

DAIVA ŠVEIKAUSKIENĖ
Lietuvių kalbos institutas

LIETUVIŲ KALBOS SINTAKSINĖ ANALIZĖ

1. Įvadas

1942 m. Harvardo universitete buvo sukurtas pirmasis pasaulyje kompiuteris MARK I [Schwanke 1991: 69], ir labai greitai tapo aišku, kad kompiuteriai gali apdoroti ne tik skaičius, bet ir visus kitus simbolius, pvz., raides, taigi – jie gali apdoroti ir kalbas.

Apie 6-ąjį praeito amžiaus dešimtmetį, atsiradus galimybei automatiškai (kompiuteriu) versti tekstus iš vienos kalbos į kitą, kalbininkai susidomėjo formaliomis gramatikomis. Daugelis lingvistų pradėjo kurti formalius kalbų aprašus, kad ir joms būtų galima panaudoti kompiuterių pajėgumus [Winograd 1983: 23]. Buvo bandoma sudaryti matematinius kalbų modelius – taip gimė formalios kalbos bei formalios gramatikos idėja. Tik panaudota ji buvo programavimui daug anksčiau negu tautų kalboms, matyt, todėl, kad formalų aprašą lengviau taikyti dirbtiniam daiktui, t. y. programavimo kalbai. Tautų kalbas formalizuoti yra žymiai sunkiau. Pirmoji kalba, kuriai buvo taikoma formali gramatika, – tai anglų kalba. Reikia pasakyti, kad ji iš tiesų kompiuterizuota daugiausiai ir geriausiai. Lietuvių kalba priklauso mažiausiai kompiuterizuotų Europos kalbų grupei [Vaišnienė, Zabarskaitė 2012: 35].

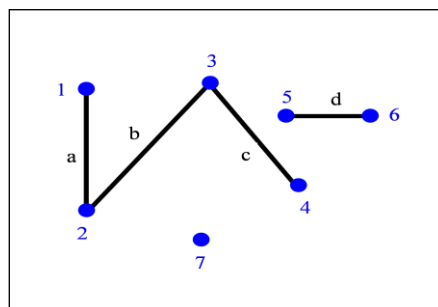
Lietuvių kalbos institute pradedami sintaksinio tekstyno, t. y. tekstyno, kuriame sakiniai yra sintaksiškai anotuoti, kūrimo darbai, ir šis straipsnis skiriamas lietuvių kalbos sakinių sintaksinei struktūrai aptarti.

2. Du būdai sintaksinei struktūrai pavaizduoti

Labiausiai paplitęs sakinio struktūros pavaizdavimo būdas yra grafas, tiksliau, medis [Allen 1987: 41]. Kalbininkams tai nėra lingvistinės, kasdien vartojamos sąvokos, todėl bus pateikta informacijos ir iš grafų teorijos.

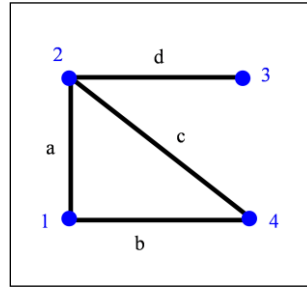
2.1. Medis grafų teorijoje

Pagrindinis medžio apibrėžimas skamba taip: „medis – tai susietas grafas be ciklų“ [Свами, Тхуласираман 1984: 33]). Šiame apibrėžime yra keturi terminai: *grafas*, *ciklas*, *medis* ir *susietas*. Iš pradžių reikėtų išsiaiškinti, kas gi vadinama *grafu*. Grafas – tai grupė objektų, sujungtų linijomis, vadinamomis lankais. Patys objektai vadinami viršūnėmis. Viršūnės dažniausiai numeruojamos, o lankams suteikiami raidžių pavadinimai. Grafas gali būti *susietas* arba *nesusietas*. Nesusieto grafo pavyzdys pateikiamas 1 pav. „Grafas vadinamas susietu, jei tarp bet kurių dviejų jo viršūnių egzistuoja kelias“ [Свами, Тхуласираман 1984: 18], t. y., jei iš kiekvienos viršūnės einant linijomis galima patekti į visas kitas grafo viršūnes. Pateiktas grafas (žr. 1 pav.) turi viršūnę 7, iš kurios neišeina nė viena linija, vadinasi, iš šios viršūnės negalima patekti į jokią kitą viršūnę. Iš viršūnės 5 galima patekti į viršūnę 6 einant linija d, bet iš jos (viršūnės 5) negalima patekti į viršūnę 7 ar 3, taigi šis grafas netenkina susieto grafo apibrėžimo sąlygų, todėl yra nesusietas.



1 pav. Nesusieto grafo pavyzdys

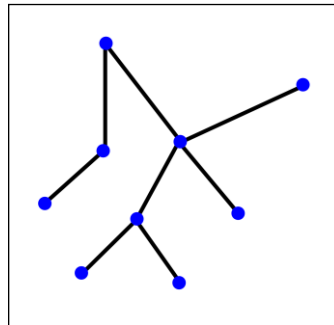
Susietas grafas parodytas 2-ame paveikslėlyje: čia iš kiekvienos viršūnės yra kelias į bet kurią kitą grafo viršūnę, t.y., jei mes esame kurioje nors viršūnėje, eidami linijomis galime pakliūti į visas kitas šio grafo viršūnes.



2 pav. Susieto grafo pavyzdys

Dar vienas kalbininkams neįprastas terminas medžio apibrėžime yra *ciklas*. Sakoma, kad grafas turi ciklą (ar kelis ciklus), jei išėjus iš kurios nors viršūnės į ją galima sugrįžti kitu keliu. Išėjus iš viršūnės 4 (žr. 2 pav.) einant linija *c* galima patekti į viršūnę 2, o sugrįžti jau kitu keliu – linijomis *a* ir *b*, vadinasi, 2-ame paveikslėlyje pavaizduotas grafas turi ciklą.

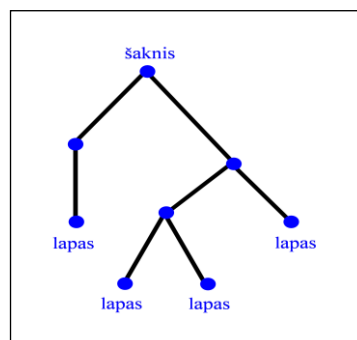
Pagal apibrėžimą medis negali turėti ciklą. Medžio pavyzdys pateiktas 3-iaame paveikslėlyje: iš bet kurios viršūnės išėjus į kurią nors kitą, sugrįžti galima tik tuo pačiu keliu, kuriuo buvo eita, kito kelio atgal nėra.



3 pav. Grafų teorijos medžio pavyzdys

Kalbininkų medžiai visada braižomi viršūne žemyn (žr. 4 pav.), t. y. jų šaknis yra viršuje, o lapai apačioje [Batori, Lenders 1989: 23]. Čia reikėtų aptarti dar du grafų teorijos terminus *šaknis* ir *lapas*. Šaknis yra tokia grafo viršūnė, į kurią neateina jokia linija, iš jos linijos gali tik išeiti. Lapas – tai tokia viršūnė, į kurią linija ateina, bet iš jos toliau nebeišeina nė viena linija.

Lingvistikoje žinomi du iš principo skirtingi medžio sudarymo metodai: frazių ir priklausomybių. Frazių gramatika atėjo iš Amerikos struktūralizmo (Bloomfield) ir daugiausia naudojama anglų kalbos sintaksinei analizei, o priklausomybių gramatika (Tesnière) labiau taikoma kitoms Europos kalboms [Batori, Lenders 1989: 625].



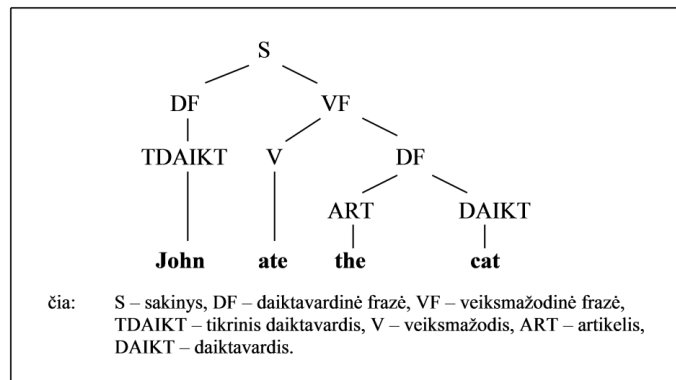
4 pav. Kalbininkų medžio pavyzdys

Abu sakinio sintaksinės struktūros pavaizdavimo metodai – ir frazių, ir priklausomybių – buvo kuriami veikiant formalioms gramatikoms, ir abu jie yra sąlygoti formalių gramatikų reikalavimo – turėti vieną pradinį simbolį. Pirmoji buvo sukurta frazių gramatika, pasirinkusi pradinį simboliu sakinį. Vėliau pasirodė priklausomybių gramatika, kuri pradinį simboliu laiko tarinį.

2.2. Frazių gramatika

Frazių gramatikoje pradinis simbolis – sakinys – iš pradžių skaidomas į dvi – veiksnio ir tarinio – grupes, kaip ir „tradicinėje gramatikoje, kurioje sakinio analizė remiasi subjekto ir predikato dichotomija“ [Holvoet 2009: 10].

Frazių gramatika buvo kuriama anglų kalbai, stengiantis maksimaliai atspindėti tos kalbos savybes. Atliekant sintaksinę analizę, čia remiamasi specifiniu anglų kalbos bruožu – griežta žodžių tvarka. Būtent ji ir yra pagrindinis kriterijus, į kurį atsižvelgiama nustatant sakinio dalis. Taigi ir pati sintaksinė struktūra susijusi su žodžių išsidėstymu sakinyje. Kaip pavyzdį galima pateikti labai paprasto anglų kalbos sakinio *John ate the cat* medį [Allen 1987: 42] (žr. 5 pav.).

