

Jurgita Vaičenonienė, Jolanta Kovalevskaitė
 Vytauto Didžiojo universitetas, Lietuva

LEKŠINĖS IR MORFOLOGINĖS VERTIMŲ KALBOS YPATYBĖS

Santrauka: Straipsnio tikslas – remiantis lyginamuoju lietuvių kalbos tekstynu, ištirti leksines ir morfologines vertimų kalbos ypatybes. Lyginamąjį lietuvių kalbos tekstyną ORVELIT sudaro 4 patekstyniai po 1 mln. žodžių: grožinės literatūros originalūs tekstai lietuvių kalba ir vertimai iš anglų į lietuvių kalbą; mokslo populiarinimo originalūs tekstai lietuvių kalba ir vertimai iš anglų į lietuvių kalbą. Morfologiškai anotuota tekstyno versija naudota tiriant morfologines ypatybes, leksinėms ypatybėms tirti naudota neanotuota versija. Tyrime taikyta tekstynais paremtų vertimo studijų metodologija, statistinis reikšmingumas vertintas panaudojant logaritminės tikimybės testą. Nustatyta, kad vertimų leksikai būdinga mažesnis informacinis tankis; didesnis tarnybinių žodžių skaičius; trumpesni sakiniai; žemesni žodingumo rodikliai ir retų žodžių vartoseną. Morfologinių kategorijų analizė rodo, kad abiejų žanrų vertimuose mažiau daiktavardžių ir būdvardžių, bet daugiau veiksmažodžių (ypač asmenuojamųjų formų), o iš tarnybinių kalbos dalių daugiau įvardžių, prielinksnių. Linksnių analizė atskleidė, kad vertimuose fiksuojama mažiau vardažodžių kilmininko, o iš veiksmažodžio kategorijų – mažiau dalyvių: grožiniuose tekstuose mažiau veikiamosios rūšies dalyvių, mokslo populiarinimo tekstuose mažiau neveikiamosios rūšies dalyvių. Nustatytos vertimų kalbos ypatybės leidžia teigti, kad vertimams iš anglų kalbos būdingos tam tikros nuo originaliosios lietuvių kalbos besiskiriančios ypatybės: šias ypatybes ir jų priežastis verta analizuoti išsamiau konkrečių atvejų tyrimuose.

Pagrindinės sąvokos: tekstynais paremtos vertimo studijos; vertimų kalba; vienakalbiai lyginamieji tekstynai; ORVELIT.

Įvadas

Viešajame ir akademiniam diskurse dažnai diskutuojama apie vertimus ir anglų kalbos įtaką lietuvių kalbai (pvz., Radvilavičiūtė, 2014; Rudaitienė, 2014). Vyraujanti nuomonė, kad vertimų kalba yra prastesnė ir kad anglų kalba skurdina lietuvių kalbą, dažnai grindžiama įvairiais leksikos skolinių, kalbos klaidų pavyzdžiais, tačiau lyginamųjų originalios ir vertimų lietuvių kalbos tyrimų, kuriuose būtų bandoma pasižiūrėti, kaip skiriasi originalioji ir vertimų kalba, iki šiol būta nedaug (Vaičenonienė, 2014). Rašant apie vertimus į lietuvių kalbą dažnai keliama klausimai apie įvairių vertimo strategijų taikymą tam tikriems leksikos, gramatikos ar sintaksės aspektams, vyrauja lyginamieji skirtingų, o ne tos pačios kalbos originalių ir verstų tekstų tyrimai.

Tekstynais paremtose vertimo studijose tyrinėjama, kokių vertimų kalbos ypatybių atsiranda dėl kalbos, iš kurios verčiama, įtakos; kiek vertimų

kalba priklauso nuo vertėjo sprendimų ir pan. Naudojant vienakalbius lyginamuosius tekstynus siekiama nustatyti, kuo ir kiek vertimų kalba skiriasi nuo tos pačios kalbos originalių tekstų. Lietuvių kalba priskirtina mažosioms kalboms (angl. *lesser-used languages*) (Marcinkevičienė et al., 2012), daugiau nei pusę lietuvių kalba leidžiamų knygų sudaro vertimai, vadinasi, mūsų kalbos vartotojai dažnai susiduria būtent su vertimais, ir būtent vertimai daro didelę įtaką lietuvių kalbos vartosenai, taigi lietuvių vertimų kalbos tyrimai yra aktualūs ir teoriniu, ir praktiniu požiūriu.

Šiame straipsnyje, remiantis lyginamuoju lietuvių kalbos tekstynu ORVELIT, aprašomos vertimų kalbos leksinės ir morfoliginės ypatybės. Lyginamąjį lietuvių kalbos tekstyną sudaro 4 patekstyniai po 1 mln. žodžių: grožinės literatūros originalūs tekstai lietuvių kalba (toliau GRO) ir vertimai (GRV) iš anglų į lietuvių kalbą; mokslo populiarinimo originalūs tekstai (MPO) lietuvių kalba ir vertimai iš anglų į lietuvių kalbą (MPV) (plačiau apie tekstyno sandarą žr. Vaičėnienė et al., 2017). Kadangi ORVELIT sudaro dviejų žanrų tekstai, jų palyginimas leis daryti patikimesnes išvadas apie leksinių ir morfoliginų ypatybių panašumus ir skirtumus originaliuose bei verstuose lietuviškose tekstuose ir galimą vertimų kalbos ypatybių varijavimą skirtinguose žanruose.

ORVELIT tekstai buvo automatiškai morfologiškai anotuoti naudojant *semantika.lt* morfoliginį anotatorių taikant *Universal Dependencies* ir *Leipcigo* glosavimo pažymas. Neanotuota tekstyno versija naudota leksinėms ypatybėms tirti, anotuota versija naudota tiriant morfologines ypatybes. Leksikos ypatybės tiriamos pagal žodingumo, vidutinio sakinių ilgio, dažniausių ir retų žodžių santykio ir reikšminių žodžių rodiklius. Vertimų kalbos morfologines ypatybes padės atskleisti savarankiškų ir tarnybinių kalbos dalių pasiskirstymas, vardažodžių linksnių, veiksmažodžių kategorijų (asmenuojamųjų ir neasmenuojamųjų formų, dalyvių) palyginimas originalioje ir vertimų kalboje. Raktinių morfoliginų žymų analizė skirta papildyti pasirinktų morfoliginų ypatybių tyrimą. Teorinis ir metodologinis tyrimo pagrindas pagrįstas tekstynų lingvistikos, deskriptyviųjų vertimo studijų ir lingvistinės stilistikos principais.

Vertimų kalbos leksikos ypatybės

Šioje dalyje apžvelgsime vieno pirmųjų anglų vertimų kalbos tyrimų (Laviosa 1998) rezultatus, paskatinusius ir kitų kalbų tyrėjus remtis panašiais metodologiniais sprendimais. Siekiant rezultatų palyginamumo, lietuvių vertimų kalbos leksikos ypatumai taip pat bus tiriami laikantis S. Laviosos darbuose taikytų metodologinių principų: aptarsime informacinį tankį ir žodingumą; palyginsime vidutinį sakinių ilgį²²; analizuosime dažniausių ir retų žodžių santykį; apžvelgsime teigiamų (angl. *positive key words*) ir neigiamų (angl. *negative keywords*) raktinių žodžių ypatumus. Dažniniai ORVELIT žodžių sąrašai ir bendroji tekstyno statistika sugeneruota naudojant įrankį *WordSmith Tools Version 7.0* (Scott, 2016), o raktiniai žodžiai – *Lancaster Stats Tools Online* (Brezina, 2018)²³. Statistiniam duomenų reikšmingumui įvertinti taikytas logaritminės tikimybės (ang. *Log Likelihood*) testas²⁴.

Lyginant vertimų kalbos ypatumus, tekstynais paremtuose vertimų tyrimuose dažniausiai minima S. Laviosos (1998) pasiūlyta metodologija. Specialiai sukurtame anglų vertimų kalbos tekстыne S. Laviosa (1998) nustatė, kad grožinės literatūros ir spaudos vertimuose tarnybinių žodžių proporcija lyginant su prasminiais žodžiais didesnė, o tai rodo mažesnę vertimų žodyno arba informacinį tankį. Apskaičiuodama prasminių ir tarnybinių žodžių santykį, S. Laviosa naudojo M. Stubbs (1986, 1996) žodyno tankio sampratą. Žodyno tankis reiškiamas procentais ir gaunamas taip: iš visų žodžių skaičiaus atimamas tarnybinių žodžių skaičius (taip gaunamas prasminių žodžių skaičius), o rezultatas vėl padalinamas iš visų žodžių skaičiaus (Laviosa, 1998, p. 10). Kadangi ORVELIT yra morfologiškai anotuotas, žodyno tankis bus paskaičiuotas naudojantis automatiškai sugeneruotais kalbos dalių dažniais padalinus prasminių žodžių skaičių iš viso atitinkamo patekstynio žodžių skaičiaus.

²² Ši ypatybė susijusi su sintakse, bet kadangi tik ji vienintelė buvo tiriama, atskiros sintaksinių ypatybių grupės neskirta ir aptarta leksikos ypatybių dalyje.

²³ Lancaster Stats Tools online. <http://corpora.lancs.ac.uk/stats/toolbox.php>

²⁴ Log-likelihood and effect size calculator. <http://ucrel.lancs.ac.uk/llwizard.html>

Kinų vertimų kalbai, kaip ir anglų, būdingas statistiškai reikšmingas žemesnis vertimų kalbos žodyno tankis ir didesnė tarnybinių žodžių proporcija lyginant su originalios kalbos prasminių ir tarnybinių žodžių santykiu.²⁵ Kitas rodiklis, nusakantis žodyno įvairovę ir žodingumą, yra standartizuotas santykis tarp tekstyno apimties ir skirtingų žodžių kiekio (angl. *standardized type token ratio*) (Xiao & Hu, 2015). Kinų vertimų kalbos tyrime nustatyta, kad vertimų žodyno įvairovė skirtinguose žanruose ne visuomet statistiškai reikšmingai mažesnė už originaliosios kinų kalbos (Xiao & Hu, 2015, p. 70).

Prasminių ir tarnybinių žodžių skaičius, informacinis tankis ir žodingumas ORVELIT tekстыne pateikiami 1-oje lentelėje.

1 lentelė.

