

ПАРАЛЕЛЬНИЙ ПЕРЕКЛАДАЦЬКИЙ КОРПУС: ПЕРЕВАГИ ТА ПРИКЛАД ЗАСТОСУВАННЯ В АНАЛІЗІ ХУДОЖНЬОЇ ЛІТЕРАТУРИ

Гриців Наталія Миколаївна,
кандидат філологічних наук, доцент,
доцент кафедри прикладної лінгвістики
Інституту комп'ютерних наук
та інформаційних технологій
Національного університету «Львівська політехніка»
nhrytsiv@yahoo.com, nataliia.m.hrytsiv@lpnu.ua
orcid.org/0000-0001-6660-7161

Кулина Ольга Іванівна,
кандидат філологічних наук,
старший викладач кафедри прикладної лінгвістики
Інституту комп'ютерних наук
та інформаційних технологій
Національного університету «Львівська політехніка»
olha.v.kulyna@lpnu.ua
orcid.org/0000-0002-2334-0660

Мета розвідки – розглянути, описати та проаналізувати взаємодію людини і комп'ютерних технологій, яка стосується опису, генерації, обробки та керування паралельним перекладацьким корпусом, з увагою до англійської та української мов. Сьогодні інструментарій корпусної лінгвістики активно застосовується для обробки даних паралельного перекладацького корпусу та статистичного аналізу отриманих даних. Актуальність цього дослідження зумовлена недостатньою увагою до практичного застосування паралельного перекладацького корпусу та лінгвостатистичних методів у перекладі художньої літератури і перекладознавстві, особливо для аналізу англійських оригінальних творів і українських перекладів з точки зору відтворення домінуючих характеристик авторського ідіостилю. Як показує досвід, відсутність зразка поетапного створення та аналізу корпусу перешкоджає науковцям, зокрема молодим ученим та студентам, активно послуговуватися перевагами корпусів у перекладі. Тому в цій статті ми ставимо собі за мету лаконічно розглянути найважливіші нюанси укладання та застосування паралельного перекладацького корпусу, подаючи їх з урахуванням рівня базових знань нашого адресата в царині філології та комп'ютерних технологій.

Застосовані в статті методи аналізу текстового корпусу дали змогу проілюструвати процес та результат розмітки, уточнити різницю між паралельним перекладацьким корпусом і корпусом паралельних текстів, окреслити етапи аналізу текстового корпусу та реструктуризацію одиниць перекладу, посутньо, унаочнити процедуру відбору мовних відповідників.

Результати та висновки впливають із висвітлених у роботі знахідок та рекомендацій, які можуть слугувати основою для кращого розуміння того, яким чином застосовувати корпуси під час вивчення авторського стилю і в який спосіб залучити отримані результати (а) до перекладознавчого аналізу, (б) якіснішого перекладу. Мовиться про те, що кількісний порівняльний аналіз у перекладі не обмежується статистичними даними для підрахунків одиниць текстового (структурного) рівня; послуговуючись кількісним порівняльним аналізом, маємо змогу різносторонньо розкривати зміст (сенс) авторських повідомлень. Окрім переваги корпусів для перекладознавчих досліджень, зокрема стосовно якості перекладу, отримані результати здатні збагатити особистий «нотатник» перекладача художньої прози.

Ключові слова: кількісний порівняльний аналіз, лінгвометрія, переклад (художній), розмітка тексту, прикладна лінгвістика.

**PARALLEL TRANSLATION CORPUS:
ADVANCES AND APPLICATION EXAMPLE IN THE ANALYSIS OF FICTION**

Hrytsiv Nataliia Mykolaivna,

*Candidate of Philological Sciences, Associate Professor,
Associate Professor at the Department of Applied Linguistics
Institute of Computer Science and Information Technology
of Lviv Polytechnic National University
nhrytsiv@yahoo.com, nataliia.m.hrytsiv@lpnu.ua
orcid.org/0000-0001-6660-7161*

Kulyna Olha Ivanivna,

*Candidate of Philological Sciences,
Senior Lecturer at the Department of Applied Linguistics
Institute of Computer Science and Information Technology
of Lviv Polytechnic National University
olha.v.kulyna@lpnu.ua
orcid.org/0000-0002-2334-0660*

The aim of the investigation is to address human and computer interaction issues related to the description, generation, application, and management of parallel translation corpus, with close attention to English and Ukrainian languages. Today, Corpus Linguistics is actively used for the processing of parallel translation corpus data and for statistical analysis of the data obtained. The topicality of this study is caused by the insufficient attention to the practical application of the parallel translation corpus and linguistic statistics in translation of fiction and translation studies, especially for the analysis of English original works and Ukrainian translations in terms of reproducing the dominant characteristics of the author's individual style. The evidence has it that the lack of illustrative sample of necessary stages of how to create corpus and analyze it prevents scholars, in particular young researchers and students, from active usage of the benefits of corpus in translation. Therefore, in this article, we aim to look in brief at the most important nuances of compiling and applying a parallel translation corpus. We consider our addressee to have a level of basic knowledge in the field of philology and computer technology.

The methods of text corpora analysis applied in the article helped to illustrate the process and result of text mark-up, clarify the difference between the parallel translation corpus and the corpus of parallel texts, outline the stages of text corpus analysis and restructuring of translation units, specify the procedure for selecting linguistic equivalents.

Results and conclusions are driven from the findings and recommendations outlined in the paper and, thus, can serve as a basis for a better understanding of how to use corpora in the study of author's individual style and how to integrate the results (a) into translation analysis, (b) better translation. It is claimed that quantitative comparative analysis in translation is not limited to statistics for calculating units of text (structural) level; by using quantitative comparative analysis, we are able to disclose the content (meaning) of the author's messages encoded in his/her creative writing. In addition to the benefits of corpora in translation studies research, in particular when it goes about translation quality assessment, the results obtained can enrich the personal notepad of a translator of prose.

Key words: quantitative comparative analysis, linguometry, translation (artistic), text markup.

