

Стаття надійшла до редакції 18.05.2024 р.

Перевірено на плагіат 19.05.2024 р.

унікальність – 100%

<https://doi.org/10.17721/StudLing2024.24.101-110>

УДК 81'25:82-31:004.8

ВИКОРИСТАННЯ ТЕХНОЛОГІЇ МАШИННОГО ПЕРЕКЛАДУ ПРИ ВІДТВОРЕННІ ХУДОЖНІХ ТЕКСТІВ МАЛОЇ ПРОЗИ

Евеліна Левонівна Саргсян, evelina.sarhsian@karazin.ua

викладач

Харківський національний університет імені В. Н. Каразіна

Ольга Андріївна Зінченко, olha.zinchenko1@gmail.com

викладач

Харківський національний університет імені В. Н. Каразіна

Системи машинного перекладу мають дотик до всіх сфер людської діяльності, і задіяні навіть у перекладах художньої літератури. Науковці зазначають, що машинний переклад має великі перспективи розвитку та є найменш трудомістким у порівнянні з роботою людини-перекладача. Саме тому з'являється потреба аналізу типових помилок машинного перекладу, що сприятиме їхньому уникненню в майбутньому на подальших етапах розвитку систем нейронного машинного перекладу.

У статті розглянуто та проаналізовано якість українськомовного перекладу іспаномовного оповідання аргентинського письменника Хуліо Кортасара "Amigos", виконаного системою машинного перекладу ДіпЛ (DeepL). Отриманий текст перекладу було порівняно з офіційним українськомовним перекладом, виконаним людиною-перекладачем. Текст перекладу було досліджено з позиції наявних помилок, а також визначено типи і кількість помилок, припущених при застосуванні системи машинного перекладу ДіпЛ (DeepL). У дослідженні були використані метод аналізу, метод синтезу, метод порівняння оригіналу та перекладу. Враховуючи кількість та значущість помилок з огляду на їх вплив на сутність тексту в межах вузького та широкого контексту, ми зробили висновок про доволі високу якість перекладу оповідання. Було визначено, що застосування машинного перекладу для перекладу творів художньої літератури в межах малої прози є потенційно можливим та дієвим при обов'язковому подальшому редагуванні отриманого тексту перекладу людиною-перекладачем. З метою поліпшення якості машинного перекладу текстів

© Саргсян Е. Л., evelina.sarhsian@karazin.ua, Зінченко О. А., olha.zinchenko1@gmail.com

Використання технології машинного перекладу при відтворенні художніх текстів малої прози (Українською)

художньої літератури розроблено рекомендації по вдосконаленню роботи машинного перекладача ДіпЛ (DeepL).

Ключові слова: машинний переклад, перекладацькі помилки, текст художньої літератури, вузький контекст, широкий контекст, людина-перекладач, система машинного перекладу DeepL.

UTILIZING MACHINE TRANSLATION TECHNOLOGY FOR REPRODUCING SHORT PROSE LITERARY TEXTS

Evelina Levonivna Sarhsian, evelina.sarhsian@karazin.ua

lecturer

V. N. Karazin Kharkiv National University

Olha Andriivna Zinchenko, olha.zinchenko1@gmail.com

lecturer

V. N. Karazin Kharkiv National University

Machine translation systems permeate all spheres of human activity, including the field of literary translation. Scholars note that machine translation holds significant potential for development and is less labor-intensive compared to the work of human translators. Consequently, there is a need to analyze common errors in machine translation to help avoid them in the future as neural machine translation systems continue to evolve.

This article examines and analyzes the quality of a Ukrainian translation of the Spanish short story “Amigos” by Argentine writer Julio Cortázar, performed by the machine translation system DeepL. The translated text was compared with the official Ukrainian translation done by a human translator. The translated text was analyzed for errors, and the types of errors made by the DeepL system were identified. Additionally, the number of errors was counted. The study employed methods of analysis, synthesis, and comparison of the original and translated texts. Considering the number and significance of the errors in terms of their impact on the essence of the text within narrow and broad contexts, we concluded that the quality of the story's translation is relatively high. It was determined that using machine translation for translating literary works, particularly short prose, is potentially feasible and effective, provided that the translated text is subsequently edited by a human translator. To improve the quality of machine translation of literary texts, recommendations were developed to enhance the performance of the DeepL machine translation system.

