

Бондаренко О. С.,*кандидат філологічних наук,**доцент кафедри перекладу, прикладної та загальної лінгвістики**Центральноукраїнського державного університету імені Володимира Винниченка***Бондаренко К. Л.,***кандидат філологічних наук,**доцент кафедри перекладу, прикладної та загальної лінгвістики**Центральноукраїнського державного університету імені Володимира Винниченка*

ПОСТРЕДАГУВАННЯ МАШИННОГО ПЕРЕКЛАДУ ЯК СКЛАДОВА ПОСЛУГИ ЛОКАЛІЗАЦІЇ

Анотація. У статті проаналізовано послугу постредагування машинного перекладу (ПМП) в контексті складової сервісу локалізації. Актуальність дослідження зумовлена недостатньою (близько 3%) представленістю ПМП в українських освітніх програмах із підготовки перекладачів. Дослідження має на меті висвітлення ролі та значення ПМП на ринку локалізаційних послуг, аналіз нормативних документів, що регулюють ПМП, та змін в операційних процесах надання лінгвістичних послуг на загальноєвропейському ринку за 2023–2024 роки. Загальноєвропейські опитування, такі як «European Language Industry Survey» (ELIS), проведені EUATC, Elia, EMT, FIT Europe, GALA, LIND і WiL, фіксують значні зміни в ринку лінгвістичних послуг. Дані загальноєвропейського дослідження ELIS (European Language Industry Survey 2024, а також зміст воркшопів у рамках ініціативи «Translating Europe» у 2023–2024 роках довів визначальну роль, що відіграє сервіс постредагування в індустрії лінгвістичних (локалізаційних) послуг. Ключові поняття, зафіксовані в інституційних документах із постредагування машинного перекладу, включають: вузькогалузеві («машинний переклад», «нередагований машинний переклад», «повне постредагування» тощо), лінгвістичні та змістові («зміст вихідною мовою», «зміст цільовою мовою» тощо), організаційні («постачальник перекладацьких послуг», «менеджер проєктів» тощо) та власне перекладові («результат перекладу», «перекладацька послуга» тощо). Цей процес має власну поняттєву систему і вимагає специфічного підходу до тексту, залучення особливих когнітивних механізмів, що відрізняються від тих, які використовуються в класичному перекладі. Хибне сприйняття постредагування як асистемного процесу, не вартого наукового вивчення, свідчить про необізнаність із сучасними тенденціями в індустрії лінгвістичних (локалізаційних) послуг, у якому постредагування вже займає значну частку ринку, і його важливість щодалі зростає.

Ключові слова: машинний переклад, постредагування, локалізація, перекладацька послуга, технологічні компетентності.

Постановка проблеми. Постредагування машинного перекладу – відносно новий вид діяльності лінгвістів, який, на жаль, майже не відбито в українських освітніх програмах підготовки перекладачів. Проведене авторами в 2023 дослідження відносно місця і ролі машинного перекладу (а також постредагування машинного перекладу) засвідчило, що, на думку викладачів

перекладацьких програм одним із чинників, що перешкоджають упровадженню відповідних курсів на рівні бакалаврату чи магістратури – є відсутність фахівців, які могли б викладати такі курси. Зауважимо, що представники українських перекладацьких та локалізаційних компаній, згідно тому ж опитуванню, проводять тренінги чи організують курси з постредагування для свого персоналу, що свідчить про те, що, вірогідно, фахівців на ринку лінгвістичних (перекладацьких та локалізаційних) послуг не бракує [1]. Припускаємо, що відсутність необхідної уваги до такого сервісу представників академічної спільноти зумовлена, скоріш за все, браком розуміння ринку лінгвістичних послуг та тих тенденцій, які він демонструє.

Аналіз останніх досліджень і публікацій. Наукове осмислення сервісу постредагування машинного перекладу у світі активно почалося приблизно 20 років тому. Переважна кількість досліджень була присвячена питанням якості продукту, згенерованого машиною та відредагованого людиною, та перекладу, повністю виконано людиною з однозначним висновком щодо кращої якості останнього [2; 3]. Згодом науковці почали зважати на вимоги ринку і концентруватися на визначенні обсягу зусиль лінгвіста з постредагування машинного перекладу [4]. На певному етапі науковці почали обговорювати той незаперечний для індустрії факт, що кінцевою метою постредагування не завжди має вбачатися досягнення якості перекладу, зробленого людиною [5], хоча в Україні дослідження з цієї проблематики все ще поодинокі [6; 7].

