

MORFOLOGIŠKAI DAUGIAREIKŠMIŲ FORMŲ ATsirADIMO PRIEŽASTYS

ESMINIAI ŽODŽIAI: morfologinis daugiareikšmiškumas, automatinė morfologinė analizė, tekstynas, homoformas, žodynas.

ANOTACIJA

Straipsnyje, remiantis duomenimis, gautais iš *Lietuvių kalbos homoformų žodyno*, aptiriamos lietuvių kalbos morfologiškai daugiareikšmių formų, straipsnyje vadinamų homoformomis, atsiradimo priežastys. Homoformas pagal atsiradimo priežastis galima suklasifikuoti į tokias grupes: labiau teorinės homoformas, atsiradusios dėl kai kurių žodynų sudarymo specifikos; homoformas, atsiradusios dėl prozodinių elementų skirtumų; homoformas, atsiradusios dėl sutampančių skirtingų žodžių; homoformas, nulemtos lietuvių kalbos morfologijos sistemos; tos pačios leksinės reikšmės homoformas, atsiradusios dėl priskyrimo skirtingoms paradigmoms.

Šios homoformų atsiradimo priežastys padeda suvokti, kodėl anotuojant morfologiniais anotatoriais atsiranda vienokių ar kitokių morfologinio daugiareikšmiškumo atvejų, kodėl negalima gauti visiškai tikslių automatinio anotavimo rezultatų.

ABSTRACT

In the article, the reasons for the homoforms of the Lithuanian language are described on the basis of data obtained from the *Dictionary of Homoforms of the Lithuanian Language*.

The author offers a classification of homoforms according to the reasons of their appearance, which is as follows: homoforms that are more theoretical in nature and appeared due to the specifics of creating dictionaries; homoforms resulting from differences in prosodic elements; homoforms resulting from different overlapping words; homoforms that arose due to the specifics of the Lithuanian morphology system; homoforms of the same meaning that appeared under different paradigms.

These reasons help us to understand why annotation by morphological annotators often result in different cases of morphological ambiguity.

IVADAS

Šiais laikais tyrinėjant kalbą, dažniausiai pasitelkiamos įvairios kompiuterinės programos. Jos suteikia galimybę apdoroti kalbą, spręsti lingvistines problemas, segmentuoti tekstą, lemuoti, klasifikuoti, klasterizuoti ar taisyti teksto klaidas, automatiškai versti. Kompiuterinėmis programomis galima atlikti ir semantinę, morfologinę, sintaksinę analizę, kurti anotuotus ar neanotuotus tekstynus, garsynus ir t. t.

Lietuvių kalba irgi kompiuterizuojama. Vienas iš lietuvių kalbos kompiuterizavimo etapų – automatinė morfologinė analizė. Jos metu išryškėja morfologinis daugiareikšmiškumas. Morfologinis daugiareikšmiškumas – tik viena nedidelė daugiareikšmiškumo dalis, kuriai priskiriama: 1) kaitomosios ir nekaitomosios žodžių formos ar žodžiai; 2) skirtingų ir tų pačių kalbos dalių žodžių formos ar žodžiai; 3) tam tikromis formomis, prozodiniais elementais besiskiriantys ir visiškai sutampantys žodžiai ar jų formos (Rimkutė 2003a: 63). Morfologiškai daugiareikšmiais žodžiai ar jų formos laikomi tada, kai tai pačiai žodžio formai pateikiamos dvi ar daugiau lemų (pvz., formai *kovų* galima pateikti lemas *kovas* ir *kova*) arba dvi ar daugiau gramatinių pažymų (pvz., *naktis* – vnsk. vardininkas arba dgsk. galininkas).

Per automatinę morfologinę analizę išryškėjęs morfologinis daugiareikšmiškumas, anot Erikos Rimkutės (2002: 86), yra labai specifinis atvejis, tačiau žmonės su daugiareikšmiais žodžiais susiduria kasdien. Dažniausiai tai nesudaro bendravimo problemų, nes jie, suvokdami kontekstą, geba gana lengvai numanyti, kuris žodis ar jo forma pavartota, o kiti variantai net neateina į galvą. Be to, dažnai morfologiškai daugiareikšmius žodžius sakytinėje kalboje atpažinti padeda skirtinga intonacija ar kirčio vieta.

Nors morfologinis daugiareikšmiškumas yra iš esmės išspręsta problema (žr. Rimkutė, Daudaravičius 2007), vis dėlto, net ir patobulinus lietuvių kalbos morfologinius anotatorius, ne visada iš kelių morfologiškai daugiareikšmių formų pavyksta pasirinkti tikėtiniausią, be to, nereti ir tokie atvejai, kai morfologinis anotatorius neteisingai parenka vieną formą iš kelių galimų.

Lietuvių kalbą analizuoti galima trimis gana išsamiai aprašytais ir viešai prieinamais morfologiniais anotatoriais – Sintaksinės ir semantinės analizės sistemoje prieinamu anotatoriumi *Semantika.lt*¹ ir anotatoriais *Lemuoklis*

¹ Žr. <https://semantika.lt/Analysis/TextAnalysis>.

bei *Morfuoklis*². Šiame straipsnyje pateikti duomenys paremti *Lemuoklio* atlikta analize, todėl šis anotatorius bus pristatytas kiek išsamiau.

Šio straipsnio tikslas – aprašyti lietuvių kalbos homoformų (šis ir kiti terminai bus paaiškinti 2 skyriuje) atsiradimo priežastis. Uždaviniai – apžvelgti su morfologiniu daugiareikšmiškumu susijusius terminus ir, remiantis išanalizuotais morfologinio daugiareikšmiškumo atvejais, juos aprašius ir suklasifikavus, aptarti homoformų atsiradimo priežastis. Analizuoti *Lietuvių kalbos homoformų žodyne* (LKHŽ)³ esantys žodžiai. Tyrimui taikyti kiekybinės analizės, tekstynų lingvistikos, aprašomasis metodai.

LKHŽ pateikti nesikartojantys 1 mln. žodžių homoformų pavyzdžiai, rasti morfologiškai anotuotame tekстыne. Iš žodyno galima pamatyti, kurios lietuvių kalbos žodžių formos yra morfologiškai daugiareikšmės. Skaitytojui patys gali padaryti išvadas, kurios formos yra tikėtinesnės ir kas tai lemia. Žodyne yra apie 36 tūkst. homoformų, jos suklasifikuotos į tris kategorijas: 1) kaitomųjų kalbos dalių homoformos; 2) nekaitomųjų kalbos dalių homoformos; 3) kaitomųjų ir nekaitomųjų kalbos dalių homoformos. Kaitomųjų kalbos dalių homoformos, taip pat kaitomųjų ir nekaitomųjų kalbos dalių homoformos smulkiau suskirstytos į 53 rūšis (jos bus minimos toliau).

E. Rimkutė, rengusi morfologiškai anotuotą lietuvių kalbos tekstyną, išsiaiškino, kad morfologinis daugiareikšmiškumas apima tiek kaitomąsias, tiek nekaitomąsias kalbos dalis (Rimkutė 2006a: 136). Anot jos, dažniausiai sutampa kaitomosios kalbos dalys, rečiausiai – kaitomosios su nekaitomosiomis. Nors žodyne nepateikiama statistinių duomenų⁴, bet pagal tai, kurioms homoformų rūšims priskiriama daugiausia žodžių formų, kaip dažnas galima paminėti šias homoformų rūšis: *(i)a* linksniuotės vardažodžių vnsk. vardininko ir *(i)o* linksniuotės vardažodžių dgsk. galininko sutapimai (pvz.: *nemažas, pripažintas, trečias*); *(i)o* ir *(i)a* linksniuotės vardažodžių vnsk. galininko sutapimai (pvz.: *abejingą, aštuntą, parduotą*); *(i)o, é* linksniuotės vardažodžių vnsk. kilmininko ir dgsk. vardininko bei šauksmininko sutapimai (pvz.: *abatijos, abejingos, abstrakčiausios*) (išsamiau apie žodyno sandarą ir rengimą žr. šio straipsnio autorės darbą: Piečytė 2023).

² Žr. <https://sitti.vdu.lt/morfuoklis/>.

³ Šis žodynas rengiamas leidybai. Kol kas viešai neskelbtas. Autorės – M. Žemrietė ir E. Rimkutė.

⁴ Dažnumui suskaičiuoti reikėtų atskiro tyrimo. Bet, paanalizavus dalį homoformų tekstynuose, matyti, kad vienos iš jų gali būti pavartotos tūkstančius kartų, kitos – tik kelis kartus.

LKHŽ ir kartu šio straipsnio šaltinis – 2000–2005 m. rengtas morfologiškai anotuotas 1 mln. žodžių lietuvių kalbos tekstynas. Jis anotuotas naudojant Vytauto Zinkevičiaus sukurtą programą *Lemuoklis*. Morfologiškai anotuotame tekстыne beveik pusė žodžių yra morfologiškai daugiareikšmiai. Šis reiškinys išryškėjo tik tada, kai mokslininkai pradėjo tvarkyti automatinės morfologinės analizės programos anotuotus failus, kai pamatė, kaip kalba traktuojama neturint kontekstinių žinių, t. y. kai kiekviena žodžio forma analizuojama kaip visiškai atskira (Rimkutė 2006a: 136). Taigi, šis straipsnis ir jame esantys paaiškinimai turėtų būti pravartūs tiems, kurie naudoja lietuvių kalbos morfologinius anotatorius. Jis turėtų padėti suprasti galimų netikslumų priežastis.

LKHŽ – pirmasis tokio tipo lietuvių kalbos žodynas. Jame matyti ne-tradicinis požiūris į kalbą, pateiktas per kompiuterinių lietuvių kalbos analizės programų prizmę. Iš šio žodyno matyti, kokią kalbą „supranta“ kompiuterinės programos. Panašų vaizdą apie lietuvių kalbą gali susidaryti ir lietuvių kalbos besimokantys užsieniečiai. Kadangi žodyną sudaro ištis didelis homoformų skaičius, jis mokslininkams atveria daugiau tyrinėjimo galimybių.

Nors, kaip minėta, Lietuvoje toks žodynas – pirmasis, tačiau panašių žodynų jau publikuota kitomis kalbomis. Pavyzdžiui, 2001 m. sukurtas rusų kalbos homoformų žodynas, autorė – Zhanna Grigoryevna Anoshkina (*Словарь омонимичных словоформ*)⁵. Šiame žodyne pateikiami ir homonimai, ir homografi (žodžiai, kurie rašomi taip pat, bet skirtingai tariami). Žodyne į kirtį neatsižvelgiama, nevartojama raidė *ě*. Žodynas sutvarkytas taip, kad žodžių formos būtų kairėje, o lekšemos – dešinėje. Kalbos dalys pateikiamos skliaustuose, plg.⁶:

ехидна	
ехидной	ехидна (с), ехидный (п)
ехидною	
ехидны	
ехидно	ехидно (н), ехидный (п)

⁵ Rusų kalbos homoformų žodynas yra prieinamas internete: <http://cfrl.ruslang.ru/homo-forms/index.htm>.

⁶ Šiuo pavyzdžiu parodomas žodžio *ехидна* linksniavimas ir kalbos dalys (jų santrumpas pateiktos skliausteliuose). Rusų kalboje šis žodis turi ne vieną reikšmę, kuri priklauso nuo konteksto. Žodis *ехидна* lietuviškai gali reikšti ežį, Australijos gyvatę ar gali būti vartojamas norint apibūdinti piktą, žiaurų, klastingą žmogų.

2008 m. buvo sukurtas anglų kalbos homofonų žodynas⁷, jo autorius – Alfredas Aloisi (plačiau apie homofonų ir homoformų santykį žr. šio straipsnio 2 sk.). Žodyno apraše teigiama, kad žodynas nuolat pildomas, tačiau trūksta informacijos, iš kur renkami žodžiai, kiek jų yra ir pan. Susisiekus su žodyno autoriumi, paaiškėjo, kad pirmoji žodyno versija buvo popierinė, tačiau greitai suprasta, kad praktiškiau sukurti internetinę, nes popierinė galėtų greitai pasenti ir prireiktų ne vieno leidimo. Žodyne yra daugiau nei 10 tūkst. homofonų, juos gali siūlyti ir žodyno vartotojai. Be to, čia pateikiamos ne tik homofonų lemos, jų kalbos dalys, bet ir apibrėžimai – tai nėra labai įprasta tokio tipo žodynams. Pateikiami skirtingai rašomi, bet vienodai tariami žodžiai, pavyzdžiui, *moo* (liet. *mū*) ir *moue* (liet. *grimasa*), plg.:

moo

1. :: *verb-intransitive*

To emit the deep, bellowing sound made by a cow; low.

