

*Бочан П. О.,
кандидат історичних наук, доцент,
доцент кафедри лінгвістики та перекладу
Чернівецького національного університету імені Юрія Федьковича
<https://orcid.org/0000-0003-1473-4411>*

ПРОБЛЕМИ МАШИННОГО ПЕРЕКЛАДУ АНГЛОМОВНОЇ ОМОНІМІЧНОЇ ЛЕКСИКИ

Анотація. Статтю присвячено дослідженню проблем машинного перекладу англomовної омонімічної лексики. У статті надаються дефініції поняттям «машинний переклад», «амбівалентна лексика», «омонім», з'ясовуються види та програми машинного перекладу, встановлюються ознаки і види омонімів, аналізують труднощі і помилки машинного перекладу англomовних омонімічних одиниць засобами української мови, описуються способи усунення подібних помилок перекладачами.

Машинний переклад є автоматизованим видом перекладу, що забезпечує міжмовну комунікацію шляхом програмної інтерпретації та іншомовної ретрансляції мовних одиниць різного рівня. Машинний переклад здійснюється за допомогою перекладацьких програм, що відрізняються за технологічним типом творення, приналежністю до певного покоління, особливостями вхідних даних, ступенем участі людини в процесі перекладу, сферою його застосування.

Машинний переклад не забезпечує повне усунення помилок при іншомовній передачі амбівалентної лексики, яка характеризується неоднозначністю змісту. До категорії подібної лексики належать і омоніми, що є лексичними одиницями, тотожними за формою, але відмінними за значенням. Омоніми не є гомогенним явищем, вони поділяють на повні та часткові. Повні омоніми співпадають графічно та фонетично. Часткова омонімія представлена омофонами, які відрізняють графічно, але є тотожними фонетично, а також омографами, тотожними за графічним оформленням і відмінними фонетично.

При машинному перекладі помилки виникають переважно при відтворенні повних омонімів та омографів, оскільки омофони більш вдало ідентифікуються перекладацькими програмами за графічним оформленням. Якість перекладів речень із вказаними одиницями залежить від типу перекладацької програми, її приналежності до старого або нового покоління перекладацьких програм, а також особливостей синтактики та семантики самих речень, в яких вживаються подібні одиниці, специфіки понять, які вони позначають.

Англomовні омоніми, як правило, або не розпізнаються перекладацькими програмами старого покоління, або обирається перший із наданих у їх англо-українському тезаурусі перекладацький відповідник омоніма. Якщо речення, в яких використовуються омоніми, не є ускладненими семантично та синтаксично, ці лексеми правильно відтворюються перекладацькими програмами. Перекладені речення з омонімами зазвичай потребують коригування з боку перекладача, що здійснюється шляхом редагування та використання перекладацьких трансформацій. Перекладацькі програми нового покоління,

що базуються на технологіях нейронного перекладу, як правило, забезпечують правильну передачу англomовних омонімів і речень із ними українською мовою. Подібні речення можуть підлягати коригуванню лише для усунення незначних недоліків і пристосування до нормативних вимог мови перекладу.

Ключові слова: англійська мова, амбівалентна лексика, машинний переклад, омонім, перекладацька трансформація, програма.

Постановка проблеми. Інформатизація та цифровізація сучасного гуманітарного знання вплинула на розвиток різних наукових дисциплін про мову. У першу чергу, це стосувалося появи та розвитку таких її міждисциплінарних галузей, як комп'ютерна лінгвістика, метою якої стала «реконструкція знань про мову і в мові, яка б забезпечувала автоматизацію інтелектуальних функцій і когнітивної діяльності людини, автоматизоване породження мовлення та його комп'ютерну обробку й розпізнавання і т. і.» [1, с. 648], а також машинний переклад, «спрямований на комп'ютерне програмування та створення автоматичних систем перекладу з однієї мови на іншу» [1, с. 685].

Аналіз останніх досліджень і публікацій. На сьогоднішній день вже існує достатня кількість наукових праць, присвячених опису теоретичних засад машинного перекладу [2; 3], а також практичних аспектів застосування його програм для перекладу текстів і одиниць різних мов [4–7]. При цьому й дотепер не приділено увагу проблемам машинного перекладу амбівалентної лексики, до якої належать й омонімічні одиниці.

