

Gramatinė morfologinių samplaikų klasifikacija

Erika Rimkutė

Anotacija. Tiek lietuvių, tiek ir kitos kalbos yra fražiškos. Tai reiškia, kad kalbėdami renkamės ne atskirus žodžius, o tų žodžių junginius ar samplaikas. Šiame straipsnyje pateikta dažnai lietuvių kalboje vartojamų leksinių vienetų, atspindinčių kalbos fražiškumą, – morfologinių samplaikų – gramatinė klasifikacija. Morfologinės samplaikos – tai iš dviejų ar daugiau kaitomų ir nekaitomų kalbos dalių sudaryti stabiliai vartojami junginiai, sudarantys sintaksiškai ir semantiškai nedalomą vienetą. Morfologinės samplaikos suskirstytos atsižvelgiant į tai, kokių kalbos dalių funkciją atlieka. Jos suskirstytos į samplaikas, atliekančias prieveiksmių, prielinksnių, dalelių, jungtukų, jaustukų, išiktukų funkcijas. Taip pat rasta samplaikinių įvardžių, kurie irgi laikomi morfologinėmis samplaikomis, ir tokių samplaikų, kurių negalima priskirti vienai kuriai nors kalbos daliai (jos vadinamos neidentifikuotomis morfologinėmis samplaikomis). Straipsnyje rašoma apie tai, kaip morfologinės samplaikos fiksuojamos žodynuose. Pastebėta, kad žodyne galima rasti nemažai morfologinių samplaikų, tačiau tik nedidelė jų dalis pateikta kaip atskiri antraštiniai žodžiai. Daugumą samplaikų žodyne galima rasti tarp pavyzdžių, iliustruojančių tam tikro žodžio vartoseną. Dėl šios priežasties žodyno naudotojams sunku rasti reikalingas morfologines samplaikas. Priede pateiktas morfologinių samplaikų sąrašas, suskirstytas pagal morfologinių samplaikų atliekamą kalbos dalių funkciją, turėtų būti pravartus besimokantiems lietuvių kalbos, atliekant automatinę morfologinę ir sintaksinę lietuvių kalbos analizę, verčiant iš lietuvių kalbos ir į lietuvių kalbą.

Reikšminiai žodžiai: morfologinės samplaikos; morfologinė analizė; sintaksinė analizė; kalbos dalis; kalbos fražiškumas.

Įvadas

Vieno iš žymiausių tekstynų lingvistikos atstovo J. Sinclairio (1996, cituojama iš Marcinkevičienė, 2000) nuomone, kalbose dominuoja fražiškumo principas: kalbėdami pasirenkame ne atskirus žodžius, o tam tikras frazes, iš kurių konstruojame sakinius. Tekstynų analizė įtikinamai atskleidė, kad apie 80 proc. teksto kuriama parenkant ne pavienius izoliuotus leksinius vienetus, bet didesnes ar mažesnes jų grupes, turinčias tam tikrą gramatinę struktūrą (Marcinkevičienė, 2000).

Šiame straipsnyje rašoma apie vienus iš dažniausių kalbos fražiškumą atskleidžiančių leksinių vienetų, kurie sudaryti iš kelių dėmenų, bet turi bendrą reikšmę. Šie vienetai vadinami morfologinėmis samplaikomis (toliau MS). Tai dvižodžiai ir ilgesni iš kaitomų ir nekaitomų kalbos dalių sudaryti sustabarėję junginiai, sudarantys sintaksiškai ir semantiškai nedalomą vienetą, pvz.: *kas nors, bet kur, anksčiau ar vėliau, taip pat, be galo* (Rimkutė, 2006).

MS kaip reikšminys išsamiai aprašytas Rimkutė, et al. (2005), Rimkutė (2006). Minėtuose darbuose paaiškinti termino *morfologinės samplaikos* pasirinkimo motyvai; pateikti penki svarbiausi MS nustatymo kriterijai (morfologinis neskaidomumas, leksinės reikšmės bendrumas, sustabarėjimo laipsnis, neapibrėžta leksinė aplinka, gramatinės sudėties pastovumas).

Kovalevskaitė, et al. (2008) straipsnyje MS bandytos klasifikuoti sintaksiškai semantiškai. Minėtame straipsnyje daugiausia dėmesio skirta MS vertimui į kelias kalbas ir samplaikiškumo nustatymui. Keliamas klausimas, kiek lietuvių kalbos MS išlieka samplaikomis ir kitose kalbose. Tarpkalbiniame kontekste išryškėjo, kad samplaikiškumo sąvoka yra itin problemiška.

Tiek šiame, tiek anksčiau minėtuose darbuose analizuojamos MS kaip dažnai kalboje vartojami leksiniai vienetai, mažai tirti ir menkai aprašyti kalbotyros darbuose. Besimokantiems lietuvių kalbos MS reikia išmokyti kaip vieną vienetą, o ne kaip atskirus žodžius, todėl reikalingas tokių samplaikų aprašas. Taip pat besimokantys turėtų žinoti, kokią funkciją atlieka samplaika, todėl šiame straipsnyje pateikiama aiškiausia morfologinė klasifikacija (iki šiol MS išsamiai suklasifikuotos morfologiškai, sintaksinė MS analizė yra gerokai sudėtingesnė, ji bus atlikta vėliau). Kita sritis, kur ypač daug dėmesio reikia skirti MS, – vertimas. Akivaizdu, kad MS turi būti verčiamos kaip vienas leksinis vienetas, bet ne visos samplaikos (ir jų atitikmenys) pateikiamos tiek lietuvių kalbos žodynuose, tiek ir dvikalbiuose žodynuose.

Šio straipsnio autorės anksčiau minėtuose darbuose daugiausia dėmesio skirta morfologinei samplaikų klasifikacijai, išsamiau aprašytos sunkiau klasifikuojamos MS, kurias galima laikyti prieveiksmiais, prielinksniais, dalelytėmis, jungtukais. Šiame straipsnyje norima pateikti išsamią MS klasifikaciją ir MS sąrašą, kuris sudarytas analizuojant šiuos leksinius vienetus kelerius metus. MS sąrašas nėra baigtinis, vis dėlto jame turėtų būti pateikta dauguma dabartinės lietuvių kalbos MS.