Informacinis tankis ir žodingumas ORVELIT tekстыne

	GRV <i>Grožinės literatūros vertimai</i>	GRO <i>Originali grožinė literatūra</i>	MPV <i>Mokslo populiarinimo vertimai</i>	MPO <i>Originali mokslo populiarinimo literatūra</i>
Prasminiai²⁶ žodžiai	630424 (62,94 %)	634619 (63,35 %)	672295 (67,67 %)	696641 (69,76 %)
Tarnybiniai žodžiai	328518 (32,80 %)	311613 (31,11 %)	291684 (29,36 %)	248608 (24,89 %)
Žodyno/informacinis tankis	65,74%	67,06%	69,74%	73,69%
Žodingumas (angl. <i>STTR</i>)	63,14%	66,02%	63,53%	62,38%

Prasminių žodžių skaičius 1-oje lentelėje rodo bendrinių ir tikrinių daiktavardžių, veiksmažodžių, būdvardžių, ir prievaismių skaičių morfologiškai anotuotoje ORVELIT tekstyno versijoje; tarnybinių žodžių suma gauta sudėjus įvardžių, skaitvardžių, prielinksnių, jungtukų, jaustukų ir dalelyčių vartojimo atvejus. Originaliuose tekstuose daugiau prasminių žodžių, o vertimuose didesnė tarnybinių žodžių proporcija ir šis skirtumas yra statistiškai reikšmingas (grožinė literatūra: LL = 249.44; mokslo populiarinimas: LL = 2644.72)²⁷. Informacinio tankio skaičiavimai taip pat rodo, kad žodyno tankis grožinės literatūros ir mokslo populiarinimo vertimuose mažesnis, o žodingumo rezultatai pasiskirstę nevienodai – nors

²⁵ Kinų vertimų kalbos tekstyną sudarė 15 skirtingų žanrų (Xiao & Hu, 2015, p. 68).

²⁶ Šiame straipsnyje sinonimiškai vartojame savarankiškų kalbos dalių terminą.

²⁷ Skirtumas laikomas reikšmingu, kai $p < 0,05$, kritinė vertė = 3,84.

originalioje grožinėje literatūroje žodyno įvairovė didesnė, abiejų mokslo populiarinimo patekstynių rodiklis tik nežymiai skiriasi. Tokie rezultatai sutampa su kinų vertimų kalbos tyrimo (Xiao & Dai, 2014, p. 7) išvadomis: informacinis verstų tekstų tankis mažesnis nei originalų, bet žodingumo rodikliai priklauso nuo konkretaus žanro.

S. Laviosos (1998, p. 5) spėjimas, kad vidutinis sakinių ilgis vertimuose bus visuomet mažesnis, nepasitvirtino. Nors spaudos žanrų tekстыne vertimų sakiniai buvo vidutiniškai trumpesni, grožinės literatūros tekстыne vidutinis vertimų sakinio ilgis viršijo originalios literatūros žodžių skaičiaus vidurkį sakinyje. Panašius rezultatus rodo ir kinų vertimų kalbos tyrimas: nors originalioje kinų kalboje sakiniai vidutiniškai ilgesni, skirtinguose žanruose šis skirtumas ne visuomet statistiškai reikšmingas (Xiao & Hu, 2015, p. 75). Taigi vidutinio sakinio ilgio parametras daugiau apibūdina tarpžanrinius tekstų skirtumus, o teiginiai, kad vertimų kalbos sakiniai ilgesni ar trumpesni dėl daugiažodiškumo ar paprastinimo, priklauso nuo įvairių aplinkybių, susijusių su konkrečiu tekstu vertimu.

2 lentelė.

Vidutinio sakinių ilgio palyginimas ORVELIT tekстыne

Patekstynis	GRV	GRO	MPV	MPO
Vidutinis sakinio ilgis žodžiais	8,75	10,85	13,36	15,55

ORVELIT tekstyno duomenys rodo, kad tiek grožinės literatūros, tiek mokslo populiarinimo tekстыuose sakiniai vidutiniškai trumpesni nei originalios kalbos patekstyniuose. Kita vertus, kaip rodo Laviosos (1998) bei Xiao ir Hu (2015, p. 75) tyrimai, šis rodiklis nėra universalus. Norint teigti, kad lietuvių kalbos vertimuose sakiniai visuomet trumpesni, reikėtų didesnės apimties skirtingų žanrų tyrimų.

S. Laviosa (1998) taip pat nustatė, kad dažniausių žodžių santykis lyginant su retais žodžiais vertimų kalbos patekstyntyje didesnis, o dažninio žodžių sąrašo viršūnė sudaro didesnę patekstytnio žodžių skaičiaus dalį. Kitaip tariant, dažniausių žodžių vertimuose daugiau nei originaliuose tekстыuose. Dažniausi žodžiai arba dažninio žodžių sąrašo viršūnė S. Laviosos (1998), R. Xiao ir X. Hu (2015) bei R. Xiao ir G. Dai (2014) tyrimuose suprantami kaip

leksikos vienetai, kurių pasikartojimo dažnis tekstyne ne mažesnis nei 0,10 %. Laviosos darbe dažninio žodžių sąrašo viršūnę sudarė pirmieji 108 vienetai, dauguma jų – tarnybinių žodžių. Retais laikytini tie žodžiai, kurių pasikartojimo dažnis tekstyne mažesnis nei 6 (Xiao & Hu, 2015, p. 72). Anglų kalbos grožinės literatūros vertimuose dažninio žodžių sąrašo viršūnėje nustatytos 82 lemos (plg. originalios kalbos – 87 lemos), kurios sudarė net 56,2% viso žodžių sąrašo (plg. originalios kalbos – 51,6 %) (Laviosa, 1998, p. 6). Kinų vertimų kalbos tyrimo rezultatai panašūs: dažninio žodžių sąrašo viršūnę sudaro 40,47 % vertimų dažninių žodžių sąrašo, o originalios kinų kalbos – 35,70 % (Xiao & Hu, 2015, p. 72). Retų žodžių (tekstyne pasikartojančių 1–5 kartus) dažnumo proporcija didesnė verstuose kinų kalbos tekstuose (Xiao & Hu, 2015, p. 73). Metodologiniu požiūriu, ORVELIT tekstyne taip pat galima palyginti retų ir dažnų žodžių proporcijas, tačiau dėl lietuvių kalbos morfologinio daugiareikšmiškumo lemų nustatymas tik iš dažninio žodžių sąrašo neanotuotame tekstyne nebūtų validus.

3 lentelė.

Dažninių žodžių sąrašų viršūnės ir reti žodžiai ORVELIT tekstyne

Patekstinis	GRV	GRO	MPV	MPO
Dažnų žodžių eiliškumas dažniniame sąrašė	1–103	1–94	1–97	1–95
Dažni žodžiai	334609 (33,35 %)	306402 (30,54 %)	294499 (29,36 %)	246680 (24,28 %)
Reti žodžiai	107760 (10,75 %)	120748 (12,05 %)	110297 (11,10 %)	112031 (11,21 %)

Nors dažninių žodžių sąrašo viršūnę, skirtingai nei anglų kalboje (Laviosa, 1998), nesudaro didžiosios žodžių sąrašo dalies, vertimuose jos proporcija didesnė. Kitaip tariant, lietuvių, kaip ir anglų bei kinų kalbų vertimuose dažniausi žodžiai vartojami dažniau nei originaliuose rašytinės kalbos tekstuose (grožinė literatūra: LL = 1243.72, mokslo populiarinimo: LL = 4485.02). Kaip minėta, nors neanotuotame žodžių sąrašė lemų suskaičiuoti neįmanoma, plika akimi matyti, kad prasminių žodžių visose sąrašų viršūnės keleriopai mažiau nei tarnybinių (dominuoja jungtukai, prielinksniai, įvardžiai). Statistiškai reikšmingi ir retų žodžių vartosenos

skirtumai: vertimuose jų mažiau nei originalioje kalboje (grožinė literatūra: LL = 737.66, mokslo populiarinimo: LL = 5.98), o tai rodo polinkį į paprastesnę kalbą.

Dar vienas R. Xiao ir G. Dai (2014) analizuojamas aspektas – raktiniai žodžiai, gauti lyginant vertimų ir originalios kalbos dažninių žodžių sąrašus (t. y. originalios kalbos tekstynas naudotas kaip lyginamasis). Teigiami raktažodžiai rodo ypač dažnus, o neigiami – ypač retai pasitaikančius leksikos vienetus. Tai leidžia daryti išvadas apie vertimuose pernelyg dažnai ar retai vartojamus kalbos elementus lyginant su originalia kalba (ibid.). Kadangi prasminiai žodžiai nusako turinį, R. Xiao ir G. Dai (2015, p. 10) daugiau dėmesio skyrė tarnybiniams žodžiams, parodantiems vartosenos skirtumus originaluose ir vertimuose. Teigiamų raktažodžių tyrimas atskleidė, kad kinų kalbos vertimuose daugiau tarnybinių žodžių, ypač asmeninių ir parodomųjų įvardžių. Neigiami raktažodžiai parodė, kad, palyginus su kinų kalbos tektais, vertimuose mažiau kultūrinės realijas nusakančių žodžių, kai kurių tarnybinių žodžių ir kt.

Naudojant *Lancaster Stats Tools Online* (Brezina, 2018) programinės įrangos *Words* funkciją sugeneruoti 10 pirmųjų teigiamų ir neigiamų grožinės literatūros ir mokslo populiarinimo patekstynių raktažodžių. Visi grožinės literatūros raktažodžiai yra daiktavardžiai ir atspindi daugiau turinio skirtumus negu vertimų kalbos ypatumus: 9 teigiami raktažodžiai yra adaptuoti angliški asmenvardžiai (pvz. *Haris, Evanas, Brunetis, Makmerfis*) ir vienas bendrinis daiktavardis – *tėvukas*. Visi 10 neigiami raktažodžiai – tai lietuviškuose kūriniuose minimi lietuviški ar kitų kalbų veikėjų vardai (pvz., *Gabrielius, Edvardas, Algė, Žygis*). Teigiamus mokslo populiarinimo vertimų raktažodžius sudaro 4 asmeniniai (*tau, tu, jūs, jums*), 2 savybiniai (*tavo, jūsų*) įvardžiai ir 3 daiktavardžiai (*Churchillis, Churchillio, pasąmonė*). Šie rezultatai siejasi su ankstesniu pastebėjimu apie didesnę tarnybinių žodžių kiekį vertimuose. Galima spėti, kad gausesnė įvardžių vartoseną rodo anglų kalbos įtaką, kai, pavyzdžiui, verčiant iš analitinės anglų kalbos į sintetinę lietuvių, vartojami įvardžiai, nors juos galima būtų praleisti (apie tai plačiau rašoma kitoje dalyje).

