УДК 811.161.1:811.161.2]'255.4
DOI https://doi.org/10.32841/2409-1154.2019.43.4.12

Гайдук Н. А.,
кандидат філологічних наук,
доцент кафедри слов'янської філології та перекладу
Маріупольського державного університету

Кокташ М. Г.,
студентка ОС «Магістр» спеціальності «Філологія. Переклад (українська, російська, польська)»
Маріупольського державного університету

ТИПОВІ ПОМИЛКИ ПІД ЧАС ВИКОРИСТАННЯ СИСТЕМИ МАШИННОГО ПЕРЕКЛАДУ "GOOGLE TRANSLATE"

Анотація. У статті проаналізовано питання використання машинного перекладу у професійній перекладацькій діяльності, що нині цікавить як науковців-лінгвістів, так і виробників програмного забезпечення, оскільки натепер машинний переклад – це цілий спектр прогресивних інформаційних технологій, що швидко розвиваються і мають широке поле застосування. Саме в умовах бурхливого розвитку науково-технічного й інформаційного прогресу особливо актуальний пошук шляхів надання максимальної допомоги професійним перекладачам. Повністю замінити людину-перекладача програмами-перекладачами нині неможливо, проте виявляється можливим хоча б максимально автоматизувати процес перекладу, що здійснюється людиною, щоб полегшити нелегку працю людини-перекладача, зробити її максимально ефективною. Здійснити таке можна лише за максимальної інтеграції зусиль спеціалістів у галузях кібернетики, програмування, психології та лінгвістики, це допоможе створити таку систему машинного перекладу, яка б дозволила отримувати адекватний переклад тексту без втручання людей-перекладачів. Досягнення цієї мети в майбутньому неможливе без ретельного аналізу причин появи типових помилок, що виникають у разі машинного перекладу. Автори статті аналізують типові помилки, що виникають під час машинного перекладу, за допомогою СМП "Google Translate", групуючи їх на лексичні (зокрема й лексикографічні), граматичні, синтаксичні помилки, та помилки, що виникають внаслідок відсутності словоформи у словнику СМП, на матеріалі перекладів російською мовою уривків зі стрічок новин сучасних українських інформаційних агентств, виконаних за допомогою СМП "Google Translate", що вивчаються, досліджуються й оцінюються з урахуванням інформації щодо алгоритмів роботи систем машинного перекладу, їхніх граматик, словників та внутрішньої побудови.

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

Постановка проблеми. XXI століття характеризується бурхливим розвитком інформаційних технологій і неухильним ростом обсягу інформації, що споживається, саме тому перекладацька діяльність на ринку праці стає все більш затребуваною. Як наслідок технічного прогресу – не лише збільшення кількості перекладачів, а й підвищення якості їхньої освіти, наявність навичок роботи із засобами технічного прогресу, зокрема системами машинного перекладу (далі – СМП). «Прогресивний розвиток інформаційних і комп'ютерних технологій та безперервне розширення меж інформаційного простору відкривають перед майбутніми фахівцями безмежний океан можливостей, про які людство навіть і не мріяло сімдесят або лише двадцять років тому» [9, с. 233].

Питання використання машинного перекладу у професійній перекладацькій діяльності нині цікавлять як науковців-лінгвістів, так і виробників програмного забезпечення, оскільки натепер машинний переклад – це цілий спектр прогресивних інформаційних технологій, що швидко розвиваються і мають широке поле застосування. «Із збільшенням об'ємів інформації і вдосконаленням технологій її машинної обробки ростуть запити щодо швидкості всіх видів перекладу, зокрема і перекладу художніх творів. У цьому контексті робота над створенням системи машинного перекладу, що могла б ефективно виконувати такі функції, є актуальним завданням прикладної лінгвістики» [7, с. 69].