Навчання принципам постредагування студентів перекладацьких програм висвітлювалось ще на початку XXI ст. [8], однак українські дидакти почали звертати увагу на цей сервіс порівняно недавно в контексті набуття здобувачами технологічних компетентностей [9]. Інколи науковці констатують «недостатню» якість згенерованого руськими перекладу без об'єктивного порівняння з продуктами «достатньої» якості і без урахування поточної ситуації з наданням такої послуги на ринку [10]. Актуальність дослідження зумовлена мінімальною представленістю сервісу постредагування машинного перекладу в українських програмах підготовки перекладачів (близько 3%), що дисонує з затребуваністю сервісу в європейській та світовій індустрії лінгвістичних (локалізаційних) послуг.

Метою нашого дослідження є висвітлення місця і ролі постредагування на ринку лінгвістичних (локалізаційних) послуг. Серед завдань: аналіз документів, які «інституцію-

ють» послугу постредагування машинного перекладу; огляд змін в операційних процесах надання лінгвістичних послуг на загальноєвропейському ринку станом на 2023–2024 роки.

Виклад основного матеріалу. Загальноєвропейські опитування відносно стану та тенденцій розвитку ринку лінгвістичних послуг, зокрема, «European Language Industry Survey», ініційовані та проведені Європейською спілкою асоціацій перекладацьких компаній (EUATC), Асоціацією європейської індустрії лінгвістичних послуг (Eliа), мережею «Європейських магістерських програм із письмового перекладу» (EMT), Європейським регіональним центром міжнародної федерації перекладачів (FIT Europe), Асоціацією «Globalization and Localization Association» (GALA), експертної групи фахівців лінгвістичної галузі Генерального Директорату Європейської Комісії з питань (LIND) та організації «Women in Localization» (WiL) вже впродовж кількох років фіксують значні зміни в виробничому процесі надання лінгвістичних послуг. Згідно з аналізом дослідження ELIS 2024, представленому в доповіді «Trends, expectations and concerns of the European language industry» у березні 2024 року, використання технологій, передусім, машинного перекладу, пік якого зафіксовано в постпандемійний рік, станом на зараз стало частиною операційного процесу компаній, що надають лінгвістичні послуги. Вперше обсяг послуги постредагування перевищив обсяги замовлень на переклад, повністю виконаний людиною.

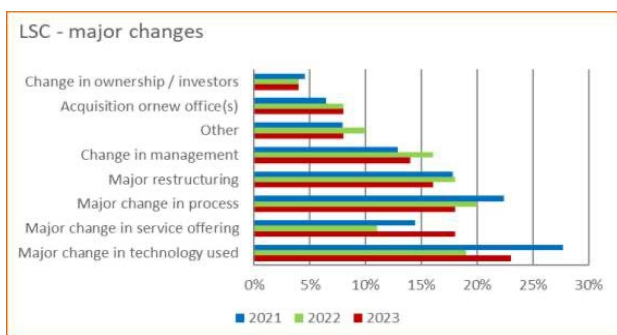


Рис. 1. Зміни в операційних процесах постачальників лінгвістичних послуг. Лінгвістичні компанії [11]

Покажемо щодо місця і ролі, які починає посідати машинний переклад у виробництві лінгвістичного продукту, є його фіксація як фактору, що спричиняє відчутні зміни в виробничій діяльності фрилансерів, – найбільшого за обсягом гравця на ринку лінгвістичних послуг. 2024 рік став першим роком, коли ці гравці офіційно зафіксували машинний переклад у якості складової виробництва лінгвістичних послуг.

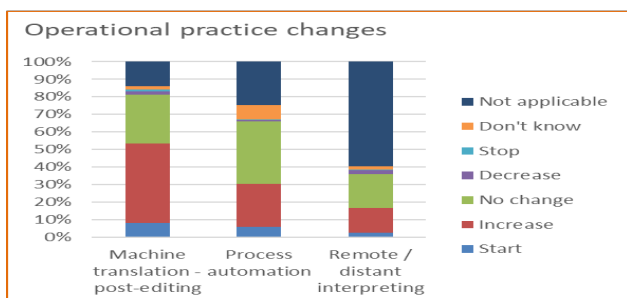


Рис. 2. Зміни в операційних процесах надання лінгвістичних послуг. Індивідуальні постачальники лінгвістичних послуг [11]

