
MORFOLOGINIO DAUGIAREIKŠMIŠKUMO TIPOLOGIJA

Erika Rimkutė

Vytauto Didžiojo universitetas, Donelaičio g. 58, LT-3000 Kaunas, Lietuva

ĮVADAS

Atsiradus naujai kalbotyros šakai – tekstynų lingvistikai – imta kitaip analizuoti kalbą. Analizės pobūdis yra nulemtas pastaruoju metu gerokai ištobulėjusių kompiuterinių technologijų, kurios pritaikomos jau ir lietuvių kalbai. Beveik dešimt metų veikiantis Vytauto Didžiojo universiteto Kompiuterinės lingvistikos centras daug prisidėjo prie lietuvių kalbos kompiuterizavimo: sudarytas ir toliau pildomas 100 mln. žodžių dabartinės lietuvių kalbos tekstynas, leidžiantis tirti iki šiol neapreptą kalbos kiekį, atlikti statistinę, kalbinę tekstų analizę. Kompiuteriai ir kompiuterinės technologijos atvėrė naujas galimybes, parodė dažnumą, reguliarius kalbos vartosenos modelius. Tai ne vien tik palengvino tyrimą, bet ir suteikė jam visai naują kokybę, atvėrė naujas tyrimo perspektyvas (Marcinkevičienė, 2000: 37).

Kompiuterinės lingvistikos centre yra sudarinėjamas vieno milijono žodžių tekstynėlis. Jame esantiems žodžiams ar žodžių formoms nurodomi antraštiniai pavidalai, t. y. lemos, ir tų formų morfologinės charakteristikos. Tam naudojama Vytauto Zinkevičiaus sukurta kompiuterinė programa „Lemuoklis“. Automatiškai sulemavus ir morfologiškai pažymėjus tekstus išryškėjo morfologinis daugiareikšmiškumas. Paaiškėjo, kad tradiciniai žodynai ir gramatikos neaprašo realios kalbos vartosenos. Susidurta su tokiais atvejais, kurie be specialios programinės įrangos nepastebimi.

Šiuo metu automatinio būdu morfologiškai anotuota ir lingvisto peržiūrėta tik nedidelė kompiuterinio lietuvių kalbos tekstyno dalis – maždaug 700 tūkst. žodžių. Per labai trumpą laiką galima automatiškai pažymėti visą 100 mln. žodžių tekstyną, tačiau tokios pažymos nėra labai naudingos dėl didelio daugiareikšmiškumo. Anotuotą tekstynėlį reikia peržiūrėti ir rankomis panaikinti daugiareikšmius variantus, įrašyti neatpažintų ar klaidingai pažymėtų žodžių lemas bei morfologines charakteristikas.

Šiame straipsnyje norima pristatyti ir apibrėžti mažai tyrinėtą morfologinį daugiareikšmiškumą, automatinio būdu morfologiškai anotuotame tekстыne nustatytus ir iki šiol nenagrinėtus pagrindinius morfologinio daugiareikšmiškumo tipus bei rūšis.

AUTOMATINĘ MORFOLOGINĘ ANALIZĘ ATLIEKANTI PROGRAMA „LEMUOKLIS“

V. Zinkevičiaus sukurta kompiuterinė programa „Lemuoklis“ (programos pavadinimas yra atsiradęs iš angliško termino *lemmatizing* (kilusio iš lotynų kalbos žodžio *lemma*, daugiskaita *lemmata* – antraštinė, žodyninė leksemos forma) (Zinkevičius, 2000: 245) gali pateikti antraštinį rašytinės žodžio formos pavidalą, vadinamąją lemą, pavyzdžiui, forma *universiteto* sulemuojama kaip daiktavardis *universitetas*. „Lemuoklis“ taip pat gali nurodyti ir morfologines charakteristikas. Minėta forma būtų laikoma vyriškosios giminės daiktavardžio vienaskaitos kilmininku.

Taigi tekstyno žymėjimą galima suskirstyti į tris etapus: pirmiausiai grynas tekstas lemuojamas – pateikiama tekстыne pavartoto žodžio antraštinė forma, t. y. lema. Antrojo etapo metu gali būti pateiktas žodžio formos morfologinis apibūdinimas. Toliau turėtų būti nustatomi kelis antraštinius pavidalus turintys žodžiai, t. y. vienareikšminama, nes, pavyzdžiui, forma *laimės* gali būti sulemuota ir kaip *laimė*, ir kaip *laimėti*. Tam tikslui reikalinga speciali programa, kurios pagrindinė funkcija būtų dviprasmybių panaikinimas (angliškas terminas – *ambiguity resolution*) (Zinkevičius, 2000: 245; Marcinkevičienė, 2000: 22). Deja, tokios programos Lietuvoje dar nėra, todėl morfologinis daugiareikšmiškumas anotuotuose tekstuose yra gana didelis¹ (žr. 2 pav.).

Šiuo metu vienintelė Lietuvoje automatinę morfologinę analizę atliekanti programa padėjo pastebėti dažną, bet iš konteksto nenustatomą daugiareikšmiškumą. Tik šios programos anotuoti tekstai gali būti morfologinio daugiareikš-

¹ Kiek ir kokius žodžių formų antraštinius pavidalus bei morfologinius požymius pateikia automatinės morfologinės analizės programa, galima spręsti iš nedidelės „Lemuoklio“ apdoroto teksto ištraukos:

vieversiai

dktv <vieversys> dktv vyr.gim dgsk V dktv vyr. gim dgsk Š

į

prln <į> prln

Dzūkiją

<??>

nebuvo

vksm <nebūti(-ėra,-ebuvo)> vksm nesngr tiesiog. nuos būt. kart. I vnsk IIIasm vksm nesngr tiesiog. nuos būt. kart. I dgsk IIIasm

taip

dll <taip> dll prvks <taip> prvks

anksti

dktv <ankstis> dktv vyr. gim vnsk Š

bdvr <ankstas> bdvr nelygin. I neįvardž vyr. gim dgsk V

bdvr <ankstus> bdvr nelygin. I neįvardž mot. gim vnsk V

prvks <anksti> prvks nelygin. I

sugrįžę

vksm <sugrįžti(-žta,-žo)> dlv nesngr veik. r būt. kart. I neįvardž vyr. gim dgsk V dlv nesngr veik. r būt. kart. I neįvardž bevrd. gim

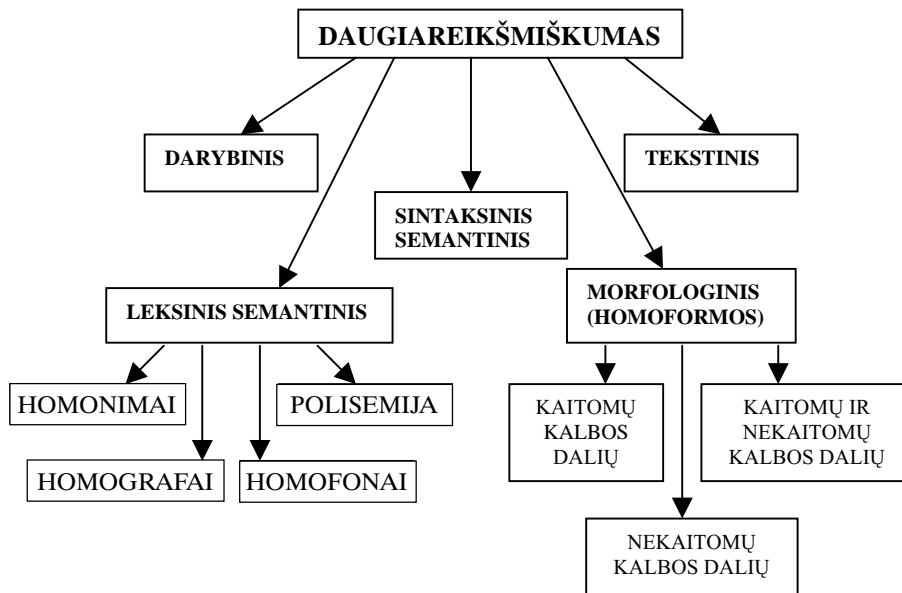
miškumo analizės objektas. Pagrindinis trūkumas yra „Lemuoklio“ nesugebėjimas analizuoti konteksto. Ši programa nenaudoja ir neturi informacijos apie semantiką. „Lemuoklis“ žodžių reikšmes sugeba nustatyti ne naudodamas kokius nors žodžių formų sąrašus su nurodytais tų formų morfologiniais apibūdinimais, o turėdamas lietuviškų žodžių šaknų sąrašą. Prie kiekvienos šaknies nurodomi naudojami skaitmeniniai kaitybos ir darybos modeliai. V. Zinkevičiaus sukurta programa taip pat nenaudoja žodžių ar žodžių formų dažninių charakteristikų ir neturi apie jas informacijos. Dėl šių priežasčių atsiranda tokių daugiareikšmiškumo atvejų, su kuriais realiai susidurti beveik neįmanoma. Tai apsunkina morfologinių formų atpažinimą, trukdo automatinei analizei, kelių kalbų lygiagrečiam nagrinėjimui. Nepaisant kai kurių trūkumų, morfologinės pažymos, kad ir ne visada tikslios, padeda kitaip analizuoti kalbą: pamatoma tai, kas atrodo visiškai neįmanoma, nors teoriškai teisinga.

