

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**

**ФАКУЛЬТЕТ ІНОЗЕМНОЇ ФІЛОЛОГІЇ
КАФЕДРА ТЕОРІЇ ТА ПРАКТИКИ ПЕРЕКЛАДУ З АНГЛІЙСЬКОЇ
МОВИ**

**Кваліфікаційна робота
магістра**

**на тему СПЕЦИФІКА МАШИННОГО ПЕРЕКЛАДУ (НА МАТЕРІАЛІ
ПУБЛІЦИСТИКИ АНГЛІЙСЬКОЮ ТА УКРАЇНСЬКОЮ МОВАМИ)**

Виконав: студент 2 курсу, групи
8.0352-а/п
спеціальності 035 Філологія
спеціалізації 035.041 германські мови
та літератури (переклад включно),
перша – англійська
освітньо-професійної програми
Переклад (англійський)
Давиденко Денис Володимирович

Керівник к.ф.н. Погонець В. В.

Рецензент к.ф.н., доц. Запольских С. П.

Запоріжжя – 2023

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет іноземної філології

Кафедра теорії та практики перекладу з англійської мови

Освітній рівень магістр

Спеціальність 035 Філологія

Спеціалізація 035.041 Германські мови та літератури (переклад включно) –
перша англійська

Освітньо–професійна програма Переклад (англійський)

ЗАТВЕРДЖУЮ

**Завідувач кафедри теорії та практики
перекладу з англійської мови**

Запольських С.П. _____

«__» _____ 2023 року

З А В Д А Н Н Я
НА КВАЛІФІКАЦІЙНУ РОБОТУ МАГІСТРА
ДАВИДЕНКО ДЕНИСУ ВОЛОДИМИРОВИЧУ

(прізвище, ім'я, по батькові)

1. Тема кваліфікаційної роботи магістра (проекту) «Специфіка машинного перекладу (на матеріалі публіцистики англійською та українською мовами)» керівник кваліфікаційної роботи Керівник Погонєць В.В., доцент кафедри теорії та практики перекладу з англійської мови, к.ф.н.

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом ЗНУ від «11» квітня 2023 року № 521-с

2. Строк подання студентом кваліфікаційної роботи (проекту) 30.11.2023 р.

3. Вихідні дані до кваліфікаційної роботи (проекту):

теоретичні засади машинного перекладу; системи машинного перекладу Meta, Google, DeepL; актуальні статті BBC про війну в Україні.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)

1) здійснити огляд теоретичних джерел щодо термінів «машинний переклад» та «переклад» в цілому; 2) описати розвиток машинного перекладу; 3) навести порівняльну характеристику зроблених перекладів; 4) проаналізувати типові помилки МП на основі відібраних текстів

5. Консультанти розділів кваліфікаційної роботи (проекту)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Вступ	Погонець В.В., доц.	09.06.2023	09.06.2023
Розділ 1	Погонець В.В., доц.	02.09.2023	02.09.2023
Розділ 2	Погонець В.В., доц.	01.10.2023	01.10.2023
Висновки	Погонець В.В., доц.	20.10.2023	20.10.2023


6. Дата видачі завдання 09.06.2023 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів виконання кваліфікаційної роботи магістра	Строк виконання етапів роботи (проекту)	Примітка
1	Пошук наукових джерел з теми дослідження, їх вивчення та аналіз; укладання бібліографії	лютий – квітень 2023	виконано
2	Добір фактичного матеріалу	травень 2023	виконано
3	Написання вступу	червень 2023	виконано
4	Написання теоретичного розділу	вересень 2023	виконано
5	Написання практичного розділу	жовтень 2023	виконано
6	Формулювання висновків	жовтень 2023	виконано
7	Проходження нормоконтролю	листопад 2023	виконано
8	Одержання відгуку та рецензії	листопад–грудень 2023	виконано
9	Захист	грудень 2023	виконано

Автор роботи несе персональну відповідальність за відсутність в роботі несанкціонованих текстових запозичень (академічного плагіату)

Магістрант



(підпис)

Д. В. Давиденко

(ініціали та прізвище)

Керівник роботи

(підпис)

В. В. Погонець

(ініціали та прізвище)

Нормоконтроль пройдено

Нормоконтролер

(підпис)

В. В. Погонець

(ініціали та прізвище)

РЕФЕРАТ

Дипломна робота – 55 стор., 59 джерел

Об’єкт дослідження: переклад, який отримується при використанні сучасних загальнодоступних систем машинного перекладу таких як Google Translate, DeepL.

Мета роботи: тлумачення терміну «машинний переклад» і порівняльний розбір перекладів зроблених за допомогою сервісів перекладу.

Теоретико-методологічні засади: як *теоретичні*, робота тісно пов’язана з існуючими *класифікаціями* помилок машинного перекладу, так і *емпіричні* методи, в роботі проведений наочний *експеримент*, де відібрані тексти перекладені системами машинного перекладу і проаналізовані.

Отримані результати: Отримані дані від дослідження свідчать про те, що стан сучасних систем машинного перекладу, таких як Google Translate, META, DeepL, є недосконалим, має як переваги так і певні недоліки. Аналіз проведених перекладів вказує на наявність значних помилок у різних аспектах, таких як пропущення змістових слів, неправильний вибір слів, граматичні та стилістичні недоліки.

Переклад, виконаний META Translate, виявив значну кількість помилок в різних категоріях, таких як пропущення змістових слів та проблеми з граматиною та лексикою.

Переклад Google Translate має деякі помилки, такі як пропущені слова, неправильний вибір слів та стилістичні помилки. Деякі з них можуть впливати на граматичну правильність та смислову цілісність речення.

У системі DeepL були виявлені проблеми з вибором слів відповідно до контексту та помилки у виборі форм слів. Проте попри це система показала найкращий результат з точки зору правильності та точності перекладу, припустивши меншу кількість помилок ніж попередні системи.

Ключові слова: переклад, машинний переклад, системи машинного перекладу, нейромережі, публіцистичний текст

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАК.....	3
ВСТУП.....	4
РОЗДІЛ 1	
ТЕОРЕТИЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ МАШИННОГО ПЕРЕКЛАДУ.....	8
1.1 Визначення понять переклад, машинний переклад та публіцистичний текст.....	8
1.2 Історія машинного перекладу.....	15
1.3 Проблеми машинного перекладу.....	20
РОЗДІЛ 2	
ОСОБЛИВОСТІ ПЕРЕКЛАДУ ПУБЛІЦИСТИЧНИХ ТЕКСТІВ СУЧАСНИМИ СИСТЕМАМИ МП.....	27
2.1 Методи оцінювання якості машинного перекладу.....	27
2.2 Класифікація і аналіз помилок МП на матеріалі відібраних текстів.....	36
ВИСНОВКИ.....	53
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	55

ПЕРЕЛІК УМОВНИХ ПОЗНАК

ВЕЛМ - велика електронна лічильна машина
ЕОМ - електронно-обчислювальна машина
МП - машинний переклад
НМП - нейронний машинний переклад
СМП - системи машинного перекладу
ССМП - системи статистичного машинного перекладу
ALPAC - Automatic Language Processing Advisory Committee
BLEU - Bilingual Evaluation Understudy
IBM - the International Business Machines Corporation
PER - Position-independent Error Rate
SER - Sentence Error Rate
WER - Word Error Rate

ВСТУП

У сучасному інформаційному суспільстві, де межі спілкування розширюються, а глобалізаційні процеси взаємодії між культурами стають все інтенсивнішими, важливим завданням стає подолання мовних бар'єрів. Одним із ефективних інструментів вирішення цього завдання є машинний переклад, який забезпечує автоматизовану конвертацію текстів з однієї мови на іншу.

У зв'язку із постійним розвитком технологій штучного інтелекту, включаючи нейромережі та глибоке навчання, в галузі машинного перекладу відбуваються значні зміни. Сучасні системи машинного перекладу, побудовані на основі глибоких нейронних мереж, демонструють вражаючі результати у порівнянні з традиційними методами.

Нові алгоритми та моделі глибокого навчання дозволяють системам машинного перекладу краще розуміти контекст, ідіоми та відтінки мови. Завдяки цьому вони можуть забезпечити більш точний та природний переклад, що стає особливо важливим у сучасному інформаційному середовищі.

Зокрема, розробка моделей, які враховують культурні відмінності та специфіку різних мов, стає тенденцією у сфері машинного перекладу. Це допомагає уникнути неточностей та забезпечує краще розуміння сутності тексту.

Із появою нових технологій важливість досліджень у галузі машинного перекладу підвищується, сприяючи подальшій оптимізації процесів перекладу та підвищенню якості отриманих результатів.

Спостерігається тенденція до визнання перекладача, а також майбутніх фахівців у галузі, як комп'ютерних експертів. Вони мають здатність використовувати не лише традиційні ресурси, але й сучасні технології для ефективного виконання перекладацьких завдань.

Наголос частіше ставиться на новому напрямку у вивченні перекладу, який підкреслює нерозривний зв'язок між новітніми технологіями і самим

процесом перекладу. Це відображає актуальність використання передових інструментів та програм у сучасній перекладацькій практиці.

Однією з ключових вимог є вміння перекладача не лише володіти майстерно рідною та іноземною мовами та глибоко розуміти предметну галузь, але також ефективно використовувати сучасні інформаційні технології для оптимізації процесу перекладу.

Відповідно до цих змін, перекладач повинен бути активно орієнтованим у світі інновацій та засобів автоматизації, що вимагає від нього постійного оновлення та удосконалення своїх навичок у контексті сучасних технологічних тенденцій.

Особливо актуальною стає проблема машинного перекладу в контексті публіцистики, оскільки цей жанр вимагає особливої точності. Публіцистичні тексти містять важливу інформацію про події, тенденції та проблеми, що відбуваються у світі, і забезпечують розуміння та осмислення подій з боку читача.

Зважаючи на значний обсяг публіцистичних матеріалів, з якими стикаються як перекладачі, так і журналісти, виникає необхідність у використанні сучасних технологій, зокрема машинного перекладу, для полегшення та оптимізації процесу перекладу публіцистичних текстів.

Актуальність даної роботи викликана потребою розглянути можливості сучасних програм-перекладачів, їх перспектив, та ролі у роботі перекладачів.

Метою даної роботи є тлумачення терміну «машинний переклад» і порівняльний розбір перекладів зроблених за допомогою сервісів перекладу.

Для досягнення даної мети, треба виконати наступні **завдання**:

1. здійснити огляд теоретичних джерел щодо термінів «машинний переклад» та «переклад» в цілому;
2. описати розвиток машинного перекладу;
3. навести порівняльну характеристику зроблених перекладів;
4. проаналізувати типові помилки МП на основі відібраних текстів.

Об'єктом дослідження є переклад, який отримується при використанні сучасних загальнодоступних систем машинного перекладу таких як Meta, Google Translate, Deepl, а **предметом** - особливості машинного перекладу.

Матеріалом дослідження стали новинні статті телерадіомовної корпорації BBC написані українською мовою та перекладені на англійську мову системами машинного перекладу Meta, Google і Deepl.

Із обраних **методів** дослідження можна виділити як *теоретичні*, робота буде тісно пов'язана з існуючими *класифікаціями* помилок машинного перекладу, так і *емпіричні* методи, в роботі буде проведений наочний *експеримент*, де відібрані тексти будуть перекладені системами машинного перекладу і згодом проаналізовані.

Практична значущість дослідження полягає у можливості використання його результатів при удосконаленні існуючих та розробці нових систем машинного перекладу та під час проведення практичних занять з практики перекладу англійської мови.

Робота пройшла **апробацію** на науково-практичній студентській конференції. Результати дослідження представлено у публікації:

Давиденко Д. В. Специфіка машинного перекладу на матеріалі публіцистичних текстів. *Збірник наукових праць студентів, аспірантів і молодих вчених «МОЛОДА НАУКА-2023» : Актуальні проблеми іноземної філології та перекладознавства* : Запоріжжя : ЗНУ, 2023. 319 с.

Структура роботи: дослідження складається зі вступу, двох розділів, висновків та списку використаних джерел.

У вступі представлена тема дослідження, його мета, актуальність, завдання, об'єкт, предмет, матеріал, методи дослідження, його практична значущість та апробація.

У першому розділі проаналізовані теоретичні засади дослідження машинного перекладу. Розібрані такі поняття як «переклад», «машинний переклад» та «публіцистичний текст»; розглянута історія розвитку машинного перекладу; розглянуті проблеми машинного перекладу.

У другому розділі розглянуті та проаналізовані особливості перекладу публіцистичних текстів сучасними системами машинного перекладу. Розглянуті методи оцінювання якості машинного перекладу; класифіковано та проаналізовано помилки машинного перекладу на матеріалі відібраних текстів.

У висновках подано узагальнені результати проведеної роботи.

Загальна кількість сторінок 55, кількість використаних джерел 59.

РОЗДІЛ 1

ТЕОРЕТИЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ МАШИННОГО ПЕРЕКЛАДУ

1.1 Визначення понять переклад, машинний переклад та публіцистичний текст

Питання перекладу та його ролі у сучасному світі завжди були предметом інтересу для вчених, лінгвістів та фахівців з мовознавства. Переклад - це складний і багатогранний процес передачі змісту та смислу тексту з однієї мови на іншу. Він вимагає вміння докладно розуміти культурні особливості обох мов, а також майстерно володіти граматиною, лексикою та стилістикою кожної з них.

У цьому розділі буде детально проаналізовано поняття «переклад», з'ясовано його сутність та ключові аспекти, що характеризують цей складний процес; основні аспекти машинного перекладу, його розвиток та поточний стан, а також переваги та недоліки цього підходу; буде розглянута також і сутність та особливості публіцистичного тексту, його роль у суспільному контексті та необхідність в контексті перекладу.