Neigiamus mokslo populiarinimo vertimų raktažodžius sudaro tikriniai (*Lietuvos, Žalgirio, Lietuvoje*) ir bendriniai (*lietuvių, mūsų, teismo, teismas, valstybės*) daiktavardžiai, santrumpa *MKP* (masinės komunikacijos priemonės), trumpinys *V* (pirma vardo raidė arba penktas amžius). Neigiami

raktažodžiai, kaip ir grožinės literatūros atveju, rodo tekstuose plėtojamas temas: natūralu, kad anglų kalbos vertimai atspindi anglo-amerikietiškos (pvz., *Churchillis*), o ne lietuviškos (pvz., abu kolokacijos *Žalgirio mūšis* dėmenys) kultūros realijas.

Apibendrinant galima pasakyti, kad vertimų kalbai būdingas mažesnis informacinis tankis, dažnesnis tarnybinių žodžių vartojimas ir trumpesni sakiniai. Žemesni žodingumo rodikliai ir retų žodžių vartoseną pastebėta gožinės literatūros vertimuose, mokslo populiarinimo tekstuose šie rodikliai tik nežymiai skiriasi. Raktažodžių analizė parodė vertimų turinio ypatumus (būdingi anglų kalbos asmenvardžiai, nebūdingos lietuviškos realijos) ir įvardžių vartosenos skirtumus mokslo populiarinimo vertimuose. Šie rezultatai patvirtina anglų (Laviosa, 1998) ir kinų (Xiao & Hu, 2015; Xiao & Dai, 2014) vertimų kalbos tyrimų išvadas, kad ne visi nustatyti ypatumai yra universalūs pobūdžio, o vertimų kalba gali varijuoti pagal skirtingus žanrus, priklauso nuo kalbos, iš kurios verčiama ar kitų intra- ir ekstratekstinių faktorių.

Morfologinės vertimų kalbos ypatybės

Yra atlikta nemažai tyrimų, kuriuose aprašomos rašytinės ir sakytinės lietuvių kalbos morfologinės ypatybės (žr. Rimkutė, 2006; Dabašinskienė, 2009 ir kt.). R. Brinkutės²⁸ tyrime lyginti gramatinių kategorijų vartosenos atskiruose funkcinuose stiliuose ir bendri rašytinės kalbos gramatinių kategorijų vartosenos duomenys; matyti, kad skirtumai nedideli, vyrauja tos pačios kalbos dalių vartojimo tam tikruose stiliuose tendencijos (Brinkutė, 2017, p. 37). Visose dalyse daugiausia vartojami daiktavardžiai, antri – veiksmažodžiai, tačiau dažniausių kalbos dalių pasiskirstymas tarp stilijų yra nevienodas: pvz., visose tekstyno dalyse daugiausia vartojami daiktavardžiai, tačiau administracinio stiliaus dalyje daiktavardžiai sudaro 45,14 %, mokslinio stiliaus – 42,6 %, publicistikos – 39,76 %, o grožinio stiliaus – tik 26,2 %; veiksmažodžių daugiausia grožiniame stiliuje – 24,69 %, kitose dalyse

²⁸ R. Brinkutės (2017) tyrime, palyginti su ankstesniu – E. Rimkutės (2006) darbu – analizuota daugiau rašytinės kalbos duomenų – naudota ne 1 mln., o 1,6 mln. MAT-o versija. MAT-as yra morfologiškai anototas lietuvių kalbos tekstynas, kuriame po automatinio anotavimo morfologinės pažymos dar peržiūrėtos lingvisto. MAT-ą sudaro grožinės ir negrožinės literatūros tekstai, publicistika, yra keli iš kitų kalbų versti tekstai. Naujausia informacija apie MAT-ą pateikiama Bielinskienė et al., 2017.

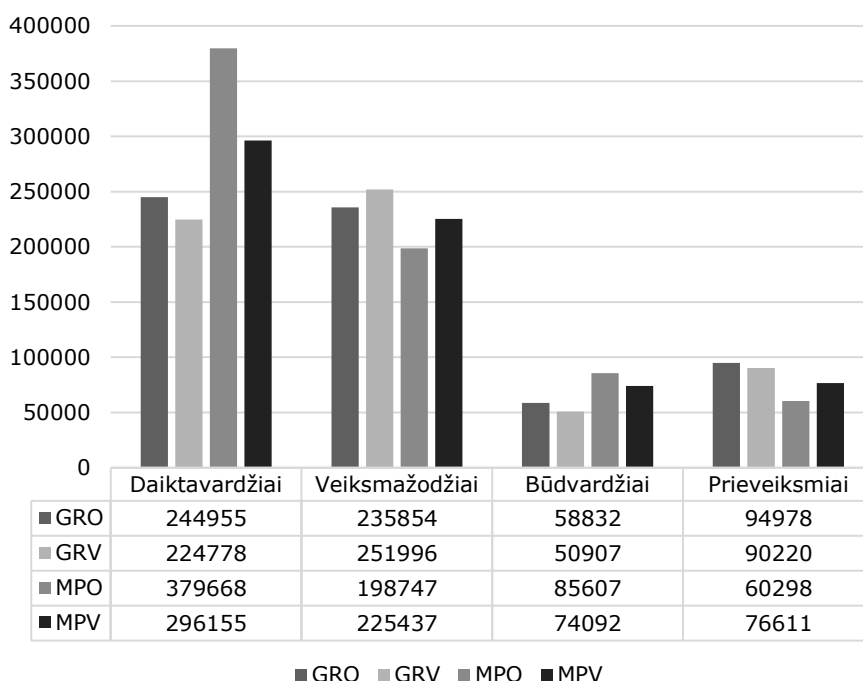
pavartota apylygiai: administraciniame stiliuje – 19,14 %, moksliniame stiliuje – 19,1 %, publicistikoje – 20,45 % (Brinkutė, 2017, p. 32–33).

ORVELIT morfologinių ypatybių analizė susideda iš dviejų dalių a) aprašomas pagrindinių gramatinių kategorijų pasiskirstymas verstuose ir originaliuose grožinės literatūros ir mokslo populiarinimo tekstuose (kalbos dalys, vardažodžių linksniai, veiksmažodžio asmenuojamosios ir neasmenuojamosios formos); b) gauti rezultatai palyginami su raktinėmis morfologinėmis žymomis. Antroji analizės dalis papildo pirmąją, nes tarp raktinių žymų gali pasitaikyti tokių morfologinių formų, kurios neišryškėjo tiriant pagrindinių morfologinių kategorijų pasiskirstymą vertimų kalboje. Raktinių žymų nustatymas iš esmės remiasi R. Xiao ir G. Dai (2014) aprašytu principu, tik vietoj būdingų ir nebūdingų žodžių čia jis pritaikytas nustatyti tam tikram tekstynui būdingas ir nebūdingas gramatines formas (Rayson et al. (2008) raktinių žodžių principo (angl. *keyword analysis*) taikymas, žr. Kunilovskaya et al. (2018). Naudojant *Lancaster Stats Tools Online* (Brezina, 2018) programinės įrangos *Words* funkciją nustatytos 10 pirmųjų teigiamų ir neigiamų anototų grožinės literatūros ir mokslo populiarinimo patekstyinių raktinių morfologinių žymų: pvz., 9-ojoje teigiamų raktinių žymų pozicijoje (mokslo populiarinimo literatūros vertimuose) yra žyma PRON.SG.GEN.~NDEF – vadinasi, MP vertimuose, palyginti su MV originalais, dažnesni yra savybiniai įvardžiai (vienaskaitos forma): *mano, tavo*.

Kaip minėta įvade, morfologinėms ypatybėms tirti naudota morfologiškai anotauta ORVELIT versija, anotauta automatiškai naudojant *semantika.lt* anotatorių. *Semantika.lt* lemų ir kalbos dalių automatinio anotavimo rodikliai yra pakankamai geri (lemų – ~98,0 proc., kalbos dalių – ~95,3 proc.), kad taip suanotuotą tekstyną būtų galima naudoti tyrimui, tiesa, morfologinių kategorijų anotavimo tikslumo rodikliai yra žemesni – ~86,8 procentų (Kapočiūtė-Dzikiėnė et al., 2017). Vis dėlto atsižvelgiant į lietuvių kalbai būdingą morfologinį daugiareikšmiškumą (1 mln. žodžių apimties lietuvių kalbos tekстыne 47 % žodžių formų yra daugiareikšmės (Rimkutė, 2006)), morfologinių kategorijų automatinio anotavimo tikslumas ~86,8 proc. yra pakankamai aukštas ir leidžiantis nustatyti morfologinių kategorijų vartojimo tendencijas ORVELIT tekстыne.

Savarankiškų ir tarnybinių kalbos dalių vartoseną originalioje ir vertimų kalboje

Toliau gilinamasi į vertimų kalbos morfologinių ypatybių panašumus ir skirtumus nuo originalo kalbos. Pirmiausia aptariamas savarankiškų ir tarnybinių kalbos dalių pasiskirstymas ORVELIT patekstyniuose, o vertimų kalbos ypatybės iliustruojamos pavyzdžiais iš ORVELIT tekstyno.