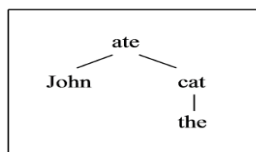


5 pav. Sakinio *John ate the cat* sintaksinė struktūra frazių gramatikos metodu

Anglų kalbos gramatika reikalauja, kad veiksnys būtų pirmoje vietoje, o tarinys antroje, todėl struktūroje pirmiau eina daiktavardinė frazė, o paskui veiksmažodinė. Toks sakinio pavaizdavimo būdas visai netiko kalboms su laisva žodžių tvarka, todėl Europoje pradėta kurti kita gramatika – priklausomybių, kuri geriau atitiko kalbų, turinčių laisvą žodžių tvarką sakinyje, specifiką.

2.3. Priklausomybių gramatika

„Priklausomybių gramatika pateikia tyrėjui universalesnį aprašymo įrankį“ [Holvoet 2009: 31]. Naudojantis ja, žinoma, galima pavaizduoti ir anglų kalbos sakinius, tačiau priklausomybių gramatikoje sakinio struktūra jau nebesusijusi su linijiniu žodžių išsidėstymu. Frazių gramatika teigia, kad sakinys yra sudarytas iš frazių ir kad ši struktūra yra svarbi perduodant reikšmę („Sentences are made up of ‚phrases‘, and this structuring is important for how they communicate meaning“) [Winograd 1983: 73]. Priklausomybių gramatikoje manoma, kad „sintaksė ne tiek turi grupuoti žodžius į frazes, kiek nustatyti tiesioginius ryšius tarp pačių žodžių“ („The job of syntax is not so much to group words into phrases, as to establish direct relationships among the words themselves.“) [Kay, Gawron, Norvig 1994: 53]. Sakinys *John ate the cat*, kurio struktūra frazių gramatikos metodu pateikta 1-ame paveikslėlyje, priklausomybių gramatikoje atrodytų taip (6 pav.):



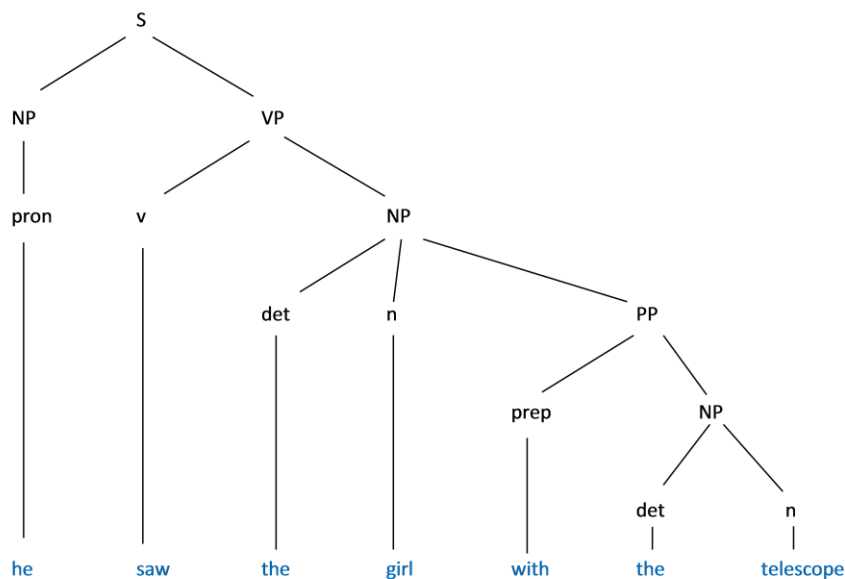
6 pav. Sakinio *John ate the cat* sintaksinė struktūra priklausomybių gramatikoje

Lietuvių kalbos, turinčios laisvą žodžių tvarką, sakinių sintaksinei struktūrai pavaizduoti geriau tinka būtent šios, priklausomybių gramatikos, metodas.

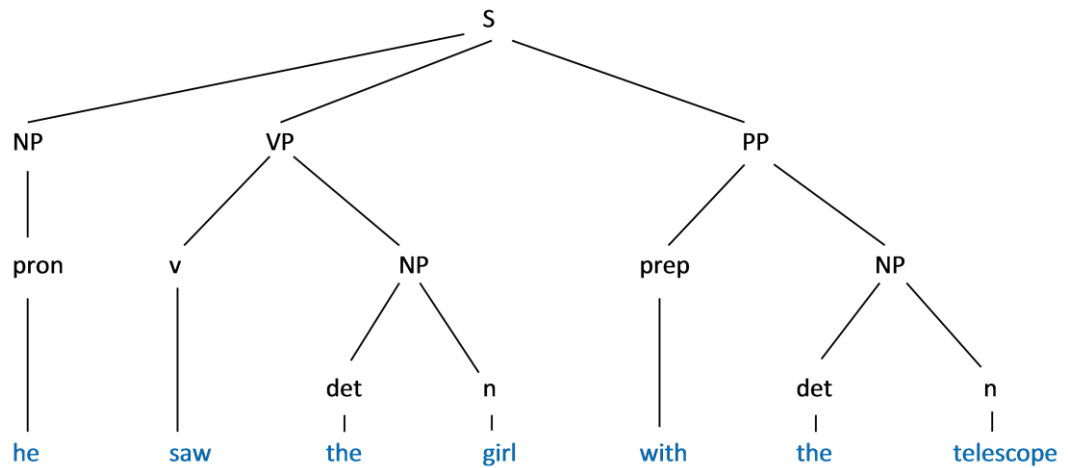
2.4. Principiniai frazių gramatikos ir priklausomybių gramatikos skirtumai

Frazių gramatikos pagrindas yra komponentų (sudedamųjų dalių) ryšiai, o ne priklausomybės santykiai, kuriais remiasi priklausomybių gramatika („Phrase structure grammars are based on the constituency relation, as opposed to the dependency relation associated with dependency grammars.“ [1 internetinė nuoroda]). Abiejų gramatikų – frazių ir priklausomybių – svarbiausias skirtumas yra ne tai, kad vienur naudojama verbocentrinė, kitur dichotominė sakinio struktūros forma. Jos abi turi vieną pradinį simbolį: frazių gramatikoje tai yra sintaksinė kategorija – sakinyš, o priklausomybių gramatikoje pradiniu simboliu laikomas tarinio funkciją atliekantis žodis. Pagrindinis skirtumas yra sakinio struktūros vaizdavimo principas, sakinio žodžių pateikimo būdas: frazių gramatikoje sakinyš pateikiamas taip, kaip mes jį matome tekste (žr. 5 pav.) ir virš sakinio žodžių nurodomos sintaksinės kategorijos. Priklausomybių gramatikoje sintaksinė struktūra nesusijusi su žodžių išsidėstymu sakinyje, mes nematome sakinio tokio, koks jis parašytas tekste, ir, jei tai mums nežinomos kalbos sakinyš, net negalėtume atkurti jo linijinės struktūros, t.y. tokio žodžių išsidėstymo, kokio reikalauja tos kalbos gramatika (žr. 6 pav.). Priklausomybių gramatikos struktūroje nėra sintaksinių kategorijų, ten vien sakinio žodžiai, sujungti linijomis.

Tai, kad sakinyš iš pradžių skaidomas į veiksnio frazę ir tarinio frazę, nėra esminė frazių gramatikos ypatybė, nes pradinis simbolis yra vis tiek vienas (sakinyš), kaip ir priklausomybių gramatikoje. Be to, frazių gramatikoje sakinyš iš pradžių nebūtinai turi būti skaidomas į dvi dalis, kartais jis iš karto skaidomas į tris, jei kokia nors (pvz., prielinksninė) konstrukcija priklauso visam sakiniui. Pagal schemą ir sprendžiama apie sakinio prasmę. Pvz., sakinio *He saw the girl with the telescope* [Hutchins, Somers 1982:61] sintaksinį daugiareikšmiškumą panaikina schema – kaip ji bus pateikta, tokia bus ir sakinio prasmė. Schemoje (žr. 7 pav.) parodomas sakinyš, kuris reiškia, kad jis matė mergaitę, laikančią savo rankose žiūroną. Toliau (žr. 8 pav.) parodyto sakinio reikšmė: jis su žiūronais stebėjo mergaitę (t.y. matė ją), vadinasi, žiūronas šį kartą buvo jau ne mergaitės, o jo rankose. Toks pavyzdys pateiktas literatūroje, nors šią sakinio prasmę turbūt galėtų turėti ir schema, kurioje PP priklausytų VP.



7 pav. Sakinio *He saw the girl with the telescope* sintaksinė struktūra



8 pav. Sakinio *He saw the girl with the telescope* sintaksinė struktūra

Frazių ir priklausomybių gramatikos skiriasi ne požiūriu į sakinio dalių svarbą (ne tai, ar tarinys, ar branduolys, t. y. veiksny su tariniu, laikomas sakinio pagrindu), bet sakinio sintaksinės struktūros pavaizdavimo būdu: priklausomybių gramatikos medyje kiekvienoje viršūnėje yra sakinio žodis („The content of each node is a word“) [Winograd 1983: 80] ir žemiau jo pateikiami jį išplečiantys žodžiai („The children of the node are its modifiers“) [Winograd 1983: 80], o frazių gramatikoje tik medžio lapuose yra žodžiai – visose kitose viršūnėse yra sintaksinės kategorijos („Leaf nodes have words and all other nodes have syntactic categories“): daiktavardinė frazė, sakinys ir pan. [Winograd 1983: 80].

