

Якість машинного перекладу усних та письмових текстів

Характерними тенденціями розвитку сучасного суспільства є посилення процесів глобалізації та інформатизації, що призводять до збільшення й спрощення інтернаціональних зв'язків і контактів. Відповіддю на соціальний запит подолання мовного бар'єру стало утворення в промислово розвинених країнах індустрії машинного перекладу – галузі комп'ютерної лінгвістики, яка формувалася з початку 70-х років. В наш час вже розроблено дуже багато СМП, які класифікуються за різними принципами. Існує така типологія СМП:

щодо участі ЕОМ у процесі перекладу, обсягу роботи, яку вона «бере на себе», системи МП поділяються на автоматизовані та автоматичні;

кількість мов, що їх «розуміє» система: двомовні та багатомовні системи;

в залежності від підходу, на основі якого виконується машинний переклад, СМП поділяються на: статистичні та традиційні системи, які ґрунтуються на використанні лінгвістичних правил[2].

Розглянувши підходи до побудови систем МП, ми отримали уявлення про те, якими принципами програми-перекладачі користуються при перекладі тексту з однієї мови іншою:

- 1) СМП, що реалізують прямий переклад (direct translation);
- 2) СМП, які базуються на трансфері (transfer approach);
- 3) СМП з мовою-посередником (interlingua approach);
- 4) статистичні СМП (statistic).

Аналізуючи ступінь адекватності перекладу, перш за все необхідно визначити чітку кореляцію понять адекватності та еквівалентності. В цілому, у цей час спостерігається тенденція до розмежування понять «адекватність» та «еквівалентність» за принципом загального та часткового. Щодо адекватності машинного перекладу, то у цьому випадку варто послуговуватися підходом, запропонованим О.Д. Швейцером та Л.С. Бархударовим, які акцентують увагу саме на функціонально-прагматичному аспекті даної категорії [1; 3].

Зазвичай, до машинного перекладу, по-перше, звертаються з метою отримати найзагальнішу інформацію про першоджерело, а саме про його жанрову та тематичну співвіднесеність. Як правило, реципієнтами перекладу є фахівці у певній області, які або не володіють іноземною мовою у достатньому обсязі, або не мають часу, щоб опрацювати значну кількість текстів. Як показав наш аналіз, машинний переклад задовольняє цю вимогу у повному обсязі. У цьому випадку ми маємо стовідсоткову адекватність. Такий тип адекватності ми будемо називати жанрово-тематичною адекватністю.

По-друге, машинний переклад потрібен, коли треба вирішити, чи є необхідність у подальшому працювати з цим текстом. У цьому випадку переклад повинен передати основну комунікативну доміную оригіналу, надати отримувачу більш або менш детальну інформацію про зміст повідомлення, що міститься у першоджерелі. Наше робоче визначення для цього типу адекватності – адекватність інформативності.

По-третє, МП можуть застосовувати перекладачі у своїй професійній діяльності. Одна з головних цілей – отримання еквівалентів термінологічних складових першоджерела для подальшого редагування усього тексту рідною мовою. Згідно нашого бачення, це – семантико-синтаксична адекватність.

Як бачимо, при визначенні ступеню адекватності машинного перекладу на першій план виходить прагматичний фактор орієнтації на реципієнта перекладу та виконання поставлених завдань. Таким чином, адекватність МП визначається нами як адекватність відтворення основної прагматичної функції оригіналу. Тому вважаємо за доцільне розробити типологію помилок, які допускають СМП і на основі цієї типології оцінити рівень функціонально-прагматичної адекватності перекладу. У своєму дослідженні ми зупинили свій вибір на таких СМП як Google Translate та Pragma 6. Такий вибір зумовлений популярністю серед користувачів та вільним доступом до програм у режимі онлайн.