Метою статті у зв'язку з цим є дослідження проблем машинного перекладу англomовної омонімічної лексики, а її завдання полягають в уточненні та наданні дефініцій поняттям «машинний переклад», «омонім», «амбівалентна лексика», з'ясуванні видів і програм машинного перекладу, встановленні ознак і видів омонімів як одиниць амбівалентної лексики, аналізі труднощів і помилок машинного перекладу англomовних омонімічних одиниць засобами української мови на конкретних прикладах, а також описі способів усунення подібних помилок перекладачами.

Виклад основного матеріалу. Аналіз запропонованих науковцями визначень поняття «машинний переклад» [2, с. 12; 3, с. 7] сприяв його уточненню та наданню йому власної дефініції. Отже, машинним перекладом вважаємо автоматизований вид перекладу, що забезпечує міжмовну комунікацію шляхом програмної інтерпретації та іншомовної ретрансляції мовних одиниць різного рівня.

На сьогоднішній день існує велика кількість програм машинного перекладу, що відрізняються за технологічним принципом (архітектурою), особливостями та типом вхідних даних тощо. Зауважимо також, що машинний переклад може здійснюватися на двох рівнях. Перший із них є дослівним або буквальним перекладом (*metaphrase translation*), що є характерним для перекладацьких програм старшого покоління. При подібному перекладі не відбувається адекватна передача змісту тексту оригіналу, оскільки буквализм, як правило, призводить до викривлення його змісту. Недослівний переклад (*paraphrase translation*) дозволяє адекватно відтворити зміст тексту оригіналу, але його передача здійснюється шляхом переформулювання, а не буквальної відповідності [3, с. 26].

Існує також машинний переклад, що здійснюється за правилами (*rule-based machine translation, RBMT*) [3, с. 27]. Хоча подібний переклад і тяжіє до буквализму, він передбачає морфологічну, синтаксичну та семантичну інформацію, на основі якої розроблено лінгвістичні правила для програми, за допомогою яких здійснюється переклад. Прикладом найпростішої із подібних програм, що може бути задіяна при перекладі англо-українських текстів, а також текстів інших європейських мов, є *Pragma*, розроблена українсько-американською компанією *Trident Software*.

Зауважимо, що наявність виключень у правилах граматики, амбівалентна лексика може потребувати додаткового редагування від перекладача перекладених програмою текстів і речень, хоча її завдання і полягають у конвертації структури (граматичної, семантичної) текстів (речень) вихідної мови. Відзначаємо також, що якість відтворення інформації за допомогою вказаної програми залежить від розбіжностей будови вихідної та цільової мов, особливостей синтаксичних зв'язків у них, сполучуваності їх мовних одиниць, ступеня складності речень і текстів, що перекладаються. Чим більшими є подібні розбіжності, тим нижчою є якість перекладу, виконаного вказаною програмою. Це можна проілюструвати на основі аналізу наведених нижче речень:

This is definitely an indication of an intelligent and sensible approach to your entry into the dental profession [8, с. 51].

Це, безумовно, показник інтелектуального та зваженого підходу до вашого вступу в професію стоматолога.

C'est certainement une indication d'une approche intelligente et sensée à votre entrée dans la profession dentaire.

Аналіз вихідного речення та його перекладів, здійснених програмою *Pragma*, вказує на вищу якість французького перекладу англійського речення. Перекладене українською мовою речення потребує подальшого коригування перекладачем. Це стосується зміни фрагмента речення «*показник інтелектуального та зваженого підходу до...*» на «*свідчить про розумний і зважений підхід до...*». Зазначене вказує вищу якість перекладів речень і текстів, що належать мовам зі схожою граматичною будовою (англійська та французька мови – аналітичні, у той час, як українська мова – синтетична).

