

УДК 81'322.4

DOI <https://doi.org/10.32782/2522-4077-2024-210-16>**ДОСЛІДЖЕННЯ ПРОБЛЕМ МАШИННОГО ПЕРЕКЛАДУ  
ЗАСОБАМИ БАГАТОМОВНИХ ПАРАЛЕЛЬНИХ КОРПУСІВ****INVESTIGATING MACHINE TRANSLATION PROBLEMS BY  
MEANS OF MULTILINGUAL PARALLEL CORPORA****Дужа-Задорожна М.П.,***orcid.org/0000-0003-4813-3294**кандидат педагогічних наук, доцент,**доцент кафедри прикладної лінгвістики**Національного університету «Львівська політехніка»***Качур Я.І.,***orcid.org/0009-0000-6164-9856**студентка II курсу магістратури кафедри прикладної лінгвістики**Національного університету «Львівська політехніка»*

Стаття присвячена виділенню та аналізу основних проблем, що виникають в результаті застосування машинного перекладу. Об'єктом дослідження стала система машинного перекладу eTranslation, що використовується на офіційному сайті Європейського Союзу. Головним інструментом проведеного дослідження є багатомовний паралельний корпус EU History, сформований нами з метою порівняння та аналізу текстів різними мовами. Він містить оригінальні статті про історію створення та розвиток Європейського Союзу англійською мовою та їхні переклади німецькою й українською мовами. Вказаний корпус дозволив виявити та проаналізувати помилки і неточності, допущені системою машинного перекладу eTranslation. У результаті дослідження було виділено кілька основних проблем машинного перекладу, серед яких неадекватне врахування полісемії та контекстних маркерів, оскільки система машинного перекладу часто не розпізнає багатозначність понять та не враховує контекст, в якому вони використовуються, що призводить до помилок у перекладі. Система машинного перекладу може також мати труднощі з перекладом складних граматичних конструкцій, аббревіатур та власних назв, що робить переклад неточним або незрозумілим. У процесі дослідження було виявлено проблему відсутності універсальних способів транскрибування з латиниці на кирилицю, що унеможливує адекватне відображення власних назв системою машинного перекладу. Згадана система не завжди дотримується правил узгодження в українській мові, що також негативно впливає на якість перекладу. Отримані результати свідчать про те, що машинний переклад на сучасному етапі ще не може повністю замінити людський переклад, особливо коли йдеться про складні тексти, що потребують глибокого розуміння контексту та знання тонкощів мови. Стаття розглядає також потенціал машинного перекладу в поєднанні з антропогенним перекладом для досягнення кращих результатів у міжмовній комунікації та доступі до інформації.

**Ключові слова:** машинний переклад (МП), система машинного перекладу (СМП), багатомовний паралельний корпус, контекстний маркер.

The article deals with the identification and analysis of the main problems that are caused by machine translation. The object of the research is the eTranslation machine translation system, which is used on the official website of the European Union. The main tool of the conducted research is the multilingual parallel corpus EU History, which we have created to compare and analyze texts in different languages. It contains original articles about the history of the creation and development of the European Union in English and their translations into German and Ukrainian. This corpus allowed us to identify and analyze errors and inaccuracies made by the eTranslation system. As a result of the study, several major machine translation problems were identified, including inadequate consideration of polysemy and contextual markers, as the machine translation system often fails to recognize the ambiguity of concepts and does not consider the context in which they are used, leading to translation errors. The machine translation system may also have difficulties translating complex grammatical constructions, abbreviations, and proper names, making the translation inaccurate or unclear. The research revealed the absence of universal methods for transcribing from Latin

to Cyrillic, which makes it impossible for the machine translation system to adequately render proper names. This system does not always adhere to the rules of agreement in the Ukrainian language, which also negatively affects the quality of translation. The results obtained indicate that machine translation at the current stage cannot fully replace human translation, especially when it comes to complex texts that require a deep understanding of context and knowledge of language nuances. The article also examines the potential of machine translation in combination with human translation to achieve better results in interlingual communication and access to information.

