

**Erika RIMKUTĖ**

*Vytauto Didžiojo universitetas*

## **HOMONIMAL, HOMOFORMOS IR HOMOGRAFAI TEKSTYNŲ LINGVISTIKOS POŽIŪRIU**

### **1. Įvadas**

Vytauto Didžiojo universiteto Kompiuterinės lingvistikos centre yra sukauptas ir toliau pildomas *Dabartinės lietuvių kalbos tekstynas*, kurį sudaro daugiau nei 100 milijonų žodžių (adresas: <http://donelaitis.vdu.lt/Tekstynas>). Taip pat sudarinėjamas ir tvarkomas 1 milijono žodžių morfologiškai anotuotas tekstynėlis. Tam naudojama V. Zinkevičiaus sukurta kompiuterinė programa *Lemuoklis* (Zinkevičius 2000: 246), galinti pateikti antraštinį rašytinės žodžio formos pavida-lą, vadinamąją lemą, taip pat nurodyti morfologines pažymas, pavyzdžiui, forma *laužo* sulemuojama kaip daiktavardis *laužas* arba kaip veiksmažodis *laužyti*, apibūdinama kaip vyriškosios giminės daiktavardžio vienaskaitos kilmininkas arba tiesioginės nuosakos esamojo laiko veiksmažodžio trečiasis asmuo. 47 proc. visų žodžių ar žodžių formų yra morfologiškai daugiareikšmiai. Morfologiškai daugiareikšmiais vadinami tie žodžiai ar žodžių formos, kuriems *Lemuoklis* automatiškai nustato dvi ar daugiau lemų arba kuriems pateikia dvi ar daugiau galimų morfologinių pažymų, pvz.: *namo* – daiktavardis irrieveiksmis (*n ā m o stogas – einu n a m ō*); *galimas* – būdvardis ir dalyvis (*g a l i m a s daiktas*); *laimės* – daiktavardis ir veiksmažodis (*l á i - m ė s kūdikis – jis l a i m ė s varžybas*) ir pan.

Automatiškai sulemuoti ir morfologiškai pažymėti tekstai dar peržiūrimi kalbininko, kuris panaikina daugiareikšmius atvejus. Tokios daugiareikšmės formos buvo analizuotos išsamiau. Tai nauja tyrimų sritis, todėl reikia pasirinkti ir apsibrėžti terminus. Minėtos formos kalbotyroje įvardijamos įvairiai: *gramatiniai / morfologiniai homonimai, homoformos, homografai, daliniai homonimai*. Šiuo straipsniu norima atskleisti esamų terminų vartojimo nenuoseklumus, pagrįsti kitokią termino *homoformos* ir naujo termino *morfologinis daugiareikšmiškumas* vartoseną.

## 2. Homonimijos sampratos

Iki šiol lietuvių kalbotyroje morfologiniam daugiareikšmiškumui nebuvo skirta daug dėmesio – išsamiausiai kalbotyroje nagrinėti tik leksiniai homonimai. Todėl ir kyla problemų, kaip įvardyti morfologinį daugiareikšmiškumą ir atskirus jo tipus.

**2.1. Homoformos ir morfologiniai homonimai.** Kalbotyroje morfologiniai / gramatiniai homonimai apibūdinami kaip fonetiškai sutampančios to paties žodžio formos, turinčios skirtingas gramatines reikšmes, pvz.: *ranka* (vns. vard.) ir *ranka* (vns. įnag.) (Jakaitienė 1980: 42; Соболева 1980: 78). Vienų kalbininkų gramatiniai homonimai neskiriami nuo homoformų. Tai gali būti to paties žodžio skirtingos formos (žr. anksčiau minėtus pavyzdžius) arba skirtingų žodžių panašiai skambančios formos, pvz.: *sakai* (dkt. vard.) ir *sakai* (vksm. esam. I vns. II asm.) (LK I 1995: 23; Barauskaitė 1979: 43–44; Drotvinas 1986: 125). F. Katamba sutampančias tų pačių ar skirtingų kalbos dalių formas vadina sinkretizmu (Katamba 1994: 25).