Всім відоме та загальноприйняте слово «переклад» зустрічається практично в будь-якій мові і використовується в низці різних контекстів. Великий обсяг різноманітних вживань слова «переклад» вказує на його універсальність та важливість в різних видах мовленнєвої діяльності. Від перекладу творів до технічних перекладів, це слово вплетено у різні аспекти культури та комунікації. Академічний Словник української мови в 11 томах дає наступні визначення слова «переклад»:

1) процес людської діяльності, спрямований на передачу інформації з мови-оригіналу на мову-перекладу;

2) результат перекладацького процесу, тобто твір, речення і т. ін., викладений усно або на папері;

3) вид людської діяльності, спрямований на відтворення певного тексту мови-оригіналу в мові-перекладу;

4) це умовна назва навчальних дисциплін, які вивчають особливості перекладацької діяльності, техніку виконання перекладу, процес перекладу, його види, способи, жанри та ін., результат діяльності перекладача [Словник української мови 1970–1980, с. 196].

Процес створення тексту на одній мові, який, в комунікативному відношенні, є еквівалентним вхідному тексту на іншій мові, формально описується як переклад. Це включає в себе не лише механічне перенесення слів, але і передачу змісту та емоційної нюансовості між двома мовами.

Комунікативна еквівалентність в даному контексті визначається як якість перекладу, яка дозволяє йому повністю замінити оригінальний текст у процесі спілкування між носіями різних мов. У своїй суті, переклад визначається як творчий процес, що вимагає від перекладача не лише майстерності у володінні двома мовами, але й глибокого розуміння культурних та соціальних контекстів, що впливають на сприйняття тексту.

Забезпечення комунікативної еквівалентності нового тексту відносно оригіналу вимагає виконання трьох основних вимог:

- текст перекладу повинен не лише передавати зміст оригіналу, але й робити це без будь-яких упущень або додаткових деталей, забезпечуючи повноту і точність інформаційного висловлення;
- текст перекладу також повинен дотримуватися мовних норм перекладу, що запобігає спотворенню або збою у сприйнятті інформації. Це означає, що перекладач повинен уникати вільного тлумачення та враховувати всі аспекти обох мов;
- крім того, обсяг тексту перекладу має бути приблизно рівним оригіналу. Це важливо для досягнення схожого стилістичного ефекту, зберігаючи лаконічність або розгорнутість виразності. Відповідне відтворення

обсягу допомагає зберігати емоційний та інформаційний зміст оригіналу без втрати в контексті перекладу [Мацак, Скоробогатова].

Процес перекладу є складним завданням, що вимагає від перекладача не лише знання двох мов, але й вміння передати смисл та контекст оригіналу з урахуванням нюансів та особливостей культур мови перекладу. Це означає, що перекладач повинен бути не лише лінгвістом, але й культурологом, що розуміє контекст і виразності обох мов.

Через постійний розвиток технологій, машинний переклад стає все більш поширеним і допомагає спростити деякі етапи процесу перекладу. Технологічні досягнення в цій галузі дозволяють автоматизувати рутинні завдання та швидко перекладати текст, але важливо враховувати, що машина не завжди здатна вірно відтворити всі відтінки смислу та стилістичні особливості оригіналу.

Однак, необхідність у кваліфікованих людських перекладачах залишається невід'ємною для досягнення високої якості комунікативної еквівалентності в перекладі публіцистичних текстів. Людина здатна враховувати відтінки мовлення, емоції автора та адаптувати текст таким чином, щоб він був адекватним та зрозумілим для читачів мови перекладу.

Такий синергетичний підхід, поєднуючи технологічні можливості машинного перекладу і людську експертизу, може допомогти в створенні якісних та точних перекладів, що відповідають високим стандартам комунікативної еквівалентності.

Завдяки стрімкому розвитку інформаційних технологій та штучного інтелекту, в останні десятиліття з'явився новий підхід до перекладу - машинний переклад. Машинний переклад - це використання спеціалізованих програмних засобів та штучного інтелекту для автоматичної перетворення тексту з однієї мови на іншу.

Машинний переклад відрізняється від традиційного перекладу, оскільки він базується на використанні комп'ютера. При графічному представленні інформації комп'ютер розпізнає мовні одиниці за їх формою, тоді як усне

сприйняття відбувається з урахуванням фізичних або акустичних характеристик мовних одиниць [Карпіловська 2006, с. 153].

Системи машинного перекладу (СМП) дають змогу швидко звертатись до великого обсягу інформації, перекладаючи її на іноземні мови. Це важливий інструмент для глобального спілкування та обміну ідеями, забезпечуючи швидку та доступну можливість зрозуміти текст на різних мовах.

Зокрема, промислові СМП активно використовуються для перекладу науково-технічних текстів, і це досягається за допомогою термінологічних баз даних, що забезпечують єдність термінології та професійної лексики. Це дозволяє забезпечувати точність та стабільність перекладу, зокрема в галузях, де важлива правильна інтерпретація термінів та спеціалізованої лексики.

Сучасні програми машинного перекладу вражають своєю швидкістю та продуктивністю, дозволяючи перекладати значні обсяги інформації за короткий проміжок часу. Це стає особливо важливим у сучасному світі, де швидкість обміну інформацією визначає ефективність ділового та наукового спілкування.

На першому етапі розвитку машинного перекладу застосовувався метод «кодування-декодування» (encoding-decoding approach), що відомий також як прямий метод машинного перекладу. Цей метод вимагав від перекладачів знання математики та програмування. Він базувався на послідовному перекладі, коли здійснювався переклад лише лексичних одиниць [Шилінська 2022, с. 150; Мірам 1998].

Проте, зростання обчислювальної потужності та розвиток алгоритмів дозволили суттєво поліпшити якість машинного перекладу. Сьогодні використовуються більш складні підходи, такі як статистичний переклад та нейронні мережі, що дозволяють досягти більш точних та зрозумілих перекладів.

Програма перекладу за допомогою мови-посередника проводить морфологічний, синтаксичний та семантичний аналіз тексту оригіналу. Далі, вона представляє текст оригіналу в опорній мові, використовуючи модуль

мови-посередника та перетворює його в текст мови перекладу з використанням семантичних, синтаксичних, лексичних та морфологічних даних словників мови перекладу [Янковець 2017].

Статистичний машинний переклад базується на порівнянні двох мовних версій та заснований на прикладах. Він використовує статистичні методи для підбору найбільш відповідних перекладів з великої кількості прикладів текстів, що допомагає покращити якість машинного перекладу.

Чим більше мовних версій доступно, тим вища якість перекладу може бути досягнута. Наявність багатомовної бази даних дозволяє комп'ютерній системі машинного перекладу враховувати різноманітність мовних конструкцій та нюансів, підвищуючи точність та природність перекладу.

У такому випадку комп'ютер не оперує лінгвістичними алгоритмами, а здійснює оцінку ймовірності використання певного слова або виразу. Іншими словами, його функція полягає у пошуку найбільш імовірного перекладу речення засобами, що ґрунтуються на існуючій базі даних двомовних корпусів текстів [Шилінська 2022, с. 150].

Машинний переклад може бути корисним інструментом для перекладачів у їх професійній діяльності. Він відкриває нові можливості для перекладачів у вигляді економії часу та ресурсів.

Однією з основних цілей цього методу перекладу є здобуття еквівалентів термінологічних частин оригінального тексту, що робить його цінним інструментом для першого етапу перекладу. Важливим аспектом є забезпечення семантичної та синтаксичної відповідності, що вимагає аналізу перекладів на різних рівнях мови [Гудманян 2019, с. 50].

Машинний переклад використовується тоді, коли виникає потреба у вирішенні питання щодо необхідності подальшої роботи з текстом. У таких випадках головною метою перекладу є передача основного комунікативного змісту оригіналу та надання отримувачу докладнішої або менш деталізованої інформації змісту повідомлення, що міститься в початковому джерелі [Гудманян 2020, с. 144; Бірюков 2008].

Машинний переклад відіграє значну роль у спілкуванні між людьми, які розмовляють різними мовами. Важливість машинного перекладу полягає в наступних аспектах:

Зближення культур: машинний переклад дозволяє людям з різних країн та культур спілкуватися, обмінюватися інформацією та розуміти один одного. Він розширює можливості комунікації та сприяє культурній обміну.

Ефективність та швидкість: машинний переклад значно прискорює процес перекладу текстів, дозволяючи людям швидко отримувати переклади іноземних матеріалів. Це особливо важливо в глобалізованому світі, де швидкий доступ до інформації є критичним.

Бізнес та торгівля: машинний переклад відіграє важливу роль у міжнародному бізнесі та торгівлі. Він допомагає підприємствам спілкуватися з партнерами з різних країн, перекладати контракти, технічну документацію, веб-сайти та інші матеріали.

Освіта та наука: машинний переклад має великий потенціал у галузі освіти та науки. Він допомагає студентам та дослідникам отримувати доступ до літератури та матеріалів, що були б недоступні через бар'єр мови. Він також допомагає сприяти обміну ідеями та співпраці між науковцями з різних країн.

Подорожі та туризм: машинний переклад дозволяє туристам спілкуватися та розуміти мову країни, в яку вони подорожують. Це полегшує комунікацію з місцевими жителями, отримання інформації про місцеві визначні місця та послуги.

Необхідно зазначити, що машинний переклад не є ідеальним і може мати певні обмеження, особливо при перекладі складних або контекстуально залежних текстів. Однак, з постійним розвитком технологій машинного навчання та нейромереж, машинний переклад стає все більш точним та надійним.

Публіцистика є особливим жанром, спрямованим на інформування та вплив на аудиторію. Публіцистичний текст - це жанр письмового висловлювання, що передає авторську позицію щодо актуальних подій,

проблем та явищ у суспільстві. Він може містити аналітичні статті, коментарі, репортажі, інтерв'ю та інші жанрові форми, спрямовані на підтримку дискусії та розвиток громадської думки.

При визначенні поняття «публіцистичний текст», важливо враховувати, що цей термін відноситься до категорії «текст» як рід і вид, тому публіцистичний текст має свої особливі текстові функції. Публіцистичні тексти спрямовані на інформування громадськості, надаючи їм змогу отримати інформацію про актуальні події, проблеми та питання, що важливі для суспільства [Щепка 2018, с. 103].

Крім того, публіцистичний стиль може використовуватися в різних медіа форматах, таких як статті, інтерв'ю, коментарі, аналітичні матеріали тощо. Це робить його універсальним інструментом для передачі інформації в різних жанрах та формах, адаптуючись до вимог конкретного медіа.

Важливим аспектом публіцистичного стилю є збалансованість і об'єктивність. Автор повинен представляти різні точки зору та аргументи, дотримуючись журналістських принципів об'єктивності та достовірності, щоб створити повну та об'єктивну картину подій.

Крім того, ефективне використання мовних засобів в публіцистичному стилі важливо для досягнення максимального впливу на аудиторію. Здатність точно та емоційно висловлювати думки та ідеї допомагає створювати сильне враження та впливати на громадську думку.

«Оціночна програма - це, по суті, провідна програма газетно-публіцистичного тексту. Той чи інший факт цікавить журналіста не сам по собі, а в аспекті соціальних оцінок. У багатьох жанрах ці оцінки даються як особистісні авторські оцінки, що включають в себе емоційну забарвленість» [Акулов 2020 с. 17; Максимов 2006, с. 84].

Слово «публіцистичність» відноситься до виразної авторської позиції, яка може бути розглянута як одна з форм прояву авторської тенденційності, що також включає поняття «публіцистика». «Публіцистика - це література з

актуальних суспільно - політичних питань сучасності, поточного життя суспільства» [Місуно 2009, с. 629].

Основна комунікативно-прагматична функція журналістських текстів полягає в пропагандистсько-агітаційному впливі, що має на меті поширення суспільно-політичної інформації та мотивацію читачів до дії, стимулюючи їхні думки та дії [Коваленко 2001].

1.2 Історія машинного перекладу

У сучасному світі, де інформаційні потоки постійно зростають і мають безмежний характер у часі та просторі, переклад має дедалі більше значення. Він здійснюється різними способами, такими як письмовий, усний, синхронний, послідовний, а також використанням машинного перекладу, який став активно використовуватися з початку 1980-х років [Чернікова 2013, с. 165; Кісіль 2011].

У XVII столітті було висловлено ідеї про механізацію перекладу, але реальні можливості з'явилися лише в 20 столітті. Цей період відзначався лише теоретичними концепціями, а реальні технічні можливості з'явилися значно пізніше.

У 1930-х роках франко-вірменський винахідник Жорж Арцруні та російський винахідник Петро Троянський подали заявки на патенти на «перекладацькі машини». З них винахід Троянського виділявся, оскільки він запропонував не лише метод перекладу, але й схему автоматичного двомовного словника, кодування міжмовних граматичних ролей і схему аналізу та синтезу. Ці ідеї визнали за інноваційні, проте вони залишилися невідомими до пізніших років.

Ідеї Троянського стали відомими лише в кінці 1950-х років, коли з'явилася перша електронна обчислювальна машина. Цей прогрес у технологіях створив фундамент для подальшого розвитку машинного перекладу і визначив новий етап в історії автоматизованого перекладу [Hutchins 2014, с. 1].

У січні 1954 року відбувся експеримент Джорджтаунського університету та ІВМ, під час якого вдалося вперше здійснити обмежений машинний переклад з російської на англійську мову. Експеримент став важливим кроком у дослідженнях у галузі машинного перекладу, оскільки вперше продемонстрував можливість перекладу між різними мовами за допомогою електронної обчислювальної машини, хоча й зазнав критики.

На основі результатів цього експерименту були сформульовані ключові концепції, які мали велике значення для подальших досліджень у цій області. Варто відзначити, що цей експеримент привернув великий інтерес до загальної проблеми машинного перекладу.

У 1956 році Інститут точної механіки та обчислювальної техніки Академії наук СРСР заявив про успішний переклад англійської мови на російську за допомогою своєї ЕОМ «ВЕЛІМ» і визнав зв'язок між своєю роботою та експериментом Джорджтаун-ІВМ.

