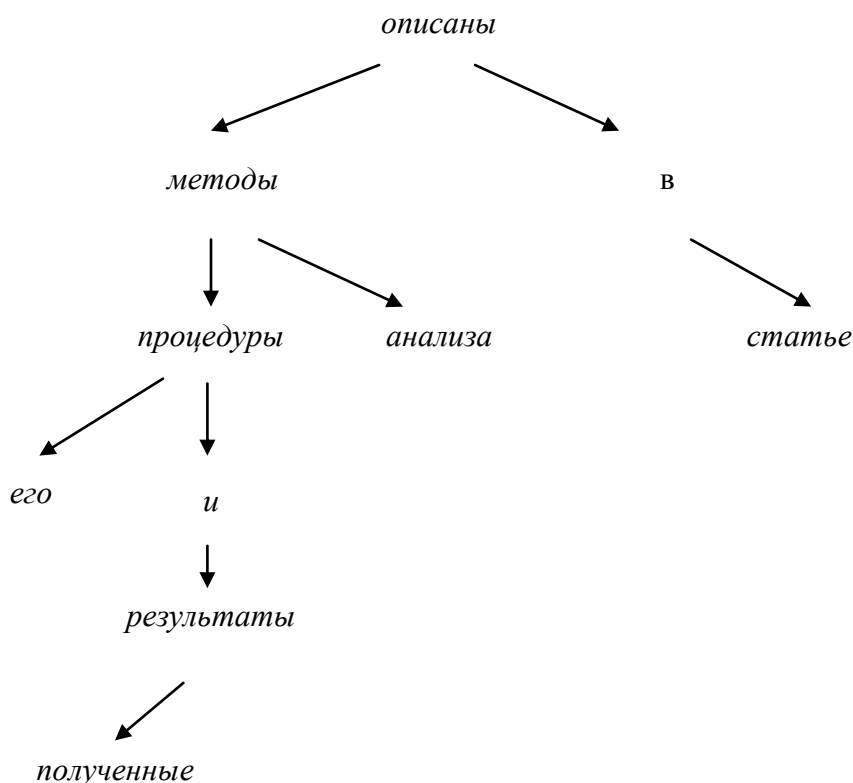


*Toks aprašas taip pat nėra Bekaus ir Nauro forma (BNF). Todėl, matyt, logiška pasekmė ir buvo tas faktas, kad naujausiose rusų kalbos automatinio vertimo sistemose sintaksinę analizę atlieka grupės algoritmu.*

Paplitus Europoje priklausomybių gramatikai, kuri gerai atspindi kalbų su laisva žodžių tvarka savybes, rusiškų sakinių struktūrai taip pat imta naudoti priklausomybių medį. Įdomu yra tikrai tai, kad rusų kalbos sintaksėje išsigalėjo tradicija sudaryti priklausomybių medį binariniu principu, t.y. su sąlyga, kad iš vienos viršūnės tegali išeiti ne daugiau, kaip du lankai. Pvz., sakinio *В статье описаны методы анализа, его процедуры и полученные результаты* sintaksinė struktūra atrodytų taip, kaip pateikta 22 pav.

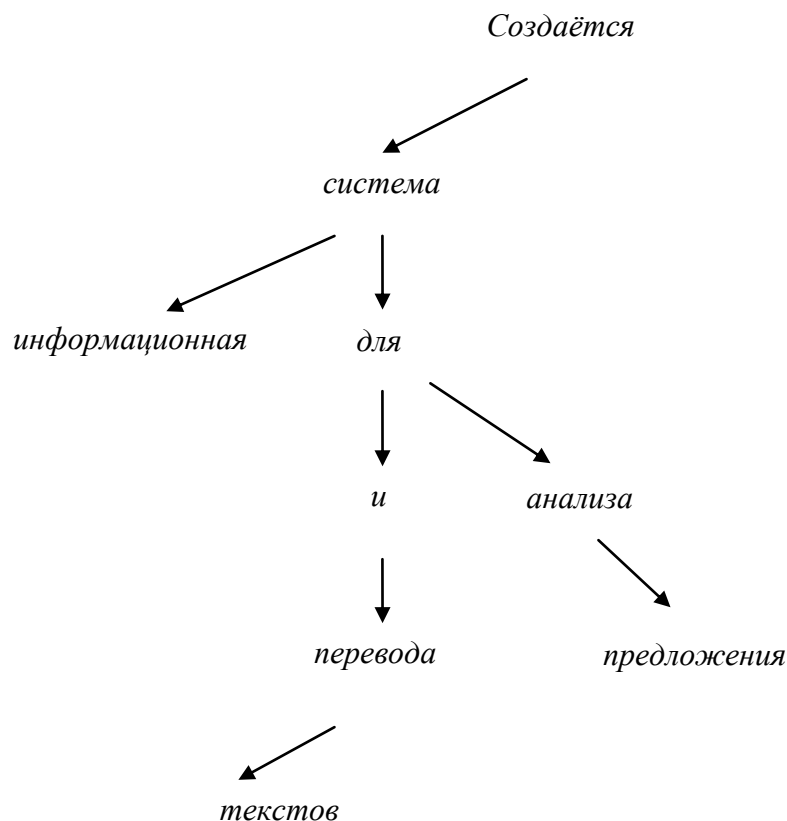
Tokiu principu braižant sakinio struktūrą, ypatingu būdu tenka traktuoti sujungiamąjį ryšį tarp žodžių (tai toks sintaksinis ryšys, kai visi jo nariai yra lygiaverčiai: pvz., sakinyje *Аš gerai prisimenu tą ankstyvą pavasario rytą* visi trys žodžiai: *тą, ankstyвą* ir *павасаріо*, yra lygiaverčiai, jie atlieka pažyminio funkciją – pažymi žodį *рытą* – taigi sakinio požiūriu jie yra to paties rango). Net sujungiamasis jungtukas (pvz., *ir*) binariniame medyje laikomas lygiaverčiu su kitais žodžiais, galinčiais eiti sakinio dalimi [Грязнухина, psl. 10].

Jei yra keli vienaarūšiai, priklausomi nuo to paties prielinksnio (sakykim, *для*) žodžiai, pvz., sakinyje *Создаётся информационная система для анализа предложения и перевода текстов*, sintaksinėje struktūroje parodomas prielinksnio (*для*) tiesioginis ryšys tik su pirmu priklausomu žodžiu (*анализа*), o antras priklausomas žodis (*перевода*) jau pateikiamas kaip susijęs ne su prielinksniu (*для*), o su jungtuku (*и*), kuris medyje tiesiogiai priklauso nuo prielinksnio (*для*) [Грязнухина, psl. 136].



**22 pav.** Sakinio *В статье описаны методы анализа, его процедуры и полученные результаты* struktūra.

Sakinio *Создаётся информационная система для анализа предложения и перевода текстов* sintaksinė struktūra pateikiama 23 pav.



23 pav. Sakinio *Создаётся информационная система для анализа предложения и перевода текстов* struktūra.

Atliekant rusų kalbos sintaksinę analizę daug problemų iškyla dėl laisvos žodžių tvarkos sakinyje [Грязнухина, psl. 6]. Ypač sunku nustatyti prielinksnio ryšių zoną, nes tiesiogiai susiję žodžiai gali būti išsidėstę labai toli vienas nuo kito. [Грязнухина, psl. 131]. Ukrainos mokslų akademijoje atliktų tyrimų metu **didžiausias užfiksuotas atstumas tarp priklausomų žodžių buvo 28, t.y. tarp dviejų tiesioginiu sintaksiniu ryšiu susietų žodžių buvo įsiterpę 28 kiti žodžiai** [Грязнухина, psl. 136].

Šios priežastys matyt ir lėmė tai, kad viena naujausių rusų kalbos sintaksinės analizės sistemų, sukurta Kijeve, nesiremia formalios nekontekstinės gramatikos aprašu ir sakinio analizę atlieka grupės algoritmų. Pvz., pirmo etapo metu šioje sistemoje dirba 8 algoritmai. Jų užduotis – surasti žodžius, pretenduojančius į veiksnio bei tarinio funkciją.

Antrame etape nustatomi ryšiai tarp pagrindinių sakinio dalių (veiksnių ir tarinių). Po to, prie segmentų, turinčių veiksnį su tariniu, prijungiami kiti segmentai, kuriuose šios sakinio dalys nebuvo rastos. Antro etapo rezultatai – nustatyta sakinio struktūra: jei randama daugiau, nei vienas veiksnys su tariniu, konstatuojama, kad tai nėra vientisinis sakinytis.

Trečias etapas skirtas apdoroti sudėtinius sakinius ir galutinai fiksuoti sakinio struktūrą.

Ketvirtame etape atliekama analizė vientisinio sakinio viduje: analizuojamos išskirtinės dalyvinės bei kitos konstrukcijos, nustatoma skyrybos ženklų funkcija (ar jie skiria vienuose sakinio dalis, ar išskirtinę dalyvinę aplinkybę ir pan.). Šio etapo rezultatas yra nagrinėjamo sakinio sintaksinę struktūrą atitinkantis priklausomybių medis. Ketvirtame etape dirba 7 algoritmai, kurių darbą koordinuoja vienas valdantis algoritmas. Jis peržiūri iš eilės visus sakinio žodžius iš kairės į dešinę ir sprendžia, kaip jie turi būti apdorojami. Sukaupia informacija apie jau išnagrinėtus žodžius panaudojama likusios sakinio dalies analizei [Грязнухина, psl. 29].

Šio etapo algoritmų atliekamo darbo pavyzdys galėtų būti žodžių, priklausomų nuo prielinksnio *между*, suradimas. Pirmą (priklausomą nuo šio prielinksnio) žodį suranda dalinis algoritmas, kuris vadinasi „priklausomo žodžio paieška“. Antro priklausomo žodžio ieško pats valdantysis ketvirto etapo algoritmas, tikrindamas kiekvieną sutiktą daiktavardį įnagininko linksnyje pagal tris hipotezes:

- 1) ar šis daiktavardis nepriklauso nuo betarpiškai jam iš kairės stovinčio veiksmažodžio (*управлять программой*);
- 2) ar jis neįeina į vardažodinę grupę, ir ar nepriklauso nuo vieno iš šios grupės narių (*организация вызова нужного алгоритма управляющей программой*);
- 3) ar kairiajame kontekste nėra daiktavardžio, priklausomo nuo prielinksnio *между*.

Pagal trečią hipotezę tikrinami tiksliai tie žodžiai, kurių eilės numeris sakinyje yra mažesnis už nagrinėjamo žodžio numerį [Грязнухина, psl. 137].

*Синтаксinės analizės realizacija algoritmu pagrindu, žinoma, gali duoti neblogus rezultatus net ir tokiais kalbais, kaip rusų kalba, kurios sintaksė pasižymi dideliu sudėtingumu („...русского языка, синтаксический строй которого характеризуется большой степенью сложности“ [Матвеева, psl. 83].*

*Tačiau, jei pavyksta programuojamą objektą aprašyti antro tipo formalia gramatika, labai supaprastėja programavimas. Todėl šiame darbe siekiama lietuvių kalbos sintaksę pateikti nekontekstinės gramatikos taisyklėmis (BNF).*

## 2.3 Statistiniai metodai

Praeito amžiaus aštunto dešimtmečio pabaigoje rusų literatūroje aprašomi bandymai atlikti sakinio automatinę sintaksinę analizę statistinių skaičiavimų pagrindu.

Rusijos mokslininkai siūlė sakinio sintaksinę struktūrą sudaryti remiantis ne lingvistiniais tyrinėjimais, o statistiniu tekstų apdorojimu. Jie teigė, kad bet kuris kalbinis vienetas (žodis, žodžio gramatinė kategorija, sintaksinė konstrukcija) sutinkamas tekste tam tikru dažniu. Pvz., anglų kalbos dažninis junglumo žodynas pateikia duomenis, kad veiksmažodžiui būdingi 195 skirtingi sintaksinių ryšių modeliai ir apie 10 dažniausiai pasitaikančių modelių sudaro 86% visų tekstuose pavartotų atvejų.

Statistiniai dėsningumai taikomi ir linijinei sakinio sandarai. Kiekviena žodžių klasė turi savo dažnį, kuriuo ji tam tikru atstumu gali būti nutolusi nuo pradinio taško. Anglų kalbos veiksmažodis su tikimybe 0,7 būna 2-5 pozicijoje nuo sakinio pradžios, o su tikimybe 0,95 sutinkamas 2-10 pozicijoje [Сердюков, psl. 124].

Remiantis šiais duomenimis sudaromas veiksmažodžio sintaksinių ryšių paieškos algoritmas. Jis dirba spėjimo principu. Atsižvelgiant į veiksmažodžio tipą bei jo aplinką, sakinio ilgumą ir kt., tam tikrose sakinio vietose (sakinio pradžioje, tarinio bei nuo jo priklausančių žodžių pozicijose) daromos prielaidos apie galimus analizės rezultatus. Tikėtini atvejai išdėstomi dažnio mažėjimo tvarka.

Jei analizės taškuose pateikti spėjimai sutampa (pvz., kai sakinyje susideda iš dviejų žodžių ir pirmoje vietoje daiktavardžio tikimybė didesnė nei veiksmažodžio, o antroje vietoje veiksmažodžio tikimybė didesnė nei daiktavardžio, daroma išvada, kad pirmasis žodis yra veiksnys, o antrasis – tarinys), išskiriama su veiksmažodžiu susijusių žodžių grupė. Po to nustatomi ryšiai tarp veiksmažodžio grupės žodžių ir likusių žodžių sakinyje. Duomenis apie sakinio dalių tikimybinį išsidėstymą galima panaudoti bet kokiame sintaksinės analizės algoritme ir tai turėtų padidinti to algoritmo efektyvumą [Сердюков, psl. 125].

*Kad tikimybinis sakinių dalių išsidėstymas paremta sintaksine analize duotų lietuvių kalbai labai gerų rezultatų yra mažai tikėtina, nes, esant laisvai žodžių tvarkai sakinyje, tikimybė, kad kuris nors žodis bus pradžioje, gale ar viduryje sakinio, turėtų skirtis nedaug, priešingai negu anglų kalboje, kur jau pačios gramatikos taisyklės reikalauja, kad tarinys būtinai stovėtų antroje vietoje, dėl ko ir tarinio išsidėstymo tikimybė nuo 2 iki 5 pozicijos čia yra labai didelė. (Antroje ar penktoje vietoje yra tarinys, priklauso nuo veiksmų išplečiančių žodžių kėkčio: artikelio, būdvardžių ir kt.)*

***Lietuvių kalboje žodžių tvarkos pakaitimas sakinyje dažnai vartojamas kaip sinoniminis variantas, siekiant išvengti monotoniškumo, todėl pagal žodžių išsidėstymą spręsti apie jų sintaksines funkcijas mažai teigama.***

Šiame darbe nebus naudojami tikimybiniai apskaičiavimai. Gramatinės analizės metodu dirbančiose automatinio vertimo sistemose modeliuojamas žmogaus smegenų darbas. Todėl ir sintaksinei analizei pasirenkama būtent ši metodika. Reikia pasakyti, kad gamtos modelių panaudojimas pasitvirtina ir kitose mokslo srityse. Pastaruoju metu buvo atkreiptas dėmesys į tai, kad technikos problemas galima spręsti pasitelkiant pavyzdžius iš gamtos [Bionik, Internetas]. Sparčiai besivystanti mokslo šaka – bionika (biologija ir technika) pateikia daugelį atvejų, kada technikos išradimai buvo atlikti atidžiai stebint gamtą. Pvz., Hanoveryje padangų gamintojai pavyzdžiu pasirinko katę. Ji gali padidinti savo letenos plotą trečdaliu ir tuo pasiekia didesnę sukibimą su paviršiumi, į kurį remiasi. Pasinaudojant šiuo principu sukurtos padangos stabdymo metu prie žemės prispaudžia didesnę gumos kiekį [Patschull, psl. 11]. Ištyrus žiurkių dantis, kurie neatšimpa graužiant ne tik medį, bet ir metalą ar betoną, buvo pradėti gaminti savaime pasigalandantys įrankiai [Deutschland, psl. 12]. Pastebėjus lotoso žiedų bei lapų savybę atstumti dulkes ir vandenį (jie visada būna stebinančiai švarūs ir nesušlampa), buvo sukurti lietaus nepraleidžiantys sienų dažai „Lotusan“ [Deutschland, psl. 13]. Galima dar būtų aprašyti ir daugelį kitų pavyzdžių, todėl šiame darbe daroma prielaida, kad ir sintaksinei analizei bei automatiniam vertimui labiau tinka žmogaus smegenų darbą modeliuojanti metodika.

## **2.4 Lietuvoje atlikti darbai**

Kauno technologijos universitete vykdomas projektas „Valdomas kompiuterinis vertimas: projektavimas, programinės įrangos kūrimas ir tyrimas“, kurio tikslas – sukurti pagalbinę kompiuterinę priemonę vertimui iš anglų kalbos į lietuvių kalbą ir atvirkščiai (Tamulynas, psl. 3). Kompiuterizuotam vertimui reikalingas dvikalbis žodynas, kuriame būtų saugoma daugiau morfologinės, sintaksinės ir semantinės informacijos, nei tradiciniuose, knygos pavidalu išleistuose, dvikalbiuose žodynuose. Todėl reikėjo sukurti žodžio „įrašo dokumentaciją“. Darbe remiamasi P. Pacevičiaus pasiūlyta žodyno įrašų struktūra, ją papildant ir išplečiant. Apibendrinta ir universali žodžio struktūra suskirstyta į tris dalis:

- 1) bendroji dalis, kurioje pateikiami bendri anglų ir lietuvių kalbų žodžiams požymiai (pvz., kokios kalbos dalys priklauso žodis, ar daiktavardis yra tikrinis, ir pan.),
- 2) anglų kalbos dalis, kurioje pateikiama informacija apie anglų kalbos žodį (pvz., daiktavardžiams nurodomos vienaskaitos ir daugiskaitos formos; būdvardžiams – nelyginamojo, aukštesniojo ir aukščiausiojo laipsnio formos; veiksmažodžiams – esamasis, būtasis laikas, dalyvis ir gerundijus (-ing galūnė) ir t.t.),
- 3) lietuvių kalbos dalis, kurioje pateikiami duomenys apie lietuvių kalbos žodį (pvz., linksniuotė, asmenuotė, valdymas, ir kt.)

Jei anglų kalbos žodis turi kelias reikšmes lietuvių kalboje, jis turi būti įtraukiamas į žodyną tiek kartu, kiek yra kitos kalbos atitikmenų. Žodyno įrašo pavyzdys pateiktas 24 pav.

Access (n/daiktav.) Access (v/veiksmaž.)	prieiga (daiktav.) prieiti (veiksmaž.) „Access Asset Management“ = „Prieiga prie turto valdymo“ „Access Accounts Receivable“ = „Prieiga prie gautinų sumų modulio“
---	---

24 pav. Žodyno įrašo pavyzdys

Žodžių daugiareikšmiškumo problema sprendžiama dviem būdais:

- 1) pateikiant vartotojui visą ar dalinę reikšmių sąrašą,
- 2) išrenkant dažniausiai vartojamą variantą.

Semantinė informacija koduojama žodyne taip, kaip parodyta 25 pav. pavyzdyje su žodžiais *vyras*, *moteris*, *berniukas* ir *mergaitė*:

<i>vyras</i>	= (+ žmogus, + vyr.gim., + suaugęs)
<i>moteris</i>	= (+ žmogus, – vyr.gim., + suaugęs)
<i>berniukas</i>	= (+ žmogus, + vyr.gim., – suaugęs)
<i>mergaitė</i>	= (+ žmogus, – vyr.gim., – suaugęs)

25 pav. Semantinės informacijos kodavimo pavyzdys

Kol kas **nėra numatyta saugoti žodyne sakinio dalis**, nes tai pareikalautų daugiau kalbininko (lingvisto) laiko sąnaudų ir padidėtų tikimybė, kad bus įvelta klaidų, kurių reikia vengti visuose kompiuterizuoto vertimo sistemos lygiuose.

Pirmąjį informacinių technologijų srities 77 terminų žodynelio variantą sudaro daiktavardžiai, būdvardžiai, veiksmažodžiai ir žodžių junginiai.

Žodynas yra realizuojamas Java kalboje, o duomenų bazė, kurioje bus saugomas visas žodynas, administruojama MySQL. Vartotojo sąsaja realizuojama naudojant standartinius Java Swing elementus. Pagrindinė sistemos funkcija yra naujų žodžių įvedimas ir senų redagavimas bei šalinimas.

Kompiuterizuoto vertimo duomenų bazės saugykla turės prieigą per Internetą, o pati kompiuterizuoto vertimo sistema bus vartotojo kompiuteryje. Tokiu būdu prie duomenų bazės galės jungtis daug vartotojų. Be to, jai esant serveryje vartotojas gali būti tikras, kad turi naujausią žodyno versiją.

Paieška žodyne trunka nuo 70 μs iki 1 ms.

Jei pateikiamas žodis, kurio žodyne nėra, sistema apie tai praneša vartotojui, pasiūlydama įvesti naują žodį. Įvedant žodį tikrinama ar nėra neleistinų simbolių. Neleistinu laikomas toks simbolis, kuris nepriklauso lietuvių ir anglų kalbų abėcėlėms; žodžio pradžioje ir pabaigoje tarpas taip pat yra neleistinas simbolis. Įvedus nekorektišką žodį, visas laukas nudažomas raudonai, o jį pataisius laukas vėl pasidaro baltas.

Reikia pasakyti, kad žodynas yra stambiausia kompiuterizuoto vertimo sistemų dalis pagal saugomos informacijos apimtį ir turi didžiausią įtaką visos sistemos funkcionavimui. Per jį vartotojas gali turėti įtakos vertimo procesui, t.y. pildyti žodžių bazę, generuoti įvairius vertimo variantus.

## 2.5 Skyriaus išvados

- Skyriuje apžvelgti darbai, susiję su automatine anglų, vokiečių ir rusų kalbų sintaksine analize.
- Nurodytos priežastys, kodėl lietuvių kalbai negalima pritaikyti jau sukurtų ir kitoms kalboms naudojamų automatinės sintaksinės analizės metodikų.
- Anglų kalbos sintaksinė analizė remiasi išimtinai vien žodžių tvarka sakinyje, todėl lietuvių kalbai negali būti taikoma, nes lietuviškame sakinyje žodžio vieta negali nulemti jo sintaksinės funkcijos ir čia daugiausia įtakos turi žodžių formos (jų galūnės).
- Vokiečių kalboje žodžių tvarka yra šiek tiek laisvesnė (papildinys gali pasikeisti vietomis su veiksmu), bet išlieka griežtas reikalavimas, kad tarinys būtų antroje vietoje, kaip ir anglų kalboje. Todėl ir vokiečių kalbos sintaksinės analizės sistemose orientuojamasi į tarinio vietą sakinyje ir pagal tai sprendžiama apie likusių žodžių sintaksines funkcijas. Lietuvių kalboje tarinys (taip pat ir jokia kita sakinio dalis) neturi fiksuotos vietos sakinio pradžios ar galo atžvilgiu, todėl ir vokiečių kalbos sintaksinės analizės metodika negali būti tiesiogiai naudojama lietuvių kalbai.
- Rusų kalba savo struktūra (kaitomumo laipsniu, žodžių tvarkos laisvumu) yra šiek tiek artimesnė mums, nei germanų kalbos, tačiau sintaksinė analizė čia atliekama algoritmų pagrindu ir nesinaudojama formaliu sintaksės taisyklių aprašu. Turint sprendžiamos problemos aprašą BNF labai supaprastėja jos programinė realizacija, todėl šiame darbe rusų kalbai taikomos metodikos taip pat atsisakoma ir keliamas uždavinys aprašyti lietuvių kalbos sintaksę nekontekstinės gramatikos taisyklėmis.
- Lietuvoje atliekamo „Valdomo kompiuterinio vertimo“ projekte nenumatytas sintaksinės informacijos pateikimas apie žodį ir tai rodo šio darbo aktualumą.



### 3 KAI KURIOS LIETUVIŲ KALBOS YPATYBĖS

Dideli skirtumai tarp lietuvių kalbos, kuri laikui bėgant mažai keitėsi (iš visų indoeuropiečių kalbų savo struktūra ji yra artimiausia mūsų prokalbei sanskritui: "...lietuvių kalbos morfologija labai artima sanskritui." [Mironas, psl. 302]), ir modernių šio meto kalbų (tokių kaip anglų kalba ir kitos) trukdo tiesiogiai perimti jau sukurtas kitoms kalboms sintaksinės analizės sistemas. Apibendrintai galima pasakyti, kad yra du pagrindiniai lietuvių kalbos bruožai, neleidžiantys mums pasinaudoti kitų šalių patyrimu – tai didelis žodžių kaitomumas ir laisva žodžių tvarka sakinyje.

Būdinga lietuvių kalbos savybė yra tai, kad ji yra kaitoma ir žodžių formų turi labai daug. Visos pagrindinės kalbos dalys turi po 24-28 formas. Pvz., angliškas žodis *two* jau vien vardininko linksnyje turi penkis atitikmenis. Žodžių junginiai *two boys*, *two girls*, *two years*, *two doors* ir *two horses* verčiami *du berniukai*, *dvi mergaitės*, *dveji metai*, *dvejos durys* ir *dvejetas arklių*. Kiekviena ši forma kaitoma šešiais linksniais – štai ir gauname 30 skaitvardį *two* atitinkančių formų. V. Labutis sako: „Žodžių formų gausumas lietuvių kalboje ir nulėmė tai, kad jos tapo svarbiausia sakinio sintaksinių ryšių raiškos priemone. Kiekviena kalba turi rinkinį gramatinių ar kitokių priemonių, kurios rodo tiesioginį prasminį žodžių ryšį. Dažniausiai tai žodžių formos, tarnybiniai žodžiai (prielinksniai, jungtukai) ir žodžių tvarka sakinyje. Skirtingose kalbose pirmenybė teikiama ne toms pačioms žodžių ryšių raiškos priemonėms. Pavyzdžiui, anglų kalbai svarbiausia žodžių tvarka, po to tarnybiniai žodžiai ir tik paskiausiai žodžių formos. Lietuvių kalboje ryšiai tarp žodžių pirmiausia rodomi žodžių formomis (ypač linksniais), po to tarnybiniais žodžiais ir tik paskutinėje vietoje žodžių tvarka“ [Labutis, 2002, psl. 25].

Lietuvių kalboje žodžių tvarkos reglamentuojamų atvejų yra labai nedaug: neiginys eina prieš neigiamą žodį (*ne visi* taip gali); prielinksnis eina prieš linksnį (mes išėjome *į mišką*); klausiamoji dalelytė visada būna sakinio pradžioje (*Ar visi susirinko?*); jei veiksnyis ir tarinio vardinė dalis išreiškiami tomis pačiomis morfologinėmis formomis (daiktavardžio vardininku, bendratimi), veiksnyis eina prieš tarinį [Ambrasas, psl. 646] – ko gero, būtų ir viskas vientisiniame sakinyje. O žodžių formomis reiškiami beveik visi sintaksiniai ryšiai.

***Automatinio vertimo sistemose, sukurtose kalboms su griežta žodžių tvarka sakinyje (anglų, vokiečių), visiškai nenumatytas sintaksinės informacijos paėmimas iš žodžių galūnių. Visus reikalingus sintaksinei analizei duomenis šios sistemos gauna iš žodžių tvarkos, kuri lietuvių kalboje kartais gali suteikti sakiniui šokią toki stilistinę atspalvį, pabrėžti vieną ar kitą žodį, tačiau jokių būdu negali nulemti jo sintaksinės funkcijos: žodžio vieta lietuviškame sakinyje beveik visiškai neturi sintaksinės informacijos.***

Ši teiginį galima pailiustruoti tokiu pavyzdžiu: du lietuviški sakiniai *Vaikas valgo obuolį* ir *Obuolį valgo vaikas* yra visiškai teisingi ir gramatiniu, ir semantiniu (reikšmės) požiūriu. Tie patys sakiniai perkelti (ne išversti!) į anglų kalbą (*The child eats the apple* ir *The apple eats the child*) yra teisingi tik gramatiškai [Labutis, 2001]. Žiūrint iš sintaksės pozicijų, šie sakiniai parodo, kokią įtaką žodžio sintaksinei funkcijai gali turėti jo vieta sakinyje. Abiejuose lietuviškuose sakiniuose žodis *vaikas* yra veiksnyis, o žodis *obuolį* papildinys. Žodis *child* pirmame angliškame sakinyje yra veiksnyis, o antrame – papildinys. Analogiškai žodis *apple* keičia savo sintaksinę funkciją iš papildinio pirmame sakinyje į veiksnį antrame sakinyje.

Ir kitų šalių literatūroje rašoma, kad sintaksinės analizės pobūdis gali būti labai skirtingas, priklausomai nuo kalbos. Rusų kalboje labai daug kas priklauso nuo daiktavardžio bei būdvardžio linksnių ir veiksmažodžio formų. Šiek tiek informacijos gali turėti ir žodžių tvarka. Anglų kalboje beveik viską nulemia žodžių tvarka, o daiktavardžio ir veiksmažodžio formų variantai sintaksinei analizei mažai tegali padėti [Henish-Dostert, Macdonald, Zarechnak, psl. 116]. Galūnių pagalba sakinio dalys nustatomos anglų kalboje tikrai labai

retais atvejis, pvz., *I know Danny and Toni knows me* ir *I know Danny and Toni know me* [Winograd, psl. 136].

Atskirai galima būtų išvardinti tokius lietuvių kalbos skirtumus nuo kitų kalbų:

1. Lietuvių kalboje veiksny, išreikštas 1-o ar 2-o asmens įvardžiu, paprastai praleidžiamas. Mes jį matome iš veiksmažodžio galūnės. *Vakar sutikau savo seną draugą* yra visai normalus lietuviškas sakiny. Vokiškas sakiny *\*Gestern habe meinem alten Freund begegnet* yra gramatiškai netaisyklingas.
2. Tarinio jungtis (esamojo laiko) lietuviškame sakinyje taip pat dažnai praleidžiama. *Jis geras mokytojas* – labai įprastas lietuviškas sakiny – anglų ir vokiečių kalbose yra negalimas (*\*Er ein guter Lehrer. \*He a good teacher.*).
3. Lietuvių kalba vienintelė iš indoeuropiečių kalbų turi išiktuką kaip atskirą, savarankišką kalbos dalį. Kitų šalių sintaksinės analizės sistemose iš viso nenumatytas tokios kalbos dalies apdorėjimas, o išiktukas gali eiti ir tariniu (*Tik šaka triokšt ir aš nukritau ant žemės.*), ir papildiniu (*Girdėjome kažkokį skambėjimą, kažkokį din din.*), ir aplinkybe (Cop cop šito darbo nepadarysi.). Reiškia, tiesiogiai perėmus kitų indoeuropiečių kalbų sintaksinės analizės sistemas, labai svarbios sakinio dalys liks neatpažintos. Yra pasaulyje kalbų, kuriose išiktukas laikomas savarankiška kalbos dalimi (pvz., tiurkų kalba), bet tai jau ne indoeuropiečių kalbos, ir bendra jų struktūra iš principo skiriasi nuo lietuvių kalbos, taigi, tų kalbų sintaksinės analizės tiesiogiai perimti tuo labiau negalima.
4. Lietuvių kalboje keli neiginiai, esantys viename sakinyje, dažniausiai nepadaro to sakinio teigiamu (išskyrus retus atvejus, kai neiginis turi du veiksmažodžiai tame pačiame sakinyje: *Aš negaliu neateiti*). Tuo tarpu sakiny *Ji nesuprato nė žodžio* reiškia, kad ji nesuprato to, kas buvo pasakyta. Vokiečių kalboje sakiny, turintis du neiginis *Sie hat kein Wort nicht verstanden* tampa teigiamu. Šis sakiny reiškia *Sie hat alles verstanden* – *Ji suprato viską*, t.y. nebuvo nė vieno žodžio, kurio ji nesuprastų.
5. Lietuvių kalboje prijungiamasis žodis gali būti šalutinio sakinio viduje. *Bėras žirgas kur stovėjo, rūteles ten sėsiu. Tėtušis kada mirė, aš dar maža buvau. Aš paežere kaiėjau, pakilo ančių pulkas.* Tiek anglų, tiek vokiečių kalbose tokia žodžių tvarka yra neleistina, todėl minėtų kalbų sintaksinės analizės sistemos tokiais atvejais negalėtų tiksliai nustatyti šalutinio sakinio ribų.
6. Lietuvių kalba neturi tokių specifinių konstrukcijų kaip “*um...zu*” grupė vokiečių kalboje ar “*There is (are) ...*” anglų kalboje.
7. Tarinys lietuvių kalboje gali būti pradžioje sakinio, visiškai nepaversdamas jo klausiamuoju ar liepiamuoju, kaip tai būna anglų ir vokiečių kalbose. Todėl perėmus šių kalbų sintaksinės analizės sistemas, daugelyje vietų būtų daromos klaidos. Visi sakiniai, kurių pirmas žodis yra veiksmažodis, būtų traktuojami kaip klausiamieji ar skatinamieji, tik jokiū būdu ne kaip tiesioginiai. Taigi, labai daug lietuviškų sakinių, prasidedančių veiksmažodžiu (*Einu namo*), būtų išnagrinėti neteisingai (*Go home* – jau yra liepiamasis sakiny).

8. Lietuvių kalba neturi griežtos, sugramatintos žodžių tvarkos [Labutis, 2002, psl. 15], kaip pvz., anglų ar vokiečių kalbos, kur tarinys tiesioginiame sakinyje būtinai turi būti antroje vietoje, klausiamajame – būtinai tik pradžioje sakinio, o vokiečių kalbos šalutiniuose sakiniuose – būtinai tiktai gale. Nė vienos sakinio dalies vieta lietuvių kalboje nėra fiksuota sakinio pradžios ar jo galo atžvilgiu. Lietuvių kalboje iš principo bet kuri sakinio dalis gali būti ir pradžioje sakinio, ir jo viduryje, ir gale.
9. Lietuvių kalboje tarp tiesioginiu sintaksiniu ryšiu susietų žodžių gali būti įsiterpę kitos sakinio dalys, kas yra neįmanoma anglų ar vokiečių kalbose (pvz., *Liūdną jis mums parnešė žinią: Mykolas sėdi kalėjime*)
10. Homonimų daugiareikšmiškumui panaikinti vokiečių ar anglų kalbose dažniausiai pakanka betarpiško konteksto. Pvz., vokiškas žodis *detaillierte* gali būti moteriškos giminės būdvardis (vienaskaitos vardininkas ir galininkas) ir veiksmožodžio trečio asmens vienaskaita. Betarpiškas kontekstas *der Vorgesetzte detaillierte die Aufgabe* arba *die detaillierte Ausarbeitung* vienareikšmiškai nusako pavartoto žodžio formą [Окулич, Петровская, psl. 187]. Lietuvių kalboje šituo pasinaudoti retai kada tegalima. (*Tada mes aukštyn lipome sunkiai. Jis jau greitai mes į viršų kamuolį.* – betarpiškas kontekstas *tada mes aukštyn* ir *greitai mes į viršų* žodžio *mes* daugiareikšmiškumo nepanaikina.

Žodžių tvarka lietuviškame sakinyje gali būti labai įvairi. Galimi visi žodžių išsidėstymo atvejai [Ambrazas, psl. 650].

Neutralia tvarka laikomi du variantai:

- a) Ta žinia labai sujaudino motiną. (SVO – subjektas, veiksmas, objektas)
- b) Ta žinia motiną labai sujaudino. (SOV)

Inversinės žodžių tvarkos galimi dar keturi atvejai.

- c) (VSO) Labai sujaudino ta žinia motiną.
- d) (VOS) Labai sujaudino motiną ta žinia.
- e) (OSV) Motiną ta žinia labai sujaudino.
- f) (OVS) Motiną labai sujaudino ta žinia.

Rusų kalba turi daugiau bendrų bruožų su lietuvių kalba (laisva žodžių tvarka sakinyje, daug linksnių), nei anglų ar vokiečių kalbos. Tačiau skirtumų čia taip pat yra pakankamai daug. Tai – trumpoji būdvardžių forma (*короткая* ir *коротка*), kurios visiškai neturi lietuvių kalba. Ši forma sakinyje eina tariniu, taigi atlieka tą pačią sintaksinę funkciją, kaip ir veiksmožodis. Todėl formalizuojant rusų kalbos sintaksę trumpoji būdvardžio forma įtraukiama į veiksmožodžio formų sąrašą. Apibrėžiant sakinio dalis, skirtumai tarp rusų ir lietuvių kalbų taip pat yra labai ryškūs. Rusų kalbos sakiniuose veiksnys gali būti išreiškiamas prielinksnine konstrukcija (*У него был...*) ar naudininko linksniu (*Ей нельзя...*) [Билан, psl. 182]. Lietuvių kalboje tokio veiksnio iš viso negali būti, o į anglų ar vokiečių kalbas jis irgi verčiamas vardininku (*У меня есть кошка. I have a cat. Ich habe eine Katze. Aš turiu katę.*).

*Taigi, net ir rusų kalbos sintaksinės analizės sistemų tiesiogiai perimti lietuvių kalbai negalima.*

Sudarant lietuviško sakinio sintaksinę struktūrą turi būti įvertintos specifinės lietuvių kalbos ypatybės. Taip elgiasi ir kitų šalių kalbininkai. Pvz., vokiečiai Engel, Eroms, Heringer savo darbuose įvertina tik vokiečių kalbos gramatikai būdingas savybes („...ziehen sie in ihren Arbeiten...nur grammatische Eigenschaften des Deutschen heran“) [Agel, psl. 92].

## 4 GRAFO PANAUDOJIMAS SAKINIO STRUKTŪRAI

Atliekant sintaksinę analizę, pavaizduoti sakinio struktūrą medžiu labai patogu vokiečių bei anglų kalbose. Pagrindine sakinio dalimi laikomą tarinį nesunku surasti kompiuteriu dėl jo pastovios vietos sakinyje. Nors vokiečių kalbos žodžių tvarka šiek tiek laisvesnė, nei anglų kalbos (vokiškame sakinyje veiksnyis gali pasikeisti vietomis su papildiniu), tačiau griežtas tarinio antroje vietoje reikalavimas išlieka. Todėl ir sintaksinė analizė pradedama remiantis šia kalbos ypatybe. Pagrindiniu sakinio žodžiu laikomas tarinys ir nuo jo pradedama formuoti viso sakinio struktūra. Lietuvių kalbai perimti tokią metodiką tiesiogiai nėra pati geriausia išeitis, nes lietuvių kalbos sakiniams nebūdinga pastovi tarinio vieta. Be to, medis negali parodyti visų sintaksinių ryšių tarp žodžių lietuviškame sakinyje.

### 4.1 Tarinio pažyminio problema

Vienas iš specifinių lietuvių kalbos bruožų yra tarinio pažyminio traktavimas kaip savarankiškos sakinio dalies. Tarininis pažyminis rodo veiksmu ar papildiniu einančio daikto ypatybę, kurią jis turi (ar kuri yra aktuali) tik to veiksmo metu, pvz., pasakius *Tėvas grįžo namo piktas* požymis *piktas* subjektui *tėvas* galioja tik grįžimo momentu ir nereiškia, kad jis visą laiką būna piktas. Tuo tarininis pažyminis skiriasi nuo atributinio pažyminio, nes, pvz., sakinyje *Piktas vilkas atsėlino prie trobelės* žodis *piktas* rodo jau pastovią vilko ypatybę, t.y., kad vilkas visada būna piktas. Papildinio atveju, pvz., *pamačiau motulę pareinančią* – papildinys *motulę* turi savybę *pareinanti* tik pamatymo metu.

Ne visos kalbos turi tokią sakinio dalį kaip tarininis pažyminis. Pvz., anglų kalbos ypatybės (mažas kaitomumas, griežta žodžių tvarka sakinyje) nulemia tai, kad tarinio pažyminio šioje kalboje būti negali ir anglų kalbos gramatika iš viso neaprašo tokios sakinio dalies [Greenbaum].

Labiau kaitomose kalbose išskiriamas tarininis pažyminis. Pvz., vokiečių kalbos gramatika jį nagrinėja ir aptaria jo ryšius tiek su tariniu, tiek su veiksmu ar papildiniu [Helbig, Buscha, psl. 339]. Tačiau, formalizuojant sintaksę, terodomas ryšys tikrai su tariniu. U. Engel, kuris kalbininko Tesnière pasiūlytą priklausomybių gramatiką pritaikė vokiečių kalbai, savo knygoje “Syntax der deutschen Gegenwartssprache” [Engel] tarinio pažyminio atskirai neaprašo ir priklausomybių medyje dvigubų ryšių nerodo. Antro ryšio ignoravimą galima būtų paaiškinti tuo, kad vokiečių kalboje šis klausimas nėra labai aktualus. Nors atributinis pažyminis ir derinamas su pažymimuoju žodžiu (*schöner Tag*, *schöne Katze*, *schönes Mädchen*), tačiau tarininis pažyminis formalaus derinimo su veiksmu ar papildiniu neturi. Vokiečių kalbos gramatikoje sakoma, kad tarininis pažyminis formaliai sutampa su aplinkybe: „...das prädikative Attribut mit der Adverbialbestimmung ... in der äusseren Form und Position im konkreten Satz übereinstimmt.” [Helbig, Buscha, psl. 556]. Taigi, tarinio pažyminio atveju skirtumas tarp jo ryšio su veiksmu ir jo ryšio su papildiniu lieka tik prasminis. Todėl nėra ir būtinybės priklausomybių medyje žymėti tą ryšį. Vokiečių mokslininkas P. Hellwig abiejų tarinio pažyminio ryšių parodymą sakinyje vadina intuityviu vaizdavimu („...eine intuitive Illustration”) [Hellwig, 2004].

*Lietuvių kalboje terodyti kūrį nors vieną iš šių ryšių, neprarandant sintaksinės informacijos, negalima, nes abu jie reiškiami formaliai. Todėl ne visų lietuvių kalbos sakinių sintaksinę struktūrą galima pavaizduoti medžiu.*

## 4.2 Būtinumas panaudoti grafą

*Lietuvių kalboje negalima pasinaudoti nė vienu iš dviejų paplitusių sakinio sintaksinės struktūros pavaizdavimo būdų. Nei frazių principu sudaromas medis, nei priklausomybių gramatikos siūlomas medis neatspindi visos sintaksinės informacijos, esančios lietuviškame sakinyje, t.y. jis negali parodyti visų ryšių tarp lietuviško sakinio žodžių. Lietuvių kalbai negalioja priklausomybių gramatikos metodą pasiūliusio kalbininko Tesnière teiginys, kad pats žodis tegali priklausyti tik nuo vieno žodžio, o priklausomų žodžių gali turėti kelis [Glück, psl. 128]. Lietuvių kalboje tarininis pažyminy formaliai priklauso ir nuo tarinio, ir nuo veiksnio ar papildinio.*

Jeigu sakinio struktūroje parodomi abu šie ryšiai, tai gaunamas grafas, kuris nebeatitinka medžio apibrėžimo (medis – susietas grafas be ciklų [Swamy, Thulasiraman, psl. 33]), nes atsiranda uždaras kelias. O jei bent vienas iš šių ryšių sakinio struktūroje nebus pavaizduotas, vertimo metu bus daromos klaidos.

Šiuos teiginius galima pailiustruoti tokiu pavyzdžiu. Vokiečių kalboje, kuri yra mažiau kaitoma, nei lietuvių kalba (t.y. turi mažiau žodžių formų su skirtingomis galūnėmis), tarininis pažyminy formaliai sutampa su aplinkybe ir čia nėra galūnių derinimo su ta sakinio dalimi, kurią jis pažymi. Pvz., sakiniuose:

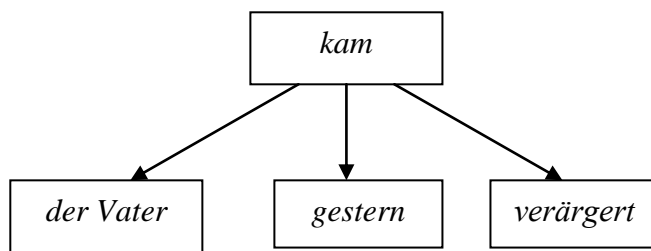
*Der Vater kam gestern verärgert.  
Die Mutter kam gestern verärgert.  
Die Brüder kamen gestern verärgert.  
Die Schwestern kamen gestern verärgert.*

tarininio pažyminio forma lieka ta pati (*verärgert*), nepriklausomai nuo veiksnio giminės ir skaičiaus, t.y. tarininis pažyminy su veiksnium formaliai nederinamas. To negalima pasakyti apie lietuvių kalbą. Lietuviškame šių sakinių vertime žodis *verärgert* turės keturis atitikmenis, kurių galūnes nulemia veiksnys. Tai rodo, kad lietuvių kalboje tarininis pažyminy formaliai (galūnėmis) derinamas su veiksnium:

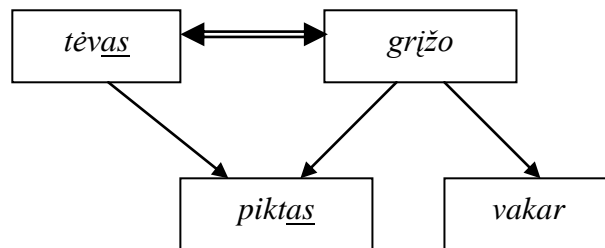
*Tėvas vakar grįžo piktas.  
Motina vakar grįžo pikta.  
Broliai vakar grįžo pikti.  
Seserys vakar grįžo piktos.*

Todėl vokiško sakinio sintaksinė struktūra, pateikta 26 pav. (psl. 38), verčiant šį sakinį į lietuvių kalbą turi būti keičiama lietuviško sakinio sintaksine struktūra, parodyta 27 pav. (psl. 38).

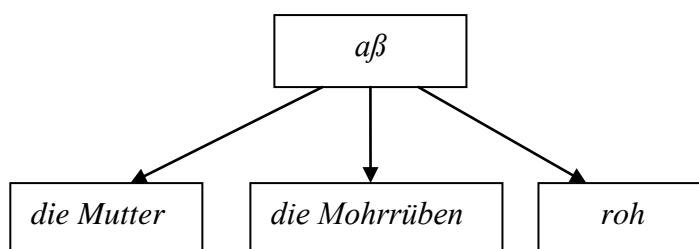
Informacijos trūkumas ypač gerai matomas tų sakinių medžio struktūroje, kurie turi tarininį pažyminį ir papildinį, nes tarininis pažyminy gali priklausyti tiek nuo veiksnio, tiek nuo papildinio. Taigi, pasidaro visai nebeaišku, su kuriuo sakinio žodžiu turi būti derinama tarininio pažyminio galūnė. Pvz., sakinį *Die Mutter aß die Mohrrüben roh* pavaizdavus kaip parodyta 28 pav. (psl. 38), neaišku, kas gi ten buvo žalias: morkos ar motina. O nuo to priklauso žodžio *žalias* galūnė lietuviškame sakinyje: ar *\*Motina morkas valgė žalia* (taip pat kaip ir *Motina vakar grįžo pikta*), ar *Motina morkas valgė žalias*. Norint sugeneruoti išverstą į lietuvių kalbą sakinį be klaidų, ši struktūra vertimo metu turi būti keičiama į lietuviško sakinio struktūrą, parodytą 29 pav. (psl. 38).



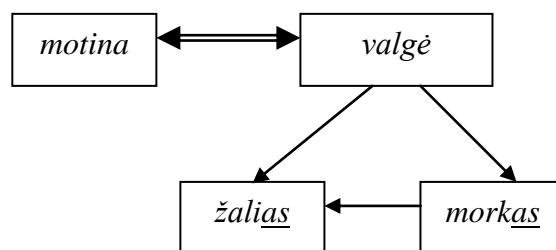
26 pav. Vokiško sakinio *Der Vater kam gestern verärgert* sintaksinė struktūra.



27 pav. Lietuviško sakinio *Tėvas vakar grįžo piktas* sintaksinė struktūra.



28 pav. Vokiško sakinio *Die Mutter aß die Mohrrüben roh* sintaksinė struktūra.



29 pav. Lietuviško sakinio *Motina morkas valgė žalias* sintaksinė struktūra.

*Iš šių dviejų pavyzdžių gerai matyti, kad vieną ir tą patį grafinį vaizdą turinčias vokėčių sakinių struktūras, pateiktas 26 pav. ir 28 pav., psl. 38 (grafas susideda iš keturių viršūnių ir trijų lankų, kurie išeina iš vienos viršūnės ir eina į likusias tris, neturinčias tarpusavyje ryšio) reikia keisti dviem skirtingomis lietuviškų sakinių struktūromis. Vadinas, vokėčių sakinių grafiniame vaizde esančios informacijos neužtenka, jei norima be klaidų sugeneruoti išverstą į lietuvių kalbą sakinį.*

*Jeigu bandytume žūtbut išlaikyti medžio grafinį vaizdą (nes kompiuteriu lengviau apdoroti medį, negu grafą), ir terodyti tik tą ryšį, kur vyksta galūnių derinimas, t.y. tik ryšį tarp tarininio pažyminio ir veiksnio arba papildinio, neparodant jo priklausomybės nuo tarinio, vertime atsirastų kito tipo klaidos – žodžio vietos sakinyje neteisingas parinkimas.*

Tarininis pažyminytis būtinai turi būti atskirtas nuo veiksnio kitais žodžiais, ar eiti po jo. Tai vienas iš labai nedaugelio atvejų lietuvių kalboje, kada žodžių tvarka turi įtakos nustatant sakinio dalis. Čia nepažeidžiamas laisvos žodžių tvarkos principas lietuvių kalbos sakiniuose. Tarininis pažyminytis gali būti ir pradžioje, ir viduryje, ir gale sakinio (*Piktas grįžo tėvas į namus. Tėvas grįžo piktas į namus. Tėvas į namus grįžo piktas.*). Tačiau tarininis pažyminytis niekada negali būti prieš veiksnį, nes tokiu atveju jis tampa atributiniu pažyminiu. *Piktas tėvas grįžo į namus* – tai jau pastovi veiksnio ypatybė, o tuo (ypatybės pastovumu) ir pasireiškia esminis skirtumas tarp atributinio ir tarininio pažyminių. Su papildiniu susijęs tarininis pažyminytis būtinai turi būti po papildinio (Parėjusi radau motulę *susikrimtusią*), kitaip jis tampa atributiniu pažyminiu (*Susikrimtusią motulę* ištiesai lydėjo nesėkmės).

*Terodant sintaksinėje struktūroje tik tą ryšį, kur vyksta galūnių derinimas, t.y. tarininį pažyminį traktuojant kaip atributinį, būtų daromos klaidos ir atliekant vertimus kita kryptimi – iš lietuvių kalbos į vokėčių kalbą, nes atributinis pažyminytis vokėčių kalboje derinamas galūnėmis su išplečiamu žodžiu taip pat kaip ir lietuvių kalboje, o tarininis pažyminytis turi aplinkybės formą ir nederinamas su tuo žodžiu (veiksniu ar papildiniu), kurį jis išplečia. Taigi, neparodžius lietuviškame sakinyje tarininio pažyminio ryšio su tariniu, vokėčių sakinių sintezės metu bus neteisingai nustatoma tarininio pažyminio galūnė.*

*Automatinio vertimo sistemose išverstas sakinytis generuojamas pasinaudojant sintaksinės analizės rezultatais [Hutchins, psl. 53]. Neparodžius lietuviško sakinių sintaksinėje struktūroje tarininio pažyminio ryšių su veiksniu ar papildiniu, morfologinės sintezės metu trūktų informacijos apie tai, kokią turi būti jo galūnė, o neparodžius ryšio su tariniu, sintaksinės sintezės metu nebūtų galima teisingai nustatyti jo vietos sakinyje. Todėl lietuvių kalbos sakiniams struktūriškai pavaizduoti yra būtina naudoti grafą, nes tik jis gali tiksliai atspindėti lietuviško sakinių savybes.*

### 4.3 Apibendrinta lietuvių kalbos sakinių sintaksinė struktūra

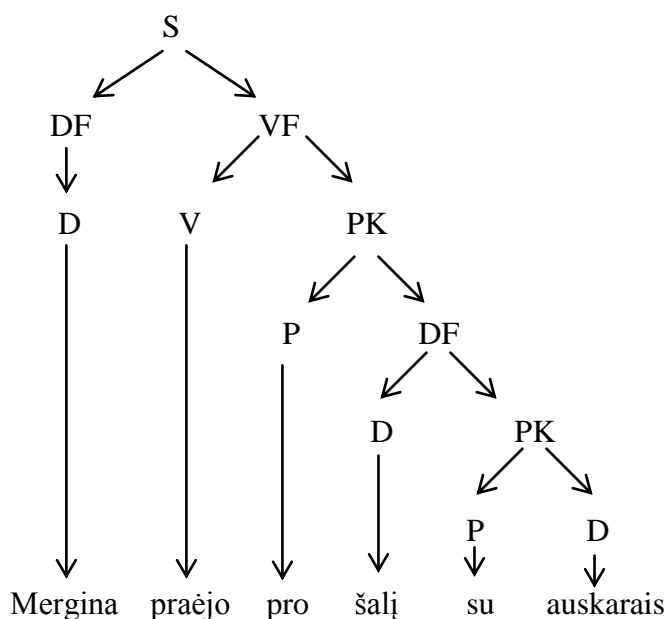
Atliekant sintaksinę analizę automatiškai (kompiuteriu), patogų turėti apibendrintą struktūrą, kuri apimtų visus galimus lietuviško sakinių atvejus, t.y. būtų bendra visiems lietuvių kalbos vientisiniams sakiniams: konkretus sakinytis turėtų aktyvuoti toje struktūroje kurį nors vieną kelią. Kiekvienam bet kokio sakinių žodžiui schemoje turi būti numatyta vieta, į kurią jis bus talpinamas analizės metu. Tokios schemas grafinis vaizdas pateiktas 33 pav. (psl. 44) Grafo viršūnėms suteikti sakinių dalių pavadinimai rodo, kur turi būti dedamas žodis, pagal nustatytą jo sintaksinę funkciją bei ryšius su kitais žodžiais. Taigi, apibendrinta struktūra tai tarsi kontūrai. Surašius visus sakinių žodžius į jiems tinkamas viršūnes, nepanaudota grafo dalis atmetama, o gautas pografis pateikiamas kaip nagrinėjamo sakinių sintaksinė struktūra.

Formuojant tokią schemą iš dalies galima pasiremti kitų šalių patirtimi, tačiau tiesiogiai perimti kitoms kalboms taikomų metodikų – frazių metodo ar priklausomybių gramatikos – nevertėtų, nes abi jos turi ir plusų, ir minusų. Pasinaudoti reiktų tikrai tuo, kas tinka lietuvių kalbai, stengiantis išvengti minėtų metodų trūkumų.

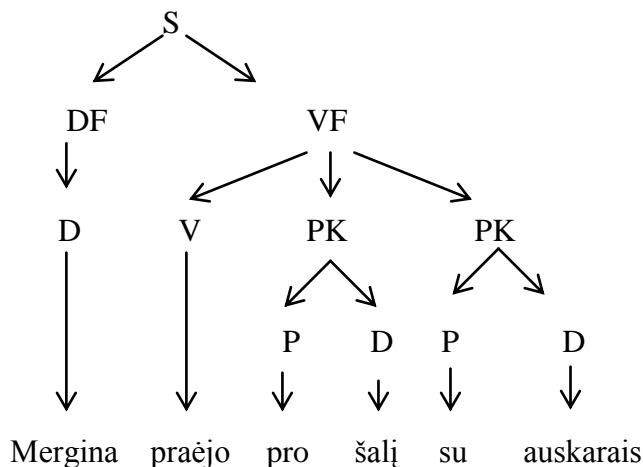
### 4.3.1 Frazių metodo trūkumai

Pagrindinis frazių metodo trūkumas, kurį nurodo jau net patys anglų kalbininkai, yra tai, kad analizė atliekama remiantis žodžių tvarka, ir tokiaime sakinyje kaip *A man walked by wearing earrings* nėra galimybės susieti žodį *man* su žodžiais *wearing earrings* [Winograd, psl. 135]. *Lietuvių kalboje, turinčioje laisvą žodžių tvarką, panašūs sakiniai gali pasitaikyti dažnai.*

Sakinio *Mergina praėjo pro šalį su auskarais*, kuris lietuvių kalboje yra visai geras ir suprantamas, sintaksinę struktūrą frazių metodu tegalima pavaizduoti dviem būdais (30 pav. ir 31 pav.):



**30 pav.** Sakinio *Mergina praėjo pro šalį su auskarais* struktūra, pavaizduota frazių metodu, traktuojant, kad mergina praėjo pro šalį, kuri buvo su auskarais, t.y. šalis buvo su auskarais.



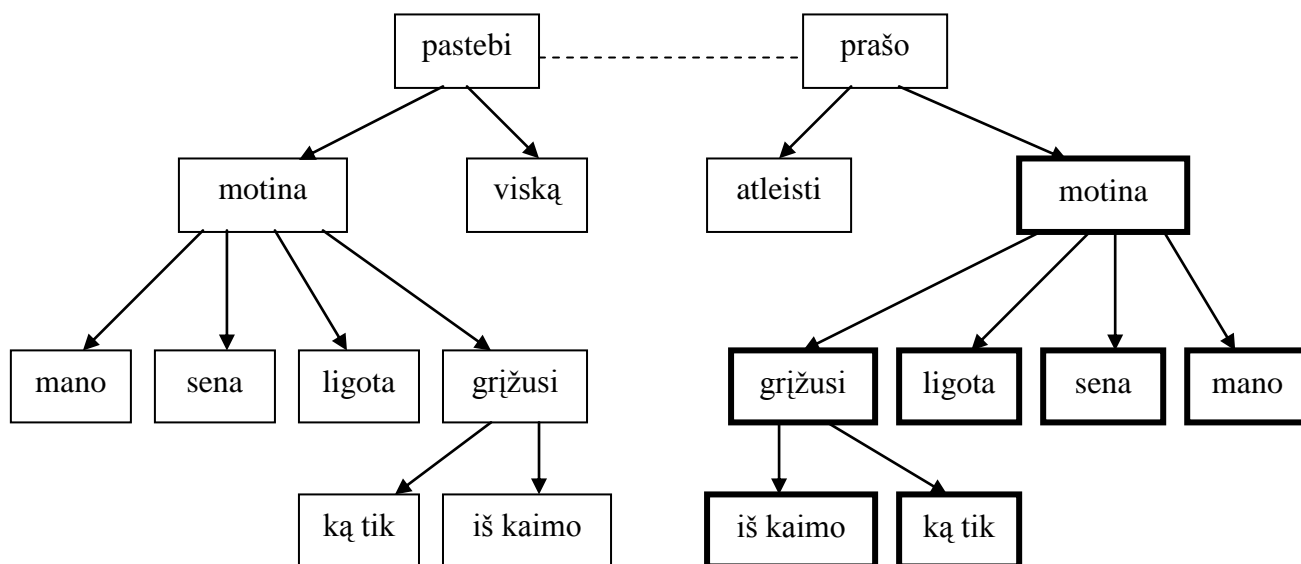
**31 pav.** Sakinio *Mergina praėjo pro šalį su auskarais* struktūra, pavaizduota frazių metodu, traktuojant, kad mergina praėjo su auskarais, t.y. ji pro šalį ėjo su auskarais, taip kaip, pvz., važiavo su mašina.



