

УДК 81'33

Д 79

Дубова О. А.

доктор філологічних наук, професор кафедри прикладної лінгвістики,
Навчально-науковий гуманітарний інститут Національного університету
кораблебудування імені адмірала Макарова (Україна, Миколаїв),
helen.evolucia@gmail.com

Юрченко І. В.

студентка спеціальності Філологія (Прикладна лінгвістика),
Навчально науковий гуманітарний інститут Національного університету
кораблебудування імені адмірала Макарова (Україна, Миколаїв),
uira99@gmail.com

ВИКОРИСТАННЯ І КОРИГУВАННЯ МАШИННОГО ПЕРЕКЛАДУ

Стаття присвячена розглядові шляхів ефективного виправлення найбільш частотних помилок, що здійснюються сервісом машинного перекладу Google Translate на матеріалі мовної пари «англійська→українська». Висвітлено зміст поняття «машинний переклад». Пояснено доцільність застосування систем машинного перекладу під час роботи з текстами різних стилів. Описано алгоритм постредагування отриманого перекладу користувачами систем машинного перекладу.

Ключові слова: машинний переклад, системи машинного перекладу, Google Translate, редагування.

Постановка проблеми. Проблема поліпшення швидкості та якості перекладу є особливо актуальною з огляду на постійне зростання інформаційних потоків. Однією з можливостей економії часових та матеріальних ресурсів і подолання мовного бар'єру в спілкуванні є застосування систем машинного перекладу (далі – СМП) як засобів підвищення продуктивності перекладацької діяльності. Однак, незважаючи на простоту у використанні і широкі функціональні можливості, машинний переклад (далі – МП) все одно вимагає постредагування, його якості і адекватності часто залишаються незадовільними.

Аналіз останніх досліджень і публікацій. СМП знаходяться в полі зору лінгвістики вже декілька десятиліть. Теоретичне обґрунтування поняття «машинного перекладу», історія його розвитку подано в працях [1, 2, 3, 12, 15]. Багато українських та зарубіжних науковців досліджували механізми МП з подальшою розробкою алгоритмів та інструментів автоматичних оцінок якості МП [4, 6, 16, 17]. Проблемі застосування МП присвячено також праці таких дослідників [9, 10, 13]. Однак дослідження, присвячені комплексному дослідженню механізмів МП, аналізу типових помилок, яких припускається онлайн-перекладач Google на базі мовної пари «англійська→українська» та визначенні шляхів їх усунення й редагування отриманого перекладу користувачами СМП, відсутні.

Мета нашого дослідження полягає у розгляді шляхів ефективного виправлення найбільш частотних помилок, що здійснюються сервісом МП Google Translate на матеріалі мовної пари «англійська→українська».

Виклад основного матеріалу дослідження. Для здійснення нашого дослідження передусім необхідно визначити зміст поняття машинний переклад (*machine translation, MT*).

Дослідники у галузі МП досі розходяться в поглядах щодо визначення поняття машинний переклад [5, 7, 8, 11, 14]. Однак вважаємо найбільш оптимальним визначення машинного перекладу як такого виду перекладу, що виконується на комп'ютері шляхом перетворення тексту з однієї природної мови в еквівалентний за змістом текст на іншій мові, а також результат такої дії.

Підкреслимо, що вимоги до аналізу перекладу, який виконує машина не повинні бути такими ж, як і до перекладу, що здійснюється людиною. Переклад, виконаний СМП, має велику кількість помилок на різних мовних рівнях. Він потребує редагування і коригування людиною, яка володіє базовими знаннями з мови перекладу та мови оригіналу і орієнтується в предметній галузі, до якої належить інформація, що перекладається. СМП переважно дозволяють лише зрозуміти суть документа, який перекладається [12].

Розглянемо детальніше деякі фактори, що впливають на якість МП у межах англійсько-української мовної пари. Відповідно до генеалогічної класифікації мов, англійська мова належить до германської гілки індоєвропейської родини мов, а українська – до слов'янської. Відповідно до морфологічної класифікації, ці мови належать і до різних структурних типів. Українська мова за своєю граматичною структурою належить до синтетичних мов, англійська – до аналітичних. До синтетичних способів вираження граматичних значень в українській мові належать афіксація, фонетичні зміни і суплетивізм. Аналітичним способом вираження граматичних значень в англійській мові є аналітичні формативи («службові слова»). Англійській мові властиві численні вияви аморфізму, в українській мові, навпаки, вияви аморфізму незначні. Саме розбіжності в будові мов, відмінності у складі

морфологічних категорій та побудові речень і становлять першочергові граматичні труднощі під час МП.