E. Jakaitienė homoformų netapatina su morfologiniais homonimais (Jakaitienė 1980: 42). Homoformas apibrėžia kaip atskiras, atsitiktinai sutampančias bendrinėje kalboje savo fonetine struktūra visiškai skirtingų žodžių formas, pvz.: *kiškis* (dkt. vns. vard.) ir *kiškis* (sngr. vksm. liep. nuos. vns. II asm.). Homoformomis laikytini ir tokie žodžiai, kurių fonetiškai sutampa ne viena, o kelios formos. Pavyzdžiui: a) veiksmazodžiai, kurių fonetiškai sutampa tik bendratis ir iš jos padarytos formos: *gesti*, *gęsta*, *geso* „ugniai, šviesai nykti, blėsti“ ir *gesti*, *genda*, *gedo* „darytis netinkamam“; b) daiktavardžiai ir būdvardžiai ar dalyviai (skiriasi jų vienaskaitos naudininko ir vietininko, daugiskaitos vardininko ir naudininko formos): *bendras* „priklausomas visiems“ ir *bendras* „bičiulis, draugas“ (Jakaitienė 1980: 43). LK I tokie atvejai priskirti prie dalinių homonimų, tiesa, ten minimi tik veiksmazodžiai (LK I 1995: 24).

LK I prie homoformų aptarti fonetiškai sutampantys linksniai (LK I 1995: 23), bet kyla klausimas, kuo laikytini sutampantys linksniai, kurie skiriasi tarimu, pvz.: *n a u j ō s parduotuvės atidarymas – atidarytos dvi n a ū j o s parduotuvės*. V. Drotvinas tų pačių kalbos dalių ir vienodai, ir skirtingai tariamas formas vadina *reguliariosiomis homo-*

*formomis* (Drotvinas 1986: 93). Kiti kalbininkai neužsimena apie to paties žodžio skirtingai tariamas formas.

**2.2. Homoformos ir daliniai homonimai.** E. Jakaitienė dalinius homonimus apibrėžia kaip žodžius, iš kurių vieno žodžio visos formos sutampa su dalimi kito žodžio formų. Dažniausiai vienas tokios homonimų poros žodis priklauso prie kaitomųjų kalbos dalių, o kitas – nekaitomųjų, pvz.: *greta* – daiktavardis irrieveiksmis, *deja* – daiktavardis ir jaustukas (Jakaitienė 1980: 43).

Iš LK I pateikto dalinių homonimų apibrėžimo (LK I 1995: 24) galima spręsti, kad ši homonimijos rūšis sutampa su anksčiau jau E. Jakaitienės minėtomis homoformomis (Jakaitienė 1980: 43), pvz.: *rausti, rausia, rausė* „kasti, knisti“ ir *rausti, rausta, raudo* „darytis raudonam“.

KTŽ homoformomis laikomos vienodai tariamos ir rašomos skirtingų žodžių formos (KTŽ 1990: 79). Vadinasi, prie homoformų nepriskiriami keli sutampantys to paties žodžio linksniai. Šitaip apibrėžus homoformas, skirtumas tarp homoformų ir dalinių homonimų pasidaro neaiškus, nes pastarieji apibūdinti kaip „homonimai, kurių ne visos morfemos visiškai sutampa ar kurie priklauso skirtingoms kalbos dalims“ (KTŽ 1990: 80). Ši homonimų tipą iliustruoja pavyzdys *greta* – daiktavardis irrieveiksmis, o prie homoformų pateiktas pavyzdys *varu* –rieveiksmis ir daiktavardžio *varas* įnagininkas (KTŽ 1990: 79–80), todėl nematyti skirtumo tarp dalinių homonimų ir homoformų.

**2.3. Homoformos ir homografai.** Daugiausia neaiškumų šio straipsnio autorei kilo analizuojant tų pačių ar skirtingų žodžių formas, kurios skiriasi prozodiniais elementais, t. y. kirčio vieta, priegaide ar balsio kiekybe, – kaip jas pavadinti: *homografais* ar *homoformomis*? Homografai dažniausiai apibrėžiami kaip skirtingi žodžiai ar jų formos, vienodai rašomi, bet tariami skirtingai (skiriasi arba priegaide, arba kirčio vieta), pvz.: *áušti* „vėsti, darytis nekarštam“ ir *aūšti* „švisti, brėkšti“ (Jakaitienė 1980: 44; LK I 1995: 23; KTŽ 1990: 23–24; Palmer 1981: 101; Drotvinas 1986: 91–92; Barauskaitė 1979: 44). Taigi atrodytų, kad prozodiniais elementais besiskiriantys nevienareikšmiai atvejai turėtų būti vadinami homografais, bet LK I rašoma, kad „yra ir homoformų, kurios vienodai rašomos, bet skiriasi kirčiu“ (LK I

