

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені В. Н. КАРАЗІНА
ФАКУЛЬТЕТ ІНОЗЕМНИХ МОВ
Кафедра перекладознавства імені Миколи Лукаша

Рекомендовано до захисту
Протокол засідання кафедри № ____
від « ____ » листопада 2024 р.
Завідувач кафедри Олександр РЕБРІЙ

(підпис)

КВАЛІФІКАЦІЙНА МАГІСТЕРСЬКА РОБОТА
**ПРОБЛЕМИ СИСТЕМ МАШИННОГО ПЕРЕКЛАДУ
GOOGLE TRANSLATE ТА MICROSOFT TRANSLATE
ПРИ ПЕРЕДАЧІ УКРАЇНСЬКОЮ МОВОЮ
СПЕЦІАЛІЗОВАНИХ АНГЛОМОВНИХ ТЕКСТІВ**

Виконавець:

студент II курсу магістратури, групи АМП-61
Шевцов Данило Сергійович
(прізвище, ім'я, по батькові)

Керівник роботи:

Черноватий Леонід Миколайович
доктор пед. наук, проф.
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Підсумкова оцінка:

за національною шкалою:

кількість балів: _____

Підпис керівника

Кваліфікаційну магістерську роботу захищено на засіданні екзаменаційної комісії

Протокол № ____ від « ____ » _____ 2024 р.

Голова екзаменаційної комісії _____

(підпис)

(прізвище та ініціали)

Харків — 2024

СПИСОК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ

МП — машинний переклад

ОТ — оригінальний текст

ТП — текст перекладу

GT — Google Translate

MT — Microsoft Translator

ЗМІСТ

ВСТУП	4
РОЗДІЛ 1. СИСТЕМИ МАШИННОГО ПЕРЕКЛАДУ	8
1.1. Історія розвитку систем машинного перекладу.....	8
1.2. Порівняльні характеристики систем машинного перекладу <i>Google Translate</i> та <i>Microsoft Translate</i>	14
Висновки до розділу 1	15
РОЗДІЛ 2. ЕФЕКТИВНІСТЬ СИСТЕМИ МАШИННОГО ПЕРЕКЛАДУ <i>GOOGLE TRANSLATE</i>	16
Висновки до розділу 2.....	27
РОЗДІЛ 3. ЕФЕКТИВНІСТЬ СИСТЕМИ МАШИННОГО ПЕРЕКЛАДУ <i>MICROSOFT TRANSLATE</i>	29
Висновки до розділу 3.....	38
РОЗДІЛ 4. ПОРІВНЯЛЬНА ЕФЕКТИВНІСТЬ СИСТЕМ МАШИННОГО ПЕРЕКЛАДУ <i>GOOGLE TRANSLATE</i> та <i>MICROSOFT TRANSLATE</i>	40
Висновки до розділу 4.....	49
ЗАГАЛЬНІ ВИСНОВКИ	50
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	54
SUMMARY	58

ВСТУП

Актуальність теми дипломної роботи полягає в тому, що в умовах активного використання комп'ютерних технологій і процесів глобалізації машинні перекладачі, серед яких найпопулярніші — *Google Translate* та *Microsoft Translate*, перетворюються у важливі механізми для забезпечення зручного та швидкого здійснення якісного перекладу текстів на різні напрямки та тематики. Але під час здійснення перекладу на українську мову профільних англійських текстів виникає низка завдань, пов'язаних з якістю та точністю їх перекладу. Особливо актуально це для текстів технічної, юридичної, медичної та інших спеціалізацій, які включають в себе досить специфічну термінологію та контекстуальні нюанси. Тому дослідження даних труднощів та пошук методів їх подолання має важливе значення з точки зору як підвищення рівня оперативності машинного перекладу, так і правильності та відповідності змісту перекладів спеціалізованих текстів. Тема роботи є **актуальною** як для поліпшення роботи вже наявних систем автоматизованого перекладу, а також для вдосконалення лінгвістичних технологій в Україні.

Об'єкт дослідження — системи машинного перекладу *Google Translate* та *Microsoft Translate*. **Предмет дослідження** — проблеми та помилки, що допускаються при перекладі спеціалізованих англійських текстів (наукових) на українську мову.

Мета дослідження аналіз та виявлення основних проблем, з якими стикаються системи машинного перекладу, зокрема *Google Translate* та *Microsoft Translate*, при перекладі спеціалізованих англійських текстів українською мовою.

Відповідно до мети ставим наступні **завдання дослідження**:

1. Проаналізувати та охарактеризувати еволюцію машинного перекладу, зосереджуючись на основних етапах розвитку технологій.

2. Порівняти системи *Google Translate* та *Microsoft Translate* – виконати аналіз двох провідних систем машинного перекладу, враховуючи їх функціональні можливості, точність та особливості використання.
3. Оцінка ефективності *Google Translate* для різних типів текстів, оцінюючи точність та відповідність вихідного тексту перекладеному.
4. Оцінка ефективності *Microsoft Translate* та його здатність точно відтворювати зміст різних текстів.
5. Виконати порівняльний аналіз ефективності *Google Translate* та *Microsoft Translate*: висвітлити основні переваги та недоліки кожної системи, порівнюючи їх ефективність у різних контекстах.

Матеріалом дослідження слугували дві програми машинного перекладу (*Google Translate* та *Microsoft Translate*), англomовна стаття у галузі перекладознавства [27] та тексти її перекладу згаданими програмами машинного перекладу.

Для досягнення мети дослідження були застосовані такі **методи**: добір матеріалів для аналізу, порівняння та аналіз текстів перекладу, класифікація виявлених типів помилок, систематизація теоретичних даних у вигляді таблиць, групування помилок за їхнім впливом на точність передачі змісту оригінального тексту, а також узагальнення отриманих результатів у висновках.

Положення, що виносяться на захист:

1. Загальна якість перекладу програмою *Google Translate*, якщо порівнювати її з перекладами студентів текстів такого ж обсягу (понад кілька тисяч друкованих знаків), є на дуже низькому рівні – 14,25 штрафних бали. Тобто, за шкалою оцінювання, яка використовується на кафедрі перекладознавства імені Миколи Лукаша, такий результат відповідає оцінці 1,5 за п’ятибальною шкалою.

2. Середнє значення показника штрафних балів (12,75) вказує на те, що якість перекладу програми МП *Microsoft Translate* також знаходиться на низькому рівні. Якщо порівняти середнє значення показника штрафних балів програми

Microsoft Translator із згаданою шкалою, можна зробити висновок, що цей переклад відповідає оцінці за п'ятибальною шкалою («незадовільно»).

3. Програми *Google Translate* та *Microsoft Translate* відносяться до різних типів – на основі правил та нейронний відповідно – і тексти перекладів обох програм є не надто схожими – дослівний збіг згаданих текстів коливається від до 73% (середнє значення – 61%). Незважаючи на це, їхня ефективність є приблизно однаковою і на досить низькому рівні. Хоча що за загальною кількістю штрафних балів незначну перевагу має програма *Microsoft Translate*, однак різниця є несуттєвою (1,5 штрафного балу). Таким чином, тексти перекладу обох програм потребують суттєвого постредагування, витрати часу на яке можуть не виправдати їхнього застосування.

Наукова новизна роботи полягає в комплексному аналізі перекладу спеціалізованих англійських текстів (наукових, технічних, медичних) на українську мову за допомогою систем машинного перекладу. Вперше проведено порівняння якості перекладів *Google Translate* і *Microsoft Translate* з акцентом на багатозначність, термінологічну точність та передачу складних граматичних конструкцій.

Теоретична цінність дослідження полягає в аналізі особливостей машинного перекладу спеціалізованих текстів, зокрема наукових, технічних, медичних і юридичних, що мають специфічну термінологію. Дослідження дає змогу оцінити, як *Google Translate* і *Microsoft Translate* обробляють контекст і термінологію, а також сприяє розвитку теорії машинного перекладу, зокрема в питаннях багатозначності, контекстуального перекладу та адаптації моделей до мовних пар, як-от англійська–українська.

Практична цінність дослідження полягає в удосконаленні машинного перекладу спеціалізованих англійських текстів на українську, таких як юридичні, медичні, технічні та наукові. Результати можуть сприяти підвищенню точності перекладу, зменшенню кількості помилок та оптимізації роботи *Google Translate* і

Microsoft Translate з вузькотермінологічними текстами. На основі виявлених недоліків можна розробити рекомендації для перекладачів і користувачів, що допоможуть уникати типових помилок і краще адаптувати переклад під особливості української мови.

Апробація отриманих результатів проводилася на студентській науковій конференції кафедри перекладознавства імені Миколи Лукаша Харківського національного університету імені В.Н. Каразіна 20 листопада 2024 року. Результати дослідження викладено у зданій до друку статті в збірнику студентських праць кафедри перекладознавства імені Миколи Лукаша Харківського національного університету імені В.Н.Каразіна *In Statu Nascendi* (випуск 24).

Структура магістерського дослідження. Кваліфікаційна робота складається зі вступу, одного теоретичного розділу та трьох практичних, висновків до кожного розділу та загальних висновків, списку використаних джерел та анотації. Загальна кількість сторінок роботи становить – 65 сторінок. Основний зміст роботи викладено на 58 сторінках. В роботі використано 30 джерел літератури, з яких іноземною мовою – 12.

РОЗДІЛ 1

СИСТЕМИ МАШИННОГО ПЕРЕКЛАДУ

1.1. Історія розвитку систем машинного перекладу

Ключовими поняттями в галузі перекладацьких систем є автоматизований переклад (Computer-Aided Translation / Computer-Assisted Translation, CAT) або машинний переклад (Machine Translation, MT) [3]. Метою автоматизованого перекладу є надання перекладачу допомоги в його роботі за допомогою автоматизованих програм, а машинний переклад — це повний і частковий переклад тексту за допомогою комп'ютера без активної участі людини [3]. У ширшому контексті можна розглядати машинний переклад як галузь наукових досліджень, що сполучає у собі лінгвістику, математику з кібернетикою і ставить за мету створення системи автоматизації перекладацького процесу [25].

Розвиток машинного перекладу бере свій початок із середини 20-го століття, коли зародилась концепція автоматичного перекладу текстів з різних мов без втручання людини [30]. Незважаючи на те, що завдання створення системи ідеального перекладу ще не вирішено, розвиток галузі набирає суттєвих темпів, та й самі спроби її вирішення сприяли появі нових лінгвістичних підходів і теорій.

Одною з основних подій на шляху розвитку машинного перекладу є формулювання його концепції в 1947 році Уорреном Вівером, директором департаменту з питань природничих наук Фонду Рокфеллера [24]. Свій подальший розвиток ця теорія одержала у відомій праці, в котрій він звернув свою увагу на можливості використання в перекладі математичних і лінгвістичних методів [29].

Досягнення сучасного машинного перекладу, є істотним проривом, але, ці системи не можуть поки остаточно усунути ручний переклад. Не зважаючи на

велику точність перекладу загальних текстів, у таких спеціалізованих напрямках, як медицина чи юриспруденція, все ще мають місце похибки, пов'язані з невірною інтерпретацією термінології та контексту [4]. Подальший розвиток машинного перекладу вимагає ретельнішого врахування особливостей мов, а також регулярного вдосконалення алгоритмів функціонування штучного інтелекту.

Становлення машинного перекладу (МП) пройшло тривалий і значний етап еволюції, починаючи з найперших систем, що функціонували виключно на рівні словників, та закінчуючи найсучаснішими технологіями, які дозволяють глибоко досліджувати зміст перекладеного тексту. Ця динаміка відображається в кількох етапах розвитку систем машинного перекладу.

У першому поколінні систем, що існували майже до середини 1960-х років, застосовувався механізм машинного перекладу, який базувався на принципах дослівного перекладу. Кожен вираз у тексті мав відповідник на основі словникових даних, що знаходився у словнику іншої мови. Втім, такі системи були не здатні впоратися з багатоваріантністю значень слів та не виконували лінгвістичного аналізу. Як наслідок, часто залишалася невисока якість перекладу, оскільки часто ігнорувалися важливі аспекти контексту [15].

Друге покоління систем машинного перекладу, яке з'явилося в середині 1960-х і в 1970-х роках, було трохи складнішим. Ці системи були створені на морфологічних зв'язках між мовами вхідного та вихідного текстів, що гарантувало кращу якість перекладу, особливо при перекладі між спорідненими мовами. Ці «морфологічні системи» могли краще відображати структуру тексту, хоча вони все одно були досить обмеженими з точки зору багатозначності та синтаксичного аналізу [15].

На основі концепції Уоррена Вівера був розроблений принцип машинного перекладу (МП), який був заснований на принципі «інтерлінгвізму» (interlingua) [30]. Ця модель припускала двоетапний процес виконання перекладу: перший

етап — переклад тексту на спеціальну мову—посередник (на основі спрощеної англійської мови), а вже потім переклад здійснювався на кінцеву мову перекладу.

У тому ж 1947 році Ендрю Бут і Дональд Бріттен розробили детальний «код» для дослівного машинного перекладу. Вже наступного року Реджинальд Річенс пропонує розділити словоформи на основу та закінчення для полегшення перекладу.

У той час комп'ютери значно різнилися від сучасних — це були об'ємні, дорогі девайси, вони займали величезні приміщення та для їх обслуговування потрібен був цілий штат фахівців з технічного обслуговування. Головне їх використання зводилося лише до виконання математичних розрахунків, переважно для військових потреб і наукових інституцій, які займаються дослідженнями, пов'язаними з оборонною галуззю. Як наслідок, ранні розробки в області машинного перекладу активно фінансувалися військовими відомствами. У США найбільший акцент було зроблено на перекладі з російської на англійську, а в СРСР, навпаки, з англійської на російську.

Найпершою вагомою подією в історії МП є конференція в Массачусетському технологічному інституті в 1952 році. А вже в 1954 році в Нью-Йорку компанія ІВМ у співпраці з Джорджтаунським університетом презентували свою першу систему машинного перекладу ІВМ Mark II, яка увійшла в історію під назвою «Джорджтаунський експеримент» [1]. Така система займалася лише перекладом з російської на англійську мову і була обмеженою у словниковому запасі — близько 250 слів і лише 6 граматичних правил. Хоча на той час видавалося, що створення високоякісного автоматичного автоперекладача — справа лише кількох років, ця технологія була досить складною. На той час розробники робили ставку саме на повністю автоматичні системи, вбачаючи в участі людини лише тимчасові компроміси. Це стало причиною занепокоєння з боку професійних перекладачів, котрі почали хвилюватися, чи не призведе їх автоматизація до звільнення працівників із роботи.