Kaip matyti iš brėžinių, nė viena iš pateiktų struktūrų neatitinka tikrosios sakinio reikšmės. Kadangi lietuvių kalboje toks žodžių išsidėstymas yra leistinas (tiesiogiai priklausomi žodžiai gali būti perskirti kitomis sakinio dalimis, pvz., pažyminys nuo savo pažymimojo žodžio gali būti atskirtas veiksmu ar tariniu kaip kad sakiniuose: *Geras gi tu, Simanai, akis turi, kad galėjai ją [žvaigždę] pamatyti* [Ulvydas, psl. 405]. *Ir ėmė abi rodyti susirinkusiems, kuri geresnius turi nasrus* [Ulvydas, psl. 228]), *todėl frazių metodas ne visada tinka lietuvių kalbos sakiniams pavaizduoti*.

### 4.3.2 Priklausomybių gramatikos trūkumai

Tradicinėje gramatikoje sakinio branduoliu laikomas veiksnys su tariniu ir sakoma, kad tik jie abu gali sudaryti sakinį, t.y. atlikti pagrindinę sakinio funkciją – komunikatyvinę (perteikti informaciją). Priklausomybių gramatikoje sakinio centras yra vien tarinys, o veiksnys traktuojamas kaip ir kitos antrininkės sakinio dalys, t.y. jis pavaldus tariniui. Tokio požiūrio susiformavimą greičiausiai lėmė tai, kad pagrindinėse Europos kalbose (vokiečių, anglų, prancūzų) tarinys turi fiksuotą vietą sakinyje – antrą nuo pradžios; vadinasi, jį labai nesunku rasti kompiuteriu ir nuo jo pradėti viso sakinio analizę. Tačiau čia iškyla problemos, kai bandoma pavaizduoti sakinius su keliais vienarūšiais tariniais. Kada veiksnys pavaldus tariniui, nebelyka galimybės parodyti medyje tiesioginių ryšių tarp veiksnio ir kelių tarinių. Tokio tipo problemas (kai žodis priklauso nuo kelių aukštesnio rango žodžių), siūloma spręsti įtraukiant tą patį žodį į schemą kelis kartus, kad kiekvienai viršūnei būtų parodyti visi jos ryšiai [Schubert, psl. 108]. Šį sprendimą galima pailiustruoti lietuviško sakinio *Mano sena, ligota, ką tik iš kaimo grįžusi motina viską pastebi ir prašo atleisti* pavyzdžiu. Pakartojant veiksnio šaką, t.y. veiksnį su visais jį išplečiančiais žodžiais, prie kiekvieno tarinio, į schemą turi būti įtraukiami papildomi žodžiai, kurių sakinyje iš viso nėra, t.y. gaunama schema, turinti daugiau žodžių, nei jų yra pačiame sakinyje (32 pav.).



32 pav. Sakinio su vienarūšiais tariniais schema.

Schemoje atsiranda septynios papildomos viršūnės (jos paryškintos), vaizduojančios sakinyje nesančius žodžius. Ir visas šis balastas vertimo metu turi būti apdorojamas kartu su naudinga informacija, o po to, prieš pat pateikiant išverstą sakinį, vėl atmetamas.

Visų šių problemų galima išvengti, jei sakinio struktūra vaizduojama remiantis tradicine gramatika, t.y. sakinio branduoliu laikant veiksnių ir tarinių. Tai, kad priklausomybių gramatikoje medžio viršūnių kartais reikia daugiau, nei sakinytis turi žodžių, byloja, kad ši struktūra neypatingai jau gerai atitinka sakinio esmę bei prigimtį. Tradicinės gramatikos principu pavaiždavus priklausomybių medį, jokių papildomų daiktų į schemą įrašyti nereikia, nes ji pati labai tiksliai atspindi sakinį. Šios gramatikos teisingumą patvirtina ir tas faktas, kad visais atvejais, kai sakinį sudaro vienas žodis (vien veiksmažodis, ar vien daiktavardis), labai nesunkiai galima atstatyti trūkstamą branduolio narį. Vien iš tarinio sudarytas sakinytis būna visiškai suprantamas, t.y. atlieka komunikatyvinę funkciją (perteikia informaciją) tiksliai tada, kai veiksnys labai aiškus iš veiksmažodžio reikšmės, pvz., *sninga, lyja*. Snigti gali tik sniegas, lyti – tik lietus. Ir tik ši reikšmė perteikiama sakiniu, sudarytu vien iš tarinio. Jei reikėtų pasakyti, kad *sninga* ar *lyja* saldainiai (*Apsižergus greitą šluotą / Raganėlė lėks į puotą / O kad niekas neliūdėtų / Ten išburs saldainių lietu.* [Jurgaitis, psl. 1]), veiksnys jau būtinai turėtų būti sakinyje, nes jį praleidus, ši informacija nebus perteikta.

Kitas atvejis, kada vienas bet koks žodis gali sudaryti sakinį, yra kontekstas (aplinkinių sakinių teikiama informacija). Jei iš konteksto aiški praleistų žodžių reikšmė, bet kokia kalbos dalis (ne tik veiksmažodis ar daiktavardis) gali sudaryti sakinį (*Palangoj gražu. Ypač vasarą. | Padangėje klykteli gervės. Taip verkmingai ir graudžiai. | Šeima nedidelė. Tik keturi.*) [Labutis. 2002, psl. 140]. Tačiau be konteksto šie (pabraukti) sakiniai darosi nesuprantami, vadinasi, ir informacijos jie neperteikia. Geriausiai galima patikrinti ar sakinytis atlieka komunikatyvinę funkciją ar ne – tai išimti jį iš konteksto. Tam labai gerai tinka pasakojimo ar pastraipos pradžia, nes ten niekada nebūna jokio konteksto. Pirmas sakinytis visada turi būti aiškus be aplinkos, todėl juo ir galima patikrinti, ar žodis sudaro sakinį ar ne. Yra keli veiksmažodžiai, kuriais pradėjus pasakojimą, jis lieka visiškai suprantamas, pvz., *Aušta. Senas alkanas šuo sunkiai keliasi iš savo guolio išgirdęs nepažįstamus balsus ir trumpai amteli.* Tačiau sunkiai galim įsivaizduoti pasakojimą, kuris prasidėtų žodžiu *sustojo*. Iš vieno šio žodžio neaišku nei kas, nei kur, nei kodėl *sustojo*. Toks sakinytis neatrodo užbaigtas ir geras. Čia aiškiai matomas informacijos trūkumas, vadinasi, komunikatyvinės funkcijos toks žodis neatlieka, ir todėl sakiniu negali būti laikomas. Taigi, neįtikinamai atrodo teiginys, kad bet koks veiksmažodis vienas gali suformuoti sakinį t.y. visada perteikti informaciją. Net ir realizavęs savo būtinąjį valentingumą, su papildiniu, vienas veiksmažodis sudaro žodžių junginį, o ne sakinį, pvz., *rašo laišką* [Ulvydas, psl. 9], o kartu su veiksmiu, tarinytis suformuoja sakinio branduolį net ir tais atvejais, kai jo stiprusis valdymas lieka nerealizuotas: žodis *rašyti* stipriai valdo galininką, tačiau sakinytis *Tėvas sėdi ir rašo* visiškai atlieka komunikatyvinę funkciją – jis yra aiškus be konteksto, juo gali prasidėti pasakojimas.

Panašu, kad priklausomybių gramatikoje perdėtai iškeliamas tarinio reikšmė formuojant sakinį. Negalima sutikti su teiginiu, kad veiksmažodis nulemia visų sakinio žodžių formą bei sintaksinę funkciją, ir nurodo, kokios sakinio dalys gali būti tame sakinyje („...bildet das Verb das Zentrum, das alle weiteren Satzglieder bezüglich ihrer Form und Funktion bestimmt. ... Das Verb bestimmt, welche Satzglieder im Satz erscheinen dürfen.“ [Müller, psl. 134]. Tarinytis nulemia tik savo tiesiogiai, betarpiškai valdomų žodžių formą bei jų buvimą sakinyje. Tačiau jei veiksnio pažyminiu einantis būdvardis reikalauja kokio nors linksnio, tai šiam valdymui tarinytis jokios įtakos neturi, pvz., sakinyje *Kenksmingos sveikatai dujos pasklido po visą gamyklą* veiksmažodis *pasklido* niekaip nevaldo daiktavardžio naudininko ir nenulemia jo buvimo. Jei veiksnio *dujos* pažyminys bus kitas, šio linksnio sakinyje gali iš viso nelikti, nors veiksmažodis ir būtų tas pats – *Panašios į baltą rūką dujos pasklido po visą gamyklą*; tiek šiame sakinyje žodžių *į rūką*, tiek praeitame sakinyje žodžio *sveikatai* formą (ir buvimą sakinyje) nulemia veiksnio *dujos* pažyminys (*kenksmingos*-kam? ir *panašios*-į ką?), o ne tarinytis. Tarinytis tegali nulemti tik betarpiškai nuo jo priklausomus žodžius, kaip ir kiekvienas valdymą turintis žodis, pvz., daiktavardis: *vaistai* – nuo ko?; skaitvardis: *šimtas* – ko?; *ketvertas* – ko?, būdvardis: *naudingas* – kam?, *pilnas* – ko?, *įdomus* – kuo? ir pan.

Galima pasakyti dar daugiau: veiksnio traktavimas kaip sakinio dalies, pavaldžios tariniui ir lygiavertės su kitomis antrininkėmis sakinio dalimis, neatitinka tikrovės, nes tarinys ne tik nulemia veiksnio formos taip, kaip jis nulemia papildinio formą, bet dargi atvirkščiai – veiksnys apsprendžia tarinio galūnę. Sakinyje *Aš noriu obuolio* pakeitus veiksmažodį, keičiasi papildinio galūnę, t.y. į tarinio vietą atėjusi kita leksema keičia papildinio formą (jo linksnį) – *Aš matau obuolį*; tuo tarpu veiksniai ji jokios įtakos nepadaro (veiksnys lieka tas pats – *aš*). Tačiau pakeitus veiksnį sakinyje, tarinio forma jau bus kita, nors likę sakinio žodžiai ir būtų tie patys: *Mes norime obuolio. Mes matome obuolį. | Jis nori obuolio. Jis mato obuolį*. Vadinasi, veiksnys formaliai valdo tarinį panašiai, kaip tarinys valdo papildinį, taigi, ne veiksmažodis nulemia veiksnio formą, kaip teigia vokiečių autoriai, bet veiksnys – tarinio. Nelabai logiška atrodo, kad pavaldus žodis apsprendžia valdančiojo žodžio galūnę (jo formą).

Jei sakinyje turi du ar daugiau vienaarūšių tarinių, jame gali būti tokie linksniai, kurių pasirodymą draudžia pirmas tarinys, bet reikalauja antras. Taip pat nėra vienas veiksmažodis nerodo, kad turi būti arba negali būti kitas, vienaarūšis su juo pačiu, veiksmažodis. Taigi, antras tarinys yra tokia sakinio dalis, kurios pirmasis tarinys nenusako niekaip. Vadinasi, ne visas sakinio dalis apibrėžia tarinys.

Kaip matyti iš aprašytų faktų, neatitikimų su tikrove priklausomybių gramatikoje yra. *Taigi, ir apie priklausomybių medį negalima pasakyti, kad jis idealiai tikėtų lietuvių kalbos sakiniams pavaizduoti.*

#### 4.3.3 Apibendrintos sakinio struktūros formavimas

Aptarus abiejų metodikų trūkumus, lietuvių kalbos sakinio struktūrai reikia perimti tai, kas jose yra gera. Priklausomybių gramatikos teigiama savybė, kuri yra ypač palanki lietuvių kalbai – tai sugebėjimas pateikti sakinio struktūrą, nesusijusią su žodžių tvarka. Todėl apibendrinta struktūra ir bus formuojama priklausomybių gramatikos principu. Tačiau sakinio branduoliu bus laikomas veiksnys su tariniu, nes toks traktavimas geriau atitinka sakinio esmę ir leidžia išvengti priklausomybių gramatikos trūkumų.

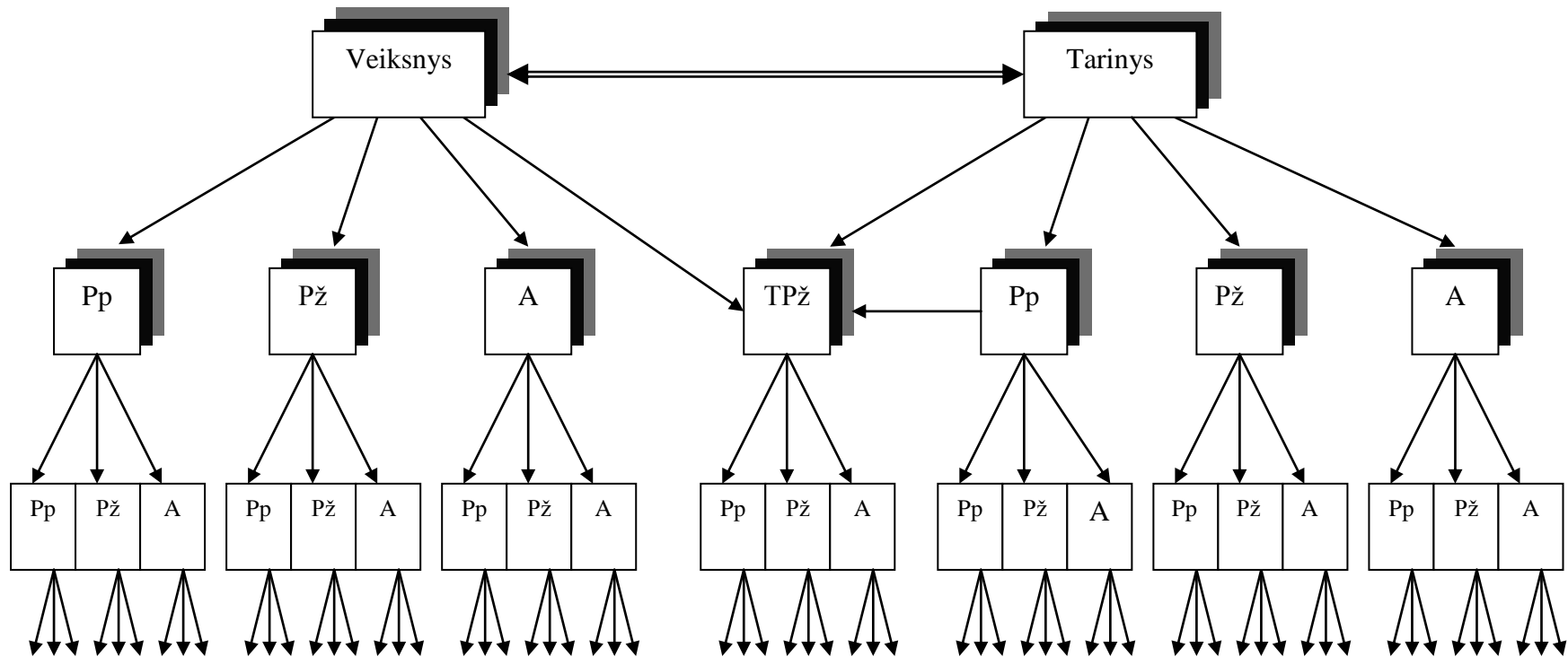
Įvertinant tai, kad medis negali apimti visos sintaksinės informacijos, esančios lietuviškame sakinyje, sintaksinei jo struktūrai pavaizduoti pasirenkamas grafai, suformuoti pagal priklausomybių gramatikos principą.

Apibendrinta lietuviško vientisinio sakinio struktūra pateikiama 33 pav. (psl. 44). Visos penkios sakinio dalys (veiksnys, tarinys, papildinys, pažymins ir aplinkybė) gali būti išplėtos pažyminiiais, papildiniais ir aplinkybėmis. Tačiau nėra viena iš jų negali būti išplėsta veiksniais ar tariniais. Tai ir atsispindi schemeje.

Dvikryptę rodyklę, jungiančią veiksnį su tariniu, galima paaiškinti tuo, kad abi šios sakinio dalys veikia viena kitą: tarinys reikalauja veiksnio kaip savo valentinio palydovo [Sližienė, psl. 5], t.y. veiksnys funkcionuoja kaip veiksmažodžio valentingumo sudėtinė dalis (todėl ir brėžiama rodyklė, nukreipta nuo tarinio į veiksnį), o kryptis nuo veiksnio į tarinį rodo, kad veiksnys nulemia tarinio formą (jo galūnę).

Iš veiksnio išeinančios rodyklės į visas tris antrininkes sakinio dalis, iš pirmo žvilgsnio, gali atrodyti prieštaraujantis lietuvių kalbos gramatikai atvejis, nes joje sakoma, kad papildinys gali jungtis tik prie veiksmažodžio, o pažymins – tik prie daiktavardžio. Tipinė veiksnio forma yra daiktavardžio vardininkas. Tokiu atveju jis išplečiamas pažyminiu (*mažas vaikas*). Kai veiksnys išreikštas bendratimi, jis gali būti išplečiamas papildiniu ar aplinkybe (Versti tekstus automatiškai nėra lengva).

## Lietuvių kalbos vientisinio sakinio apibendrinta struktūrinė schema



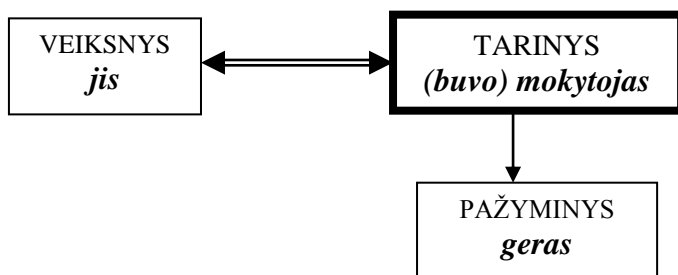
Pž – pažymins, Pp – papildins, A – aplinkybė, TPž – tarininis pažymins.

**33 pav.** Apibendrinta lietuvių kalbos sakinio struktūrinė schema.

Analogiškai galima būtų paaiškinti ir tarinio jungimą su pažyminiu. Tarinys, išreikštas tipine forma – veiksmažodžiu, negali būti išplėstas pažyminiu (kaitomas veiksmažodis nederinamas su vardažodžio linksniu), tačiau, kai tarinį sudaro jungtis su vardine dalimi, išreikšta daiktavardžiu, tai tas daiktavardis gali turėti pažyminį.

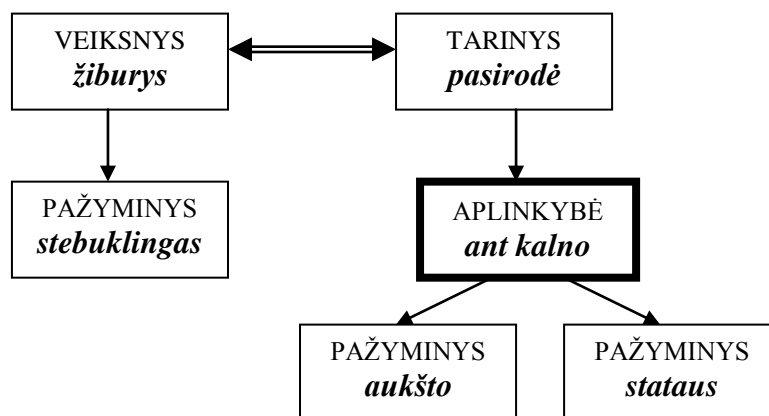
Esamojo laiko jungtis lietuvių kalboje dažniausiai būna nulinė, t.y. ji praleidžiama, taigi, ir visas tarinys labai dažnai tesusideda vien iš daiktavardžio. Pvz., sakiniuose *Jis buvo mokytojas*, ir *Jis buvo geras mokytojas*, taip pat *Jis – geras mokytojas* daiktavardis yra tarinys (34 pav.).

Kadangi jungtis išreiškiama veiksmažodžiu, neturinčiu savarankiškos reikšmės (taigi, ir negalinčiu valdyti kitų žodžių), sakinio struktūroje ji pateikiama kartu su vardine dalimi (vienoje pozicijoje), nes iš jos niekada neišeis atskiros rodyklės į kitus žodžius, nesusijusius su vardažodžiu. Vokiečių kalbininkai sako, kad semantiniu požiūriu jungtis yra tuščia („...*die Kopulaverben semantisch relativ leer sind*“) [Helbig, Buscha, psl. 539]. Matyt todėl lietuvių kalbos esamojo laiko jungtis dažniausiai ir būna praleidžiama.



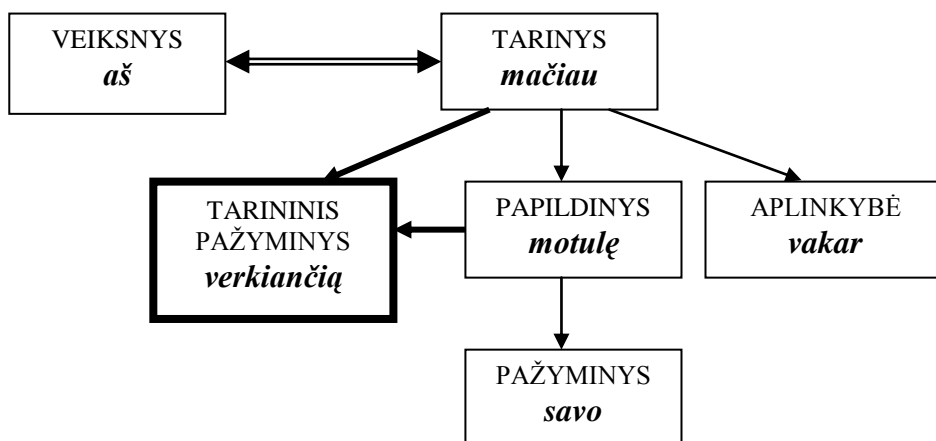
34 pav. Sakinio *Jis (buvo) geras mokytojas* sintaksinė struktūra.

Nesavarankiškas kalbos dalis (dalelytė, prielinksni) taip pat tikslinga patalpinti į vieną poziciją su žodžiu, kurį jos išplečia. Kitų šalių kalbininkai irgi kartais sudeda po du žodžius į vieną viršūnę. Pvz., vokiečių kalbos veiksmažodžio analitinės formos (*habe gesehen*) rašomos kartu ir traktuojamos kaip vienas žodis, tiksliau kaip žodžio ekvivalentas [Agel, psl. 34]. Be to, atskirose kalbose prielinksniai gali skirtis, t.y. su keliais žodžiais naudojamas tas pats prielinksnis vienoje kalboje, gali turėti kelis atitikmenis kitoje, priklausomai nuo to žodžio, su kuriuo jis siejamas, reikšmės (pvz., vokiškai *glauben an* ir *arbeiten an*, o angliškame vertime *work at*, bet *believe in*; vokiškai *blicken auf* ir *weisen auf*, o angliškai *look at*, bet *point to* ir pan. [Hutchins, Sommers, psl. 37]), o kartais prielinksninė konstrukcija gali būti verčiama netgi vien linksniu, be prielinksnio, pvz., žodis *panašus* lietuvių kalboje reikalauja galininko su prielinksniu *į*, o vokiškas jo atitikmuo valdo naudininką be jokio prielinksnio (*panašus į tėvą – ähnlich dem Vater*). Taigi, vertimo metu yra patogiau prielinksni apdoroti kartu su linksniu. Sakinio su prielinksniu pavyzdys (*Ant aukšto stataus kalno pasirodė stebuklingas žiburys*) pateikiamas 35 pav.



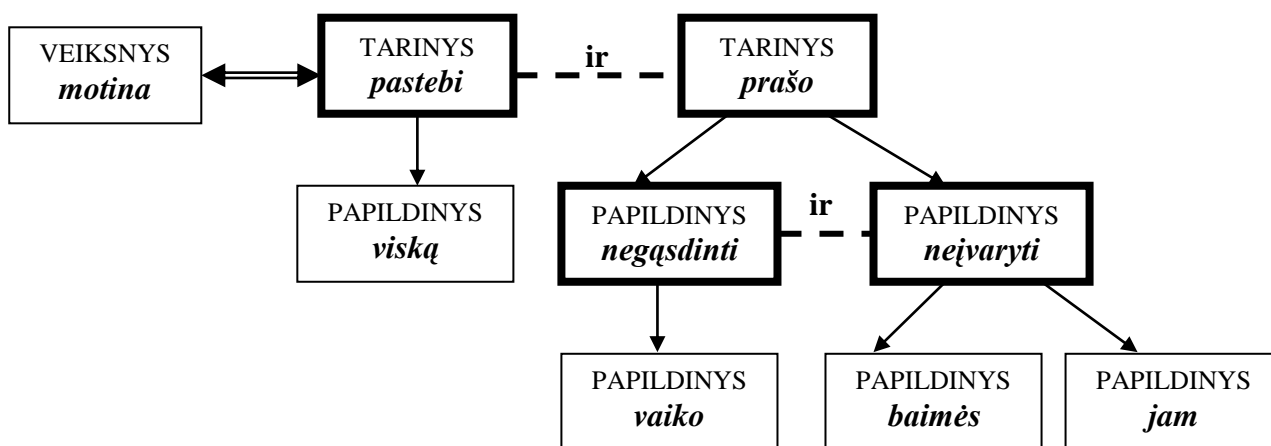
35 pav. Sakinio *Ant aukšto stataus kalno pasirodė stebuklingas žiburys* sintaksinė struktūra.

Tarininis pažymynys turi dvigubus sintaksinius ryšius. Jis formaliai derinamas su veiksmiu ar papildiniu ir kartu šliejamas prie veiksmažodžio. Todėl apibendrintoje schemoje (žr. 33 pav., psl. 44) ir parodytos į jį ateinančios trys rodyklės. Ryšys su tariniu bus kiekviename sakinyje, turinčiame tarininį pažyminį, o kuri kita rodyklė bus naudojama (ar ateinanti nuo veiksnio, ar – nuo papildinio), nulems konkretaus sakinio žodžiai. Pvz., sakinio *Vakar aš mačiau savo motulę verkiančią* struktūra atrodytų taip (36 pav.):



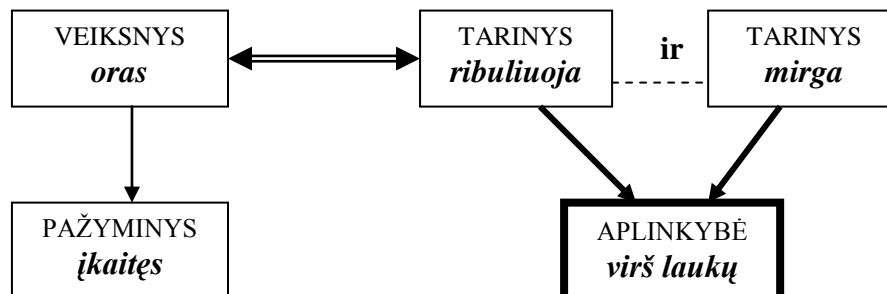
36 pav. Sakinio *Vakar aš mačiau savo motulę verkiančią* sintaksinė struktūra.

Šešėliai ant blokelių, atitinkančių sakinio dalis (žr. 33 pav., psl. 44), žymi galimus vienaarūšių sakinio dalių atvejus. Vienarūšės sakinio dalys siejamos sujungiamuoju ryšiu. V. Labutis sako, jog tai yra santykis tarp lygiaverčių žodžių, kurie formaliai vienas kito neveikia, ir nedominuoja vienas kito atžvilgiu. Juos lygiaverčiais daro jų vienodas santykis su kokiu nors kitu sakinio žodžiu [Labutis, 2002, psl. 27]. Todėl schemoje vienaarūšės sakinio dalys jungiamos linija be rodyklės, o virš tos linijos, kaip neorientuoto lanko žymė, pateikiamas jungtukas, jei jis sakinyje yra. Kadangi jungtukas nėra pavaldus kuriam nors sakinio žodžiui, ir yra susijęs su abiem jungiamais žodžiais vienodai, jis neįrašomas į vieną poziciją su kuria nors sakinio dalimi ir vertimo metu bus apdorojamas atskirai. 37 pav. parodyta sakinio (*Motina viską pastebi ir prašo negąsdinti vaiko ir neįvartyti jam baimės*), turinčio vienaarūšių sakinio dalių, struktūra.



37 pav. Sakinio *Motina viską pastebi ir prašo negąsdinti vaiko ir neįvartyti jam baimės* sintaksinė struktūra.

Jei kuris nors žodis išplečia kelias vienaarūšes sakinio dalis, į jį ateis kelios rodyklės iš visų išplečiamų žodžių. Tokio sakinio pavyzdys (*Įkaitęs oras ribuliuoja ir mirga virš laukų*) [Ulvydas, psl. 480], pateikiamas 38 pav., kur viena aplinkybė išplečia du vienaarūšius tarinius. Tai dar vienas atvejis, kai priklausomybių medis neturi priemonių pavaizduoti sakinio struktūros su visais sintaksiniais ryšiais. Vienaarūšes sakinio dalys paprastai yra susijusios su vienu ir tuo pačiu žodžiu, kurį jos išplečia. Jei dar ir jas išplečiantis žodis yra bendras, atsiranda uždaras kelias.



38 pav. Sakinio *Įkaitęs oras ribuliuoja ir mirga virš laukų* sintaksinė struktūra.

Apibendrintoje schemoje (žr. 33 pav., psl. 44) parodyti visi galimi lietuviško sakinio variantai. Kiekvienu konkrečiu atveju bus naudojamos to sakinio struktūrą atitinkančios grafo viršūnės bei lankai.

Reikia pasakyti, kad kiekviena šio skyriaus pavyzdžiuose pateikta schema yra apibendrintos struktūros dalis, t.y. apibendrinto lietuvių kalbos sakinio atskiras atvejis.

Apibendrintoje lietuvių kalbos sakinio sintaksinėje struktūroje neparodytas sakinio dalių išskaidymas pagal jų rūšis. Formaliai aprašant sakinio dalis tarinys skaidomas į vientisinį ir sudėtinį, o pažyminio pateikiami net penki tipai: derinamasis, nederinamasis, aplinkybinis, papildininis ir tarininis. Lituanistų darbuose labai akcentuojami tiesioginio ir netiesioginio papildinio skirtumai, tačiau šiame darbe jie apdorojami kartu, t.y. BNF apraše neišskiriami į atskirus atvejus, nes formaliai jų vartoseną niekuo nesiskiria – skirtumus tarp jų nulemia veiksmažodžio valtingumas, jo reikšmė, o ne paties papildinio savybės; t.y. ar papildinys tiesioginis ar netiesioginis priklauso nuo jį valdančio veiksmažodžio savybių ir tas pats žodis su vienu veiksmažodžiu gali būti tiesioginis, su kitu netiesioginis papildinys (*Visiems vadovavo Petras* – tiesioginis papildinys, *Saulė šviečia visiems* – netiesioginis, nes veiksmažodis *vadovauti* stipriai valdo naudininką, o veiksmažodis *šviešti* stipriai valdo šio linksnio valdymo neturi). Kaip kontrastą galima būtų pateikti derinamojo ir nederinamojo pažyminio skirtumus: jie negali būti apdorojami pagal tą pačią metodiką, nes pasikeitus pažymimojo žodžio linksniui, derinamasis pažyminys irgi keičia linksnį, o nederinamasis pažyminys visada išlieka toks pat: kilmininkas (*mano aukštas namas, mano aukšto namo, mano aukštam namui* ir t.t.) ar bendratis (*noras dirbti, noro dirbti* ir pan.). Taigi, skirtumą tarp šių dviejų pažyminio rūšių lemia paties pažyminio savybės, o ne daiktavardžio, kurį jie išplečia, reikšmė. Ar žodis eina derinamuoju pažyminiu ar ne, visai nepriklauso nuo to žodžio, kurį jis pažymi.

#### 4.4 Skyriaus išvados

- Iš šiuo metu naudojamų sintaksinės struktūros pavaizdavimo būdų lietuvių kalbos sakiniams geriau tinka priklausomybių metodas, nes jis sakinio struktūros nesiejia su žodžių tvarka sakinyje ir todėl gerai atspindi kalbų, turinčių laisvą žodžių tvarką (taigi ir lietuvių kalbos) savybes.
- Esant dideliame lietuvių kalbos kaitomumui (yra daug žodžių formų su skirtingomis galūnėmis) išskyla problemų pavaizduojant kai kurias sakinio dalis. Ypatingą vietą užima tarininis pažyminy. Išskirtinis jo bruožas yra tai, kad jis formaliai priklauso nuo dviejų sakinio dalių: nuo tarinio ir nuo veiksnio ar papildinio. Ryšys su tariniu apsprendžia jo vietą kitų sakinio žodžių atžvilgiu, o ryšys su veiksniu ar papildiniu nulemia jo galūnę.
- Priklausomybių medyje tegalima parodyti kiekvienos sakinio dalies priklausomybę tik nuo vieno žodžio. Lietuvių kalbos sakiniuose yra būtina parodyti tarinio pažyminio priklausomybę ir nuo tarinio, ir nuo veiksnio ar papildinio, nes priešingu atveju (terodant ryšį tik su viena šių sakinio dalių) bus prarandama sintaksinė informacija ir vertimo metu dėl to gali atsirasti klaidų. Todėl medis iš principo negali atspindėti visos sintaksinės informacijos, esančios lietuviškame sakinyje. Parodžius abu tarinio pažyminio ryšius, gaunamas uždaras kelias ir toks grafas jau nebėra medis.
- Modifikuotas priklausomybių gramatikoje naudojamas sakinio sintaksinės struktūros pateikimo būdas, papildant jį dvigubais sintaksiniais ryšiais, kuriuos turi tarininis pažyminy.



## 5 SINTAKSĖS TAISYKLIŲ UŽRAŠYMAS BNF

Jau sukurtas pasaulyje automatinio vertimo sistemas galima suskirstyti į tris tipus pagal jų sintaksinės analizės etapo ypatybes [Кунице, psl. 30]:

- 1) sistemos, kuriose iš viso nėra sintaksinės analizės;
- 2) sistemos, kuriose išskiriamas grynai sintaksinės analizės etapas (prieš jį eina morfologinė analizė, o po jo seka tam tikros semantinės operacijos, pagerinančios sintaksinės analizės darbo rezultatus);
- 3) sistemos, kuriose sintaksinė ir semantinė analizė sujungtos į vieną sudėtingą procedūrą, ir sintaksės bei semantikos sąveika užtikrina greitą ir gerą analizę.

Pirmajam tipui priskiriamos tiesioginio vertimo sistemos. Jose šaltinio teksto apdorojimas veda tiesiai į išverstą tekstą ir tokių sistemų vertimo kokybė labai prasta.

Antro tipo sistemų kokybė geresnė, bet čia gaunama daug perteklinių variantų, t.y. lieka pakankamai didelis daugiareikšmiškumas. Pvz., sakinio sintaksinės struktūros dažnai būna nevienareikšmės dėl pavartotų prielinksnių konstrukcijų, kurios gali būti labai sudėtingos. Prielinksninė konstrukcija susideda iš prielinksnio ir daiktavardinės frazės išreikštos kartais net ir kitomis prielinksninėmis konstrukcijomis. Kai sakinyje jų susikaupia daug, pasidaro sunku atsekti, kuris žodis kurį pažymi [Lehrberger, Bourbeau, psl. 100]. Didėjant prielinksnių konstrukcijų kiekiui sakinyje, galimų sintaksinių struktūrų skaičius auga eksponentiškai (katalano skaičiais  $(2n)!/((n+1)!n!)$ ). Iš 2 lentelės, kurioje pateikta galimų sakinio interpretacijų kiekio priklausomybė nuo jame pavartotų prielinksnių konstrukcijų skaičiaus, matyti, kad esant 6 prielinksniams sintaksinę struktūrą galima sudaryti 429 būdais [Kay, Gawron, Norvig, psl. 35].

**2 lentelė** Galimų sakinio struktūros variantų skaičiaus priklausomybė nuo pavartotų prielinksnių konstrukcijų kiekio.

Prielinksnių konstrukcijų skaičius	Galimų sakinio interpretacijų kiekis	Prielinksnių konstrukcijų skaičius	Galimų sakinio interpretacijų kiekis
0	1	5	132
1	2	6	429
2	5	7	1430
3	14	8	4862
4	42	9	16796

Konkretus pavyzdys, pateikiamas literatūroje [Kelly, psl. 15], būtų sakinių *The coastguard observed the yacht in the harbour with binoculars* ir *The gold watch was sold by the jeweller to a man with a beard* daugiareikšmiškumas. Sintaksinės priemonės čia jau nebegali nuspręsti, su kuriuo žodžiu siejasi prielinksnio *with* konstrukcija. Pirmuoju atveju žodžių junginys *with binoculars* susijęs su veiksmožodžiu *observe*, antruoju atveju *with a beard* pažymi daiktavardį *man*. Tai galima nustatyti tik semantinės informacijos pagalba, kuri žodį *binoculars*, turintį požymį „instrumentas“, leidžia susieti su žmogaus veiklos (taigi, ir suvokimo) veiksmožodžiais, tokiais kaip *observe*, o žodį *beard* draudžia laikyti tokio veiksmožodžio kaip *sell* papildiniu. Todėl naujausiose automatinio vertimo sistemose sintaksė ir semantika sudaro vieningą visumą.

Šiame darbe nauja yra tai, kad visi trys antrame punkte išvardinti etapai (psl. 48) sujungiami į vieną. **Sintaksinė analizė atliekama neatsietai nuo morfologijos ir, dar prieš sakinio apdorojimą sintaksiniu požiūriu, jo žodžiams nurodomi kai kurie semantiniai požymiai. Tokia metodika iš principo skiriasi nuo jau sukurtų sistemų strategijos, kai sintaksiniam daugiareikšmiškumui panaikinti pasitelkiama tik semantika.**

Automatinis vertimas, taigi, ir sintaksinė analizė priklauso dirbtinio intelekto sferai, todėl kuriant automatinę sintaksinę analizę būtina atsižvelgti į tai, kaip mąsto žmogus, skaitydamas vienos ar kitos kalbos sakinį. Norint atrinkti, kas su kuo yra susiję lietuviškame sakinyje, daugiausia dėmesio reikia skirti žodžių galūnėms, priešingai nei anglų kalboje, kur skaitant sakinį labiausiai žiūrima, kurioje vietoje stovi žodis, ir pagal tai sprendžiama apie jo vaidmenį sakinyje. Būtent tokiu principu ir pagrįsta yra anglų kalbos sintaksinė analizė. Lietuvių kalbos sakiniai turi būti analizuojami įvertinant specifines lietuvių kalbos savybes – didelį kaitomumą ir laisvą žodžių tvarką sakinyje, Taigi, čia beveik visiškai negalima remtis žodžio vieta ir būtina įvertinti daugybę galūnių, kurios ir turi pagrindinį sintaksinės informacijos krūvį.

Kuriant formalią gramatiką lietuvių kalbos sintaksei aprašyti, visos sintaksinės funkcijos diferencijuojamos pagal jomis einančių žodžių morfologines kategorijas, kad tiesioginiu sintaksiniu ryšiu susietus žodžius būtų galima rasti pagal morfologinių kategorijų atitikimą. Pvz., jei sakinyje yra daiktavardžio vardininkas, vienaskaita, moteriška giminė, tai sutikus vienaskaitos moteriškosios giminės būdvardžio vardininką, bus galima daryti išvadą, kad jis yra to daiktavardžio pažymyns.

Formalus lietuvių kalbos sintaksės aprašas pateikiamas BNF. Jis sudarytas iš dviejų dalių. Pirmą dalį skirtą sintaksinių funkcijų ir morfologinių kategorijų atitikimui, t.y. kiekvienai sintaksinei funkcijai nurodoma, kokiomis morfologinėmis formomis ji gali būti išreikšta (sakinio schemeje žr. 33 pav., psl. 44, tai atitiktų grafo viršūnes). Antroje dalyje aprašomi sintaksiniai ryšiai tarp žodžių (grafo lankai, jungiantys tas viršūnes), t.y. formalizuojamos žodžių junginių sudarymo taisyklės [Žandaris, psl. 284]. Čia ir įvertinama laisva žodžių tvarka lietuvių kalbos sakiniuose, kur kiekviena sakinio dalis gali būti ir pradžioje sakinio, ir jo viduryje, ir gale. Be to, tarp junginių sudarančių žodžių gali būti įsiterpę kitos, junginiui nepriklausančios sakinio dalys. Todėl būtina papildomai fiksuoti informaciją apie tiesioginiu sintaksiniu ryšiu susietų žodžių tarpusavio padėtį.

Formalios gramatikos neterminaliniai simboliai (plačiau žr. psl. 7) bus rašomi kampiniuose skliaustuose didžiosiomis raidėmis jungiant žodžius brūkšneliais. Tie žodžiai – tai sutrumpinti sintaksinių ir morfologinių kategorijų pavadinimai. Po jų pasvirusiu šriftu gali būti pateikiami semantiniai požymiai, kurių pagalba bus tikslinama žodžio sintaksinė funkcija, t.y. sprendžiama, kokiai sakinio daliai priskirti morfologinę formą. Pvz., daiktavardžio galininkas šiaip jau yra laikomas papildiniu (*dirbti žemę*), tačiau galininkas, turintis laiko požymį, eina laiko aplinkybe (*dirbti naktį*). Kaip semantiniai požymiai šiame darbe bus traktuojami ir valentiniai linksniai, t.y. tie linksniai, kurių buvimo sakinyje reikalauja žodžio reikšmė.

Terminaliniai simboliai rašomi mažosiomis raidėmis ir juose žodžiai jungiami pabraukimu. Terminaliniai simboliai susideda tik iš morfologinių kategorijų; semantiniai požymiai, taip pat kaip ir neterminaliniuose simboliuose pateikiami pasvirusiu šriftu. Terminaliniai simboliai ir yra pradiniai duomenys sintaksinei analizei atlikti.

## 5.1 Grafo viršūnės

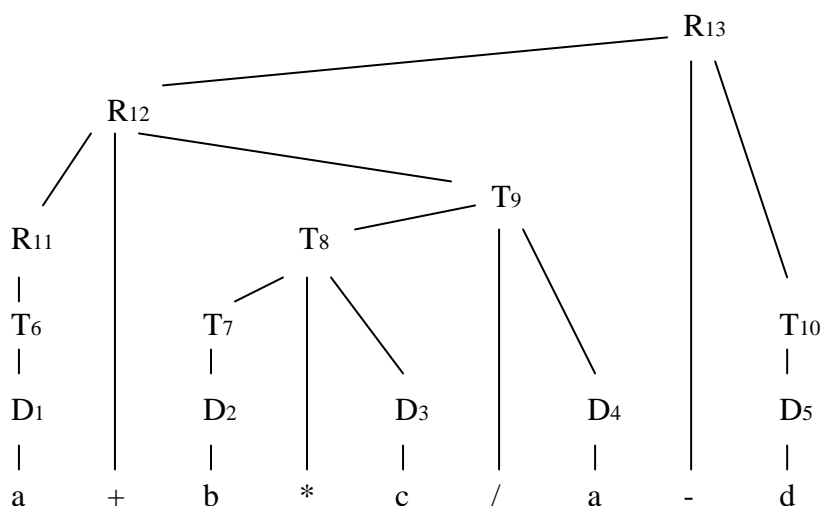
Sudarant BNF aprašą pagrindiniu informacijos šaltiniu apie lietuvių kalbos sintaksę laikomas „Lietuvių kalbos gramatikos“ trečiasis tomas [Ulvydas, psl. 279-566] ir, atsižvelgiant į ten lituanistų pateiktą informaciją apie sakinio dalis, formuojamas lietuvių kalbos sintaksės taisyklių aprašas BNF. Aptariant kiekvieną sakinio dalį, išvardijami minėtoje knygoje pateikti variantai ir pagrindžiama, kodėl atmetami kai kurie atvejai, jei tokių aprašomoje sakinio dalyje yra.

Sudarant kiekvienos sakinio dalies BNF aprašą parodomas labai detalus, laipsniškas perėjimas nuo bendresnių sąvokų prie konkretesnių, todėl kai kuriose taisyklėse vieni neterminaliniai simboliai keičiami kitais neterminaliniais simboliais. (Toks aprašymo būdas buvo pasirinktas dėl dviejų priežasčių: pirma, aprašas turi būti kuo lengviau suprantamas žmogui, ir antra, siekiama, kad būtų išvengta daugiareikšmiškumo pačioje formalioje gramatikoje.)

Iš pirmo žvilgsnio didelis neterminalinių simbolių kiekis gali atrodyti visai nereikalingas, nes vieno neterminalinio simbolio pakeitimas kitu iš esmės nieko neduoda. Tačiau taip elgiamasi ir matematikoje, aprašant aritmetinius reiškinius, kur daugiareikšmiškumas yra neleistinas. Pvz., formalioje gramatikoje, aprašančioje reiškinį  $a+b*c/a-d$  (jos taisyklės pateiktos 39 pav.), įvesti papildomi neterminaliniai simboliai (terminas, daugiklis) reikalingi tam, kad gramatika vienareikšmiškai nustatytų operacijų atlikimo tvarką. Ji parodyta 40 pav. Skaitmenys šalia raidžių medyje rodo operacijos eilės numerį.

$R \rightarrow R + T \mid R - T \mid T$	R – reiškinys
$T \rightarrow T * D \mid T / D \mid D$	T – terminas
$D \rightarrow a \mid b \mid c \mid d$	D – daugiklis

39 pav. Formali gramatika reiškinio  $a+b*c/a-d$  reikšmei apskaičiuoti.



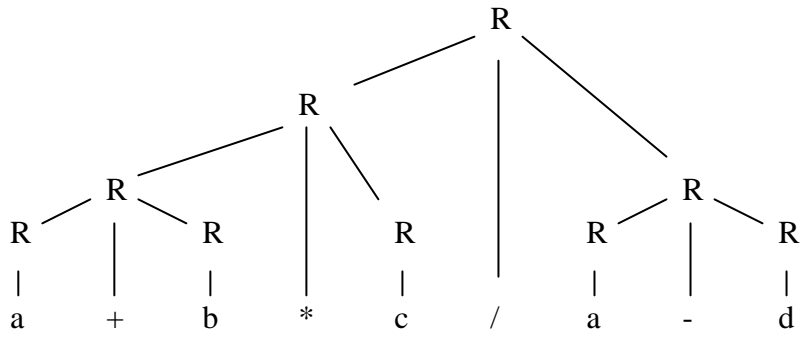
40 pav. Medis, pagal kurį skaičiuojama reiškinio  $a+b*c/a-d$  reikšmė.

Pradžioje visiems terminaliniams simboliams priskiriamas neterminalinis simbolis D. Tada žiūrima, ar iš gautos sekos galima kurioje nors vietoje suformuoti terminą. Kadangi du daugikliai sujungti operacijos ženklu negali būti terminas, todėl pasinaudojama trečia termino alternatyva ir vietoje pirmo daugiklio įrašoma T. Tačiau terminas ir daugiklis, sujungti pliuso

ženklų, negali sudaryti termo, todėl vėl imama trečia termo alternatyva ir antras daugiklis  $D_2$  keičiamas į termą  $T_7$ . Gautoje sekoje yra kombinacija terminas, daugybos ženklas ir daugiklis, o tai jau atitinka pirmąją termo alternatyvą, todėl ji ir pritaikoma šiems simboliams (gaunamas  $T_8$ ) ir t.t. Jei gramatika neturėtų papildomų neterminalinių simbolių (41 pav.), su ja nebūtų galima teisingai apskaičiuoti šio reiškinio reikšmės. Tai gerai matyti jo generavimo medyje, kuris parodytas 42 pav.

$R \rightarrow a \mid b \mid c \mid d \mid R + R \mid R - R \mid R * R \mid R / R$	$R$ – reiškinys
--	-----------------

41 pav. Gramatika, neturinti papildomų neterminalinių simbolių.



42 pav. Reiškinių  $a+b*c/a-d$  medis, sudarytas pagal gramatiką, neturinčią papildomų neterminalinių simbolių.

Kad būtų išvengta panašaus tipo klaidų, atliekant lietuvių kalbos sintaksinę analizę visos sintaksinės funkcijos pradžioje diferencijuojamos pagal kalbos dalis, kurios gali tą funkciją atlikti. Vėliau kiekviena kalbos dalis dar skirstoma pagal jai būdingas morfologines kategorijas. Pvz., veiksnio apraše iš pradžių nurodoma, kad tai gali būti veiksnys, išreikštas daiktavardžiu, veiksnys, išreikštas įvardžiu (daiktavardiškuoju), ir veiksnys, išreikštas bendratimi. Vėliau veiksnys, išreikštas daiktavardžiu, dar dalinamas į veiksnį, išreikštą daiktavardžio vardininku, vienaskaitos vyriškąją gimine; daiktavardžio vardininku, vienaskaitos moteriškąją gimine ir t.t. Veiksnys, išreikštas bendratimi, skaidomas pagal veiksmazodžio valtingumą, t.y. veiksnys, išreikštas bendratimi, nereikalaujančia jokio linksnio, veiksnys išreikštas bendratimi, reikalaujančia kilmininko, bendratimi, reikalaujančia audinininko, galinininko ir t.t. Šiame darbe semantiniai požymiai naudojami tik sakinių dalių nustatymo reikmėms, todėl valdomi linksniai, kurie tarnauja būtent šiam tikslui, taip pat priskiriami prie jų.

### 5.1.1 Veiksnių aprašas.

Lituanistų darbuose nurodoma, kad būdingiausia veiksnio forma yra vardininko linksnis. Jie pateikia net penkias kalbos dalis, kurių vardininkas gali eiti veiksniumi, tai – daiktavardis, įvardis, skaitvardis, būdvardis ir dalyvis. Daiktavardis ir įvardis įtraukiami į veiksnio aprašą BNF. Būdvardžio, dalyvio ir skaitvardžio atsisakoma dėl aiškiai praleisto jų pažymimo daiktavardžio, kuris iš esmės ir turėtų būti veiksnys. Tokių sakinių apdorojimas atidedamas ateičiai, nes jų vartoseną yra retoka ir jie bus įtraukti kartu su elipsėmis.

Lietuvių kalbos gramatikoje rašoma, kad yra dar vienas linksnis, kurį galima laikyti veiksmiu, tai – kilmininkas, ir pateikiami pavyzdžiai iš grožinės literatūros [Ulvydas, psl. 290]:

*Cinoko dar neišvirė, o Baltaragio jau susėdo valgyti.* (1)

Toliau toje pat knygoje pasakyta „*Kai šalia kilmininko iš konteksto lengvai numanomas pažymimasis žodis, kilmininkas laikomas ne veiksmiu, bet pažyminiu.*“ ir duodamas toks pavyzdys:

*Žiūri, aukšti jo dvarai, o berno – dar aukštesni.* (2)

*Tačiau (1) sakinyje pažymimasis žodis (šeima) nėra kįk ne sunkiau numanomas, negu (2) sakinyje – (dvarai).*

Kitas veiksnio kilmininko pavyzdys pateikiamas sakinyje:

*Kaimynų jau ir linus nurovė, o mūsų dar vasarojus ant lauko.* (3)

Šiandien lietuvis, skaitydamas tokį sakinį, suprastų, kad ne “kaimynai savo linus nurovė”, o “kaimynų linus kažkas kitas nurovė” (vagys ar pan.). *Panašu, kad toks veiksnys jau yra tapęs archaizmu ir šių dienų literatūroje nebėra aktualus, todėl kilmininko linksnis į veiksnio aprašą neįtraukiamas.*

Dar viena forma, laikoma veiksmiu, yra veiksmožodžio bendratis. Ji ir šiandieniniuose tekstuose vartojama gana dažnai, todėl įtraukiama į veiksnio aprašą, kaip trečia alternatyva. Tuo šis aprašas ir baigiamas, atmetant dar du literatūroje nurodomus atvejus: padalyvį ir bet kokius žodžius kabutėse. Padalyvio atsisakoma dėl jo nesistemingumo. Yra tik keli žodžiai lietuvių kalboje, kuriems galima pritaikyti tokį traktavimą. Gramatikoje pateikiamas sakinytis [Ulvydas, psl. 292]:

*Buvo girdėti griaudžiant.* (4)

Gal dar galėtų būti *buvo matyti lyjant* ar pan., tačiau tokių pavyzdžių nebus labai daug. Kiti žodžiai panašiose konstrukcijose iš viso nevartojami. Pvz., \* *buvo rašyti kalbant*, \* *buvo skubėti vėluojant* yra negalimi sakiniai lietuvių kalboje. Taigi, veiksnio padalyvis yra greičiau išimtis, o ne taisyklė, todėl sintaksinės analizės metu jis nebus apdorojamas.

Antro atvejo – bet kokių žodžių kabutėse – atsisakoma siekiant supaprastinti sintaksinės analizės realizaciją. Šis atvejis yra priimtinas ir ateityje, tobulinant ir plečiant sintaksinę analizę, greičiausiai bus įtrauktas.

Taigi, pradinė veiksnio aprašo forma BNF būtų tokia:

<VEIKSNYS> ::= <VEIKSN-DAIKT> | <VEIKSN-ĮVARD-DAIKT> | <VEIKSN-BENDR>;

Tai reiškia, kad sintaksinės analizės metu bus traktuojama, jog veiksnys gali būti išreikštas arba daiktavardžiu, arba įvardžiu, arba bendratimi (prie įvardžio kaip semantinis požymis nurodyta santrumpa *DAIKT* reiškia, kad veiksmiu gali eiti tik daiktavardiškieji įvardžiai).

Kiekvienas dešinėje pusėje esantis neterminalinis simbolis toliau detalizuojamas pagal morfologines kategorijas. Linksniuojamoms kalbos dalims nurodomi keturi vardinininko atvejai: vienaskaitos bei daugiskaitos vyriškoji ir moteriškoji giminė. Daiktavardžių bendroji giminė nebus įtraukiama dėl retos šių žodžių vartosenos. Veiksnys, išreikštas bendratimi diferencijuojamas pagal veiksmožodžio valentingumą, t.y. pagal tai, kokių linksnių reikalauja ar kokius gali jis prisijungti. Skirtingi veiksmožodžiai prisijungia ne tuos pačius linksnius, pvz., žodis *norėti* reikalauja kilmininko (*noriu ko? – obuolio*), žodis *matyti* – galinininko (*matau ką? – žvaigždę*), žodis *dovanoti* prisijungia du linksnius – galinininką ir naudinininką (*dovanoti kam? ir ką? – gėles mokytojai*), o žodis *miegoti* nereikalauja jokio linksnio. Visas veiksnio aprašas BNF atrodo taip:

<VEIKSNYS> ::= <VEIKSN-DAIKT> | <VEIKSN-[VARD-DAIKT]> | <VEIKSN-BENDRAT>;

<VEIKSN-DAIKT> ::= <VEIKSN-DAIKT-VARD-VNS-VYRG> |  
<VEIKSN-DAIKT-VARD-VNS-MOTG> |  
<VEIKSN-DAIKT-VARD-DGS-VYRG> |  
<VEIKSN-DAIKT-VARD-DGS-MOTG>;

<VEIKSN-[VARD-DAIKT]> ::= <VEIKSN-[VARD-VARD-VNS-VYRG-DAIKT]> |  
<VEIKSN-[VARD-VARD-VNS-MOTG-DAIKT]> |  
<VEIKSN-[VARD-VARD-DGS-VYRG-DAIKT]> |  
<VEIKSN-[VARD-VARD-DGS-MOTG-DAIKT]> |  
<VEIKSN-[VARD-BEVG]>;

<VEIKSN-BENDRAT> ::= <VEIKSN-BENDR> |  
<VEIKSN-BENDR-KILM> |  
<VEIKSN-BENDR-NAUD> |  
<VEIKSN-BENDR-GAL> |  
<VEIKSN-BENDR-[NAG]> |  
<VEIKSN-BENDR-VIET>;

<VEIKSN-DAIKT-VARD-VNS-VYRG> ::= daikt\_ward\_vns\_vyrg;  
<VEIKSN-DAIKT-VARD-VNS-MOTG> ::= daikt\_ward\_vns\_motg;  
<VEIKSN-DAIKT-VARD-DGS-VYRG> ::= daikt\_ward\_dgs\_vyrg;  
<VEIKSN-DAIKT-VARD-DGS-MOTG> ::= daikt\_ward\_dgs\_motg;

<VEIKSN-[VARD-VARD-VNS-VYRG-DAIKT]> ::= ivard\_ward\_vns\_vyrg\_daikt;  
<VEIKSN-[VARD-VARD-VNS-MOTG-DAIKT]> ::= ivard\_ward\_vns\_motg\_daikt;  
<VEIKSN-[VARD-VARD-DGS-VYRG-DAIKT]> ::= ivard\_ward\_dgs\_vyrg\_daikt;  
<VEIKSN-[VARD-VARD-DGS-MOTG-DAIKT]> ::= ivard\_ward\_dgs\_motg\_daikt;  
<VEIKSN-[VARD-BEVG]> ::= ivard\_bevg;

<VEIKSN-BENDR> ::= bendr;  
<VEIKSN-BENDR-KILM> ::= bendr\_kilm;  
<VEIKSN-BENDR-NAUD> ::= bendr\_naud;  
<VEIKSN-BENDR-GAL> ::= bendr\_gal;  
<VEIKSN-BENDR-[NAG]> ::= bendr\_inag;  
<VEIKSN-BENDR-VIET> ::= bendr\_viet;

## 5.1.2 Tarinio aprašas

Kompiuteriu apdorojant tarinį, jis suskirstomas (kaip ir lietuvių kalbos gramatikoje) į vientisinį ir sudėtinį pagal tai, iš kelių žodžių jis sudarytas. Vientisinį tarinį gali sudaryti asmenuojamosios ir neasmenuojamosios veiksmažodžio formos. Asmenuojamos formos BNF apraše pateikiamos visos. Neasmenuojamų formų apdorojimas atidedamas ateičiai, kai į sintaksinę analizę bus įtraukti ir sudėtiniai sakiniai, nes padalyviu išreikštas tarinys būna sakiniuose, kurie savo konstrukcija labai panašūs į sudėtinio prijungiamojo sakinio šalutinius dėmenis (*Kur čia kokį įnagį suradus. Kaip čia tave į miestą nugabenus. Kad bent greičiau į vietą nukeliavus.*). Netiesioginės nuosakos sakiniuose tariniu einantis dalyvis šiandien vartojamas labai retai (*Sakoma, kad Jūraitis nieko nesakęs tik juokijasis*). Dėl šios priežasties jo apdorojimas bus įtrauktas vėliau, kaip ir dalelytės, jaustuko bei ištiktuko, pateiktų gramatikoje. Vientisiniu tariniu einanti bendratis labai aiškiai rodo praleistą modalinį veiksmažodį (*Vilkas bėgt, meška vyti* – ėmė bėgt, pradėjo vyti. | *Tiek šieno grėbti, vežti, o ir dobilai jau pjaunami* – reikia grėbti, vežti), todėl tokio tipo tarinių apdorojimą numatoma įtraukti į analizę ateityje kartu su elipsiniais (nepilniaisiais, pvz., *Ir saulė ta pati pas juos ten. Ir ta pati pas mus.*) sakiniiais.