Straipsnyje pateiktos samplaikos yra išrinktos iš *Dabartinės lietuvių kalbos tekstyno* (žr. <http://donelaitis.vdu.lt>; taip pat iš šio tekstyno morfologiškai anotuotos dalies (plačiau žr. Rimkutė, 2006), *Dabartinės lietuvių kalbos žodyno* (žr. <http://www.autoinfa.lt/webdic/>). Pateiktame sąrašė nėra MS, rinktų iš sakytinės kalbos.¹ Žinoma, dauguma

¹ Priede pateiktame MS sąrašė yra jaustukų, kurie dažniausiai vartojami sakytinėje kalboje, grožinėje literatūroje. Nors sakytinės kalbos tekstų *Dabartinės lietuvių kalbos tekstyne* beveik nėra (tam tikra sakytinės kal-

čia pateiktų MS vartojama ir sakytinėje, ir rašytinėje kalboje, nes šie leksiniai vienetai pasižymi dažna vartoseną ir nepriklausymu nuo konteksto, tačiau sakytinėje kalboje, pvz., šnekamosios lietuvių kalbos tekстыne², gali išryškėti ir kitokių samplaikų, kurias būtų galima tirti išsamiau.

Gramatinė MS klasifikacija

Kaip minėta įvade, iki šiol MS išsamiai išanalizuotos morfologiškai: suskirstytos pagal tai, kokių kalbos dalių funkciją gali atlikti. MS pirmiausia suklasifikuotos morfologiškai, nes tokia informacija buvo būtina rengiant morfologiškai anotuotą lietuvių kalbos tekstyną (plačiau žr. Rimkutė, 2006). Iš pradžių automatiškai morfologiškai anotuojant tekstus MS buvo išskaidytos ir analizuotos kaip atskiri žodžiai, tačiau pamačius šių junginių sintaksinį ir semantinį bendrumą, taip pat atsižvelgiant į tai, kad dalis MS (tiesa, taip neįvardytos) pateiktos lietuvių kalbos žodynuose, buvo sudarytas MS sąrašas, o pačios MS suklasifikuotos morfologiškai ir analizuotos kaip tam tikra viena kalbos dalis (plg. čia ir sakytinės kalbos tekstyno rengimo praktiką, žr. Kamandulytė et al., 2008).

MS suskirstytos į 10 grupių: 1) samplaikas, laikytinas prieveiksmais (jų rasta ir vartojama daugiausia (žr. 1 priedą)), pvz.: *iš anksto, be galo, daugių daugiausia*; 2) samplaikas, laikytinas dalelytėmis, pvz.: *tarp kitko, vargu ar*; 3) samplaikas, laikytinas jungtukais, pvz.: *ne tik, bet ir, vis dėlto*; 4) samplaikas, laikytinas prielinksniais, pvz.: *nuo pat, prieš pat*; 5) samplaikas, laikytinas ir prieveiksmais, ir dalelytėmis (rasta viena tokia MS – *taip pat*); 6) samplaikinius įvardžius, pvz.: *kas nors, tam tikras, bet kas*; 7) samplaikas, laikytinas jaustukais, pvz.: *po galais, labas vakaras*; 8) samplaikas, laikytinas ištiktukais (pateikta tik viena tokia MS *čiūčia liūlia*, ji įtraukta iš *Dabartinės lietuvių kalbos žodyno*); 9) samplaikinius būdvardžius, pvz.: *vienui vienas*, ir 10) neidentifikuotas samplaikas, pvz., *be to, ir t. t., ir pan.*, kurių negalima priskirti vienai kuriai nors kalbos daliai.

Visas samplaikų sąrašas pateiktas 2 straipsnio priede. Šalia samplaikos nurodytas jos pavartojimo dažnumas 100 mln. žodžių *Dabartinės lietuvių kalbos tekstynė*.

Nustatant, kurios kalbos dalies funkciją atlieka tam tikra MS, samplaikas bandyta pakeisti vienu žodžiu, kurio kalbos dalį nesudėtinga nustatyti. Pavyzdžiui, MS *iš tolo* galima pakeisti vienu žodžiu *tolį*, kuris laikomas aiškiu prieveiksmiu. MS *iš tolo* atlieka tą pačią funkciją kaip ir *tolį*, taigi ši MS laikytina prieveiksmiu. Dalelytė *pat* dažnai įeina į reikšmės požiūriu neskaidomų junginių sudėtį, pvz., *prieš pat, iki pat*, bet nepakeičia kategorinės reikšmės, taigi anksčiau minėtos MS laikomos prielinksniais.

Dažniausiai visi MS dėmenys yra nekaitomi žodžiai, pvz.: *bet kada, vos tik, kada nors*. Retesnės iš kaitomų ir nekai-

tomų žodžių sudarytos MS, pvz.: *be galo, iš tiesų, nuo pradžių*. Pačios rečiausios iš abiejų kaitomų dėmenų sudarytos MS, pvz.: *šia prasme, nori nenori, vienui vienas*. *Dabartinės lietuvių kalbos tekstyno* duomenys parodė, kad vienos kurios nors kalbos dalies funkciją atliekančios MS pasižymi tomis pačiomis gramatinėmis ypatybėmis, pvz., *be galo* ir *kur nors* abu prieveiksmai; abi šios samplaikos atlieka aplinkybių funkciją, abi valdomos veiksmažodžių. Taigi nesvarbu, ar MS sudarytos iš kelių prieveiksminių, dalelyčių ar jungtukų, ar jas sudaro skirtingų kalbų žodžiai – jų gramatinės funkcijos yra panašios.

Straipsnio autorės manymu, tiek besimokantiems lietuvių kalbos, tiek verčiantiems į lietuvių kalbą ar iš jos, tiek analizuojant kalbą automatiškai, reikalingiausia yra morfologinė MS klasifikacija. Sintaksinė MS analizė yra sudėtingesnė, klasifikuojant sintaksiškai galima neišvengti subjektyvumo (subjektyvumo gali būti neišvengta ir pateiktoje morfologinėje klasifikacijoje), nes sunku apsispręsti dėl tinkamų sintaksinės klasifikacijos kriterijų, be to, ta pati samplaika gali atlikti kelias funkcijas.