Зауважимо також, що визначена перекладацька програма поступається за якістю виконання перекладів сучасним нейронним перекладацьким програмам (*neural machine translators*), що є програмами нового покоління, які забезпечують більш природні формулювання різнорівневих мовних одиниць при іншомовному відтворенні, оскільки володі-

ють здатністю до їх глибокого контекстного моделювання [3, с. 28]. Прикладами подібних програм є *Google Translator*, *DeepL Translator*, *Amazon Translator*, *Microsoft Translator*, *ChatGPT*, що активно вживаються перекладачами нині. Види машинного перекладу та програми його реалізації розрізняються також за ступенем участі людини в цьому процесі, сферою застосування тощо.

Попри переваги машинного перекладу та здатність його програм перекладати значний обсяг текстів з високою швидкістю, він не може замінити перекладача на етапі підготовки їх для перекладу, а також на етапі перевірки та постредагування перекладених текстів, що полягає в усуненні граматичних, синтаксичних, стилістичних та інших помилок і недоліків машинного перекладу. Вважається, що чим більше машинний переклад наближається до професійного людського, тим кращим він є [2, с. 11]. При цьому не тільки тексти, але й окремі групи їх одиниць, можуть відтворюватися з помилками та вимагати подальшого коригування перекладачем.

Зокрема йдеться про омоніми, поширені не тільки в європейських, але й інших мовах світу. Омніми характеризують як «лексичні одиниці, що мають однаковий звуковий склад і однакове графічне оформлення» [9, с. 11], «лексичні одиниці, тотожні за вираженням та різні за змістом» [10, с. 25]. У англійських словниках надаються наступні визначення омонімам:

«each of two or more words having the same spelling or pronunciation but different meanings and origins» [11, с. 839];

«each of two words having the same pronunciation but different meanings, origins, or spelling» [12, с. 1102].

До ознак вираження омонімів належить, відповідно, співпадіння слів у написанні та вимові, у той час, як в плані змісту вони можуть відрізнятися за своїм лексичним і граматичним значенням. Це дозволяє стверджувати, що омоніми поділяються на лексичні, граматичні та лексико-граматичні. Подібний розподіл омонімів є важливим, оскільки ознака розбіжності граматичних значень також вважається ознакою омонімії, поряд із лексичними [13, с. 82]. Відповідно, омоніми можуть володіти однією або обома із вказаних ознак. Омніми відносять до одиниць амбівалентної лексики, що характеризується своєю неоднозначністю у розумінні [14, с. 78].

Встановлення ознак омонімів, а також аналіз їх україномовних та англійських дефініцій, надає можливість запропонувати власне визначення цих лексичних одиниць: дві чи більше одиниці мови, що мають власні лексичні та / або граматичні значення в плані змісту, але є тотожними за звуковим та / або графічним оформленням. На основі запропонованого та інших дефініцій омонімів, а також наукових праць, присвячених проблемам омонімії [9; 10; 13; 14], розрізняємо такі види омонімів: повні омоніми, що співпадають графічно та фонетично; омофони, що є частковими омонімами, які відрізняють графічно, але є тотожними фонетично; омографи, що є частковими омонімами, тотожними за графічним оформленням і відмінними фонетично.

Зауважимо, що недоліками та помилками характеризується, у першу чергу, машинний переклад текстів і речень із повними омонімами та омографами, оскільки омофони у більшій мірі розрізняються перекладацькими програмами за графічним оформленням. Проаналізуємо, як визначені одиниці та речення із ними відтворюються перекладацькими програмами *Pragma* та *Microsoft Translator* (версія 25.10.02):

It was done within a changed structure in order to make it easier for the reader to understand and to better accommodate to future changes [15, с. 127].

Це було зроблено в межах зміненої структури, для того, щоб зробити її простіше для читача, щоб зрозуміти та адаптувати до майбутніх змін.

Це зроблено в межах зміненої структури, щоб спростити її розуміння читачем і **краще** адаптувати до майбутніх змін.

Зауважимо, що лексема *better*, вжита у вищенаведеному англійському реченні, є повним омографом. Вона використовується для позначення вищого ступеня прикметника *good*, дієслова зі значеннями *перевершувати*, *перевищувати*, а також іменника зі значенням *той, хто заключає пари* [16, с. 58], утвореного суфіксальним способом від дієслова *to bet*.