Lietuvių kalbos gramatikoje aprašomas dviejų tipų sudėtinis tarinys: suvestinis ir sudurtinis. Abu jie sudaromi iš dviejų žodžių, kurie gali būti nutolę vienas nuo kito, todėl BNF aprašas formuojamas tais pačiais principais, kaip ir žodžių junginiams.

Sudėtinis suvestinis tarinys pateikiamas kaip žodžių junginys, susidedantis iš modalinio veiksmažodžio ir bendraties. Suvestinio tarinio su dalyviu apdorojimas atidedamas vėlesniam laikui dėl retos jo vartosenos (*Jis girdavosi padedąs gėles globoti. Jie tikėjosi išgausią per mane žinias. Žmonės nustojo kalbėje.*) BNF apraše modaliniams veiksmažodžiams nurodomas semantinis požymis *MODALV*. Vokiečių autoriai teigia, kad modaliniai veiksmažodžiai neturi savarankiško valentingumo („*Modalverben ... bringen keine eigene Valenz mit.*“) [Agel, psl. 95]. Todėl modaliniais šiame darbe bus vadinami tie veiksmažodžiai, kurie neturi atskiro valdymo. Pvz., sakinyje *Aš noriu išmokti tą eilėrašį* tik žodis *išmokti* valdo žodį *eilėrašį*. Žodis *noriu* jokių kitų žodžių (be žodžio *išmokti*) atskirai nevaldo. Taigi, žodžiai *noriu* ir *išmokti* bus traktuojami kaip vienas tarinys ir sakinio struktūroje pateikiami vienoje pozicijoje. Sakinyje *Aš prašiau seserį išmokyti mane vokiečių kalbos* žodis *prašiau* ir žodis *išmokyti* valdo skirtingus žodžius (*prašiau seserį; išmokyti mane, išmokyti kalbos*), todėl jie negali būti surašyti į vieną poziciją, nes nuo kiekvieno iš jų turi eiti atskiros rodyklės į skirtingus žodžius. Todėl šiame sakinyje tariniu bus laikomas tik žodis *prašiau*, o žodis *išmokyti* jau bus traktuojamas kaip papildinys. Galima sakyti, kad modaliniais bus laikomi reikalaujantys tik vieno papildymo veiksmažodžiai.

Sudėtinis sudurtinis tarinys susideda iš jungties ir vardinės dalies. Jungtimi paprastai eina veiksmažodžio *būti* formos. Tas pats semantinis požymis *JUNGT* bus priskiriamas ir vadinamosios negrynosios jungties veiksmažodžiams (*likti, tapti, pasidaryti, atrodyti, jaustis* ir kt.). Sudarant vardinės dalies aprašą apsiribojama tik daiktavardžio ir būdvardžio vardininku (*Jie buvo tavo draugai. Aš esu laimingas.*). Visų kitų formų kol kas atsisakoma dėl retos vartosenos (Jo žiedas [*yra*] aukso. Jis [*yra*] pačiam smagume. ir pan.).

Taigi pradžioje tarinio aprašas skaidomas į dvi rūšis: vientisinį ir sudėtinį:

<TARIN> ::= <VIENT-TARIN> | <SUDĖT-TARIN>;

Veiksmažodžių valentingumui įvertinti linksniai, galintys eiti papildiniu, traktuojami kaip semantiniai požymiai. Jie pateikiami 3 lentelėje. Nors kai kurie veiksmažodžiai stipriai valdo vietininką (dalyvauti – kur?) tačiau jis visada eina aplinkybe, todėl šioje lentelėje neparodytas.

**3 lentelė** Linksniai, kurie traktuojami, kaip veiksmažodžių semantiniai požymiai.

Eil.nr.	Linksniai	Pavyzdžiai
1.	Kilmininkas	bijoti
2.	Naudininkas	pritarti
3.	Galininkas	matyti
4.	Įnagininkas	stebėtis

Veiksmažodžių, valdančių kelis linksnius ar prielinksnines konstrukcijas, apdorojimas nesiskiria nuo vieną linksnį valdančių veiksmažodžių, nes tiesioginio sintaksinio ryšio visada bus ieškoma tarp dviejų žodžių ir, jei veiksmažodis reikalauja kelių linksnių, sintaksinių ryšių bus surandama tiek, kiek sakinyje pavartota to veiksmažodžio valentinių palydovų.

Prielinksninės konstrukcijos (t.y. linksnis su prielinksniu) galinčios eiti papildiniu, taip pat įtraukiamos į semantinių požymių sąrašą. 4 lentelėje (psl. 56) pateikiami linksniai su prielinksniais, kurie bus traktuojami kaip veiksmažodžio semantiniai požymiai. Kol kas įtraukiami tik pirminiai prielinksniai, išvestinių prielinksnių apdorojimas atidedamas ateičiai, kai bus plečiama sintaksinė analizė.

4 lentelė Linksniai su prielinksniais, kurie traktuojami, kaip veiksmažodžių semantiniai požymiai.

Eil. nr.	Linksnis	Prielinksnis	Pavyzdžiai
1.	kilmininkas	ant	supykti <u>ant</u> kaimyno
2.	kilmininkas	be	apsieiti <u>be</u> pinigų
3.	kilmininkas	dėl	tartis <u>dėl</u> atlyginimo
4.	kilmininkas	iš	tyčiotis <u>iš</u> draugų
5.	kilmininkas	nuo	gintis <u>nuo</u> priešų
6.	kilmininkas	prie	prisitaikyti <u>prie</u> aplinkos
7.	galininkas	apie	pasakoti <u>apie</u> kelionę
8.	galininkas	į	atsakyti <u>į</u> klausimą
9.	galininkas	per	trenkti <u>per</u> ausį,
10.	galininkas	prieš	kovoti <u>prieš</u> engėjus
11.	galininkas	už	bausti <u>už</u> nusikaltimus
12.	įnagininkas	su	susitikti <u>su</u> draugais

Vientisinio tarinio aprašas BNF apima veiksmažodžio formą su jo valdomais linksniais, ar prielinksninėmis konstrukcijomis, kurie pateikiami kaip semantiniai požymiai. Tačiau kartais ir veiksmažodžio reikšmė gali turėti įtakos sakinio dalių nustatymui, pvz., bendratis, išplečiantis savarankišką veiksmažodį, dažniausiai eina papildiniu, tačiau bendratis, turinti slinkties požymį, eina aplinkybe, o ne papildiniu, todėl BNF apraše šis požymis nurodomas kartu su linksniais, kaip semantinis. Tos tarinio alternatyvos, kurios neturi semantinių požymių, aprašo jokie linksnio nereikalaujančius veiksmažodžius. Taigi, vientisinio tarinio aprašas būtų toks:

```

<VIENT-TARIN> ::= <VIENT-TARIN-VEIKSM-VNS-1A> |
<VIENT-TARIN-VEIKSM-VNS-2A> |
<VIENT-TARIN-VEIKSM-VNS-3A> |
<VIENT-TARIN-VEIKSM-DGS-1A> |
<VIENT-TARIN-VEIKSM-DGS-2A> |
<VIENT-TARIN-VEIKSM-DGS-3A> |
<VIENT-TARIN-VEIKSM-VNS-1A-KILM> |
<VIENT-TARIN-VEIKSM-VNS-2A-KILM> |
<VIENT-TARIN-VEIKSM-VNS-3A-KILM> |
<VIENT-TARIN-VEIKSM-DGS-1A-KILM> |
<VIENT-TARIN-VEIKSM-DGS-2A-KILM> |
<VIENT-TARIN-VEIKSM-DGS-3A-KILM> |
<VIENT-TARIN-VEIKSM-VNS-1A-NAUD> |
<VIENT-TARIN-VEIKSM-VNS-2A-NAUD> |
<VIENT-TARIN-VEIKSM-VNS-3A-NAUD> |
<VIENT-TARIN-VEIKSM-DGS-1A-NAUD> |
<VIENT-TARIN-VEIKSM-DGS-2A-NAUD> |
<VIENT-TARIN-VEIKSM-DGS-3A-NAUD> |
<VIENT-TARIN-VEIKSM-VNS-1A-GAL> |
<VIENT-TARIN-VEIKSM-VNS-2A-GAL> |
<VIENT-TARIN-VEIKSM-VNS-3A-GAL> |
<VIENT-TARIN-VEIKSM-DGS-1A-GAL> |
<VIENT-TARIN-VEIKSM-DGS-2A-GAL> |
<VIENT-TARIN-VEIKSM-DGS-3A-GAL> |
<VIENT-TARIN-VEIKSM-VNS-1A-INAG> |
<VIENT-TARIN-VEIKSM-VNS-2A-INAG> |
<VIENT-TARIN-VEIKSM-VNS-3A-INAG> |
<VIENT-TARIN-VEIKSM-DGS-1A-INAG> |
<VIENT-TARIN-VEIKSM-DGS-2A-INAG> |
<VIENT-TARIN-VEIKSM-DGS-3A-INAG> |
<VIENT-TARIN-VEIKSM-VNS-1A-SLINKT> |
<VIENT-TARIN-VEIKSM-VNS-2A-SLINKT> |
<VIENT-TARIN-VEIKSM-VNS-3A-SLINKT> |
<VIENT-TARIN-VEIKSM-DGS-1A-SLINKT> |
<VIENT-TARIN-VEIKSM-DGS-2A-SLINKT> |
<VIENT-TARIN-VEIKSM-DGS-3A-SLINKT> |

```



<VIENT-TARIN-VEIKSM-VNS-1A-ANTKILM> |  
<VIENT-TARIN-VEIKSM-VNS-2A-ANTKILM> |  
<VIENT-TARIN-VEIKSM-VNS-3A-ANTKILM> |  
<VIENT-TARIN-VEIKSM-DGS-1A-ANTKILM> |  
<VIENT-TARIN-VEIKSM-DGS-2A-ANTKILM> |  
<VIENT-TARIN-VEIKSM-DGS-3A-ANTKILM> |  
<VIENT-TARIN-VEIKSM-VNS-1A-BEKILM> |  
<VIENT-TARIN-VEIKSM-VNS-2A-BEKILM> |  
<VIENT-TARIN-VEIKSM-VNS-3A-BEKILM> |  
<VIENT-TARIN-VEIKSM-DGS-1A-BEKILM> |  
<VIENT-TARIN-VEIKSM-DGS-2A-BEKILM> |  
<VIENT-TARIN-VEIKSM-DGS-3A-BEKILM> |  
<VIENT-TARIN-VEIKSM-VNS-1A-DÉLKILM> |  
<VIENT-TARIN-VEIKSM-VNS-2A-DÉLKILM> |  
<VIENT-TARIN-VEIKSM-VNS-3A-DÉLKILM> |  
<VIENT-TARIN-VEIKSM-DGS-1A-DÉLKILM> |  
<VIENT-TARIN-VEIKSM-DGS-2A-DÉLKILM> |  
<VIENT-TARIN-VEIKSM-DGS-3A-DÉLKILM> |  
<VIENT-TARIN-VEIKSM-VNS-1A-IŠKILM> |  
<VIENT-TARIN-VEIKSM-VNS-2A-IŠKILM> |  
<VIENT-TARIN-VEIKSM-VNS-3A-IŠKILM> |  
<VIENT-TARIN-VEIKSM-DGS-1A-IŠKILM> |  
<VIENT-TARIN-VEIKSM-DGS-2A-IŠKILM> |  
<VIENT-TARIN-VEIKSM-DGS-3A-IŠKILM> |  
<VIENT-TARIN-VEIKSM-VNS-1A-NUOKILM> |  
<VIENT-TARIN-VEIKSM-VNS-2A-NUOKILM> |  
<VIENT-TARIN-VEIKSM-VNS-3A-NUOKILM> |  
<VIENT-TARIN-VEIKSM-DGS-1A-NUOKILM> |  
<VIENT-TARIN-VEIKSM-DGS-2A-NUOKILM> |  
<VIENT-TARIN-VEIKSM-DGS-3A-NUOKILM> |  
<VIENT-TARIN-VEIKSM-VNS-1A-PRIEKILM> |  
<VIENT-TARIN-VEIKSM-VNS-2A-PRIEKILM> |  
<VIENT-TARIN-VEIKSM-VNS-3A-PRIEKILM> |  
<VIENT-TARIN-VEIKSM-DGS-1A-PRIEKILM> |  
<VIENT-TARIN-VEIKSM-DGS-2A-PRIEKILM> |  
<VIENT-TARIN-VEIKSM-DGS-3A-PRIEKILM> |  
<VIENT-TARIN-VEIKSM-VNS-1A-APIEGAL> |  
<VIENT-TARIN-VEIKSM-VNS-2A-APIEGAL> |  
<VIENT-TARIN-VEIKSM-VNS-3A-APIEGAL> |  
<VIENT-TARIN-VEIKSM-DGS-1A-APIEGAL> |  
<VIENT-TARIN-VEIKSM-DGS-2A-APIEGAL> |  
<VIENT-TARIN-VEIKSM-DGS-3A-APIEGAL> |  
<VIENT-TARIN-VEIKSM-VNS-1A-[GAL]> |  
<VIENT-TARIN-VEIKSM-VNS-2A-[GAL]> |  
<VIENT-TARIN-VEIKSM-VNS-3A-[GAL]> |  
<VIENT-TARIN-VEIKSM-DGS-1A-[GAL]> |  
<VIENT-TARIN-VEIKSM-DGS-2A-[GAL]> |  
<VIENT-TARIN-VEIKSM-DGS-3A-[GAL]> |  
<VIENT-TARIN-VEIKSM-VNS-1A-PERGAL> |  
<VIENT-TARIN-VEIKSM-VNS-2A-PERGAL> |  
<VIENT-TARIN-VEIKSM-VNS-3A-PERGAL> |  
<VIENT-TARIN-VEIKSM-DGS-1A-PERGAL> |  
<VIENT-TARIN-VEIKSM-DGS-2A-PERGAL> |  
<VIENT-TARIN-VEIKSM-DGS-3A-PERGAL> |  
<VIENT-TARIN-VEIKSM-VNS-1A-PRIEŠGAL> |  
<VIENT-TARIN-VEIKSM-VNS-2A-PRIEŠGAL> |  
<VIENT-TARIN-VEIKSM-VNS-3A-PRIEŠGAL> |  
<VIENT-TARIN-VEIKSM-DGS-1A-PRIEŠGAL> |  
<VIENT-TARIN-VEIKSM-DGS-2A-PRIEŠGAL> |  
<VIENT-TARIN-VEIKSM-DGS-3A-PRIEŠGAL> |  
<VIENT-TARIN-VEIKSM-VNS-1A-UŽGAL> |  
<VIENT-TARIN-VEIKSM-VNS-2A-UŽGAL> |  
<VIENT-TARIN-VEIKSM-VNS-3A-UŽGAL> |  
<VIENT-TARIN-VEIKSM-DGS-1A-UŽGAL> |  
<VIENT-TARIN-VEIKSM-DGS-2A-UŽGAL> |  
<VIENT-TARIN-VEIKSM-DGS-3A-UŽGAL> |  
<VIENT-TARIN-VEIKSM-VNS-1A-SU[NAG]> |  
<VIENT-TARIN-VEIKSM-VNS-2A-SU[NAG]> |  
<VIENT-TARIN-VEIKSM-VNS-3A-SU[NAG]> |  
<VIENT-TARIN-VEIKSM-DGS-1A-SU[NAG]> |  
<VIENT-TARIN-VEIKSM-DGS-2A-SU[NAG]> |  
<VIENT-TARIN-VEIKSM-DGS-3A-SU[NAG]>;

<VIENT-TARIN-VEIKSM-VNS-1A> ::= veiksm\_vns\_1a;  
 <VIENT-TARIN-VEIKSM-VNS-2A> ::= veiksm\_vns\_2a;  
 <VIENT-TARIN-VEIKSM-VNS-3A> ::= veiksm\_vns\_3a;  
 <VIENT-TARIN-VEIKSM-DGS-1A> ::= veiksm\_dgs\_1a;  
 <VIENT-TARIN-VEIKSM-DGS-2A> ::= veiksm\_dgs\_2a;  
 <VIENT-TARIN-VEIKSM-DGS-3A> ::= veiksm\_dgs\_3a;  
 <VIENT-TARIN-VEIKSM-VNS-1A-KILM> ::= veiksm\_vns\_1a\_kilm;  
 <VIENT-TARIN-VEIKSM-VNS-2A-KILM> ::= veiksm\_vns\_2a\_kilm;  
 <VIENT-TARIN-VEIKSM-VNS-3A-KILM> ::= veiksm\_vns\_3a\_kilm;  
 <VIENT-TARIN-VEIKSM-DGS-1A-KILM> ::= veiksm\_dgs\_1a\_kilm;  
 <VIENT-TARIN-VEIKSM-DGS-2A-KILM> ::= veiksm\_dgs\_2a\_kilm;  
 <VIENT-TARIN-VEIKSM-DGS-3A-KILM> ::= veiksm\_dgs\_3a\_kilm;  
 <VIENT-TARIN-VEIKSM-VNS-1A-NAUD> ::= veiksm\_vns\_1a\_naud;  
 <VIENT-TARIN-VEIKSM-VNS-2A-NAUD> ::= veiksm\_vns\_2a\_naud;  
 <VIENT-TARIN-VEIKSM-VNS-3A-NAUD> ::= veiksm\_vns\_3a\_naud;  
 <VIENT-TARIN-VEIKSM-DGS-1A-NAUD> ::= veiksm\_vns\_1a\_naud;  
 <VIENT-TARIN-VEIKSM-DGS-2A-NAUD> ::= veiksm\_vns\_2a\_naud;  
 <VIENT-TARIN-VEIKSM-DGS-3A-NAUD> ::= veiksm\_vns\_3a\_naud;  
 <VIENT-TARIN-VEIKSM-VNS-1A-GAL> ::= veiksm\_vns\_1a\_gal;  
 <VIENT-TARIN-VEIKSM-VNS-2A-GAL> ::= veiksm\_vns\_2a\_gal;  
 <VIENT-TARIN-VEIKSM-VNS-3A-GAL> ::= veiksm\_vns\_3a\_gal;  
 <VIENT-TARIN-VEIKSM-DGS-1A-GAL> ::= veiksm\_vns\_1a\_gal;  
 <VIENT-TARIN-VEIKSM-DGS-2A-GAL> ::= veiksm\_vns\_2a\_gal;  
 <VIENT-TARIN-VEIKSM-DGS-3A-GAL> ::= veiksm\_vns\_3a\_gal;  
 <VIENT-TARIN-VEIKSM-VNS-1A-INAG> ::= veiksm\_vns\_1a\_inag;  
 <VIENT-TARIN-VEIKSM-VNS-2A-INAG> ::= veiksm\_vns\_2a\_inag;  
 <VIENT-TARIN-VEIKSM-VNS-3A-INAG> ::= veiksm\_vns\_3a\_inag;  
 <VIENT-TARIN-VEIKSM-DGS-1A-INAG> ::= veiksm\_vns\_1a\_inag;  
 <VIENT-TARIN-VEIKSM-DGS-2A-INAG> ::= veiksm\_vns\_2a\_inag;  
 <VIENT-TARIN-VEIKSM-DGS-3A-INAG> ::= veiksm\_vns\_3a\_inag;  
 <VIENT-TARIN-VEIKSM-VNS-1A-SLINKT> ::= veiksm\_vns\_1a\_slinkt;  
 <VIENT-TARIN-VEIKSM-VNS-2A-SLINKT> ::= veiksm\_vns\_2a\_slinkt;  
 <VIENT-TARIN-VEIKSM-VNS-3A-SLINKT> ::= veiksm\_vns\_3a\_slinkt;  
 <VIENT-TARIN-VEIKSM-DGS-1A-SLINKT> ::= veiksm\_vns\_1a\_slinkt;  
 <VIENT-TARIN-VEIKSM-DGS-2A-SLINKT> ::= veiksm\_vns\_2a\_slinkt;  
 <VIENT-TARIN-VEIKSM-DGS-3A-SLINKT> ::= veiksm\_vns\_3a\_slinkt;  
 <VIENT-TARIN-VEIKSM-VNS-1A-ANTKILM> ::= veiksm\_vns\_1a\_antkilm;  
 <VIENT-TARIN-VEIKSM-VNS-2A-ANTKILM> ::= veiksm\_vns\_2a\_antkilm;  
 <VIENT-TARIN-VEIKSM-VNS-3A-ANTKILM> ::= veiksm\_vns\_3a\_antkilm;  
 <VIENT-TARIN-VEIKSM-DGS-1A-ANTKILM> ::= veiksm\_dgs\_1a\_antkilm;  
 <VIENT-TARIN-VEIKSM-DGS-2A-ANTKILM> ::= veiksm\_dgs\_2a\_antkilm;  
 <VIENT-TARIN-VEIKSM-DGS-3A-ANTKILM> ::= veiksm\_dgs\_3a\_antkilm;  
 <VIENT-TARIN-VEIKSM-VNS-1A-BEKILM> ::= veiksm\_vns\_1a\_bekilm;  
 <VIENT-TARIN-VEIKSM-VNS-2A-BEKILM> ::= veiksm\_vns\_2a\_bekilm;  
 <VIENT-TARIN-VEIKSM-VNS-3A-BEKILM> ::= veiksm\_vns\_3a\_bekilm;  
 <VIENT-TARIN-VEIKSM-DGS-1A-BEKILM> ::= veiksm\_dgs\_1a\_bekilm;  
 <VIENT-TARIN-VEIKSM-DGS-2A-BEKILM> ::= veiksm\_dgs\_2a\_bekilm;  
 <VIENT-TARIN-VEIKSM-DGS-3A-BEKILM> ::= veiksm\_dgs\_3a\_bekilm;  
 <VIENT-TARIN-VEIKSM-VNS-1A-DÉLKILM> ::= veiksm\_vns\_1a\_dèlkilm;  
 <VIENT-TARIN-VEIKSM-VNS-2A-DÉLKILM> ::= veiksm\_vns\_2a\_dèlkilm;  
 <VIENT-TARIN-VEIKSM-VNS-3A-DÉLKILM> ::= veiksm\_vns\_3a\_dèlkilm;  
 <VIENT-TARIN-VEIKSM-DGS-1A-DÉLKILM> ::= veiksm\_dgs\_1a\_dèlkilm;  
 <VIENT-TARIN-VEIKSM-DGS-2A-DÉLKILM> ::= veiksm\_dgs\_2a\_dèlkilm;  
 <VIENT-TARIN-VEIKSM-DGS-3A-DÉLKILM> ::= veiksm\_dgs\_3a\_dèlkilm;  
 <VIENT-TARIN-VEIKSM-VNS-1A-IŠKILM> ::= veiksm\_vns\_1a\_iškilm;  
 <VIENT-TARIN-VEIKSM-VNS-2A-IŠKILM> ::= veiksm\_vns\_2a\_iškilm;  
 <VIENT-TARIN-VEIKSM-VNS-3A-IŠKILM> ::= veiksm\_vns\_3a\_iškilm;  
 <VIENT-TARIN-VEIKSM-DGS-1A-IŠKILM> ::= veiksm\_dgs\_1a\_iškilm;  
 <VIENT-TARIN-VEIKSM-DGS-2A-IŠKILM> ::= veiksm\_dgs\_2a\_iškilm;  
 <VIENT-TARIN-VEIKSM-DGS-3A-IŠKILM> ::= veiksm\_dgs\_3a\_iškilm;  
 <VIENT-TARIN-VEIKSM-VNS-1A-NUOKILM> ::= veiksm\_vns\_1a\_nuokilm;  
 <VIENT-TARIN-VEIKSM-VNS-2A-NUOKILM> ::= veiksm\_vns\_2a\_nuokilm;  
 <VIENT-TARIN-VEIKSM-VNS-3A-NUOKILM> ::= veiksm\_vns\_3a\_nuokilm;  
 <VIENT-TARIN-VEIKSM-DGS-1A-NUOKILM> ::= veiksm\_dgs\_1a\_nuokilm;  
 <VIENT-TARIN-VEIKSM-DGS-2A-NUOKILM> ::= veiksm\_dgs\_2a\_nuokilm;  
 <VIENT-TARIN-VEIKSM-DGS-3A-NUOKILM> ::= veiksm\_dgs\_3a\_nuokilm;  
 <VIENT-TARIN-VEIKSM-VNS-1A-PRIEKILM> ::= veiksm\_vns\_1a\_priekilm;  
 <VIENT-TARIN-VEIKSM-VNS-2A-PRIEKILM> ::= veiksm\_vns\_2a\_priekilm;  
 <VIENT-TARIN-VEIKSM-VNS-3A-PRIEKILM> ::= veiksm\_vns\_3a\_priekilm;  
 <VIENT-TARIN-VEIKSM-DGS-1A-PRIEKILM> ::= veiksm\_dgs\_1a\_priekilm;  
 <VIENT-TARIN-VEIKSM-DGS-2A-PRIEKILM> ::= veiksm\_dgs\_2a\_priekilm;  
 <VIENT-TARIN-VEIKSM-DGS-3A-PRIEKILM> ::= veiksm\_dgs\_3a\_priekilm;  
 <VIENT-TARIN-VEIKSM-VNS-1A-APIEGAL> ::= veiksm\_vns\_1a\_apiegal;  
 <VIENT-TARIN-VEIKSM-VNS-2A-APIEGAL> ::= veiksm\_vns\_2a\_apiegal;  
 <VIENT-TARIN-VEIKSM-VNS-3A-APIEGAL> ::= veiksm\_vns\_3a\_apiegal;

<VIENT-TARIN-VEIKSM-DGS-1A-APIEGAL> ::= veiksm\_vns\_1a\_apiegal;  
 <VIENT-TARIN-VEIKSM-DGS-2A-APIEGAL> ::= veiksm\_vns\_2a\_apiegal;  
 <VIENT-TARIN-VEIKSM-DGS-3A-APIEGAL> ::= veiksm\_vns\_3a\_apiegal;  
 <VIENT-TARIN-VEIKSM-VNS-1A-I GAL> ::= veiksm\_vns\_1a\_igal;  
 <VIENT-TARIN-VEIKSM-VNS-2A-I GAL> ::= veiksm\_vns\_2a\_igal;  
 <VIENT-TARIN-VEIKSM-VNS-3A-I GAL> ::= veiksm\_vns\_3a\_igal;  
 <VIENT-TARIN-VEIKSM-DGS-1A-I GAL> ::= veiksm\_vns\_1a\_igal;  
 <VIENT-TARIN-VEIKSM-DGS-2A-I GAL> ::= veiksm\_vns\_2a\_igal;  
 <VIENT-TARIN-VEIKSM-DGS-3A-I GAL> ::= veiksm\_vns\_3a\_igal;  
 <VIENT-TARIN-VEIKSM-VNS-1A-PER GAL> ::= veiksm\_vns\_1a\_pergal;  
 <VIENT-TARIN-VEIKSM-VNS-2A-PER GAL> ::= veiksm\_vns\_2a\_pergal;  
 <VIENT-TARIN-VEIKSM-VNS-3A-PER GAL> ::= veiksm\_vns\_3a\_pergal;  
 <VIENT-TARIN-VEIKSM-DGS-1A-PER GAL> ::= veiksm\_vns\_1a\_pergal;  
 <VIENT-TARIN-VEIKSM-DGS-2A-PER GAL> ::= veiksm\_vns\_2a\_pergal;  
 <VIENT-TARIN-VEIKSM-DGS-3A-PER GAL> ::= veiksm\_vns\_3a\_pergal;  
 <VIENT-TARIN-VEIKSM-VNS-1A-PRIEŠ GAL> ::= veiksm\_vns\_1a\_priesgal;  
 <VIENT-TARIN-VEIKSM-VNS-2A-PRIEŠ GAL> ::= veiksm\_vns\_2a\_priesgal;  
 <VIENT-TARIN-VEIKSM-VNS-3A-PRIEŠ GAL> ::= veiksm\_vns\_3a\_priesgal;  
 <VIENT-TARIN-VEIKSM-DGS-1A-PRIEŠ GAL> ::= veiksm\_vns\_1a\_priesgal;  
 <VIENT-TARIN-VEIKSM-DGS-2A-PRIEŠ GAL> ::= veiksm\_vns\_2a\_priesgal;  
 <VIENT-TARIN-VEIKSM-DGS-3A-PRIEŠ GAL> ::= veiksm\_vns\_3a\_priesgal;  
 <VIENT-TARIN-VEIKSM-VNS-1A-UŽ GAL> ::= veiksm\_vns\_1a\_uzgal;  
 <VIENT-TARIN-VEIKSM-VNS-2A-UŽ GAL> ::= veiksm\_vns\_2a\_uzgal;  
 <VIENT-TARIN-VEIKSM-VNS-3A-UŽ GAL> ::= veiksm\_vns\_3a\_uzgal;  
 <VIENT-TARIN-VEIKSM-DGS-1A-UŽ GAL> ::= veiksm\_vns\_1a\_uzgal;  
 <VIENT-TARIN-VEIKSM-DGS-2A-UŽ GAL> ::= veiksm\_vns\_2a\_uzgal;  
 <VIENT-TARIN-VEIKSM-DGS-3A-UŽ GAL> ::= veiksm\_vns\_3a\_uzgal;  
 <VIENT-TARIN-VEIKSM-VNS-1A-SU I NAG> ::= veiksm\_vns\_1a\_suinag;  
 <VIENT-TARIN-VEIKSM-VNS-2A-SU I NAG> ::= veiksm\_vns\_2a\_suinag;  
 <VIENT-TARIN-VEIKSM-VNS-3A-SU I NAG> ::= veiksm\_vns\_3a\_suinag;  
 <VIENT-TARIN-VEIKSM-DGS-1A-SU I NAG> ::= veiksm\_vns\_1a\_suinag;  
 <VIENT-TARIN-VEIKSM-DGS-2A-SU I NAG> ::= veiksm\_vns\_2a\_suinag;  
 <VIENT-TARIN-VEIKSM-DGS-3A-SU I NAG> ::= veiksm\_vns\_3a\_suinag;

Sudėtinis tarinys aprašomas BNF kaip žodžių junginys, t.y. kaip GIJA tarp dviejų žodžių, nurodant jų tarpusavio padėtį sakinyje ir galimus intarpo atvejus. Todėl jo aprašas bus pateiktas kitame skyriuje kartu su grafo lankais (psl. 73, skyriuje 5.2).

### 5.1.3 Pažyminio aprašas

Pažyminių lietuvių kalboje yra penkios rūšys:

- 1) derinamasis pažyminis,
- 2) nederinamasis pažyminis,
- 3) papildininis pažyminis,
- 4) aplinkybinis pažyminis,
- 5) tarininis pažyminis.

Ši informacija BNF būtų užrašoma taip:

<PAŽYMINYS> ::= <DERIN-PAŽYMINYS> | <NEDERIN-PAŽYMINYS> | <PAPILD-PAŽYMINYS> |  
 <APLINKYB-PAŽYMINYS> | <TARININ-PAŽYMINYS>;

Derinamojo pažyminio giminė, skaičius ir linksnis visada sutampa su pažymimojo žodžio linksniu, skaičiumi ir gimine (*senas namas*, *seną namą*, *seniems namams*). Juo gali eiti būdvardžiai, dalyviai ir skaitvardžiai. Lietuvių kalbos gramatikoje sakoma, kad derinamuoju pažyminiu gali eiti visi vyriškosios ir moteriškosios giminės būdvardžiai [Ulvydas, psl. 406], todėl bevardės giminės būdvardžių BNF pažyminio apraše nėra. Kadangi dalyvio laikas bei rūšis žodžio sintaksinei funkcijai įtakos neturi, todėl dalyviai BNF apraše skaidomi pagal tas pačias kategorijas kaip ir būdvardžiai – giminė, skaičius ir linksnis (pasikeitus dalyvio laikui ar rūšiai jo sintaksinė funkcija lieka ta pati, ko negalima pasakyti, pvz., apie linksnio kategoriją – pasikeitus daiktavardžio linksniui jo sintaksinė funkcija gali keistis: *Vilkas mato*

*kiški* – papildinys, *Kiški ausys kyšojo iš už kelmo* – pažymins). Visos sakinio dalys, kurios yra išreikštos daiktavardžiu, gali turėti derinamąjį pažyminį, todėl jo aprašas turi apimti visus šešis linksnius.

Nederinamuoju pažyminiu gali būti tikrai vienaskaitos arba daugiskaitos kilmininkas ir jis nėra derinamas su pažymimuoju žodžiu (*brolio namas, brolio namą, brolio namams; beržų viršūnės, beržų viršūnėse; mūsų vaikas, mūsų vaikui, mūsų vaikams* ir t.t.). Nederinamuoju pažyminiu taip pat laikoma veiksmazodžio bendratis (*noras dirbti, noro dirbti; viltis sugrižti, vilties sugrižti; teisė gyventi, teisę gyventi*).

Papildininiu pažyminiu laikomas įnagininko linksnis, kuris siejamas su veiksmazodiniais daiktavardžiais (*tikėjimas ateitimi, važiavimas dviračiu*). Naudininko linksnis, siejamas su daiktavardžiu, taip pat eina papildininiu pažyminiu (*tinkamumas darbu, drabužiai vaikams*). Papildiniam pažyminiui paprastai galima iškelti abu klausimus – ir papildinio ir pažyminio (*tikėjimas ateitimi: tikėjimas koks? ir tikėjimas kuo?; drabužiai vaikams: drabužiai kokie? ir drabužiai kam?*).

Aplinkybiniu pažyminiu laikomi daiktavardžių, turinčių laiko požymį, vienaskaitos galininkas ir daugiskaitos įnagininkas (*atostogos vasarą, darbas naktimis*). Vietininko linksnis, siejamas su daiktavardžiu, taip pat laikomas aplinkybiniu pažyminiu (*gyvenimas kaime, ežeras miške*). Aplinkybiniu pažyminiu gali eiti irrieveksmiai (*kelias atgal*).

Tarininis pažymins rodo ypatybę daikto, kuris eina veiksmu ar papildiniu. Tarininiu pažyminiu einančio įnagininko pavyzdžių gramatikoje nepateikiama, todėl BNF apraše bus tik keturi linksniai (su veiksmu susijusiam tariniam pažyminiui – vardininkas ir su papildiniu susijusiam – kilmininkas, naudininkas ir galininkas). Vietininkas tegali eiti aplinkybe, todėl į tarininio pažyminio aprašą jis neįtraukiamas. Tarinį pažyminį sudaro daiktavardžiai, būdvardžiai, būdvardiškieji įvardžiai bei skaitvardžiai ir dalyviai.

BNF apraše derinamasis, nederinamasis ir tarininis pažyminiai skaidomi pagal kalbos dalis, o aplinkybiniu ir papildininiu pažyminiu tegali eiti tik daiktavardis. Išskaidytas pagal kalbos dalis ir jų morfologines kategorijas pažyminio aprašas BNF atrodo taip:

```
<DERIN-PAŽYMINYS> ::= <DERIN-PAŽYM-BŪDV> |
<DERIN-PAŽYM-ĮVARD-BŪDV> |
<DERIN-PAŽYM-SKAITV-BŪDV> |
<DERIN-PAŽYM-DALYV>;
```

```
<DERIN-PAŽYM-BŪDV> ::= <DERIN-PAŽYM-BŪDV-VARD-VNS-VYRG> |
<DERIN-PAŽYM-BŪDV-VARD-VNS-MOTG> |
<DERIN-PAŽYM-BŪDV-VARD-DGS-VYRG> |
<DERIN-PAŽYM-BŪDV-VARD-DGS-MOTG> |
<DERIN-PAŽYM-BŪDV-KILM-VNS-VYRG> |
<DERIN-PAŽYM-BŪDV-KILM-VNS-MOTG> |
<DERIN-PAŽYM-BŪDV-KILM-DGS-VYRG> |
<DERIN-PAŽYM-BŪDV-KILM-DGS-MOTG> |
<DERIN-PAŽYM-BŪDV-NAUD-VNS-VYRG> |
<DERIN-PAŽYM-BŪDV-NAUD-VNS-MOTG> |
<DERIN-PAŽYM-BŪDV-NAUD-DGS-VYRG> |
<DERIN-PAŽYM-BŪDV-NAUD-DGS-MOTG> |
<DERIN-PAŽYM-BŪDV-GAL-VNS-VYRG> |
<DERIN-PAŽYM-BŪDV-GAL-VNS-MOTG> |
<DERIN-PAŽYM-BŪDV-GAL-DGS-VYRG> |
<DERIN-PAŽYM-BŪDV-GAL-DGS-MOTG> |
<DERIN-PAŽYM-BŪDV-ĮNAG-VNS-VYRG> |
<DERIN-PAŽYM-BŪDV-ĮNAG-VNS-MOTG> |
<DERIN-PAŽYM-BŪDV-ĮNAG-DGS-VYRG> |
<DERIN-PAŽYM-BŪDV-ĮNAG-DGS-MOTG> |
<DERIN-PAŽYM-BŪDV-VIET-VNS-VYRG> |
<DERIN-PAŽYM-BŪDV-VIET-VNS-MOTG> |
<DERIN-PAŽYM-BŪDV-VIET-DGS-VYRG> |
<DERIN-PAŽYM-BŪDV-VIET-DGS-MOTG>;
```

<DERIN-PAŽYM-IVARD> ::= <DERIN-PAŽYM-IVARD-VARD-VNS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-VARD-VNS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-VARD-DGS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-VARD-DGS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-KILM-VNS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-KILM-VNS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-KILM-DGS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-KILM-DGS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-NAUD-VNS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-NAUD-VNS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-NAUD-DGS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-NAUD-DGS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-GAL-VNS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-GAL-VNS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-GAL-DGS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-GAL-DGS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-[NAG]-VNS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-[NAG]-VNS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-[NAG]-DGS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-[NAG]-DGS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-VIET-VNS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-VIET-VNS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-VIET-DGS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-IVARD-VIET-DGS-MOTG-BŪDV>;

<DERIN-PAŽYM-SKAITV> ::= <DERIN-PAŽYM-SKAITV-VARD-VNS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-VARD-VNS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-VARD-DGS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-VARD-DGS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-KILM-VNS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-KILM-VNS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-KILM-DGS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-KILM-DGS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-NAUD-VNS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-NAUD-VNS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-NAUD-DGS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-NAUD-DGS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-GAL-VNS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-GAL-VNS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-GAL-DGS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-GAL-DGS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-[NAG]-VNS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-[NAG]-VNS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-[NAG]-DGS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-[NAG]-DGS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-VIET-VNS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-VIET-VNS-MOTG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-VIET-DGS-VYRG-BŪDV> |  
 <DERIN-PAŽYM-SKAITV-VIET-DGS-MOTG-BŪDV>;

<DERIN-PAŽYM-DALYV> ::= <DERIN-PAŽYM-DALYV-VARD-VNS-VYRG> |  
 <DERIN-PAŽYM-DALYV-VARD-VNS-MOTG> |  
 <DERIN-PAŽYM-DALYV-VARD-DGS-VYRG> |  
 <DERIN-PAŽYM-DALYV-VARD-DGS-MOTG> |  
 <DERIN-PAŽYM-DALYV-KILM-VNS-VYRG> |  
 <DERIN-PAŽYM-DALYV-KILM-VNS-MOTG> |  
 <DERIN-PAŽYM-DALYV-KILM-DGS-VYRG> |  
 <DERIN-PAŽYM-DALYV-KILM-DGS-MOTG> |  
 <DERIN-PAŽYM-DALYV-NAUD-VNS-VYRG> |  
 <DERIN-PAŽYM-DALYV-NAUD-VNS-MOTG> |  
 <DERIN-PAŽYM-DALYV-NAUD-DGS-VYRG> |  
 <DERIN-PAŽYM-DALYV-NAUD-DGS-MOTG> |  
 <DERIN-PAŽYM-DALYV-GAL-VNS-VYRG> |  
 <DERIN-PAŽYM-DALYV-GAL-VNS-MOTG> |  
 <DERIN-PAŽYM-DALYV-GAL-DGS-VYRG> |  
 <DERIN-PAŽYM-DALYV-GAL-DGS-MOTG> |  
 <DERIN-PAŽYM-DALYV-[NAG]-VNS-VYRG> |  
 <DERIN-PAŽYM-DALYV-[NAG]-VNS-MOTG> |  
 <DERIN-PAŽYM-DALYV-[NAG]-DGS-VYRG> |  
 <DERIN-PAŽYM-DALYV-[NAG]-DGS-MOTG> |  
 <DERIN-PAŽYM-DALYV-VIET-VNS-VYRG> |  
 <DERIN-PAŽYM-DALYV-VIET-VNS-MOTG> |  
 <DERIN-PAŽYM-DALYV-VIET-DGS-VYRG> |  
 <DERIN-PAŽYM-DALYV-VIET-DGS-MOTG>;

<NEDERIN-PAŽYMINYS> ::= <NEDERIN-PAŽYM-DAIKT> |  
<NEDERIN-PAŽYM-[VARD-DAIKT]>;

<NEDERIN-PAŽYM-DAIKT> ::= <NEDERIN-PAŽYM-DAIKT-KILM-VNS-VYRG> |  
<NEDERIN-PAŽYM-DAIKT-KILM-VNS-MOTG> |  
<NEDERIN-PAŽYM-DAIKT-KILM-DGS-VYRG> |  
<NEDERIN-PAŽYM-DAIKT-KILM-DGS-MOTG>;

<NEDERIN-PAŽYM-[VARD-DAIKT]> ::= <NEDERIN-PAŽYM-[VARD-KILM-VNS-VYRG-DAIKT]> |  
<NEDERIN-PAŽYM-[VARD-KILM-VNS-MOTG-DAIKT]> |  
<NEDERIN-PAŽYM-[VARD-KILM-DGS-VYRG-DAIKT]> |  
<NEDERIN-PAŽYM-[VARD-KILM-DGS-MOTG-DAIKT]>;

<PAPILD-PAŽYMINYS> ::= <PAPILD-PAŽYM-DAIKT-NAUD-VNS-VYRG> |  
<PAPILD-PAŽYM-DAIKT-NAUD-VNS-MOTG> |  
<PAPILD-PAŽYM-DAIKT-NAUD-DGS-VYRG> |  
<PAPILD-PAŽYM-DAIKT-NAUD-DGS-MOTG> |  
<PAPILD-PAŽYM-DAIKT-[NAG-VNS-VYRG]> |  
<PAPILD-PAŽYM-DAIKT-[NAG-VNS-MOTG]> |  
<PAPILD-PAŽYM-DAIKT-[NAG-DGS-VYRG]> |  
<PAPILD-PAŽYM-DAIKT-[NAG-DGS-MOTG]>;

<APLINKYB-PAŽYMINYS> ::= <APLINKYB-PAŽYM-DAIKT-GAL-VNS-VYRG-LAIKP> |  
<APLINKYB-PAŽYM-DAIKT-GAL-VNS-MOTG-LAIKP> |  
<APLINKYB-PAŽYM-DAIKT-GAL-DGS-VYRG-LAIKP> |  
<APLINKYB-PAŽYM-DAIKT-GAL-DGS-MOTG-LAIKP> |  
<APLINKYB-PAŽYM-DAIKT-[NAG-VNS-VYRG-LAIKP]> |  
<APLINKYB-PAŽYM-DAIKT-[NAG-VNS-MOTG-LAIKP]> |  
<APLINKYB-PAŽYM-DAIKT-[NAG-DGS-VYRG-LAIKP]> |  
<APLINKYB-PAŽYM-DAIKT-[NAG-DGS-MOTG-LAIKP]> |  
<APLINKYB-PAŽYM-DAIKT-[NAG-VNS-VYRG-VIETP]> |  
<APLINKYB-PAŽYM-DAIKT-[NAG-VNS-MOTG-VIETP]> |  
<APLINKYB-PAŽYM-DAIKT-[NAG-DGS-VYRG-VIETP]> |  
<APLINKYB-PAŽYM-DAIKT-[NAG-DGS-MOTG-VIETP]> |  
<APLINKYB-PAŽYM-DAIKT-VIET-VNS-VYRG> |  
<APLINKYB-PAŽYM-DAIKT-VIET-VNS-MOTG> |  
<APLINKYB-PAŽYM-DAIKT-VIET-DGS-VYRG> |  
<APLINKYB-PAŽYM-DAIKT-VIET-DGS-MOTG> |  
<APLINKYB-PAŽYM-PRIEV>;

<TARININ-PAŽYMINYS> ::= <TARININ-PAŽYM-DAIKT-VARD-VNS-VYRG> |  
<TARININ-PAŽYM-DAIKT-VARD-VNS-MOTG> |  
<TARININ-PAŽYM-DAIKT-VARD-DGS-VYRG> |  
<TARININ-PAŽYM-DAIKT-VARD-DGS-MOTG> |  
<TARININ-PAŽYM-DAIKT-KILM-VNS-VYRG> |  
<TARININ-PAŽYM-DAIKT-KILM-VNS-MOTG> |  
<TARININ-PAŽYM-DAIKT-KILM-DGS-VYRG> |  
<TARININ-PAŽYM-DAIKT-KILM-DGS-MOTG> |  
<TARININ-PAŽYM-DAIKT-NAUD-VNS-VYRG> |  
<TARININ-PAŽYM-DAIKT-NAUD-VNS-MOTG> |  
<TARININ-PAŽYM-DAIKT-NAUD-DGS-VYRG> |  
<TARININ-PAŽYM-DAIKT-NAUD-DGS-MOTG> |  
<TARININ-PAŽYM-DAIKT-GAL-VNS-VYRG> |  
<TARININ-PAŽYM-DAIKT-GAL-VNS-MOTG> |  
<TARININ-PAŽYM-DAIKT-GAL-DGS-VYRG> |  
<TARININ-PAŽYM-DAIKT-GAL-DGS-MOTG> |  
<TARININ-PAŽYM-BŪDV-VARD-VNS-VYRG> |  
<TARININ-PAŽYM-BŪDV-VARD-VNS-MOTG> |  
<TARININ-PAŽYM-BŪDV-VARD-DGS-VYRG> |  
<TARININ-PAŽYM-BŪDV-VARD-DGS-MOTG> |  
<TARININ-PAŽYM-BŪDV-KILM-VNS-VYRG> |  
<TARININ-PAŽYM-BŪDV-KILM-VNS-MOTG> |  
<TARININ-PAŽYM-BŪDV-KILM-DGS-VYRG> |  
<TARININ-PAŽYM-BŪDV-KILM-DGS-MOTG> |  
<TARININ-PAŽYM-BŪDV-NAUD-VNS-VYRG> |  
<TARININ-PAŽYM-BŪDV-NAUD-VNS-MOTG> |  
<TARININ-PAŽYM-BŪDV-NAUD-DGS-VYRG> |  
<TARININ-PAŽYM-BŪDV-NAUD-DGS-MOTG> |  
<TARININ-PAŽYM-BŪDV-GAL-VNS-VYRG> |  
<TARININ-PAŽYM-BŪDV-GAL-VNS-MOTG> |  
<TARININ-PAŽYM-BŪDV-GAL-DGS-VYRG> |  
<TARININ-PAŽYM-BŪDV-GAL-DGS-MOTG> |

<TARININ-PAŽYM-[IVARD-VARD-VNS-VYRG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-VARD-VNS-MOTG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-VARD-DGS-VYRG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-VARD-DGS-MOTG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-KILM-VNS-VYRG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-KILM-VNS-MOTG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-KILM-DGS-VYRG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-KILM-DGS-MOTG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-NAUD-VNS-VYRG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-NAUD-VNS-MOTG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-NAUD-DGS-VYRG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-NAUD-DGS-MOTG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-GAL-VNS-VYRG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-GAL-VNS-MOTG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-GAL-DGS-VYRG-BŪDV]> |  
 <TARININ-PAŽYM-[IVARD-GAL-DGS-MOTG-BŪDV]> |  
 <TARININ-PAŽYM-SKAITV-VARD-VNS-VYRG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-VARD-VNS-MOTG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-VARD-DGS-VYRG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-VARD-DGS-MOTG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-KILM-VNS-VYRG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-KILM-VNS-MOTG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-KILM-DGS-VYRG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-KILM-DGS-MOTG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-NAUD-VNS-VYRG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-NAUD-VNS-MOTG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-NAUD-DGS-VYRG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-NAUD-DGS-MOTG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-GAL-VNS-VYRG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-GAL-VNS-MOTG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-GAL-DGS-VYRG-BŪDV> |  
 <TARININ-PAŽYM-SKAITV-GAL-DGS-MOTG-BŪDV> |  
 <TARININ-PAŽYM-DALYV-VARD-VNS-VYRG> |  
 <TARININ-PAŽYM-DALYV-VARD-VNS-MOTG> |  
 <TARININ-PAŽYM-DALYV-VARD-DGS-VYRG> |  
 <TARININ-PAŽYM-DALYV-VARD-DGS-MOTG> |  
 <TARININ-PAŽYM-DALYV-KILM-VNS-VYRG> |  
 <TARININ-PAŽYM-DALYV-KILM-VNS-MOTG> |  
 <TARININ-PAŽYM-DALYV-KILM-DGS-VYRG> |  
 <TARININ-PAŽYM-DALYV-KILM-DGS-MOTG> |  
 <TARININ-PAŽYM-DALYV-NAUD-VNS-VYRG> |  
 <TARININ-PAŽYM-DALYV-NAUD-VNS-MOTG> |  
 <TARININ-PAŽYM-DALYV-NAUD-DGS-VYRG> |  
 <TARININ-PAŽYM-DALYV-NAUD-DGS-MOTG> |  
 <TARININ-PAŽYM-DALYV-GAL-VNS-VYRG> |  
 <TARININ-PAŽYM-DALYV-GAL-VNS-MOTG> |  
 <TARININ-PAŽYM-DALYV-GAL-DGS-VYRG> |  
 <TARININ-PAŽYM-DALYV-GAL-DGS-MOTG>;

<DERIN-PAŽYM-BŪDV-VARD-VNS-VYRG> ::= būdv\_ward\_vns\_vyrg;  
 <DERIN-PAŽYM-BŪDV-VARD-VNS-MOTG> ::= būdv\_ward\_vns\_motg;  
 <DERIN-PAŽYM-BŪDV-VARD-DGS-VYRG> ::= būdv\_ward\_dgs\_vyrg;  
 <DERIN-PAŽYM-BŪDV-VARD-DGS-MOTG> ::= būdv\_ward\_dgs\_motg;  
 <DERIN-PAŽYM-BŪDV-KILM-VNS-VYRG> ::= būdv\_kilm\_vns\_vyrg;  
 <DERIN-PAŽYM-BŪDV-KILM-VNS-MOTG> ::= būdv\_kilm\_vns\_motg;  
 <DERIN-PAŽYM-BŪDV-KILM-DGS-VYRG> ::= būdv\_kilm\_dgs\_vyrg;  
 <DERIN-PAŽYM-BŪDV-KILM-DGS-MOTG> ::= būdv\_kilm\_dgs\_motg;  
 <DERIN-PAŽYM-BŪDV-NAUD-VNS-VYRG> ::= būdv\_naud\_vns\_vyrg;  
 <DERIN-PAŽYM-BŪDV-NAUD-VNS-MOTG> ::= būdv\_naud\_vns\_motg;  
 <DERIN-PAŽYM-BŪDV-NAUD-DGS-VYRG> ::= būdv\_naud\_dgs\_vyrg;  
 <DERIN-PAŽYM-BŪDV-NAUD-DGS-MOTG> ::= būdv\_naud\_dgs\_motg;  
 <DERIN-PAŽYM-BŪDV-GAL-VNS-VYRG> ::= būdv\_gal\_vns\_vyrg;  
 <DERIN-PAŽYM-BŪDV-GAL-VNS-MOTG> ::= būdv\_gal\_vns\_motg;  
 <DERIN-PAŽYM-BŪDV-GAL-DGS-VYRG> ::= būdv\_gal\_dgs\_vyrg;  
 <DERIN-PAŽYM-BŪDV-GAL-DGS-MOTG> ::= būdv\_gal\_dgs\_motg;  
 <DERIN-PAŽYM-BŪDV-[NAG-VNS-VYRG]> ::= būdv\_inag\_vns\_vyrg;  
 <DERIN-PAŽYM-BŪDV-[NAG-VNS-MOTG]> ::= būdv\_inag\_vns\_motg;  
 <DERIN-PAŽYM-BŪDV-[NAG-DGS-VYRG]> ::= būdv\_inag\_dgs\_vyrg;  
 <DERIN-PAŽYM-BŪDV-[NAG-DGS-MOTG]> ::= būdv\_inag\_dgs\_motg;  
 <DERIN-PAŽYM-BŪDV-VIET-VNS-VYRG> ::= būdv\_viet\_vns\_vyrg;  
 <DERIN-PAŽYM-BŪDV-VIET-VNS-MOTG> ::= būdv\_viet\_vns\_motg;  
 <DERIN-PAŽYM-BŪDV-VIET-DGS-VYRG> ::= būdv\_viet\_dgs\_vyrg;  
 <DERIN-PAŽYM-BŪDV-VIET-DGS-MOTG> ::= būdv\_viet\_dgs\_motg;

<DERIN-PAŽYM-IVARD-VARD-VNS-VYRG-BŪDV> ::= įvard\_ward\_vns\_vyrg\_būdv;  
 <DERIN-PAŽYM-IVARD-VARD-VNS-MOTG-BŪDV> ::= įvard\_ward\_vns\_motg\_būdv;  
 <DERIN-PAŽYM-IVARD-VARD-DGS-VYRG-BŪDV> ::= įvard\_ward\_dgs\_vyrg\_būdv;  
 <DERIN-PAŽYM-IVARD-VARD-DGS-MOTG-BŪDV> ::= įvard\_ward\_dgs\_motg\_būdv;  
 <DERIN-PAŽYM-IVARD-KILM-VNS-VYRG-BŪDV> ::= įvard\_kilm\_vns\_vyrg\_būdv;  
 <DERIN-PAŽYM-IVARD-KILM-VNS-MOTG-BŪDV> ::= įvard\_kilm\_vns\_motg\_būdv;  
 <DERIN-PAŽYM-IVARD-KILM-DGS-VYRG-BŪDV> ::= įvard\_kilm\_dgs\_vyrg\_būdv;  
 <DERIN-PAŽYM-IVARD-KILM-DGS-MOTG-BŪDV> ::= įvard\_kilm\_dgs\_motg\_būdv;  
 <DERIN-PAŽYM-IVARD-NAUD-VNS-VYRG-BŪDV> ::= įvard\_naud\_vns\_vyrg\_būdv;  
 <DERIN-PAŽYM-IVARD-NAUD-VNS-MOTG-BŪDV> ::= įvard\_naud\_vns\_motg\_būdv;  
 <DERIN-PAŽYM-IVARD-NAUD-DGS-VYRG-BŪDV> ::= įvard\_naud\_dgs\_vyrg\_būdv;  
 <DERIN-PAŽYM-IVARD-NAUD-DGS-MOTG-BŪDV> ::= įvard\_naud\_dgs\_motg\_būdv;  
 <DERIN-PAŽYM-IVARD-GAL-VNS-VYRG-BŪDV> ::= įvard\_gal\_vns\_vyrg\_būdv;  
 <DERIN-PAŽYM-IVARD-GAL-VNS-MOTG-BŪDV> ::= įvard\_gal\_vns\_motg\_būdv;  
 <DERIN-PAŽYM-IVARD-GAL-DGS-VYRG-BŪDV> ::= įvard\_gal\_dgs\_vyrg\_būdv;  
 <DERIN-PAŽYM-IVARD-GAL-DGS-MOTG-BŪDV> ::= įvard\_gal\_dgs\_motg\_būdv;  
 <DERIN-PAŽYM-IVARD-INAG-VNS-VYRG-BŪDV> ::= įvard\_inag\_vns\_vyrg\_būdv;  
 <DERIN-PAŽYM-IVARD-INAG-VNS-MOTG-BŪDV> ::= įvard\_inag\_vns\_motg\_būdv;  
 <DERIN-PAŽYM-IVARD-INAG-DGS-VYRG-BŪDV> ::= įvard\_inag\_dgs\_vyrg\_būdv;  
 <DERIN-PAŽYM-IVARD-INAG-DGS-MOTG-BŪDV> ::= įvard\_inag\_dgs\_motg\_būdv;  
 <DERIN-PAŽYM-IVARD-VIET-VNS-VYRG-BŪDV> ::= įvard\_viet\_vns\_vyrg\_būdv;  
 <DERIN-PAŽYM-IVARD-VIET-VNS-MOTG-BŪDV> ::= įvard\_viet\_vns\_motg\_būdv;  
 <DERIN-PAŽYM-IVARD-VIET-DGS-VYRG-BŪDV> ::= įvard\_viet\_dgs\_vyrg\_būdv;  
 <DERIN-PAŽYM-IVARD-VIET-DGS-MOTG-BŪDV> ::= įvard\_viet\_dgs\_motg\_būdv;

