

УДК 81'25

DOI <https://doi.org/10.32782/2522-4077-2025-213-27>

УПРАВЛІННЯ ЯКІСТЮ ПОСТРЕДАГУВАННЯ МАШИННОГО ПЕРЕКЛАДУ У ВИРОБНИЧОМУ ПРОЦЕСІ ЛОКАЛІЗАЦІЇ

MANAGING THE QUALITY OF MACHINE TRANSLATION POST-EDITING IN THE LOCALISATION WORKFLOW

Бондаренко О. С.,**orcid.org/0000-0002-3148-8753**

кандидат філологічних наук,

доцент кафедри перекладу, прикладної та загальної лінгвістики

Центральноукраїнського державного університету імені Володимира Винниченка

Негрієнко М. С.,**orcid.org/0009-0007-3373-7542**

студентка I курсу другого

(магістерського) рівня вищої освіти

Центральноукраїнського державного університету імені Володимира Винниченка

Статтю присвячено аналізу перекладацьких помилок у професійному середовищі та трансформації методів їхнього оцінювання під впливом новітніх лінгвістичних сервісів, зокрема постредагування машинного перекладу (ПРМП). Предметом дослідження є метрики якості перекладу, що застосовуються в індустрії лінгвістичних послуг, зокрема LISA QA Model, яку розглянуто як одну з найуживаніших моделей на ринку лінгвістичних послуг. Автори детально аналізують структуру метрики LISA, що включає категорії, підкатегорії та рівні критичності помилок. На прикладі реального проєкту – локалізованого продукту Google Search Ads 360 – продемонстровано практичне застосування метрики та наведено приклади типових помилок у таких категоріях, як *Точність*, *Мовна відповідність*, *Термінологія*, *Стиль*, *Технічна помилка*, *Відповідність стандартам локалі*, *Відповідність вимогам до проєкту*.

У статті обґрунтовано необхідність системного впровадження метрик помилок для управління якістю на всіх етапах перекладацького процесу. Традиційні моделі, створені для оцінювання антропогенного перекладу, для ПРМП мають бути адаптовані. На основі проведеного аналізу встановлено, що більшість категорій метрики LISA залишаються релевантними для оцінювання якості ПРМП, однак ступінь їх застосовності істотно варіюється. Найвищу вірогідність виявлення помилок мають підкатегорії відповідності термінології, стилю, вимогам до проєкту та стандартам локалі. Це пов'язано з непередбачуваністю результатів роботи рушіїв машинного перекладу навіть за умов використання однакових промптів чи довідкових ресурсів. Також варто зауважити, що помилки типу *Додавання/Випущення* чи *Неправильне значення* залишаються актуальними через тенденцію рушіїв МП до генералізації або неповного передавання змісту. Водночас, деякі підкатегорії (наприклад, *Орфографія*, *Форматування*, *Відтворення посилко/тегів*) мають низьку вірогідність застосування, оскільки сучасні МП-системи досить добре справляються з такими технічними аспектами.

Ключові слова: локалізація, управління якістю перекладу, виробничий процес перекладу, індустрія лінгвістичних послуг, постредагування машинного перекладу, машинний переклад.

The article is dedicated to the analysis of translation errors in a professional environment and the transformation of evaluation methods under the influence of modern linguistic services, particularly post-editing machine translation (PEMT). The focus of the study is on translation quality metrics used in the language services industry, particularly LISA QA Model, which is examined as one of the most widely used models on the market. The authors provide a detailed analysis of the structure of the LISA metric, which includes categories, subcategories, and levels of error severity. Using the real-world project of the localized product *Google Search Ads 360* as an example, the article demonstrates the practical application of the metric and presents examples of typical errors in categories such as *Accuracy*, *Language*, *Terminology*, *Style*, *Functional*, *Regional Standards Compliance*, and *Project Requirements Compliance*. The article substantiates the need for a systematic implementation of error metrics to manage quality at all stages of the translation process. Traditional models developed for the evaluation of human translation must be adapted for PEMT. Based on the analysis, it is established that most LISA metric categories remain relevant for evaluating PEMT quality, although their applicability varies

significantly. The highest likelihood of error detection occurs in subcategories related to terminology, style, compliance with project requirements, and locale standards. This is due to the unpredictability of machine translation engine outputs, even when identical prompts or reference materials are used. It is also noted that error types such as Addition/Omission and Incorrect Meaning remain relevant due to the tendency of MT engines to generalize or incompletely render content. Meanwhile, some subcategories (e.g., Spelling, Formatting, Tags/Links) have a low likelihood of errors, as modern MT systems handle such technical aspects fairly well.