DAUGIAREIKŠMIŠKUMO SAMPRATA

Šiame straipsnyje vartojamas terminas *daugiareikšmiškumas*, kuris, atsižvelgiant į pasirinktą aspektą, gali būti suvokiamas įvairiai. Galima nagrinėti kalbos, kaip sistemos, daugiareikšmiškumą: kas nulemia, kad vienais atvejais kalba yra beveik nedaugiareikšmė, o kitais išryškėja didžiulis daugiareikšmiškumas, kiek tai priklauso nuo analizės priemonių. Taigi įmanoma tirti kalbos daugiareikšmiškumą pagal tai, kas ją analizuoja – žmogus ar žmogus, padedamas kompiuterio, ir kaip analizuoja – susietus ar pavienius žodžius. Ko gero, daugiareikšmiškumas akivaizdžiausias, kai analizuojama pagal kalbos lygmenis. Daugiareikšmiškumas išryškėja netgi darybiniame lygmenyje. Ypač daug neaiškumų atsiranda leksiniame semantiniame, morfologiniame, taip pat ir sintaksiniam semantiniame bei tekstiniame kalbos lygmenyje (1 pav.). Tik automatinės morfologinės analizės metu paaiškėjo, kad galima analizuoti daugiareikšmiškumą pagal tikėtinumą: jis gali būti tikėtinas ir neįtikėtinas.

Morfologiškai anotuotas tekstynas – dar visai netyrinėtas reiškinys, todėl reikia apsibrėžti ir pagrindinius, su automatine morfologine analize susijusius terminus. Iki šiol buvo vartotas terminas *daugiareikšmiškumas*, kuris suprantamas kiek kitaip nei įprasta: čia kalbama ne apie žodžių (leksemų), o apie morfologinių formų daugiareikšmiškumą. Taigi analizuojamas morfologinis daugiareikšmiškumas, o ne polisemija. Pasirinktas būtent šis terminas, nes jis apima tokius kalbos reiškinys, kuriems sunku rasti vienareikšmį, aiškų pavadinimą. Anglų kalboje daugiareikšmiai vienetai trumpai ir aiškiai vadinami *ambiguity*, o jų panaikinimas – *ambiguity resolution*.

Iki šiol lietuvių kalbotyroje morfologiniam daugiareikšmiškumui nebuvo skirta daug dėmesio – išsamiausiai kalbotyroje nagrinėti tik leksiniai homonimai. Todėl ir kyla problemų, kaip įvardyti morfologinį daugiareikšmiškumą bei atskirus jo tipus. Be to, nėra visiškai aiški lietuvių kalbotyroje homonimiją įvardijanti terminija (plg. Jakaitienė, 1980: 42–44; LK I, 1995: 23–24; Drotvinas, 1983: 91–92, 125; KTŽ, 1990: 79–80; Urbutis, 1956: 4; LKE, 1999: 247–249), todėl ir pasirinktas tekstynų lingvistikos sričiai naujas, bendresnis terminas. Tikimasi, kad ilgainiui jis prigis.



1 pav. Daugiareikšmiškumas pagal kalbos lygmenis

Terminą *daugiareikšmiškumas* galima vartoti kalbant ne tik apie darybinį, morfologinį bei leksinį semantinį, bet ir apie sintaksinį semantinį, tekstinį daugiareikšmiškumą bei kitus į klasifikacijos rėmus sunkiai išspraudžiamus atvejus. Taigi *daugiareikšmiškumas* suvokiamas kaip reiškinys, apimantis: a) kaitomas ir nekaitomas; b) skirtingų ir tų pačių kalbos dalių; c) tam tikromis formomis, kirčio vieta, priegaide ar balsio kiekybe besiskiriančias ir visiškai sutampančias; d) skirtingai rašomas, bet vienodai tariamas; e) skirtingai tariamas, bet vienodai rašomas žodžių formas ar žodžius; f) nevienareikšmes sudedamąsias žodžių dalis; g) skirtingas poliseminių ir homoniminių žodžių reikšmes ar šiaip dėl neaiškių žodžių ryšių, platesnio konteksto nežinojimo kylančias dviprasmybes. Tai bendriausias, įvairius daugiareikšmiškumo lygmenis apimantis apibrėžimas, vartojamas platesne reikšme negu *homonimija*.

Morfologinis daugiareikšmiškumas – tai tik viena nedidelė daugiareikšmiškumo dalis, kuriai priskiriamos: a) kaitomos ir nekaitomos; b) skirtingų ir tų pačių kalbos dalių; c) tam tikromis formomis, prozodiniais elementais besiskiriančios ir visiškai sutampančios žodžių formos ar žodžiai.

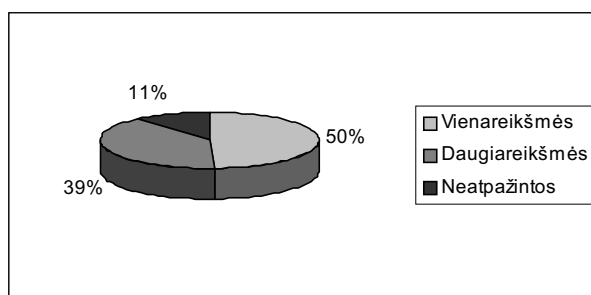
Morfologinį daugiareikšmiškumą sudaro homoformos, kurios suskirstytos pagal tai, ar sutampa kaitomos, ar kaitomos ir nekaitomos, ar tik nekaitomos kalbos dalys (žr. 1 pav.). Lingvistai nesutaria, ką reikėtų laikyti homoformomis. Daug neaiškumų kyla, jei analizuojamos prozodiniais elementais, t. y. kirčio vieta, priegaide ar balsio kiekybe, besiskiriančios homoformos, pavyzdžiui, *neturiu katės – kieme pešasi dvi kėtės*. Nesutariama, ar tai homoformos, ar homografa (šiam straipsnyje laikoma homoformomis).

Šiame tyrinėjime per daug nesigilinama į terminijos nenuoseklumus – tai ne šio straipsnio objektas. Didžiausias dėmesys skiriamas autorės apibrėžtam

morfologiniam daugiareikšmiškumui ir pagrindinėms jo rūšims. Todėl kai kurie daugiareikšmiškumo įvardijimai skiriasi nuo tradicinių kalbotyroje vartojamų terminų. Pagrindinis skirtumas yra tas, kad homoformomis laikomi ir tradiciškai homografams priskiriami daugiareikšmiai atvejai, t. y. tokie, kurie tariami skirtingai. Toks analizės aspektas pasirinktas dėl terminijos nevienodumo ir dėl to, kad tiriamos rašytinės nekirčiuotos formos. Šie kriterijai lėmė tai, kad neanalizuoti fonetiniai skirtumai.