1. Вступ

На наше переконання, застосування паралельного перекладацького корпусу в перекладі, як і вивчення цього явища, є актуальною, продуманою та своєчасною практикою і відповідає вимогам, продиктованим сучасними потребами різних предметних галузей, адже сьогодні спостерігаємо стрімкий розвиток корпусних студій у світовій науці, що підтверджує чимала кількість досліджень у цій галузі та різноманіття корпусів, які постійно з'являються. Попри беззаперечну наукову та методичну цінність цих видань у вивченні загальної теорії корпусної лінгвістики (Advances in Intelligent Systems, 2019), взаємозв'язок корпусної лінгвістики та інших філологічних дисциплін (Levchenko, 2019), типології корпусів і методики різноманітного залучення корпусних даних для аналізу (Altenberg, 2000; Corpus-based Language Studies, 2002; Hunston, 2002), розроблення загальних засад створення та вивчення природних мов (Кай, 1993; Sinclair 1991), вони зазвичай орієнтовані на системотехнічні та структурні характеристики лінгвістичних корпусів. Водночас перекладацькі теоретико-критичні студії (Baker, 1993; Corpus-based Translation Studies, 2002; Simard 1993; Tymoczko, 1998), які є важливою складовою частиною сучасного перекладознавства, здебільшого уміщують, проте не обмежують описово-аналітичні конструкти та не достатньо відображають прикладний аспект корпусів. Приємним винятком є спорадичні праці (Kenny, 2001; Munday, 1998; Ulrych, 1997), націлені на творчу зорієнтованість у роботі з паралельним перекладацьким корпусом.

З погляду дидактики в останні десятиліття спостерігаємо перехід до компетентнісного підходу в навчанні філологів (Бесєдіна, 2020; Ольховська, 2015; Черноватий, 2013), що сприяє новим науковим пошукам у питанні міжкультурних та міждисциплінарних досліджень, пов'язаних з назрілими викликами освіти в сучасному суспільстві. Увагу зосереджено на спробі впровадити сучасну інтеграційну модель навчання різних міжпредметних галузей з метою об'єднання заради побудови добротного фахового портфоліо перекладача-початківця.

Перелік умовних скорочень:

ТО – текст оригіналу, вихідний текст (у цьому випадку – англійський).

ТП – текст перекладу, цільовий текст, перекладений текст (тут – український).

ППК – паралельний перекладацький корпус.

КПА – кількісний порівняльний аналіз.

TEI – text encoding initiative / набори тегів та правила, додаток мови розширюваної розмітки (XML) для кодування текстів.

XML – Extensible Markup Language / розширювана мова розмітки.

NLP – natural language processing / обробка природної мови.

2. Теоретико-методологічні основи

У широкому сенсі термін «паралельний корпус» позначає тексти одних і тих же доменів, отриманих з різних мов, при цьому отримані взірці-приклади не обов'язково повинні бути перекладом (у класичному розумінні) оригінального тексту. Оскільки його не потрібно конфігурувати між двома мовами, наприклад, в індійському контексті, він може поширюватися на декілька мов, де основним критерієм є підтримка паралелізму текстових доменів або зразків тексту.

Натомість термін «перекладацький корпус» позначає тексти, що супроводжуються їхніми перекладами однією або декількома мовами. Він складається з оригінального тексту з мови джерела та його перекладу цільовою мовою. Іншими словами, перекладацький корпус повинен містити тексти та їхні переклади. Однак, хоча він зазвичай складається із двомовних та багатомовних текстів, він може містити будь-який текст, який знаходиться під поняттям «переклад» (Zanettin, 2000: 102).

У цьому дослідженні за своєю структурою ППК складається з двох частин: (а) тексту, взятого з мови джерела; та (б) її переклад із цільової мови.

Інструментарій корпусної лінгвістики в перекладознавстві: вступні зауваги. З перекладознавчого погляду, лінгвістичний корпус – у поєднанні з кількісним порівняльним аналізом – доречно застосувати на різних етапах: (а) до початку процесу перекладу; (б) аналізуючи безпосередній процес перекладу; (в) критично оцінюючи перекладений продукт із погляду структурної та образної відповідності. Нині в діючих системах штучного інтелекту зазначені застосування (з теоретичного, методологічного, практичного і прикладного ракурсів) вивчаються і використовуються в паралельному перекладацькому корпусі (далі – ППК). Саме ця ланка здатна вдовольнити горизонт очікування перекладача і перекладознавця, тому є напрочуд цікавою.

У розвідці наведено приклади розмітки/маркування тексту та вирівнювання тексту в ППК, подано різницю між ППК і корпусом паралельних текстів, пунктирно окреслено методи аналізу текстового корпусу, висвітлено процес реструктуризації одиниць перекладу, унаочнено етап вилучення мовних відповідників, подано фрагмент створення двомовної лексичної бази даних та ділянок, де ППК та отримані на його основі результати можуть найоптимальніше залучатися до перекладознавчих досліджень та оцінки якості перекладу; виокремлення перекладацьких помилок; з'ясування зв'язності тематичних та інформаційних структур, що увиразнює змістове наповнення оригінального твору, яке ілюструє мережу текстових взаємозв'язків згуртованих чи згрупованих в оригіналі та його (не)відтворення в перекладі.

Обґрунтування доцільності дослідження. Для ефективних перекладознавчих досліджень доцільно використовувати корпусо-орієнтований та описовий підходи поряд із деякими статистичними методами для зіставлення та аналізу відповідників у перекладі. Окрім того, можна скористатися цифровим роззначенням оригінального та перекладеного текстів. Таке рішення доволі зручне, зокрема працюючи з об'ємними текстами, у нашому випадку – це роман Донни Тартт «Щиголь» (Tartt, 2013; Tartt, 2016), який налічує понад 800 сторінок. Роман «Щиголь», написаний сучасною новелісткою, володаркою Пулітцерівської премії, Донною Тартт, вибрано для аналізу, оскільки його ще не досліджували в англо-українській паралелі з погляду ППК та з ракурсу побудови статистичного портфоліо автора та перекладача. Щобільше, роман «Щиголь» удостоївся медалі Ендрю Карнегі за досконалість у художній літературі та був екранізований у 2019 році.