<DERIN-PAŽYM-SKAITV-VARD-VNS-VYRG-BŪDV> ::= skaitv\_ward\_vns\_vyrg\_būdv;  
 <DERIN-PAŽYM-SKAITV-VARD-VNS-MOTG-BŪDV> ::= skaitv\_ward\_vns\_motg\_būdv;  
 <DERIN-PAŽYM-SKAITV-VARD-DGS-VYRG-BŪDV> ::= skaitv\_ward\_dgs\_vyrg\_būdv;  
 <DERIN-PAŽYM-SKAITV-VARD-DGS-MOTG-BŪDV> ::= skaitv\_ward\_dgs\_motg\_būdv;  
 <DERIN-PAŽYM-SKAITV-KILM-VNS-VYRG-BŪDV> ::= skaitv\_kilm\_vns\_vyrg\_būdv;  
 <DERIN-PAŽYM-SKAITV-KILM-VNS-MOTG-BŪDV> ::= skaitv\_kilm\_vns\_motg\_būdv;  
 <DERIN-PAŽYM-SKAITV-KILM-DGS-VYRG-BŪDV> ::= skaitv\_kilm\_dgs\_vyrg\_būdv;  
 <DERIN-PAŽYM-SKAITV-KILM-DGS-MOTG-BŪDV> ::= skaitv\_kilm\_dgs\_motg\_būdv;  
 <DERIN-PAŽYM-SKAITV-NAUD-VNS-VYRG-BŪDV> ::= skaitv\_naud\_vns\_vyrg\_būdv;  
 <DERIN-PAŽYM-SKAITV-NAUD-VNS-MOTG-BŪDV> ::= skaitv\_naud\_vns\_motg\_būdv;  
 <DERIN-PAŽYM-SKAITV-NAUD-DGS-VYRG-BŪDV> ::= skaitv\_naud\_dgs\_vyrg\_būdv;  
 <DERIN-PAŽYM-SKAITV-NAUD-DGS-MOTG-BŪDV> ::= skaitv\_naud\_dgs\_motg\_būdv;  
 <DERIN-PAŽYM-SKAITV-GAL-VNS-VYRG-BŪDV> ::= skaitv\_gal\_vns\_vyrg\_būdv;  
 <DERIN-PAŽYM-SKAITV-GAL-VNS-MOTG-BŪDV> ::= skaitv\_gal\_vns\_motg\_būdv;  
 <DERIN-PAŽYM-SKAITV-GAL-DGS-VYRG-BŪDV> ::= skaitv\_gal\_dgs\_vyrg\_būdv;  
 <DERIN-PAŽYM-SKAITV-GAL-DGS-MOTG-BŪDV> ::= skaitv\_gal\_dgs\_motg\_būdv;  
 <DERIN-PAŽYM-SKAITV-INAG-VNS-VYRG-BŪDV> ::= skaitv\_inag\_vns\_vyrg\_būdv;  
 <DERIN-PAŽYM-SKAITV-INAG-VNS-MOTG-BŪDV> ::= skaitv\_inag\_vns\_motg\_būdv;  
 <DERIN-PAŽYM-SKAITV-INAG-DGS-VYRG-BŪDV> ::= skaitv\_inag\_dgs\_vyrg\_būdv;  
 <DERIN-PAŽYM-SKAITV-INAG-DGS-MOTG-BŪDV> ::= skaitv\_inag\_dgs\_motg\_būdv;  
 <DERIN-PAŽYM-SKAITV-VIET-VNS-VYRG-BŪDV> ::= skaitv\_viet\_vns\_vyrg\_būdv;  
 <DERIN-PAŽYM-SKAITV-VIET-VNS-MOTG-BŪDV> ::= skaitv\_viet\_vns\_motg\_būdv;  
 <DERIN-PAŽYM-SKAITV-VIET-DGS-VYRG-BŪDV> ::= skaitv\_viet\_dgs\_vyrg\_būdv;  
 <DERIN-PAŽYM-SKAITV-VIET-DGS-MOTG-BŪDV> ::= skaitv\_viet\_dgs\_motg\_būdv;

<DERIN-PAŽYM-DALYV-VARD-VNS-VYRG> ::= dalyv\_ward\_vns\_vyrg;  
 <DERIN-PAŽYM-DALYV-VARD-VNS-MOTG> ::= dalyv\_ward\_vns\_motg;  
 <DERIN-PAŽYM-DALYV-VARD-DGS-VYRG> ::= dalyv\_ward\_dgst\_vyrg;  
 <DERIN-PAŽYM-DALYV-VARD-DGS-MOTG> ::= dalyv\_ward\_dgst\_motg;  
 <DERIN-PAŽYM-DALYV-KILM-VNS-VYRG> ::= dalyv\_kilm\_vns\_vyrg;  
 <DERIN-PAŽYM-DALYV-KILM-VNS-MOTG> ::= dalyv\_kilm\_vns\_motg;  
 <DERIN-PAŽYM-DALYV-KILM-DGS-VYRG> ::= dalyv\_kilm\_dgst\_vyrg;  
 <DERIN-PAŽYM-DALYV-KILM-DGS-MOTG> ::= dalyv\_kilm\_dgst\_motg;  
 <DERIN-PAŽYM-DALYV-NAUD-VNS-VYRG> ::= dalyv\_naud\_vns\_vyrg;  
 <DERIN-PAŽYM-DALYV-NAUD-VNS-MOTG> ::= dalyv\_naud\_vns\_motg;  
 <DERIN-PAŽYM-DALYV-NAUD-DGS-VYRG> ::= dalyv\_naud\_dgst\_vyrg;  
 <DERIN-PAŽYM-DALYV-NAUD-DGS-MOTG> ::= dalyv\_naud\_dgst\_motg;  
 <DERIN-PAŽYM-DALYV-GAL-VNS-VYRG> ::= dalyv\_gal\_vns\_vyrg;  
 <DERIN-PAŽYM-DALYV-GAL-VNS-MOTG> ::= dalyv\_gal\_vns\_motg;  
 <DERIN-PAŽYM-DALYV-GAL-DGS-VYRG> ::= dalyv\_gal\_dgst\_vyrg;  
 <DERIN-PAŽYM-DALYV-GAL-DGS-MOTG> ::= dalyv\_gal\_dgst\_motg;  
 <DERIN-PAŽYM-DALYV-INAG-VNS-VYRG> ::= dalyv\_inag\_vns\_vyrg;  
 <DERIN-PAŽYM-DALYV-INAG-VNS-MOTG> ::= dalyv\_inag\_vns\_motg;  
 <DERIN-PAŽYM-DALYV-INAG-DGS-VYRG> ::= dalyv\_inag\_dgst\_vyrg;  
 <DERIN-PAŽYM-DALYV-INAG-DGS-MOTG> ::= dalyv\_inag\_dgst\_motg;  
 <DERIN-PAŽYM-DALYV-VIET-VNS-VYRG> ::= dalyv\_viet\_vns\_vyrg;  
 <DERIN-PAŽYM-DALYV-VIET-VNS-MOTG> ::= dalyv\_viet\_vns\_motg;  
 <DERIN-PAŽYM-DALYV-VIET-DGS-VYRG> ::= dalyv\_viet\_dgst\_vyrg;  
 <DERIN-PAŽYM-DALYV-VIET-DGS-MOTG> ::= dalyv\_viet\_dgst\_motg;



<NEDERIN-PAŽYM-DAIKT-KILM-VNS-VYRG> ::= daikt\_kilm\_vns\_vyrg;  
<NEDERIN-PAŽYM-DAIKT-KILM-VNS-MOTG> ::= daikt\_kilm\_vns\_motg;  
<NEDERIN-PAŽYM-DAIKT-KILM-DGS-VYRG> ::= daikt\_kilm\_dgs\_vyrg;  
<NEDERIN-PAŽYM-DAIKT-KILM-DGS-MOTG> ::= daikt\_kilm\_dgs\_motg;

<NEDERIN-PAŽYM-IVARD-KILM-VNS-VYRG-DAIKT> ::= ivard\_kilm\_vns\_vyrg\_daikt;  
<NEDERIN-PAŽYM-IVARD-KILM-VNS-MOTG-DAIKT> ::= ivard\_kilm\_vns\_motg\_daikt;  
<NEDERIN-PAŽYM-IVARD-KILM-DGS-VYRG-DAIKT> ::= ivard\_kilm\_dgs\_vyrg\_daikt;  
<NEDERIN-PAŽYM-IVARD-KILM-DGS-MOTG-DAIKT> ::= ivard\_kilm\_dgs\_motg\_daikt;

<PAPILD-PAŽYM-DAIKT-NAUD-VNS-VYRG> ::= daikt\_naud\_vns\_vyrg;  
<PAPILD-PAŽYM-DAIKT-NAUD-VNS-MOTG> ::= daikt\_naud\_vns\_motg;  
<PAPILD-PAŽYM-DAIKT-NAUD-DGS-VYRG> ::= daikt\_naud\_dgs\_vyrg;  
<PAPILD-PAŽYM-DAIKT-NAUD-DGS-MOTG> ::= daikt\_naud\_dgs\_motg;  
<PAPILD-PAŽYM-DAIKT-INAG-VNS-VYRG> ::= daikt\_inag\_vns\_vyrg;  
<PAPILD-PAŽYM-DAIKT-INAG-VNS-MOTG> ::= daikt\_inag\_vns\_motg;  
<PAPILD-PAŽYM-DAIKT-INAG-DGS-VYRG> ::= daikt\_inag\_dgs\_vyrg;  
<PAPILD-PAŽYM-DAIKT-INAG-DGS-MOTG> ::= daikt\_inag\_dgs\_motg;

<APLINKYB-PAŽYM-DAIKT-GAL-VNS-VYRG-LAIKP> ::= daikt\_gal\_vns\_vyrg\_laiqp;  
<APLINKYB-PAŽYM-DAIKT-GAL-VNS-MOTG-LAIKP> ::= daikt\_gal\_vns\_motg\_laiqp;  
<APLINKYB-PAŽYM-DAIKT-GAL-DGS-VYRG-LAIKP> ::= daikt\_gal\_dgs\_vyrg\_laiqp;  
<APLINKYB-PAŽYM-DAIKT-GAL-DGS-MOTG-LAIKP> ::= daikt\_gal\_dgs\_motg\_laiqp;  
<APLINKYB-PAŽYM-DAIKT-INAG-VNS-VYRG-LAIKP> ::= daikt\_inag\_vns\_vyrg\_laiqp;  
<APLINKYB-PAŽYM-DAIKT-INAG-VNS-MOTG-LAIKP> ::= daikt\_inag\_vns\_motg\_laiqp;  
<APLINKYB-PAŽYM-DAIKT-INAG-DGS-VYRG-LAIKP> ::= daikt\_inag\_dgs\_vyrg\_laiqp;  
<APLINKYB-PAŽYM-DAIKT-INAG-DGS-MOTG-LAIKP> ::= daikt\_inag\_dgs\_motg\_laiqp;  
<APLINKYB-PAŽYM-DAIKT-INAG-VNS-VYRG-VIETP> ::= daikt\_inag\_vns\_vyrg\_vietp;  
<APLINKYB-PAŽYM-DAIKT-INAG-VNS-MOTG-VIETP> ::= daikt\_inag\_vns\_motg\_vietp;  
<APLINKYB-PAŽYM-DAIKT-INAG-DGS-VYRG-VIETP> ::= daikt\_inag\_dgs\_vyrg\_vietp;  
<APLINKYB-PAŽYM-DAIKT-INAG-DGS-MOTG-VIETP> ::= daikt\_inag\_dgs\_motg\_vietp;  
<APLINKYB-PAŽYM-DAIKT-VIET-VNS-VYRG> ::= daikt\_viet\_vns\_vyrg;  
<APLINKYB-PAŽYM-DAIKT-VIET-VNS-MOTG> ::= daikt\_viet\_vns\_motg;  
<APLINKYB-PAŽYM-DAIKT-VIET-DGS-VYRG> ::= daikt\_viet\_dgs\_vyrg;  
<APLINKYB-PAŽYM-DAIKT-VIET-DGS-MOTG> ::= daikt\_viet\_dgs\_motg;  
<APLINKYB-PAŽYM-PRIEV> ::= priev;

<TARININ-PAŽYM-DAIKT-VARD-VNS-VYRG> ::= daikt\_vard\_vns\_vyrg;  
<TARININ-PAŽYM-DAIKT-VARD-VNS-MOTG> ::= daikt\_vard\_vns\_motg;  
<TARININ-PAŽYM-DAIKT-VARD-DGS-VYRG> ::= daikt\_vard\_dgs\_vyrg;  
<TARININ-PAŽYM-DAIKT-VARD-DGS-MOTG> ::= daikt\_vard\_dgs\_motg;  
<TARININ-PAŽYM-DAIKT-KILM-VNS-VYRG> ::= daikt\_kilm\_vns\_vyrg;  
<TARININ-PAŽYM-DAIKT-KILM-VNS-MOTG> ::= daikt\_kilm\_vns\_motg;  
<TARININ-PAŽYM-DAIKT-KILM-DGS-VYRG> ::= daikt\_kilm\_dgs\_vyrg;  
<TARININ-PAŽYM-DAIKT-KILM-DGS-MOTG> ::= daikt\_kilm\_dgs\_motg;  
<TARININ-PAŽYM-DAIKT-NAUD-VNS-VYRG> ::= daikt\_naud\_vns\_vyrg;  
<TARININ-PAŽYM-DAIKT-NAUD-VNS-MOTG> ::= daikt\_naud\_vns\_motg;  
<TARININ-PAŽYM-DAIKT-NAUD-DGS-VYRG> ::= daikt\_naud\_dgs\_vyrg;  
<TARININ-PAŽYM-DAIKT-NAUD-DGS-MOTG> ::= daikt\_naud\_dgs\_motg;  
<TARININ-PAŽYM-DAIKT-GAL-VNS-VYRG> ::= daikt\_gal\_vns\_vyrg;  
<TARININ-PAŽYM-DAIKT-GAL-VNS-MOTG> ::= daikt\_gal\_vns\_motg;  
<TARININ-PAŽYM-DAIKT-GAL-DGS-VYRG> ::= daikt\_gal\_dgs\_vyrg;  
<TARININ-PAŽYM-DAIKT-GAL-DGS-MOTG> ::= daikt\_gal\_dgs\_motg;  
<TARININ-PAŽYM-BÜDV-VARD-VNS-VYRG> ::= būdv\_vard\_vns\_vyrg;  
<TARININ-PAŽYM-BÜDV-VARD-VNS-MOTG> ::= būdv\_vard\_vns\_motg;  
<TARININ-PAŽYM-BÜDV-VARD-DGS-VYRG> ::= būdv\_vard\_dgs\_vyrg;  
<TARININ-PAŽYM-BÜDV-VARD-DGS-MOTG> ::= būdv\_vard\_dgs\_motg;  
<TARININ-PAŽYM-BÜDV-KILM-VNS-VYRG> ::= būdv\_kilm\_vns\_vyrg;  
<TARININ-PAŽYM-BÜDV-KILM-VNS-MOTG> ::= būdv\_kilm\_vns\_motg;  
<TARININ-PAŽYM-BÜDV-KILM-DGS-VYRG> ::= būdv\_kilm\_dgs\_vyrg;  
<TARININ-PAŽYM-BÜDV-KILM-DGS-MOTG> ::= būdv\_kilm\_dgs\_motg;  
<TARININ-PAŽYM-BÜDV-NAUD-VNS-VYRG> ::= būdv\_naud\_vns\_vyrg;  
<TARININ-PAŽYM-BÜDV-NAUD-VNS-MOTG> ::= būdv\_naud\_vns\_motg;  
<TARININ-PAŽYM-BÜDV-NAUD-DGS-VYRG> ::= būdv\_naud\_dgs\_vyrg;  
<TARININ-PAŽYM-BÜDV-NAUD-DGS-MOTG> ::= būdv\_naud\_dgs\_motg;  
<TARININ-PAŽYM-BÜDV-GAL-VNS-VYRG> ::= būdv\_gal\_vns\_vyrg;  
<TARININ-PAŽYM-BÜDV-GAL-VNS-MOTG> ::= būdv\_gal\_vns\_motg;  
<TARININ-PAŽYM-BÜDV-GAL-DGS-VYRG> ::= būdv\_gal\_dgs\_vyrg;  
<TARININ-PAŽYM-BÜDV-GAL-DGS-MOTG> ::= būdv\_gal\_dgs\_motg;

<TARININ-PAŽYIM-IVARD-VARD-VNS-VYRG-BŪDV> ::= ivard\_var\_dgs\_vyrg\_būdv;  
 <TARININ-PAŽYIM-IVARD-VARD-VNS-MOTG-BŪDV> ::= ivard\_var\_vns\_motg\_būdv;  
 <TARININ-PAŽYIM-IVARD-VARD-DGS-VYRG-BŪDV> ::= ivard\_var\_dgs\_vyrg\_būdv;  
 <TARININ-PAŽYIM-IVARD-VARD-DGS-MOTG-BŪDV> ::= ivard\_var\_dgs\_motg\_būdv;  
 <TARININ-PAŽYIM-IVARD-KILM-VNS-VYRG-BŪDV> ::= ivard\_kilm\_vns\_vyrg\_būdv;  
 <TARININ-PAŽYIM-IVARD-KILM-VNS-MOTG-BŪDV> ::= ivard\_kilm\_vns\_motg\_būdv;  
 <TARININ-PAŽYIM-IVARD-KILM-DGS-VYRG-BŪDV> ::= ivard\_kilm\_dgs\_vyrg\_būdv;  
 <TARININ-PAŽYIM-IVARD-KILM-DGS-MOTG-BŪDV> ::= ivard\_kilm\_dgs\_motg\_būdv;  
 <TARININ-PAŽYIM-IVARD-NAUD-VNS-VYRG-BŪDV> ::= ivard\_naud\_vns\_vyrg\_būdv;  
 <TARININ-PAŽYIM-IVARD-NAUD-VNS-MOTG-BŪDV> ::= ivard\_naud\_vns\_motg\_būdv;  
 <TARININ-PAŽYIM-IVARD-NAUD-DGS-VYRG-BŪDV> ::= ivard\_naud\_dgs\_vyrg\_būdv;  
 <TARININ-PAŽYIM-IVARD-NAUD-DGS-MOTG-BŪDV> ::= ivard\_naud\_dgs\_motg\_būdv;  
 <TARININ-PAŽYIM-IVARD-GAL-VNS-VYRG-BŪDV> ::= ivard\_gal\_vns\_vyrg\_būdv;  
 <TARININ-PAŽYIM-IVARD-GAL-VNS-MOTG-BŪDV> ::= ivard\_gal\_vns\_motg\_būdv;  
 <TARININ-PAŽYIM-IVARD-GAL-DGS-VYRG-BŪDV> ::= ivard\_gal\_dgs\_vyrg\_būdv;  
 <TARININ-PAŽYIM-IVARD-GAL-DGS-MOTG-BŪDV> ::= ivard\_gal\_dgs\_motg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-VARD-VNS-VYRG-BŪDV> ::= skaitv\_var\_vns\_vyrg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-VARD-VNS-MOTG-BŪDV> ::= skaitv\_var\_vns\_motg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-VARD-DGS-VYRG-BŪDV> ::= skaitv\_var\_dgs\_vyrg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-VARD-DGS-MOTG-BŪDV> ::= skaitv\_var\_dgs\_motg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-KILM-VNS-VYRG-BŪDV> ::= skaitv\_kilm\_vns\_vyrg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-KILM-VNS-MOTG-BŪDV> ::= skaitv\_kilm\_vns\_motg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-KILM-DGS-VYRG-BŪDV> ::= skaitv\_kilm\_dgs\_vyrg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-KILM-DGS-MOTG-BŪDV> ::= skaitv\_kilm\_dgs\_motg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-NAUD-VNS-VYRG-BŪDV> ::= skaitv\_naud\_vns\_vyrg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-NAUD-VNS-MOTG-BŪDV> ::= skaitv\_naud\_vns\_motg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-NAUD-DGS-VYRG-BŪDV> ::= skaitv\_naud\_dgs\_vyrg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-NAUD-DGS-MOTG-BŪDV> ::= skaitv\_naud\_dgs\_motg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-GAL-VNS-VYRG-BŪDV> ::= skaitv\_gal\_vns\_vyrg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-GAL-VNS-MOTG-BŪDV> ::= skaitv\_gal\_vns\_vyrg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-GAL-DGS-VYRG-BŪDV> ::= skaitv\_gal\_dgs\_vyrg\_būdv;  
 <TARININ-PAŽYIM-SKAITV-GAL-DGS-MOTG-BŪDV> ::= skaitv\_gal\_dgs\_motg\_būdv;  
 <TARININ-PAŽYIM-DALYV-VARD-VNS-VYRG> ::= dalyv\_var\_vns\_vyrg;  
 <TARININ-PAŽYIM-DALYV-VARD-VNS-MOTG> ::= dalyv\_var\_vns\_motg;  
 <TARININ-PAŽYIM-DALYV-VARD-DGS-VYRG> ::= dalyv\_var\_dgs\_vyrg;  
 <TARININ-PAŽYIM-DALYV-VARD-DGS-MOTG> ::= dalyv\_var\_dgs\_motg;  
 <TARININ-PAŽYIM-DALYV-KILM-VNS-VYRG> ::= dalyv\_kilm\_vns\_vyrg;  
 <TARININ-PAŽYIM-DALYV-KILM-VNS-MOTG> ::= dalyv\_kilm\_vns\_motg;  
 <TARININ-PAŽYIM-DALYV-KILM-DGS-VYRG> ::= dalyv\_kilm\_dgs\_vyrg;  
 <TARININ-PAŽYIM-DALYV-KILM-DGS-MOTG> ::= dalyv\_kilm\_dgs\_motg;  
 <TARININ-PAŽYIM-DALYV-NAUD-VNS-VYRG> ::= dalyv\_naud\_vns\_vyrg;  
 <TARININ-PAŽYIM-DALYV-NAUD-VNS-MOTG> ::= dalyv\_naud\_vns\_motg;  
 <TARININ-PAŽYIM-DALYV-NAUD-DGS-VYRG> ::= dalyv\_naud\_dgs\_vyrg;  
 <TARININ-PAŽYIM-DALYV-NAUD-DGS-MOTG> ::= dalyv\_naud\_dgs\_motg;  
 <TARININ-PAŽYIM-DALYV-GAL-VNS-VYRG> ::= dalyv\_gal\_vns\_vyrg;  
 <TARININ-PAŽYIM-DALYV-GAL-VNS-MOTG> ::= dalyv\_gal\_vns\_motg;  
 <TARININ-PAŽYIM-DALYV-GAL-DGS-VYRG> ::= dalyv\_gal\_dgs\_vyrg;  
 <TARININ-PAŽYIM-DALYV-GAL-DGS-MOTG> ::= dalyv\_gal\_dgs\_motg;

#### 5.1.4 Papildinio aprašas

Papildiniu eina keturių linksnių (kilmininko, naudininko, galininko ir įnagininko) formos ir kai kurios prielinksninės konstrukcijos. Lituaniai sako, kad papildinio funkciją gali atlikti:

- 1) daiktavardis,
- 2) įvardis,
- 3) būdvardis,
- 4) dalyvis,
- 5) skaitvardis,
- 6) bendratis.

Papildiniu einančio būdvardžio, dalyvio ir skaitvardžio kol kas atsisakoma dėl tos pačios priežasties, kaip ir veiksniu apraše (žr. psl. 52).

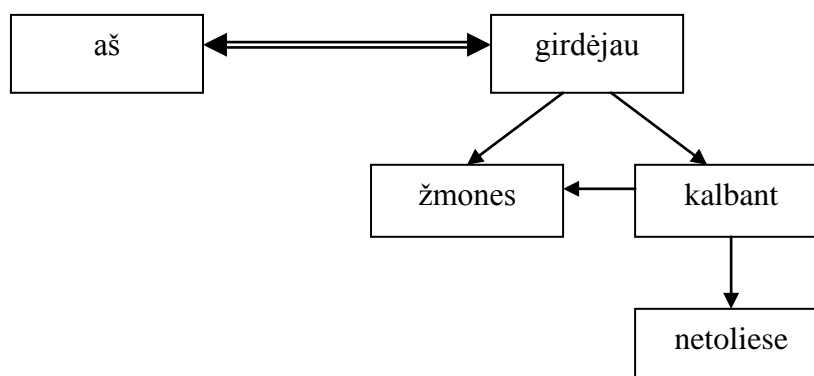
Gramatikoje dar pateikiami papildinio, išreikšto vardininku pavyzdžiai [Ulvydas, psl. 467]:

*Mažam ne taip lengva linai iš žemės rauti.* (5)

*Gaila man miškas kirsti.* (6)

Tačiau šių dienų vartosenos požiūriu tai jau yra nenatūrali lietuvių kalba; tokie sakiniai skamba dirbtinai, neįprastai. Todėl vardininko linksnis kaip papildinio forma į BNF aprašą neįtraukiamas.

Lietuvių kalbos gramatikoje aprašomi padalyvis, išiktukas, jaustukas ir prieveiksmis, kaip galintys eiti papildiniu. Visų šių formų laikinai atsisakoma dėl retos vartosenos. Atmetant padalyvį tuo pačiu nebus apdorojamas ir suvestinis papildinys, kurį sudaro daiktavardis su padalyviu. Čia būtų dar vienas atvejis, kai lietuvių kalbos sakinio struktūros negalima pavaizduoti medžiu. Suvestinio papildinio daiktavardis priklauso ir nuo tarinio ir nuo papildiniu einančio padalyvio. Pvz., sakinio *Aš girdėjau netoliese žmones kalbant* sintaksinė struktūra būtų tokia (43 pav.):



43 pav. Sakinio *Aš girdėjau netoliese žmones kalbant* sintaksinė struktūra.

Tačiau tokioje pozicijoje atskirti kuo eina padalyvis – papildiniu ar aplinkybe – galima tik semantikos pagalba. Jokios formalios priemonės šito skirtumo nerodo; pvz., sakiniuose *Mačiau skrendant paukštį danguje* ir *Mačiau lyjant paukštį danguje* papildinys *skrendant* ir aplinkybė *lyjant* formaliai nesiskiria niekuo. Todėl, kol neparuošta lietuvių kalbos automatinė semantinė (žodžių reikšmės) analizė, tokio tipo papildinių apdorojimas atidedamas ateičiai.

Taigi, pradinis papildinio aprašas BNF būtų toks:

<PAPILDINYS> ::= <PAPILD-DAIKT> | <PAPILD-IVARD-DAIKT> | <PAPILD-BENDRAT> | <PAPILD-PIELKONS>;

Toliau papildinio daiktavardis ir įvardis skaidomi pagal giminę, skaičių ir linksnį, bendratis – pagal valentingumą. Prielinksninės konstrukcijos aprašomos tuo pačiu būdu kaip žodžių junginiai ir jų aprašas pateikiamas kartu su grafo lankais psl. 73, skyriuje 5.2. Papildinio aprašas, išskaidytas pagal morfologines kategorijas atrodytų taip:

<PAPILD-DAIKT> ::= <PAPILD-DAIKT-KILM-VNS-VYRG> | <PAPILD-DAIKT-KILM-VNS-MOTG> | <PAPILD-DAIKT-KILM-DGS-VYRG> | <PAPILD-DAIKT-KILM-DGS-MOTG> | <PAPILD-DAIKT-NAUD-VNS-VYRG> | <PAPILD-DAIKT-NAUD-VNS-MOTG> | <PAPILD-DAIKT-NAUD-DGS-VYRG> | <PAPILD-DAIKT-NAUD-DGS-MOTG> | <PAPILD-DAIKT-GAL-VNS-VYRG> | <PAPILD-DAIKT-GAL-VNS-MOTG> | <PAPILD-DAIKT-GAL-DGS-VYRG> | <PAPILD-DAIKT-GAL-DGS-MOTG> | <PAPILD-DAIKT-ŪNAG-VNS-VYRG> | <PAPILD-DAIKT-ŪNAG-VNS-MOTG> | <PAPILD-DAIKT-ŪNAG-DGS-VYRG> | <PAPILD-DAIKT-ŪNAG-DGS-MOTG>;

<PAPILD-IVARD-DAIKT> ::= <PAPILD-IVARD-KILM-VNS-VYRG-DAIKT> |  
 <PAPILD-IVARD-KILM-VNS-MOTG-DAIKT> |  
 <PAPILD-IVARD-KILM-DGS-VYRG-DAIKT> |  
 <PAPILD-IVARD-KILM-DGS-MOTG-DAIKT> |  
 <PAPILD-IVARD-NAUD-VNS-VYRG-DAIKT> |  
 <PAPILD-IVARD-NAUD-VNS-MOTG-DAIKT> |  
 <PAPILD-IVARD-NAUD-DGS-VYRG-DAIKT> |  
 <PAPILD-IVARD-NAUD-DGS-MOTG-DAIKT> |  
 <PAPILD-IVARD-GAL-VNS-VYRG-DAIKT> |  
 <PAPILD-IVARD-GAL-VNS-MOTG-DAIKT> |  
 <PAPILD-IVARD-GAL-DGS-VYRG-DAIKT> |  
 <PAPILD-IVARD-GAL-DGS-MOTG-DAIKT> |  
 <PAPILD-IVARD-INAG-VNS-VYRG-DAIKT> |  
 <PAPILD-IVARD-INAG-VNS-MOTG-DAIKT> |  
 <PAPILD-IVARD-INAG-DGS-VYRG-DAIKT> |  
 <PAPILD-IVARD-INAG-DGS-MOTG-DAIKT>;

<PAPILD-BENDRAT> ::= <PAPILD-BENDR> |  
 <PAPILD-BENDR-KILM> |  
 <PAPILD-BENDR-NAUD> |  
 <PAPILD-BENDR-GAL> |  
 <PAPILD-BENDR-INAG> |  
 <PAPILD-BENDR-VIET>;

<PAPILD-PRIELKONS> ::= <GJA-ANTKILM #PRIEL-ANT+KILM> |  
 <GJA-BEKILM #PRIEL-BE+KILM> |  
 <GJA-DÉLKILM #PRIEL-DÉL+KILM> |  
 <GJA-IŠKILM #PRIEL-IŠ+KILM> |  
 <GJA-NUOKILM #PRIEL-NUO+KILM> |  
 <GJA-PRIEKILM #PRIEL-PRIE+KILM> |  
 <GJA-APIEGAL #PRIEL-APIE+GAL> |  
 <GJA-[GAL #PRIEL-[-+GAL> |  
 <GJA-PERGAL #PRIEL-PER+GAL> |  
 <GJA-PRIŠGAL #PRIEL-PRIŠ+GAL> |  
 <GJA-UŽGAL #PRIEL-UŽ+GAL> |  
 <GJA-SUINAG #PRIEL-SU+INAG>;

<GJA-ANTKILM #PRIEL-ANT+KILM> ::= <GJA-ANT-DAIKT-KILM-VNS-VYRG #PRIEL-ANT+DAIKT-KILM-VNS-VYRG> |  
 <GJA-ANT-DAIKT-KILM-VNS-MOTG #PRIEL-ANT+DAIKT-KILM-VNS-MOTG> |  
 <GJA-ANT-DAIKT-KILM-DGS-VYRG #PRIEL-ANT+DAIKT-KILM-DGS-VYRG> |  
 <GJA-ANT-DAIKT-KILM-DGS-MOTG #PRIEL-ANT+DAIKT-KILM-DGS-MOTG> |  
 <GJA-ANT-IVARD-KILM-VNS-VYRG #PRIEL-ANT+IVARD-KILM-VNS-VYRG> |  
 <GJA-ANT-IVARD-KILM-VNS-MOTG #PRIEL-ANT+IVARD-KILM-VNS-MOTG> |  
 <GJA-ANT-IVARD-KILM-DGS-VYRG #PRIEL-ANT+IVARD-KILM-DGS-VYRG> |  
 <GJA-ANT-IVARD-KILM-DGS-MOTG #PRIEL-ANT+IVARD-KILM-DGS-MOTG>;

<GJA-BEKILM #PRIEL-BE+KILM> ::= <GJA-BE-DAIKT-KILM-VNS-VYRG #PRIEL-BE+DAIKT-KILM-VNS-VYRG> |  
 <GJA-BE-DAIKT-KILM-VNS-MOTG #PRIEL-BE+DAIKT-KILM-VNS-MOTG> |  
 <GJA-BE-DAIKT-KILM-DGS-VYRG #PRIEL-BE+DAIKT-KILM-DGS-VYRG> |  
 <GJA-BE-DAIKT-KILM-DGS-MOTG #PRIEL-BE+DAIKT-KILM-DGS-MOTG> |  
 <GJA-BE-IVARD-KILM-VNS-VYRG #PRIEL-BE+IVARD-KILM-VNS-VYRG> |  
 <GJA-BE-IVARD-KILM-VNS-MOTG #PRIEL-BE+IVARD-KILM-VNS-MOTG> |  
 <GJA-BE-IVARD-KILM-DGS-VYRG #PRIEL-BE+IVARD-KILM-DGS-VYRG> |  
 <GJA-BE-IVARD-KILM-DGS-MOTG #PRIEL-BE+IVARD-KILM-DGS-MOTG>;

<GJA-DÉLKILM #PRIEL-DÉL+KILM> ::= <GJA-DÉL-DAIKT-KILM-VNS-VYRG #PRIEL-DÉL+DAIKT-KILM-VNS-VYRG> |  
 <GJA-DÉL-DAIKT-KILM-VNS-MOTG #PRIEL-DÉL+DAIKT-KILM-VNS-MOTG> |  
 <GJA-DÉL-DAIKT-KILM-DGS-VYRG #PRIEL-DÉL+DAIKT-KILM-DGS-VYRG> |  
 <GJA-DÉL-DAIKT-KILM-DGS-MOTG #PRIEL-DÉL+DAIKT-KILM-DGS-MOTG> |  
 <GJA-DÉL-IVARD-KILM-VNS-VYRG #PRIEL-DÉL+IVARD-KILM-VNS-VYRG> |  
 <GJA-DÉL-IVARD-KILM-VNS-MOTG #PRIEL-DÉL+IVARD-KILM-VNS-MOTG> |  
 <GJA-DÉL-IVARD-KILM-DGS-VYRG #PRIEL-DÉL+IVARD-KILM-DGS-VYRG> |  
 <GJA-DÉL-IVARD-KILM-DGS-MOTG #PRIEL-DÉL+IVARD-KILM-DGS-MOTG>;

<GJA-IŠKILM #PRIEL-IŠ+KILM> ::= <GJA-IŠ-DAIKT-KILM-VNS-VYRG #PRIEL-IŠ+DAIKT-KILM-VNS-VYRG> |  
 <GJA-IŠ-DAIKT-KILM-VNS-MOTG #PRIEL-IŠ+DAIKT-KILM-VNS-MOTG> |  
 <GJA-IŠ-DAIKT-KILM-DGS-VYRG #PRIEL-IŠ+DAIKT-KILM-DGS-VYRG> |  
 <GJA-IŠ-DAIKT-KILM-DGS-MOTG #PRIEL-IŠ+DAIKT-KILM-DGS-MOTG> |  
 <GJA-IŠ-IVARD-KILM-VNS-VYRG #PRIEL-IŠ+IVARD-KILM-VNS-VYRG> |  
 <GJA-IŠ-IVARD-KILM-VNS-MOTG #PRIEL-IŠ+IVARD-KILM-VNS-MOTG> |  
 <GJA-IŠ-IVARD-KILM-DGS-VYRG #PRIEL-IŠ+IVARD-KILM-DGS-VYRG> |  
 <GJA-IŠ-IVARD-KILM-DGS-MOTG #PRIEL-IŠ+IVARD-KILM-DGS-MOTG>;

<GJJA-NUOKILM #PRIEL-NUO+KILM> ::= <GJJA-NUO-DAIKT-KILM-VNS-VYRG #PRIEL-NUO+DAIKT-KILM-VNS-VYRG> |  
 <GJJA-NUO-DAIKT-KILM-VNS-MOTG #PRIEL-NUO+DAIKT-KILM-VNS-MOTG> |  
 <GJJA-NUO-DAIKT-KILM-DGS-VYRG #PRIEL-NUO+DAIKT-KILM-DGS-VYRG> |  
 <GJJA-NUO-DAIKT-KILM-DGS-MOTG #PRIEL-NUO+DAIKT-KILM-DGS-MOTG> |  
 <GJJA-NUO-[VARD-KILM-VNS-VYRG #PRIEL-NUO+[VARD-KILM-VNS-VYRG]> |  
 <GJJA-NUO-[VARD-KILM-VNS-MOTG #PRIEL-NUO+[VARD-KILM-VNS-MOTG]> |  
 <GJJA-NUO-[VARD-KILM-DGS-VYRG #PRIEL-NUO+[VARD-KILM-DGS-VYRG]> |  
 <GJJA-NUO-[VARD-KILM-DGS-MOTG #PRIEL-NUO+[VARD-KILM-DGS-MOTG]>;

<GJJA-PRIEKILM #PRIEL-PRIE+KILM> ::= <GJJA-PRIE-DAIKT-KILM-VNS-VYRG #PRIEL-PRIE+DAIKT-KILM-VNS-VYRG> |  
 <GJJA-PRIE-DAIKT-KILM-VNS-MOTG #PRIEL-PRIE+DAIKT-KILM-VNS-MOTG> |  
 <GJJA-PRIE-DAIKT-KILM-DGS-VYRG #PRIEL-PRIE+DAIKT-KILM-DGS-VYRG> |  
 <GJJA-PRIE-DAIKT-KILM-DGS-MOTG #PRIEL-PRIE+DAIKT-KILM-DGS-MOTG> |  
 <GJJA-PRIE-[VARD-KILM-VNS-VYRG #PRIEL-PRIE+[VARD-KILM-VNS-VYRG]> |  
 <GJJA-PRIE-[VARD-KILM-VNS-MOTG #PRIEL-PRIE+[VARD-KILM-VNS-MOTG]> |  
 <GJJA-PRIE-[VARD-KILM-DGS-VYRG #PRIEL-PRIE+[VARD-KILM-DGS-VYRG]> |  
 <GJJA-PRIE-[VARD-KILM-DGS-MOTG #PRIEL-PRIE+[VARD-KILM-DGS-MOTG]>;

<GJJA-APIEGAL #PRIEL-APIE+GAL> ::= <GJJA-APIE-DAIKT-GAL-VNS-VYRG #PRIEL-APIE+DAIKT-GAL-VNS-VYRG> |  
 <GJJA-APIE-DAIKT-GAL-VNS-MOTG #PRIEL-APIE+DAIKT-GAL-VNS-MOTG> |  
 <GJJA-APIE-DAIKT-GAL-DGS-VYRG #PRIEL-APIE+DAIKT-GAL-DGS-VYRG> |  
 <GJJA-APIE-DAIKT-GAL-DGS-MOTG #PRIEL-APIE+DAIKT-GAL-DGS-MOTG> |  
 <GJJA-APIE-[VARD-GAL-VNS-VYRG #PRIEL-APIE+[VARD-GAL-VNS-VYRG]> |  
 <GJJA-APIE-[VARD-GAL-VNS-MOTG #PRIEL-APIE+[VARD-GAL-VNS-MOTG]> |  
 <GJJA-APIE-[VARD-GAL-DGS-VYRG #PRIEL-APIE+[VARD-GAL-DGS-VYRG]> |  
 <GJJA-APIE-[VARD-GAL-DGS-MOTG #PRIEL-APIE+[VARD-GAL-DGS-MOTG]>;

<GJJA-[GAL #PRIEL-]+GAL> ::= <GJJA-[DAIKT-GAL-VNS-VYRG #PRIEL-]+DAIKT-GAL-VNS-VYRG> |  
 <GJJA-[DAIKT-GAL-VNS-MOTG #PRIEL-]+DAIKT-GAL-VNS-MOTG> |  
 <GJJA-[DAIKT-GAL-DGS-VYRG #PRIEL-]+DAIKT-GAL-DGS-VYRG> |  
 <GJJA-[DAIKT-GAL-DGS-MOTG #PRIEL-]+DAIKT-GAL-DGS-MOTG> |  
 <GJJA-[VARD-GAL-VNS-VYRG #PRIEL-]+VARD-GAL-VNS-VYRG> |  
 <GJJA-[VARD-GAL-VNS-MOTG #PRIEL-]+VARD-GAL-VNS-MOTG> |  
 <GJJA-[VARD-GAL-DGS-VYRG #PRIEL-]+VARD-GAL-DGS-VYRG> |  
 <GJJA-[VARD-GAL-DGS-MOTG #PRIEL-]+VARD-GAL-DGS-MOTG>;

<GJJA-PERGAL #PRIEL-PER+GAL> ::= <GJJA-PER-DAIKT-GAL-VNS-VYRG #PRIEL-PER+DAIKT-GAL-VNS-VYRG> |  
 <GJJA-PER-DAIKT-GAL-VNS-MOTG #PRIEL-PER+DAIKT-GAL-VNS-MOTG> |  
 <GJJA-PER-DAIKT-GAL-DGS-VYRG #PRIEL-PER+DAIKT-GAL-DGS-VYRG> |  
 <GJJA-PER-DAIKT-GAL-DGS-MOTG #PRIEL-PER+DAIKT-GAL-DGS-MOTG> |  
 <GJJA-PER-[VARD-GAL-VNS-VYRG #PRIEL-PER+[VARD-GAL-VNS-VYRG]> |  
 <GJJA-PER-[VARD-GAL-VNS-MOTG #PRIEL-PER+[VARD-GAL-VNS-MOTG]> |  
 <GJJA-PER-[VARD-GAL-DGS-VYRG #PRIEL-PER+[VARD-GAL-DGS-VYRG]> |  
 <GJJA-PER-[VARD-GAL-DGS-MOTG #PRIEL-PER+[VARD-GAL-DGS-MOTG]>;

<GJJA-PRIEŠGAL #PRIEL-PRIEŠ+GAL> ::= <GJJA-PRIEŠ-DAIKT-GAL-VNS-VYRG #PRIEL-PRIEŠ+DAIKT-GAL-VNS-VYRG> |  
 <GJJA-PRIEŠ-DAIKT-GAL-VNS-MOTG #PRIEL-PRIEŠ+DAIKT-GAL-VNS-MOTG> |  
 <GJJA-PRIEŠ-DAIKT-GAL-DGS-VYRG #PRIEL-PRIEŠ+DAIKT-GAL-DGS-VYRG> |  
 <GJJA-PRIEŠ-DAIKT-GAL-DGS-MOTG #PRIEL-PRIEŠ+DAIKT-GAL-DGS-MOTG> |  
 <GJJA-PRIEŠ-[VARD-GAL-VNS-VYRG #PRIEL-PRIEŠ+[VARD-GAL-VNS-VYRG]> |  
 <GJJA-PRIEŠ-[VARD-GAL-VNS-MOTG #PRIEL-PRIEŠ+[VARD-GAL-VNS-MOTG]> |  
 <GJJA-PRIEŠ-[VARD-GAL-DGS-VYRG #PRIEL-PRIEŠ+[VARD-GAL-DGS-VYRG]> |  
 <GJJA-PRIEŠ-[VARD-GAL-DGS-MOTG #PRIEL-PRIEŠ+[VARD-GAL-DGS-MOTG]>;

<GJJA-UŽGAL #PRIEL-UŽ+GAL> ::= <GJJA-UŽ-DAIKT-GAL-VNS-VYRG #PRIEL-UŽ+DAIKT-GAL-VNS-VYRG> |  
 <GJJA-UŽ-DAIKT-GAL-VNS-MOTG #PRIEL-UŽ+DAIKT-GAL-VNS-MOTG> |  
 <GJJA-UŽ-DAIKT-GAL-DGS-VYRG #PRIEL-UŽ+DAIKT-GAL-DGS-VYRG> |  
 <GJJA-UŽ-DAIKT-GAL-DGS-MOTG #PRIEL-UŽ+DAIKT-GAL-DGS-MOTG> |  
 <GJJA-UŽ-[VARD-GAL-VNS-VYRG #PRIEL-UŽ+[VARD-GAL-VNS-VYRG]> |  
 <GJJA-UŽ-[VARD-GAL-VNS-MOTG #PRIEL-UŽ+[VARD-GAL-VNS-MOTG]> |  
 <GJJA-UŽ-[VARD-GAL-DGS-VYRG #PRIEL-UŽ+[VARD-GAL-DGS-VYRG]> |  
 <GJJA-UŽ-[VARD-GAL-DGS-MOTG #PRIEL-UŽ+[VARD-GAL-DGS-MOTG]>;

<GJJA-SUINAG #PRIEL-SU+INAG> ::= <GJJA-SU-DAIKT-INAG-VNS-VYRG #PRIEL-SU+DAIKT-INAG-VNS-VYRG> |  
 <GJJA-SU-DAIKT-INAG-VNS-MOTG #PRIEL-SU+DAIKT-INAG-VNS-MOTG> |  
 <GJJA-SU-DAIKT-INAG-DGS-VYRG #PRIEL-SU+DAIKT-INAG-DGS-VYRG> |  
 <GJJA-SU-DAIKT-INAG-DGS-MOTG #PRIEL-SU+DAIKT-INAG-DGS-MOTG> |  
 <GJJA-SU-[VARD-INAG-VNS-VYRG #PRIEL-SU+[VARD-INAG-VNS-VYRG]> |  
 <GJJA-SU-[VARD-INAG-VNS-MOTG #PRIEL-SU+[VARD-INAG-VNS-MOTG]> |  
 <GJJA-SU-[VARD-INAG-DGS-VYRG #PRIEL-SU+[VARD-INAG-DGS-VYRG]> |  
 <GJJA-SU-[VARD-INAG-DGS-MOTG #PRIEL-SU+[VARD-INAG-DGS-MOTG]>;

<PAPILD-DAIKT-KILM-VNS-VYRG> ::= daikt\_kilm\_vns\_vyrg;  
 <PAPILD-DAIKT-KILM-VNS-MOTG> ::= daikt\_kilm\_vns\_motg;  
 <PAPILD-DAIKT-KILM-DGS-VYRG> ::= daikt\_kilm\_dgs\_vyrg;  
 <PAPILD-DAIKT-KILM-DGS-MOTG> ::= daikt\_kilm\_dgs\_motg;  
 <PAPILD-DAIKT-NAUD-VNS-VYRG> ::= daikt\_naud\_vns\_vyrg;  
 <PAPILD-DAIKT-NAUD-VNS-MOTG> ::= daikt\_naud\_vns\_motg;  
 <PAPILD-DAIKT-NAUD-DGS-VYRG> ::= daikt\_naud\_dgs\_vyrg;  
 <PAPILD-DAIKT-NAUD-DGS-MOTG> ::= daikt\_naud\_dgs\_motg;  
 <PAPILD-DAIKT-GAL-VNS-VYRG> ::= daikt\_gal\_vns\_vyrg;  
 <PAPILD-DAIKT-GAL-VNS-MOTG> ::= daikt\_gal\_vns\_motg;  
 <PAPILD-DAIKT-GAL-DGS-VYRG> ::= daikt\_gal\_dgs\_vyrg;  
 <PAPILD-DAIKT-GAL-DGS-MOTG> ::= daikt\_gal\_dgs\_motg;  
 <PAPILD-DAIKT-INAG-VNS-VYRG> ::= daikt\_inag\_vns\_vyrg;  
 <PAPILD-DAIKT-INAG-VNS-MOTG> ::= daikt\_inag\_vns\_motg;  
 <PAPILD-DAIKT-INAG-DGS-VYRG> ::= daikt\_inag\_dgs\_vyrg;  
 <PAPILD-DAIKT-INAG-DGS-MOTG> ::= daikt\_inag\_dgs\_motg;

<PAPILD-IVARD-KILM-VNS-VYRG> ::= ivard\_kilm\_vns\_vyrg;  
 <PAPILD-IVARD-KILM-VNS-MOTG> ::= ivard\_kilm\_vns\_motg;  
 <PAPILD-IVARD-KILM-DGS-VYRG> ::= ivard\_kilm\_dgs\_vyrg;  
 <PAPILD-IVARD-KILM-DGS-MOTG> ::= ivard\_kilm\_dgs\_motg;  
 <PAPILD-IVARD-NAUD-VNS-VYRG> ::= ivard\_naud\_vns\_vyrg;  
 <PAPILD-IVARD-NAUD-VNS-MOTG> ::= ivard\_naud\_vns\_motg;  
 <PAPILD-IVARD-NAUD-DGS-VYRG> ::= ivard\_naud\_dgs\_vyrg;  
 <PAPILD-IVARD-NAUD-DGS-MOTG> ::= ivard\_naud\_dgs\_motg;  
 <PAPILD-IVARD-GAL-VNS-VYRG> ::= ivard\_gal\_vns\_vyrg;  
 <PAPILD-IVARD-GAL-VNS-MOTG> ::= ivard\_gal\_vns\_motg;  
 <PAPILD-IVARD-GAL-DGS-VYRG> ::= ivard\_gal\_dgs\_vyrg;  
 <PAPILD-IVARD-GAL-DGS-MOTG> ::= ivard\_gal\_dgs\_motg;  
 <PAPILD-IVARD-INAG-VNS-VYRG> ::= ivard\_inag\_vns\_vyrg;  
 <PAPILD-IVARD-INAG-VNS-MOTG> ::= ivard\_inag\_vns\_motg;  
 <PAPILD-IVARD-INAG-DGS-VYRG> ::= ivard\_inag\_dgs\_vyrg;  
 <PAPILD-IVARD-INAG-DGS-MOTG> ::= ivard\_inag\_dgs\_motg;

<PAPILD-BENDR> ::= bendr;  
 <PAPILD-BENDR-KILM> ::= bendr\_kilm;  
 <PAPILD-BENDR-NAUD> ::= bendr\_naud;  
 <PAPILD-BENDR-GAL> ::= bendr\_gal;  
 <PAPILD-BENDR-INAG> ::= bendr\_inag;  
 <PAPILD-BENDR-VIET> ::= bendr\_viet;

### 5.1.5 Aplinkybės aprašas

Lietuvių kalbos gramatikoje sakoma, kad aplinkybės funkciją gali atlikti daiktavardžio, įvardžio, tam tikrais atvejais ir skaitvardžio linksnis ar linksnis su prielinksniu,rieveiksmis, veiksmažodžio bendratis, siekinys, lyginamasis posakis ir frazeologizmas [Ulvydas, 482]. Į BNF aprašą įtraukiami tik daiktavardžiai, įvardžiai,rieveiksmiai ir bendratis. Skaitvardžio ir siekinio atsisakoma dėl retos vartosenos. Pusedalyvinės, padalyvinės ir dalyvinės aplinkybės (o taip pat ir lyginamieji posakiai) savo sandara artimos sudėtinio sakinio šalutiniams dėmenims [Ulvydas, psl. 527]. Todėl jas numatoma apdoroti kartu su sudėtiniais sakiniais. Lyginamieji posakiai artimi šalutiniams sakiniais todėl juos numatoma apdoroti kartu su sudėtiniais sakiniais, o frazeologizmai, kaip ir visa idiomatika atidedama ateičiai, kai bus plečiama sintaksinė analizė.

Prielinksninės konstrukcijos, kurios neina papildiniu, priskiriamos aplinkybėms. Jei tas pats prielinksnis su linksniu (pvz., iš + kilmininkas) gali eiti ir papildiniu ir aplinkybe (*grįžti iš miško* – aplinkybė, *tyčiotis iš draugų* – papildinys), kuriai sakinio daliai priskirti prielinksninę konstrukciją, sprendžiama semantinių požymių pagalba. Prielinksninės konstrukcijos, kurios įtraukiamos į aplinkybės aprašą BNF pateiktos 5 lentelėje, psl. 71.

5 lentelė Prielinksninės konstrukcijos, galinčios eiti aplinkybe.

Eil. Nr.	Linksnis	Prielinksnis	Pavyzdžiai
1.	kilmininkas	Ant	kaboti ant vinies
2.	kilmininkas	Dėl	sielotis dėl nesėkmės
3.	kilmininkas	Iki	tęstis iki vakaro
4.	kilmininkas	Iš	grįžti iš miesto
5.	kilmininkas	Nuo	nukristi nuo medžio
6.	kilmininkas	Prie	pririšti prie tvoros
7.	kilmininkas	Tarp	stovėti tarp durų
8.	kilmininkas	Po	ateiti po valandos
9.	kilmininkas	Virš	pakilti virš debesų
10.	kilmininkas	Už	pasislėpti už namo
11.	galininkas	Apie	apeiti apie namą
12.	galininkas	Į	nueiti į mišką
13.	galininkas	Pagal	eiti pagal sieną
14.	galininkas	Per	peršokti per griovį
15.	galininkas	Prieš	plaukti prieš srovę
16.	galininkas	Pro	praeiti pro medį
17.	galininkas	Po	bėgioti po kiemą
18.	galininkas	Už	bausti už nusikaltimą
19.	įnagininkas	Po	palįsti po stalu
20.	įnagininkas	Su	susipykti su draugu
21.	įnagininkas	Ties	sustoti ties riba

Pagal reikšmę aplinkybės gramatikoje skirstomos į septynias rūšis [Ulvydas, 483]:

- 1) vietos,
- 2) laiko,
- 3) būdo,
- 4) kiekybės,
- 5) priežasties,
- 6) tikslo,
- 7) lyginamoji.

Kol nėra kompiuterizuota lietuvių kalbos semantinė (žodžių reikšmės) analizė, suskaidyti aplinkybių į tiek daug rūšių nėra galimybės. Todėl sakinio struktūroje jos, kaip ir kitos sakinio dalys, į rūšis neskaidomos. Vidiniame formate galima išskirti trijų rūšių aplinkybes, pagal tai, kokius semantinius požymius turi aplinkybėmis einantys žodžiai – laiko, vietos ar kiekio, nes būtent šie požymiai lemia kai kurių linksnų priskyrimą tam tikrai sakinio daliai t.y. papildiniui ar aplinkybei (ji *sveria vaiką* – papildinys, ji *sveria kilogramą* – kiekio aplinkybė). Pradžioje aplinkybė, kaip ir visos sintaksinės funkcijos skaidoma į kalbos dalis:

<APLINKYBĖ> ::= <APLINKYB-DAIKT> |  
 <APLINKYB-[IVARD-ASM]> |  
 <APLINKYB-BENDR> |  
 <APLINKYB-PRIEV> |  
 <APLINKYB-PIELKONS>;

Tipiškiausia aplinkybės forma yra prievėiksmis.

Bendratis gali būti naudojama tikslo aplinkybei, bet tik tai tuose sakiniuose, kur tariniu eina slinkties veiksmažodis (*pasiuntė atnešti, nuėjo maudytis*).

Tik asmeniniai įvardžiai gali eiti vietos aplinkybe ir jų ši sintaksinė funkcija geriau matyti tik iš konteksto (*Plačiai buvo išsikeroję ažuolo šakos. Paukščiai suko jose lizdus ir linksmi čiułbėdavo rytais*).

Prielinksninės konstrukcijos aprašomos BNF kaip žodžių junginiai ir pateikiamos skyriuje 5.2., psl. 73. Prielinksninės konstrukcijos priskyrimą papildiniui ar aplinkybei lemia veiksmažodis. Tais atvejais, kai ta pati prielinksninė konstrukcija (pvz., *į + galininkas*) gali eiti ir papildiniu, ir aplinkybe (*atsakyti į klausimą – papildinys, nuvažiuoti į Kauną – aplinkybė*), tik semantinių požymių pagalba galima nustatyti kuo ji eina nagrinėjamame sakinyje. Veiksmažodžio semantinių požymių lauke nurodomos tos prielinksninės konstrukcijos, kurios gali eiti papildiniu. Visos kitos sakinyje esančios prielinksninės konstrukcijos laikomos aplinkybėmis, jei jos nenustatytos kaip aplinkybiniai pažyminiai. Būtent todėl veiksmažodžio valdymas apdorojimo metu ir nagrinėjamas pats paskutinis.

Daiktavardžio vietininkas visada eina aplinkybe. Kitų linksnų ir prielinksninių konstrukcijų priskyrimą aplinkybėms lemia žodžio reikšmė, todėl BNF aplinkybės apraše nurodomi laiko, vietos, kiekio požymiai, tai reiškia, kad tik juos turintys žodžiai gali eiti aplinkybe. Jei veiksmažodis neturi semantinio požymio, kad sakinyje pavartoto linksnio reikalauja kaip papildinio ir daiktavardis neturi požymio, leidžiančio jam eiti aplinkybe, traktuojama, kad toks sakinyje yra negalimas (pvz., *\*nuvažiuoti į klausimą – nuvažiuoti nevaldo galininko su prielinksniumi į kaip papildinio ir žodis klausimą neturi vietos požymio*).