2. :: *noun*

The lowing of a cow or a similar sound.

moue

1. :: *noun*

A small grimace; a pout.

Anglų kalbos homofonai pateikti ir popieriniuose žodynuose, pavyzdžiui, 1997 m. publikuotas Leslie Pressono homofonų žodynas (*Dictionary of Homophones*). Jis skirtas besimokantiems anglų kalbos, turi daugiau nei 600 homofonų porų. 2012 m. sudarytas A. D. Reed žodynas *Reed's Homophones: A Comprehensive Book of Sound-alike Words* (žodyne pateikiami amerikiečių anglų kalbos homofonai, jų porų yra daugiau nei 1 000. Prie kiekvieno žodžio pateikiamas apibrėžimas, kalbos dalis ir vartojimo pavyzdys).

Lietuvoje taip pat parengta knygų, vadovėlių, mokslinių straipsnių panašiomis temomis. Jonas Šukys 2004 m. išleido knygą *Panašiai skambantys žodžiai: normos ir klaidos*, kurioje išsamiai analizuojami paronimai (terminas bus paaiškintas 2 sk.). Knygoje mokslininkas paaiškina sugretintų žodžių normines reikšmes, nurodo ir pataisytą klaidingą vartoseną, be to, aptaria paronimų ir homografų vartosenos aspektus.

Dalis homoformų atsiranda dėl atsitiktinai sutampančių skirtingų reikšmių žodžių. Tokių žodžių analizei skirtas Rūtos Marcinkevičienės (2011) vadovėlis *Žodžio reikšmė. Žodynai ir tekstynai*, kurio viename skyriuje rašoma apie polisemiją, čia išsamiai aprašyti homonimai, pateikta jų klasifikacija.

⁷ Anglų kalbos homofonų žodynas prieinamas internete: <https://www.homophone.com>.

Ypač daug darbų apie morfologinį daugiareikšmiškumą publikavo E. Rimkutė. Galima paminėti jos disertaciją (Rimkutė 2006a), keletą mokslinių straipsnių (Rimkutė 2002, 2003a, b), kur plačiai aprašytos įvairios daugiareikšmiškumo rūšys, ypač daug dėmesio skirta morfologiniam daugiareikšmiškumui bei homoformoms.

Šis straipsnis susideda iš įvado (pirmoji dalis), dviejų teorinių dalių, dviejų tiriamųjų dalių ir apibendrinimo. Darbo gale pateikiamas visų cituotų literatūros šaltinių sąrašas.

1. MORFOLOGIŠKAI ANOTUOTAS TEKSTYNAS IR JAME IŠRYŠKĖJĘS MORFOLOGINIS DAUGIAREIKŠMIŠKUMAS

Pirmasis morfologiškai anototas lietuvių kalbos tekstynas – Vidos Žilinskienės ir Laimos Grumadienės sukurta elektroninė duomenų bazė. Ji naudota specialiai dažniniams žodynams rengti (Rimkutė 2006b: 34). V. Žilinskienės ir L. Grumadienės *Dabartinės rašomosios lietuvių kalbos dažninio žodyno elektroninis tekstynas* (2002) sudarytas iš 1,2 mln. žodžių, jame esantys tekstai – originalūs (ne verstiniai), priklauso publicistiniam, moksliniam, beletristiniam arba kanceliariniam stiliui. Šie tekstai buvo anotuoti, o vėliau pagal juos sukurti žodynai – *Dažninis dabartinės rašomosios lietuvių kalbos žodynas* (1997) ir *Dažninis dabartinės rašomosios lietuvių kalbos žodynas* (1998). Vytauto Didžiojo universiteto Kompiuterinės lingvistikos centre 2000 m. pradėtas kurti įvade minėtas morfologiškai anototas tekstynas⁸. Jis sudarytas pusiau automatiškai, pritaikius automatinės morfologinės analizės programą *Lemuoklis*. Pirmąją jo versiją sudarė apie 1 mln. žodžių, tačiau vėliau tekstynas buvo papildytas, pataisyti netikslumai⁹. Tekstynė publicistikos tekstai sudaro 36 proc., mokslinės literatūros tekstai – 24 proc., grožinės literatūros tekstai – 19 proc., administracinio stiliaus tekstai – 2,8 proc., Lietuvos Respublikos Seimo stenogramos – 6,8 proc.

Tekstynas sudarytas pusiau automatiškai. Tuo metu, kai buvo sudaromas šis tekstynas, V. Zinkevičiaus parengta morfologinės analizės programa

⁸ Šis tekstynas viešai prieinamas internete: <https://clarin.vdu.lt/xmlui/handle/20.500.11821/33>.

⁹ Lietuvių kalbos homoformų žodynui duomenys imti iš šios pirmosios tekstyno versijos, nes norėta užfiksuoti kuo daugiau lietuvių kalbos morfologinio daugiareikšmiškumo atvejų. Vėliau *Lemuoklyje* įdiegti statistiniai vienareikšminimo moduliai (Rimkutė, Daudaravičius 2007), dėl to ši programa pradėjo veikti tiksliau, pateikdavo daug mažiau morfologiškai vienareikšmių formų.

Lemuoklis, pateikianti lemas ir morfologines pažymas, neatsižvelgdavo į kontekstą, šioje programoje nebuvo įdiegta informacijos apie semantiką, žodžių formos dažniausiai būdavo nustatomos turint lietuviškų žodžių šaknų sąrašą, o ne naudojantis kokiais nors žodžių formų sąrašais su nurodytomis tų formų gramatinėmis pažymomis (Zinkevičius 2000: 30). Tai reiškia, kad programa žodžių formoms pateikdavo visas galimas gramatines pažymas ir visas galimas lemas (antraštines formas) (žr. įvade pateiktus pavyzdžius). Dar paminėtina, kad *Lemuoklis* analizavo rašytines, be kirčio ženklų formas. Būtent dėl to homoformos galėjo skirtis ir prozodiniais elementais (Rimkutė, Grybinaitė 2004: 75–76).

Šiuo metu kalbos technologijų tyrimuose dažniau naudojamas anotatorius *Semantika.lt*, nes jis tikslesnis; gali būti ir tvarkomas, ir atnaujinamas, ir papildomas (Bielinskienė ir kt. 2017: 2). Tačiau jis irgi ne visuomet anotuoja taisyklingai – neretai neatpažįsta svetimybių, žargono, trumpinių. Nematant visos kalbos ar konteksto taip, kaip mato žmogus, daroma morfologiškai daugiareikšmių žodžių anotavimo klaidų (jei vienas žodis turi kelias skirtingas gramatines reikšmės, morfologiniai anotatoriai kartais pasirenka neteisingą gramatinę pažymą). Visai neseniai, 2023 m. pabaigoje, pradėjo veikti ir nauja žodžių ir teksto analizės programa *Morfuoklis* anotatoriaus *Semantika.lt* pagrindu. Su šia programa galima ir sintezuoti, t. y. pagal nurodytas gramatines pažymas sugeneruoti norimą žodį ar kelis.

Daugiausia anotavimo klaidų pasitaiko dėl morfologinio daugiareikšmiškumo: nors programos ir tobulinamos, vis dar labai dažnai neatpažįsta konteksto (arba jį atpažįsta netinkamai). Pastaruoju metu skirta nemažai dėmesio šiai lietuvių kalbos automatinės analizės problemai spręsti – sukurta įvairių įrankių, kurie gali riboti semantinį, sintaksinį ir morfologinį daugiareikšmiškumą, sudaryta taisyklių, padedančių vienareikšminti tekstus, išanalizuotas ir suklasifikuotas lietuvių kalbos morfologinis daugiareikšmiškumas (Rimkutė, Grigonytė 2006: 32).

Homoformos dažnos ne tik lietuvių kalboje. Panaši situacija išryškėjo ir nagrinėjant kitų kalbų tekstynus, pavyzdžiui, čekų kalboje morfologiškai daugiareikšmės formos sudaro 46 proc., slovėnų kalboje – apie 8,01 proc., rumunų – 40 proc., anglų – 38,65 proc., vengrų – 21,58 proc. (Hajič 2004: 173, cituojama iš Rimkutė 2006a: 47). Remiantis estų mokslininkų pateiktais duomenimis (Puolakainen 2012: 194), galima teigti, kad beveik pusė estų kalbos žodžių yra morfologiškai daugiareikšmiai.

Evalda Jakaitienė (2009: 87) teigė, kad būta svarstymų, ar toks reiškinys kaip žodžių daugiareikšmiškumas apskritai gali egzistuoti – esą jeigu viena

garsų seka turi dvi skirtingas reikšmes, tokiu atveju tai ir turėtų būti du atskiri žodžiai, o ne vienas daugiareikšmis. Svarbu paminėti, kad mokslininkė čia kalba apie leksinį, o ne morfologinį daugiareikšmiškumą. Tačiau tokia nuomonė nėra populiari – pasak Skaistės Aleksandravičiūtės (2008: 275), daugiareikšmiškumas yra sistemingas kalbos reiškiny, nes sietinas su lietuvių kalbos morfologinės sistemos ypatybėmis ir linksnių sinkretizmu, be to, susijęs su kalbos ekonomija ir ribotumu.

2. SVARBIAUSI MORFOLOGINIO DAUGIAREIKŠMIŠKUMO TERMINAI

Kalbotyroje morfologinis daugiareikšmiškumas dažniausiai suprantamas ir įvardijamas keliais skirtingais terminais. Toliau pateikiama diskusija dėl terminų, t. y. kaip pavadinti morfologiškai daugiareikšmes formas: homoformomis, homonimais, homografais ar kitaip (neaptariamais terminais, susiję su kitais daugiareikšmiškumo lygmenimis, toliau didžiausias dėmesys skiriamas tik morfologinio daugiareikšmiškumo atvejams).

Plačiausios reikšmės terminas, kurį iš dalies būtų galima vartoti šiame straipsnyje, yra *homonimai*, jie dažniausiai apibrėžiami kaip „žodžiai, turintys tą pačią raišką ir visiškai skirtingas reikšmes“ (Jakaitienė 2009: 101), kaip „vienodai skambantis su kitu žodžiu, tai pačiai kalbos daliai priklausantis, bet skirtingą reikšmę turintis žodis“ (Gaivenis, Keinys 1990) ar kaip „vienodai tariami žodžiai (būna ir homoniminių formų, morfemų)“ (Šukys 2004: 9).

Kaip minėta anksčiau, morfologiškai daugiareikšmėmis vadinamos kaitomosios ir nekaitomosios, vienodai ir skirtingai tariamos formos. Analizuojant morfologinį daugiareikšmiškumą, leksinė reikšmė nėra svarbi (tai dažniausiai akcentuojama aprašant homonimus), todėl akivaizdu, kad terminas *homonimai* yra per platus kalbant apie morfologiškai daugiareikšmes formas.

Terminas *homoformos* kelia nemažai neaiškumų lingvistams. Kai kurie iš jų (Barauskaitė 1979: 43–44; Barauskaitė ir kt. 1995: 23) homoformomis laiko skirtingas to paties žodžio formas arba skirtingų žodžių panašiai skambančias formas, pvz., *aprašai* – ir daiktavardžio *aprašas* dgsk. vardininkas, ir veiksmažodžio *aprašyti* esamojo laiko antrojo asmens forma. Nikolajus Pavlovičius Kolesnikovas homoformas dar vadina paradigmniais homonimais (Kolesnikov 1987: 7, cituota iš Nevzorova ir kt. 2005: 231), o Francis Katamba sutampančias tų pačių ar skirtingų kalbos dalių formas vadina sinkretizmu (Katamba 1994: 15).

Šiame straipsnyje morfologiškai daugiareikšmės formos vadinamos homoformomis. Kaip paaiškės iš toliau pateikiamos šio termino vartosenos apžvalgos, terminas *homoformos* šiame straipsnyje vartojamas platesne, nei įprasta kalbotyros darbuose, reikšme. Todėl toliau, išsamiau nesigilinant į visas homonimijos rūšis, aprašoma homoformų samprata kalbotyros darbuose.