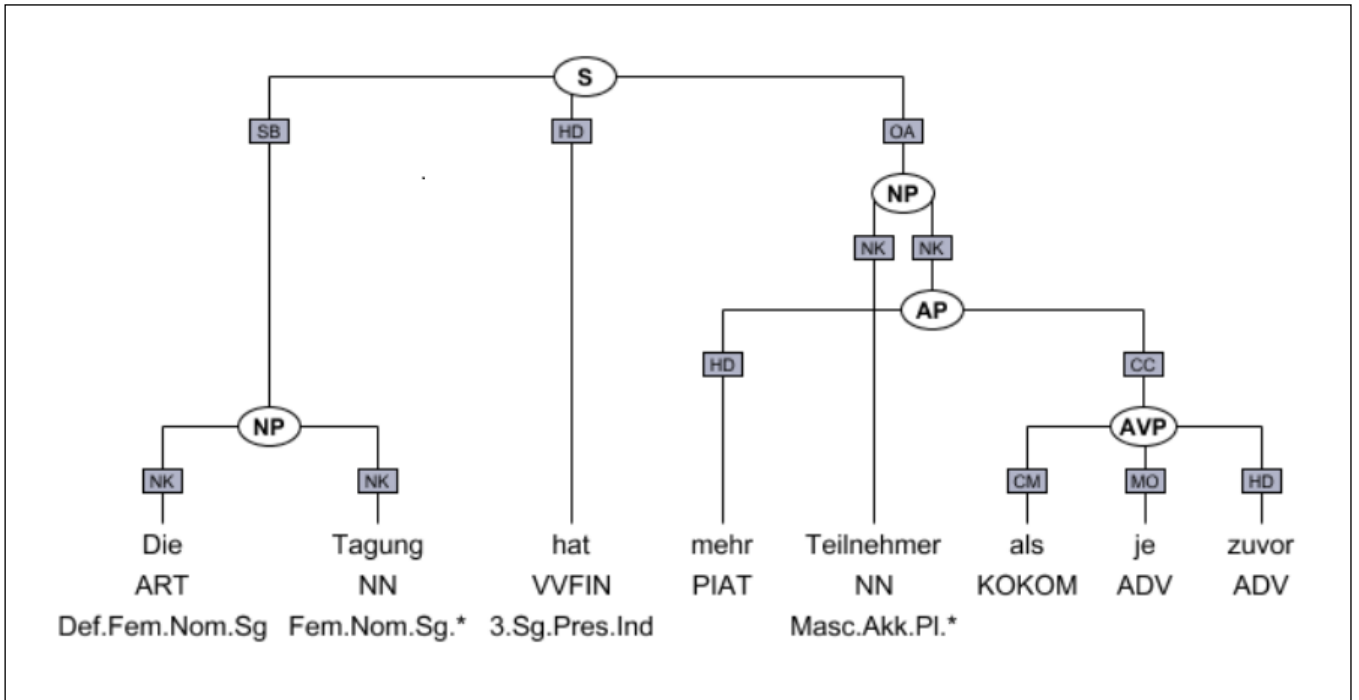
3. Lietuvių kalbos sakinių sintaksinė struktūra

Pradedant sintaksiškai anoutuoti lietuvių kalbos tekstyną, reikia pasirinkti formą, kuria bus pateikiama sintaksinė informacija apie lietuvių kalbos sakinius. Svarbiausia, žinoma, yra pati sakinio sintaksinė struktūra. Ji turi būti pateikta taip, kad kuo geriau atitiktų lietuvių kalbos ypatybes. Lietuvių kalbos specifiką atspindinčiais sakiniais parodžius abiejų metodikų – frazių ir priklausomybių – skirtumą, geriau matyti jų tinkamumas lietuvių kalbai. Tipiniai pavyzdžiai su anglų kalbos žodžių tvarka (pirmoje vietoje veiksny, antroje – tarinys, trečioje – papildinys ar aplinkybė) išsamios informacijos nepateikia. Tokie atvejai yra jau išnagrinėti, jiems sukurtos programos, gerai atliekančios tos kalbos sintaksinę analizę, kuriai buvo kuriamos, t. y. anglų. Deja, jos visai netinka lietuvių kalbos sakiniams analizuoti. Čia turi būti sukurtos būtent lietuvių kalbos specifiką atspindinčios programos. Ir pagrindas, žinoma, yra sakinio sintaksinė struktūra.

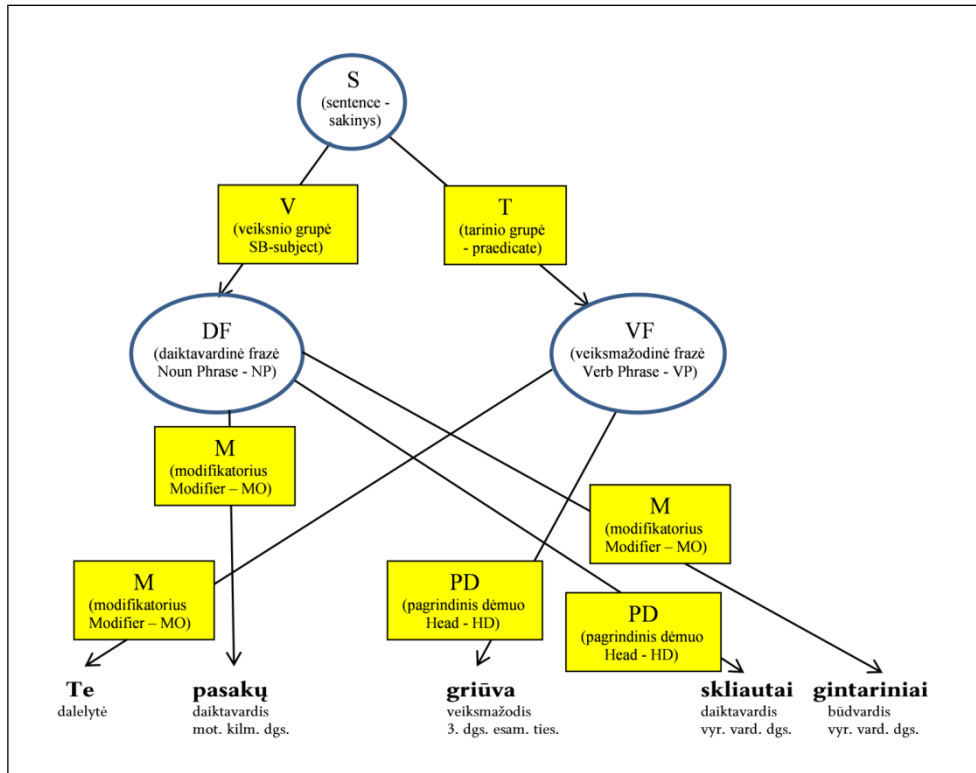
3.1. Frazių gramatika

Kadangi lietuvių kalbos žodžių tvarka yra laisva, vaizduojant sakinius frazių gramatikos metodu sintaksinė struktūra būtų labai paini ir sudėtinga. Ši problema pasireiškia jau vokiečių kalboje, kurios žodžių tvarka dar gana griežta, nors ir laisvesnė negu anglų kalboje. Vokiečių kalboje veiksny gali pasikeisti su papildiniu vietomis, t. y. papildinys gali būti pirmoje vietoje, o veiksny tada nueina į trečią, bet griežtas tarinio antroje vietoje reikalavimas išlieka. Internete pateikto vokiško sakinio struktūroje jau matome linijų susikirtimą (žr. 9 pav.) [2 internetinė nuoroda], o, vaizduojant kai kuriuos lietuvių kalbos sakinius šiuo metodu, jų būtų tikra raizgalynė

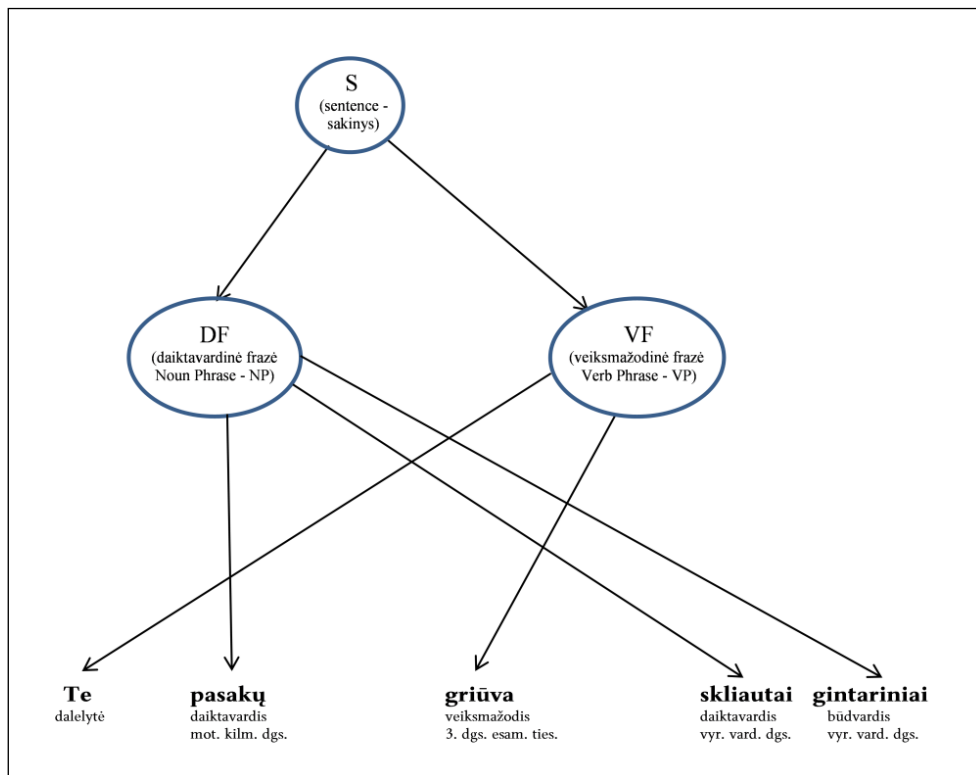
(žr. pav.: 10, 11 ir 12). Šie sakiniai atitinkamai pavaizduoti neparodant sintaksinio ryšio tipo (žr. pav.: 10a, 11a ir 12a), tačiau struktūra lieka vis tiek labai paini.