1995: 23). Pateikiami tokie pavyzdžiai: „*báltų* ir *baltų* (*báltai* ir *báltas*), *dōvanos* ir *dovanōs* (*dovanà* ir *dovanóti*), *māno* ir *māno*“ (LK I 1995: 23). V. Drotvinas prie homoformų nurodo pavyzdžius *supánčios* (vksm. būs. I. III asm.) ir *sūpančios* (dlv. mot. gim. vns. kilm. arba dgs. vard.) (Drotvinas 1986: 91). Šie pavyzdžiai taip pat skiriasi kirčio vieta, nors priskiriami prie homoformų.

Kirčio vieta besiskiriančių homoformų samprata atrodo logiška, nes analizuojant morfologiškai anotuotus tekstus pastebėta nemažai formų, kurios skiriasi tarimu (apie 20 proc. homoformų skiriasi prozodiniais elementais) (žr. Rimkutė 2002: 91–92). Bet kyla klausimas, kokiais kriterijais remiantis galima skirti homoformas nuo homografų, jei homoformoms nebūtinai fonetinis tapatumas. Galbūt tarimu besiskiriančiomis homoformomis derėtų laikyti tik to paties žodžio formas, pvz.: *mamà* (vns. vard.) ir *māma* (vns. šauksm.), *galì* (veiksmazodžio *galėti* esam. I. vns. II asm.) ir *gāli* (veiksmazodžio *galėti* esam. I. III asm.), o homografais – skirtingos reikšmės žodžius, pvz.: *kártis* „pagalys“ ir *kařtis* „kartumas“ ar *kařtis* „arklio sprandas; vieta, kur auga karčiai“. Tačiau LK I ir V. Drotvino pateikti pavyzdžiai nėra skirtingos to paties žodžio prozodiniais elementais besiskiriančios formos, o skirtingų reikšmių žodžiai. Be to, ne visada aišku, ar tai to paties, ar skirtingų žodžių formos.

Taigi atrodytų, kad skirtingi žodžiai ar žodžių formos, kurie skiriasi prozodiniais elementais, gali būti laikomi homoformomis. Dar vienas kriterijus, paaiškinantis tokią homoformų sampratą, yra tai, kad homografai dažnai apibūdinami kaip skirtingos reikšmės žodžiai, bet tą pabrėžia ne visi kalbininkai, pavyzdžiui, LKE rašoma, kad homografai – tai vienodai rašomi žodžiai arba žodžių formos (LKE 1999: 248). Dažnai sutampa keli to paties vardažodžio linksniai. Savaiame suprantama, kad tai nėra skirtingos reikšmės žodžiai. Tad neaišku, prie kokios homonimijos dalies priskirtinos panašios formos.

O. S. Achmanova tuos atvejus, kai skirtingai tariamos to paties žodžio formos, vadina *homografinėmis žodžių formomis* ir pabrėžia, kad tai nėra homografai (Ахманова 1986: 436). V. Drotvinas homoformas, homografus ir homofonus siūlo vadinti *homoniminėmis žodžių formomis* (Drotvinas 1986: 91). Toks terminas atrodo priimtinas, todėl tas žodžių formas, kurios skiriasi prozodiniais elementais, būtų

galima vadinti *homografinėmis* ar *homoniminėmis* žodžių formomis, bet vietoj ilgo termino pasirinktas trumpesnis – tiesiog *homoformos*. Jau minėta, kad analizė iš dalies yra nulemta tiriamo objekto – automatiniu būdu morfologiškai anotuoto tekstinio, todėl į fonetinius homoformų skirtumus per daug nesigilinama, nes tam reikalingi išsamesni tyrimai. Šie kriterijai lėmė, kad šiek tiek pakeista termino *homoformos* vartosena.