Kovalevskaitė et al. (2008) straipsnyje MS pagal atliekamą sintaksinę funkciją suskirstytos į 9 grupes: 1) sudedamąsias, 2) priešpriešines, 3) laiko, 4) vietos, 5) būdo, 6) priežasties ir tikslo, 7) sąlygos ir nuolaidos, 8) tikslinimo, 9) kita.

MS skirstytos remiantis semantiniais sintaksiniais požymiais. Šitai klasifikuojant MS (kaip ir morfologinės analizės atveju) dažniausiai remtasi pakeitimu: samplaikas bandyta pakeisti aiškios gramatinės reikšmės žodžiais ir remiantis tais žodžiais priskirti vienam ar kitam tipui, pvz., *tu tarpu* galima pakeisti žodžiais *o, bet, tačiau*, vadinasi, tokia reikšmė vartojama samplaika *tu tarpu* priskirtina priešpriešinį ryšį žyminčių samplaikų tipui. *Kur kas* galima pakeisti *daug daugiau*, taigi tai bus būdo prieveiksmis, *galų gale* galima pakeisti *pagaliau* – tai būdo aplinkybė.

Tos pačios MS gali būti priskiriamos skirtingoms grupėms, atsižvelgiant į jų tekste atliekamą funkciją, ypač variantiškumu pasižymi MS, atliekančios prielinksnių funkciją, pvz., *prieš pat, ligi pat*, gali įgyti ir vietos, ir laiko reikšmę (išsamiau žr. Kovalevskaitė et al., 2008).

Kaip jau minėta, dėl sintaksinės analizės sudėtingumo čia trumpai pristatytą sintaksinę MS klasifikaciją būtina tikslinti. Taip pat reikėtų ištirti, kokiomis sintaksinėmis funkcijomis pasižymi įvairiomis kalbos dalimis laikytinos MS. Ši tyrimą planuojama atlikti ateityje.

MS pateikimas *Dabartinės lietuvių kalbos žodyne*

Šiame skyriuje apžvelgta, kiek tekstynė vartojamų MS pateikiama *Dabartinės lietuvių kalbos žodyne* (toliau DLKŽ; naudotasi 2000 m. laida ir elektroniniu žodyno variantu, žr. <http://www.autoinfa.lt/webdic/>), kaip jos suklasifikuotos, kiek toks MS pateikimas parankus lietuvių kalbos vartotojams.

DLKŽ rastos 217 MS⁴, o *Dabartinės lietuvių kalbos tekstynė* rastos 367 samplaikos (žr. 1 priedą), taigi akivaizdžiai

bos atmaina galima laikyti Lietuvos Respublikos Seimo stenogramas), bet tekstynė nemažai grožinės literatūros tekstų. Greičiausia juose ir buvo pavartoti šiame sąraše pateikti samplaikiniai jaustukai.

² Žr. http://www.vdu.lt/LTcourses/?pg=41&lang=1&menu_id=112.

³ Morfologinės samplaikos, tokios kaip *kas nors, kažin kas, šioks toks ir pan.*, vadinamos ne samplaikomis, laikytinomis įvardžiais, o samplaikiniais įvardžiais, nes toks terminas jau yra nusistovėjęs kalbotyros darbuose (plg. Rosinas, 1996).

⁴ DLKŽ rastos MS išsamiai aprašytos V. Levanavičiūtės 2008 m. apgintame bakalauro darbe *Morfologinė „Dabartinės lietuvių kalbos žodyno“ morfologinių samplaikų, modalinių žodžių ir įterpinių analizė*.

matyti, kad žodyne neužfiksuota daug dabartinėje rašytinėje lietuvių kalboje vartojamų MS. Kita vertus, nemaža dalis DLKŽ rastų samplaikų pateiktas tik pavyzdžiuose, iliustruojančiuose tam tikro žodžio vartoseną (apie tai rašoma toliau), taigi žodyne nelengva rasti MS.

Iš DLKŽ išrinktos MS buvo suklasifikuotos atsižvelgiant į tai, kokios kalbos dalies funkciją atlieka. 151 iš DLKŽ pateiktų MS gali būti laikomos prieveiksmais, pvz.: *per amžius, iš pasalu, vienu žodžiu*; 28 samplaikos laikytinos įvardžiais, pvz.: *bet kas, šioks toks*; 21 MS vartojama kaip jautukai, pvz.: *ačiū Dievui, aš tau*; 14 MS atlieka dalelių funkciją, dėl to priskirtinos šiai kalbos daliai, pvz.: *tai va, tai kad*; penkios samplaikos laikytinos jungtukais, pvz.: *kaip kad, kad ir kada*; penkios samplaikos laikytinos prielinksniais, pvz.: *ligi pat, iki pat*; žodyne rastas vienas samplaikinis ištiktukas (*čiučia liūlia*) ir viena samplaika, kuri atlieka tiek prieveiksmino, tiek dalelytės funkciją (*taip pat*).

Dauguma iš žodyno išrinktų samplaikų rasta ir *Dabartinės lietuvių kalbos tekстыne*, tačiau DLKŽ buvo ir tokių samplaikų, kurios nebuvo įtrauktos į anksčiau remiantis tekстыno duomenimis sudarytą MS sąrašą. Į pirminį sąrašą kai kurios MS (pvz.: *per amžius, iš pasaly, trumpai drūtai, čiūčia liūlia*) nebuvo įtrauktos, nes iš pradžių abejota dėl jų samplaikiškumo, nebuvo ištirtas dažnumas (vienas iš samplaikų nustatymo kriterijų); kai kurių junginių iš viso nerasta tekстыne (žr. 1 priedą).