**Key words:** machine translation (MT), machine translation system (MTS), multilingual parallel corpus, contextual marker.

**Постановка проблеми.** Глобалізація, світова співпраця, відкриті кордони – все це веде до тісного зв'язку між людьми з різних країн та регіонів. Крім вказаних позитивних чинників, які спричиняють великі хвилі міграцій, на жаль, мусимо назвати також війни та природні катаклізми. У більшості випадків саме переклад є єдиним шляхом для вирішення проблем, що виникають під час міжнародної комунікації. Туристи, біженці, працівники великих корпорацій, волонтери, студенти потребують допомоги в різних ситуаціях від оформлення документів до короткої розмови біля каси, проте забезпечити кожного кваліфікованим перекладачем не можливо, а самостійне вивчення мови – дуже довгий і складний процес. Саме в таких випадках у гру вступає машинний переклад (МП). Передача інформації з однієї мови на іншу ще ніколи не була такою швидкою та легкою, але різні культурні традиції й менталітет, явище полісемії, відмінності граматичних мовних систем та інші аспекти стоять на заваді адекватності, ефективності та якості МП. На сучасному етапі розвитку комп'ютерні технології ще не готові представити інструмент, який був би здатний врахувати всі можливі варіанти та видати ідеальний результат перекладу.

Виходячи з вищесказаного, вважаємо, що виділення та аналіз основних проблем МП є надзвичайно важливими для розробки інноваційних рішень, які допоможуть вдосконалити процес перекладу.

**Аналіз останніх досліджень і публікацій.** Відповідно до всебічного аналізу наукових джерел, можемо стверджувати, що на даний момент МП все ще залишається відносно новою сферою досліджень, хоча деякі її проблеми висвітлюють науковці як в Україні так і за кордоном. Зокрема українська лінгвістка Ю. Стахмич [1] акцентує свою увагу на поняттях еквівалентності та адекватності в межах комп'ютерної лінгвістики; мовознавиці Г. Веселовська та С. Радецька [2] розглядають можливі класифікації машинного перекладу, його переваги та недоліки, виокремлюють чітку типологію МП; науковець Л. Черноватий [3] фокусується на проблемах МП і розглядає його з точки зору способу навчання майбутніх перекладачів; дослідниця Л. Галій [4] працює з типовими помилками систем машинного перекладу (СМП); нігерійська лінгвістка М. Окпор [5] досліджує різні підходи до вивчення МП, а ірландські науковці М. Гірін та А. Вей [6] розглядають статистичний машинний переклад як домінуючий підхід у дослідженнях МП. Однак поза увагою дослідників залишився аналіз використання багатомовних паралельних корпусів як інструменту машинного перекладу.

**Метою дослідження** є виділення та аналіз основних проблем, що виникають в результаті застосування машинного перекладу з використанням паралельних корпусів. Головним інструментом проведеного дослідження є багатомовний паралельний корпус EU History, який ми створили спеціально для досягнення поставленої цілі. Матеріалом дослідження стали статті про історію Європейського Союзу англійською, німецькою та українською мовами з офіційного сайту ЄС.

**Виклад основного матеріалу дослідження.** Всебічне дослідження наукових джерел щодо аналізу машинного перекладу дозволило нам сформулювати дефініцію МП та визначити поняття еквівалентності змісту оригінального тексту та адекватності МП.

Під машинним перекладом більшість дослідників розуміють технологію автоматизованого перекладу текстів з однієї мови на іншу за допомогою комп'ютерної програми, яку можна

визначити як систему машинного перекладу (СМП). Спочатку згадана система аналізує вихідну мову та кодує семантичні й граматичні компоненти тексту. Далі отриманий результат перетворюється відповідно до правил функціонування цільової мови після чого генерується текст перекладу [7]. Сучасні СМП враховують дані на різних рівнях мови: морфологічному, семантичному, синтаксичному, але все ж частотність граматичних та лексичних помилок є досить високою, що в свою чергу піднімає питання їх класифікації [4].