Переклад, здійснений програмою *Pragma*, свідчить про вилучення аналізованого омоніма, що зумовлено варіативністю його відповідників в мові перекладу та неможливістю ідентифікації його правильного значення у вказаному реченні. Безумовно, подібне вилучення не є виправданим, оскільки воно спотворює зміст вихідного речення при перекладі. Про недосконалість і необхідність коригування речення, перекладеного вказаною програмою, свідчить і його граматична структура та семантична організація, зокрема його фрагменти *щоб ...*, *щоб...*.

Аналіз сполучуваності одиниць, їх значення (в тому числі, й перекладеного омоніма *better*), а також структурної, граматичної та семантичної організації речення, відтвореного програмою *Microsoft Translator*, засвідчує правильність і адекватність його передачі засобами української мови. Для цього в перекладене речення було внесено такі зміни, як перестановка слів та їх заміна, що корелює із контекстом їх вживання не тільки в мові оригіналу, але й в мові перекладу, зокрема: *in order to make it easier for the reader to understand* – для того, щоб зробити її простіше для читача. Вказані зміни стали можливими завдяки застосуванню перекладацьких трансформацій перестановка слів і заміна.

Звертаємося до аналізу іншого приклада речення з омонімом, що є скороченою одиницею (аббревіатурою), і його україномовних перекладів, здійснених програмами *Pragma* та *Microsoft Translator*:

Caution should be exercised when using frequencies above the MUF at night in tropical regions as long-distance 'chordal' or transequatorial propagation may be stimulated [17, с. 27].

Слід виявляти обережність при використанні частот вище *MUF* вночі в тропічних регіонах, оскільки далекодіюче хордове або трансекваторіальне поширення може бути стимульоване.

В тропічних регіонах слід виявляти обережність, оскільки використання частот вище *MPC* може активізувати режим дальнього хордового або трансекваторіального поширення радіохвиль.

Зауважимо, що у вихідному реченні використовується омонімічна аббревіатура *MUF*, повними відповідниками якої в англійській мові є наступні терміносполуки: *Maximum Usable Frequency*, *Model Uncertainty Factor*, *Multi-User Forth* [18, с. 330]. Перша із зазначених сполук належить сферам радіофізики та радіоелектроніки, друга терміносполука вживається у математичному моделюванні, ризик-аналізі, кліматології. Що стосується третьої терміносполуки, то її

використання здійснюється у сферах програмування та конструювання операційних систем. Саме контекст вживання визначає вибір правильного відповідника визначеної аббревіатури в українській мові, яким можуть бути такі сполуки, як: *максимально придатна частота (MPC)*, *фактор невизначеності моделі*, *багатокористувацький Forth* (мова програмування) [19, с. 199].

Вочевидь, перший із наведених перекладів вказує на те, що омонімічна аббревіатура не була дешифрована програмою *Pragma* відповідно до контексту її використання. В зв'язку з цим її було відтворено в україномовному перекладі в оригінальній графічній формі *MUF*. На протипагу зазначеному способу передачі аналізованого омоніма, його україномовний переклад, здійснений програмою *Microsoft Translator* (*MPC*, тобто *максимально придатна частота*), є адекватним контексту використання визначеної одиниці, оскільки в реченні, де вона вживається, йдеться про режими поширення радіохвиль.

Зазначене речення є граматично правильним, пристосованим до норм і правил української мови. Воно адекватно відтворює зміст вихідного англійського речення, чому сприяє вживання перекладацьких трансформацій перестановка слів, додавання слова *радіохвиль* (*Caution should be exercised ... in tropical regions* – В тропічних регіонах слід виявляти обережність ...; *long-distance 'chordal' or transequatorial propagation may be stimulated* – режим дальнього хордового або трансекваторіального поширення радіохвиль), лексико-граматичні заміни (*when* – оскільки, *may be stimulated* – поширення). У свою чергу, речення, перекладене за допомогою програми *Pragma*, тяжіє до буквализму та характеризується граматичною та синтаксичною незграбністю, стилістичними огріхами (*long-distance 'chordal'* – дальнє хордове ... поширення або далеке хордове ...поширення, а не далекодіюче хордове ... поширення) тощо. Безумовно, перекладене у подібний спосіб речення, потребує подальшого коригування перекладачем.