Симптоматичним в цьому контексті є намагання відрефлексувати зміну ролі та статусу машинного перекладу в індустрії лінгвістичних послуг на професійних зустрічах, конференціях тощо. Аналіз змісту 30 локальних воркшопів для практиків та представників університетів в рамках ініціативи «Translating Europe», проведених у 2023–2024 роках, доводить, що 42% (12 заходів) були присвячені проблемам (постредагування) машинного перекладу, як-от: «Myths and truths about machine translation in 2023» (Орхус (Данія), 25.05.2023), «Unlocking Borders: Post-editing machine translation nowadays and beyond» (Клуж-Напока (Румунія), 15.03.2023), «Machine translation in specialized contexts» (Дублін (Ірландія), 11.06.2024), «AI translation vs human translation: training perspectives and the impact on the translation profession» (Керкіра (Греція), 24.06.2024) тощо [12].

Індустрія лінгвістичних послуг намагається проаналізувати вплив машин на виробничий процес глибше: так, наприклад, темою найбільшої індустріальної конференції GALA 2024, яку щорічно проводить «The Globalisation and Localisation Association» стало «Navigating the Rise of AI: Dos and Don'ts, Myths and Truths for a New AI-Assisted Localization Industry». Принагідно зауважимо, що AI (Штучний інтелект – далі ШІ) в контексті лінгвістичних послуг розуміється як автоматизовані системи, які генерують машинний переклад. Здебільшого в дискусіях практиків ідеться про масиви згенерованих машинною текстами, із яким згодом працюють лінгвісти, адаптуючи під потреби клієнтів. Глобальну тему представницької зустрічі реалізовано в низці секцій, присвяченим, наприклад: питанням забезпечення носіїв непоширених мов перекладеним машиною і відредагованою людиною контентом («AI for the Next 4 Billion People»); етичності і вразливості використання налаштованих під клієнта рушіїв і результатів генерації інструментів, побудованих із застосуванням великих мовних моделей («Ethics of AI in the Translation Industry – Aren't We the Most Impacted Field?»). Значна кількість засідань присвячена вже не факту використання ШІ на ринку лінгвістичних послуг, а окремим сферам чи видам діяльності, де використання цих інструментів може становити загрозу для якості виробленого продукту чи ефективності виробництва. Так, наприклад, підбір персоналу та взаємодія з виконавцями залишається людині, оскільки великі мовні моделі, що живлять штучний інтелект, можуть містити гендерні або вікові перестороги, що призводять до упередженості при підборі виконавців (секція «Rage Against the Machine: AI, Stay Away From People Management») [13].

На особливу увагу заслуговують прагнення ключових гравців ринку унормувати процес взаємодії з продуктами, згенерованими штучним інтелектом, представлені майже одночасно на круглому столі «Technology, Ethics and Quality in the Translation Industries» Другого міжнародного конгресу «Translation and Cultural Sustainability. Challenges and New Avenue» та на засіданні GALA 2024 «Fit for the Future: Knowing which Standards Matter and Why You Should Care». Наразі стандартизування та унормування надання такої послуги набуває особливої ваги в зв'язку з тим, що клієнти та кінцеві споживачі мають бути обізнані з походженням продукту (лінгвістичного, в нашому випадку), який вони замовляють або споживають. Користувач має розуміти, чи має він справу з продуктом, повністю згенерованим штучним інтелектом без подальшої адаптації людиною (навіть з теперішнім розвитком систем машинного перекладу такий варіант може бути цілком прийнятним для

споживача) або з гібридом – машинним перекладом, відредагованим людиною. Слідування постачальників лінгвістичних послуг стандартам – запорука сталої якості послуги і стабільної задоволеності клієнтів.

Наявний на ринку з 2017 року Стандарт із постредагування «Translation services – Post-editing of machine translation output – Requirements» описує базові складові послуги [14]. Враховуючи те, що стандарти потенційно мають бути згармонізованими з уже існуючими, поряд із усталеним в індустрії номінаціями наводитимемо ті, які використано в стандарті ISO 17100: 2015, який Україна ухвалила методом підтвердження англійською мовою як ДСТУ EN ISO 17100: 2017 з наданням йому чинності з 01.10.2017, і технічних умовах ISO/TS 11669:2012 [15; 16]. До ключових для постредагування/післяредагування понять відносяться: 1) вузькоспеціалізовані поняття машинного перекладу; 2) поняття лінгвістичного та змістового характеру; 3) поняття організаційного характеру; 4) поняття, пов'язані з перекладом.