Саме в умовах бурхливого розвитку науково-технічного й інформаційного прогресу особливо актуальний пошук шляхів надання максимальної допомоги професійним перекладачам. Повністю замінити людину-перекладача програмами-перекладачами нині неможливо, особливо це стосується художнього перекладу. «Переклад художніх творів вимагає від перекладача не тільки знання мови оригіналу і навичок перекладу, а також глибокого рівня фонових знань, уміння працювати зі стилем і абстрактного мислення. На сучасному етапі розвитку науки і техніки всі ці вимоги важко перенести на комп'ютерну програму або вебзастосунок, що виконує функції системи машинного перекладу. У результаті, машинний переклад у цій сфері не застосовується загалом або застосовується тільки як допоміжний засіб, водночас основні функції виконує людина-перекладач» [7, с. 69]. Проте виявляється можливим хоча б максимально автоматизувати процес перекладу, що здійснюється людиною, щоб, з одного боку, максимально полегшити нелегку працю людини-перекладача, а з іншого – зробити цю працю максимально ефективною. Здійснити це можна лише за максимального інтегрування зусиль спеціалістів у галузях кібернетики, програмування, психології та лінгвістики, це, можливо, дозволить створити таку систему машинного перекладу, яка б дозволила отримувати адекватний переклад тексту без втручання людей-перекладачів. Досягнення цієї мети в майбутньому неможливе без ретельного аналізу причин появи типових помилок, що виникають під час машинного перекладу, отже, тему дослідження можна вважати актуальною.

Об'єктом дослідження є переклади російською й українською мовами уривків зі стрічок новин сучасних українських

інформаційних агентств, виконані за допомогою СМП "Google Translate", що вивчаються, досліджуються й оцінюються з урахуванням інформації щодо алгоритмів роботи систем машинного перекладу, їхніх граматик, словників та внутрішньої побудови.

Предметом дослідження є типові помилки, що виникають під час автоматичного перекладу за допомогою СМП до втручання у переклад професійного перекладача.

Мета статті полягає в систематизації типових помилок, що виникають під час автоматичного машинного перекладу, а також в узагальненні проблем, з якими стикаються системи машинного перекладу, для досягнення функціонально-прагматичної адекватності.

Аналіз останніх досліджень і публікацій. Уперше думку про можливості машинного перекладу висловив Чарльз Беббідж, який розробив проєкт цифрової аналітичної машини – механічного прототипу електронних цифрових обчислювальних машин, що з'явилися через сто років [1]. У XVIII та XIX століттях було розроблено багато універсальних мов, зокрема й есперанто. Відомо, що ідея універсальної мови для машинного перекладу ніяк не проявилася на початкових етапах розвитку цієї технології. Замість неї розглядалися тільки пари мов. Однак протягом 1950–1960-х років XX століття дослідники в Кембриджі, на чолі з Маргарет Мастерман, у Ленінграді, на чолі з Миколою Андрєєвим, у Мілані, на чолі із Сільвіо Чеккато почали роботу в цій галузі [1].

Наше дослідження спирається на праці з машинного перекладу М. Блехмана [2; 3], В. Войнова [3], Ю. Марчука [5], Г. Ситар [6], Ф. Бацевича [1], М. Періжняка [7], А. Янківця [9], роботи з теорії перекладу В. Коміссарова [4], О. Швейцера [8].

Виклад основного матеріалу. Класифікація, систематизація та вивчення помилок СМП виникли із практичних потреб редагування перекладу і вдосконалення СМП. Помилки СМП можна згрупувати в кілька типів. Покажемо це на прикладах перекладів російською мовою уривків зі стрічок новин сучасних українських інформаційних агентств, виконаних за допомогою СМП "Google Translate". За основу класифікації помилок, що виникають під час перекладу за допомогою СМП, взяли класифікацію Г. Ситар [6].

Отже, розглянемо типові помилки, що виникають під час перекладу за допомогою СМП "Google Translate":

1. Досить поширеним явищем у текстах, перекладених за допомогою СМП "Google Translate", є наявність неперекладених слів у тексті перекладу, коли програма просто не може знайти еквівалентний відповідник у мові перекладу.

Основних причин для цього кілька:

1.1. Відсутність слова або форми слова в електронному словнику (неповнота словника), здебільшого це стосується неологізмів або ж власних назв, цю проблему можна вирішити додаванням потрібних слів до словника користувача), наприклад:

Текст оригіналу (далі – ТО):

«19 листопада 2018-го на 97-му році помер професор Ілько Корунець. Його називають фундатором сучасної української школи перекладознавства» [12].

Варіант перекладу за допомогою СМП "Google Translate" (машинний переклад (далі – МП)): *«19 ноября 2018-го на 97-м году умер профессор Илья Корунець. Его называют основателем современной украинской школы перекладознавства».*

Запропонуємо відредагований варіант перекладу даного прикладу (текст перекладу (далі – ТП)):

«19 ноября 2018-го на 97-м году умер профессор Илья Корунец. Его называют основателем современной украинской школы переводоведения».