MORFOLOGINIS DAUGIAREIKŠMIŠKUMAS ANOTUOTAME TEKSTYNE

Morfologiniam daugiareikšmiškumui nustatyti buvo išanalizuota 10 000 morfologiškai anotuotų žodžių formų, pavartotų viename dienraščio „Lietuvos rytas“ numeryje. Iš jų beveik pusė buvo vienareikšmės, t. y. joms „Lemuoklis“ pateikdavo tik vieną antraštinę lemą ir tik vieną morfologinę charakteristiką. Visiškai neatpažintų formų buvo apie 11%. Tai dažniausiai tikriniai daiktavardžiai, užsienio kalbų žodžiai ar sutrumpėjusios lietuviškų žodžių formos. Rasta beveik 40% daugiareikšmių žodžių formų ar žodžių (2 pav.). Panašus daugiareikšmių formų skaičius randamas ir Browno tekстыne – čia apie 40% žodžių būdingas morfologinis daugiareikšmiškumas. Kai kurios formos gali turėti net septynias pažymas – toks yra žodelis *still* (Charniak, 1993: 49–50).



2 pav. Vienareikšmių ir daugiareikšmių žodžių ar žodžių formų santykis

Daugiareikšme laikoma ta tekste pavartota žodžio forma ar žodis, kuriam „Lemuoklis“ pateikia: a) bent dvi skirtingas lemas, pavyzdžiui, *laužo* – daiktavardžio *laužas* vnsk K ir veiksmažodžio *laužyti* esam. I III asm²; b) ar bent du skirtingus tos pačios lemos morfologinius požymius, pavyzdžiui, *mama* – sutampa šio daiktavardžio vienaskaitos vardininkas ir šauksmininkas. Dažnai automatinio būdu morfologiškai anotuojant tekstą tai pačiai žodžio formai pateikiamos ir skirtingos lemos, ir skirtingos kelių lemu morfologinės charakteristikos, pavyzdžiui:

galima

bdvr <galimas> bdvr nelygin. I neįvardž mot. gim vnsk V bdvr nelygin. I neįvardž mot. gim vnsk Įn bdvr nelygin. I neįvardž bevrd. gim

vksm <galėti(-i,-ėjo)> dlv nesngr neveik.r esam.l neįvardž mot.gim vnsk V dlv nesngr neveik.r esam.l neįvardž mot.gim vnsk Įn dlv nesngr neveik.r esam.l neįvardž bevrđ.gim

Taigi forma *galima* priskiriama trims daugiareikšmiškumo rūšims: 1) veiksmožodžių ir būdvardžių homoformoms, 2) dalyvių ir 3) būdvardžių linksnių sutapimui. Kaip matyti antrajame paveiksle, 39% yra morfologiškai daugiareikšmių formų, bet jų morfologinių požymių skaičius yra gerokai didesnis (morfologiškai daugiareikšmių formų rasta 3935, o jų morfologinių požymių – 5986 atvejai).

1 lentelė. Daugiareikšmių formų morfologinių požymių statistika						
Bendras daugiareikšmių formų morfologinių požymių skaičius	Kaitomų kalbos dalių homoformos		Kaitomų ir nekaitomų kalbos dalių homoformos		Nekaitomų kalbos dalių homoformos	
	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.
5986	4228	70,6	464	7,8	1294	21,6

Homoformų tipai

Kaip minėta, pagal kalbos dalių sutapimus skirti trys homoformų tipai: kaitomų, kaitomų ir nekaitomų bei tik nekaitomų kalbos dalių homoformos. Homoformų tipas – tai tam tikros sutampančios kalbos dalys.

Kaitomų kalbos dalių homoformos – tai tų pačių ar skirtingų kalbos dalių morfologiškai sutampančios formos, pavyzdžiui, *sukalė šuniui būdą – išsiugdė gerą būdą*. Sutampa forma *būdą*, kuri gali būti daiktavardžių *būda* arba *būdas* vienskaitos galininkas. Taigi sinkretiškos tai pačiai kalbos daliai priskiriamų žodžių formos. Sutampančias skirtingas kalbos dalis iliustruoja toks pavyzdys: *jo kalbà trumpa – jis kalba garsiai*. Šiuo atveju vienodai rašomos kai kurios daiktavardžio *kalba* ir veiksmožodžio *kalbėti* formos.

Kaitomų kalbos dalių homoformos sudaro didžiausią morfologinio daugiareikšmiškumo dalį (70,6%) (1 lentelė). Jos smulkiau skirstomos į tų pačių kaitomų ir skirtingų kaitomų kalbos dalių homoformas (2 lentelė).

2 lentelė. Kaitomų kalbos dalių homoformos				
Bendras skaičius	Sutampančios tos pačios kalbos dalys		Sutampančios skirtingos kalbos dalys	
	Sk.	Proc.	Sk.	Proc.
4228	3297	78	931	22

² Šiame straipsnyje vartojamos kiek kitokios santrumpos, nes tokius sutrumpinimus pateikia automatinę morfologinę analizę vykdanči programa „Lemuoklis“ (žr. santrumpų sąrašą).

Sutampančios tos pačios kaitomos kalbos dalys – tai tokios homoformos, kurias sudaro tos pačios kalbos dalies morfologiškai daugiareikšmės formos, pavyzdžiui, daugiareikšmės to paties arba dviejų (kartais ir daugiau) skirtingų daiktavardžių formos (žr. anksčiau minėtą pavyzdį – *būda*).

Ši morfoliginio daugiareikšmiškumo dalis sudaro 78% visų kaitomų kalbos dalių homoformų. Joms priskirtos daiktavardžių, veiksmažodžių, būdvardžių, įvardžių ir skaitvardžių homoformos. Nagrinėti sutampantys tų kalbos dalių linksniai, asmenys ar kitos kaitybinės formos. Iš šių homoformų dažniausios daiktavardžių ir veiksmažodžių homoformos (3 lentelė). Šioms homoformoms priskirtinas ir linksnių sinkretizmas, būdingas daiktavardžiams, būdvardžiams, įvardžiams, skaitvardžiams bei dalyviams, pavyzdžiui, *pakviestōs jaunōs mamōs vaikas – susirinko pākviestos jāunos māmōs*.

3 lentelė. Sutampančių tų pačių kaitomų kalbos dalių homoformos										
Bendras skaičius	Dktv		Vksm		Bdvr		Įvrd		Sktv	
	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.
3297	1405	42,6	1196	36,3	355	10,8	243	7,4	98	2,9

Antrasis kaitomų kalbos dalių homoformų tipas – **skirtingų kaitomų kalbos dalių** sutapimai, t. y. kai ortografiškai sutampa skirtingoms kalbos dalims priskiriamų žodžių formos, pavyzdžiui, daiktavardis ir įvardis: *šis vyras neturi pačios – tai mano pačios darbas*. Šios homoformos sudaro 22% morfoliginio daugiareikšmiškumo atvejų (2 ir 4 lentelės).

Skirtingų kalbos dalių homoformos gali būti dvinarės arba daugianarės priklausomai nuo to, kiek kalbos dalių sutampa.

Dvinarės homoformos – dviem skirtingoms kaitomoms kalbos dalims priskiriamų žodžių atsitiktinai rašytiniu pavidalu sutampančios tam tikros formos. Tai daiktavardžių ir veiksmažodžių, daiktavardžių ir įvardžių, daiktavardžių ir skaitvardžių, veiksmažodžių ir įvardžių homoformos (4 lentelė), pavyzdžiui, sutampa kai kurios daiktavardžio ir veiksmažodžio formos: *laimės kūdikis – jis laimės varžybas*.