Keywords: *machine translation, translation errors, literary text, narrow context, broad context, human translator, DeepL machine translation system.*

Вступ. Щороку, у контексті стрімкого технологічного розвитку, усе більша увага науковців фокусується на технології машинного перекладу (МП).

До того ж, практика застосування автоматизованого перекладу стає все більш поширеною і в діяльності професійних перекладачів, які нерідко вдаються до використання МП при перекладі текстів різного рівня складності та тематичного спрямування, що пояснюється значним підвищенням ефективності роботи через її полегшення та прискорення безпосередньо процесу відтворення текстів іншою мовою. Менше з тим, питання якості вихідного тексту, виконаного за допомогою МП, усе ще залишається відкритим. Зокрема чимало лінгвістів вдається до порівняльного аналізу перекладу як продукту автоматизованого перекладу та перекладу, виконаного людиною-перекладачем [English Language Teaching 2016].

Однак треба зазначити, що автоматизований переклад художніх текстів є відносно новим напрямком перекладацьких досліджень, адже протягом багатьох років переклад творів художньої літератури вважався слабким місцем машинного перекладу, оскільки для них характерний творчий підхід, варіативність стилістичних фігур та прийомів виразності; до того ж, вихідний текст повинен віддзеркалювати досвід читання тексту оригіналу [Гербероф і Торал 2020, Кенні 2022].

Поява сучасної парадигми машинного перекладу у межах штучного інтелекту, який забезпечує підвищення плавності перекладу та відносного відходу від буквализму при відтворенні тексту оригіналу, спонукає дослідників проаналізувати потенціал технології автоматизованого перекладу на сучасному етапі її розвитку в контексті здійснення саме художнього перекладу [Toral, Wieling and Way 2018]. Можливості застосування МП при відтворенні літературних творів з теоретичної точки зору все частіше розглядається в перекладацьких студіях [Abdulfattah Omar and Yasser Gomaа 2020; Antonio Toral and Andy Way 2014; Dorothy Kenny and Marion Winters 2020]. Проте, з іншого боку, бракує розвідок, присвячених практичному аспекту цього питання, зокрема аналізу якості перекладу художніх текстів з іспанської на українську, що обґрунтовує актуальність нашого дослідження.

Як зазначають Антоніо Торала та Енді Уей, «незважаючи на суттєвий прогрес у галузі автоматизованого перекладу в період останніх двох десятиліть, наразі немає доказів перспектив того, що машинні перекладачі зможуть бути корисними при перекладі художніх текстів» [Toral, Antonio 2018; Way, Andy 2018, 12-14]. І хоча поява систем нейронного машинного перекладу сприяла відносному покращенню якості перекладу та сприяла його більшій гармонійності, надмірна дослівність [Кенні 2022, 23-49], яка становить велику

проблему для відтворення текстів художньої літератури, все ще залишається суттєвою вадою машинного перекладу.

Наразі системи машинного перекладу найкраще виконують переклад недвозначних коротких текстів, що й вимагає майже буквально відтворення, на відміну від творчого перекладу [Moorkens et al. 2018; Moorkens, Joss 2018]. При перекладі літературних творів процес значно ускладнюється необхідністю комплексного підходу до нього, адже в цьому випадку суттєву роль відіграють не лише лексичні складові, але й ті елементи шару підтексту, що неможливо відтворити без детального ознайомлення та аналізу тексту оригіналу. Отже, цілком природньо, що збереження досвіду читання постає невід'ємною частиною процесу перекладу. Проте думку про те, що МП зовсім не має перспектив у галузі перекладу художніх творів, вважаємо неправильною. Тому необхідним є аналіз типових помилок МП, яких ймовірно можна буде уникати в майбутньому на подальших етапах розвитку систем нейронного машинного перекладу.