У даному прикладі бачимо дві помилки: неперекладеними залишилися прізвище *«Корунець»* і термін *«перекладознавство».*

1.2. Неправильне написання слова (в аналізованих нами стрічках новин помилок не знаходили). Теоретично перед застосуванням СМП людина-перекладач повинна ретельно перевірити текст, що підлягає перекладу, для запобігання виникненню помилок такого характеру. Це стосується і наступного виду помилок СМП.

1.3. Неправильне оформлення переносу слів (перенос, поставлений вручну, насправді є дефісом, у комп'ютерному словнику немає таких слів), наприклад: «*пояс-нення», «вихова-ність»* тощо. В аналізованих нами стрічках новин виставлених вручну переносів також не знаходили.

Варто наголосити, що помилкою саме СМП є лише помилка, що ми класифікували в пункті 1.1, а помилки, зазначені в пунктах 1.2 та 1.3, – це помилки не програми, а людини, яка припустилася помилок у тексті оригіналу, або ж помилки людини-перекладача, що готувала текст до перекладу і підготувала його на неналежному рівні. Саме тому не зупиняємося на цих двох видах помилок.

2. Лексичні (зокрема, й лексикографічні) помилки:

2.1. Вибір словникового еквівалента (значення слова), що не відповідає контексту (неврахування багатозначності слова), наприклад:

ТО: *«Одного дня їм цілком випадково вдається відбити масовану ворожу атаку»* [10].

МП: *«Однажды им совершенно случайно удается отбить массированную враждебную атаку».*

Неправильно обраний еквівалент слова *«ворожий».* Обрано значення *«враждебный»,* тоді як правильний варіант – *«вражеский».*

ТП: *«Однажды им совершенно случайно удается отбить массированную вражескую атаку».*

2.2. Неправильний вибір словникового еквівалента внаслідок орфографічної помилки (за умови збігу неправильного написання з іншим словом), наприклад: *«сумма» (рос.) та «сума» (рос.) – укр. «сума», або «торба».* Як уже було зазначено вище, у стрічках новин помилок не знаходимо, але теоретично перекладач повинен враховувати таку ситуацію.

2.3. Неврахування омонімії, наприклад:

ТО: *«Наші країни й народи – це не просто сусіди. Ми брати й сестри. І про це ми ніколи не повинні забувати», – заявив глава держави»* [11].

МП: *«Наши страны и народы – это не просто соседи. Мы брать и сестры. И об этом мы никогда не должны забывать», – заявил глава государства».*

ТП: *«Наши страны и народы – это не просто соседи. Мы братья и сестры. И об этом мы никогда не должны забывать», – заявил глава государства».*

У даному прикладі маємо справу з омонімією *«брати»* – дієслово й іменник множини.

2.4. Неврахування загального / спеціального характеру слова в певному контексті (тобто загальновживане це слово чи

термін), що може пояснюватися тим, що у СМП словник націлений на переклад текстів загальної, а не спеціальної тематики, досить часто СМП обирає загальну лексику замість медичної, хімічної тощо, наприклад:

ТО: *«Унаслідок інциденту жінка отримала забій м'яких тканин та струс мозку»* [12].

МП: *«В результате инцидента женщина получила забой мягких тканей и сотрясение мозга».*

ТП: *«В результате инцидента женщина получила ушиб мягких тканей и сотрясение мозга».*

СМП сприйняла медичний термін *«забій»* у значенні *«поверхня корисної копалини (гірської породи), яка виймається під час видобувних робіт або при проведенні гірничої виробки».*

2.5. Неврахування сполучуваності слів, буквальний поелементний переклад стійкої сполуки, наприклад:

ТО: *«Напомним, заместитель главы НБУ Дмитрий Сологуб заявил, что колебания на валютном рынке Украины свидетельствуют о том, что в стране в действительности гибкий плавающий обменный курс»* [11].

МП: *«Нагадаємо, заступник глави НБУ Дмитро Сологуб заявив, що коливання на валютному ринку України свідчать про те, що в країні в дійсності гнучкий плаваючий обмінний курс».*

ТП: *«Нагадаємо, заступник глави НБУ Дмитро Сологуб заявив, що коливання на валютному ринку України свідчать про те, що в країні насправді гнучкий плаваючий обмінний курс».*

Маємо справу з помилкою *«в действительности»* – *«у дійсності»* (правильно – *«насправді»*).

