



faculty of social  
sciences, arts  
and humanities

42/2023

Research Journal  
Studies about Languages  
pp. 105-118

ISSN 1648-2824 (print)

ISSN 2029-7203 (online)

DOI 10.5755/j01.sal.1.42.33269

## TRANSLATION / VERTIMAS

Besimokančiųjų vertimo studentų tekstynas MUST-LT:  
kolokacijų vertimo atvejo analizė

Received 03/2023

Accepted 05/2023

**HOW TO CITE:** Vaičenonienė, J., (2023). Besimokančiųjų vertimo studentų tekstynas MUST-LT: kolokacijų vertimo atvejo analizė. *Studies about Languages / Kalbų studijos*, 42, 105–118. <https://doi.org/10.5755/j01.sal.1.42.33269>

# Besimokančiųjų vertimo studentų tekstynas MUST-LT: kolokacijų vertimo atvejo analizė

Student translation corpus MUST-LT:  
a case analysis of collocation translation

JURGITA VAIČENONIENĖ, Vytauto Didžiojo universitetas, Lietuva

## Santrauka

Kolokacijų vertimas – mažai nagrinėta tema vertimo studijų darbuose; daugiau dėmesio tenka profesionalių vertėjų darbų tyrimams, o besimokančiųjų vertėjų tekstai ligi šiol plačiau netyrinėti (Granger ir Lefer, 2018; 2020). Kita vertus, kaip teigia Granger ir Lefer (2018), besimokančiųjų vertėjų darbų analizė gali suteikti daug žinių ir pedagoginiams, ir tyrimų tikslams. Šio straipsnio tikslas – pristatyti tarptautinį MUST projektą Lietuvos vertimo studijų tyrėjų auditorijai ir parodyti jame kaupiamo MUST-LT tekstyno pritaikymo galimybes kaip atvejo analizę pasitelkiant kolokacijų vertimų tyrimą. Medžiaga, t. y. 32 publicistinio žanro tekstų vertimai, tyrimui paimta iš anglų–lietuvių kalbų MUST-LT patekstynio (46299 žodžiai). Tyrimui atrinktos 6 kolokacijos (iš viso 192 vertimai), po dvi kiekvienai iš šių vertimo kategorijų: kai vertimas (1) gali būti arba (2) neturėtų būti pažodinis dėl konteksto, stilistinės raiškos, įprastinės lietuvių kalbos vartosenos arba (3) tarpkalbinės raiškos skirtumų. Paaiškėjo, kad kaip ir ankstesniuose kolokacijų vertimo tyrimuose, studentų išverstose kolokacijose, kurių dėmenims neturėtų būti taikomas pažodinis vertimas, gali būti jaučiama originalo kalbos įtaka, stilistinės raiškos paprastinimas, o vertimo sprendimai, lyginant su kolokacijomis, kurioms pažodinis vertimas tinka, ne tokie sėkmingi. Tokie tyrimai gali padėti formuojant vertėjams skirtų kursų, juose aptariamų aspektų ar atliekamų užduočių spektrą, pateikti duomenys gali būti perpanaudoti kitiems ir kitų klaidų tyrimams, pratimams generuoti.

**RAKTINIAI ŽODŽIAI:** MUST-LT, Hypa4MUST, besimokančiųjų versti tekstynas, kolokacijų vertimas.

## Įvadas

Kolokacijų vertimas – gana reta tema deskriptyviųjų vertimo studijų darbuose, daugiau dėmesio tenka profesionalių vertėjų darbų tyrimams, o besimokančiųjų vertėjų tekstai beveik netyrinėti (Granger ir Lefer, 2018; 2020). Kita vertus, kaip teigia Granger ir Lefer (2018), besimokančiųjų vertėjų darbų analizė gali suteikti daug žinių ir pedagoginiams, ir tyrimų tikslams. Vienas tokio pobūdžio projektų, sujungiantis tekstynais grįstą vertimo studijų ir tekstynais grįstą mokinių kalbos tyrimus, yra Louvain-la-Neuve

universiteto *Anglų kalbos tekstynų lingvistikos centro*<sup>1</sup> profesorių Granger ir Lefer 2016 m. inicijuotas Daugiakalbio besimokančiųjų versti tekstyno projektas MUST, kuriame dalyvauja ir Lietuvos tyrėjai.

Šio straipsnio tikslas – pristatyti tarptautinį MUST projektą Lietuvos vertimo studijų tyrėjų auditorijai ir parodyti jame kaupiamo MUST-LT tekstyno pritaikymo galimybes kaip atvejo analizę pasitelkiant kolokacijų vertimų tyrimą. Medžiaga, t. y. 32 publicistinio žanro teksto vertimai, tyrimui paimta iš anglų–lietuvių kalbų MUST-LT patekstinio (plačiau 4-oje dalyje). Tyrimui atrinktos 6 kolokacijos (iš viso 192 vertimai), po dvi kiekvienai iš šių vertimo kategorijų: kai vertimas (1) gali būti arba (2) neturėtų būti pažodinis dėl konteksto, stiliaus, įprastinės lietuvių kalbos vartosenos arba (3) tarpkalbinės raiškos skirtumų. Svarbu pažymėti, kad, analizuojant studentų vertimus, daugiau dėmesio bus skiriama leksinei, semantinei ir stilistinei kolokacijos vertei, jos perteikimui. Tai yra gramatinis vertimo tikslumas nebus vertinamas, paliekant tokią diskusiją kitiems tyrimams. Deskriptyviaja analize siekiama ne akcentuoti vertimo klaidas, bet aprašyti esamą situaciją, parodyti, kas daugiau ar mažiau būdinga pasirinktų atvejų vertimams. Šis tyrimas metodologiškai yra lyginamasis, duomenimis grįstas, jo rezultatai aptariami kokybiškai ir kiekybiškai.

Toliau pristatomas MUST projektas ir darbo eiga jo aplinkoje. Poskyris *Lietuvos tyrėjų dalyvavimas MUST projekte* skirtas Lietuvos tyrėjų dalyvavimo projekte apžvalgai, o poskyris *Vertimų anotavimo sistema „TAS“* – vertimo anotavimo sistemos TAS1 aptarimui. Vėliau apžvelgiami kolokacijų vertimo tyrimai deskriptyviųjų vertimo studijų tyrėjų darbuose, analizuojami kolokacijų vertimo atvejai studentų vertimuose, pateikiami apibendrinimai ir ateities tyrimų gairės.

## MUST projektas ir darbo aplinka Hypal4MUST

Tarptautinį Daugiakalbio besimokančiųjų versti tekstyno projektą MUST (*The Multilingual Student Translation Corpus*) 2016 metais inicijavo Granger ir Lefer (Louvain-la-Neuve universitetas, Belgija). Iniciatyvos bendruomenė kasmet plečiasi ir šiuo metu apima 38 mokslo institucijas iš 19 šalių. Detalus projekto aprašymas, dalyviai ir veiklos pristatomi projekto tinklalapyje<sup>2</sup>, *ResearchGate*<sup>3</sup> puslapyje ir projekto vadovijų publikacijose (pvz., Granger ir Lefer, 2020). Pagrindinis projekto tikslas – kaupti studentų (tiek studijuojančių vertimą, tiek besimokančių užsienio kalbų) vertimų tekstyną. Tekstynas kaupiamas specialioje Hypal4MUST<sup>4</sup> aplinkoje, kurią išvystė čekų tyrėjas Obrusnik'as (2014), ji pasižymi ypač turtingais kelių pakopų metaduomenimis (žr. Granger ir Lefer, 2020, p. 1188–1192):

- 1 Kaupiami įvairių žanrų tekstai, kuriuos įkelia projekto dalyviai ir iš kurių formuojamos vertimo užduotys. Pateikiant ar vėliau atliekant paiešką, tekstus galima rinktis pagal tokius kriterijus, kaip žanras / požanris, kalba, ištrauka ar visas tekstas, pavadinimas, autorius, autoriaus gimtoji kalba / institucija, žodžių skaičius, raktiniai žodžiai, publikavimo statusas, formatas, paskirtis, tikslinė auditorija, ar tai originalo kalba rašytas tekstas, ar prieinamas profesionalus teksto vertimas.
- 2 Studentai pasirašo dalyvavimo sutikimus ir, jungdamiesi prie Hypal4MUST aplinkos, suveda duomenis, kurie matomi tik tos institucijos projekto koordinatoriams ir anonimizuojami atliekant tyrimus ar kitaip naudojant duomenis (pvz., lytis, metai, gimtoji / svetimoji kalbos, dėstymo kalbos, dabartinė ir ankstesnė studijų kryptys, užsienio kalbos žinių įsivertinimas, vertimo patirtis, ar verčiama į gimtąją ar ne gimtąją kalbą).
- 3 Pasirinkus norimą tekstą, formuojama vertimo užduotis, kuri aprašoma šiais metaduomenimis: pavadinimas, apibūdinimas, atlikimo terminas, prieigos kodas, studento statusas (besimokantis vertimo, besimokantis užsienio kalbos, kt.), studijų pakopa, vertimo kalba, užduoties tipas, vertinimas, trukmė, įrankiai ir ištekčiai užduočiai atlikti, grįžtamasis ryšys ir pataisytas variantas, vertimo instrukcija.

