

Vartosenos modeli# analiz# mokomojoje leksikografijoje: #valgomasis tyrimas lietuvi# kalbos veiksm#od#i# pavyzd#iu

Kovalevskait#, Jolanta; Jancait#, Laima

Vartosenos modeli# analiz# mokomojoje leksikografijoje: #valgomasis tyrimas lietuvi# kalbos veiksm#od#i# pavyzd#iu

Taikomoji kalbotyra, vol. 12, 2019

Vilniaus Universitetas, Lituania

Disponible en: <https://www.redalyc.org/articulo.oa?id=694674844005>

DOI: <https://doi.org/10.15388/TK.2019.17235>



Esta obra está bajo una Licencia Creative Commons Atribución 4.0 Internacional.

Vartosenos modelių analizė mokomojoje leksikografijoje: žvalgomasis tyrimas lietuvių kalbos veiksmazodžių pavyzdžiu

Corpus Pattern Analysis for Learner Lexicography: A Pilot-
study of Lithuanian Verbs

Jolanta Kovalevskaitė jolanta.kovalevskaitė@vdu.lt

Vytauto Didžiojo universitetas, Lituania

Laima Jancaitė lai.jancaite@gmail.com

Vytauto Didžiojo universitetas, Lituania

Taikomoji kalbotyra, vol. 12, 2019

Vilniaus Universitetas, Lituania

Recepción: 14 Junio 2019

Publicación: 11 Octubre 2019

DOI: <https://doi.org/10.15388/TK.2019.17235>

Redalyc: <https://www.redalyc.org/articulo.oa?id=694674844005>

Summary: Šio straipsnio tikslas – pristatyti žvalgomąjį tyrimą, kuriame siekta išbandyti vartosenos modelių analizės metodą (angl. Corpus Pattern Analysis, Hanks 2004), pavyzdžiu paėmus du dažnus lietuvių kalbos veiksmazodžius iš pagrindinio žodyno. Šis metodas paremtas tekstynų inspiruotos (angl. corpus-driven) lingvistikos principais, kai į vartoseną žiūrima kaip į neatsiejamą reikšmės dalį, o vartosenoje rasti leksiniai ir gramatiniai dėsningumai yra pagrindas reikšmėms skirti. Susipažinęs ar supažindintas su šiais dėsningumais, besimokantis lietuvių kalbos galėtų geriau suvokti reikšmių skirtumus, o tai būtų svarbu ir suvokiant kalbą, ir ją produkuojant. Išdėsius vartosenos modelių analizės principus, straipsnyje aprašomas dviejų lietuvių kalbos veiksmazodžių tyrimas, parodant, kaip taikytas šis analizės metodas; taip pat atskleidžiama, kokios yra šiuo metodu surinktų duomenų naudojimo galimybės mokomuosiuose žodynuose. Be to, aptariami ir probleminiai aspektai: vartosenos modelių atpažinimas, aprašymas ir pateikimas mokomuosiuose žodynuose.

Keywords: lietuvių kalba, tekstynų lingvistika, vartosenos modelių analizė, vartosenos modelis, mokomieji žodynai.

Abstract: The aim of this paper is to present a pilot study which applies the framework of *Corpus Pattern Analysis* (CPA, Hanks 2004) to analyse some Lithuanian verbs which form part of the basic vocabulary. CPA draws on the insights of the corpus-driven language analysis and contextual and functional theory of meaning: a meaning of a word is associated with a specific lexical and grammatical environment, e.g. corpus patterns which represent an interconnection of lexical and grammatical elements. The CPA procedure is one of the several corpus-driven methods differing from the pattern grammar (Hunston, Francis 2000) in the way that CPA not only uses typical grammatical categories (e.g. word classes) but also introduces semantic values (e.g. semantic types) to distinguish different senses of a word. Semantic types are often the main separator of meanings, especially when two verb senses are associated with the same grammatical pattern. Concerning learners' dictionaries, CPA could provide learners with more detailed usage data, and this could lead to a better understanding of meaning differences, important both for language reception and language production. After introducing the CPA methodology, we present the CPA analysis of two Lithuanian verbs, namely, the inductive procedure followed to observe and define meaning-related patterning. We also discuss the problematic issues related to the application of CPA as identified in this study and mentioned by other CPA practitioners. First, observing and defining corpus patterns is a challenging task for lexicographers, especially because of the pattern / meaning division and generalizations related to semantic types. The second problematic aspect is automatization in the process

of pattern recognition. The third issue relates to foreign language learners as a target group: meaning-related patterning observed in the data has to be presented in a learner dictionary in a user-friendly way.

Keywords: Lithuanian language, corpus linguistics, corpus pattern analysis (CPA), pattern, learner dictionaries.

Summary: Straipsnio tikslas – pristatyti įvalgomąjį tyrimą, kuriame siekta išbandyti vartosenos modelių analizės metodą (angl. Corpus Pattern Analysis, CPA, Hanks 2004) lietuvių kalbai, pavyzdžiu paėmus du dažnus veiksmažodžius iš pagrindinio lietuvių kalbos žodyno. Vartosenos modelių analizė paremta tekstynų inspiruotos (angl. corpus-driven) lingvistikos principais, kai į vartoseną žiūrima kaip į neatsiejamą reikšmės dalį, o vartosenoje nustatyti leksiniai ir gramatiniai dėšningumai yra pagrindas reikšmėms skirti. Šis metodas yra vienas iš būdų tirti ir aprašyti konkrečių žodžių vartoseną, nuo kitų straipsnyje aptariamų metodų (pavyzdžiui, pattern grammar, Hunston ir Francis 2000) jis skiriasi tuo, kad modeliai apima ne tik gramatinį, bet ir leksinį junglumą, ypatingą dėmesį skiriant semantinėms grupėms, nes, kaip matyti iš tyrimo, būtent iš jų atsiskleidžia reikšmių skirtumai, kai reikšmės realizuojamos vienodais gramatiniais modeliais. Mokomojoje leksikografijoje vartosenos modelių analizė gali būti naudota tam, kad besimokantysis būtų geriau supažindintas su žodžio vartoseną, vadinasi, galėtų geriau suvokti reikšmių skirtumus, o tai būtų svarbu ir suvokiant kalbą, ir ją produkuojant. Išdėščius vartosenos modelių analizės principus, straipsnyje aprašomas dviejų lietuvių kalbos veiksmažodžių tyrimas, po kurio išryškėja ir šio metodo taikymo probleminių aspektų: pirma, vartosenos dėšningumų interpretavimas dažnai yra sudėtingas uždavinys, ypač dėl modelių / reikšmių skyrimo ir semantinių grupių apibrėžimo; antra, siekiant dirbti šiuo metodu su daugiau duomenų, reikalinga procesą automatizuoti ir nuspręsti, kaip vartosenos modelių analizės metodu sukauptus duomenis patogiai ir aiškiai pateikti mokomuosiuose žodynuose.

Keywords: lietuvių kalba, tekstynų lingvistika, vartosenos modelių analizė, vartosenos modelis, mokomieji žodynai.

1. Įvadas

Straipsnyje pristatomas vartosenos modelių analizės (angl. *Corpus Pattern Analysis*, CPA, Hanks 2004) metodas, paremtas tekstynų lingvistikos principais, kai į vartoseną žiūrima kaip į neatsiejamą reikšmės dalį, o vartosenoje rasti leksiniai ir gramatiniai dėšningumai yra pagrindas leksinio vieneto reikšmėms skirti. Vartosenos modeliuose yra ir leksinių, ir gramatinių elementų, nes leksika ir gramatika yra neatsiejamos. Tuo remiantis įrodoma viena svarbiausių tekstynų lingvistikos teorinių nuostatų apie leksikos ir gramatikos vienovę (Sinclair 1991). Leksiniai ir gramatiniai dėšningumai gali būti ištiriami analizuojant leksinius vienetus tekstynuose, tam tikru metodu aprašomi ir tam tikra forma pateikiami kalbos besimokantiems vartotojams: paminėtinas *Cobuild* žodyno (1987) projektas, *Cobuild* žodyno rengimo patirtimi paremta modelių gramatika¹ (angl. *Pattern Grammar*, Hunston ir Francis 2000), taip pat *FrameNet* metodas (Atkins ir kt. 2003). Patricko Hankso pasiūlytas vartosenos modelių analizės metodas yra vienas iš būdų supaprastinti ir formalizuoti Johno Sinclairio (1996, 1999) aprašytų leksinio vieneto vartosenos dėšningumų (išplėstinio reikšmės vieneto) tyrimą (DiMuccio-Faila ir Giacomini 2017: 441).

Taikant Hankso aprašytą vartosenos modelių analizę, vartoseną tiriama jau nustatant reikšmes, kitaip sakant, vartoseną reikalinga ne tik apibrėžtoms reikšmėms iliustruoti, bet ir jas atpažinti, todėl metodas

atitinka tekstynų inspiruotos (angl. *corpus-driven*) tekstynų lingvistikos kryptį (Sinclair 1991; lietuvių kalba Marcinkevičienė 2000). Šis metodas pagrįstas mintimi, kad skirtingos reikšmės susijusios su skirtingais vartosenos modeliais. Parodyti formos ir turinio sąsajas ypač aktualu mokomuosiuose žodynuose, nes tai galėtų padėti kitakalbiams geriau suvokti reikšmių skirtumus. Hanksas (2004: 91) teigia, kad paprastai iš žodynų informacijos neaišku, kaip atskirti vieną reikšmę nuo kitos, nes pateikiama per mažai informacijos apie reikšmei būdingą vartosenos kontekstą. Vartosenos informacija itin svarbi ir produkuojant kalbą: suvokiant, kokie leksiniai, gramatiniai elementai susiję su tam tikra reikšme, besimokančiajam gali būti aiškiau, kaip vartoti žodį šia reikšme. Pavyzdžiui, dėstytojai, naudojantys tekstynus kalbai mokytį, siūlo tokias užduotis, kurios skatina studentus aiškintis formos ir turinio sąsajas: tyrinėdami tam tikro žodžio vartosenos dėsningumus tekстыne, studentai mato, kaip žodžio reikšmė priklauso nuo konteksto (Tognini Bonelli 2000: 209).

Šis tyrimas kilo iš poreikio ieškoti būdų, kaip išsamiau aprašyti leksinio vieneto vartoseną lietuvių kalbos besimokantiems kitakalbiams skirtuose žodynuose. Metodai, kuriuos taikant aprašomi vartosenos dėsningumai, yra aktualūs siekiant mokomuosiuose žodynuose išsamiau atspindėti tam tikro leksinio vieneto vartoseną ir parodyti, kaip yra susijusios reikšmė ir forma.

Šio straipsnio tikslas – pristatyti pirmuosius bandymus taikyti vartosenos modelių analizę (angl. *Corpus Pattern Analysis*, Hanks 2004) lietuvių kalbai, pavyzdžiu paėmus du dažnus pagrindinio žodyno veiksmazodžius – *kilti* ir *žiūrėti*: šie veiksmazodžiai patenka tarp 500 dažniausių žodžių (Utka 2009) ir pateikiami kitakalbiams skirtose metodinėse priemonėse (Ramonienė, Pribušauskaitė, Vilkienė 2016a, 2016b). Žvalgomajam tyrimui naudotas bandomasis 1 mln. žodžių tekstynas, kurio tekstais siekta reprezentuoti kitakalbiams studentams aktualią vartoseną: pasirinkti lietuvių kalbos besimokantiesiems skirti tekstai iš vadovėlių ir gimtakalbių vartoseną atskleidžiantys tekstai iš periodikos. Vadovėlių tekstai (0,5 mln. žodžių) buvo paimti iš rengiamo *Mokomojo tekstyno*², periodikos tekstai (apie 0,5 mln. žodžių) buvo paimti iš *Morfologiškai anotuoto tekstyno* (MAT-o)³. Kad ir nedidelis, toks bandomasis 1 mln. žodžių tekstynas tiko žvalgomojo tyrimo tikslams, nes siekta analizuoti tipinę pasirinktų veiksmazodžių vartoseną, kuri paprastai jau išryškėja ir iš palyginti nedidelės imties.