2. 6. Переклад власної назви як загальної, наприклад:

ТО: *«Печерский районный суд города Киева восстановил прежнего ректора Национального медицинского университета имени Богомольца Екатерину Амосову в должности руководителя вуза. Об этом Амосова сообщила в Facebook"* [11].

МП: *«Печерський районний суд міста Києва відновив колишнього ректора Національного медичного університету імені прочанина Катерину Амосову на посаді керівника вищого навчального закладу. Про це Амосова повідомила в Facebook".*

ТП: *«Печерський районний суд міста Києва відновив колишнього ректора Національного медичного університету імені Богомольця Катерину Амосову на посаді керівника вищого навчального закладу. Про це Амосова повідомила у в Facebook".*

2.7. Неправильний переклад скорочень за умови, що вони збігаються з іншим словом або формою слова, наприклад:

ТО: *«На Филиппинах от коронавируса из Китая умер 44-летний пациент. Об этом сообщает филиппинское отделение Всемирной организации здравоохранения (ВОЗ) на своей странице в Twitter"* [11].

МП: *«На Філіппінах від коронавируса з Китаю помер 44-річний пацієнт. Про це повідомляє філіппінське відділення Всесвітньої організації охорони здоров'я (ВІЗ) на своїй сторінці в Twitter".*

ТП: *«На Філіппінах від коронавіруса з Китаю помер 44-річний пацієнт. Про це повідомляє філіппінське відділення Всесвітньої організації охорони здоров'я (ВООЗ) на своїй сторінці у Twitter".*

У даному разі маємо справу не лише з неправильним перекладом скорочення, крім того, у словнику СМП немає слова *«коронавірус».*

3. Граматичні помилки:

3.1. Неправильне визначення частиномового статусу слова, унаслідок цього добір іншого еквівалента (неврахування міжчастиномовної омонімії, тобто збігу форм різних слів), наприклад:

ТО: *«Не можна попрацювавши з тілом не повпливати на тонку структуру, але можна це робити намірено, через тіло. Що я, власне, і роблю»* [12].

МП: *«Нельзя поработав с телом не повлиять на тонкую структуру, но можно это делать намеренно, через тело. Что я, собственное, и делаю».*

ТП: *«Нельзя поработав с телом не повлиять на тонкую структуру, но можно это делать намеренно, через тело. Что я, собственно, и делаю».*

У даному разі маємо справу зі збігом форми прикметника та прислівника *«власне».*

3.2. Порушення керування й узгодження граматичних форм на рівні словосполучення, що пов'язано переважно із двома причинами, а саме: зміною граматичних категорій опорного слова (наприклад, категорії роду); збігом різних відмінкових форм у мові оригіналу (морфологічна омонімія), наприклад:

ТО: *«Головним героєм фільму є воєнний підрозділ із трьох чоловіків, який утримує одну доволі тиху ділянку фронту»* [10].

МП: *«Главным героем фильма является военный подразделение из трех мужчин, содержащий одну довольно тихую участок фронта».*

ТП: *«Главным героем фильма является военное подразделение из трех мужчин, удерживающее один довольно тихий участок фронта».*

Як бачимо, у цьому коротенькому реченні декілька помилок, два порушення керування й узгодження на рівні словосполучення, крім того, лексична помилка, про яку йшлося вище.

3.3. Неправильне творення граматичної форми слова внаслідок неврахування омонімії, приклад можемо навести з пункту 2.3.

3. 4. Помилки в перекладі прийменників, наприклад:

ТО: *«Ті, хто не уклав декларації з сімейним лікарем, терапевтом, у звичайному порядку можуть звернутися до терапевтів за місцем проживання»* [10].

МП: *«Те, кто не заключил декларации посемейным врачом, терапевтом, в обычном порядке могут обратиться к терапевтам по месту жительства».*

ТП: *«Те, кто не заключил декларации с семейным врачом, терапевтом, в обычном порядке могут обратиться к терапевтам по месту жительства».*

4. Синтаксичні помилки, що базуються здебільшого на порушенні норм побудови речень, наприклад:

ТО: *«Представитель Института электросварки внес ряд предложений для решения проблемы с мостом Патона: очистить все металлические конструкции от мусора; восстановить проходы; загерметизировать все протекающие участки, покрасить (для предотвращения коррозии в дальнейшем)»* [11].