Tokia metaduomenų gausa, sukaupus pakankamos apimties vertimų tekstyną, leidžia atlikti daugialypę duomenų paiešką, gali būti pritaikoma analizuojant vertimus, formuojant grįžtamąjį ryšį, užduotį, atliekant lyginamuosius tyrimus ir kt. Hypal4MUST aplinka turi atskirus prisijungimus projekto koordinatoriams, projekte dalyvaujantiems

<sup>1</sup> Centre for English Corpus Linguistics, UC Louvain. <https://uclouvain.be/en/research-institutes/ilc/cecl>

<sup>2</sup> Centre for English Corpus Linguistics, UC Louvain. <https://uclouvain.be/en/research-institutes/ilc/cecl/must.html>

<sup>3</sup> MUST-Multilingual-Student-Translation-corpus. <https://www.researchgate.net/project/MUST-Multilingual-Student-Translation-corpus>

<sup>4</sup> Hypal4MUST. <https://hypal4must.eu/hypal/>

dėstytojams ir studentams. Studentų darbo eigą sudaro šie žingsniai: pasirašius dalyvavimo sutikimą, suteikiamas prisijungimo kodas; užsiregistravus duotu adresu, įrašomas užduoties kodas, suvedami asmeniniai metaduomenys, laikantis aprašytos ar paskaitos metu aptartos vertimo instrukcijos, iki nurodyto termino įkeliamas vertimas. Anotuotą vertimą su pastabomis studentas gauna nurodytu el. paštu. Priklausomai nuo užduoties, vertimas redaguojamas ir įkeliamas dar kartą, aptariamas paskaitoje ar kt. Tas pats studentas semestro arba studijų metu gali atlikti kelias užduotis, o tai ypač naudinga atliekant ilgalaikę stebėseną, teikiant individualų grįžtamąjį ryšį.

Moderatorių darbo eiga Hypal4MUST aplinkoje susideda iš tekstyno sukūrimo, teksto pasirinkimo, užduoties formulavimo, studentų tekstų lygiagretinimo, anotavimo (ir vertimų tyrimų). Projekto dalyviai gauna išsamią instrukciją<sup>5</sup>, joje nuosekliai aprašomi visi proceso žingsniai ir metaduomenys. Visų prима, prisijungus prie Hypal4MUST aplinkos, moderatorius susikuria norimam tekstynui pavadinimą, jame nurodo iš kokių kalbų ir į kokias kalbas bus verčiama. Kitame žingsnyje iš įkeltų tekstų duomenų bazės pasirenkamas norimas tekstas arba užpildžius metaduomenis, įkeliamas naujas. Toliau formuojama vertimo užduotis – jai suteikiamas pavadinimas, užpildomi metaduomenys, studentams nusiunčiamas prisijungimo kodas. Suformuluotos užduoties parametrus galima koreguoti – pavyzdžiui, jeigu ji duodama ir kitų metų studentų grupei, tereikia pakeisti užduoties atlikimo datą. Kai studentai įkelia vertimus, moderatorius turi juos peržiūrėti ir, jeigu reikia, perlygiuoti paragrafo ir sakinio lygmenimis. Pasirinktinai vertimus galima anotuoti naudojant įdiegtą vertimų anotavimo sistemą (TAS – *Translation Annotation System*) (plačiau žr. poskyrį *Vertimų anotavimo sistema „TAS“*; ji sukurta vykdant MUST projektą). Sulygiagretintą ir suanotuotą tekstą pažymėjus kaip baigtą, jis automatiškai nusiunčiamas studentui nurodytu el. paštu. Sukauptus tekstus galima parsisiųsti ir analizuoti su tekstynams skirta programine įranga (pvz., *AntConc*), arba atlikti pasirinktų žodžių, klaidų, individualių studentų ar visos grupės anotavimo statistikos paiešką Hypal4MUST aplinkoje integruotose tekstyno paieškos funkcijose (žr. **Pav. 1**).

**Pav. 1.** Darbo Hypal4MUST aplinkoje pavyzdys<sup>6</sup>

MUST projekto bendruomenės nariams skirtingų institucijų renkami duomenys tampa prieinami surinkus ne mažiau negu 100 studentų vertimų, tačiau patys duomenys nėra atviros prieigos visiems vartotojams, jais gali naudotis ir keistis tik projekto dalyviai. Tai yra, norint gauti prieigą prie konkrečios kalbos studentų vertimų, reikėtų kreiptis į projekto vadovę ir prisijungti prie MUST iniciatyvos.

<sup>5</sup> Granger, S., & Lefer, M.A. (2019). *Hypal4MUST\_user-guide*.

<sup>6</sup> Spalvomis paryškinti žodžiai rodo skirtingų klaidų tipų anotavimą. <https://hypal4must.eu/hypal/>

## Lietuvos tyrėjų dalyvavimas MUST projekte

Lietuvos tyrėjai į projektą įsitraukė nuo pat pradžių 2016 metais, dalyvavo visuose projekto organizuotuose seminaruose ir prisidėjo prie metaduomenų ir anotavimo gairių kūrimo, projekto metu kaupiami duomenys aprašyti publikacijose (Kovalevskaitė ir Vaičenonienė, 2015; Vaičenonienė et al., 2021), integruoti į vertimo besimokantiems skirtą mokomąją priemonę (Kovalevskaitė et al., 2022)<sup>7</sup>. Šiuo metu projekte dalyvauja dvi institucijos: Vytauto Didžiojo universitetas ir Vilniaus universitetas.

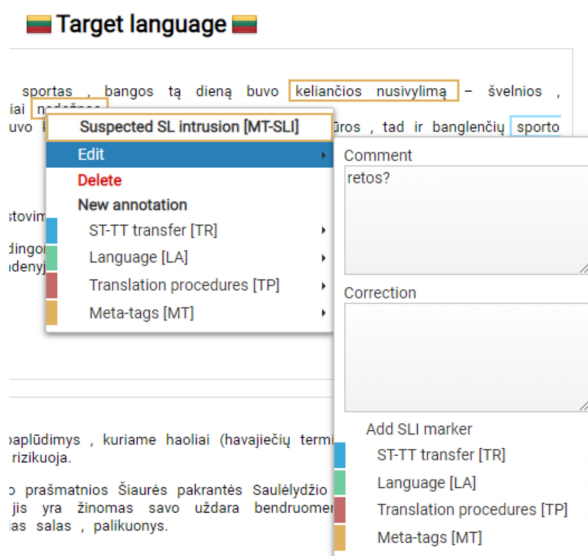
Vytauto Didžiojo universitete kaupiamas MUST-LT tekstynas apima tokias kalbų poras: anglų–lietuvių (46299 žodžiai); vokiečių–lietuvių (22981 žodžiai); prancūzų–lietuvių (6098 žodžiai); italų–lietuvių (10649 žodžiai). Iš viso tekstyną sudaro 86027 žodžiai, 353 vertimai, kuriuos atliko 199 studentai. Verčiami skirtingų žanrų tekstai: pavyzdžiui, anglų–lietuvių MUST-LT patekstinį šiuo metu sudaro 5 skirtingi tekstai, kurie neviršija 1000 žodžių apimties. Juos išvertė Anglų filologijos bakalauro, Taikomosios anglų kalbotyros ir Moderniosios lingvistikos magistro studijų programų studentai, lankantys vertimo pagrindų ar privalomuosius vertimo studijų kursus.