До першої групи відносять вузькоспеціалізовані поняття, як-от *машинний переклад/перекладання; нередатований машинний переклад*, який означає згенерований рушієм машинного перекладу результат цільовою/перекладовою мовою; *система машинного перекладу; повне (часткове) постредагування/післяредагування*). Друга група містить поняття лінгвістичного та змістового характеру: *контент/зміст* (тексту, аудіо, відео), *вихідна/першотвірна мова, зміст вихідною мовою/першотвірний зміст, цільова/перекладова мова, зміст цільовою мовою/зміст перекладовою мовою, текст, природня мова, контрольована мова, сегмент*. До третьої групи відносяться поняття організаційного характеру: *клієнт, перекладач, вивірляльник, постредактор/післяредагувальник, постачальник перекладацьких послуг/ надавач перекладчих послуг, менеджер/ керівник проєктів*. До четвертої групи віднесено поняття, пов'язані з перекладом/перекладанням: *перекладати, переклад/перекладання, переклад людиною/антропогенний переклад* [9], *результат перекладу/перекладання, перекладацька/перекладча послуга, пам'ять перекладів, вичитування/звіряння, verification (підтвердження керівника проєкту*.

Як зазначалось, розвиток технологій здатний призводити до того, що певні поняття стандарту, не встигнувши закріпитися інституційно, ідуть в небуття із зникненням потреби, що яскраво демонструє сервіс «часткового» та «повного» постредагування. Часткове постредагування – це процес редагування машинного перекладу, спрямований на створення зрозумілого тексту, без прагнення досягти якості, подібної до людського перекладу [14]. Найбільші гравці ринку лінгвістичних послуг визначають якість такої послуги як таку, що є «достатньої якості» («good enough/fit for purpose»). Процес постредагування на цьому етапі передбачає мінімальне виправлення граматичних та орфографічних помилок; редагування речень, що можуть увести в оману; виправлення типових помилок машинного перекладу; видалення зайвих варіантів перекладу, згенерованих машиною [17]. Серед проблем, із якими на цьому етапі стикаються недосвідчені лінгвісти, – невміння планувати час та зусилля з урахуванням специфіки сервісу, яке реалізовано в надмірному виправленні згенерованого машиною продукту в прагненні досягнути якості перекладу, зробленого повністю людиною. Незважаючи на те, що надання цього сервісу передбачає виправлення граматичних, синтаксичних, пунктуаційних та семантичних помилок, перевіряти узгодженість термінології

тощо, ключовим залишається максимально ефективно використання результату машинного перекладу, а не «переписування» перекладу.

Зауважимо, що проблема розмежування та навчання частковому та повному постредагуванню машинного перекладу, так достатньо і недорефлексована науковою спільнотою, може незабаром стати неактуальною: сервіс часткового (light) редагування, зважаючи на результати машинного перекладу, який завдяки тренуванню рушіїв на великих мовних моделях стає кращим, перестає цікавити замовників – відтак потреба у «частковому постредагуванні» як у окремій лінгвістичній послугі виникає в індустрії перекладу все рідше.

При цьому сам сервіс постредагування машинного перекладу слід відмежовувати від перекладу людиною: він має власну поняттєву систему, його виконання потребує особливого підходу до тексту, навіть когнітивні механізми, які забезпечують взаємодію зі згенерованим рушієм продуктом – відмінні від тих, які застосовуються у класичному перекладі. Сприйняття сервісу як «недоперекладу», а зусиль, що докладаються на постредагування машинного перекладу, занадто асистемними, аби ставати предметом наукових розвідок, – хибне, оскільки свідчить про повну необізнаність із тим, з чого вже майже наполовину складається індустрія лінгвістичних (локалізаційних) послуг.