E. Jakaitienė (2009: 103) homoformas apibrėžia kaip atskiras, savo garsine sandara atsitiktinai sutampančias visai nehomonimiškų žodžių poras. Kaip pavyzdį mokslininkė pateikia tokį sakinį: *Mama nuo vaiko muses vaiko*, kuriame sutampa daiktavardžio *vaikas* vnsk. kilmininkas ir veiksmažodžio *vaikyti* esamojo laiko trečiojo asmens forma. Remiantis mokslininke, gali sutapti ir kai kurių žodžių kelios gramatinės formos, pvz.: *lipti, lipa, lipo* („judėti aukštyr ar žemyn“) ir *lipti, limpa, lipo* („kibti, klijuotis“). Čia sutampa veiksmažodžių bendratis ir iš jos padarytos formos. Problemiška tai, kad tokius atvejus kiti priskiria ne homoformoms, o daliniams homonimams (Barauskaitė ir kt. 1995: 24).

Kazimiero Gaivenio ir Stasio Keinio *Kalbotyros terminų žodyne* (KTŽ) homoformomis laikomos vienodai tariamos ir rašomos skirtingų žodžių formos (Gaivenis, Keinys 1990). Vadinas, homoformoms nepriskiriami linksnių sinkretizmo atvejai, kai sutampa keli to paties žodžio linksniai. Tačiau toks apibrėžimas gali kelti neaiškumų skiriant homoformas ir dalinius homonimus. Žodyne pastarieji apibrėžti kaip „homonimai, kurių ne visos morfemos visiškai sutampa ar kurie priklauso skirtingoms kalbos dalims“ (Gaivenis, Keinys 1990: 80). E. Jakaitienė dalinius homonimus apibrėžia kaip tokius žodžius, „kurių arba vieno visos gramatinės formos sutampa su dalimi kito žodžio formų, arba kurių sutampa bent po kelias formas“ (Jakaitienė 2009: 101–102). Pasak jos, dažniausiai tokiu atveju vienas žodis būna kaitomasis, o kitas nekaitomasis, pavyzdžiui, *gaila* – tai ir daiktavardis, reiškiantis gailestį, gailėjimąsi, irrieveiksmis (pavyzdžiai iš morfologiškai anotuoto teksto).

Pasak E. Jakaitienės, analitinėse kalbose daliniais homonimais laikomi žodžiai, kurie gali priklausyti skirtingoms kalbos dalims, pavyzdžiui, angliškas žodis *tender* gali būti ir daiktavardis, reiškiantis pasiūlymą, ir būvardis, reiškiantis švelnų, jautrų, ir veiksmažodis, reiškiantis „pasiūlyti, pateikti“ (Jakaitienė 2009: 102). Kiti mokslininkai (Kaljanov 2023: 27) tokius atvejus vadina kiek kitaip – gramatiniais homonimais. Tai fonetiškai su-

tampančios to paties žodžio formos, turinčios skirtingas gramatines reikšmes, pavyzdžiui, *ranka* (vnsk. vardininkas) ir *ranka* (vnsk. įnagininkas).

Homoformų apibrėžimą galima rasti ir *Lietuvių kalbos enciklopedijoje* (LKE). Čia rašoma, kad homoformomis laikomos kai kurios sutampančios skirtingoms kalbos dalims priklausančių žodžių formos, pavyzdžiui, *grūdo* – daiktavardžio *grūdas* vnsk. kilmininkas ir veiksmažodžio *grūsti* būtojo kartinio laiko trečiojo asmens forma (LKE 1999: 247–248). Iš tokio apibrėžimo neaišku, ar homoformomis galima laikyti sutampančias tos pačios kalbos dalies žodžių formas, pavyzdžiui, *tapo* (šiuo atveju skiriasi kitos pagrindinės formos: *tapyti*, *tapė* ir *tapti*, *tampa*).

Morfologiškai anotuotame tekстыne pateikiamos tik rašytinės ir neskirčiuotos formos. Dažnai aktualu tokių formų kirčio vieta ar priegaidė – neretai morfologiškai daugiareikšmės formos skiriasi prozodija, yra skirtingai tariamos. Lingvistai dažnai tokius atvejus vadina homografais. KTŽ (1999: 23–24) rašoma, kad homografai yra „vienodai rašomi, bet skirtingai tariami žodžiai: *līngė* (automobilio detalė) ir *liņgė* (paukštis)“, o E. Jakaitienė šią sąvoką apibrėžia taip: tai „tokie žodžiai (ar jų formos), kurie vienodai rašomi, bet ištariami skirtingai (skiriasi a) kirčio vieta ar b) priegaide)“ (Jakaitienė 2009: 103).

Gali būti sunku suvokti ir kuo homografai skiriasi nuo homonimų ir homoformų. Remiantis LKE (1999: 248) pateikta Vinco Urbučio nuomone ir čia patektu homografų apibrėžimu, homografai yra taip pat rašomi žodžiai ar jų formos. Teigiama, kad homografai lietuvių kalboje dažniausiai kartu yra ir homonimai (pvz., forma *kartą* gali būti ir žodžio *karta* (t. y. *generacija*), ir žodžio *kartas* (t. y. *sykis*) galininkas). Enciklopedijoje homografais laikomi ir tokie atvejai, kai sutampa skirtingoms kalbos dalims priklausančios žodžių formos (LKE 1999: 248), pavyzdžiui, *aprašai* (ši forma gali būti ir daiktavardžio *aprašas* dgsk., ir veiksmažodžio *aprašyti* esamojo laiko trečiasis asmuo – pavyzdžiai paimti iš morfologiškai anotuoto lietuvių kalbos tekstyno). Tačiau bendro susitarimo nėra. J. Šukys rašo, kad „žodžių, kurie šiaip jau yra vienodi ir kuriuos DLKŽ numeruoja kaip homonimus, bet kurie skiriasi kirčio vieta, kirčiuote, kirčiuoto skiemens priegaide, tikraisiais homonimais galima nelaikyti, o skirti prie homoformų. Homoformų (kitai – homoniminių formų) kalboje pasitaiko daug daugiau nei tikrųjų homonimų, ir čia dažnai būna visiškai skirtingos kalbos dalys: *maišo* (daiktavardžio *maišas* kilmininkas) – *maišo* (būtojo kartinio laiko trečiojo asmens veiksmažodžio *maišyti* forma); *jos* (įvardžio

ji dgsk. vardininko bei vnsk. kilmininko forma) – jos (veiksmažodžio *joti* būsimąjo laiko trečiojo asmens forma)“ (Šukys 2004: 9). Šiai nuomonei antrina ir kiti mokslininkai, teigdami, kad „yra ir homoformų, kurios vienodai rašomos, bet skiriasi kirčiu“, pvz., *báltų* ir *baltų* (*báltai* ir *báltas*) (Barauskaitė ir kt. 1995: 248).

Klausimų kelia ir N. P. Kolesnikovo pateiktas homografų apibrėžimas (Kolesnikov 1987: 6, cituota iš Nevzorova ir kt. 2005: 231). Pasak jo, homografams gali būti priskirtos: 1) skirtingą gramatinę ir leksinę reikšmę turinčios, bet vienodai rašomos žodžių formos (kaip pavyzdį galima pateikti lietuvišką žodį *balti*, t. y. sutampa veiksmažodžio bendratis ir vyr. g. būdvardžio *baltas* dgsk. vardininko forma); 2) skirtingą leksinę, bet vienodą gramatinę reikšmę turinčios taip pat rašomos žodžių formos (pvz., forma *žiedas* gali būti trys skirtingi žodžiai, turintys skirtingas leksines reikšmes, t. y. žodis gali reikšti ir augalo dalį, ir papuošalą, ir kelio rūšį); 3) skirtingą gramatinę, bet vienodą leksinę reikšmę turinčios žodžių formos (plg. lietuviško veiksmažodžio *turėti* esamojo laiko antrojo asmens vnsk. formą *turì* ir esamojo laiko trečiojo asmens formą *tùri*). Galima suprasti, kad mokslininkas supina homografų ir homoformų terminus, kadangi šioje klasifikacijoje nagrinėjami ir skirtingų, ir tų pačių reikšmių žodžiai bei jų formos. Tačiau taip suprasti homoformas atrodytų gana logiška, nes morfologiškai anotuotuose tekstuose apie 20 proc. daugiareikšmių žodžių skiriasi būtent prozodiniais elementais, t. y. tarimu (Rimkutė 2002: 7).

Jurgis Pakerys vartoja dar vieną terminą pavadinti tiems atvejams, kai sutampa skirtingai tariamos to paties žodžio formos, – *homografinės žodžių formos*. Pabrėžiama, kad tai skiriasi nuo homografų (Pakerys 2009: 130). O štai Vincentas Drotvinas homoformas, homografus ir homofonus vadiną dar vienu terminu – *homoniminės žodžių formos* (Drotvinas 1986: 91, cituota iš Rimkutė 2006a: 43). Toks pavadinimas pasirodė priimtinas pirmajai morfologinio daugiareikšmiškumo tyrėjai E. Rimkutei, bet, siekdamas išvengti tokio ilgo termino vartojimo, ji nutarė tiesiog pasirinkti *homoformų* terminą (Rimkutė 2003b: 36). Tokios sampratos laikosi ir šio straipsnio autorė.

Morfologiškai daugiareikšmėms formoms pavadinti skiriamas ir dar vienas terminas – *paronimai*. Jais vadinami „panašiai tariami arba rašomi žodžiai, kurių reikšmė dažniausiai skirtinga; kartais ji gali būti ir šiek tiek artima, bet paprastai nesinonimiška (tokiu atveju dažnai sakoma, kad žodžius sieja semantinis bendrumas, kitaip – tas pats semantinis laukas)“ (Šukys 2004:

9–10). J. Šukys aiškina, kad paronimams rašytinėje kalboje galima priskirti ir homografus, t. y. vienodai rašomus, bet skirtingai tariamus žodžius, pvz.: *draugė* – *draugė* („draugija, būrys“), o sakytinėje kalboje galima priskirti homofonus, t. y. vienodai tariamus, bet skirtingai rašomus žodžius, pvz., *rūgti* – *rūkti*. Vis dėlto *paronimo* terminas kalbotyroje turi gana aiškiai apibrėžtą reikšmę, todėl nuspręsta išplėsti *homoformos* reikšmę ir toliau vartoti būtent šį terminą. Be to, šio straipsnio autorė laikosi nuostatos, kad tai pačiai ar skirtingoms kalbos dalims priskiriamos kaitybinės formos, nesutampančios kirčio vieta, priegaide ar balsio kiekybe, gali būti ir *homoformos*, ir *homografai*. Kitaip tariant, kai kurie homografai gali būti ir *homoformos*, o kai kurios *homoformos* – *homografai* (kai atsižvelgiama į tarties aspektus ir analizuojama morfologiškai).

Negana to, lingvistų darbuose vis dar nėra apibrėžti kai kurie konkretūs atvejai, pavyzdžiui, neaišku, kaip vadinti dviejų ar daugiau nekaitomųjų kalbos dalių sutapimus, todėl, remiantis E. Rimkute (2006a: 44), tokius sutapimus šiame tyrime nutarta priskirti nekaitomųjų kalbos dalių *homoformoms*.

Kituose skyriuose rašant apie *homoformas* laikomasi anksčiau išdėstytos *homoformų* sampratos. Ketvirtame skyriuje pateikiama *homoformų*, t. y. morfologinio daugiareikšmiškumo vienetų, klasifikacija, remiantis morfologiškai anotuoto teksto duomenimis. Straipsnyje vartojami daugiausia *morfologiškai daugiareikšmių formų* ir *homoformų* terminai.

3. LIETUVIŲ KALBOS HOMOFORMŲ KLASIFIKACIJA PAGAL SUTAMPANČIAS KALBOS DALIS

Kaip minėta anksčiau, toms pačioms ar skirtingoms kalbos dalims priklausančios sutampančios formos šiame straipsnyje vadinamos *homoformomis*. E. Rimkutė (2006a: 40) išskyrė tris jų tipus: 1) kaitomųjų kalbos dalių *homoformos*; 2) kaitomųjų ir nekaitomųjų kalbos dalių *homoformos*; 3) nekaitomųjų kalbos dalių *homoformos*.