Remiantis anglų kalbos mokomąja leksikografija (žodynais, skirtais svetimkalbiams anglų kalbos vartotojams), straipsnyje pirmiausia aptariama keletas metodų vartosenos dėsningumams (angl. *patterning*) analizuoti (žr. antrą dalį). Trečioje dalyje išdėstomi vartosenos modelių analizės (Hanks 2004) principai. Ketvirtoje dalyje pateikiama dviejų lietuvių kalbos veiksmazodžių vartosenos modelių analizė: parodoma, kaip šis metodas taikytas reikšmėms atskirti, taip pat aprašomos vartosenos modelių ir reikšmių sąsajos. Paskutinėje dalyje diskutuojama dėl probleminių aspektų, susijusių su metodo taikymu, ir aptariamos

šiuo metodu surinktų duomenų naudojimo galimybės mokomuosiuose žodynuose.

2. Vartosenos dėsningumų tyrimo ir aprašymo metodai

Kaip jau minėta įvade, iš tekstynų lingvistikos išryškėjo vartosenos svarba tiriant žodžio reikšmę, leksikos ir gramatikos vienovę (apie šias Sinclairio (1991) idėjas apžvalgą lietuvių kalba plačiau žr. Marcinkevičienė 2000). Sinclairio (1991) teigimu, reikšmės raiška yra ne pavieniai žodžiai, o žodžiai su artimiausiu kontekstu, vadinasi, kalba veikia pagal fraziškumo principą, kai reikšmė kuriama vartojant didelį kiekį gatavų ar pusiau gatavų frazių, todėl realiau kalbos sandaros vienetu laikyti ne žodį ar sakinį, bet tarpinį variantą – vadinamąjį išplėstinį reikšmės vienetą (Sinclair 1996: 114, cit. iš Marcinkevičienė 2000: 42). Išplėstiniam reikšmės vienetams tirti svarbūs keturi lygmenys: leksinis lygmuo – kolokacijos, gramatinis lygmuo – koligacijos, t. y. gramatinė kolokacijų raiška, semantinis lygmuo – bendras semantinis artimųjų ir tolimųjų kolokatų laukas, pragmatinis lygmuo – semantinė prozodija, t. y. vertinamieji konteksto elementai (Marcinkevičienė 2000: 43). Tiesa, Sinclairio aprašytas išplėstinio reikšmės vieneto tyrimas vis dar kelia praktinių iššūkių leksikografui: nors dabar yra nemažai galimybių leksinio vieneto vartosenos dėsningumus (jau ir didelės apimties tekstynuose) analizuoti automatizuotai ir taip paspartinti ir palengvinti leksikografo darbą, vis dėlto esminis darbas – tuos dėsningumus sisteminti, apibendrinti ir suprantamai pateikti vartotojui – išlieka sudėtingas.

Pirmajame tekstynu paremtame mokomajame žodyne *Cobuild* (išsamiai apie projektą – Moon 2007) pradėta teikti kitaip apibrėžtas reikšmes ir remtasi vartosenos modeliais, juos siejant su leksiniu vienetu ar leksinio vieneto reikšme. Tai, kaip teigia Elena Tognini-Bonelli (2000: 219), buvo naujas būdas, kaip gali būti nustatoma ir apibrėžiama reikšmė. Pagrindinis *Cobuild* sumanytojas Sinclairis kritikavo kitus, anksčiau pasirodžiusius, žodynus, nes, jo nuomone, šie neskirdavo reikiamo dėmesio sintaksei (Sinclair 1984: 4, cit. iš Moon 2007: 167). Būtent leksikos ir gramatikos sąveikoje kuriama reikšmė (Sinclair 2000, cit. iš Marcinkevičienė 2000: 45), taigi, tiriant šias kalbotyros sritis kartu, galima išvelgti modelius, būdingus tam tikroms reikšmėms. Kaip rašo Rosamunda Moon (2007: 168), rengiant *Cobuild* žodyną siekta, tiriant tekstyno duomenis, nustatyti vartosenos modelius ir susieti juos su daugiareikšmių žodžių reikšmėmis.

Leksikos ir gramatikos vienovė *Cobuild* projekte atsispindi ir iš vadinamųjų kontekstualizuotų apibrėžčių (angl. *full-sentence explanations*, plačiau Moon 2007, Moon 2018): kiekvieną apibrėžtį sudaro sakinytis, kuriame siekta atskleisti ir aprašomam leksiniam vienetui būdingą leksinį bei gramatinį junglumą. Beje, *Cobuild* projekte šita strategija buvo taikoma visais atvejais, t. y. nebūtinai tik tada, kai aprašomam vienetui buvo būdingas specifinis leksinis ar gramatinis junglumas. Rašant apibrėžtis tokiu principu, pačiose apibrėžtyse

naudotos tam tikros formulės: *if / when, you say... you mean* ir kt., palyginkime, pavyzdžiui, veiksmažodžio *develop*⁴ reikšmes:

1. *When something develops, it grows or changes over a period of time into a better, more advanced, or more complete form;*
2. *To develop an area of land means to build houses or factories on it;*
3. *If someone develops a new machine, they produce it by improving the original design;*
4. *To develop a characteristic, illness, or fault means to begin to have it.*

Kaip matyti, jau iš apibrėžties gaunama informacija, pavyzdžiui, su kokių kitų leksinių semantinių grupių žodžiais tą reikšmę galima vartoti. Kitaip sakant, žodyno naudotojas gauna vartosenos pavyzdžių jau iš apibrėžties, vadinasi, taip yra pateikiama svarbi informacija kalbai produkuoti, o būtent tokios informacijos trūksta daugumoje mokomųjų žodynų (Ch. Fillmore 2003: 268, cit. iš Atkins, Rundell 2008: 409)⁵. Rengiant *Cobuild* žodyną taikyta tekstynų inspiruota (angl. *corpus-driven*) metodika – vartosenos dėsningumai nustatyti tekстыne, tačiau, pasak Hankso (2008: 109), žodyninis straipsnis čia vis dėlto rašomas pagal reikšmes, o ne pagal vartosenos modelius (angl. *the entry structure... is meaning-driven rather than pattern-driven*): pavyzdžiui, *proposing an idea* ir *proposing a theory* yra laikomos atskiromis reikšmėmis, nors jų vartosenos modeliai (V n) vienodi, o tai būtų pagrindas šias reikšmes jungti.⁶

Cobuild žodyno gramatiniai modeliai buvo ta medžiaga, kurios pagrindu parašyta Susanos Hunston ir Gillos Francis *Pattern Grammar* (2000)⁷. Iš tekstynų tyrimų sukaupta medžiaga atskleidė modelių ir reikšmių sąsajas: „vartosenos modelis išryškėja tada, kai žodžiai tam tikrame junginyje vartojami palyginti dažnai, parenkami dėsningai ir perteikia tam tikrą reikšmę“⁸ (Hunston ir Francis 2000: 37). Pavyzdžiui, palyginus, kokie žodžiai dažnai kartojami su veiksmažodžiu *afraid*, nustatyti tokie dėsningumai: reikšmė „išsigąsti“ realizuojama bent keliais modeliais, o kita reikšmė – „atsiprašyti dėl kažko“ – realizuojama tik dviem modeliais: V *that* (*I'm afraid that I just wasn't ready*) ir V *to-inf* (*I'm afraid to say that...*). Vadinasi, pasak Hunston ir Francis (2000: 37), „žodžio modelis yra tokia struktūra, kuri siejasi su tuo žodžiu ir yra to žodžio reikšmės dalis“.⁹

Tiriant leksinių vienetų vartosenos modelius galima dirbti dviem būdais (Hanks 2008: 111): rašyti gramatiką arba rašyti žodyną. Hunston ir Francis (2000) darbas yra gramatika, parašyta pakeitus perspektyvą: nuo žodžio pereita prie modelio, kuris būdingas žodžių grupei (-ėms). Šis požiūris aiškinamas taip: „jeigu žodis gali būti vartojamas keliuose skirtinguose modeliuose, tai ir koks nors konkretus modelis gali būti būdingas įvairiems žodžiams“¹⁰ (Hunston ir Francis 2000: 43). Taigi modelių gramatikoje aprašomi ne žodžiai ir jų vartosenos modeliai, o vartosenos modeliai ir jų žodžiai. Pastebėjus, kad bendri modeliai yra būdingi panašios reikšmės žodžiams ir, remiantis šia įtvalga, parengta tam tikrų daiktavardžių, veiksmažodžių ir būdvardžių grupių modelių

analizė, pavyzdžiui, daiktavardžiams būdingi modeliai su būdvardžiu (adj N); modeliai su kitu daiktavardžiu (n N); modeliai su prielinksniais (from N, on N, to N...); taip pat modeliai, kuriuose priklausantys komponentai (bendratis, tam tikras jungtukas) eina po daiktavardžio (N to-Inf, N that...). Tįnoma, modelių analizė svarbi ir tarnybinėms kalbos dalims, pasiūminčioms funkcijų margumu, plg. Anglų kalbos pavyzdžius all, both, every, such: akivaizdu, kad, kai both perteikiamos skirtingos funkcijos, jam būdingi ir skirtingi modeliai (tr. Hunston 2004: 103).

Hunston (2004) mini, kad žodynuose, skirtuose besimokantiems svetimųjų kalbų, svarbu pateikti modelius, nes modeliai, o ne kalbos dalys suteikia daugiau aiškumo: pavyzdžiui, modeliai rodo, kad ta pačia kalbos dalimi laikomi žodžiai gali pasižymėti nevienoda vartoseną (plg., pavyzdžiui, Hunston (2004: 104) apie būdvardžiais laikomus *every* ir *interesting*). Vadinasi, jeigu kalbos dalį leksiniam vienetui priskirtume ištyrę jo vartosenos dėsningumus, modelis būtų patogus būdas parodyti studentui, kaip veikia gramatika tos kalbos, kurios jis mokosi (Hunston 2004: 103–104). Vis dėlto, tiriant vartosenos modelius, gramatikos kategorijos, pavyzdžiui, kalbos dalys, yra per abstrakčios ir jų ne visada pakanka (žr. modelių gramatikos kritiką Hunston 2004: 111; Hanks 2008: 115). Hunston (2004: 105) nuomone, siekiant aprašyti būdingą vartoseną, svarbu parodyti leksiką. Būtent leksika yra svarbi vartosenos modelių analizės dalis (Hanks 2004, 2008), čia leksinės semantinės grupės¹¹ yra priemonė tiksliau aprašyti kiekvieno žodžio vartosenos modelius, pavyzdžiui (Hanks 2008: 113):

Veiksmažodis	Modelių gramatikos modelis	Modelių žodyno (CPA) modelis	Komentaras
lead	V to n	[[Anything]] lead [[Human]] to [[Belief]]	Skiriasi nuo tokio modelio kaip [[Route]] lead {to [[Location]]}

Vartosenos modelių analizės paskirtis – atskirti reikšmes pagal leksinius ir gramatinius dėsningumus. Kitame skyriuje pristatomi pagrindiniai vartosenos modelių analizės principai.