Література:

- Bondarenko K., Bondarenko O. Machine translation impact on translation education and professional practice in Ukraine. *Abstracts of 2nd International Congress «Translation and Cultural Sustainability. Challenges and New Avenues»* (2024, April 17-19). 2024. P. 129.
- Fiederer R., O'Brien S. Quality and Machine Translation: A realistic objective? *The Journal of Specialised Translation*. 2009. № 11. P. 52-74.
- Krings H.P. Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes (Geoffrey S. Koby, ed.). *Translators' Journal*. 2001. № 47. P. 435-436.
- Martinez L.G. Human translation versus machine translation and full post-editing of raw machine translation output: dissertation of MA in Translation Studies: 18.08.2003. Dublin, 2003. 137 p.
- O'Curran E. Translation Quality in Post-Edited versus Human-Translated Segments: A Case Study. *Proceedings of the Third Workshop on Post-editing Technology and Practice (WPTP-3)*. 2014. P. 113-118.
- Бондаренко О. Оцінка якості перекладу: дидактичний та галузевий підходи. *Актуальні проблеми романо-германської філології та прикладної лінгвістики*. 2018. № 1(15). С. 47-49.
- Каложний О.М. Особливості технологій машинного перекладу та їх вплив на продуктивність роботи перекладача: Кваліфікаційна робота другого (магістерського) рівня вищої освіти. Центральнороманський державний педагогічний університет імені Володимира Винниченка. 2020. 112 с.
- Shei C. Teaching MT through pre-editing: three case studies. *Proceedings of the 6th EAMT Workshop Teaching Machine Translation*. 2002. P. 89-98
- Черноватий Л. Проблеми машинного перекладу та його застосування у навчанні майбутніх перекладачів. *Наукові записки. Серія: Філологічні науки*. 2023. Вип. 202. С. 84-93.
- Гудманян А.Г., Сітко А.В., Струк І.В. Функціонально-прагматична адекватність машинного перекладу публіцистичних текстів. *Львівський філологічний часопис*. 2019. Вип. 5. С. 48-54.
- ELIS (European Language Industry Survey). Trends, expectations and concerns of the European language industry. 2024. Режим доступу: <https://elis-survey.org/>

12. Translating Europe Workshops. Режим доступу: https://commission.europa.eu/about-european-commission/departments-and-executive-agencies/translation/translating-europe/translating-europe-workshops-tew_en
13. GALA 2024. Program. Retrieved from <https://www.gala-global.org/conference/program>
14. ISO 18587: 2017 Translation services – Post-editing of machine translation output– Requirements.
15. ISO 17100:2015 [Translation services – Requirements for translation services: ISO 17100:2015. First edition, 2015-05-01]. VI, 20 p. (Перекладча послуга. – Вимоги до перекладчих послуг).
16. ISO/TS 11669:2012 [Translation projects – General guidance: ISO/TS 11669:2012. First edition, 2012-05-15]. VI, 36 p. (Перекладчий проект. – Загальні поради).
17. RWS Alpha. What You Need to Know About Light and Full Post-editing. 2020. Режим доступу: <https://alpha.rws.com/blog/what-you-need-to-know-about-light-and-full-post-editing>.

Bondarenko O., Bondarenko K. Post-editing machine translation as an element of localisation service

Summary. Post-editing of machine translation (PEMT) is a relatively new type of activity for linguists. The relevance of the study is due to the insufficient (approximately 3%) representation of PEMT in Ukrainian educational programs for translators. The research aims to highlight the role and significance of PEMT in the localization services market, analyze regulatory documents regulating PEMT, and changes in operational processes of providing linguistic services in

the overall European market for 2023-2024. European surveys such as the European Language Industry Survey (ELIS), conducted by EUATC, Elia, EMT, FIT Europe, GALA, LIND, and WiL, record significant changes in the linguistic services market. Data from the European Language Industry Survey 2024, as well as the content of workshops within the «Translating Europe» initiative in 2023-2024, have demonstrated the decisive role played by post-editing services in the linguistic (localization) services industry. Key concepts recorded in institutional documents on post-editing of machine translation include concepts related to machine translation («machine translation», «machine translation output», «full post-editing», «light post-editing»), linguistic and content-related concepts («content», «source language», «source language content», «target language», etc.), organizational concepts («translation service provider» and «project manager»), and translation-specific concepts («translation output», «translation service», etc.). This process has its own conceptual system and requires a specific approach to the text, as well as special cognitive mechanisms that differ from those used in traditional translation. Misconceptions about post-editing as a non-systematic process, not worthy of scientific study, indicate a lack of awareness of modern trends in the linguistic (localization) services industry, where post-editing already occupies a significant share of the market, and its importance continues to grow.

Key words: machine translation, post-editing, localisation, translation service, technological competence.