Головна мета цієї статті полягає в тому, щоб провести порівняльний аналіз перекладів іспаномовного тексту художньої літератури малої прози, виконаних людиною–професійним перекладачем, та системою автоматизованого перекладу українською мовою.

У ході роботи задля досягнення поставленої мети було виконано наступні завдання:

- аналіз якості українськомовного перекладу іспаномовного оповідання з точки зору наявних помилок;
- визначення типів наявних помилок, припущених при застосування системи машинного перекладу ДіпЛ;
- підрахунок помилок МП.

Матеріалом нашої роботи слугує оповідання відомого аргентинського письменника Хуліо Кортасара *Amigos* (укр. «Друзі», мова оригіналу – іспанська) та переклади зазначеного твору українською мовою. У ході роботи було проаналізовано наявний сертифікований україномовний переклад, виконаний професійним перекладачем Юрієм Покальчуком, та переклад, отриманий при застосуванні машинного перекладача ДіпЛ. Вибір тексту аналізу обумовлений популярністю творів зазначеного іспаномовного письменника та його невеликим обсягом. Зважаючи на те, що наразі системи МП краще працюють саме з короткими текстами, вважаємо доцільним аналіз творів малої прози та вбачаємо в перекладі невеликих за обсягом оповідань більше потенціалу для ефективного застосування МП, принаймні найближчим

часом. Вибір системи МП обумовлено його популярністю та доволі високими показниками якості при відтворенні текстів різного тематичного спрямування.

Нами були застосовані метод аналізу, метод синтезу, метод порівняння оригіналу та перекладу.

Під час роботи було встановлено, що з позиції відтворення тексту художньої літератури іншою мовою за допомогою МП, необхідним є виявлення тих типів помилок, які були припущені системою МП. Помилки при відтворенні тексту художньої літератури, у свою чергу, було встановлено у вузькому (у межах словосполучення чи речення) та широкому (у межах тексту в цілому) контексті.

Також маємо зауважити, що текст перекладу, отриманий за допомогою МП, та текст, який виконано людиною-перекладачем, було порівняно шляхом використання програми автоматичного порівняння текстів. Відповідно до отриманих даних, збіг становить 77 %, що є дуже високим показником збігу та свідчить про доволі гідний рівень якості перекладу автоматизованої системи ДіпЛ.

Нами були виокремлені лексико-граматичні, лексико-семантичні та узусні помилки у тексті перекладу, виконаного системою автоматизованого перекладу ДіпЛ.

Наведемо таблицю з деякими прикладами встановлених типів помилок.

Таблиця 1

Помилки МП

Типи помилок	МП	Переклад українською людиною-перекладачем	Оригінал
лексико-граматичні	– про каву і машину – собі, – все закінчилося б за мить	– відносно кафе й автомобіля – собі, – усе повинно скінчитися за одну мить	– del café y del auto – él mismo— todo quedaría despachado en un momento
узусні	– повільно думати – думаючи про обличчя Ромеро	– обміркувати – розмірковуючи, який вираз обличчя матиме Ромеро	– pensar despacio – pensando en la cara que pondría Romero

лексико-семантичні	– за сірою губною помадою – про каву і машину – спокійно, не втрачаючи ні секунди – Номеру Один спало на думку вбити Ромеро – о п'ятій-сьомій	– по сірому капелюсі – відносно кафе й автомобіля – повагом, але не втрачаючи жодної миті – Номер Перший наказав йому вбити Ромеро – за п'ять хвилин до сьомої	– por el chambergo gris – del café y del auto – tranquilo pero sin perder un instante – al Número Uno se le hubiera ocurrido hacer matar a Romero – a las siete menos cinco
--------------------	---	---	---

Згідно даних, наведених у Таблиці 1, у зазначених прикладах лексико-семантичні помилки виявилися найбільш суттєвими, бо вони значною мірою змінили зміст тексту оригіналу. Наприклад, іспанська лексична одиниця *chambergo*, що є певним видом капелюху, була спотворена при перекладі системою МП та відтворена як «губна помада».