У червні 1956 року Джорджтаунський університет отримав значний грант від Національного наукового фонду для проведення інтенсивних досліджень з перекладу російських наукових матеріалів на англійську мову.

Грант було подовжено на другий рік, що свідчило про успіх та важливість проведених досліджень. Додатковий час дозволив збільшити обсяг та глибину аналізу перекладу, а також зосередити увагу на подальшому вдосконаленні технологій машинного перекладу [Dostert 1957, с. 4].

У 1964 році був створений комітет ALPAC (Automatic Language Processing Advisory Committee) за ініціативою американського уряду з метою дослідження перспектив та можливостей машинного перекладу. Це було важливою подією для розвитку цієї галузі.

Двома роками пізніше, у 1966 році, ALPAC опублікував свою відому доповідь, в якій висловив думку про те, що роботи в галузі машинного перекладу є марнотратством часу і грошей. Доповідь стала визначальною подією, яка суттєво вплинула на сприйняття машинного перекладу в науковому співтоваристві.

Доповідь ALPAC викликала сумніви щодо ефективності машинного перекладу та викликала обговорення його недоліків. Висловлені в ній тези призвели до переосмислення та переоцінки досягнень у цій галузі, а також до зменшення інвестицій та розвитку нових проєктів [Chéragui 2012, с.161].

Розвиток перекладу в 1970-х роках рухався у трьох ключових напрямках: розробка комп'ютерних інструментів для перекладачів, створення операційних систем машинного перекладу з різноманітною підтримкою для людини та проведення фундаментальних теоретичних досліджень для вдосконалення методів машинного перекладу.

Поява інтерактивних комп'ютерних середовищ у реальному часі з 1960-х років і подальший розвиток текстових редакторів у 1970-х та мікрокомп'ютерів у 1980-х надали можливість розробляти перекладацькі інструменти і надалі. Розвиток перекладацьких інструментів на основі інтерактивних комп'ютерних середовищ та текстових редакторів сприяв створенню перекладацьких платформ у подальшому.

Важливо зауважити, що розробка перекладацьких інструментів, хоча й відбувалася в неакадемічному середовищі, не виявилася настільки складною задачею з точки зору комп'ютерної лінгвістики, а тому це відкрило нові можливості для перекладачів та дослідників цієї галузі [Hutchins 1995].

В 1980-х роках, сфера машинного перекладу відзначилася швидким розвитком на різних фронтах. Нові операційні системи, які з'явилися в цей період, створили комфортні умови для інновацій та вдосконалення в галузі машинного перекладу.

Спостерігалось значне зростання нових операційних систем, і ринок комерційних МПсистем поширився на всі види перекладу. Дослідження в

сфері МП також стали диверсифікованішими, охоплюючи різні напрямки. Це свідчило про широкий інтерес та попит на автоматизовані засоби перекладу як у науковому, так і у комерційному використанні.

Найбільш активний розвиток у сфері машинного перекладу відбувся в Японії, де комп'ютерні компанії стали відзначатися високим рівнем зацікавленості та ефективною роботою над новаторським програмним забезпеченням для автоматизованого перекладу.

Більшість комп'ютерних компаній в країні активно працювали над розробкою програмного забезпечення для автоматизованого перекладу, переважно спрямованого на ринки японсько-англійського та англо-японського перекладу [Hutchins 1995].

У 1990-х роках вчені з ІВМ вирішили повернутися до досліджень у сфері машинного перекладу і в 1991 році представили одну з перших систем статистичного машинного перекладу (ССМП). Це визначалося прагненням до новаторства та пошуку ефективних методів в галузі автоматизованого перекладу.

Технологія статистичного машинного перекладу, яку вчені ІВМ представили у 1991 році, відрізнялася від попередніх підходів до машинного перекладу. Основна особливість полягала в тому, що ця система не потребувала ручного введення правил для кожної мови, що визначало її прогресивність та ефективність.

ССМП використовували двомовний корпус текстів для виявлення закономірностей між мовами, які можна було використовувати для створення статистичних моделей перекладу.

З поширенням електронних комп'ютерів, які стали неодмінною частиною повсякденного життя, системи машинного перекладу почали активно розвиватися, адаптуючись до потреб користувачів у сучасному інформаційному середовищі.

У 1997 році компанія SYSTRAN вперше вивела на ринок веб-інструмент для машинного перекладу, відкривши доступ до цієї технології не лише

фахівцям, але й звичайним користувачам, які не мали спеціальних знань у галузі мовознавства. Це позначилося на розширенні кола користувачів та популяризації машинного перекладу серед широкого загалу.

Практично через десять років, у 2006 році, компанія Google випустила свій власний інструмент машинного перекладу - Google Translate. Сервіс, базований на статистичному машинному перекладі, відкрив нові можливості для користувачів і швидко завоював популярність, стаючи однією з ключових платформ перекладу в період з 2007 до 2016 року [Jakubková 2023].

На сьогоднішній день, не зважаючи на значний прогрес, досягнутий у галузі машинного перекладу, ця технологія все ще далека від ідеальності та не може замінити якість роботи професійного перекладача, оскільки переваги людського фахівця полягають у глибокому розумінні культурних, ідіоматичних та контекстуальних нюансів мови.

Це твердження отримало переконливе підтвердження під час великого змагання, організованого Кібернетичним університетом Седжона та Міжнародною асоціацією усних і письмових перекладачів Кореї в лютому 2017 року. У змаганні взяли участь чотири професійних перекладачі та три програми машинного перекладу, зокрема Google Translate, перекладацька програма Systran та додаток Papago від Naver.

При порівнянні результатів змагання виявилось, що хоча машини зуміли перекласти чотири тестові документи швидше, професійні перекладачі демонстрували значно вищий рівень точності та глибини розуміння тексту, отримавши 49 балів із 60 можливих.

Особливу увагу звернули на той факт, що люди вирізнялися у передачі виразності мови, логіки та організації тексту, що визначило важливість людського фактору в процесі перекладу. Найвищий бал серед машинних перекладачів, отриманий Google Translate, становив лише 28 балів, що свідчить про те, що навіть найбільш розвинені програми машинного перекладу поки що не можуть конкурувати з ефективністю та точністю людських фахівців [Richard, 2020].

1.3 Проблеми машинного перекладу

Розвиток технологій машинного перекладу приніс значні досягнення, але водночас відкрив ряд невирішених проблем, які впливають на якість та точність перекладу текстів. Серед цих викликів виділяються проблеми, які стосуються не лише точності перекладу слів і фраз, але й здатності передавати контекст, інтонацію та авторський стиль.

На початку 20-го століття стало очевидним, що повністю автоматизований і високоякісний переклад є недосяжною метою, і необхідне як передперекладове, так і післяперекладове редагування текстів [Hutchins 2010].

Серед основних проблем машинного перекладу, які були відзначені Достертом, важливо відзначити те, що одне слово мови оригіналу може мати кілька можливих еквівалентів у мові перекладу, що створює труднощі у виборі найбільш точного перекладу. Крім того, порядок речень або одиниць мови в оригінальному тексті може не відповідати тому, що є оптимальним для передачі на мову перекладу, що може викликати непорозуміння або втрату сенсу в перекладі [Dostert 1957].

Ці проблеми викликають необхідність вдосконалення алгоритмів та підходів до машинного перекладу, а також підкреслюють важливість ролі людського редактора і перекладача для досягнення високої якості перекладу. Взаємодія між технологіями та професійними знаннями є ключовим фактором для ефективного розвитку перекладацької галузі.

Кузьменко у своєму дослідженні зазначає проблеми, пов'язані з граматичними та синтаксичними зв'язками у реченні, а також з відтворенням термінології. Оскільки художня література має образний характер та стилістичне навантаження, машинний переклад таких текстів є неможливим завданням [Кузьменко 2009].

В основному, машинний переклад застосовується для автоматичного перекладу різних видів документації, такої як науково-технічні матеріали,

офіційні інструкції, листи електронної пошти та веб-сторінки, щоб забезпечити доступність інформації для аудиторії різних мов.

Також, використання машинного перекладу може великою мірою спростити рутинну роботу перекладачів, які можуть використовувати автоматизовані інструменти для створення «чернетки» перекладу. Ця «чернетка» є основою, яку можна легко відредагувати та удосконалити, враховуючи контекст та вимоги замовника [Смірнова 2011; Кузьменко 2009].

У процесі оцінки перекладацьких помилок, які виникають через можливі проблеми, важливо враховувати їх кількість та ступінь серйозності. Зокрема, мовні помилки можуть бути оцінені за наступною системою:

- Переклад демонструє майстерне володіння граматиною, правописом та пунктуацією мови перекладу. Декілька помилок або помилки відсутні взагалі.
- Переклад демонструє вміле володіння граматиною, правописом та пунктуацією мови перекладу. Невелика кількість незначних помилок.
- Переклад демонструє слабке володіння граматиною, правописом та пунктуацією мови перекладу. Велика кількість незначних помилок.
- Переклад демонструє часткову відсутність навичок володіння граматиною, правописом та пунктуацією мови перекладу. Велика кількість помилок.
- Переклад демонструє повну відсутність навичок володіння граматиною, правописом та пунктуацією мови перекладу. Велика кількість серйозних помилок [Гудманян 2020, с.148].

Сучасні системи машинного перекладу стають залежними від точності та стандартизованості вхідного тексту. Особливості цих систем полягають у розпізнаванні лише стандартних слів та правильно написаних текстів за допомогою правил та лексичних баз даних.

Навіть у випадку стандартизованих текстів, системам машинного перекладу не уникнути помилок, особливо коли мови мають структурні

відмінності. Завдяки цим системам можна зрозуміти суть тексту, але важко очікувати ідеально правильний переклад.

Особливі виклики виникають з нестандартними текстами, такими як електронні листи, де може бути велика різноманітність стилів та форм висловлення, а також можливість включення неофіційних слів та висловів. Користувачі, що використовують дві мови, можуть змішувати їх, цитувати попередні статті, використовувати слова з інших мов та допускати орфографічні та граматичні помилки через мовну інтерференцію [Climent 2003].

Після проведення аналізу різних проблем, пов'язаних із нейронним машинним перекладом (НМП) та порівняння його з традиційним статистичним машинним перекладом, Філіп Коен та Ребекка Ноулз з факультету комп'ютерних наук Університету Джона Гопкінса прийшли до висновку, що:

- НМП-системи мають меншу якість перекладу поза предметною областю, іноді жертвуючи адекватністю на користь плавності перекладу.
- Навчання НМП вимагає більшого обсягу навчальних даних, що призводить до гіршої якості в умовах з обмеженими ресурсами, але кращої продуктивності з високими ресурсами.
- Використання НМП-систем на рівні підслівників (наприклад, з кодуванням попарно) поліпшує переклад низькочастотних слів, але все ще показує слабкість у перекладі слів з високим ступенем флексії, таких як дієслова.
- НМП має нижчу якість перекладу дуже довгих речень, але краще справляється з реченнями довжиною до 60 слів.
- Модель уваги для НМП не завжди виконує роль моделі вирівнювання слів, але фактично може значно відрізнятись.
- Застосування декодування з пошуком по променю поліпшує якість перекладу тільки при використанні вузьких променів, але якість знижується при збільшенні простору пошуку [Koehn 2017, с.28].

Вивчаючи різні підходи до машинного перекладу, Маргарет Думебі Окпор прийшла до висновку, що цей процес є вкрай складним завданням, в основному через велику складність природних мов. Причина складнощів полягає в тому, що слова в мові можуть мати багато значень та потенційних перекладів, що ускладнює точне відтворення сенсу в іншій мові.

Речення можуть тлумачитися різними способами, і іноді правильний переклад вимагає глибокого розуміння контексту та зв'язків між різними частинами тексту. Окрім того, для вдалого перекладу іноді потрібні широкі знання про світ, що ускладнює завдання систем машинного перекладу.

З великою кількістю лінгвістичних залежностей і їхньою складністю важко врахувати всі аспекти в системі машинного перекладу. У результаті система машинного перекладу змушена працювати з неповними знаннями, що може впливати на точність та повноту перекладу.

Існує декілька підходів до розв'язання цієї проблеми, проте жоден із них не є ідеальним. Дослідження показали, що деякі підходи мають мінімальні проблеми, а універсальне рішення відсутнє. Тому комбінування найкращих аспектів декількох підходів для створення гібридного підходу може допомогти вирішити проблеми, пов'язані з машинним перекладом [Okpor 2014, с.165].

Спираючись на працю про лінгвістичні проблеми, пов'язані з машинним перекладом Єгошуа Бар-Гіллеля, можна розглянути чотири основні проблеми, пов'язані з машинним перекладом з лінгвістичної точки зору. Розглянемо кожна з них:

Операційний синтаксис: Одним із найважливіших кроків у машинному перекладі є визначення синтаксичної структури речення мовою оригіналу. Цей процес вимагає чіткості і є складним завданням для лінгвістів-структуралістів. Якщо програма перекладу покладається на розуміння тексту машиною, то відсутність семантичного органу в сучасних машинах може перешкоджати процесу перекладу. Бар-Гіллель припускає, що розробка операційного синтаксису для мов була б корисною.

Взаємоперекладність природних мов: Теза про взаємоперекладність усіх природних мов обговорюється лінгвістами та філософами. Бар-Гіллель стверджує, що ця теза є неоднозначною і може бути зрозумілою в різних сенсах. В одному сенсі, коли розглядається замкнута мова з жорсткими правилами і фіксованим словником, теза є хибною. Однак, в іншому сенсі, теза може бути тривіально правильною. Бар-Гіллель досліджує можливість нетривіальних смислів, в яких теза може бути істинною [Bar-Hillel 1953].