Kaitomųjų kalbos dalių *homoformos* – tai tų pačių ar skirtingų kalbos dalių morfologiškai sutampančios formos (Rimkutė 2006a: 52). E. Rimkutė kaip pavyzdžius pateikia tokius sakinius: *lįsk į būdą* (daiktavardžio *būda* vnsk. galininkas) ir *sėdėk* vs. *atrado naują gydymo būdą* (daiktavardžio *būdas* vnsk. galininkas). Kaip matyti, čia sutampa daiktavardžių *būda* ir *būdas* vnsk. galininko formos. Tačiau nebūtinai turi sutapti tos pačios kalbos

dalis skirtingų žodžių formos – gali sutapti ir dviejų skirtingų kaitomųjų kalbos dalių formos ar to paties žodžio dvi skirtingos formos. E. Rimkutė pateikia tokį pavyzdį: *anglų kalbà* (daiktavardis) vs. *kai kurie Lietuvos politikai nuolat kalba* (veiksmažodis) *absurdus*. Kaip matyti, šiuose sakiniuose sutampa kai kurios daiktavardžio *kalba* formos ir veiksmažodžio *kalbėti* esamojo laiko trečiojo asmens forma.

E. Rimkutei (2006a: 48) išanalizavus 10 000 morfologiškai anotuotų žodžių formų, esančių morfologiškai anotuotame tekстыne, paaiškėjo, kad kaitomųjų kalbos dalių homoformos sudaro didžiausią morfoliginio daugiareikšmiškumo dalį – net 76,2 proc.

Kaitomųjų ir nekaitomųjų kalbos dalių homoformos – dviejų ar daugiau skirtingų kalbos dalių, iš kurių viena kaitomoji, o kita ar kitos nekaitomosios, sutampančios formos (Rimkutė 2006a: 59). Šiai kategorijai priskiriamos daiktavardžių,rieveiksmių, prielinksnių, jungtukų ar dalelių; būdvardžių irrieveiksmių; veiksmažodžių,rieveiksmių, prielinksnių arba dalelių ir kitų sutampančių kalbos dalių homoformos. Jos sudaro 6,3 proc. viso morfoliginio daugiareikšmiškumo atvejų morfologiškai anotuotame tekстыne. Kaip pavyzdį mokslininkė pateikia tokius žodžių junginius: *nāmo* (daiktavardis) *stogas* vs. *einu namō* (rieveiksmis). Galima matyti, kad čia sutampa daiktavardis surieveiksmiu. Kaitomųjų ir nekaitomųjų kalbos dalių homoformos dažniausiai sudarytos iš trijų ar daugiau sutampančių kalbos dalių, t. y. yra daugianarės.

Nekaitomųjų kalbos dalių homoformos – tai sutampančios dvi arba daugiau nekaitomųjų kalbos dalių. Į šią kategoriją patenka jungtukų, dalelių,rieveiksmių ir pan. sutapimai. Nekaitomųjų kalbos dalių homoformos yra antroji pagal dažnumą homoformų rūšis, ji sudaro 17,5 proc. visų homoformų morfologiškai anotuotame tekстыne (Rimkutė 2006a: 63).

Rengiant LKHŽ, šios trys kategorijos suskirstytos į smulkesnes 53 rūšis. Iš esmės remtasi E. Rimkutės (2006a) siūloma klasifikacija, bet ši klasifikacija išplėsta.

LKHŽ ištrauka pateikta 1 paveiksle. Žodyne sutampantis žodis ar žodžio forma rašoma pajuodintai, kitoje eilutėje nurodoma, kokiomis kalbos dalimis laikytina ši forma; kampiniuose skliaustuose pateikiama antraštinė forma, toliau pateikiamos išsamios gramatinės pažymos. Nekaitomosios kalbos dalys pateikiamos abėcėlės tvarka, kaitomosios, taip pat kaitomosios ir nekaitomosios – pagal kalbos dalių sutapimus, o viduje – pagal abėcėlę.

1. KAITOMŲJŲ KALBOS DALIŲ HOMOFORMOS

1.1. (*i*)o, *ė* linksn. dkt., bdv., įv, sktv., dlv. vns. K. + (*i*)o, *ė* linksn. dkt., bdv., įv, sktv., dlv. dgs. V. (dgs. Š.)

abatijos

dkt. <abatija>, dkt. mot. g. vns. K., dkt. mot. g. dgs. V., dkt. mot. g. dgs. Š.

abejingos

bdv. <abejingas>, bdv. teig. nelygin. neįvardž. mot. g. dgs. V., bdv. teig. nelygin. neįvardž. mot. g. vns. K.

abejonės

dkt. <abejonė>, dkt. mot. g. vns. K., dkt. mot. g. dgs. V., dkt. mot. g. dgs. Š.

abejotinos

vksm. <abejoti(-ja,-jo)>, dlv. teig. nesngr. reik. neįvardž. mot. g. dgs. V., dlv. teig. nesngr. reik. neįvardž. mot. g. vns. K.

abėcėlės

dkt. <abėcėlė>, dkt. mot. g. dgs. V., dkt. mot. g. vns. K., dkt. mot. g. dgs. Š.

1 PAV. Lietuvių kalbos homoformų žodyno ištrauka

Kai kurios homoformos gali būti priskirtos kelioms grupėms, pavyzdžiui, *vieši* – tai būdvardžio *viešas* dgsk. vardininko ir veiksmažodžio *viešėti* homoformos, taip pat veiksmažodžio *viešėti* esamojo laiko vnsk. antrojo ir trečiojo asmens sutapimas; *pirmiausia* – tai būdvardžio ir skaitvardžio, taip pat ir tų kalbos dalių mot. g. vnsk. vardininko, įnagininko ir bevardės giminės sutapimas. Todėl tas pats žodis žodyne gali būti pakartotas kelis kartus prie skirtingų homoformų rūšių. Vadinasi, prie konkrečios formos žodyne ne išsyk pateiktos visos galimos gramatinės pažymos, o tik tos, kurios tinka konkrečiai sutapimų rūšiai, pavyzdžiui, daiktavardžių ir būdvardžių sutapimams ar pan. Vis dėlto daugiausia žodyne tokių morfologinio daugiareikšmiškumo atvejų, kurie gali būti priskirti tik vienai homoformų rūšiai.

LKHŽ kaitomųjų kalbos dalių sutapimai suskirstyti į 41 rūšį. Kaitomųjų ir nekaitomųjų kalbos dalių homoformos žodyne suskirstytos į 11 rūšių. Nekaitomųjų kalbos dalių sutapimai žodyne smulkiau neskirstyti, kadangi sudarant šį žodyną buvo remtasi būtent E. Rimkutės (2006a) klasifikacija. Ji visų nekaitomųjų kalbos dalių sutapimus priskyrė bendrai rūšiai.

4. LIETUVIŲ KALBOS HOMOFORMŲ KLASIFIKACIJA PAGAL ATSIDIRADIMO PRIEŽASTIS

Kadangi LKHŽ sudaro ištis didelis morfologiškai daugiareikšmių žodžių skaičius (apie 36 tūkst.), galima mėginti nustatyti lietuvių kalbos homoformų atsiradimo priežastis. Atsižvelgiant į duomenis homoformų žodyne ir į *Lemuoklio* siūlomas morfologines anotacijas, lietuvių kalbos homoformas pagal atsiradimo priežastis galima suklasifikuoti į penkias grupes. Toliau visos šios grupės išsamiai pristatomos.

4.1. Teoriškai galimos homoformas, atsiradusios dėl žodynų sudarymo specifikos

Kalbos morfologinį daugiareikšmiškumą pradėjus analizuoti kompiuteriu, ėmė aiškėti vienas svarbus skirstymas: morfologinis daugiareikšmiškumas gali būti tikėtinas ir netikėtinas. Tikėtinas – toks, kokį žmogus dažniausiai suvokia. Netikėtinas daugiareikšmiškumas gali būti dviejų tipų – morfologinio anotatoriaus sugeneruoti dariniai ir retai dabartinėje kalboje vartojami (arba visai nebevartojami) žodžiai.

Remdamasis savo morfologiniais ištekliais, *Lemuoklis* pateikia realybėje nevartojamų žodžių, kurių kai kurios formos sutampa su lietuvių kalboje dažnai vartojamomis kalbos dalimis ar jų formomis. Taip atsitinka dėl to, kad žodynuose pateikiami ne visi net ir reguliariai sudaromi dariniai. E. Rimkutė, 2006 m. tvarkydama morfologiškai anotuotą tekstyną, sudarė sąrašą morfologinio anotatoriaus sugeneruotų nerealių žodžių. Šis sąrašas įtrauktas į morfologinį anotatorių kaip ignoruotini žodžiai, todėl dabar problema, kai *Lemuoklis* pats generuoja žodžius, iš esmės išspręsta.

Lemuoklio leksikos pagrindą sudaro žodžiai, esantys *Dabartinės lietuvių kalbos žodyne* (taip pat ir kituose žodynuose), kuriame pateiktas ne vienas dabar ypač retai vartojamas žodis, pvz.: *rata*, *girdas*, *metinys*, *ložėti*, *baisis*, *graudis*, *vajys* ir t. t. Jei šių žodžių formos sutampa su įprastai kalboje vartojamų žodžių formomis, *Lemuoklis* nurodo visus variantus, o tai lingvistams labai apsunkina tekstyno tvarkymą.

Antroji problema, kai archajiški žodžiai sutampa su įprastais, liko iš dalies neišspręsta. Tokių homoformų daugiausia tarp kaitomųjų ir nekaitomųjų kalbos dalių sutapimų, pavyzdžiui, *Lemuoklio* analizėje žodžių formos *bravo*, *ate*, *arba* laikomos homoformomis, nurodoma, kad šie žodžiai – ir daiktavardžiai, nes tokią informaciją galima rasti ir žodynuose (interneti-

niame *Tarptautinių žodžių žodyne* (TŽŽ) žodis *bravas* apibrėžiamas kaip „XVII–XVIII a. Italijoje – samdomasis žudikas, avantiūristas, galintis panaudoti bet kokią smurtą“, *atas* – „Laoso piniginių vienetų kipo šimtoji dalis“, *arba* – „1. dengtas medinis dviratis vežimas, vartotas Vid. Azijos tautų; vart. Afganistane sunkiems kroviniams vežioti; 2. keturratis vežimas javams vežti; vartotas Kaukaze ir P. Ukrainoje“.

Tokius žodžius paprastai užtrunka atskirti nuo netikėtinų homoformų, nes kiekvieną jų reikia tikrinti įvairiuose žodynuose ir tekstynuose. Sunku nuspręsti ir tai, ar reikėtų įtraukti juos į LKHŽ, nes vien kitų žodynų duomenimis remtis nebeužtenka. Iš šios homoformų atsiradimo priežasties matyti, kad susiduriama su dilema: kuo didesnis kalbos analizės programos leksikonas, tuo daugiau programa turėtų atpažinti žodžių ir jų formų, bet kartu tai reiškia, kad gali atsirasti nenumatytų sutapčių – morfologinio daugiareikšmiškumo atvejų. Taigi, siekiant kuo tikslesnių rezultatų, reikia gerai apgalvoti kompiuterinėse programose naudojamų žodžių atrankos principus.

4.2. Homoformas, atsiradusios dėl prozodinių elementų skirtumų

Homoformų dėl prozodinių elementų skirtumų (pvz., kirčio vietos, priegaidės, balsio ilgumo) atsiranda tik rašytinėje kalboje. Kadangi lietuvių kalbos morfologiniai anotatoriai analizuoja rašytines, be kirčio ženklų žodžių formas, homoformas neretai skiriasi prozodika.