Keletas žodyne rastų samplaikų yra labai retos, dabartinėje kalboje beveik nebevartojamos, todėl tokios MS į straipsnyje pateiktą MS sąrašą neįtrauktos. Tai samplaikos *kadaĩ* *kadẽs*; *kadaĩ* *kadžiũ*; *ką tik ne*; *ko tik ne*; *kur ten kas*; *iš paskutinio*; *nẽ štai*, *nẽ antai*.

Jau minēta, ka DLKŽ nevarojamas terminas *morfologi-
nēs samplaikos*. Kartais DLKŽ MS pateikamas kaip atski-
ras antraštinis žodis, kuriam nurodoma viena kalbos dalis,
pvz.: *kur nors* – prieveiksmis, *bet kuris* – įvardis ir pan.
Dažniausiai samplaikų galima rasti antraštinį žodį iliust-
rojančiuose pavyzdžiuose, pvz.: *anksčiau ar vėliau, be
perstojo*⁵. Kai kurios DLKŽ pateiktos ir į šio straipsnio
priedą įtrauktos MS laikomos frazeologizmais. Tai sam-
plaikos *nuo amžių, per amžius, antra vertus, iš bėdos, iš
teisybės, iš tiesų, vienu žodžiu, savo ruožtu, šiaip sau, toli
gražų, vargai negalais, vienas kitas* ir kt.⁶ MS *tiesą sakant*
žodyne laikoma įterpiniu ir tik pačiam kalbos vartotojui
reikia nuspręsti, kurios kalbos dalies funkciją ši samplaika
atlieka. Panaši situacija ir su DLKŽ pateiktais modaliniais
žodžiais, kuriuos galima laikyti MS, pvz., *iš viso*. DLKŽ

tik nurodoma, kad tai modalinis žodis, bet nepateikiama jo kalbos dalis.⁷

Taigi iš anksčiau pateiktų pavyzdžių matyti, kad MS gali būti vaizdingos, vadinasi, dalį MS galima laikyti frazeologizmais. MS gali būti vartojamos kaip įterpiniai ar kaip modaliniai žodžiai. Šiame straipsnyje laikomasi nuomonės, kad nepaisant MS sintaktinių ir semantinių ypatybių, MS galima ir būtina suklasifikuoti morfologiškai.

Apibendrinant MS pateikimą žodyne, galima teigti, kad nemažai šiame darbe morfologinėmis samplaikomis vadinamų junginių galima rasti ir DLKŽ, vis dėlto dauguma dabartinėje lietuvių kalboje vartojamų MS nėra pateiktos DLKŽ. Žodyne pasigesta, pvz., tokių dabartinėje lietuvių kalboje dažnai vartojamų MS: *iš tikro, iš tikrųjų, šiaip ne taip, trumpai sakant, viena vertus* ir pan. Be to, net ir rastos žodyne MS pateikiamos gana nenuosekliai: vienur prie atitinkamo antraštinio žodžio iliustruojančiuose pavydžiuose, kitur – atskiru žodyniniu straipsniu, dar kitos samplaikos pateikiamos kaip frazeologizmai, įterpiniai ar modaliniai žodžiai.

Tikimasi, kad šiam darbe pateiktas MS sąrašas ir MS pavartojimo *Dabartinės lietuvių kalbos tekстыne* skaičiai bus pravartūs būsimų žodynų sudarytojams. Į samplaikas, kaip į dažnai vartojamą specifinį dabartinės kalbos leksinį vienetą, vertėtų atkreipti dėmesį tiek vienalabių, tiek ir dvikalbių žodynų sudarytojams.

Išvados

Straipsnyje aprašyti dažnai dabartinėje lietuvių kalboje vartojami leksiniai vienetai – morfologinės samplaikos. Tai keliažodžiai iš kaitomų ir nekaitomų kalbos dalių sudaryti sustabarę junginiai, sudarantys sintaksiškai ir sėmantiškai nedalomą vieneta.

Pateikta išsami morfologinė MS klasifikacija ir pirmieji sintaksinės MS klasifikacijos siūlymai. Atsižvelgiant į tai, kokių kalbos dalių funkciją atlieka MS, jos suskirstytos į samplaikas, laikytinas prieveiksmais, prielinksniais, jungtukais, dalelytėmis, jastukais. Pateikta ir samplaikinių ištiktukų bei įvardžių. MS rinktose iš įvairių šaltinių: tekstyne, žodyne. Šis MS sąrašas, nors ir nėra baigtinis, bet turėtų apimti didžiąją dabartinės lietuvių kalbos MS dalį.

Šiuo straipsniu norėta atkreipti dėmesį į MS kaip į dažnai vartojamą dabartinės lietuvių kalbos leksinį vienetą, kuriam dar trūksta dėmesio kalbotyros darbuose; reikalingas nuoseklesnis MS pateikimas *Dabartinės lietuvių kalbos žodyne*.

Priede pateiktas MS sąrašas turėtų būti pravartus besimokantiems lietuvių kalbos, automatiškai (pirmiausia morfo-

⁵ Jeigu MS pateikiama tik iliustruojant kurio nors žodžio vartoseną, susiduriama su problema, kaip surasti tą samplaiką, pvz., iškantys MS *ankščiau ar vėliau* ras ją tik prie žodžio *ankščiau*, o prie žodžių *vėlai*, *vėliau* ši samplaika nenurodyta. Taigi DLKŽ MS paieška yra sudėtinga. Net ir naudojantis elektroniniu žodyno variantu negalima rasti MS, jei jos pateiktos tik prie antraštinį žodį iliustruojančių pavyzdžių.

⁶ Šiame darbe nebus gilnamasi į vaizdingumo kriterijų, tačiau norėtųsi atkreipti dėmesį, kad DLKŽ vaizdingumo kriterijus nėra taikomas nuosekliai. Kyla klausimas, jei frazeologizmais laikomos tokios samplaikos, kaip *antra vertus*, *iš tiesų*, tai kodėl frazeologizmu nelaikomas *trumpai drūtai*?

Beje, DLKŽ prie minėtų frazeologizmų nurodomos tik jų reikšmės, bet nepaaiškinama, kokia kalbos dalimi laikytini tie vaizdingi junginiai.