Типологічні та прикладні характеристики аналізованого корпусу. *Анотований:* текстові дані предметної області роззначаються до рівня фраземи. *Двомовний:* відповідні тексти англійською та українською мовами. *Повний текст:* побудований із цілих текстів оригіналу та перекладу. *Корпус-взірець* (sample corpus): взірці оригіналу та їхні переклади уміщено до корпусу. *Ілюстративний:* розроблений для порівняння мовних особливостей оригіналу та перекладу. *Літературний корпус:* особливим різновидом корпусу-взірця є літературний корпус, який може бути додатково розділений на основі типу тексту, що входить до нього: (а) конкретний автор (Донна Тартт); (б) тип тексту одного автора (роман зрілого періоду творчості Донни Тартт); (в) конкретний текст одного автора (корпус роману Донни Тартт «Щиголь»); (г) особливий жанр тексту (роман); (д) особливий період (американський роман двадцять першого століття); (е) певна група (сучасний американський містичний та психологічний трилер); (є) певна тема (емоційний реєстр: емоції, зосереджені навколо болю, втрати, відчуження). *Паралельний:* текст англійською мовою разом з його перекладом українською мовою. *Корпус-взірець:* містить невелику колекцію текстів, які вибираються з особливою ретельністю та увагою, і їх слід детально вивчати. *Статичний:* не передбачає постійного поповнення багатьох текстів корпусу. *Письмовий:* друкований текстовий документ.

3. Процедури та результати

Дослідження проводилося в межах наукової теми «Застосування новітніх технологій для оптимізації інформаційних процесів природною мовою» кафедри прикладної лінгвістики Національного університету «Львівська політехніка»; куратор проекту – І. М. Кульчицький (к.т.н., доц.). Аналіз проведено на основі певних етапів: *Eman 1:* Створення електронних версій ТО (тексту оригіналу) та ТП (тексту перекладу); *Eman 2:* Роззначення текстів: структурне та з виокремленням лінгвальних ознак (ословеснення образу болю); *Eman 3:* Автоматичне опрацювання даних за допомогою прицільно створеного комп'ютерного програмного забезпечення.

Роззначення текстів: структурне та з виокремленням лінгвальних ознак (ословеснення образу болю).

Роззначення полягає у наданні текстам чи окремим текстовим компонентам спеціальних позначок (тегів). Це можуть бути позамовні відомості про автора та вибраний текст, структурні особливості тексту (розділ/підрозділ, абзац, речення, лексема, графема); питома лінгвальні ознаки (лексичні, граматичні, стилістичні, дискурсні прикмети складників тексту). Поряд із цим індивідуальний елемент розмітки (*тег*, від англ. *tag*) дає певне поле для маневрування з наміром прицільно вдовольнити особистісні наукові потреби перекладавця чи перекладача. Ця практика стала невід’ємною складовою частиною переважаючих новочасних досліджень і здатна суттєво доповнити теоретико-методологічні й практичні аспекти наукової розвідки. Це у свою чергу дозволяє пошукувцю панорамно бачити власне наукове дослідження, найперше з ракурсу прикладного застосування та доречності в сучасному мультимедійному середовищі. Також імпонує використання ППК та авторських розробок визначеного шаблону маркування текстів на комп’ютерній платформі, яке, безумовно, сприяє пошуковій діяльності, зокрема для великих за обсягом творів (текстів).

Позначки роззначення тексту:

<p> </p> – позначка абзацу

<s> </s> – позначка речення

<pain> </pain> – позначення мегаобразу *біль*

<loss> </loss> – позначення образу (втрата рідної людини)

<addiction> </addiction> – позначення образу (залежність: наркотична, алкогольна, від куріння)

<anxiety> </anxiety> – позначення образу (страх втратити друга)

<punishment> </punishment> – позначення образу (страх звинувачення і схоплення поліцією за крадіжку картини)

<rambling> </rambling> – позначення образу (сиротинство/бездомність)

<guilt> </guilt> – позначення образу (відчуття провини за крадіжку картини і продаж підробок)

<suicide> </suicide> – позначення образу (думки про самогубство/депресія)

<abandonment> </abandonment> – позначення образу (нерозділене кохання)

<disappointment> </disappointment> – позначення образу (байдужість батька та розчарування в ньому)

<obsession> </obsession> – позначення образу (неспроможність часто спостерігати за картиною)

<failure> </failure> – позначення образу (нереалізованість кар’єри батька)

<con n=1 > </con n=1 > – приклад ословеснення концептуалізованого образу в тексті (n=1; n – нумерація, 1 – порядковий номер).

Фрагмент роззначення оригінального тексту з увагою до мегаобразу БІЛЬ¹

<p><pain><loss><s><con n=1 >Thanks to my fever I had a lot of weird and extremely vivid dreams, <con>sweats where I thrashed around<con> hardly knowing if it was day or night, but on the last and worst of these nights I dreamed about my mother: </con n=1 > a quick, mysterious dream that felt more like a visitation. </s> <s> I was in Hobie’s shop– or, more accurately, some haunted dream space staged like a sketchy version of the shop–when she came up suddenly behind me so I saw her reflection in a mirror. </s> <s> <con n=2 > At the sight of her I was paralyzed with happiness; </con n=2 > it was her, down to the most minute detail, the very pattern of her freckles, she was smiling at me, more beautiful and yet not older, black hair and funny upward quirk of her mouth, not a dream but a presence that filled the whole room: a force all her own, a living otherness. </s> <s> <con n=3 > And as much as I wanted to, I knew I couldn’t turn around, that to look at her directly was to violate the laws of her world and mine; </con n=3 > she had come to me the only way she could, and our eyes met in the glass for a long still moment; but just as she seemed about to speak–with what seemed a combination of amusement, affection, exasperation–a vapor rolled between us and I woke up.</s></loss></pain></p>

<p><s>ii.</s> </p>

Фрагмент роззначення перекладеного тексту з увагою до мегаобразу БІЛЬ

<p><pain><loss><s><con n=1 >Через лихоманку я пережив безліч дивних і надзвичайно яскравих марень, обливаючись потом, не знаючи, котра година, але в останню й найгіршу з тих ночей мені наснилася мати: </con n=1 > то було швидке й таємниче сновидіння, схоже на несподіваний візит із потойбіччя. </s> <s> Я був у крамниці Гобі чи, точніше кажучи, в якомусь примарному просторі, радше схематично подібному до тієї крамниці, коли вона раптом зупинилася позаду мене так, що я побачив її віддзеркалення. </s> <s> <con n=2 >Помітивши її, я заціпенів від щастя; </con n=2 > то була вона, до найдрібніших деталей, з усім своїм ластовинням, мама всміхалася мені, навіть вродливіша, але ще не постаріла, зі своїм чорним волоссям і кумедним вигином рота, не сон, а її присутність заповнювала кімнату: то була її власна сила, жива інакшість. </s> <s> <con n=3 >І хоч як мені хотілося обернутися, я знав, що подивитися на неї очі в очі було б порушенням законів її світу і мого світу; </con n=3 > вона прийшла до мене в єдиний спосіб, у який могла прийти, і наші очі зустрілися в дзеркалі на тривалу мить, коли все завмерло; та коли, як мені здалося, вона хотіла заговорити – з виразом обличчя, що, схоже, був сумішшю радощів, прихильності й роздратування, – на нас накопився туман, і я прокинувся. </s></loss></pain></p>