Tokių homoformų pavyzdžių daugiausia tarp kaitomųjų kalbos dalių sutapimų, jų yra didžiojoje dalyje smulkesnių rūšių. Dažnai tai būna sutapusios dvi ar daugiau vieno žodžio formos. Tokie pavyzdžiai žodyne yra 11 rūšių: *(i)o* linksniuotės vardažodžių vnsk. vardininko sutapimai su vnsk. įnagininku ir šauksmininku (pvz.: *agentūrà* vs. *agentūra*, *mamà* vs. *māma*); *(i)o*, *ė* linksniuotės vardažodžių vnsk. kilmininko ir dgsk. vardininko sutapimai (pvz.: *absoliučiõs* vs. *absoliučios*; *adekvačiõs* vs. *adekvācijas*; *áikštės* vs. *aikštės*); *ė* linksniuotės daiktavardžių vnsk. įnagininko ir šauksmininko¹⁰ sutapimai (pvz.: *katė* vs. *kāte*, *dumblė* vs. *duṁble*); bendraties ir neveikiamosios rūšies būtojo laiko vyr. g. dalyvių dgsk. vardininko sutapimai (pvz.:

¹⁰ Atkreiptinas dėmesys, kad šauksmininkas daug kur galimas tik teoriškai. Sudarant homoformų žodyną, neatsižvelgta į daiktavardžių semantiką, laikyta, kad galimos visų daiktavardžių šauksmininko formos.

apdirbti vs. *apdirbti*; *apeiti* vs. *apeiti*); tariamosios nuosakos ir neveikiamosios rūšies būtojo laiko vyr. g. dalyvių dgsk. kilmininko sutapimai (pvz.: *aplýtų* vs. *aplytų*; *deñgtų* vs. *dengtų*); (i)a linksniuotės daiktavardžių vnsk. vietininko ir šauksmininko sutapimai (pvz.: *aidè* vs. *áide*; *garažè* vs. *garāže*); (i)u ir i linksniuotės daiktavardžių vnsk. vardininko ir dgsk. galininko sutapimai (pvz.: *apimtīs* vs. *āpimtis*; *āpskritīs* vs. *apskritīs*); esamojo laiko vnsk. trečiojo ir antrojo asmenų sutapimai (pvz.: *neprigīrði* vs. *neprigirdi*; *negālì* vs. *negali*); įvardžių mot. g. vnsk. naudininko ir bevardės giminės sutapimai (pvz., *tái* vs. *taĩ*, *šítai* vs. *šitaĩ*; žinoma, *šítai* gali būti ir dalelytė, tačiau tokie sutapimai žodyne patenka jau į kitą kategoriją – kaitomųjų ir nekaitomųjų kalbos dalių sutapimų kategoriją, nes žodyne, kaip minėta, pateikiamos tik konkrečiai rūšiai tinkančios gramatinės pažymos); (i)u linksniuotės vardažodžių vnsk. galininko ir dgsk. kilmininko¹¹ sutapimai (pvz.: *ālų* vs. *alų*; *liétų* vs. *lietų*); veiksmažodžių bendraties ir esamojo laiko antrojo asmens sutapimai (pvz.: *iñkšti* vs. *inkšti*; *klýsti* vs. *klysti*).

Prozodinių elementų skirtumų nulemtų homoformų atsiranda ir sutapus dviem ar daugiau skirtingų kaitomųjų kalbos dalių. Tokio tipo homoformų žodyne yra 7 rūšys: daiktavardžių ir veiksmažodžių sutapimai (pvz.: *Adomo* – sutampa tikrinio daiktavardžio *Adōmas*¹² vnsk. kilmininko ir neveikiamosios rūšies dalyvio *ādomas* vnsk. kilmininko formos; *apsaugos* – sutampa daiktavardžio *apsauga* vnsk. kilmininko, dgsk. galininko ir veiksmažodžio *apsaugoti* būsimosios laiko formos, plg.: *apsaugōs* vs. *apsáugos*); daiktavardžių ir būdvardžių sutapimai (pvz., *atlaidus* – sutampa daiktavardžio *atlaidai* dgsk. galininko ir būdvardžio *atlaidus* vnsk. vardininko formos, plg. *ātlaidus* vs. *atlaidūs*); veiksmažodžių ir būdvardžių sutapimai (pvz., *añtrinę*, *asmėninę* vs. *antrinę*, *asmeninę* – sutampa veiksmažodžių *antrinti*, *asmeninti* veikiamosios rūšies būtojo kartinio laiko dalyvių ir būdvardžių *antrinis*, *-ė*, *asmeninis*, *-ė* mot. g. vnsk. galininko formos); veiksmažodžių ir įvardžių sutapimai (pvz.: *kitaĩ* vs. *kítai*; *kùria* vs. *kurià* – sutampa veiksmažodžių *kisti*, *kurti* ir įvardžių *kitas*, *kuri* formos); daiktavardžių, būdvardžių ir veiksmažodžių sutapimai (pvz.: *balti* – daiktavardžio *baltis* vnsk. šauksmininko forma *bal̃ti*; būdvardžio *baltas* dgsk. vardininko forma *balti*; bendratis *bálti*); daiktavardžių ir skaitvardžių ar įvardžių sutapimai (pvz.: *Anos* – sutampa tikrinio daiktavardžio

¹¹ Kartais (kaip ir toliau pateiktais atvejais) dgsk. kilmininkas galimas tik teoriškai.

¹² Tiesa, anotatorius geba skirti didžiąsias raides nuo mažųjų. Vis dėlto morfologinis daugiareikšmiškumas tokiu atveju galėtų susidaryti, jei ir dalyvis būtų rašomas didžiąja raide, pavyzdžiui, sakinio pradžioje.

Ana ir įvardžio *ana* formos, plg. *Ānos* vs. *anōs*; *kelias* – sutampa daiktavardis *kēlias* ir įvardis *keliās*); daiktavardžių, veiksmažodžių ir įvardžių ar būdvardžių, ar skaitvardžių sutapimai (pvz.: *dvejės* – sutampa skaitvardis *dvėjės*, daiktavardžio *dveja* vnsk. kilmininko forma *dvejōs* ir dgsk. vardininko forma *dvėjės*, taip pat būsimąjo laiko veiksmažodis *dvejōs*).

Tokių sutapimų taip pat pastebėta tarp kaitomųjų ir nekaitomųjų kalbos dalių homoformų, jie žodyne yra 8 rūšių: veiksmažodžių, būdvardžių irrieveiksmių sutapimai (pvz., *blogai* – sutampa veiksmažodžio *blogti* vnsk. antrojo asmens būtojo kartinio laiko forma *blōgai*, būdvardžio *blogas*, *-a* mot. g. vnsk. naudininko forma *blōgai* irrieveiksmis *blogai*); būdvardžių irrieveiksmių sutapimai (pvz., *aklai*, *efektyviai* – sutampa būdvardžių mot. g. naudininko formos *āklai*, *efektīviai* irrieveiksmiai *aklai*, *efektyviai*); daiktavardžių irrieveiksmių sutapimai (pvz., *seniai* – sutampa daiktavardžio *senis* dgsk. vardininko forma *sēniai* irrieveiksmis *seniai*); daiktavardžių, būdvardžių ar įvardžių irrieveiksmių sutapimai (pvz.: *aiškiai*, *aukštai* – sutampa daiktavardžių *aiškus*, *aukštas* dgsk. vardininko formos *aiškiai*, *aūkštai*, būdvardžių *aiškus*, *-i*, *aukštas*, *-a* mot. g. naudininko formos *áiškiai*, *áukštai* irrieveiksmiai *áiškiai*, *aukštai*); daiktavardžių, veiksmažodžių, būdvardžių irrieveiksmių sutapimai (pvz., *bukai* – sutampa daiktavardžio *bukas* dgsk. vardininko ir vnsk. šauksmininko formos *būkai*, veiksmažodžio *bukti* būtojo kartinio laiko vnsk. antrojo asmens forma *bukaĩ*, būdvardžio *bukas* mot. g. vnsk. naudininko forma *būkai*,rieveiksmis *bukai*); skaitvardžių, prielinksnių irrieveiksmių sutapimai (pvz., *pirma* – sutampa skaitvardžio *pirmas*, *-a* mot. g. vnsk. vardininkas ir įnagininkas, bevardės giminės formos *pirmà*, *pĩrma*, ir prielinksnis beirieveiksmis *pĩrma*); veiksmažodžių irrieveiksmių sutapimai (pvz., *apmaudžiau* – veiksmažodžio *apmaudyti* būtojo kartinio laiko vnsk. pirmojo asmens forma *apmáudžiau* ir aukštesniojo laipsniorieveiksmis *apmaudžiaũ*), daiktavardžių, įvardžių irrieveiksmių ar dalelyčių sutapimai (pvz., *Ana* – sutampa tikrinio daiktavardžio *Ana* vnsk. vardininkas, įnagininkas bei šauksmininkas, įvardžio ir dalelytės *ana* formos, plg. *Anà*, *Āna* ir *anà*).

Jei *Lemuoklis* ir kiti morfologiniai anotatoriai gebėtų analizuoti sukirčiuotus tekstus¹³, tikėtina, morfologinis daugiareikšmiškumas gerokai sumažėtų.

¹³ Tą padaryti automatiškai nėra sudėtinga, pavyzdžiui, programa *Kirčiuoklis*, ji prieinama <https://kalbu.vdu.lt/mokymosi-priemonės/kirciuoklis>. Kita vertus, ši programa remiasi automatinės morfologinės analizės programos rezultatais, todėl susidaro užburtas ratas: nuo to, kiek gerai veikia morfologinis analizatorius, priklauso kirčiavimo programos kokybė. Paminėtina, kad *Kirčiuoklyje* morfologiškai daugiareikšmėms formoms pateikiami keli kirčiavimo variantai, taigi tai ne itin padėtų spręsti morfologinio daugiareikšmiškumo problemą.

Ir anksčiau buvo bandymų suklasifikuoti homoformas pagal prozodinius elementus. E. Rimkutė (2002: 91–98) išanalizavo 10 tūkst. gramatiškai anotuotų žodžių, pavartotų *Lietuvos ryto* straipsniuose. Ištirta, kad 83 proc. homoformų yra fonetiškai tapačios, o 17 proc. skiriasi kirčio vieta. Homoformas, kurios nesiskiria kirčio vieta, mokslininkė dar skirstė į dvi rūšis: besiskiriančias balsio kiekybe (pvz., veiksmažodis *māno* ir įvardis *māno*) ir besiskiriančias priegaide (pvz., veiksmažodžiai *klaūse* (lema – *klausyti*) ir *kláuse* (lema – *klausti*)). Anksčiau pateiktais pavyzdžiais aptartos ir kirčio vieta, ir priegaide besiskiriančios homoformos. Pastarojoje kategorijoje kartu aptarti ir tie atvejai, kurie skiriasi balsio kiekybe, tik jie atskirai neišskirti.

4.3. Homoformas, atsiradusios dėl sutampančių skirtingų žodžių

Į šią kategoriją patenka dalis kaitomųjų kalbos dalių homoformų, kaitomųjų ir nekaitomųjų kalbos dalių homoformos. Teoriškai visi kaitomųjų ir nekaitomųjų kalbos dalių homoformų kategorijos atvejai taip pat yra skirtingų reikšmių žodžiai, nes turi skirtingas lemas. Bet taip teigti ne visuomet teisinga, kadangi, pavyzdžiui, tarprieveiksmių ir būdvardžių sutapimų yra gana daug semantinio bendrumo, viena forma dažnai būna padaryta iš kitos. Todėl homoformas, atsiradusias dėl daugiareikšmių žodžių, būtų galima skirti dar į dvi smulkesnes kategorijas – skirtingos leksinės reikšmės homoformas ir dalinės leksinės reikšmės homoformas.