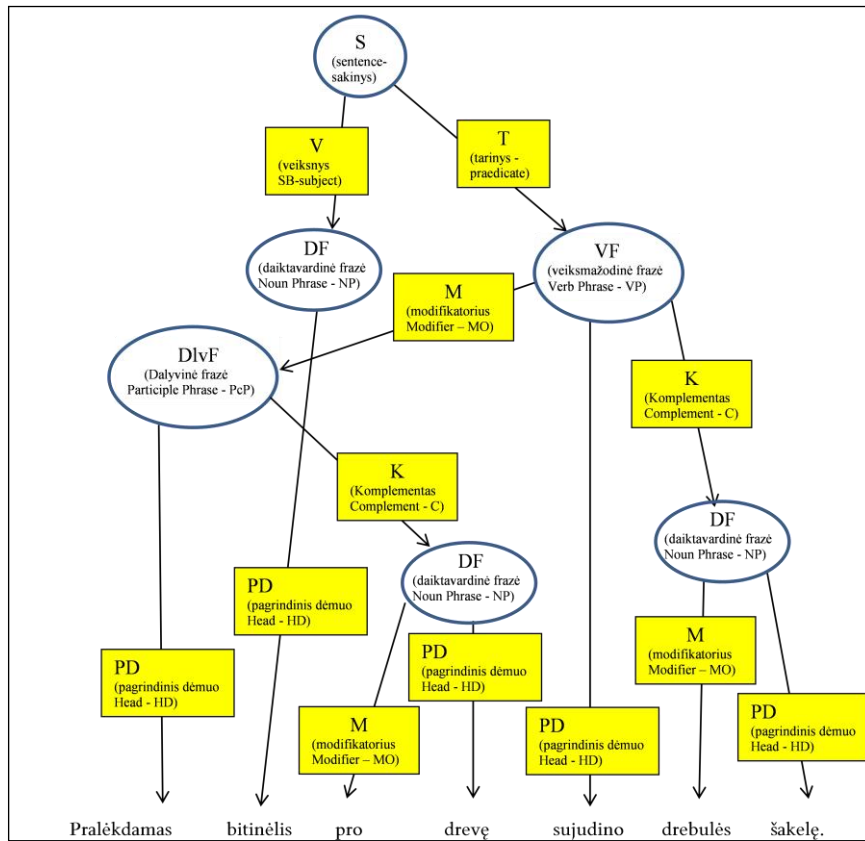
9 pav. Vokiško sakinio *Die Tagung hat mehr Teilnehmer als je zuvor* struktūra frazių metodu, turinti susikertančių linijų



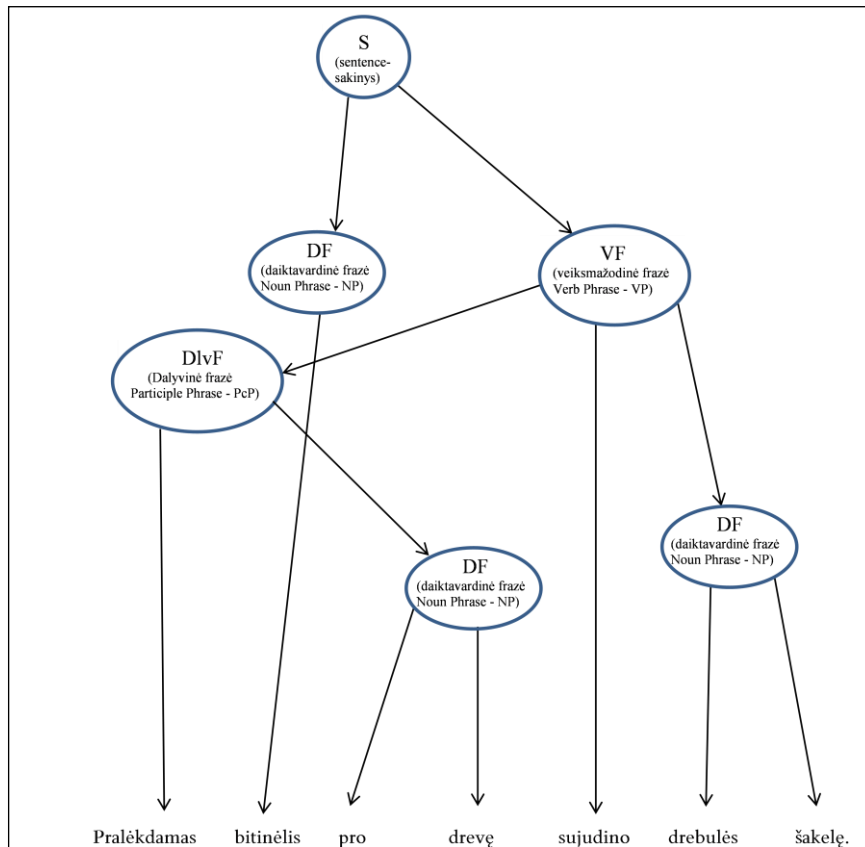
10 pav. Sakinio *Te pasakų griūva skliautai gintariniai* sintaksinė struktūra frazių metodu



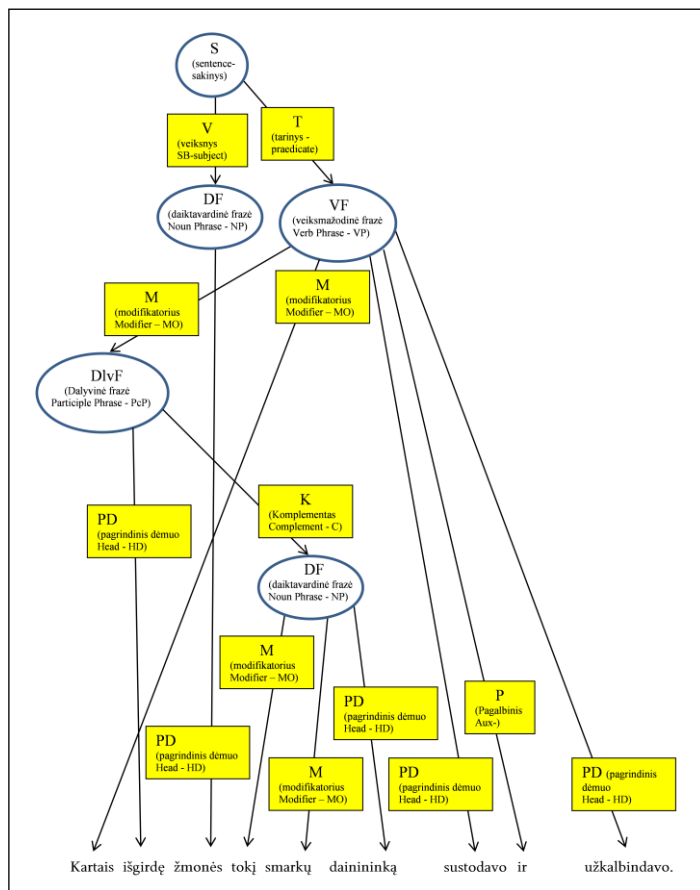
10a pav. Sakinio *Te pasakų griūva skliautai gintariniai* sintaksinė struktūra frazių metodu neparodant sintaksinio ryšio tipo tarp žodžių



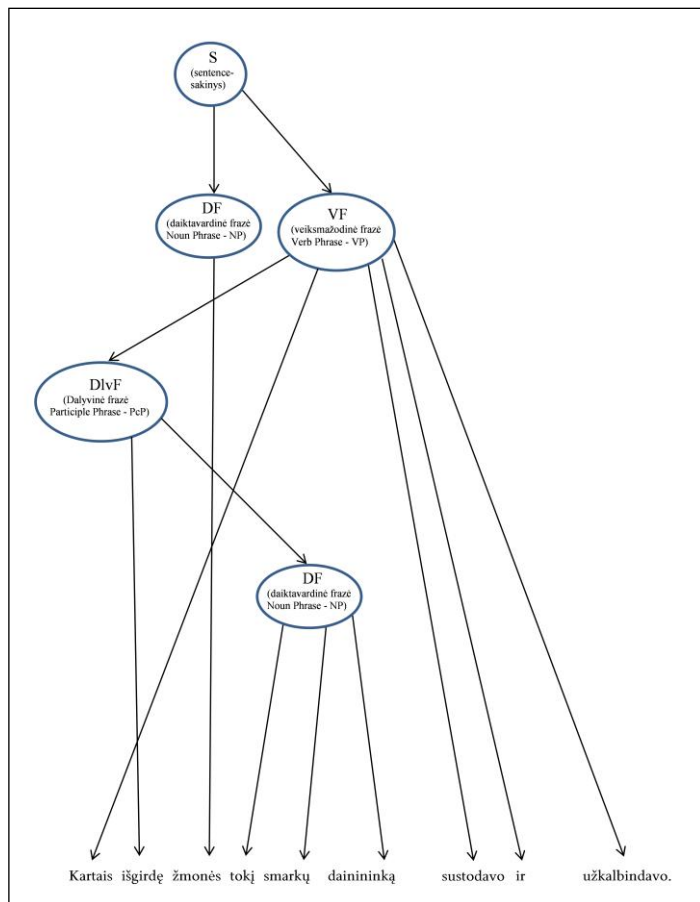
11 pav. Sakinio *Pralėkdamas bitinėlis pro drevę sujudino drebulės šakelę* sintaksinė struktūra frazių metodu



11a pav. Sakinio *Pralėkdamas bitinėlis pro drevę sujudino drebulės šakelę* sintaksinė struktūra frazių metodu neparodant ryšio tipo



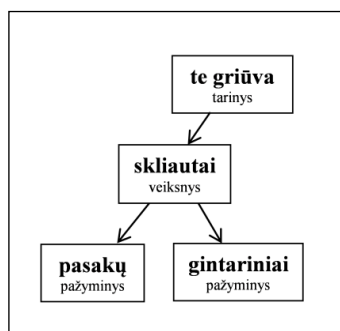
12 pav. Sakinio *Kartais išgirdę žmonės tokį smarkų dainininką sustodavo ir užkalbindavo* sintaksinė struktūra frazių metodu



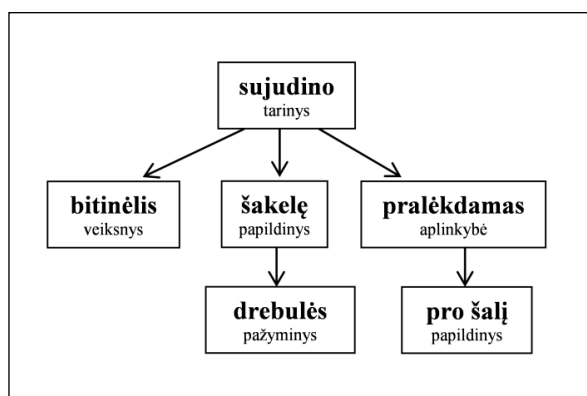
12a pav. Sakinio *Kartais išgirdę žmonės tokį smarkų dainininką sustodavo ir užkalbindavo* sintaksinė struktūra frazių metodu neparodant ryšio tipo