7 Čia išryškėja svarbi DLKŽ problema, kai ne prie visų žodžių ir žodžių junginių nurodomos kalbos dalys. Plg. šiuos modalius žodžius ir įterpinius: *rodos, sakykime, žinia, žinoma, tikriausiai, kaip antai, antra vertus, tiesą sakant, tiesa, pavyzdžiui*. Nediskutuotina, kad šie žodžiai dažnai tekste įgyja tam tikrų modaliųjų atspalvių, taigi vartojami kaip modaliniai žodžiai, arba įterpiami norint ką nors pabrėžti, paaiškinti. Vis dėlto kyla klausimas, ar tokių modaliųjų žodžių ir įterpinių negalima analizuoti morfologiškai. Žodyno naudotojams sintaksinės pažymos yra pravarčios, bet vis tiek reikėtų laikytis nuoseklos ir prie visų antraštinį žodžių pateikti ir morfologines pažymas.

logiškai ir sintaksiškai) analizuojant lietuvių kalbą, verčiant tekstus iš lietuvių kalbos ir į ją.

Literatūra

1. Kamandulytė, L., Savickienė, I., 2008. The Corpus of Spoken Lithuanian: Methodology and Development. *The Third Baltic Conference on Human Language Technologies*, pp.127–133.
2. Kovalevskaitė, J., Rimkutė, E., 2008. Morfologinių samplaikų struktūros ypatumai: kelių kalbų palyginimas. *Darbai ir dienos*, nr. 50, pp.117–152.
3. Levanavičiūtė, V., 2008. *Morfologinė „Dabartinės lietuvių kalbos žodyno“ morfologinių samplaikų, modalinių žodžių ir įterpinių analizė*: bakalauro darbas. Kaunas: Vytauto Didžiojo universitetas.
4. Marcinkevičienė, R., 2000. Tekstynų lingvistika (teorija ir praktika). *Darbai ir dienos*, nr. 24, pp.7–63.
5. Rimkutė, E., Jarašiūnaitė, G., Homola, P., 2005. Morfologinių samplaikų atpažinimas ir klasifikavimas. *Lituanistica*, nr. 2, pp.58–75.
6. Rimkutė, E., 2006. *Morfologinio daugiareikšmiškumo ribojimas kompiuteriniame tekste*: daktaro disertacija. Kaunas: Vytauto Didžiojo universitetas.
7. Rosinas, A., 1996. *Lietuvių bendrinės kalbos įvardžiai*. Vilnius: Mokslo ir enciklopedijų leidykla.

Erika Rimkutė

Grammatical Classification of Morphological Multiword Units

Summary

The article deals with morphological multiword units that are frequent in modern Lithuanian and reflect the phraseological tendency of the language. A comprehensive morphological classification of these lexical units is provided. Morphological multiword units are constantly used combinations that are made up of two or more inflective or non-inflective parts of speech and have a unified common meaning. Morphological multiword units are classified according to the part of speech function they perform in a sentence. They are subdivided into combinations that perform the functions of adverbs, prepositions, particles, conjunctions and exclamations. Complex pronouns that are considered as morphological multiword units, as well as multiword units that cannot be treated as any part of speech (they are called unidentified morphological multiword units) were also identified. The article also discusses the way morphological multiword units are recorded in dictionaries. It has been noticed, that a lot of morphological multiword units are included in dictionaries, but only a part of them is presented as separate dictionary entries. Most morphological multiword units can be found among the examples that illustrate the usage of a certain word. For this reason, dictionary users have difficulty in finding the necessary morphological multiword units. The list of morphological multiword units classified by the part of speech function is provided in the appendix. It should be useful to the learners of Lithuanian, in automatic morphological and syntactic analysis of the Lithuanian language and in translation to and from Lithuanian.

Straipsnis įteiktas 2009 01
Parengtas spaudai 2009 04

Apie autorę

Erika Rimkutė, dr., Vytauto Didžiojo universiteto Humanitarinių mokslų fakulteto Kompiuterinės lingvistikos centro mokslo darbuotoja, Lietuvių kalbos katedros lektorė.

Mokslinės veiklos sritys: tekstynų lingvistika, kompiuterinė lingvistika, automatinė morfologinė analizė ir sintezė, morfologinis daugiareikšmiškumas, automatinė sintaksinė analizė.

Adresas: Vytauto Didžiojo universitetas, Humanitarinių mokslų fakultetas, Kompiuterinės lingvistikos centras, K. Donelaičio g. 52–206, 44244 Kaunas.

El. paštas: e.rimkute@hmf.vdu.lt

1 PRIEDAS. Morfologinė MS klasifikacija

Kalbos dalis, kurios funkciją atlieka MS	Bendras MS skaičius	MS, pavartotos ir žodyne, ir tekстыne	MS, pavartotos tik tekстыne	MS, pavartotos tik žodyne
Prieveiksmiai	257 (67,99 %)	142	109	6
Įvardžiai	32 (8,47 %)	23	4	5
Jaustukai	31 (8,20 %)	21	10	-
Dalelytės	19 (5,03 %)	9	10	-
Jungtukai	16 (4,23 %)	5	11	-
Neidentifikuotos MS	9 (2,38 %)	-	9	-
Prielinksniai	9 (2,38 %)	5	4	-
Būdvardžiai	3 (0,79 %)	2	1	-
Prieveiksmiai / dalelytės	1 (0,26 %)	1	-	-
Išiktukai	1 (0,26 %)	1	-	-
Iš viso	378 (100,00 %)	209	158	11

2 PRIEDAS. Morfologinių samplaikų sąrašas

Kalbos dalis, kuria laikytinos MS	Pavyzdžiai
1. Prieveiksmiai	- aiškių aiškiausiai (36)⁸ - anksčiau ar vėliau (1344) - antra vertus (3597)⁹ - aukštyn žemyn (271) - be abejo (7814) - be galo (4246) - be galo be krašto (48) - be jokio abejo (1) - be niekur nieko (70) - be perstojo (582) - be reikalo (2315) - bent kiek (1821) - bet kada (1393) - bet kaip (2722) - bet kiek (379) - bet kuomet - bėgant laikui (78) - bet kur (1230) - čia pat (4563) - dar tik (2366) - dar vis (634) - daugiau ar mažiau (1576) - daugiau daugiausia (88) - daugiau daugiausiai (11) - daug kaip (1) - daug kur (894) - dėl ko (3340) - dėl to (17000*)¹⁰ - dėl visa ko (176) - dėl viso pikto (1) - diena į dieną (2) - diena iš dienos (171) - dieną naktį (320) - galų gale (4301) - iki kol (46) - iki kilei (4) - iki šiol (17000*) - iki šiolei (428) - iki tol (3280) - iki tolei (34)