Prieš integruojant MUST projektą į mokymą, vertimo užduotys formuojamos taip, kad derėtų su kurso programa<sup>8</sup>. Pavyzdžiui, dėstant dalyką *Kontrastinė stilistika*, per paskaitas studentams pristatomos vertimo klaidų ir vertimo klaidų tipologijos (pvz., *Mellange*<sup>9</sup>, TAS), kartu analizuojami besimokančiųjų versti tekstai, studentai bando juos anotuoti taikydami pasirinktą anotavimo tipologiją. Kito susitikimo metu pristatomas MUST projektas, surenkami studentų sutikimai ir skiriama vertimo užduotis Hypa4MUST aplinkoje. Kurso metu studentai mokosi skirtingų žanrų tekstams būdingų kalbos ypatumų ir lyginti įvairių žanrų vertimus su originaliais pasirinktos kalbos tekstais. Projekto užduotis leidžia įvertinti, kaip studentai praktiškai taiko įgytas teorines žinias versdami konkretaus žanro tekstą. Žinoma, projekto integravimas ir pritaikymas studijose varijuoja ir priklauso nuo dalyko turinio, tyrimo tikslo, studentų poreikių ir kitų dedamųjų.

### Vertimų anotavimo sistema „TAS“

Remiantis ankstesnių tyrimų patirtimi (Granger ir Lefer, 2018), Hypa4MUST aplinkai išvystyta vertimų anotavimo sistema (2016–2019 m. integruota pirmą versija, o nuo 2019 m. testuojama antra versija). Kadangi šiuo metu daugiausia duomenų MUST-LT tekстыne sukaupta ir anotuota TAS1 versijoje, būtent ją detaliau pristatysime. Iš viso ją sudaro maždaug 60 žymų, kuriomis skirtingais hierarchiniais lygmenimis galima anotuoti, kiek originalo turinys perteikia ir atitinka vertimo instrukciją, kalbos ypatumus ir vertimo procedūras (Granger ir Lefer, 2018). Be žymų, anotatorius gali naudotis laukeliais, kuriuose galima įrašyti komentarą ar siūlomą vertimo variantą (žr. Pav. 2).

Anotatorius gali pasirinkti, koku hierarchijos lygmeniu žymėti klaidas (pvz., gramatinė klaida / gramatinio junglumo klaida / būdvardžio valdymo klaida), taisyti, komentuoti jas ar ne, žymėti visus netikslumus ar tik pasirinktas kategorijas ir kt., priklausomai nuo užduoties ar atliekamo tyrimo. TAS1 taip pat yra galimybė pažymėti ne tik klaidingus, bet ir gerus vertimo sprendimus („Plus“ metažymė), arba siunčiant studentui ištaisytą



Pav. 2. Vertimų anotavimo sistemos TAS1 pavyzdys

<sup>7</sup> Kovalevskaitė, J., Rimkutė, E., & Vaičenonienė, J. (2022). *Lietuvių kalbos kolokacijos: vartojimas, mokymas(is) ir vertimas. Mokomoji priemonė*. Kaunas: Vytauto Didžiojo universitetas. <https://resursai.pastovu.vdu.lt/mokomoji-priemone/apie>, <https://doi.org/10.7220/9786094675249>

<sup>8</sup> Kurso aprašo, kuriame integruotas MUST projekto turinys pavyzdys: *Kontrastinė stilistika*. <https://www.vdu.lt/lt/study/subject/494/>

<sup>9</sup> *The MeLLANGE Learner Translator Corpus (LTC)*. <http://corpus.leeds.ac.uk/mellange/ltc.html> (žiūrėta 2023 01 13).

tekstą, kartu pridėti ir profesionalaus vertėjo vertimą palyginimui (jeigu toks yra), anotatoriaus komentarus ir pastabas. Tas pats segmentas gali būti anotuojamas keliomis žymomis, o tai ypač aktualu, nes ta pati klaida gali būti kelių vertimo procesų rezultatas, klaidos gali persidengti, kartais skirtingiems to paties vertimo vieneto segmentams gali būti priskirtos skirtingos žymos. Reikėtų pastebėti, kad TAS1 versija buvo laikoma darbine ir nuolat keičiama, tobulinama, dėl to kai kurių žymų TAS2 versijoje neliko, o kai kurios buvo praplėstos naujomis. Taigi, į toliau pateiktą TAS1 (žr. **1 lentelę**) reikėtų žiūrėti kaip į tarpinį variantą, o ne rekomenduojamą tipologiją vertimo klaidoms žymėti (atliekant žymų vertimą, remtasi Hypal4MUST aplinka). Pavyzdžiui, anglų–lietuvių kalbų porų MUST-LT patekstyne šiuo metu iš 211 įkeltų tekstų 147 yra sulygiagretinti paragrafo ir sakinio lygmeniu ir turi 446 anotavimo žymes: tekstai studentams anotuoti per vykčius studijų dalykus neturint mintyje konkretaus tyrimo tikslo, todėl negalima kalbėti apie kokias nors bendras vertimo tendencijas ar dažniausias / rečiausias klaidas. Anotavimo tikslas iki šiol buvo pateikti studentams grįžtamąjį ryšį ir išbandyti TAS1 sistemą, kuri, kaip minėta, buvo nuolat keičiama, vystoma, pildoma, ir kol kas ne visos paieškos ar statistikos generavimo galimybės buvo prieinamos, todėl, sulygiagretinus tekstus pasakiniui, patogiausia toliau juos analizuoti parsisiuntus pasirinktu konkordansų generavimo įrankiu.

**1 lentelė.** Vertimo anotavimo sistema TAS1 (Granger ir Lefer, 2018; 2022)

I Originalaus teksto perteikimas vertimo kalba	
1 Turinio perteikimas	<ul style="list-style-type: none"> <li>a praleidimas</li> <li>b pridėjimas</li> <li>c iškrypimas</li> <li>d neapsisprendimas</li> </ul>
2 Leksika	<ul style="list-style-type: none"> <li>a išverstas žodis / frazė, kurios nereikia versti</li> <li>b neišverstas žodis / frazė, kurią reikia išversti</li> <li>c terminas išverstas neterminu</li> <li>d neterminas išverstas terminu</li> </ul>
3 Diskursas / pragmatika	<ul style="list-style-type: none"> <li>a jungtukai</li> <li>b tema-rema</li> </ul>
4 Registas ir kultūra	<ul style="list-style-type: none"> <li>a neišlaikomas originalo kalbos registas</li> <li>b neadaptuota kultūrinė realija</li> </ul>
5 Vertimo instrukcija	<ul style="list-style-type: none"> <li>a nesivadovaujama pateiktų terminų sąrašu</li> <li>b nesilaikoma nurodytų formaliųjų teksto rengimo reikalavimų</li> </ul>
II Kalba	
1 Gramatika	<ul style="list-style-type: none"> <li>a kaitybos klaidos</li> <li>b netinkamos veiksmažodžio laikų ir nuosakų formos</li> <li>c netinkamos veiksmažodžio rūšies formos</li> <li>d netinkama žodžių tvarka</li> <li>e netinkami įvardžiai, artikkeliai ir kt. pažymimieji žodžiai</li> <li>f klaidingas, perteklinis arba praleistas prielinksnis</li> <li>g klaidos, apimančios visas įvardžių kategorijas</li> <li>h derinimo klaidos</li> <li>i gramatinio junglumo klaidos <ul style="list-style-type: none"> <li>1 būdvardžių valdymas</li> <li>2 daiktavardžių valdymas</li> <li>3 veiksmažodžių valdymas</li> <li>4rieveiksmių valdymas</li> </ul> </li> </ul>

2 Leksika ir terminija	<ul style="list-style-type: none"> <li>a vienažodžiai neterminai               <ul style="list-style-type: none"> <li>1 daryba</li> <li>2 netikri atitikmenys</li> </ul> </li> <li>b vienažodžiai terminai (nepakankamos specialiosios terminijos srities žinios)               <ul style="list-style-type: none"> <li>1 netinkama termino forma (darybiškai)</li> <li>2 klaidingas termino atitikmuo</li> </ul> </li> <li>c keliažodžiai neterminai               <ul style="list-style-type: none"> <li>1 netiksliai išverstas sudurtinių žodžių arba junginių elementas</li> <li>2 netinkamai išverstos kolokacijos</li> <li>3 perkeltinės vientisos reikšmės pastoviųjų junginių (idiomų arba frazeologizmų) vertimo klaidos</li> </ul> </li> <li>d klaidingai išversti keliažodžiai terminai</li> </ul>
3 Kohezija (teksto rišlumas)	<ul style="list-style-type: none"> <li>a siejamieji žodžiai (jungtukai,rieveksmiai ir pan.)</li> <li>b netikslūs anaforiniai ryšiai, perteikiami įvardžiais</li> </ul>
4 Formatavimas	<ul style="list-style-type: none"> <li>a rašyba</li> <li>b skyryba</li> <li>c matmenys, datos, skaičiai</li> </ul>
5 Stilius	<ul style="list-style-type: none"> <li>a griozdiškas stilius</li> <li>b daugiažodžiavimas, pertekliniai žodžiai ir pan.</li> <li>c netinkamai parinkti regioniniai, atmainų, stiliaus ir kt. variantai</li> <li>d verčiant pavartojama neformali leksika, nors pagal registrą būtų tinkamesnis formalinis variantas</li> </ul>