Daugianarės homoformos – iš trijų ar daugiau kaitomų skirtingų kalbos dalių sudarytos homoformos: daiktavardžių, būdvardžių irrieveiksmių; veiksmažodžių, būdvardžių irrieveiksmių bei kitų kalbos dalių homoformos, pavyzdžiui, *aplankiau Vieną – palikau ją vieną – tai atsitiko vieną dieną – pametė vieną pirštinę*. Čia sutampa net 4 kalbos dalys: tikrinis daiktavardis, būdvardis, įvardis ir skaitvardis.

Kaitomų ir nekaitomų kalbos dalių homoformos – tai dviejų ar daugiau skirtingų kalbos dalių, iš kurių viena kaitoma, o kita ar kitos nekaitomos, sutampančios formos. Joms priskirtinos daiktavardžių,rieveiksmių, kartais ir prielinksnių, jungtukų ar dalelyčių; būdvardžių irrieveiksmių; veiksmažodžių,rieveiksmių, prielinksnių arba dalelyčių bei kitų sutampančių kalbos dalių homoformos (5 lentelė). Jos sudaro 7,8% viso morfoliginio daugiareikšmiškumo atvejų

4 lentelė. Skirtingų kaitomų kalbos dalių homoformos

Bendras skaičius	Dktv ir vksm		Dktv ir sktv		Dktv ir bdvr		Vksm ir įvrd		Vksm ir bdvr		Įvrd, dktv ir vksm		Bdvr, įvrd, dktv ir sktv		Dktv, bdvr ir vksm	
	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.
931	364	39,1	5	0,5	190	20,4	86	9,3	186	20	2	0,2	44	4,7	54	5,8

5 lentelė. Kaitomų ir nekaitomų kalbos dalių homoformos

Bendras skaičius	Dktv, bdvr ir prvks		Dktv, prvks, prln, įngt, dll		Bdvr, prvks, prln		Vksm, bdvr, prvks, dll, būdn, ištk		Įvrd, prvks, dll, įngt		Sktv, prln ir prvks		Dktv, vksm, bdvr, prvks, prln		Vksm, dktv ir prvks	
	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.	Sk.	Proc.
464	27	5,8	105	22,6	107	23,1	147	31,7	56	12,1	1	0,2	16	3,4	5	1,1

(žr. 1 lentelę). Šį daugiareikšmiškumo tipą iliustruoja toks pavyzdys: *nāmo stogas – einu namō*, t. y. sutampa daiktavardis ir prieveiksmis. Kaitomų ir nekaitomų kalbos dalių homoformos beveik visada yra daugianarės, t. y. sudarytos iš trijų ar daugiau sutampančių kalbos dalių.

Nekaitomos kalbos dalys sutampa gana dažnai (sudaro 21,6% homoformų) (žr. 1 lentelę). Dažnai nelengva nustatyti, kuriai kalbos daliai priskirtinas daugiareikšmiškas tarnybinis žodelis, pavyzdžiui, *ir* gali būti jungtukas, dalelytė ir prieveiksmis.

Morfologinio daugiareikšmiškumo rūšys

Ankstesniame skyriuje aptarta morfologinio daugiareikšmiškumo klasifikacija, pagrįsta kalbos dalių sutapimais ir struktūrine analize. Analizuojant automatiniu būdu morfologiškai anotuotą tekstyną pastebėta, kad skirtingoms homoformų grupėms priskirtiems daugiareikšmiams atvejams būdingos tam tikros *daugiareikšmiškumo rūšys* – pagal tam tikrą bendrą morfologinį požymį sugrupuotų morfologiškai daugiareikšmių formų visuma, pavyzdžiui, dažnai sutampa (*i*)o linksniuotės daiktavardžių, būdvardžių, įvardžių, skaitvardžių ir dalyvių vienaskaitos kilmininkas, daugiskaitos vardininkas bei šauksmininkas; minėtų kalbos dalių vienaskaitos vardininkas, įnagininkas bei šauksmininkas ir pan. Buvo nustatytos 42 pagrindinės morfologinio daugiareikšmiškumo rūšys. Didžioji dalis šių rūšių iki šiol buvo

neaptikta ir nenagrinėta kalbininkų, nes tik automatinė morfologinė analizė parodė tikrąjį morfologinio daugiareikšmiškumo mastą lietuvių rašomojoje kalboje. Toliau bus smulkiau aprašomos ir analizuojamos pagrindinės morfologinio daugiareikšmiškumo rūšys, pateikiami jų statistiniai duomenys. Jau minėta, kad kaitomos kalbos dalys apima didžiąją morfologinio daugiareikšmiškumo dalį, todėl ir dauguma rūšių priklauso kaitomų kalbos dalių homoformoms (6 lentelė).

6 lentelė. Pagrindinės morfologinio daugiareikšmiškumo rūšys					
Rūšies Nr.	Kaitomų kalbos dalių homoformos	Kaitomų ir nekaitomų kalbos dalių homoformos	Nekaitomų kalbos dalių homoformos	Sk.	Proc.
1	2	3	4	5	6
1			Visiškas visų formų sutapimas	1294	21,62
2	(i)o, é linksn. dktv, bdvr, įvrd, sktv ir dlv vnsk K, dgsk V (dgsk Š) sutapimas			647	10,8
3	(i)o, é linksn. dktv, bdvr, įvrd, sktv ir dlv vnsk V, Įn (Š ir bevrđ.gim) sutapimas			404	6,75
4	Atsitiktinai sutampančios skirtingos leksinės reikšmės dktv, vksm ir bdvr formos			386	6,45
5	Kai kurių dktv linksnių ir asmenuojamųjų bei neasmenuojamųjų vksm formų sutapimas			364	6,08
6	Vksm bndr ir neveik. r būt. l vyr. gim dlv dgsk V sutapimas			310	5,18
7	Visų linksn. bdvr, įvrd, sktv ir dlv vyr. bei mot.gim dgsk K sutapimas			259	4,33
8	Skirtingos tos pačios reikšmės dktv, bdvr, vksm ir įvrd paradigmos			232	3,87
9	(i)a linksn. dktv dgsk V ir Š sutapimas			224	3,74

6 lentelė (tęsinys)					
1	2	3	4	5	6
10	(i)a linksn. bdvr, įvrd, sktv ir dlv vnsk V bei (i)o linksn. tų pačių kalbos dalių dgsk G sutapimas			193	3,23
11	Dktv ir bdvr formų sutapimas			190	3,17
12	Vksm ir bdvr formų sutapimas			186	3,11
13		Vksm, bdvr ir prvks, kartais dll, būdn, ištk sutapimas		147	2,46
14	(i)o ir (i)a linksn. bdvr, įvrd, sktv ir dlv vnsk G sutapimas			121	2,02
15		Bdvr ir prvks ar prln sutapimas		107	1,79
16		Kai kurių dktv formų ir prvks, prln, jngt ar dll sutapimas		105	1,75
17	<i>Substantiva mobilia</i> dktv homoformos			97	1,62
18	Veik. r vyr. gim dlv dgsk V ir bevrđ. gim formų sutapimas			94	1,57
19	Vksm ir įvrd formų sutapimas			86	1,44
20	Tariam. nuos vksm III asm ir neveik. r būt. l vyr. bei mot. gim dlv dgsk K sutapimas			56	0,94
21		Kai kurių įvrd formų ir prvks, dll ar jngt sutapimas		56	0,94
22	Dktv, bdvr ir vksm formų sutapimas			54	0,9
23	(i)a linksn. dktv vnsk Vt ir Š sutapimas			51	0,85
24	Bdvr, įvrd, dktv ir sktv formų sutapimas			44	0,74