Але в 1950-х роках науково-дослідні організації в США та Європі, що займалися питаннями машинного перекладу, зіткнулися з серйозними проблемами. Основні причини низької якості перекладів були в обмежених технічних можливостях тогочасних комп'ютерів: малий обсяг вбудованої пам'яті, слабка продуктивність та відсутність можливості застосовувати мови програмування високого рівня. Також, значною перешкодою стала відсутність належної теоретичної бази щодо вирішення складних лінгвістичних проблем. Дуже часто переклад текстів проводився без урахування його синтаксичної структури та контексту, а це значно погіршувало якість перекладу [15].

Важливою сходинкою в розвитку машинних перекладачів було запровадження підходу статистичного машинного перекладу (SMT — Statistical Machine Translation, SMT) [21]. Цей напрям здобув широку популярність ще наприкінці 1980-х — на початку 1990-х років. В його основу покладено статистичне оцінювання імовірності різних можливих перекладацьких варіантів на підставі великих корпусів двомовних текстів за спеціальними статистичними моделями [21]. Головною ідеєю цього методу є покладання на статистичні закономірності, встановлені в паралельних текстових базах даних, на протипагу лінгвістичним правилам, які розробляються вручну [29].

Одними з фундаментальних наукових розробок, які поклали початок розвитку статистичного машинного перекладу, була робота П. Ф. Брауна та його співавторів [21]. У своїй роботі вони використали метод статистичної теорії прийняття рішень для аналізу мови, що стало значущим кроком у застосуванні статистичних методів у царині машинного перекладу [29]. Підхід цих авторів базується на обчисленнях вірогідності виконання перекладу відповідно до оригіналу для того, щоб вибрати якомога оптимальніший його варіант.

При машинному перекладі мовна модель має ключове значення для забезпечення природності й точності перекладу [24]. Ця модель визначає можливість використання певних словосполучень і речень, дозволяючи програмі

спрогнозувати правильну послідовність слів в тексті оригіналу і перекладу. Вихідна модель мови допомагає зрозуміти текст, тоді як модель мови перекладу дає змогу згенерувати потрібний переклад.

Разом з ними працюють перекладацькі моделі, котрі обчислюють ймовірність передачі слів або словосполучень з мови оригіналу на мову перекладу. Для цього використовуються статистичні дані з корпусів двомовних матеріалів, що дозволяє системі встановлювати найкращі відповідники між мовами, знаходячи закономірності в паралельних текстах [12].

Машинний статистичний переклад (SMT) дозволив суттєво збільшити якість перекладу порівняно з підходом, що ґрунтувався виключно «на правилах», оскільки він дозволяє системам навчатися на базі даних і легше працювати з лінгвістичними варіаціями. Утім, такий вид перекладу має певні недоліки, включно з проблемами розшифрування та дотримання правильності порядку слів у перекладі [13]. Незважаючи на це досягнення, у галузі машинного перекладу сталася певна зміна моделі з виходом нейронного машинного перекладу (НМП), у якому використовують штучні нейронні мережі, що дозволяє набагато збільшити якість і швидкість перекладу.

За допомогою нейронного машинного перекладу здійснюється перетворення вихідного речення на переклад цільовою мовою за участю штучних нейронних мереж, завдяки цьому забезпечується більш точний і відповідний до змісту контексту переклад.

Головний привід для створення НМПроз – виток технологій глибинного пізнання та створення нейронних мереж [2]. На початку 2010-х років глибинне навчання, початково застосоване в ігрових технологіях, отримало велику популярність в різноманітних програмах опрацювання природної мови [24].

Справжній прорив у цій царині був здійснений з запровадженням архітектур Sequence-to-sequence, які також відомі як кодувальні-декодувальні алгоритми (encoder-decoder architecture). Вони застосовують зворотні нейронні

мережі (RNN) для шифрування вихідного речення, а тоді декодування його в перекладений текст. Моделі продемонстрували свою ефективність у процесі написання зв'язних і контекстуально вірних перекладів, а також використовуються для розпізнавання тексту і мовлення.

Одна з визначних дат розвитку нейронного машинного перекладу — запуск в 2007 році сервісу *Google Translate*. Від початку система була заснована лише на статистичних моделях, що використовують дослівний переклад. Але у 2016 році *Google* перейшов на нейронні моделі, що, за допомогою глибокого навчання та нейронних мереж, істотно підвищило якість перекладу. Нейронні моделі, на відміну від статистичних, безпосередньо вивчають співвідношення між оригінальними та перекладними реченнями, завдяки чому значно спрощується процес перекладу.

Із плином часу нейрологічні програми довели спроможність ліпше сприймати контекст і розпізнавати складні мовні структури, що дозволило суттєво підвищити точність перекладів. Завдяки успіху сегментних моделей Seq2seq і розробці нових складних архітектур вдалося досягти стрімкого прогресу в області нейронного машинного перекладу, що зробило його основним підходом у цій галузі.

1.2. Порівняльні характеристики систем машинного перекладу *Google Translate* та *Microsoft Translate*

Нині є безліч різних систем машинного перекладу, кожна з яких має власні недоліки та переваги. До найбільш поширених та відомих відносяться *Google Translate* та *Microsoft Translator*. Ці системи використовують нейронні мережі та алгоритми машинного навчання при перекладі текстів з однієї мови на іншу. Наведемо нижче їхні ключові переваги та недоліки у форматі порівняльної таблиці [14].

Табл. 1.1. Порівняння систем машинного перекладу *Google Translate* та *Microsoft Translate*

Критерій	<i>Google Translate</i>	<i>Microsoft Translate</i>
Підтримка мов	Підтримує понад 100 мов, що дозволяє перекладати тексти, відео, аудіо та зображення.	Підтримує понад 60 мов, також дозволяє перекладати тексти, відео та аудіо.
Якість перекладу	Висока якість перекладу для популярних мов (англійська, французька, німецька).	Висока точність для основних мов, але може бути менш точною для рідкісних мов.
Функціональність	Дозволяє перекладати документи, веб-сторінки, зображення. Є режим перекладу в реальному часі.	Також підтримує переклад документів, веб-сторінок та режим реального часу для розмов.
Доступність	Безкоштовний доступ до більшості функцій.	Безкоштовний доступ, але деякі функції, як переклад веб-сторінок та документів, доступні лише у платній версії.
Інтеграція з продуктами	Інтегрується з <i>Chrome</i> , <i>Gmail</i> , <i>Android</i> та іншими продуктами <i>Google</i> .	Інтегрується з продуктами <i>Microsoft</i> , такими як <i>Office</i> , <i>Skype</i> , <i>Edge</i> та інші.
Точність перекладу менш популярних мов	Менш точний переклад для деяких мов, що не є поширеними.	Теж має нижчу точність перекладу для малопоширених мов, але підтримка мов обмеженіша, ніж у <i>Google Translate</i> .
Офлайн-режим	Підтримує офлайн переклад для обмеженої кількості мов.	Обмежений офлайн-переклад, що підтримує менше мов, ніж онлайн-версія.
Додаткові можливості	Можливість перекладу зображень, текстів, аудіо та відео, функція «Розмова» для перекладу в реальному часі.	Також має функцію «Розмова» для перекладу в реальному часі, але деякі можливості обмежені у безкоштовній версії.

ВИСНОВКИ ДО 1 РОЗДІЛУ

Розвиток систем машинного перекладу пройшов значний шлях від перших спроб дослівного перекладу на базі словникових даних до сучасних нейронних мереж, що забезпечують якісний і точний переклад. Зародження машинного перекладу почалося в середині ХХ століття, а ключовою подією стало формулювання концепції Уорреном Вівером у 1947 році [22]. Історія розвитку цієї галузі включає кілька етапів, які позначені вдосконаленням методів перекладу: від механізмів дослівного перекладу до статистичних моделей та впровадження штучних нейронних мереж.

Значний прорив у якості перекладу відбувся з розвитком статистичного машинного перекладу (SMT), а пізніше — з появою нейронного машинного перекладу (NMT), який заснований на глибокому навчанні. НМП дозволив досягти високого рівня адекватності перекладу завдяки здатності аналізувати контекст і адаптуватися до багатозначності мовних структур [1].

Однак, незважаючи на суттєвий прогрес, машинний переклад все ще стикається з викликами, такими як точність перекладу спеціалізованих текстів у сферах медицини чи юриспруденції [3]. Подальший розвиток цієї галузі залежить від удосконалення алгоритмів штучного інтелекту, врахування специфіки мов та адаптації до різних контекстів.

РОЗДІЛ 2

ЕФЕКТИВНІСТЬ СИСТЕМИ МАШИННОГО ПЕРЕКЛАДУ GOOGLE TRANSLATE

У другому розділі досліджується аналіз перекладу тексту, виконаного за допомогою програми *Google Translate (GT)* [*Google Translate*]. Оригінальним текстом (ОТ) була англomовна стаття [27], присвячена дослідженню оперативної пам'яті в умовах синхронного та послідовного перекладу, загальним обсягом понад кілька тисяч друкованих знаків.

На першому етапі стаття була перекладена українською мовою за допомогою автоматизованого перекладача *Google Translate*. Далі цей текст перекладу (ТП) був відредагований автором даної роботи, і на третьому етапі обидві версії — машинний переклад (ТП МП) та його відредагована версія — були порівняні з метою виявлення та класифікації помилок, допущених у машинному перекладі.

Для аналізу помилок використовувалася типологія, запропонована в науковій літературі [16], яка включає чотири основні типи помилок:

- Перший тип (1 штрафний бал) — це помилки, що спричиняють втрату або спотворення основної інформації оригінального тексту.
- Другий тип (0,5 штрафного бала) — це неясності, які можуть спричинити нерозуміння, особливо якщо читач не має необхідних предметних знань.
- Третій тип (0,25 штрафного бала) — це термінологічні помилки, які можуть вплинути на розуміння тексту, якщо читач не володіє певною термінологією.
- Четвертий тип (0,1 штрафного бала) — це мовні помилки (граматичні, лексичні, стилістичні), які не впливають на загальне розуміння тексту.

Саме третій тип помилок — термінологічні (0,25 штрафного бала) – було введено нами під час проведення наукового дослідження. Ці помилки виникали тоді, коли автоматичний перекладач використовував загальноживані слова або інші терміни замість конкретних термінів, що призводило до збереження змісту, але порушувало термінологічну точність.

Для полегшення порівняння, ТО було розділено на фрагменти, ТП яких послідовно аналізувалися. Далі (див. табл. 2.1) подано ТО фрагмента 1 (1648 друкованих знаків без пробілів), текст його перекладу програмою МП *GT* та редагований його варіант.

Таблиця 2.1. Класифікація помилок у перекладі фрагмента 1 тексту оригіналу програмою *Google Translate*.

Текст оригіналу Фрагмент 1	Текст перекладу МП (Google Translate)	Постредагований переклад
<p>Working memory in simultaneous and consecutive interpreting Barbara Moser-Mercer</p> <p>Introduction The concept of working memory (WM) supposes a dedicated limited capacity system underlying human thought that processes, maintains and stores information during a short period of time and which provides an interface between perception, long-term memory (LTM) and action (Baddeley 2003). It has been considered a limited capacity system from the early days of its conception when in the 1960s and 1970s linear flow models of information processing described the transfer of verbal information through rehearsal from a limited-capacity short term memory store to long-</p>	<p>Оперативна пам'ять при синхронному та послідовному перекладі Барбара Мозер-Мерсер</p> <p>Вступ Концепція робочої пам'яті (WM) передбачає спеціальну систему обмеженої ємності, що лежить в основі людської думки, яка обробляє, підтримує та зберігає інформацію протягом короткого періоду часу і яка забезпечує інтерфейс між сприйняттям, довготривалою пам'яттю (LTM) і діями (Baddeley 2003). Її вважали системою з обмеженою ємністю з перших днів її задуму, коли в 1960-х і 1970-х роках моделі лінійного потоку обробки інформації описували передачу словесної інформації шляхом репетиції з короткочасної пам'яті обмеженої ємності в довготривалу пам'ять. Сьогодні</p>	<p>Оперативна пам'ять у (1) синхронному та послідовному перекладі Барбара Мозер-Мерсер</p> <p>Вступ Концепція робочої пам'яті (РП) (2) розглядає (3) її як спеціалізовану (4) систему обмеженої ємності, що забезпечує підґрунтя (5) людського мислення (6), яка обробляє, підтримує та зберігає інформацію протягом короткого періоду часу (7) і яка забезпечує зв'язок (8) між сприйняттям, довготривалою пам'яттю (LTM) (9) і діями (Baddeley 2003). Її вважали системою з обмеженою ємністю з перших днів введення в науковий обіг цього поняття (10), коли в 1960-х і 1970-х роках лінійно-потоківі моделі (11) обробки інформації описували передачу (шляхом повторення</p>

<p>term memory. Researchers today describe working memory as a dynamic system that includes both maintaining and manipulating information through a series of interactive components that include executive control and attentional resources and the processing of both verbal and visual material (Wingfield 2016). These limitations are of particular importance in spoken language comprehension, especially when the stimuli have complex linguistic structures or listening effort is increased by poor acoustic quality or reduced hearing acuity (Wingfield 2016). The term working memory itself seems to have been coined by Miller, Galanter and Pribram (1960) who described the dynamic role of a transient memory system they termed working memory; the term did not attract much attention until Baddeley and Hitch (1974) provided convincing experimental evidence of working memory as a central component of human cognition. As definitions of working memory have evolved over the years, there appears to be consensus on its central characteristic, that of retaining information in conscious awareness by way of rehearsal for a short period of time when it is not or no longer present in the environment, of manipulating that information and of using it to guide behavior (Postle 2006).</p>	<p>дослідники описують робочу пам'ять як динамічну систему, яка включає як збереження інформації, так і маніпулювання нею за допомогою ряду інтерактивних компонентів, які включають виконавчий контроль і ресурси уваги, а також обробку словесного та візуального матеріалу (Wingfield 2016). Ці обмеження мають особливе значення для розуміння розмовної мови, особливо коли стимули мають складну лінгвістичну структуру або зусилля на прослуховування збільшуються через погану акустичну якість або знижену гостроту слуху (Wingfield 2016). Сам термін робоча пам'ять, здається, був винайдений Міллером, Галантером і Прибрамом (1960), які описали динамічну роль тимчасової системи пам'яті, яку вони назвали робочою пам'яттю; цей термін не привернув особливої уваги, доки Баддлі та Хітч (1974) не надали переконливих експериментальних доказів того, що робоча пам'ять є центральним компонентом людського пізнання. Оскільки визначення робочої пам'яті змінювалися протягом багатьох років, здається, існує консенсус щодо її центральної характеристики, збереження інформації у свідомому стані шляхом репетиції протягом короткого періоду часу, коли її немає або більше немає в навколишньому середовищі, маніпулювання цією інформацією та використання її для керування поведінкою (Postle 2006).</p>	<p>(12) вербальної (13) інформації з короткочасної пам'яті обмеженої ємності в довгочасну (14) пам'ять. Сьогодні дослідники описують робочу пам'ять як динамічну систему, яка включає як збереження інформації, так і маніпулювання нею за допомогою ряду інтерактивних компонентів, які включають виконавчий контроль і ресурси уваги, а також обробку вербального (15) та візуального матеріалу (Wingfield 2016). Ці обмеження мають особливе значення для розуміння усної (16) мови, особливо коли стимули мають складну мовну (17) структуру або зусилля для аудіювання (18) збільшуються через погану акустичну якість або знижену гостроту слуху (Wingfield 2016). Сам термін робоча пам'ять, здається, був запропонований (19) Міллером, Галантером і Прибрамом (1960), які описали динамічну роль перехідної (20) системи пам'яті, яку вони назвали робочою пам'яттю; цей термін не привернув особливої уваги, доки Баддлі та Хітч (1974) не надали переконливих експериментальних доказів того, що робоча пам'ять є центральним компонентом людського пізнання. Оскільки дефініції (21) робочої пам'яті змінювалися протягом багатьох років, здається, існує консенсус щодо її головної (22) характеристики, тобто утримання (23) інформації у світлій плямі свідомості (24) шляхом її повторення (25) протягом короткого періоду (26), коли вона вже (27)</p>
---	--	---

		відсутня (28) у довідці (29), а також (30) маніпулювання цією інформацією та використання її для керування поведінкою (Postle 2006).
--	--	--

Загалом, з наведеної таблиці впливає наявність 30 неточностей допущених програмою *Google Translate (GT)*.