Key words: localization, translation quality management, translation production process, language services industry, post-editing of machine translation, machine translation.

Постановка проблеми. Якість лінгвістичного продукту – ключове питання для індустрії лінгвістичних послуг. Як і в будь-якому сервісі, забезпечення сталої якості продукту – запорука успішності бізнесу. Системний підхід до якості, яка має бути такою, яка задовольнить клієнта – складова виробничого процесу локалізації.

Вже понад два десятиліття типовим для індустрії лінгвістичних послуг є застосування до оцінки якості перекладу так званих метрик помилок, використання яких гравцями на ринку дозволяє зробити оцінювання якості максимально об'єктивним. Питання ефективності метрик до оцінювання якості, зокрема, у зіставному аспекті, не часто розглядається вітчизняними науковцями, оскільки їхнє застосування переважно обмежується індустрією лінгвістичних послуг і не часто поширюється на науково-освітнє середовище, в якому навчають перекладачів [1]. Розробка та адаптація метрик оцінювання якості має неабияке прикладне значення зокрема й тому, що з появою нового сервісу – постредагування машинного перекладу (далі – ПРМП) – питання якості лінгвістичного продукту стало ще гостріше, адже ефективність метрик, які були розроблені для оцінки антропогенного перекладу, може не бути такою стосовно гібридного продукту, яким є постредагований людиною машинний переклад.

Аналіз досліджень і публікацій. Дослідження ефективності метрик помилок системно здійснюється переважно європейськими науковцями, які мають тісні зв'язки з ринком перекладу та локалізації. Системні наукові розвідки з якості перекладу здійснювалися Дж. Кобі, П. Філдсом, Д. Хейгом, А.Ломмелем, А.Мелбі [2], деякі з яких згодом працювали над стандартом ISO 17100:2015 Translation Services-Requirements for Translation Services. Предметом наукових досліджень ставали особливості застосування метрик оцінювання в окремих організаціях [3; 4]. Дослідники обговорювали необхідність системності термінології в метриках оцінювання [5]. Останніми роками тенденції досліджень змістились від визначення окремих принципів оцінювання до системних оглядів і визначення перспектив їхнього застосування, зокрема, й відносно нових лінгвістичних послуг [6; 7].

Питання якості перекладу в контексті професійного оцінювання в індустрії лінгвістичних послуг переважно розглядається дослідниками, безпосередньо дотичними до цієї галузі в якості перекладачів, менеджерів з оцінки якості [8].

На наш погляд, численні публікації, присвячені оцінці нерередагованого машинного перекладу, – продукту непрофесійних версій рушіїв машинного перекладу (як-от безкоштовний сервіс Google Translate – на противагу платному Google Cloud Translation або системи ШІ, які від початку не розроблялись для використання у виробничих ланцюжках професійного перекладу (наприклад, ChatGPT) [9; 10; 11] не можуть уважатися такими, що мають прикладну цінність для ринку лінгвістичних послуг. Першою з таких причин є те, що постачальники лінгвістичних послуг переважно користуються професійними інструментами машинного перекладу, які очікувано видають переконливіший результат. Другою причиною є те, що критерії оцінювання науковців, які застосовуються до нерередагованого машинного перекладу, переважно є такими, які не мають нічого спільного з індустріальними метриками оцінювання якості.

Метою дослідження є визначення особливостей професійної метрики помилок LISA у контексті перспектив їхнього використання для визначення якості лінгвістичного продукту, у тому числі, того, що є результатом постредагованого машинного перекладу.

Виклад основного матеріалу. Перша модель забезпечення якості LISA QA Model була розроблена в 1995 році асоціацією Localization Industry Standards Association (LISA) [12] для використання в галузі локалізації програмного й апаратного забезпечення. Хоча з 2011 року розробник LISA припинив випуск оновлень, запропоновані асоціацією методи стандартизації помилок все ще широко використовуються для оцінювання якості перекладу. Окрім метрики LISA QA Model є й інші рішення, як-от SAE J2450, MQM-DQF, ASTM4639 тощо [13]. Найпопулярнішими серед них є уніфікована метрика помилок MQM-DQF, інтегрована в основні CAT-інструменти, хоча LISA QA (TQI) все ще залишається популярною.

Метрики помилок, як правило, містять градацію за фактором критичності помилки для якості готового продукту. За LISA QA Model визначаються *critical* / *критичні*, *major* / *грубі* і *minor* / *незначні* помилки. Критична помилка – це помилка найбільшої ваги, оскільки може призвести до буквальних утрат (фізичних, фінансових, репутаційних тощо) постачальника або користувача перекладеного або локалізованого продукту. Критичною помилкою може бути, наприклад, неправильна локалізація одиниць вимірювання при відтворенні обсягу дози ліків: така помилка може призводити до критичних наслідків для хворого.