3. Vartosenos modelių analizės principai

Vartosenos modelių analizės metodas pagrįstas mintimi, kad skirtingos to paties žodžio reikšmės atspindi skirtingus joms būdingus vartosenos modelius (DiMuccio-Failla ir Giacomini 2017: 438). Hanksas (2004: 90) teigia, kad, remiantis tekstynais, galima įrodyti, jog daugelis žodžių stipriai susiję tik su labai mažai tam tikrų tipinių vartosenos modelių. Teorija apie **tipinę ir netipinę vartoseną** (angl. *Theory of Norms and Exploitations*, Hanks 2004) yra svarbi metodo dalis. Netipine leksinio vieneto vartoseną laikytina neįprasta reikšmė, netaisyklinga ar neįprastai pavartota sintaksinė konstrukcija, reta perkeltinė vartoseną, pavyzdžiui,

naujos metaforos (El Maarouf 2013: 125), apibendrintai – vartosenos išimtys. Pasak Hankso (2004: 90), leksikografiškai turėtų aprašyti tipinę vartosena, o ne visus įmanomus vartosenos atvejus, kurie gali būti tiek tipiniai, tiek netipiniai, – atitinkamai, kalbos vartotojams bus lengviau suprasti, kaip žodis paprastai vartojamas, kokios jo vartosenos normos.

Kitas svarbus vartosenos modelių analizės principas – norint atpažinti modelius, kuriais remiantis galima atskirti vieną reikšmę nuo kitos, reikia atsižvelgti į leksinio vieneto **leksinę ir gramatinę aplinką**. Neužtenka tirti vieną kurį aspektą – leksiką arba gramatiką – reikia analizuoti juos abu, kaip leksikos ir gramatikos vienybę. Tiriant gramatinio junglumo ypatybes, nustatomi ir tam tikri tam tikroje pozicijoje dažnai vartojami žodžiai, taip apibendrinant tipinį leksinį junglumą, o atrasti leksiniai dėsningumai paskui apibendrinami semantinėmis grupėmis. Taikant vartosenos modelių analizę, pradedama nuo formos (struktūros) ir pirmiausia tiriamas gramatinis junglumas¹².

Modelis apibrėžiamas kaip sintaksiniais ryšiais susieta elementų struktūra, kur elementų **semantines grupes** (angl. *semantic types*, žymima [[...]]) sudaro tam tikri **leksiniai rinkiniai** (angl. *lexical sets*, žymima {...}). Aprašant veiksmažodžių modelius tiriamas valentingumas, nustatomi būtinieji valentiniai palydovai (argumentai) ir nurodoma, kokiomis semantinėmis grupėmis (ir jas sudarančiais leksiniais rinkiniais) realizuojami šie argumentai (Jezek ir Hanks 2010: 8).

Pavyzdys iš *Pattern Dictionary of English Verbs*:

[[Human]] **attend** [[Event]]

Lexical set [[Event]] = {meeting, conference, funeral, ceremony, course, school, seminar, dinner, reception, workshop, wedding, concert, premiere ...}

Atpažįstant vartosenos modelius, žodžiai analizuojami tekstyne, siekiant nustatyti leksines ir gramatinės žodžio vartosenos ypatybes. Tiriant gramatiką, kurios elementai įtraukiami į vartosenos modelius, svarbios kalbos dalys, formos, vyraujantys sintaksiniai dėsningumai (El Maarouf 2013: 122). Tiriant reikia atkreipti dėmesį į tam tikros kalbos gramatikos ypatybes, nors yra ir universalijų. Tarkim, atliekant anglų kalbos veiksmažodžių vartosenos modelių analizę nustatomos sintaksinės funkcijos subjektas, objektas, predikatas, atributas, aplinkybės, t. y. taikomas SPOCA (*subject, predicator, object, complement, adverbial*) principas (El Maarouf 2013: 134). Aprašant vartosenos modelį anglų kalba, subjektas būna veiksmažodžio kairėje, o objektas – dešinėje. Taip pat vartosenos modelis gali būti sudarytas iš vienos ar kelių sintaksinių konstrukcijų, tai priklauso nuo konkrečių atvejų. Aprašant vartosenos modelio gramatiką, reikia atkreipti dėmesį į tai, kad kai kurie sintaksės dalykai yra pasirinktini, o kai kurie – būtini (ir kai kurie būtini elementai gali būti numanomi). Juos būtina atskirti. Pavyzdžiui, kai kurios aplinkybės tam tikrame vartosenos modelyje yra būtinas elementas; kitos aplinkybės arba kitame vartosenos modelyje bus vartojamos pasirinktinai (El Maarouf 2013: 138). Vartosenos modelyje turėtų būti nustatomi būtini elementai. Bene sunkiausias uždavinys – tinkamai apibendrinti leksinius rinkinius ir priskirti jiems semantinių grupių pavadinimus.

Svarbu nustatyti tinkamą leksinių rinkinių apibendrinimo lygį (Hanks 2004: 88). Ketvirtoje šio straipsnio dalyje ši problema aptariama plačiau.

Modelių nustatymo etape naudojamos kategorijos ir struktūros, stengiantis kaip įmanoma tiksliau perteikti rastus dėsningumus. Hanksas (2007: 11) pateikia etapus, kaip turėtų būti atliekama vartosenos modelių analizė:

1) iš tekstyno medžiagos sudaromas analizuojamo leksinio vieneto konkordansas. Pavyzdžiui, naudojant 100 mln. žodžių tekstyną, rekomenduojama, kad konkordansą sudarytų bent 250 vartosenos pavyzdžių;

2) nustatomi tipiniai sintagminiai modeliai (angl. syntagmatic patterns). Čia svarbu atsižvelgti į leksiką ir gramatiką. Svarbu nustatyti tipines sintaksines struktūras, jas fiksuoti. Šiame tyrimo etape taip pat reikia atpažinti leksinius rinkinius (angl. lexical sets), t. y. kokie todėl dažnai vartojami šalia analizuojamo žodžio, suskirstyti juos į semantines grupes (angl. semantic types), t. y. apibendrinti semantiškai. Ontologijoje, kuria naudotasi kuriant anglų kalbos veiksmažodžių vartosenos modelių tydyną, yra pateiktas visų semantinių grupių sąrašas¹³;

3) tuomet kiekviena konkordanso eilutė priskiriama vienam iš nustatytų sintagminių modelių. Šiame tyrimo etape būtina atkreipti dėmesį, kuriose eilutėse atsispindi įprasta, tipinė vartoseną (angl. norms), o kuriose atsispindi neįprasta, reta, nauja vartoseną (angl. exploitations). Taip pat reikalinga atmesti tikrinius vardus, atvejus, kur žodis tik paminėtas, sunkiai interpretuojamus atvejus ir pan.;

4) tada, jeigu reikia, ištiriamos papildomos konkordanso imtys (pavyzdžiui, jeigu konkordanso eilučių trūksta, kad tinkamai būtų aprašytas vartosenos modelis, ar dėl kitų priežasčių);

5) nustatytiems vartosenos modeliams priskiriamos reikšmės ir parengiamas modelio reikšmės aiškinimas (angl. implicature), kurio esmė – paaiškinti žodžio reikšmę visu sakiniu, kur kartu atsispindėtų reikšmės kontekstas (iš pavyzdžių matyti, kad ir paaiškinimuose yra tos pačios semantinės grupės, žymimos dvigubais laužtiniais skliaustais, ir atsispindi sintaksinė struktūra). Palyginus su išplėstinio leksinio vieneto teorija (Sinclair 1996), vartosenos modeliuose retai nustatomi vertinamieji (pragmatikos) elementai, nors minima, kad pragmatika kai kuriais atvejais aprašoma reikšmių aiškinimų papildymuose (angl. secondary implicatures, Hanks 2008: 119).

Po visų vartosenos modelių analizės etapų parengiamas tam tikro veiksmažodžio modelių aprašas, pavyzdžiui, veiksmažodžiui *attend* nustatyti 3 vartosenos modeliai (žr. *Pattern Dictionary of English Verbs*)¹⁴:

1. *Pattern: [[Human]] attends ([[Event]] or [[Location]])*

Implicature: [[Human]] is present at [[Event]] or [[Location]]

Example: The 18-year-olds are both attending the optional competition class at the college.

2. *Pattern: [[Human]] attends to [[Anything]]*

Implicature: *[[Human]]* pays attention to and deals as necessary with *[[Anything]]*

Example: *the baby checks to see if mother is attending to the pointing gesture.*

3. *Pattern:* *[[Human 1]]* attends *[[Human 2]]*

Implicature: *[[Human 1]]* looks after *[[Human 2]]*

Example: *On April 15 a group of doctors threatened to strike against the curfew, which they alleged was preventing them from attending patients.*

Iš vartosenos modelių matyti sintaksinės ypatybės (subjektas, objektas); antras vartosenos modelis nuo pirmo ir trečio dar skiriasi ir gramatiniu junglumu (prielinksnis *to*). Be to, skiriasi ir vartosenos modelių leksika. Nors kaip subjektas visur pateikiama semantinė grupė *[[Human]]*, objekto leksinė raiška skiriasi: pirmame modelyje kaip objektas pateikiamos semantinės grupės *[[Event]]* ir *[[Location]]*, trečiame – *[[Human 2]]* (tokia pat semantinė grupė, kaip ir *to* paties vartosenos modelio subjekto). Semantinę grupę *[[Event]]* sudaro leksinis rinkinys {meeting, conference, demonstration, election, appointment...}; semantinę grupę *[[Location]]* sudaryta apibendrinus leksinį rinkinį {school, museum, camp, college...}. Antro modelio objekto pozicijoje nurodyta semantinė grupė *[[Anything]]* yra abstraktesnė už semantines grupes *[[Human]]*, *[[Event]]*, *[[Location]]*. Vartosenos modelių analizę taikantys leksikografai nurodo, kad tokios semantinės grupės, kaip *[[Anything]]* („bet kas“) arba *[[Animate]]* („gyvas padaras“), yra tokio bendro pobūdžio, tiek daug aprėpiančios, kad kartais iš jų nematyti veiksmažodžių kolokacinių polinkių. Tai gali būti problema mokomojoje leksikografijoje: kai vartosenos modelių analizę naudojama mokomajam žodynui sudaryti, leksinę aplinką svarbu atspindėti kuo tiksliau (tai mini ir DiMuccio-Failla ir Giacomini 2017).

Mokomuosiuose žodynuose vyraujanti tendencija yra atsakyti į klausimą „Kiek reikšmių turi kiekvienas žodis ir kaip aprašyti kiekvieną reikšmę?“, o kuriant vartosenos modelių žodyną keliamas klausimas – „Į kokius vartosenos modelius įeina žodis ir kokios yra tų modelių reikšmės?“ (Hanks 2008: 104). Vadinasi, taikant šį metodą, pirmiausia siekiama tirti pačią vartoseną, joje atrasti tai, kas yra tipiška, ir suteikti tai tipinei vartosenai modelių pavidalą, o tada susieti modelį su reikšme. Modelių ir reikšmių siejimo etape atsispindi vienas esminių metodo principų – reikšmės čia siejamos ne su atskirais veiksmažodžiais, o su veiksmažodžių modeliais (Jezek ir Hanks 2010: 8).

4. Vartosenos modelių analizės taikymas lietuvių kalbos veiksmažodžiams *žiūrėti* ir *kilti*

Šiame skyriuje paaiškinama, kaip vartosenos modelių metodas buvo išbandytas analizuojant lietuvių kalbos veiksmažodžius, ir parodoma dviejų lietuvių kalbos veiksmažodžių vartosenos modelių analizė¹⁵.