Помилка 1 мовна граматична, (тип 4) — невірний вибір прийменника. *GT* переклав прийменник «in» як «при», що є не дослівною передачею, натомість має бути дослівно перекладений прийменник «у», який набагато чіткіше передає розуміння змісту. В даному випадку маємо 0,1 штрафний бал.

Помилки 2 і 9 – це неясність (тип 2) ми бачимо використання аббревіатури мовою оригіналу без пояснення, що ускладнює розуміння для читачів, які не знайомі з термінологією оригіналу. Тобто, залишення аббревіатури LTM без її розшифрування у тексті перекладу, а також переклад WM може заплутати людину, яка не спеціалізується на відповідній термінології. В сумі тут маємо 1 штрафний бал, так як маємо 2 помилки другого типу. Помилки 3 та 4 – це теж неясність (тип 2) — використання двозначних або неточних слів. Це може спантеличити читача, тому що вибрані слова мають кілька значень, і контекст не завжди однозначно вказує на правильне тлумачення. Наприклад, «передбачає» може бути зрозумілим як «прогнозує», тоді як в даному контексті краще вжити «розглядає». Також «спеціальна» має бути замінена на «спеціалізована», щоб точно передати функцію робочої пам'яті. В даному випадку маємо теж 1 бал (2 по 0,5) так як це другий тип помилки.

Помилка 5 мовна лексична, (тип 4) – тут вибір невдалого словосполучення «лежить в основі», що не відповідає контексту. Це може призвести до втрати точності перекладу або навіть викривлення сенсу, тож тут краще вжити «забезпечує підґрунтя», що відповідає також стилістичним нормам наукового тексту. Тому в даному випадку рахуємо 0,1 штрафний бал.

Помилка 6 термінологічна, (тип 3) — замінено термін на загальноповживане слово, тобто «мислення» краще замінити на «думку», що більше відповідає термінологічному контексту. Тут маємо 0,25 штрафного бала.

Помилка 7 мовна стилістична, (тип 4) – тавтологія, тобто повторення слова або ідеї, що є зайвим. Так, «протягом короткого періоду часу» — слово «часу» є зайвим і не додає нової інформації. Тут маємо 0,1 штрафний бал.

Помилка 8 термінологічна, (тип 3) — використання терміна «інтерфейс» замість більш зрозумілого «зв'язок», що спрощує і робить текст зрозумілішим. Зараховуємо 0,25 штрафного бала.

Помилка 10 мовна стилістична, (тип 4) — дослівний переклад, що не враховує особливості мови перекладу. Такий підхід може призвести до неестетичного і важкого для розуміння тексту, так як слово «задум» не передає думку до кінця, краще вжити «ведення в науковий обіг цього поняття». Тут маємо 0,1 штрафний бал.

Помилка 11 термінологічна, (тип 3) — невдала модель утворення терміна. Це стосується некоректного утворення нового терміна, який не відповідає загальноприйнятим правилам або логіці, тобто «моделі лінійного потоку» варто замінити на більш логічний варіант «лінійно-потоківі моделі». В даному випадку зараховуємо 0,25 штрафного балу.

Помилка 12 інформаційна, (тип 1) — нерозуміння терміна або його викривлення. Це може призвести до того, що перекладена інформація стає неточною або повністю змінює сенс оригіналу, бо слово «репетиція» є абсолютно не доречним, замість нього варто вживати більш підходяще «повторення». Зараховуємо 1 штрафний бал.

Помилки 13 та 15 є термінологічні, (тип 3) — використання загальноповживаних слів замість спеціалізованих термінів. Тобто, заміна терміна «вербальний» на «словесний» спрощує текст, але втрачає наукову точність. Тут зараховуємо 0,5 штрафного балу. Помилка 14 термінологічна, (тип 3) —

використання неточних термінів, що не відображають точного значення термінології, що ускладнює розуміння тексту фахівцями в цій сфері, так замінено «короткочасна» на «довготривала пам'ять». Маємо 0,25 штрафного балу. Помилка 16 термінологічна, (тип 3) — використання неточного терміну «розмовної мови», де краще вжити більш конкретне слово «усної» мови, що спрощує текст, але робить його менш науковим і точним. Зараховуємо 0,25 штрафного балу. Помилки 17 та 18 термінологічні, (тип 3) – використання загальноживаних слів замість термінів «лінгвістичну» варто замінити на «мовну», а «прослуховування» на «аудіювання». Тут рахуємо 0,25 штрафного балу. Помилка 19 термінологічна, (тип 3) – використання недостовірного терміна «винайдений», замість якого має бути вжитий «запропонований», що набагато більше підходить по стилю. Рахуємо 0,25 штрафного балу. Помилка 20 термінологічна (тип 3) – використання неточного терміна «тимчасової» потрібно замінити на більш точне слово «перехідної». Маємо 0,25 штрафного балу. Помилка 21 термінологічна, (тип 3) – використано загальноживане слово «визначення» замість терміна «дефініція». Маємо 0,25 штрафного балу.

Помилки 22 та 23 мовні, (тип 4) – відбувається порушення сполучуваності слів і використання неточного слова, що порушує зв'язність тексту та викликає неточності в сенсі. Так в 22 вжито «центральної» замість «головної», а в 23 вжито неточне відтворення контексту «збереження» замість «утримання». Тому тут рахуємо 0,1 та 0,1 штрафних балів. Помилка 24 термінологічна, (тип 3) – вживання неточного терміну, що спотворює зміст і робить його менш зрозумілим для наукового контексту. Тобто, замість «у свідомому стані» коректніше вжити «у світлій плямі свідомості». Рахуємо 0,25 штрафного балу.

Помилка 25 інформаційна, (тип 1) – тут відбувається викривлення інформації адже слово «репетиція» не має в даному контексті нічого спільного із словом «повторення». Рахуємо 1 штрафний бал.

Помилка 26 мовна, (тип 4) – в даному випадку маємо тавтологію «період часу» варто просто замінити на «період». Тут маємо 0,1 штрафного балу. Помилки 27 та 29 мовні, (тип 4) — використано зайві слова, які не додають нової інформації і створюють зайве навантаження на текст. Тобто «навколишнє середовище» є надто громіздким для даного тексту, тому доречно замінити його на слово «довкілля». Помилки 28 та 30 мовні, (тип 4) — відхилення від академічного стилю і додавання слів для забезпечення когезії речення, що іноді призводить до ускладнення тексту або робить його менш чітким.

Числові результати проведеного аналізу наводимо в таблиці 2.2:

Таблиця 2.2. Кількість та питома вага помилок в ТП *Google Translate* за фрагментами і типами (кількість помилок/питома вага (% від загальної кількості помилок)).

Система МП <i>Google Translate</i>	Типи помилок				Усього помилок	Штрафних балів
	I тип	II тип	III тип	IV тип		
Фрагмент 1	2/7	4/13	13/43	11/37	30	8,35

Як впливає з табл. 2.2, загальна кількість помилок у фрагменті 1, є в кількості 30, з яких помилки 1 типу становлять 7%, 2 типу — 13%, 3 типу – 43% та 4 типу – 37%. Загальна оцінка ТП МП цього фрагмента складає 8,35 штрафних балів.

Далі подано результати аналізу перекладу фрагмента 2 (див. Табл. 2.3.).

Таблиця 2.3. Класифікація помилок у перекладі фрагмента 2 тексту оригіналу програмою *Google Translate*.

Текст оригіналу Фрагмент 2	Текст перекладу МП (<i>Google Translate</i>)	Постредагований переклад
WM models and their link to interpreting Linear-stage or information processing	Моделі WM та їх зв'язок з інтерпретацією Лінійно-стадійні або моделі обробки інформації	Моделі ВМ(1) та їх зв'язок з інтерпретацією Лінійно-стадійні моделі або моделі обробки інформації

<p>models</p> <p>Early references to short-term memory (STM) made it clear that it was designed to temporarily hold a limited amount of information for a brief period and for immediate recall before the information would be transferred to long-term memory (LTM) using rehearsal processes. Miller (1956) speculated that the STM capacity limitation could be enhanced through chunking incoming information, which was to be accomplished through connecting incoming information with information already stored in LTM. Waugh and Norman (1965) added to the discussion that STM was not only constrained by the passage of time, but also by interference from newly incoming information.</p> <p>Linear-stage or information processing models dominated the next two decades. These models illustrated how information would flow from one processing stage to the next and how the information itself would change in the process. STM occupied a central position in these models and was described as a sort of buffer in which information would be temporarily stored, for at most 18 seconds, before being semantically coded and moved on to LTM (Cowan 2008). Without</p>	<p>Ранні згадки про короткочасну пам'ять (STM) показали, що вона була розроблена для тимчасового утримання обмеженої кількості інформації протягом короткого періоду часу та для негайного виклику, перш ніж інформація буде передана в довготривалу пам'ять (LTM) за допомогою репетиційних процесів. Міллер (1956) припустив, що обмеження пропускної здатності STM може бути посилено шляхом поділу вхідної інформації на фрагменти, що мало бути досягнуто шляхом з'єднання вхідної інформації з інформацією, яка вже зберігається в LTM. Во і Норман (1965) додали до обговорення, що STM був обмежений не лише плином часу, але й перешкодами від нової інформації, що надходить.</p> <p>Лінійно-стадійні або моделі обробки інформації домінували в наступні два десятиліття. Ці моделі ілюстрували, як інформація буде проходити від одного етапу обробки до наступного і як сама інформація буде змінюватися в процесі. STM займав центральне місце в цих моделях і описувався як свого роду буфер, у якому інформація тимчасово зберігалася протягом щонайбільше 18 секунд, перш ніж її семантично закодувати та перемістити до LTM</p>	<p>Перші (2) згадки про короткочасну пам'ять (КНП)(3) давали зрозуміти (4), що вона призначена(5) для тимчасового зберігання(6) обмеженої кількості інформації протягом короткого періоду і для негайного згадування(7) перед тим, як(8) інформація буде перенесена(9) в довготривалу пам'ять (ДНП) за допомогою процесів повторення (10). Міллер (1956) припустив, що обмеження ємності STM (11) може бути збільшене за рахунок розбиття вхідної інформації на частини (12), що має бути досягнуто шляхом з'єднання вхідної інформації з інформацією, яка вже зберігається в LTM. Во і Норман (Waugh and Norman, 1965) додали до дискусії, що STM обмежується не тільки плином часом(13), але й інтерференцією (14) від нової інформації, що надходить.</p> <p>Протягом наступних двох десятиліть домінували лінійно-стадійні моделі (15) обробки інформації. Ці моделі ілюстрували, як інформація переходить (16) від одного етапу обробки до іншого і як змінюється сама інформація в процесі обробки. STM займав центральне місце в цих моделях і описувався як своєрідний буфер, в якому інформація тимчасово зберігається, щонайбільше 18 секунд, перш ніж вона буде семантично кодована</p>
--	---	--

<p>vocal or sub-vocal rehearsal, however, information would decay and no longer be available for further processing. Influential information processing models at the time were those proposed by Treisman (1964), Massaro (1975), and Neisser (1967), and Atkinson and Shiffrin (1968). The latter was among the most influential information processing models and already suggested, without specifying this any further, that a short-term store (STS) would act as a more dynamic structure that supported complex cognitive processes. Broadbent's (1974) filter theory supplemented these linear stage models with the concept of selective filtering, whereby information was not only processed in attended channels, but also in the unattended channel as long as the acoustic trace had not yet faded from the echoic or auditory sensory information store.</p>	<p>(Cowan 2008). Однак без репетиції вокалу чи підголосу інформація занепадає й більше не буде доступною для подальшої обробки. Впливовими моделями обробки інформації на той час були моделі, запропоновані Трейсманом (1964), Массаро (1975) і Нейссером (1967), а також Аткинсоном і Шиффріном (1968). Остання була однією з найвпливовіших моделей обробки інформації і вже припустила, не уточнюючи це далі, що короткострокове сховище (STS) буде діяти як більш динамічна структура, яка підтримує складні когнітивні процеси. Теорія фільтра Бродбента (1974) доповнила ці моделі лінійного етапу концепцією вибіркової фільтрації, згідно з якою інформація оброблялася не лише в каналах, які перебувають під наглядом, але й у каналі без нагляду, доки акустичний слід ще не зник із луни або слухових сенсорів. сховище інформації.</p>	<p>(17) і передана до LTM (Cowan 2008). Однак без вокальної або підголоскової репетиції (18) інформація розпадається і стає недоступною для подальшого опрацювання. Впливовими моделями обробки інформації на той час були моделі, запропоновані Трайсманом (1964), Масаро (1975) і Нейссером (1967), а також Аткинсоном і Шиффріном (1968). Остання була однією з найвпливовіших моделей обробки інформації і вже тоді припускала, не уточнюючи цього далі, що короткочасна пам'ять (19) (STS) буде діяти як більш динамічна структура, яка підтримує складні когнітивні процеси. Теорія фільтрів Бродбента (20) (1974) доповнила ці лінійні стадійні моделі (21) концепцією селективної фільтрації, згідно з якою інформація обробляється не лише в активних каналах, але й у неактивних (22), доки акустичний слід ще не зник з ехо- або слухового сховища (23) сенсорної інформації.</p>
--	--	--

Помилки 1 та 3 неясність, (тип 2) – «КНП». Збереження абрєвіатури мовою оригіналу (STM і LTM) створює неясність для читача. Маємо 1 штрафний бал.