### III Vertimo procedūros (TAS2)

#### IV Metažymės

1 Puikus vertimas	
2 Originalo kalbos įtaka	

Kaip minėta, šiuo metu testuojama TAS2 versija skiriasi. Pavyzdžiui, detalesnės turinio perteikimo žymos, gramatikos žyma pervadinta į „gramatikos ir sintaksės“ žymą, metažymose atsirado „pasikartojančios klaidos“ pasirinkimo galimybė, galima atlikti tikslią konkrečių žodžių (lemų, formų), junginių paiešką naudojantis CQL operatoriais ir kt. Ateityje, dirbant su galutine TAS2 versija, planuojama tekstus rinkti ir anotuoti ne tik teikiant grįžtamąjį ryšį studentams, bet ir keliant įvairius vertimo studijose aktualius klausimus. Norėdami parodyti MUST-LT teksto pritaikymo galimybes, ketvirtoje dalyje kaip atvejo analizę pristatysime kolokacijų vertimo tyrimą, todėl trečioje trumpai apžvelgsime kolokacijų vertimo tyrimus tekstynais grįstose vertimo studijose.

#### Kolokacijų vertimo tyrimų apžvalga

Vertimo studijose kolokacijų raiška ir jų vertimas – aktuali tema. Deskriptyviųjų vertimo studijų kryptyje atkreipiamas dėmesys į tai, kad to paties žodžio kolokatų laukas gali sutapti, skirtis ar net neegzistuoti skirtingose kalbose (Baker, 2011, p. 47–49) arba apsiribojama dažnumo požiūriu į kolokacijas (Bernardini, 2007, p. 1), kuomet statistiniais ir tekstynų lingvistikos metodais siekiama kolokacijas nustatyti, o ne jas klasifikuoti. Tokio pobūdžio darbų nėra gausu, juos persveria konceptualiųjų metaforų vertimo tyrimai, kuriuose kolokacijos dažniausiai minimos kaip lingvistinė metaforų raiška (pvz., Newmark, 1988; Schaeffner, 2004; Steen, 2014).

Baker (2011) diskusija apie kolokacijų vertimo problemas dažnai minima deskriptyviųjų vertimo studijų tyrėjų darbuose (pvz., Øverås, 1998). Baker (2011, p. 54–63) aptaria kolokacijų vertimo problemas, kylančias dėl kolokacijų raiškos skirtumų skirtingose kalbose, tokias kaip originalo kalbos įtaka; klaidingai suprasta ir išversta originalo kalbos kolokacija dėl vertimo kalbos įtakos; sankirta tarp tikslaus prasmės perteikimo ir natūralumo; kultūros ypatumų, skirtingai realizuojamų vertimo kalboje perteikimo problemos; stilistiškai žymėtų kolokacijų perteikimas. Gregorio-Godeo ir Molina (2011), remdamiesi Baker (2011) kolokacijų vertimo problemų taksonomija, tyrinėja, kaip kolokacijos verčiamos BBC žinių portale iš anglų į ispanų kalbą. Autoriai pastebi dvikalbių kolokacijų žodynų trūkumą, kurie padėtų vertėjams greičiau parinkti vertimo ekvivalentus ir pristato DUFIE projekto rengiamą anglų–ispanų frazeologinių vienetų žodyną. Kolokacijos suprantamos kaip pasikartojantys žodžių junginiai, kurių reikšmė negali būti suprantama iš jų sudedamųjų dalių (Gregorio-Godeo ir Molina, 2011, p. 137). Tyrimui pasirinktas angliškąs publicistinis tekstas ir jo vertimas į ispanų kalbą. Išrinktų anglišku kolokacijų vertimai į ispanų kalbą, patikrinti ispanų kalbos tekстыne, parodė įvairias kolokacijų vertimo klaidas dėl originalo ar vertimo kalbų įtakos, tikslios reikšmės ar idiomatiškumo perteikimo problemas, kitus neatitikimus, kurių galima būtų išvengti naudojantis dvikalbiu kolokacijų žodynu.

Kolokacijų vertimas publicistiniame stiliuje šiek tiek nagrinėtas ir lietuvių kalboje. Iš mokslo populiarinimo tekstų išrinkę 250 anglišku dvižodžių kolokacijų ir jų lietuviškų vertimų Stulpinaitė et al. (2016, p. 41) sugrupavo jas pagal Catford'o (1965) vertimo transformacijų taksonomiją. Paaiškėjo, kad vertimuose „dalis junginių išliko kolokacijomis, tačiau nemaža jų dalis po vertimo transformacijų tapo savarankiškais leksiniais vienetais ar laisvaisiais žodžių junginiais“ (Stulpinaitė et al., 2016, p. 41).

Viena ryškiausių deskriptyviųjų vertimo studijų kryptių – tekstynais grįsti vertimo kalbos tyrimai, nagrinėjantys vertimo kalbos universalijas ir ypatumus, lemiamus socialinių, kultūrinių, pragmatinių ar kognityvių faktorių. Naudojant lygiagrečiuosius ir lyginamuosius tekstynus, ieškoma kalbos ypatybių būdingų tik verstiems tekstams. Kita vertus, kolokacijų, kaip tyrimo objekto, vertimo ar lyginimo originaliuose ir verstuose tam tikros kalbos tekstuose dėl tekstynų trūkumo, metodologinių ar kitokių iššūkių nėra daug. Galima išskirti Øverås (1998), Kenny (2001), Bernardini (2007, 2015), Feng et al. (2018) darbus apie kolokacijų raišką verstuose tekstuose. Øverås (1998) tyrime analizuojama, ar versti anglų ir norvegų kalbų tekstai (2,6 mln. žodžių apimties lygiagrečiame dvikrypčiame anglų–norvegų kalbų grožinės ir negrožinės literatūros tekстыne) pasižymi didesniu eksplicitiškumu negu originalūs šių kalbų tekstai. Tai yra aiškinamasi, ar versti tekstai yra rišlesni, jiems būdingos tam tikros perteklinės, originalo kalboje tik implikuojamos struktūros. Øverås (1998) atkreipia dėmesį, kad originalo kalbos kolokacijos verčiant gali būti premodifikuojamos tam tikrais aiškinamaisiais elementais (pvz., *the psychological damage* (originalo kalba (OK)) – *serious psychological damage* (vertimo kalba (VK)), neįprastos kolokacijos verčiamos įprastomis ir kt.

Tam tikras polinkis normalizuoti neįprastas (*ad hoc*) kolokacijas nustatytas ir Kenny (2001) tyrime. Iš 2 mln. žodžių apimties lygiagrečiojo vokiečių–anglų kalbų grožinės literatūros tekстыno žodžių sąrašo rankiniu būdu išrinktos neįprastos kolokacijos (30) su žodžiu *akis* – žodis buvo tolygiai pavartotas daugelio autorių (1159 lemos). Nustatyta, kad 22 proc. neįprastų kolokacijų vertimuose normalizuotos, išverstos stilistiškai neutraliais žodžių junginiais. Vaičėnienė (2012) tyrime iš lygiagretaus lietuvių–anglų kalbų grožinės literatūros tekстыno (1,7 mln. žodžių) išrinktos 48 kūrybiškos kolokacijos su dažnomis ir tolygiai įvairių autorių kūrinuose pasiskirsčiusiomis lemomis *žmogus*, *akis*, *galva* bei abstraktais *laikas*, *gyvenimas*, *mintis*. Lyginant su Kenny (2001) tyrimu, paaiškėjo, kad kūrybiškos lietuviškos kolokacijos neišlaikomos dvigubai dažniau – net 44 proc. atveju jos buvo normalizuotos, jų stilistinis išskirtinumas angliškuose vertimuose prarastas.