Kadangi tyrimas buvo žvalgomas, pasirinktiems veiksmažodžiams tirti atsitiktinės atrankos būdu suformuota imtis – apie 150 eilučių

iš kiekvieno veiksmažodžio konkordanso. Pirmiausia buvo išnagrinėti gramatinio junglumo dėsningumai, t. y. nustatyti gramatiniai modeliai¹⁶. Vėliau buvo išanalizuotas leksinis junglumas: nustatyti kolokatai, pagal reikšmę susiję kolokatai sujungti į leksinius rinkinius, o šie apibendrinti semantinėmis grupėmis¹⁷ – taip sudaryti vartosenos modeliai, kuriuose matyti tipinė leksinė ir gramatinė žodžio aplinka. Tipinė vartoseną nuo netipinės vartosenos atskirta pagal dažnumą. Jeigu tam tikros sintaksinės struktūros pavartojimo atvejų skaičius labai mažas (čia – mažiau nei 5 kartai), laikyta, kad struktūra nėra tipinė ir neaktuali apibendrinant semantines grupes. Pavyzdžiui, jeigu semantinei grupei [žmogus] priklausantys žodžiai tam tikroje pozicijoje pasitaikė mažiau nei 5 kartus, ši semantinė grupė neįtraukta į vartosenos modelį, nes vartoseną laikyta netipine. Pagal dažnumą atskirti ir kolokatai: jeigu tam tikrame kontekste tam tikras žodis pasitaikė mažiau nei 4 kartus, jis nelaikytas kolokatu ir neįtrauktas į modelį, nes pavartotas per retai. Taigi žodžius nuspręsta laikyti kolokatais, jei jie pasikartojo bent po 4 kartus tam tikroje pozicijoje. Nors žvalgomajame tyrime to nedaryta, bet, taikant šį metodą, būtų svarbu nustatant kolokatus remtis ne tik paprastu dažnumu tekstyne, bet ir statistiniais įverčiais.

Žiūrėti ir kilti – dažni lietuvių kalbos veiksmažodžiai, *Mokomajame lietuvių kalbos žodyne* aprašyti kaip daugiareikšmiai. Po vartosenos modelių analizės tekstyne šiems veiksmažodžiams nustatyti keli modeliai, atitinkamai ir kelios reikšmės (žr. 1 ir 2 lentelę). Pagal vartosenos modelių analizės metodą nustatytus modelius susiejus su reikšmėmis, kartu su modeliu pateikiamas ir reikšmės aiškinimas. Modelių seka atspindi jų dažnumą tekstyne. Jei prie vartosenos modelio pateiktas vienas skaitmuo, vadinasi, šis vartosenos modelis susietas su viena reikšme. Jei vartosenos modelis skaidomas (1.1., 1.2.), vadinasi, du modeliai susieti su ta pačia (viena) reikšme. Analizuojant veiksmažodžius *žiūrėti* ir *kilti*, nustatytos trys reikšmės, kurios susietos su dviem vartosenos modeliais (klausias, ar visais atvejais reikėtų skirti du modelius, aptariamą skyriaus pabaigoje). Pavyzdžiui, ir tuomet, kai veiksmažodis *žiūrėti* prisijungia objektą su prielinksniu (1.1. modelis), ir tuomet, kai prisijungia šalutinį sakinį (1.2. modelis), perteikiama ta pati reikšmė (žr. *Aiškinimą*).

Kaip jau minėta ankstesniame skyriuje, vartosenos modeliams aprašyti yra taikoma tam tikra struktūra – atitinkamai vaizduojami modelio gramatiniai ir leksiniai komponentai: žodžio formos, kalbos dalys, semantinės grupės, kolokatai. Hanksas (2008: 117) nurodo, kad vartosenos modelį gali sudaryti frazės, semantiškai susijusių žodžių grupės (pavyzdžiui, semantinę grupę [žmogus] sudarantys daiktavardžiai), tam tikri žodžiai (ypač pastoviųjų junginių atveju¹⁸). Žodžio forma yra svarbi informacija ir apie patį analizuojamą žodį, ir apie kitus modelio komponentus, ypač objektus. Pavyzdžiui, iš vartosenos matyti, kad žodis *kilti* 2 ir 4 modelyje vartojamas veikiamosios rūšies būtojo kartinio laiko dalyvio forma, o 1.1., 1.2. ir 3 modeliams tai nėra būdinga. Vadinasi, šią vartosenos ypatybę svarbu atspindėti būtent 2 ir 4 modelyje. Tais atvejais, kai pirmenybė kuriai nors formai (-oms) vartosenoje neišryškėjo, modeliuose parodytos neutralios formos – veiksmažodis

pateiktas tiesiogine nuosaka, III a. es. I. (plg. veiksmo žodžio žiūrėti modelius). Be tiriamojo žodžio, į modelius gali prireikti įtraukti ir kitų žodžių, veiksmo žodžių atveju tai gali būti prielinksniai (pavyzdžiui, žiūrėti 2.1. modelyje – prielinksnis į, kilti 2 modelyje – prielinksnis iš). Nurodant semantines grupes ir kolokatus, kartu pateikiama ir gramatinė informacija: pavyzdžiui, 3 žiūrėti modelyje objektas pateiktas galininko linksniu – [reginį | televizijos programą], nes šiuo atveju su žodžiu žiūrėti vartojamas daiktavardis galininko linksniu.

Daugumai su skirtingomis reikšmėmis susietų veiksmo žodžių modelių būdingas skirtingas gramatinis junglumas. Iš žiūrėti 1.1. modelio matyti, kad šiuo atveju žodis žiūrėti prisijungia prielinksniui. (netiesioginį objektą), o iš 3 modelio matyti, kad žodis visada prisijungia galininką (tiesioginį objektą). Abu vartosenos modeliai susieti su skirtingomis reikšmėmis. Vis dėlto dažnai būtent pagal semantines grupes galima atskirti modelius, kuriems būdingas vienodas gramatinis junglumas (plg. Hanks 2008: 116). Kartais, kai gramatinis junglumas sutampa, atskirti reikšmes galima remiantis leksiniu junglumu: plg. kilti 1.1. ir 3 modelį, čia matyti vienodas gramatinis junglumas, bet 3 modelio atveju žodis kilti prisijungia fizinį, t. y. apčiuopiamą, objektą; 1.1. modelio atveju prisijungiami abstraktūs objektai. Ispanų kalbos žodyno kūrėjai, atlikę sangražinių veiksmo žodžių vartosenos modelių analizę, nustatė, kad kartais gramatinis junglumas gali atspindėti daugiau nei vieną reikšmę – taigi, daroma išvada, kad gramatinio aprašo neužtenka ir labai svarbi tiek leksika, tiek gramatika (Renau ir Battaner 2012: 353). Kiekvieno žodžio atvejis yra individualus ir reikia atskirai vertinti, kuriais tam tikro žodžio elementais, labiau leksine ar labiau gramatine aplinka, reikia remtis atskiriant vieną vartosenos modelį nuo kito ir kartu vieną reikšmę nuo kitos.

1 lentelė

Veiksmo žodžio žiūrėti vartosenos modeliai

1.1. Modelis: [Žmogus] žiūri į [žmogų] [materialinį objektą] Aiškinimas: [Žmogus] stebi, yra nukreipęs akis į [žmogų] [materialinį objektą] Pavyzdžiai: Į veidrodį žiūrėti negalėjau – man tai būdavo trauma. Lopšinės dukterį nuramina, ji žiūri į mane dainuojančią ir šypsosi.
1.2. Modelis: [Žmogus] žiūri + ŠALUTINIS SAKINYS Aiškinimas: [Žmogus] stebi akimis + ŠALUTINIS SAKINYS Pavyzdys: Atbėgusi mūsų „auklė“ ilgai dairėsi, žiūrėjo, kur mes.
2.1. Modelis: [Žmogus] žiūri į [žmogų] [abstraktą] KAIP Aiškinimas: [Žmogus] vertina [žmogų] [abstraktą] KAIP KAIP: {neigiamai, entuziastingai, abejingai, atsargiai, su ironija...} Pavyzdžiai: Pradžioje gyventojai gana skeptiškai žiūrėjo į savo naujuosius kaimynus. Liberaldemokratai į „darbiečių“ pasiūlymą žiūri gana entuziastingai.
2.2. Modelis: [Žmogus] žiūri į [žmogų] [abstraktą] kaip į [objektą] Aiškinimas: [Žmogus] vertina [žmogų] [abstraktą] kaip [objektą] Pavyzdžiai: Į sveikatą žiūriu kaip į didžiausią vertybę... viename mieste šalies šiaurėje į mus žiūrėjo kaip į muziejinius eksponatus.
3. Modelis: [Žmogus] žiūri [reginį] [televizijos programą] Su kolokatais: [Žmogus] žiūri (televizorių, filmą...) Aiškinimas: [Žmogus] stebi dėl malonumo [reginį] [televizijos programą] Pavyzdžiai: Mano vyras turi silnybę – žiūrėti erotinius filmus. Jums atrodo, kad žiūrite spektaklį iš pirmojo rango ložės.

Dirbant vartosenos modelių analizės metodu svarbu ne tik nustatyti, kad, vartojant tam tikrą žodį tam tikra reikšme, reikia objekto, išreikšto, pavyzdžiui, galininku, bet taip pat parodyti, kokie žodžiai dažniausiai vartojami toje objekto pozicijoje ir, jeigu įmanoma, sugrupuoti juos pagal reikšmę įsemantines grupes. Semantinės grupės nurodomos laužtiniuose skliaustuose „[]“, jei daugiau negu viena, atskiriama „|“, plg. [Žmogus] žiūri į [žmogų] | [abstraktą] kaip į [objektą]. Aprašant veiksmažodžio *žiūrėti* 3 vartosenos modelį, semantinė grupė [reginys] buvo sudaryta iš leksinio rinkinio {kolekcija, spektaklis, kūriniai...}; aprašant veiksmažodžio *žiūrėti* 1.1. vartosenos modelį, semantinė grupė [materialinis objektas] sudaryta apibendrinus leksinį rinkinį {veidrodis, lėkštė, vonia...}. Sisteminimo sunkumų kyla siekiant apibendrinti abstraktesnius reiškinius: tarkim, aprašant veiksmažodžio *žiūrėti* 2.1. ir 2.2. modelį, leksinis rinkinys {nužudymas, sveikata, liga, pasiūlymas, idėja, pastangos...} apibendrintas semantine grupe [abstraktas], o veiksmažodžio *kilti* 1 modelio semantinė grupė [nesutarimai] sudaryta iš leksinio rinkinio {konfliktinės situacijos, ginčai, susirėmimai...}. Veiksmažodžio *kilti* 2 modelio semantinė grupė [fizinis objektas] pasirinkta apibendrinus leksinį rinkinį {žmogus, lėktuvas, gervė...}, kuriame matyti palyginti skirtingų žodžių, tad šiuo atveju visi laikyti fiziniiais objektais, o atskiros semantinės grupės [žmogus], [gyvūnas] neskirtos dėl to, kad šių semantinių grupių elementų vartosenoje pasitaikė mažiau nei 5 atvejai, taigi nepakankamai, kad šiame tyrime būtų priskirta tipinei vartosenai.