4.3.1. Skirtingos leksinės reikšmės homoformas

Į šią kategoriją patenka sutampančios formos, kurios darybiškai nėra susijusios. Tarp jų – 8 kaitomųjų ir nekaitomųjų kalbos dalių homoformų rūšys: veiksmažodžių, būdvardžių irrieveiksmių homoformos, pavyzdžiui, *griežtai* – šiuo atveju sutampa dalyvio *griežtas*, *-a* (lema – *griežti*) mot. g. vnsk. naudininko forma su visai kitą reikšmę turinčiurieveiksmiu *griežtai* ir būdvardžio *griežtas*, *-a* mot. g. vnsk. naudininko forma. Vadinas, šiuo atveju dalyvį surieveiksmiu ir būdvardžiu sieja skirtinga leksinė reikšmė, o būdvardį surieveiksmiu – iš dalies sutampanti, nes jie padaryti vienas iš kito (išsamiau žr. 4.3.2 sk.); daiktavardžių irrieveiksmių ar prielinksnių, jungtukų, dalelių, jaustukų, ištiktukų homoformos (pvz.: *dėlei* – sutampa daiktavardžio *dėlė* vnsk. naudininko ir prielinksnio *dėlei* formos; *opa* – sutampa daiktavardžio *opa* vnsk. vardininko, šauksmininko, įnagininko formos ir ištiktukas); įvardžių irrieveiksmių ar dalelių, jungtukų homoformos

(pvz., *kam* –rieveiskmio ir įvardžio sutapimas); daiktavardžių,rieveiskmių ir būdvardžių ar įvardžių homoformos (pvz., mot. g. daiktavardžio *kandis*, t. y. „maža peteliškė, kurios vikšrai kapoja drabužius, ėda javus, miltus, augalus“ (LKŽ 2005), vnsk. naudininkas *kandžiai* ir vyr. g. daiktavardžio *kandis*, t. y. „kandimas, įkandimas, įkasta vieta“ (LKŽ 2005), dgsk. vardininkas *kandžiai* sutampa su būdvardžio *kandus*, *-i* mot. g. vnsk. naudininku irrieveiskmiu *kandžiai* (tarp būdvardžio irrieveiskmio leksinės reikšmės yra dalinis sutapimas, o su daiktavardžiais *kandis* juos jau sieja skirtinga leksinė reikšmė); daiktavardžių, veiksmažodžių, būdvardžių,rieveiskmių ir prielinksnių homoformos (pvz., *arti* – sutampa daiktavardžio *artis* vnsk. šauksmininko, veiksmažodžio *árti* bendraties, būdvardžio *artus*, *-i* mot. g. vnsk. vardininko,rieveiskmio ir prielinksnio *artì* formos); veiksmažodžių, daiktavardžių irrieveiskmių homoformos (pvz., *kartais* – sutampa dalyvio *kártas* dgsk. įnagininko, daiktavardžio *kařtas* dgsk. įnagininko irrieveiskmio *kartais* formos); veiksmažodžių ir prielinksnių ar dalelyčių, jungtukų, jaus-tukų, ištiktukų, santrumpų homoformos (pvz.: *ak* – sutampa jaustukas ir daiktavardžio *akti* trečiojo asmens vnsk. liepiamosios nuosakos forma; *užvis* – sutampa veiksmažodžio *užvisti* trečiojo asmens būsimojo laiko irrieveiskmio *užvis* formos; *vysk* – sutampa žodžio *vyskupas* grafinės santrumpos ir veiksmažodžio *vysti* liepiamosios nuosakos trečiojo asmens vnsk. formos); daiktavardžių, įvardžių irrieveiskmių ar dalelyčių homoformos (pvz., *ana* – sutampa tikrinio daiktavardžio *Ana*, įvardžio ir dalelytės *ana* formos).

Skirtingos leksinės reikšmės homoformų galėtų būti ir tarp kaitomųjų kalbos dalių homoformų. LKHŽ jos sudaro 8 rūšis: skirtingos leksinės reikšmės daiktavardžių, veiksmažodžių ir būdvardžių homoformos (pvz., *anglių* – sutampa daiktavardžių *anglys* ir *anglė* dgsk. kilmininko formos); daiktavardžių ir veiksmažodžių homoformos (pvz., *adomą* – sutampa dalyvio *adomas* vnsk. galininko ir tikrinio daiktavardžio *Adomas* vnsk. galininko formos); veiksmažodžių ir įvardžių ar skaitvardžių homoformos (pvz.: *jos* – sutampa veiksmažodžio *joti* trečiojo asmens būsimojo laiko ir įvardžio *ji* dgsk. vardininko ir vnsk. kilmininko formos; *kitai* – sutampa veiksmažodžio *kisti* būtojo kartinio laiko antrojo asmens ir įvardžio *kita* vnsk. naudininko formos; *trejos* – sutampa veiksmažodžio *trejoti* būsimojo laiko trečiojo asmens ir skaitvardžio *treji*, *-os* vardininko formos); daiktavardžių, būdvardžių ir veiksmažodžių homoformos (pvz.: *kaltų* – sutampa veiksmažodžio *kalti* tariamosios nuosakos trečiojo asmens forma, dalyvio *káltas*, *-a* mot. ir vyr. g. dgsk. kilmininko forma, būdvardžio *kaľtas*, *-a* mot. ir vyr. g. dgsk. kilmininko forma ir

daiktavardžio *káltas* dgsk. kilmininko formos; *kartus* – sutampa dalyvio *kártas* vyr. g. dgsk. galininko, būdvardžio *kartus* vnsk. vardininko ir daiktavardžio *kartas* dgsk. galininko formos); būdvardžių, įvardžių, skaitvardžių ir daiktavardžių homoformos (pvz., *viena* – sutampa tikrinio daiktavardžio *Viena* ir skaitvardžio, įvardžio bei būdvardžio *vienas*, *-a* vnsk. vardininko, įnagininko, šauksmininko formos); daiktavardžių, veiksmažodžių ir įvardžių, skaitvardžių homoformos (pvz.: *dvejos* – sutampa daiktavardžio *dveja* vnsk. kilmininko, dgsk. vardininko, veiksmažodžio *dvejoti* trečiojo asmens būsimąjo laiko ir skaitvardžio *dvejos* formos; *maną* – sutampa iš veiksmažodžio *manyti* padaryto veikiamosios rūšies esamojo laiko dalyvio dgsk. vardininko, įvardžio *manà* mot. r vyr. g. vnsk. galininko ir daiktavardžio *māna* vnsk. galininko formos); būdvardžių ir įvardžių homoformos (pvz., *juodu* – sutampa įvardžio *jis* dviskaitos vardininko, galininko ir būdvardžio *juodas* vnsk. kilmininko formos).

4.3.2. Homoformos su iš dalies sutampančia leksine reikšme

Tokios homoformos turi skirtingas lemas, tačiau jos semantiškai ir darybiškai susijusios, pvz.: *absoliutiškai* – sutamparieveiksmio *absoliutiškai* ir būdvardžio *absoliutiškas* formos; *ilgai* – sutamparieveiksmio *ilgai* ir būdvardžio *ilgas* formos. Tai dažniausiairieveiksmiai, sudaryti iš būdvardžių, t. y. tos pačios šaknies žodžiai, ir tik iš dalies juos galima laikyti skirtingos reikšmės žodžiais (jei žiūrima ne formaliai morfologiškai, o semantiškai).

Dalinės leksinės homoformos žodyne sudaro 8 kaitomųjų kalbos dalių homoformų rūšis, konkrečiau: *substantiva mobilia* daiktavardžių, pavyzdžiui, *administratorių* (sutampa žodžių *administratorius* bei *administratore* dgsk. kilmininko formos), *bėglių* (sutampa žodžių *bėglys* ir *bėglė* dgsk. kilmininko formos) homoformos; daiktavardžių ir veiksmažodžių homoformos (pvz.: *apkaustai* – sutampa veiksmažodžio *apkaustyti* esamojo laiko antrojo asmens ir daiktavardžio *apkaustai* vardininko formos; *atsakai* – sutampa daiktavardžio *atsakas* dgsk. vardininko ir veiksmažodžio *atsakyti* esamojo laiko antrojo asmens formos); daiktavardžių ir būdvardžių homoformos (pvz., *absolutus* – sutampa daiktavardžio *absolutas* dgsk. vardininko ir būdvardžio *absolutus* vnsk. vardininko formos; *abstraktu* – sutampa daiktavardžio *abstraktas* vnsk. įnagininko ir būdvardžio *abstraktus*, *-i* bevardės giminės formos); veiksmažodžių ir būdvardžių homoformos (pvz.: *magnetinę* – sutampa dalyvio *magnetinės* būtojo kartinio laiko dgsk. vardininko ir būdvardžio *magnetinis*, *-ė* mot. g. vnsk. galininko formos; *masinę* – sutampa dalyvio *masinės* vyr. g.

dgsk. galininko ir būdvardžio *masinis*, *-i* mot. g. vnsk. galininko formos); daiktavardžių, būdvardžių ir veiksmažodžių homoformos (pvz., *kaltą* – sutampa būdvardžio *kaltas* ir dalyvio *káltas* vyr. ir mot. g. vnsk. galininko ir daiktavardžio *káltas* vnsk. galininko formos, tiesa, pastaruoju atveju susipina dvi skirtingos priežastys. Tarp dalyvio ir daiktavardžio išties esama dalinio leksinio sutapimo, kadangi jų reikšmė susijusi, o štai su būdvardžiu *káltas* dalyvis ir daiktavardis *káltas* sutampa dėl prozodinių priežasčių); būdvardžių ir įvardžių homoformos (pvz., *pirmiausia* – sutampa būdvardžio *pirmiausias*, *-ia* bevardės giminės ir mot. g. vnsk. vardininko, įnagininko, šauksmininko ir skaitvardžio *pirmas* formos); daiktavardžių ir skaitvardžių ar įvardžių homoformos (pvz.: *aštuonetą* (sutampa daiktavardis ir skaitvardis), *maniškė* (sutampa daiktavardis ir įvardis)); daiktavardžių, veiksmažodžių ir įvardžių ar būdvardžių, skaitvardžių homoformos (pvz., *niūri* – sutampa daiktavardžio *niūris* vnsk. šauksmininko, būdvardžio *niūrus*, *-i* mot. g. vnsk. vardininko ir veiksmažodžių *niūrėti* bei *niurti* antrojo ir trečiojo asmens esamojo laiko formos).

Dalinio leksinio daugiareikšmiškumo pavyzdžių galima rasti ir tarp kaitomųjų bei nekaitomųjų kalbos dalių sutapimų – veiksmažodžių, būdvardžių irrieveiksmių homoformų (pvz., *aklinai* – sutampa veiksmažodžio *aklinti* būtojo kartinio laiko vnsk. antrojo asmens,rieveiksmio *aklinai* ir būdvardžio *aklinas*, *-a* mot. g. vnsk. naudininko formos), būdvardžių irrieveiksmių homoformų (*abejingai*, *aistringai*, *cikliškai*); daiktavardžių irrieveiksmių homoformų (pvz.: *dovanai*, *laukan*, *pusbalsiu*); daiktavardžių,rieveiksmių ir būdvardžių ar įvardžių homoformų (pvz., *bailiai* – sutampa daiktavardžio *bailys* dgsk. vardininko, būdvardžio *bailus*, *-i* mot. g. vnsk. naudininko irrieveiksmio *bailiai* formos, nors skiriasi daiktavardžio *bailiai* ir būdvardžio beirieveiksmio *baīliai* kirčiavimas. Jei dėmesio į kirčiavimą nekreiptume, visas tris formas sietų iš dalies sutampanti leksinė reikšmė.

Dar paminėtinos daiktavardžių, veiksmažodžių irrieveiksmių arrieveiksmių homoformos (pvz., *uždarai* – sutampa iš veiksmažodžio *uždaryti* padarytos dalyvio formos *uždaras*, *-a*, vnsk. naudininko, daiktavardžio *uždaras* dgsk. vardininko ir būdvardžio *uždaras*, *-a* mot. g. vnsk. naudininko irrieveiksmio *uždarai* formos; vėlgi skiriasi kai kurių žodžių kirčiai, plg. *ūždarai* ir *uždarai*), veiksmažodžių, daiktavardžių irrieveiksmių homoformos (pvz., *apsiaustai* – sutampa daiktavardžio *apsiaustas* dgsk. vardininko, dalyvio *apsiaustas*, *-a* mot. g. vnsk. naudininko irrieveiksmio *apsiaustai* formos); skaitvardžių irrieveiksmių homoformos (pvz.: *keturiomis*, *pirmiau*),

veiksmazodžių irrieveiksmių homoformos (pvz.: *apibrėžtai*, *apibendrintai* – sutampa dalyvių *apibrėžtas*, *-a*, *apibendrintas*, *-a* mot. g. vnsk. naudininko formos surieveiksmiais *apibrėžtai*, *apibendrintai*), daiktavardžių, įvardžių irrieveiksmių homoformos (pvz., *nieko* – sutampa daiktavardžio ir įvardžio *niekas* formos surieveiksmiu *nieko*), būdvardžių, skaitvardžių irrieveiksmių homoformos (pvz., *pirmiausiai* – šis žodis, priklausomai nuo konteksto, gali būti tiek būdvardis, tiekrieveiksmis, tiek skaitvardis).

4.4. Homoformos, nulemtos lietuvių kalbos morfologinės sistemos

Kai kurias homoformas lemia lietuvių kalbos morfologinė sistema. Vardažodžių sutapimus lemia sinkretiški tam tikrų linksniuotės linksniai, sutampančios kai kurių linksnių skirtingų giminių ar skaičiaus formos (pvz., *naujų* – tiek mot., tiek vyr. g. dgsk. kilmininko forma); gali sutapti veiksmazodžių asmuo, skaičius, laikas ir t. t. Tokių žodžių leksinė reikšmė yra tokia pat.