Помилка 2 мовна граматична, (тип 4) – «ранні згадки» у контексті наукового тексту не може використовуватись, краще вживати «перші згадки». Рахуємо 0,1 штрафний бал. Помилка 4 мовна (лексична) помилка (тип 4) – «давати зрозуміти» — вираз краще замінити на «показали». Тут маємо 0,1 штрафний бал.

Помилка 5 термінологічна помилка (тип 3) – «призначена» краще замінити на «розроблена». Маємо 0,25 штрафного балу.

Помилка 6 мовна стилістична, (тип 4) – замість виразу «тимчасове зберігання» краще сказати «зберігання обмеженої кількості інформації протягом короткого періоду». Рахуємо 0,1 штрафний бал.

Помилка 7 неясність (тип 2) – «і для негайного згадування» краще використовувати термін «виклик», який краще передає процес витягування інформації з пам'яті. Тут рахуємо 0,5 штрафного балу.

Помилка 8 мовна лексична, (тип 4) – «перед тим, як» краще використовувати більш формальний варіант, наприклад, «до того, як». Тут маємо 0,1 штрафного балу. Помилка 9 мовна граматична, (тип 4) – «буде перенесена» варто замінити на «переведена». Маємо 0,1 штрафний бал.

Помилка 10 неясність (тип 2) – «процеси репетиції» більш точним було б використання «процес повторення». Також маємо 0,1 штрафний бал.

Помилка 11 термінологічні (тип 3) – «обмеження пропускну́ї здатності STM» або ж краще вжити «обмеження ємності STM». Рахуємо 0,25 штрафних балів.

Помилка 12 мовна лексична (тип 4) — «поділ вхідної інформації на фрагменти», тоді як «частини» звучить менш формально. Помилка 13 мовна граматична (тип 4) – «плином часу» формулювання є більш точним, оскільки підкреслює процес і є ближчим до оригіналу «over time».

Помилка 14 термінологічна (тип 3) – «інтерференція» використовується замість «перешкоди».

Помилка 15 неясність (тип 2) – «домінували в наступні два десятиліття» або ж краще «протягом наступних двох десятиліть домінували».

Помилка 16 мовна лексична (тип 4) – «проходити від одного етапу обробки до наступного», що звучить менш природно, ніж «переходити від одного етапу до іншого».

Помилка 17 термінологічна (тип 3) — «кодована», що краще передає зміст терміну «encoded», тоді як «закодувати» має дещо технічний відтінок. Помилка 18 термінологічна (тип 3) – «репетиція вокалу чи підголосу» більш точний варіант – «вокальна або підголоскова репетиція», оскільки це термінологічне поняття, яке описує процес повторення інформації для її збереження в пам'яті.

Помилка 19 термінологічна (тип 3) — терміни «короткострокове сховище» і «сховище інформації», спрощує оригінальну термінологію, знижуючи точність. У другому фрагменті коректно вжито термін короткочасна пам'ять та сховище сенсорної інформації. Помилка 20 термінологічна (тип 3) – «фільтра Бродбента» у першому фрагменті порушує логіку термінології (має бути «теорія фільтрів Бродбента» у другому фрагменті). Помилка 21 термінологічна (тип 3) – «лінійного етапу» у першому фрагменті менш зрозуміло у порівнянні з «лінійними стадійними моделями» у другому.

Помилка 22 мовна граматична (тип 4) — у першому фрагменті є фраза «перебувають під наглядом», що виглядає неточним перекладом з оригінальної мови, а в другому фрагменті використано відповідне слово «активні канали».

Помилка 23 неясність (тип 2) – «слухових сенсорів» у першому фрагменті є неточним формулюванням, правильніше використано у другому як «слухове сховище сенсорної інформації».

Числові результати проведеного аналізу наводимо в таблиці 2.4:

Таблиця 2.4. Кількість та питома вага помилок в ТП *Google Translate* за фрагментами і типами (кількість помилок/питома вага (% від загальної кількості помилок)).

Система МП <i>Google Translate</i>	Типи помилок				Усього помилок	Штрафних балів
	I тип	II тип	III тип	IV тип		
Фрагмент 2	0/0	6/26	8/34	9/39	23	5,9

Як впливає з табл. 2.4, загальна кількість помилок у фрагменті 2, є в кількості 23, з яких помилки 1 типу становлять 0%, 2 типу — 26%, 3 типу – 34% та 4 типу – 39%. Загальна оцінка ТП МП цього фрагмента складає 5,9 штрафних балів.

ВИСНОВКИ ДО РОЗДІЛУ 2

У ході дослідження, присвяченого аналізу перекладу спеціалізованих англійських текстів на українську мову за допомогою автоматизованого перекладача *Google Translate (GT)*, було проведено ретельне порівняння двох фрагментів оригінального тексту (*OT*), машинного перекладу (*МП*) та його постредагованої версії. Результати аналізу дозволили виявити різні типи помилок, характерних для перекладу, здійсненого програмою *GT*, та оцінити якість і точність перекладу з погляду лексичної, граматичної, термінологічної та стилістичної правильності.

Так у першому фрагменті, найбільш вагомі помилки стосувалися термінологічної неточності, що свідчить про недосконалість *Google Translate* у передачі спеціалізованої лексики. Зокрема, програма часто не враховувала контекст при виборі загальноживаних слів замість фахових термінів, що призводило до термінологічних втрат. Окрім того, значна кількість помилок другого типу була пов'язана з невдалим перекладом складних граматичних структур, що створювало неясності для читача.

В другому фрагменті перекладу основні помилки також були пов'язані з термінологією та граматичними конструкціями. При цьому найбільше помилок другого та третього типів виникло через автоматизований переклад складних термінів, що зумовило порушення правильності перекладу та утворення двозначних конструкцій.

У результаті загального аналізу встановлено, що машинний переклад *Google Translate* є цілком ефективним при роботі з універсальними текстами, але виявляє серйозні помилки у роботі зі спеціалізованими науково-технічними матеріалами. Так, найбільша кількість помилок стосується термінології та передачі складних граматичних конструкцій. Розглянутий вище перекладач має схильність до застосування загальноприйнятих слів, чим дуже часто порушує коректність відтворення наукової інформації. Також програма не завжди грамотно обробляє терміни, залишаючи їх у вигляді аббревіатур або ж неправильно інтерпретуючи значення термінів відповідно до контексту.

Порівняння результатів двох фрагментів дозволило зробити висновок, що кількість термінологічних та мовних помилок (типи 3 та 4) є найпоширенішими для обох частин тексту. Помилки першого типу, що призводять до значної втрати інформації, виявлялися рідше, але їх наявність свідчить про потенційну загрозу точності перекладу наукових текстів. Проблеми з передачею термінології можуть мати серйозні наслідки для розуміння складних наукових робіт, особливо якщо користувач покладається на GT як на основний інструмент перекладу.

РОЗДІЛ 3

ЕФЕКТИВНІСТЬ СИСТЕМИ МАШИННОГО ПЕРЕКЛАДУ

MICROSOFT TRANSLATE

Розділ 3 описує аналіз ТП програми МП *Microsoft Translate (MS)* [*Microsoft Translate*]. Текстом оригіналу (ТО) послуговувала раніше описувана у 2 розділі англomовна стаття [27]. Результати аналізу ТП фрагмента 1, перекладеного за допомогою МП *MS*, який проводився за тими ж принципами, що і в розділі 2, подано далі (див. табл. 3.1).

Таблиця 3.1. Класифікація помилок у перекладі фрагмента 1 тексту оригіналу програмою *Microsoft Translate*.

Текст оригіналу Фрагмент 1	Текст перекладу МП (Microsoft Translate)	Постредагований переклад
<p>Working memory in simultaneous and consecutive interpreting Barbara Moser-Mercer Introduction The concept of working memory (WM) supposes a dedicated limited capacity system underlying human thought that processes, maintains and stores information during a short period of time and which provides an interface between perception, long-term memory (LTM) and action (Baddeley 2003). It has been considered a limited capacity system from the early days of its conception when in the 1960s and 1970s linear flow models of information processing described the transfer of verbal information through rehearsal from a limited-capacity shortterm memory store to long-term memory.</p>	<p>Робоча пам'ять при синхронному та послідовному перекладі Барбара Мозер-Мерсер Введення Концепція робочої пам'яті (WM) передбачає спеціальну систему обмеженої ємності, що лежить в основі людського мислення, яка обробляє, підтримує та зберігає інформацію протягом короткого періоду часу та забезпечує інтерфейс між сприйняттям, довготривалою пам'яттю (LTM) та дією (Baddeley 2003). Вона вважалася системою обмеженої ємності з перших днів її зародження, коли в 1960-х і 1970-х роках лінійні потокові моделі обробки інформації описували передачу вербальної інформації за допомогою репетиції від короткочасного сховища обмеженої ємності до довготривалої пам'яті. Сьогодні дослідники описують робочу</p>	<p>Оперативна пам'ять (1) у синхронному та послідовному перекладі Барбара Мозер-Мерсер Вступ (2) Концепція робочої пам'яті (РП) розглядає її як спеціалізовану систему обмеженої ємності, що забезпечує підґрунтя (3) людського мислення, яка обробляє, підтримує та зберігає інформацію протягом короткого періоду часу і яка (4) забезпечує зв'язок (5) між сприйняттям, довготривалою пам'яттю (LTM) і діями (Baddeley 2003). Її вважали системою з обмеженою ємністю з перших днів введення в науковий обіг (6) цього поняття, коли в 1960-х і 1970-х роках лінійно-потокові (7) моделі обробки інформації описували передачу (шляхом повторення (8) вербальної інформації з короткочасної пам'яті (9) обмеженої ємності в</p>

<p>Researchers today describe working memory as a dynamic system that includes both maintaining and manipulating information through a series of interactive components that include executive control and attentional resources and the processing of both verbal and visual material (Wingfield 2016). These limitations are of particular importance in spoken language comprehension, especially when the stimuli have complex linguistic structures or listening effort is increased by poor acoustic quality or reduced hearing acuity (Wingfield 2016). The term working memory itself seems to have been coined by Miller, Galanter and Pribram (1960) who described the dynamic role of a transient memory system they termed working memory; the term did not attract much attention until Baddeley and Hitch (1974) provided convincing experimental evidence of working memory as a central component of human cognition. As definitions of working memory have evolved over the years, there appears to be consensus on its central characteristic, that of retaining information in conscious awareness by way of rehearsal for a short period of time when it is not or no longer present in the environment, of manipulating that information and of using it to guide behavior (Postle 2006).</p>	<p>пам'ять як динамічну систему, яка включає як підтримання, так і маніпулювання інформацією за допомогою серії інтерактивних компонентів, які включають виконавчий контроль та ресурси уваги, а також обробку як вербального, так і візуального матеріалу (Wingfield, 2016). Ці обмеження мають особливе значення для розуміння розмовної мови, особливо коли стимули мають складну лінгвістичну структуру або зусилля на слухання посилюються через погану акустичну якість або знижену гостроту слуху (Wingfield 2016). Сам термін «робоча пам'ять», мабуть, був введений Міллером, Галантером і Прібрамом (1960), які описали динамічну роль тимчасової системи пам'яті, яку вони назвали робочою пам'яттю; цей термін не привертав особливої уваги до тих пір, поки Бедделі і Хітч (1974) не надали переконливі експериментальні докази робочої пам'яті як центрального компонента людського пізнання. У міру того, як визначення робочої пам'яті еволюціонували протягом багатьох років, здається, існує консенсус щодо її центральної характеристики — збереження інформації у свідомому свідомості шляхом репетиції протягом короткого періоду часу, коли вона відсутня або вже не присутня в середовищі, маніпулювання цією інформацією та використання її для керування поведінкою (Postle 2006).</p>	<p>довгочасну пам'ять. Сьогодні дослідники описують робочу пам'ять як динамічну систему, яка включає як збереження (10) інформації, так і маніпулювання нею за допомогою ряду інтерактивних компонентів, які включають виконавчий контроль і ресурси уваги (11), а також обробку вербального та (12) візуального матеріалу (Wingfield 2016) (13). Ці обмеження мають особливе значення для розуміння усної мови (14), особливо коли стимули мають складну мовну структуру (15) або зусилля для аудіювання (16) збільшуються через погану акустичну якість або знижену гостроту слуху (Wingfield 2016). Сам термін робоча пам'ять, здається (17), був запропонований Міллером, Галантером і Прібрамом (1960), які описали динамічну роль перехідної системи пам'яті (18), яку вони назвали робочою пам'яттю; цей термін не привертнув особливої уваги, доки (19) Бадделі та Хітч (1974) (20) не надали переконливих експериментальних доказів того, що робоча пам'ять є центральним компонентом людського пізнання. Оскільки дефініції робочої пам'яті (21) змінювалися протягом багатьох років, здається, існує консенсус щодо її головної характеристики, тобто утримання інформації у свідомій пам'яті (22) шляхом її повторення (23) протягом короткого періоду, коли вона вже відсутня у довгій пам'яті (24), а також маніпулювання цією інформацією та використання її для керування поведінкою (Postle 2006).</p>
--	---	---

Помилка 1 інформаційна (1 тип) – «робоча пам'ять» є більш науковим терміном, тоді як «оперативна пам'ять» спотворює основну інформацію. Рахуємо 1 штрафний бал.

Помилка 2 мовна стилістична (тип 4) – виконана стилістична заміна, яка не впливає на загальне розуміння тексту. Рахуємо 0,1 штрафного бала.

Помилка 3 неясність (тип 2) — «підґрунтя» замість «системи обмеженої ємності» створює неясність для читача. Рахуємо 0,5 штрафного бала.

Помилка 4 мовна стилістична (тип 4) — «і яка» замість «та яка», не є суттєвою та не впливає на розуміння, але порушує стиль. Рахуємо 0,1 штрафного бала.

Помилка 5 інформаційна (тип 1) — заміна «інтерфейс» на «зв'язок» змінює точне значення терміну, що впливає на точність передачі інформації. Рахуємо 1 штрафний бал.

Помилка 6 мовна стилістична (тип 4) – «запровадження» в наукових текстах краще вживати «введення в науковий обіг». Рахуємо 0,1 штрафного бала. Помилка 7 мовна стилістична (тип 4) — зміна «лінійно-потоківі» замість «лінійно потоківі» — не впливає на зміст, але порушує стиль. Рахуємо 0,1 штрафного бала.