Ідіоми: Ідіоматичні вирази становлять значну проблему для машинного перекладу. Якимось чином можна уявити, як машина може виконувати дослівний переклад, але саме цей тип перекладу зазнає краху, коли стикається з ідіомою, яка за визначенням, а саме визначенням 3 Колегіального словника Вебстера 1951 року, є «висловом у мові, який властивий тільки йому самому або за граматичною конструкцією, або за значенням, яке не може бути виведене як ціле зі сполучених значень його елементів».

Універсальні синтаксичні категорії: Бар-Гіллель згадує, що запровадження синтаксичних категорій універсально для всіх мов без жодних методологічних втрат видається правдоподібним. Однак автор також визнає, що не в змозі надати універсальне визначення терміна «речення» і що в цій царині потрібне подальше уточнення.

Бар-Гіллель також згадує про можливість існування інших універсальних або специфічних категорій на додаток до речення і власної назви. Автор припускає, що, хоча було б корисно мати загальний набір лінгвістичних категорій для багатьох цілей, можуть бути й недоліки, пов'язані зі створенням універсальної системи категорій. Загалом, зауваження автора щодо універсальних синтаксичних категорій є попередніми і визнають складність і потенційні компроміси, пов'язані з їх запровадженням у всіх мовах [Bar-Hillel 1953].

Певну схожість у висловлюваннях і думках ми можемо знайти в роботах Гаррі Гарріса, який стверджує, що машинний переклад, процес автоматичного

перекладу тексту з однієї мови на іншу за допомогою комп'ютерних алгоритмів, стикається з кількома проблемами. Ці проблеми включають в себе

Неоднозначність: У мовах часто зустрічаються слова або фрази з багатозначним значенням, і без контексту визначити правильний переклад може бути складно. Системи машинного перекладу намагаються точно передати значення багатозначних слів або фраз.

Ідіоматичні вирази: Мови містять ідіоматичні вирази, які мають унікальні значення, що не можуть бути перекладені дослівно. Системам машинного перекладу може бути складно зрозуміти й точно перекласти такі вирази, що призводить до незручних або неправильних перекладів.

Граматика та синтаксис: Різні мови мають різні граматичні структури та синтаксичні правила. Системи машинного перекладу повинні розуміти і правильно застосовувати ці правила, щоб створювати граматично правильні переклади. Однак граматика й синтаксис можуть бути складними й відрізнятися в різних мовах, що ускладнює завдання машинних перекладачів, які постійно створюють точні переклади.

Культурні особливості: Мови тісно пов'язані з культурою, і певні фрази або поняття можуть не мати прямих еквівалентів в інших мовах. Системам машинного перекладу буває складно врахувати ці культурні нюанси, що призводить до того, що в перекладах бракує потрібного культурного контексту або чутливості [Harry Harris 2010].

Згідно з висловленою думкою Дієго Муссаллема, Маттіаса Вауера та Аксея-Сиріла Нгонга Нгомо з Data Science Group Падерборнського університету, є декілька невирішених викликів у машинному перекладі, які стосуються певних проблем:

Велика складність семантичної неоднозначності: Ця проблема виникає з-за багатозначності та омонімії слів. Вирішення цього завдання потребує великої кількості паралельних даних для точного перекладу.

Різниця у структурі речень: Перебудова структури - це реорганізація синтаксичних елементів мови, а це стає ключовим аспектом процесу перекладу.

Лінгвістичні властивості/особливості: Велика кількість мов має складну систему часових форм. Коли системи машинного перекладу стикаються з реченнями з таких мов, їм може бути важко розпізнати поточний вхідний час і перекласти вхідне речення у правильний час цільової мови.

Ці виклики змушують розглядати нові підходи та розв'язання для покращення точності і ефективності машинного перекладу [Diego 2019, с.3].

РОЗДІЛ 2

ОСОБЛИВОСТІ ПЕРЕКЛАДУ ПУБЛІЦИСТИЧНИХ СТАТЕЙ СУЧАСНИМИ СИСТЕМАМИ МП

2.1 Методи оцінювання якості машинного перекладу

В оцінці якості результатів машинного перекладу застосовують різноманітні методи, щоб забезпечити об'єктивність і точність вимірювань. Процес оцінювання машинного перекладу включає в себе як ручні, так і автоматичні підходи, спрямовані на визначення ефективності програм-перекладачів.

Один з ключових показників автоматичної оцінки - це коефіцієнт помилок на слово (WER), який розглядається на рівні окремих слів. Визначаючи різниці між оригінальним і перекладеним реченнями, WER визначає відсоток слів, які потребують корекції для досягнення відповідності бажаному результату. Треба врахувати, що такий метод має свої обмеження, зокрема, необхідність визначення еталонного речення. Ця метрика ураховує лише один правильний переклад, у той час як існує можливість кількох адекватних варіантів перекладу для одного й того ж речення.

Крім того, на шляху вдосконалення процесу оцінювання машинного перекладу виникає необхідність враховувати інші аспекти, такі як контекстуальна адаптабельність, спрямованість на специфічні галузі або мовна різноманітність.

Позиційно-незалежний коефіцієнт помилок (PER) спрямований на подолання обмеження, яке виникає при використанні WER і пов'язане з порядком слів [Tillmann 1997]. PER розглядає еталон та гіпотезу як множини слів, що дає можливість вирівнювати слова з гіпотези зі словами в еталоні незалежно від їхньої позиції. Завдяки цьому підходу PER вихідного тексту

виконаного МП гарантовано буде меншим або дорівнюватиме WER. Важливою рисою цього методу є можливість ігнорування позиції слів, але, на жаль, це також може призвести до неможливості відрізнити правильний переклад від перекладу, де слова були зашифровані або введені невірно [Dorr 2010, с.816].

Коефіцієнт помилок у реченні (SER) є важливим показником, який допомагає оцінювати якість машинного перекладу. Його основна мета - виміряти ступінь розбіжності між перекладеним реченням та еталонним реченням. При цьому, SER враховує всі можливі зміни, які були внесені в текст перекладу, щоб зробити його більш відповідним еталону. Однак слід враховувати, що вимірювання SER може бути більш складним завданням, особливо при наявності різних можливих перекладів для одного і того ж еталонного речення. Також враховуючи мовні особливості та нюанси, деякі зміни можуть бути важко визначити як помилки, оскільки вони можуть бути обумовлені вибором між різними варіантами виразів [Khenglawt 2018, с.144].

BLEU [Papineni 2002] - це метрика, яку часто використовують дослідники машинного перекладу і яка має тісний зв'язок з метрикою NIST [Doddington 2002, с.141]. Вона стала однією з перших, що показала високу кореляцію з оцінкою перекладу людиною, і в подальшому породила багато різних варіантів для точнішої оцінки.

Суть BLEU полягає в оцінці окремих речень порівняно з набором високоякісних еталонних перекладів. Потім ці оцінки усереднюються, щоб отримати кінцевий показник для системи машинного перекладу. Іншими словами, BLEU визначає, наскільки близько результат перекладу до еталону, створеного людиною. Така міра дозволяє отримати оцінку, яка відображає якість роботи системи МП [Tomedes 2022].

Важливо зауважити, що хоча BLEU є корисною метрикою, вона також має свої обмеження. Вона може не завжди точно враховувати всі нюанси перекладу, особливо коли різні варіанти перекладу можуть бути вірними. Також враховуйте, що метрики, включаючи BLEU, є лише інструментами для

оцінки, і їх слід доповнювати аналізом та експертною оцінкою для отримання більш повного розуміння якості перекладу.

Деякі види текстів показали високу надійність при використанні МП для перекладу. Це стосується технічних, юридичних, маркетингових і менеджментових матеріалів, а також текстів з областей туризму та громадського харчування. Також МП добре справляється з перекладом посібників, інструкцій, директив ЄС, нормативних актів, страхових контрактів і багато іншого [Štefčík 2015, с.140].

Часто переклад публіцистичних текстів оцінюється лише кількісно, особливо це стосується наукових статей. Для цього використовуються автоматичні методи оцінювання, такі як метрика Meteor [Denkowski 2010, с.339] Вона порівнює переклад машини з одним або кількома еталонними перекладами, створеними людиною. Проте важливо зауважити, що ці метрики дають лише кількісні результати і не розкривають докладно сильні та слабкі сторони системи перекладу [Stymne 2011, с.56].

Сеунджун Лі та його колеги розподілили метрики автоматичної оцінки, що використовуються в машинному перекладі, на кілька типів в залежності від методу застосування. Ці варіанти включають лексичні метрики, метрики на основі вбудовування та контрольовані метрики.

Лексичні метрики, для початку, вимірюють подібність між гіпотезою та еталоном за допомогою математичних формул або евристичних підходів. Важливо відмітити, що ці метрики оперують без використання алгоритмів машинного навчання. Залежно від одиниць порівняння, їх поділяють на ті, що базуються на словах, та ті, що використовують символи для оцінки схожості.

Метрики на основі вбудовування, з іншого боку, використовують алгоритми машинного навчання або глибокого навчання для вимірювання подібності. Вони використовують гіпотези та еталони, представлені у вигляді щільних векторів, для глибшого аналізу змісту та контексту.

Контрольовані метрики вивчають регресійні моделі, використовуючи марковані дані, які відзначені коментарями людей. Це дозволяє зважено оцінювати якість перекладу з урахуванням людської експертизи.

До того ж, можна класифікувати метрики автоматичної оцінки за призначенням завдання навчання. Так, співставлення, регресія, ранжування та генерація відповідей стають окремими категоріями в завданні оцінювання [Seungjun Lee 2023].

При оцінці великого обсягу машинних перекладів було створено небагато схем класифікації помилок машинного перекладу, і більшість з них були створені для конкретних цілей. SAE J2450 QualityMetric, розроблена Співтовариством автотранспортних інженерів [SAE 2001], і Framework For Standard Error Marking, створений Американською асоціацією перекладачів [ATA 2017], зорієнтовані на оцінку людського перекладу і не мають достатньої деталізації.

Типологія помилок, що часто виникають у автоматичному перекладі, створена Корреа, також не повністю відповідає потребам, головним чином через те, що вона розроблена для оцінки методом «білого ящика» під час створення системи [Correa 2003]. Типології помилок для машинного перекладу такі як Лоффлера-Лауріана, що заснована на лінгвістичних аспектах, які виникають під час редагування [Loffler-Laurian 1996] можуть слугувати корисною основою для подальших досліджень. Проте вони все ще не надто докладні і вимагає доступу до вихідного тексту для застосування, тому наступну представлену Деббі Еліотом та його колегами класифікацію можна вважати більш доречною [Elliott 2004, с. 67] (див. Рис. 2.1).

Схема категоризації помилок за Еліотом

<i>Part of speech</i>	Noun: Compound	Noun string or Named entity	Inappropriate	Meaning clear/unclear	
			Untranslated	Part meaning clear/unclear	
					Untranslated
					Part untranslated
	Noun: Acronym	Incorrect			
	Noun: Pronoun	Inappropriate	Incorrect anaphor		
			Other		
		Untranslated			
		Unnecessary			
		Omitted	Direct object pronoun		
		Relative pronoun			
		Other			
	Noun: Common or Adjective or Adverb or Conjunction	Inappropriate	Meaning clear/unclear		
			Outrageous		
		Untranslated			
		Unnecessary			
		Omitted			
	Preposition	Inappropriate	With noun/verb/adjective		
		Untranslated			
		Unnecessary			
		Omitted			
	Determiner	Inappropriate			
		Untranslated			
Unnecessary		Definite article			
		Indefinite article			
		Other			
Omitted	Definite article				
	Indefinite article				
	Other				
Verb	Inappropriate	Meaning clear/unclear			
		Outrageous			
		Multiword verb structure			
	Untranslated				
	Unnecessary				
	Omitted				
<i>Tense or conjugation</i>	Tense or mood				
	Conjugation				
<i>Incorrect position</i>	Acronym / pronoun / common noun / adjective / adverb / conjunction / preposition / determiner / verb / negator / noun string appendage				
	Compound noun sequence	Noun string or Named entity	Word order		
Arrangement					
<i>Other</i>	Part of speech incorrect / inelegant or inappropriate style / incomprehensible expression / spelling error / incorrect negation / ordinal number untranslated / qualifier unnecessary				
	Number	Singular should be plural / plural should be singular			
	Case	Upper case required / lower case required			

Fig. 1. Fluency Error Categorization Scheme

Одночасно треба враховувати вартість, швидкість і обсяг словникового запасу, оскільки ці фактори мають великий вплив на якість машинного перекладу. Фланаган виділяє декілька чинників, які ускладнюють оцінку якості машинного перекладу:

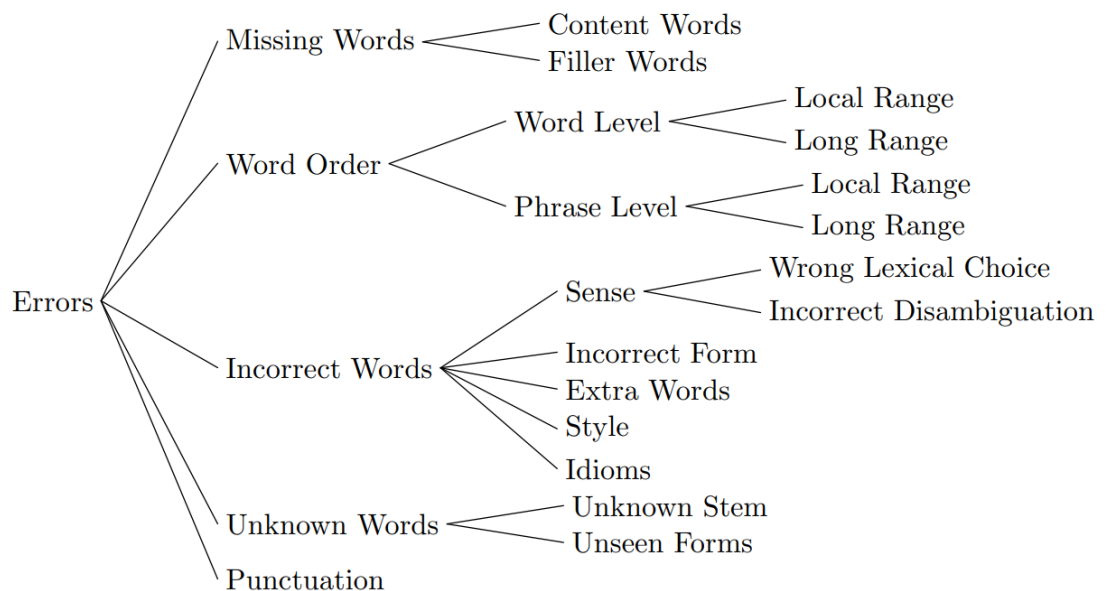
- в одного тексту може бути декілька правильних перекладів;
- помилки можуть виникати не тільки на рівні окремих слів, але й в цілих фразах, зв'язках між реченнями або порядку слів, що, як відзначають Вілар та інші дослідники, робить некоректним простий підрахунок кількості помилкових слів;
- одна помилка може викликати інші.