Feng et al. (2018) aiškinasi, ar verslo kalbos vertimams iš kinų į anglų kalbą būdinga skurdesnė kolokacijų raiška ir eksplicitiškumas. Kolokacijų palyginimui originaliuose ir verstuose tekstuose pasirinkti du lingvistiniai indikatoriai – kolokabilumas ir deleksikalizacija. Kolokacijos skirstytos į tris grupes: laisvuosius junginius (kiekvienas dėmuo išlaiko savo tiesioginę reikšmę pvz., *check details / patikrinti detales*); riboto junglumo pastoviuosius žodžių junginius (pvz., *budget deficit / biudžeto deficitas*); idiomą (sustabarėję, nedalomi junginiai, pvz., *catch eye / patraukti akį*). Vertėjai kolokacijas turi atpažinti kaip visumą, o ne skaidyti į sudedamąsias dalis. Taigi, daroma prielaida, kad dėl perteklinio laisvųjų žodžių junginių ir ne tokio dažno pastoviųjų junginių ir idiomų vartojimo vertimų kalba atrodo paprastesnė, skurdesnė lyginant su originaliais kinų kalbos tekstais (Feng et al., 2018,

p. 680). Deleksikalizaciją Feng et al. (2018) supranta kaip semantiškai išblukusią žodžio reikšmę – ją galima suprasti tik junginiuose su kitais leksiniais vienetais (pvz., *take / imti – take note / pastebėti*). Deleksikalizuotomis kolokacijos laikytos, jeigu: 1) nors viena kolokacijos dedamoji negali būti vartojama savarankiškai (pvz., *make a decision / priimti sprendimą*); 2) tiesioginė nors vienos kolokacijos dedamosios reikšmė kolokacijoje kinta ar prarandama (pvz., *zip code / pašto kodas vs. zip / užtrauktukas*). Spėta, kad mažiau deleksikalizuotų kolokacijų vertimuose rodo polinkį į eksplicitiškumą – perteklinį aiškinimą. Tyrimo rezultatai pagrindė šias prielaidas – verslo vertimų kalba paprastesnė, eksplicitiškensė negu originali angliška vartosena.

Bernardini (2007, 2015) tyrimuose laikomasi dažnumu grįsto požiūrio į kolokacijas apsiribojant dvižodžiais junginiais. Remdamasi italų ir anglų kalbų porų lyginamaisiais ir lygiagrečiaisiais originalios ir verstos grožinės literatūros tekstynais, Bernardini (2007) nustatė, kad vertimuose daugiau daiktavardžio + prielinksnio / prielinksnio + daiktavardžio junginių negu originaliuose italų kalbos tekstuose. Kitame tyrime aiškintasi, kaip skiriasi kolokacijų raiška originaliuose angliškuose ir iš italų kalbos verstuose finansų srities tekstuose (Bernardini, 2015). Pabrėžiama, kad kolokacijų išmanymas – ypač svarbi užsienio kalbos ir vertimo išmanymo kompetencija, bet dažnai probleminė net verčiant į gimtąją kalbą (Bernardini, 2015, p. 526). Iš tekstynų nustatyti ir palyginti 8 tipų žodžių junginiai: bdv. + bdv.; bdv. + dkt.; prv. + bdv.; prv. + vksm.; dkt. + dkt.; dkt. + vksm.; vksm.+ dkt.; vksm.+ prv.; bdv + prl./jng. + bdv.; dkt. + prl./jngt. +dkt.; vksm. + det./ prl. + dkt.; dkt. + [prl.] + det./jv./sktv. + dkt; vksm.+ [dal./prl./jv.] + [det./jv.] + dkt. Išvadose teigiama, kad versti finansų srities tekstai pasižymi žemesniu kolokacijų skaičiumi negu originalūs tos srities angliški tekstai, o akcinių laiškų patekstynyje pastebima atvirkštinė tendencija: vertimuose kolokacijų skaičius didesnis negu originaliuose angliškuose laiškuose. Verčiant technines kolokacijas, jaučiama OK interferencija, matomas vertėjų polinkis normalizuoti kūrybinę, neįprastą OK raišką renkantis įprastas VK frazes (Bernardini, 2015, p. 535).

Iš šios apžvalgos matome, kad deskriptyviųjų vertimo studijų kryptyje kolokacijos suvokiamos visų pirma kaip dažni, pasikartojantys žodžių junginiai. Remiantis tekstynų lingvistikos metodais, tiriamos vertimų kalbos universalijos, kolokacijų dažnis originalioje ir vertimų kalboje, stilistiškai žymėtų kolokacijų perteikimas įvairiuose žanruose. Vis tik kolokacijų vertimo tyrimų nėra gausu, ypač trūksta besimokančiųjų versti tekstų tyrimų.

### Atvejo analizė: kolokacijų vertimas besimokančiųjų versti tekстыne MUST-LT

Šiam tyrimui naudosisime anglų–lietuvių kalbų MUST-LT patekstynio duomenis, konkrečiai – publicistinio žanro teksto vertimus<sup>10</sup>, kuriuose analizuoti pasirinktų kolokacijų vertimai, galimai rodantys anglų kalbos interferenciją arba stilistinę paprastinimą studentų vertimuose. Per 2018–2021 m. iš viso surinkti 32 šio teksto studentų vertimai (juos sudaro 8817 lietuviškų žodžių). Ši užduotis skirta vertimo besimokantiems magistro studijų studentams kaip namų darbų užduotis, jai atlikti buvo galima naudoti tam tinkančius išteklius ir įrankius, vertimo instrukcija nedetalizuota. Studentų metaduomenys: amžius 22–30 m., gimtoji kalba – lietuvių, pirmoji užsienio kalba – anglų, verčiama į gimtąją kalbą, daugumos vertimo studijų patirtis – vieneri studijų metai, o bakalau-ro studijų kryptis – anglų arba lietuvių filologijos studijos. Kolokacijų vertimų atitikmenims generuoti taikyta *AntConc* programine įranga (Anthony, 2022). Tyrimo naudoti duomenys yra atviros priegos, juos galima parsisiųsti ir perpanaudoti<sup>11</sup>. Pasirinktos trys kolokacijų vertimo situacijos:

- 1 angliška kolokacija gali būti verčiama pažodžiui (*to grab the board; to spin the board*);
- 2 angliška kolokacija gali būti verčiama pažodžiui, bet toks vertimas stilistiškai, turinio arba vartosenos dažnumo prasme nėra taikliausias vertimo pasirinkimas duotame kontekste ir geriau būtų taikyti ne pažodinę vertimo strategiją (*takeoff zone, annoyingly infrequent*);
- 3 angliška kolokacija į lietuvių kalbą neturėtų būti verčiama pažodžiui, nes skiriasi tarpkalbinė / kultūrinė raiška (*gray ponytail; to hit the ocean*).

<sup>10</sup> Lancaster, J. (2015). *Pure Hawaiian*. *National Geographic* (2). (žiūrėta 2023 01 13).

<sup>11</sup> Vaičėnienė, J. (2023). *MUST-LT\_studentų vertimai (v.1) [Data set]*. Zenodo. <https://doi.org/10.5281/zenodo.7543488>



Pasirinktų kolokacijų kontekstas originaliame tekste ir profesionalaus vertėjo sprendimai parodyti pavyzdyje:

*In the islands where surfing began, the waves on that particular day were a disappointment—mushy, chest high, and annoyingly infrequent (1). Still, Hawaiians have never needed much of an excuse to grab a board (2) and hit the ocean (3), and the takeoff zone (4) was packed. Teens on shortboards. Moms on longboards. Grade-schoolers on bodyboards. A guy with a gray ponytail (5) on a stand-up paddleboard. <...> I spun my board (6) toward the beach and paddled hard. <...> (Pure Hawaiian, Lancaster 2015)*

*Salose, kur gimė banglenčių sportas, tądien bangos nebuvo įspūdingos: putojo, vos siekė krūtinę, jų tekdavo gerokai palūkėti (1). Tačiau norėdami su banglente (2) pulti į vandenį (3) havajiečiai nelaukia ypatingų sąlygų, todėl starto zona (4) buvo sausakimša. Paaugliai su trumposiomis lentomis, mamos – su ilgosiomis, vaikai – su plaukimo lentomis. Žilus plaukus į uodegą susirišęs (5) vyrukas stovėdamas ant banglentės irklavo. <...> Atgrėžiau lentą (6) į paplūdimį ir ėmiau sparčiai irtis<sup>12</sup>. (Tikrieji Havajai, Lancaster 2015)*