Зразки фрагментів створеного корпусу оригінального тексту з увагою до мегаобразу БІЛЬ

<?xml version="1.0" encoding="UTF-8" standalone="yes"?>

<entry xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">

<pain>

¹ Роззначення цих фрагментів тексту виконав П. Вус – випускник кафедри прикладної лінгвістики НУ «Львівська політехніка» під керівництвом авторів статті та куратора проекту І. Кульчицького.

<lg>eng</lg>
 <nu>61</nu>
 _{abandonment}

(а) Фрагмент 1

<tx>Every moment of their visit had been torture.</tx>

</pain>
 < pain n>
 <lg>eng</lg>
 <nu>62</nu>
 _{abandonment}

(б) Фрагмент 2

<tx>as hard as I'd tried, I hadn't been able to stay away from Pippa, I'd hovered obtrusively and hated myself for it, so painfully excited had I been by her nearness:</tx>

</ pain>
 < pain>
 <lg>eng</lg>
 <nu>63</nu>
 _{abandonment}

(в) Фрагмент 3

<tx>The agony of her hand on my sleeve. </tx>

</ pain>
 < pain>
 <lg>eng</lg>
 <nu>63</nu>
 _{abandonment}

4. Результати КПА

Відтінок створеного паралельного перекладацького корпусу, з акцентуацією авторської домінанти *біль*

ТО			ТП		
1	loss	Thanks to my fever I had a lot of weird and extremely vivid dreams, sweats where I thrash	1	loss	Через захворювання я перешкодив бачити дитинку і надзвичайно жаркою мовчю, обливаючись потом, не знав
2	loss	At the sight of her I was paralyzed with happiness,	2	loss	Покитивши її, я задивлено від щастя;
3	loss	And as much as I wanted to, I knew I couldn't turn around, that to look at her directly was	3	loss	І хоч як мені хотілося обернутися, я знав, що подивитися на неї очі в очі було б порушенням законів її сві
4	loss	I felt a petty, irrational pang of disappointment	4	loss	А проте я відчував легкий ірраціональний біль розчарування
5	loss	skull-cracking headache, practically seeing double	5	loss	голова в мене розколювалася від болю, речі двоїлися перед очима
6	disappointment	To think of something happening to my mother was especially frightening because my dad	6	disappointment	Думати, що з моєю матір'ю щось статися ввався лячно, мені було особливо страшно, адже сподіватися на
7	disappointment	Apart from this daily awkwardness, I didn't see him much. He didn't eat dinner with us or at	7	disappointment	Окрім цих повсякденних незручностей, бачив я його нечасто. Він не об'їдав із нами й не бував у моєї сям
8	disappointment	towards the end of the week, a sketchy note from my dad arrived (postmarked Newark, N	8	disappointment	натроєнка тижня, від батька надійшло коротке повідомлення (повідомлення Ньюарком, штат Нью-Джерсі),
9	rambling	Even though the apartment was rent stabilized, getting by without my dad's salary was a	9	rambling	Хоч ми сподівалися стабільно комерце за наше помешкання, було надзвичайно складно щомісяць об'їда
10	rambling	cold, wet, with a tooth-crunching headache—it struck me that with my dad gone, no one in	10	rambling	змараній, мокрої, з болем, від якого розколювалася голова, — до мене раптом дійшло, що, позаяк нащ
11	loss	The living room—normally so airy and open, buoyant with my mother's presence—had sh	11	loss	Вітальня — за нормальних обставин така свіжа й відкрита, ниттерадно завдяки присутності моєї матері
12	loss	I could feel my heart beating,	12	loss	Я чув, як біжить моє серце;
13	loss	I missed her so much I wanted to die: a hard, physical longing, like a craving for air underw	13	loss	Я так сумував за нею, що мені хотілося померти: то було тяжка фізична туга, схожа на страждання без по
14	loss	Lying awake, I tried to recall all my best memories of her—to freeze her in my mind so I w	14	loss	Не в змозі заснути, я намагався пережити всі свої найкращі спогади про неї — заморозити її у своїй свід
15	loss	drifting uneasily between dreaming and sleep	15	loss	важно переключити від думки до сну.
16	rambling	My panic must have been written plainly on my face.	16	rambling	На моєму обличчі було, мабуть, написано, що я перебуваю в стані паніки.
17	rambling	I was a minor, and that my parent or guardian had to be present at an official interview	17	rambling	в неповнолітній, і один із моїх батьків або опікунів повинен був бути присутнім на офіційному інтерв'ю
18	punishment	I would simply have to admit it and face the consequences, no matter what the consequ	18	punishment	Тоді мені довелося просто визнати свою провину й змириться з наслідками, хоч би якіми вони були
19	punishment	At a couple of points, I was on the verge of blurting it out, in my terror.	19	punishment	Десь разів за два я мало не зізнався в усьому, опанований страхом.
20	punishment	It was complicated. Every time I thought of it my stomach squirmed,	20	punishment	Проблема було для мене тяжкою. Щоразу, коли я думав про неї, вніт мені судилося,
21	punishment	What if someone found the picture? What would happen to me?	21	punishment	А що буде, як хтось знайде картину? Що станеться зі мною?
22	punishment	Yet in my homeless limbo, it seemed insane to step up and admit to what I knew a lot of pe	22	punishment	Але в моєму бездомному становищі мені здавалося надто ризикованим розкрити рот і зізнатися в тому
23	punishment	Where the hell was the painting? Though I was nearly blind with panic, I didn't stop but kep	23	punishment	Де, в біса, картина? Хоч я майже осліп від паніки, та не зупинився, в біже діль, який коридором, на авто
24	loss	My mother's bedroom was dark and cool, and the faint, just-detectable smell of her perfum	24	loss	У спальні матері було темно й прохолодно, і ледь-чутні пахощі її парфумів були чимось більшим, ніж в як
25	loss	a million other heart-piercing sights.	25	loss	мільйона інших дитячих, що розривали мені серце.
26	failure	It's one of those lifelong regrets. I would have loved to do something with my gift but I did	26	failure	Це той наль, який, мабуть, буде вічним. </s> </s> Я хотів би якось використати свій дар, але мені не поща
27	failure	Then—after he'd married my mother—it had all fizzled out. He had a long list of reasons w	27	failure	Потім, після того як він одружився з моєю матір'ю, його ентузіазм пропав. Він називав довгий список пр
28	punishment	While we were waiting to get our boarding passes I was stiff with fear, fully expecting Secu	28	punishment	Потім ми стояли в черзі, щоб одержати посадкові талони, в весь злякан від страху, чекуючи, що працівни
29	punishment	I felt closed-in and terrified in the bright press of strangers—conscious too, as if everyo	29	punishment	я з жакон відчував, як мене затискає в тюрму колі незнайомиця, — я стовав наче тольк, і всі на-бто див