Помилка 8 неясність (тип 2) — заміна слова «репетиція» на «повторення» може спричиняє певну неясність через зміну терміну, що має специфічне значення. Рахуємо 0,5 штрафного бала.

Помилка 9 інформаційна (тип 1) — заміна «сховище» на «пам'ять» змінює науковий термін. Рахуємо 1 штрафний бал.

Помилка 10 мовна стилістична (тип 4) – «підтримання» краще замінити на більш стилістично правильний варіант «збереження». Рахуємо 0,1 штрафного бала. Помилка 11 мовна стилістична (тип 4) – «виконавчий контроль та ресурси уваги» та «виконавчий контроль і ресурси уваги» тут вжито незначну стилістичну зміну у використанні сполучників. Рахуємо 0,1 штрафного бала.

Помилка 12 неясність (тип 2) — відсутня фраза «а також обробку як вербального, так і візуального матеріалу» при перекладі може створити неясність. Рахуємо 0,5 штрафного бала.

Помилка 13 мовна стилістична (тип 4) — різниця у використанні «і» замість «та» — не впливає суттєво на зміст, але порушує стиль. Рахуємо 0,1 штрафного бала.

Помилка 14 інформаційна (тип 1) — заміна «розмовна мова» відрізняється від «усної мови» в контексті мовлення і слухання. Рахуємо 1 штрафний бал.

Помилка 15 мовна стилістична (тип 4) — «мовна структура» замість «лінгвістична структура» — порушує стилістичну точність. Рахуємо 0,1 штрафного бала.

Помилка 16 неясність (тип 2) — використання терміну «аудіювання» замість «слухання» є менш поширеним і може спричинити незрозуміння. Рахуємо 0,5 штрафного бала. Помилка 17 неясність (тип 2) — «здається» замість «мабуть» створює незначну неясність у тексті. Рахуємо 0,5 штрафного бала.

Помилка 18 інформаційна (тип 1) — «тимчасова» і «перехідна» мають різні значення, що призводить до спотворення інформації. Рахуємо 1 штрафний бал.

Помилка 19 мовна стилістична (тип 4) — «не привертав особливої уваги до тих пір» та «не привернув особливої уваги, доки», тут вжита стилістична заміна прислівників, яка не впливає на загальне розуміння. Рахуємо 0,1 штрафного бала.

Помилка 20 термінологічна (тип 3) — неправильне використання назви «Баддлі» замість «Бадделі». Рахуємо 0,25 штрафного бала.

Помилка 21 неясність (тип 2) — термін «дефініції» є менш поширеним спричинює неясність у даному контексті. Рахуємо 0,5 штрафного бала.

Помилка 22 інформаційна (тип 1) — фраза «світла пляма свідомості» є незвичною і змінює науковий контекст порівняно з «свідоме свідомості», що спотворює зміст. Рахуємо 1 штрафний бал.

Помилка 23 мовна стилістична (тип 4) – «репетиції» краще замінити на «повторення». Рахуємо 0,1 штрафний бал. Помилка 24 мовна стилістична (тип 4) — різниця «в довкіллі» замість «середовищі», суттєво не впливає на зміст, але порушує стиль. Рахуємо 0,1 штрафного бала.

Числові результати проведеного аналізу наводимо в таблиці 3.2:

Таблиця 3.2. Кількість та питома вага помилок в ТП *Microsoft Translate* за фрагментами і типами (кількість помилок/питома вага (% від загальної кількості помилок)).

Система МП <i>Microsoft Translate</i>	Типи помилок				Усього помилок	Штрафних балів
	I тип	II тип	III тип	IV тип		
Фрагмент 1	6/25	6/25	1/4	11/46	24	7,35

Як впливає з табл. 3.2, загальна кількість помилок у фрагменті 1, є в кількості 24, з яких помилки 1 типу становлять 25%, 2 типу — 25%, 3 типу – 4% та 4 типу – 46%. Загальна оцінка ТП МП цього фрагмента складає 7,35 штрафних балів.

Таблиця 3.3. Класифікація помилок у перекладі фрагмента 2 тексту оригіналу програмою *Microsoft Translate*.

Текст оригіналу Фрагмент 2	Текст перекладу МП (Microsoft Translate)	Постредагований переклад
WM models and their link to interpreting Linear-stage or information processing models Early references to short-term memory (STM) made it clear that it was designed to temporarily hold a limited amount of information for a brief period and for immediate recall before the information	Моделі ЗМ та їх зв'язок з усним перекладом Лінійно-ступінчасті або моделі обробки інформації Ранні згадки про короткочасну пам'ять (STM) дали зрозуміти, що вона була розроблена для тимчасового зберігання обмеженого обсягу інформації протягом короткого періоду часу і для негайного	Моделі ВМ та їх зв'язок з інтерпретацією Лінійно-стадійні моделі або моделі обробки інформації Перші згадки про короткочасну пам'ять (КНП) (1) давали зрозуміти, що вона призначена для тимчасового зберігання обмеженої кількості інформації протягом короткого періоду і для негайного згадування (2)

<p>would be transferred to long-term memory (LTM) using rehearsal processes. Miller (1956) speculated that the STM capacity limitation could be enhanced through chunking incoming information, which was to be accomplished through connecting incoming information with information already stored in LTM. Waugh and Norman (1965) added to the discussion that STM was not only constrained by the passage of time, but also by interference from newly incoming information. Linear-stage or information processing models dominated the next two decades. These models illustrated how information would flow from one processing stage to the next and how the information itself would change in the process. STM occupied a central position in these models and was described as a sort of buffer in which information would be temporarily stored, for at most 18 seconds, before being semantically coded and moved on to LTM (Cowan 2008). Without vocal or sub-vocal rehearsal, however, information would decay and no longer be available for further processing. Influential information processing models at the time were those proposed by Treisman (1964),</p>	<p>запам'ятовування, перш ніж інформація буде передана в довготривалу пам'ять (LTM) за допомогою репетиційних процесів. Міллер (1956) припустив, що обмеження пропускної здатності STM може бути посилено за рахунок фрагментації вхідної інформації, що повинно бути досягнуто шляхом з'єднання вхідної інформації з інформацією, що вже зберігається в LTM. Во і Норман (1965) додали до дискусії, що STM обмежується не тільки плином часу, але і перешкодами від нової інформації, що надходить. Лінійно-ступінчасті або моделі обробки інформації домінували протягом наступних двох десятиліть. Ці моделі ілюструють, як інформація буде перетікати з одного етапу обробки на інший і як буде змінюватися сама інформація в процесі. STM займав центральне місце в цих моделях і був описаний як свого роду буфер, в якому інформація тимчасово зберігалася, максимум 18 секунд, перш ніж була семантично закодована і передана до LTM (Cowan 2008). Однак без вокальної або підвокальної репетиції інформація розкладалася б і більше не була б доступна для подальшої обробки. Впливовими моделями обробки інформації в той час були моделі,</p>	<p>перед тим, як інформація буде перенесена (3) в довготривалу пам'ять (ДНП) за допомогою процесів репетиції. Міллер (1956) припустив, що обмеження ємності STM може бути збільшене за рахунок розбиття вхідної інформації на частини (4), що має бути досягнуто шляхом з'єднання вхідної інформації з інформацією, яка вже зберігається в LTM (5). Во (6) і Норман (Waugh and Norman, 1965) додали до дискусії, що STM обмежується не тільки плином часом, але й інтерференцією від нової інформації (7), що надходить. Протягом наступних двох десятиліть домінували лінійно-стадійні (8) моделі обробки інформації (9). Ці моделі ілюстрували (10), як інформація переходить (11) від одного етапу обробки до іншого (12) і як змінюється (13) сама інформація в процесі обробки. STM займав центральне місце в цих моделях і описувався як своєрідний (14) буфер, в якому інформація тимчасово зберігається, щонайбільше 18 секунд, перш ніж вона буде семантично кодована (15) і передана до LTM (Cowan 2008). Однак без вокальної або підголоскової (16) репетиції інформація розпадається (17) і стає недоступною (18) для подальшого опрацювання (19). Впливовими моделями обробки інформації на той</p>
--	--	--

<p>Massaro (1975), and Neisser (1967), and Atkinson and Shiffrin (1968). The latter was among the most influential information processing models and already suggested, without specifying this any further, that a short-term store (STS) would act as a more dynamic structure that supported complex cognitive processes. Broadbent's (1974) filter theory supplemented these linear stage models with the concept of selective filtering, whereby information was not only processed in attended channels, but also in the unattended channel as long as the acoustic trace had not yet faded from the echoic or auditory sensory information store.</p>	<p>запропоновані Трейсманом (1964), Массаро (1975) і Нейссером (1967), а також Аткинсоном і Шиффріном (1968). Остання була однією з найвпливовіших моделей обробки інформації і вже припускала, не уточнюючи це далі, що короткострокове сховище (STS) буде діяти як більш динамічна структура, яка підтримує складні когнітивні процеси. Теорія фільтрів Бродбента (1974) доповнила ці лінійні стадійні моделі концепцією селективної фільтрації, згідно з якою інформація оброблялася не тільки в відвідуваних каналах, але і в необслуговуваному каналі до тих пір, поки акустичний слід ще не зник з ехогенного або слухового сховища сенсорної інформації.</p>	<p>час (20) були моделі, запропоновані Трайсманом (21) (1964), Масаро (22) (1975) і Нейссером (1967), а також Аткинсоном і Шиффріном (1968). Остання була однією з найвпливовіших моделей обробки інформації і вже тоді припускала (23), не уточнюючи цього далі (24), що короткочасна пам'ять (25) (STS) буде діяти як більш динамічна структура, яка підтримує складні когнітивні процеси. Теорія фільтрів Бродбента (1974) доповнила ці лінійні стадійні моделі (26) концепцією селективної фільтрації, згідно з якою інформація обробляється (27) не лише в активних каналах (28), але й у неактивних (29), доки акустичний слід ще не зник з ехо- або слухового сховища (30) сенсорної інформації.</p>
--	---	---

Помилка 1 та 5 термінологічні (тип 3) – першому варіанті перекладу використовуються англійські аббревіатури «STM», «LTM», у другому варіанті — українські «КНП», «ДНП». Рахуємо 0,5 штрафного бала (двічі по 0,25 штрафного бала).

Помилка 2 мовна лексична (тип 4) — використано термін «запам'ятовування», що є некоректним, правильніше — «згадування», оскільки йдеться про короткочасну пам'ять. Рахуємо 0,1 штрафного бала. Помилка 3 мовна стилістична (тип 4) — фраза «перенесена» є менш природною в науковому контексті порівняно з «передана». Рахуємо 0,1 штрафного бала.

Помилка 4 термінологічна (тип 3) — термін «фрагментація» є точнішим, тому «розбиття на частини» є менш правильним варіантом. Рахуємо 0,25

штрафного бала. Помилка 6 термінологічна (тип 3) – ім'я не перекладене, як «Во», що створює неузгодженість. Рахуємо 0,25 штрафного бала. Помилка 7 термінологічна (тип 3) – у першому варіанті використано менш точний термін «перешкоди», а в другому варіанті використано правильний термін «інтерференція», який є технічно вірнішим для опису зазначеного явища. Рахуємо 0,25 штрафного бала.

Помилка 8 термінологічна (тип 3) — «лінійно-ступінчасті» і «лінійно-стадійні» моделі можуть бути сприйняті як різні поняття, але в контексті інформації краще вжити «стадійні». Рахуємо 0,25 штрафного бала.

Помилка 9 мовна стилістична (тип 4) — порядок слів у першому варіанті перекладу є більш природним і стилістично правильним для української мови. Рахуємо 0,1 штрафного бала. Помилки 10, 11, 13 мовні граматичні (тип 4) — у першому варіанті перекладу дієслова використовуються в майбутньому часі, а у другому варіанті — у минулому та теперішньому часі, що є правильнішим. Рахуємо 0,3 штрафних бали (тричі по 0,1).

Помилка 12 термінологічна (тип 3) — слово «перетікати» менш точне, ніж «переходить». Рахуємо 0,25 штрафного бала.

Помилка 14 мовна стилістична (тип 4) — «свого роду» є менш підходящою, ніж «своєрідний». Рахуємо 0,1 штрафного бала. Помилка 15 мовна граматична (тип 4) — «буде кодована» більш коректна з огляду на узгодження, ніж «була закодована». Рахуємо 0,1 штрафного бала.

Помилка 16 термінологічна (тип 3) — термін «підголоскова» є неточним перекладом терміну, більш правильним є «підвокальна». Рахуємо 0,25 штрафного бала.

Помилка 17 мовна граматична (тип 4) — використано умовну форму «розкладалася б», яка правильна, на відміну від вживання теперішнього часу у слові «розпадається». Рахуємо 0,1 штрафного бала. Помилка 18 мовна стилістична

(тип 4) — «більше не була б доступна» , краще стосується контексту ніж «стає недоступною», яке не передає поступовості процесу. Рахуємо 0,1 штрафного бала.

Помилка 19 термінологічна (тип 3) — слово «опрацювання» є менш точним, ніж «обробка», яка є загальноприйнятим терміном. Рахуємо 0,25 штрафного бала.

Помилка 20 мовна стилістична (тип 4) — правильно використаний варіант перекладу «на той час» замість «в той час». Рахуємо 0,1 штрафного бала.

Помилка 21 термінологічна (тип 3) — неправильне написання імені «Трайсман» замість вірного «Трейсман». Рахуємо 0,25 штрафного бала. Помилка 22 термінологічна (тип 3) – ім'я «Масаро» подане з однією «с», хоча вірна транскрипція має бути «Массаро». Рахуємо 0,25 штрафного бала.

Помилка 23 мовна стилістична (тип 4) — слово «тоді» робить вислів точнішим. Рахуємо 0,1 штрафного бала. Помилка 24 мовна граматична (тип 4) — форма «це» є менш правильною, ніж «цього». Рахуємо 0,1 штрафного бала.

Помилка 25 термінологічна (тип 3) – «короткочасна пам'ять» є загальноприйнятим терміном, а отже правильним. Рахуємо 0,25 штрафного бала. Помилка 26 термінологічна (тип 3) – не вірно вжито термін «сценічні», тоді як правильний термін — «стадійні». Рахуємо 0,25 штрафного бала.

Помилка 27 мовна граматична (тип 4) — у контексті правильніше використовувати теперішній час, оскільки мова йде про універсальні процеси. Рахуємо 0,1 штрафного бала.

Помилка 28 термінологічна (тип 3) — термін «відвідуваних каналах» є некоректним, краще вжити «активні канали». Рахуємо 0,25 штрафного бала.

Помилка 29 термінологічна (тип 3) — термін «необслуговуваному каналі» є менш точним у порівнянні з терміном «неактивних». Рахуємо 0,25 штрафного бала.