Отже, розрізняють такі типи помилок:

- Змістові помилки (перекручення інформації, неповнота або надлишковість, неперекладені фрагменти, втрата послідовності викладу).
- Термінологічні помилки (порушення єдності термінології, невідповідність термінів глосарію, словникам, галузевим стандартам).
- Мовні помилки (граматичні, лексичні, орфографічні, пунктуаційні).
- Стилістичні помилки (стильова невідповідність дібраних мовних засобів, незв'язність тексту, невідповідність структури речень, усталених фраз певному жанру).

Перекладацькі помилки є об'єктивним показником якості перекладу. Для аналізу нами було обрано публіцистичні статті загальним обсягом 16 533 знаків. Шляхом статистичного аналізу було визначено, що найпоширенішими є мовні помилки – близько 55%, стилістичні помилки – 20%, змістові помилки – 15% та термінологічні – 10%. Розглянемо помилки допущені при машинному перекладі за допомогою перекладачів Pragma 6 та Google Translate більш детально.

Показовим прикладом можуть слугувати такі речення:

The news was reported by Bloomberg, quoting diplomats involved in discussing the new sanctions, according to Ukrainian online newspaper Ukrainska Pravda [6].

За допомогою перекладача Pragma 6 отримуємо такий переклад:

Новини був повідомлений Bloomberg, вказуючи дипломатів, залучених в обговорення нових санкцій, згідно з Українською газетою Онлайн-правдою Ukrainska [5].

З поданого прикладу бачимо викривлення змісту на граматичному та стилістичному рівнях. Фразу *The news was reported by Bloomberg* перекладач Pragma 6 відтворив буквально, вдаючись до послівного перекладу: *Новини був повідомлений Bloomberg*. Як наслідок, спостерігаємо і граматичні помилки: відсутнє узгодження між членами речень; у ТП використовується пасивний стан дієслова, що не є характерним для норм української мови. Структура речення є абсолютно неадаптованою і чужорідною. Те ж саме спостерігаємо при перекладі такої частини речення: *quoting diplomats involved in discussing the new sanctions*. У ТП читаємо: *вказуючи дипломатів, залучених в обговорення нових санкцій*. Буквальний переклад лексичних елементів створює термінологічні та стилістичні помилки, оскільки відповідно до норм публіцистичного стилю доцільніше використовувати усталені фрази, на кшталт *з посиланням на дипломатів, близьких до переговорів*. Особливої уваги заслуговує переклад останньої частини речення. Англійське *according to Ukrainian online newspaper Ukrainska Pravda* в українській мові, завдяки перекладачеві Pragma 6, стало справжнім мовним покручем: *згідно з Українською газетою Онлайн-правдою Ukrainska*. Машинний перекладач повністю спотворив текст, перекутивши зміст повідомлення. Таким чином людина, незнайома з українським газетним виданням, не зможе вірно трактувати прочитане. На нашу думку, допущено і стилістичну помилку, замість ужитого слова *газета* доцільно було б використати словосполучення *інтернет-видання*.

Значно менше помилок допущено при використанні перекладача Google Translate:

Новина повідомив Bloomberg з посиланням на дипломатів, що беруть участь в обговоренні нових санкцій, відповідно до української інтернет-газета Українська правда [4].

Як і в попередньому випадку, текст характеризується наявністю помилок різного типу. Усупереч чинним граматичним нормам спостерігаємо порушення узгодження між словами. Так само як і Pragma 6, Google вжив лексему *газета*. Проте, варто відзначити і значну позитивну відмінність, за якої перекладач Google зберіг підрядну структуру речення, характерну нормам української мови.

Таким чином, бачимо, що СМП Pragma послуговується підходом буквального перекладу, тоді як Google Translator зміг адаптувати повідомлення, не спотворивши при цьому основного змісту.

Проблема вибору коректного еквіваленту терміну стає очевидна у такому прикладі:

Officials from the EU and the United States are meeting to discuss this issue in Brussels on Thursday [6]. – *Посадовці з ЄС та США засідання, щоб обговорювати цей випуск у Брюссель у четвер* [4].