3.1. Priklausomybių gramatika

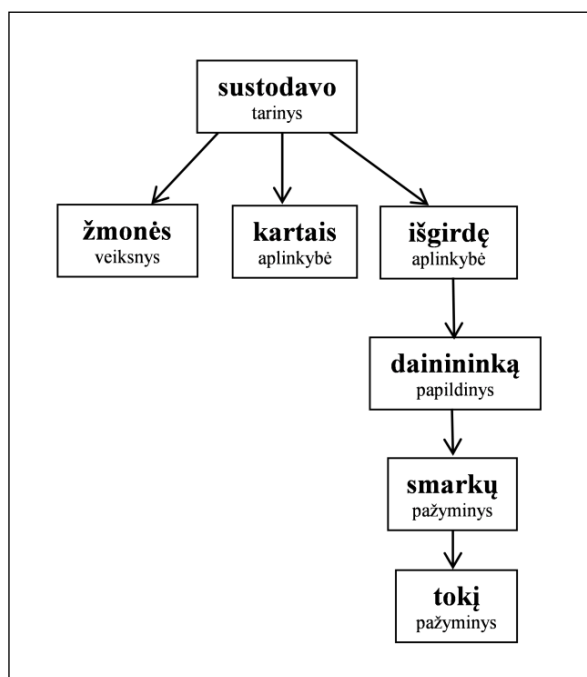
Ankstesniame skyriuje pateikti sakiniai (žr. 10–12a pav.), pavaizduoti priklausomybių metodu, atrodytų taip (žr. pav.: 13, 14 ir 15):



13 pav. Sakinio *Te pasakų griūva skliautai gintariniai* sintaksinė struktūra priklausomybių metodu

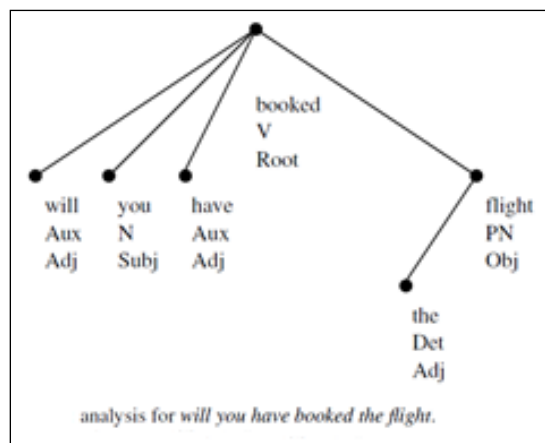


14 pav. Sakinio *Pralėkdamas bitinėlis pro drevę sujudio drebulės šakelę* sintaksinė struktūra priklausomybių metodu



15 pav. Sakinio *Kartais išgirdę žmonės tokį smarkų dainininką sustodavo ir užkalbindavo* sintaksinė struktūra priklausomybių metodu

Matyti, kad priklausomybių metodu pavaizduotų sakinių struktūros yra daug paprastesnės ir suprantamesnės, todėl anotuojant tekstyną bus naudojama priklausomybių gramatika. Aišku, kad ji mums tinka labiau, tačiau ir jos nevertėtų perimti akiai, kaip vienintelės galimos ir neklystančios. Pateiktas Pensilvanijos tekstyno pavyzdys (žr. 16 pav.), kuriame sakinys pavaizduotas priklausomybių gramatikos metodu [3 internetinė nuoroda]. Nekelia pasitikėjimo tokia schema, kurioje pagalbinis veiksmažodis *will*, tarnaujantis laiko formai sudaryti (ir atliekantis iš esmės tą pačią funkciją, kaip mūsų būtojo dažninio laiko priesaga *-dav-*), laikomas to paties rango žodžiu kaip ir veiksnys.



16 pav. Pensilvanijos tekstyno pavyzdys

Parenkant lietuviško sakinio sintaksinę struktūrą turėtų būti įvertintos specifinės lietuvių kalbos ypatybės, ir iš visų pasaulyje sukurtų metodikų paimama tai, kas mums tikrai reikalinga. Taip elgiasi ir kitų šalių kalbininkai. Pavyzdžiui, vokiečiai Engel, Eroms, Heringer savo darbuose įvertina tik vokiečių kalbos gramatikai būdingas ypatybes („...ziehen sie in ihren Arbeiten...nur grammatische Eigenschaften des Deutschen heran“) [Agel 2000: 92].

3.2. Sintaksinės struktūros forma lietuvių kalbos sakiniams

Parenkant būdą, kuriuo bus vaizduojama lietuvių kalbos sakinių sintaksinė struktūra, vertėtų paimti iš abiejų gramatikų – ir frazių, ir priklausomybių – tai, kas geriausiai tinka lietuvių kalbai. Priklausomybių gramatikos vaizdavimo principas mums tinka labiau nei frazių, nes geriau atspindi kalbų, turinčių laisvą žodžių tvarką sakinyje, ypatybes. Tai labai gražiai ir aiškiai parodyta knygoje „Dabartinės lietuvių kalbos gramatika“ [Ambrazas, red. 1997: 478].

Frazių gramatikos bruožas palankus lietuvių kalbai – tai tarinio ir veiksnio frazių lygiavertiškumas. Lietuvių kalbos sakinių struktūroje frazių analogai galėtų būti veiksnio ir tarinio grupės, kurios nėra tikslus frazių atitikmuo, nes čia ir baigiasi frazių gramatikos panaudojimas lietuvių kalbos sakiniams. Toliau sakinyje vaizduojamas pagal priklausomybių gramatikos principus.

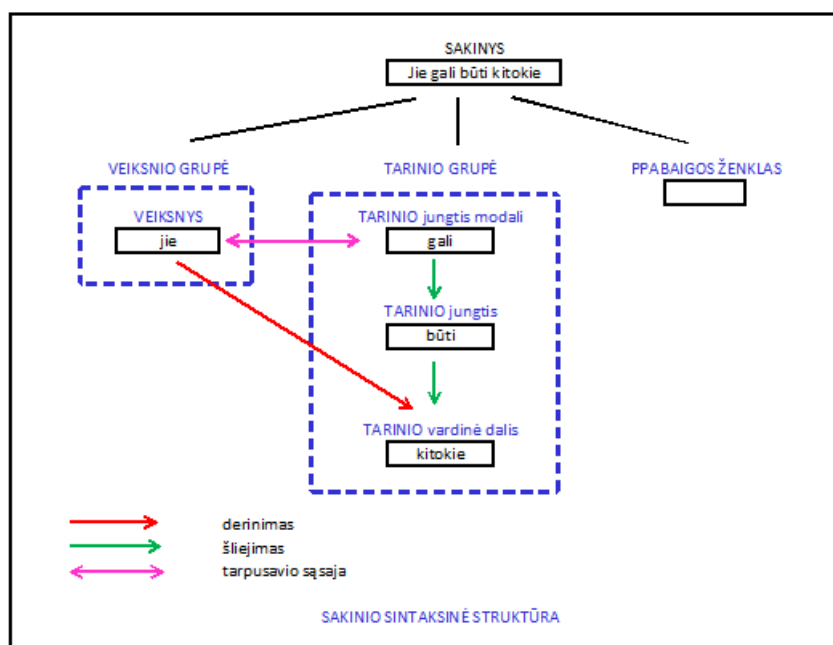
3.2.1. Priklausomybių gramatikos silpnosios pusės

Naujausioje lituanistinėje literatūroje kritikuojamas DLKG autorių tarpusavio sąsajos apibrėžimas nurodant kaip trūkumą, kad „...tipiniu raiškos būdu paskelbta tarpusavio morfologinė priklausomybė faktiškai naudojama kaip vienintelis skiriamasis kriterijus“ [Holvoet, Mikulskas, red. 2009: 27]. Galima pasakyti, kad tarpusavio sąsaja nėra vienintelis argumentas, rodantis veiksnio ir tarinio skirtumą nuo antrininkų sakinio dalių. Dar vienas būtų toks: visas penkias sakinio dalis – ir veiksnį, ir tarinį, ir papildinį, ir pažyminį, ir aplinkybę – galima išplėsti papildiniu, pažyminiu ir aplinkybe, tačiau nė vienos iš antrininkų sakinio dalių – nei papildinio, nei pažyminio, nei aplinkybės – negalima išplėsti veiksmiu ar tariniu. Argumentų galima rasti ir daugiau: veiksnys su tariniu gali sudaryti sakinį, tarinys su papildiniu, ar tarinys su aplinkybe, ar veiksnys su pažyminiu sakinių paprastai nesudaro, nebent išimtis

– eliptiniai sakiniai ir pan. Visais atvejais, kai sakinių sudaro vienas žodis (vien veiksmazodis ar vien daiktavardis), labai nesunkiai galima atkurti trūkstantį branduolio narį. Tipiškas sakinyvis yra dvinaris, jis mums lyg ir aiškesnis. Juolab, kad kai kurios kitos kalbos iš viso neturi vienanarių sakinių: net ir gamtos reiškiniams aprašyti prideda formalų veiksnį, pvz., *es* (vok.iečių *Es regnet.*), *it* (anglų *It rains. It is raining.*) ir kt.

Tokių argumentų galima būtų pateikti ir daugiau. Jie visi rodo, kad „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 1998: 202].

Pirmi sintaksiškai anotuoti sakiniai iš lietuvių kalbos tekstyno labai aiškiai parodo problemas, su kuriomis būtų susiduriama laikant veiksnį žemesnio rango sakinio dalimi nei tarinys. Pateiktame sintaksiškai anotuoto sakinio (*Jie gali būti kitokie*) pavyzdyje [5 internetinė nuoroda] (žr. 17 pav.) matome, kad veiksnys nulemia tarinio vardinės dalies galūnę. Pavaizdavus šį sakinį laikantis verbocentrinio požiūrio, rodyklė, žyminti derinimo sintaksinį ryšį, turėtų eiti į viršų, taigi žemesnio rango žodis turėtų nulemti aukštesnio rango žodžio formą. Tai neatrodo labai logiška.