За метриками помилок, як правило, визначаються категорія і вага помилки. За LISA QA Model виділяють 7 категорій помилок: 1) *Accuracy* / *Точність*; 2) *Language* / *Мовна відповідність*; 3) *Terminology* / *Термінологія*; 4) *Style* / *Стиль*; 5) *Functional* / *Технічна помилка*; 6) *Regional* / *Відповідність стандартам локалі*; 7) *Compliance* / *Відповідність вимогам до проєкту*. У свою чергу зазначені типи помилок діляться на підтипи, які будуть розглянуті нижче.

Помилка категорії *Accuracy* / *Точність* констатується, коли зміст цільового продукту не вповні відповідає змісту вихідного. Це може відбуватися, коли лінгвіст не відтворив посилки, припустившись помилки підкатегорії *Cross references* / *Перехресні посилки* на розділи або компоненти продукту, як-от, програмні застосунки, посібники, документацію, довідковий матеріал тощо.

Помилки підкатегорії *Addition/Omission* / *Додавання/Випущення* так само можуть призводити до неточності передавання змісту, оскільки обсяг значення вихідного не відповідний обсягу значення цільового тексту.

Найтиповішим різновидом категорії *Accuracy* / *Точність* є помилка, яку відносять до підкатегорії *Incorrect Meaning* / *Неправильне значення*. Цю помилку констатують, коли перекладач відтворив зміст вихідного тексту неточно або неправильно, що призводить до некоректної інтерпретації цільового тексту користувачем. Так, наприклад, відтворення “*then you can use your account*” як “*тепер ви можете використовувати цей обліковий запис*” трактується як таке, що неправильно передає значення, і призводить до потенційно невірної інтерпретації, оскільки в перекладеному фрагменті акцент на персоналізації продукту зміщено. Такого б не відбулося, якби перекладач використав варіант: “*тепер ви можете використовувати свій обліковий запис*”, хоча така версія суперечить загальноприйнятому негативному ставленню до тавтології в традиційній теорії перекладу, яка здебільшого ґрунтується на аналізі літературних творів, для яких повторення займенників (*ви-свій*) може вважатися вадою. До різновиду помилок категорії *Accuracy* / *Точність* відносять підкатегорію *Unlocalized text* / *Нелокалізований уривок тексту*, коли частина помилково (не за вимогою клієнта) залишається неперекладеною.

До помилок категорії *Language* / *Мовна відповідність* в індустрії відносять неправильне використання граматики, стилістики, некоректну орфографію або синтаксис, які можуть ускладнити розуміння тексту. У прикладі “*In the toolbar above the statistics table, click Edit, and then click Create an automated rule*” – “У розділі “Змінити макс. ціни за клік”, виберіть “Призначити нові ставки” кома після “Змінити макс. ціни за клік” – зайва, адже така пунктуація відповідає нормам вихідної, а не цільової мови. Схожа помилка фіксується в наступному прикладі: “*Under “Change max. CPC bids”, you can select “Set new bids”, “Increase bids” or “Decrease bids”.* – У розділі “Змінити макс. ціни за клік”, виберіть “Призначити нові ставки”, “Збільшити ставки” або “Зменшити ставки”, бо кома перед виберіть – зайва, оскільки йдеться про

звичайну обставину місця, логічно пов'язану з дією. Принагідно відзначимо, що до підкатегорії *Punctuation* / Пунктуаційна помилка в індустрії відносять також порушення правил переносів і інтервалів. До підкатегорії *Spelling/Typo* / Орфографія/одрук можуть належати окрім очевидних одруків варіанти, які уможливають неоднозначність через варіативність наголосів. Інколи такі помилки можна трактувати як незначні, як-от, у формі слова “телеон” (замість “телефон”). У деяких випадках помилка має більшу вагу і спричинити введення в оману або спантеличення клієнта: *During registration, click “City” and enter its name – Під час реєстрації натисніть “Міст” й укажіть його назву.* Одрук *міст* замість *місто* може спричинити проблеми з реєстрацією і (потенційно!) до відмови від користування продуктом. Як помилки підкатегорії *Grammar/Syntax* / Граматика/ Синтаксис трактують неправильне утворення відмінкових форм, ступенів порівняння прикметників або прислівників, уживання ненормативних форм дієслів, нечіткі смислові зв'язки, дієприслівниковий зворот, не пов'язаний із присудком однією дійовою особою тощо.