Išskaidytas pagal morfologines kategorijas aplinkybės aprašas atrodytų taip:

```
<APLINKYB-DAIKT> ::= <APLINKYB-DAIKT-VIET-VNS-VYRG> |
<APLINKYB-DAIKT-VIET-VNS-MOTG> |
<APLINKYB-DAIKT-VIET-DGS-VYRG> |
<APLINKYB-DAIKT-VIET-DGS-MOTG> |
<APLINKYB-DAIKT-NAUD-VNS-VYRG-LAIKP> |
<APLINKYB-DAIKT-NAUD-VNS-MOTG-LAIKP> |
<APLINKYB-DAIKT-NAUD-DGS-VYRG-LAIKP> |
<APLINKYB-DAIKT-NAUD-DGS-MOTG-LAIKP> |
<APLINKYB-DAIKT-GAL-VNS-VYRG-LAIKP> |
<APLINKYB-DAIKT-GAL-VNS-MOTG-LAIKP> |
<APLINKYB-DAIKT-GAL-DGS-VYRG-LAIKP> |
<APLINKYB-DAIKT-GAL-DGS-MOTG-LAIKP> |
<APLINKYB-DAIKT-GAL-VNS-VYRG-KIEKP> |
<APLINKYB-DAIKT-GAL-VNS-MOTG-KIEKP> |
<APLINKYB-DAIKT-GAL-DGS-VYRG-KIEKP> |
<APLINKYB-DAIKT-GAL-DGS-MOTG-KIEKP> |
<APLINKYB-DAIKT-ĪNAG-VNS-VYRG-LAIKP> |
<APLINKYB-DAIKT-ĪNAG-VNS-MOTG-LAIKP> |
<APLINKYB-DAIKT-ĪNAG-DGS-VYRG-LAIKP> |
<APLINKYB-DAIKT-ĪNAG-DGS-MOTG-LAIKP> |
<APLINKYB-DAIKT-ĪNAG-VNS-VYRG-VIETP> |
<APLINKYB-DAIKT-ĪNAG-VNS-MOTG-VIETP> |
<APLINKYB-DAIKT-ĪNAG-DGS-VYRG-VIETP> |
<APLINKYB-DAIKT-ĪNAG-DGS-MOTG-VIETP>;
```

```
<APLINKYB-ĪVARD-ASM> ::= <APLINKYB-ĪVARD-VIET-VNS-VYRG-ASM> |
<APLINKYB-ĪVARD-VIET-VNS-MOTG-ASM> |
<APLINKYB-ĪVARD-VIET-DGS-VYRG-ASM> |
<APLINKYB-ĪVARD-VIET-DGS-MOTG-ASM>;
```

```
<APLINKYB-PRIELKONS> ::= <GĪA#PRIEL-ANT+KILM> |
<GĪA#PRIEL-DĒL+KILM> |
<GĪA#PRIEL-ĪKĪ+KILM> |
<GĪA#PRIEL-ĪŠ+KILM> |
<GĪA#PRIEL-NUO+KILM> |
<GĪA#PRIEL-PRIE+KILM> |
<GĪA#PRIEL-TARP+KILM> |
<GĪA#PRIEL-PO+KILM> |
<GĪA#PRIEL-VIRŠ+KILM> |
<GĪA#PRIEL-UŽ+KILM> |
```



```

<GIJA#PRIEL-APIE+GAL> |
<GIJA#PRIEL-Į+GAL> |
<GIJA#PRIEL-PAGAL+GAL> |
<GIJA#PRIEL-PER+GAL> |
<GIJA#PRIEL-PO+GAL> |
<GIJA#PRIEL-PRIEŠ+GAL> |
<GIJA#PRIEL-PRO+GAL> |
<GIJA#PRIEL-PO+ĮNAG> |
<GIJA#PRIEL-TIES+ĮNAG>;

```

```

<APLINKYB-DAIKT-VIET-VNS-VYRG> ::= daikt_viet_vns_vyrg;
<APLINKYB-DAIKT-VIET-VNS-MOTG> ::= daikt_viet_vns_motg;
<APLINKYB-DAIKT-VIET-DGS-VYRG> ::= daikt_viet_dgs_vyrg;
<APLINKYB-DAIKT-VIET-DGS-MOTG> ::= daikt_viet_dgs_motg;
<APLINKYB-DAIKT-NAUD-VNS-VYRG-LAIKP> ::= daikt_naud_vns_vyrg_laikp;
<APLINKYB-DAIKT-NAUD-VNS-MOTG-LAIKP> ::= daikt_naud_vns_motg_laikp;
<APLINKYB-DAIKT-NAUD-DGS-VYRG-LAIKP> ::= daikt_naud_dgs_vyrg_laikp;
<APLINKYB-DAIKT-NAUD-DGS-MOTG-LAIKP> ::= daikt_naud_dgs_motg_laikp;
<APLINKYB-DAIKT-GAL-VNS-VYRG-LAIKP> ::= daikt_gal_vns_vyrg_laikp;
<APLINKYB-DAIKT-GAL-VNS-MOTG-LAIKP> ::= daikt_gal_vns_motg_laikp;
<APLINKYB-DAIKT-GAL-DGS-VYRG-LAIKP> ::= daikt_gal_dgs_vyrg_laikp;
<APLINKYB-DAIKT-GAL-DGS-MOTG-LAIKP> ::= daikt_gal_dgs_motg_laikp;
<APLINKYB-DAIKT-GAL-VNS-VYRG-LAIKP> ::= daikt_gal_vns_vyrg_kiekp;
<APLINKYB-DAIKT-GAL-VNS-MOTG-LAIKP> ::= daikt_gal_vns_motg_kiekp;
<APLINKYB-DAIKT-GAL-DGS-VYRG-LAIKP> ::= daikt_gal_dgs_vyrg_kiekp;
<APLINKYB-DAIKT-GAL-DGS-MOTG-LAIKP> ::= daikt_gal_dgs_motg_kiekp;
<APLINKYB-DAIKT-ĮNAG-VNS-VYRG-LAIKP> ::= daikt_inag_vns_vyrg_laikp;
<APLINKYB-DAIKT-ĮNAG-VNS-MOTG-LAIKP> ::= daikt_inag_vns_motg_laikp;
<APLINKYB-DAIKT-ĮNAG-DGS-VYRG-LAIKP> ::= daikt_inag_dgs_vyrg_laikp;
<APLINKYB-DAIKT-ĮNAG-DGS-MOTG-LAIKP> ::= daikt_inag_dgs_motg_laikp;
<APLINKYB-DAIKT-ĮNAG-VNS-VYRG-VIETP> ::= daikt_inag_vns_vyrg_vietp;
<APLINKYB-DAIKT-ĮNAG-VNS-MOTG-VIETP> ::= daikt_inag_vns_motg_vietp;
<APLINKYB-DAIKT-ĮNAG-DGS-VYRG-VIETP> ::= daikt_inag_dgs_vyrg_vietp;
<APLINKYB-DAIKT-ĮNAG-DGS-MOTG-VIETP> ::= daikt_inag_dgs_motg_vietp;

```

```

<APLINKYB-ĮVARD-VIET-VNS-VYRG-ASM> ::= įvard_viet_vns_vyrg_asm;
<APLINKYB-ĮVARD-VIET-VNS-MOTG-ASM> ::= įvard_viet_vns_motg_asm;
<APLINKYB-ĮVARD-VIET-DGS-VYRG-ASM> ::= įvard_viet_dgs_vyrg_asm;
<APLINKYB-ĮVARD-VIET-DGS-MOTG-ASM> ::= įvard_viet_dgs_motg_asm;

```

```

<APLINKYB-BENDR> ::= bendr;

```

```

<APLINKYB-PRIEV> ::= priev;

```

## 5.2 Grafo lankai

Antra BNF aprašo dalis sudaroma remiantis apibendrinta lietuvių kalbos sakinio sintaksine struktūra (žr. 33 pav., psl. 44). Kad būtų nuosekliai aprašyti visi galimi lietuvių kalboje sintaksiniai ryšiai (jie atitinka grafo lankus), kiekvienam iš viršūnės išeinančiam lankui sudaroma po atskirą formalios gramatikos taisyklę. Šios taisyklės turi aprašyti dviejų žodžių junginį, todėl neterminaliniam simboliui suteikiamas pavadinimas GIJA. Pradedant nuo veiksnio viršūnės sudaromos pirmosios gijos. Kadangi grafo viršūnių apraše veiksnys jau buvo išskaidytas pagal kalbos dalis (žr. psl. 54), gijos sudaromos atsižvelgiant į kiekvienai kalbos daliai būdingus sintaksinius ryšius.

```

<GIJA#VEIKSN-DAIKT+PAŽYMINYS>;
<GIJA#VEIKSN-DAIKT+TARINYS>;
<GIJA#VEIKSN-ĮVARD-DAIKT+TARINYS>;
<GIJA#VEIKSN-BENDR+PAPILDINYS>;
<GIJA#VEIKSN-BENDR+APLINKYBĖ>;
<GIJA#VEIKSN-BENDR+TARINYS>;

```

Įvardis pažyminių turėti negali, todėl BNF apraše ir nėra gijos, jungiančios įvardžiu išreikštą veiksnį su pažyminiu.

Kadangi pažyminy skaidomas į rūšis BNF apraše, todėl pradžioje veiksnio, išreikšto daiktavardžiu, aprašas išplečiamas pagal pažyminio rūšis:

```
<GJA#VEIKSN-DAIKT+PAŽYMINYS> ::=
    <GJA#VEIKSN-DAIKT+DERIN-PAŽYM> |
    <GJA#VEIKSN-DAIKT+NEDERIN-PAŽYM> |
    <GJA#VEIKSN-DAIKT+APLINKYB-PAŽYM> |
    <GJA#VEIKSN-DAIKT+PAPILD-PAŽYM> |
    <GJA#VEIKSN-DAIKT+TARIN-PAŽYM>;
```

Suskirsčius kiekvieną pažyminio rūšį pagal jai būdingas kalbos dalis, pvz.,

```
<GJA#VEIKSN-DAIKT+DERIN-PAŽYM> ::=
    <GJA#VEIKSN-DAIKT+DERIN-PAŽYM-BŪDV> |
    <GJA#VEIKSN-DAIKT+DERIN-PAŽYM-[IVARD-BŪDV]> |
    <GJA#VEIKSN-DAIKT+DERIN-PAŽYM-SKAITV> |
    <GJA#VEIKSN-DAIKT+DERIN-PAŽYM-DALYV>;
```

toliau BNF aprašas tęsiamas skaidant kalbos dalis pagal jų morfologines kategorijas:

```
<GJA#VEIKSN-DAIKT+DERIN-PAŽYM-BŪDV> ::=
    <GJA#VEIKSN-DAIKT-VARD-VNS-VYRG+DERIN-PAŽYM-BŪDV-VARD-VNS-VYRG> |
    <GJA#VEIKSN-DAIKT-VARD-VNS-MOTG+DERIN-PAŽYM-BŪDV-VARD-VNS-MOTG> |
    <GJA#VEIKSN-DAIKT-VARD-DGS-VYRG+DERIN-PAŽYM-BŪDV-VARD-DGS-VYRG> |
    <GJA#VEIKSN-DAIKT-VARD-DGS-MOTG+DERIN-PAŽYM-BŪDV-VARD-DGS-MOTG>;
```

Analogiškai aprašomos ir gijos, apimančios veiksnio ryšį su pažyminiaus, išreikštais įvardžiu, skaitvardžiu ar dalyviu.

GJA turi įvertinti laisvą žodžių tvarką lietuvių kalboje, t.y. susieti priklausomybių medį su linijiniu žodžių išsidėstymu sakinyje. Išskaidyto pagal morfologines kategorijas gijos aprašo BNF dešinėje taisyklės pusė turi tris pozicijas. Pirmoje ir trečioje – yra sakinio dalys, tarp kurių ieškomas sintaksinis ryšys, o viduryje – neterminalinis simbolis, kuris vadinamas intarpu tarp aprašomų sakinio dalių, pvz.:

```
<GJA#VEIKSN-DAIKT-VARD-VNS-VYRG+DERIN-PAŽYM-BŪDV-VARD-VNS-VYRG> ::=
    <DERIN-PAŽYM-BŪDV-VARD-VNS-VYRG>
    <INTARPAS#VEIKSN-DAIKT-VARD-VNS-VYRG+DERIN-PAŽYM-BŪDV-VARD-VNS-VYRG>
    <VEIKSN-DAIKT-VARD-VNS-VYRG>;
```

Dešinėje BNF taisyklės pusėje fiksuojama informacija apie žodžių išsidėstymą. Pačiame gijos neterminaliniame symbole žodžiai pateikiami pavaldumo tvarka, t.y. pradžioje nurodomas pagrindinis junginio žodis, o antroje vietoje pateikiamas jį išplečiantis ar priklausomas žodis. Jei išplečiantis žodis gali būti ir prieš ir po pagrindinio, tokiais atvejais tam pačiam žodžiui junginiui sudaromos dvi atskiros gijos. Formaliai gijoje žodžių tvarką atspindi pliuso ženklas arba žvaigždutė jungiantys žodžių aprašus. Pavartotas „+“ ženklas rodo, kad sakinyje išplečiantis žodis eina prieš pagrindinį junginio žodį. Priešingam atvejui naudojamas ženklas „\*“. Gija tarp veiksnio, išreikšto daiktavardžio vardininku, vienaskaitos vyriškąja gimine ir jo derinamojo pažyminio, išreikšto būdvardžio vardininku, vienaskaitos vyriškąja gimine, esant inversinei žodžių tvarkai, atrodytų taip:

```
<GJA#VEIKSN-DAIKT-VARD-VNS-VYRG*DERIN-PAŽYM-BŪDV-VARD-VNS-VYRG> ::=
    <VEIKSN-DAIKT-VARD-VNS-VYRG>
    <INTARPAS#VEIKSN-DAIKT-VARD-VNS-VYRG*DERIN-PAŽYM-BŪDV-VARD-VNS-VYRG>
    <DERIN-PAŽYM-BŪDV-VARD-VNS-VYRG>;
```

Prielinksninės konstrukcijos taip pat aprašomos, kaip GJA. P.vz., prielinksninė konstrukcija, sudaryta iš prielinksnio *po* ir įnagininko linksnio BNF apraše atrodytų taip:

```
<GJA#PRIEL-PO+[NAG]> ::=
    <GJA#PRIEL-PO+DAIKT-[NAG-VNS-VYRG]> |
    <GJA#PRIEL-PO+DAIKT-[NAG-VNS-MOT]> |
    <GJA#PRIEL-PO+DAIKT-[NAG-DGS-VYRG]> |
    <GJA#PRIEL-PO+DAIKT-[NAG-DGS-MOT]> |
    <GJA#PRIEL-PO+[IVARD-[NAG-VNS-VYRG-DAIKT]> |
    <GJA#PRIEL-PO+[IVARD-[NAG-VNS-MOT-DAIKT]> |
    <GJA#PRIEL-PO+[IVARD-[NAG-DGS-VYRG-DAIKT]> |
    <GJA#PRIEL-PO+[IVARD-[NAG-DGS-MOT-DAIKT]>;
```

Prielinksnis visada eina prieš linksnį, todėl taisyklė, įvertinanti inversinę žodžių tvarką, šiuo atveju nerašoma.

Įvardis negali būti niekuo išplečiamas, todėl jo prielinksninės konstrukcijos aprašas atrodytų taip:

<GIIA#PRIEL-PO+[IVARD-**INAG**-VNS-VYRG-DAIKT]> ::= <PRIEL-PO> <[IVARD-**INAG**-VNS-VYRG-DAIKT]>;

Kiekvienai daiktavardžio alternatyvai rašoma BNF taisyklė, nurodanti, kas gali būti įsiterpę tarp daiktavardžio ir prielinksnio.

<GIIA#PRIEL-PO+DAIKT-**INAG**-VNS-VYRG> ::= <PRIEL-PO>  
<INTARPAS# PRIEL-PO+DAIKT-**INAG**-VNS-VYRG>  
<DAIKT-**INAG**-VNS-VYRG>;

Jei įsiterpti gali keli tą pačią sintaksinę funkciją atliekantys žodžiai, intarpo apraše nurodomas pasikartojimo simbolis – riestiniai skliaustai (plačiau apie simbolių reikšmę BNF žr. psl. 7). Derinamųjų pažyminių, išreikštų būdvardžiu ar dalyviu gali būti keli, todėl šios sakinio dalys pateikiamos riestiniuose skliaustuose.

<INTARPAS# PRIEL-PO+DAIKT-**INAG**-VNS-VYRG> ::=  
 [{<DERIN-PAŽYM-BŪDV-**INAG**-VNS-VYRG>}]  
 [{<DERIN-PAŽYM-DALYV-**INAG**-VNS-VYRG>}]  
 [<DERIN-PAŽYM-[IVARD-**INAG**-VNS-VYRG-BŪDV]>]  
 [<DERIN-PAŽYM-SKAITV-**INAG**-VNS-VYRG-BŪDV>]  
 [<NEDERIN-PAŽYM-DAIKT> | <NEDERIN-PAŽYM-[IVARD-DAIKT]>]  
 [<GIIA#NEDERIN-PAŽYM-DAIKT+NEDERIN-PAŽYM-DAIKT>]  
 [<GIIA#NEDERIN-PAŽYM-DAIKT+DERIN-PAŽYM-BŪDV>]  
 [<GIIA#NEDERIN-PAŽYM-DAIKT+DERIN-PAŽYM-DALYV>]  
 [<GIIA#NEDERIN-PAŽYM-DAIKT+DERIN-PAŽYM-[IVARD-BŪDV]>]  
 [<GIIA#NEDERIN-PAŽYM-DAIKT+DERIN-PAŽYM-SKAITV-BŪDV>]  
 [<GIIA# DERIN-PAŽYM-BŪDV-**INAG**-VNS-VYRG+APLINKYB-PRIEV>]  
 [<GIIA# DERIN-PAŽYM-DALYV-**INAG**-VNS-VYRG+APLINKYB-PRIEV>]  
 [<GIIA# DERIN-PAŽYM-DALYV-**INAG**-VNS-VYRG+APLINKYB-DAIKT>]  
 [<GIIA# DERIN-PAŽYM-DALYV-**INAG**-VNS-VYRG+APLINKYB-PRIELKONS>]  
 [<GIIA# DERIN-PAŽYM-DALYV-**INAG**-VNS-VYRG+PAPILD-DAIKT>]  
 [<GIIA# DERIN-PAŽYM-DALYV-**INAG**-VNS-VYRG+PAPILD-[IVARD-DAIKT]>]  
 [<GIIA# DERIN-PAŽYM-DALYV-**INAG**-VNS-VYRG+PAPILD-PRIELKONS>];

Čia reiktų pateikti ir sudėtinio tarinio aprašą. Kurie sudėtinio tarinio atvejai įtraukiami į BNF aprašą detaliau pateikta skyriuje 5.1.2., psl. 54. Sudėtinis suvestinis tarinys – tai gija, kurią sudaro modalinį požymį turintis veiksmazodis ir bendratis, o sudėtinis sudurtinis tarinys apima jungtį ir vardažodį. Modalinio veiksmazodžio požymis suteikiamas žodžiams, kurių leksinės reikšmės nepakanka sakinio minčiai išsakyti (*galiu eiti*, *turi dirbti*, *pradėjo skaityti*). Jungtį sudaro žodžio *būti* ar kito egzistavimo reikšmės veiksmazodžio formos. Taigi, pradžioje sudėtinis tarinys išskaidomas į dvi dalis:

<SUDĖT-TARIN> ::= <SUDĖT-SUVEST-TARIN> | <SUDĖT-SUDURT-TARIN>;

Toliau jau pateikiamas kiekvienos tarinio rūšies aprašas kaip gijos:

<SUDĖT-SUVEST-TARIN> ::= <GIIA#VEIKSM-VNS-1A-MODALV+BENDR> |  
 <GIIA#VEIKSM-VNS-2A-MODALV+BENDR> |  
 <GIIA#VEIKSM-VNS-3A-MODALV+BENDR> |  
 <GIIA#VEIKSM-DGS-1A-MODALV+BENDR> |  
 <GIIA#VEIKSM-DGS-2A-MODALV+BENDR> |  
 <GIIA#VEIKSM-DGS-3A-MODALV+BENDR> |  
 <GIIA#VEIKSM-VNS-1A-MODALV+BENDR-**KILM**> |  
 <GIIA#VEIKSM-VNS-2A-MODALV+BENDR-**KILM**> |  
 <GIIA#VEIKSM-VNS-3A-MODALV+BENDR-**KILM**> |  
 <GIIA#VEIKSM-DGS-1A-MODALV+BENDR-**KILM**> |  
 <GIIA#VEIKSM-DGS-2A-MODALV+BENDR-**KILM**> |  
 <GIIA#VEIKSM-DGS-3A-MODALV+BENDR-**KILM**> |  
 <GIIA#VEIKSM-VNS-1A-MODALV+BENDR-**NAUD**> |  
 <GIIA#VEIKSM-VNS-2A-MODALV+BENDR-**NAUD**> |  
 <GIIA#VEIKSM-VNS-3A-MODALV+BENDR-**NAUD**> |  
 <GIIA#VEIKSM-DGS-1A-MODALV+BENDR-**NAUD**> |  
 <GIIA#VEIKSM-DGS-2A-MODALV+BENDR-**NAUD**> |  
 <GIIA#VEIKSM-DGS-3A-MODALV+BENDR-**NAUD**> |

<GIJA#VEIKSM-VNS-1A-MODALV+BENDR-GAL> |  
 <GIJA#VEIKSM-VNS-2A-MODALV+BENDR-GAL> |  
 <GIJA#VEIKSM-VNS-3A-MODALV+BENDR-GAL> |  
 <GIJA#VEIKSM-DGS-1A-MODALV+BENDR-GAL> |  
 <GIJA#VEIKSM-DGS-2A-MODALV+BENDR-GAL> |  
 <GIJA#VEIKSM-DGS-3A-MODALV+BENDR-GAL> |  
 <GIJA#VEIKSM-VNS-1A-MODALV+BENDR-~~NAG~~> |  
 <GIJA#VEIKSM-VNS-2A-MODALV+BENDR-~~NAG~~> |  
 <GIJA#VEIKSM-VNS-3A-MODALV+BENDR-~~NAG~~> |  
 <GIJA#VEIKSM-DGS-1A-MODALV+BENDR-~~NAG~~> |  
 <GIJA#VEIKSM-DGS-2A-MODALV+BENDR-~~NAG~~> |  
 <GIJA#VEIKSM-DGS-3A-MODALV+BENDR-~~NAG~~> |  
 <GIJA#VEIKSM-VNS-1A-MODALV+BENDR-VIET> |  
 <GIJA#VEIKSM-VNS-2A-MODALV+BENDR-VIET> |  
 <GIJA#VEIKSM-VNS-3A-MODALV+BENDR-VIET> |  
 <GIJA#VEIKSM-DGS-1A-MODALV+BENDR-VIET> |  
 <GIJA#VEIKSM-DGS-2A-MODALV+BENDR-VIET> |  
 <GIJA#VEIKSM-DGS-3A-MODALV+BENDR-VIET>;

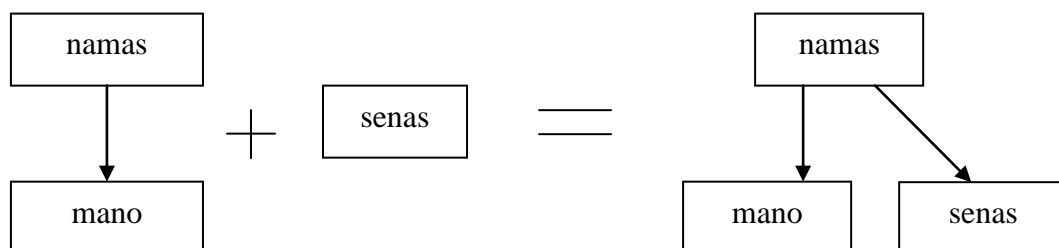
<SUDĒT-SUDURT-TARIN> ::= <GIJA#VEIKSM-VNS-1A-JUNGT+DAIKT> |  
 <GIJA#VEIKSM-VNS-2A-JUNGT+DAIKT> |  
 <GIJA#VEIKSM-VNS-3A-JUNGT+DAIKT> |  
 <GIJA#VEIKSM-DGS-1A-JUNGT+DAIKT> |  
 <GIJA#VEIKSM-DGS-2A-JUNGT+DAIKT> |  
 <GIJA#VEIKSM-DGS-3A-JUNGT+DAIKT> |  
 <GIJA#VEIKSM-VNS-1A-JUNGT+BŪDV> |  
 <GIJA#VEIKSM-VNS-2A-JUNGT+BŪDV> |  
 <GIJA#VEIKSM-VNS-3A-JUNGT+BŪDV> |  
 <GIJA#VEIKSM-DGS-1A-JUNGT+BŪDV> |  
 <GIJA#VEIKSM-DGS-2A-JUNGT+BŪDV> |  
 <GIJA#VEIKSM-DGS-3A-JUNGT+BŪDV> |  
 <GIJA#VEIKSM-VNS-1A-JUNGT+DALYV> |  
 <GIJA#VEIKSM-VNS-2A-JUNGT+DALYV> |  
 <GIJA#VEIKSM-VNS-3A-JUNGT+DALYV> |  
 <GIJA#VEIKSM-DGS-1A-JUNGT+DALYV> |  
 <GIJA#VEIKSM-DGS-2A-JUNGT+DALYV> |  
 <GIJA#VEIKSM-DGS-3A-JUNGT+DALYV> |  
 <GIJA#VEIKSM-VNS-1A-JUNGT+SKAITV> |  
 <GIJA#VEIKSM-VNS-2A-JUNGT+SKAITV> |  
 <GIJA#VEIKSM-VNS-3A-JUNGT+SKAITV> |  
 <GIJA#VEIKSM-DGS-1A-JUNGT+SKAITV> |  
 <GIJA#VEIKSM-DGS-2A-JUNGT+SKAITV> |  
 <GIJA#VEIKSM-DGS-3A-JUNGT+SKAITV> |  
 <GIJA#VEIKSM-VNS-1A-JUNGT+PRIEV>;

Toliau kiekviena alternatīva skaidoma pagal tai kalbos daļai būdingas morfoloģines kategorijas, pvz.:

<GIJA#VEIKSM-VNS-1A-JUNGT+DAIKT> ::= <GIJA#VEIKSM-VNS-1A-JUNGT+DAIKT-VARD-VNS-VYRG> |  
 <GIJA#VEIKSM-VNS-1A-JUNGT+DAIKT-VARD-VNS-MOTG> |  
 <GIJA#VEIKSM-VNS-1A-JUNGT+DAIKT-VARD-DGS-VYRG> |  
 <GIJA#VEIKSM-VNS-1A-JUNGT+DAIKT-VARD-DGS-MOTG> |  
 <GIJA#VEIKSM-VNS-2A-JUNGT+DAIKT-VARD-VNS-VYRG> |  
 <GIJA#VEIKSM-VNS-2A-JUNGT+DAIKT-VARD-VNS-MOTG> |  
 <GIJA#VEIKSM-VNS-2A-JUNGT+DAIKT-VARD-DGS-VYRG> |  
 <GIJA#VEIKSM-VNS-2A-JUNGT+DAIKT-VARD-DGS-MOTG> |  
 <GIJA#VEIKSM-VNS-3A-JUNGT+DAIKT-VARD-VNS-VYRG> |  
 <GIJA#VEIKSM-VNS-3A-JUNGT+DAIKT-VARD-VNS-MOTG> |  
 <GIJA#VEIKSM-VNS-3A-JUNGT+DAIKT-VARD-DGS-VYRG> |  
 <GIJA#VEIKSM-VNS-3A-JUNGT+DAIKT-VARD-DGS-MOTG> |  
 <GIJA#VEIKSM-VNS-1A-JUNGT+DAIKT-VARD-VNS-VYRG> |  
 <GIJA#VEIKSM-VNS-1A-JUNGT+DAIKT-VARD-VNS-MOTG> |  
 <GIJA#VEIKSM-VNS-1A-JUNGT+DAIKT-VARD-DGS-VYRG> |  
 <GIJA#VEIKSM-VNS-1A-JUNGT+DAIKT-VARD-DGS-MOTG> |  
 <GIJA#VEIKSM-VNS-2A-JUNGT+DAIKT-VARD-VNS-VYRG> |  
 <GIJA#VEIKSM-VNS-2A-JUNGT+DAIKT-VARD-VNS-MOTG> |  
 <GIJA#VEIKSM-VNS-2A-JUNGT+DAIKT-VARD-DGS-VYRG> |  
 <GIJA#VEIKSM-VNS-2A-JUNGT+DAIKT-VARD-DGS-MOTG> |  
 <GIJA#VEIKSM-VNS-3A-JUNGT+DAIKT-VARD-VNS-VYRG> |  
 <GIJA#VEIKSM-VNS-3A-JUNGT+DAIKT-VARD-VNS-MOTG> |  
 <GIJA#VEIKSM-VNS-3A-JUNGT+DAIKT-VARD-DGS-VYRG> |  
 <GIJA#VEIKSM-VNS-3A-JUNGT+DAIKT-VARD-DGS-MOTG>;

### 5.2.1 Laisvos žodžių tvarkos įvertinimas

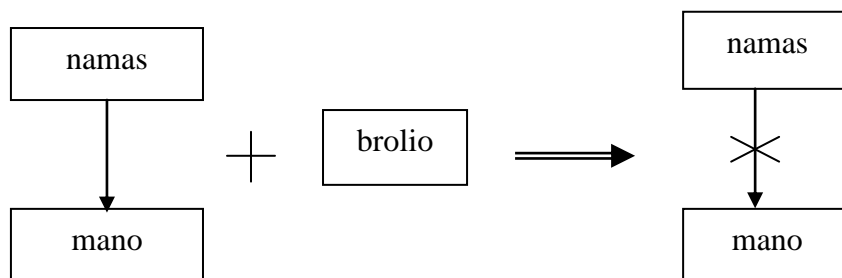
Intarpas turėtų įvertinti laisvą žodžių tvarką lietuvių kalbos sakiniuose, t.y. jis turėtų nurodyti, kokios kitos sakinio dalys gali būti įsiterpusios tarp dviejų tiesioginiu sintaksiniu ryšiu susietų žodžių. Žodžių tvarka lietuvių kalboje yra visiškai laisva tik sakinyje. Žodžių junginiuose jau galima rasti ją reglamentuojančių dalykų, kurie lituanistų darbuose gal ir nebuvo detaliau aptarti. Pvz., tarp veiksnio ir jo nederinamojo pažyminio negali būti įsiterpęs kitas nederinamasis pažyminys, nes jis suardo tą junginį: tarp žodžių *mano namas* gali būti įsiterpęs tik derinamasis pažyminys (pvz., *senas*), kuris nepadaro įtakos tam junginiui: *mano senas namas* lieka vis tiek *mano namas* (44 pav.):



44 pav. Derinamojo pažyminio įsiterpimas į žodžių junginį *mano namas*.

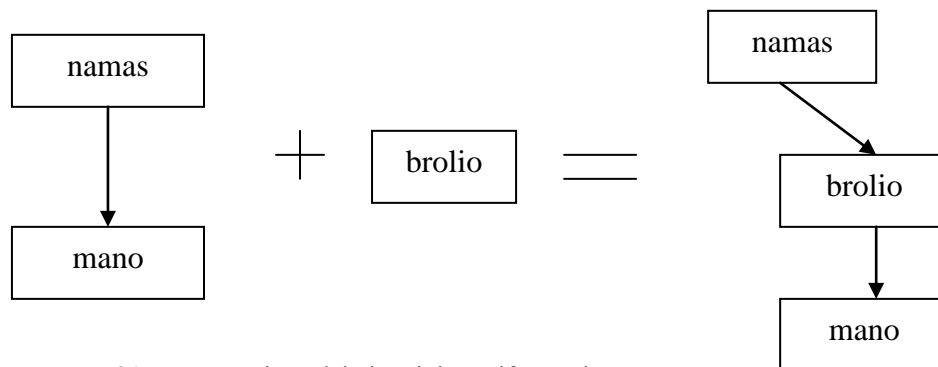
Atsiradęs naujas junginys *senas namas* nesuardo pirmojo ir sakinyje atsistoja šalia jo, t.y. vietoj vieno junginio *mano namas* atsiranda du junginiai *mano namas* ir *senas namas*, vadinasi, pradinis žodžių junginys lieka nepakitęs, tik prie jo prisideda dar vienas papildomas.

Tačiau, jei tarp žodžių *mano namas* įsiterpia žodis *brolio* (nederinamasis pažyminys), pradinio žodžių junginio nebelieka: *mano brolio namas* – tai jau nebe *mano namas* (45 pav.):



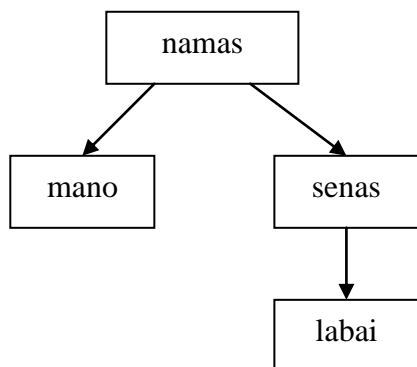
45 pav. Nederinamojo pažyminio įsiterpimas į žodžių junginį *mano namas*.

Įsiterpus žodžiui *brolio* atsiranda du visiškai nauji žodžių junginiai: *mano brolio* ir *brolio namas* (46 pav.):



46 pav. Naujų žodžių junginių susiformavimas.

Vadinasi BNF apraše ir turi būti nurodyta, kad veiksnio ir jo nederinamojo pažyminio intarpas negali būti nederinamasis pažyminys. Šis intarpas gali būti tik derinamasis pažyminys ar jo gija, t.y. derinamasis pažyminys su jį išplečiančiais žodžiais, pvz., *mano labai senas namas* (47 pav.):



47 pav. Intarpas, išreikštas derinamojo pažyminio gija.

BNF apraše ši informacija turėtų atsispindėti taip:

```

<INTARPAS#VEIKSN-DAIKT-VARD-VNS-VYRG+NEDERIN-PAŽYM-IVARD-KILM-VNS-VYRG> ::=
  <DERIN-PAŽYM-BŪDV-VARD-VNS-VYRG> |
  <DERIN-PAŽYM-SKAITV-VARD-VNS-VYRG> |
  <DERIN-PAŽYM-DALYV-VARD-VNS-VYRG> |
  <GIJA#DERIN-PAŽYM-BŪDV-VARD-VNS-VYRG+APLINKYB-PRIEV> |
  <GIJA#DERIN-PAŽYM-DALYV-VARD-VNS-VYRG+APLINKYB-PRIEV>;
  
```

Tai reiškia, kad intarpas tarp veiksnio vardininko ir jo nederinamojo pažyminio, kuris išreiškiamas tik kilmininko linksniu, gali būti įsiterpę trijų rūšių derinamieji pažyminiai, kuriais eina atitinkamų morfologinių kategorijų būdvardis, dalyvis ar skaitvardis, taip pat gijos, sudarytos iš būdvardžio ar dalyvio su juos pažyminčiomis aplinkybėmis.

Pačiame intarpo neterminaliniam simboliuje nurodoma ta pati informacija, kuri būna pateikta jį apimančioje gijoje, t.y. tiesioginiu sintaksiniu ryšiu susieti žodžiai su jų morfologinėmis kategorijomis.

### 5.2.2 Intarpo vaidmuo formalioje gramatikoje

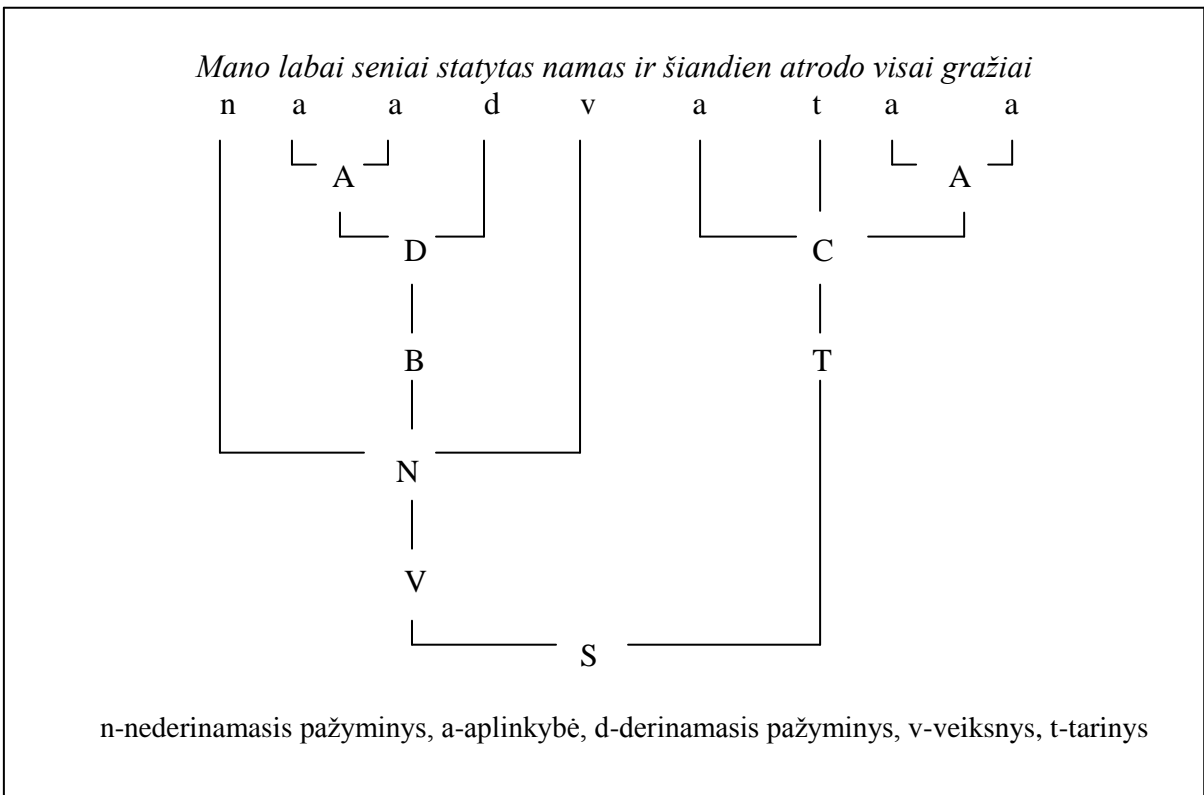
Sakinio sintaksinė analizė – tai yra iš esmės žodžių junginių ieškojimas. Galutinis rezultatas turėtų būti tiesioginiu sintaksiniu ryšiu susietų žodžių poros, nurodant abiejų žodžių sintaksines funkcijas. Kiekviena tokia pora ir aprašoma BNF parametru GIJA. Turint informaciją apie žodžių tarpusavio ryšius jau galima formuoti sakinio sintaksinės struktūros grafinį vaizdą, t.y. priklausomybių medį ar grafą.

Kaip vyks analizė kompiuteryje, galima pailustruoti pavyzdžiu. Kad būtų akivaizdžiau ir suprantamiau, parašomas nesudėtingas BNF taisyklių rinkinys (48 pav., psl. 79), su kuriuo tegalima išnagrinėti tik nedidelę sakinių grupę. Pradiniu simboliu laikomas simbolis *S* (apie simbolių reikšmę BNF plačiau žr. psl. 7) reiškia visą sakinį. Jis skaidomas į veiksnio grupę *V* ir tarinio grupę *T*. Kartais sakinį gali sudaryti vien veiksnio ar vien tarinio grupė.

Veiksnių grupė išreiškiama veiksniu  $v$  arba veiksniu su jį išplečiančiais žodžiais, t.y. gija  $N$ , jungiančia veiksnių  $v$  su jo nederinamuoju pažyminiu  $n$ . Ši gija ( $N$ ) gali susidėti iš nederinamojo pažyminio  $n$ , po kurio seka veiksnių  $v$ , arba tarp jų dar gali būti įsiterpęs intarpas  $B$ . Tas intarpas – tai derinamasis pažyminys  $d$  arba derinamojo pažyminio gija  $D$ , kuri išreiškiama derinamuoju pažyminiu  $d$  su prieš jį einančia aplinkybe  $a$ , ar aplinkybės gija  $A$ . Aplinkybės gija  $A$  tegali būti tik dvi, viena po kitos einančios aplinkybės  $a$ . Tarinio grupę  $T$  gali sudaryti tarinys  $t$  arba gija  $C$  siejanti tarinį  $t$  su aplinkybe  $a$ , ar aplinkybės gija  $A$ . Su šiuo taisyklių rinkiniu galima išnagrinėti, pvz., sakinį *Mano labai seniai statytas namas ir šiandien atrodo visai gražiai* (49 pav.):

$S ::= V[T] \mid T[V];$	* V – veiksnio grupė, T – tarinio grupė *
$V ::= v \mid N;$	* N – GIJA veiksnių + nederinamasis pažyminys *
$N ::= nv \mid nBv;$	* B – INTARPAS tarp veiksnio ir nederinamojo pažyminio *
$B ::= d \mid D;$	* D – GIJA derinamasis pažyminys + aplinkybė *
$D ::= ad \mid Ad;$	* A – GIJA aplinkybė + aplinkybė *
$A ::= aa;$	* C – GIJA tarinys + aplinkybė *
$T ::= t \mid C;$	
$C ::= at \mid At \mid ta \mid tA \mid AtA;$	

48 pav. BNF taisyklių rinkinio pavyzdys.



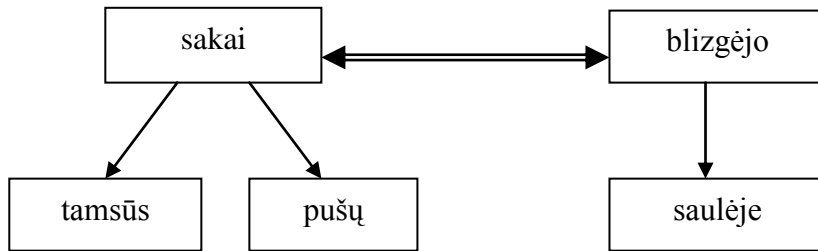
49 pav. Sintaksiškai išnagrinėto sakinio pavyzdys.

### 5.2.3 Morfologinio daugiareikšmiškumo panaikinimas

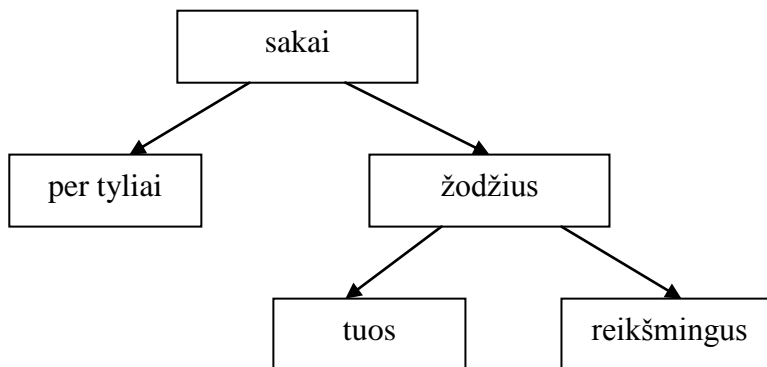
Pasirinkus morfologiniu požiūriu daugiareikšmi žodį, pvz., *sakai* galima stebėti, kaip sintaksinės analizės metu kartais gali būti panaikinamas jo daugiareikšmiškumas. Kiekvienam variantui paimama po sakinį.

*Tamsūs pušų sakai blizgėjo saulėje.*  
*Per tyliai sakai tuos reikšmingus žodžius.*

Šių abiejų sakinių sintaksinės struktūros atrodytų taip (50 pav., ir 51 pav.):

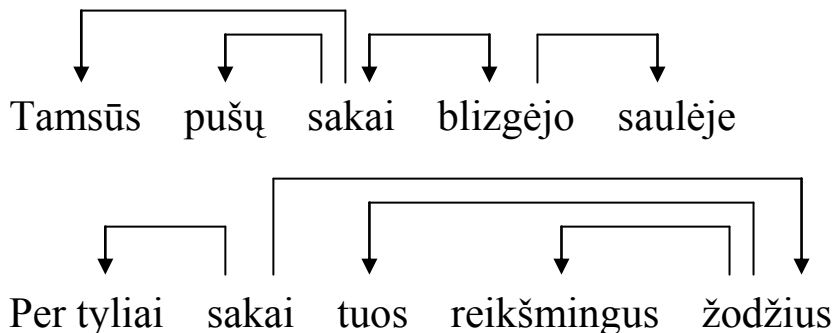


50 pav. Sakinio *Tamsūs pušų sakai blizgėjo saulėje* sintaksinė struktūra.



51 pav. Sakinio *Per tyliai sakai tuos reikšmingus žodžius* sintaksinė struktūra.

Lankus tarp grafo viršūnių, t.y. sintaksinius ryšius tarp žodžių, galima parodyti ir linijinėje sakinio struktūroje, t.y. pačiame sakinyje, kuri mes matome parašytą (52 pav.). Vokiečių autoriai teigia, kad „...*sakinys tai yra suplota ir ištempta į vielą struktūrinė schema*“ [Agel, psl. 34].

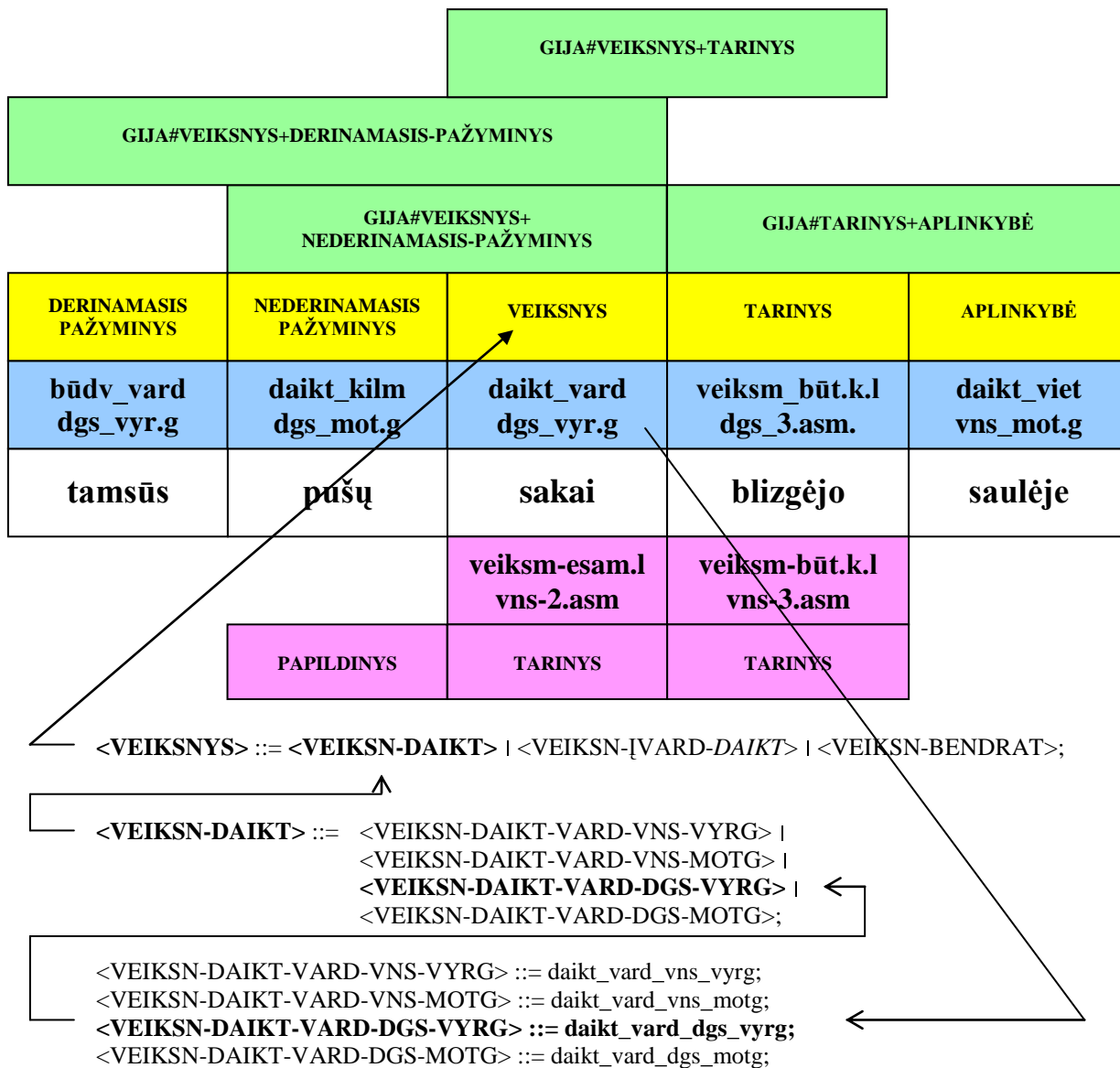


52 pav. Sintaksiniai ryšiai tarp žodžių, pavaizduoti linijinėje sakinio struktūroje.

Formalus parametras GIJA iš esmės ir atitinka šias rodykles virš sakinio žodžių.



Sintaksinė analizė prasideda nuo to, kad kiekvienam sakinio žodžiui nurodoma morfologinė informacija apie jį (pirma eilutė virš sakinio, 53 pav.). Tai atliks Vytauto Zinkevičiaus sukurta lietuvių kalbos morfologinių modelių programinė įranga. Toliau galima parodyti kelią, kaip iš morfologinių kategorijų gaunamos sintaksinės. Funkcijos priskyrimas žodžiui vyksta pradėdant nuo apačios, t.y. nuo morfologinių kategorijų (nuo terminalinių simbolių BNF apraše). Veiksnyš šiame sakinyje nustatomas taip (53 pav.):



**53 pav.** Veiksnyš suradimo kelias sakinyje *Tamsūs pušų sakai blizgėjo saulėje*.

GIJA tarp veiksnio ir jo derinamojo pažyminio apima schemoje trijų žodžių pozicijas. Šiuo atveju tarpas yra nederinamasis pažyminys. Ši informacija atsispindi BNF apraše: tarp veiksnio ir jo derinamojo pažyminio gali būti nederinamasis pažyminys (apatinė eilutė, ji pabraukta).

$\langle \text{INTARPAS\#VEIKSN-DAIKT-VARD-DGS-VYRG+DERIN-PAŽYM-BŪDV-VARD-DGS-VYRG} \rangle ::=$   
 $\langle \text{DERIN-PAŽYM-BŪDV-VARD-DGS-VYRG} \rangle \mid$   
 $\langle \text{DERIN-PAŽYM-DALYV-VARD-DGS-VYRG} \rangle \mid$   
 $\langle \text{NEDERIN-PAŽYM-DAIKT-KILM-VNS-VYRG} \rangle \mid$   
 $\langle \text{NEDERIN-PAŽYM-DAIKT-KILM-VNS-MOTG} \rangle \mid$   
 $\langle \text{NEDERIN-PAŽYM-DAIKT-KILM-DGS-VYRG} \rangle \mid$   
 $\langle \text{NEDERIN-PAŽYM-DAIKT-KILM-DGS-MOTG} \rangle;$

Po sakinio žodžiais *pušų, sakai* ir *blizgėjo* pateiktos jų sintaksinės alternatyvos (53 pav., psl. 81) atmetamos, nes jos šiame sakinyje nesudaro junginių (gijų). Veiksmažodžio *blizgėjo* vienaskaitos trečiam asmeniui sakinyje nėra veiksnio, išreikšto vienaskaitos vardininku, veiksmažodžiui *sakai* lieka nerealizuotas jo valentingumas (*sakyti* reikalauja galininko, kurio sakinyje nėra), o žodis *pušų* negali eiti papildiniu, nes sakinyje esantis tarinys (*blizgėjo*) nereikalauja kilmininko linksnio, t.y. tariniu einantis veiksmažodis neturi semantinio požymio, kad jį būtina išplėsti kilmininku.

Antro sakinio analizė atrodytų taip (54 pav.):

GIJA#TARINYS+PAPILDINYS				
		GIJA#PAPILDINYS+DERINAMASIS-PAŽYMINYS		
GIJA#TARINYS+APLINKYBĖ			GIJA#PAPILDINYS+PAPILDINIO- DERINAMASIS-PAŽYMINYS	
APLINKYBĖ	TARINYS	DERINAMASIS PAŽYMINYS	DERINAMASIS PAŽYMINYS	PAPILDINYS
prieveiksm	veiksm-esam.1 vns-2.asm	įvard-gal dgs-vyr.g	būdv-gal dgs-vyr.g	daikt-gal dgs-vyr.g
per tyliai	sakai	tuos	reikšmingus	žodžius
daikt-vard dgs-vyrg				
VEIKSNYS				

54 pav. Sakinio *Per tyliai sakai tuos reikšmingus žodžius* sintaksinė struktūra.

Šiame sakinyje žodis *sakai* negali būti daiktavardis, nes nėra tarinio, išreikšto veiksmažodžio trečiuoju asmeniu. Įvertinant tarinio su jungtimi atvejį, kur jungtis galėtų būti praleista, čia nėra tarinio vardinės dalies, išreikštos daugiskaitos vardininku.

Šiuose brėžiniuose gijos parodomos panašiai kaip ir frazės, analizuojant sakinį frazių metodu. *Esminis skirtumas tarp šių atvaizdavimo būdų yra tai, kad frazių gramatikoje frazę žymintis keturkampis gaunamas iš po ja stovinčių žodžių (plačiau žr. psl. 18). Gija rodo ryšį tarp dviejų kraštinių jos apimtų žodžių; visi likusieji jos viduje esantys žodžiai traktuojami kaip intarpas ir gijos aprašytam žodžių junginiui nepriklauso. Tik tokiu būdu galima įvertinti laisvą žodžių tvarką sakinyje ir parodyti tiesioginį sintaksinį ryšį tarp žodžių, kurie sakinyje nėra šalia vienas kito.*

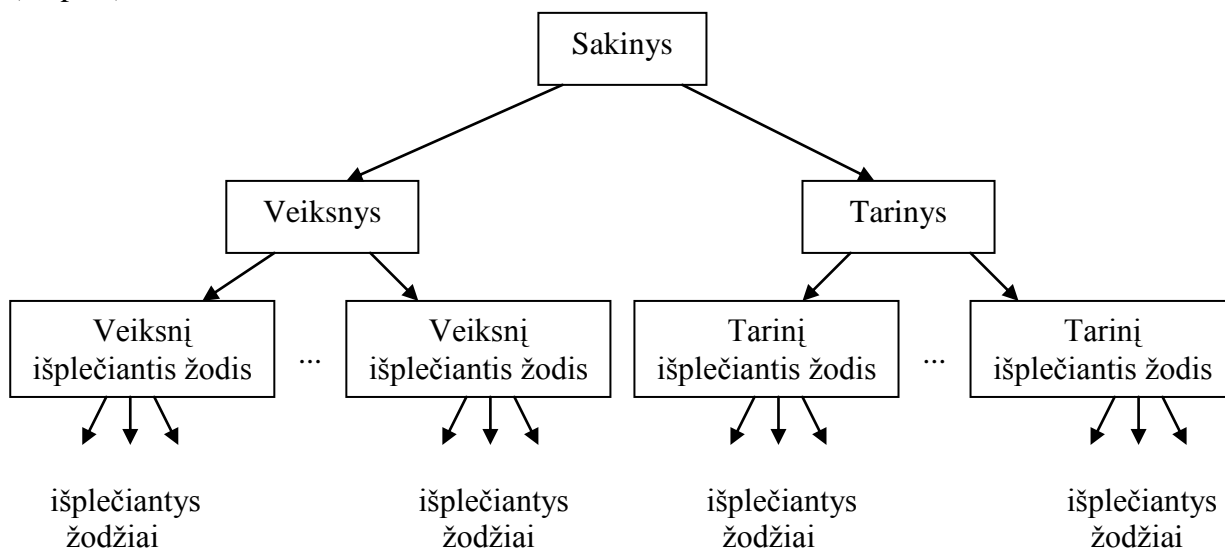
### 5.3 Skyriaus išvados

- Šiuolaikinėse automatinės sintaksinės analizės sistemose rezultatų pagerinimui naudojami semantiniai duomenys. Šiame darbe nauja yra tai, kad sintaksinė analizė atliekama neatsietai nuo morfologijos, t.y. sintaksinės funkcijos skaidomos pagal morfologines kategorijas, siekiant maksimaliai atsižvelgti į specifinę lietuvių kalbos savybę – didelį kaitomumą.
- Sudarant formalų lietuvių kalbos sintaksės taisyklių aprašą BNF, visos sakinio dalys skaidomos pagal morfologinius kriterijus: pradžioje pagal kalbos dalis, kurios gali eiti ta sakinio dalimi, o vėliau pagal kiekvienai kalbos daliai būdingas morfologines kategorijas.
- Visas BNF taisykles galima suskirstyti į dvi grupes. Viena grupė aprašo sakinio dalis, t.y. nurodo, kokiomis morfologinėmis kategorijomis jos gali būti išreikštos. Sintaksinės struktūros grafe tai atitiktų viršūnes. Kita grupė skirta grafo lankams. Šios taisyklės nurodo, kokie žodžiai gali sudaryti junginius. Jos taip pat įvertina žodžių tarpusavio padėtį sakinyje.
- Kiekvienos taisyklės dešinioji pusė sudaryta iš trijų neterminalinių simbolių. Pirmoje ir trečioje pozicijoje yra sakinio dalys, tarp kurių ieškomas sintaksinis ryšys, o viduryje parašomas dar vienas neterminalinis simbolis, kuris vadinasi intarpu ir nurodo kokie kiti žodžiai gali būti įsiterpę tarp aprašomų sakinio dalių.
- Sudaryto BNF aprašo pagalba galima paruošti programinę įrangą, kuri sintaksiškai išnagrinėtų lietuvių kalbos vientisinius sakinius.
- Parodyta, kad kai kuriais atvejais automatinė sintaksinė analizė gali panaikinti žodžių morfologinį daugiareikšmiškumą.

## 6 SINTAKSINĖS ANALIZĖS ALGORITMAS

Visos lietuvių kalbos sintaksės taisyklės šiame darbe aprašytos BNF, o tai yra generuojanti gramatika. Kadangi į BNF taisykles jokių papildomų apribojimų ar nurodymų įdėti negalima, todėl labai svarbu, kaip bus parašytas analizatorius. Kad būtų lengviau skirti operacijų atlikimo prioritetus, visos alternatyvos BNF taisyklėse yra išdėstomos prioritetų mažėjimo tvarka. Tačiau jos iš principo negali nurodyti, kokia tvarka tomis alternatyvomis turi būti naudojamos. Tai jau yra transliatoriaus (analizatoriaus) užduotis. Jis turi būti rašomas taip, kad antra BNF taisyklės alternatyva būtų imama tik tuo atveju, jeigu nepasitvirtina pirmoji. Pvz., apdorojant gijas, radus du žodžius sakinyje, kurie gali sudaryti jos kraštinius narius, kaip pirma alternatyva turi būti imamas intarpas, ir tik tuo atveju, jei jis nerandamas, turėtų būti fiksuojamas žodžių junginys, sudarytas iš greta stovinčių žodžių.

Šiame darbe sakinio centru laikomos dvi sakinio dalys – veiksnys ir tarinys (taip sakinyje traktuojamas tradicinėje gramatikoje), todėl sintaksinėje struktūroje jie abu išdėstomi pačiame viršuje, kaip to paties rango žodžiai. Tačiau ruošiant analizatorių schema, turinti dvi pradines viršūnes, nėra patogi, nes taip pavaizdavus sakinį jo negalima aprašyti formaliomis gramatikomis. Kiekviena formali gramatika privalo turėti keturias dalis (žr. psl. 7) ir viena iš jų yra pradinis neterminalinis simbolis S, nuo kurio pradedamas eilučių generavimas, arba kuris turi būti gautas analizės metu. Taigi, sintaksinėje struktūroje turi būti viena viršūnė, nuo kurios prasidėtų viso grafo formavimas. Ši problema sprendžiama sujungiant abu sakinio struktūros pavaizdavimo būdus: frazių metodą ir priklausomybių gramatiką. Iš frazių metodo bus paimta tikta pradinė viršūnė – sakinyje, kuris skaidomas į dvi grupes: veiksnio ir tarinio. Tuo frazių metodo panaudojimas ir baigiamas. Toliau sakinyje bus formuojamas priklausomybių gramatikos principu: žemyn nuo veiksnio ir tarinio išdėstomi juos išplečiantys žodžiai. Apibendrintas sakinio vaizdas formalios gramatikos požiūriu būtų toks (55 pav.):



55 pav. Apibendrintas sakinio vaizdas formalios gramatikos požiūriu.

Pavaizdavus tokiu būdu sakinį, jo branduolys traktuojamas pagal tradicinę gramatiką ir tuo pačiu struktūra lieka nesusieta su žodžių išsidėstymu sakinyje. Taigi, lietuvių kalbos sakinio struktūrą galima būtų apibūdinti taip: tik vienoje jos viršūnėje (šaknyje) yra sintaksinė kategorija (sakinyje), pagal analogiją su frazių metodo medžio šaknėmis, o visose kitose viršūnėse – sakinio žodžiai, kaip ir priklausomybių gramatikoje. Toks sakinio traktavimo būdas neprieštaruoja ir naujausiems lituanistų darbams. V. Labutis savo knygoje “Lietuvių kalbos sintaksė” rašo: „Visą sakinį pirmiausia galima struktūriškai skaidyti į veiksnio ir tarinio grupes, kurių vienos centre – veiksnys, o kitos – tarinys. Pagal svarbą ir struktūrinį vaidmenį sakinyje yra pagrindo veiksnį ir tarinį laikyti pirmojo rango sakinio dalimis, o visos kitos sakinio dalys būtų žemesnių rangų.“ [Labutis, 2002, psl.202].