Šios kategorijos visumą sudaro vien kaitomųjų kalbos dalių sutapimai. Vardažodžių sutapimų daugiausia, juos galima skirti į skaičiaus, giminės ir linksnio sinkretizmo atvejus.

Linksnių sutapimų atvejai yra 7 rūšių: (*i*)o linksniuotės vardazodžių vnsk. vardininko, įnagininko, šauksmininko ir bevardės giminės sutapimai (pvz.: daiktavardžiai *abatija*, *data*, *emulsija*; būdvardžiai *faktiška*, *fatališka*, *galinga*); *ė* linksniuotės daiktavardžių vnsk. įnagininko ir šauksmininko sutapimai (pvz., daiktavardžiai *beteise*, *dėde*, *bedante*); (*i*)a, (*i*)u linksniuotės daiktavardžių dgsk. vardininko ir šauksmininko sutapimai (pvz., daiktavardžiai *abiturientai*, *akiniai*, *angelai*); (*i*)a linksniuotės daiktavardžių vnsk. vietininko ir šauksmininko sutapimai (pvz.: *albumė*, *bate*, *čempionė*); skaitvardžių ir įvardžių vardininko ir galininko sutapimai (pvz., įvardžiai *judu*, *mužu*, skaitvardžiai *abi*, *dvi* (verta paminėti, kad *Bendrinės lietuvių kalbos žodyne* (BLKŽ) žodis *abi* laikomas įvardžiu); *i* linksniuotės daiktavardžių ar sudaiktavardėjusių dalyvių, būdvardžių dgsk. vardininko ir šauksmininko sutapimai (pvz.: *akmenys*, *antys*, *baltieji*, *kaltinamieji*); skaitvardžių ar įvardžių vardininko, galininko ir įnagininko sutapimai (pvz., skaitvardžiai *aštuoniolika*, *dvylika*, *trylika* ir įvardis *keliolika*).

Giminės sutapimai žodyne yra 7 rūšių: vardazodžių vyr. ir mot. g. dgsk. kilmininko sutapimai (pvz.: dalyviai *abejojančių*, *akcentuojančių*, *apibendrinančių*, būdvardžiai *apynaujų*, *birių*, *blogų* – sutampa vyr. ir mot. g.); (*i*)a

linksniuotės vardažodžių vyr. g. vnsk. vardininko ir (i)o linksniuotės vardažodžių mot. g. dgsk. galininko sutapimai (pvz.: būdvardžiai *abejingas, brokuotas, dieviškas*; dalyviai *diferencijuotas, forsutas, gautas*); (i)o ir (i)a linksniuotės vardažodžių vnsk. galininko sutapimai (pvz.: būdvardžiai *abejingq, aklq, darganq*; dalyviai *dažomq, draudžiamq, gaunamq*); veikiamosios rūšies dalyvių vyr. g. dgsk. vardininko ir bevardės giminės sutapimai (pvz.: *abejojq, amžinq, apdarq*); skaitvardžių vardininko mot. ir vyr. g. sutapimai (pvz., *trys*); skaitvardžių galininko mot. ir vyr. g. sutapimai (pvz., *tris*); skaitvardžių ar įvardžių naudininko ir (ar) įnagininko vyr. ir mot. g. sutapimai (pvz.: *dviem, judviem*); daiktavardžių mot. ir vyr. g. sutapimai su bendrąja gimine (pvz.: *giminė, mazgotė* – bendroji giminė sutampa su moteriškąja; *ligonis* – bendroji giminė sutampa su vyriškąja);

Kaip minėta, gali sutapti ir kaitomųjų kalbos dalių skaičius. Tokių atvejų žodyne yra 4 rūšių: (i)o, *ė* linksniuotės vardažodžių vnsk. kilmininko ir dgsk. vardininko ir šauksmininko sutapimai (pvz.: daiktavardžiai *abejonės, abėcėlės, baterijos*; būdvardžiai *basos, birios, dabartinės*; dalyviai *decentralizuotos, degeneruotos, eiliuotos*); (i)u ir *i* linksniuotės vardažodžių vnsk. vardininko ir dgsk. galininko sutapimai (pvz.: *administratorius, gentis, kalendariorius*); (i)u linksniuotės daiktavardžių vnsk. galininko ir dgsk. kilmininko sutapimai (pvz.: *agresorių, derlių, išminčių*); (i) linksniuotės vyr. g. būdvardžių vnsk. šauksmininko ir dgsk. vardininko sutapimai (pvz.: *artimesni, garsesni, mažesni*).

Neretai sutampa ir įvairios veiksmožodžių formos. Tokių sutapimų žodyne – 7 rūšys. Tai: veiksmožodžių bendraties ir neveikiamosios rūšies dalyvių būtojo laiko vyr. g. dgsk. vardininko homoformos (pvz.: *absorbuoti, išardyti, nulakuoti*); tariamosios nuosakos trečiojo asmens veiksmožodžių ir neveikiamosios rūšies būtojo laiko dalyvių dgsk. kilmininko homoformos (pvz.: *adaptuotų, gamintų, girdėtų*); esamojo laiko veiksmožodžių trečiojo ir antrojo asmenų homoformos (pvz.: *atsispindi, įsidėmi, krūpteli*); dalyvių sutapimai su padalyviais ar pusdalyviais (pvz.: *adoruojantis, bijantis* – šiuo atveju sangražiniai padalyviai sutampa su nesangražinio esamojo laiko vyr. g. vnsk. vardininko dalyvio forma; pusdalyvis *duodamas* sutampa su esamojo laiko dalyviais – tiek su mot. g. dgsk. galininku, tiek su vyr. g. vnsk. vardininku); bendraties sutapimai su veiksmožodžių vnsk. antruoju asmeniu (pvz.: *neklusti, nuslysti, trykšti*); veiksmožodžių asmenuojamosios ir neasmenuojamosios formų sutapimai (pvz.: *dirbtinai, imtinai, maldautinai* – mot. g. dalyvių vnsk. naudininko formos sutampa su būdiniais; *sugrįžta, tirta* – veiks-

mažodžių esamojo laiko trečiojo asmens formos sutampa su bevardės ir mot. g. dalyviais – tiek su vardininku, tiek su įnagininku); esamojo ir būsimojų laiko padalyvių homoformos (pvz.: *apklusiant, pratęsiant, tręšiant* – sutampa būsimojų ir esamojo laiko padalyvių formos).

4.5. Tos pačios reikšmės homoformos, atsiradusios dėl priskyrimo skirtingoms paradigmoms

Kai kurie daiktavardžiai, būdvardžiai, veiksmažodžiai, įvardžiai irrieveiksmiai, turintys tą pačią reikšmę, gali priklausyti dviem skirtingoms paradigmoms, pvz.: *Audrių* galėtų būti tiek formos *Audrius*, tiek formos *Audrys* vnsk. galininkas; *būčiau* galėtų priklausyti tiek *būti, yra, buvo* paradigmai, tiek *būti, būna, buvo*, tiek ir *būti, būva, buvo*; *griuvus* – galėtų priklausyti tiek *griūti, griūna, griuvo*, tiek *griūti, griūva, griuvo* paradigmoms; formos *stebukladario* lema galėtų būti tiek *stebukladaris*, tiek *stebukladarys*; *mano* – tiek įvardžio su lema *mano*, tiek įvardžio su lema *manas* forma. Tokiems atvejams LKHŽ skirta viena atskira rūšis.

4.6. Priežasčių sąsajos

Paminėtos penkios atsiradimo priežastys gali būti susijusios tarpusavyje, o be to, kai kurių žodžių homoformų atsiradimą gali lemti ne viena priežastis, pavyzdžiui, *abstrākčiai* vs. *abstrakčiai* – čia daugiareikšmiškumą lemia tiek prozodinių elementų skirtumai (skiriasi kirčio vieta), tiek dalinis reikšmės sutapimas (vienas žodis yra padarytas iš kito). Kitas pavyzdys – būdvardžio *kaļtas* ir daiktavardžio *kāļtas* sutapimai. Šie žodžiai galėtų patekti tiek į prozodinių elementų, tiek į skirtingos leksinės reikšmės nulemtų homoformų kategorijas.

Be to, vieną rūšį žodyne sudarančios skirtingos homoformos nebūtinai lemiamos tų pačių priežasčių. Pavyzdžiui, vieni tariamosios nuosakos ir neveikiamosios rūšies dalyvių sutapimai (pvz.: *adaptuotų, brandintų*) susidaro dėl prozodinių elementų skirtumų (plg., *aplýtų* ir *aplytų*), o kiti – dėl lietuvių kalbos morfologinės sistemos, t. y. tiesiog sutampa dvi veiksmažodžio formos (pvz.: *automatizuotų, derėtų*).

Penkios homoformų susidarymo priežastys padeda dar geriau suvokti, kodėl anotuojant *Lemuokliu* (ar kitais morfologiniais anotatoriais) atsiranda vienokių ar kitokių morfologinio daugiareikšmiškumo atvejų. Papildžius

šià informacijà statistiniais duomenimis, bûtų galima įvardyti, kokia priežastis tikėtiniausia. Taip bûtų galima patobulinti ir LKHŽ. Suvokiant, kas lemia morfologinį daugiareikšmiškumą, į žodyną bûtų galima įtraukti naujų duomenų, pavyzdžiui, galėtų atsirasti pažymos, kurios rodytų daugiareikšmiškumo tikimybę ar kt. Be to, analizuojant išsamiau, pateikiant statistinius duomenis, galėtų paaiškėti ir daugiau homoformų atsiradimo priežasčių. Tai – galimas ateities tyrimas.

APIBENDRINIMAS

Straipsnyje aprašytas vienas iš lietuvių kalbos kompiuterizavimo etapų – automatinė morfologinė analizė. Jos metu išryškėja morfologinio daugiareikšmiškumo problema.

Remiantis rengiamu LKHŽ, atsirado galimybė išryškinti priežastis, kodėl lietuvių kalboje atsiranda homoformų. Pagal atsiradimo priežastis homoformas galima klasifikuoti taip: 1) teorinės homoformos, atsiradusios dėl kai kurių žodynų sudarymo specifikos; 2) homoformos, atsiradusios dėl prozodinių elementų skirtumų; 3) homoformos, atsiradusios dėl sutampančių skirtingų žodžių; 4) homoformos, nulemtos lietuvių kalbos morfologinės sistemos; 5) tos pačios reikšmės homoformos, atsiradusios dėl priskyrimo skirtingoms paradigmoms.

Teorinės homoformos apsunkina morfologinę analizę tada, kai archajiški ar dabartinėje lietuvių kalboje (beveik) nebevartojami žodžiai sutampa su dažniais žodžiais. Ši problema parodo, kad morfologiniams anotatoriams reikalingi kitaip sudaryti žodžių sąrašai.

Homoformos, nulemtos prozodinių elementų, atsiranda tik rašytinėje kalboje. Jų pavyzdžių daugiausia tarp kaitomųjų kalbos dalių sutapimų, dažnai tai būna sutapusios vieno žodžio formos (pvz.: vnsk. vardininkas *mamà* ir šauksmininkas *māma*; vnsk. vardininkas *apimtìs* ir dgsk. galininkas *āpimtis*), tačiau gali sutapti ir dvi skirtingos kalbos dalys (pvz., *apsaugos* – sutampa daiktavardžio *apsauga* vnsk. kilmininko, dgsk. vardininko ir veiksmožodžio *apsaugoti* trečiojo asmens būsimąjo laiko formos). Tokių sutapimų taip pat pastebėta tarp kaitomųjų ir nekaitomųjų kalbos dalių homoformų, pavyzdžiui, veiksmožodžių, būdvardžių irrieveiksmių sutapimų (pvz., *blogai* – sutampa veiksmožodžio *blogti* būtojo kartinio laiko antrojo asmens vnsk., būdvardžio *blogas*, *-a* mot. g. vnsk. naudininko irrieveiksmio *blogai* formos).

Homoformos, atsiradusios dėl sutampančių skirtingų žodžių, gali būti skirtingų leksinių reiškinių (sutampantys žodžiai, kurie darybiškai nėra susiję, pvz., *arti* – sutamparieveiksmis, prielinksnis, veiksmazodžio bendratis ir dalyvis, plg. *arti* ir *árti*) arba iš dalies sutampančių reiškinių (turi skirtingas lemas, tačiau jos semantiškai ir darybiškai susijusios, pvz.: *absoliutiškai* – sutamparieveiksmio *absoliutiškai* bei būdvardžio *absoliutiškas*, –a mot. g. vnsk. naudininko formos).