Kaip matyti iš 1 ir 2 lentelėse pateiktų modelių, kartais semantinių grupių tam tikroje pozicijoje gali būti nurodytos kelios (pavyzdžiui, *žiūrėti* 1.1., 2.1. modeliai). Atskiriant tokias semantines grupes kaip [žmogus] ir [abstraktas], kelios semantinės grupės yra informatyvios, nes tarp jų yra aiškus skirtumas, tačiau kitais atvejais (plg. *kilti* 1.1., 1.2. modeliai), kai greta pateikiama daug panašių semantinių grupių, semantinių grupių grandinėle panėšėja į kolokatų grandinėle, pavyzdžiui: [jausmas | problema | grėsmė | mintis | diskusija | skandalas | nesutarimas | neramumai | fizinis negalavimas] (žr. *kilti* 1.1. modelis). Iš dalies tokių sprendimų atsiranda dėl to, kad dėl kolokatų gausos ir semantinės įvairovės sudėtinga juos apibendrinti keliomis didesnėmis semantinėmis grupėmis. Kartais šių vartosenos situaciją galima parodyti tiesiog pateikiant dažniausius kolokatus (kolokatai nurodomi skliaustuose „{ }“). Tarkim, prie veiksmažodžio *žiūrėti* 3 vartosenos modelio viena iš semantinių grupių nurodyta [televizijos programa], bet ji ne visai tiksliai apibendrina žodžius iš leksinio rinkinio {filmas, kanalas, laida...}. Kaip pažymi Hanksas (2004: 91), kartais naudingiau nurodyti patį leksinį rinkinį, jo neapibendrinant semantine grupe: jis pateikia veiksmažodžio *toast* (liet. skrudinti) pavyzdį daiktavardžius, vartojamus su šiuo žodžiu, galima apibendrinti vienu žodžiu *food* (liet. maistas); tačiau tai būtų netikslinga, nes yra daug maisto, kuris nėra skrudinamas. Aptarti pavyzdžiai atspindi vieną iš problemiškesnių semantinių grupių nustatymo aspektų: „vieną ar kitą to paties veiksmažodžio argumentų poziciją užpildančius paradigminius žodžių rinkinius ne visada galima

apibendrinti tiksliais semantinėmis grupėmis, nes tuose rinkiniuose dažnai pasitaiko skirtingų semantinių grupių žodžių“¹⁹ (Jezek ir Hanks 2010: 8).

2 lentelė
Veiksmazodžio kilti vartosenos modeliai

<p>1.1. <i>Modelis: Kyla</i> [jausmas problema grėsmė mintis diskusija skandalas nesutarimas neramumai fizinis negalavimas] <i>Su kolokatais: kyla</i> {klausimas, mintis, pavojus} <i>Aiškinimas:</i> Atsiranda [jausmas problema grėsmė mintis diskusija skandalas nesutarimas neramumai fizinis negalavimas] <i>Pavyzdžiai:</i> Gali kilti pagunda ieškoti gyvenimo meilės. Kilus problemai, visi svarstome, kaip ją išspręsti. ... svajotojams nuo šio pusbrangio akmens kyla pavojus susirgti šizofrenija. ... artėjant rinkimams į Seimą vėl pagrįstai kyla klausimas: kaip ir kokį kandidatą rinkti ... tarp Vytauto ir Vokiečių ordino pasiuntinių kilo dėmesio vertas disputas. ... teismo teisėjas Valdas Kairys, kilus skandalui dėl bylų wilkinimo, atsistatydino. Aš matau, kad kyla labai produktyvus ginčas. ... pareigūnai surado pagrobėjus ir juos apsupo, kilo susišaudymas. Kartais gali kilti infekcija arba imti kraujuoti.</p>
<p>1.2. <i>Modelis: Kyla</i> [problemų minčių nesutarimų] <i>Su kolokatu: kyla</i> {problemų} <i>Aiškinimas:</i> Atsiranda [problemų minčių nesutarimų] <i>Pavyzdžiai:</i> Rimtesnių sunkumų kyla apylinkės laukuose. Kils įvairių kūrybinių sumanymų... .. kils nemažai ginčų su verslo partneriais.</p>
<p>2. <i>Modelis:</i> [žmogus] kilęs iš [teritorijos žmonių] <i>Aiškinimas:</i> [žmogus] gavęs kilmę iš [teritorijos žmonių] <i>Pavyzdžiai:</i> Danielis Dei-Liuisas (Daniel Day-Lewis) kilęs iš garsios Anglijoje šeimos. Pirmosios knygos autorius buvo kilęs iš žemaičių tarmės ploto...</p>
<p>3. <i>Modelis: Kyla</i> [fizinis objektas] <i>Aiškinimas:</i> Juda į viršų [fizinis objektas]</p>
<p><i>Pavyzdys:</i> ... kas minutę kur nors leidžiasi ar kyla lėktuvas.</p>
<p>4. <i>Modelis:</i> [žodžiai] kilę iš [šaltinio] <i>Aiškinimas:</i> [žodžiai] atsiradę, gavę kilmę iš [šaltinio] <i>Pavyzdys:</i> Galvojama, kad Europoje populiariausia pavardė kilusi iš kalvio amato pavadinimo.</p>

Tiriant tam tikro žodžio vartosenos modelius, išryškėja atvejų, kai tam tikri elementai gali būti praleidžiami, numanomi. Pavyzdžiui, veiksmazodžio *žiūrėti* 1.1. modelio komponentas (į [žmogų | materialų objektą]) yra pateikiamas skliaustuose, taip parodant, kad šiuo modeliu realizuojamos reikšmės objektas dažnai numanomas arba išreiškiamas kitaip. Šiuo atveju *žiūrėti* 1.1. modelis gali prisijungti prielinksnių . ir galininko linksnio daiktavardį, vis dėlto iš konkordanso pastebėta, kad objektas nebūtinai vartojamas su prielinksniu . – kartais jis gali būti išreiškiamas kitaip, numanomas arba nesukonkretintas, pavyzdžiui: *Užvertęs galvą, žiūrėjo aukštyn...*(objektas yra tai, kas matyti užvertus galvą į viršų, *aukštyn*). Šiuo atveju objektas neišreikštas prielinksniu . ir galininku, todėl žymint pirmąjį vartosenos modelį objektas su prielinksniu . nurodytas paprastuose skliaustuose: (į [žmogų | materialų objektą]). Tai reiškia, kad vartosenoje ne visada pasitaiko prielinksnis .:

analizuotame tekстыne daugiau nei 5 kartus objektas buvo numanomas arba išreiškiamas kitaip, todėl modelyje ši vartosenos ypatybė ir buvo atspindėta.

Apibendrinant galima teigti, kad, remiantis vartosenos modeliais, galima atskirti reikšmes vieną nuo kitos. Vis dėlto reikia turėti omenyje, kad analizuoti veiksmažodžiai turi savarankišką reikšmę. Neanalizuoti pagalbiniai ir pusiau pagalbiniai veiksmažodžiai (tokie, kaip *likti, tapti, imti* ir pan.), kurių leksinė reikšmė nėra savarankiška, dažnai nublukusi (Ramonienė, Pribušauskaitė 2003: 27), tokių veiksmažodžių vartosenos modelius gali būti aktualu tirti dėl to, kad šių veiksmažodžių reikšmė labai priklausoma nuo konteksto, kaip ir apibendrinamųjų daiktavardžių (pavyzdžiui, *dalykas, reikalas, esmė* ir kt.), kurių reikšmė bendra ir neapibrėžta (Drūlienė 2014). Tognini-Bonelli (2000: 229) pabrėžia, kad vadovėliuose, skirtuose negimtakalbiams, nedaug teminima apie desemantizacijos (reikšmės išblukimo) reiškinį, kai leksinis vienetas ilgesniame junginyje įgyja naują reikšmę ar funkciją. Tačiau, mokant kalbų, desemantizacija tam tikrų leksinių vienetų pavyzdžiu turėtų būti aptariama dažniau, nes nuo to priklauso leksinio vieneto funkcija, kuri susijusi su tam tikrais vartosenos dėsningumais. Be to, į desemantizaciją yra linkę būtent dažnesni žodžiai (Sinclair 1997, cituojama iš Tognini Bonelli 2000: 229), t. y. kaip tik tie, kurie sudaro svarbią pagrindinės leksikos dalį.

5. Vartosenos modelių analizės naudojimo galimybės ir probleminiai aspektai

Vartosenos modelių analizė jau taikoma mokomųjų žodynų aprašams kurti (plg. italų kalbos (DiMuccio-Failla ir Giacomini 2017), ispanų kalbos (Renau ir Battaner 2012) žodynus). Šiame skyriuje pirmiausia paminėsime vartosenos modelių analizės metodu surinktų duomenų naudojimo galimybes mokomuosiuose žodynuose, o antrojoje dalyje aptarsime probleminius aspektus, išryškėjusius taikant vartosenos modelių analizę lietuvių kalbos veiksmažodžiams.

5.1. Vartosenos modelių analizės metodu surinktų duomenų naudojimo galimybės

Taikant vartosenos modelių analizę, galima lengviau atskirti daugiareikšmių žodžių reikšmes: tiriant vartoseną, galima nustatyti, kokie modeliai būdingi vienai, o kokie – kitai daugiareikšmio žodžio reikšmei. Susiedamas reikšmę su vartosenos modeliu (-iais), leksikografas gali aiškiau parodyti reikšmę, o, matant vartosenos modelį, besimokančiajam lengviau suprasti, kokia reikšmė žodis pavartotas arba kaip jam pačiam pavartoti tą žodį konkrečia reikšme. Pavyzdžiui, iš ankstesnėje dalyje aptarto veiksmažodžio *žiūrėti* vartosenos modelių 2.1., 2.2. galima atskirti reikšmę „vertinti, turėti nuomonę“ nuo 1.1., 1.2. modeliais perteikiamos pagrindinės šio veiksmažodžio reikšmės.

Vartosenos modeliais apibendrinta informacija gali būti naudinga leksikografui ir rašant apibrėžtis, t. y. aiškinant reikšmes, nes, apibendrinant leksinius dėsningumus semantinėmis grupėmis, modeliuose fiksuojama semantinė informacija. Apibrėžtyje galima arba pateikti semantines grupes, arba paminėti konkrečius žodžius iš tų grupių leksinių rinkinių. Pavyzdžiui, apibūdinant žodžio *žiūrėti* reikšmę, „vertinti, turėti nuomonę“, galima formuluoti: *kai žmogus žiūri į žmogų, pasiūlymą, idėją, pastangas... palankiai, jis vertina tą dalyką gerai*. Taip jau apibrėžtyje galima pateikti ne tik informaciją, ką žodis reiškia, bet ir vartosenos informaciją, šiuo atveju – leksinį junglumą parodant būdingus kolokatus. Toks principas artimas antrame skyriuje paminėtai *Cobuild* žodyno kontekstualizuotų apibrėžčių strategijai.