Основними недоліками СМП слід вважати відсутність точності (використання систем МП все ще не дозволяє якісно перекласти тексти на основі літературної природної мови) та, як наслідок, обмеженість використання (існують певні типи документів і стилі перекладу, які недоцільно перекладати за допомогою електронних ресурсів МП).

Окремі труднощі МП пов'язані з необхідністю визначення прийомів перекладу безеквівалент-

них форм, ідіом і словосполучень, назв, власних імен, аббревіатур, умовних позначок, усунення омонімії на різних рівнях, а також необхідністю залучення у процес перекладу екстралінгвістичних знань.

У своєму дослідженні ми перевірили загальнодоступну та найбільш популярну серед інших он-лайн систем МП (Yandex та Pragma), – систему Google Translate на наявність граматичних та лексичних помилок при перекладі з англійської на українську мову та на наявність вищезазначених недоліків (див. табл. 1).

Таблиця 1. – Аналіз перекладу Google Translate

Оригінал	Переклад Google Translate	Постредагування
1. Head of IUPAC Liza Hope said...	Керівник IUPAC Ліза Хоуп сказала...	Голова ІЮПАК (Міжнародного союзу теоретичної і прикладної хімії) Лайза Гоуп сказала ...
2. Variable geometry wings constituted the main feature of the plane.	Крила змінної геометрії становили головну особливість площини.	Головною особливістю цього літака були крила змінної конфігурації.
3. The signal output is shown to be directly coupled to the transformer input.	Показано, що вихідний сигнал безпосередньо з'єднано з входом трансформатора.	Показано, що вихід сигналу безпосередньо з'єднано з входом трансформатора.
4. The book earned him his M.I.T. Ph.D. in 1959.	Книга принесла йому ступінь M.I.T. Доктор філософії у 1959 році.	Завдяки цій книзі він отримав ступінь доктора у Массачусетському технологічному університеті в 1959 році.

Це дослідження було здійснено за допомогою таких лінгвістичних прийомів, як прийом зіставного аналізу текстів оригіналу та перекладів українською мовою.

Результати дослідження свідчать про те, що машинний переклад, перш за все, відтворює формальну структуру оригіналу, забезпечуючи дослівний переклад, що полягає в стилізації за законами мови оригіналу, відтворенні мовних форм мови оригіналу та вираженні стилістичних структур мови оригіналу і мови перекладу аналогічними засобами, що у свою чергу зумовлює спотворення, неточності, неясності в тексті перекладу.

З огляду на вищевикладене ми здійснили спробу розробити й описати алгоритм постредагування, який потенційні користувачі СМП та студенти можуть застосовувати для роботи з текстами, що є результатом МП «англійська→українська». Цей алгоритм включає такі етапи:

1. лексико-семантична правка машинного перекладу:

- 1.1. редагування багатозначних слів;
- 1.2. редагування омонімів;
- 1.3. редагування аббревіатур;

2. граматична правка:

- 2.1. морфологічне редагування:
 - 2.1.1. редагування перекладу пасивних конструкцій;
 - 2.1.2. застосування граматичної трансформації додання при перекладі часових форм;
 - 2.1.3. перевірка правильності перекладу артикль;

2.2. синтаксичне редагування:

- 2.2.1. перевірка правильності вживання прийменників;

2.2.2. перевірка правильності керування дієслів іменниками;

2.2.3. перевірка правильності узгодження в роді, числі і відмінку підпорядкованих прикметників з головним іменником;

3. стилістична правка:

3.1. редагування перекладу текстів науково-технічного стилю;

3.2. редагування перекладу текстів публіцистичного стилю;

3.3. редагування перекладу текстів художнього стилю;

4. пунктуаційна правка:

4.1. перевірка вживання коми у порівнюваних мовах;

4.2. перевірка вживання тире у порівнюваних мовах;

4.3. перевірка вживання лапок у порівнюваних мовах.

Зважаючи на той факт, що СМП найчастіше припускаються помилок у виборі правильного еквівалента перекладному слову або словосполученню, що призводить до спотворення або навіть втрати загального змісту перекладного тексту, під час редагування МП у першу чергу необхідно здійснити лексико-семантичну правку результатів перекладу.