Важко зрозуміти причину помилок у виводі машинного перекладу, так як оцінювач зазвичай не може слідкувати за всіма тестами і діями в програмному забезпеченні, що ускладнює визначення, що саме спричинило невдалий переклад тексту [Flanagan 1994; Vilar 2006]

У наступній таблиці представлена таксономія, запропонована Віларом, Сюй, Д'Харо та Неєм для ідентифікації помилок у машинному перекладі [Vilar 2006, с. 699] (див. Рис. 2.2).

Рисунок 2.2

Класифікація помилок перекладу за Віларом, Сю, Д'Харо та Неєм



Наступна класифікація розглядає помилки через призму мовних одиниць, на які вони впливають. Грубі категорії - Орфографія, Лексика, Граматика, Семантика і Дискурс - вказують на рівень мови, на якому відбувається помилка.

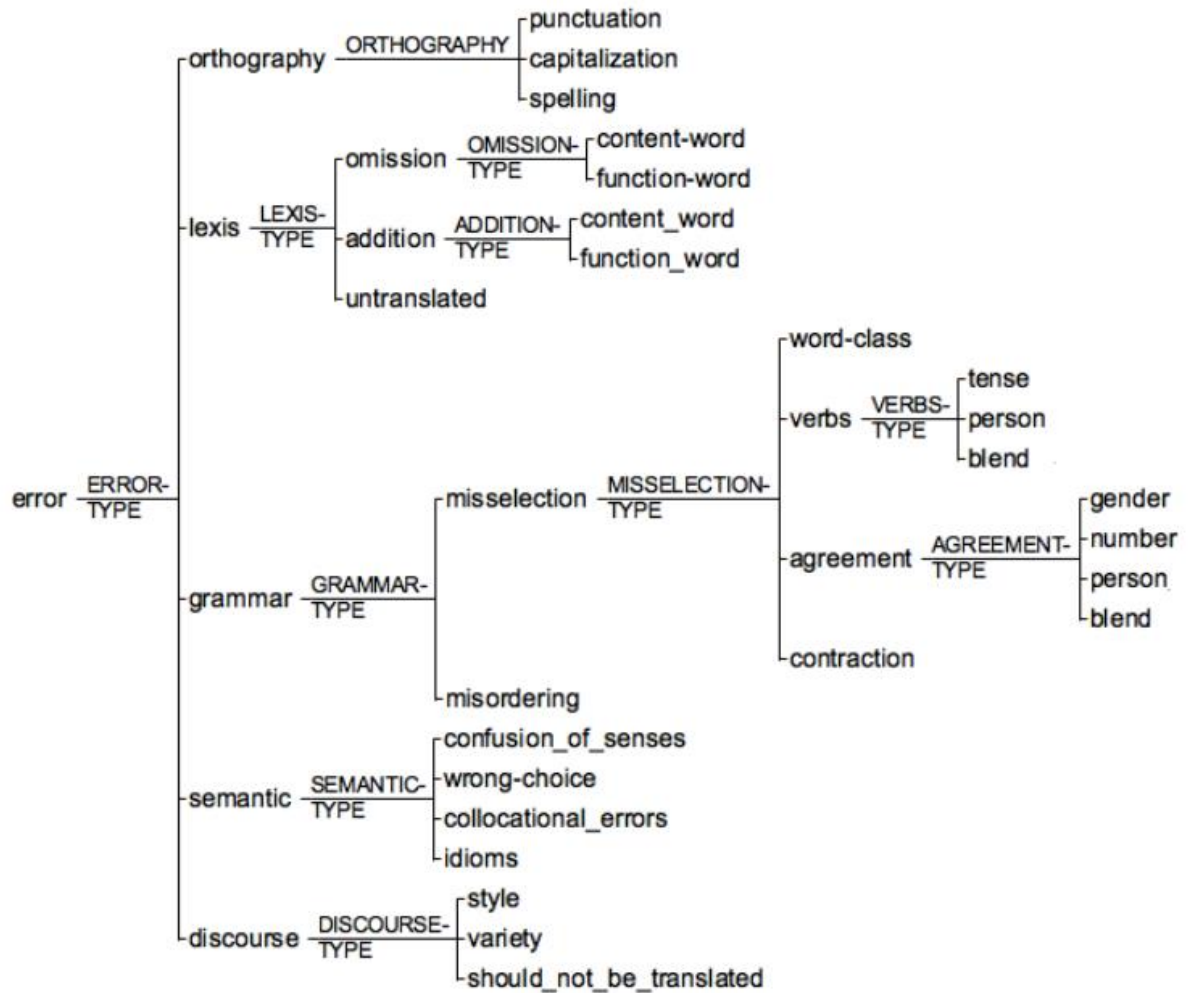
Орфографічні помилки охоплюють всі неточності, пов'язані з неправильним використанням розділових знаків та написанням слів. Дослідником Костою на ряду з його колегами було виділено три види орфографічних помилок: помилки в пунктуації, використання великої літери та правопис.

До категорії лексичного рівня вони включили всі помилки, що стосуються виразів. Важливо розрізнити, що, на відміну від орфографічних помилок, що торкаються символів у слові, лексичні помилки відносяться до загального способу перекладу слова. Таким чином, на рівні лексики враховуються такі помилки: відсутність перекладу, опущення та додавання.

Помилки на рівні граматики виявляються у морфологічних та синтаксичних аспектах мови. Коста разом з колегами виділили два типи помилок: помилки неправильного вибору та помилки в порядку слів.

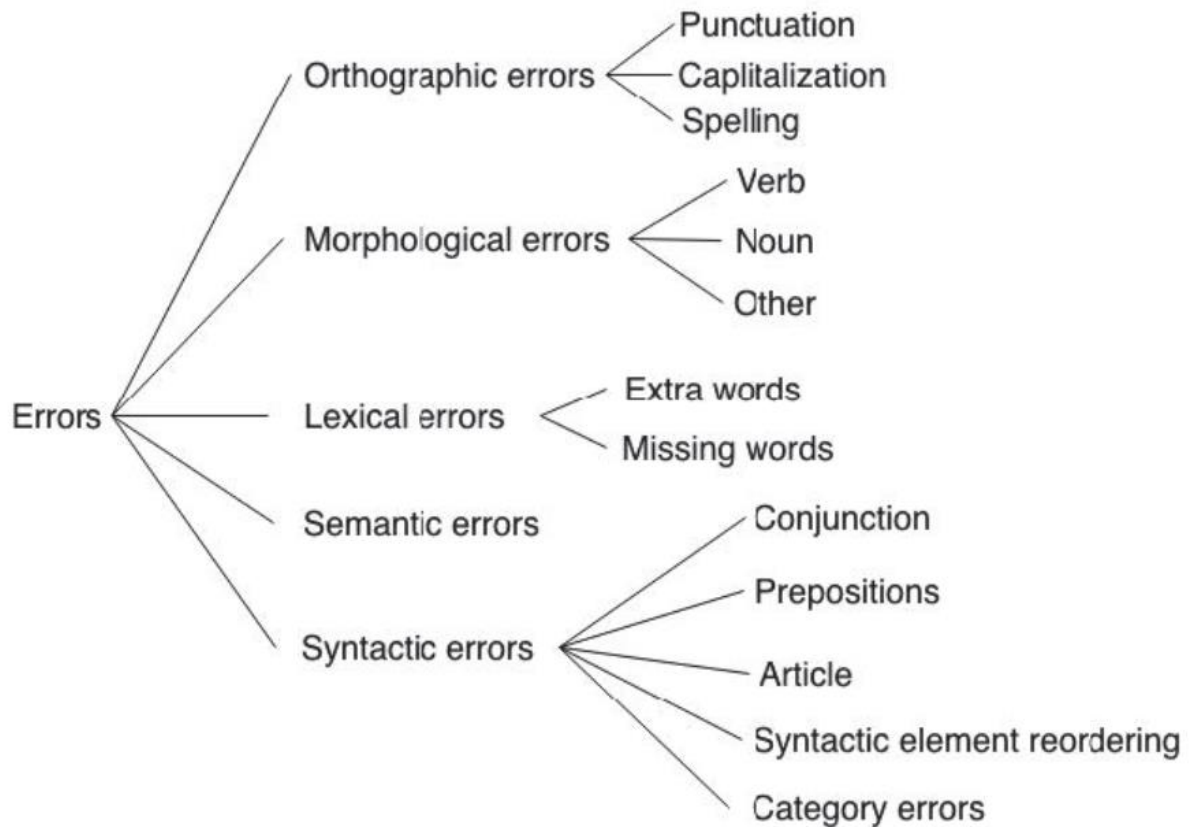
Семантичні помилки пов'язані з невірним вибором слів і значень. Коста виділив три види помилок: плутанина значень, неправильний вибір, а також помилки в колокаціях та ідіомах.

Помилки на рівні дискурсу стосуються варіацій у використанні мови, які не є стандартними, але також не вважаються помилками. Тут можна виділити три ситуації на рівні дискурсу: стиль, різноманітність та випадки, коли переклад не потрібен [Costa, с. 6] (див. Рис. 2.3).



Інше ж дослідження провели Нурелдін з колегою Абдулхаліком, які аналізували дані за допомогою моделі Hsu. У зв'язку з цим Hsu надав більш детальну типологію таких помилок, як показано на таб. 3 нижче. Ця модель класифікує помилки на орфографічні, морфологічні, лексичні, семантичні та синтаксичні [Noureldin 2020, с.410] (див. Рис. 2.4).

Класифікація помилок МП (запозичена з Hsu, 2014)



Classification of MT Errors (adopted from Hsu, 2014)

Для вимірювання якості перекладу помилки було перекласифіковано за допомогою системи багатовимірних показників якості (Multidimensional Quality Metrics - MQM) та системи оцінки якості локалізації (Localization Quality Evaluation - LQE), згідно з якою помилки було поділено на три типи:

1. Незначні помилки: це помилки, які не впливають на розуміння змісту, але впливають на вільне володіння мовою.
2. Значні помилки: це помилки, які ускладнюють розуміння тексту перекладу, але при цьому загальний зміст передається.
3. Критичні помилки: це помилки, які змінюють зміст вихідного тексту і роблять його або незрозумілим, або спотвореним [Noureldin 2020, с.411].

2.2 Класифікація і аналіз помилок МП на матеріалі відібраних текстів

У цьому підрозділі буде розглянуто переклади статей служби новин BBC, виконаних системами машинного перекладу. Мета - визначити якість цих перекладів, використовуючи раніше описану класифікацію помилок за Віларом, Сю, Д'Харо та Неєм.

Будуть проаналізовані переклади текстів, опублікованих на веб-сайті BBC, зосереджуючись на різних типах помилок, визначених за класифікацією Вілара. Ця класифікація дозволяє структуровано виявити та оцінити проблеми, що виникають у процесі машинного перекладу, та надати рекомендації щодо їх виправлення.

Вибір даної класифікації для аналізу перекладів текстів з української мови на англійську системами машинного перекладу був обумовлений наступним рядом причин:

1. Класифікація охоплює широкий спектр потенційних помилок, які можуть виникнути під час машинного перекладу, що гарантує врахування всіх аспектів якості перекладу. Такий комплексний підхід дає змогу ретельно оцінювати системи машинного перекладу.
2. Система класифікації поділяє помилки на окремі типи, що полегшує виявлення та аналіз конкретних проблем. Ця класифікація допомагає дослідникам і розробникам визначити, які саме сфери потребують удосконалення.
3. Класифікація призначає пріоритети різним типам помилок, визнаючи, що не всі помилки однаково важливі. Наприклад, помилки, пов'язані з пропуском слів змісту, вважаються більш критичними, оскільки вони можуть суттєво вплинути на зміст перекладеного тексту. Така розстановка пріоритетів допомагає вирішити, які проблеми слід вирішувати в першу чергу.

4. Класифікація визнає, що помилки не є ізольованими, а взаємопов'язаними. Наприклад, неправильний переклад слова може призвести до проблем із порядком слів. Усвідомлення цієї взаємозалежності допомагає зрозуміти першопричини помилок у перекладі.
5. Класифікацію можна адаптувати до різних мовних пар і використовувати як основу для оцінювання систем машинного перекладу для різних мов. Така адаптивність забезпечує її актуальність і застосовність у ширшому контексті.

Спираючись на запропоновану класифікацію у роботі будуть братися до уваги п'ять типів помилок, а саме *пропущені слова*, що у свою чергу включають в себе як *змістові слова*, так і *слова-заповнювачі*; помилки *порядку слів*, що можуть стосуватися *рівня слів і фразового рівня*; *неправильні слова*, що являють собою найбільшу групу помилок, включаючи в себе *сміслові помилки, неправильні форми, зайві слова, стиль і ідіоми; невідомі слова з невідомою основою та невидимими формами; пунктуація*.