Mokant kalbą, dažnai diskutuojama apie gerų pavyzdžių svarbą, pabrėžiama, kad dažniau būtent keli geri pavyzdžiai, o ne apibrėžtys padeda suprasti, kaip naujas žodis turi ar gali būti vartojamas (Frankenberg-Garcia 2015). Atrenkant gerus pavyzdžius, siūloma orientuotis į tokius, kurie būtų įprasti, informatyvūs ir suprantami (Atkins ir Rundell 2008). Vartosenos modeliuose susisteminta informacija apie vartosenos dėsningumus yra geras pagrindas įprastiems pavyzdžiams atrinkti. Iš vartosenos modelių matyti, kokią informaciją svarbiausia pateikti besimokančiajam: iš modelių matyti tam tikros žodžio vartosenos dėsningumai, o besimokančiajam naudinga į visa tai atkreipti dėmesį, kad galėtų lengviau suprasti ir produkuoti kalbą. Pavyzdžiui, atrenkant pavyzdžius *kilti* vartosenos modeliui². [*žmogus*] *kilęs iš [teritorijos | žmonių]* aišku, kad svarbūs tokie sakiniai, kuriuose būtų atspindėti žodžiai, nurodantys šeimą, giminę, nes tokios yra modelyje užfiksuotos semantinės grupės, pavyzdžiui:

*Danielis Dei-Liuisas (Daniel Day-Lewis) kilęs iš garsios Anglijoje šeimos.
Lotynų Amerikoje, iš kurios kilęs Chulijas, žmonės visai kitaip bendrauja.
Tikrai esam kilę iš vokiečių (...).*

Dirbant šiuo metodu, atrenkamus pavyzdžius būtų galima susieti su konkrečiais vartosenos modeliais, o būtent tai ir yra rekomenduojama, kai rašoma apie mokomųjų žodynų kokybę: pasak Anos Frankenberg-Garcia (2014: 141), žodyno straipsnyje besimokantysis turėtų matyti pakankamai tam tikro žodžio vartosenos dėsningumų, kad galėtų juos suprasti ir tokiu būdu geriau juos įsisavinti.

Apibendrinant galima teigti, kad vartosenos modeliai mokomajame žodyne naudingi dėl kelių priežasčių: juos naudojant galima atskirti reikšmes vieną nuo kitos (tokiu atveju besimokantysis geriau suprastų kalbą) ir suteikti informacijos apie vartoseną (tokiu atveju besimokančiajam būtų lengviau produkuoti kalbą). Kaip jau minėta trečiame skyriuje, taikant vartosenos modelių analizės metodą, jau nustatant reikšmes, reikėtų remtis modeliais: pradėti vartosenos dėsningumų tyrimą atpažįstant modelius ir žiūrėti, kokia reikšmė perteikiama modeliu, t. y. koks turinys perteikiamas kokia forma. Kita vertus, vartosenos modeliai gali būti naudoti ir vėlesniuose etapuose – rašant apibrėžtį, atrenkant pavyzdžius.

5.2. Vartosenos modelių analizės probleminiai aspektai

Apie probleminius aspektus jau buvo užsiminta trečiame skyriuje, čia jie aprašyti plačiau ir papildyti susitelkiant į vartosenos modelių

- 1) atpažinimą (vartosenos duomenų interpretavimą);
- 2) aprašymą;
- 3) naudojimą (mokomojo žodyno naudotojų perspektyva).

Kaip ir kiekviename tokio pobūdžio tyrime, svarbus yra tekstyno reprezentatyvumo ir dydžio klausimas. Pavyzdžiui, tiriant veiksmažodžio *kilti* vartoseną, šiame bandomajame 1 mln. tekстыne nebuvo išskirtas vartosenos modelis *kyla kainos*: pasitaikė mažiau nei 5 atvejai, vadinasi, jų nepakako atskiram modeliui išskirti. Vis dėlto jeigu šį veiksmažodį analizuotume dideliame bendrajame tekстыne (pavyzdžiui, *Dabartinės lietuvių kalbos tekстыne*), šis modelis būtų dažnesnis. Nors tipinė vartoseną paprastai jau išryškėja ir nedideliame tekстыne, kaip rodo aptariamas pavyzdys, tipinės vartosenos vaizdas mažame tekстыne ne visada gali būti išsamus, nes dalį tam tikro žodžio tipinės vartosenos gali būti sunkiau nustatyti dėl per mažo (nepakankamo) dažnumo. Kita vertus, jeigu orientuojamasi į kitakalbius ir naudojamas tekстыnas sudarytas būtent iš A1–B2 lygio mokiniams aktualių tekstų, galima sakyti, kad, nors ir mažesnis, toks tekстыnas yra pakankamas tai įprastai vartosenos daliai atskleisti, kuri sudaro pagrindinį mokinių žodyną.

Analizuojant didelį konkordansą, tyrimas gali užimti daug laiko. Šiame tyrime vartosenos modelių nustatymas buvo minimaliai automatizuotas: naudojantis tekstyno analizės įranga *AntConc* (Anthony 2018) nustatyti dažniausi kolokatai. Tačiau, dirbant šiuo metodu su daugiau duomenų, reikėtų galvoti, kaip kuo daugiau etapų automatizuoti: pavyzdžiui, minėtame italų mokomojo žodyno projekte DiMuccio-Failla ir Giacomini (2017) naudojama tekstyno analizės sistema *SketchEngine*, kad leksikografas gautų pirminę automatiškai apdorotą medžiagą, kokiomis leksinio ir gramatinio junglumo ypatybėmis pasižymi analizuojamas žodis ²⁰. Sudarant didesnius bendruosius žodynus, taikomi metodai, kuriais automatiškai surenkama informacija iš tekstyno apie tam tikro vieneto (žodžio, frazės) leksinę ir gramatinę aplinką (pavyzdžiui, Storjohann 2010), vadinasi, leksikografas gauna atrinktą pirminę medžiagą, kurią analizuoja ir redaguoja toliau, jeigu prireikia, pasitikrindamas konkordanse. Taigi analizuojant dėsningumus, pirminė automatinė duomenų atranka yra reikalingas žingsnis, apie kurį jau minėjo ir modelių gramatikos metodu dirbę lingvistai (Hunston 2004: 112).

Nors, taikant vartosenos modelių analizę, reikšmės skiriamos remiantis ne intuicija, o tekstynais (El Maarouf 2013: 119), vis dėlto introspekcijos vaidmuo lieka svarbus, todėl, kaip teigia tekstynų lingvistai, analizuojant tekstynus svarbu remtis ir intuicija, ir introspekcija, ir duomenų analize (Moon 1987, Sinclair 2004, cit. iš Dimuccio-Failla ir Giacomini 2017: 438). Tačiau vartosenos dėsningumų interpretavimas yra sudėtingas uždavinys. Apibendrinama modelių gramatikoje taikomus

principus, dažnumą Hunston (2004) nurodo kaip vieną iš modelių atpažinimo kriterijų. Pagal dažnumą galima atpažinti, kurios sąsajos tarp analizuojamo žodžio ir greta jo vartojamų kategorijų, tikėtina, yra svarbios, tačiau vis tiek palyginti sudėtinga interpretuoti, kas laikytina dėsningumu ir kaip kas turėtų būti siejama (Hunston 2004: 106-107). Pavyzdžiui, dažnai sudėtinga nustatyti, kurie elementai laikytini modelio dalimi, o kurie nepriskirtini modeliui, nors dažnai pasikartoja šalia analizuojamo žodžio. Taigi, pasak Hunston (2004: 110), dažnis neturėtų būti laikomas absoliučiu modelių komponentų (ir paties modelio) nustatymo kriterijumi: „tai, kas dažna, nustatant modelius neabejotinai yra svarbu, bet ne visos kolokacijos ir koligacijos pasižymi dažna vartosena, vadinasi, net ir tai, kas tekstyne pavartota rečiau, gali būti tipinė vartosena.“ Modelių atpažinimo klausimas sietinas su modelio ribų klausimu: taip, žodžio vartosenoje išryškėja dėsningumai, bet ne visų žodžių atveju vartosenos modelio ribos yra ryškios (Jezek ir Hanks 2010: 8).

Kitas dažnas klausimas dirbant su modeliais (apie tai ne kartą mini ir Hanksas, žr. Hanks 2004, 2008) – sprendimas, kiek modelių reikia smulkinti ar stambinti. Pavyzdžiui, dabar *žiūrėti* 1.2. modelis nuo 1.1. modelio skiriasi šalutiniu sakiniu; galbūt būtų galima parodyti šalutinį sakinį kaip variantą (statmenu brūkšniu), bet taip modelis pasidarytų labai sudėtingas skaityti. *Kilti* atveju 1.2. modelis išskirtas tik dėl to, kad kaip įprasta vartosena nustatytas subjekto vartojimas dalies kilmininku. Bet abu modeliai perteikia tą pačią reikšmę, taigi galėtų būti sujungti į vieną sudėtingesnę. Kitas atvejis, kur galima būtų pasvarstyti apie vieną modelį, yra susijęs su veiksmažodžio *žiūrėti* 2.1. ir 2.2. modeliais:

2.1. Modelis: [Žmogus] žiūri | [žmogų] | [abstraktą] KAIP.

Pavyzdys: *Pradžioje gyventojai gana skeptiškai žiūrėjo į savo naujuosius kaimynus.*

2.2. Modelis: [Žmogus] žiūri | [žmogų] | [abstraktą] kaip | [objektą]

Pavyzdys: *... viename mieste šalies Šiaurėje į mus žiūrėjo kaip į muziejinius eksponatus.*

Abiejuose modeliuose vienas iš komponentų yra būdo aplinkybė, tačiau jos raiška yra skirtinga: modelyje tai įprastai yrarieveiksmiai (KAIP) {neigiamai, entuziastingai, abejingai, atsargiai...}. modelyje būdo aplinkybė reiškiamas konstrukcija *kaip | [objektą]*, taigi čia aprašant būdo aplinkybę reikalingas objektas (jo semantika nebuvo sukonkretinta dėl to, kad leksiniame rinkinyje {tėvas, vertybė, faktas, eksponatas, situoktinis...} neišryškėjo semantinės grupės). Vartosenoje matyti: kai būdo aplinkybė reiškiamasrieveiksmiu, nevartojama konstrukcija *kaip | [objektą]*, ir atvirkščiai.

Dažnas klausimas dirbant praktiškai – kiek smulkinti ar stambinti pačias reikšmes. Su *kilti* 2 ir 4 modeliais susijęs klausimas, ar verta skirti dvi atskiras reikšmes, plg. abiejų modelių pavyzdžius 2 lentelėje. Semantinė grupė [žmogus] (2 modelyje) atskiriama nuo semantinės grupės [žodis] (4 modelyje), taip pat skiriasi ir aplinkybės pozicijoje esančios semantinės grupės, plg. [teritorija], [žmonės] (2 modelyje) ir [šaltinis] (4 modelyje).

Kadangi reikšmių aiškinimai panašūs, tai galėtų būti pagrindas pateikti vieną reikšmę, bet nurodyti du jos realizavimo modelius.

Trečiame skyriuje jau buvo minėta, kad vienas iš sunkiausių vartosenos modelių analizės metodo etapų – leksinių rinkinių apibendrinimas semantinėmis grupėmis. Abstrahavimo klausimas rengiant *Pattern Dictionary of English Verbs* yra sprendžiamas naudojant ontologiją, kurioje įtrauktos semantinės grupės nuosekliai pildomos leksiniai rinkiniai. Šiame tyrime jokia ontologija nebuvo naudota, taigi semantinių grupių parinkimo klausimas iškildavo palyginti dažnai, ypač su abstraktais. Pavyzdžiui, veiksmažodžio *kilti* 1.1. vartosenos modelis yra:

kyla [jausmas | problema | grėsmė | mintis | diskusija | skandalas | nesutarimas | neramumai | fizinis negalavimas]

Šiuo atveju visas semantines grupes būtų galima apibendrinti viena grupe – [abstraktas], tačiau toks apibendrinimas turbūt nebūtų tikslus, nes apimtų per daug įvairių objektų, kurie neaktualūs šiam modeliui. Pabandžius apibendrinti modeliui būdingų abstraktų grupes, viena vertus, pasirinktos konkrečios semantinės grupės, kita vertus, tų semantinių grupių su abstraktais susidarė labai daug – tai gali būti sunkiai suprantama informacija pradedantiesiems kalbos vartotojams. Kai modelyje tokia didelė įvairovė, jeigu yra galimybė, galima skaidyti vartosenos modelį į kelis pagal leksinį ar gramatinį junglumą. Galima išėitis – dabartines semantines grupes jungti į platesnes grupes, pavyzdžiui, žiūrint į dabar išskirtas grupes *jausmas* | *problema* | *grėsmė* | *mintis* | *diskusija* | *skandalas* | *nesutarimas* | *neramumai* | *fizinis negalavimas* matyti, kad dauguma jų apibūdina neigiamos reikšmės abstraktą (problema, grėsmė, skandalas, nesutarimas, neramumai), taigi galima būtų skirti tris semantines grupes: [jausmas], [neigiamas reiškiny] ir [fizinis negalavimas], o tada pateikti dažniausių kolokatų rinkinius prie kiekvienos grupės.