6 lentelė (tęsinys)					
1	2	3	4	5	6
25	(i)u ir i linksn. dktv vnsk V bei dgsk G sutapimas			42	0,71
26	é linksn. dktv ir bdvr vnsk Įn bei Š sutapimas			38	0,63
27	Vksm esam. I vnsk II ir III asm sutapimas			35	0,58
28	Įvrd mot. gim vnsk N ir bevrd. gim sutapimas			33	0,55
29		Dktv, bdvr ir prvks formų sutapimas		27	0,45
30	(i)u linksn. dktv vnsk G ir dgsk K sutapimas			26	0,43
31		Dktv, vksm, bdvr, prvks ar prln sutapimas		16	0,27
32	Sutampa visi nekaitomo dktv ar bdvr linksniai			12	0,2
33	Kiekinių sktv mot. gim V ir G sutapimas			11	0,18
34	Dalyvių ir padalyvių sutapimas			7	0,12
35	Sudaiktavardėjusių dlv abiejų giminių vnsk bei dgsk V ir Š sutapimas			7	0,12
36		Vksm, dktv ir prvks sutapimas		5	0,08
37	Kai kurių dktv ir sktv formų sutapimas			5	0,08
38	Kiekinių sktv vyr. ir mot. gim V sutapimas			5	0,08
39	Kiekinių sktv vyr. ir mot. gim G sutapimas			4	0,07
40	Kiekinių sktv vyr. ir mot. gim N ir Įn sutapimas			3	0,05
41	Kai kurių dktv, vksm ir įvrd formų sutapimas			2	0,03
42		Sutampa sktv, prln ir prvks		1	0,02

Toliau šiame straipsnyje kiek plačiau bus pakomentuotos 6 lentelėje išdėstytos pagrindinės morfologinio daugiareikšmiškumo rūšys, pateikiami pavyzdžiai, iliustruojantys tų rūšių vartoseną. Morfologinio daugiareikšmiškumo rūšys aprašomos pagal anksčiau pateiktą kaitomų, kaitomų ir nekaitomų bei vien tik nekaitomų kalbos dalių homoformų skirstymą (žr. homoformų tipai).

Nekaitomų kalbos dalių homoformų rūšis yra pati dažniausia, pavyzdžiui, *kur buvai? – kur tau!* Šiame pavyzdyje visiškai sutamparieveiksmis ir dalelytė. Tokie nekaitomų kalbos dalių sutapimai, kaip galima spręsti iš statistikos, yra labai dažni, bet jų daugiareikšmiškumas sunkiausiai panaikinamas. Beje, tai vienintelė nekaitomoms kalbos dalims priskiriamų homoformų rūšis. Skirti smulkesnes rūšis nėra prasmės, nes sutampa visos formos.

Kaitomų kalbos dalių homoformų rūšys. Kaitomų kalbos dalių sutapimai yra dažniausi automatinio būdu morfologiškai anotuotame tekстыne, todėl daugiausiai rūšių (net 33) sudaro šio tipo homoformas.

Iš kaitomų kalbos dalių homoformoms priskirtų sinkretiškų atvejų dažniausios daiktavardžių, būdvardžių, įvardžių, skaitvardžių ir dalyvių homoformos, pavyzdžiui, *atkurtos žinomos partijos atstovas – susirinko atkurtos žinomos partijos* (2-a daugiareikšmiškumo rūšis). Iš pateikto pavyzdžio matyti, kad susiduriama su (*i*)o, *é* linksniuočių vnsk K, dgsk V (dgsk Š) sinkretizmu. Ši morfologinio daugiareikšmiškumo rūšis – taip pat viena sunkiausiai panaikinamų. Reikėtų paminėti, kad analizuojamos minėtų linksniuočių vnsk K ir dgsk V formos dažnai skiriasi prozodiniais elementais, t. y. kirčio vieta, priegaide ar balsio kiekybe. Tad pagal tradicines lingvistines nuostatas tokios formos turėtų būti laikomos homografais, bet antroje šio straipsnio dalyje jau minėta, kad lingvistai ne visai sutaria dėl homonimijos įvardijimų. Todėl analizuojant daiktavardžių ir kitų kaitomų kalbos dalių homoformas neatsižvelgta į jų tarimo skirtumus. Tai ir nulėmė, kad nagrinėjamos tos homoformos, kurios gali skirtis prozodiniais elementais.

Trečiai rūšiai taip pat priskiriamas linksnių sinkretizmas. Tai daiktavardžių, būdvardžių, įvardžių, skaitvardžių ir dalyvių homoformos, kurių sutampa (*i*)o linksniuotės vnsk V, Įn (Š ir beverd. gim), pavyzdžiui, *sėdi jaunà mamà – atėjo su jáuna mamà – māmà, kur tu!* Kaip ir anksčiau minėtai morfologinio daugiareikšmiškumo rūšiai, pastarajai būdingi fonetiniai skirtumai.

Ketvirta daugiareikšmiškumo rūšis – atsitiktinai sutampančios skirtingos leksinės reikšmės daiktavardžių, veiksmažodžių ir būdvardžių formos, pavyzdžiui, *yra megz-tinis – šiandien yra šalta; važiauvau ilgu keliu – vienai darosi ilgù.*

Toliau pagal dažnumą minėtinas ir asmenuojamųjų, ir neasmenuojamųjų veiksmažodžių formų bei daiktavardžių sutapimas (5-a rūšis), pavyzdžiui, *gera tartis – reikia ilgai tartis; laužo liepsna – vėjas laužo šakas.*

Labai dažnas (pagal dažnumą 6-a rūšis) sinkretizmas tarp veiksmažodžių bendraties ir neveikiamosios rūšies būtojo laiko vyriškosios giminės dalyvių daugiskaitos vardininko, pavyzdžiui, *turiu palikti namus – verkia palikti vaikai.*

Dažnokas visų būdvardžių, įvardžių, skaitvardžių ir dalyvių linksniuočių vyr. ir mot. gim dgsk K sutapimas, pavyzdžiui, *nemėgsta tų dviejų perstatytų aukštų namų – nemyli tų dviejų aukštų pagyvenusių moterų* (7-a rūšis). Šios daugiareikšmės formos taip pat laikomos kaitomų kalbos dalių homoformomis.

Automatiniu būdu morfologiškai anotuojant tekstus pateikiama skirtingų tų pačių reikšmių daiktavardžių, būdvardžių, veiksmažodžių ar įvardžių lemu, kurių kai kurios formos visiškai sutampa, pavyzdžiui, *krito sniegas* (skiriasi pagrindinės formos: *kristi, krinta, krito* ir *kristi, krenta, krito*), *stebukladario* (žodyne pateikiamos dvi lemos: *stebukladaris* ir *stebukladarys*). Dėl tokių skirtingų žodžių paradigmu atsirandančių homoforum pasitaiko gana dažnai (pagal dažnumą tai 8-a rūšis), bet tik nuo tekstą peržiūrinčio lingvisto priklauso, kuri lema bus pasirinkta. Žinoma, vieni žodžiai būna dažniau vartojami, įprastesni, bet nemaža ir tokių formų, kurioms automatiškai nustatytos lemos yra vienodai įprastos ir taip pat dažnai vartojamos.

Tik daiktavardžiams būdingos (*i*)a linksniuotės dgsk V ir Š homoformos, pavyzdžiui, *dirba vyrai – vyrai, palaukit!* Tekstyne toks linksnių sinkretizmas gana dažnas (9-a daugiareikšmiškumo rūšis), bet automatiškai sunkiai panaikinamas, todėl panašius daugiareikšmiškumo atvejus gali pašalinti tik lingvistas peržiūrėdamas sulemuotą ir anotuotą tekstą.

Kaitomų kalbos dalių linksnių sinkretizmui priskiriamos būdvardžių, įvardžių, dalyvių ir skaitvardžių homoformos (10-a rūšis), pavyzdžiui, *atėjo tas laukiamas svečias – pakvietė tas laukiamas viešniams*. Čia sutampa (*i*)a linksniuotės vnsk V ir (*i*)o linksniuotės dgsk G.