Помилка «Пропущене слово» може виникнути у реченні, коли воно не містить якогось слова. Ця помилка може бути різних видів, в залежності від того, наскільки це пропущене слово важливе для змісту речення.

Перше - це «Змістові слова». Це важливі слова, які необхідні для вираження основного змісту речення. Часто це іменники, дієслова та інші слова, які надають конкретність і зміст реченню. Відсутність таких слів може змінити зміст речення або зробити його незрозумілим.

Другий тип – «Слова-заповнювачі». Це слова, які не несуть основного значення в реченні і можуть бути пропущені без втрати основного змісту. Вони використовуються, головним чином, для того, щоб сформувавши граматично правильне речення. В їх відсутності речення все ще може бути зрозумілим, але воно може бути граматично некоректним.

Зазвичай помилки першого типу, пов'язані з відсутністю «головних слів», є більш важливими і потребують негайного виправлення. Але не слід

ігнорувати і помилки другого типу, оскільки вони також можуть впливати на зрозумілість та прийнятність речення.

Тепер розберемо наступний важливий аспект машинного перекладу, а саме – «Порядок слів» у перекладеному реченні. Ця категорія помилок може включати в себе перестановку на рівні окремих слів або на рівні словосполучень.

У випадку перевпорядкування на рівні слів ми можемо створити правильне речення, переміщуючи окремі слова незалежно одне від одного, але є ще один аспект - перевпорядкування на рівні словосполучень. Тут нам потрібно переміщувати блоки послідовних слів, щоб сформувавши правильний переклад на основі згенерованої гіпотези.

Отже, контекст та правильний порядок слів грають величезну роль у якості машинного перекладу, і помилки на цьому рівні можуть значно вплинути на зрозумілість та коректність перекладу.

Найширшою категорією помилок у цій класифікації є «Неправильні слова». Вони виникають у тих випадках, коли система не може знайти правильний переклад для конкретного слова. Ця категорія помилок поділяється на п'ять підкатегорій для більшої деталізації.

Перша підкатегорія – «Смислові помилки». У цьому випадку проблемне слово порушує загальний зміст речення. Такі помилки виникають, коли система вибирає неправильний переклад для слова. Або ж, це може статися, коли система не може визначити правильне значення початкового слова в даному контексті. Це може призвести до серйозних порушень у змісті тексту та спотворення смислу.

Наступною підкатегорією категорії «Неправильні слова» є помилки «Неправильної форми». Цей вид помилок виникає, коли система не може вірно сформувавши вірну форму слова, хоча переклад базової форми був правильним. Ця проблема особливо актуальна для флективних мов, де велика різноманітність словесних форм створює важкість для машинного перекладу. Спосіб аналізу цих помилок значно відрізняється в залежності від конкретної

пари мов, яку ми розглядаємо. Наприклад, для іспанської мови, яка має багато флексій, важливо розрізняти неправильні часи дієслів і проблеми узгодження між іменниками, прикметниками та артиклями.

Ще однією підкатегорією помилок є «Зайві слова». Ці помилки виникають, коли в перекладеному реченні з'являються зайві слова. Хоч цей вид помилок особливо важливий при перекладі прямої мови, оскільки особливості мовлення однієї людини можуть призвести до надлишкового використання слів у сформованому реченні, брати цю підкатегорію до уваги все ж важливо.

Останні дві підкатегорії помилок не є настільки важливими. «Стилістичні помилки» виникають при невдалому виборі слова при перекладі речення. Хоч при цьому загальний зміст і залишається збереженим, але переклад в такому випадку не можна вважати правильним. Прикладом цього може бути невдале повторення слова в близькому контексті. У такому випадку людина, ймовірно, вибрала би синонім, щоб уникнути повторень.

Друга підкатегорія помилок стосується ідіом, які машинна система не розпізнає і намагається перекласти буквально. Зазвичай такі вирази не можна передати дослівно, що призводить до додаткових неточностей в перекладі.

Невідомі слова також є джерелом помилок. Тут ми можемо додатково розрізняти справді невідомі слова (або основи) та невидимі форми відомих основ.

Ця категорія особливо важлива в контексті китайсько-англійського перекладу. Для більшості європейських мов, або інших мов, що мають спільний алфавіт, невідомі власні назви можуть бути легко «перекладені» шляхом простого копіювання вихідного слова в згенерований переклад без подальшої обробки. У випадку китайських ієрогліфів справа ускладнюється, оскільки їх неможливо просто перекласти на англійську мову, адже це потребує особливого процесу перетворення, іноді з урахуванням вимови. Тому для цієї мовної пари ми розрізняємо невідому особу, місцевість, організацію та інші власні назви.

Нарешті, можуть виникати «Пунктуаційні помилки», але для сучасного машинного перекладу вони становлять лише незначні перешкоди для мов без фіксованих правил пунктуації.

Важливо відзначити, що ці типи помилок не завжди відокремлені один від одного. Часто трапляються випадки, коли один тип помилок призводить до виникнення іншого. Наприклад, неправильний переклад слова може спричинити неправильний порядок слів у створеному реченні [Vilar 2006, с. 698].

Перейдемо безпосередньо до аналізу самих проблем, що можуть виникати у процесі перекладу статей українською мовою на англійську засобами машинного перекладу. Почнемо з сервісу META, який пропонує нам компанія METAPORT MEDIA LIMITED, а у якості тексту, що буде аналізуватися візьмемо фрагмент статті «"Канабіс допоміг пережити найтяжчі дні". Як ветерани війни з Росією тамують біль» опублікований службою новин BBC 27 вересня 2023 [Adams 2023; Адамс 2023] (див. Табл. 2.1).

Таблиця 2.1

Фрагмент тексту «"Канабіс допоміг пережити найтяжчі дні". Як ветерани війни з Росією тамують біль»

Український варіант	Англійський варіант	Meta Translate
У скромній кийвській квартирі солдати, які зараз не на фронті, збираються, щоб покурити канабіс і на мить забути про те, що бачили.	In a modest Kyiv apartment, off-duty soldiers meet to smoke cannabis and forget, for a moment, the things they have seen.	To the modest Kyiv apartment soldiers that now not at the front going, to smoke канабіс and on an instant to forget that saw.
Вони не хочуть, щоб їх ідентифікували. Канабіс серед військових не схвалюють навіть тут, далеко від фронту.	They don't want to be identified. The drug is frowned upon in the military, even here, far from the front.	They do not want, that they were identified. Канабіс among soldiery disapprove even here, far from front.

Пропущені слова (змістові слова):

- Український варіант: «збираються, щоб покурити»
- Англійський варіант: "meet to smoke"
- Meta Translate: "going, to smoke"

Відсутність допоміжного дієслова в цьому контексті вважається пропуском слова змісту, а не слова-заповнювача. Допоміжні дієслова, такі як "is", "am", "are", "was", "were", "will be" тощо, необхідні для вираження часу та способу в англійській мові. Тому, коли допоміжне дієслово відсутнє, це може суттєво вплинути на зміст і граматичну правильність речення. У цьому випадку пропуск "are" впливає на час ("present continuous") і вважається

помилкою. Заповнювачі, з іншого боку, зазвичай включають такі слова, як "and", "the", "a", "an", "in" тощо, які часто не є важливими для передачі основного змісту речення, але є необхідними для граматичної правильності.

- *Український варіант: «солдати, які зараз не на фронті»*
- *Англійський варіант: “off-duty soldiers”*
- *Meta Translate: “soldiers that now not at the front”*

Як і в попередньому випадку, при прямому перекладі фрази нульовою трансформацією було пропущено граматично значуще дієслово “to be”, що вважається помилкою.

- *Український варіант: «забути про те, що бачили.»*
- *Англійський варіант: “forget the things they have seen.”*
- *Meta Translate: “to forget that saw.”*

У наступній фразі спостерігається якнайменше пропущення важливого як з граматичної точки зору, так і з точки зору передачі смислу слово, а саме займенника “they”. Відсутність займенника “they” в машинному перекладі призводить до граматично неправильного та неповного речення.

- *Український варіант: «Вони не хочуть, щоб їх ідентифікували.»*
- *Англійський варіант: “They don't want to be identified.”*
- *Meta Translate: “They do not want, that they were identified.”*

У цьому контексті пропуск дієслова "to be" впливає на повноту і точність речення, що робить його критичною помилкою, оскільки воно впливає на загальний зміст і граматичну правильність перекладу як і в випадках наведених раніше.

- *Український варіант: «Канабіс серед військових не схвалюють»*
- *Англійський варіант: “The drug is frowned upon in the military”*
- *Meta Translate: “Cannabis among soldiery disapprove”*

Ще один випадок з упущенням граматично значущого дієслова "to be" можна простежити у цьому рядку. Як і в попередніх прикладах, обмежуючись

лише нульовою трансформацією при перекладі, система МП припустилась помилки пропустивши змістове слово.

Порядок (на рівні словосполучень):

- *Український варіант: «солдати, які зараз не на фронті»*
- *Англійський варіант: “off-duty soldiers”*
- *Meta Translate: “soldiers that now not at the front”*

Фраза "off-duty soldiers" у еталонному варіанті тексту правильно доносить суть, що ці солдати зараз не перебувають на передовій. Однак у машинному перекладі слова переставляються місцями і використовується фраза "soldiers that now not at the front", що не є звичним способом вираження тієї самої ідеї англійською мовою.

Неправильні слова (сміслові помилки):

- *Український варіант: «У скромній київській квартирі солдати...»*
- *Англійський варіант: “In a modest Kyiv apartment...”*
- *Meta Translate: “To the modest Kyiv apartment...”*

Машинний переклад неправильно переклав “У” як “То” замість “In”, що призвело до смислової помилки.

- *Український варіант: «і на мить забути про те, що бачили.»*
- *Англійський варіант: “and forget, for a moment, the things they have seen.”*
- *Meta Translate: “and on an instant to forget that saw.”*

Машинний переклад дає фразу "on an instant", яка, хоча і є дослівним перекладом оригінального тексту, але не точно передає його зміст англійською мовою.

Неправильні слова (неправильної форми):

- *Український варіант: «і на мить забути про те, що бачили.»*
- *Англійський варіант: “and forget, for a moment, the things they have seen.”*
- *Meta Translate: “and on an instant to forget that saw.”*

В даному фрагменті помилка з неправильною формою виникає через те, що слово "saw" неправильно відмінюється в минулому часі. У контексті

речення доречніше було б перекласти як "had seen", щоб передавати правильний час і значення. "Saw" - це простий минулий час дієслова "see", тоді як "have seen" представляє теперішній доконаний час в англійській мові.

В еталонному реченні "have seen" передає смисл того, що солдати були свідками подій у минулому, і цей досвід є актуальним для теперішнього моменту, оскільки вони намагаються забути цей досвід. Машинний переклад, в якому використано слово "saw", є неправильним, оскільки в ньому відсутній час дієслова, необхідний для даного контексту.

- *Український варіант: «не схвалюють навіть тут»*
- *Англійський варіант: “is frowned upon in the military”*
- *Meta Translate: “disapprove even here”*

Слово "disapprove" у цьому випадку має стояти у формі минулого часу, тобто "disapproved" (враховуючи наявність передуючого йому дієслова "is"), щоб речення було граматично правильним і зберігало узгодженість у часі дієслова.

Неправильні слова (стилістичні помилки):

- *Український варіант: «забути про те, що бачили.»*
- *Англійський варіант: “forget the things they have seen.”*
- *Meta Translate: “to forget that saw.”*

Машинний переклад повторює слово "to" перед словами "to smoke" та "to forget". Цей повтор порушує плавний хід речення і є стилістично неправильним. В англійській мові звичніше сказати "to smoke cannabis" і "forget what they have seen", не повторюючи "to".

Невідомі слова:

- *Український варіант: «покурити канабіс»*
- *Англійський варіант: “smoke cannabis”*
- *Meta Translate: “smoke канабіс”*

Ця помилка виникає, коли машинний переклад стикається зі словом або терміном, якого немає в його словнику або який він не може адекватно

перекласти. У цьому випадку машинний переклад, схоже, сприйняв «канабіс» як невідоме слово і спробував перекласти його безпосередньо як "канабіс".

- *Український варіант: «Канабіс серед військових не схвалюють»*
- *Англійський варіант: “The drug is frowned upon in the military”*
- *Meta Translate: “Cannabis among soldiery disapprove”*

Як і в попередньому випадку система машинного перекладу не змогла знайти відповідник до слова “канабіс” чим спричинила появу помилки не переклавши це слово.

Пунктуаційні помилки:

- *Український варіант: «збираються, щоб покурити»*
- *Англійський варіант: “meet to smoke”*
- *Meta Translate: “going, to smoke”*

Програма, перекладаючи фрагмент нульовою трансформацією, скопіювала положення розділових знаків з української мови, що спричинило пунктуаційну помилку в англійській мові.

- *Український варіант: «не схвалюють навіть тут, далеко від фронту.»*
- *Англійський варіант: “is frowned upon in the military, even here, far from the front.”*
- *Meta Translate: “disapprove even here, far from front.”*

Ще одну пунктуаційну помилку система МП припустилася пропустивши кому перед словом “even”.

На прикладі цього фрагменту можна прийти висновку, що переклад виконаний системою META Translate є неякісним зі значною кількістю помилок в різних категоріях. Найпоширеніша з них - пропущення змістових слів. У сукупності ці помилки призводять до того, що переклад стає не тільки неточним, але й складним для розуміння.

Машинний переклад не передає нюансів оригінального українського тексту і близько не відповідає еталонному перекладу, який вважається

безпомилковим. Він має проблеми з базовою граматикою та лексикою, що призводить до втрати сенсу та зв'язності перекладу.

Загалом, виходячи з цього прикладу, можна зробити висновок, що машинний переклад у цьому контексті не підходить для створення якісних перекладів, особливо коли йдеться про складні речення та мови зі структурними відмінностями.