У дослідженні було виділено 11 периферійних образів мегаобразу *біль*, які ослівеснено в романі Донни Тартт «Щиголь» на основі таких семантичних полів: втрата рідної людини (loss), залежність – наркотична, алкогольна, від куріння (addiction), страх втратити друга (anxiety), страх звинувачення і схоплення поліцією за крадіжку картини (punishment), сиротинство/бездомність (rambling), відчуття провини за крадіжку картини і продаж підробок (guilt), думки про самогубство/депресія (suicide), нерозділене кохання (abandonment), байдужість батька та розчарування в ньому (disappointment), неспроможність часто спостерігати за картиною (obsession), нереалізованість кар'єри батька (failure). На основі отриманих даних проведено детальний аналіз кожного образу та з'ясовано спосіб концептуалізації: експліцитний (очевидний) та імпліцитний (прихований).

Загальна кількість образів – 107 прикладів. Вони обрамлюють цілісний мегаобраз відчуття *болу*, що його ослівеснено в романі Донни Тартт «Щиголь». Найчастіше в романі зустрічається образ *abandonment* – 23 приклади.

Найменш частотний серед усіх образів – *failure*. Попри те, що його вжито тільки двічі, він несе потужне семантичне наповнення; нереалізованість кар’єри батька, як її змальовано у творі, – це своєрідне протиставлення двох начал – материнського (довкола цього розлого описано ніжні до матері відчуття та пекельний біль від її втрати) та батьківського (із загостреною увагою до деструктивних стосунків батька та сина, неспроможність сина скористатися сильним батьківським плечем, так би мовити, та його підтримкою в особливо скрутний час). Стосовно решти вибраних для аналізу образів маємо такі результати: *loss* – 15 взірців, *addition* – 14 взірців, *anxiety* – 12 взірців, *punishment* – 15 взірців, *rambling* – 7 взірців, *guilt* – 6 взірців, *suicide* – 4 взірці, *disappointment* – 3 взірці, *obsession* – 6 взірців. Для того щоб визначити, якими засобами ословеснено *біль* у виокремлених корпусом прикладах, було створено окрему таблицю, в якій ми виділили дві колонки: перша колонка – винятково для прикладів, у яких вжито безпосередню лексему *біль* та її семантико-граматичні похідні (позначено одиницею); друга колонка призначена для тих прикладів, у яких спостерігаємо синоніми до слова *біль* без безпосереднього врахування слова *біль* і його похідних (позначено двома одиницями). Остання колонка вказує на частоту слова.

Після проведення статистики експліцитного ословеснення *болю*, всі лексеми ТО та ТП було окремо переміщено у дві таблиці програми Excel. Таким чином, підраховано кількість усіх слів окремо для оригінального тексту та для перекладу. Під час даного процесу було також підраховано частоту всіх слів у відповідних текстах. Перша колонка – це порядковий номер конкретного слова, друга колонка вказує на частоту вибраного слова.

ТО				ТП					
685	sorrow	1	1	3	953	туга	1	1	2
737	terror	1	1	1	954	туги	1	1	1
777	tooth-crunching	1	1	1	996	хвороба	1	1	1
779	torment	1	1	1	1031	шок	1	1	1
781	torture	1	1	1	43	болею	1		3
5	aching	1		1	44	болю	1		4
345	headache	1		2	45	болючий	1		1
380	hurt	1		1	46	болючими	1		1
					47	боляче	1		1

Рис. 1. Вибірка за ознакою в таблиці Excel

З цього отримуємо співвідношення ословеснення *болю* та його семантико-граматичних різновидів (боліти, боляче, болючий). Спостерігаємо таке: у ТО та ТП експліцитні ословеснення *болю* майже ідентичні та відрізняються несуттєво, натомість синоніми піддаються певній трансформації в перекладі. У ТО мегаобраз *біль*, поряд із його похідними словами, становить 45,78%. Синонімічний ряд поняття *біль* становить 13,22%. У ТП лексема *біль* та його похідні форми становлять 59,79%, а синоніми до *болю* – 16,21%.

У підсумку: ТО налічує 294 619 слів, ТП – 268 409 слів. У вибраних підтекстах (корпусі) в оригіналі роману «Щиголь» авторка оригіналу послуговується 2248 словами, в перекладі відповідного підтексту маємо 2031 слово. Експліцитна концептуалізація *болю* налічує 58 прикладів у ТО і 75 прикладів у ТП цього твору. Уживаність безпосередньої лексеми *біль* та її семантико-граматичні деривативи становлять 13 позицій, у мові перекладу – 16. Відповідно, синоніми до цього поняття вжито 45 разів у ТО та 59 разів у ТП роману «Щиголь».