Toliau minėtinos daiktavardžių ir būdvardžių homoformos, pavyzdžiui, *užgrobė valdžią – turi valdžią žmoną* (11-a rūšis) ir veiksmažodžių bei būdvardžių sinkretizmas (12-a rūšis), pavyzdžiui, *atėjo su įtariu drauge – įtaria, kad jis meluoja; žinomas žmogus*. Būdvardžių ir neveikiamosios rūšies bei reikiamybės dalyvių sinkretizmas dažnas, bet tai sunkiai skiriama homoforum rūšis. Paskutiniajame pavyzdyje visiškai sutampa visos būdvardžio ir dalyvio *žinomas* formos, todėl neaišku, ar tai būdvardis, ar dalyvis, kaip jie skirtini.

Kita pagal dažnumą kaitomų kalbos dalių homoforum rūšis apima būdvardžių, įvardžių, skaitvardžių ir dalyvių (*i*)o bei (*i*)a linksniuotės vnsk G sinkretizmą, pavyzdžiui, *viršijo antrą nustatytą limitą – peržengė antrą nustatytą ribą* (14-a rūšis).

Substantiva mobilia daiktavardžių homoformos apima tokius pavyzdžius: *Jonas yra geras mokytojas – atleido geras mokytojas; atvyko žinomi rašytojai – žinomi rašytojai trūksta nuoširdumo* (17-a morfologinio daugiareikšmiškumo rūšis). *Substantiva mobilia* daiktavardžių sutampa abiejų giminių dgsk K; vnsk G; vyr. gim dgsk V ir mot. gim vnsk N; vyr. gim vnsk V ir mot. gim dgsk G; vyr. gim vnsk Vt ir Š bei mot. gim vnsk Įn ir Š.

Veiksmažodžiams būdinga, kad gali sutapti veik. r vyr. gim dalyvių dgsk V ir bevrđ. gim formos, pavyzdžiui, *jie buvę parodoje – pasakoję, kad buvę šilta* (18-a rūšis). Tokie sinkretizmo atvejai teoriškai yra gana dažni (žr. 6 lentelę), bet pastebėta, kad dažniausiai vartojamos vyr. gim dalyvių formos.

Devyniolikta kaitomų kalbos dalių homoforum rūšis apima asmenuojamųjų veiksmažodžių formų ir įvardžių sutapimą, pavyzdžiui, *įdomu, ką jis mąno – mąno namas*. Nustatyti, kuri daugiareikšmė forma iš tikrųjų pavartota tekste, padėtų gretimų žodžių analizė. Deja, kol kas tą turi padaryti po automatinės morfologinės analizės tekstą peržiūrintis lingvistas.

Tekstyne randama sutampančių tariam. nuos veiksmazodžių III asm ir neveik. r būt. I dalyvių vyr. bei mot. gim dgsk K formų, pavyzdžiui, *jis viską dėl tavęs padarytų – pasigendama padarytų darbų / užduočių*. Daug dažniau vartojamos tariamosios nuosakos formos negu dalyviai, tačiau norint panaikinti šią daugiareikšmiškumo rūšį reikalinga išsami konteksto analizė. Ši dvidešimta morfologinio daugiareikšmiškumo rūšis taip pat priskiriama kaitomų kalbos dalių homoformoms.

Dvidešimt antrai morfologinio daugiareikšmiškumo rūšiai priskiriamos daiktavardžių, būdvardžių ir veiksmazodžių homoformos, pavyzdžiui, *kalėjime praleido trejus metus – tvarte stovi metus arklys – metus rūkyti būna sunku*.

(i)a linksniuotės daiktavardžių vnsk Vt ir Š sutapimą iliustruoja toks pavyzdys: *mokausi universitetė – mano universitete!* (23-a rūšis). Akivaizdu, kad šauksmininkas būdingas gyvas būtybes reiškiantiems daiktavardžiams, nors kartais personifikacijos atveju ir negyvas būtybes įvardijantys žodžiai gali būti pavartoti šauksmininko linksniu. Tai aišku kalbą analizuojantiems žmonėms, bet automatinę morfologinę analizę atliekančiai kompiuterinei programai svarbus tik formalus raidžių virtinėlės sutapimas, todėl ir neskiriama, kai pavartotas vietininkas, kai šauksmininkas, net jei tas šauksmininkas ir visai nerealus, pavyzdžiui, *eksteritorialume, antropocentrizme*.

Kita morfologinio daugiareikšmiškumo rūšis yra gana dažna (24-a rūšis), bet nepasižyminti formų įvairove. Jis apima sutampančias būdvardžių, įvardžių, skaitvardžių ir kartais daiktavardžių formas: *palikau ją vieną – tai atsitiko vieną dieną – pametė vieną pirštinę – aplankiau Vieną*.

Dvidešimt penktai rūšiai priskiriamos sutampančios (i)u ir i linksniuotės vnsk V bei dgsk G formos, pavyzdžiui, *jaunas sūnūs – atsivedė abu sūnus; svarbi paslaptis – išdavė paslaptis*.

Dvidešimt šeštai morfologinio daugiareikšmiškumo rūšiai priskirtas é linksniuotės daiktavardžių ir būdvardžių vnsk In bei Š sinkretizmas, pavyzdžiui, *atėjo su didele katė – didele kate, nebėk!* Dažnai toks įnagininko ir šauksmininko sutapimas yra galimas tik teoriškai – iš tikrųjų dažniausiai vartojamas įnagininkas, bet morfologiškai anotuotuose tekstuose pateikiami visi, kad ir menkiausiai tikėtini variantai.

Retai randamos sinkretiškos veiksmazodžių esam. I vnsk II ir III asmenų formos, pavyzdžiui, *tu gali tai padaryti – jis gali tai padaryti* (27-a rūšis). Kaip ir daugelį kitų prieš tai minėtų kaitomų kalbos dalių rūšių, pastarąją šnekamojoje kalboje galima atskirti iš skirtingo tarimo; remiantis kontekstu rašytinėje kalboje taip pat nelieka neaiškumų, jei analizuoja žmogus. Deja, „Lemuoklis“ apibūdina tik užrašytą formą (ši programa kol kas nesugeba atpažinti kirčio ženklų) neatsižvelgdamas į gretimus žodžius.

Tik įvardžių homoformoms būdingas mot. gim vnsk N ir beverd. gim sutapimas (28-a rūšis), pavyzdžiui, *įteik gėles tai moteriai – tai atsitiko seniai*. Ši rūšis nėra dažna, be to, nepasižymi didele formų įvairove. Dažniausiai morfologiškai daugiareikšme laikoma forma *tai*.

Viena retesnių daiktavardžių homoformų rūšių (30-a) – (i)u linksniuotės vnsk G ir dgsk K sutapimas, pavyzdžiui, *pamačiau inspektorių – trūksta inspektorių*.

Trisdešimt antra morfologinio daugiareikšmiškumo rūšis apima tiek kaitomas, tiek nekaitomas kalbos dalis. Nekaitomų daiktavardžių ir būdvardžių sutampa visi linksniai, pavyzdžiui, *stovi bordo taksi – laukiu bordo taksi – matau bordo taksi – sėdžiu bordo taksi*. Nors daiktavardis ir būdvardis laikomi kaitomomis kalbos dalimis, tačiau kai kurie daiktavardžiai ir būdvardžiai turi tik vieną formą, kuri atitinka visus linksnius. Lemuodamas nekaitomą daiktavardį „Lemuoklis“ jam automatiškai generuoja 14 hipotetinių morfologinių reikšmių: 7 vienaskaitos ir 7 daugiskaitos linksnius, o nekaitomiems būdvardžiams – net 28 galimas formas, t. y. 7 vienaskaitos, tiek pat daugiskaitos linksnių ir visa tai dukart – atskirai vyriškajai ir moteriškajai giminei. Todėl automatiškai anotuojant tekstus, „pagal nutylėjimą“ paliekama tik viena lema – nenurodomos skaičiaus, giminės ir linksnio kategorijos (Zinkevičius, 2000: 263–264).