Помилки категорії *Terminology* / Термінологія свідчать про недотримання лінгвістом правил відтворення термінів. Така помилка констатується при наданні неправильного термінологічного відповідника, неузгодженості термінів чи недотриманні глосарію клієнта. До підкатегорії *Industry Standard Terminology* / Загальногалузева термінологія відносять випадки залучення термінів іншої фахової мови. Така помилка може призвести до неправильного розуміння тексту, спотворення змісту й втрати смислу. У підкатегорії *Inconsistency* / Єдність термінології фіксуються випадки, коли лінгвіст використовує декілька відповідників термінам чи терміносполукам замість одного, що може спричинити нерозуміння клієнтом змісту. Так, наприклад, термін *account* в одній частині проєкту відтворено як *обліковий запис*, а в іншій – як *акаунт*: *To create an account, click “Sign Up”, enter your email, set a password. Make sure to secure your account by enabling Two-Factor Authentication and using a strong password. – “Щоб створити обліковий запис, натисніть “Зареєструватися”, введіть свою електронну пошту й встановіть пароль. Обов'язково захистіть свій акаунт, увімкнувши двофакторну автентифікацію й використовуючи надійний пароль”.* Помилки підкатегорії *Glossary* / Відповідність глосарію констатуються, коли лінгвіст використовує відповідники термінів, відмінні від наданих клієнтом в глосарії (термінологічній базі).

Помилки категорії *Style* / Стиль – найскладніші для визначення, особливо на пострадянському просторі, де стиль і стильові особливості розуміються в контексті функційних стилів, серед яких традиційно виділяють стиль художнього мовлення, офіційно-ділового, наукового, публіцистичного, конфесійного тощо. Узагалі помилки стилю є неоднозначними, оскільки стиль не завжди має чітко визначені правила, тому такі помилки часто можуть визначатися суб'єктивно. Треба точно визначити, чи припустився перекладач стильової помилки чи все ж таки це є преференційною правкою клієнта або мовного експерта. Певним орієнтиром можуть уважатися так звані настанови зі стилю, які деякі клієнти розробляють із метою реконструкції особливого “лінгвального образу” або “голосу продукту”, який має відтворюватися у всіх локалях, для яких його адаптують [14]. До помилок підкатегорії *General Style* / Відповідність загальному стилю відносять неправильний регістр, невідповідний рівень офіційності (звертання на “ти” чи на “ви”), недотримання норм певного стилю тощо. До помилок підкатегорії *Style Guide* / Відповідність настановам зі стилю відносять випадки неузгодженості цільового продукту настановам, наданим клієнтом. Як правило, такі настанови надаються разом із файлами перекладу і містять рекомендації і правила щодо пунктуації, правопису певних слів, створення списків тощо. До помилок підкатегорії *Readability* / Читабельність відносять випадки, коли цільовий текст, незважаючи на відтворений зміст, за формою не дозволяє читачеві-носію мови повноцінно декодувати зміст. Так, наприклад, відповідник англійському: *If you're viewing your manager account, the Overview cards display performance summaries for all of the sub-manager accounts, campaigns and ad groups. – Якщо ви в керуючому обліковому записі,*

на картках “Огляд” відображаються підсумки ефективності для всіх підпорядкованих керуючих облікових записів, кампаній і груп оголошень не дозволяє одразу зрозуміти перспективи доступності певного функціоналу продукту відповідно до режиму користувача (що передбачалось у вихідному фрагменті).

Категорію *Functional / Технічна помилка* в теорії перекладознавства зазвичай не помічають, вважаючи, що лінгвіст несе відповідальність лише за лінгвальні помилки. Для галузі локалізації такі вади мають критичне значення, особливо якщо клієнт у вимогах до проекту наголошував на дотриманні певних технічних чи технологічних вимог. Серед підкатегорій відзначимо *Format / Форматування*, коли перекладач, наприклад, не виділяє напівгрубим шрифтом те, що виділено таким чином у вихідному тексті. Помилка підкатегорії *Tags/Links / Відтворення посилоч/тегів* може набути статус критичної, якщо посилки в цільовому тексті не працюють належним чином, чи змінено порядок тегів. До підкатегорії *Technical procedures / Технічне оформлення* відносять такі вади, як, наприклад, неправильне (невідповідне нормі локалі чи вимогам клієнта) оформлення списків тощо.