Помилка 30 термінологічна (тип 3) — термін «ехогенне» є менш точним порівняно з «ехо-сховище». Рахуємо 0,25 штрафного бала.

Числові результати проведеного аналізу наводимо в таблиці 3.4:

Таблиця 3.4. Кількість та питома вага помилок в ТП *Microsoft Translate* за фрагментами і типами (кількість помилок/питома вага (% від загальної кількості помилок)).

Система МП <i>Microsoft Translate</i>	Типи помилок				Усього помилок	Штрафних балів
	I тип	II тип	III тип	IV тип		
Фрагмент 2	0/0	0/0	16/53	14/47	30	5,4

Як випливає з табл. 3.4, загальна кількість помилок у фрагменті 2, є в кількості 30, з яких помилки 1 типу становлять 0%, 2 типу — 0%, 3 типу – 53% та 4 типу – 47%. Загальна оцінка ТП МП цього фрагмента складає 5,4 штрафних балів.

ВИСНОВКИ ДО РОЗДІЛУ 3

У висновку до даного розділу можна зазначити, що аналіз перекладу, виконаного за допомогою програми машинного перекладу *Microsoft Translate*, виявив низку помилок різного характеру. Найбільша кількість помилок була пов'язана з точністю передачі інформації (тип 1), що свідчить про те, що програмі не завжди вдається правильно передати спеціалізовані терміни або зберегти наукову точність тексту. Особливу увагу слід звернути на те, що неправильний переклад термінів («робоча пам'ять» замість «оперативна пам'ять», «інтерфейс» замість «зв'язок» тощо) впливає на розуміння тексту, що є критичним для наукових статей. Також було виявлено певну кількість стилістичних помилок (тип 4), які хоч і не впливають на точність передачі змісту, порушують загальний стиль тексту.

Крім того, помилки типу «неясність» (тип 2) ускладнювали розуміння тексту через використання менш точних або малопоширених термінів, таких як «підгрунтя» замість «система обмеженої ємності», «аудіювання» замість «слухання». Це все, в свою чергу, свідчить про необхідність постредагування

автоматичних перекладів, особливо при роботі з науковими або технічними текстами.

Загалом, машинний переклад *Microsoft Translate* продемонстрував можливості перекладу на загальному рівні, проте для досягнення високої якості перекладу та збереження точності, необхідне професійне постредагування, особливо для спеціалізованих наукових матеріалів.

РОЗДІЛ 4

**ПОРІВНЯЛЬНА ЕФЕКТИВНІСТЬ СИСТЕМ МАШИННОГО ПЕРЕКЛАДУ
GOOGLE TRANSLATE та MICROSOFT TRANSLATE**

У розділі 4 проводиться порівняльний аналіз ТП програм МП *Google Translate (MT)* та *Microsoft Translate (MS)* з метою встановлення їхньої порівняльної ефективності та особливостей функціонування. Аналіз проводиться за фрагментами, а потім узагальнюється за текстом загалом.

Результати порівняльного аналізу перекладу фрагмента 1 подано в таблиці 4.1.

Таблиця 4.1. Порівняльні тексти перекладу фрагмента 1 програмами Google Translate (GT) і Microsoft Translate (MT) та їхні постредаговані версії.

Текст перекладу МП (<i>Google Translate</i>)	Пост-редагований текст перекладу (<i>Google Translate</i>)	Текст перекладу МП (<i>Microsoft</i>)	Пост-редагований текст перекладу (<i>Microsoft</i>)
<p>Оперативна пам'ять при синхронному та послідовному перекладі Барбара Мозер-Мерсер</p> <p>Вступ Концепція робочої пам'яті (WM) передбачає спеціальну систему обмеженої ємності, що лежить в основі людської думки, яка обробляє, підтримує та зберігає інформацію протягом короткого періоду часу і яка забезпечує інтерфейс між сприйняттям, довготривалою пам'яттю (LTM) і діями (Baddeley 2003). Її вважали</p>	<p>Оперативна пам'ять у (1) синхронному та послідовному перекладі Барбара Мозер-Мерсер</p> <p>Вступ Концепція робочої пам'яті (РП) (2) розглядає (3) її як спеціалізовану (4) систему обмеженої ємності, що забезпечує підґрунтя (5) людського мислення (6), яка обробляє, підтримує та зберігає інформацію протягом короткого періоду часу (7) і яка забезпечує зв'язок (8) між сприйняттям, довготривалою пам'яттю (LTM) (9) і діями (Baddeley 2003). Її вважали системою з</p>	<p>Робоча пам'ять при синхронному та послідовному перекладі Барбара Мозер-Мерсер</p> <p>Введення Концепція робочої пам'яті (WM) передбачає спеціальну систему обмеженої ємності, що лежить в основі людського мислення, яка обробляє, підтримує та зберігає інформацію протягом короткого періоду часу та забезпечує інтерфейс між сприйняттям, довготривалою пам'яттю (LTM) та дією (Baddeley 2003). Вона вважалася</p>	<p>Оперативна пам'ять (1) у синхронному та послідовному перекладі Барбара Мозер-Мерсер</p> <p>Вступ (2) Концепція робочої пам'яті (РП) розглядає її як спеціалізовану систему обмеженої ємності, що забезпечує підґрунтя (3) людського мислення, яка обробляє, підтримує та зберігає інформацію протягом короткого періоду часу і яка (4) забезпечує зв'язок (5) між сприйняттям, довготривалою пам'яттю (LTM) і діями (Baddeley 2003). Її вважали системою з обмеженою ємністю з</p>

<p>системою з обмеженою ємністю з перших днів її задуму, коли в 1960-х і 1970-х роках моделі лінійного потоку обробки інформації описували передачу словесної інформації шляхом репетиції з короткочасної пам'яті обмеженої ємності в довготривалу пам'ять. Сьогодні дослідники описують робочу пам'ять як динамічну систему, яка включає як збереження інформації, так і маніпулювання нею за допомогою ряду інтерактивних компонентів, які включають виконавчий контроль і ресурси уваги, а також обробку словесного та візуального матеріалу (Wingfield 2016). Ці обмеження мають особливе значення для розуміння розмовної мови, особливо коли стимули мають складну лінгвістичну структуру або зусилля на прослуховування збільшуються через погану акустичну якість або знижену гостроту слуху (Wingfield 2016). Сам термін робоча пам'ять, здається, був винайдений Міллером,</p>	<p>обмеженою ємністю з перших днів введення в науковий обіг цього поняття (10), коли в 1960-х і 1970-х роках лінійно-потокові моделі (11) обробки інформації описували передачу (шляхом повторення (12) вербальної (13) інформації з короткочасної пам'яті обмеженої ємності в довгочасну (14) пам'ять. Сьогодні дослідники описують робочу пам'ять як динамічну систему, яка включає як збереження інформації, так і маніпулювання нею за допомогою ряду інтерактивних компонентів, які включають виконавчий контроль і ресурси уваги, а також обробку вербального (15) та візуального матеріалу (Wingfield 2016). Ці обмеження мають особливе значення для розуміння усної (16) мови, особливо коли стимули мають складну мовну (17) структуру або зусилля для аудіювання (18) збільшуються через погану акустичну якість або знижену гостроту слуху (Wingfield 2016). Сам термін робоча пам'ять, здається, був запропонований (19) Міллером, Галантером і Прібрамом (1960), які описали динамічну</p>	<p>системою обмеженої ємності з перших днів її зародження, коли в 1960-х і 1970-х роках лінійні потокові моделі обробки інформації описували передачу вербальної інформації за допомогою репетиції від короткочасного сховища обмеженої ємності до довготривалої пам'яті. Сьогодні дослідники описують робочу пам'ять як динамічну систему, яка включає як підтримання, так і маніпулювання інформацією за допомогою серії інтерактивних компонентів, які включають виконавчий контроль та ресурси уваги, а також обробку як вербального, так і візуального матеріалу (Wingfield, 2016). Ці обмеження мають особливе значення для розуміння розмовної мови, особливо коли стимули мають складну лінгвістичну структуру або зусилля на слухання посилюються через погану акустичну якість або знижену гостроту слуху (Wingfield 2016). Сам термін «робоча пам'ять», мабуть, був введений Міллером, Галантером і Прібрамом (1960),</p>	<p>перших днів введення в науковий обіг (6) цього поняття, коли в 1960-х і 1970-х роках лінійно-потокові (7) моделі обробки інформації описували передачу (шляхом повторення (8) вербальної інформації з короткочасної пам'яті (9) обмеженої ємності в довгочасну пам'ять. Сьогодні дослідники описують робочу пам'ять як динамічну систему, яка включає як збереження (10) інформації, так і маніпулювання нею за допомогою ряду інтерактивних компонентів, які включають виконавчий контроль і ресурси уваги (11), а також обробку вербального та (12) візуального матеріалу (Wingfield 2016) (13). Ці обмеження мають особливе значення для розуміння усної мови (14), особливо коли стимули мають складну мовну структуру (15) або зусилля для аудіювання (16) збільшуються через погану акустичну якість або знижену гостроту слуху (Wingfield 2016). Сам термін робоча пам'ять, здається (17), був запропонований Міллером, Галантером і Прібрамом (1960), які описали динамічну роль перехідної системи пам'яті (18),</p>
--	--	--	--

<p>Галантером і Прібрамом (1960), які описали динамічну роль тимчасової системи пам'яті, яку вони назвали робочою пам'яттю; цей термін не привернув особливої уваги, доки Баддлі та Хітч (1974) не надали переконливих експериментальних доказів того, що робоча пам'ять є центральним компонентом людського пізнання. Оскільки визначення робочої пам'яті змінювалися протягом багатьох років, здається, існує консенсус щодо її центральної характеристики, збереження інформації у свідомому стані шляхом репетиції протягом короткого періоду часу, коли її немає або більше немає в навколишньому середовищі, маніпулювання цією інформацією та використання її для керування поведінкою (Postle 2006).</p>	<p>роль перехідної (20) системи пам'яті, яку вони назвали робочою пам'яттю; цей термін не привернув особливої уваги, доки Баддлі та Хітч (1974) не надали переконливих експериментальних доказів того, що робоча пам'ять є центральним компонентом людського пізнання. Оскільки дефініції (21) робочої пам'яті змінювалися протягом багатьох років, здається, існує консенсус щодо її головної (22) характеристики, тобто утримання (23) інформації у світлій плямі свідомості (24) шляхом її повторення (25) протягом короткого періоду (26), коли вона вже (27) відсутня (28) у довкіллі (29), а також (30) маніпулювання цією інформацією та використання її для керування поведінкою (Postle 2006).</p>	<p>які описали динамічну роль тимчасової системи пам'яті, яку вони назвали робочою пам'яттю; цей термін не привертав особливої уваги до тих пір, поки Бедделі і Хітч (1974) не надали переконливі експериментальні докази робочої пам'яті як центрального компонента людського пізнання. У міру того, як визначення робочої пам'яті еволюціонували протягом багатьох років, здається, існує консенсус щодо її центральної характеристики — збереження інформації у свідомому свідомості шляхом репетиції протягом короткого періоду часу, коли вона відсутня або вже не присутня в середовищі, маніпулювання цією інформацією та використання її для керування поведінкою (Postle 2006).</p>	<p>яку вони назвали робочою пам'яттю; цей термін не привернув особливої уваги, доки (19) Баддлі та Хітч (1974) (20) не надали переконливих експериментальних доказів того, що робоча пам'ять є центральним компонентом людського пізнання. Оскільки дефініції робочої пам'яті (21) змінювалися протягом багатьох років, здається, існує консенсус щодо її головної характеристики, тобто утримання інформації у світлій плямі свідомості (22) шляхом її повторення (23) протягом короткого періоду, коли вона вже відсутня у довкіллі (24), а також маніпулювання цією інформацією та використання її для керування поведінкою (Postle 2006).</p>
--	---	--	---

Числові параметри ТП обох програм, тобто порівняльна кількість та питома вага помилок в текстах перекладу програм *GT* і *MS* фрагменту 1 за типами помилок подано в табл. 4.2.

Таблиця 4.2. Порівняльна кількість та питома вага помилок в текстах перекладу *GT* і *MS* за фрагментами і типами (кількість помилок/питома вага (%) від загальної кількості помилок). Фр – номер фрагмента, % - відсоток збігу текстів перекладу програм *GT* і *MS*, *m* – середнє значення.

Фр	%	Типи помилок								Усього помилок		Штрафних балів ТП	
		I тип		II тип		III тип		IV тип					
		GT	MS	GT	MS	GT	MS	GT	MS	GT	MS	GT	MS
1	61	4/7	6/25	4/13	6/25	13/43	1/4	11/37	11/46	30	24	8,35	7,35

Як ясно з результатів табл. 4.2, ТП фрагмента 1 двома програмами є досить схожими і дослівно збігаються у 79% випадків. Однак загальна кількість помилок (30) у ТП програми *GT* перевищує відповідний показник програми *MS* на 6 помилок, проте це не суттєва різниця. Тут варто зазначити, що в обох перекладах кількість мовних помилок однакова, але різна у відсотковому значенні, враховуючи загальну кількість помилок. Помилки I та II типу у ТП *GT* та *MS* мають незначну різницю (2 помилки), а вже помилка III типу ТП *MS* суттєво відрізняється від перекладу *GT*, де кількість таких помилок становить 13, натомість *MS* має лиш 1 таку помилку.

Обидві системи перекладу коректно переклали терміни «робоча пам'ять» та «довготривала пам'ять», проте, термін «вербальна інформація» перекладається подібно, але з різним контекстуальним підходом до структури речення. Найбільша кількість помилок спостерігалася в тому, що речення перекладені доволі дослівно, з типовими для машинного перекладу «буквальними» побудовами, що знижує їх природність і зрозумілість. У випадку «зусилля для аудіювання» обидві системи не враховували контекст, що призводить до неточного перекладу.

GT частіше використовує складніші мовні звороти, що загалом наближає текст до наукового стилю, тоді як *MS* у деяких місцях перекладає простіше, що робить текст більш читабельним, але менш точним. *GT* більш коректно переклав технічні терміни, такі як «виконавчий контроль», тоді як у перекладі *MS* вони перекладені менш точно. В цілому, *MS* схильний до спрощення стилістики

перекладу, тоді як *GT* намагається передати складність тексту, іноді жертвуючи зрозумілістю.

Незважаючи на відмінності в деталях, загальна структура та зміст текстів, перекладених обома системами, дуже схожі. Обидва переклади дотримуються основної ідеї вихідного тексту, зберігаючи ключові терміни та загальну логіку речень. Обидві системи перекладу демонструють схожі проблеми, пов'язані з багатозначністю, перекладом термінів та синтаксичними структурами. Зокрема, у всіх варіантах перекладу зберігається однакова термінологія щодо «робочої пам'яті», її функцій та обмежень. Обидві системи стикаються з труднощами передачі багатозначності деяких слів, таких як «аудіювання», що свідчить про загальні проблеми алгоритмів машинного перекладу з розумінням контексту.