Машинні перекладачі нездатні розуміти контекст, щоб правильно перекладати полісемантичні слова. Тому, на нашу думку, людині-перекладачу варто починати коригування тексту МП з редагування помилок, зумовлених багатозначністю слів. Перекладач, редагуючи текст, має виявити слова, вжиті з такими порушеннями, і семантичні невідповідності, а далі працювати з ними, застосовуючи

тлумачні словники англійської та української мов, перекладні англо-українські словники, словники синонімів української мови. СМП також нездатні розрізняти контекстуальні значення або стилістичні відтінки омонімів, правильно обирати значення омонімічного терміну з суміжної галузі науки або техніки. Тому після редагування перекладу багатозначних слів перекладачу варто перейти до редагування омонімів.

Через розбіжності в граматичних структурах мови оригіналу і мови перекладу другими за частотністю помилками для програм МП є різноманітні граматичні помилки, які можуть серйозно ускладнювати розуміння тексту перекладу.

Першим етапом граматичної правки МП вбачаємо редагування перекладу пасивних конструкцій. Зважаючи на той факт, що граматичні категорії виду та часу в порівнюваних нами мовах мають різний морфологічний зміст, СМП часто неадекватно чи зовсім неправильно перекладають часові форми. У зв'язку з цим при англо-українському перекладі виникає необхідність заміни граматичного часу. Іншою проблемою СМП, що потребує редагування, є неправильний переклад артиклів, які є засобом вираження категорії означеності / неозначеності англійської мови.

Необхідність синтаксичного редагування машинного перекладу зумовлена тим, що більшість англійських та українських речень мають різну синтаксичну структуру відмінності в порядку розташування складників простих і складних речень, а тому при перекладі необхідно застосовувати граматичні трансформації для адекватної передачі смислу оригіналу з урахуванням норм мови перекладу.

Третій етап редагування машинного перекладу передбачає усунення неточностей в передачі стилістичного забарвлення. У результаті дослідження з'ясовано, що задовільної якості можна очікувати від МП формалізованих технічних текстів, тоді як літературний переклад виходить за рамки МП. Сучасні СМП не можуть забезпечити семантико-синтаксичну адекватність перекладу текстів художнього та публіцистичного стилів: спотворюється зміст оригіналу, і отриманий переклад не задовольняє повною мірою вимоги користувача щодо отримання релевантної інформації. Останнім етапом редагування машинного перекладу є пунктуаційна правка.

Висновок із цього дослідження і перспективи подальших розвідок. Механізми МП онлайн-перекладача Google та аналіз типових помилок, яких припускається цей перекладач на базі мовної пари «англійська→українська», є недостатньо дослідженими. Незважаючи на заявлене розробниками постійне вдосконалення систем МП, зокрема сервісу Google Translate, МП слугують лише допоміжним ресурсом і не можуть повністю замінити особу фахового перекладача.

У результаті дослідження було виявлено необхідність втручання людини в безпосередню

роботу СМП з огляду на те, що ці системи наразі не вміють точно розпізнавати частини мови, трансформувати складні граматичні конструкції, співвідносити семантику слова з широким контекстом, враховувати культурну специфіку оригіналу, адекватно передавати стилістичні засоби.

Перший досвід дослідження шляхів алгоритмізації постредагування МП Google Translate на матеріалі мовної пари «англійська→українська» показав, що розроблена послідовність дій може допомогти користувачам сервісів МП досягти семантико-синтаксичної адекватності під час перекладу текстів художнього, публіцистичного та наукового стилів. Розроблений алгоритм може бути використаний під час удосконалення програм МП. В подальшому варто розробити детальний алгоритм для текстів кожного конкретного стилю.