Також слід відзначити, що певні помилки були викликані й багатозначністю лексичних одиниць. Наприклад, іспанське *café* контекстуально позначало «кафе», що стає зрозумілим у межах тексту (широкий контекст). Але системою МП було запропоновано інший відповідник – «кава», що порушило логічний зв'язок у межах цілого тексту.

До того ж, були наявні й помилки логічного узгодження: внаслідок опускання при перекладі сполучника *але*, у «своєму» тексті МП було створено алогічний вислів *спокійно, не втрачаючи ні секунди*, де адвербіальна одиниця *спокійно* логічно не узгоджена з дієприслівниковим зворотом *не втрачаючи ні секунди*.

Схожа проблема виникла й у наступному прикладі при втраті одиниці *hacer*, що контекстуально позначає «змусити», «наказати». Через відсутність цієї одиниці в реченні стає не зрозумілим, хто саме має вбити Ромеро.

Ще однією помилкою є неправильне відтворення позначення часу 18:55, що при перекладі МП («о п'ятій-сьомій») певним чином розширила часові межі на період з 17:00 до 19:00.

Натомість лексико-граматичні помилки виявилися найменш серйозними, адже не вплинули на розуміння словосполучення та, відповідно, і

речення в цілому (порушення вживання сполучників і/й та чергування літер у/в).

Узусні помилки також не стали перешкодою для сприйняття загального значення речення, хоча й вплинули на загальне враження від тексту. У цьому випадку, система машинного перекладу вдалася до буквализму, проте той вираз, що звучить іспанською мовою цілком природньо (*pensar despacio* – «повільно думати»), не є характерним для української мови. У свою чергу, людина-перекладач доволі вдало відтворила цей вислів дієсловом «обмірковувати», що часто застосовується в повсякденному житті носіями української мови. Також штучно звучить у наведеному контексті вираз «думати про обличчя», якому бракує іменника «вираз», і людина-перекладач скористалася прийомом перекладацької трансформації лексичного додавання, що є доречною та бажаною в межах цього словосполучення.

У рамках широкого контексту також маємо відзначити зміну, до якої вдалася система МП, і яка не становить граматичної помилки, але порушує цілісність оповідання, руйнуючи його кільцеву структуру. На початку твору автор зазначає: “*En ese entonces Romero era un tal Romero, y él un tal Beltrán*”. (укр. «Тоді Ромеро був такий собі Ромеро, а Белтран був такий собі Белтран»; укр. МП «Тоді Ромеро був таким собі Ромеро, а він таким собі Бельтраном»). У кінці письменник повертається до неї, зазначаючи: “*La última visión de Romero había sido la de un tal Beltrán*” (укр. «Останнє, що побачив Ромеро у своєму житті, був такий собі Белтран»). Цей повтор є ключовим для фінального враження та довершує цілісність оповідання. Тим не менш, у перекладі системи МП ми бачимо змінений варіант: «Останнім, кого бачив Ромеро, був якийсь Белтран», що, безумовно, порушило структуру художнього тексту.

Загальна кількість помилок вузького контексту становить 20 одиниць (15 % – лексико-граматичні, 20 % – узусні, 65 % – лексико-семантичні). Загальна кількість помилок широкого контексту становить 2 одиниці. Також маємо зазначити невеликий обсяг тексту (оригінал – 628 слів, офіційний переклад українською – 563 слова, МП – 524 слова).

Проте отриманий показник не є поганим, і дає підстави вважати, що застосування МП для перекладу творів художньої літератури, принаймні в межах малої прози, є потенційно можливим та дієвим при обов’язковому подальшому редагуванні отриманого тексту перекладу людиною-перекладачем.

Маємо також зазначити, що загальне враження від прочитання отриманого тексту перекладу за допомогою системи МП відрізняється від

враження, отриманого при читанні офіційного перекладу. Людина-перекладач вдало відтворила стиль автора, на що поки не здатна автоматизована система перекладу. І попри правильність відтворення певних одиниць у лексико-граматичному чи лексико-семантичному відношенні, іноді втрачається експресивність художнього твору, що є його основним достоїнством (наприклад: офіційний переклад – «в ньому закипає злість», МП – «це його злило»).