Vienas iš semantinių grupių nustatymo probleminių aspektų – skirtingų žodžių polinkis jungtis tik su tam tikrais tos pačios semantinės grupės žodžiais, nes „tos pačios semantinės grupės sandara priklauso nuo konkretaus veiksmažodžio“²¹ (Jezek ir Hanks 2010: 8), pavyzdžiui, su veiksmažodžiu *attend* iš semantinės grupės [Activity] yra pasirenkami {meeting, lecture, funeral, coronation...}, bet su kitais daiktavardžiais iš [Activity] grupės (pavyzdžiui, {route, march, stroll...}) bus vartojamas ne *attend*, o kiti veiksmažodžiai (plačiau žr. Jezek ir Hanks 2010: 11–20). Kitaip sakant, nors semantinėmis grupėmis galima apibendrinti žodžio leksinę ir semantinę aplinką, vis dėlto kiekvieno žodžio kolokatai yra būdingi tik jam, vadinasi, skirtingų žodžių modeliuose gali būti nurodomos tos pačios semantinės grupės, bet iš tiesų vartojami žodžiai iš tų grupių skirsis priklausomai nuo žodžio ir jo reikšmės. Pavyzdžiui, veiksmažodžio *kilti* 2 modelio atveju objekto pozicijoje vartojami žodžiai iš semantinės grupės [žmonės] yra tokie: {giminė, šeima, vokiečiai...}. Ši semantinė grupė pateikta daugiskaita ir tai yra svarbu, nes įprastai čia šnekama ne apie vieną žmogų, o apie giminystės ar tautybės ryšiais

susijusių žmonių grupę. Veiksmožodžio *žiūrėti* vartosenos modelyje 2.1. (pavyzdžiui, *Pradžioje Torgau gyventojai gana skeptiškai žiūrėjo į savo naujuosius kaimynus.*) žodis *kaimynai* taip pat atspindi semantinę grupę [žmogus], tačiau *kaimynas*, *kaimynai* nėra būdingas semantinės grupės [žmonės] kolokatas veiksmožodžio *kulti 2* modelyje.

Kalbant apie darbą su semantinėmis grupėmis, reikia atkreipti dėmesį į tai, kad jas pavadinantys žodžiai turi būti informatyvūs, be to, šie pavadinimai turi būti žinomi besimokančiajam: pavyzdžiui, tokie semantinių grupių pavadinimai, kaip *neramumai* ar *fizinis negalavimas*, gali būti per sudėtingi.

Žodžio *reginys* (3 veiksmožodžio *žiūrėti* modelis) besimokantysis gali nežinoti, nes tai nėra dažnas žodis. Ieškant sprendimo, čia galima svarstyti pateikti leksinio rinkinio ištrauką su besimokančiajam žinomais žodžiais arba visai nepateikti semantinės grupės pavadinimo, jeigu jis sudėtingas, o iš karto pateikti leksinio rinkinio ištrauką (plg. Hanks 2004: 91).

Dar viena svarbi vartosenos modelių analizės dalis – vartosenos modelių aprašymas, tiksliau – to aprašymo techniniai aspektai. Žvalgomajame tyrime su lietuvių kalbos veiksmožodžiais taikyti pagrindiniai metodo principai, tačiau naudoti kitokie darbo įrankiai nei tie, kuriuos naudoja *Pattern Dictionary of English Verbs* žodyno rengėjai: *SketchEngine*, CPA ontologija, CPA modelių redaktorius (plačiau žr. Baisa ir kt. 2015). Taigi, dirbant tokiu metodu kaip vartosenos modelių analizė, svarbu turėti arba adaptuoti reikalingus įrankius kiekvieno tyrimo etapo užduotims: kad būtų galima automatizuotai parengti žodžio vartosenos santrauką, t. y. esminę informaciją apie žodžio leksinį ir gramatinį junglumą, kad leksikografas turėtų patogias priemones (redaktorių, žodyno rengimo sistemą ar pan.) vartosenos modeliams nustatyti, redaguoti, pildyti ir, be abejo, kad būtų tinkama duomenų bazė, kurioje visa leksikografo sukaupta informacija būtų saugoma ir analizuojama.

6. Apibendrinamosios pastabos

Šiame straipsnyje dviejų lietuvių kalbos veiksmožodžių pavyzdžiu pristatytas vartosenos modelių analizės metodas, kuris naudojamas atskirti ir aprašyti reikšmes pagal leksinius ir gramatinius dėsningumus. Nors taikant šį metodą kol kas daugiausia įdirbio sukaupta aprašant veiksmožodžius (žr. *Pattern Dictionary of English Verbs*), pasak Hankso (2008), šį metodą tinka taikyti ir būdvardžiams, su tam tikrais apribojimais – ir daiktavardžiams (plačiau žr. Hanks 2008: 117). Mokomojoje leksikografijoje šis metodas gali būti taikomas siekiant parodyti formos ir turinio sąsajas, t. y. kad su tam tikra reikšme sietini ir tam tikri leksiniai, gramatiniai, galbūt ir pragmatiniai elementai. Susipažinęs ar supažindintas su šiais dalykais besimokantysis galės geriau suvokti reikšmių skirtumus, o tai svarbu ir suvokiant kalbą, ir ją produkuojant. Vartosenos modelių analizė yra vienas iš būdų tirti ir aprašyti konkrečių žodžių leksinius ir gramatinius dėsningumus: nuo vieno iš straipsnyje aptariamų metodų (*pattern grammar*, Hunston ir

Francis 2000) šis metodas skiriasi tuo, kad, viena vertus, skirtas žodynams, o ne leksinėms gramatikoms rengti, antra vertus, čia modeliuose susitelkiama ne tik į gramatines kategorijas, – kartu tiriama, kokiomis semantinėmis grupėmis tos kategorijos realizuojamos. Iš žvalgomojo tyrimo matyti, kad kai kurios reikšmės, pasižyminčiomis vienodu gramatiniu modeliu, atskiriamos būtent pagal semantines grupes.

Po žvalgomojo tyrimo paaiškėjo ne tik metodo galimybės, bet ir probleminiai aspektai, susiję su vartosenos modelių atpažinimu (interpretavimu), aprašymu ir pateikimu mokomuosiuose žodynuose. Skirtingoms reikšmėms būdinga tam tikra leksinė ir gramatinė aplinka, kurią galima perteikti modeliu, tačiau analizuojant dažnai kildavo klausimas, kuria strategija – modelių stambinimo ar smulkinimo – remtis. Pasirinkta strategija lemia ir modelių skaičių – ar tam tikrus dėsningumus pateikti vienu ar dviem modeliais. Interpretuoti vartosenos dėsningumus yra sudėtinga, reikalingi kuo tikslesni kriterijai, kaip interpretuoti vartosenos duomenis, analizuojant reikšminga ir intucija, ir introspekcija. Dirbant vartosenos modelių analizės metodu, leksikografui reikia nustatyti, kokių reikšmių žodžiai vartojami tam tikroje junginio pozicijoje (kaip subjektai, objektai ir pan.), ir, jeigu čia išryškėja dėsningumų, apibendrinti semantinėmis grupėmis. Šiame etape leksikografui kylantys klausimai: kiek ir kaip apibendrinti nustatytus leksinius rinkinius ir kaip pateikti semantines grupes, atsižvelgiant į mokomojo žodyno tipą. Vienas iš galimų sprendimų – naudoti ontologiją, kaip tai daroma *Pattern Dictionary of English Verbs*.

Tam, kad darbas šiuo metodu leksikografui būtų patogus, reikalingi pagalbiniai įrankiai ir dalinis vartosenos dėsningumų analizės automatizavimas. Vartosenos modelių analizės metodu sukaupti duomenys iš tekstyno gali būti naudojami labai įvairiai: ne tik nustatant daugiareikšmių žodžių reikšmes, bet ir rašant reikšmių aiškinimus, atrenkant tinkamus pavyzdžius. Vartosenos modelių analizė pristatoma kaip supaprastintas būdas Sinclairio (1991; 1999) aprašytam išplėstiniam leksiniam vienetui tirti, bet čia svarbu kreipti dėmesį ne tik į leksikografą, bet ir į mokomojo žodyno naudotoją. Leksikografui šis metodas yra būdas, kaip nustatyti ir pateikti aprašomo žodžio vartosenos dėsningumus. Kad besimokantysis galėtų pasinaudoti visa šiuo metodu leksikografo sukaupta informacija, ji turi būti suprantamai pateikiama. Ar aprašyti veiksmažodžių vartosenos modeliai būtų suprantami besimokantiesiems? Ar į tuos modelius veikiau turėtume žiūrėti kaip į duomenis, kuriuos naudoja tik leksikografas, o besimokantysis (žodyno naudotojas) juos pamato tik apibrėžtyje ir / ar bapavyzdžiuose? Jeigu siekiama mokomajame žodyne parodyti ir pačius vartosenos modelius, atskira užduotis yra rasti būdą, kaip tuos modelius aiškiai ir patogiai pateikti mokomųjų žodynų naudotojams (apie panašių uždavinių sprendimą mokomojoje leksikografijoje plg. Lemmens ir Wekker 1991; Abel ir kt. 2003; Dziemianko 2015).