⁸ Atkreiptinas dėmesys, kad ne visada čia pateikti junginiai vartojami kaip MS, plg. šiuos sakinius: *Geriausia, kai gali užmigti bet kur ir bet kada* vs. *Bet kur jie dės Lietuvoje?* Pirmame pavyzdyje matyti, kad *bet kur* ir *bet kada* yra MS, nes galima pakeisti kitais aiškiais prieveiksmiais *visur* ir *visada*. Antrame sakinyje *bet* yra jungtukas, *kur* – prieveiksmis ir šie žodžiai tokia kontekste nesudaro MS.

⁹ Kursyvu pažymėtos tos MS, kurios rastos ir DLKŽ, ir *Dabartinės lietuvių kalbos tekстыne*. Pabrauktos tos MS, kurios rastos DLKŽ, bet nepavartotos tekстыne, todėl prie jų nenurodytas dažnumas. Paprastu šriftu parašytos tos MS, kurios rastos tik *Dabartinės lietuvių kalbos tekстыne*.

¹⁰ MS vartojimo dažnumas buvo skaičiuojamas su programa *Word Smith Tools*. Dėl techninių galimybių ši programa gali suskaičiuoti dažnumą iki 17 000. Taigi prie tų MS, kurių tikslaus dažnumo nepavyko suskaičiuoti, parašyta 17 000 ir žvaigždutė.

Kalbos dalis, kuria laikytinos MS	Pavyzdžiai
	- iki valiai (65) - iki valios (51) - ilgą laiką (2454) - iš anksto (8358) - iš antros pusės (45) - iš arčiau (432) - iš arti (648) - iš bėdos (177) - iš dalies (9600) - iš eilės (3935) - iš esmės (14538) - iš karto (9763) - iš kitos pusės (988) - iš lengvo (27) - iš lėto (904) - iš mažens (166) - iš naujo (5447) - iš palengvo (9) - iš paniūrų (2) - iš pasalių (204) - iš paskos (1428) - iš paskutiniųjų (575) - iš pat mažens (31) - iš pat pradžių (765) - iš pat pradžios (22) - iš pradžios (93) - iš pradžių (12617) - iš reto (23) - iš seno (139) - iš teisybės (51) - iš tiesų (12409) - iš tikro (1699) - iš tikrųjų (16894) - iš tyko (4) - iš toli (768) - iš toliau (335) - iš tolo (1970) - iš vienos pusės (853) - iš viso (14151) - į valias (148) - jokiu būdu (1715) - juo labiau (4508) - ką tik (5578) - ką tiktai (82) - kada (5) - kada kas (1) - kada ne kada (42) - kada nors (4333) - kada tik (180)

¹¹ *Kaip mat* rašymas dviem žodžiais laikomas rašybos klaida, vis dėlto ši samplaika gana dažna *Dabartinės lietuvių kalbos tekстыne*.

¹² Dviem žvaigždutėmis pažymėti tie atvejai, kai MS vartojimo dažnumą reikėjo skaičiuoti atsižvelgiant į tai, ar yra tarpas tarp samplaikos dėmenų, ar jo nėra. Savaiame suprantama, kad tiek *t.y.*, tiek ir *t. y.* yra viena ir ta pati samplaika, vis dėlto norint gauti tikslų pavartojimo dažnumą, reikia atsižvelgti ir į jų rašybą.

Kalbos dalis, kuria laikytinos MS	Pavyzdžiai
	90. <i>kai kada</i> (1822) 91. <i>kai kur</i> (2818) 92. <i>kai tik</i> (3105) 93. <i>kai tiktai</i> (57) 94. <i>kaip kada</i> (28) 95. <i>kaip mat</i> (164) ¹¹ 96. <i>kaip matai</i> (6) 97. <i>kaip ne kaip</i> (19) 98. <i>kaip niekur nieko</i> (126) 99. <i>kaip nors</i> (2468) 100. <i>kaip reikiant</i> (831) 101. <i>kaip tiktai</i> (190) 102. <i>kaip žinia</i> (943) 103. <i>kairėn dešinėn</i> (13) 104. <i>kas be ko</i> (271) 105. <i>kas dieną</i> (299) 106. <i>kaži kada</i> (8) 107. <i>kaži kaip</i> (61) 108. <i>kaži kiek</i> (7) 109. <i>kaži kodėl</i> (18) 110. <i>kaži kur</i> (54) 111. <i>kažin kada</i> (49) 112. <i>kažin kaip</i> (166) 113. <i>kažin kiek</i> (208) 114. <i>kažin kodėl</i> (72) 115. <i>kažin kur</i> (177) 116. <i>kiek nors</i> (125) 117. <i>kita vertus</i> (7607) 118. <i>kitaip kalbant</i> (22) 119. <i>kitaip sakant</i> (2139) 120. <i>kitaip tariant</i> (3075) 121. <i>kol kas</i> (17000*) 122. <i>kolei kas</i> (67) 123. <i>kuo labiau</i> (736) 124. <i>kur kas</i> (10427) 125. <i>kur ne kur</i> (194) 126. <i>kur nors</i> (3037) 127. <i>kur ten kas</i> 128. <i>kurį laiką</i> (4612) 129. <i>kuris laikas</i> (547) 130. <i>kuriam laikui</i> (725) 131. <i>laikui bėgant</i> (990) 132. <i>lyg niekur nieko</i> (345) 133. <i>lig šiol</i> (1387) 134. <i>lig šiolei</i> (105) 135. <i>lig tol</i> (150) 136. <i>lig tolei</i> (11) 137. <i>lig valiai</i> (9) 138. <i>lig valios</i> (4) 139. <i>ligi kol</i> (2) 140. <i>lig kolei</i> 141. <i>ligi kolei</i> 142. <i>ligi šiol</i> (1643) 143. <i>ligi šiolei</i> (96) 144. <i>ligi tol</i> (264) 145. <i>ligi tolei</i> (3) 146. <i>ligi valiai</i> (9) 147. <i>ligi valios</i> (2) 148. <i>maža ką</i> (89) 149. <i>maža to</i> (1092) 150. <i>mažų mažiausiai</i> (411) 151. <i>nei šiaip, nei taip</i> (51) 152. <i>ne ką</i> (1652) 153. <i>ne kažin kiek</i> (2) 154. <i>ne kažin kur</i> 155. <i>ne kiek</i> (1029) 156. <i>ne motais</i> (382) 157. <i>negana to</i> (1428) 158. <i>nepriklausomai nuo</i> (2989) 159. <i>nieku būdu</i> (66) 160. <i>nieku gyvu</i> (179) 161. <i>nori nenori</i> (427) 162. <i>norint nenorint</i> (3) 163. <i>norom nenorom</i> (425) 164. <i>noromis nenoromis</i> (44)