17 pav. Sintaksiškai anotuoto tekstyno sakinio pavyzdys

3.2.2. Savitos lietuvių kalbos ypatybės

Viena iš savitų lietuvių kalbos ypatybių, kurios negalima ignoruoti, ir kurios anglų kalba iš viso neturi [Greenbaum 1996], o kitoms kalboms ji ne tokia aktuali kaip mums – tai tarininis pažyminyvis. Sakinio sintaksinėje struktūroje būtina parodyti abu jo ryšius: ir su tariniu, ir su veiksmiu ar papildiniu. Nors tai padarius bus sugriautas sakinio sintaksinės struktūros pavaizdavimo medžiui principas, nes gautas grafas jau nebeatitinka medžio apibrėžimo dėl atsiradusio uždaro kelio – ciklo, lietuvių kalboje terodyti kurį nors vieną iš šių ryšių, neprarandant sintaksinės informacijos, negalima todėl, kad abu jie reiškiami formaliai: jei bent vienas iš šių ryšių sakinio struktūroje nebus pavaizduotas, automatinio vertimo į kitas kalbas metu bus daroma klaidų.

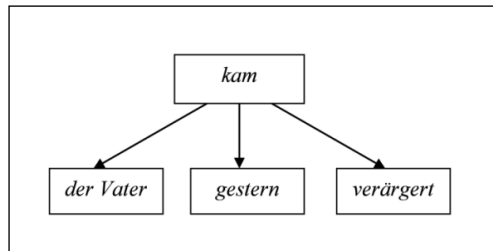
Šį teiginį galima pailustruoti tokiu pavyzdžiu [Šveikauskienė 2005: 411–416]: vokiečių kalboje, kuri turi mažiau žodžių formų su skirtingomis galūnėmis, tarininis pažyminyvis formaliai sutampa su aplinkybe („...*das prädikative Attribut mit der Adverbialbestimmung ... in der äußeren Form und Position im konkreten Satz übereinstimmt.*“ [Helbig, Buscha 1989: 556]), ir čia nėra galūnių derinimo su ta sakinio dalimi, kurią jis pažymi. Pvz.:

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

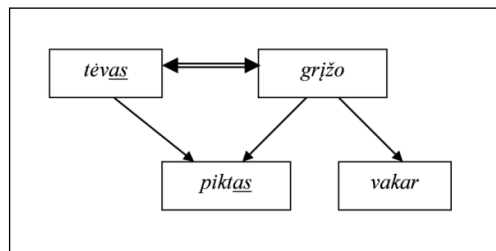
Sakiniuose tarininio pažyminio forma lieka ta pati (*verärgert*) nepriklausomai nuo veiksnio giminės ir skaičiaus, t.y. tarininis pažyminys su veiksmu formaliai nederinamas. To negalima pasakyti apie lietuvių kalbą. Lietuviškame šių sakinių vertime žodis *verärgert* turės keturis atitikmenis, kurių galūnes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

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

Todėl vokiško sakinio *Der Vater kam gestern verärgert* sintaksinė struktūra (žr. 18 pav.), verčiant šį sakinį į lietuvių kalbą, turi būti keičiama lietuviško sakinio *Tėvas vakar grįžo piktas* sintaksine struktūra (žr. 19 pav.).

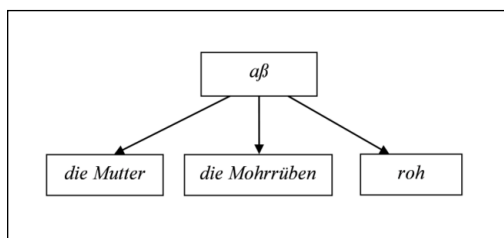


18 pav. Vokiško sakinio *Der Vater kam gestern verärgert* sintaksinė struktūra

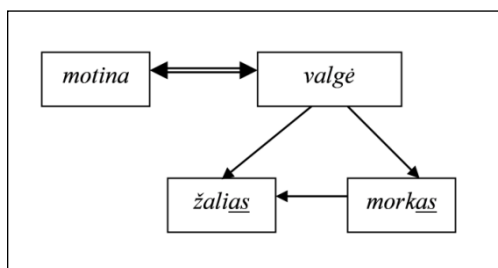


19 pav. Lietuviško sakinio *Tėvas vakar grįžo piktas* sintaksinė struktūra

Informacijos trūkumas ypač gerai matomas tų sakinių struktūroje, kurie turi tarininį pažyminį ir papildinį, nes tarininis pažyminys 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 20-ame paveikslėlyje, neaišku, kas 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 *Motina morkas valgė žalias* struktūrą (žr. 21 pav.).



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



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

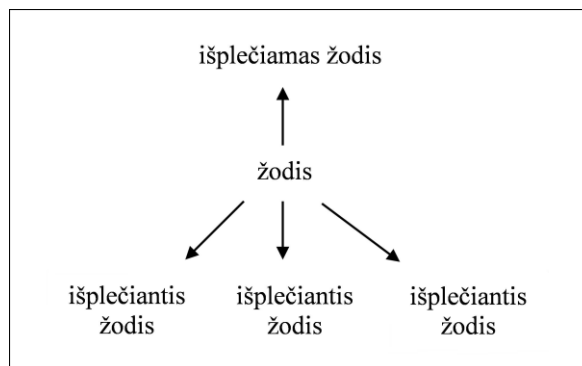
Iš šių dviejų pavyzdžių gerai matyti, kad vieną ir tą patį grafinį vaizdą turinčias vokiškų sakinių struktūras, pateiktas 18-ame ir 20-ame paveikslėliuose (grafas susideda iš keturių viršūnių ir trijų linijų, kurios išeina iš vienos viršūnės ir eina į likusias tris, neturinčias tarpusavyje ryšio), atitinka dvi skirtingos lietuviškų sakinių struktūros (žr. 19 pav. ir 21 pav.). Vadinas, vokiško sakinio grafiniame vaizde (priklausomybių medyje) esančios informacijos neužtenka, jei norima be klaidų sugeneruoti į lietuvių kalbą išverstą sakinį.

Jeigu bandytume žūtbūt išlaikyti medžio grafinį vaizdą 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 atsiras kito tipo klaidų – žodžio vietos sakinyje neteisingas parinkimas.

Tarininis pažyminys būtinai turi būti atskirtas nuo veiksnio kitais žodžiais ar eiti po jo. Tai vienas iš labai nedaugelio atvejų lietuvių kalboje, kai žodžių tvarka turi įtakos nustatant sakinio dalis. Čia nepažeidžiamas laisvos žodžių tvarkos principas lietuvių kalbos sakiniuose. Tarininis pažyminys gali būti ir pradžioje, ir viduryje, ir sakinio gale (*Turtinga parvažiuos dukra iš Amerikos. Dukra parvažiuos turtinga iš Amerikos. Dukra iš Amerikos parvažiuos turtinga.*). Tačiau tarininis pažyminys niekada negali būti prieš veiksni, nes tokiu atveju jis tampa atributiniu pažyminiu. *Turtinga dukra parvažiuos iš Amerikos* – tai jau pastovi veiksnio ypatybė, o tuo (ypatybės pastovumu) ir pasireiškia esminis skirtumas tarp atributinio ir tarininio pažyminių. (Reikėtų pridurti, kad tarininis pažyminys gali būti atkeliamas į atributinio pažyminio vietą sakytinėje kalboje, kai jo predikatiškumą rodo intonacija, tačiau tekste jokių duomenų apie intonaciją nebūna ir į ją galima atsižvelgti tik apdorojant garsynus, bet šiame straipsnyje jie nebus plačiau aprašomi.) Su papildiniu susijęs tarininis pažyminys būtinai turi būti po papildinio (*Seserį ji prisimena jauną, gražią ir turtingą.*), kitaip jis tampa atributiniu pažyminiu (*Ji dažnai lankydavo savo jauną, gražią ir turtingą seserį.*). Terodant sintaksinėje struktūroje tik tą ryšį, kur vyksta galūnių derinimas, t.y. tarininį pažyminį traktuojant kaip atributinį, būtų daroma klaidų ir atliekant vertimus kita kryptimi, pavyzdžiui, iš lietuvių kalbos į vokiečių kalbą, nes atributinis pažyminys vokiečių kalboje derinamas galūnėmis su išplečiamu žodžiu (*schöne Mädchen, schöne Katze, schöner Tag*) taip pat, kaip ir lietuvių kalboje, o tarininis pažyminys turi aplinkybės formą ir nederinamas galūnėmis su tuo žodžiu (veiksniu ar papildiniu), kurį jis išplečia. Taigi, neparodžius lietuviškame sakinyje tarininio pažyminio ryšio su tariniu, vokiško sakinio sintezės metu bus neteisingai nustatoma tarininio pažyminio galūnė. Automatinio vertimo sistemose išverstas sakiny generuojamas pasinaudojant sintaksinės analizės rezultatais [Hutchins 1986: 53]. Neparodžius lietuviško sakinio sintaksinėje struktūroje tarininio pažyminio ryšio su veiksniu ar papildiniu, morfologinės sintezės metu trūktų informacijos apie tai, kokia

turi būti jo galūnė, o, neparodžius ryšio su tariniu, sintaksinės sintezės metu nebūtų galima teisingai nustatyti jo vietos sakinyje.

Todėl sudarant lietuvių kalbos sakinių sintaksinę struktūrą tarininiam pažyminiui bus pateikiami du žodžiai, nuo kurių jis priklauso, tuo pažeidžiant priklausomybių gramatikos principą, kur teigiama, kad kiekvienas žodis gali turėti kelis nuo jo priklausomus žodžius, bet pats tegali priklausyti tik nuo vieno. Apibendrinta priklausomybių gramatikos medžio viršūnė pateikiama 22-ame paveikslėlyje. Kiekvienoje viršūnėje yra žodis, kuris gali turėti vieną ar kelis jį išplečiančius žodžius, tačiau pats jis tegali išplėsti tik vieną virš jo esantį žodį [4 internetinė nuoroda].