Наступною системою машинного перекладу, чії помилки будуть проаналізовані, став відомий по всьому світу Google Translate, що пропонує нам однойменна компанія Google. У якості тексту, що буде аналізуватися візьмемо фрагмент статті «Росіяни пішли в масштабний наступ на Авдіївку. Що там відбувається» опублікований службою новин BBC 13 жовтня 2023 [Abdurasulov 2023; Абдурасулов 2023] (див. Табл. 2.2).

Таблиця 2.2

Фрагмент тексту «Росіяни пішли в масштабний наступ на Авдіївку.

Що там відбувається»

Український варіант	Англійський варіант	Google Translate
Російські війська розпочали масштабний наступ на місто Авдіївка на сході України. За повідомленнями, до атаки у цьому районі залучили три батальйони чисельністю близько 2 000 військовослужбовців, десятки одиниць бронетехніки, а також літаки.	Russian troops have launched a major offensive on the town of Avdiivka in eastern Ukraine. Three battalions involving some 2,000 troops, dozens of armoured vehicles as well as jets have been reportedly involved in what is described as the largest-scale attack in this area.	Russian troops launched a large-scale offensive on the city of Avdiyivka in eastern Ukraine. According to reports, three battalions of about 2,000 troops, dozens of armored vehicles, and aircraft were involved in the attack in the area.

Пропущені слова (змістові слова):

- *Український варіант: «Російські війська розпочали»*
- *Англійський варіант: “Russian troops have launched”*
- *Google Translate: “Russian troops launched”*

У контексті речення наявність слова "have" необхідна для граматичної правильності та точної передачі дії. У даному випадку "have" виступає допоміжним дієсловом, яке вказує на теперішній доконаний час, що означає, що дія запуску відбулася в якийсь момент у минулому і має відношення до теперішнього часу. Без "have" речення здається неповним і не відповідає очікуваній граматичній структурі англійської мови. Це суттєва помилка, оскільки вона впливає на правильність перекладу.

Неправильні слова (сміслові помилки):

- *Український варіант: «на місто Авдіївка»*
- *Англійський варіант: “on the town of Avdiivka”*
- *Google Translate: “on the city of Avdiyivka”*

Використання слова "city" для опису Авдіївки в контексті цього речення є недоречним, оскільки Авдіївка не є великим містом. Некоректне використання слова "city" вносить смислову помилку в переклад, оскільки воно неправильно відображає масштаб і розмір Авдіївки. Для забезпечення точного перекладу було б доречніше використовувати термін "town" для опису Авдіївки, що відповідає її фактичним розмірам та статусу.

- *Український варіант: «а також літаки.»*
- *Англійський варіант: “as well as jets”*
- *Google Translate: “and aircraft”*

Неправильний переклад "aircraft" замість "jets" у рядку є насамперед смисловою помилкою. Це пов'язано з тим, що "aircraft" - термін широкий, він може охоплювати різні типи літальних апаратів, включаючи як літаки, так і гелікоптери.

Неправильні слова (неправильної форми):

- *Український варіант: «десятки одиниць бронетехніки»*
- *Англійський варіант: “dozens of armoured vehicles”*
- *Google Translate: “dozens of armored vehicles”*

Неправильний переклад слова "armored" замість "armoured" є хоч і не такою критичною, але все ж помітною помилкою. "Armored" - це варіант написання в американській англійській, тоді як "armoured" - у британській. Система машинного перекладу Google Translate не має уявлення про реципієнтів перекладеного тексту і за замовчуванням надає переклад американським варіантом англійської мови, а тому припускається помилки у випадках, коли потребується британський варіант.

- *Український варіант: «залучили»*
- *Англійський варіант: “have been reportedly involved”*
- *Google Translate: “were involved”*

Використання "were" замість "have been" у рядку є помилкою, пов'язаною з часом дієслова. Вживання "have been" необхідне для дотримання правильного узгодження часових форм дієслів у реченні. Речення описує дію, що триває, а "have been" - це теперішній доконаний час, який зазвичай використовується для позначення дій, що почалися в минулому і продовжуються в теперішньому часі. Участь літаків в атаці є безперервною дією, тому "have been" є більш доречним.

Неправильні слова (стилістичні помилки):

- *Український варіант: «до атаки у цьому районі»*
- *Англійський варіант: “the largest-scale attack in this area.”*
- *Google Translate: “in the attack in the area.”*

Використання "this" замість другого "the" у рядку призвело б до більш доречної та менш повторюваної конструкції. Зокрема, це питання невдалого вибору слів і повторення артиклів. Використання слова "this" або перефразування речення, щоб уникнути повторення артикля "the", покращило б загальну читабельність перекладу.

Невідомі слова:

- *Український варіант: «Авдіївка»*
- *Англійський варіант: “Avdiivka”*
- *Google Translate: “Avdiyivka”*

У цьому випадку система МП не змогла розпізнати правильну назву населеного пункту і надала некоректний переклад. І хоча ця помилка є незначною в перекладі, все ж важливо дотримуватися правильного написання географічних назв при перекладі. Хоча ця помилка не суттєво впливає на загальний зміст речення, для забезпечення точності перекладу рекомендується використовувати правильне написання "Avdiivka".

Аналіз машинного перекладу Google Translate показує наявність декількох помилок, включаючи пропущені слова, неправильний вибір слів, помилки в написанні, а також стилістичні помилки. Помилки в перекладі та виборі слів впливають на точність та зрозумілість тексту. Деякі з цих помилок є критичними, так як вони впливають на граматичну правильність та смислову цілісність речення.

Покращення системи машинного перекладу повинно включати можливість правильної обробки допоміжних дієслів та обов'язкових елементів мови, що використовуються для правильної граматичної структури.

Машинний переклад має потенціал для поліпшення, і в подальшому можливо досягнення кращої якості перекладу завдяки розвитку технологій та вдосконаленню алгоритмів. Для досягнення цієї мети, важливо постійно працювати над усуненням виявлених помилок, зокрема тих, що стосуються граматики, семантики та стилю тексту.

Наостанок розглянемо доволі перспективну та більш сучасну системою машинного перекладу DeepL Translate, що пропонує нам компанія DeepL SE. DeepL використовує нову поліпшену архітектуру нейронних мереж, що призводить до більш природного звучання перекладів, ніж у конкуруючих сервісів. У якості тексту, що буде аналізуватися візьмемо фрагмент статті «Полювання на дрони. Як обмеження Китаю на продаж безпілотників

вплинули на війну в Україні» опублікований службою новин BBC 22 жовтня 2023 [Shevchenko 2023; Шевченко 2023] (див. Табл. 2.3).

Таблиця 2.3

Фрагмент тексту «Полювання на дрони. Як обмеження Китаю на продаж безпілотників вплинули на війну в Україні»

Український варіант	Англійський варіант	Deepl Translate
<p>Безпілотники мали величезний вплив на війну в Україні, їх у великих кількостях використовували обидві сторони. Проте крок Китаю щодо встановлення обмежень на експорт викликав занепокоєння, що можуть виникнути проблеми з постачанням на фронт.</p>	<p>Drones have had a profound effect on the war in Ukraine, used in great quantities by both sides. China's move to place restrictions on exports, however, has led to concerns that there could be a problem with supplies.</p>	<p>Drones have had a huge impact on the war in Ukraine, with both sides using them in large numbers. However, China's move to impose export restrictions has raised concerns that there could be problems with supplies to the frontline.</p>

Неправильні слова (сміслові помилки):

- Український варіант: «Безпілотники мали величезний вплив»
- Англійський варіант: “Drones have had a profound effect”
- Deepl Translate: “Drones have had a huge impact”

"Profound" є більш доречним словом у цьому контексті, щоб передати глибину впливу, тоді як "huge" не передає той самий рівень значущості.

- *Український варіант: «Безпілотники мали величезний вплив»*
- *Англійський варіант: “Drones have had a profound effect”*
- *Deepl Translate: “Drones have had a huge impact”*

Слово "impact" часто вживається в рамках значення «впливу» в той час як слово "effect" вживається у значенні «результат». "Effect" - більш доцільне слово в цьому контексті, адже робота безпілотників вже відома своїм хорошим результатом.

Неправильні слова (неправильної форми):

- *Український варіант: «їх у великих кількостях використовували обидві сторони.»*
- *Англійський варіант: “used in great quantities by both sides.”*
- *Deepl Translate: “with both sides using them in large numbers.”*

Оскільки, згідно з переліком правил вживання теперішнього тривалого часу опублікований British Council, ми можемо використовувати цей час, щоб говорити про минуле, коли ми розказуємо історію або підсумовуємо книгу, фільм чи п'єсу [British Council]. Вживання "using" в теперішньому часі в машинному перекладі не відповідає минулому часу в оригінальному перекладі. Правильною формою має бути "used".

На основі проведеного аналізу, можна зробити наступні висновки про цю систему машинного перекладу Deepl:

Система машинного перекладу має проблеми з правильним вибором слів відповідно до контексту. Були вибрані слова, які не відповідають контексту і не можуть передати бажане значення.

Система також допускає помилки у виборі форм слів. Ці помилки порушують правильний відтінок часу в перекладі.

Важливо відзначити, що система не завжди враховує семантичні відтінки слів і не завжди вибирає найбільш точні та адекватні слова для передачі змісту.

Для покращення системи DeepL та підвищення якості машинного перекладу варто розробити більш точні алгоритми для вибору слів, які враховують контекст і семантику тексту. Важливо, щоб система відповідала не лише за словниковими значеннями слів, але і за їхніми семантичними відтінками в конкретному контексті.

Важливо також і враховувати правильний порядок слів та відтінок часу в перекладі. Покращення граматичної коректності допоможе зробити переклад більш природним та зрозумілим.

Отже, система машинного перекладу DeepL потребує подальшого вдосконалення для забезпечення більш точних та відповідних перекладів, зокрема у виборі слів і форм слова відповідно до контексту.

ВИСНОВКИ

Отримані дані дослідження свідчать про те, що стан сучасних систем машинного перекладу, таких як Google Translate, META, DeepL, є недосконалим, має як переваги так і певні недоліки. Аналіз проведених перекладів вказує на наявність різноманітних помилок у певних аспектах, таких як пропущення змістових слів, неправильний вибір слів, граматичні та стилістичні недоліки.

Особливості машинного перекладу, виявлені під час дослідження, полягають у тому, що системи часто не враховують семантичні відтінки слів і мають проблеми з контекстним вибором слів. Граматичні та структурні помилки також знижують якість перекладу, роблячи його не тільки неточним, але і важким для розуміння. Зазначені особливості вказують на те, що машинний переклад на сьогоднішній день не завжди може замінити людський переклад в усіх ситуаціях, особливо при роботі з складними конструкціями мови та текстами із значними семантичними відмінностями.

У роботі проведено порівняльний аналіз трьох систем машинного перекладу: META Translate, Google Translate та DeepL. Отримані результати дозволяють виділити деякі ключові відмінності між ними:

Переклад, виконаний META Translate, виявив значну кількість помилок в різних категоріях, таких як пропущення змістових слів та проблеми з граматикою та лексикою.

Переклад Google Translate має деякі помилки, такі як пропущені слова, неправильний вибір слів та стилістичні помилки. Деякі з них можуть впливати на граматичну правильність та смислову цілісність речення. Проте Google має потенціал для поліпшення, особливо з розвитком технологій та алгоритмів. Він залишається однією з найпоширеніших та найвикористовуваних систем машинного перекладу, здатною швидко перекладати тексти.

У системі DeepL були виявлені проблеми з вибором слів відповідно до контексту та помилки у виборі форм слів. Система не завжди враховує семантичні відтінки слів і може вибирати не найточніші та адекватні слова для передачі змісту. Проте попри це система показала найкращий результат з точки зору правильності та точності перекладу, припустивши меншу кількість помилок ніж попередні системи.

Роль машинного перекладу у сучасному світі залишається важливою, оскільки ці технології дозволяють швидко перекладати текст з однієї мови на іншу, сприяючи спілкуванню та обміну інформацією в умовах глобалізації. Однак необхідно розуміти його обмеження та враховувати, що для деяких завдань та контекстів краще використовувати людський перекладач.

Щодо перспектив розвитку машинного перекладу, важливо продовжувати вдосконалювати алгоритми та моделі, збільшувати обсяги навчальних даних та розробляти нові методи оцінки якості перекладу. Також важливо акцентувати увагу на розробці технологій, які дозволять системам машинного перекладу краще розуміти контекст та семантику, щоб підняти рівень точності та природності перекладу.

У кінцевому висновку можна підкреслити, що машинний переклад є важливим інструментом, але йому ще потрібно подолати численні труднощі, щоб наблизити якість перекладу до людського рівня. Такий підхід до дослідження машинного перекладу дозволяє краще зрозуміти його обмеження та сприяє подальшому вдосконаленню цієї важливої технології.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

СПИСОК ТЕОРЕТИЧНИХ ДЖЕРЕЛ

1. Акулов Є. І. Англомовний публіцистичний текст: семантичні та прагматичні аспекти перекладу : дисертація. *Київський політехнічний інститут імені Ігоря Сікорського*. 2020. 96 с.
2. Бірюков А. В. Оцінка якості машинного перекладу : автореф. ... канд. філол. наук : 10.02.16. К., 2008. 17 с.
3. Гудманян А. Г. Прагматична адекватність машинного перекладу. In: *Social communications in the innovative educational space: theoretical and applied aspects. National Technical University of Ukraine Kyiv Polytechnic Institute*. 2020. С. 142-152.
4. 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» № 5*. 2019. С. 48-54.
5. Карпіловська Є. А. Вступ до прикладної лінгвістики: комп'ютерна лінгвістика : підручник. Донецьк : ТОВ «Юго-Восток, Лтд», 2006. 188 с.
6. Кісіль М. В. До проблеми машинного перекладу. *Актуальні проблеми соціальної комунікації* : збірник матеріалів конференції. Івано-Франківськ : В-во ІФНГУНГ, 2011. С. 97-103.
7. Коваленко А. Я. Науково-технічний переклад: [навч. пос. для шкіл з поглибл. вивч. англ. мови, ліцеїв, гімназій, коледжів]. Тернопіль : Видавництво Карп'юка, 2001. 284 с.
8. Кузьменко Д. М. Комп'ютерна лінгвістика і Web 2.0. *Studia Linguistica*. Випуск 2. 2009. 250 с.
9. Максимов С. Є. Практичний курс перекладу (англійська та українська мови) Теорія та практика перекладацького аналізу тексту для студентів факультету перекладачів та факультету заочного та вечірнього навчання : Навчальний посібник : Ленвіт, 2006. 157 с.