Kalbos dalis, kuria laikytinos MS	Pavyzdžiai
	165. <i>nuo amžių</i> (259) 166. <i>nuo anksto</i> 167. <i>nuo mažens</i> (542) 168. <i>nuo mažo</i> (3) 169. <i>nuo mažumės</i> (142) 170. <i>nuo pat mažens</i> (104) 171. <i>nuo pat mažumės</i> (47) 172. <i>nuo pat pradžių</i> (569) 173. <i>nuo pradžios</i> (401) 174. <i>nuo pradžių</i> (128) 175. <i>nuo seno</i> (1621) 176. <i>nuo šiol</i> (1938) 177. <i>nuo šiolei</i> (35) 178. <i>nuo tol</i> (148) 179. <i>nuo tolei</i> (3) 180. <i>per amžius</i> (655) 181. <i>per daug</i> (9323) 182. <i>per mažai</i> (1876) 183. <i>pirmyn atgal</i> (122) 184. <i>po laiko</i> (32) 185. <i>po šiai dienai</i> (123) 186. <i>po to</i> (17000*) 187. <i>po truputį</i> (1577) 188. <i>po visam</i> (15) 189. <i>prieš tai</i> (5697) 190. <i>priešingu atveju</i> (1584) 191. <i>priklausomai nuo</i> (3520) 192. <i>savo ruožtu</i> (4229) 193. <i>šen bei ten</i> (201) 194. <i>šen ir ten</i> (74) 195. <i>šen ten</i> (36) 196. <i>šiaip ar taip</i> (1842) 197. <i>šiaip jau</i> (1507) 198. <i>šiaip ne taip</i> (260) 199. <i>šiaip sau</i> (1118) 200. <i>šiaip taip</i> (843) 201. <i>šiek tiek</i> (16180) 202. <i>t. y.** (6951)</i> ¹² 203. <i>t. y.** (17000*)</i> 204. <i>taip sakant</i> (909) 205. <i>tai šen, tai ten</i> (140) 206. <i>ten pat</i> (3445) 207. <i>ties ir ties</i> (179) 208. <i>ties pat</i> (3324) 209. <i>ties to</i> (182) 210. <i>tiesą kalbant</i> (8) 211. <i>tiesą pasakius</i> (357) 212. <i>tiesą sakant</i> (2119) 213. <i>tiesiai šviesiai</i> (26) 214. <i>tik ką</i> (1103) 215. <i>tik tik</i> (79) 216. <i>tokiu būdu</i> (5217) 217. <i>toli gražu</i> (2771) 218. <i>trumpą laiką</i> (1252) 219. <i>trumpai drūtai</i> (154) 220. <i>trumpai kalbant</i> (44) 221. <i>trumpai sakant</i> (91) 222. <i>trumpai tariant</i> (550) 223. <i>tuo atveju</i> (3893) 224. <i>tuo būdu</i> (1198) 225. <i>tuo labiau</i> (3389) 226. <i>tuo laiku</i> (1252) 227. <i>tuo metu</i> (17000*) 228. <i>tuo pačiu</i> (3128) 229. <i>tuo pat metu</i> (3315) 230. <i>tuo tarpu</i> (17000*) 231. <i>tuo tikslu</i> (705) 232. <i>tuoj pat</i> (4454) 233. <i>tuoj tuoj</i> (199) 234. <i>tuojau pat</i> (1470) 235. <i>už dyką</i> (377) 236. <i>vargais negalais</i> (458) 237. <i>vėl ir vėl</i> (243) 238. <i>viena vertus</i> (1186) 239. <i>vienai ar kitaip</i> (1810)