Література

1. Андреева А. Д., Меньшиков И. Л., Мокрушин А. А. Обзор систем машинного перевода. *Молодой ученый*. Казань, 2013. №12 (59). С. 64–66.
2. Анохін В. М. Вимоги до сучасних електронних навчальних видань і можливості їх реалізації у середовищі Adobe Captivate. *Вісн. нац. ун-ту «Львів. політехніка»*. Серія: Інформатизація вищого навчального закладу. 2012. № 731. С. 71–76.
3. Бірюков А. В. Система оцінки якості автоматизованого перекладу. *Вісн. Сумського держ. ун-ту*. Серія: Філологічні науки. 2004. № 3 (62). С. 42–48.
4. Воронович В. В. Машинный перевод: конспект лекций для студентов 5-го курса специальности «Современные иностранные языки». Минск: Белорусский государственный университет, 2013. 39 с.
5. Гудманян А. Г., Сітко А. В., Струк І. В. Функціонально-прагматична адекватність машинного перекладу публіцистичних текстів. *Науковий журнал Львівського державного університету безпеки життєдіяльності «Львівський філологічний часопис»*: зб. наук. пр. 2019. № 5. С. 19.
6. Денисова С. П. Стан та перспективи перекладознавчих досліджень. *Вісн. Київського нац. лінгв. ун-ту*. Серія: Філологія. 2014. № 1. С. 146–160.
7. Кісіль М. В. До проблеми машинного перекладу. *Актуальні проблеми соціальної комунікації*: матеріали II Всеукр. наук.-практ. конф., 18-23 трав. 2011 р., Івано-Франківськ, Україна. Івано-Франківськ: ІФНГУНГ, 2011. С. 97–103.
8. Колесник А. С. Огляд систем машинного перекладу. *Інтелектуальні системи та прикладна лінгвістика*: матеріали II Всеукр. наук.-практ. конф., 28 бер. 2013 р., Харків. Харків: НТУ «ХПІ», 2013. С. 41–44.

9. Ляндебурська К. І., Гудманян А. Г. Якість машинного перекладу усних та письмових текстів. *Фаховий та художній переклад: теорія, методологія, практика*: збірник наукових праць / за заг. ред. А. Г. Гудманяна, С. І. Сидоренка. Київ: Аграр Медіа Груп, 2015. С. 72–77.
10. Міщенко А. Л. Машинний переклад у контексті сучасного науковотехнічного перекладу. *Вісн. Харківського нац. ун-ту ім. В. Н. Каразіна. Серія: Романо-германська філологія. Методика викладання іноземних мов*. Харків, 2013. № 1051. С. 172.
11. Нелюбин Л. Л. Толковий переводоведческий словарь. Москва: ФЛИНТА, 2016. URL: <https://rucont.ru/efd/244470> (дата обращения 12.01.21).
12. Стахмич Ю. С. Адекватність та еквівалентність перекладу в контексті комп'ютерної лінгвістики. *Вісн. Житомирського держ. ун-ту ім. Івана Франка*: наук. журн. / відп. ред. Н. А. Сейко. Житомир: ЖДУ ім. І. Франка, 2013. № 66. С. 224–225.
13. Тарануха В. Ю. Интеллектуальная обработка текстов. Київ: Нац. авіац. ун-т, 2014. 80 с.
14. Щипицина Л. Ю. Информационные технологии в лингвистике: учеб. пособие. Москва: ФЛИНТА: Наука, 2013. 128 с.
15. Янковець А. В. Особливості застосування сучасних систем машинного перекладу в умовах підготовки майбутніх перекладачів у вищому військовому навчальному закладі. *Вісн. Національної академії Державної прикордонної служби України*. 2012. Вип. 5. С. 180.
16. Chesterman A. The Name and Nature of Translator Studies. *Hermes: Journal of Language and Communication Studies*, 2018. № 13. 64 p.
17. Snover M. A Study of Translation Edit Rate with Targeted Human. *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas*. Cambridge, 2016. 230 p.

Dubova O. A., Doctor of Philology, Professor of the Department of Applied Linguistics, Educational and Scientific Humanitarian Institute of the Admiral Makarov National University of Shipbuilding (Ukraine, Mykolaiv), helen.evolucia@gmail.com

Yurchenko I. V., student specialty Philology (Applied linguistics) of Applied Linguistics, Educational and Scientific Humanitarian Institute of the Admiral Makarov National University of Shipbuilding (Ukraine, Mykolaiv), uira99@gmail.com

THE USE AND CORRECTION OF MACHINE TRANSLATION

The research paper is aimed at defining ways to effectively correct the most frequent errors made by the machine translation service Google Translate on the material of the language pair "English→Ukrainian". We highlighted the content of the concept of "machine translation". We explained the expediency of using machine translation systems when translating texts of different styles. We also developed an algorithm for post-editing the received translation for users of machine translation systems.

Key words: machine translation, machine translation systems, Google Translate, editing.