Angliškam tekstui būdinga ekspresyvi stilistinė raiška – pasirinktuose duomenyse ji atsiskleidžia veiksmožiais (*to grab* (2), *to hit* (3)), būdvardžiais (*annoyingly* (1)), šis ekspresyvumas turėtų būti išlaikomas ir vertime. Duotame profesionalaus vertimo pavyzdyje matome, kad išskirtais atvejais taikomas praleidimas (*su banglente* (2)), stilistinio ekvivalento parinkimas (*pulti* (3)) perfravavimas (žr., *tekdavo gerokai palūkėti* (1)). Nors kai kurie pasirinkimai nutolina nuo gramatinės originalo raiškos, stilistinė teksto vertė daugeliu atveju lieka nepakitusi, kolokacija matoma kaip vienas nedalomas vienetas, mažiau dėmesio skiriant struktūriniam informacijos perkėlimui. Analizuojant studentų vertimus, nesiekiami kategoriškai siūlyti konkretaus vertimo ekvivalento, kadangi jų gali būti įvairių: kaip matome profesionaliame vertime, kai kuriais atvejais gramatinis praleidimas ar perfravavimas pagrindžiamas (net jeigu ir pažodinis vertimas įmanomas) jeigu išlaikoma stilistinė originalo vertė.

Pirmosios kategorijos žodžių junginiuose (*to grab the board* (2), *to spin the board* (6)) kiekvienas dėmuo išlaiko savo reikšmę, turi tiesioginį atitikmenį lietuvių kalboje ir gali būti verčiami pažodžiui. Nors abi pasirinktos kolokacijos gali būti verčiamos pažodžiui, jos skiriasi stilistine verte: pirmosios veiksmožodis (*to grab*) yra ekspresyvesnis, o antrosios neutralė (*to spin*). Galima būtų manyti, kad tokiu atveju ir lietuviški vertimo atitikmenys turėtų būti parenkami teikiant pirmenybę ekspresyvumo kriterijui pirmu atveju ir renkantis neutralų ekvivalentą antruoju. Pažodinis vertimas gali būti vienoks, jeigu kolokacija verčiama be konteksto, ir kitoks išliame tekste. Atmetus kontekstui netinkamas žodžio *board* reikšmes, kadangi rašoma apie banglenčių sportą, abiejų kolokacijų daiktavardis į lietuvių kalbą galėtų būti perteikiamas hiperonimu *lenta* arba hiponimu *banglente*. Veiksmožodžių atitikmenys gali būti įvairūs, juos galima pasirinkti naudojantis įvairiais vienkalsiais, dvikalsiais žodynais, tekstytais, automatinį vertimo sistemomis, vertimo atmintimis, kitais duomenynais ir įrankiais, pavyzdžiui:

**2 lentelė.** Pažodinio vertimo atitikmenų paieška

Vienkalbio žodyno duomenys <sup>13</sup>	Dvikalbio žodyno duomenys <sup>14</sup>	Google vertėjo duomenys <sup>15</sup>
<i>to grab</i> staigiu ar smarkiu judesiu ką nors suimti	griebti, sučiupti	patraukti, griebti, užgrobti, pačiupti, sučiupti, pasisavinti
<i>to spin</i> labai greitai ap(s)ukti, sukti(s)	sukti, suktis	suktis, sukti

<sup>12</sup> *Tikrieji Havajai*. [https://www.nationalgeographic.lt/lt/zurnalu\\_archyvas/2015\\_vasaris/tikrieji\\_havajai/](https://www.nationalgeographic.lt/lt/zurnalu_archyvas/2015_vasaris/tikrieji_havajai/) (žiūrėta 2023 01 13).

<sup>13</sup> *Longman Dictionary of Contemporary English | LDOCE*. <https://www.ldoceonline.com/dictionary/grab> (žiūrėta 2023 01 11).

<sup>14</sup> *Anglų–lietuvių kalbų žodynas*. <https://ekalba.lt/anglu-lietuviu-kalbu-zodynas/> (žiūrėta 2023 01 11).

<sup>15</sup> *Google vertėjas*. <https://translate.google.com/?hl=lt&sl=lt&tl=en&text=patraukti&op=translate> (žiūrėta 2023 01 11).

Įvertinus kontekstui tinkamiausius atitikmenis, galimas pažodinis kolokacijų vertimas būtų *griebti lentą / banglentę* ir *apsukti lentą / banglentę*. Gramatinė raiška sakinyje gali kisti priklausomai nuo konkretaus vertimo (pvz., *griebti / pagriebti / griebė / griebdami / pagriebė / griebę*, etc.). Besimokančiųjų versti pasirinkimus galima sugrupuoti pagal pasirinktą veiksmožodžio vertimo būdą:

- 1 Pažodinis vertimas stilistiškai ekspresyviu ekvivalentu arba sinonimu, turinčiu panašią stilistinę vertę (28 atvejų): *griebti banglentę* (6)<sup>16</sup>, *pasigriebs banglentę*, *griebti lentą* (5), *pasigriebti banglentę* (3), *pagriebtų banglentę* (2), *pagriebti banglentę*, *griebtų banglentę* (2), *griebtis banglentės*, *pagriebti lentą*; *čiupti lentą* (2), *čiupti banglentę*, *stverti lentą*, *pasičiupti banglentes*, *stverti banglentę*.
- 2 Perfrazavimas, praleidimas, stilistinės vertės neišlaikymas (4 atvejai): *neieškojo rimtos progos, kad galėtų nerti į vandenyną*; *norėdami plaukti banglente*; *atsisakyti pasiplaukiojimo su banglente*, *pasiimti banglentę*.

Kaip matome iš pavyzdžių, stilistinė vertė išlaikyta 88 proc. Jei palygintume antrą žodžių junginį *to spin the board*, matyti, kad jis taip pat daugiausia verstas pažodžiui renkantis artimiausią atitikmenį (26 atvejai arba 81 proc.) (*pasukau savo banglentę* (9), *apsukau savo banglentę* (5), *pasukau savo lentą* (3), *apsukau savo lentą* (2), *su banglente pasisukau*, *pasukau lentą* (2), *sukau lentą* (2), *pasukau banglentę*, *apsukau lentą*) arba perfrazuotas parenkant ne artimiausią sinonimą, neišlaikant originalo raiškos (6 atvejai) (*nusukau savo banglentę*, *pakreipiau savo banglentę*, *pastūmiau savo lentą*, *pasileidau bėgti su banglente*, *nukreipiau banglentę*, *pasukau link paplūdimio*).

Antrosios kategorijos kolokacijų bent vienas dėmuo neturėtų būti verčiamas pažodžiui nes kolokacija turi specifinę reikšmę kontekste (*takeoff zone* (4)) arba pažodinis vertimas būtų ne toks įprastas lietuvių kalbos vartosenai, stilistiškai sunkus (*annoyingly infrequent* (1)). Angliškas žodžių junginys *takeoff zone* bekontekstėje situacijoje galėtų būti įvairiai interpretuojamas<sup>17</sup>. Pavojus kyla, jeigu *takeoff* bus traktuojamas ne banglenčių sporto kontekste, o išverčiamas pažodžiui, kilimo aukštyn prasme (\**kilimo zona*), dažna oro transporto kontekste. Profesionalaus vertėjo minima *starto zona* – tinkamas, nors ir ne vienintelis įmanomas pasirinkimas (plg. *starto aikštelė*, *starto vieta*, etc.). Kita kolokacija, nors ir galėtų būti verčiama pažodžiui kaip *erzinančiai nedažnos*, stilistiškai sklandesnis ir įprastesnis lietuvių kalbos vartosenai<sup>18</sup> būtų *erzinančiai retos* junginio pasirinkimas. Taigi, galime daryti prielaidą, kad tais atvejais, kai bus verčiama pažodžiui, galima anglų kalbos interferencija. Skirtingai negu pirmosios kolokacijų grupės vertimuose, matome, kad tinkamiau parinktų lietuviškų atitikmenų procentas studentų vertimuose ženkliai mažesnis:

- 1 Vertimas kontekstiniu lietuvišku atitikmeniu (8): *starto zona (-oje)* (6), *starto aikštelė*, *starto vieta*;
- 2 Vertimas neįprastu / nesklaidžiu lietuvišku atitikmeniu (7): *banglenčių zona*, *banglentininkų plaukiojimo vieta*, *pasiruošimo zona*, *banglentininkų zona*, *paplūdimio zona*, *išplaukimo zona*, *banglenčių sporto vieta*;
- 3 Perfrazavimas (7): *pakrantė* (3), *paplūdimys(-iai)* (2), *krantas*, *vandenyno krantas*;
- 4 Praleidimas (6): *bangų laukė gausa žmonių*, *pasigriebtų banglentę ir perskrostų vandenyną*, *bangų laukdavo būrys žmonių*, *netrukdė pasičiupti banglentes ir įlįsti į vandenyną*, *žmonių buvo daugybė*, *nereikėjo rimtos priežasties griebti lentą ir pulti link vandenyno*;
- 5 Anglų kalbos interferencija (4): *pakilimo zona (-oje)* (3), *kilimo zona*.