5. Прикладне застосування ППК вибраного підтексту

ТО	ТП
Thanks to my fever I had a lot of weird and extremely vivid dreams, sweats where I thrashed around hardly knowing if it was day or night, but on the last and worst of these nights I dreamed about my mother	Через лихоманку я пережив безліч дивних і надзвичайно яскравих марень, обливаючись потом, не знаючи, котра година, але в останню й найгіршу з тих ночей мені наснилася мати
at the sight of her I was paralyzed) [...] as much as I wanted to, I knew I couldn't turn around, that to look at her directly was to violate the laws of her world and mine	помітивши її, я заціпенів [...] хоч як мені хотілося обернутися, я знав, що подивитися на неї очі в очі було б порушенням законів її світу і мого світу
Sometimes, in the night, i woke up wailing	Іноді я прокидався вночі, виючи
I felt a petty, irrational pang of disappointment	я почував легкий ірраціональний біль розчарування
Boris never seemed annoyed or even very startled when I woke him, as if he came from a world where there was nothing so unusual in a nocturnal howl of pain	Борис ніколи не дратувався й навіть не дуже дивувався, коли я його будив так, ніби прийшов сюди зі світу, в якому прокидатися вночі, виючи від болю, не вважалось чимось незвичним
My mother's bedroom was dark and cool, and the faint, just-detectable smell of her perfume was almost more than I could bear	У спальні матері було темно й прохолодно, і ледь чутні пахощі її парфумів були чимось більшим, ніж я міг витримати

a million other heart-piercing sights	мільйона інших деталей, що розривали мені серце
The living room—normally so airy and open, buoyant with my mother’s presence —had shrunk to a cold, pale discomfort, like a vacation house in winter	Вітальня – за нормальних обставин така свіжа й відкрита, життєрадісна завдяки присутності моєї матері – тепер здавалася холодною, блідою й незатишною, схожа на дачний будинок узимку
I could feel my heart beating	Я чув, як гупає моє серце
I missed her so much I wanted to die: a hard, physical longing, like a craving for air underwater	Я так сумував за нею, що мені хотілося померти: то була тяжка фізична туга, схожа на страждання без повітря під водою
Lying awake, I tried to recall all my best memories of her —to freeze her in my mind so I wouldn’t forget her [...] drifting uneasily between dreaming and sleep.	Не в змозі заснути, я намагався пережити всі свої найкращі спогади про неї – заморозити її у своїй свідомості, щоб ніколи не забути про неї) [...] важко переходячи від дрімоти до сну

Вживання ключових слів. На цьому етапі в нагоді стають генератори вживання ключових слів. Таких платформ вільного онлайн-доступу маємо сьогодні чимало.

Ми розділили отримані ключові слова на відповідні Блоки.

Блок 1 зумовлено нормами англійської мови, коли іменник вимагає обов’язкового вживання присвійного займенника (*my, his, her* тощо) або артикля (означеного *the* чи неозначеного *a/an*). Ця ознака аломорфна у відношенні до української мови. Тому відмінності у Блоку 1 в зіставленні ТО та ТП зумовлені об’єктивною необхідністю граматичних особливостей мовної пари *англійська :: українська*.

Для нас цікавий Блок 3, оскільки в ньому зафіксовано опорні смислові слова, що їх вжито на позначення мегаобразу *біль*.

Це та варіативна робоча платформа, якою може скористатися як перекладач на етапі саморедагування та кінцевого вичитування, так і перекладознавець для перекладознавчих спостережень стосовно якості перекладу.

Зосереджуємо увагу на головних частинах мови (дієслово, іменник, прикметник). Частково розглядаємо їхні деривативи (дієприкметник).

	ТО	ТП
Блок 1	her – 8; and – 7; the – 6; was – 6;	
Блок 2		бути – 3; світ – 3;
Блок 3	wanted – 2; had – 2; woke – 2; mother’s – 2; like – 2; could – 2; where – 2; world – 2; around – 2; much – 2; night – 2;	виючи – 2; чимось – 2; серце – 2; матері – 2; вночі – 2; очі – 2;
Блок 4		витримати – 1; міг – 1; ніж – 1; більший – 1; через – 1; парфумів – 1; чутні – 1; ледь – 1; прохолодно – 1; темно – 1; спальні – 1; незвичним – 1; вважати – 1; пахощі – 1; інший – 1; мільйона – 1; здавалася – 1; чув – 1; узимку – 1; будинок – 1; дачний – 1; схожий – 1;

Ключові слова ТО вибраного підтексту: *іменники* – mother, night, world; *дієслова* – wanted, woke. Ключові слова ТП вибраного підтексту: *іменники* – матір, серце, очі, ніч (вночі), світ; *дієприслівник* – виючи.

Розбіжності – очевидні, проте саме вони – найцікавіші. Отже, помітні такі рішення перекладача:

Для лексеми ТО *wailing* та *howl* запропоновано один відповідник *виючи*. Послугуючись компонентним аналізом, бачимо, що англійська лексема *to wail* має семему *weeping* з українським відповідником *ридання* (рідше – *схлипування*), проте лексеми ТО *howl* відтворено прямим українським відповідником *вити*. Уривок *Sometimes, in the night, i woke up wailing* можна передати так: *Іноді я прокидався вночі, ридуючи*. Така альтернатива жодним чином не порушує «комунікативної динаміки» (у термінології М. Бейкер) висловлювання, напрочуд, таке рішення збагатило б коефіцієнт ексклюзивності ТП.

Генератор ключових слів не розпізнає словосполучення, такі як *heart-piercing*, як окремі значущі одиниці. Тому її було вилучено з автоматичного розрахунку. Вибір перекладача зумовлений аломорфною особливістю англійської та української мов, тобто відсутністю конструкцій *Participle* в українській мові. Таким чином, щоб надати висловлюванню природнього звучання, партиципну конструкцію ТО *a million other heart-piercing sights* складно відтворити іншим способом, ніж *мільйона інших деталей, що розривали мені серце*. Вибір перекладача цілком умотивований.

Фразу TO *to look at her directly* відтворено за допомогою фраземи *подивитися на неї очі в очі* (зворотній переклад: *to look at her eyes to eyes*). Такий вибір певним чином увиразнює локалізацію в перекладі на рівні концептуалізації та символічної образності цільової (української) культури. Знаємо, що *очі* в українській етнокультурній традиції – символ багатий та експресивний (*очі – дзеркало душі; її очі, мов волошки; карі очі, чорні брови*) (Словник символів культури України: 152–153).

Часто вживані випадки з надмірною експресивністю перекладу можуть слугувати плідною платформою для вивчення індивідуального стилю (почерку) перекладача. Зауважимо, що шукати такі випадки вручну – не надто доцільно, оскільки праця копітка та затратна з погляду часу і зусиль. Натомість за умови правильного підходу до формування ППК, такі дані отримуємо, натиснувши одну кнопку.