Далі розглядаємо аналіз наступного фрагмента в табл. 4.3:

Таблиця 4.1. Порівняльні тексти перекладу фрагмента 2 програмами *Google Translate (GT)* і *Microsoft Translate (MT)* та їхні постредаговані версії.

Текст перекладу МП (<i>Google Translate</i>)	Пост-редагований текст перекладу (<i>Google Translate</i>)	Текст перекладу МП (<i>Microsoft</i>)	Пост-редагований текст перекладу (<i>Microsoft</i>)
<p>Моделі WM та їх зв'язок з інтерпретацією Лінійно-стадійні або моделі обробки інформації</p> <p>Ранні згадки про короткочасну пам'ять (STM) показали, що вона була розроблена для тимчасового утримання обмеженої кількості інформації протягом короткого періоду</p>	<p>Моделі VM(1) та їх зв'язок з інтерпретацією Лінійно-стадійні моделі або моделі обробки інформації</p> <p>Перші (2) згадки про короткочасну пам'ять (КНП)(3) давали зрозуміти (4), що вона була призначена(5) для тимчасового зберігання(6) обмеженої кількості інформації протягом короткого періоду і для негайного згадування(7) перед</p>	<p>Моделі 3M та їх зв'язок з усним перекладом</p> <p>Лінійно-ступінчасті або моделі обробки інформації</p> <p>Ранні згадки про короткочасну пам'ять (STM) дали зрозуміти, що вона була розроблена для тимчасового зберігання обмеженого обсягу інформації протягом короткого періоду часу і для негайного запам'ятовування,</p>	<p>Моделі VM та їх зв'язок з інтерпретацією Лінійно-стадійні моделі або моделі обробки інформації</p> <p>Перші згадки про короткочасну пам'ять (КНП) (1) давали зрозуміти, що вона призначена для тимчасового зберігання обмеженої кількості інформації протягом короткого періоду і для негайного згадування (2) перед тим, як інформація</p>

<p>часу та для негайного виклику, перш ніж інформація буде передана в довготривалу пам'ять (LTM) за допомогою репетиційних процесів. Міллер (1956) припустив, що обмеження пропускної здатності STM може бути посилено шляхом поділу вхідної інформації на фрагменти, що мало бути досягнуто шляхом з'єднання вхідної інформації з інформацією, яка вже зберігається в LTM. Во та Норман (1965) додали до обговорення, що STM був обмежений не лише плином часу, але й перешкодами від нової інформації, що надходить. Лінійно-стадійні або моделі обробки інформації домінували в наступні два десятиліття. Ці моделі ілюстрували, як інформація буде проходити від одного етапу обробки до наступного і як</p>	<p>тим, як(8) інформація буде перенесена(9) в довготривалу пам'ять (ДНП) за допомогою процесів повторення (10). Міллер (1956) припустив, що обмеження ємності STM (11) може бути збільшене за рахунок розбиття вхідної інформації на частини (12), що має бути досягнуто шляхом з'єднання вхідної інформації з інформацією, яка вже зберігається в LTM. Во і Норман (Waugh and Norman, 1965) додали до дискусії, що STM обмежується не тільки плином часом(13), але й інтерференцією (14) від нової інформації, що надходить. Протягом наступних двох десятиліть домінували лінійно-стадійні моделі (15) обробки інформації. Ці моделі ілюстрували, як інформація переходить (16) від одного етапу обробки до іншого і як змінюється сама інформація в процесі обробки. STM займав центральне місце в цих моделях і описувався як своєрідний буфер, в</p>	<p>перш ніж інформація буде передана в довготривалу пам'ять (LTM) за допомогою репетиційних процесів. Міллер (1956) припустив, що обмеження пропускної здатності STM може бути посилено за рахунок фрагментації вхідної інформації, що повинно бути досягнуто шляхом з'єднання вхідної інформації з інформацією, що вже зберігається в LTM. Во і Норман (1965) додали до дискусії, що STM обмежується не тільки плином часу, але і перешкодами від нової інформації, що надходить. Лінійно-ступінчасті або моделі обробки інформації домінували протягом наступних двох десятиліть. Ці моделі ілюструють, як інформація буде перетікати з одного етапу обробки на інший і як буде змінюватися сама інформація в процесі. STM займав центральне</p>	<p>буде перенесена (3) в довготривалу пам'ять (ДНП) за допомогою процесів репетиції. Міллер (1956) припустив, що обмеження ємності STM може бути збільшене за рахунок розбиття вхідної інформації на частини (4), що має бути досягнуто шляхом з'єднання вхідної інформації з інформацією, яка вже зберігається в LTM (5). Во (6) і Норман (Waugh and Norman, 1965) додали до дискусії, що STM обмежується не тільки плином часом, але й інтерференцією від нової інформації (7), що надходить. Протягом наступних двох десятиліть домінували лінійно-стадійні (8) моделі обробки інформації (9). Ці моделі ілюстрували (10), як інформація переходить (11) від одного етапу обробки до іншого(12) і як змінюється (13) сама інформація в процесі обробки. STM займав центральне місце в цих моделях і описувався як своєрідний (14) буфер, в якому</p>
--	---	--	---

<p>сама інформація буде змінюватися в процесі. STM займав центральне місце в цих моделях і описувався як свого роду буфер, у якому інформація тимчасово зберігалася протягом щонайбільше 18 секунд, перш ніж її семантично закодувати та перемістити до LTM (Cowan 2008). Однак без репетиції вокалу чи підголосу інформація занепадає й більше не буде доступною для подальшої обробки. Впливовими моделями обробки інформації на той час були моделі, запропоновані Трайсманом (1964), Масаро (1975) і Нейссером (1967), а також Аткинсоном і Шиффріном (1968). Остання була однією з найвпливовіших моделей обробки інформації і вже припустила, не уточнюючи це далі, що короткострокове сховище (STS) буде діяти як більш динамічна</p>	<p>якому інформація тимчасово зберігається, щонайбільше 18 секунд, перш ніж вона буде семантично кодована (17) і передана до LTM (Cowan 2008). Однак без вокальної або підголоскової репетиції (18) інформація розпадається і стає недоступною для подальшого опрацювання. Впливовими моделями обробки інформації на той час були моделі, запропоновані Трайсманом (1964), Масаро (1975) і Нейссером (1967), а також Аткинсоном і Шиффріном (1968). Остання була однією з найвпливовіших моделей обробки інформації і вже тоді припускала, не уточнюючи цього далі, що короткочасна пам'ять (19) (STS) буде діяти як більш динамічна структура, яка підтримує складні когнітивні процеси. Теорія фільтрів Бродбента (20) (1974) доповнила ці лінійні стадійні моделі (21) концепцією</p>	<p>місце в цих моделях і був описаний як свого роду буфер, в якому інформація тимчасово зберігалася, максимум 18 секунд, перш ніж була семантично закодована і передана до LTM (Cowan 2008). Однак без вокальної або підвокальної репетиції інформація розкладалася б і більше не була б доступна для подальшої обробки. Впливовими моделями обробки інформації в той час були моделі, запропоновані Трайсманом (1964), Масаро (1975) і Нейссером (1967), а також Аткинсоном і Шиффріном (1968). Остання була однією з найвпливовіших моделей обробки інформації і вже припускала, не уточнюючи це далі, що короткострокове сховище (STS) буде діяти як більш динамічна структура, яка підтримує складні когнітивні процеси. Теорія фільтрів</p>	<p>інформація тимчасово зберігається, щонайбільше 18 секунд, перш ніж вона буде семантично кодована (15) і передана до LTM (Cowan 2008). Однак без вокальної або підголоскової (16) репетиції інформація розпадається (17) і стає недоступною (18) для подальшого опрацювання (19). Впливовими моделями обробки інформації на той час (20) були моделі, запропоновані Трайсманом (21) (1964), Масаро (22) (1975) і Нейссером (1967), а також Аткинсоном і Шиффріном (1968). Остання була однією з найвпливовіших моделей обробки інформації і вже тоді припускала (23), не уточнюючи цього далі (24), що короткочасна пам'ять (25) (STS) буде діяти як більш динамічна структура, яка підтримує складні когнітивні процеси. Теорія фільтрів Бродбента (1974) доповнила ці лінійні стадійні моделі (26) концепцією селективної</p>
--	--	--	---

структура, яка підтримує складні когнітивні процеси. Теорія фільтра Бродбента (1974) доповнила ці моделі лінійного етапу концепцією вибіркової фільтрації, згідно з якою інформація оброблялася не лише в каналах, які перебувають під наглядом, але й у каналі без нагляду, доки акустичний слід ще не зник із луни або слухових сенсорів. сховище інформації.	селективної фільтрації, згідно з якою інформація обробляється не лише в активних каналах, але й у неактивних (22), доки акустичний слід ще не зник з ехо- або слухового сховища (23) сенсорної інформації.	Бродбента (1974) доповнила ці лінійні сценічні моделі концепцією селективної фільтрації, згідно з якою інформація оброблялася не тільки в відвідуваних каналах, але і в необслуговуваному каналі до тих пір, поки акустичний слід ще не зник з ехогенного або слухового сховища сенсорної інформації.	фільтрації, згідно з якою інформація обробляється (27) не лише в активних каналах (28), але й у неактивних (29), доки акустичний слід ще не зник з ехо- або слухового сховища (30) сенсорної інформації.
---	--	---	--

Числові параметри ТП обох програм, тобто порівняльна кількість та питома вага помилок в текстах перекладу програм *GT* і *MS* фрагменту 2 за типами помилок подано в табл. 4.4.

Таблиця 4.4. Порівняльна кількість та питома вага помилок в текстах перекладу *GT* і *MS* за фрагментами і типами (кількість помилок/питома вага (%) від загальної кількості помилок). Фр – номер фрагмента, % — відсоток збігу текстів перекладу програм *GT* і *MS*, *m* – середнє значення.

Фр	%	Типи помилок								Усього помилок		Штрафних балів ТП	
		I тип		II тип		III тип		IV тип					
		GT	MS	GT	MS	GT	MS	GT	MS	GT	MS	GT	MS
2	73	0/0	0/0	6/26	0/0	8/34	16/53	9/39	14/47	23	30	5,9	5,4

Як ясно з результатів табл. 4.4, ТП фрагмента 2 двома програмами є досить схожими і дослівно збігаються у 73% випадків. Однак загальна кількість помилок (23) у ТП програми *GT* є меншою за відповідний показник програми *MS* на 7 помилок, проте це не суттєва різниця. Тут варто зазначити, що в обох перекладах

кількість мовних помилок однакова, але різна у відсотковому значенні, враховуючи загальну кількість помилок. Помилки I та II типу у ТП *GT* та *MS* мають незначну різницю (6 помилок), а вже помилка III типу ТП *MS* суттєво відрізняється від перекладу *GT*, де кількість таких помилок становить 8, що є вдвічі меншим від ТП *MS* (який має 16 таких помилок).

Обидві системи машинного перекладу мають труднощі з правильним перекладом термінів, таких як «репетиція» та «обробка інформації», мається на увазі, що в обох випадках фрази можуть бути перекладені занадто буквально, що робить текст важким для розуміння в науковому контексті. Як *GT*, так і *MS* здійснюють дослівний переклад, не враховуючи контексту або стилістики, що робить текст менш природним. Наприклад, у перекладі фрази «буфер, в якому інформація тимчасово зберігається» обидві системи зберігають таку саму структуру, яка звучить громіздко. Також вони демонструють труднощі з багатозначними словами, наприклад, слово «репетиція» в одному контексті означає повторення інформації для запам'ятовування, але це вжито без контекстуального уточнення.

Окрім того, *MS* перекладає «короткострокове сховище» як «короткочасна пам'ять», тоді як *GT* може використовувати варіант «короткострокове зберігання».

У перекладі речення про «впливові моделі обробки інформації» *MS* робить речення коротшим і простішим, тоді як *GT* відтворює складнішу структуру.

В описі «вокальна або підвокальна репетиція» переклад *MS* може бути менш точним і спрощеним, що впливає на зрозумілість термінів.

В цілому обидва переклади зберігають загальну структуру і зміст тексту. Усі основні терміни та концепції, такі як «короткочасна пам'ять», «лінійно-стадійні моделі» і «репетиційні процеси», присутні в обох перекладах і відображають той самий зміст. Обидві системи передають ключові ідеї щодо процесу обробки інформації та ролі короткочасної пам'яті в цих моделях.

ВИСНОВКИ ДО РОЗДІЛУ 4

У даному розділі було здійснено порівняння машинних перекладів двох програм — *Google Translate* та *Microsoft Translate* — з пост-редагованими версіями, щоб проаналізувати їхню ефективність і специфіку роботи.

В ході дослідження було виявлено схожість структури перекладу, так як обидві програми виконують переклад тексту з науковою термінологією на належному рівні, зберігаючи основну інформацію. Однак у перекладах є неточності в деталях, які потребують корекції, що й було зазначено в пост-редагованому варіанті перекладу.

Лексичні та граматичні відмінності, наприклад, у перекладах *Microsoft Translate* частіше спостерігається правильна передача складних термінів, тоді як *Google Translate* має деякі неточності в передачі таких термінів.

Стиль і синтаксис *Google Translate* демонструє більш «дослівний» підхід, часто не враховуючи стиль тексту, тоді як *Microsoft Translate* намагається передати текст у більш природному стилі, але іноді робить це на шкоду точності. Після редагування обох перекладів основні помилки були виправлені, що покращило їх якість і зробило переклад більш наближеним до оригіналу.

ЗАГАЛЬНІ ВИСНОВКИ

Розвиток систем машинного перекладу пройшов значний шлях від перших спроб дослівного перекладу на базі словникових даних до сучасних нейронних мереж, що забезпечують якісний і точний переклад. Зародження машинного перекладу почалося в середині ХХ століття, а ключовою подією стало формулювання концепції Уорреном Вівером у 1947 році [22, с. 146]. Історія розвитку цієї галузі включає кілька етапів, які позначені вдосконаленням методів перекладу: від механізмів дослівного перекладу до статистичних моделей та впровадження штучних нейронних мереж.

Значний прорив у якості перекладу відбувся з розвитком статистичного машинного перекладу (SMT), а пізніше — з появою нейронного машинного перекладу (NMT), який заснований на глибокому навчанні. НМП дозволив досягти високого рівня адекватності перекладу завдяки здатності аналізувати контекст і адаптуватися до багатозначності мовних структур.