1 pav. Savarankiškų kalbos dalių pasiskirstymas ORVELIT patekstyniuose

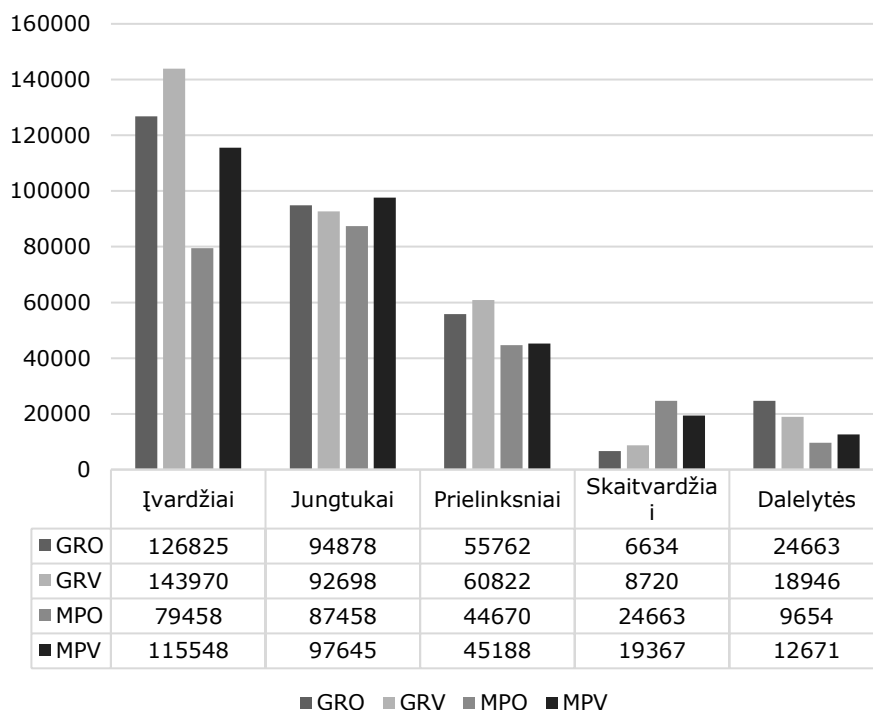
Tyrimo rezultatai rodo, kad visų savarankiškų kalbos dalių dažnių skirtumai originaluose ir vertimuose yra statistiškai reikšmingi. Teigiama, kad lietuvių kalboje yra įprastesnė veiksmažodinė, o anglų kalboje – daiktavardinė raiška, todėl redaktoriai pataria lietuvių kalbos nedaiktavardinti, išlaikyti lietuvių kalbai būdingesnę raišką (Pažūsis 2014, p. 380), pvz., denominalizacijos atvejais, veiksmažodinį daiktavardį keičiant veiksmažodžiu. Tačiau po šio tyrimo paaiškėjo, kad ORVELIT abiejų žanrų

vertimuose mažiau daiktavardžių (GRV: LL = 865.47; MPV: LL = 9913.66) ir būdvardžių (GRV: LL = 572.25; MPV: LL = 771.81), bet daugiau veiksmažodžių (GRV: LL = 535.37; MPV: LL = 1823.12), beje, atotrūkis tarp MPO ir MPV daiktavardžių ir veiksmažodžių skaičiaus didesnis nei tarp GRO ir GRV (plg. 1 pav.).

ORVELIT-o GRO ir MPO dalyje atsispindi lietuvių rašytinės kalbos tyrimų rezultatai (Rimkutė, 2006; Brinkutė, 2017): abiejuose patekstyniuose pirmi – daiktavardžiai, antri – veiksmažodžiai, bet GRO dalyje atotrūkis tarp veiksmažodžių ir daiktavardžių labai mažas, o MPO daiktavardžių gerokai daugiau nei veiksmažodžių. ORVELIT-o MPV patekstynyje, kaip ir MPO, taip pat daugiausia pavartota daiktavardžių, antri pagal dažnumą – veiksmažodžiai. Tokie rezultatai nestebina, nes, tikėtina, mokslo populiarinimo srityje daiktavardžiais perteikiami pavadinimai, terminai. Tačiau MPO daiktavardžių gerokai daugiau nei MPV patekstynyje, vadinasi, vertimuose daiktavardžių fiksuojama daug mažiau. Galima paminėti, kad MPV dalyje, palyginti su MPO dalimi, daug mažiau vartojama ir vardažodžių kilmininko formų (žr. kitą poskyrį), o juk kilmininku lietuvių kalboje reiškiami rūšiniai santykiai, pvz., terminuose: *apygardos teismas*, *apeliacinio teismo pirmininkas*. Vadinasi, nors MPO ir MPV patekstyniuose daiktavardžių pavartota daugiau nei veiksmažodžių, MPV daiktavardžių skaičius mažesnis. Raktinių morfologinių žymų duomenys rodo, kad tikrinių daiktavardžių GRV ir MPV dalyse palyginti su GRO ir MPO dalimis pavartota mažiau.

O štai pasižiūrėjus į būdvardžių irrieveiksmių pasiskirstymą matyti, kad GRV patekstynyje po veiksmažodžių ir daiktavardžių trečioje vietoje pagal dažnumą yrarieveiksmiai, po jų – būdvardžiai; MPV patekstynyje būdvardžių irrieveiksmių skaičius apylygis – jie čia užima trečiąją vietą po daiktavardžių ir veiksmažodžių. Palyginus su originalios kalbos patekstyniais,rieveiksmių mažiau grožinės literatūros (LL = 121.99), bet daugiau mokslo populiarinimo (LL = 2034.71) vertimuose. Apskritai GROrieveiksmių pavartota gerokai daugiau nei būdvardžių, ir ta pati tendencija išlieka ir GRV patekstynyje, - tai galime vertinti kaip žanro ypatybes. Mokslo populiarinimo tekstuose būdvardžių, kaip ir daiktavardžių, gerokai daugiau, bet vertimuose būdvardžių, kaip ir daiktavardžių, taip pat užfiksuota mažiau. Grožinėje literatūroje pastebėtos tokios pačios koreliacijos, taigi sumažėjus daiktavardžių sumažėja ir būdvardžių. Mokslo populiarinimo vertimuose matyti sąsajų tarp

veiksmazodžių irrieveiksmių – čia daugiau ir veiksmazodžių, irrieveiksmių. Grožinės literatūros vertimuose veiksmazodžių pavartota daugiau, irrieveiksmių – gerokai mažiau. Tolimesniuose tyrimuose lygiagrečiamame tekstyne būtų galima tirti, kiek verčiant tekstus iš anglų į lietuvių kalbą išlaikoma veiksmazodinė raiška ten, kur ji turėtų būti išlaikoma.



2 pav. Tarnybinių kalbos dalių pasiskirstymas ORVELIT patekstyniuose

Nors tarnybinių kalbos dalių proporcija lyginant su savarankiškomis kalbos dalimis vertimuose yra didesnė, ne visos tarnybinių žodžių kategorijos vertimuose dažnesnės (žr. 2 pav.). Kaip ir savarankiškų kalbos dalių atveju, skirtumai tarp visų tarnybinių kalbos dalių kategorijų vertimuose ir originaluose yra statistiškai reikšmingi. Galima teigti, kad GRV ir MPV vienareikšmiškai daugiau įvardžių (GRV: LL = 1087.50; MPV: LL = 6908.19), o grožinės literatūros vertimuose gerokai daugiau pavartota ir prielinksnų (GRV: LL = 220.05). Pagal dažnį GRO ir GRV tarnybinės dalys vertimuose

išsidėsto taip: įvardžiai-jungtukai-prielinksniai-dalelytės-skaitvardžiai. Nors įvardžiai padeda atpažinti grožinės literatūros tekstus, nes jais reiškama naratyvumo funkcija (Utka, 2006), vis dėlto statistiškai reikšmingai daugiau jų pavartota ne tik GRV, bet ir MPV. MPO pagal dažnumą pirmoje vietoje yra jungtukai, įvardžiai eina po jų, po to – prielinksniai, skaitvardžiai, dalelytės. MPV dalyje pirmoje vietoje jau matome įvardžius (MPV: įvardžiai-jungtukai-prielinksniai-skaitvardžiai-dalelytės). MPV atveju mes negalime pateikti argumento, kad įvardžių dažnumas sietinas su registro ypatybėmis, be to, kaip jau minėta, įvardžių dažnio atotrūkis tarp MPO ir MPV – itin didelis (LL = 6908.19). Vadinasi, abiejų ORVELIT patekstyinių duomenys rodo aiškią tendenciją, kad vertimuose įvardžiai yra vartojami dažniau nei originaliojoje lietuvių kalboje.

Ši tendencija gali būti siejama su vienu iš dažnų kalbos dalies keitimų verčiant, vadinamąja pronominalizacija (daiktavardžio keitimas įvardžiu), tačiau neatitinka kitų tyrėjų pastebėtų tendencijų, grįstų daugiausia kokybinėmis, ne kiekybinėmis išvalgomis, pavyzdžiui, L. Pažūsis teigia, kad:

<...> vertimuose iš anglų kalbos į lietuvių kalbą, atrodo, daug dažniau pastebime atvirkštinį (pa)keitimą, t. y. OK daiktavardiškųjų įvardžių (pa)keitimą VK daiktavardžiais ar daiktavardiniais žodžių junginiais, vadinamąją depronominizaciją. Depronominizacija yra dažnesnė negu pronominalizacija, matyt <...> dėl to, kad įvardžio, kaip vienos abstrakčiausių kalbos dalių, konkretus leksinis turinys paaiškėja tik kontekste, o įvardžių gausa tekstą daro neaiškų, sunkiau suprantamą. (Pažūsis, 2014, p. 370)

Kituose tyrimuose (Seguinot, 1988; Weissbrod, 1992, p. 153; Goethals, 2007, p. 91, cit. iš Pažūsis, 2014, p. 370) taip pat apibendrinama, kad originalo tekste pasitaikantį įvardį vertimo tekste dažniau keičia daiktavardis ar daiktavardinis junginys, o ne atvirkščiai.