Категорія *Regional / Відповідність стандартам локалі* застосовна до помилок, які є наслідком неврахування особливостей цільової локалі. У межах цієї категорії виділяють кілька підкатегорій. До підкатегорії *Regional/Country Standards / Відповідність місцевим/національним стандартам документації* відносять помилки відтворення дат, одиниць виміру, валют, адрес, номерів телефонів, поштових індексів тощо: наприклад, в англійських країнах дати оформлюють як: *місяць/число/рік*, а в українській локалі – *число/місяць/рік*. До підкатегорії *Local Market Suitability / Відповідність місцевому ринку* відносять помилки, пов’язані з неврахуванням особливостей законодавства, усталених традицій, назв посад або відповідальних осіб певної локалі.

Категорія *Compliance / Відповідність вимогам до проекту* застосовується до випадків недотримання інструкцій, отриманих від менеджерів проекту або клієнта. У підкатегорії *Instructions / Відповідність інструкціям* фіксуються, наприклад, транслітерації власних назв, хоча клієнт наполягав на їхній незмінності. Підкатегорія *Feedbacks / Врахування відгуків клієнта* призначається варіанту, який є наслідком недотримання попередньо висловлених зауважень або настанов.

На наш погляд, застосування проаналізованої моделі відносно антропогенного перекладу та постредагування машинного перекладу виявляє різні ступені ефективності. Говорити про появу нової сталої метрики оцінювання постредагування машинного перекладу поки зарано, але вже очевидним є втрата актуальності деяких категорій у метриці, оскільки таких помилок не припускаються навіть рушії машинного перекладу.

Висновки. Як бачимо, традиційні метрики оцінювання антропогенного перекладу все ще переважно застосовні для продукту рушіїв машинного перекладу. Натомість визначення незастосовних категорій має певну цінність для редакторів чи менеджерів із якості, оскільки оптимізує метрику для їхніх потреб. Водночас ті категорії, де рушій може видавати непередбачуваний варіант (наприклад, може врахувати довідкові матеріали, а може й ні), потребують додаткової уваги постредакторів-фрилансерів або менеджерів, які розраховують час і бюджет для доопрацювання машинного перекладу до якості, необхідної для клієнта.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Бондаренко О. Оцінка якості перекладу: дидактичний та галузевий підходи. *Актуальні проблеми романо-германської філології та прикладної лінгвістики*. 2018. № 1(15). С. 47–49.
2. Koby G. S., Fields P., Hague D. R., Lommel A., Melby A. Defining translation quality. *Tradumàtica*. 2014. № 12. С. 0413–0420.
3. Dragan J. Quality in professional translation: Assessment and improvement. Vol. 9. London : A&C Black, 2013. 272 p.

4. Snow T. A. Establishing the viability of the multidimensional quality metrics framework. Provo : Brigham Young University, 2015. 156 p.
5. Brunette L. Towards a terminology for translation quality assessment: A comparison of TQA practices. *Evaluation and Translation*. London : Routledge, 2014. P. 169–182.
6. Moorkens J., Castilho S., Gaspari F., Doherty S. Translation quality assessment. Machine translation: Technologies and applications series. Cham : Springer International Publishing, 2018. Vol. 1. 299 p.
7. Han C. Translation quality assessment: a critical methodological review. *The Translator*. 2020. Vol. 26, № 3. P. 257–273.
8. Іващенко Р. В. Особливості локалізації програмного забезпечення на матеріалі відеохостингу YouTube: кваліфікаційна робота на правах рукопису. Кропивницький: Центральноукраїнський державний університет імені Володимира Винниченка, 2023. 123 с.
9. Мишко А., Гудманян А., Брай А. Оцінка адекватності машинного перекладу письмових спеціалізованих текстів. *Advanced Linguistics*. 2023. № 12. С. 66–77.
10. Ольховська А., Хоменко А. Вплив використання нейронної системи машинного перекладу на якість перекладу текстів у галузі економіки. *Молодь і ринок*. 2020. № 5(184). С. 30–35.
11. Ayvazyan N., Torres-Simón E., Pym A. Translation students' trust in machine translation: Too good to be true? *Revolutionizing Translation Studies: Synthesizing Translation with AI and IT innovation*. Berlin : Springer, 2024. P. 9–30.
12. LISA QA Model 3.1: Assisting the localization development, production and quality control processes for global product distribution. Romainmôtier : LISA, 2006. (Press release).
13. Lommel A., Uszkoreit H., Burchardt A. Multidimensional quality metrics (MQM): A framework for declaring and describing translation quality metrics. *Tradumàtica*. 2014. № 12. P. 0455–0463.
14. Bondarenko K. Ukrainian localization: Cultural aspects. Traducción e interpretación como mediación (inter)cultural: Visiones y perspectivas de futuro / eds. M. F. de Casadevante Mayordomo, A. C. Sánchez López. Madrid : Comares, 2024. P. 229–245.