Kai kurie lingvistai mano, kad nekaitomų daiktavardžių giminę, skaičių ir linksnį padeda nustatyti greta pavartoti žodžiai (plg. Paulauskienė, 1994: 155–156; 2001: 5–6; Judžentis, 2002: 44); kitų teigimu, nekaitomi daiktavardžiai yra vienos kurios nors giminės (Balkevičius, 1976: 24). Dėl dabartinio „Lemuoklio“ nesugebėjimo analizuoti kontekstą neatsižvelgiama į gretimus žodžius, todėl negalima nustatyti nekaitomų žodžių morfologinių kategorijų. Dėl šios priežasties nekaitomiems daiktavardžiams ir būdvardžiams nėra pateikiamos giminės, linksnio ir skaičiaus kategorijos.

Iš kiekinių skaitvardžių dažniausiai sutampa (33-ia rūšis) mot. gim V ir G, pavyzdžiui, *skrenda dvi gulbės – pamačiau dvi gulbes*.

Pačios rečiausios sutampančios veiksmažodžių formos – dalyviai ir padalyviai (34-a rūšis), pavyzdžiui, *visus lankantis žmogus – mirė lankantis kunigui*.

Sudaiktavardėjusių dalyvių sutampa vyr. ir mot. gim vnsk bei dgsk V ir Š (35-a rūšis), pavyzdžiui, *eina teisiamasis – teisiamasis, prašom atsistoti!*

Anotuotame tekstyne rasti vos keli sutampančių daiktavardžių ir skaitvardžių formų pavyzdžiai (37-a rūšis), pavyzdžiui, *neliko jokių dvejų – neišgyveno nė dvejų metų*. Dažniausiai su skaitvardžiais sutampa koks nors retas, nebevertojamas, bet „Dabartinės lietuvių kalbos žodyne“ dar nurodytas daiktavardis, pvz., *dveja, pirmoji*.

Tokio paties dažnumo (0,08%) kiekinių skaitvardžių vyr. ir mot. gim vardininko homoformos, pavyzdžiui, *trys vyrai – trys moterys* (38-a rūšis). Šiek tiek retesnis kiekinių sktv vyr. ir mot. gim G sinkretizmas, pavyzdžiui, *pakviečiau tris vyrus ir tris moteris* (39-a rūšis).

Kiekinių skaitvardžių vyr. ir mot. gim N ir Įn formų sutapimas (40-a rūšis) – vienas rečiausių, pavyzdžiui, *pasiskundė dvim draugams / draugėms – atėjo su dvim draugais / draugėmis*.

Patį rečiausia kaitomų kalbos dalių morfologinio daugiareikšmiškumo rūšis (41-a) – kai kurių daiktavardžių, veiksmažodžių ir įvardžių homoformos, pavyzdžiui, *keli sunkų maišą – atėjo keli žmonės*. Dar reikia paminėti daiktavardžio *kelis* vienaskaitos šauksmininką. Akivaizdu, kad pastaroji forma įmanoma tik teoriškai.

Kaitomų ir nekaitomų kalbos dalių homoformų rūšys. Išanalizavus automatiniu būdu morfologiškai anotuotame tekстыne rastas daugiareikšmes formas, pastebėta, kad kaitomos ir nekaitomos kalbos dalys sutampa daug rečiau nei kaitomos

kalbos dalys, todėl rasta mažiau kaitomų bei nekaitomų kalbos dalių homoformų rūšių: iš 42 nustatytų morfologinio daugiareikšmiškumo rūšių minėtoms homoformoms priskirtos tik 8. Iš jų dažniausias veiksmažodžių, būdvardžių,rieveksmių, kartais dalelyčių, jungtukų, būdinių ar ištiktukų sutapimas (13-a rūšis), pavyzdžiui, *sunkiau bulves – gyventi darosi sunkiau*. Aukštesniojo laipsniorieveksmiai savo forma sutampa su aukštesniojo laipsnio bevardės giminės būdvardžiais. Nežinia, kokiais kriterijais remiantis skirtinos šios formos. DLKG rašoma, kad minėtos formos skiriasi sintaksinėmis funkcijomis (DLKG, 1996: 415), tačiau kol kas atliekama automatinė morfologinė analizė, todėl neįmanoma automatiškai skirti sutampančių būdvardžių irrieveksmių formų. Tai nelengva padaryti ir po automatinės analizės tekstą peržiūrinčiam lingvistui.

Kaip minėta, asmenuojamosios ir neasmenuojamosios veiksmažodžių formos dar gali sutapti su dalelytėmis, pavyzdžiui, *jie esą gerai žmonėmis – esą niekas čia nebuvo*; prielinksniais, pavyzdžiui, *ji ties rankšluostį ant žolės – susitikome ties tavo namais*.

Sutampančias būdvardžių irrieveksmių formas iliustruoja šis pavyzdys: *nesiruošė staigiai kovai – staigiai puolė* (15-a rūšis). Nors tokia homoformų rūšis gana dažna, ateityje, kai automatiškai bus analizuojamas kontekstas, bus galima nustatyti, kuri sinkretiška kalbos dalis pavartota: būdvardis arrieveksmis. Beveik visada po būdvardžio vartojamas daiktavardis, o porieveksmio dažniau randamas veiksmažodis. Dar būdvardis gali sutapti ir su prielinksniu, pavyzdžiui, *netoli (netolima) distancija – netoli namų stovi senas beržas*.

Dažniausiai daiktavardžiai sutampa su kaitomomis kalbos dalimis, bet pasitaiko ir daiktavardžių beirieveksmių ar prielinksnių homoformų (16 rūšis), pavyzdžiui, *nāmo stogas – einu namō; atvyko su mano priešais – stovi priešais mane – priešais matau mišką*.

Kai kurios įvardžių formos sutampa surieveksmiais, dalelytėmis ar jungtukais, pavyzdžiui, *stebėjosi juo – kai nieko nedėdi ant danties, juo skauda – juo labiau – juo daugiau, juo geriau* (21-a rūšis). Šios homoformos nepasižymi įvairumu.

Automatiniu būdu morfologiškai anotuotame tekstynėlyje rasta sinkretiškų kai kurių daiktavardžių, būdvardžių irrieveksmių formų (29-a rūšis), pavyzdžiui, *skirtingi žinių lygiai – paruošti lygiai vietai reikėjo daug pastangų – ilgai plušo, kol lygiai nuobliavo lentą*.

Dažnai sutampa daugiau negu dvi tiek kaitomos, tiek nekaitomos kalbos dalys, pavyzdžiui, *atėjau kartū su draugais – negaliu valgyti, nes kartū – džiaugiausi tuo kartū – visi baisėjosi kardu, bet nepakartu žmogumi*. Šiame pavyzdyje sutamparieveksmis, būdvardis, daiktavardis ir dalyvis (31-a rūšis).

Trisdešimt šešta kaitomų bei nekaitomų kalbos dalių rūšis – tai sutampančios daiktavardžių, veiksmažodžių irrieveksmių formos, pavyzdžiui, *anais kartais tu buvai linksmesnė – vaikas stebėjosi kartais, bet nepakartais vyrais – kartais būna liūdna*.

Patį rečiausia šios rūšies homoformoms priskirta rūšis (42) – tai skaitvardžių, prielinksnių irrieveksmių sutapimas, pavyzdžiui, *pirma, reikia pabaigti visus darbus – pirmà buvo geriau gyventi – jis eina pirmà manęs*.