У ході аналізу ми також зіткнулися й з протилежною ситуацією, де власне система МП впоралася дещо краще за людину-перекладача. У тексті оригіналу наведені дві реальні локації: “*salió del café de Corrientes y Libertad*” та “*en el café de Cochabamba y Piedras*”. І хоча автор не надає слово «вулиця» у реченнях, з граматичної структури (через прийменник *de* – показник приналежності) витікає те, що автор указує на кафе, що знаходяться на розі двох вулиць. Система МП відтворила українською ці вирази відповідно: «вийшов з кафе на розі Коррієнтес і Лібертад» та «в кав'ярні на розі Кочабамба і П'єдрас». Натомість, людина-перекладач узагальнила назви вулиць, перетворивши їх на назви кафе: «вийшов з кафе “Коррієнтес і Лібертад”» та «в кафе “Кочабамба і П'єдрас”». Для українськомовного читача така неточність не є суттєвою та не дуже впливає на сприйняття, але тут ми бачимо неадекватне відтворення оригіналу. Автор зазначає як локацію кав'ярні чотири вулиці, що дійсно існують та знаходяться в Аргентині, у місті Буенос-Айрес, та дійсно відповідно перетинаються.

Застосування системи машинного перекладу для текстів художньої літератури малої прози може виявитися відносно корисним при подальшому розробленні нейронних систем перекладу, адже вже зараз ця система демонструє достатній рівень якості перекладу, проте, звісно, все ще не може повністю замінити роботу людини-перекладача. Вважаємо, що найближчим часом доречним є розгляд потенційного використання МП саме в межах перекладу коротких оповідань, адже помилки широкого контексту, які все ще наявні при автоматизованому перекладі, можуть значно ускладнити процес редагування людиною-перекладачем тексту, отриманого після застосування МП.

Перспективою дослідження є залучення більш репрезентативного емпіричного матеріалу та проведення порівняльного аналізу текстів перекладів, виконаних різними системами МП, з їхнім подальшим зіставленням з офіційними перекладами, виконаними людиною-перекладачем.

Література:

1. Кортасар, Х. «Темна зброя. Збірка». *Дніпро* (1983): 21-23.
2. Abdulfattah, O.; Yasser, G. «The Machine Translation of Literature: Implications for Translation Pedagogy». *In International Journal of Emerging Technologies in Learning vol. 15(11) (2020): 228-235.* URL: <https://online-journals.org/index.php/i-jet/article/view/13275/715>.
3. Benjamins, J. «Creativity in translation». *John Benjamins Publishing Company* (2022): 184-212. URL: <https://www.jbe-platform.com/>.
4. Kenny, D.; Winters, M. «Machine translation, ethics and the literary translator's voice». *In Translation Spaces, vol. 9(1) (2020): 123-149.* URL: https://pure.hw.ac.uk/ws/portalfiles/portal/41374556/Kenny_Winters_Translation_Spaces_Article_accepted_24_April.pdf.
5. Kenny, D. «Human and machine translation. Machine translation for everyone: Empowering users in the age of artificial intelligence». *Berlin: Language Science Press* (2022): 23-49.
6. Macan J.; Mitkov R., Olaf-Michael, S. «Translating and The Computer», *In J. Esteves-Ferreira vol. 36* (2014). URL: <https://aclanthology.org/2014.tc-1.23.pdf>.
7. Moorkens, J. «What to expect from Neural Machine Translation: a practical in class translation evaluation exercise». *The Interpreter and Translator Trainer, v. 12, n. 4,* (2018): 375–387. URL: <https://doi.org/10.1080/1750399X.2018.1501639>.
8. Toral, A.; Way, A. 2014. «Is machine translation ready for literature?» *In J. Esteves-Ferreira* (2014). URL: <https://aclanthology.org/2014.tc-1.23.pdf>.
9. Toral, W.; Way A., Toral, A.; Wieling, M. «Post-editing Effort of a Novel with Statistical and Neural Machine Translation». *Frontiers in Digital Humanities v. 5* (2018): 1–11. URL: <https://doi.org/10.3389/fdigh.2018.00009>.