Vadinasi, įvardžius vertimuose lietuvių kalba verta patyrinėti plačiau, juo labiau, kad jų perviršį vertimuose rodo ir raktinės morfologinės žymos: ir GRV dalyje, ir MPV dalyje daugiau vartojama asmeninių įvardžių (*aš, tu, mes, jūs...*), o MPV dalyje dažniau nei MPO pasirodo ir savybiniai įvardžiai (*mano, tavo, savo*). Įdomu pasižiūrėti šių tarnybinių žodžių pasiskirstymą ir kiekvieno patekstyinio dažninio sąrašo viršūnėje. 4 lentelėje pateikiame iš dažninių žodžių sąrašų viršūnių išrinktus pirmuosius 10 įvardžių (formų):

Pirmi 10 įvardžių (formų) dažniniuose žodžių sąrašuose

GRV	GRO	MPV	MPO
JIS (11594)	JIS (7092)	SAVO (6252)	SAVO (3769)
AŠ (8979)	AŠ (6106)	JIS (4780)	JŪ (3175)
JI (7427)	JI (5602)	JIE (3425)	JIS (2753)
MAN (5117)	JO (4471)	JŪ (3075)	JO (2575)
JO (4598)	SAVO (4129)	AŠ (2933)	JOS (2006)
SAVO (3697)	JOS (3678)	JO (2809)	JIE (1833)
JŪ (3558)	MAN (3523)	JI (2638)	JI (1287)
MANO (3500)	TU (2772)	JŪSŪ (2405)	MŪSŪ (1208)
JOS (3364)	MANO (2535)	JOS (2112)	JŪ (1095)
JIE (3201)	JŪ (2329)	MANO (1847)	AŠ (946)

Kaip rodo duomenys, ir originaluose, ir vertimuose vartojama daug tų pačių įvardžių, tik skiriasi jų pozicija dažniniame sąraše ir, žinoma, dažnis, plg.: *jis*: GRV – 11594, o GRO – 7092, MPV – 4780, o MPO – 2753. Taigi vertimuose tų pačių įvardžių dažnis didesnis, išskyrus *jo*, *savo*, *jos*, *tu* (dažnesni GRO patekstinyje); *jū*, *mūsų*, *ji* (dažnesni MPO patekstinyje). Taip pat matyti žanrų skirtumai, kai, pavyzdžiui, mokslo populiarinimo vertimuose net dvigubai dažnesnis savybinis įvardis *savo* yra retesnis grožinės literatūros vertimuose; literatūros vertimuose apskritai didesnė įvardžių asmenų įvairovė dėl veikėjų dialogų (pvz., daugiau antrojo asmens įvardžių) ar pasakotojo pozicijos (pvz., ar naratyvas pirmojo ar trečiojo asmens). Tai jau nustatyta kitų tyrimų, pvz., teksto funkcijų nustatymo metodikoje (Utka, 2006) grožinės literatūros tekstas atpažįstamas pagal jam tipišką naratyvumo funkciją, kurios rodikliai – 3-iojo asmens įvardžių (pvz., *jis*, *ji*) bei erdvės santykius reiškiančių prielinksnių dominavimas.

Pastebėtos asmeninių įvardžių vartojimo tendencijos pagrindžia prielaidas dėl anglių, kaip analitinės kalbos, įtakos lietuvių, kaip sintetinei, kalbai. Žinoma, kad asmeniniai įvardžiai veiksnio pozicijoje (pvz. (*Aš*) *einu namo*) lietuvių kalboje gali būti praleidžiami, nes implikuojami veiksmožodžio

galūne (Ambrasas, 2006, p. 189). Asmeninis įvardis *aš* grožinės literatūros vertimuose pavartotas 8979, o originalioje – 6106 kartus. Viena vertus, šiuos skirtumus gali lemti patekstylių turinio ypatybės, kai vertimų patekstylių kūriniuose daugiau pirmojo asmens raiškos. Kita vertus, gali būti, kad analitinei anglų kalbai privalomas veiksnys vertimuose išlaikomas dažniau negu originalioje lietuvių kalboje, plg. šiuos pavyzdžius iš GRV patekstylio (įvardis, kuris galėtų būti praleistas, paryškintas):

*Nieko kito **aš** nesitikėjau, bet ...
Aš spoksojau į jungiklius ir ...
Kaip žinote, mokykloje **aš** esu jos klasės auklėtoja.*

Kaip nurodo L. Pažūsis (2014, p. 601), anglų kalboje nėra pasirinkimo, vartoti asmeninius įvardžius ar ne, todėl jie tokie dažni vertimuose, daug dažnesni negu originaliuose lietuvių kalbos tekstuose (žr. Buitkienė, 2010, p. 30).

Pastebėta ir dar viena įdomi tendencija – perteklinė savybinio įvardžio *savo* vartoseną mokslo populiarinimo vertimuose (plg. MPO – 3769, o MPV – 6252). Manoma, kad taip išaugusį šio įvardžio dažnumą gali lemti anglų kalbai būdingos struktūros „savybinis įvardis + daiktavardis“ (pvz., *my/your/his/her/its/our/your/their bag*) pažodinis vertimas į lietuvių kalbą junginiu „*savo* + daiktavardis“, nors dažnu atveju pakaktų tik daiktavardžio. Plg. Tekorienė (2000, p. 72) teigia, kad kalbant apie žmogaus kūno dalis, jo drabužius ir kitus jam priklausančius daiktus, anglų kalboje daug dažniau negu lietuvių vartojami savybiniai įvardžiai, ir pateikia tokius pavyzdžius: *She took off **her** gloves = Ji nusiėmė pirštines; She shook **her** head = Ji papurtė galvą* (ibid.). Toliau pateikiame keletą pavyzdžių iš MPV patekstylio, kuriuose savybinis įvardis galėtų būti praleistas:

*Kiekvieną kartą, kai ją apimdavo pagunda pagalvoti „Prarandu **savo** atmintį“...
Atiduočiau **savo** dešinę ranką, jei mano duktė pasveiktų.
... atsakė, kad krištolinis rutulys seniai garsėja **savo** magiškomis galiomis.*

Perteklinė savybinio įvardžio „savo“ vartoseną galima ir tais atvejais, kai verčiant pažodžiui kopijuojama angliška struktūra „priel. „with“ + savybinis įv. + daikt.“, pavyzdžiui:

*... pasiųs dar pas kitą žmogų, taip pat nežinantį, kur turėtų kreiptis su **savo** skundu.
Tėvai negali suprasti, kodėl jų draugai su **savo** mažyliais gyvena normalų ramų gyvenimą...
Kolumbas su **savo** ispanais išsilaipino arčiau Amerikos branduolio...
Su **savo** lankais, strėlėmis ir ietimis karybą jie pavertė meno forma.*

Išsamesni prielinksnių vartosenos tyrimai taip pat leistų patikrinti, ar (tam tikro / tam tikrų) prielinksnių dažnumą lemia anglų kalbos įtaka. Žinia, kad angliškus prielinksnius lietuvių kalboje atitinka linksniai ir prielinksniai (Tekorienė, 2000, p. 203), prielinksniai vartojami su trimis linksniais – kilmininku (daugiausia), galininku ir įnagininku (tik keletas) (Ambrasas, 2006, p. 439). Pavyzdžiui, lietuvių kalboje įnagininkui reikšti vartojami keturi prielinksniai: *su*, *sulig*, *ties* ir *po* (Ambrasas, 2006, p. 437–453). Kai įnagininku reiškiamas veiksmo įrankis, dažniau vartojamas pats įnagininkas (pvz., *rašo (su) pieštuku*) (ibid). Pažvelgę į ORVELIT konkordansus matome, kad reiškiant veiksmo įrankį ar kitas funkcijas pasirinkta prielinksninė konstrukcija ten, kur galėtų būti vien tik įnagininkas:

*Jam teko šokti **su** parašiotu virš džiuuglėmis užėjusiu Naujosios Gvinėjos kalnų.
Kai kurie berniukai **su** juodom kepurėm, ...
...jos akį patraukė gražus ispaniškas odos rankinukas **su** ilgu dirželiu.*

Sprendžiant iš didesnio prielinksnio *su* dažnumo GRV (įvairiomis reikšmėmis prielinksnis *su* GRO dalyje pavartotas 6352 kartus, o GRV – 6816 kartų), tikėtina, kad vertimuose vartoseną su prielinksniu yra dažnesnė. Žinoma, norint išsiaiškinti, kiek anglų kalbos įtaka reikšminga vertimuose vartojant prielinksnius, reikėtų daug gilesnės kokybinės analizės, aprėpiančios visas prielinksnių kategorijas ir jų funkcijas.

Kitų tarnybinių žodžių vartojimas vertimuose skiriasi priklausomai nuo žanro: grožinės literatūros vertimuose matomas mažesnis

jungtukų (LL = 25.19) ir dalelyčių (LL = 751.24), o mokslo populiarinimo tekstuose atitinkamai dažnesnis jų vartojimas (LL = 615.52; LL = 424.91). Ir atvirkščiai: grožinės literatūros vertimai pasižymi gausniu skaitvardžių (LL = 284.43) ir jaustukų (LL = 42.11) vartojimu lyginant su originalia literatūra; mokslo populiarinimo vertimuose šių tarnybinių kalbos dalių yra mažiau (LL = 611.13; LL = 526.91).

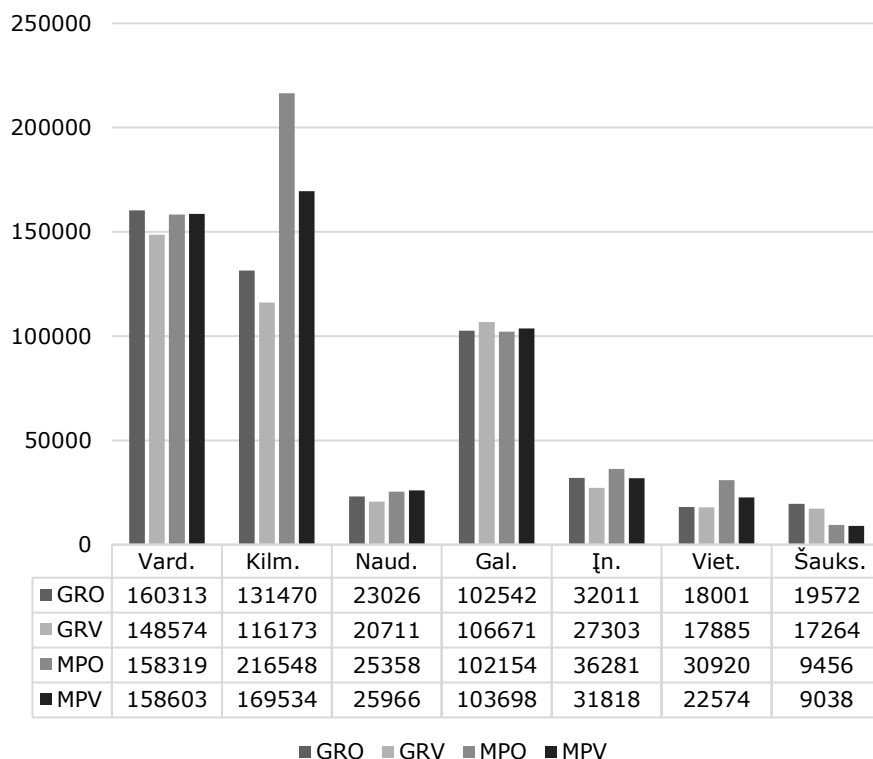
Vardažodžių²⁹ linksniai

Iš R. Brinkutės (2017) tyrimo rezultatų matyti, kad publicistikoje, kaip ir moksliniame bei administraciniame stiliuje, dažniausiai vartojamas vardažodžių vienaskaitos ir daugiskaitos kilmininkas (34,05 proc.), antras pagal dažnumą vardažodžių vienaskaitos ir daugiskaitos vardininkas (32,11 proc.), o trečias pagal dažnumą – vardažodžių vienaskaitos ir daugiskaitos galininkas (16,6 proc.). ORVELIT MPO dalyje linksniai pagal dažnumą išsidėsto panašiai: kilmininkas (21,68 proc.), vardininkas (15,85 proc.), galininkas (10,23 proc.). Morfologiškai anotuoto tekstyno grožinėje literatūroje vardininkas – pirmoje vietoje (38,84 proc.), o kilmininko ir galininko yra fiksuojama apylygiai (atitinkamai 21,99 proc. ir 21,37 proc.) (Brinkutė, 2017, p. 66). ORVELITO GRO patekstyvio pirmoje vietoje taip pat matome vardininką (16 proc.), toliau eina nedaug vienas nuo kito atitrūkę kilmininkas (13,13 proc.) ir galininkas (10,24 proc.).