На сучасному етапі не існує чітко прописаних стандартів, які б вказували єдино правильний шлях перекладу і відповідно до яких можна було б встановити точні критерії для його оцінки. Необхідно зважати також на те, що для сфери перекладу неможливо визначити універсальні правила, адже мови постійно змінюються та розвиваються і в найближчому майбутньому МП не зможе функціонувати без участі людини. Незважаючи на це, головною метою будь-якого перекладу залишається еквівалентність змісту оригінального тексту та його адекватність, проте трактування цих термінів науковцями є неоднозначним.

У нашому дослідженні ми послуговуємося визначенням вказаних понять українською лінгвісткою Ю. Стахмич, яка зазначає, що «під адекватністю машинного перекладу ми розуміємо забезпечення прагматичних засад перекладу відповідно до поставленої мети на максимально можливому рівні еквівалентності, зі збереженням форми оригіналу та якнайменшим порушенням норми мови перекладу» [1, с. 235]. А еквівалентність вона визначає як «семантичну та структурну відповідність одиниць основних рівнів мови (лексико-семантичного, морфологічного і синтаксичного) у текстах оригіналу та перекладу» [1, с. 236].

Варто зазначити, що МП все ще не може зрівнятися у якості з антропогенним перекладом, який також зі свого боку не може досягти абсолютної еквівалентності. Таке порівняння насправді не є доречним та їх протиставлення не є продуктивним підходом для розвитку галузі перекладу загалом. Натомість інтеграція цих двох видів перекладу в гіперантропогенний може бути ключем для покращення насамперед таких показників як швидкість та якість [3].

Для дослідження машинного перекладу доцільно використовувати багатомовні паралельні корпуси, які на сучасному етапі є одним з головних викликів у галузі лінгвістики та перекладознавства. Вони є цінними ресурсами та багатофункціональними інструментами для різноманітних застосувань в тому числі і для машинного перекладу. Англійська мовознавиця М. Бейкер виділяє 3 основні типи корпусів, які широко застосовують у сфері перекладу: порівняльний, паралельний та багатомовний [8]. Порівняльний корпус складається з колекції текстів однією мовою, які об'єднані спільною тематикою. Вони можуть бути корисними для виділення спільних та відмінних рис або ж для збору статистичних даних. Паралельний корпус — це збірка текстів, де один з них є оригіналом, а інші — версіями перекладу того самого оригінального тексту. Багатомовний корпус — набір двох або більше одномовних корпусів різними мовами. Загалом, з точки зору перекладознавства, дослідники мають справу з паралельними корпусами, які також можна визначити як багатомовні паралельні корпуси, оскільки вони завжди включають принаймні дві мови.

Для проведення нашого дослідження ми уклали багатомовний паралельний корпус EU History. Незважаючи на те, що паралельні корпуси мають свої відмінні риси, не існує конкретного способу їх створення. Цей процес залежить від різних факторів: мети, для якої формується корпус, сфери його застосування, типу корпусу тощо. Згідно з аналізом різноманітних наукових праць щодо побудови корпусів, зокрема паралельних [6; 9; 10; 11] ми виділили три основні етапи: 1) проектування корпусу; 2) збір та попередня обробка мовних даних; 3) побудова корпусу.

Враховуючи важливість та актуальність отримання інформації про історію ЄС для українців, ми запланували на першому етапі створення багатомовного паралельного корпусу англійською, німецькою та українською мовами, метою якого було б дослідження якості машинного перекладу та порівняння інформації, представленої різними мовами. Основні характеристики корпусу:

– Тип – багатомовний паралельний корпус, який включає англійську (як вихідну мову), німецьку (перша цільова мова; антропогенний переклад) та українську (друга цільова мова; машинний переклад).

- Джерело даних – офіційний сайт ЄС.
- Тематична спрямованість – історія Європейського Союзу.
- Розмір – обмежений наявністю відповідних паралельних текстів, представлених на сайті.
- Проблеми, які він може вирішити – оцінка якості машинного перекладу (МП).