Norint išvengti painiavos, analizuojant morfologiškai anotuotame tekстыne esančias morfologiškai daugiareikšmes formas, laikomasi nuostatos, kad skirtingoms ar toms pačioms kalbos dalims priskiriamos skirtingos kaitybinės formos, nesutampančios kirčio vieta, priegaide ar balsio kiekybe, gali būti ir homoformos, ir homografai, pvz.: *k a r ā l i a u s s o s t a s – k a r a l i a ū s i l g a i, n ā m o s t o g a s – e i n u n a m ō, s ē d i s ē n i a i – t a i a t s i t i k o s e n i a ĩ*. Taigi kai kurios homoformos – sutampantys vardažodžių linksniai gali būti homografai, jei žiūrima tarties aspektu, o homografai – homoformos ar skirtingos tų pačių žodžių linksnių formos nagrinėjant morfologiškai. Pasirinktas paprasčiausias variantas – visos morfologiškai daugiareikšmės formos, nepaisant to, ar skiriasi prozodiniais elementais, ar ne, laikomos homoformomis, o terminas *homografai* apskritai nevartojamas.

**2.4. Kiti homonimijos atvejai.** Kai kuriuose kalbiniuose darbuose ne visai aiški homoformų apimtis: ar homoformas sudaro tik tų pačių, ar ir skirtingų kalbos dalių žodžių formos. Remiantis V. Urbučio ir LKE pateiktais pavyzdžiais, galima spręsti, kad homoformomis laikomos kai kurios sutampančios, tik skirtingoms kalbos dalims priklausančių žodžių formos, pvz.: *grūdo* – daiktavardžio *grūdas* vns. kilm. ir *grūdo* – veiksmožodžio *grūsti* būt. kart. I. III asm. (Urbutis 1956: 4; LKE 1999: 247–248), todėl neaišku, ar homoformomis galima laikyti sutampančias tos pačios kalbos dalies žodžių formas, pavyzdžiui, *tapo* (skiriasi kitos pagrindinės formos: *tapyti, tapė* ir *tapti, tampa*).

Iš LKE pateikto homografų apibrėžimo neaišku, kuo homografai skiriasi nuo homonimų ir homoformų, nes čia rašoma: „homografai – įprastine rašyba vienodai rašomi kalbos elementai (žodžiai ar jų formos). Lietuvių kalbos homografai dažniausiai kartu yra ir homonimai“, pvz.: *kasa* „pynė“ – *kasa* „piniginių operacijų vieta“ (kitų kal-

bininkų supratimu, *kasa* yra tikrasis / absoliutusias homonimas; šis pavyzdys galėtų būti laikomas homografu, jei sutaptų daiktavardžio *kasa* ir veiksmažodžio *kasti* esam. l. III asm. formos). Daiktavardžio *varpas* vns. kilm. forma *varpo* ir veiksmažodžio *varpyti* esam. l. III asm. forma *varpo* taip pat laikomos homografais (kituose kalbiniuose darbuose – homoformomis). LKE dar rašoma, kad skirtingai tariami homografai taip pat nėra jokia retenybė (pvz.: *áušti* – *aūšti*, *péréjimas* – *perėjimas*, *girià* – *gìria*) (LKE 1999: 248).

Kalbininkų darbuose dar trūksta tokių atvejų aptarimo, kai sutampa dvi ar daugiau nekaitomų kalbos dalių, pvz.: *kaip*, *ir*, *tik*, *iki*; ar jos taip pat priskirtinos prie homoformų, ar turėtų būti laikomos koku nors kitu daugiareikšmiškumo atveju (šio straipsnio autorė tokius atvejus priskiria prie nekaitomų kalbos dalių homoformų (žr. 3. *Morfologinis daugiareikšmiškumas*).

### 3. Morfologinis daugiareikšmiškumas

Iš to, kas išdėstyta, aišku, kad skiriasi kalbiniuose darbuose pateikiama terminija. Kad nekiltų neaiškumų, kas vis dėlto analizuojama – homoformos, homografai ar daliniai homonimai, – pasirinktas bendresnis terminas *morfologinis daugiareikšmiškumas*. Tai reikškinys, apimantis: 1) kaitomas ir nekaitomas formas; 2) skirtingų ir tų pačių kalbos dalių žodžių formas; 3) tam tikromis formomis, prozodiniais elementais besiskiriančias ir visiškai sutampančias žodžių formas ar žodžius.