Джерела ілюстративного матеріалу:

10. <https://cdn.aaai.org/ocs/7864/7864-36722-1-PB.pdf>
11. ERIC - EJ1089886 - A Comparative Study of "Google Translate" Translations: An Error Analysis of English-to-Persian and Persian-to-English Translations, English Language Teaching, 2016 (ed.gov)
12. Automatic Translations Versus Human Translations in Nowadays World - ScienceDirect
13. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781315648934-18/issues-human-automatic-translation-quality-assessment-stephen-doherty>
14. Microsoft Word - V2S1 English papers (researchgate.net) [2105.04616] Automatic Classification of Human Translation and Machine Translation: A Study from the Perspective of Lexical Diversity (arxiv.org)
15. <https://narrativabreve.com/2014/09/cuento-de-julio-cortazar-los-amigos.html>
16. <https://bul.in.ua/k/kortasar-hulio/druzi-2/>
17. <https://www.deepl.com/translator>
18. <https://gotranscript.com/text-compare#diff>

References:

1. Cortázar, J. «Tayemna zbroya. Zbirka [Secret weapon. Collection]». Dnipro (1983): 21-23.
2. Abdulfattah, O.; Yasser, G. «The Machine Translation of Literature: Implications for Translation Pedagogy». In *International Journal of Emerging Technologies in Learning* vol. 15(11) (2020): 228-235. URL: <https://online-journals.org/index.php/ijet/article/view/13275/715>.
3. Benjamins, J. «Creativity in translation». *John Benjamins Publishing Company* (2022): 184-212. URL: <https://www.jbe-platform.com/>.
4. Kenny, D.; Winters, M. «Machine translation, ethics and the literary translator's voice». In *Translation Spaces*, vol. 9(1) (2020): 123-149. URL: https://pure.hw.ac.uk/ws/portalfiles/portal/41374556/Kenny_Winters_Translation_Spaces_Article_accepted_24_April.pdf.
5. Kenny, D. «Human and machine translation. Machine translation for everyone: Empowering users in the age of artificial intelligence». *Berlin: Language Science Press* (2022): 23-49.
6. Macan J.; Mitkov R., Olaf-Michael, S. «Translating and The Computer», In *J. Esteves-Ferreira* vol. 36 (2014). URL: <https://aclanthology.org/2014.tc-1.23.pdf>.
7. Moorkens, J. «What to expect from Neural Machine Translation: a practical in class translation evaluation exercise». *The Interpreter and Translator Trainer*, v. 12, n. 4, (2018): 375–387. URL: <https://doi.org/10.1080/1750399X.2018.1501639>.
8. Toral, A.; Way, A. 2014. «Is machine translation ready for literature?» In *J. Esteves-Ferreira* (2014). URL: <https://aclanthology.org/2014.tc-1.23.pdf>.
9. oral, W.; Way A., Toral, A.; Wieling, M. «Post-editing Effort of a Novel with Statistical and Neural Machine Translation». *Frontiers in Digital Humanities* v. 5 (2018): 1–11. URL: <https://doi.org/10.3389/fdigh.2018.00009>.
10. <https://cdn.aaai.org/ocs/7864/7864-36722-1-PB.pdf>
11. .ERIC - EJ1089886 - A Comparative Study of "Google Translate" Translations: An Error Analysis of English-to-Persian and Persian-to-English Translations, English Language Teaching, 2016 (ed.gov)
12. Automatic Translations Versus Human Translations in Nowadays World - ScienceDirect
13. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781315648934-18/issues-human-automatic-translation-quality-assessment-stephen-doherty>
14. Microsoft Word - V2S1 English papers (researchgate.net) [2105.04616] Automatic Classification of Human Translation and Machine Translation: A Study from the Perspective of Lexical Diversity (arxiv.org)
15. <https://narrativabreve.com/2014/09/cuento-de-julio-cortazar-los-amigos.html>
16. <https://bul.in.ua/k/kortasar-hulio/druzi-2/>
17. <https://www.deepl.com/translator>
18. <https://gotranscript.com/text-compare#diff>