Другий етап передбачав збір даних та попередню обробку текстів. Спочатку було перенесено необхідні текстові дані у таблицю в середовищі Excel. Щоб переконатися, що корпус працюватиме належним чином і уникнути додаткової роботи з форматами даних пізніше, найкращим рішенням є узгодження всіх текстів. Тому тексти було розбито на фрагменти: окремі речення або короткі абзаци не більше двох-трьох речень. Далі проведено перевірку, чи всі фрагменти знаходяться на одному рівні в таблиці для всіх мов. Таким чином на даному етапі ми вже отримали чорновий варіант корпусу.

На третьому етапі для побудови корпусу було використано сервіс ScetchEngine. Після реєстрації, обрання опції *New corpus*, введення імені для нашого корпусу та обрання його типу, а саме *Multilingual corpus*, сервіс запропонував нам завантажити файли з даними. Позначення *Aligned documents* означає, що користувачі попередньо узгодили тексти вручну та представляють їх у табличному форматі, де кожен сегмент тексту мовою оригіналу відповідає тому самому сегменту мовою перекладу. Якщо ж вказано *Non-aligned documents*, то система працює з форматами простого тексту та передбачає, що текст представлено в одному стовпці для кожної мови, а після завантаження даних сервіс ScetchEngine автоматично вирівняє тексти на рівні речень та перетворить їх у корпус. Цей процес не вимагає втручання користувача, але помилки такого узгодження неможливо виправити в SketchEngine. Тож, щоб уникнути помилок в разі некоректної роботи сервісу, ми завантажили попередньо узгоджені тексти. Після цього система ScetchEngine створила три окремі корпуси відповідно до кожної мови. Їх можна було досліджувати як окремі одномовні корпуси, проте щоб отримати паралельний корпус ми використали інструмент *Parallel Concordance*. Він одразу показує вирівняний та узгоджений текст, в якому зручно досліджувати фрагменти одночасно трьома мовами, крім того, дозволяє здійснити пошук необхідних елементів і також показує результат одночасно для всіх мов. Основною мовою офіційного сайту Європейського Союзу є англійська, проте більшість інформації, над укладанням якої працювали реальні люди, представлена мовами країн-членів ЄС, серед яких є й німецька. Українська мова відображається у розділі «інші мови», проте переклад здійснюється сервісом Європейської комісії *eTranslation*, який є передовою системою машинного перекладу.

В результаті ми отримали корпус текстів, які містять інформацію про засновників ЄС та його історію. В процесі проведеного дослідження ми визначили кілька найпоширеніших груп помилок, які виникали в результаті машинного перекладу. Варіант перекладу німецькою мовою використали для глибшого розуміння змісту текстів та отримання альтернативного погляду на можливий для української мови варіант перекладу.

Згідно з дослідженням нами було встановлено, що доволі часто виникають *проблеми перекладу явища полісемії, стилістичного та контекстуального маркерів* тощо. Наведемо приклади:

*Many of them worked to end the horrors of two world wars and to promote peace and solidarity* [12].

*Багато з них працювали над тим, щоб покласти край жахам двох світових воєн і сприяти миру і солідарності.*

Поняття *жах* є прямим відповідником до слова *horror*, але, якщо дослідити ширший контекст, а особливо новинний дискурс, можна помітити, що в схожих мовленнєвих ситуаціях



ми частіше зустрінемо слово *жахіття*, яке насправді є абсолютним синонімом до слова *жах*. Крім цього, всі члени речення є правильно узгоджені між собою, проте саме речення наслідує структуру англійського і тому звучить неприродньо для української мови. Загалом переклад передає основний зміст, але форма недопрацьована.

Пропонуємо варіант адекватного перекладу: *Багато з них працювали над тим, щоб покласти край жахіттям, які виникли внаслідок світових воєн, а також, щоб сприяти розвитку миру та солідарності.*

Наступний приклад детальніше висвітлює проблему полісемії:

*His leadership led to strong ties developing between Italy and the United States, no mean feat considering that Italy had one of the largest communist parties in Western Europe at the time* [12].