Terminas *daugiareikšmiškumas* vartotinas ne tik analizuojant morfologiniais požymiais besiskiriančias homoformas, bet ir kalbant apie darybinį (pvz., *ištvirkėlis* – *dobilėlis*: tas pats darybos formantas suteikia skirtingas reikšmes), leksinį semantinį (pvz., tikrasis homonimas *žiedas*), sintaksinį semantinį (pvz., junginyje *mokytojo pakvietimas* dėl skirtingų sintaksinių ryšių neaišku, ar kviečia mokytojas, ar jis pats yra kviečiamas), tekstinį (pvz., *jis čia išsėdėjo visą vakarą*: norint suprasti deiktikus *jis* ir *čia*, reikalingas platesnis kontekstas) daugiareikšmiškumą.

Būtina nurodyti, kad morfologinio daugiareikšmiškumo analizė yra gana specifinė – analizuojamos rašytinės, be kirčio ženklų žodžių

formas. Automatinę morfologinę analizę atliekanti programa *Lemuoklis* neturi jokių žinių apie analizuojamų žodžių semantiką, jų vartojimo dažnumą, todėl žodžiai analizuojami neatsižvelgiant į kontekstą. Dėl to kartais randama tokių homoformų, kurios galimos tik teoriškai, pavyzdžiui, forma *kokio* gali būti morfologiškai apibūdinta kaip daiktavardis *kokis*, reiškiantis „kokybė“, ir įvardis *koks*.

Morfologiškai anotuotame tekстыne skiriamos trys homoformų rūšys: 1) kaitomųjų, 2) kaitomųjų ir nekaitomųjų ir 3) vien tik nekaitomųjų kalbos dalių homoformos. Pirmąją homoformų rūšį iliustruoja tokie pavyzdžiai: *sukalė šuniui būdą – išsiugdė gerą būdą* (sutampa forma *būda*, kuri gali būti daiktavardžių *būda* arba *būdas* vns. gal.); *svarbi paslaptis – išdavė paslaptis* (sutampa daiktavardžio *paslaptis* vns. vard. ir dgs. gal. linksniai). Antruoju atveju sutampa, pavyzdžiui, daiktavardžio forma,rieveksmis ir prielinksnis: *atvyko su mano priešais – priešais matau mišką – stovi priešais mane*. Prie nekaitomųjų kalbos dalių homoformų priskirtinas žodelis *ir* – jis gali būti jungtukas, dalelytė irrieveksmis.

Kaip matyti iš pateiktų pavyzdžių, homoformomis laikomi ir tie žodžiai ar žodžių formos, kurie tariami skirtingai. O morfologinis daugiareikšmiškumas, kaip reiškinys, apima visas šias homoformų rūšis. Taigi automatinio būdu morfologiškai anotuotame tekstynėlyje analizuojamos homoformos, o ne homografai ar homonimai. Dėl to, kad analizuojamos rašytinės žodžių formos, daugiausia dėmesio skirta morfologiniams, o ne fonetiniams ar semantiniams skirtumams.

Dėl analizės pobūdžio pakeista homoformų samprata, todėl jos iš dalies apima ir homografus, ir dalinius homonimus. Straipsnio autorė mano, kad geriau pasirinkti vieną aiškų terminą, o ne kelis tarpusavyje susijusius homonimijos terminus. Tikimasi, kad tokia homoformų samprata analizuojant morfologiškai daugiareikšmes formas ilgainiui prigis.

#### 4. Išvados

Straipsnyje nagrinėjama, kaip kalbininkų vartojami su homonimija susiję terminai ir kaip turėtų būti įvardijamos iš automatinės morfologinės analizės išaiškėjusios morfologiškai daugiareikšmės formos.

1. Kalbiniuose darbuose nustatyta keletas skirtingų to paties termino sampratų:

1) homoformos ir morfologiniai / gramatiniai homonimai yra laikomi ta pačia arba skirtingomis homonimijos dalimis;

2) iš esamų homonimijos apibrėžimų neaišku, ar homoformas sudaro:

a) tų pačių ar skirtingų kalbos dalių žodžių formos,

b) fonetiškai vienodos ar ir skirtingos to paties žodžio formos;

3) vienų kalbininkų tos pačios formos įvardijamos kaip homoformos, kitų – kaip daliniai homonimai;

4) vienur teigiama, kad homoformos vienodai tariamos, kitur – kad gali skirtis tarimu;

5) netiksliai apibrėžiami kriterijai, kaip skirtinos homoformos, homonimai ir homografai;

6) neaišku, kuo laikomi tarnybinių kalbos dalių sutapimai.