22 pav. Priklausomybių medžio viršūnės ryšiai su kitomis viršūnėmis

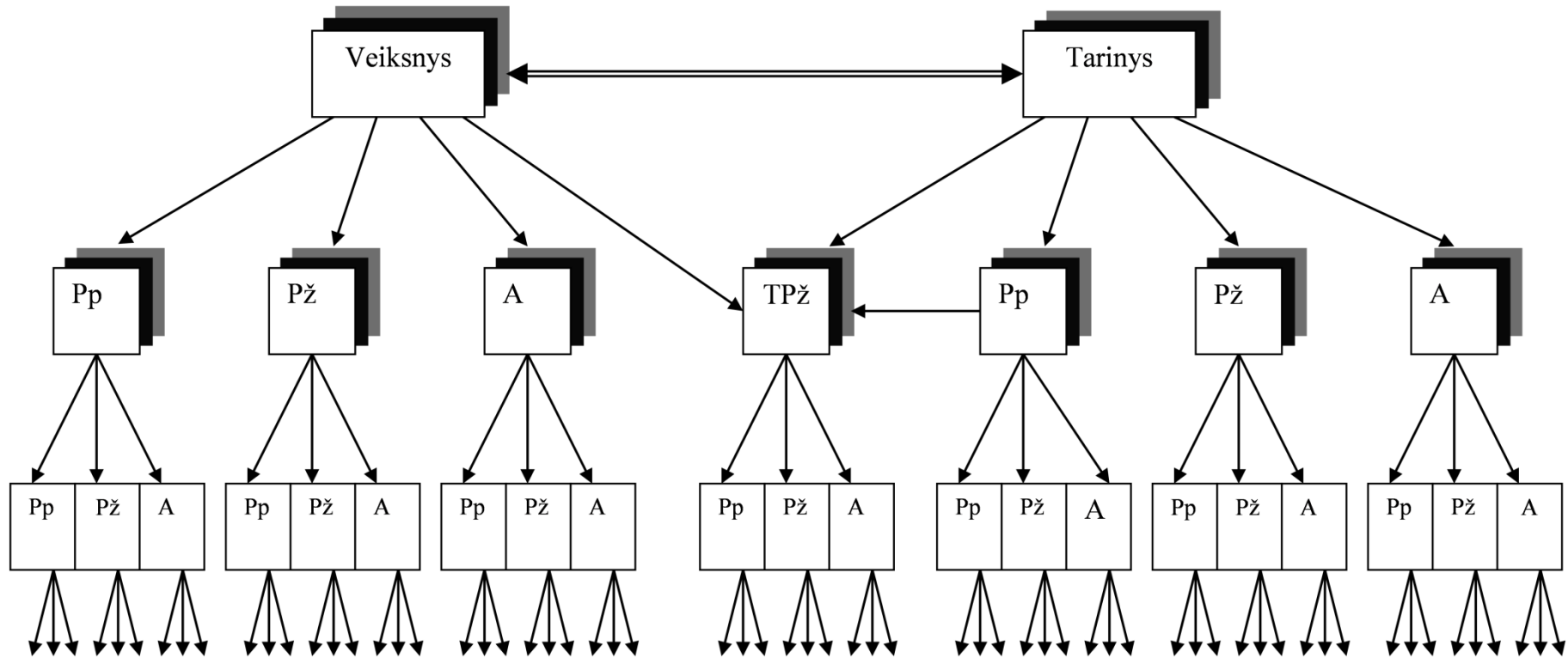
Sintaksinės analizės užduotis surasti kiekvienam sakinio žodžiui visus išplečiančius žodžius ir nustatyti, kurį žodį jis pats išplečia.

3.2.3. Apibendrinta sakinio sintaksinė struktūra

Lietuvių kalbos sakiniams struktūriškai pavaizduoti būtina naudoti grafą, nes medis iš principo negali atspindėti visos sintaksinės informacijos, esančios lietuviškame sakinyje. Atliekant sintaksinę analizę automatiškai (kompiuteriu), patogiu turėti grafu pavaizduotą apibendrintą sintaksinę struktūrą, kuri apimtų visus galimus lietuviško sakinio atvejus, t. y. būtų bendra visiems lietuvių kalbos vientisiniams sakiniams [Šveikauskienė 2005: 407–418]: konkretus sakinyje turėtų aktyvuoti toje struktūroje kurį nors vieną pografą [Свами, Тхуласираман 1984: 13]. Kiekvienam bet kokio vientisinio sakinio žodžiui scheme turi būti numatyta vieta, į kurią jis bus dedamas analizės metu. Tokios schemas grafinis vaizdas pateiktas 23-iame paveikslėlyje. Grafo viršūnėms suteikti sakinio dalių pavadinimai rodo, kur turi būti dedamas žodis pagal nustatytą jo sintaksinę funkciją bei ryšius su kitais žodžiais. Taigi, apibendrinta struktūra – tai tarsi kontūrai. Kai turime konkretų sakinį, dėliojame jo žodžius į jiems tinkamas vietas (grafo viršūnes) pagal sintaksinę funkciją bei ryšius su kitais žodžiais. Surašius visus sakinio žodžius, nepanaudota grafo dalis atmetama – iš schemas pašalinamos tos grafo viršūnės, kurios lieka tuščios, t. y. į kurias neįdėtas joks žodis, o likusi grafo dalis pateikiama kaip nagrinėjamo sakinio sintaksinė struktūra.

Apibendrintoje lietuvių kalbos vientisinio sakinio struktūroje (žr. 23 pav.) iš veiksnio išeinančios rodyklės į visas tris antrines sakinio dalis, iš pirmo žvilgsnio, gali atrodyti prieštaraujantis lietuvių kalbos gramatikai atvejis, nes joje sakoma, kad papildinys gali jungtis tik prie veiksmažodžio, o pažyminys – 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*). Tarininis pažyminys turi dvigubus sintaksinius ryšius. Jis formaliai derinamas su veiksmiu ar papildiniu ir kartu šliejamas prie veiksmažodžio. Todėl apibendrintoje scheme 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. Šešėliai ant blokelių, atitinkančių sakinio dalis, žymi galimus vienaarūšių sakinio dalių atvejus.

Lietuvių kalbos vietišinio sakinio apibendrinta struktūrinė schema



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

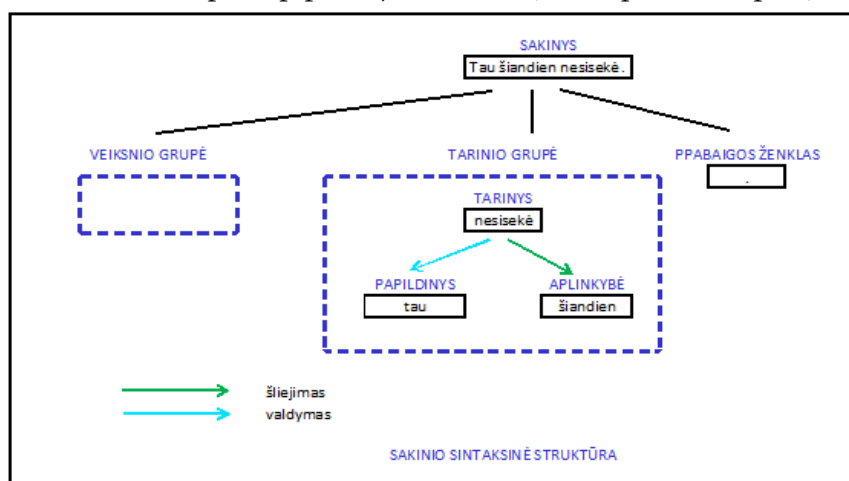
23 pav. Apibendrinta lietuvių kalbos vietišinio sakinio struktūrinė schema

3.2.4. Sakinio struktūra sintaksiniame tekстыne

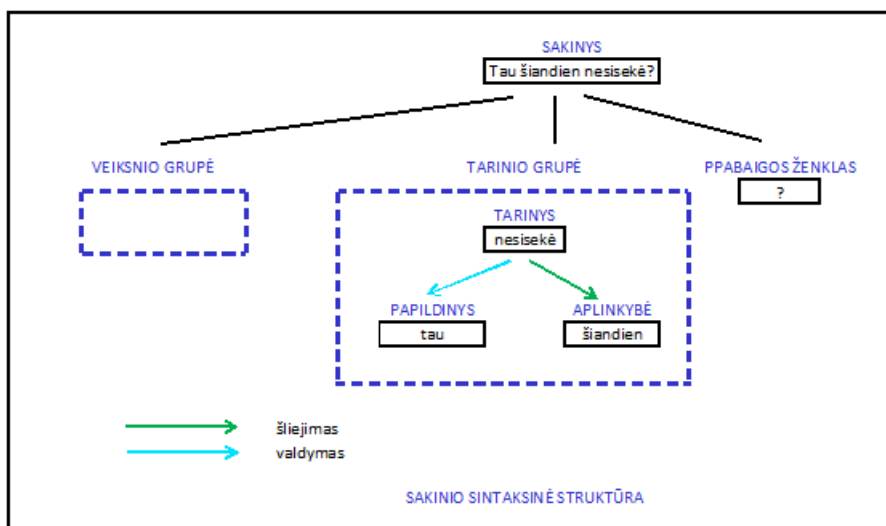
Belieka aptarti, kokia bus lietuvių kalbos sakinio sintaksinė struktūra pateikiant ją anotuotame tekстыne. Lietuvių kalboje žodžių tvarka yra laisva, todėl sakinio struktūros vaizdavimo principas pasirenkamas priklausomybių. Dichotominis požiūris naudingas lietuvių kalbos sakiniams. Taigi sintaksinės struktūros pagrindas turėtų būti apibendrintoje lietuvių kalbos sakinio schemoje pateiktas vaizdas (žr. 23 pav.).

Sakinio branduolio išskyrimas sintaksinėje struktūroje būtinas dėl to, kad geriau būtų matyti specifinė lietuvių kalbos savybė – galimybė praleisti tarinio jungtį, išreikštą veiksmažodžio *būti* esamojo laiko 3-uoju asmeniu, ir veiksni, išreikštą 1-ojo ar 2-ojo asmens įvardžiu. Ši informacija būtų ypač naudinga automatiniam vertime į kalbas (pvz., vokiečių, anglų), kuriose sakiny be tarinio jungties ar veiksnio yra netaisyklingas. Tuščia vieta jungties ar veiksnio pozicijoje atkreiptų dėmesį ir praneštų sistemai, kad šio žodžio iš paties sakinio ji negaus ir kad būtinai turi jį susigeneruoti pagal taisykles papildomai (veiksniui įvardį parinkti pagal tarinio galūnę, t.y. pagal jo asmenį ir skaičių, o tarinio formą, t.y. jo galūnę – pagal veiksnio asmens atitikmenį).