*Його керівництво призвело до міцних зв'язків між Італією і Сполученими Штатами, не маючи на увазі, що Італія мала одну з найбільших комуністичних партій в той час в Західній Європі.*

У наведеному прикладі слово *керівництво* є еквівалентом до слова *leadership* і вони можуть означати керівництво як процес і як осіб керуючої ланки. В той же час використання прямого відповідника *лідерство* могло б усунути дану проблему, адже це слово вказує лише на процес.

Також слово *mean* має цілу низку значень і СМП переклала його як *мати на увазі*, що є одним з найпоширеніших значень. Прямим відповідником *mean* є *означати*, але, якщо поглянути на наступне слово *feat*, яке можна перекласти як *подвиг*, то стає зрозумілим, що в даному випадку *mean* не може бути дієсловом. Серед інших значень цього слова також є такі прикметники як *пересічний, середній, посередній*. Тож, на нашу думку, значення фрази *no mean feat* можна трактувати як *непересічний виклик*.

Якщо у випадку з *leadership* міг виникнути легкий дисонанс, який не кожен читач би помітив, то неправильне відтворення слова *mean* призвело до повного спотворення змісту.

Далі проаналізуємо німецький варіант наведеного прикладу:

*Unter seiner Führung entwickelte sich eine enge Zusammenarbeit zwischen Italien und den Vereinigten Staaten. Das war keine leichte Aufgabe, wenn man bedenkt, dass es in Italien zu dieser Zeit eine der größten kommunistischen Parteien Westeuropas gab* [12].

Ми бачимо, що перекладач в обох випадках успішно розв'язав проблему полісемії. Замість того, щоб буквально відтворити оригінал (*Seine Führung führte dazu*) перекладач вжив фразу *unter seiner Führung*, що означає *під його керівництвом/проводом* і чітко вказує саме на процес керівництва. Також, для того щоб уникнути нагромадження інформації, перекладач розділив одне речення на два повноцінних, а фразу *непересічний виклик* переклав як *keine leichte Aufgabe*, що буквально означає *нелегке завдання* і є хорошим еквівалентом, який відтворює суть оригінальної фрази.

Пропонуємо такі приклади адекватного перекладу згаданого речення:

1) *Під його керівництвом розвинувся тісний зв'язок між Італією і Сполученими Штатами, що було нелегким завданням/непересічним викликом, адже Італія мала одну з найбільших комуністичних партій на той час в Західній Європі.*

2) *Його лідерство призвело до міцних зв'язків між Італією і Сполученими Штатами, незважаючи на те, що Італія мала одну з найбільших комуністичних партій на той час в Західній Європі.*

Іншою групою помилок, які виникають при МП є *проблема перекладу власних назв та аббревіатур*. На жаль, на даний момент не існує цілісної системи правил передачі власних назв записаних латинкою засобами українського алфавіту. Головною умовою є якнайближче відтворення звукової оболонки слів шляхом транскрибування.

Наприклад, ім'я італійського діяча *Alcide De Gasperi* при кожній згадці було відтворене СМП по різному: *Альцид де Гаспері, Алькіде де Гаспері, Альчиде де Гаспері*. У перших двох випадках виникла проблема з відтворенням звука, який позначає латинська літера «с». В останньому

варіанті система правильно транскрибувала цей звук, оскільки ім'я походить з Італії і потрібно відтворювати саме італійську манеру звучання. Проте остаточний і єдино правильний варіант українською мовою визначити важко саме через відсутність системи передачі власних назв. Якщо орієнтуватися на звукову оболонку, то правильним є варіант *Альчиде де Гаспері*. Якщо враховувати правила функціонування звуків в українській мові, *Альчиде де Гаспері* також можливий варіант. А найчастіше на різноманітних ресурсах можна зустріти варіант *Альчиде де Гаспері*, оскільки літера «Г» лише нещодавно почала повертатися до активного вжитку і ще не вкорінилася в системі транскрибації.