Однак, незважаючи на суттєвий прогрес, машинний переклад все ще стикається з викликами, такими як точність перекладу спеціалізованих текстів у сферах медицини чи юриспруденції. Подальший розвиток цієї галузі залежить від удосконалення алгоритмів штучного інтелекту, врахування специфіки мов та адаптації до різних контекстів.

У ході дослідження, присвяченого аналізу перекладу спеціалізованих англійських текстів на українську мову за допомогою автоматизованого перекладача *Google Translate (GT)*, було проведено ретельне порівняння двох фрагментів оригінального тексту (*OT*), машинного перекладу (*MT*) та його постредагованої версії. Результати аналізу дозволили виявити різні типи помилок, характерних для перекладу, здійсненого програмою *GT*, та оцінити якість і точність перекладу з погляду лексичної, граматичної, термінологічної та стилістичної правильності.

Так у першому фрагменті, найбільш вагомі помилки стосувалися термінологічної неточності, що свідчить про недосконалість *Google Translate* у передачі спеціалізованої лексики. Зокрема, програма часто не враховувала контекст при виборі загальноживаних слів замість фахових термінів, що призводило до термінологічних втрат. Окрім того, значна кількість помилок другого типу була пов'язана з невдалим перекладом складних граматичних структур, що створювало неясності для читача.

В другому фрагменті перекладу основні помилки також були пов'язані з термінологією та граматичними конструкціями. При цьому найбільше помилок другого та третього типів виникло через автоматизований переклад складних термінів, що зумовило порушення правильності перекладу та утворення двозначних конструкцій.

У результаті загального аналізу встановлено, що машинний переклад *Google Translate* є цілком ефективним при роботі з універсальними текстами, але виявляє серйозні помилки у роботі зі спеціалізованими науково-технічними матеріалами. Так, найбільша кількість помилок стосується термінології та передачі складних граматичних конструкцій. Розглянутий вище перекладач має схильність до застосування загальноприйнятих слів, чим дуже часто порушує коректність відтворення наукової інформації. Також програма не завжди грамотно обробляє терміни, залишаючи їх у вигляді аббревіатур або ж неправильно інтерпретуючи значення термінів відповідно до контексту.

Порівняння результатів двох фрагментів дозволило зробити висновок, що кількість термінологічних та мовних помилок (типи 3 та 4) є найпоширенішими для обох частин тексту. Помилки першого типу, що призводять до значної втрати інформації, виявлялися рідше, але їх наявність свідчить про потенційну загрозу точності перекладу наукових текстів. Проблеми з передачею термінології можуть мати серйозні наслідки для розуміння складних наукових робіт, особливо якщо користувач покладається на GT як на основний інструмент перекладу.

У висновку до третього розділу можна зазначити, що аналіз перекладу, виконаного за допомогою програми машинного перекладу *Microsoft Translate*, виявив низку помилок різного характеру. Найбільша кількість помилок була пов'язана з точністю передачі інформації (тип 1), що свідчить про те, що програмі не завжди вдається правильно передати спеціалізовані терміни або зберегти наукову точність тексту. Особливу увагу слід звернути на те, що неправильний переклад термінів («робоча пам'ять» замість «оперативна пам'ять», «інтерфейс» замість «зв'язок» тощо) впливає на розуміння тексту, що є критичним для наукових статей. Також було виявлено певну кількість стилістичних помилок (тип 4), які хоч і не впливають на точність передачі змісту, порушують загальний стиль тексту.

Крім того, помилки типу «неясність» (тип 2) ускладнювали розуміння тексту через використання менш точних або малопоширених термінів, таких як «підгрунття» замість «система обмеженої ємності», «аудіювання» замість «слухання». Це все, в свою чергу, свідчить про необхідність постредагування автоматичних перекладів, особливо при роботі з науковими або технічними текстами.

Загалом, машинний переклад *Microsoft Translate* продемонстрував можливість перекладу на загальному рівні, проте для досягнення високої якості перекладу та збереження точності, необхідне професійне постредагування, особливо для спеціалізованих наукових матеріалів.

У четвертому розділі було здійснено порівняння машинних перекладів двох програм — *Google Translate* та *Microsoft Translate* — з пост-редагованими версіями, щоб проаналізувати їхню ефективність і специфіку роботи.

В ході дослідження було виявлено схожість структури перекладу, так як обидві програми виконують переклад тексту з науковою термінологією на належному рівні, зберігаючи основну інформацію. Однак у перекладах є

неточності в деталях, які потребують корекції, що й було зазначено в пост-редагованому варіанті перекладу.

Лексичні та граматичні відмінності, наприклад, у перекладах *Microsoft Translate* частіше спостерігається правильна передача складних термінів, тоді як *Google Translate* має деякі неточності в передачі таких термінів.

Стиль і синтаксис *Google Translate* демонструє більш «дослівний» підхід, часто не враховуючи стиль тексту, тоді як *Microsoft Translate* намагається передати текст у більш природному стилі, але іноді робить це на шкоду точності. Після редагування обох перекладів основні помилки були виправлені, що покращило їх якість і зробило переклад більш наближеним до оригіналу.

У підсумку можна зазначити, що результати аналізу машинних перекладів *Google Translate* та *Microsoft Translate* вказують на значний потенціал цих систем для роботи з загальними текстами, але також підтверджують їхні обмеження при перекладі спеціалізованих текстів, які містять складну термінологію та контекстуальні особливості. Дослідження виявило, що обидві програми здатні передати основний зміст наукових текстів, але потребують значного постредагування для досягнення точності, необхідної у спеціалізованих галузях. Хоча *Microsoft Translate* продемонстрував дещо кращу передачу термінології, обидві програми часто допускають помилки, пов'язані з неправильним вибором термінів, стилістичними невідповідностями та синтаксичними неточностями, що свідчить про важливість залучення фахівців до перевірки та редагування таких перекладів. Зважаючи на вищезазначене, можна зробити висновок, що машинний переклад має обмеження в точності та адекватності передачі інформації в науково-технічних текстах, а отже, для забезпечення якості перекладу необхідне професійне постредагування.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Гречуха Л.О., Кузєбна В.В. Системи машинного перекладу: оглядовий аналіз. *Молодий вчений*. 2017. № 2 (42). С. 372–374.
2. Гудманян А. Г., Сітко А. В., Струк І. В. Функціонально-прагматична адекватність машинного перекладу публіцистичних текстів. *Львівський філологічний часопис*. 2019. № 5. С. 48–52.
3. Данилов Г. Машинний переклад, системи машинного перекладу та їх специфіка. *Науковий Вісник ПНПУ Ім. К. Д. Ушинського*. 2021. № 33. С. 293-311.
4. Енциклопедія перекладознавства: в 4 т. / за ред. О. А. Калиниченка та Л. М. Черноватого. Вінниця : Нова книга, 2020. Т. 1. 552 с.
5. Ігнатенко В. Д. Використання сучасних інформаційних технологій у підготовці майбутніх філологів. *Іноземні мови*. 2020. № 1(101). С. 37–42.
6. Ігнатенко В.Д. Місце машинного перекладу у діяльності сучасного перекладача: сучасні дискусії та бачення. *Modern philology: theory, history, methodology*. Riga : Baltija Publishing, 2024. P. 523-534.
7. Корольова Т.М., Жмаєва Н.С., Колчаг Ю.І. Постредагування при машинному перекладі. *Науковий вісник ПНПУ ім. К.Д. Ушинського*. 2020. № 30. С. 102–119.
8. Кочеткова Н.С., Ревіна О.В. Особливості машинного перекладу. *Філологічні науки. Питання теорії і практики*. Київ : Грамота. 2017. № 6(72). С. 106-112.
9. Лазєбна Н. В. Семантична неоднозначність урбаністичної термінології (українсько-англійський переклад Google Translate vs людський переклад). *Нова філологія*. 2019. № 76. С. 61–64.

- 10.Мацак Ж. І., Скоробогатого Т. І. Машинний переклад та його специфіка – URL: http://www.rusnauka.com/31_ONBG_2009/Philologia/54653.doc.ht (дата звернення 22.09.2024)
- 11.Міщенко А. Машинний переклад у контексті сучасного науково-технічного перекладу. *Вісник ХНУ ім. В.Н. Каразіна. Серія «Романо-германська філологія. Методика викладання іноземних мов»*. 2013. № 1051. С. 172–180.
- 12.Рейтинг найкращих онлайн перекладачів по версії GDE SAAS URL: <https://gde-saas.ru/top/rejtingluchshih-onlajn-perevodchikov/>(дата звернення: 05.10.2024).
- 13.Романюк Л.В. До проблеми використання систем машинного перекладу студентами немовних спеціальностей. *Одеський лінгвістичний вісник*. 2017. том 2. № 10. С. 143–146.
- 14.Смірнова Т. В. Переваги та недоліки застосування машинного перекладу. URL: <http://readera.org/article/perevahye-ta-nedolikyе-zastosuvannjamashyennoho-perekladu-10182877.html> (дата звернення 08.10.2024)
- 15.Чернікова Л.Ф. Лінгвістична якість машинного перекладу. *Культура народів Причорномор'я*. 2012. № 248. С. 165–168.
- 16.Черноватий Л. М. Методика викладання перекладу. Вінниця : Нова книга, 2013. 376 с.
- 17.Черноватий Л. М. Психолінгвістичні основи теорії педагогічної граматики. Харків : Основа, 2006. 245 с.
- 18.Черноватий Л. Проблеми машинного перекладу та його застосування у навчанні майбутніх перекладачів. *Наукові записки. Серія: Філологічні науки*. 2022. № 202. С. 84–93.
- 19.Bennett W. S. Machine Translation in North America. *Concise History of the Language Sciences : from the Sumerians to the cognitivists*. 1995. P. 445–451.

20. Google Translate official web page. URL : <https://translate.google.com/about/?hl=uk>
21. Hearne M., Way A. Statistical Machine Translation: A Guide for Linguists and Translators. *Language and Linguistics Compass*. 2011. Vol. 5. N. 5. P. 205–226.
22. Hutchins W. J. Machine translation: a concise history. ResearchGate. URL: https://www.researchgate.net/publication/250888028_Machine_translation_a_concise_history (date of access: 21.10.2024)
23. Hutchins W. J. Warren Weaver and the launchng of MT. *Studies in the History of the Language Sciences*. Amsterdam : Benjamins. 2000. P. 17-34.
24. Jurafsky D. Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition. 2008. Upper Saddle River, N.J : Pearson Prentice Hall. 54 p.
25. Koponen M. Assessing Machine Translation Quality with Error Analysis. URL: https://www.sktl.fi/@Bin/40701/Koponen_MikaEL2010.pdf (дата звернення: 21.09.2024)
26. Microsoft Translator official web page. URL: <https://www.microsoft.com/uk-ua/translator/apps/>
27. Moser-Mercer, B. Working memory in simultaneous and consecutive interpreting. *The Routledge Handbook of Translation, Interpreting and Bilingualism* / ed. A. Ferreira, J. W. Schwieter. 2023. London and New York : Routledge. P. 280-297.
28. Okpor M. Machine translation approaches: issues and challenges. *International Journal of Computer Science Issues*. 2014. Vol. 11. N 5. P. 159–165.
29. Tan Zh., Wang Sh., Yang Z., Chen G., Huang X., Sun M., Liu Y. Neural machine translation: A review of methods, resources, and tools. *AI Open*. 2020. Vol. 1. P. 5-21.

30. Vigier-Moreno F. J., Pérez-Macías. L. Assessing Neural Machine Translation of Court Documents: a Case Study on the Translation of a Spanish Remand Order into English. *Revista de Llengua i Dret, Journal of Language and Law*. 2022. № 78. P. 73–91.

SUMMARY

The relevance of the thesis topic is due to the rapid development of machine translation technologies, which have become important tools for the rapid translation of texts of various subjects. However, when translating specialized English-language texts into Ukrainian, such as technical, legal and medical texts, there are difficulties with the accuracy of terminology and contextual nuances. Existing research shows that machine translation systems such as *Google Translate* and *Microsoft Translate* work well with everyday texts, but have limitations in terms of accuracy when translating specialized materials. Studying these problems and finding ways to overcome them is important for improving the quality of machine translation and developing linguistic technologies in Ukraine.

The aim of the study is to analyze and identify the main problems faced by machine translation systems, in particular *Google Translate* and *Microsoft Translate*, when translating specialized English texts into Ukrainian.

By this goal, we set the following *research objectives*:

1. To analyse and characterize the evolution of machine translation, focusing on the main stages of technology development.
2. To compare *Google Translate* and *Microsoft Translate* — to analyze the two leading machine translation systems, taking into account their functionality, accuracy and usage features.
3. Evaluate the effectiveness of *Google Translate* for different types of texts, assessing the accuracy and consistency of the source text with the translated text.
4. Evaluate the effectiveness of *Microsoft Translate* and its ability to accurately reproduce the content of various texts.
5. Perform a comparative analysis of the effectiveness of *Google Translate* and *Microsoft Translate*: highlight the main advantages and disadvantages of each system, comparing their effectiveness in different contexts.

To achieve the aim of the study, the following **methods** were used: selection of materials for analysis, comparison and analysis of translation texts, classification of the identified types of errors, systematization of theoretical data in the form of tables, grouping of errors by their impact on the accuracy of the original text, and summarization of the results in the conclusions.

The scientific novelty of the work lies in a comprehensive analysis of the translation of specialized English texts (scientific, technical, medical) into Ukrainian using machine translation systems. For the first time, a comparison of the quality of *Google Translate* and *Microsoft Translate* translations was conducted with a focus on ambiguity, terminological accuracy and the transfer of complex grammatical structures.

The theoretical value of the study lies in the analysis of the peculiarities of machine translation of specialized texts, including scientific, technical, medical and legal texts with specific terminology. The study allows us to evaluate how *Google Translate* and *Microsoft Translate* handle context and terminology, and contributes to the development of machine translation theory, in particular in the areas of ambiguity, contextual translation, and model adaptation to language pairs such as English-Ukrainian.

The practical value of the study is to improve machine translation of specialized English texts into Ukrainian, such as legal, medical, technical and scientific texts. The results can help improve the accuracy of translation, reduce the number of errors and optimize the work of *Google Translate* and *Microsoft Translate* with highly technical texts. Based on the identified shortcomings, recommendations for translators and users can be developed to help avoid common mistakes and better adapt translation to the peculiarities of the Ukrainian language.

The results were tested at the student scientific conference of the Mykola Lukash Department of Translation Studies of V.N. Karazin Kharkiv National University on 20 November 2024. The results of the study are presented in an article submitted for publication in the collection of student papers of the Mykola Lukash Department of

Translation Studies of V. N. Karazin Kharkiv National University *In Statu Nascendi*
(issue 24).