МП: *«Представник Інституту електрозварювання вніс ряд пропозицій для вирішення проблеми з мостом Патона: очистити всі металеві конструкції від сміття; відновити проходи; загерметизувати всі протікають ділянки, пофарбувати (для запобігання корозії в подальшому)».*

ТП: *«Представник Інституту електрозварювання вніс низку пропозицій для вирішення проблеми з мостом Патона:*

очистити всі металеві конструкції від сміття; відновити проходи; загерметизувати всі ділянки, що протікають, пофарбувати (для запобігання корозії в подальшому)».

Висновки. Розглянувши типові помилки, що виникають під час перекладу за допомогою СМП "Google Translate", можемо констатувати, що для одержання правильного перекладу за допомогою СМП потрібне беззаперечне втручання людини, і не лише на етапі постредагування (під час якого людина може вибрати потрібне значення чи форму слова), але і до початку здійснення перекладу, коли треба належним чином підготувати текст, усунувши орфографічні й інші помилки, правильно обрати стилістику та специфіку мови оригіналу та перекладу, за потреби додати специфічні терміни до словничка користувача тощо. Тому попереднє редагування – це обов'язкова процедура МП на сучасному етапі. До того ж аналіз помилок СМП доводить, що переклад повністю за допомогою СМП, без втручання людини, на даному етапі розвитку СМП неможливий, навіть для медіатекстів загальної тематики.

Література:

1. Бацевич Ф. Основи комунікативної лінгвістики. Київ : Вид. центр «Академія», 2004. С. 304–342
2. Блехман М. Машинный перевод: история и реалии. *Компьютерное обозрение.* 1996. № 5 (29). С. 22–23.
3. Блехман М., Войнов В. Методы построения локальных АСНТИ. Харьков : ХГУ, 1989. 106 с.
4. Комиссаров В. Смысловая стратификация текста как переводческая проблема. *Текст и перевод.* Москва : Наука, 1988. С. 32–56.
5. Марчук Ю. Проблемы машинного перевода. Москва : Наука, 1983. 232 с.
6. Островський Д., Ситар Г. Способи відтворення англійських власних назв українською та російською мовами (на матеріалі роману Дж.К. Ролінґ «Гаррі Поттер і філософський камінь»). *Вісник студентського наукового товариства Донецького національного університету імені Василя Стуса.* 2017. Т. 1. № 9. С. 58–62.
7. Періжняк М. Основні проблеми використання машинного перекладу для перекладу художніх творів. *Інтелектуальні системи та прикладна лінгвістика :* матеріали Всеукраїнської науково-практичної конференції. Харків, 2012. С. 69.
8. Швейцер А. Теорияперевода: статус, проблемы, аспекты. Москва : Наука,1988. 214 с.
9. Янковець А. Особливості використання машинного перекладу в умовах професійної підготовки майбутніх перекладачів. *Науковий вісник Ужгородського університету*. Серія «Педагогіка. Соціальна робота» / гол. ред. І. Козубовська. Ужгород : Говерла, 2012. Вип. 25. С. 233–236.
10. URL: https://www.ukrinform.ua/.
11. URL: https://www.rbc.ua/ukr.
12. URL: https://gazeta.ua/.

Gaiduk N., Koktash M. Typical errors when using a machine translation system "Google Translate"

Summary. The article deals with the use of machine translation in professional translation, which is currently of interest to both linguists and software developers, as today machine translation is a whole range of advanced, rapidly evolving information technologies which are widely used.

The search for ways to provide maximum assistance to professional translators is especially important in the conditions of rapid development of scientific, technical, and informational progress. It is not currently possible to replace completely a human translator by translating programs, but it is at least possible to automate the translation process performed by a person to facilitate the difficult work of a translator and make it as efficient as possible. This can only be done by integrating the efforts of cybernetics, programming, psychology and linguistics specialists. It will create a system of machine translation that allows to receive adequate text translation without the intervention of human translators. Achieving this goal is impossible without a thorough analysis of the causes of typical errors that occur with machine translation. The authors of the article analyze the typical errors that occur with machine translation using the system of machine translation "Google Translate", grouping them into lexical (including lexicographic), grammatical, syntactic errors, and errors that result from the lack of word forms in the dictionary of machine translation systems "Google Translate" in the translation material in Russian excerpts from newsfeeds of modern Ukrainian agencies, executed with the help of machine translation system "Google Translate", which are studied, researched and evaluated taking into account information about the working algorithm of machine translation systems, their grammars, dictionaries and internal construction.

Key words: information technologies, machine translation, machine translation systems, Google Translate.