Homoformas, atsiradusias dėl lietuvių kalbos morfologinės sistemos, lemia lietuvių kalbos gramatika, pavyzdžiui, sinkretiniai linksniai, sutampančios skirtingų giminių ar skaičiaus formos (pvz., *naujų* – tiek mot., tiek vyr. g. forma); gali sutapti veiksmazodžių asmuo, skaičius, laikas ir t. t. Tokių žodžių leksinė reikšmė yra tokia pat, kitaip sakant, sutampa skirtingos to paties žodžio kaitybinės formos. Šios kategorijos visumą sudaro vien kaitomųjų kalbos dalių sutapimai.

Homoformų atsiradimą gali lemti ir tai, kad tas pats žodis priskiriamas skirtingoms paradigmoms, pavyzdžiui, *būčiau* – ši forma galėtų priklausyti tiek paradigmai *būti, yra, buvo*, tiek *būti, būna, buvo*, tiek ir *būti, būva, buvo* paradigmoms.

Šios priežastys padeda suvokti, kodėl automatiškai anotuojant atsiranda vienokių ar kitokių morfologinio daugiareikšmiškumo atvejų. Be to, remiantis šia informacija, būtų galima patobulinti LKHŽ. Suvokiant, kas lemia morfologinį daugiareikšmiškumą, į žodyną būtų galima įtraukti papildomą informaciją, pavyzdžiui, galėtų atsirasti pažymų, kurios rodytų daugiareikšmiškumo tikimybę ar kitus duomenis.

LITERATŪRA

- Aleksandravičiūtė Skaistė 2008: Paieškos sistemos „Google“ naudojimo galimybės tiriant kalbos reiškinius. – *Bendrinė kalba* 81, 266–282.
- Barauskaitė Janina 1979: *Leksikologijos pradmenys. Leksinė semantika*, Vilnius: Lietuvos TSR aukštojo ir specialiojo vidurinio mokslo ministerija.
- Barauskaitė Jūratė, Čepaitienė Giedrė, Mikulėnienė Danuolė 1995: *Lietuvių kalba. Leksikologija, fonetika, akcentologija, dialektologija, rašyba* 1, Vilnius, Mokslo ir enciklopedijų leidykla.
- Bielinskienė Agnė, Boizou Loïc, Rimkutė Erika 2017: Lietuvių kalbos morfologiškai ir sintaksiškai anotuoti tekstynai. – *Bendrinė kalba* 90, 1–30.

- BLKŽ 2022: *Bendrinės lietuvių kalbos žodynas* (rengiamas), vyr. red. D. Liutkevičienė, Vilnius: Lietuvių kalbos institutas. Prieiga internete: <https://ekalba.lki.lt/bendrinės-lietuvių-kalbos-zodynas>.
- DLKŽ 2021: *Dabartinės lietuvių kalbos žodynas*, vyriausiasis redaktorius S. Keinys, 8-as patais. ir papild. leidimas (elektroninis variantas), Vilnius: Lietuvių kalbos institutas. Prieiga internete: <https://ekalba.lt>.
- Gaivenis Kazimieras, Keinys Stasys 1990: *Kalbotyros terminų žodynas*, Kaunas: Šviesa.
- Grumadienė Laima 2002: Dabartinės rašomosios lietuvių kalbos dažninis žodynas ir jo bazė. – *Litanistika* 46, 19–37.
- Jakaitienė Evalda 2009: *Leksikologija*, Vilnius: Vilniaus universiteto leidykla.
- Kaljanov Ali 2023: Grammatical Homonyms in Modern Karakalpak Language. – *The International Journal of Language and Cultural* 5(1), 26–31. Prieiga internete: <https://www.growingscholar.org/journal/index.php/TIJOLAC/article/view/322>.
- Katamba Francis 1994: *English Words*, London: Routledge.
- LKE 1999: *Lietuvių kalbos enciklopedija*, Vilnius: Mokslo ir enciklopedijų leidybos institutas.
- LKŽ 2005: *Lietuvių kalbos žodynas*, t. 1–20 (elektroninis variantas), vyr. red. G. Naktinienė, Vilnius: Lietuvių kalbos institutas, 2005. Prieiga internete: <https://ekalba.lt/lietuvių-kalbos-zodynas>.
- Marcinkevičienė Rūta 2011: *Žodžio reikšmė. Žodynai ir tekstynai*, Kaunas: Vytauto Didžiojo universitetas.
- Nevzorova Olga, Zin'kina Julia, Pjatkin Nicolaj 2005: Applied Problems of Functional Homonymy Resolution for Russian Language. – *International Journal ITA* 12(3), 231–235. Prieiga internete: <http://www.foibg.com/ijita/vol12/ijita12-3-p05.pdf>.
- Pakerys Jurgis 2009: Būtinųjų predikatyvų žymėjimas lietuvių kalbos veiksmo pavadinimo konstrukcijose. – *Gramatinių funkcijų prigimtis ir raiška* (Acta Salensia 1), Vilnius: Vilniaus universitetas, 125–145.
- Piečytė Miglė 2023: *Morfologinio daugiareikšmiškumo analizė ir klasifikacija*: baigiamasis magistro darbas, Kaunas: Vytauto Didžiojo universitetas.
- Puolakainen Tiina 2012: How Does the Choise of Morphological Analyser Influence the Quality of Syntactical Analysis? – *Human Language Technologies – The Baltic Perspective: Proceedings of the Fifth International Conference Baltic HLT*, Amsterdam, Berlin, Tokyo, Washington, DC: IOS Press, 193–200.
- Rimkutė Erika 2002: Homoformos dabartinės lietuvių kalbos tekstyne. – *Lituanistika* 2, 86–101.

- Rimkutė Erika 2003a: Morfologinio daugiareikšmiškumo tipologija. – *Lituanistika* 4, 60–78.
- Rimkutė Erika 2003b: Homonimai, homoformos ir homografaai tekstynų lingvistikos požiūriu. – *Terminologija* 10, 33–41.
- Rimkutė Erika 2006a: *Morfologinio daugiareikšmiškumo ribojimas kompiuteriniame tekстыne*: daktaro disertacija, Kaunas: Vytauto Didžiojo universitetas.
- Rimkutė Erika 2006b: Dabartinės lietuvių kalbos gramatinių formų vartoseną morfologiškai anotuotame tekстыne. – *Lituanistika* 2, 34–55.
- Rimkutė Erika, Grybinaitė Rasa 2004: Dažniausios lietuvių kalbos morfoliginio daugiareikšmiškumo rūšys ir jų automatinis vienareikšminimas. – *Kalbų studijos* 5, 74–78. Prieiga internete: <https://www.vdu.lt/cris/handle/20.500.12259/53446>.
- Rimkutė Erika, Grigonytė Gintarė 2006: Automatizuotas lietuvių kalbos morfoliginio daugiareikšmiškumo ribojimas. – *Lituanistika* 9, 30–37.
- Rimkutė Erika, Daudaravičius Vidas 2007: Morfoliginis „Dabartinės lietuvių kalbos tekstyno“ anotavimas. – *Kalbų studijos* 11, 30–35.
- Šukys Jonas 2004: *Panašiai skambantys žodžiai: normos ir klaidos*, Vilnius: Gimtasis žodis.
- TŽŽ 2023: *Tarptautinių žodžių žodynas* (elektroninis variantas). Prieiga internete: <https://tzz.lt/>.
- Zinkevičius Vytautas 2000: Lemuoklis – morfologinei analizei. – *Darbai ir dienos* 24, 246–273.

Įteikta 2023 04 05

Priimta 2023 06 29

MIGLĖ ŽEMRIETĖ

THE REASONS FOR THE MORPHOLOGICALLY AMBIGUOUS FORMS

S u m m a r y

Although the problem of morphological ambiguity in the Lithuanian language has been solved for all practical purposes morphological annotators, even when are improved, are not always capable of choosing the most probable form from several morphologically polysemous forms. Moreover, such cases are not uncommon when the morphological annotator incorrectly chooses one form from several possible ones. Therefore, a decision was made to conduct a study that would analyze and classify cases of morphological polysemy, hereinafter referred to as homoforms of the Lithuanian language. The aim of the study is to discuss their

possible causes based on the classified homoforms through the application of quantitative analysis, textual linguistics, and descriptive.

When describing the reasons for homoforms, reference is made to the *Dictionary of Homoforms* of the Lithuanian language, prepared by the author of this article and Erika Rimkutė.

Since the *Dictionary of Homoforms of the Lithuanian Language* consists of a really large number (about 36,000) of morphologically polysemous words, it becomes possible to define the reasons for the appearance of homoforms in the Lithuanian language. Homoforms can be classified according to the reasons for their appearance as follows: (1) homoforms that are rather theoretical in nature and appeared in the process of the compilation of dictionaries; (2) homoforms resulting from differences in prosodic elements; (3) homoforms resulting from different overlapping words; (4) homoforms that came into existence due to the specifics of the Lithuanian morphology system; (5) homoforms of the same meaning that appeared under different paradigms.

These five reasons help us to better understand why automatic annotation often produces a kind of morphological ambiguity. Based on this information, it would be possible to improve the *Dictionary of Homoforms of the Lithuanian Language* on its own right. By understanding what determines morphological polysemy, new data could be added to the dictionary, such as notations that indicate the probability of ambiguity or other data.

KEYWORDS: morphological ambiguity, automatic morphological analysis, corpus, homoforms, dictionary.

MIGLĖ ŽEMRIETĖ

Lietuvių kalbos institutas

Petro Vileišio g. 5, LT-10308 Vilnius

piecyte@gmail.com

MIGLĖ ŽEMRIETĖ

MORFOLOGIŠKAI DAUGIAREIKŠMIŲ
FORMŲ ATsirADIMO PRIEŽASTYS

S a n t r a u k a

Nors morfologinis daugiareikšmiškumas yra iš esmės išspręsta problema, net ir patobulinus lietuvių kalbos morfologinius anotatorius, jiems ne visada iš kelių morfologiškai daugiareikšmių formų pavyksta pasirinkti tikėtiniausią. Be to, nereti ir tokie atvejai, kai morfologinis anotatorius neteisingai parenka vieną formą

iš kelių galimų. Todėl nutarta išanalizuoti ir suklasifikuoti morfologinio daugiareikšmiškumo atvejus, juos vadinant lietuvių kalbos homoformomis. Tyrimo tikslas – remiantis suklasifikuotomis homoformomis, aptarti galimas jų priežastis. Taikyti kiekybinės analizės, tekstynų lingvistikos, aprašomasis metodai.

Aprašant homoformų priežastis, remiamasi *Lietuvių kalbos homoformų žodynu*. Jį sudaro apie 36 tūkst. homoformų, todėl galima apibrėžti homoformų lietuvių kalboje atsiradimo priežastis. Šiame straipsnyje homoformos pagal atsiradimo priežastis suklasifikuotos į penkias grupes: 1) teorinės homoformos, atsiradusios dėl kai kurių žodynų sudarymo specifikos; 2) homoformos, atsiradusios dėl prozodinių elementų skirtumų; 3) homoformos, atsiradusios dėl sutampančių skirtingų žodžių; 4) homoformos, nulemtos lietuvių kalbos morfologijos sistemos; 5) tos pačios reikšmės homoformos, atsiradusios dėl priskyrimo skirtingoms paradigmoms.

Ši klasifikacija padeda geriau suvokti, kodėl morfologiniams anotatoriams anotuojant atsiranda vienokių ar kitokių morfologinio daugiareikšmiškumo atvejų. Remiantis šia informacija, būtų galima patobulinti ir patį *Lietuvių kalbos homoformų žodyną*. Suvokiant, kas lemia morfologinį daugiareikšmiškumą, į žodyną būtų galima įtraukti naujų duomenų, pavyzdžiui, galėtų atsirasti pažymos, kurios rodytų daugiareikšmiškumo tikimybę ar kitus duomenis.

ESMINIAI ŽODŽIAI: morfologinis daugiareikšmiškumas, automatinė morfologinė analizė, tekstynas, homoformos, žodynas.

MIGLĖ ŽEMRIETĖ
Lietuvių kalbos institutas
Petro Vileišio g. 5, LT-10308 Vilnius
piecyte@gmail.com