- 10.Мацак Ж. Г., Скоробогатова Т. І. Машинний переклад та його специфіка.
URL: https://www.rusnauka.com/30_NIEK_2009/Philologia/54186.doc.htm
(дата звернення: 16.10.2023)
- 11.Мірам Г. Алгоритми перекладу: вступний курс з формалізації перекладу
(англ. мовою) / за ред. М. Даймонда. Київ: Твім інтер, 1998. 176 с.
- 12.Місуно Е. А. Переклад з англійської мови на російську мову: практикум:
навч. посібник / І.В. Шаблігіна. Мінськ : Аверсев, (Підручник вищої
школи), 2009. 255 с.
- 13.Словник української мови : в 11 тт. / АН УРСР. Інститут мовознавства;
за ред. І. К. Білодіда. К. : Наукова думка, 1970–1980. 699 с.
- 14.Смірнова Т. Машинний переклад у сучасному суспільстві. 2011. URL:
<https://er.nau.edu.ua/bitstream/NAU/47336/> (дата звернення: 30.10.2023).
- 15.Чернікова Л. Ф. Лінгвістична якість машинного перекладу. *Питання
духовної культури Філологічні науки*. 2013. С. 165-168.
- 16.Шилінська І. Ф. Використання систем машинного перекладу під час
навчання перекладу науково-технічних текстів М. З. Кузів *Науковий
вісник Міжнародного гуманітарного університету*. Серія : Філологія.
2022. Вип. 53(2). С. 149-152. URL:
[http://nbuv.gov.ua/UJRN/Nvmgu_filol_2022_53\(2\)_36](http://nbuv.gov.ua/UJRN/Nvmgu_filol_2022_53(2)_36) (дата звернення:
17.10.2023).
- 17.Щепка О. А. Сучасний публіцистичний текст: кваліфікаційні та
диференційні ознаки. *Науковий вісник Міжнародного гуманітарного
університету*. Сер.: Філологія. № 34 том 2. 2018. С. 103-106.
- 18.Янковець А. Особливості використання машинного перекладу в умовах
професійної підготовки майбутніх перекладачів. *Науковий вісник
Ужгородського національного університету*. Серія «Педагогіка.
Соціальна робота». Вип. 25. 2017. С. 233–236.
- 19.Abdujalil A. Ukraine war: Russia attacks Avdiivka stronghold in eastern
Ukraine. *BBC*. 12 October 2023. URL: <https://www.bbc.com/news/world-europe-67095103> (accessed: 18.10.2023).

20. Bar-Hillel Y. Some Linguistic Problems Connected with Machine Translation. *Philosophy of Science*, Vol. 20, No. 3. Jul. 1953. P. 217-225.
21. Brooks R. A Brief History of Machine Translation Technology, K International, November 25, 2020, URL: <https://www.k-international.com/blog/history-machine-translation/> (accessed: 17.10.2023).
22. Chérâgui M. A. Theoretical Overview of Machine translation. *University Ahmed Draia, Adrar, Algeria*. 2012. P. 160-169.
23. Climent, S., Moré, J., Oliver, A., Salvatierra, M., Sànchez, I., Taulé, M., & Vallmanya, L. Bilingual newsgroups in Catalonia: A challenge for machine translation. *Journal of Computer-Mediated Communication* 9 (1). 2003. URL: <http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2003.tb00360.x/full> (accessed: 18.11.2023)
24. Correa N. A Fine-grained Evaluation Framework for Machine Translation System Development. In: *Proceedings of MT Summit IX: Papers*, New Orleans, Louisiana. 2003. URL: <https://aclanthology.org/2003.mtsummit-papers.7.pdf> (accessed: 17.10.2023)
25. Denkowski M., Lavie A. METEOR-NEXT and the METEOR paraphrase tables: Improved evaluation support for five target languages. In Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR, *Association for Computational Linguistics*. 2010. P. 339–342.
26. Doddington G. Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In Proceedings of the second international conference on Human Language Technology Research (HLT '02). *Morgan Kaufmann Publishers Inc.*, San Francisco, CA, USA, 2002. P. 138–145.
27. Dorr B., Snover M., Madnani N. Part 5: Machine Translation Evaluation Chapter 5.1 Introduction. *GALE book*. 2010. P. 801-895
28. Dostert L. Brief History Of Machine Translation Research. *Eighth Annual Round Table Meeting on Linguistics and Language Studies, Georgetown University*. 1957. P. 3-10.

29. Dostert, L. Research in machine translation : Report of the 8th Annual Round table meeting on Linguistics and language studies. Washington, D.C. : Georgetown University Press, 1957. 193 p.
30. Flanagan, M. Error classification for MT evaluation. In *Proceedings of the First Conference of the Association for Machine Translation in the Americas*, Columbia, Maryland, USA, 1994. P. 65-72.
31. Framework for Standard Error Marking. *American Translators Association*. 2017. URL: <https://www.atanet.org/certification/how-the-exam-is-graded/error-marking/> (accessed: 14.11.2023)
32. Harris H. Machine translations revisited: issues and treatment protocol. *The Language Teacher » Readers' Forum*, May / June 2010. P. 25-29.
33. Hutchins J. Machine translation: a brief history. Concise history of the language sciences: from the Sumerians to the cognitivists. *Oxford: Pergamon Press*, 1995. P. 431-445.
34. Hutchins J. Machine translation: a concise history. *Journal of Translation Studies* 13 (1 & 2), 2010. P. 29-70.
35. Hutchins J. Reflections on the history and present state of machine translation. 1995. URL: <https://aclanthology.org/1995.mtsummit-1.6.pdf> (accessed: 17.10.2023).
36. Hutchins J. The history of machine translation in a nutshell. 2014. URL: <https://aclanthology.org/www.mt-archive.info/10/Hutchins-2014.pdf> (accessed: 07.09.2023).
37. Jakubková N., Warner A. A history of machine translation: from the typewriter to chatgpt in one century. *Nimdzi Insights*. 22 March 2023. URL: <https://www.nimdzi.com/history-of-machine-translation/> (accessed: 17.10.2023).
38. Khenglawt V. Machine translation and its approaches. *Advances in Engineering Research*, volume 178. *Mizoram Science Congress*. 2018. P.141-145.

39. Koehn P., Knowles R. Proceedings of the First Workshop on Neural Machine Translation, *Association for Computational Linguistics*. August 4, 2017. P. 28–39.
40. Lee S., Lee J., Moon H., Park C., Seo J., Eo S., Koo S., Lim H. A Survey on Evaluation Metrics for Machine Translation. *Department of Computer Science and Engineering, Korea University*. 2023. 22 p.
41. Loffler-Laurian A-M. Typologie des erreurs. In: *La Traduction Automatique. Presses Universitaires du Septentrion*. 1996. P. 95-107.
42. Moussallem D., Wauer M. & Axel-Cyrille Semantic Web for Machine Translation: Challenges and Directions. *Data Science Group, University of Paderborn, Germany*. 2019. P. 1-9.
43. Okpor M. D. Machine Translation Approaches: Issues and Challenges *International Journal of Computer Science Issues, Vol. 11, Issue 5, No 2, September 2014*. P.159-165.
44. Papineni K., Roukos S., Ward T., Zhu W-J. BLEU: A method for automatic evaluation of machine translation. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL 02), *Watson Research Center, Yorktown Heights*. 2002. P. 311–318.
45. Paul Adams Ukraine war: Cannabis in focus for veterans treatment. *BBC*. 27 September 2023. URL: <https://www.bbc.com/news/world-europe-66877717> (accessed: 30.09.2023).
46. Present continuous In: *English Grammar. British Council*. URL: <https://learnenglish.britishcouncil.org/grammar/english-grammar-reference/present-continuous> (accessed: 07.11.2023).
47. SAE J2450: Translation Quality Metric, *Society of Automotive Engineers, Warrendale, USA*, 2001. P. 1-12.
48. Shevchenko V. Ukraine fears drone shortages due to China restrictions. *BBC*. 22 October 2023. URL: <https://www.bbc.com/news/world-europe-67078089> (accessed: 23.10.2023).

49. Štefčík J. Evaluating machine translation quality: a case study of a translation of a verbatim transcription from Slovak into German. *Vertimo Studijos*. 2015. P. 139-153.
50. Stymne S. Blast: A tool for error analysis of machine translation output. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Systems Demonstrations. *Association for Computational Linguistics*. 2011. P. 56-61.
51. Tillmann C., Vogel S., Ney H., Zubiag A., Sawaf H.. Accelerated DP Based Search For Statistical Translation. *European Conference on Speech Communication and Technology*. 1997.
52. Tomedes Machine Translation Evaluation: The Ultimate Guide. 2022. URL: <https://www.machinetranslation.com/blog/machine-translation-evaluation-ultimate-guide> (accessed: 14.10.2023).

СПИСОК ІЛЮСТРАТИВНОГО МАТЕРІАЛУ

53. Абдуджаліл А. Росіяни пішли в масштабний наступ на Авдіївку. Що там відбувається. *BBC*. 13 жовтня 2023. URL: <https://www.bbc.com/ukrainian/articles/cje959wnpqpo> (дата звернення: 18.10.2023).
54. Пол Адамс "Канабіс допоміг пережити найтяжчі дні". Як ветерани війни з Росією тамують біль. *BBC*. 27 вересня 2023. URL: <https://www.bbc.com/ukrainian/features-66932898> (дата звернення: 30.09.2023).
55. Шевченко В. Полювання на дрони. Як обмеження Китаю на продаж безпілотників вплинули на війну в Україні. *BBC*. 22 жовтня 2023. URL: <https://www.bbc.com/ukrainian/articles/c16721l8egno> (дата звернення: 23.10.2023).

56. Costa, Angela & Ling, Wang & Luís, Tiago & Correia, Rui & Coheur, Luisa. A linguistically motivated taxonomy for Machine Translation error analysis. 2015. P. 1-34.
57. Elliott, D., Hartley, A., & Atwell, E. A fluency error categorization scheme to guide automated machine translation evaluation. In *Proceedings of the 6th Conference of the Association for Machine Translation in the Americas: Technical Papers*, Washington, USA. Springer, 2004. P. 64-73.
58. Noureldin A. & Abdulkhaliq A. Machine Translation: The Case of Arabic-English Translation of News Texts. In: *Theory and Practice in Language Studies*, Vol. 10, No. 4. *Academy Publication*. 2020. P. 408-418.
59. Vilar D., Xu J., d'Haro L. F., Ney H. Error analysis of statistical machine translation output. In: *Proceedings of the 5th international Conference of Language resources and Evaluation. European Language Resources Association*. 2006. P. 697-702.

SUMMARY

The presented paper is dedicated to the analysis of mistakes made by machine translation systems.

The object of the study is the translation obtained by using modern publicly available machine translation systems such as Meta, Google Translate and Deepl, and the subject is the peculiarities of machine translation.

The aim of this research paper is to interpret the term "machine translation" and to compare translations made with the help of translation services.

In order to achieve this goal, the following tasks need to be accomplished:

1. to analyze the terms "machine translation" and "translation" in general
2. describe the development of machine translation
3. to provide a comparative description of the translations made
4. summarize the collected results and draw up a conclusion

The data obtained from the study shows that the state of modern machine translation systems, such as Google Translate, Meta, and Deepl, is far from perfect. The analysis of the translations shows that there are significant errors in various aspects, such as omission of meaningful words, incorrect word choice, grammatical and stylistic flaws.

The peculiarities of machine translation identified in the study are that the systems often do not take into account the semantic connotations of words and have problems with contextual word choice. Grammatical and structural errors also reduce the quality of translation, making it not only inaccurate but also difficult to understand.

The practical significance of the study lies in the possibility of using its results to improve existing and develop new machine translation systems and to conduct practical English classes.


Keywords: *translation, machine translation, machine translation systems, neural networks, publicistic text.*

Декларація
академічної доброчесності
здобувача ступеня вищої освіти ЗНУ

Я, Давиденко Денис Володимирович, студент 2 курсу магістратури, форми навчання денної, факультету іноземної філології, спеціальність 035 Філологія, освітньо-професійна програма переклад (англійський), адреса електронної пошти limix.show@gmail.com - підтверджую, що написана мною кваліфікаційна робота на тему «СПЕЦИФІКА МАШИННОГО ПЕРЕКЛАДУ (НА МАТЕРІАЛІ ПУБЛІЦИСТИКИ АНГЛІЙСЬКОЮ ТА УКРАЇНСЬКОЮ МОВАМИ)» відповідає вимогам академічної доброчесності та не містить порушень, що визначені у ст. 42 Закону України «Про освіту», зі змістом яких ознайомлений/ознайомена;

- заявляю, що надана мною для перевірки електронна версія роботи є ідентичною її друкованій версії;

- згоден/згодна на перевірку моєї роботи на відповідність критеріям академічної доброчесності у будь-який спосіб, у тому числі за допомогою Інтернет-системи, а також на архівування моєї роботи в базі даних цієї системи.

Дата 04.12.2023 Підпис  ПІБ (студент) Давиденко Д. В.