## 6.1 Sintaksinės analizės metodika

Dėl didelių skirtumų tarp lietuvių kalbos ir kitų kalbų, kurioms automatinio vertimo sistemos jau yra sukurtos, tiesiogiai pasinaudoti kitų šalių patirtimi negalima, todėl buvo būtina sukurti naują metodą, kuris gerai įvertintų specifinius lietuvių kalbos bruožus – didelį kaitomumą ir laisvą žodžių tvarką sakinyje. Taigi, norint atlikti lietuvių kalbos vientisinių sakinių automatinę sintaksinę analizę, reikia naudoti tokį metodą:

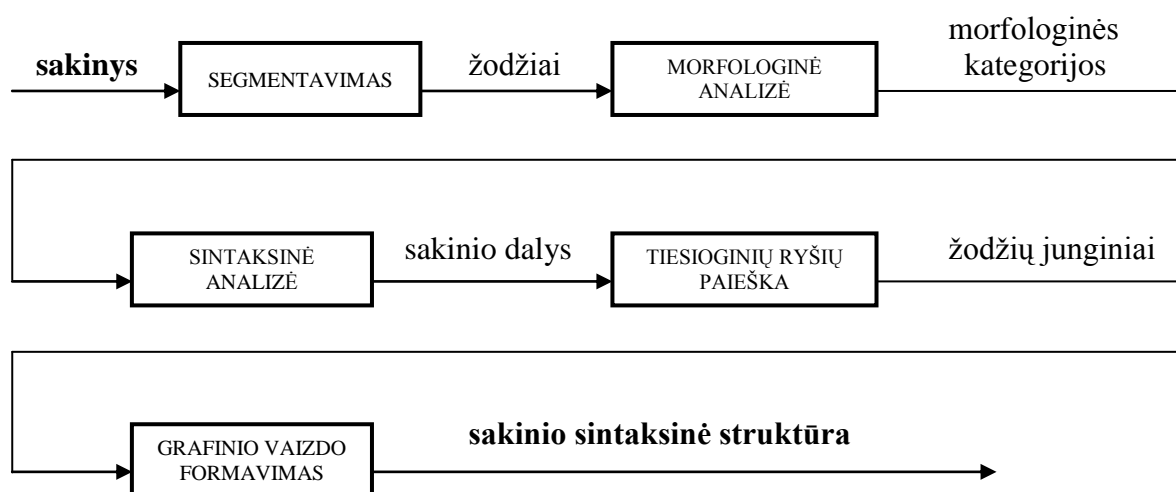
1. Surinkti iš gramatikų informaciją apie visas sakinio dalis ir išsiaiškinti, kokiomis morfologinėmis žodžio formomis gali būti išreikšta kiekviena jų.
2. Sintaksines funkcijas išskaidyti pagal kalbos dalis, kuriomis jos gali būti reiškiamos, pvz., pažyminys, išreikštas būdvardžiu; pažyminys, išreikštas dalyviu; pažyminys, išreikštas skaitvardžiu, ir t.t.
3. Toliau testuoti skaidymą pagal kiekvienai kalbos daliai būdingas morfologines kategorijas, pvz., veiksny, išreikštas daiktavardžiu, toliau turi būti skaidomas į veiksni, išreikštą daiktavardžio vardininku, vienaskaitos vyriškąją gimine; daiktavardžio vardininku, vienaskaitos moteriškąją gimine ir t.t.
4. Išsiaiškinti, kokie semantiniai požymiai gali nulemti morfologinės formos priskyrimą tam tikrai sintaksinei funkcijai (tais atvejais, kai viena morfologinė forma gali atlikti kelias sintaksines funkcijas, tik semantinių požymių pagalba galima vienareikšmiškai nustatyti kokios sakinio dalimi eina žodis nagrinėjamame sakinyje; plačiau žr. psl. 50).
5. Gautą informaciją (1-4 punktai) užrašyti BNF (Bekaus ir Nauro forma, plačiau žr. psl. 7).
6. Turint formaliai užrašytą informaciją apie sintaksinių funkcijų ir morfologinių kategorijų atitikimą, reikia sudaryti kitą BNF taisyklių rinkinį, nurodantį, kaip suieškomi tiesioginiu sintaksiniu ryšiu susiję žodžiai. Tam reikia išsiaiškinti, kokie žodžiai gali būti įsiterpę tarp tiesiogiai susijusių žodžių, ir kokių žodžių tarp jų būti negali. Pvz., tarp nederinamojo pažyminio ir pažymimojo žodžio negali būti įsiterpęs kitas nederinamasis pažyminys, nes jis pažymi tik pirmą po jo einantį daiktavardį, o tarp prielinksnio ir linksnio gali būti įsiterpę ne tik pažyminiai, bet ir tuos pažyminius išplečiantys žodžiai, taigi, ir kitos prielinksninės konstrukcijos, pvz., *Ant aukšto prieš kelis šimtmečius supilto kalno stovėjo pilis*.
7. Rašant žodžių junginių taisykles, kairėje pusėje turi būti neterminalinis simbolis, nurodantis du žodžius, tarp kurių ieškomas ryšys. Dešinė taisyklės pusė turi apimti tris pozicijas: į pirmą ir trečią – junginio žodžius reikia surašyti tokia tvarka, kokia jie gali būti išsidėstę sakinyje, o viduryje turi būti pateiktas neterminalinis simbolis, žymintis informaciją apie tai, kokie žodžiai gali būti įsiterpę tarp tiesioginiu sintaksiniu ryšiu susietų žodžių; jei žodžių tvarkos gali būti abu variantai, pvz., jei papildinys gali būti ir prieš tarinį ir po jo (*Vaikas greitai susikrovė knygas į kuprinę. Knygas į kuprinę vaikas greitai susikrovė.*), kiekvienam atvejui turi būti parašyta atskira taisyklė, atspindinti žodžių išsidėstymą.
8. Intarpą aprašančios BNF taisyklės dešinėje pusėje alternatyvomis nurodomi žodžiai, galintys įsiterpti į aprašomą žodžių junginį. Kai kurios alternatyvos paaimamos į laužtinius ir riestinius skliaustus (apie skliaustų reikšmę BNF plačiau žr. psl. 7), įvertinant tuos atvejus, kai tokio tipo intarpo iš viso nėra arba kai jų gali būti keli. Kaip viena iš intarpo alternatyvų taip pat gali būti ir neterminalinis simbolis, aprašantis ryšį tarp kitų dviejų žodžių t.y. tarp tiesioginiu ryšiu susijusių žodžių gali būti įsiterpęs ir kitas žodžių junginys.
9. Sudarant algoritmą pirmiausia reikia įvertinti, kad mažiausiai variantų turi pagrindinės sakinio dalys, todėl jos turi būti ir surandamos pačios pirmos. Po to aplink veiksni turi būti ieškoma jo pažyminių, įvertinant visus galimus intarpus, ir kai jie bus surasti, pereiti tarinio išplėtimo, t.y. prie papildinių ir aplinkybių nustatymo. Sprendžiant, kuo

eina linksnis svarbiausias kriterijus turėtų būti semantiniai požymiai. Semantinių požymių faile veiksmažodžiams turi būti pateikti tik tie linksniai, kurie eina papildiniu. Linksnis ar prielinksninė konstrukcija, neįeinanti į valdomų linksnų sąrašą, gali būti nustatoma kaip aplinkybė, jei turi reikiamus semantinius požymius. Tais atvejais, kai linksnis negali būti nei papildiniu nei aplinkybe, sakinių reikia traktuoti kaip negalimą. Ryšio tarp linksnio ir prielinksnio ieškoma pagal tą pačią metodiką kaip ir žodžių junginių, t.y. įvertinama, kad tarp žodžio ir jo prielinksnio gali būti įsiterpę kiti žodžiai. Žodžių junginių tarp tarinio ir papildinio pradeda ieškoti nagrinėjant veiksmažodžio valentingumą, t.y. sugebėjimą prisijungti tam tikrus linksnius.

10. Jei visiems žodžiams surandami sintaksiniai ryšiai, konstatuojama, kad tai teisinga sakinio struktūra. Jei baigus apdoroti visus sakinio žodžius bent vienas jų lieka be ryšio su kitu žodžiu, laikoma, kad toks sakiny yra negalimas.

## 6.2 Algoritmo blokinė schema

Lietuvių kalbos vientisinių sakinių sintaksinės analizės algoritmas pateiktas 56 pav. Pradinius duomenis sintaksinei analizei sudaro sakiny. Pirmo etapo metu jis suskaidomas į žodžius ir po to kiekvienam žodžiui surandama jo morfologinė forma. Toliau, ieškoma morfologinių kategorijų sutapimo ir taip surandami tiesiogiai priklausomi žodžiai (žodžių junginiai). Jei visiems žodžiams pavyksta surasti sintaksinius ryšius (t.y. nelieka nė vieno žodžio, neturinčio ryšio su kitu žodžiu), suformuojama sakinio sintaksinė struktūra ir pateikiama grafinio vaizdo lange. Jei bent vienas žodis lieka be ryšio su kitu žodžiu, laikoma, kad sakiniui struktūros sudaryti nepavyko.



56 pav. Algoritmo blokinė schema.

## 6.3 Algoritmo iliustracija pavyzdžiu morfologiniam daugiareikšmiškumui panaikinti

Sintaksinės analizės darbo algoritmą galima pailustruoti sakiniu *Tamsūs pušų sakai blizgėjo saulėje*. 57 pav. (psl. 87) detaliau parodyta pirma blokinės schemos (56 pav.) eilutė, t.y. segmentavimo ir morfologinės analizės blokelių darbo rezultatai. Dėl morfologinio daugiareikšmiškumo gaunama po du žodžių *sakai* ir *blizgėjo* variantus.

SAKINYS	Tamsūs pušų sakai blizgėjo saulėje.				
SEGMENTAI	tamsūs	pušų	sakai	blizgėjo	saulėje
MORFOLOGINĖ ANALIZĖ	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Daikt- <b>vard</b> -dgs-vyrg	Veiksm-vns-3a	Daikt- <b>viet</b> -vns-motg
				Veiksm-vns-2a	Veiksm-dgs-3a

57 pav. Segmentavimo ir morfologinės analizės iliustracija pavyzdžiu.

Reikia įvertinti, kad bet kuri pateikta žodžio morfologinė forma gali būti teisinga sakinyje. Todėl būtina peržiūrėti visas galimas kombinacijas, t.y. sudaryti tiek sakinių, su informacija apie žodžių morfologines kategorijas, kiek yra galimų skirtingų variantų. Šiame pavyzdyje jų yra keturi ir pateikiami 58 pav. Tai iš esmės būtų tas pats sakinys, tik tai žodžiai čia pakeisti jų morfologinėmis kategorijomis. Gautus variantus galima pavadinti morfologiniais sakiniiais.

1.	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Daikt- <b>vard</b> -dgs-vyrg	Veiksm- <b>vns</b> -3a	Daikt- <b>viet</b> -vns-motg
2.	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Veiksm-vns-2a	Veiksm- <b>vns</b> -3a	Daikt- <b>viet</b> -vns-motg
3.	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Daikt- <b>vard</b> -dgs-vyrg	Veiksm- <b>dgs</b> -3a	Daikt- <b>viet</b> -vns-motg
4.	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	V-eiksm-vns-2a	Veiksm- <b>dgs</b> -3a	Daikt- <b>viet</b> -vns-motg

58 pav. Morfologinių sakinių variantai.

Sintaksinės analizės blokelių (56 pav., psl. 86) darbas detaliau parodytas 59 pav. Žodis *pušų* (daugiskaitos kilmininkas) gali atlikti dvi sintaksines funkcijas – papildinio ir pažyminio. Čia taip pat reikia įvertinti, kad bet kuri jų gali būti teisinga pateiktame sakinyje. Todėl iš viso gaunami 8 galimi pateikto sakinio variantai. Visi jie parodyti 60 pav. (psl. 88) Sakinys, kuriame žodžiai pakeisti jų sintaksinėmis funkcijomis, šiame darbe vadinamas sintaksiniu sakiniu.

SINTAKSINĖ ANALIZĖ	DERIN-PAŽYM	NEDERIN-PAŽYM	VEIKSN	TARIN	APLINKYB
	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Daikt- <b>vard</b> -dgs-vyrg	Veiksm-vns-3a	Daikt- <b>viet</b> -vns-motg
	PAPILD				
	Daikt- <b>kilm</b> -dgs-motg				

59 pav. Sintaksinės analizės iliustracija pavyzdžiu.

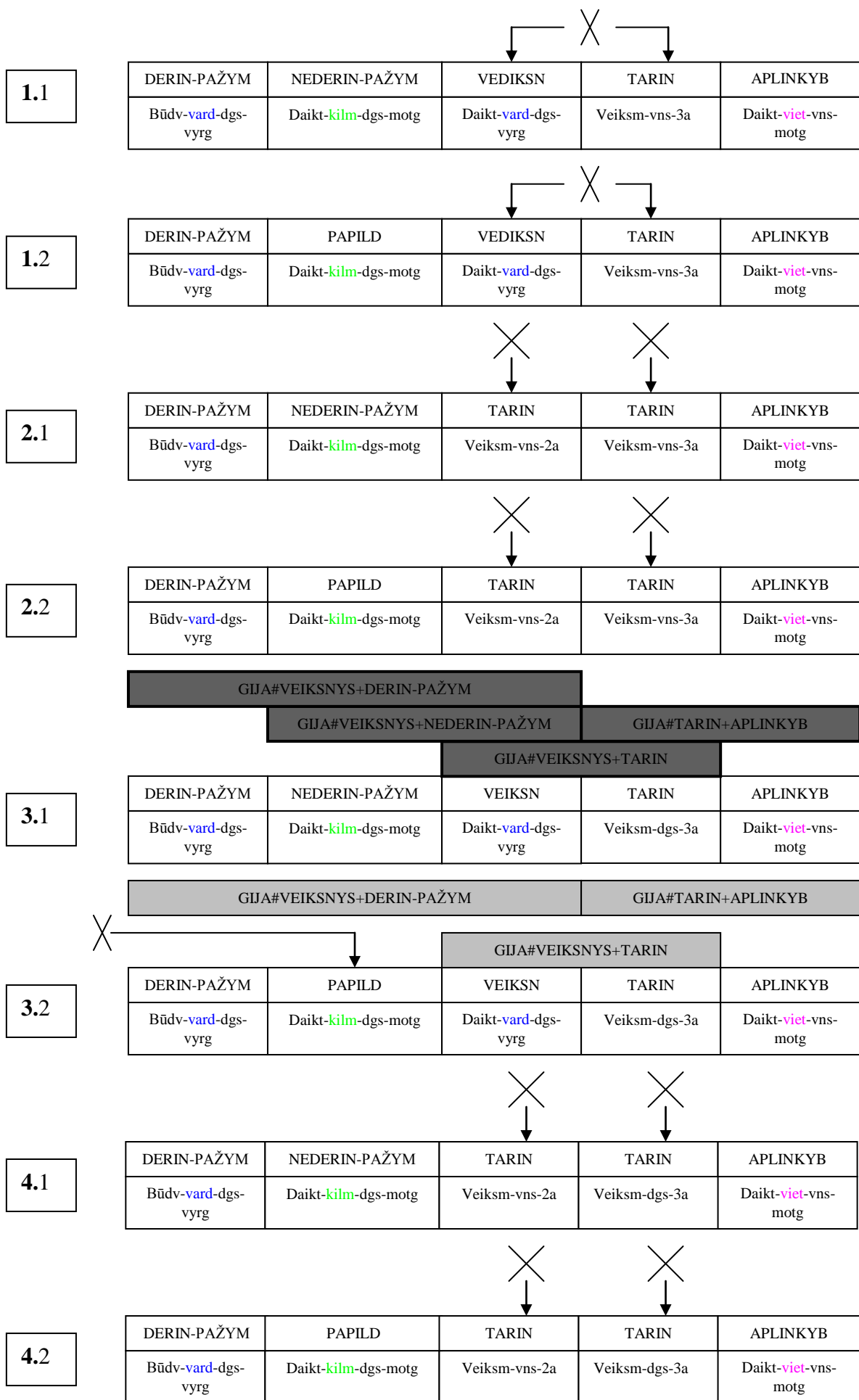
1.1	DERIN-PAŽYM	<b>NEDERIN-PAŽYM</b>	VEIKSN	TARIN	APLINKYB
	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Daikt- <b>vard</b> -dgs-vyrg	Veiksm- <b>vns-3a</b>	Daikt- <b>viet</b> -vns-motg
1.2	DERIN-PAŽYM	<b>PAPILD</b>	VEIKSN	TARIN	APLINKYB
	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Daikt- <b>vard</b> -dgs-vyrg	Veiksm- <b>vns-3a</b>	Daikt- <b>viet</b> -vns-motg
2.1	DERIN-PAŽYM	<b>NEDERIN-PAŽYM</b>	TARIN	TARIN	APLINKYB
	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Veiksm- <b>vns-2a</b>	Veiksm- <b>vns-3a</b>	Daikt- <b>viet</b> -vns-motg
2.2	DERIN-PAŽYM	<b>PAPILD</b>	TARIN	TARIN	APLINKYB
	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Veiksm- <b>vns-2a</b>	Veiksm- <b>vns-3a</b>	Daikt- <b>viet</b> -vns-motg
3.1	DERIN-PAŽYM	<b>NEDERIN-PAŽYM</b>	VEIKSN	TARIN	APLINKYB
	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Daikt- <b>vard</b> -dgs-vyrg	Veiksm- <b>dgs-3a</b>	Daikt- <b>viet</b> -vns-motg
3.2	DERIN-PAŽYM	<b>PAPILD</b>	VEIKSN	TARIN	APLINKYB
	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Daikt- <b>vard</b> -dgs-vyrg	Veiksm- <b>dgs-3a</b>	Daikt- <b>viet</b> -vns-motg
4.1	DERIN-PAŽYM	<b>NEDERIN-PAŽYM</b>	TARIN	TARIN	APLINKYB
	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Veiksm- <b>vns-2a</b>	Veiksm- <b>dgs-3a</b>	Daikt- <b>viet</b> -vns-motg
4.2	DERIN-PAŽYM	<b>PAPILD</b>	TARIN	TARIN	APLINKYB
	Būdv- <b>vard</b> -dgs-vyrg	Daikt- <b>kilm</b> -dgs-motg	Veiksm- <b>vns-2a</b>	Veiksm- <b>dgs-3a</b>	Daikt- <b>viet</b> -vns-motg

60 pav. Sintaksinių sakinių variantai.

Tiesioginių ryšių paieškos blokelis detaliau parodytas 61 pav. (psl. 89) Ieškant sintaksinių ryšių pagal morfologinių kategorijų atitikimą, atmetami visi atvejai, kurie pateiktam sakiniui netinka. Pavyzdyje parodyta, kaip iš gautų aštuonių variantų paliekamas vienas galimas. Kombinacijos 1.1 ir 1.2 atmetamos dėl morfologinių kategorijų nesutapimo: veiksnys išreikštas daiktavardžio daugiskaitos vardininku, o tarinys – veiksmožodžio vienaskaita. Kombinacijose 2.1 ir 2.2 o taip pat 4.1 ir 4.2 tariniams nesurandami veiksniai, t.y. nenustatomas sakinio branduolys. Kombinacijoje 3.2 vienas žodis lieka be ryšio su kitais žodžiais, nes tariniu einantis veiksmožodis nevaldo kilmininko t.y. tarinys nereikalauja vardažodžio (daiktavardžio, įvardžio) kilmininko linksnyje kaip papildinio, todėl šis variantas taip pat traktuojamas kaip negalimas sakiny. Taigi, tik kombinacija 3.1 laikoma teisinga, t.y. atitinkančia pateiktą sakinį.

Nagrinėjant kitą sakinį, gali skirtis morfologinių variantų ir sintaksinių variantų kiekis, taigi, ir morfologinių bei sintaksinių sakinių kiekis. Tačiau galutinis variantas vis tiek būtų tas pats: iš visų gautų variantų bus paliekamas vienas galimas.





61 pav. Tiesioginių sintaksinių ryšių nustatymas.

## 6.4 Skyriaus išvados

- Sukurtas lietuvių kalbos vientisinių sakinių automatinės sintaksinės analizės metodas, įvertinantis specifinius lietuvių kalbos bruožus – didelį kaitomumą ir laisvą žodžių tvarką sakinyje.
- Sudarytas algoritmas lietuvių kalbos vientisinių sakinių automatinei sintaksinei analizei atlikti, grindžiamas formaliu lietuvių kalbos sintaksės aprašu BNF (Bekaus ir Nauro forma).
- Algoritmo veikimas pailiustruotas sakinio pavyzdžiu, kuriame atsispindi sintaksinės analizės programinės įrangos sugebėjimas panaikinti morfologinį daugiareikšmiškumą.

## 7 SINTAKSINĖS ANALIZĖS ALGORITMO PROGRAMINĖ ĮRANGA

Lietuvių kalbos sakinių sintaksinės analizės programinė įranga rašoma Visual Basic'6 kalba, kuri priklauso objektinio programavimo kalbų grupei. Būtent ši kalba buvo pasirinkta atsižvelgiant į tai, kad automatinė morfologinė analizė, t.y. lietuvių kalbos morfologijos modelių programinė įranga jau yra parašyta ta kalba, ir todėl bus paprastesnis duomenų perėmimas, nes morfologinės analizės darbo rezultatai naudojami kaip pradiniai duomenys sintaksinei analizei.

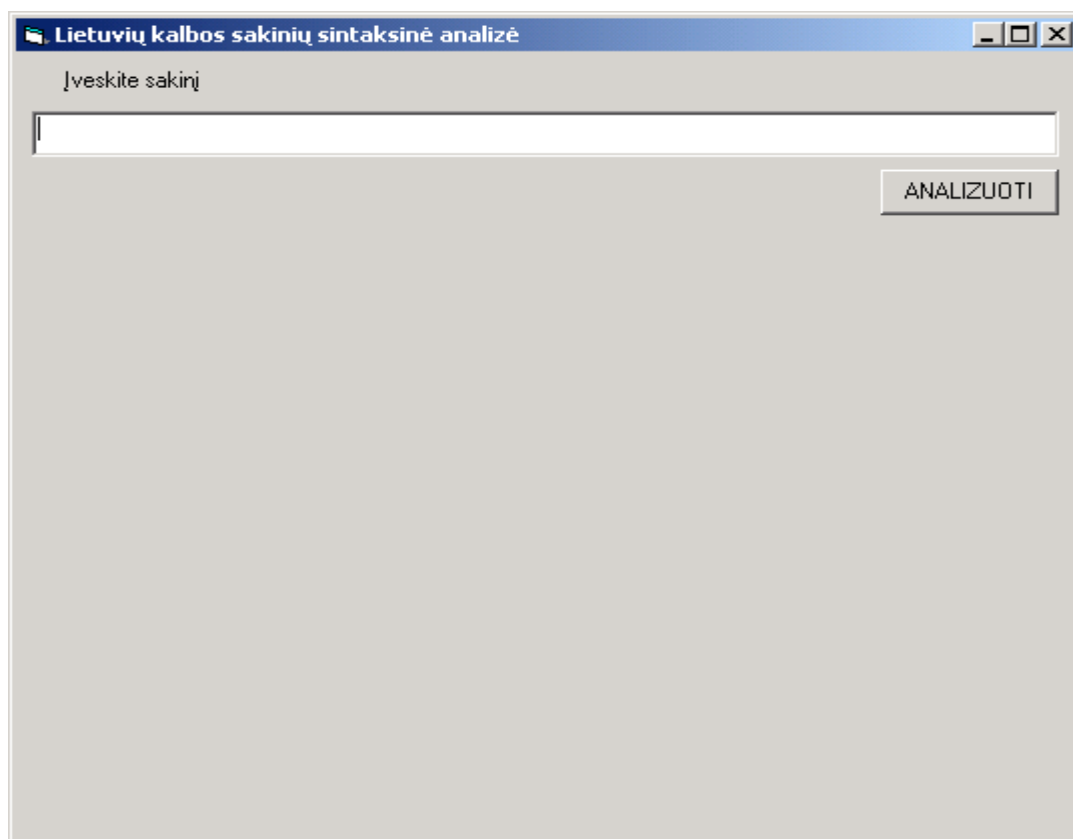
Visual Basic programa susideda iš nedidelių loginių blokų – paprogramių. Taip suskirstytą programą lengviau derinti, nes ją galima suskaidyti į nedidelius uždavinius ir paprogramių kodus rašyti bei jas derinti nepriklausomai viena nuo kitos [Šulcas, psl. 43]. Paprogramės Visual Basic'6 kalboje būna dviejų tipų:

- 1) bendrosios, kurios yra iškviečiamos iš paleidimo paprogramės; bendrųjų paprogramių šiame darbe yra 31;
- 2) paleidimo paprogramė, kuri pradeda savo darbą, kai vartotojo lange paveikiamas koks nors objektas.

Valdančiosios (paleidimo) paprogramės darbas inicijuojamas mygtuko Command1 paspaudimu (\_click). Šis mygtukas objektų lange pažymimas žodžiu ANALIZUOTI.

Pradiniai duomenys sintaksinei analizei yra lietuvių kalbos sakiny, kurį numatoma įvesti klaviatūros pagalba arba perkelti pele nukopijuotą. Informacijai įvesti iš klaviatūros naudojamas teksto laukelis [Starkus, psl. 55], todėl pradiniam lange patalpinamas laukelis su užrašu virš jo „Įveskite sakinį“. Kai sakiny, kurio sintaksinę struktūrą norima gauti, patalpinamas į šį laukelį, mygtuku ANALIZUOTI paleidžiamas programos darbas.

Pradinio lango grafinis vaizdas pateikiamas 62 pav.



62 pav. Pradinio lango grafinis vaizdas.

Visą programą sudaro 41543 eilutės Visual Basic'6 kalboje. Programoje panaudotos 7 Visual Basic funkcijos.

## 7.1 Programinės įrangos blokinė schema

Sintaksinės analizės programinės įrangos blokinė schema pateikta 63 pav.(psl. 94).

Atliekant sintaksinę analizę pradinis duomenis sudaro lietuvių kalbos sakiny. Jis įvedamas iš teksto laukelio kaip simbolių eilutė ir priskiriamas kintamajam Sakinys atmetant (funkcijos *Trim* pagalba) visus tarpus prieš pirmą žodį ir už paskutinio simbolio laukelyje. Jei sakinio pabaigoje nebus taško, jis vis tiek bus nagrinėjamas.

Pradžioje sakiny turi būti suskaidytas į žodžius. Tai atlieka paprogramė SEGMENTAVIMAS. Duomenų pasikeitimas vyksta naudojant tris parametrus. Sakiny paprogramei perduodamas kaip eilutės tipo kintamasis, o grąžinami du dydžiai: kintamasis ŽodžiųSkaičius, kuriame yra informacija apie tai, iš kelių žodžių sudarytas sakiny ir vienmatis masyvas segmentai, kurio kiekvienas elementas atitinka vieną sakinio žodį. Žodžių ribos nustatomos tokiu būdu: po vieną tikrinami eilutės tipo kintamojo Sakinys simboliai ir žiūrima ar tai nėra skyriklis (tarpas, skyrybos ženklas, skaičius ir pan.); visi skyriklių aibei nepriklausantys simboliai, kaupiami kintamajame Segmentas. Sutikus skyriklį fiksuojama, kad kintamojo Segmentas reikšmė yra žodis, ir jo turinys priskiriamas segmentų masyvo elementui. Masyvo elementai pradedami numeruoti nuo 1 ir elemento indeksas atitinka žodžio vietą sakinyje, t.y masyve žodžiai išdėstomi ta pačia tvarka, kaip jie buvo pavartoti sakinyje.

Toliau kreipiamasi į paprogramę MORFOLOGIJA, perduodant jai per parametrus žodžių skaičių ir suskaidytą į žodžius sakinį, t.y. masyvą segmentai. Šiame darbe buvo pasitelkta lietuvių kalbos morfologijos modelių programinė įranga, kurią sukūrė Vytautas Zinkevičius (plačiau apie tai rašoma literatūroje [Zinkevičius]). Ji prieinama kaip dinaminių saitų bibliotekos [Kudirka, psl.42] (DLL – Dynamic Link Library) failas.

Lietuvių kalboje nėra tokios aštrios daugiareikšmiškumo problemos, kokią turi anglų kalba, kur beveik kiekvienas veiksmažodis yra kartu ir daiktavardis. Tačiau kartais gali būti kelios to paties žodžio morfologinės reikšmės. Morfologinės analizės metu kartais gaunami keli galimi variantai. Beveik pusė (47%) lietuvių kalbos žodžių morfologiniu požiūriu yra daugiareikšmiai [Rimkutė, psl.46], t.y. tas pats žodis gali atitikti kelias morfologines formas. Visada sutampa veiksmažodžių vienaskaitos ir daugiskaitos 3-io asmens formos. Dažniausiai sutinkamos sutampančios daiktavardžių formos yra: moteriškos giminės vienaskaitos vardininkas ir įnagininkas (kas? – *knyga* ir kuo? *knyga*), vienaskaitos kilmininkas ir daugiskaitos vardininkas (ko? – *gėlės* ir kas? – *gėlės*). Kartais gali sutapti ir skirtingų kalbos dalių atskiros formos, pvz., veiksmažodžio bendratis *rašyti* ir būtojo laiko dalyvis *rašyti* (*jam rašyti laiškeliai gulėjo neišsiūsti*) ir t.t. Todėl morfologinės analizės paprogramė grąžina jau nebe vienmatį, o trimatį (x,y,z) masyvą. Kiekvienam žodžiui numatoma 100 galimų reikšmių, t.y. masyvo elementų skaičius z-kryptimi yra 100; y-kryptimi pateikiami sakinio žodžiai (t.y. elemento indeksas rodo žodžio eilės numerį sakinyje), o x-kryptimi išdėstomi 8 elementai, turintys informaciją apie žodį morfologiniu, sintaksiniu ir semantiniu aspektais. Jų reikšmė pateikiama 6 lentelėje, psl. 93.

Jei kalbos dalis neturi kokios nors morfologinės kategorijos (pvz., veiksmažodis nekaitomas linksniais, daiktavardis – asmenimis), į atitinkamą masyvo elementą neįrašoma nieko t.y. jis traktuojamas kaip tuščia eilutė ir apdorojimo metu programa nesikreipia į šį elementą. Morfologijos paprogramė surašo informaciją į pirmus šešis ir į aštuntąjį elementą, kuriuose yra morfologiniai ir semantiniai duomenys; lieka tik septintasis, sintaksinę funkciją rodantis elementas, kurį užpildo sintaksės paprogramė. Pirmiems penkiems ir aštuntajam elementui informacija paimama iš dinaminių saitų bibliotekos (DLL – Dynamic Link Library) failo, naudojantis morfologinės analizės funkcija. Kiekvieną kartą kreipiantis į šią biblioteką galima gauti gramatinę informaciją apie vieną žodį. Kreipimosi į DLL metu morfologinės analizės funkcijai reikia perduoti tris parametrus:

- simbolių eilutė (žodį, kurį reikia išnagrinėti gramatiškai),
- atminties adresą, kur turės būti įrašytas rezultatas,
- atminties srities, skirtos rezultatui, dydį (reikalingą tam, kad DLL funkcija, įrašydama atsakymą, neviršytų jam skirtos atminties dydžio).

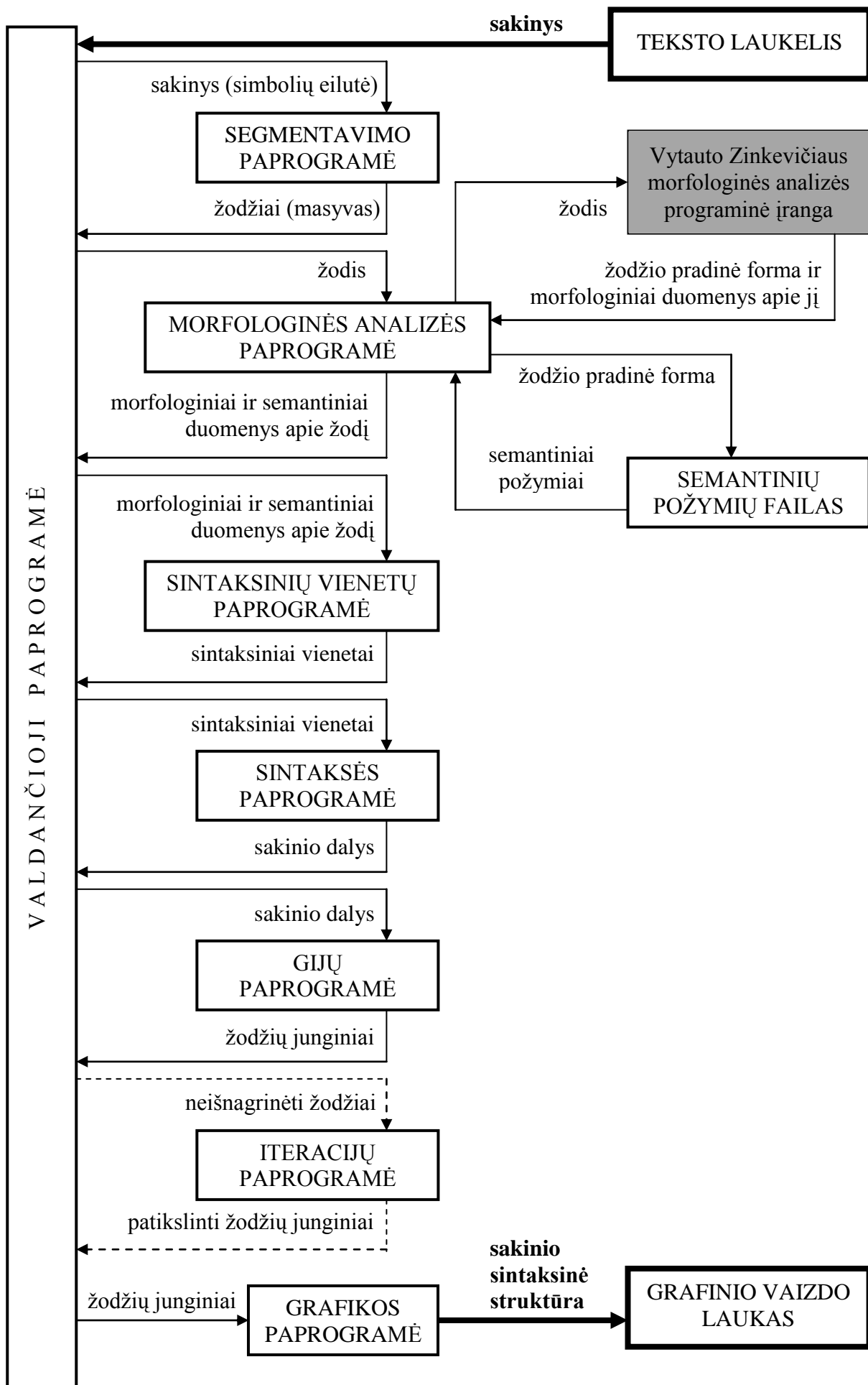
**6 lentelė** Masyvo elementuose (x-kryptimi) saugomos informacijos apie žodį pobūdis.

Elemento eilės nr. (x-kryptimi)	Saugomos informacijos pobūdis
1	kalbos dalis
2	linksnis
3	skaičius
4	giminė
5	asmuo
6	semantiniai požymiai
7	sintaksinė funkcija
8	įvardžiutinumumas

Morfologinės analizės funkcija pateikia informaciją apie žodį gramatinėmis kategorijomis. Pvz., simbolių eilutei *namas* ji nustato: „daiktavardis, vyriškoji giminė, vienaskaita, vardininkas“. Kaip tai atliekama (kokie lietuvių kalbos leksikos bei gramatikos žinias atspindintys duomenys tam naudojami ir kt.) smulkiau aprašyta literatūroje [Zinkevičius]. Jei žodis atitinka kelias formas, funkcija nurodo jas visas. Pvz., simbolių eilutei *mes* pateikiama:

- „įvardis, daugiskaita, vardininkas“,
- „veiksmažodis, nesangražinis, tiesioginė nuosaka, būsimasis laikas, vienaskaita, III asmuo“,
- „veiksmažodis, nesangražinis, tiesioginė nuosaka, būsimasis laikas, daugiskaita, III asmuo“.

Sėkmingo funkcijos įvykdymo atveju (t.y. jei gautą simbolių eilutę pavyksta atpažinti kaip lietuvių kalbos žodį) į rezultato sritį įrašoma informacija apie vieną ar kelias galimas to žodžio gramatines formas. Kiekvienos formos duomenis sudaro lema (pradinė forma – bendratis veiksmažodžiams; vienaskaitos vardininko linksnis daiktavardžiams, būdvardžiams, skaitvardžiams ir t.t.) ir 13-kos skaičių kodas, kuriame yra informacija apie kalbos dalį, sangražišumą, veikslą, nuosaką, laiką, laipsnį, įvardžiutinumą, giminę, skaičių, linksnį, asmenį ir kt. Ne visa ši informacija naudojama sintaksinei analizei, nes kai kurie duomenys neturi įtakos sintaksinės funkcijos nustatymui, pvz., veiksmažodžių laikas ir kt.



63 pav. Programinės įrangos blokinė schema.

Po to prie gautų morfologinių duomenų prijungiama semantinė informacija – duomenys apie žodį, kurie gali turėti įtakos nustatant jo sintaksinę funkciją. Ši informacija saugoma atskirame faile. Tai tekstinis failas. Jo įrašai nuskaitomi po eilutę. Kiekvienoje eilutėje talpinama žodžio pradinė forma ir po jos tarpu atskirti to žodžio semantiniai požymiai. Šiame darbe naudojamų požymių sąrašas būtų toks:

- laiko požymis daiktavardžiams, kuris sutrumpintai užrašomas *laikp*; šį požymį turi tokie žodžiai kaip *savaitė, vakaras, ruduo* ir t.t.
- vietos požymis daiktavardžiams sutrumpintai žymimas *vietp*;
- asmens požymis; jis suteikiamas įvardžiams, kurie nurodo asmenis, nes morfologinės analizės programa šios informacijos neduoda; sutrumpintai žymima taip: *asm*;
- tikriniams daiktavardžiams nurodomi požymiai *vardas, pavardė, vietovardis, valstybė* ir k.t.
- socialiniai požymiai – *įstaiga, pareigybė* ir t.t.
- valdomų linksnių sąrašas veiksmažodžiams; sutrumpintai linksniai žymimi taip: *kilm, naud, gal, įnag*; jei veiksmažodis valdo kelis linksnius, pateikiami visi, atskiriant juos tarpu;
- valdomos prielinksninės konstrukcijos, kurios gali eiti papildiniu taip pat pateikiamos veiksmažodžių semantiniame lauke, jos žymimos be tarpo rašant prielinksniį ir jo valdomą linksniį, pvz.: *antkilm, bekilm, dėlkilm, iškilm, nuokilm, priekilm, apiegal, įgal, pergal, priešgal, užgal, sujnag* ir t.t.

Kiekvienai morfologinės analizės metu gautai pradinei formai atliekama paieška semantinių požymių faile. Žodžio pradinė forma naudojama kaip filtras. Operatoriumi *Like* ji lyginama su eilutės pradžia (t.y. su simbolių seka iki pirmo tarpo). Suradus eilutę, kurios pradžia sutampa su filtru, likusi eilutės dalis perrašoma į šeštąjį (semantinių požymių) elementą duomenų apie žodį masyve (plačiau žr. 6 lentelė, psl. 93). Šie duomenys vėliau naudojami sprendžiant, kokią sintaksinę funkciją gali atlikti tas žodis. Semantinių požymių failą galima redaguoti nepriklausomai nuo programos, todėl visada bus galima įtraukti naują žodį bei naują požymį, taigi, yra numatytos sistemos išplėtimo galimybės.

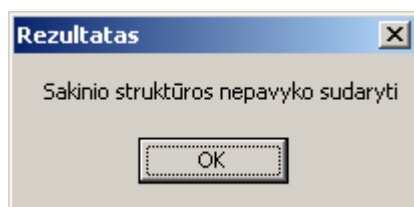
Šiame darbe į minėtą failą įrašoma tik ta semantinė informacija, kuri naudojama BNF apraše t.y. yra būtina sintaksinei analizei atlikti. Todėl reikia pasakyti, kad tai nėra išsamus semantinis žodžio aprašas, nes teapima tik sintaksės reikmėms tarnaujančius požymius, t.y. tuos, kurie gali nulemti žodžio sintaksinę funkciją ar jo ryšį su kitu žodžiu. Pvz., daiktavardžio galininkas paprastai eina papildiniu (*skaito knygą*), jei veiksmažodis stipriai valdo galininką, bet tais atvejais, kai daiktavardis turi laiko požymį, jis eina aplinkybe, net ir esant stipriam galininko valdymui (*skaito naktį*), ir jei abu tokie galininkai būna viename sakinyje (*Visą naktį ji skaitė šią knygą* ir *Visą knygą ji perskaitė šią naktį*), tik semantiniai požymiai nulemia, kuriai sakinio daliai priskirti žodį. Semantinė informacija pradiniam variante pateikiama tik apie daiktavardžius, įvardžius ir veiksmažodžius iš pirmų 2000 žodžių pagal lietuvių kalbos dažninį žodyną. Dažninis žodynas padeda atrinkti reikalingiausius žodžius. Apie tai kalba skaičiai: dešimt dažniausių žodžių užima net 12.33% tekstų, nors patys tesudaro vos 0,03% visų tekstuose pasitaikančių žodžių; penkiasdešimt dažniausių žodžių užima 24,27% teksto, penki šimtai žodžių – daugiau kaip 50%, o du tūkstančiai – net 75,50% tekstų [Grumadienė, Žilinskienė, psl. 32]. Taigi tikėtina, kad dažninio žodyno pagalba bus gautas labiausiai vartojamų žodžių sąrašas. Vėliau, naudojant sintaksinės analizės programą, sutikus žodį, apie kurį nėra duomenų semantinių požymių faile, bus galima juos įtraukti. Baigusi darbą paprogramė MORFOLOGIJA gražina trimatį masyvą „morfologijos variantai“, kuriame yra informacija apie visas galimas kiekvieno žodžio morfologines formas ir vienmatį masyvą „morfologijos variantų skaičius“, kuriame kiekvienam žodžiui nurodytas

galimų jo formų kiekis. Masyvo elemento indeksas rodo žodžio eilės numerį sakinyje, o elemento įrašas atitinka to žodžio morfologinių formų skaičių.

Valdančioji paprogramė suformuoja morfologinį sakinį, t.y. pirmą pateiktą morfologinių kategorijų variantą. Tai dvimatis masyvas (x, y). Kadangi šio masyvo dydis priklauso nuo žodžių skaičiaus sakinyje, todėl jis skelbiamas kaip dinaminis, t.y. kintamo dydžio. Visual Basic'6 kalboje galima keisti tik paskutinį masyvo matavimą [Šulcas, psl. 34], todėl žodžių skaičius nurodomas kaip antrasis matavimas. Pirmas matavimas yra 8 ir šia kryptimi išdėstoma gramatinė informacija apie žodį (plačiau žr. psl. 93, 6 lentelė). Kurio žodžio informacija pateikta, žymi antrasis masyvo indeksas.

Paskui morfologinis sakinytis perduodamas paprogramei SINTAKSĖ. Ji pagal kiekvieno žodžio morfologines kategorijas nustato galimas jo sintaksines funkcijas remiantis sakinio dalių aprašu BNF (Bekaus ir Nauro forma). Čia daugiareikšmiškumo taip pat neišvengiama, nes ta pati morfologinė forma gali atlikti kelias sintaksines funkcijas, pvz., daiktavardžio kilmininkas gali eiti ir nederinamuoju pažyminiu (*vilko kailis*), ir papildiniu (*bijau vilko*), bendratis gali eiti tiek veiksniu (*dirbti šachtoje buvo sunku*), tiek papildiniu (*dainuoti ji mėgdavo*), tiek pažyminiu (*noras gyventi buvo didelis*); kai kurios prielinksninės konstrukcijos (pvz., prielinksnis *iš* su kilmininku) gali eiti ir pažyminiu, ir papildiniu, ir aplinkybe (*namas iš plytų* – pažyminys, *išsiskirti iš bendraamžių* – papildinys, *grįžti iš miesto* – aplinkybė); prieveiksmis gali eiti ir aplinkybe (*gražiai dainuoja*), ir nederinamuoju pažyminiu (*kelias atgal*) ir pan.

Ši daugiareikšmiškumą panaikina paprogramė GIJOS, sudarydama sakinio sintaksinę struktūrą. Pradžioje surandamas sakinio branduolys, po to pradedama ieškoti veiksnį ir tarinį išplečiančių žodžių pagal morfologinių kategorijų atitikimą, kartu įvertinant semantinius požymius. Tais atvejais, kai analizuojamam sakiniui pavyksta sudaryti sintaksinę struktūrą, t.y. jei surandamos gijos (tiesioginiai sintaksiniai ryšiai) visiems žodžiams ir nė vienas žodis nelieka neapdorotas (t.y. nelieka nė vieno žodžio be ryšio su kitu žodžiu), konstatuojama, kad surasta teisinga sakinio struktūra ir informacija apie šį faktą nurodoma gražinant loginio tipo kintamąjį „analizės rezultatas“. Kartu gražinamas ir dvimatis masyvas „sintaksiniai ryšiai“. Jame kiekvienam žodžiui nurodomi visi žodžiai, su kuriais jis turi tiesioginius sintaksinius ryšius. Jei nepavyksta sudaryti sakinio struktūros, t.y. jei bent vienas žodis lieka be ryšio su kitu žodžiu, sakinytis perduodamas paprogramei ITERACIJOS. Ši paprogramė neišnagrinėtam žodžiui ieško kito morfologinės informacijos varianto ir bando priskirti jam kitą sintaksinę funkciją bei surasti ryšį su koku nors žodžiu, t.y. imamas kitas morfologinis neišnagrinėto žodžio variantas ir sakinytis pradedamas nagrinėti iš naujo (daroma iteracija). Jei pavyksta gauti teigiamą analizės rezultatą, išnagrinėtas sakinytis perduodamas grafinio vaizdo formavimo paprogramei ir jo sintaksinė struktūra pateikiama ekrane. Jei vis dėlto nepavyksta rasti sintaksinių ryšių visiems sakinio žodžiams, konstatuojama, kad toks sakinytis yra negalimas ir informacija apie tai pateikiama pranešimų kortelėje (Message Box) panaudojant standartinę funkciją MsgBox. Pranešimo kortelės vaizdas pateikiamas 64 pav.

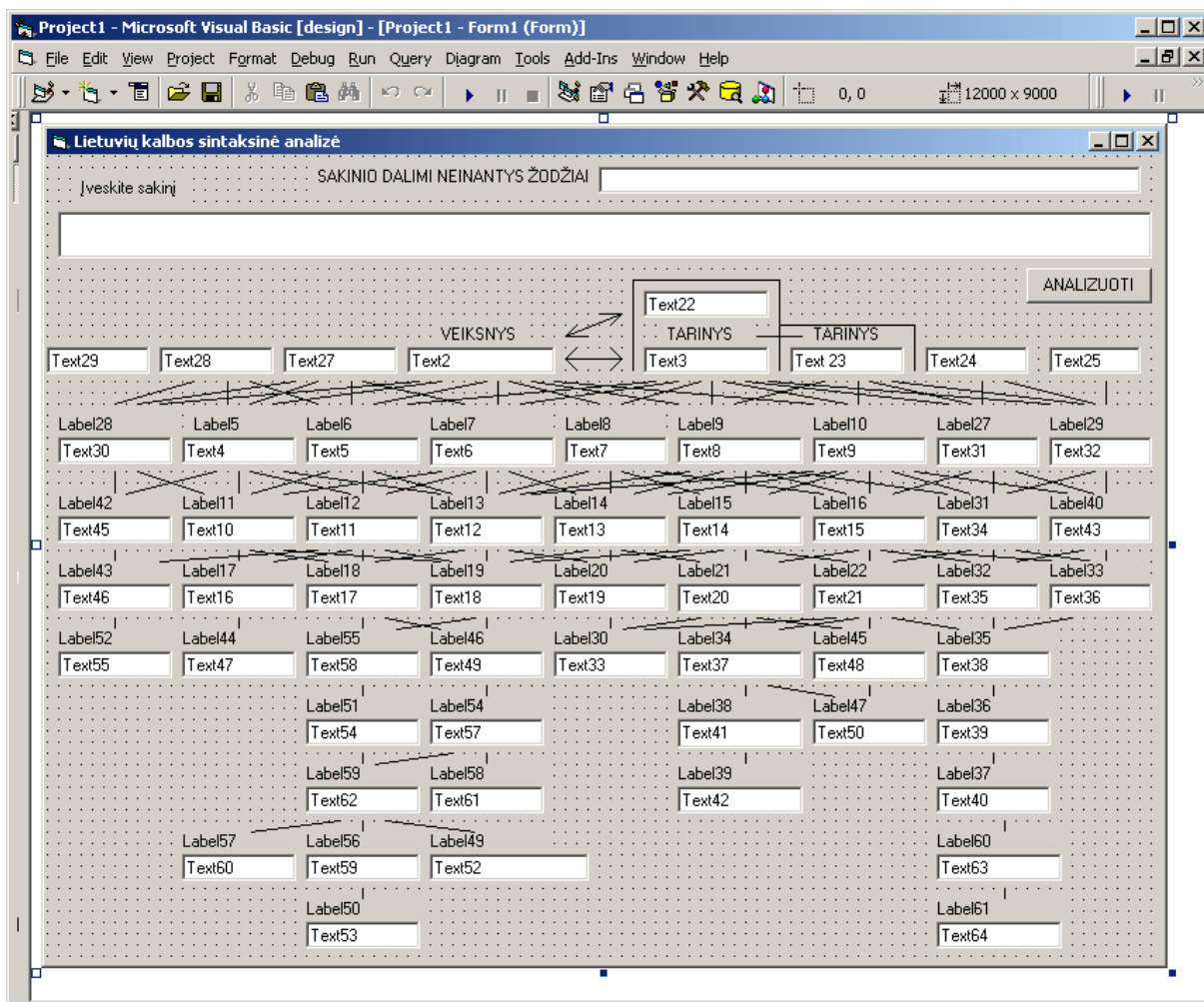


64 pav. Pranešimo kortelė, informuojanti kad sakinio struktūros sudaryti nepavyko.

Kaip pagrindas grafo pavaizdavimui imama matrica, kurios elementuose talpinami sakinio žodžiai kartu su informacija apie jų sintaksinę funkciją. Eilutės atitinka žodžių rangus pagal sintaksinę priklausomybę sakinyje. Veiksnytis ir tarinytis laikomi pirmo rango sakinio dalimis ir išdėstomi grafo viršuje. Su veiksniu ar su tariniu tiesiogiai susietos sakinio dalys laikomos antro rango ir išdėstomos antroje eilutėje. Antro rango sakinio dalis išplečiantys



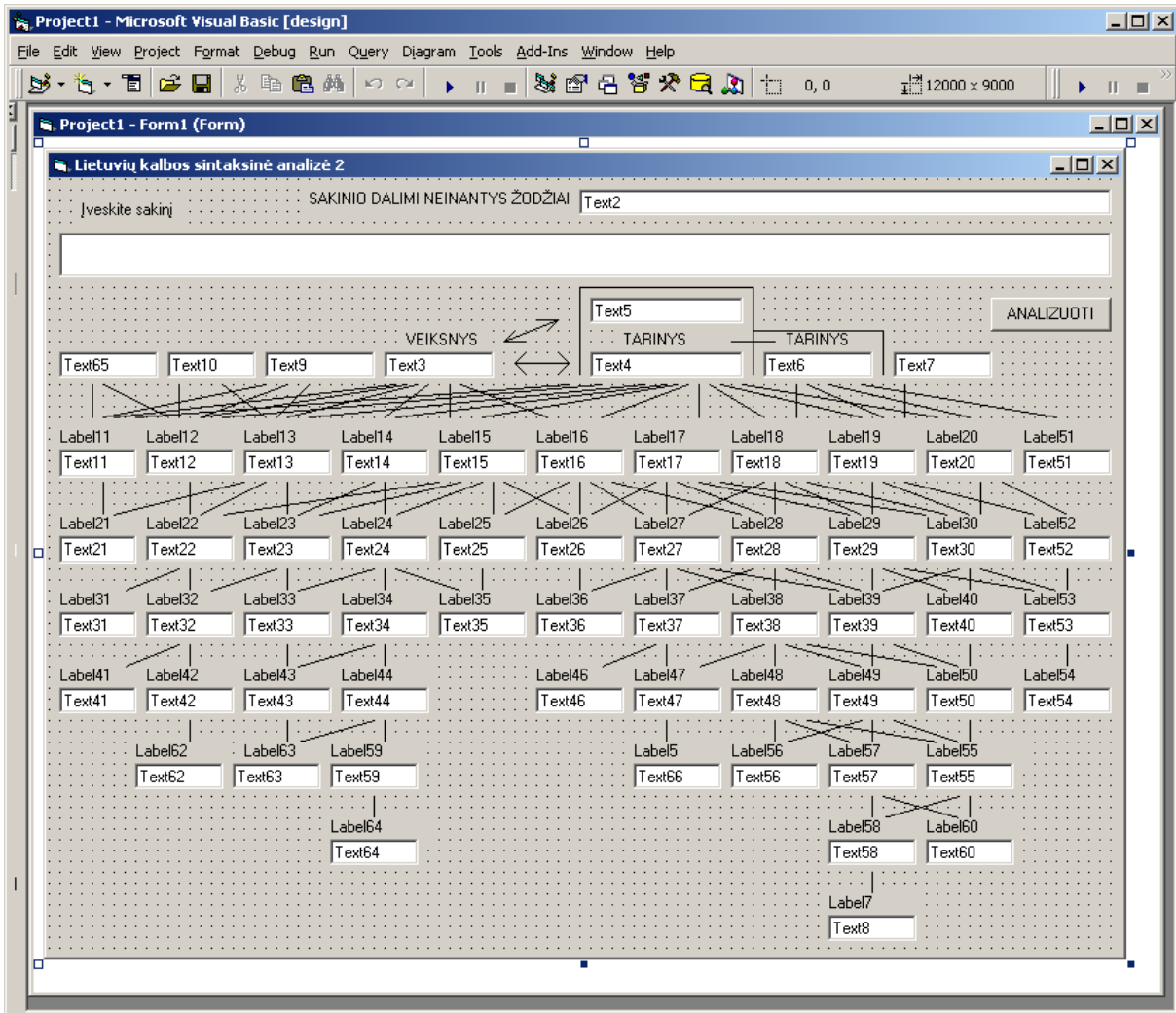
žodžiai laikomi trečio rango ir išdėstomi trečioje eilutėje ir t. t. Sakinio struktūros grafinio vaizdo kontūrai, kuriuose formuojama kiekvieno konkretaus sakinio struktūra, pateikiami 65 pav.



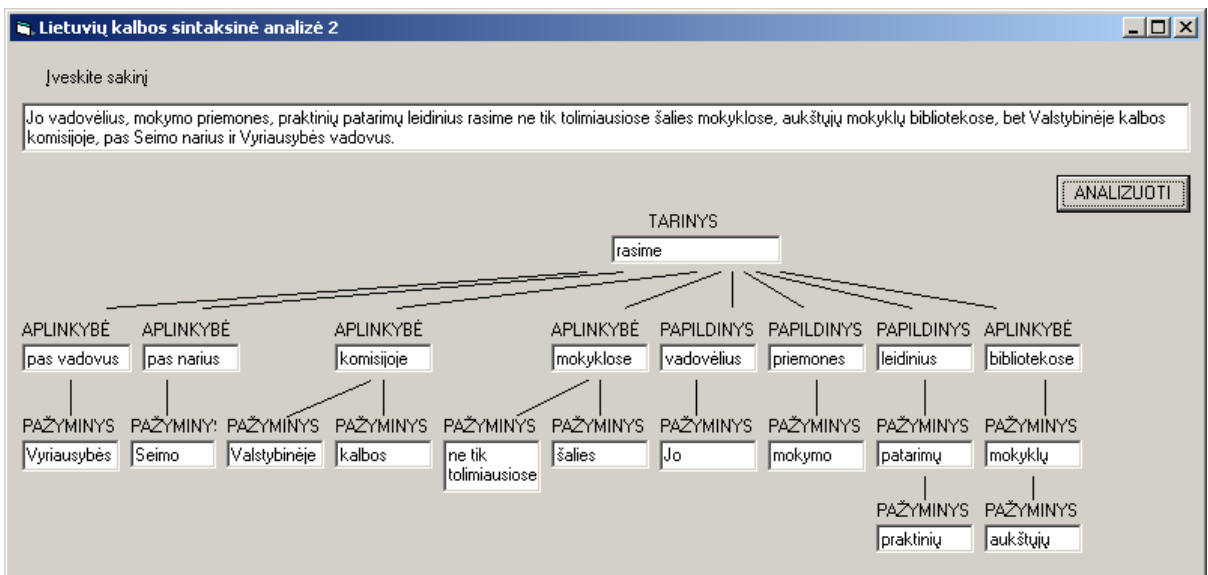
65 pav. Sakinio struktūros grafinio vaizdo kontūrai.

Atvaizduojant sakinį ekrane išryškėjo naudojamos programavimo kalbos ribotumas. Visual Basic'6 leidžia viename lange naudoti tik 256 objektus. Esant didesnei sintaksinių struktūrų įvairovei visų sakinių negalima pavaizduoti viename lange, todėl reikėjo naudoti kelis langus. Dar vieno tokio lango pavyzdys pateikiamas 66 pav., psl. 98. Dėl Visual Basic '6 ribotumo būtų tikslinga programinę įrangą perrašyti C++ programavimo kalba.

Darbo pradžioje visa ši struktūra yra nematoma (pradinis lango vaizdas pateiktas 62 pav., psl. 91). Atlikus sakinio analizę išryškinašamos, t.y. padaromos matomomis, tik tos linijos bei laukeliai, kurie atitinka išnagrinėto sakinio struktūrą. Į laukelį įrašoma tokia informacija: pats žodis talpinamas laukelio viduje, o jo sintaksinė funkcija įrašoma į laukelio žymės vietą. Išnagrinėto sakinio (*Jo vadovėlius, mokymo priemones, praktinių patarimų leidinius rasime ne tik tolimiausiose šalies mokyklose, aukštųjų mokyklų bibliotekose, bet Valstybinėje kalbos komisijoje, pas Seimo narius ir Vyriausybės vadovus.*) pavyzdys, t.y. jo struktūros grafinis vaizdas pateikiamas 67 pav. (psl. 98). Tai 9-tas sakinyš iš tekstyno Nr. 3003. Šiam sakiniui būdinga tai, kad jis yra eliptinis, t.y. neturi veiksnio. Šis sakinyš turi du vienarūšius žodžius (*vadovus* ir *narius*), einančius su tuo pačiu prielinksniu *pas*, kuris prie antro žodžio praleistas. Pavyzdyje matyti, kad programa atstato praleistą prielinksnį ir struktūroje nurodydama ryšius tarp žodžių pateikia abu linksnius su prielinksniais.



66 pav. Sakinio struktūros grafinio vaizdo kontūrai antrame lange.

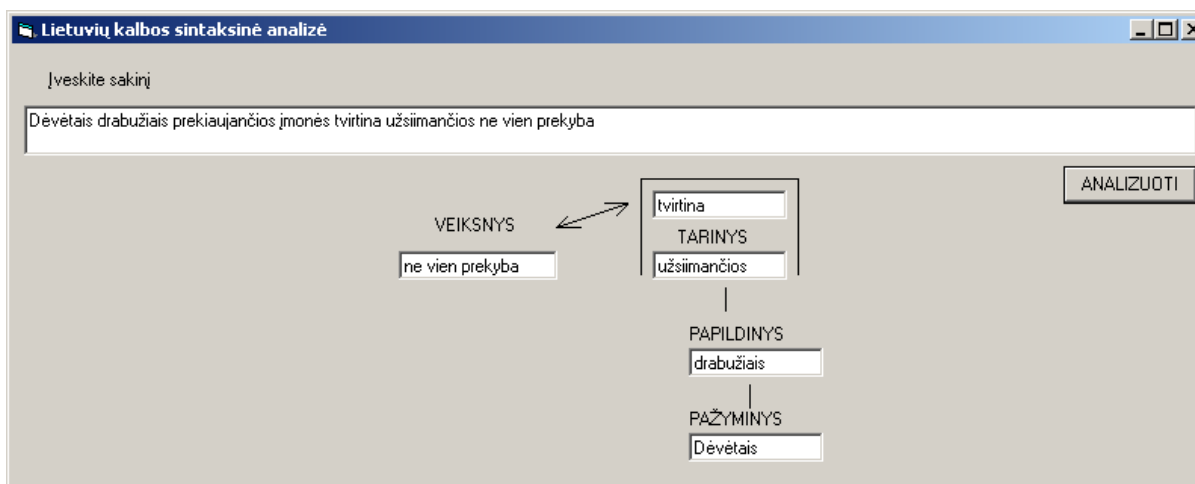


67 pav. Išnagrinėto sakinio pavyzdys.