Для терміну *officials* Pragma 6 подає варіант *посадовці*, а Google – *чиновники*.

Чиновники з ЄС та США зібралися, щоб обговорити це питання в четвер у Брюсселі [5].

Зважаючи на особливості публіцистичного стилю, на нашу думку, доцільніше було б вжити термін *представники*, оскільки словосполучення *представники ЄС та США* має конотацію офіційності. Таким чином, використання синоніму *представники* є цілком виправданим.

Зважаючи на результати аналізу перекладів програмами-перекладачами Google Translate і Pragma 6, можна стверджувати, що сучасні СМП не в змозі досягти семантико-синтаксичної адекватності і перекладачеві для редагування ТП обов'язково варто звертатися до оригінального тексту.

Якість перекладу Google Translate і Pragma 6 приблизно однакова і обидві СМП можуть виконувати переклади зі стовідсотковим рівнем жанрово-тематичної адекватності. Аналіз помилок показав, що при перекладі програми-перекладачі спотворюють зміст оригіналу на 50% і отриманий переклад не задовольняє у повній мірі вимоги реципієнта в отриманні релевантної інформації.

СМП Google Translate частіше підбирає коректні лексичні і фразеологічні еквіваленти, хоча в той же час виникають проблеми з побудовою граматично правильних речень. Це можна пояснити тим, що в базі даних цієї системи більше відповідностей, але контекст статистичної системи Google Translate становить лише 5 слів.

При перекладі довгих речень Google Translate, допустивши помилку на початку, не обов'язково допустить помилку в кінці, причому з більшою ймовірністю, ніж Pragma 6, яка в таких випадках спотворює все речення, намагаючись узгодити першу частину з другою частиною. При цьому, частіше всього, помилка переноситься також і на другу частину. Причиною більшості помилок в машинному перекладі є нерозуміння тексту системою МП, так як розуміння вихідного тексту є первинним етапом у процесі перекладу, а, відповідно, й основою правильного викладу його на іншій мові.

Так, ми бачимо, що Google Translate має засоби розуміння, що дозволяють в деякій мірі розуміти фразовий рівень мови, тоді як Pragma 6 – синтаксичний. У той же час, виходячи з результатів аналізу, обидві СМП достатньо добре здатні розуміти тільки морфологічний рівень.

Відтак, адекватність перекладу сучасними СМП – це скоріше випадковість, ніж показник здатності СМП адекватно перекладати. Первісна проблема досягнення адекватності машинного перекладу полягає в нездатності людства на даному етапі розвитку створити ІС з вищим рівнем розуміння. В першу чергу це зумовлено тим, що ще ніхто не зміг пояснити механізмів обробки і структуру зберігання інформації в людському мозку.

Література

1. Бархударов Л.С. Язык и перевод. Вопросы общей и частной теории перевода / Л.С. Бархударов. – М.: Междунар. отношения, 1975. – 238 с.
2. Кулагина О.С. Исследования по машинному переводу / О.С. Кулагина. – М.: Наука, 1979. – 320 с. 3. Швейцер А.Д. Теория перевода: проблемы, аспекты / А.Д. Швейцер. – М.: Наука, 1988. – 214 с. 4. Захід подовжив та розширив санкції проти Росії.– Перекладач Google Translate [Електронний ресурс]. – Режим доступу: <https://translate.google.com.ua/en/uk/west-may-expand-and-prolong-sanctions-against-russia.html>.
5. Захід подовжив та розширив санкції проти Росії. – Перекладач Pragma 6 [Електронний ресурс]. – Режим доступу: <http://translate.ua/us/on-line/west-may-expand-and-prolong-sanctions-against-russia.html>.
6. West may expand and prolong sanctions against Russia [Електронний ресурс]. – Режим доступу: <http://www.unian.info/world/1008916-west-may-expand-and-proong-sanctions-against-russia.html>.