ORVELITO vertimų duomenys patvirtina ir sakytinėje (Dabašinskienė, 2009), ir rašytinėje kalboje (Rimkutė, 2006; Brinkutė, 2017) nustatytas tendencijas, kad gramatiniai linksniai vartojami dažniau nei semantiniai; pagrindinę grupę sudaro vardininkas, kilmininkas, galininkas, periferinę – naudininkas, įnagininkas, vietininkas, šauksmininkas. Tačiau, kaip ir originalo kalbos patekstyniuose, GRV ir MPV gramatinių linksnių išsidėstymas pagal dažnumą skiriasi: vardininkas, kilmininkas, galininkas (GRV) ir kilmininkas, vardininkas, galininkas (MPV). Tai patvirtina ankstesnių tyrimų rezultatus (plg. Dabašinskienė, 2009, p. 10) apie rašytinę ir sakytinę kalbos atmainas: grožinis stilius artimesnis sakytinei kalbai, todėl GRV, kaip ir

²⁹ Vardažodžių linksnių analizė apima daiktavardžių, būdvardžių, įvardžių, skaitvardžių, dalyvių linksnius.

sakytinėje, vardininkas yra dažniausias linksnis; mokslo populiarinimo stilius artimesnis rašytinei kalbai, todėl MPV, kaip ir rašytinėje, kilmininkas yra dažniausias linksnis.



3 pav. Vardažodžių linknių pasiskirstymas

Vertinant pagal statistinio reikšmingumo rodiklius, paaiškėja, kad vertimuose vienareikšmiškai mažiau kilmininko ir įnagininko linksnio formų: atotrūkis tarp kilmininko dažnumo MPV ir MPO dalyse daug didesnis nei tarp GRV ir GRO (płg. GRV: LL = 944.40; MPV: LL = 5495.71), o štai įnagininkas panašiai retesnis abiejuose vertimų patekstyniuose (GRV: LL = 373.74; MPV: LL = 269.80). MPV patekstynyje vietininkas daug retesnis nei MPO, o GRV ir GRO vietininko vartosenos skirtumai nėra statistiškai reikšmingi. GRV patekstynyje, palyginti su GRO, rečiau vartojamas vardininkas (LL = 445.45), naudininkas (LL = 122.42) ir šauksmininkas (LL = 144.54). Abiejų patekstynių

vertimuose šiek tiek daugiau vartojama galininko (GRV: LL = 81.79; MPV: LL = 21.09).

Morfologinių ypatybių analizei naudotas automatiškai morfologiškai anotuotas tekstynas, todėl svarbu paminėti, kad tam tikrų formų anotavimas gali būti ne visai tikslus. Pvz., Boizou et al. (2018, p. 22) teigia, kad apie 5 proc. *Semantika.lt* anotatoriaus žymėtų šauksmininko formų turėtų būti anotuojamos kaip vardininko formos, vadinasi, vardininko pavartojimo atveju gali būti daugiau nei dabar rodoma 3 paveiksle. Kita probleminė formų pora – mot. g. vienaskaitos kilmininko formų sutapimas su mot. g. daugiskaitos vardininku, plg. *gražios dienos*. Šių formų anotavimo klaidos buvo tarp dažniausių, analizuojant mokslinių tekstų automatinio anotavimo tikslumą (Boizou et al. 2018, ten pat). Taigi reikia turėti omenyje, kad kilmininko ir vardininko dažniai, pateikiami šiame tyrime, gali būti su tam tikra paklaida.

Kaip minėta ankstesniame skyriuje, abiejų patekstynių vertimuose daiktavardžių fiksuojama mažiau. Todėl galbūt mažiau stebina gana didelis atotrūkis tarp kilmininko vartosenos MPV ir MPO dalyse (kiek mažesnis atotrūkis užfiksuotas GRV ir GRO dalyse). Atrodytų, mokslo populiarinimo tekstuose tikėtumės daug terminijos, pavadinimų, kur rūšiniai santykiai lietuvių kalboje reiškiami kilmininku, bet jeigu apskritai vertimuose pavartota mažiau daiktavardžių, tai tikėtina, kad sumažės ir kilmininko formų. Kadangi skiriamos objekto, subjekto, tikslo ir pažymimojo kilmininko reikšmės (Labutis, 1998), būtų įdomu iširti, kaip tos reikšmės atsispindi vertimuose: pavyzdžiui, galbūt retesnė yra dalies kilmininko vartoseną, t. y. gal vertimuose kilmininkas daug rečiau vartojamas objekto daliai reikšti (*pateikė pasiūlymų* (vs. *pateikė pasiūlymus*). Reikšme „neapibrėžtas kiekis“ kilmininkas sudaro priešpriešą vardininkui bei galininkui (Judžentis, 2012, p. 76). Galėtume daryti prielaidą, kad GRV kilmininko atotrūkis nuo GRO nėra toks didelis kaip tarp MPO ir MPV, nes GRV gerokai rečiau vartojamas vardininkas, o MPO ir MPV vardininko vartoseną labai panaši. Kitaip sakant, kai vardininko mažiau – daugiau erdvės kilmininkui. Be to, kaip matome, galininko vertimuose yra šiek tiek daugiau. Siekiant patikrinti šias prielaidas reikia analizuoti kilmininko reikšmes ir tikslingai tirti visų trijų linksnių vartoseną. Panašiai reikėtų analizuoti ir kitus vertimuose pastebėtus linksnių vartojimo ypatumus: vertimuose mažiau vartojamas linksnis yra įnagininkas, taip pat naudininkas (tik GRV) ir vietininkas (tik MPV).

Veiksmažodžio kategorijos

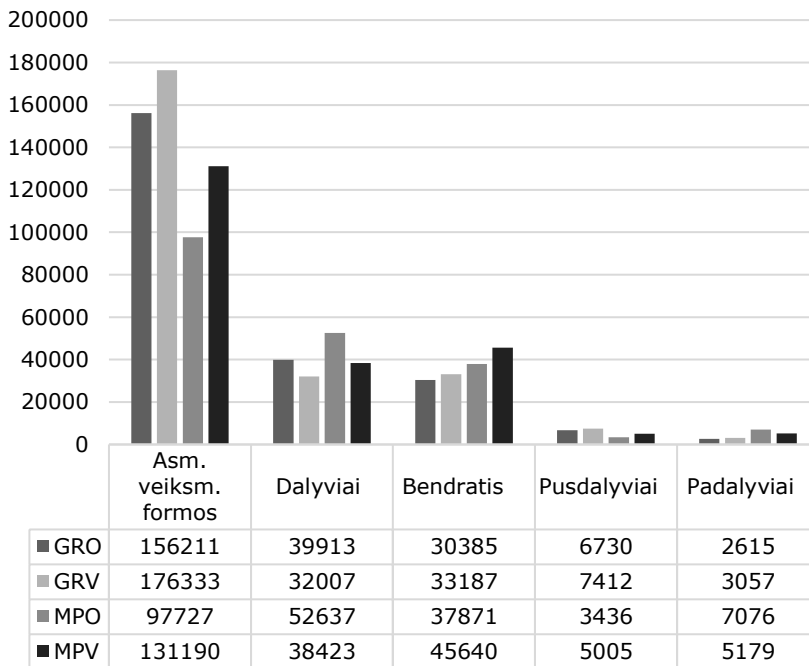
Kaip minėta, abiejų patekstyinių vertimuose (palyginti su GRO ir MPO) užfiksuota statistiškai daugiau veiksmažodžių. Šiame poskyryje aptariamas veiksmažodžio asmenuojamųjų ir neasmenuojamųjų formų, dalyvių rūšių pasiskirstymas. Iš kitų veiksmažodžio kategorijų ar formų paminime tokias, kurios buvo atpažintos kaip vertimams būdingesnės ar mažiau būdingos tiriant raktinių morfologinių žymų metodu.

GRV ir MPV pavartota statistiškai reikšmingai daugiau asmenuojamųjų veiksmažodžio formų (GRV: LL = 1219.72; MPV: LL = 5085.96). Abiejų patekstyinių vertimuose užfiksuota daugiau veiksmažodžių, todėl buvo galima tikėtis ir didesnio asmenuojamųjų formų kiekio. Apskritai, GRO patekstynyje asmenuojamosios formos užima didesnę procentinę dalį nei MPO patekstynyje (atitinkamai 15,60 proc. ir 9,79 proc.).

Pasižiūrėjus į veiksmažodžio BŪTI formų (*esu, esame, esi, esate, yra*) dažnumą, išryškėja, kad šios formos daug dažnesnės MPV dalyje, pvz.: *yra*: MPO – 5915, MPV – 7970. Konkrečiuose pavyzdžiuose, ypač tariniuose su neveikiamosios rūšies dalyviais, *yra* galima būtų praleisti, nors, žinoma, reikėtų pirmiausia ištirti, ar originaliojoje kalboje tokie praleidimai iš tikrųjų dažni, pvz.:

*...nemėgsta girdėti, deja, **yra** sakomi pernelyg dažnai <...>*
*Jukatano majų žodynuose **yra** minimi ir du šokolado <...>*

MPV statistiškai daugiau nei MPO vartojama bendraties ir pusdalyvio formų. Bendraties daugiau užfiksuota abiejuose patekstyniuose, tačiau MPV – daugiau nei GRV (GRV: LL = 123.74; MPV: LL = 765.05). Greičiausiai MP vertimuose dažnesni tariniai su bendratimi: *galėjo / turi būti, émé / pradėjo kilti, reikia / galima pasirinkti*; galbūt rečiau bendratis vartojama su daiktavardžiais, pvz.: *būdas atlikti, noras patikti, tikimybė tapti*. Vis dėlto būtų svarbu išsamesnį bendraties vartojimo tyrimą susieti su neveikiamosios rūšies būtojo laiko dalyvių vartosenos tyrimu, nes, kaip nurodo Boizou et al. (2018), šios formos yra tarp tų, dėl kurių padidėja automatinio anotavimo klaidų tikimybė (plg. *galima pateikti pavyzdį* ir *laiku pateikti klausimai*). Vadinasi, siekiant patikimesnių išvadų, reikėtų tikrinti anotavimo tikslumą.