Literatūra

- Abel, A., J. Gamper, J. Knapp, V. Weber. 2003. Describing Verb Valency in an Electronic Learner's Dictionary: Linguistic and Technical Implications. *Proceedings of Ed-Media 2003 World Conference on Educational Multimedia, Hypermedia & Telecommunications* (eds. Lassner, D., Mcnaught, C.), June 23-28, 2003, Honolulu, Hawaii. Norfolk (USA), 1202–1209.
- Atkins, B. T. S., M. Rundell. 2008. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Atkins, S., C. J. Fillmore, C. R. Johnson. 2003. Lexicographic Relevance: Selecting Information from Corpus Evidence. *International Journal of Lexicography*, volume 16, issue 3, 251–280.
- Baisa, V., I. El Maarouf, P. Rychlý, A. Rambousek. 2015. Software and Data for Corpus Pattern Analysis. *Proceedings of the Ninth Workshop on Recent Advances in Slavonic Natural Language Processing*. Brno, Tribun EU. 75–86.
- Bielinskienė, A. 2010. *Sujungiamojo ryšio semantika, pragmatika ir vartosenos ypatumai*. Daktaro disertacija. Kaunas: Vytauto Didžiojo universitetas.
- Blauzdytė, R. 2011. *Dažniausių lietuvių kalbos veiksmažodžių gramatinio junglumo modeliai*. Magistro darbas. Kaunas: Vytauto Didžiojo universitetas.
- DiMuccio-Failla, P. V., L. Giacomini. 2017. Designing a Learner's Dictionary Based on Sinclair's Lexical Units by Means of Corpus Pattern Analysis and the Sketch Engine. *Proceedings of the eLex2017 Conference: Electronic Lexicography in the 21st century*. I. Kosem, C. Tiberius, M. Jakubiček, J. Kallas, S. Krek, V. Baisa (eds.). Leiden. 437–457.
- Drūlienė, V. 2014. *Lietuvių kalbos apibendrinamieji daiktavardžiai: leksinė reikšmė ir pragmatinės funkcijos*. Daktaro disertacija. Kaunas: Vytauto Didžiojo universitetas.
- Dziemianko, A. 2015. Colours in Online Dictionaries: A Case of Functional Labels. *International Journal of Lexicography*, volume 28, issue 1, 27–61.
- El Maarouf, I. 2013. Methodological Aspects of Corpus Pattern Analysis. *ICAME Journal*, no. 37, 119–148.
- Frankenberg-Garcia, A. 2014. The Use of corpus examples for language comprehension and production. *ReCALL* 26 (2), 128–146.
- Frankenberg-Garcia, A. 2015. Dictionaries and encoding examples to support language production. *International Journal of Lexicography*, 24/4, 490–512.
- Grigonytė, G., J. Kovalevskaitė, E. Rimkutė. 2018: Linguistically Motivated Automatic Classification of Lithuanian Texts for Didactic Purposes. *Proceedings of the Eighth International Conference Baltic HLT 2018. Frontiers in Artificial Intelligence and Applications* (eds. K. Muischnek, K. Müürisep), vol. 307. Amsterdam, Berlin, Tokyo, Washington, DC: IOS Press. 38–46.
- Hanks, P. 2004. Corpus Pattern Analysis. *Euralex 2004 Proceedings*, vol. 1, 87–99.
- Hanks, P. 2007. Seminario medžiaga iš *Nordic Graduate School of Language Technology* (NGSLT) kursų *Lexicology, lexicography and corpus linguistics*,

- Ryga. Prieiga internetu: http://www.lumii.lv/ngslt/hanks/05_Riga_Wednesday.pdf [žiūrėta 2019-01-05].
- Hanks, P. 2008. Lexical Patterns: from Hornby to Hunston and Beyond. *Proceedings of the 13th EURALEX International Congress*. Barcelona, Spain, July 15-19, 2008, 89–129.
- Hunston S., G. Francis. 2000. *Pattern Grammar. A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam, Philadelphia: John Benjamins.
- Hunston, S. 2004. The Corpus, Grammar Patterns, and Lexicography. *Lexicographica*, 20, 100–113.
- Jezek, E., P. Hanks. 2010. What lexical sets tell us about conceptual categories. *Lexis. Journal in English Lexicology*, no. 4, 1–15.
- Labutis, V. 2002. *Lietuvių kalbos sintaksė*. Vilnius: Vilniaus universiteto leidykla.
- Lemmens, M., H. Wekker. 1991. On the Relationship between Lexis and Grammar in English Learners' Dictionaries. *International Journal of Lexicography*, volume 4, issue 1, 1–14.
- Marcinkevičienė, R. 2000. Tekstynų lingvistika: teorija ir praktika. *Darbai ir dienos*, t. 24, 7–64.
- Moon, R. 2007. Sinclair, lexicography, and the Cobuild Project. *International Journal of Corpus Linguistics*, volume 12, issue 2, 159–181.
- Moon, R. 2018. Explaining Meaning in Learners' Dictionaries. *The Oxford Handbook of Lexicography*. Oxford: Oxford University Press.
- Ramonienė, M., J. Pribušauskaitė. 2003. *Praktinė lietuvių kalbos gramatika*. Vilnius: Baltų lankų leidyba.
- Ramonienė, M., J. Pribušauskaitė, L. Vilkienė. 2016a. *Slenkstis*. Vilnius: VU.
- Ramonienė, M., J. Pribušauskaitė, L. Vilkienė. 2016b. *Pusiaukelė*. Vilnius: VU.
- Renau, I., P. Battaner. 2012. Using CPA to represent Spanish pronominal verbs in a learner's dictionary. *Proceedings of the XV EURALEX*, 350–361.
- Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. 1996. The Search of Units of Meaning. *Textus IX*, 75–106.
- Sinclair, J. 1999. The Lexical Item. *Contrastive Lexical Semantics* (eds. E. Weigand). *Current Issues in Linguistic Theory* 17. Amsterdam, Philadelphia: John Benjamins, 1–24.
- Storjohann, P. 2010. Colligational patterns in a corpus and their lexicographic documentation. *Proceedings of the Corpus Linguistics Conference 2009* (eds. M. Mahlberg, V. González-Díaz, C. Smith), 1–18.
- Tognini-Bonelli, E. 2000. Corpus Classroom Currency. *Darbai ir Dienos*, t. 24, 205–243.

Žodynai ir įrankiai

- Anthony, L. 2018. AntConc (3.5.2 versija) [tekstyno analizės įranga]. Tokyo: Waseda University. Prieiga internetu: <http://www.laurenceanthony.net/software/antconc/>.
- Norkaitienė, M., R. Šepetytė, Z. Šimėnaitė. 2010. Mokomas lietuvių kalbos žodynas. Vilnius: Baltos lankos.

Pattern Dictionary of English Verbs. Prieiga internetu: <http://pdev.org.uk/#browse?q=f=C> (tęstinis projektas, žiūrėta 2019-06-14).

Utkā, A. 2009. Dažninis rašytinis lietuvių kalbos žodynas: 1 milijono žodžių morfologiškai anotuoto tekstyno pagrindu. Kaunas: VDU leidykla. Prieinamas internetu CLARIN-LT saugykloje: <https://clarin.vdu.lt/xmlui/handle/20.500.11821/12> (žiūrėta 2019-08-14).

Notes

- 1 Lietuviškas termino vertimas – modelių gramatika – perimtas iš Agnės Bielinskienės (2010).
- 2 Mokomasis tekstynas kuriamas vykdant projektą „Užsienio baltistikos centrų ir Lietuvos mokslo ir studijų institucijų bendradarbiavimo skatinimas“. Tekstynas sudarytas iš patekstynių skirtingiems lietuvių kalbos mokymo lygiams (A1, A2, B1, B2) (plačiau Grigonytė ir kt. 2018): atitinkamo lygio besimokantiems suprantamų ir aktualių rašytinės kalbos tekstų dalis jau parengta, sakytinės kalbos tekstų dalis baigiama rengti.
- 3 Morfologiškai anotuotas tekstynas. Prieinamas internetu CLARIN-LT saugykloje <http://hdl.handle.net/20.500.11821/9>.
- 4 Čia naudotasi Cobuild serijos žodynu Collins Cobuild Student's Dictionary. London: HarperCollins Publishers. 1994.
- 5 Atskirai reikėtų aptarti klausimą, ar tikslinga informaciją apie vartoseną pateikti reikšmių aiškinimuose (apie Cobuild apibrėžčių formos kritiką žr. Hanks (1987), taip pat Dziemianko (2006), visi cituojami iš Frankenberg-Garcia 2015); juk apibrėžčiai keliami tokie kriterijai, kaip glaustumas, o reikšmės aiškinimas iš esmės skirtas receptijai (Atkins, Rundell 2008: 410). Produktavimui yra numatyti pavyzdžiai, kurie Cobuild žodyne taip pat atrinkti iš tekstyno.
- 6 Čia Hanksas remiasi leidimu: Sinclair J. M., Fox G., Francis G. ir kt. (1995, 2001). Collins Cobuild English Language Dictionary. London: HarperCollins.
- 7 Verta paminėti iki gramatikos parengtus leidinius: Francis G., Hunston S., Manning E (1996): Collins Cobuild Grammar Patterns 1: Verbs. London: HarperCollins; Francis G., Hunston S., Manning E (1998): Collins Cobuild Grammar Patterns 2: Nouns and Adjectives. London: HarperCollins.
- 8 A pattern can be identified if a combination of words occurs relatively frequently, if it is dependent on a particular word choice, and if there is a clear meaning associated with it (Hunston ir Francis 2000: 37).
- 9 The patterns of a word can be defined as all the words and structures which are regularly associated with the word and which contribute to its meaning (Hunston ir Francis 2000: 37).
- 10 Just as a word can have several different patterns, so a pattern can be seen to be associated with a variety of different words (Hunston ir Francis 2000: 43).
- 11 Trumpinant straipsnyje leksinės semantinės grupės toliau vadinamos semantinėmis grupėmis.
- 12 Remiantis išplėstinio leksinio vieneto teorija (Sinclair 1996), vartosenos dėsnų tyrimas pradedamas nustatant kolokacijas: pagal dažnumą nustatomi svarbūs kolokatai, toliau jie analizuojami pagal kitus parametrus (poziciją, kalbos dalį, semantines sąsajas ir pan.), o toks klasifikavimas veda prie apibendrinimų, kurie padeda susisteminti vartosenos ypatybes (Tognini Bonelli 2000: 209).
- 13 Pattern Dictionary of English Verbs ontologija: <http://pdev.org.uk/#onto>.
- 14 Svarbu paminėti, kad šalia modelio įprastai būna nurodomas ir jo dažnumas procentais (El Maarouf 2013: 122). 15 čia aptariamų veiksmažodžių analizė atlikta Laimos Jancaitės bakalauro darbe (2018) „Polisemiškų veiksmažodžių

- reikšmių nustatymas vartosenos modelių metodu“. Bakalauro darbe ištirti dar 7 veiksmažodžiai.
- 15 cia aptariamų veiksmažodžių analizė atlikta Laimos Jancaitės bakalauro darbe (2018) „Polisemiškų veiksmažodžių reikšmių nustatymas vartosenos modelių metodu“. Bakalauro darbe ištirti dar 7 veiksmažodžiai.
 - 16 Analizuojant gramatinį junglumą, naudotasi Ritos Blauzdytės (2011) magistro darbo medžiaga, kur pateikti 100 dažniausių dabartinės lietuvių kalbos veiksmažodžių junglumo modeliai, nustatyti tekstyne. Junglumas suvokiamas kaip žodžio (formos) galėjimas jungtis su kitu žodžiu (kita forma). Kadangi žodžio ar atskirų jo formų junglumu vadinamas tiek valentinių partnerių, tiek valentingumo nenumatomų elementų jungimasis prie žodžio (Labutis 2002: 22), manytume, kad junglumo sąvoka modeliams tirti yra parankesnė nei valentingumo, nes leidžia tiksliau atspindėti vartosenos dėsningumus.
 - 17 Šiam žvalgomajam tyrimui nebuvo naudota ontologija: semantinės grupės apibendrintos ir pavadintos tik šio tyrimo poreikiams, pagal tai, kokia informacija surinkta iš tyrinėtos medžiagos (kai kurie semantinių grupių pavadinimai perimti iš anglų kalbos ontologijos, naudojamos Pattern Dictionary of English Verbs).
 - 18 Vartosenos modelių analizės metodu tiriant anglų kalbos veiksmažodžius aprašomi ir tokie pastovieji junginiai, kaip idiomos (frazologizmai), fraziniai veiksmažodžiai. Šiame tyrime frazeologizmai buvo atpažinti, tačiau neįtraukti prie analizuotų veiksmažodžių vartosenos modelių.
 - 19 Paradigmatic sets of words that populate specific argument slots within the same verb sense do not map neatly onto conceptual categories, as they often include words belonging to different semantic types (Jezek, Hanks 2010: 8).
 - 20 <https://www.sketchengine.eu/> (žr. funkciją WordSketch).
 - 21 The internal composition of lexical sets changes when one moves from verb to verb (Jezek ir Hanks 2010: 8).