Vertimas kontekstiniu lietuvišku atitikmeniu (25 proc.) ir perfrazavimas (22 proc.) gali būti laikomi tinkamiausiais studentų vertimo sprendimais, tačiau net 53 proc. vertimų laikytini netiksliais, probleminiais. Kolokacijos *annoyingly infrequent* vertimus galima grupuoti taip:

- 1 Parinktas stilistiškai tinkamesnis atitikmuo *erzinančiai retos (-ai)* (17);
- 2 Perfrazavimas, parinktas sinonimas premodifikatui (8): *gana retos*, *nervinančiai retos*, *nykiai retos*, *retos*, *erzinančiai nepastovios*, *retai pasirodydavo*, *nepastovumas erzino*, *ir jų reikėjo laukti*;
- 3 Galima anglų kalbos interferencija, stilistiškai mažiau tinkamas pasirinkimas (7): *erzinančiai nedažnos* (5), *erzinančiai žemos ir nedažnos*, *buvo įkyru nes jos nedažnai kildavo*.

<sup>16</sup> Čia ir kitur nurodomas atvejų skaičius.

<sup>17</sup> Pavyzdžiui, IATE ekvivalentai daiktavardžio premodifikatui: <https://iate.europa.eu/search/result/1673599220006/1>

<sup>18</sup> Kaip, pavyzdžiui, rodo Google vertėjo paieškos rezultatai, *retas* vartoseną dažnesnė: <https://translate.google.com/?hl=lt&sl=en&tl=lt&text=infrequent&op=translate> (žiūrėta 2021 01 13).

Nors sėkmingų vertimų šiuo atveju daugiau, matome, kad apskritai antrosios kategorijos kolokacijos, kurioms pažodinis vertimas neturėtų būti taikomas dėl žinomo konteksto arba kitokios įprastinės lietuvių kalbai vartosenos, besimokantiems versti problematiškesnės lyginant su pirmosios kategorijos žodžių junginiais.

Trečiai kategorijai priskyrėme kolokacijas, kurių tarpkalbinė / kultūrinė raiška akivaizdžiai skiriasi nepriklausomai nuo konteksto ir čia pasirinkimo nelieka – bent vienas dėmuo neturėtų būti verčiamas pažodžiui (*gray ponytail; to hit the ocean*). Lietuviškai vartosenai įprasta kolokacija būtų *žili plaukai*, o pažodinis vertimas *pilki plaukai* rodytų anglų kalbos interferenciją. Tokio vertimo atvejų studentų tekstuose pasitaikė net 34 proc. (pvz., *pilkais plaukais ir kasyte, pilka kasele, pilka arkljo uodega, pilka plaukų kasa*, etc.), o likusiais atvejais angliška spalva *grey* išversta kaip *žili* arba *šviesūs* plaukai. Antrame pavyzdyje angliškas veiksmažodis *to hit* yra semantiškai platus<sup>19</sup>, jo vartoseną priklauso nuo kolokato, o šiuo atveju jis negali būti verčiamas pažodžiui, vietoj jo reikia rasti panašaus ekspresyvumo veiksmažodį, kuris padėtų išlaikyti stilistinę originalo vertę, kaip tai padaryta profesionaliame vertime (žr. *pulti į vandenį*). Vertinant veiksmažodžio vertimą, studentų pasirinkimus galima būtų taip sugrupuoti:

- 1 Verčiama panašų ekspresyvumą turinčiu lietuvišku atitikmeniu (20): *šokti į vandenyną (4), pasinerti į vandenyną (2), pulti link vandenyno (2), lėkti į vandenyną (2), pultų į vandenyną, nerti į vandenyną, šokti į jūrą, šokti į vandenį, nerti į bangas, nerti į vandenyną, nerti į jas, lėkti link vandenyno, bėgti prie vandenyno, bėgtų į vandenyną*;
- 2 Verčiama neutraliu lietuvišku atitikmeniu arba nutolstama nuo originalo reikšmės (9): *traukti prie vandenyno, keliauti prie paplūdimio, traukti link vandenyno, traukti į vandenyną, pasileisti vandenyno link, trauktų iki pat vandenyno, keliauti prie jūros; lįsti į vandenyną, perskroستų vandenyną*;
- 3 Praleidžiama (3): *neturėjo priežasčių atsisakyti pasiplaukiojimo su banglente, norėdami plaukti banglente, stverti banglentę ir paplaukioti*.

Ekspresyvumo vertė išlaikyta 62 proc. vertimų, o likę pasirinkimai per daug nutolę stilistiškai ir semantiškai nuo originalo. Įdomu, kad keliuose pavyzdžiuose matoma lietuvių kultūros interferencija – *vandenynas* verčiamas *jūra*, o tai iškreipia originalo turinį. Analizuotų pavyzdžių diskusija apibendrinta **3 lentelėje**.

**3 lentelė.** Kolokacijų vertimo palyginimas

Vertimo pobūdis	Kolokacija	Originalo turinį / stilistinę vertę perteikiantis sprendimas	Stilistinės vertės neišlaikymas / nesklandus vertimas / praleidimas	Interferencija
1 Gali būti verčiama pažodžiui	<i>to <u>grab</u> the board</i>	28 / 88 %	4 / 12 %	–
	<i>to <u>spin</u> the board</i>	26 / 81 %	6 / 19 %	–
2 Pažodinis vertimas netinka šiame kontekste / neįprastas lietuvių k. vartosenai	<i><u>takeoff</u> zone</i>	15 / 47 %	13 / 41 %	4 / 12 %
	<i><u>annoyingly infrequent</u></i>	17 / 53 %	8 / 25 %	7 / 22 %
3 Pažodinis vertimas netinka dėl skirtingos tarpkalbinės raiškos	<i><u>gray</u> ponytail</i>	20 / 63 %	1 / 3 %	11 / 34 %
	<i>to <u>hit</u> the ocean</i>	20 / 63 %	12 / 37 %	–

<sup>19</sup> Pavyzdžiui, Google vertėjo siūlomi atitikmenys: <https://translate.google.com/?hl=lt&sl=en&tl=lt&text=hit&op=translate> (žiūrėta 2023 01 13).

Iš analizuotų pavyzdžių matyti, kad geriausiai besimokantiems vertimo sekėsi versti kolokacijas, kurioms galima taikyti pažodinį vertimą, kurių dėmenys turi panašius atitikmenis anglų ir lietuvių kalbose ir nepriklauso nuo konteksto (1 kategorija). O konteksto, stiliaus ypatumai ar tarpkalbinė / kultūrinė raiška ne visuomet įvertinami, todėl pasitaiko daugiau nesklandaus, stilistiškai neutralesnio vertimo atvejų. Tai sutampa ir su Kenny (2001), Feng et al. (2018), Bernardini (2015) profesionalių vertimų tyrimais, kuriuose pastebėta skurdesnė kolokacijų raiška. Kad antros ir trečios kategorijos kolokacijos yra sudėtingesnės besimokantiems vertėjams, rodo ir pastebėta anglų kalbos interferencija, kai vietoj lietuvių kalbai įprastos vartosenos pasirinktas pažodinis angliško dėmens vertimas (taip pat žr. Gregorio-Godeo ir Molina (2011), Bernardini (2015) tyrimus). Šie pastebėjimai galėtų būti naudingi rengiant besimokančių vertėjų pratybas ar aptariant kolokacijų vertimo ypatumus, stengiantis atkreipti dėmesį į galimus vertimo iššūkius. Žinoma, duomenų aptarimas yra deskriptyvus, rekomendacinio pobūdžio, o tikslesnėms išvadoms reikėtų didesnės pavyzdžių imties.