Atiduodant duoklę formaliosioms gramatikoms ir atsižvelgiant į tai, kad turbūt visi sintaksiškai anotuoti tekstynai pateikiami su schemomis, turinčiomis vieną pradinį simbolį, lietuvių kalbos sakinų sintaksinė struktūra bus vaizduojama pradiniu simboliu laikant sakinį, ir tai būtų elementas, perimtas iš frazių gramatikos. Tik jo skaidymas į veiksnio ir tarinio grupes, kaip siūlo lietuvių kalbininkai (Sirtautas, Grenda 1988:52), bus dar papildytas viena viršūne, kuri atspindės specifinį lietuvių kalbos bruožą – laisvą žodžių tvarką. Pirmame etape sakiny bus skaidomas į tris dalis: veiksnio grupę, tarinio grupę ir pabaigos ženklą. Pabaigos ženklo viršūnė yra labai svarbus dalykas, gal net svarbesnis, negu gali atrodyti iš pirmo žvilgsnio. Anglų ar vokiečių kalbose klausiamojo sakinio tipą rodo žodžių tvarka. Lietuvių kalboje kartais klaustukas gale sakinio yra vienintelis rodiklis, žymintis sakinio klausiamąją intonaciją. Du sakiniai, turintys absoliučiai vienodus žodžius ir jų tvarką, kartais tik pagal pabaigos ženklą gali būti priskirti tiesioginio ar klausiamojo sakinio tipui, pavyzdžiui, *Tau šiandien nesisekė.* ir *Tau šiandien nesisekė?* Tik pabaigos ženklas nulems išversto į anglų ar vokiečių kalbą sakinio žodžių tvarką, todėl šie sakiniai anotavimo metu bus braižomi taip, kaip parodyta žemiau (žr. 24 pav. ir 25 pav.).

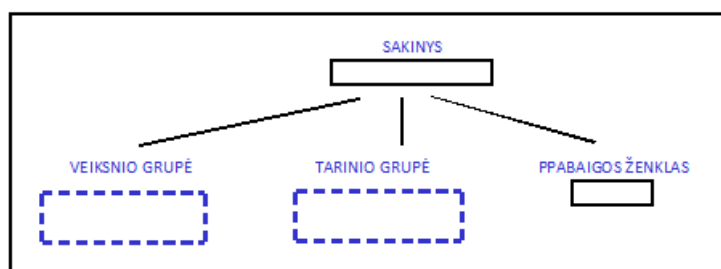


24 pav. Tiesioginio sakinio *Tau šiandien nesisekė.* schema



25 pav. Klausiamojo sakinio *Tau šiandien nesisekė?* schema

Apibendrintai galima pasakyti, kad lietuvių kalbos sakinių sintaksinė struktūra bus vaizduojama dviem etapais. Iš pradžių, pagal analogiją su frazių gramatika, sakinytis bus skaidomas į veiksnio bei tarinio grupes ir pabaigos ženklą (žr. 26 pav.). Į viršūnę, kurios žymė yra SAKINYS, įrašomi visi sakinio žodžiai taip, kaip mes juos matome tekste. Sakinio pabaigos simbolis dedamas į atskirą viršūnę, kurios žymė PABAIGOS ŽENKLAS. Veiksnių ir tarinių grupės, pažymėtos punktyrais, apima atitinkamai veiksnį ir tarinį su visais juos išplečiančiais žodžiais, kurie vaizduojami jau priklausomybių gramatikos principu. Veiksnių grupėje pradinis simbolis yra veiksnys, tarinių grupėje – tarinys.



26 pav. Pirmasis sintaksinės analizės etapas – frazių gramatikos metodas

Anotuojant lietuvių kalbos sakinius, į grafo viršūnes bus įrašomi žodžiai, o viršūnių žymės turės sakinio dalių pavadinimus. Linijos, jungiančios viršūnes, bus nuspalvintos skirtingai pagal sintaksinio ryšio tipą.

4. Išvados

Lietuvių kalbos sakiniams pavaizduoti labiau tinka priklausomybių nei frazių metodas. Frazių gramatikoje sakinio sintaksinė struktūra braižoma pateikiant sakinį tokį, koks jis yra tekste, žodžiai sugrupuojami į frazes, ir frazių pavadinimai užrašomi virš sakinio žodžių. Priklausomybių gramatikoje sintaksinė struktūra nesusijusi su žodžių tvarka sakinyje. Ten parodomi tiesioginiai sintaksiniai ryšiai tarp žodžių, nesiejant jų su linijiniu išsidėstymu. Dėl laisvos žodžių tvarkos lietuvių kalboje sintaksinė struktūra pavaizduota frazių metodu darosi labai paini ir sudėtinga, kartais gali būti daug susikertančių linijų. Priklausomybių gramatikos metodu pavaizduota struktūra yra daug aiškesnė ir suprantamesnė.

Lietuvių kalbos sintaksinei struktūrai pavaizduoti būtina naudoti grafą, nes medis iš principo negali atspindėti visos sintaksinės informacijos, esančios lietuviškame sakinyje. Savita lietuvių kalbos ypatybė yra dvigubi tarinio pažyminio sintaksiniai ryšiai. Jis priklauso ir nuo tarinio, ir nuo veiksnio ar papildinio. Su veiksnium ar papildiniu jis derinamas galūnėmis, o šliejimas prie tarinio nulemia jo vietą sakinyje, skirtingą nuo atributinio pažyminio. Pavaizdavus abu šiuos ryšius, struktūra nebeatitinka medžio apibrėžimo, nes susidaro uždaras kelias – ciklas, o pavaizdavus kurį nors vieną šių ryšių bus prarasta sintaksinė informacija.

Apibendrintai lietuvių kalbos vientisinio sakinio sintaksinei struktūrai pavaizduoti naudojama tradicinės gramatikos forma ir priklausomybių gramatikos metodas. Iš frazių gramatikos perimamas pradinis simbolis, t.y. juo laikomas sakinyje, tačiau jau pirmame etape skaidant jį į veiksnio ir tarinio grupes papildomai nurodomas pabaigos ženklas atskiroje viršūnėje. Tuo frazių gramatikos panaudojimas ir baigiamas. Toliau veiksnio ir tarinio grupės vaizduojamos priklausomybių metodu pradiniais simboliais laikant atitinkamai veiksnį ir tarinį.

Šaltiniai

1 internetinė nuoroda:

http://en.wikipedia.org/wiki/Phrase_structure_grammar#Constituency_relation

2 internetinė nuoroda

<http://www.ims.uni-stuttgart.de/projekte/TIGER/TIGERSearch/doc/html/TigerXML.html>

3 internetinė nuoroda:

<ftp://ling.upenn.edu/studentpapers/creswell/rambow-et-al.pdf>

4 internetinė nuoroda:

<http://www.cl.uni-heidelberg.de/~hellwig/dug-2003.pdf>

5 internetinė nuoroda:

http://www.lki.lt/LKI_LT/index.php?option=com_content&view=article&id=803&Itemid=153

DLKG – Ambrazas, V. (red.) 1997: *Dabartinės lietuvių kalbos gramatika*, Vilnius: Mokslo ir enciklopedijų leidybos institutas.

Literatūra

1. Agel, V. 2000: *Valenztheorie*, Tübingen: Narr.
2. Allen, J. 1987: *Natural Language Understanding*, Amsterdam: The Benjamin / Cummings Publishing Company.
3. Ambrazas, V. (red.) 1997: *Dabartinės lietuvių kalbos gramatika*, Vilnius: Mokslo ir enciklopedijų leidybos institutas.
4. Batori, I. S., Lenders, W. 1989: *Computerlinguistik: Ein internationales Handbuch zur computergestützten Sprachforschung und ihrer Anwendung*, Berlin: Walter de Gruyter.
5. Greenbaum, S. 1996: *Oxford English Grammar*, Oxford: Oxford University Press.
6. Helbig, G., Buscha, J. 1989: *Deutsche Grammatik: Ein Handbuch für Ausländer*, Leipzig: VEB Verlag Enzyklopädie.
7. Holvoet, A. 2009: Argumentų hierarchijos ir gramatinės funkcijos. R. Mikulskas (red.), *Gramatinių funkcijų prigimtis ir raiška*, Vilnius: Academia Salensis.
8. Holvoet, A. 2009: *Bendrosios sintaksės pagrindai*. Vilnius: Academia Salensis.
9. Hutchins, W. J. 1986: *Machine Translation: Past, Present, Future*, Chichester: Ellis Harwood Limited.
10. Hutchins, W. J., Sommers, H. L. 1982: *An Introduction to Machine Translation*. London: Academic Press.
11. Kay, M., Gawron, J. M., Norvig, P. 1994: *VerbMobil: A Translation System for Face-to-Face Dialog*, Stanford: CSLI.
12. Köhler, R., Altmann, G., Grzybek, P. 2012: *Quantitative Syntax Analysis*, Berlin / Boston: Walter de Gruyter.

13. Labutis, V. 1998: *Lietuvių kalbos sintaksė*, Vilnius: Vilniaus universiteto leidykla.
14. Müller, H. M. 2002: *Arbeitsbuch: Linguistik*, Paderborn: Verlag Ferdinand Schöningh.
15. Schwanke, M. 1991: *Maschinelle Übersetzung: Ein Überblick über Theorie und Praxis*, Berlin: Springer-Verlag.
16. Šveikauskienė, D. 2005: Graph Representation of the Syntactic Structure of the Lithuanian Sentence. *Informatica*. Vol. 16, Nr. 3, p. 407–418.
17. Vaišnienė, D., Zabarskaitė, J. 2012: Lietuvių kalba skaitmeniniame amžiuje. G. Rehm, H. Uzhkoreit (red.), *Baltųjų knygų serija*, Heidelberg, New York, Dordrecht, London: Springer.
18. Winograd, T. 1983: *Language as a Cognitive Process. Volume I: Syntax*, London: Addison-Wesley Publishing Company.
19. Свами, М., Тхуласираман, К. 1984: *Графы, сети и алгоритмы*, vertimas į rusų k. Москва: Мир.

Gauta: 2013 vasario 4 d.

Priimta: 2013 kovo 22 d.

DAIVA ŠVEIKAUSKIENĖ
Lietuvių kalbos institutas
Vileišio g. 5–216
LT-10308 Vilnius
[daiva.fmf@gmail.com]