## 7.2 Iteracinio metodo panaudojimas

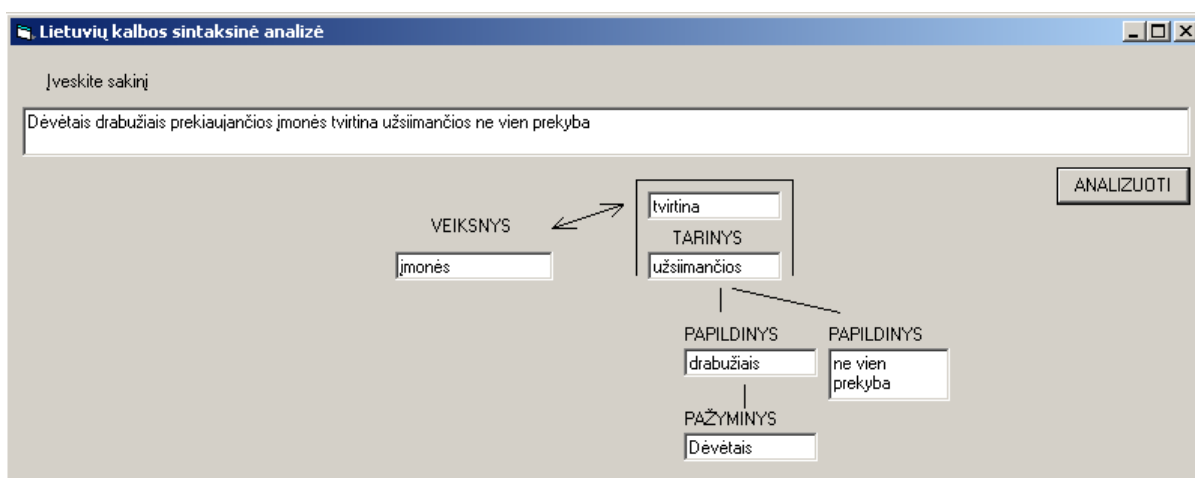
Sintaksinės analizės rezultatų pagerinimui naudojamas iteracinis metodas. Sintaksinės analizės programa, gavusi neigiamą analizės rezultatą (t.y. kai sakinyje lieka neišnagrinėtų žodžių), bando taisyti savo klaidas – ieško kito galimo morfologinio varianto tiems žodžiams, kuriems nebuvo surasti sintaksiniai ryšiai, ir analizuoja sakinį dar kartą. Kaip pavyzdį galima pateikti sakinio *Dėvėtais drabužiais prekiaujančios įmonės tvirtina užsiimančios ne vien prekyba* analizės detalią eigą. Šis sakinytis paimtas iš tekstyno Nr.1107, sakinio eilės numeris – 45.

Morfologinė analizė pirmiau pateikia vienaskaitą o po to daugiskaitą; linksnius – tokia tvarka, taip kaip jie išdėstyti lietuvių kalbos gramatikose, t.y. pirmiau vardininką, paskui kilmininką ir kitus linksnius iš eilės. Taigi šiame sakinyje žodžiui *įmonės* buvo nurodytas vienaskaitos kilmininkas, o žodžiui *prekyba* – vardininkas, todėl būtent jam buvo priskirta veiksnio sintaksinė funkcija ir gauta sakinio struktūra, kuri parodyta 68 pav.



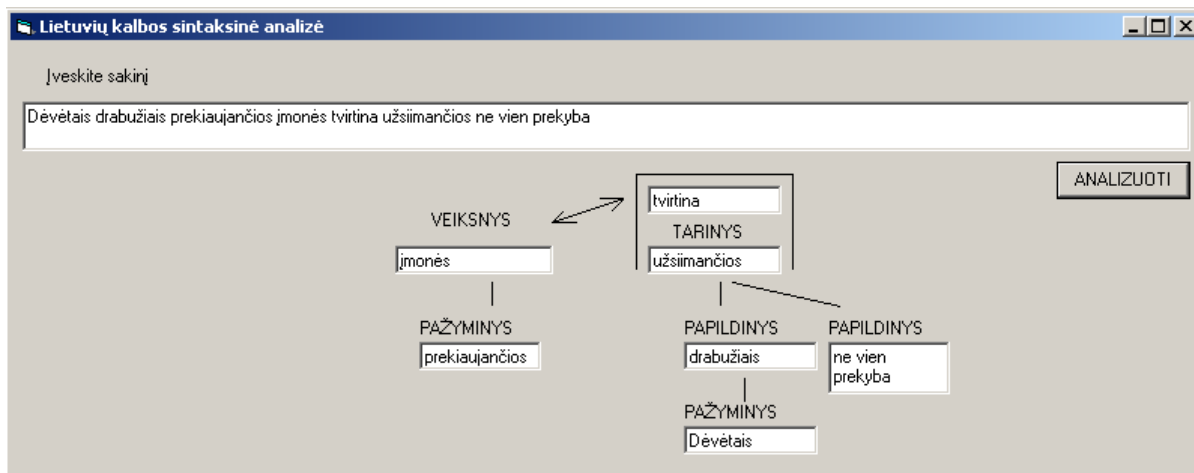
68 pav. Sakinio struktūra, gauta naudojant pirmą morfologinės informacijos variantą.

Bet lieka neišnagrinėtas žodis *įmonės*. Tada sakinytis perduodamas iteracijų paprogramei ir imamas kitas šio žodžio morfologinis variantas – daugiskaitos vardininkas. Taigi žodis *įmonės* įrašomas į veiksnio vietą ir imamas kitas žodžio *prekyba* variantas – įnagininkas. Kadangi tarinys (*užsiimančios*) stipriai valdo įnagininką, žodis *prekyba* padaromas papildiniu (69 pav.).

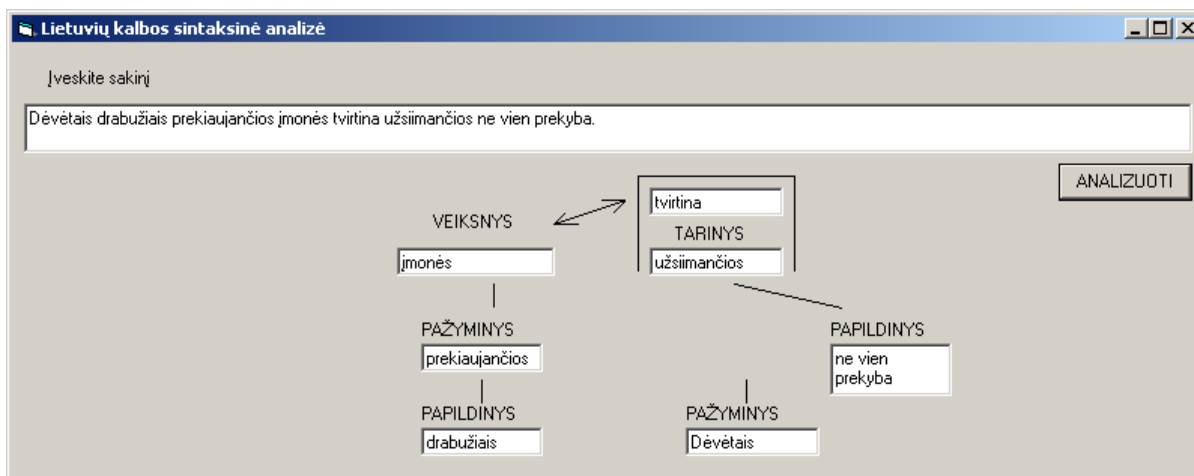


69 pav. Sakinio struktūra, atstačius teisingą veiksnį.

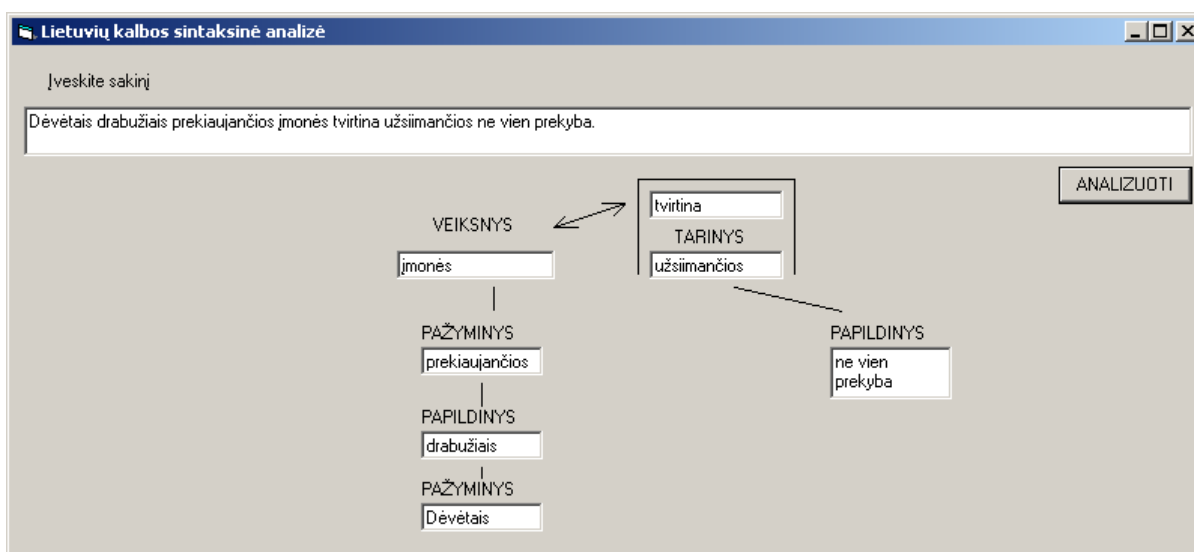
Toliau vėl pradedama nagrinėti sakinį nuo pradžių, t.y. ieškoti pagrindines sakinio dalis (veiksni ir tarinį) išplečiančių žodžių. Taip surandamas naujo veiksnio pažymins *prekiaujančios* (70 pav.), o vėliau ir jo išplėtimas *dėvėtais drabužiais* (71 pav. ir 72 pav.). Taip gaunama teisinga sakinio sintaksinė struktūra.



70 pav. Sakinio struktūra, nustačius veiksnio pažyminį.



71 pav. Sakinio struktūra, suradus veiksnio pažyminio išplėtimą.



72 pav. Galutinė teisingai išnagrinėto sakinio struktūra.

### 7.3 Skyriaus išvados

- Parašyta sintaksinės analizės algoritmą realizuojanti programinė įranga.
- Programinei įrangai sukurti buvo naudojama Visual Basic'6 programavimo kalba, kuri priklauso objektinio programavimo kalbų grupei, ir Vytauto Zinkevičiaus sukurta lietuvių kalbos morfologinių modelių programine įranga.
- Aprašytas iteracijų metodo panaudojimas sintaksinės analizės rezultatų pagerinimui.
- Programą sudaro 41543 eilutės Visual Basic'6 kalboje.
- Programa susideda iš valdančiosios (įvykio apdorojimo) paprogramės ir 31 bendrosios paprogramės.
- Panaudotos 7-ios Visual Basic funkcijos. Iškilusios problemos pavaizduojant sakinį ekrane rodo, kad ruošiant komercinį variantą reiktų naudoti naujesnę programavimo kalbą (pvz., C++).

## 8 SINTAKSINĖS ANALIZĖS EKSPERIMENTAS

Sukurtos sistemos efektyvumui patikrinti naudojami testai. Pirmosioms sintaksinės analizės sistemoms paprastai būdavo sudaromi dviejų tipų testai [Марчук, psl. 91]:

- 1) sakiniai, paimti iš publikacijų,
- 2) specialiai sudaryti sakiniai, atsižvelgiant į algoritmo darbą.

Realūs tekstai geriau parodo programos tinkamumą praktiniam naudojimui, tačiau jie turi vieną trūkumą: jie išsamiai nepatikrina algoritmo, t.y. jie gali praeiti ne visomis algoritmo šakomis. Pvz., jei numatytas 11 prielinksnių apdorojimas, o iš publikacijų paimtuose sakiniuose jų pasitaikys tik 6, reiškia 5 algoritmo šakos liks nepatikrintos ir, jei jose bus klaidų, tos klaidos liks nepastebėtos. Specialiai sudaryti testai apima visus numatytus algoritme teksto apdorojimo atvejus, t.y. jie praeina visais algoritmo keliais. Tai žinoma nereiškia, kad specialiai sudarytas testas visada duos teigiamą rezultatą. Jei sistema dirba blogai, ji ir specialiai sudarytą sakinių išnagrinės neteisingai. Realūs tekstai duoda tikrai bendrą įspūdį apie sistemos atitikimą gyvenimo reikmėms. Specialiai sudaryti testai išsamiau patikrina sistemos darbą.

Šiuo metu Internete pateiktame tekстыne visus žodžius galima rasti pavartotus publikacijose. Todėl šiam darbui testai buvo sudaryti kitokiu principu, ir pats eksperimentas buvo išskaidytas į du etapus.

### 8.1 Eksperimento pirmasis etapas

Įvertinant tai, kad lietuvių kalba yra labai sudėtinga morfologiniu požiūriu, ir pradinio etapo metu paruošti programą, atliekančią absoliučiai bet kokio lietuviško sakinio analizę yra labai sunku, todėl metodikos patikrinimui paruoštoje programoje pradžia įvedami tam tikri testinių sakinių apribojimai. Jų sąrašas pateikiamas 73 pav. psl. 103. Įvedamus apribojimus galima suskirstyti į du tipus:

1. apribojimai, kuriuos sąlygoja nepakankamas lietuvių kalbos kompiuterizavimo lygis,
2. apribojimai, kurie leidžia susiaurinti veikimo sritį (t.y. apdorojamų sakinių įvairovę), kad būtų lengviau praktiškai realizuoti programinę įrangą.

Pirmo tipo apribojimu siekiama atmesti tas klaidas, kurias iš anksto galima numatyti. Tai tokios klaidos, kurių atsiradimą nulemia automatinės semantinės analizės nebuvimas. Pavyzdžiui, kuo eina padalyvis sakiniuose *Mačiau skrendant paukštį danguje* ir *Mačiau lyjant paukštį danguje*, galima pasakyti tik semantinės informacijos pagalba. Formaliai šie sakiniai nesiskiria niekuo ir tik nuo žodžio reikšmės priklauso, kad *skrendant* eina papildiniu, o *lyjant* – aplinkybe. Todėl siekiant išvengti tokių klaidų, kurios iš anksto žinomos, ir geriau atskleisti pačios programos veikimą, t.y. paliekant tokias klaidas, kurių numatyti negalima, į apribojimų sąrašą įtraukiamos kai kurios kalbos dalys: padalyvis, pusdalyvis. Tai yra 7 apribojimas (73 pav. psl. 103).

Antro tipo apribojimais (1-6 punktai 73 pav. psl. 103) siekiama supaprastinti automatinės sintaksinės analizės programinę realizaciją, nes, esant labai sudėtingai lietuvių kalbos morfologinei struktūrai, uždavinys, parengti programą, galinčią išnagrinėti bet kokį lietuvių kalbos sakinių yra labai sunkus ir kad galima būtų greičiau paruošti veikiantį pavyzdį, įvedami kai kurie apribojimai dar ir sakinio dalims..

Šio etapo metu veikiančios programos testiniam patikrinimui panaudoti sakiniai, tenkinantys šiuos reikalavimus:

1. Sistema apdoros tik vientisinius veiksmažodinius asmeninius sakinius, t.y. nenumatytas sudėtinių sujungiamųjų bei sudėtinių prijungiamųjų sakinių, elipsių ir neveikiamosios rūšies sakinių (pasyvo) apdorojimas.
2. Neįvertinama inversinė žodžių tvarka, t.y. neapdorojami tokie atvejai: *Namas mano buvo didelis.*
3. Sakinio dalys gali būti tik vientisinės ir pažyminiai – tik atributiniai.
4. Neįvertinamas daiktavardžių, būdvardžių ir dalyvių valdymas
5. Nenumatytas apdorojimas tokių kalbos dalių kaip prielinksnis ir išiktukas.
6. Sakinio viduje neturi būti skaitmenų, žodžių kabutėse, tarptautinės rašybos žodžių (*taksi, bolero* ir pan.), archaizmų, priedėlių, įterpinių, kreipinių.
7. Apdorojami tik tie dalyviai, kurie eina pažyminiu, nes „*jų aplinkybinės reikšmės neturi formalių požymių, ir visų pirma priklauso nuo sakinio leksinės sudėties*“ (Ulvydas, psl. 527), taip pat nebus apdorojami padalyviai bei pusdalyviai, nes nėra galimybės kompiuteriu vienareikšmiškai nustatyti jų sintaksinę funkciją, kol neparuošta lietuvių kalbos automatinė semantinė analizė.

**73 pav.** Programinei įrangai taikomų apribojimų sąrašas.

Tačiau šie apribojimai jokių būdu netaikomi visai metodikai. Darbe pateikta sintaksinės analizės metodika apima bet kokią vientisinę lietuvių kalbos sakinį. Kad programa galėtų apdoroti visus vientisinius lietuvių kalbos sakinius, reikėtų remiantis ta pačia metodika parašyti dar papildomas BNF taisykles ir jas atitinkantį programos kodą.

### 8.1.1 Testų tipai

Kadangi tekстыne nenumatyta galimybė atlikti paiešką pagal sintaksinius požymius, todėl surasti sakinius, kurie atitiktų programai taikomus apribojimus, teko rankiniu būdu, taigi negalima pasakyti, kad buvo atrinkti patys tinkamiausi sakiniai. Paieška buvo vykdoma pagal žodžio formą ir žiūrima, kurie iš pateiktų 200 sakinių atitinka apribojimus, taigi iš tekstyno paimti sakiniai sudaro atskiras grupes, kuriose pasikartoja koks nors vienas žodis. Dėl to testas darosi monotoniškas.

Sintaksinės analizės programos darbo patikrinimui buvo sudaryti trijų rūšių testai ir naudotos šešios imtys:

- a) pirmam testui sakiniai buvo renkami iš visų galimų tekstų; taigi, pirmą imtį sudaro 100 sakinių: 50 sakinių paimta iš tekstyno (jie buvo atrenkami iš konkordansų, kurių pagalba galima gauti reikalingą informaciją, vertingus pavyzdžius [Marcinkevičienė 2002, psl. 18]), o 50 sakinių – iš „Lietuvių kalbos gramatikos“ III tomo [Ulvydas], nes ten yra specialiai lituanistų atrinkti sakiniai iš publikuotų leidinių, kurie gerai iliustruoja atskiras sakinio dalis.
- b) antrą testą sudaro keturios imtys po 100 sakinių. Pirmos dvi imtys atrinktos padalinus tekstyną į dvi nepersidengiančias dalis, jų struktūra pateikiama 74 pav.:

I imtis	II imtis
Respublikinė periodika	Specializuota periodika
Vietinė periodika	Grožinė literatūra (knygos)
Populiarioji periodika	Negrožinė literatūra (knygos)
	Memuarai

**74 pav.** Tekstyno dalių sąrašas kiekvienai iš pirmų dviejų imčių.

Trečią ir ketvirtą šio testo imtis (taip pat po 100 sakinių) sudaro sakiniai iš spausdintų leidinių kurie taip pat suskirstyti į dvi nepersidengiančias grupes: knygas ir periodiką.

- c) trečias testas buvo paruoštas naudojantis sakinių modeliais. Litanistai yra sudarę 21 modelį lietuvių kalbos asmeniniams veiksmažodiniams sakiniams [Ambrazas, psl. 589-603]. Kadangi programoje nenumatytas prielinksnių konstrukcijų apdorojimas, todėl į testą neįtraukiami 4 modeliai, turintys prielinksnines konstrukcijas. Kiekvienam iš likusių 17-kos modelių buvo paimta po 10 sakinių ir tokiu būdu sudarytas 170-ties sakinių testas. Sakiniai atrinkti iš gramatikose pateikiamų pavyzdžių ir iš N. Sližienės „Veiksmažodžių junglumo žodyno“. Šis testas ypatingai gerai atspindi visą lietuvių kalbos struktūrą, nes modeliai kiekvienoje kalboje yra skirtingi. Modelį sudaro sakiniui gramatiškai būtinos žodžių formos, kurias nulemia veiksmažodžio (tarinio) valentingumas. Modelių skaičius kiekvienoje kalboje yra ribotas, bet pagal juos galima sudaryti neapibrėžtą ir neribotą kiekį konkrečių sakinių – tiek teapimančių modelyje nurodytas žodžių formas, tiek papildytų kitais, sakinio sandarai nebūtinais žodžiais.

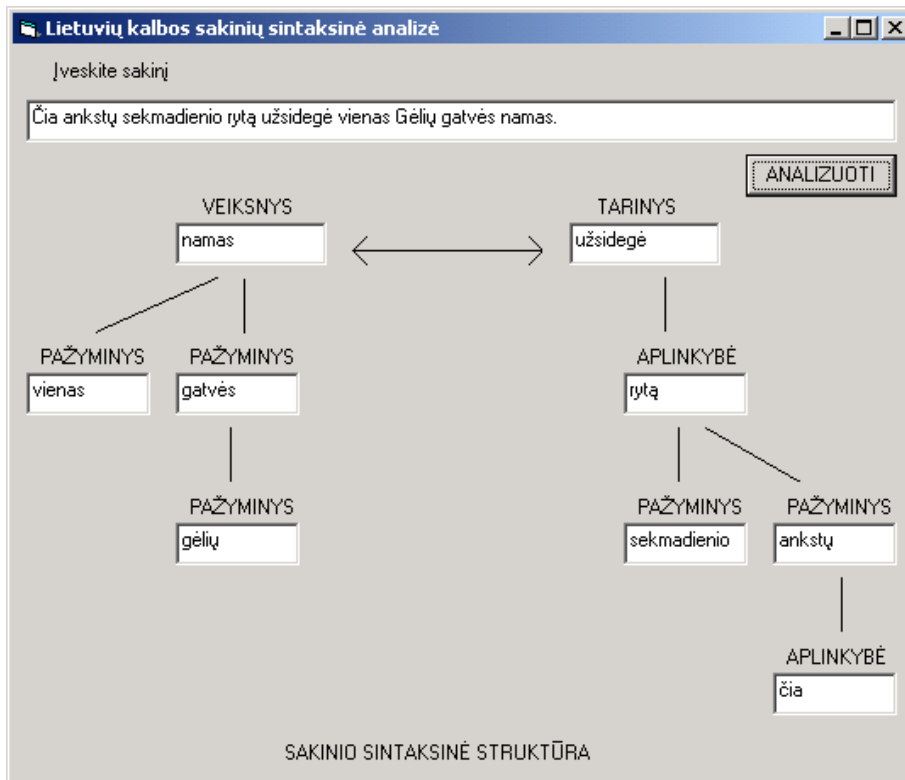
Tokiu būdu buvo gautas programos testas, susidedantis iš 670 sakinių – viena imtis (100 sakinių) apima visas teksto rūšis; keturios nepersidengiančios imtys po 100 sakinių (400 sakinių) sudarytos skaidant tekstyną į dvi dalis pagal tematiką ir spausdintus leidinius – į knygas bei periodiką, o viena imtis (170 sakinių) surinkta pagal sakinių modelius. Testams naudotų sakinių sąrašas pateikiamas priede A, pridedamoje prie disertacijos kompaktinėje plokštelėje.

### 8.1.2 Eksperimento pirmo etapo rezultatai

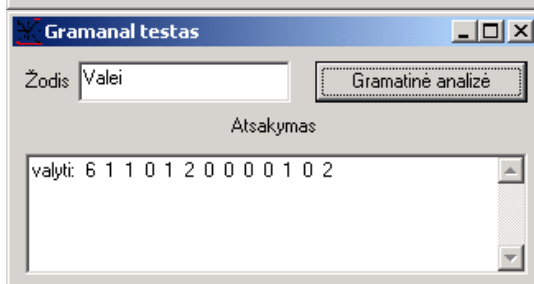
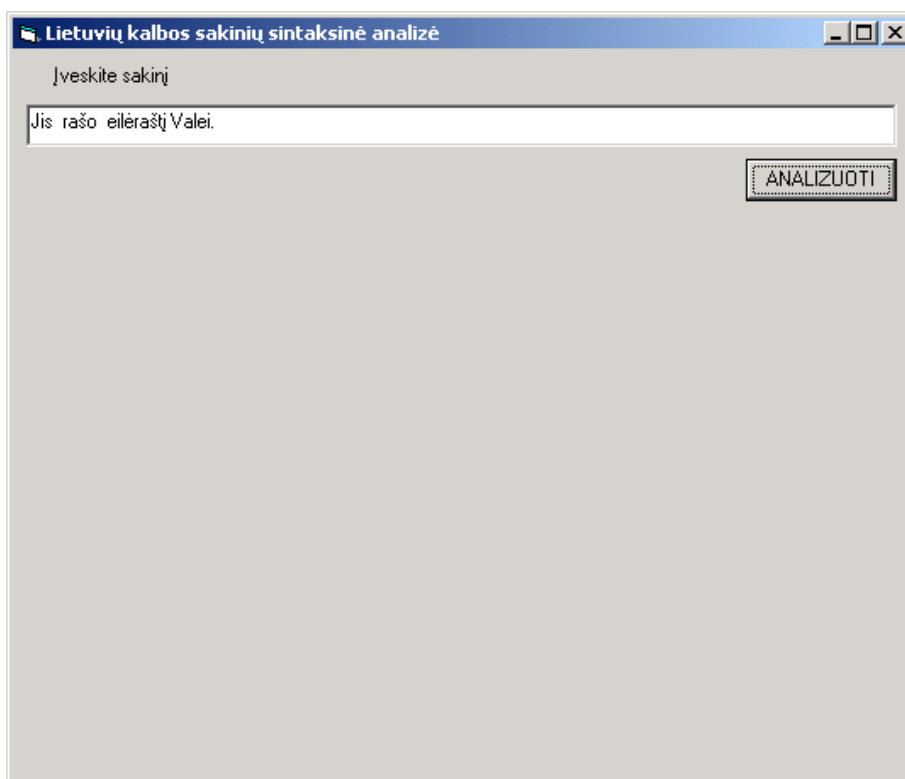
Patikrinus lietuvių kalbos automatinės sintaksinės analizės darbą, buvo gauti tokie rezultatai: iš 670 sakinių teisingai buvo išnagrinėta 629. Jų sintaksinės struktūros teisingumą aprobavo lituanistė, humanitarinių mokslų daktarė E. Valiulytė. 41 sakiniui sintaksinės struktūros sudarytos klaidingai. Analizės klaidas galima suskirstyti į 3 tipus:

1. Klaidos, kurios atsiranda dėl semantinės informacijos trūkumo, pvz., aplinkybinisrieveiksmis traktuojamas kaip būdvardžio, o ne tarinio aplinkybė (75 pav., psl. 105). Tokio tipo klaidos buvo 8 sakiniuose. Šitokių klaidų bus galima išvengti, kai bus sukurta lietuvių kalbos automatinė semantinė analizė, t.y. kai turėsime informaciją, pvz., apie tai, kad vietosrieveiksmis negali išplėsti laiko požymį turinčio būdvardžio ir pan.
2. Klaidos, kurios atsiranda dėl pradinių duomenų neoptimalaus pateikimo, t.y. kai morfologinė analizė pirmoje vietoje nurodo itin retai vartojamą žodžio formą. Pvz., žodžiui *žalius* (sakinyje *Plieninis viesulas nudraskė tau žalius lapus*) kaip pirma pradinės formos alternatyva nurodomas daiktavardis *žalis*. Kad tai nėra sintaksinės analizės metu atsirandanti klaida, rodo tas faktas, jog pakeitus šį žodį žodžiu *geltonus*, klaidos jau nebelieka. Analogiškai išnyksta klaidos sakinyje *Berniukas vienai mergaitei paduoda obuolį* pakeitus žodį *vienai* (kurio pirma pradinės formos alternatyva yra daiktavardis) žodžiu *dviem* ar sakinyje *Jis rašo labai paprastai žodį itin*. Aprašyti atvejai pateikiami priede B, pridedamoje prie disertacijos kompaktinėje plokštelėje. O sakiniui *Jis rašo eilėraštį Valei* programa iš viso nepateikia jokios struktūros, nes žodis *Valė* nėra įtrauktas į lietuvių kalbos morfologinius modelius, ir kaip pradinė forma šiuo atveju pateikiama tik veiksmažodžio bendratis (76 pav., psl. 105). Dėl neoptimaliu būdu pateiktų morfologinių duomenų klaidos buvo 24 sakinių struktūrose.



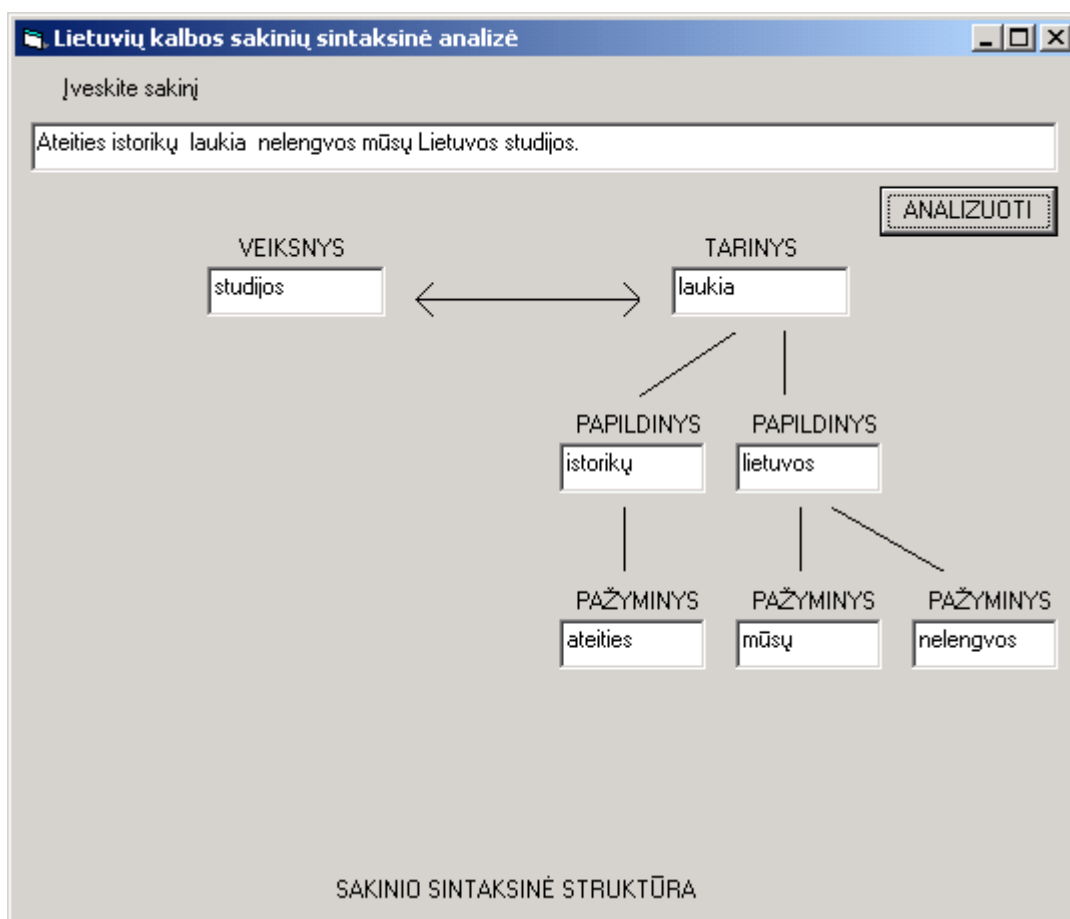


75 pav. Klaida, atsirandanti dėl semantinės informacijos trūkumo.



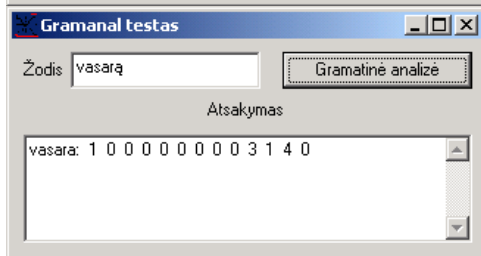
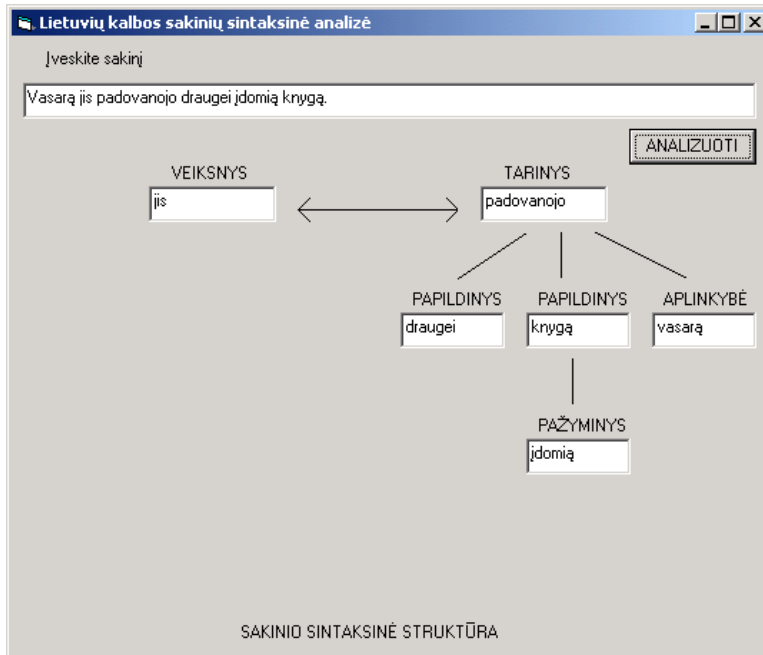
76 pav. Klaida, atsirandanti dėl neoptimalaus morfologinių duomenų pateikimo.

3. Klaidos dėl pačios sintaksinės analizės programos kaltės buvo 9 sakiniuose. Jos atsiranda tada, kai pavartojamas vienas iš homografų (vienodai rašomi, bet skirtingi savo morfologine forma žodžiai), pvz., kai sutampa daiktavardžių ar būdvardžių moteriškos giminės vienaskaitos kilmininko ir daugiskaitos vardininko formos. Tokios klaidos pavyzdys pateiktas 77 pav., sakinyje *Ateities istorikų laukia nelengvos mūsų Lietuvos studijos*. Apdorojant šnekamąją kalbą, t.y. jei sakinys analizei būtų pateikiamas ne tekstu, raidėmis, o garsais per mikrofoną ar telefoną, šios klaidos nebeliktų, nes tada pagal žodžio kirtį galima būtų vienareikšmiškai nustatyti jo formą. Kitas kelias išvengti šio tipo klaidų vėlgi lieka semantinė analizė. Turint informaciją apie tai, kad lengvos ar nelengvos gali būti studijos, o ne Lietuva, šios klaidos taip pat neliktų. Dar vienas būdas šiai problemai spręsti būtų dažninis žodžių junginių žodynas, kuris pasakytų, kad *nelengvos studijos* yra labiau tikėtinas žodžių junginys, nei *nelengva Lietuva*.

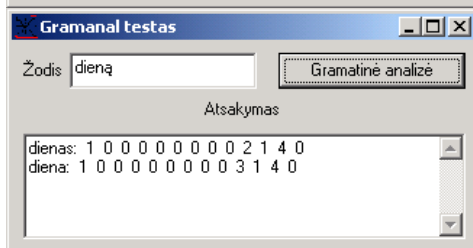
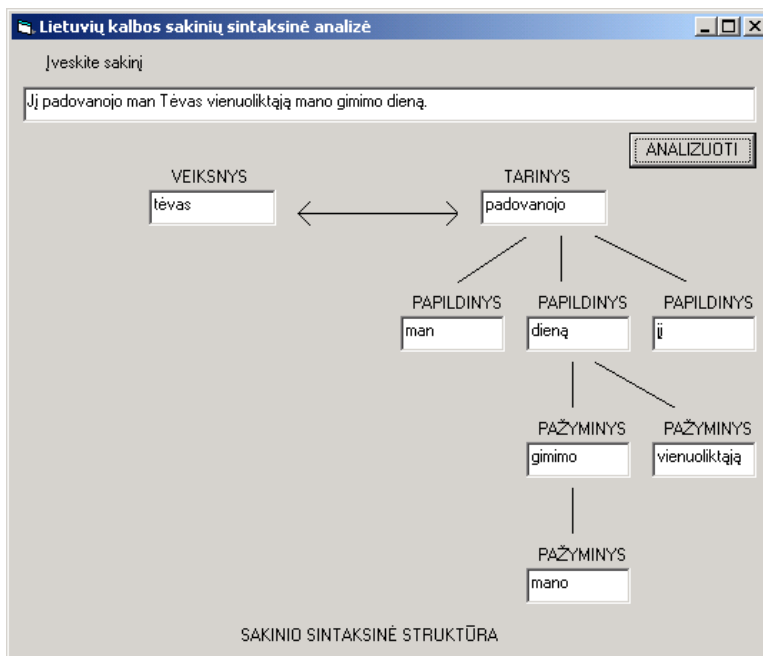


77 pav. Klaida, atsirandanti dėl sintaksinės analizės programos kaltės.

Reikia pabrėžti, kad semantiniai požymiai labai padeda sintaksinei analizei. Pvz., kai sakinyje pavartoti du daiktavardžio galininkai, kurie formaliai niekuo nesiskiria, t.y. morfologinė analizė pateikia jiems lygiai tuos pačius duomenis, tik semantikos pagalba galima nustatyti, kad vienas jų yra papildinys, o kitas aplinkybė. Tokio sakinio (*Vasarą jis padovanojo draugei įdomią knygą*) pavyzdys pateikiamas 78 pav., psl. 107. Bet semantiniai požymiai padeda išanalizuoti sakinį tik tuo atveju, jei sintaksinės analizės programa gauna teisingus pradinius duomenis iš morfologijos. Žodžiui *vasarą* pateikiamas vienintelis galimas variantas. Tačiau net ir semantiniai požymiai negali padėti teisingai išanalizuoti sakinio, jei morfologinė analizė pateikia duomenis ne pačiu geriausiu būdu. Pvz., žodžiui *dieną* kaip pirma pradinės formos alternatyva nurodomas žodis *dienas*, kuris yra chemijos terminas ir todėl jam nesuteiktas laiko požymis. Taigi sakinyje *Jį man padovanojo Tėvas vienuolikąją mano gimimo dieną* abu galininkai (daiktavardžio ir įvardžio) laikomi papildiniais (79 pav., psl. 107). Šio tipo klaidoms panaikinti reikėtų morfologinę analizę jungti su automatiniu terminų atpažinimo programomis [Zeller].



78 pav. Sakinys, kuriame semantinis laiko požymis padeda atskirti papildinį nuo aplinkybės, turinčius tas pačias morfologines kategorijas.



79 pav. Sakinys, kuriame atsiranda klaida dėl morfologinių duomenų neoptimalaus pateikimo.

Testų rezultatai pagal atskiras imtis pateikiami 7 lentelėje.

7 lentelė Testų rezultatai.

Testo Nr.	Imties kiekis sakiniams	Teisingai išnagrinėta	Klaidos dėl:			Viso klaidų	Tikslumas %
			morfo- logijos	seman- tikos	sintak- sės		
I	100	92	5	2	1	8	92
II. 1	100	92	3	0	5	8	92
II. 2	100	90	8	1	1	10	90
II. 3	100	89	5	4	2	11	89
II. 4	100	96	3	1	0	4	96
III	170	170	0	0	0	0	100
<b>Viso:</b>							
	670	629	24	8	9	41	93,88

Apibendrinant galima pasakyti, kad sistema dirba 93,88% tikslumu. Priede B, pridedamoje prie disertacijos kompaktinėje plokštelėje, duotos kai kurių teisingai išnagrinėtų sakinių struktūros. Dėl sintaksinės analizės programos kaltės klaidingai išnagrinėta 1,34% sakinių. Kitos klaidos buvo arba dėl semantinės informacijos nepakankamumo (1,2% sakinių) arba dėl morfologinės analizės netikslumų (3,58% sakinių). Jų galima būtų išvengti patobulinus morfologinę analizę, ir sukūrus lietuvių kalbos automatinę semantinę analizę. Būtent semantinės analizės automatizavimas turėtų būti dabar svarbiausias lietuvių kalbos kompiuterizavimo uždavinys.

**Reikia pabrėžti, kad ypatingai gerai įvertintas specifinis lietuvių kalbos bruožas – laisva žodžių tvarka sakinyje. Priede C (pridedamoje prie disertacijos kompaktinėje plokštelėje) pateiktas pavyzdys, kai to paties sakinio žodžiai sukeičiami vietomis įvairiais būdais ir programa vis tiek suformuoja tą pačią sintaksinę struktūrą. Ji sėkmingai išnagrinėja sakinį net ir tuo atveju, kai pirmoje vietoje yra papildinys, antroje – aplinkybė, trečioje – dar vienas papildinys, ketvirtoje – veiksnyis ir penktoje – tarinys. Sakinio su tokia žodžių tvarka negalėtų išnagrinėti anglų ar vokiečių kalbų sintaksinės analizės sistemos, nes jose iš viso nenumatytas tokių atvejų apdorojimas, kai tarinys yra toliau, nei antroje vietoje.**

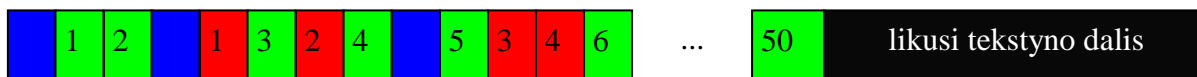
## 8.2 Eksperimento antrasis etapas

Atsižvelgiant į tai, kad pavieniai sakiniai, paimti kad ir iš labai įvairių rūšių tekstų, gali pilnai neatspindėti programos veikimo kokybės rišliame tekste, buvo atliktas antras eksperimentas, kuriame testui panaudoti eksperto (doc. G.Raškinio) atrinkti sakiniai iš 8 skirtingų “Dabartinės lietuvių kalbos tekstyno” sričių, su reikalavimu imti iš eilės visus vientisinius rišlaus (vientiso turinio arba vientisos minties) teksto sakinius, nepraleidžiant nė vieno. Iš visų aštuonių tekstyno dalių testui buvo imama po 50 pirmų vientisinių sakinių.

Dar vienas antro etapo bruožas yra tai, kad buvo nuimti visi apribojimai, t.y. sintaksinės analizės programa imasi nagrinėti bet kokį lietuvių kalbos vientisinį sakinį.

Programos veikimo sričiai tekste atspindėti buvo paskaičiuota, kurią dalį tekstyno sudaro vientisiniai sakiniai. Eksperto duotos tekstyno dalys buvo apdorotos tokiu būdu (80 pav.):

- a) Visi vientisiniai rišlaus teksto sakiniai sunumeruoti iš eilės iki 50 ir pažymėti žalia spalva,
- b) Visi sudėtiniai sakiniai, sutikti iki 50-tojo vientisinio taip pat sunumeruoti iš eilės ir pažymėti raudona spalva,
- c) Nerišlus tekstas (kuris nėra vientisos minties arba vientiso turinio tekstas) pažymėtas mėlyna,
- d) Likusi tekstyno dalis, po 50-tojo vientisinio sakinio, nebuvo tyrinėjama ir palikta juodu šriftu.



80 pav. Tekstyno apdorojimo pavyzdys.

Sužymėtos tekstyno dalys pateikiamos priede D pridedamoje prie disertacijos kompaktinėje plokštelėje.

Suskaičiavus sudėtinius sakinius iki 50-tojo vientisinio, buvo nustatytas vientisinių sakinių procentas rišliame tekste. Analogiškai buvo apdorotos minėtos tekstyno dalys ir iki 100-tojo vientisinio sakinio. Kiekvienam tekstynui atskirai paskaičiuota, kurią dalį rišlaus teksto sudaro vientisiniai sakiniai ir rezultatai pateikti 8 lentelėje.

8 lentelė Vientisinių sakinių dalis rišliame lietuvių kalbos tekste.

Bandymo Nr.	Tekstyno Nr.	iki 50-tojo vientisinio		Vientisinių % iki 50-tojo	iki 100-tojo vientisinio		Vientisinių % iki 100-tojo
		Sudėtiniai	Visi		Sudėtiniai	Visi	
1	1107	54	104	48.0	95	195	51,3
2	2038	25	75	66,7	71	171	58,5
3	3003	29	79	63.2	63	163	61,3
4	4032	43	93	54.3	88	188	53.5
5	5014	50	100	50.0	116	216	46,3
6	6009	19	69	72.5	40	140	71.4
7	8000	23	73	68.5	37	137	73.0
8	10012	52	102	49.0	112	212	47.2
<b>Viso:</b>		<b>294</b>	<b>694</b>	<b>57.6</b>	<b>622</b>	<b>1422</b>	<b>56.3</b>

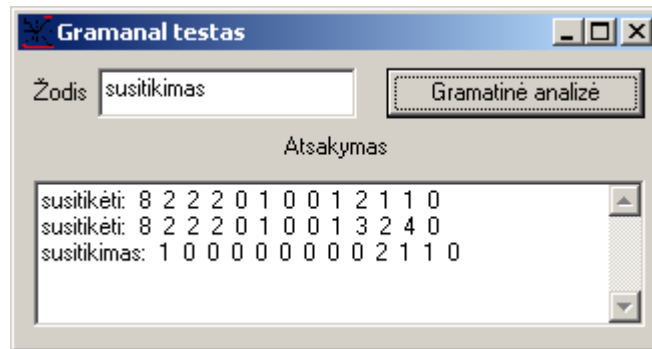
Apskaičiuojant vientisinių sakinių dalį rišliame tekste, iškilo probleminė situacija dėl vientisinio sakinio ribų nustatymo. Neaišku, kaip traktuoti sakinių tuo atveju, kai vientisinio sakinio viduje yra keli sakiniai, pvz., netiesioginės kalbos atveju, arba išvardinant ką nors punktais. Prieš punktus dedamas dvitaškis, kuris pagal gramatikos taisykles nėra sakinio pabaiga, po jo išdėstomi punktai kuriuose gali būti po kelis ir vientisinius, ir sudėtinius sakinius. Kur tada vientisinio sakinio ribos? Jei visą konstrukciją laikytume sudėtinu sakiniu ir jo nenagrinėtume, tai kaip tuomet įvykdyti reikalavimą “analizuoti visus vientisinius sakinius, nepraleidžiant nė vieno”, jeigu tos konstrukcijos viduje lieka net keli vientisiniai sakiniai.

Tai rodo, kad lietuvių kalba dar nėra pakankamai lituanistų ištyrinėta teoriniu požiūriu ir kai kuriais klausimais trūksta informacijos.

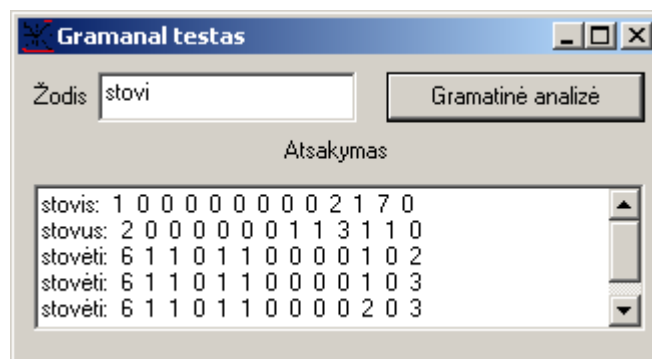
## 8.2.1 Morfologinių duomenų patikslinimas

Atliekant eksperimentą, pirmo etapo metu daugiausia klaidų buvo dėl neoptimalaus morfologinių duomenų pateikimo. Todėl antrame etape nuspręsta įvesti morfologijos korekciją, kuri eliminuotų visas klaidas atsirandančias dėl morfologijos kaltės. Taip bus geriau atspindėtas pačios sintaksinės analizės darbas. Galima išskirti dvi patikslinimo sritis:

- 1) Lietuvių kalbos morfologinių modelių programinė įranga pateikia informaciją apie žodį nurodydama jo pradinę formą ir visus galimus pateiktos formos variantus. Variantai išdėstomi gramatikose naudojama tvarka: pradžioje daiktavardis, paskui būdvardis, ir kitos kalbos dalys; pradžioje vienaskaita, po to daugiskaita, pradžioje vardininkas, po to kilmininkas ir kiti linksniai. Todėl kartais į pirmą vietą patenka itin retai vartojami žodžiai, pvz., žodžiui susitikimas kaip pirma alternatyva nurodomas dalyvio vardininkas, kurio pradinė forma yra bendratis susitikėti (81 pav.); žodžiui stovi kaip pirmas variantas pateikiamas daiktavardžio stovis šauksmininko linksnis (82 pav.), kuris formaliai gal ir gali būti lietuvių kalbos žodis, tačiau jo pavartojimą labai sunku įsivaizduoti. Šiame eksperimento etape variantai buvo sukeisti vietomis taip, kad pirmas būtų dažniausiai vartojamas atvejis.

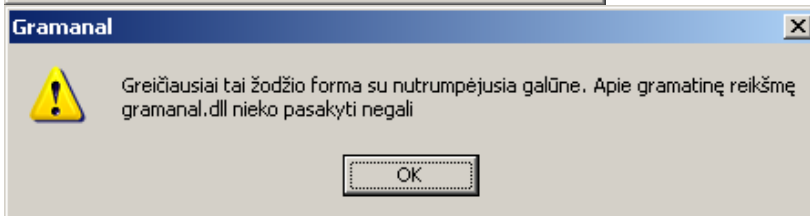
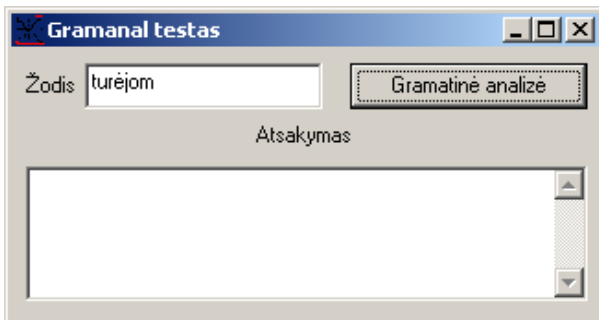


81 pav. Žodžio *susitikimas* morfologiniai duomenys.

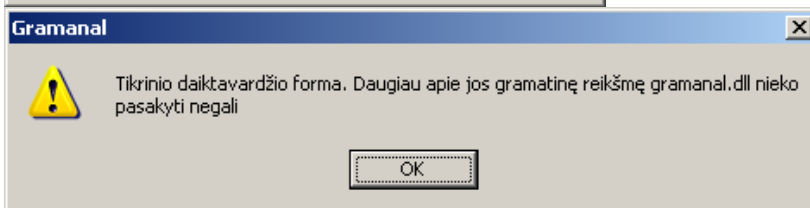
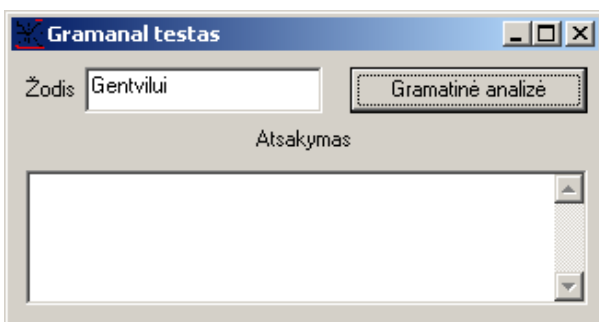


82 pav. Žodžio *stovi* morfologiniai duomenys.

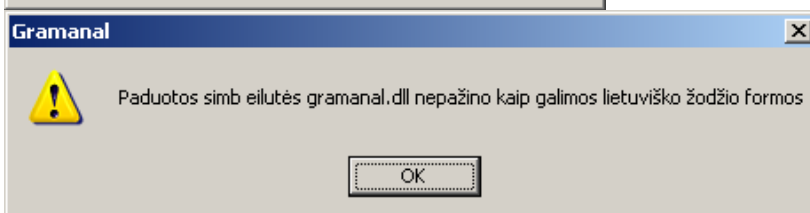
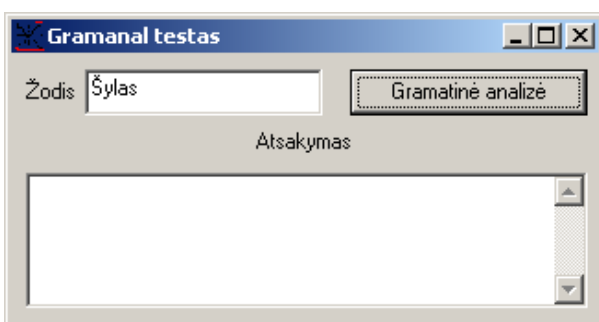
- 2) papildomai nurodomi duomenys apie tas lietuvių kalbos žodžių formas, kurių morfologinės analizės sistema iš viso neatpažįsta arba nepateikia apie jas išsamios informacijos (tai žodžiai su nutrumpėjusiomis galūnėmis, kai kurios lietuviškos pavardės ir kt.). Tokių atvejų pavyzdžiai parodyti 83 pav., psl. 111.



Tekstyno Nr. 5014 sakiny s Nr.36: *Turėjom vieną sūnų.*



Tekstyno Nr. 1107 sakiny s Nr.2: *Tą jis patvirtino vakar Mažeikių naftos perdirbimo gamykloje viešėjusiam premjerui Rolandui Paksui bei finansų ministrui Jonui Lionginui ir paskirtajam Ūkio ministerijos vadovui Eugenijui Gentvilui.*



Tekstyno Nr. 1107 sakiny s Nr.1: *Bendrovės "Mažeikių nafta" generalinis direktorius Džeimsas Šylas tikisi per 3-6 mėnesius pasirašyti ilgalaikę naftos tiekimo sutartį su Rusijos naftos bendrovėmis.*

83 pav. Žodžių, kuriems nepateikta išsami informacija pavyzdžiai.

### 8.2.2 Gauti rezultatai

Pirmi 50 vientisinių sakinių iš kiekvieno tekstyno buvo išnagrinėti panaudojant automatinės sintaksinės analizės programinę įrangą. Gautos sintaksinės struktūros surinktos į atskirus failus, kurie pateikiami priede E pridedamoje prie disertacijos kompaktinėje plokštelėje, o analizės tikslumas parodytas 9 lentelėje (psl. 113). Atlikus testą gauti tokie rezultatai: iš 400 testui naudotų sakinių teisingai buvo išnagrinėta 368. Klaidingos buvo 32 sakinių struktūros. Pasitaikiusias klaidas galima suskirstyti į tris tipus:

1. neteisingai nustatyta sakinio dalis (tokios klaidos buvo 5 sakiniuose); šio tipo klaidos pavyzdys pateikiamas 88 pav., psl. 115.
2. neteisingai surasti ryšiai tarp žodžių (tokių klaidų buvo 16-koje sakinių); šio tipo klaidų pavyzdžiai pateikiami 89 pav., psl. 115. ir 90 pav., psl. 116.
3. iš viso nesudaryta sakinio struktūra (tokios klaidos buvo 11 sakinių); šio tipo klaidos pavyzdys pateiktas 91 pav., psl. 116.

Klaidų kiekis kiekviename tekстыne pateiktas 10 lentelėje (psl. 113). Daugiausiai klaidų buvo periodikos tekstuose, mažiausiai – grožinėje literatūroje.

Teisingai išnagrinėtų sakinių pavyzdžiai pateikiami:

- 84 pav., psl. 114 – sakiny *Atsiradus daugiau erdvės vidaus rinkoje, vietos gamintojai galėtų gaminti pigesnes prekes, dažniau nukainoti turinčias mažesnę paklausą.* Tai 37 sakiny iš tekstyno Nr. 1107. Šiame sakinyje matyti, kad nuimtas apribojimas „sakinio dalys gali būti tik vientisinės“, kuris buvo taikomas sakiniams eksperimento pirmo etapo metu. Tariny su negrynąja jungtimi (*galėtų*) laikomas sudėtinu.
- 85 pav., psl. 114 – sakiny *Jis atkreipdavo dėmesį ir atsiliepdavo į Edmundo Husserlio, Moritzo Geigerio, Maxo Schelerio, Martino Heideggerio, Edithos Stein ir kitų naujausius veikalus, vos jiems pasirodžius.* Tai 6 sakiny iš tekstyno Nr. 10012. Šiam sakiniui būdinga tai, kad vienas žodis išplėstas labai dideliu kiekiu vienaarūšių sakinio dalių – papildiny *į veikalus* turi net septynis pažyminius.
- 86 pav., psl. 114 – sakiny *Šeimos, kaip santykinai autonomiško ir uždaro kultūros subjekto ir instituto, egzistavimą užtikrina jo specifinė vidinė struktūra ir funkcijos: šeimos narių saugumo, individualaus vystymosi, etnoso reprodukcijos ir kt.* Tai 13 sakiny iš tekstyno Nr. 4032. Šiame sakinyje yra du vienaarūšiai veiksniai.
- 87 pav., psl. 115 – sakiny *Apie egzistencinę filosofiją pirmasis pradėjo rašyti Jonas Repšys (1930-1976).* Tai 22 sakiny iš tekstyno Nr. 10012. Jame yra tarininis pažyminys (plačiau žr. Skyr. 4.1, psl. 36).

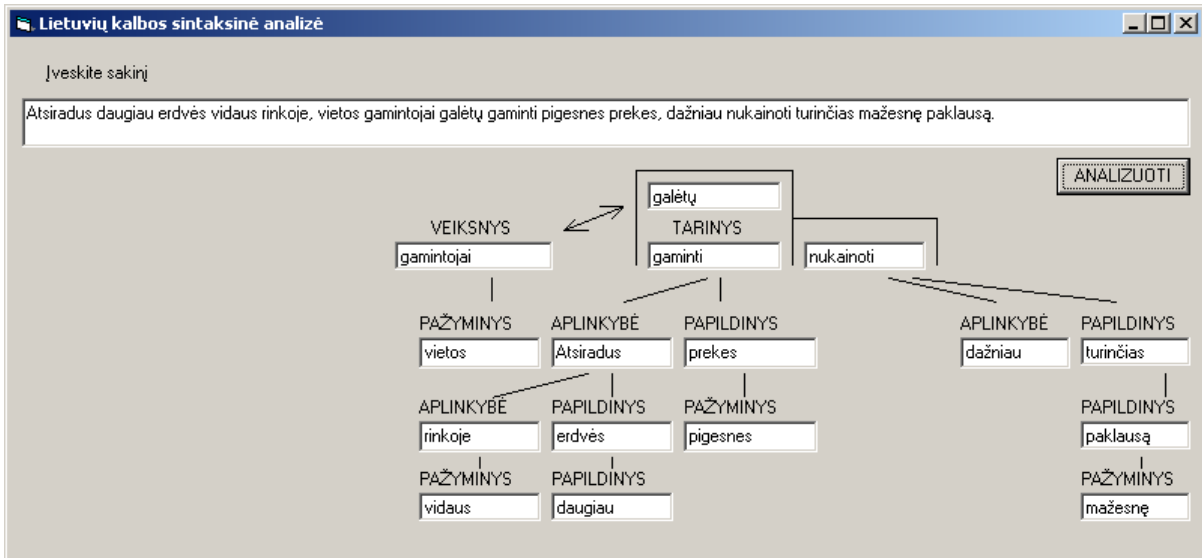


9 lentelė Automatinės sintaksinės analizės tikslumas.