Ще однією проблемою передачі власних імен є те, що вони можуть співпадати з звичайними широкоживаними словами, які мають значення, наприклад, французьке ім'я *Simone Veil*, яке вживалося як заголовок до біографії, тобто поза контекстом, СМП переклала як *Симона вуаль*. Пізніше в контексті речень ми зафіксували ще два варіанти: *Симоне Вейл* та *Симона Вейль*. Якщо брати особливості французької звукової системи, то правильним є варіант *Симон Вейль*, але знову ж таки, якщо дещо адаптувати ім'я під українську систему, то правильним може бути і *Симона Вейль*.

Щодо відтворення абрєвіатур та назв спілок, то СМП в більшості випадків справляється добре. Наведемо приклади: *Christian Democratic Union (CDU)* – Християнсько-демократичний союз (ХДС); *Federal Republic of Germany* – ФРН, *NATO* – НАТО, *United Nations* – ООН; *EU – ЄС*; *European Economic Community (EEC)* – Європейське економічне співтовариство (ЄЕС). Проте з назвами, що не є загальновідомими та вживаються рідше все ж виникають неточності: *European Coal and Steel Community (ECSC)* – Європейське співтовариство вугілля і сталі (ЕСС), тоді як має бути *ЄСВС*. Хоча також тут варто зазначити, що часто в українських текстах автори чи перекладачі дозволяють собі використовувати англійський варіант абрєвіатур, проте це може призвести до певних непорозумінь з боку читачів.

Наше дослідження показало, що доволі часто зустрічаються помилки при перекладі конструкцій, прийменників, порушення узгодження чи керування та інших граматичних аспектів. Найпоширенішою виявилася проблема порушення узгодження. Наведемо приклади:

*He, along with other political prisoners, drew up the manifesto while incarcerated by the Italian fascist regime on the island of Ventotene* [12].

*Він разом з іншими політичними в'язнями склав маніфест під час ув'язнення італійського фашистського режиму на острові Венотене.*

З даного речення зрозумілим є те, що хтось ув'язнив італійський фашистський режим, а не те, що представники даного режиму ув'язнили політика. В оригінальному тексті для того, щоб показати хто кого поневолив використовується пасив, який не притаманний українській мові. Крім того, сполучник *while* наголошує на часовому проміжку.

Німецький варіант: *Während seiner Inhaftierung auf der Insel Ventotene durch das italienische faschistische Regime schrieb er das Manifest zusammen mit anderen politischen Gefangenen* [12].

Тут використовується прийменник *durch*, який вказує на причину ув'язнення об'єкта, а також прийменник *während*, який вказує на конкретний часовий проміжок.

Український переклад можна представити просто шляхом узгодження всіх членів речення відповідно: *Він разом з іншими політичними в'язнями склав маніфест під час ув'язнення італійським фашистським режимом на острові Венотене*. Або для наголошення того коли саме відбувається дія та для легшого сприйняття речення ми можемо розширити його та змінити порядок слів: *Під час його ув'язнення представниками італійського фашистського режиму на острові Венотене, він разом з іншими політичними в'язнями склав маніфест*.

Переклад граматичних конструкцій з *-ing* також є викликом для перекладача. Наприклад:

*He was the man behind Italy joining the American Marshall Plan and becoming part of NATO* [12].

*Він був людиною, яка стоїть за Італією, приєдналась до американського плану Маршалла і стала частиною НАТО.*

З наведеного речення ми можемо зрозуміти, що людина стоїть, приєдналася і стала частиною, що не має сенсу. В англійській мові закінчення *-ing* класично використовується для утворення дієслів, але в деяких випадках і для позначення процесів. Таким чином тут слова *joining* та *becoming* вживаються в ролі іменників.

У німецькому варіанті перекладач вирішив побудувати складне речення, в якому слово *Italia* стало окремим підметом:

*Er trug maßgeblich dazu bei, dass Italien am US-amerikanischen Marshallplan teilnahm und der NATO beitrug* [12].