4 pav. **Asmenuojamosios ir neasmenuojamosios formos**

Pusedalyvių ir padalyvių vartojimo situacija gana skirtinga: pusedalyviai būdingi GRO kaip stiliaus ypatybė (pasakojimuose), MPO jų dvigubai mažiau; su padalyviais – atvirkščiai: jie gerokai būdingesni kaip MPO stiliaus ypatybė, GRO jų vartojama mažiau. Vertimuose pusedalyvių užfiksuota daugiau abiem atvejais, tiesa, MPV – gerokai daugiau nei GRV (GRV: LL = 32.95; MPV: LL = 301.62). Tačiau įdomesnė yra padalyvių vartosena vertimuose: palyginus su GRO dalimi GRV dalyje padalyvių daugiau, ir tas padidėjimas, nors nežymus, yra statistiškai reikšmingas (LL = 34.51), o MPV padalyvių vartojama gerokai mažiau (LL = 284.98). Padalyvio kaip gramatinės kategorijos anglų kalboje nėra, verčiant padalyvius iš lietuvių į anglų kalbą vartojamos kitos priemonės: šalutiniai įvairių aplinkybių sakiniai, prielinksninės daiktavardžio konstrukcijos, gerundijaus konstrukcijos ir kt. (plačiau Tekorienė, 2011, p. 6–7). Todėl galima prielaida, kad verčiant iš anglų kalbos į lietuvių padalyviai nėra tiek dažni, kiek originaliojoje lietuvių kalboje, nes padalyvius atitinkančioms anglų kalbos konstrukcijoms vertėjas galbūt renkasi tokias priemones, kurios panašesnės į anglų kalbą: pvz., verčiant šalutinius laiko aplinkybės sakinius su

when, while, since, išlaikoma ta pati konstrukcija – šalutinis sakinyss jungtuku *kai*; verčiant prielinksninės daiktavardžio konstrukcijas (*after some time*), pasirenkamas ne padalyvis (būtų: *praėjus kuriam laikui*), o prielinksninė konstrukcija (būtų: *po kurio laiko*). Įdomu iširti, galbūt padalyvis visada ar dažniausiai pasirenkamas kaip atitikmuo konkrečiais atvejais, tarkim, verčiant gerundijų, plg.: *Mes be niekur nieko matome, dėl ko, perpiovus magneta, atsiranda du nauji poliai. We see immediately why cutting a magnet causes two new poles to appear on the new ends* (pavyzdžiai iš Tekorienė, 2011, p. 12).

5 lentelė.

Dalyvių rūšys

	GRV	GRO	MPV	MPO
Veikiamoji	21430 (2,14%)	26577 (2,65%)	17954 (1,81%)	19736 (1,98%)
Neveikiamoji	10396 (1,04%)	13150 (1,31%)	19964 (2,01%)	32358 (3,24%)

Kitas svarbus aspektas yra dalyvių dažnumas vertimuose: GRV dalyvių vartojama statistiškai reikšmingai mažiau (LL = 870.28), MPV dalyvių skaičiaus sumažėjimas taip pat labai reikšmingas: nuo 52637 (5.27 proc.) MPO patekstynyje iki 38423 (3.87 proc.) MPV, plg. LL = 2154.04. Dėl šių rezultatų buvo svarbu pasižiūrėti ir į dalyvių rūšių pasiskirstymą (žr. 5 lentelę). GRV mažiau pavartota veikiamosios rūšies dalyvių (pvz., *plintančia propaganda*), ta pati tendencija pastebima ir MPV. Neveikiamosios rūšies dalyvių (pvz., *nepagailėta karčių žodžių*) GRV vartojama mažiau (LL = 322.65) nei GRO, o MPV rezultatai, palyginti su MPO, dar iškalbingesni (LL = 2899.44). Taigi iš esmės matyti tendencija, kad vertimuose yra vartojama mažiau dalyvių: abiejų žanrų tekstuose, bet ypač grožiniuose tekstuose – mažiau veikiamosios rūšies dalyvių. Raktinės žymos rodo, kad MPV tarp neigiamų raktinių žymų yra neveikiamosios rūšies esamojo laiko mot. g. daugiskaitos dalyvių formos (pvz., *taikomos taisyklės*). Gali būti, kad retas šių formų vartojimas MPV yra susijęs su dažnu asmenuojamųjų formų vartojimu tame pačiame MPV (4 pav.), kai originaluose linkstama pasirinkti raišką su dalyviu, o vertimuose – su asmenuojamąja forma.

Spėjama, kad verčiant šalutinius pažyminio sakinius su *which* iš anglų kalbos pernelyg dažnai pasirenkama versti pavartojant įvardį *kuris* + asmenuojamąją veiksmažodžio formą. Plg.: GRV įvardžio *kuris* dažnumas – 920, o GRO – 804 kartai; MPV įvardžio *kuris* dažnumas – 1254, o MPO – 911 kartų. Pateikiame kelis pavyzdžius iš ORVELIT MPV patekstinio, kur galima būtų vartoti dalyvius (*pirkėjo, įsigijusio...; jei įsigijau/įsigijęs neveikiantį kompiuterį; berniuką, norintį daug saldainių*):

*... kiek nutolę nuo pirkėjo, kuris įsigyja mūsų gaminį vėliau...
...jei įsigijau kompiuterį, kuris neveikia, aš norėsiu, kad...
...įsivaizduokite dvimetį berniuką, kuris nori daug šokoladinių sausainių...*

Kita priežastis, dėl kurios vertimuose matome didesnę asmenuojamųjų formų skaičių, gali būti sietina su taprkultūriniais žanrų ypatumais. MPV dalyje daug populiariosios psichologijos, pedagogikos, patarimų knygų, kur su skaitytoju bendraujama kreipiantis „tu“, „Jūs“ arba dalinantis savo patirtimi („aš“ pozicija): *pagalvokite, kaip... turbūt niekada neįsivaizdavote...* ir pan. Raktinės morfologinės žymos rodo, kad MPV daugiau nei MPO pavartota liepiamosios nuosakos formų (*liaukis, įsivaizduokite*), tariamosios nuosakos (*įsitikintum*), esamojo (*prisimenate, mokaisi*) ir būsimojų laiko (*mokysitės*) formų. Vadinas, MP vertimų kalboje išryškėja dialoginis stilius, orientacija į skaitytoją, ir tai turbūt yra mažiau būdinga originaliems lietuvių autorių rašytiems mokslo populiarinimo tekstams, kur daug didesnis neveikiamosios rūšies dalyvių skaičius, o pastarieji gali būti vartojami sudėtinuose tariniuose (*buvo įkurtas, buvo kalbama*). Žinoma, vienas iš veiksmių gali būti tematika.

Raktinės morfologinės žymos išryškino tam tikras laiko ir nuosakų formas: kaip minėta, MPV dažnesnės tam tikros liepiamosios ir tariamosios nuosakos formos, tam tikros esamojo ir būsimojų laiko formos. GRV dalyje palyginti su GRO dalimi daugiau pavartota būtent 1-ojo asmens vienaskaitos arba daugiskaitos būtojo laiko formų (*padariau, atsidūrėme, neradome*), tariamosios nuosakos formų (*būtume*). Visos šios formos yra naratyvui būdingos būtojo laiko veiksmažodžių formos, tačiau reikia pasakyti, kad greičiausiai šios formos rodo ne vertimo ir originalo skirtumus, o žanrinius skirtumus, todėl šiam tyrimui nėra tokios aktualios.

Apibendrinamosios pastabos

Šiame tyrime, naudojant ORVELIT tekstyną, bandyta nustatyti, kokios leksinės ir morfologinės ypatybės išryškėja vertimų kalboje, kuo ji skiriasi nuo originaliųjų lietuviškųjų tekstų kalbos. Tyrimo rezultatai rodo, kad vertimai iš anglų kalbos turi jiems būdingų bendrųjų ypatybių, taip pat konkretaus žanro ar kitų veiksmų nulemtų ypatybių. Leksinių ypatybių analizė atskleidžia, kad tiek grožinės literatūros, tiek mokslo populiarinimo vertimų kalba pasižymi mažesne žodyno ir informacinės įvairovės. ORVELIT vertimams būdinga mažesnis informacinis tankis; didesnis tarnybinių žodžių skaičius; trumpesni sakiniai; mažesni žodingumo rodikliai ir retų žodžių vartoseną (ypač grožinės literatūros vertimuose). Raktinių žodžių palyginimas atskleidė, kad vertimuose daugiau skolinių, žodžių, susijusių su anglų kalbos kultūrinėmis realijomis, o mokslo populiarinimo tekstuose dažnesni kai kurių kategorijų asmeniniai ir savybiniai įvardžiai.

Tiriant morfologines ypatybes nustatyta, kad ORVELIT abiejų žanrų vertimuose mažiau daiktavardžių ir būdvardžių, bet daugiau veiksmažodžių. Linksnių analizė atskleidė, kad vertimuose fiksuojama mažiau vardažodžių kilmininko: dėl sumažėjusio daiktavardžių skaičiaus, tikėtina, yra mažiau ir sudėtinių pavadinimų, terminijos, o dėl to retesnis ir kilmininkas. Veiksmažodžio formų analizė atskleidė, kad vertimuose daugiau pavartota asmenuojamųjų veiksmažodžio formų. Pavyzdžiui, bendraties daugiau užfiksuota abiejuose patekstyniuose, tačiau MPV – daug daugiau nei GRV, tai turbūt lemia žanro ypatybės. Vis dėlto reikėtų išsamesnio bendraties tyrimo, kuris būtų susietas su neveikiamosios rūšies dalyvių vartosenos tyrimu, nes dalis formų sutampa dėl morfologinio daugiareikšmiškumo (tai pasakytina ir apie kai kuriuos linksnius, pvz., vardininką ir šauksmininką). Vertimuose užfiksuota daugiau pusdalyvių, tiesa, mokslo populiarinimo vertimuose – gerokai daugiau, nors pagal žanrą pusdalyviai būdingesni grožinei literatūrai. Padalyvių pastebėta daugiau grožinės literatūros vertimuose ir mažiau mokslo populiarinimo vertimuose. Aiškiai matyti, kad vertimuose yra vartojama ir mažiau dalyvių: grožiniuose tekstuose mažiau veikiamosios rūšies dalyvių, mokslo populiarinimo tekstuose mažiau neveikiamosios rūšies dalyvių. Mažesnis dalyvių skaičius vertimuose, manytume, gali būti susijęs su pastebėta

tendencija, kad vertimuose daugiau vartojama asmenuojamųjų veiksmažodžio formų, ypač MPV dalyje.

