

КОГНІТИВНО-СТИЛІСТИЧНІ ПАРАМЕТРИ ВІДТВОРЕННЯ АНГЛОМОВНОГО ІНШОМОВНОГО ДИСКУРСУ В СИСТЕМАХ APERTIUM TA DEEPL: ПОРІВНЯЛЬНИЙ АСПЕКТ

Дарина ПРАЧУК

*здобувачка вищої освіти першого (бакалаврського) рівня вищої освіти
факультету іноземної філології*

Волинського національного університету імені Лесі Українки

Науковий керівник:

Олена КАРПІНА

кандидат філологічних наук, доцент

факультету іноземної філології

Волинського національного університету імені Лесі Українки, м. Луцьк

Комп'ютерна лінгвістика, а також нейронні мережі досягли високого рівня розвитку у 2026 році, проте проблеми лінгвістичної точності та керованості алгоритмів перекладу потребують доопрацювання. Актуальність цього дослідження полягає у перевірці точності перекладу іноземного дискурсу в системах, що мають різну архітектуру: системи, базовані на правилах (Apertium) та нейронні системи перекладу (DeepL). Коли перед нами стоїть завдання опрацювати тексти, що репрезентують різні функціональні стилі, то дуже важливо розуміти функціонал цих систем, адже вимоги можуть бути різними, як-от суворя термінологічна відповідність чи художня адаптація підтексту.

Щодо історіографії дослідження, то варто відзначити таких дослідників як Н. П. Дарчук, В. А. Широков та В. Старко, чії праці заклали фундамент для аналізу українського середовища в автоматизованих системах [1; 2; 3]. Однак, нейронні моделі швидко змінюються, а це, своєю чергою зумовлює постійну перевірку та зміну критеріїв оцінки їхньої ефективності, порівняно із загальноприйнятими підходами [5; 6].

Apertium (система, базована на лінгвістичних правилах – RBMT) влаштована за моделлю «білої скриньки» (White Box), адже вона запрограмована на завершених автоматах – формалізованих моделях для поетапної обробки мовних одиниць, та експертних лінгвістичних правилах. Це дає змогу поступово відстежувати морфологічний аналіз та структурний трансфер, що необхідно для наукового моделювання та навчальних цілей [1; 4]. Як зазначає М. Форкада, архітектура Apertium

забезпечує передбачуваність результату, що є її головною перевагою [4].

Натомість DeepL (NMT) створений на архітектурі Transformer використовує механізми «уваги» (Attention mechanisms). Це дозволяє системі глибоко розуміти контекст і враховувати статистичні ймовірності, проте логіка прийняття рішень не розкривається («чорна скринька») [5; 6].

Своєю чергою, Apertium досить точний у відтворенні термінологічних систем та збереженні формальної структури документа (нумерація, абзаци, модальність). Оскільки ці типи дискурсу дуже чіткі, жорстка алгоритмізація Apertium запобігає появі семантичних спотворень («галюцинацій»), притаманних нейронним мережам [3, с. 115]. В офіційно-ділових текстах Apertium виглядає досить стабільним, але потрібно задавати уточнюючі питання для перекладу модальних дієслів.

Публіцистика потребує врахування певних специфічних аспектів, а саме передачі іронії, метафор та гри слів. Apertium у багатьох випадках прямо перекладає текст, що руйнує його прагматику. Наприклад, заголовки новин містять велику кількість ідіом, які у системі Apertium змінюють своє первинне значення. Натомість DeepL досить точно передає зміст ідіом та добирає їхні змістові відповідники, що дає доволі точні результати в перекладі [2; 6].

Нейронні моделі стали повністю домінувати в літературному перекладі. Система DeepL генерує синтаксичні конструкції, адаптовані до норм української мови, що полегшує сприйняття тексту. Apertium поступається у цьому аспекті, адже значно гірше обробляє фрази [1, с. 44].

Дослідження показало важливу особливість: показники BLEU для системи Apertium часто занижені. Це пов'язано з тим, що метрика запрограмована на формальний збіг n-грамів і не розпізнає синоніми та схожі зазначенням слова чи фрази, які з точки зору граматики коректні, але відсутні в оригіналі [5]. Натомість DeepL отримує кращу оцінку BLEU, оскільки його функції орієнтовані на найбільш ймовірні (і, відповідно, наявні в еталонах) слова. Отже, автоматизовані метрики потребують доповнення експертним лінгвістичним аналізом помилок [1, с. 150].

Отже, проведене порівняння доводить, що Apertium залишається доцільним інструментом для опрацювання структурованих текстів, де помилка може призвести до спотворення змісту. DeepL демонструє високу ефективність у сфері медіа-комунікацій та літератури, де

важливо відтворювати природну мову. Перспективним напрямом розвитку перекладацьких технологій є інтеграція нейронних моделей із технологіями, базованими на правилах для забезпечення термінологічної та структурної точності.

ЛІТЕРАТУРА

1. Дарчук Н. П. *Комп'ютерна лінгвістика: автоматичне опрацювання тексту : підручник*. Київ : ВПЦ "Київський університет", 2024. 256 с.
2. Крижанівський А., Старко В. *Формальні моделі української морфології в системах автоматичного перекладу*. Філологічні студії. 2025. № 18. С. 45–58.
3. Широков В. А. *Комп'ютерна лінгвографія*. Київ : Наукова думка, 2024. 420 с.
4. Forcada M. L. *Making sense of machine translation*. Palgrave Macmillan, 2024. 210 p.
5. Jurafsky D., Martin J. H. *Speech and Language Processing*. Stanford University, 2025. URL: <https://web.stanford.edu/~jurafsky/slp3/>
6. Macketanz V., Burchardt A. *Evaluating Machine Translation Quality with Test Suites*. Springer Nature, 2024. 315 p.