Kalbos dalis, kuria laikytinos MS	Pavyzdžiai
	240. <i>viens du</i> (75) 241. <i>vienu metu</i> (1579) 242. <i>vienur kitur</i> (153) 243. <i>vienu žodžiu</i> (282) 244. <i>vis dar</i> (9914) 245. <i>viso labo</i> (779) 246. <i>vis tiek</i> (8586) 247. <i>vis tik</i> (845) 248. <i>vis tiktai</i> (117) 249. <i>visur kur</i> (49) 250. <i>vis vien</i> (1957) 251. <i>vis viena</i> (274) 252. <i>visą laiką</i> (5412) 253. <i>visų pirma</i> (7293) 254. <i>vos ne vos</i> (126) 255. <i>vos tik</i> (1709) 256. <i>vos tiktai</i> (40) 257. <i>vos vos</i> (677)
2. Dalelytės	1. <i>be kita ko</i> (620) 2. <i>gali būti</i> (5) 3. <i>kaip ir</i> (17000*) 4. <i>kažin ar</i> (2066) 5. <i>ko gera</i> (133) 6. <i>ko gero</i> (6100) 7. <i>lyg ir</i> (7532) 8. <i>net ir</i> (12655) 9. <i>šia prasme</i> (1404) 10. <i>ta prasme</i> (1121) 11. <i>taip kad</i> (119) 12. <i>tarp kita ko</i> (80) 13. <i>tarp kitko</i> (1105) 14. <i>vargu ar</i> (5709) 15. <i>vargu bau</i> (106) 16. <i>vien tik</i> (5544) 17. <i>vien tiktai</i> (353) 18. <i>tai va</i> (347) 19. <i>taip jau</i> (182)
3. Jungtukai	1. <i>kaip antai</i> (765) 2. <i>bent jau</i> (7339) 3. <i>bet ir</i> (17000*) 4. <i>kad ir kada</i> (11) 5. <i>kad ir kaip</i> (2372) 6. <i>kad ir kiek</i> (390) 7. <i>kad ir kur</i> (260) 8. <i>kad tik</i> (5077) 9. <i>kaip kad</i> (63) 10. <i>kaip tik</i> (13558) 11. <i>lig tik</i> (5) 12. <i>ne tik</i> (17000*) 13. <i>ne tiktai</i> (1067) 14. <i>ne vien</i> (9195) 15. <i>nors ir</i> (12727) 16. <i>vis dėlto</i> (17000*)
4. Prielinksniai	1. <i>į pat</i> (2) 2. <i>iš pat</i> (1195) 3. <i>iki pat</i> (3564) 4. <i>ligi pat</i> (328) 5. <i>nuo pat</i> (4844) 6. <i>prie pat</i> (1617) 7. <i>prieš pat</i> (2172) 8. <i>priešais pat</i> (5) 9. <i>ties pat</i> (19)
5. Prieveiksmiai / dalelytės	1. <i>taip pat</i> (17000*)
6. Įvardžiai	1. <i>bet kas</i> (9) 2. <i>bet katras, -a</i> 3. <i>bet koks, -ia</i> (102) 4. <i>bet kuris, -i</i> (114) 5. <i>daug kas</i> (3652) 6. <i>kai kas</i> (45) 7. <i>kai katras, -a</i> 8. <i>kai koks, -ia</i> 9. <i>kai kuris, -i</i> (335) 10. <i>kas kita</i> (394) 11. <i>kas ne kas</i> (67)

Kalbos dalis, kuria laikytinos MS	Pavyzdžiai
	12. <i>kas nors</i> (949) 13. <i>kažin kas</i> (205) 14. <i>kaži kas</i> (78) 15. <i>kažin katras, -a</i> 16. <i>kaži katras, -a</i> 17. <i>kažin koks, -ia</i> (185) 18. <i>kaži koks, -ia</i> (76) 19. <i>kažin kuris, -i</i> (9) 20. <i>kaži kuris, -i</i> (3) 21. <i>ko jau ko</i> (78) 22. <i>kuris ne kuris</i> (34) 23. <i>maža kas</i> (75) 24. <i>nė vienas, -a</i> (1807) 25. <i>šis tas</i> (339) 26. <i>šioks, -ia toks, -ia</i> (2400) 27. <i>tam tikras, -a</i> (1929) 28. <i>vienas, -a kitas, -a</i> (2234) 29. <i>visa kas</i> (37) 30. <i>visa kita</i> (1429) ¹³ 31. <i>visa tai</i> (1677) 32. <i>visa šitai</i> (25)
7. Jaustukai	1. <i>ačiū Dievui</i> (989) 2. <i>aš tau</i> (4) 3. <i>čia tai bent</i> (67) 4. <i>dėkui Dievui</i> (129) 5. <i>laba diena</i> (392) 6. <i>labą dieną</i> (57) 7. <i>labas rytas</i> (291) 8. <i>labą rytą</i> (43) 9. <i>labas vakaras</i> (157) 10. <i>labą vakarą</i> (20) 11. <i>labos nakties</i> (36) 12. <i>nieko sau</i> (153) 13. <i>po galais</i> (86) 14. <i>po paraliais</i> (31) 15. <i>po perkūnais</i> (27) 16. <i>po velnių</i> (328) 17. <i>po šimts perkūnų</i> (15) 18. <i>po šimts pyptių</i> (19) 19. <i>po šimts velnių</i> (15) 20. <i>po šimts</i> (35) 21. <i>šit kaip</i> (55) 22. <i>štai kaip</i> (56) 23. <i>štai tau</i> (167) 24. <i>tai bent</i> (126) 25. <i>tai tau</i> (28) 26. <i>tiek jau to</i> (87) 27. <i>tiek to</i> (135) 28. <i>Viešpatie tu mano</i> (8) 29. <i>viso geriausio</i> (52) 30. <i>viso gero</i> (105) 31. <i>viso labo</i> (7)
8. Išiktukai	1. <i>čiūčia liūlia</i> (40)
9. Būdvardžiai	1. <i>vienui vienas, -a</i> ¹⁴ (267) 2. <i>ne kažin kas</i> (53) 3. <i>ne kažin koks</i> (75)
10. Neidentifikuotos samplaikos	1. <i>be kita ko</i> (1906) 2. <i>be to</i> (17000*) 3. <i>ir kt.</i> (17000*) 4. <i>ir pan.</i> (12306) 5. <i>ir panašiai</i> (1527) 6. <i>ir t. t.**</i> (3211) 7. <i>ir t. t.**</i> (11269) 8. <i>ir taip toliau</i> (838) 9. <i>nepaisant to</i> (2255)

¹³ MS laikomos tik samplaikinių įvardžių *visa kita, visa tai, visa šitai* bevardės giminės formos, nes jos vartojamos daug dažniau nei atitinamos vyriškosios ar moteriškosios giminės formos, be to, šios formos yra stabilios, o tai ir yra vienas iš MS nustatymo kriterijų.

¹⁴ Samplaikinių būdvardžių galima rasti ir daugiau, pvz.: *mažut mažutėlis, gražut gražutėlis*. Tačiau tokių būdvardžių konstrukcijų gali būti labai įvairių, kai kurios vartojamos retai, todėl neįtrauktos į MS sąrašą.

DOI: 10.5755/j01.sal.1.14.43333