2. Dėl terminijos nevienodumo ir gana specifinio analizės pobūdžio pasirinktas terminas *morfologinis daugiareikšmiškumas*, kuris suvokiamas kaip reiškinys, apimantis: 1) kaitomas ir nekaitomas formas; 2) skirtingų ir tų pačių kalbos dalių žodžių formas; 3) tam tikromis formomis, prozodiniais elementais besiskiriančias ir visiškai sutampančias žodžių formas ar žodžius.

3. Automatiniu būdu morfologiškai anotuotame tekstynėlyje daugiau kaip 47 proc. formų yra daugiareikšmės. Šioms konkrečioms formoms įvardyti pasirinktas terminas *homoformos*. Skiriamos 1) kaitomųjų, 2) kaitomųjų ir nekaitomųjų, 3) vien tik nekaitomųjų kalbos dalių homoformos, kurios gali skirtis prozodiniais elementais, t. y. kirčio vieta, priegaide ar balsio kiekybe. Terminas *homoformos* apima ir dalį tradiciškai homografais ar daliniais homonimais laikomų formų.

#### LITERATŪRA

- А х м а н о в а О. С. 1986: *Словарь омонимов русского языка*, Москва.  
B a r a u s k a i t ė J. 1979: *Leksikologijos pradmenys (Leksinė semantika)*, Vilnius.  
D r o t v i n a s V. 1986: *Pagrindinės leksikologijos sąvokos*, Vilnius.  
J a k a i t i e n ė E. 1980: *Lietuvių kalbos leksikologija*, Vilnius.



- K a t a m b a F. 1994: *English Words*, London and New Your.
- KTŽ 1990: K. Gaivenis, S. Keinys. *Kalbotyros terminų žodynas*, Kaunas.
- LK I 1995: *Lietuvių kalba* 1, Vilnius.
- LKE 1999: *Lietuvių kalbos enciklopedija*. Vilnius.
- P a l m e r F. R. 1981: *Semantics*, Cambridge.
- R i m k u t ė E. 2002: Homoformos dabartinės lietuvių kalbos tekstyne. – *Lituanistica* 2 (50), 86–101.
- С о б о л е в а П. А. 1980: *Словообразовательная полисемия и омонимия*, Москва.
- U r b u t i s V. 1956: *Lietuvių kalbos leksikos homonimų susidarymo būdai* (Filologijos mokslų kandidato disertacijos autoreferatas), Vilnius.
- Z i n k e v i č i u s V. 2000: *Lemuoklis* – morfologinei analizei. – *Darbai ir Dienos* 24, 246–273.

### **HOMONYMS, HOMOFORMS AND HOMOGRAPHS FROM THE VIEWPOINT OF CORPUS LINGUISTICS**

#### *Summary*

The article deals with a small part of the corpus of the Lithuanian language, which was automatically tagged. The corpus with morphological tags has shown a high level degree of ambiguity of the language: about 47 percent of word forms are ambiguous.

In the Lithuanian linguistics traditionally the greatest attention has been devoted to the lexical homonyms. Other levels of ambiguity have not been analysed. For this reason it was necessary to define the main terms of the morphological ambiguity and to solve the problematic issues of the relevant terminology.

It was difficult to define what are the main differences between homographs, homonyms and homoforms. Traditionally in linguistics it is affirmed that homoforms must coincide phonetically, while homographs are pronounced differently. This article shows that some homoforms can differ by the place of accent, the quality of vowels or intonation. Homoforms were analysed instead of homographs or homonyms, because the main attention was devoted to morphological but not phonetic or semantic differences.

Erika RIMKUTĖ  
 Vytauto Didžiojo universitetas  
 Donelaičio g. 58, Kaunas  
 E. paštas Erika\_Rimkute@fc.vdu.lt

Gauta 2003-03-18