У наступному аналізованому уривкові *світ* є складником образу *болю*, оскільки його метафоризовано засобами психологічних та емоційних відчуттів мук, страждань та болю, приміром: *I knew I couldn't turn around, that to look at her directly was to violate the laws of her world and mine* (протиставлення світу живих та мертвих) and *world where there was nothing so unusual in a nocturnal howl of pain*. Цю ознаку збережено в перекладі, її частково увиразнено, я знав, що подивитися на неї очі в очі було б порушенням законів її *світу* і мого *світу* and *ніби прийшов сюди зі світу, в якому прокидатися вночі, виючи від болю, не вважалося чимось незвичним*.

Попри невелику кількість проілюстрованих прикладів, ці кілька зразків показують, що інструменти NLP (опрацювання природної мови) – це вагомні ресурси для аналізу перекладів. Поряд із лінгвістично орієнтованими підходами вони корисні на різних етапах процесу перекладу: (а) доперекладацький аналіз, статистичне профілювання оригіналу автора, визначення авторського ідіостилу за допомогою лінгвотриї та генераторів ключових слів, роззначення домінуючих лексичних та стилістичних особливостей – поряд із навиками та досвідом перекладача, – безсумнівно, сприятимуть успішному та якісному перекладу; (б) процес перекладу: інструменти NLP можуть допомогти виявити динамічні та сталі елементи в тексті; вирішити, які стратегії залучати (експлікацію чи компенсацію), зосібна, мати можливість їх збалансувати; це полегшує лексичний вибір для перекладача; невідповідність стане швидше усвідомленим рішенням перекладача, а не помилкою чи недбалістю; (в) перегляд готового перекладу: з погляду оцінки якості перекладу та критики перекладу, інструменти NLP, а також результати ППК дають змогу бачити цілісну картину, простежити об'єктивні та суб'єктивні зсуви в перекладі (граматичні, лексичні, текстові), увиразнити когезію та когерентність у ТП. Визначивши одиницю перекладу, маємо змогу визначити механізми та вивчити структуру транзитивності, з'ясувати, чи витримано в перекладі ідейно-тематичну послідовність, задекларовану в оригіналі.

6. Висновки

Перекладацький паралельний корпус – це невід'ємна сьогодні в науці цифрова робоча панель, своєрідна пластична платформа, покликана вдовольнити науково-пошукові потреби науковців філологічного, зокрема, перекладознавчого профілю. Авторська розмітка допомагає вибрати найоптимальніший спосіб кодування та роззначення первинних даних у залежності від цільової потреби та дозволяє адаптувати принципи ТЕІ до власного дослідження, яке водночас не йде в розріз із існуючими стандартами. У нашому випадку авторську розмітку адаптовано для увиразнення переважаючого образу американського роману Донни Тартт «Щиголь» та його українського перекладу. Це дало змогу опрацювати чималий текстовий матеріал та отримати численні дані, якими можемо послуговуватися для різного призначення: вивчати стилі автора та перекладача – у сукупності та відділено, аналізувати перекладацькі рішення, простежувати зсуви та невідповідності, ілюструвати матеріал під час викладання перекладу та перекладознавства.

Серед переваг – зручність зберігання об'ємних даних та легка доступність до різномірних обчислень, які посутньо можемо долучати до різних аналізів (статистичних, мовознавчих, літературознавчих, перекладознавчих). Текст можна опрацювати поодиноці та групою попередньо натренованих учасників проекту. Отримані дані легко накопичувати, доповнювати й удосконалювати, які згодом стануть добротною візитівкою перекладача чи науковця. На часі – створення спільної всеукраїнської платформи ППК для взірцевих зразків художнього перекладу українською мовою. Така амбіція уможливує створення українського перекладацького еталону в різних жанрах: прозі, поезії, драмі.

Література:

1. Advances in Intelligent Systems and Computing III: Selected papers from the International Conference on Computer Science and Information Technologies, CSIT 2018, September 11-14 Lviv, N. Shakhovska and M. Medykovskyy (eds), Ukraine. Springer: Springer Nature Switzerland, 2019.
2. Altenberg B., Aijmer K. The English-Swedish parallel corpus: A resource for contrastive research and translation studies. In *Corpus Linguistics and Linguistics Theory*, ed. C. Mair, and M. Hundt, 2000. P. 15–33.
3. Baker M. Corpus linguistics and translation studies: Implications and applications. In *Text and Technology: In Honour of John Sinclair*, ed. M. Baker, G. Francis, and E. Tognini-Bonelli, Amsterdam: John Benjamins, 1993, 233–250.
4. *Corpus-based Language Studies: An Advanced Resource Book*, ed. T. McEnery, R. Xiao, Y. Tono, 2006, Routledge.
5. *Corpus-based Translation Studies: Theory, Findings, Applications*, ed. S. Laviosa, 2002, Rodopy.
6. Dash N.S., Arulmozi S., *History, Features, and Typology of Language Corpora*, Springer: Springer Nature Singapore, 2018.
7. Hunston S. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press. Isabelle, P., Dymetman, M., Foster, G., Jutras, J.-M., Macklovitch, E., Perrault, F., Ren, X. and M., 2002.