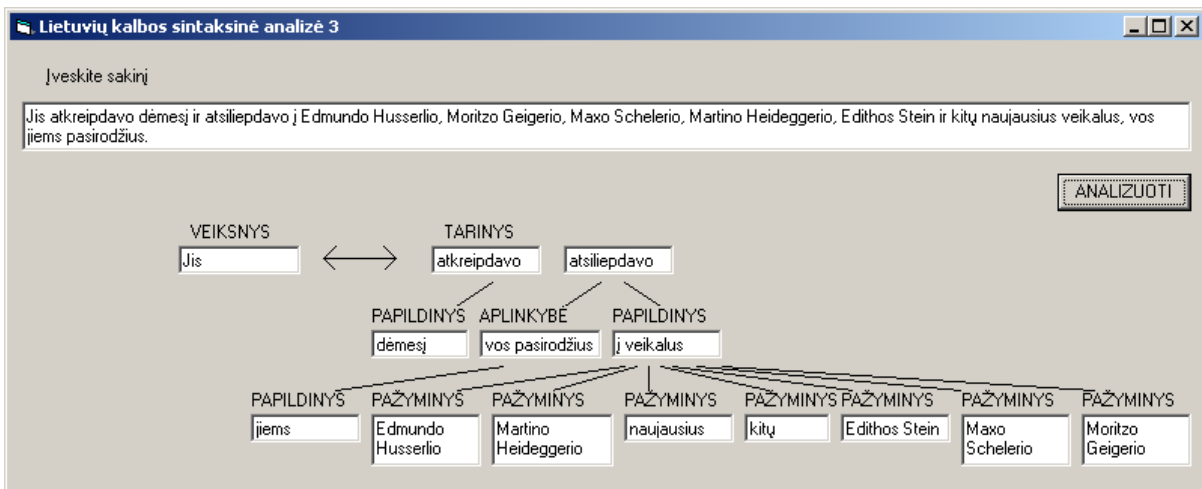
Bandymo Nr.	Tekstyno Nr. ir teksto pobūdis	Teisingai išnagrinėti sakiniai	Klaidos	Tikslumas %
1	1107 (respublikinė periodika)	46	4	92
2	2038 (vietinė periodika)	43	7	86
3	3003 (bendrojo pobūdžio periodika)	44	6	88
4	4032 (mokslinė periodika)	44	6	88
5	5014 (legendos, sakmės)	49	1	98
6	6009 (grožinė literatūra)	50	0	100
7	8000 (vyriausybės nutarimai)	47	3	94
8	10012 (filosofijos vertimai)	45	5	90
<b>Viso:</b>		<b>369</b>	<b>32</b>	<b>92.00</b>

10 lentelė Klaidų tipai.

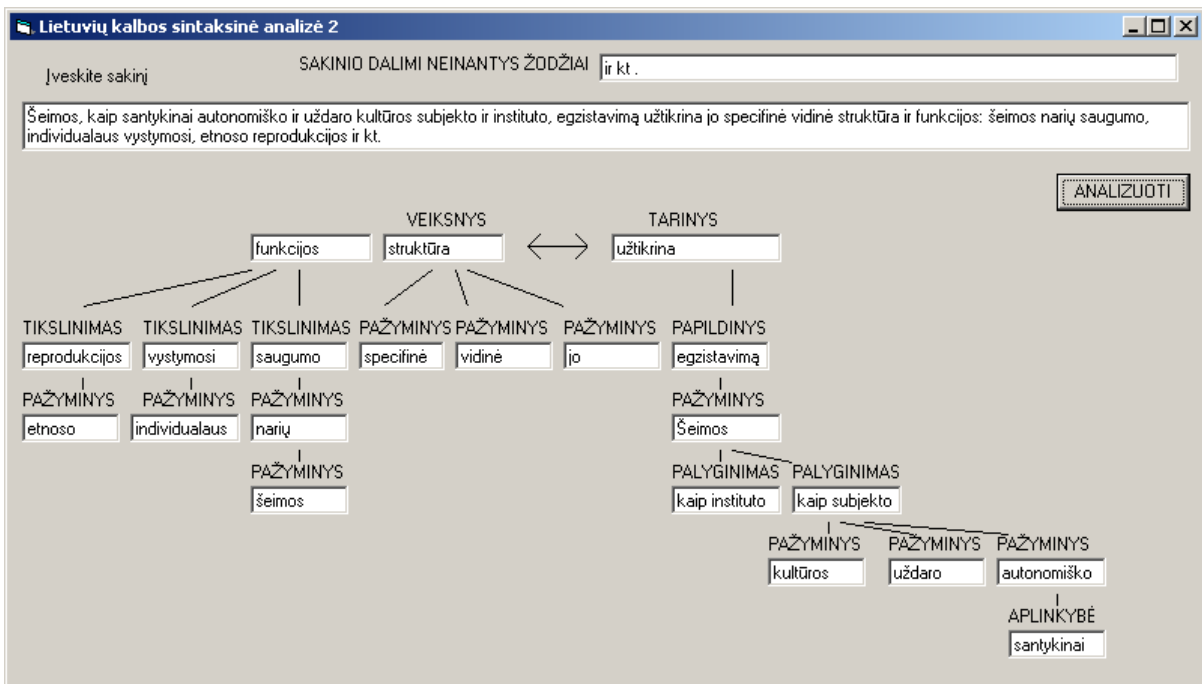
Bandymo Nr.	Tekstyno Nr. ir teksto pobūdis	Klaidos	Nepateikta struktūra	Klaidinga	
				Ryšiai	Sakinio dalis
1	1107 (respublikinė periodika)	4	0	2	2
2	2038 (vietinė periodika)	7	1	5	1
3	3003 (bendrojo pobūdžio periodika)	6	2	3	1
4	4032 (mokslinė periodika)	6	5	1	0
5	5014 (grožinė literatūra)	1	1	0	0
6	6009 (grožinė literatūra)	0	0	0	0
7	8000 (vyriausybės nutarimai)	3	1	2	0
8	10012 (filosofijos vertimai)	5	1	3	1
<b>Viso:</b>		<b>32</b>	<b>11</b>	<b>16</b>	<b>5</b>



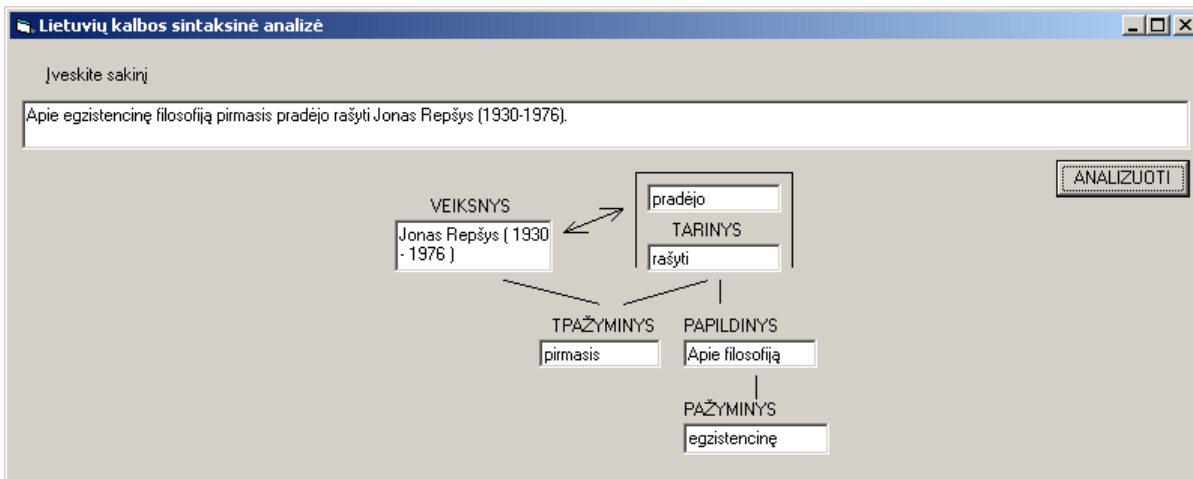
84 pav. Teisingai išnagrinėto sakinio pavyzdys (37 sakiny iš tekstyno Nr. 1107).



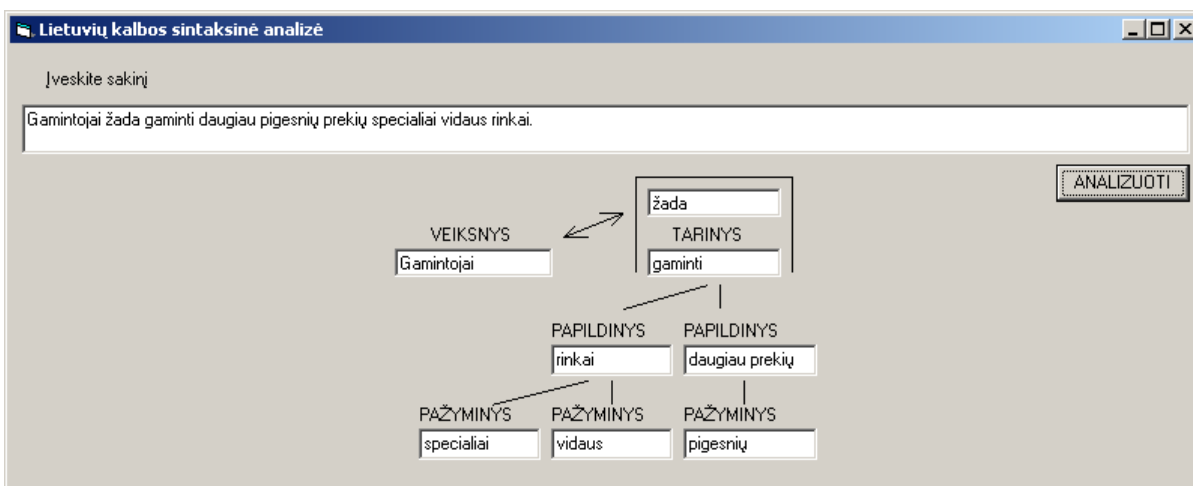
85 pav. Teisingai išnagrinėto sakinio pavyzdys (6 sakiny iš tekstyno Nr. 10012).



86 pav. Teisingai išnagrinėto sakinio pavyzdys (13 sakiny iš tekstyno Nr. 4032).

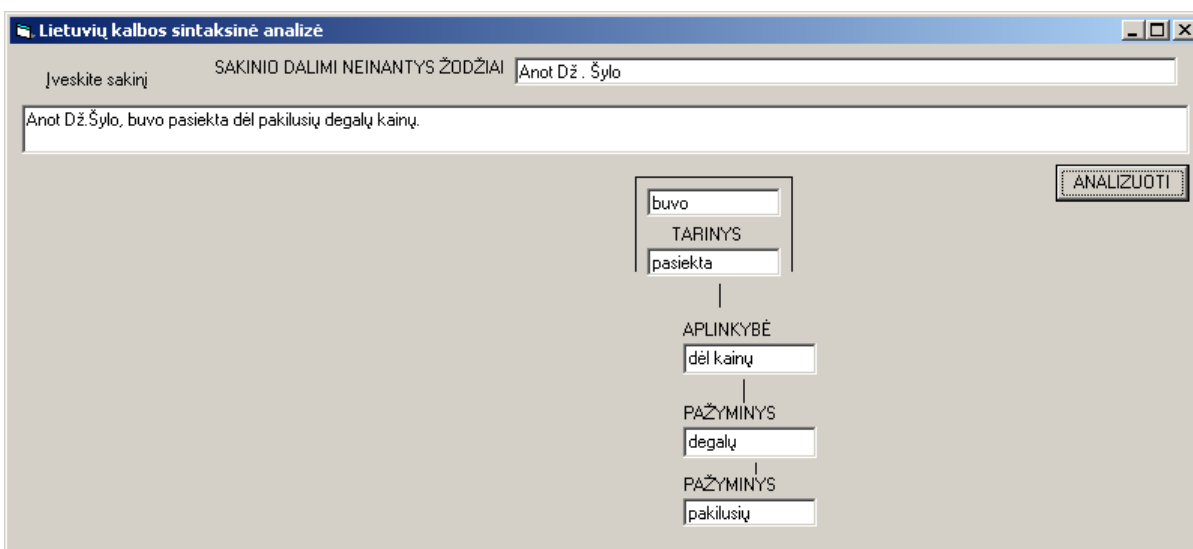


87 pav. Teisingai išnagrinėto sakinio pavyzdys (22 sakiny iš tekstyno Nr. 10012).



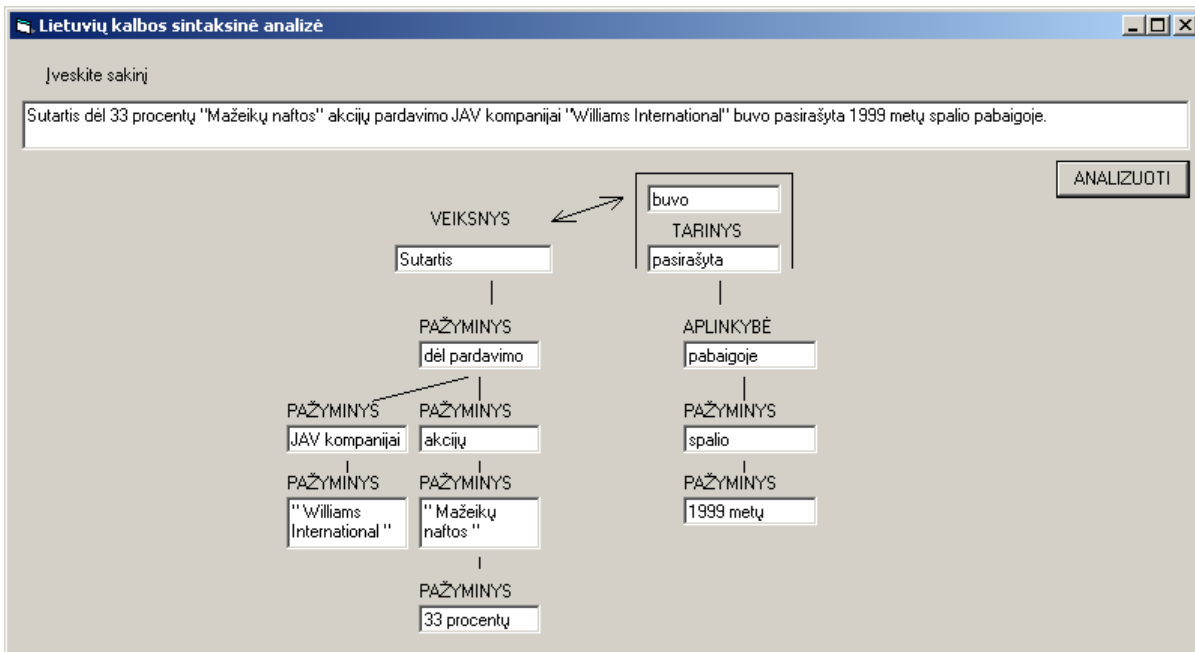
88 pav. Klaidingai išnagrinėtas sakiny (50 sakiny iš tekstyno Nr. 1107); klaidos tipas – neteisingai nustatyta sakinio dalis.

Sakinyje *Gamintojai žada gaminti daugiau pigesnių prekių specialiai vidaus rinkai* pažyminiu neteisingai laikomas žodis *specialiai*. Klaidos priežastis – sutampančios žodžių formos. Kaip pirmas žodžio *specialiai* morfologinis variantas pateikiamas būdvardžio naudininkas, todėl jis jungiamas su daiktavardžio naudininku *rinkai*. Žodis *specialiai* šiame sakinyje eina aplinkybe. Šio tipo klaidos padėtų išvengti kirčiavimas, nesrieveiksmis ir būdvardis kirčiuojami skirtingai.



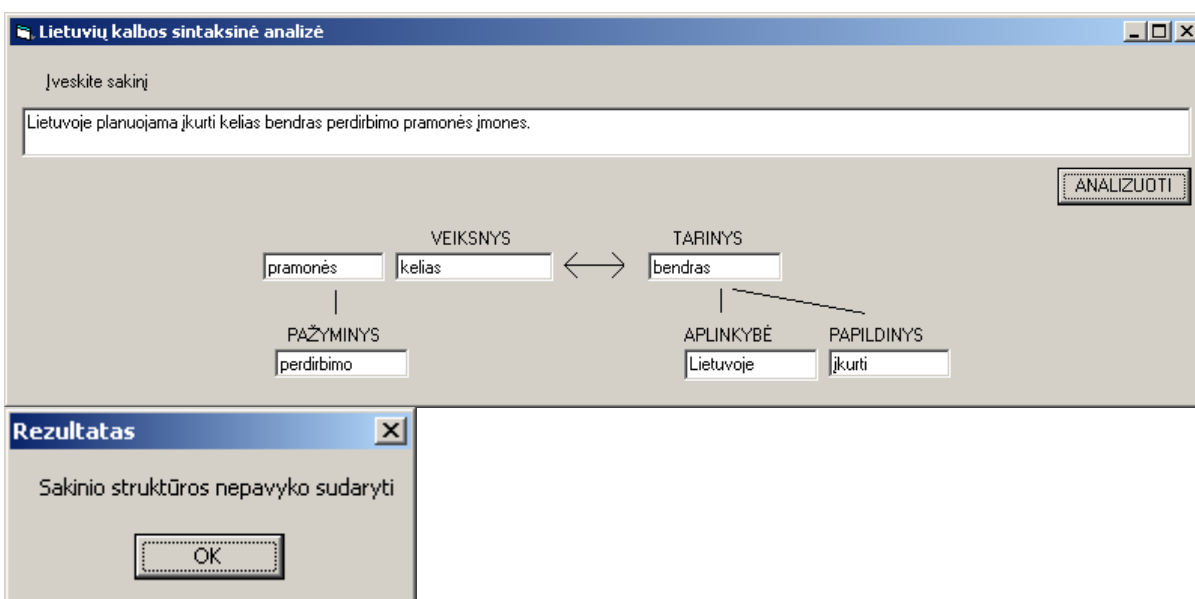
89 pav. Klaidingai išnagrinėtas sakiny (29 sakiny iš tekstyno Nr. 1107); klaidos tipas – neteisingai surasti ryšiai tarp žodžių.

Sakinyje *Anot Dž.Šylo, buvo pasiekta dėl pakilusių degalų kainų neteisingai nustatytas sintaksinis ryšys tarp žodžių pakilusių ir degalų. Klaidos priežastis – sutampančios žodžių formos. Žodis pakilusių turėtų pažymėti aplinkybę dėl kainų. Šios klaidos priežastis – sutampančios žodžių formos. Čia padėti galėtų dažninis žodžių junginių žodynas, kuriame būtų informacija apie tai, kad pakilusios kainos pasitaiko dažniau, negu pakilę degalai.*



90 pav. Klaidingai išnagrinėtas sakiny (31 sakiny iš tekstyno Nr. 1107); klaidos tipas – neteisingai surasti ryšiai tarp žodžių.

Sakinyje *Sutartis dėl 33 procentų "Mažeikių naftos" akcijų pardavimo JAV kompanijai "Williams International" buvo pasirašyta 1999 metų spalio pabaigoje pažyminys 33 procentų neteisingai susietas su "Mažeikių naftos". Klaidos priežastis – semantinės informacijos trūkumas. Turėtų būti duomenys elektronine forma apie tai, kad žodis akcijos gali sietis su procentiniu dydžiu, o pavadinimas negali turėti ryšio su procentais.*



91 pav. Klaidingai išnagrinėtas sakiny (23 sakiny iš tekstyno Nr. 2038); klaidos tipas – nesudaryta sakinio struktūra.

Sakiniui *Lietuvoje planuojama įkurti kelias bendras perdirbimo pramonės įmones* iš viso nepavyko sudaryti struktūros. Šiame sakinyje buvo keli žodžiai kuriems morfologinė informacija pateikta neatitinkanti šio sakinio varianto. Jei tai atsitinka vienam žodžiui, iteracijų paprogramė dar bando ieškoti kito varianto ir pradeda sakinio analizę iš naujo, tačiau jei viename sakinyje susikaupia daugiau žodžių su neteisingais morfologiniais duomenimis, dažniausiai sakinio išnagrinėti nebepavyksta.

Klaidų šaltinius galima išskirti į dvi rūšis:

- klaidos dėl semantinės informacijos trūkumo (dėl lietuvių kalbos automatinės semantinės analizės nebuvimo);
- klaidos dėl lietuvių kalbos morfologinių formų sutapimo.

Pirmoji priežastis lieka atviras klausimas – automatinės semantinės analizės sukūrimas laikytinas šiuo metu svarbiausiu uždaviniu atliekant lietuvių kalbos kompiuterizavimo darbus.

Automatinės semantinės analizės sukūrimas turėtų padėti panaikinti ir kitą priežastį – klaidas dėl sutampančių morfologinių formų.

Reikia pabrėžti, kad čia parodyta, kiek maksimaliai galima pasiekti lietuvių kalbos sintaksinės analizės automatizavimo srityje, šios dienos sąlygomis, t.y. turint dabartinį lietuvių kalbos kompiuterizavimo lygį. Sintaksinės analizės programinėje įrangoje papildomai morfologiškai apdorojami tie žodžiai, kuriems duomenų nepateikia lietuvių kalbos morfologinių modelių programinė įranga, tai – nelietuviškos pavardės (pvz. *Dž. Šylas* ir pan.), žodžiai su nutrumpėjusiomis galūnėmis (pvz., *turėjom* – pilna forma *turėjome*), kurių morfologinės analizės programinė įranga neatpažįsta kaip lietuviško žodžio ir kt. Be to, naudojamas semantinių požymių failas (automatinė semantinė analizė lietuvių kalbai kol kas dar nesukurta). Tačiau jis teapima tik tuos požymius, kurie tarnauja sintaksės reikmėms. Semantinių požymių faile daugiausia nurodyta: informacija apie veiksmažodžių valdymą, t.y. kokio linksnio tarinys reikalauja kaip papildinio; laiko požymis, nes jis nulemia, kuo eina sakinyje galininkas – papildiniu ar laiko aplinkybe, ir pan. Bet tai nėra išsamūs lietuvių kalbos semantinė analizė. Taigi aišku, kad analizuojant tekstus, kuriuose bus kitos nelietuviškos pavardės ar žodžių formos, tokios aukštos kokybės nebus galima gauti. Tačiau labai svarbi šios programinės įrangos savybė yra tai, kad čia paliktos visiškai atviros išplėtimo galimybės. Bet kuriuo metu į sistemą galima įtraukti naujus žodžius ar formas ir tai atlikus programa turėtų išanalizuoti sakinius apytikriai tokiu tikslumu, koks pasiektas šiame darbe.

Kad programa galėtų nagrinėti visus sakinius, būtina turėti semantinę (žodžio reikšmės) informaciją apie kiekvieną lietuvių kalbos žodį, t.y reikia sukurti semantinę lietuvių kalbos duomenų bazę. Morfologijos išplėtimo failas turi būti nuolat pildomas naujai pasirodžiusiomis pavardėmis ar pavadinimais, o informacija apie visus dabar vartojamus žodžius turi būti sukaupta lietuvių kalbos morfologinėje duomenų bazėje, kurią taip pat būtina sukurti. Ją kuriant galima pasinaudoti lietuvių kalbos morfologinių modelių programine įranga, tačiau jos duomenis reiktų papildyti tomis formomis, kurių ši programinė įranga neatpažįsta kaip lietuviško žodžio formos (kai kurios lietuviškos pavardės, nelietuviški vardai ir kt.), arba nepateikia išsamios informacijos (žodžiai su nutrumpėjusiomis galūnėmis, kai kurie tikriniai daiktavardžiai).

Sumažinti klaidų kiekį padėtų žodžių junginių žodynas. Lietuvių kalbai jis dar nėra sudarytas, taigi reiktų paruošti žodyną elektronine forma, kuriame būtų informacija apie tai, kad to paties žodžio junginys su vienu žodžiu yra labiau tikėtinas, negu su kitu.

### 8.3 Skyriaus išvados

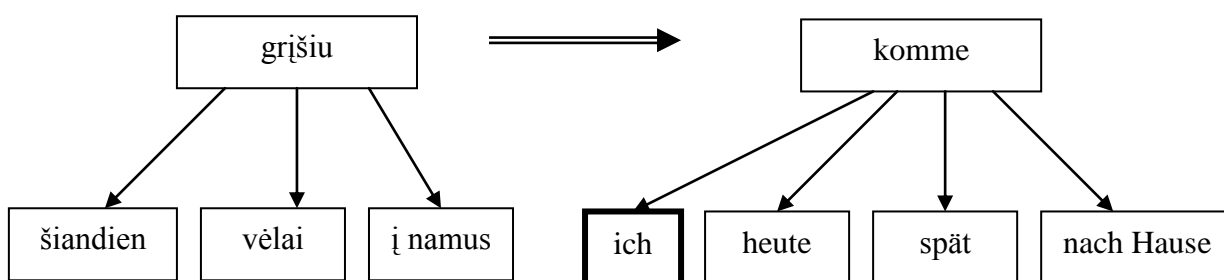
- Pirmo etapo metu gauti rezultatai rodo, kad programa dirba 93,88% tikslumu. Atskirose imtyse tikslumas svyruoja nuo 89% iki 100%.
- Antro etapo metu gauta 92% teisingų sintaksinių struktūrų.
- Klaidų analizė parodė, kad klaidų sumažinimui būtina sukurti:
  - a) lietuvių kalbos morfologinę duomenų bazę
  - b) lietuvių kalbos sintaksinę duomenų bazę
  - c) dažninių žodžių junginių žodyną

## 9 SINTAKSINĖS ANALIZĖS TAIKYMAS

Automatinė sintaksinė analizė gali būti naudojama įvairiose srityse. Jos pagalba kartais gali būti panaikinamas morfolginis daugiareikšmiškumas (žr. skyr. 6.3, psl. 86). Sintaksinę analizę galima panaudoti gramatinių klaidų tikrintuvo darbo pagerinimui, kad būtų rodomos ne tik morfolginės, bet ir sintaksinės klaidos. Dar viena automatinės sintaksinės analizės pritaikymo sritis galėtų būti paieška tekстыne pagal sakinio struktūrą ar žodžių sintaksinius požymius. Sintaksinė analizė gali būti panaudojama įvairiose dirbtinio intelekto srityse: dialogų sistemose, informacijos išgavimo iš teksto programose ir mokymo sistemose, teksto analizės sistemose ir pan. Svarbesnės sintaksinės analizės taikymo sritis vertėtų aptarti smulkiau – tai automatinis vertimas, ištisinės šnekamosios kalbos atpažinimas, klaidų suradimas tekстыne.

### 9.1 Automatinis vertimas

Vienu pagrindinių taikymų galima laikyti automatinį vertimą. Turint automatinio būdu sudarytą sakinio sintaksinę struktūrą jau galima keisti analogiška kitos kalbos struktūra. Tiesiogiai pažodžiui versti negalima, nes visų pirma, pažodinį vertimą sunku skaityti („*пословно-пооборотный перевод читается с известным напряжением*“ [Билан, Кричевич, psl. 83]). Be to, daugeliu atvejų bus gaunami gramatiškai netaisyklingi sakiniai kitose kalbose (pvz., lietuviškas sakiny *Einu namo* perkeltas pažodžiui į vokiečių kalbą *Gehe nach Hause* bus gramatiškai netaisyklingas ir vokiečių kalbos speleriai rodys sintaksines klaidas), arba jie gali būti neteisingai išversti (angliškame pažodiniame vertime gramatinių klaidų gal ir nebus, tačiau sakinio reikšmė bus perteikiama ne ta: *Go home* yra jau liepiamasis sakiny, kurio lietuviškas atitikmuo būtų *Eik namo*, o tai yra visai ne tas pats kaip *Einu namo*). Įvertinant tai, kad lietuvių kalboje pirmo ir antro asmens įvardžiai paprastai praleidžiami (ir tai daroma dažniausiai stilistiniais sumetimais, kad nesusidarytų informacijos perteklius – mes juos matome iš veiksmažodžio galūnių), o germanų kalbos netoleruoja sakinių be veiksnio ar tarinio, minėtos sakinio dalys turės būti atstatytos, atliekant automatinį vertimą į vokiečių ir anglų kalbas. Tai turės būti atliekama keitimo etape sintaksinių struktūrų lygmenyje. Pvz., sakinio *Šiandien grįšiu į namus vėlai* struktūra verčiant šį sakinį į vokiečių kalbą turėtų būti keičiama taip (92 pav.):



92 pav. Praleisto veiksnio atstatymas.

Tokių atvejų yra ir daugiau: esamojo laiko jungtis (žodis *yra*) lietuviškame sakinyje paprastai praleidžiama. Ją būtina atstatyti atliekant vertimus į anglų bei vokiečių kalbas, nes šiose kalbose sakinių be veiksmažodžio iš viso būti negali.

## 9.2 Ištinės šnekos atpažinimas

Kita automatinės sintaksinės analizės taikymo sritis – ištinės šnekos atpažinimas, t.y. kada per mikrofoną ar telefoną įvesti į kompiuterį žmonių pokalbiai verčiami tekstu (raidėmis). Šnekos atpažinimo technologijų tikslas yra sukurti mašinas kurios galėtų paimti žodžiu sakomą informaciją apdoroti [Deller, Hansen, Proakis, psl. 601]. Daugumai žmonių natūraliausias ir veiksmingiausias būdas pasikeisti informacija yra pokalbis, t.y sakinė kalba. Todėl dirbtinio intelekto tyrinėtojai siekia sukurti mašinas, kurios galėtų “girdėti”, “suprasti” ir “apdoroti” pasakytą informaciją bei “kalbėti” kad pasikeitimas informacija būtų užbaigtas. Šiuo metu veikiančios šnekos atpažinimo sistemos skirstomos į tris kategorijas:

1. Sistemos, apimančios nedidelį žodžių kiekį (~10-100 žodžių);
2. Sistemos, kuriose žodžiai tariami izoliuotai (žodynas gali siekti 10000 žodžių);
3. Ištinės šnekos atpažinimo sistemos, dirbančios ribotos tematikos srityse, pvz., įstaigų korespondencija (žodynas ~1000-5000 žodžių)

Trečios kategorijos atpažinimo sistemose naudojami lingvistiniai duomenys, tame tarpe ir sintaksinė analizė, nes šiose sistemose nebėra laiko ribų tarp akustinių signalų ir todėl žodžių atskyrimui vienas nuo kito pasitelkiama gramatika. Traktuojant sakinį kaip pagrindinį pranešimo vienetą, atpažinimo metu naudojamos žinios apie tai, kaip iš fonemų sudaromi žodžiai (leksiniai ryšiai) ir kaip iš žodžių sudaromi sakiniai (sintaksiniai ryšiai).

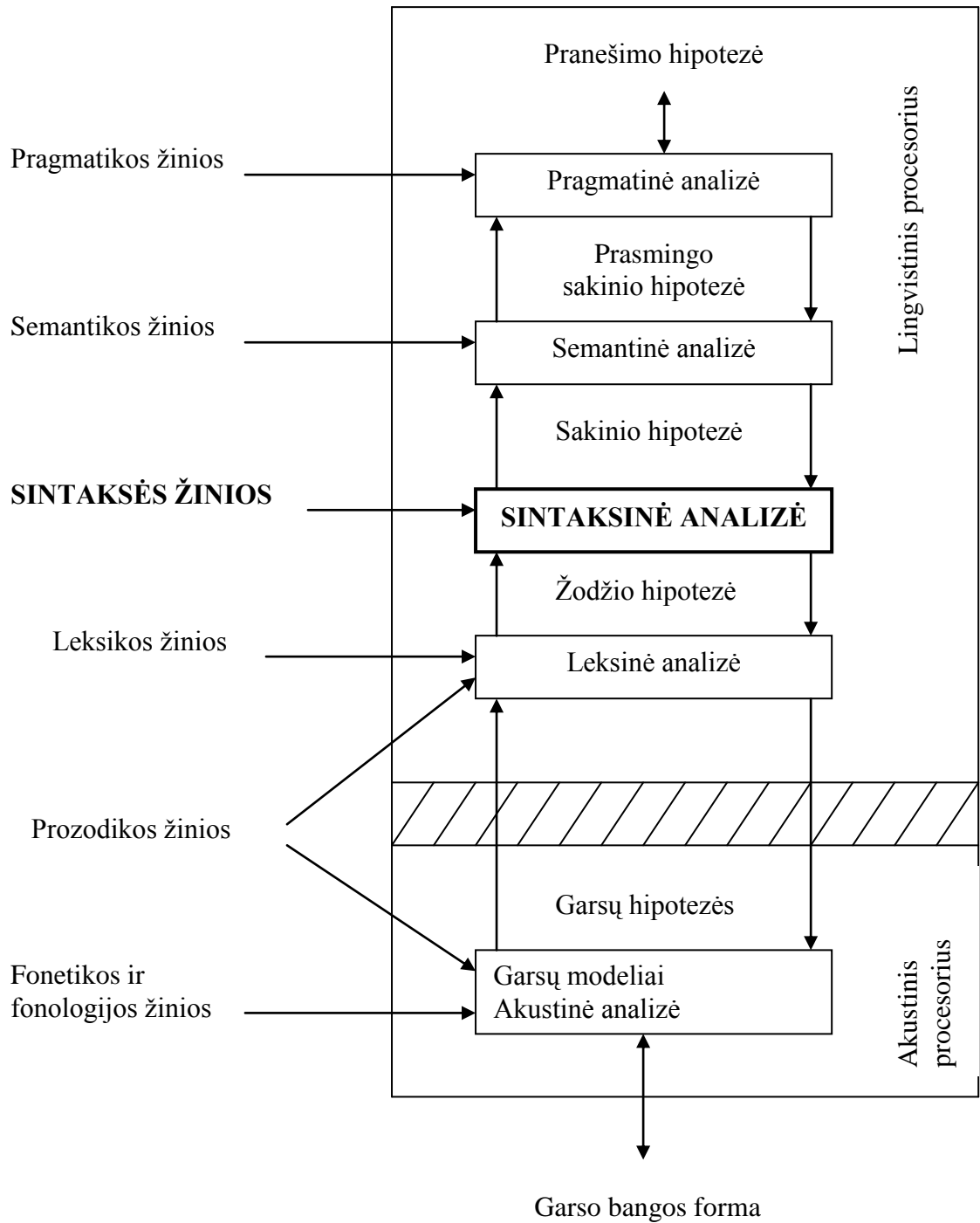
93 pav. (psl. 121) pateikiama apibendrinta ištinės šnekamosios kalbos atpažinimo blokinė schema [Deller, Hansen, Proakis, psl. 618]. Čia pavaizduoti du procesoriai: akustinis ir lingvistinis (kalbos). Kairėje pusėje pateikti žinių šaltiniai. Pavaizduota sistema naudoja garsus kaip elementarius kalbos simbolius.

Gimtąja kalba šnekantis žmogus nesąmoningai naudoja žinias apie kalbą, aplinką, kontekstą tam, kad suprastų sakinį. Žinios apima garsų savybes (fonetika), tarimo įvairovę (fonologija), kirčiavimą ir intonaciją (prozodika), žodžių garsinius modelius (žodyną), gramatinę kalbos struktūrą (sintaksę), žodžių ir sakinių reikšmę ir pokalbio kontekstą (pragmatika). Ištinės šnekamosios kalbos atpažinimo schemeje visi šie žinių šaltiniai išdėstyti hierarchine tvarka. Akustinis procesorius laikomas ta atpažinimo sistemos dalimi, kuri susieja garso bangos formą su „intelektiniu“ kalbos sektoriumi, pateikdama garso bangos formą kaip parametrus ar požymius.

Ištinės šnekamosios kalbos apibendrintoje schemeje (93 pav., psl. 121) rodyklėmis parodytos dvi „informacijos srauto“ kryptys, kurios žymi dvi pagrindines kategorijas, pagal kurias klasifikuojamos sukurtos šnekos atpažinimo sistemos. Jei akustinis procesorius naudojamas iškelti hipotezei apie įvairius žodžius, garsus ir t.t., ir tada šios hipotezės perduodamos apdoroti aukšty, lingvistiniam (kalbiniam) procesoriui, kad būtų patikrinta ar jie gali būti sujungti pagal “aukštesnio lygio” taisykles, tai tokia sistema vadinama dirbančia „iš apačios į viršų“.

Apdorojimas „iš viršaus į apačią“ prasideda nuo sakinio hipotezės, kuri iškeliamą aukščiausiam lingvistinio procesoriaus lygmenyje. Šios hipotezės tikrinamos kiekviename iš žemesnių lygmenų pagal galimumą reikšti pasakymą. Kiekvienas lygmuo, kreipdamasis į žemesnį lygmenį pateikia jam informaciją, kurią jis naudojo, atlikdamas savo įvertinimą. Į akustinį procesorių kreipiamasi, kad jis nustatytų ar akustika atitinka žemiausiam lingvistinio procesoriaus lygmenyje gautus simbolius, pvz., garsų rinkinį.





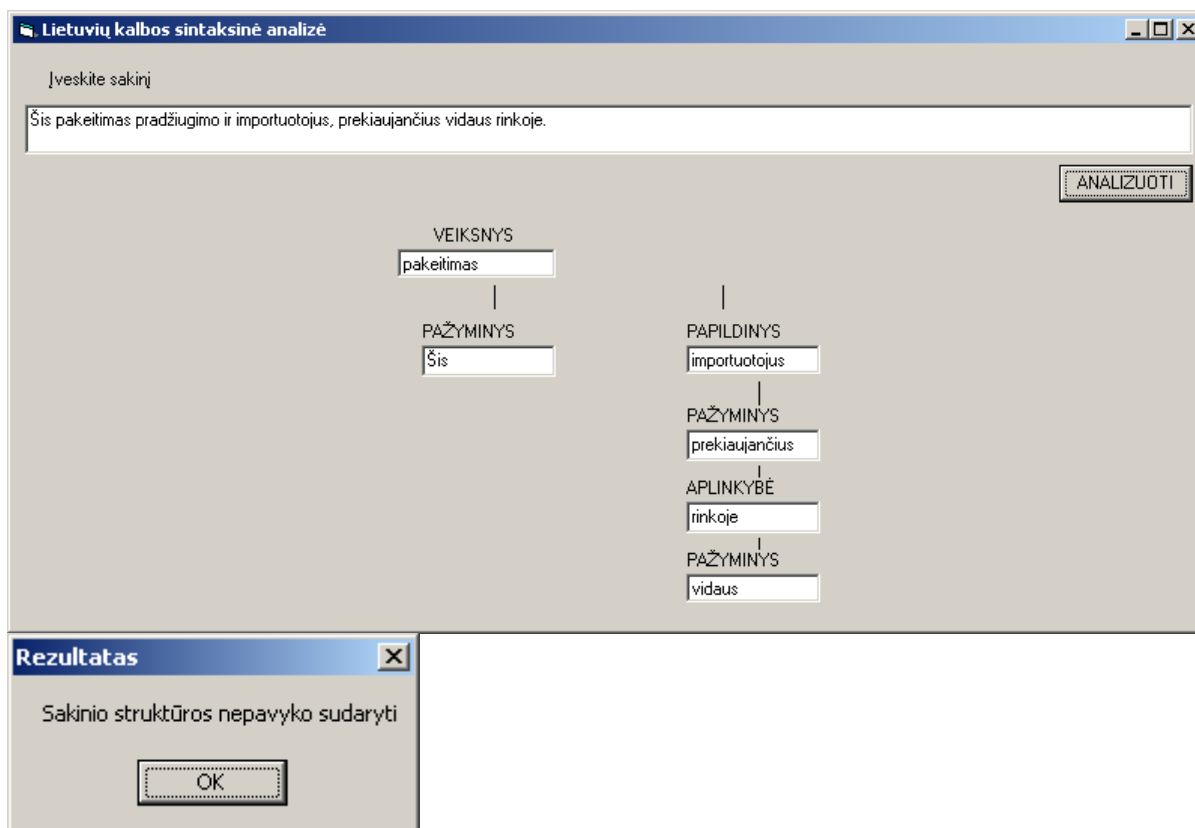
93 pav. Ištisinės šnekamosios kalbos atpažinimo apibendrinta blokinė schema.

93 pav. parodytoje ištisinės šnekos atpažinimo sistemos schemoje paryškintai nurodytos tos vietos, kuriose gali būti naudojama automatinė sintaksinė analizė.

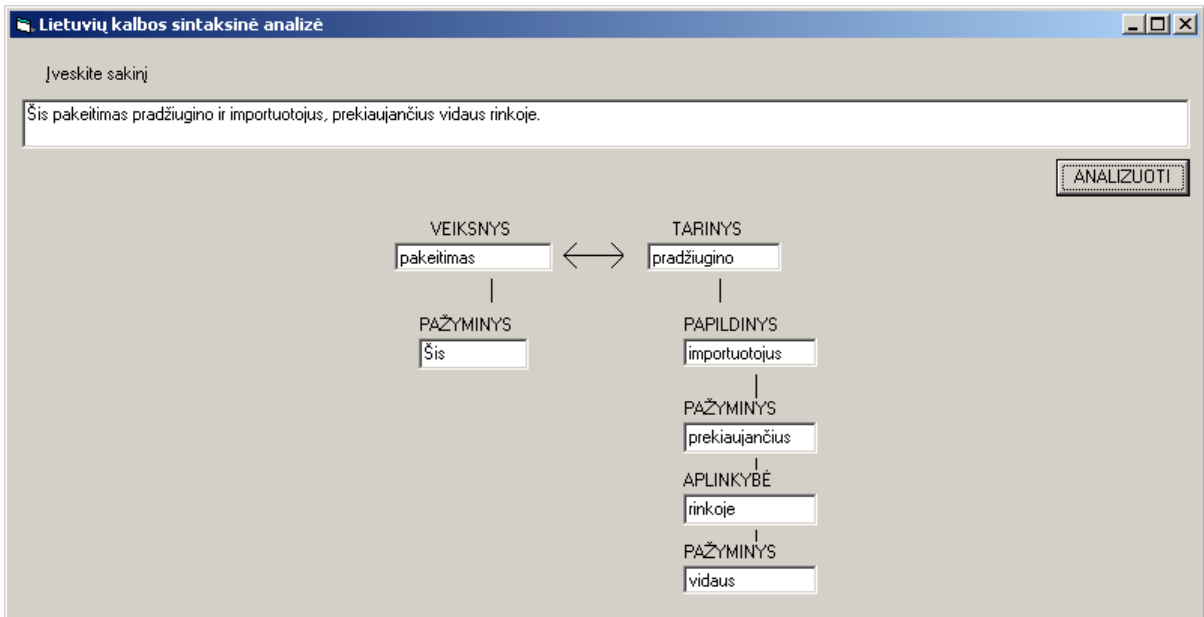
### 9.3 Klaidų suradimas tekстыne

Automatinės sintaksinės analizės programinė įranga leidžia užfiksuoti tokias klaidas, kurių neparodo morfologinės analizės sistema. Jei sakinyje neteisingai pavartotas šiaip jau leistinas lietuvių kalboje žodis, morfologinė analizė nerodo klaidos. Nagrinėjant žodį sintaksiškai, tokiu atveju paprastai gaunamas neigiamas analizės rezultatas, nes klaidingai pavartotam žodžiui nesurandami sintaksiniai ryšiai tame sakinyje. Todėl pateikiamas pranešimas, kad sakinio struktūros nepavyko sudaryti ir tuo atkreipiamas dėmesys į tą sakinį, t.y. parodoma, kad jame gali būti klaida. Kaip pavyzdžių galima pateikti iš eksperimento antrame etape naudotų sakinių. Sakinys Nr.48 iš tekstyno Nr.1107 pačiam tekстыne pateiktas taip: *Šis pakeitimas pradžiugimo ir importuotojus, prekiaujančius vidaus rinkoje.* Žodis *pradžiugimo* šiaip jau yra galimas lietuvių kalboje ir morfologinė analizė jam pateikia tokią informaciją: daiktavardis, vienaskaitos kilmininkas, pradinė forma *pradžiugimas*. Ir tik sintaksinė analizė parodo, kad šiame sakinyje gali būti klaida, t.y. šiam sakiniai ji pateikia pranešimą „Sakinio struktūros nepavyko sudaryti“ (94 pav.), nes daiktavardžio kilmininkui *pradžiugimo* nesurandamas žodis (kitas daiktavardis), kurio pažyminiu jis galėtų būti pagal savo poziciją sakinyje.

Ištaisius klaidą – žodį *pradžiugimo* pakeitus žodžiu *pradžiugino* – gaunama teisinga sakinio struktūra (95 pav., psl. 123).



94 pav. Sakinio su korektūros klaida analizė.



95 pav. Sakinio su ištaisyta klaida struktūra.

## BENDROSIOS IŠVADOS

1. Apžvelgtos anglų, vokiečių ir rusų kalbų automatinės sintaksinės analizės sistemos ir parodyta, kad lietuvių kalbos sintaksinei analizei tiesiogiai perimti jau sukurtų kitoms kalboms sistemų negalima dėl didelių skirtumų tarp kalbų.
2. Modifikuotas priklausomybių gramatikoje naudojamas sakinio sintaksinės struktūros pavaizdavimo metodas, papildant priklausomybių medį dvigubais sintaksiniais ryšiais, kuriuos turi lietuvių kalbos tarininis pažyminyš. Parodyta, kad ne visus lietuvių kalbos sakinius galima pavaizduoti priklausomybių medžiu, todėl lietuvių kalbai naudojamas priklausomybių grafas, turintis uždara kelią.
3. Pirmą kartą sudaryta formali gramatika lietuvių kalbos sintaksei aprašyti, pateikiant ją BNF (Bekaus ir Nauro forma).
4. Sukurtas naujas metodas sintaksinei analizei atlikti, kuris gerai įvertina specifinius lietuvių kalbos bruožus – didelį kaitomumą ir laisvą žodžių tvarką sakinyje.
5. Sėkmingai išspręstos sakinio dalių nustatymo bei ryšių tarp jų suradimo problemos, susijusios su laisva žodžių tvarka.
6. Eksperimentas, atspindintis lietuviškų tekstų įvairovę, parodė, kad automatinė sintaksinė analizė davė 92% tikslumą.
7. Tešiant lietuvių kalbos sintaksinės analizės automatizavimo darbus reikia:
  - a) Sukurti patogesnę programinės įrangos variantą praktinėms lietuvių kalbos vientisinių sakinių automatinės sintaksinės analizės reikmėms;
  - b) Išplėsti automatinę sintaksinę analizę sudėtiniams lietuvių kalbos sakiniams;
8. Visiems lietuvių kalbos sakiniams analizuoti ir klaidoms sumažinti būtina atlikti tokius darbus:
  - a) sukurti lietuvių kalbos morfologinę duomenų bazę (morfologyną), kurioje būtų sukaupta morfologinė informacija elektronine forma apie visų lietuvių kalbos žodžių visas formas. Dalį šiai duomenų bazei reikalingos informacijos galima gauti panaudojant Vytauto Zinkevičiaus lietuvių kalbos morfologijos modelių programinę įrangą
  - b) sukurti lietuvių kalbos semantinę duomenų bazę (semantyną), kurioje būtų sukaupta informacija elektronine forma apie kiekvieną lietuvių kalbos žodžio reikšmę.
  - c) Sudaryti lietuvių kalbos dažninių žodžių junginių žodyną.

## LITERATŪROS SĄRAŠAS

1. **Agel, V.** *Valenztheorie*. Tübingen: Narr, 2000.
2. **Allen, J.** *Natural Language Understanding*. Amsterdam: The Benjamin/Cummings Publishing Company, 1987.
3. **Ambrasas, V. (red.)**. *Dabartinės lietuvių kalbos gramatika*. Vilnius: Mokslo ir enciklopedijų leidybos institutas, 1997.
4. **Arnold, D., Balkan, L.** *Machine Translation: An Introductory Guide*. Cambridge: Blackwell Publishers, 1994.
5. **Batori, I.S., Lenders, W.** *Computerlinguistik: Ein internationales Handbuch zur computergestützten Sprachforschung und ihrer Anwendung*. Berlin: Walter de Gruyter, 1989.
6. **Bionik.** Informacija Internete  
<http://zdf.de/ZDF.de/INHALT/23/0,1872,2291543,00.html>; 2005.
7. **Block, H.U.** *Machinelle Übersetzung französischer Nominalsyntaxen ins Deutsche. – Linguistische Arbeiten 149. Dissertation, Universität Köln*. Tübingen: Max Niemeyer Verlag, 1984.
8. **Chabris, Ch.F.** *Artificial Intelligence & Turbo C*. Homewood: Dow Jones-Irwin, 1989.
9. **Cormen, T.H., Leiserson, Ch.E., Rivest, R.L.** *Introduction to Algorithms*. Cambridge/London: The MIT Press, 1990.
10. **Dagienė, V., Grigas, G.** *Programavimo kalbos*. Planuojamos išleisti knygos rankraštis.
11. **Daudaravičius, V.**, *Syntactical Analyzer of Lithuanian Sentence*. VDU Informatikos fakulteto magistro darbas, Kaunas, 2002.
12. **Deller, J.R., Hansen, J.H.L., Proakis, J.G.** *Discrete-Time Processing of Speech Signals*. New York: IEEE Press, 2000
13. **Engel, U.** *Syntax der deutschen Gegenwartssprache*. Berlin: Erich Schmidt Verlag, 1994.
14. **Glück, H. (Hrsg.)** *Metzler-Lexikon: Sprache*. Stuttgart: Verlag J.B. Metzler, 1993.
15. **Greenbaum, S.** *Oxford English Grammar*. Oxford: Oxford University Press, 1996.
16. **Grumadienė, L., Žilinskienė, V.** *Dažninis dabartinės rašomosios lietuvių kalbos žodynas*. Vilnius, 1997.
17. **Guilbaud, J-P.**, *Principles and Results of a French MT System at Grenoble University (Geta)*. Michaelson, S., Wilks, Y.(series Ed.) *Machine Translation Today: The State of The Art. – Information Technology Series*. Edinburgh: Edinburgh University Press, 1987, p. 278-318.
18. **Helbig, G., Buscha, J.** *Deutsche Grammatik: Ein Handbuch für Ausländer*. Leipzig: VEB Verlag Enzyklopädie, 1989.
19. **Hellwig, P.** <http://www.cl.uni-heidelberg.de/~hellwig/dug-2003.pdf>
20. **Hellwig, P.** Rankraštis (laiškas elektroniniu paštu). 2004.
21. **Henisz-Dostert, B., Macdonald, R.R., Zarechnak, M.** *Machine Translation. – Trends in Linguistics. Studies and Monographs 11*. Ed. W. Winter. New York: Mouton Publishers, 1979.
22. **Hutchins, W.J.** *Machine Translation: Past, Present, Future*. Chichester: Ellis Harwood Limited, 1986.
23. **Hutchins, W.J., Sommers, H.L.** *An Introduction to Machine Translation*. London: Academic Press, 1982.
24. **Jensen, K., Heidorn, G.E., Richardson, S.D.** *Natural language processing: The PLNLP Approach*. Boston/London: Kluwer Academic Publishers, 1993.
25. **Jurgaitis, R.** *Raganėlė ir kiti. Spalvinimo knygelė*. Vilnius: leidykla "Jūsų flintas, 2004.
26. **Kay, M., Gawron, J.M., Norvig, P.** *VerbMobil: A Translation System for Face-to-Face Dialog*. Stanford: CSLI, 1994.
27. **Kelly, J.D.K.** *Progress in Machine Translation*. Wilmslow, (UK): Sigma Press, 1989.
28. **Kudirka, Z.**, *Informatika – keturkalbis terminų žodynas*. Vilnius: Matematikos ir informatikos institutas, 1999.
29. **Labutis, V.** Vilniaus Universitete skaityto kurso "Lietuvių kalbos sintaksė" paskaitų konspektas, 2001.
30. **Labutis, V.** *Lietuvių kalbos sintaksė*. Vilnius: Vilniaus universiteto leidykla, 2002.
31. **Laffling, J.** *Towards high-precision machine translation: based on contrastive textology*. Berlin/New York: Foris Publications, 1991.
32. **Langer, S.** *Selektionsklassen und Hyponymie im Lexikon: semantische Klassifizierung von Nomina. Dissertation*. München: Universität München, 1996.

33. **Lehrberger, J., Bourbeau, L.** *Machine Translation: linguistic characteristics of MT systems and general methodology of evaluation*. Amsterdam/Philadelphia: John Benjamins publishing company, 1988.
34. **Marcinkevičienė, R.** *Tekstynų lingvistika ir lietuvių kalbos vartoseną*. Habilitacinis darbas, VDU, Kaunas, 2002.
35. **Mironas, R.** *Pančatantra*. Vertimas iš sanskrito kalbos. Iš serijos *Raštijos paminklai*. (Red. D. Navickaite) Vilnius: Mintis, 1990.
36. **Müller, H. M.** *Arbeitsbuch: Linguistik*. Paderborn: Verlag Ferdinand Schöningh, 2002.
37. **Nirenburg, S.** *Machine Translation: Theoretical and Methodological Issues*. London: Cambridge university Press, 1987.
38. **Patschull, K.** *Bionik*. Deutschland, Nr.3, 2005.
39. **Rimkutė, E.**, *Morfologinio daugiareikšmiškumo ribojimas kompiuteriniame tekстыne*. Daktaro disertacija, VDU Lietuvių kalbos institutas, Kaunas, 2006.
40. **Schubert, K.** *Metataxis: Contrastive Dependency Syntax for Machine Translation*. Dordrecht: Foris Publications, 1987.
41. **Slizienė, N.** *Lietuvių kalbos veiksmažodžių junglumo žodynas*. Vilnius: Mokslo ir enciklopedijų leidykla, 1994.
42. **Starkus, B.** *Visual Basic 6 jūsų kompiuteryje*. Kaunas: Smaltija, 2000.
43. **Šulcas, V.** *Visual Basic 6. Gramatika. I tomas*. Kaunas: Smaltija, 2003.
44. **Šveikauskienė, D.** *Automatinio vertimo apžvalga. – Informacinės technologijos '98. Konferencijos pranešimų medžiaga*. Kaunas: Technologija, 1998. psl. 191-194.
45. **Swamy, M. N. S., Thulasiraman, K.** *Graphs, Networks and Algorithms*. New York: John Wiley & Sons, 1981. Vertimas į rusų k. Москва: Мир, 1984.
46. *Tekstynas*. Informacija Internete. <http://donelaitis.vdu.lt>
47. **Ulvydas, K. (red.)**. *Lietuvių kalbos gramatika. Tomas III – Sintaksė*. Vilnius: Mokslo, 1976.
48. **Winograd, T.** *Language as a Cognitive Process. Volume I: Syntax*. London: Addison- Wesley Publishing Company, 1983.
49. **Zeller, I.**, *Automatinis terminų atpažinimas ir apdorojimas*. Daktaro disertacija. VDU Lietuvių kalbos institutas, Kaunas, 2005.
50. **Zinkevičius, V.** *Lemuoklis - morfologinei analizei. - Darbai ir dienos 24*, Kaunas: Vytauto Didžiojo universitetas, 2000. psl. 245-273
51. **Žandaris, A.** *Formaliosios gramatikos ir lietuviškas sakinytis. – Kompiuterininkų dienos '97. Birštonas*. Vilnius: Žara, 1997. psl. 281-284.
52. **Белецкая, И.П.** *Деревья зависимостей как инструмент синтаксического анализа текста. – Международный семинар по машинному переводу*. Москва: ВЦП, 1983.
53. **Билан, В. Н.** *Семантико-синтаксический модуль воспроизводящей инженерно-лингвистической модели 'Переводчик'*. – *Проблемы внутренней динамики речевых норм*. Минск: МГПИИЯ, 1982. с. 174-185.
54. **Билан, В.Н., Крисевич, В.С. и др.** *Семантико-синтаксический перевод в группе Статистика речи. – Автоматический анализ текста*. Минск: МГПИИЯ, 1976.
55. **Гарина, Т.М.** *Программирование синтаксических преобразований – Международный семинар по машинному переводу*. Москва: ВЦП, 1979. с. 43-44.
56. **Грязнухина, Т.А.** *Синтаксический анализ научного текста на ЭВМ*. Киев: Наукова думка, 1999.
57. **Ермаков, А.Е.** *Автоматическая обработка текста; Синтаксический анализ*. Informacija Internete <http://www.aot.ru/docs/synan.html> 2007.
58. **Истомин, С.А., Петровская, В.М., Печников, А.П.** *Представление структуры предложения на основе словосочетаний и использование этого принципа при АП. – Переработка текста методами инженерной лингвистики. Тезисы докладов конференции*. Минск: МГПИИЯ, 1982. с. 80-83.
59. **Кунце, Ю.** *Введены семантических критериев в синтаксические правила. – Научно-технические исследования (6)*, 1981. с.30-34.
60. **Марчук, Ю. Н.** *Опыт машинного семантико-синтаксического анализа текста для перевода. – Семантика текста и проблемы перевода*. Москва: Институт языкознания, 1984. с. 90-97.
61. **Матвеева, Е.И.** *Автоматический синтаксический анализ с применением сетевой грамматики. – Известия АН Кирг. ССР: Общественные науки*. Фрунзе: изд. АН Кирг. ССР, 1987. № 2, С. 83-88.

62. **Окулич, Н.Э., Петровская, В.М.** *Автоматический валентный анализ синтаксической структуры немецкого предложения. – Романское и германское языкознание. Впуск 1. Вопросы экспериментальной фонетики и прикладной лингвистики.* Минск: МГПИИЯ. 1978. С. 185-192.
63. **Сердюков, П. И.** *Оптимизация алгоритма поиска синтаксических связей. – Международный семинар по машинному переводу.* Москва: ВЦП, 1979. С. 123-125.

Daiva Šveikauskienė

LIETUVIŲ KALBOS VIENTISINIŲ SAKINIŲ  
AUTOMATINĖ SINTAKSINĖ ANALIZĖ

Daktaro disertacija

Spausdino VGTU leidykla „Technika“  
Saulėtekio al. 11, LT-10223 Vilnius  
Tiražas 20 egz.