Tarnybinių kalbos dalių proporcijos originalioje ir vertimų kalboje taip pat skiriasi. Abiejų žanrų vertimuose vienareikšmiškai daugiau įvardžių: grožinėje literatūroje įvardžiai yra žanro ypatybė, todėl faktas, kad mokslo populiarinimo vertimuose įvardžių irgi gerokai daugiau, padeda pripažinti perteklinį įvardžių vartojimą vertimų kalbos ypatybe. Grožinės literatūros vertimuose dažnesni asmeniniai, o mokslo populiarinimo – savybiniai įvardžiai, būtent šių grupių įvardžių vertimas, tikėtina, yra susijęs su anglų, kaip analitinės kalbos, įtaka lietuvių, kaip sintetinei, kalbai. Tikėtina, kad verčiant dažnai perkeliama anglų kalbai privalomos, o lietuvių kalbai pasirenkamos konstrukcijos, ir tai matyti iš ORVELIT randamų pavyzdžių, kai išlaikomi asmeniniai įvardžiai veiksnio pozicijoje; savybiniai įvardžiai perteikiant anglišką „priel. „with“ + savybinis įv. + daikt.“ struktūrą; tariniuose su neveikiamosios rūšies dalyviais paliekama lietuvių kalbai pasirenkama, o anglų kalbai privaloma veiksmažodžio „būti“ forma; automatiškai perkeliama šalutinių pažyminio sakinių „kuris + asmenuojama veiksmažodžio forma“ struktūra, nors galima būtų versti vartojant dalyvines formas.

Reikėtų pastebėti, kad turint tik kiekybinius morfologiškai anotuoto tekstyno duomenis gana sunku vertinti nustatytų vertimų ypatybių priežastis, todėl būtina tolesnė kokybinė ORVELIT tekstyno duomenų analizė, kai kuriais atvejais naudojant ir atitinkamą lygiagretųjį tekstyną. Tikimės, kad nustatytos vertimų kalbos tendencijos, pateiktos įžvalgos ir iškelti klausimai bus naudingi tolesniems vertimų kalbos tyrimams.

Padėka. Straipsnio autorės dėkoja dr. Loic Boizou ir doc. dr. Erikai Rimkutei už pagalbą ir konsultacijas apdorojant morfologiškai anotuotus duomenis.

Literatūros sąrašas

- Ambrazas, V. (2006). Lithuanian grammar. Vilnius: Baltos lankos.
- Bielinskienė, A., Boizou, L., Rimkutė, E. (2017). Lietuvių kalbos morfologiškai ir sintaksiškai anotuoti tekstynai. Bendrinė kalba [elektroninis išteklius], 90, 1–30.

- Boizou, L., Kapočiūtė-Dzikiene, J., Rimkutė, E. (2018). Deeper error analysis of Lithuanian morphological analyzers. *Human language technologies - the Baltic perspective: proceedings of the 8th international conference Baltic HLT*. Amsterdam/ Berlin: IOS Press, 18–25.
- Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge: Cambridge University Press.
- Brinkutė R. (2017). Gramatinių kategorijų pasiskirstymas morfologiškai anotuotame lietuvių kalbos tekste. Magistro darbas. Kaunas: VDU.
- Dabašinskienė, I. (2009). Šnekamosios lietuvių kalbos morfoliginės ypatybės. *Acta Linguistica Lithuanica LX (90)*, 1–15.
- Judžentis A. (2012). *Lietuvių kalbos gramatinės kategorijos*. Vilnius: VU leidykla.
- Kapočiūtė-Dzikiene, J., Rimkutė, E., & Boizou, L. (2017). A Comparison of Lithuanian morphological analyzers. K. Ekšteina, & V. Matoušek, the *Proceedings of the TSD 2017: Text, Speech, and Dialogue: 20th International Conference*. Springer, 47–56.
- Kunilovskaya M., Morgoun N., Pariy A. (2018). Learner vs. professional translations into Russian: Lexical profiles. *The International Journal for Translation & Interpreting Research*, 10 (1), 33–52.
- Laviosa, S. (1998). Core patterns of lexical use in a comparable corpus of English narrative prose. *Meta*, 43(4), 557–570.
- Marcinkevičienė, R., Kovalevskaitė, J., Utkas, A., & Vaičenonienė, J. (2012). An overview of multilingual processing for Lithuanian. Red. Cristina Vertan, Walther v. Hahn. *Multilingual processing in Eastern and Southern EU languages: low-resourced technologies and translation / Newcastle: Cambridge Scholars Publishing*, 119–143.
- Pažūsis L. (2014). *Kalba ir vertimas*. Vilniaus universitetas.
- Radvilavičiūtė, G. (2014). Draugiška aplinka anglų kalbai lietuvių galvose ir burnose. Prieiga per internetą: <http://literaturairmenas.lt/2014-09-05-nr-3486/1963-publicistika/3093-giedra-radvilaviciute-draugiska-aplinka-anglu-kalbai-lietuviu-galvose-ir-burnose>
- Rimkutė, E. (2006). *Morfologinio daugiareikšmiškumo ribojimas kompiuteriniame tekste*. Daktaro disertacija. Kaunas: Vytauto Didžiojo universitetas.

- Rudaitienė, V. (2014). Kalbų kontaktai ir lietuvių kalbos tradicija. *Žmogus ir žodis*, 16(3), 111–121. Prieiga per internetą: <http://www.zmogusirzodis.leu.lt/index.php/zmogusirzodis/article/viewFile/88/85>
- Scott, M. (2016). *WordSmith Tools Version 7*, Stroud: Lexical Analysis Software.
- Stubbs, M. (1986). Lexical density: A technique and some findings. M. Coulthard, *Talking about text. Discourse Analysis*, 13, *English Language Research* (pp. 27–42). Birmingham: University of Birmingham.
- Stubbs, M. (1996). *Text and corpus analysis: Computer-aided studies of language and culture*. Oxford and Cambridge: Blackwell.
- Tekorienė, D. (2000). *Anglų kalbos gramatika vidurinėms mokykloms*. Vilnius: Vaga.
- Tekorienė D. (2011). *Lietuvių kalbos padalyvių ir pusedalyvių atitikmenys anglų kalboje*. Vilnius: VU leidykla. Prieiga per internetą: http://www.esparama.lt/es_parama_pletra/failai/ESFproduktai/2011_Mokomoji_knyga_Pusedalyvio_ir_padalyvio_konstrukciju_vertimas.pdf
- Vaičėnienė, J., Kovalevskaitė, J., & Ringailienė, T. (2017). Tekstynais paremti vertimų kalbos tyrimai ir šaltiniai. *Kalbų studijos*, 30, 42–55.
- Vaičėnienė, J. (2014). How relevant is the distinction between the original and translated texts in general corpora? *Human Language Technologies – The Baltic Perspective* (pp. 248–256). Amsterdam/Berlin: IOS Press.
- Xiao, R. & Hu, X. (2015). *Corpus-based studies of translational Chinese in English-Chinese translation*. New Frontiers in Translation Studies. Shanghai Jiao Tong University Press and Springer-Verlag Berlin Heidelberg.
- Xiao, R. & Dai, G. (2014). Lexical and grammatical properties of translational Chinese: Translation universal hypotheses reevaluated from the Chinese perspective. *Corpus Linguistics and Linguistic Theory*, 10(1), 11–55. Prieiga per internetą: http://eprints.lancs.ac.uk/64330/1/02_Xiao_and_Dai.pdf

Jurgita Vaičenonienė,

Vytautas Magnus University, Lithuania; jurgita.vaicenoniene@vdu.lt

Jolanta Kovalevskaitė

Vytautas Magnus University, Lithuania; jolanta.kovalevskaite@vdu.lt

LEXICAL AND MORPHOLOGICAL FEATURES OF TRANSLATIONAL LITHUANIAN

Summary. In Lithuanian public and academic discourse, discussions about the influence of English have received considerable attention. Much has been written on the English borrowings in Lithuanian or various translation strategies applied at word, phrase or syntactic levels of language, whereas there have been only few attempts to investigate how Lithuanian translated from English differs from original language. This is why we found it interesting to investigate lexical and morphological features of translated Lithuanian applying the methods of corpus linguistics. For research purposes, we used a morphologically annotated comparable 4 mln. word corpus of original and translated fiction and popular science literature ORVELIT. It has been found that translations deviate in certain ways from original Lithuanian. Translated Lithuanian has: a lower lexical density, higher proportion of function words, shorter sentences, and higher proportion of list heads; translated fiction has a lower lexical variability and smaller proportion of low frequency words, whereas in popular science translations, these differences are less evident. Keyword analysis has shown content differences in originals and translations and the overuse of personal and possessive pronouns in popular science translations. The distribution of content and function words differs in originals and translations and in different registers. Translated Lithuanian has: more verbs (especially finite forms and adverbial participles), but less nouns and adjectives; fiction translations have less and popular science more adverbs than originals; there are more pronouns and prepositions in both popular science and fiction translations; depending on the register, there are higher or lower numbers of conjunctions, particles and interjections. Some of the differences may be explained by the English language interference as: the overuse of the optional 1st person pronoun in subject position, the overuse of optional preposition "su" with instrumental case, or the overuse of optional link verb in the complex predicate. In other words, the English influence is seen in transferring certain features obligatory for analytical language where omission would be a more natural choice in original Lithuanian. These findings in most cases agree with the previous research on translationese of other languages. It is hoped that the identified tendencies to over- or under-use certain lexical and morphological features as a result of English language interference might appear to be useful when editing and translating.

Keywords: Corpus-based translation studies; comparable corpora; original and translated Lithuanian; ORVELIT.