Зауважимо, що частиномовна приналежність, стилістичний реєстр, структура омонімів не є визначальними факторами правильності іншомовної передачі цих одиниць при машинному перекладі. Ключове значення має вид та характеристики перекладацької програми, що застосовується, а також структура та семантика речення, що є контекстом вживання омоніма, його місце в реченні, специфіка понять, які він позначає, зокрема:

Appropriate measures should be taken to prevent exhaust steam to any crane [20, с. 17].

Відповідні заходи повинні бути вжити, щоб запобігти відпрацьованому пару потрапити до будь-якого крана.

Необхідно вжити відповідних заходів для запобігання потрапляння відпрацьованого пару до будь-якого вантажонідиального крана.

Омонімічна лексема *crane* вживається в англійській мові не тільки для позначення видів птахів (*журавель*, *цапля*), але й для найменування технічного поняття (*вантажонідиальний*) кран [16, с. 70]. Омонім було правильно відтворено і програмою *Pragma*, і програмою *Microsoft Translator*. Цьому сприяло вживання визначеної одиниці у простому неускладненому англійському реченні у кінцевій позиції. Як і у попередньо наведених прикладах, речення із аналізованим омонімом, від-

творене програмою Прагма, потребує постредагування перекладачем, на відміну від речення, перекладеного програмою Microsoft Translator. В останнє із зазначених речень можна внести лише окремі зміни, враховуючи загальний контекст його застосування.

Висновки та перспективи дослідження. Таким чином, омоніми є лексичними одиницями, що характеризуються тотожністю графічного та / або фонетичного оформлення та різницею значення. Труднощі машинного перекладу вказаних одиниць зумовлені не тільки їх неоднозначністю, що виявляє їх приналежність до категорії амбівалентної лексики, але й якістю перекладацьких програм, задіяних для їх іншомовного відтворення. Машинний переклад англійських речень і текстів із вказаними одиницями може вимагати коригування та усунення помилок, зумовлених неможливістю перекладацьких програм правильно ідентифікувати та відтворити значення визначених одиниць. **Перспективи подальших досліджень** вбачаємо в аналізі труднощів машинного перекладу інших амбівалентних одиниць англійської мови засобами української, а також з'ясуванні способів усунення помилок і недоліків перекладених речень із ними.

Література:

1. Селіванова О. О. Сучасна лінгвістика : напрями та проблеми. Полтава : Довкілля-К, 2008. 712 с.
2. Hutchins W. J., Somers H. L. An Introduction to Machine Translation. London: Academic Press, 1992. 362 p.
3. Asscher O. Machine Translation and Translation Theory. London : Routledge, 2025. 178 p.
4. Гудманян А., Мишко А., Брай А. Оцінка адекватності машинного перекладу письмових спеціалізованих текстів. *Advanced Linguistics*. 2023. № 12. С. 66–77. DOI: 10.20535/2617-5339.2023.12.291604
5. Гудманян А. Г., Сітко А. В., Струк І. В. Функціонально-прагматична адекватність машинного перекладу публіцистичних текстів. *Львівський філологічний часопис*. 2019. Т. 5. С. 48–52. DOI <https://doi.org/10.32447/2663-340X-2019-5-8>
6. Jurafsky D., Martin J. H. *Speech and Language Processing* / 2-nd ed. Upper Saddle River : Prentice Hall, 2009. 988 p.
7. Navigli R. Word Sense Disambiguation: A Survey. *ACM Computing Surveys*. 2009. Vol. 41. № 2. P. 1–69. DOI <https://doi.org/10.1145/1459352.1459355>
8. Fejerskov O., Kidd E. *Dental Caries: The Disease and Its Clinical Management* / 3-d ed. Oxford : Wiley Blackwell, 2015. 640 p.
9. Шипнівська О. О. Міжчастиномовна морфологічна омонімія сучасної української мови монографія. Київ : Центр учбової літератури, 2017. 192 с.
10. Tulloch A. *Understanding English Homonyms: Their Origins and Usage*. Hong Kong : Hong Kong University Press, 2017. 164 p.
11. Hornblower S., Spawforth A., Eidinow E. *The Oxford Classical Dictionary* / 4-th ed. Oxford : Oxford University Press, 2012. 1592 p.
12. *Longman Dictionary of Contemporary English* / ed. by D. Summers. Harlow: Pearson Education Limited, 2008. 1949 p.
13. Кійко С. В. Німецько-українська міжмовна омонімія : монографія. Чернівці : Видавничий дім «Родовід», 2014. 434 с.
14. Lein J. D., Berriman E. *Mostly Homonyms: A Whimsical Perusal of Those Words That Sound Alike*. Lanham : University Press of America, 2012. 166 p.
15. Swales J. M., Feak C. B. *Academic Writing for Graduate Students: Essential Tasks and Skills* / 3rd ed. Ann Arbor : University of Michigan Press, 2012. 320 p.
16. Гороть С. І. *Англо-український словник*. Вінниця : Нова книга, 2006. 1700 с.
17. Klobuchar J. A., Rawer K., Ratcliffe J. A. *Ionospheric Radio Propagation*. New York : Wiley, 1980. 412 p.
18. *Abbreviations Dictionary* / ed. by J. Lachlan. New York : Random House, 2016. 830 p.
19. *Англо-український словник аббревіатур. English-Ukrainian Abbreviations Dictionary* / уклад. : Богданова Л., Іваненко О., Петренко М. Київ : Академія, 2019. 432 с.
20. Occupational Safety and Health Administration (OSHA). *Safety Standards for Steam-Powered Equipment*. Washington D.C. : OSHA, 2015. 68 p.

Bochan P. Problems of machine translation of English homonymous vocabulary

Summary. The article studies the problems of machine translation of English homonyms. The article provides definitions of such notions as «machine translation», «homonym», «ambivalent lexicon», clarifies types and programs of machine translation, determines characteristics and types of homonyms, analyzes difficulties and errors in machine translation of English homonyms by means of the Ukrainian language, and describes ways of correcting such errors by translators.

Machine translation is an automated type one. It enables interlingual communication through software interpretation and foreign-language retransmission of language units of different levels. Machine translation is carried out with the help of translating programs different in technological approach, generation (old or new), input data characteristics, degree of human involvement in the translation process and the sphere of its application.

Machine translation does not completely eliminate errors in the foreign-language retransmission of ambivalent vocabulary, characterized by ambiguity of meanings. This category of vocabulary also includes homonyms. These are lexical units identical in form but different in meanings. Homonyms are not a homogeneous phenomenon; they are divided into full and partial ones. Full homonyms are identical in their spelling and pronunciation. Partial homonymy includes homophones that differ graphically but are identical phonetically, and homographs, identical graphically but differ phonetically.

Machine translation errors tend to occur in the processes of interlingual reproducing of full homonyms and homographs, since homophones are more successfully identified by translating programs due to their spelling. The quality of the translation of sentences containing such units depend on the type of the translation program used in the process, its belonging to the old or new generation of translating programs, as well as the syntactic and semantic features of the sentences with homonyms, and the specificity of the notions they designate.

English homonyms are usually either not recognized by older generation translating programs, or the first translating variant of a homonym provided by their English-Ukrainian thesaurus is chosen. If sentences with homonyms are not semantically and syntactically complicated, these lexemes

are usually correctly reproduced by translating programs. Translated sentences with homonyms usually require correction by the translator, carried out through editing and using translating transformations. New-generation translating programs based on neural translating technologies usually ensure the correct reproduction of English homonyms and sentences with them into the Ukrainian language. Such sentences may undergo revision to eliminate minor deficiencies and to ensure compliance with the normative standards of the target language.

Key words: English language, ambivalent vocabulary, machine translation, homonym, translating transformation, translating program.

Дата першого надходження статті
до видання: 19.02.2026
Дата прийняття статті до друку
після рецензування: 17.03.2026
Дата публікації (оприлюднення)
статті: 29.04.2026