8. Kay M., Röscheisen M. Text-translation alignment. *Computational Linguistics*, 1993, 191: 13–27.
9. Kenny D. *Lexis and Creativity in Translation. A corpus-based study*. Routledge, 2001.
10. Levchenko O., Dilai M. Attitudes Toward Feminism in Ukraine: A Sentiment Analysis of Tweets. In *Advances in Intelligent Systems and Computing III: Selected papers from the International Conference on Computer Science and Information Technologies*, CSIT 2018, September 11-14 Lviv, N. Shakhovska and M. Medykovskyy (eds), Ukraine. Springer: Springer Nature Switzerland, 2019. P. 119–131.
11. Munday J.A Computer-assisted approach to the Analysis of Translation Shifts, *Meta*, XLIII, 4, 1998.
12. Simard M. Translation analysis and translation automation. Proceedings of the Fifth International Conference on Theoretical and Methodological Issues in Machine Translation TMI 93: MT in the Next Generation, Kyoto, Japan, pp. 1993, p. 12–20.
13. Sinclair J. *Corpus, concordance, collocation*, Oxford : Oxford University Press, 1991, 137 p.
14. Tartt D. *The Goldfinch*, Little Brown Publishing, 2013. 771 p.
15. Tymoczko M. Computerized corpora and the future of translation studies, *Meta*, 1998. 434. P. 652–659.
16. Ulrych M. The impact of multilingual parallel concordancing on translation. In *Practical Applications in Language Corpora*, ed. B.L. Tomaszczyk, and P.J. Melia, Lodz: Lodz University Press, 1997. P. 421–436.
17. Zanettin F. Parallel corpora in translation studies: Issues in corpus design and analysis. In *Intercultural Faultlines. Research Models in Translation Studies I: Textual and Cognitive Aspects*, ed. M. Olohan, Manchester: St. Jerome, 2000. P. 105–118.
18. Беседина Є.В. Концептуальна модель професійної підготовки конкурентоспроможних перекладачів в умовах диверсифікації вищої освіти. URL: http://www.nbuv.gov.ua/portal/Soc_Gum/Pfto/2009_3/files/ped_2009_03_07_Besedina.pdf (дата звернення: 15.04.2020).
19. Ольховська А.С. Системи автоматизованого перекладу як складова фахового навчання: переваги та недоліки. *Філологічні трактати*. 2015. Т. 7. № 2. С. 102–108.
20. Словник символів культури України. URL: https://library.udpu.edu.ua/library_files/397017.pdf (дата звернення: 11.04.2020).
21. Тартт Д. Щиголь. Переклад з англійської: Віктор Шовкун. Харків : КСД, 2016. 816 с.
22. Черноватий Л.М. Методика викладання перекладу. Вінниця : Нова Книга, 2013, 376 с.

References:

1. *Advances in Intelligent Systems and Computing III: Selected papers from the International Conference on Computer Science and Information Technologies*, CSIT 2018, September 11-14 Lviv, N. Shakhovska and M. Medykovskyy (eds), Ukraine. Springer: Springer Nature Switzerland, 2019.
2. Altenberg B., and K. Aijmer. The English-Swedish parallel corpus: A resource for contrastive research and translation studies. In *Corpus Linguistics and Linguistics Theory*, ed. C. Mair, and M. Hundt, 2000, pp. 15–33.
3. Baker M. Corpus linguistics and translation studies: Implications and applications. In *Text and Technology: In Honour of John Sinclair*, ed. M. Baker, G. Francis, and E. Tognini-Bonelli, Amsterdam: John Benjamins, 1993, 233–250.
4. Besyedina Ye. V. (2009). Kontseptual'na model' profesinyoi pidhotovky konkurentospromozhnykh perekladachiv v umovakh dyversyfikatsiyi vyshchoyi osvity. [The conceptual model of competitive translators' professional training in the conditions of diversification of higher education]. Available at: http://www.nbuv.gov.ua/portal/Soc_Gum/Pfto/2009_3/files/ped_2009_03_07_Besedina.pdf (in Ukrainian)
5. Chernovatyy L. M. (2013). *Metodyka vykladannya perekladu*. [Methods of teaching translation]. Vinnytsya : Nova Knyha.
6. *Corpus-based Language Studies: An Advanced Resource Book*, ed. T. McEnery, R. Xiao, Y. Tono, 2006, Routledge
7. *Corpus-based Translation Studies: Theory, Findings, Applications*, ed. S. Laviosa, 2002, Rodopy
8. Dash N. S., S. Arulmozi, *History, Features, and Typology of Language Corpora*, Springer: Springer Nature Singapore, 2018.
9. Hunston S. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press. Isabelle, P., Dymetman, M., Foster, G., Jutras, J-M., Macklovitch, E., Perrault, F., Ren, X. and M., 2002.
10. Kay M., Röscheisen M. Text-translation alignment. *Computational Linguistics*, 1993, 191: 13–27.
11. Kenny D. *Lexis and Creativity in Translation. A corpus-based study*. Routledge, 2001
12. Levchenko O., Dilai M. Attitudes Toward Feminism in Ukraine: A Sentiment Analysis of Tweets. In *Advances in Intelligent Systems and Computing III: Selected papers from the International Conference on Computer Science and Information Technologies*, CSIT 2018, September 11-14 Lviv, N. Shakhovska and M. Medykovskyy (eds), Ukraine. Springer: Springer Nature Switzerland, 2019, pp. 119 – 131.
13. Munday J. A Computer-assisted approach to the Analysis of Translation Shifts, *Meta*, XLIII, 4, 1998.
14. Ol'khov's'ka A. S. Systemy avtomatyzovanoho perekladu iak skladova fakhovoho navchannia: perevahy ta nedoliky [The systems of automated translations as a component of professional translator training: advantages and disadvantages]. *Filolohichni traktaty*. 2015. Т. 7, № 2. С. 102–108.
15. Simard M. Translation analysis and translation automation. Proceedings of the Fifth International Conference on Theoretical and Methodological Issues in Machine Translation TMI 93: MT in the Next Generation, Kyoto, Japan, pp. 1993, p. 12–20.
16. Sinclair J. *Corpus, concordance, collocation*, Oxford : Oxford University Press, 1991, 137 p.

17. Slovnyk symboliv kul'tury Ukrainy [The dictionary of symbols of Ukraine] URL : https://library.udpu.edu.ua/library_files/397017.pdf (дата звернення: 11.04.2020).
18. Tartt D. Schyhol' [The Goldfinch]. Pereklad z anhlijs'koi [Translation from English: Viktor Shovkun, Kharkiv: KSD, 2016, 816 stor.
19. Tartt D. The Goldfinch, Little Brown Publishing, 2013, 771 p.
20. Tymoczko M. Computerized corpora and the future of translation studies, *Meta*, 1998, 434, pp. 652–659.
21. Ulrych M. The impact of multilingual parallel concordancing on translation. In *Practical Applications in Language Corpora*, ed. B.L. Tomaszczyk, and P.J. Melia, Lodz: Lodz University Press, 1997, pp. 421–436.
22. Zanettin F. Parallel corpora in translation studies: Issues in corpus design and analysis. In *Intercultural Faultlines. Research Models in Translation Studies I: Textual and Cognitive Aspects*, ed. M. Olohan, Manchester: St. Jerome, 2000, pp. 105–118.

Стаття надійшла до редакції 28.10.2020
The article was received 28 October 2020