На нашу думку, адекватним варіантом перекладу може бути такий:

*Він був відповідальний за приєднання Італії до американського плану Маршалла та вступ до НАТО.*

**Висновки і перспективи подальших досліджень.** Таким чином, у процесі дослідження, проведеного на основі створеного нами паралельного корпусу *EU History*, що містить оригінальні тексти англійською мовою та їхні переклади німецькою й українською мовами, було виокремлено та проаналізовано основні проблеми машинного перекладу. Було встановлено, що серед помилок і неточностей, допущених системою МП доволі часто зустрічаються: неврахування явища полісемії, стилістичного та контекстуального маркерів; помилки під час перекладу різноманітних конструкцій та аббревіатур; порушення правил узгодження чи керування та інших граматичних аспектів. Під час дослідження було виявлено також проблему відсутності універсальних способів транскрибування з латиниці на кирилицю, що унеможлиблює адекватне відображення власних назв системами машинного перекладу.

Узагальнюючи вищенаведене, можемо стверджувати, що МП запропонований сервісом eTranslation можна використовувати лише для поверхневого ознайомлення з інформацією, він ще не може повністю замінити антропогенний переклад, особливо коли йдеться про складні тексти, що потребують глибокого розуміння контексту та знання тонкощів мови. Згідно з проведеним дослідженням, вважаємо, що паралельні корпуси є цінним знаряддям для роботи з перекладами, зокрема й з МП.

Перспективи подальших досліджень вбачаємо у вивченні впливу різних систем машинного перекладу на переклад текстів в інших галузях та виявлення інших проблем і труднощів при застосуванні МП.

#### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Стахмич Ю. С. Адекватність та еквівалентність перекладу в контексті комп'ютерної лінгвістики. *Вісник Житомирського державного університету імені Івана Франка*. 2012. № 66. С. 235–238.
2. Веселовська Г., Радецька С. Машинний переклад: його типологія, переваги та недоліки. *Актуальні питання гуманітарних наук*. 2021. Вип 35, том 7. С. 23–27.
3. Черноватий Л. Проблеми машинного перекладу та його застосування у навчанні майбутніх перекладачів. *Наукові записки. Серія: Філологічні науки*. 2022. №202. С. 84–93.
4. Галій Л. Г., Шачковська, Т. Ю. Типові помилки систем машинного перекладу. 2017. С. 1–8.
5. Okpor M. D. Machine translation approaches: issues and challenges. *International Journal of Computer Science Issues (IJCSI)*. 2014. 159 p.
6. Hearne M., Way A. Statistical machine translation: a guide for linguists and translators. *Language and Linguistics Compass*. 2011. pp. 205–226.
7. Lifeng H., Jones G., Smeaton A. AlphaMWE: Construction of multilingual parallel corpora with MWE annotations. arXiv preprint arXiv. 2011.03783, 2020.
8. Baker M. Corpora in Translation Studies. *An Overview and Suggestions for Future Research*. 1995. pp. 223–243.
9. Choudhary N., Jha G. N. Creating multilingual parallel corpora in Indian languages. *Human Language Technology Challenges for Computer Science and Linguistics: 5th Language and Technology Conference, LTC 2011, Poznań, Poland, November 25–27, 2011, Revised Selected Papers 5*. Springer International Publishing, 2014.

10. Rosen A., Vavřín M. Building a multilingual parallel corpus for human users. *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*. 2012. 2447–2452 p.
11. Widdows D., Dorow B., Chan C-K. Using Parallel Corpora to enrich Multilingual Lexical Resources. *In Proceedings of the Third International Conference on Language Resources and Evaluation (LREC'02)*. 2002. URL: <http://www.lrec-conf.org/proceedings/lrec2002/pdf/103.pdf> (дата звернення: 20.07.2024)
12. Website of the European Union. URL: [https://european-union.europa.eu/principles-countries-history/history-eu\\_en](https://european-union.europa.eu/principles-countries-history/history-eu_en) (дата звернення: 20.07.2024)