## Apibendrinimas

Straipsnyje buvo pristatytas besitęsiantis tarptautinis vertimo studijų srities projektas MUST, jo darbo aplinka Hypal4MUST (Obrusnik, 2014), rengiamos vertimų anotavimo gairės (TAS1/2), kaupiamų tekstynų (MUST-LT) taikymo ir tyrimo galimybės. Kaip atvejo analizė nagrinėtos skirtingo vertimo pobūdžio kolokacijos besimokančiųjų vertėjų tekstuose. Paaiškėjo, kad kaip ir ankstesniuose kolokacijų vertimo tyrimuose, studentų išverstose kolokacijose kurių dėmenims neturėtų būti taikomas pažodinis vertimas, gali būti jaučiama originalo kalbos įtaka, stilistinės raiškos paprastinimas, o vertimo sprendimai, lyginant su kolokacijomis, kurioms pažodinis vertimas tinka, ne tokie sėkmingi. Tokie tyrimai gali padėti formuojant vertėjams skirtų studijų dalykų, juose aptariamų aspektų ar atliekamų užduočių spektrą, pateikti duomenys gali būti perpanaudoti kitiems ir kitų klaidų tyrimams, pratimams generuoti, o pristatyta klaidų tipologija, anotuoti pasirinktiems vertimo aspektams. Ateityje MUST-LT tekstynas bus plečiamas, duomenys anotuojami taikant TAS2 anotavimo sistemą, plėtojami besimokančiųjų versti tyrimai.

**Padėka.** Dėkoju visiems MUST-LT projekto dalyviams prisidedantiems prie duomenų kaupimo ir jų tyrimo (Jolantai Kovalevskaitei, Aurelijai Leonavičienei, Jurgitai Macijauskaitei-Bondai, Erikai Rimkutei, Sidonai Žvaliauskienei) ir studentams, be kurių nuoširdaus dalyvavimo nebūtų ir tekstyno.

## Literatūros sąrašas

- 1 Baker M. (2011). In Other Words. A Coursebook on Translation. London/ New York: Routledge. <https://doi.org/10.4324/9780203832929>
- 2 Bernardini, S. (2015). Translation. The Cambridge Handbook of English Corpus Linguistics. Biber, D., Reppen, R (Eds.), Cambridge: CUP, 515-536.<https://doi.org/10.1017/CBO9781139764377.029>
- 3 Bernardini, S. (2007). Collocations in translated language: Combining parallel, comparable and reference corpora. Proceedings of the Corpus Linguistics Conference (CL2007), Birmingham: University of Birmingham, 1-16. <https://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2007/15Paper.pdf>.
- 4 Feng, H., Crezee, I., & Grant, L. (2018). Form and meaning in collocations: a corpus-driven study on translation universals in Chinese-to-English business translation. *Perspectives*, 26(5), 677-690. <https://www.tandfonline.com/doi/abs/10.1080/0907676X.2018.1424222>.<https://doi.org/10.1080/0907676X.2018.1424222>
- 5 Granger, S., & Lefer, M.A. (2018). MUST: A collaborative corpus collection initiative for translation teaching and research. In Granger, S., Lefer, M.-A. & Aguiar de Souza Penha Marion, L. (Eds.) *Book of Abstracts. Using Corpora in Contrastive and Translation Studies Conference (5th ed)*. CECL Papers, 1. Louvain-la-Neuve: Centre for English Corpus Linguistics/Université catholique de Louvain, 72-73.
- 6 Granger, S., & Lefer, M.A. (2020). The Multilingual student translation corpus: a resource for translation teaching and research. *Language Resources and Evaluation*, 54, 1183-1199. <https://rdcu.be/b0MkU>. <https://doi.org/10.1007/s10579-020-09485-6>

- 7 Gregorio-Godeo, E., & Molina, S. (2011). Collocation and the translation of new: an English-Spanish electronic dictionary of multi-word combinations as a translation tool. *Perspectives: Studies in Translatology* 19(2), 135-152. <https://doi.org/10.1080/0907676X.2010.544747>
- 8 Keny, D. (2001). *Lexis and Creativity in Translation. A Corpus-based Study*. Manchester: St. Jerome.
- 9 Kovalevskaitė, J., & Vaičenonienė, J. (2015). Technologijos ir praktinės užduotys vertimo mokyme: studentų nuomonės tyrimas. *Darnioji daugiakalbiystė / Sustainable Multilingualism*, 7, Vilnius: Versus Aureus, 160-180. <https://doi.org/10.7220/2335-2027.7.9>
- 10 Newmark, P. (1988). *A Textbook of Translation*. New York / London: Prentice Hall.
- 11 Obrusnik, A. (2014). Hypal: A user-friendly tool for automatic parallel text alignment and error tagging. Eleventh International Conference Teaching and Language Corpora, Lancaster, 20-23 July 2014, 67-69.
- 12 Øverås, L. (1998). In Search of the third code: an investigations of norms in literary translation. *Meta*, 43(4), 557-570. <https://www.erudit.org/fr/revues/meta/1998-v43-n4-meta169/003775ar.pdf>. <https://doi.org/10.7202/003775ar>
- 13 Schaeffner, Ch. (2004). Metaphor and translation: some implications of a cognitive approach. *Journal of Pragmatics*, 36, 1253-1269. <http://docenti.unimc.it/gaia.aragrande/teaching/2017/18210/files/lecture-per-esame-par-te-teorica/cognitive-approach-to-the-translation-of-metaphor-schaffner-2004>. <https://doi.org/10.1016/j.pragma.2003.10.012>
- 14 Steen, G. (2014). Translating metaphor: what's the problem? Miller, D. R., & Monti, E., (Eds.). *Translating Figurative Language*, Bologna, CeSLiC, 11-24.
- 15 Stulpinaitė, M., Horbačasienė, J., & Kasperavičienė, R. (2016). Issues in translation of linguistic collocations. *Kalbų Studijos / Studies about Languages* 29, 31-41. <https://doi.org/10.5755/j01.sal.0.29.15056>
- 16 Vaičenonienė, J. (2012). Nejprastų kolokacijų vertimas grožinės literatūros kūrinuose: tekstynų paremtas tyrimas. *Darbai ir dienos*, 58, Kaunas: VDU, 117-136. <https://doi.org/10.7220/2335-8769.58.7>
- 17 Vaičenonienė, J., Kovalevskaitė, J., & Rimkutė, E. (2021). Kolokacijų vertimas: žvilgsnis į vertėjų virtuvę. *Gimtoji kalba*, 2, 3-13.

#### Programinė įranga.

Anthony, L. (2022). *AntConc* (Version 4.2.0) [Computer Software]. Tokyo, Japan: Waseda University. <https://www.lauranceanthony.net/software>

## Abstract

### Jurgita Vaičenonienė

#### Student translation corpus MUST-LT: a case analysis of collocation translation

The translation of collocations is an under-researched topic in the field of Translation Studies; more attention is given to professional translator output research, while the texts of trainee translators have not been extensively studied so far. On the other hand, as Granger and Lefer (2018) maintain, the analysis of the work of trainee translators can provide useful knowledge for both pedagogical and research purposes. The aim of this article is to present an international Multilingual Student Translation corpus project to Lithuanian researchers and show the application possibilities of the MUST-LT sub-corpus compiled within the project framework. A case study of collocation translation in the English-to-Lithuanian MUST-LT sub-corpus (46299 words) is presented. 6 collocations (192 translations) are analysed, taking into account if their translation (1) may or (2) may not be word for word due to the context, habitual Lithuanian usage, or (3) cross-lingual differences. The results have shown that English interference and neutralization of stylistic expression occur in collocations for which word-for-word translation should not be applied; student translation decisions are less successful compared to collocations which may be translated literally. Application of trainee translator output can be useful for developing translation studies oriented exercises, curricula, or translation research in general.

---

**Apie autorę****JURGITA VAIČENONIENĖ**

Doc. dr., Vytauto Didžiojo universitetas, Skaitmeninių išteklių ir tarpdisciplininių tyrimų institutas, Kaunas, Lietuva

**Mokslinių interesų sritys**

Tekstynų lingvistika, vertimo studijos, skaitmeninė humanitarika

**Adresas** V. Putvinskio g. 23-216, Kaunas, LT-44243, Lietuva

**El. paštas** jurgita.vaicenoniene@vdu.lt