Ir kaitomų, ir kaitomų bei nekaitomų, ir vien tik nekaitomų kalbos dalių homoformas dar būtų galima skirstyti pagal tikėtinumą. Jau buvo užsiminta, kad ne visos formos yra vienodai tikėtinos: vienu formų sutapimai yra tikrai dažni, kiti galimi tik teoriškai, bet tai turėtų būti kitas analizės lygmuo. Daugeliu atvejų automatinę morfologinę analizę atliekanti programa „Lemuoklis“ parodė, kad žmogus yra labai subjektyvus ir nepastebi dažnų homoformų. Visus, net ir menkiausiai tikėtinus, morfologinio daugiareikšmiškumo atvejus galima pastebėti tik naudojant tam tikrą programinę įrangą. Dėl to šios morfologinio daugiareikšmiškumo analizės objektas dar yra gana naujas ir neištirtas.

IŠVADOS

Straipsnyje supažindinta su dažnu, bet be specialios programinės įrangos retai pastebimu morfologiniu daugiareikšmiškumu, kuriam iki šiol lietuvių lingvistikoje nebuvo skiriama daug dėmesio. Šis kalbos reiškinys imtas analizuoti tik po to, kai kompiuterinė programa „Lemuoklis“ automatiškai sulemavo ir morfologiškai pažymėjo tekstus.

1. Morfologinis daugiareikšmiškumas – nauja tyrinėjimo sritis, todėl teko apsibrėžti pagrindinius terminus; pasirinktas terminas *daugiareikšmiškumas*. *Morfologinis daugiareikšmiškumas* – tai reiškinys, kuriam priskiriamos: a) kaitomos ir nekaitomos; b) skirtingų ir tų pačių kalbos dalių; c) tam tikromis formomis, prozodiniais elementais besiskiriančios ir visiškai sutampančios žodžių formos ar žodžiai.

2. Automatiniu būdu morfologiškai anotuotame tekstynėlyje apie 40% formų yra daugiareikšmės. Pateikta klasifikacija pagal kalbos dalių sutapimus: skirtos kaitomų, kaitomų ir nekaitomų bei tik nekaitomų kalbos dalių homoformos. Pačios dažniausios yra kaitomų kalbos dalių homoformos.

3. Smulkiau išnagrinėtos ir aprašytos 42 pagrindinės morfologinio daugiareikšmiškumo rūšys. Didžioji dalis (33) rūšių priskirta kaitomų kalbos dalių homoformoms, 8 rūšys apima kaitomų ir nekaitomų kalbos dalių sutapimus. Pati dažniausia, bet mažiausia įvairovė (pirma rūšis) pasižymi nekaitomų kalbos dalių homoformų rūšis.

Ši morfologinio daugiareikšmiškumo tipologija turbūt nėra galutinė. Reikia didesnio automatiniu būdu morfologiškai anotuoto ir lingvisto peržiūrėto teksto, kad būtų atspindėta didesnė įvairovė. Reikėtų patyrinėti kitus žanrus ir pažiūrėti, ar ten išlaikoma publicistikos tekstuose nustatyta morfologinio daugiareikšmiškumo proporcija, ar galima atrasti iki šiol dar nenustatytų rūšių.

Santrumpos

bdvr – būdvardis

beverd. gim – bevardė giminė

būdn – būdinys

būt. kart. I – būtasis kartinis laikas

dgsk – daugiskaita

mot. gim – moteriškoji giminė

N – naudininkas

nejvardž – neįvardžiutinė forma

nelygin. I – nelyginamasis laipsnis

nesngr – nesangražinė forma

dktv – daiktavardis	neveik. r – neveikiamoji rūšis
dll – dalelytė	padlv – padalyvis
dlv – dalyvis	prln – prielinksnis
esam. I – esamasis laikas	proc. – procentai
G – galininkas	prvks – prieveiksmis
I asm – pirmasis asmuo	sk. – skaičius
II asm – antrasis asmuo	sktv – skaitvardis
III asm – trečiasis asmuo	sng – sangražinė forma
Įn – įnagininkas	Š – šauksmininkas
ištk – ištiktukas	tariam. nuos – tariamoji nuosaka
įvardž – įvardžiuotinė forma	tiesiog. nuos – tiesioginė nuosaka
įvrd – įvardis	V – vardininkas
jngt – jungtukas	veik. r – veikiamoji rūšis
K – kilmininkas	vyr. gim – vyriškoji giminė
KTŽ – Kalbotyros terminų žodynas	vksm – veiksmazodis
linksn. – linksniuotė	vnsk – vienaskaita
LK I – Lietuvių kalba, I tomas	Vt – vietininkas
LKE – Lietuvių kalbos enciklopedija	

Literatūra

- Balkevičius*, 1976 Balkevičius J. Nelinksniuojamųjų daiktavardžių giminė. *Mūsų kalba*. 1976. Nr. 4. P. 22–26.
- Charniak*, 1993 Charniak E. *Statistical Language Learning*. Cambridge, Massachusetts, London: A Bradford Book The MIT Press, 1993.
- DLKG*, 1996 *Dabartinės lietuvių kalbos gramatika*. Vilnius: Mokslo ir enciklopedijų leidykla, 1996.
- Drotvinas*, 1986 Drotvinas V. *Pagrindinės leksikologijos sąvokos*. Vilnius: Lietuvos TSR aukštojo ir specialiojo vidurinio mokslo ministerija, 1986.
- Jakaitienė*, 1980 Jakaitienė E. *Lietuvių kalbos leksikologija*. Vilnius: Mokslo, 1980.
- Judžentis*, 2002 Judžentis A. Naujas žvilgsnis į lietuvių kalbos daiktavardžių giminės kategoriją. *Acta Linguistica Lithuanica*. 2002. T. 46. P. 39–47.
- KTŽ*, 1990 Gaivenis K., Keinys S. *Kalbotyros terminų žodynas*. Kaunas: Šviesa, 1990.
- LK*, 1995 *Lietuvių kalba*. Vilnius: Mokslo ir enciklopedijų leidykla, 1995. T. 1.
- LKE*, 1999 *Lietuvių kalbos enciklopedija*. Vilnius: Mokslo ir enciklopedijų leidybos institutas, 1999.
- Marcinkevičienė*, 2000 Marcinkevičienė R. Tekstynų lingvistika (teorija ir praktika). *Darbai ir Dienos*. 2002. T. 24. P. 7–64.
- Paulauskienė*, 1994 Paulauskienė A. *Lietuvių kalbos morfologija*. Vilnius: Mokslo ir enciklopedijų leidykla, 1994.
- Paulauskienė*, 2001 Paulauskienė A. Apie „nelinksniuojamųjų“ daiktavardžių giminę. *Gimtoji kalba*. 2001. Nr. 5. P. 5–6.
- Urbutis*, 1956 Urbutis V. *Lietuvių kalbos leksikos homonimų susidarymo būdai* (Filologijos mokslų kandidato laipsnio disertacijos autoreferatas). Vilnius, 1956.
- Zinkevičius*, 2000 Zinkevičius V. *Lemuoklis* – morfologinei analizei. *Darbai ir Dienos*. 2000. T. 24. P. 246–273.

Erika Rimkutė

THE TYPOLOGY OF MORPHOLOGICAL AMBIGUITY

S u m m a r y

The article deals with morphological ambiguity, a type of linguistic ambiguity which was analysed in the automatically morphologically annotated corpus of the Lithuanian language. The corpus with morphological tags has shown a large ambiguity of the language: about 40 percent of word forms are ambiguous.

In Lithuanian linguistics, traditionally the greatest attention has been devoted to lexical homonyms. Other levels of ambiguity have not been analysed. For this reason it was necessary to define the main terms and classification of morphological ambiguity. Forty-two main subtypes of morphological ambiguity were defined and analysed. These subtypes were distributed according to their part-of-speech characteristics. The homoforms were most frequent in non-inflected parts of speech and in the case of syncretism of nouns, adjectives, pronouns, participles, and numerals.