

## ПОСТРЕДАГУВАННЯ ПРИ МАШИННОМУ ПЕРЕКЛАДІ

### Тетяна Корольова

доктор філологічних наук, професор, завідувач кафедри перекладу і теоретичної та прикладної лінгвістики Державного закладу «Південноукраїнський національний педагогічний університет імені К. Д. Ушинського»  
Одеса, Україна  
e-mail: kortami863@gmail.com  
ORCID ID <https://orcid.org/0000-0003-3441-196X>

### Наталя Жмаєва

кандидат філологічних наук, доцент кафедри перекладу і теоретичної та прикладної лінгвістики Державного закладу «Південноукраїнський національний педагогічний університет імені К. Д. Ушинського»  
Одеса, Україна  
e-mail: zhmaeva@gmail.com  
ORCID ID <https://orcid.org/0000-0003-3382-0155>

### Юлія Колчаг

магістр зі спеціальності 035 Філологія Державного закладу «Південноукраїнський національний педагогічний університет імені К. Д. Ушинського»  
Одеса, Україна  
e-mail: juliaigorivna@ukr.net

### АНОТАЦІЯ

Сучасна індустрія перекладацьких послуг виділяє два рівня якості перекладу, яких можна досягти в результаті постредагування МП: «задовільна» якість («good enough» quality) передбачає відтворення основної інформації вихідного повідомлення, припускає наявність стилістичних, синтаксичних та морфологічних оєріхів; якість перекладу, наближена або рівна якості перекладу, виконаного людиною («quality similar or equal to human translation») — «парадний» варіант відредагованого тексту, придатний для публікації.

Огляд систем МП дозволяє стверджувати, що найпотужнішою системою МП на сьогоднішній день є Google Neural Machine Translation (GNMT), яка використовує найсучасніші методи навчання для досягнення найбільших удосконалень.

Під час аналізу текстів, перекладених за допомогою Google Translate, були констатовані наступні проблеми: спотворення референційного змісту вихідного повідомлення, неправильний добір варіантних відповідей, відсутність уніфікованості термінів, невідтворення абревіатур, неузгодженість мовних одиниць за особами, числами і відмінками, неправильний добір функціонального аналога при відтворенні абсолютних, герундіальних та дієприкметникових конструкцій, буквальний переклад словосполучень, відсутність застосування трансформацій граматичної структури вихідного повідомлення (додавань, перестановок).

З оглядом класифікованих проблем машинного перекладу, а також рівней якості, які висуваються до постредагування, проведено постредагування текстів, перекладених засобами машинного перекладу, та сформульовані вимоги та рекомендації до постредагування результатів машинного перекладу у межах дослідженої мовної пари з урахуванням особливостей конкретної системи МП та типу текстів, що перекладаються.

**Ключові слова:** машинний переклад, постредагування, інформативний текст, Google Neural Machine Translation, рівень якості перекладу, тип постредагування.

**Вступ.** Науково-технічний прогрес призводить до зростання потоків інформації, якою обмінюються носії різних мов, що сприяє розширенню міжмовних кордонів і збільшенню обсягу текстів, які потребують перекладу. Сучасні перекладачі вже не можуть користуватися в своїй роботі лише словниками. У перекладознавстві все більше уваги приділяється аналізу інформаційних технологій, що дозволяють оптимізувати процес перекладу.

Постредагування машинного перекладу — відносно новий вид перекладацької діяльності, який стає все більш затребуваним в останні роки. Машинний переклад (далі — МП) є перекладом тексту з однієї природної мови на іншу, що здійснюється за допомогою комп'ютерної програми.

Початок ХХІ століття ознаменувався новим витком стрімкого зростання інтересу до застосування МП в професійному перекладацькому процесі та глобальним розвитком перекладацьких технологій. Використання МП стає великою перевагою при перекладі інформативних текстів, націлених на передачу переважно фактуаль-

но-змістової інформації. МП здатний полегшити й оптимізувати роботу перекладача, забезпечуючи переклад значних обсягів тексту за меншу кількість часу і з меншими зусиллями. Незважаючи на значні досягнення МП в останні роки, появу нейронного МП, помилки й досі є неминучими, тому постредагування у більшості випадків залишається невід'ємним процесом процедури МП.

**Актуальність.** МП взагалі та постредагування зокрема не є новим об'єктом дослідження перекладознавців. Окремі аспекти МП (загальні проблеми, проблема еквівалентності та адекватності, класифікація помилок) неодноразово ставали предметом дослідження науковців (Андієнко, 2005; Нечаева, Светова, 2018; Осіпа, 2008; Стахмич, 2013), проте бракує праць, присвячених вивченню практичних, власне лінгвістичних, аспектів постредагування текстів МП. Виникає потреба класифікації проблем, які виникають під час МП, розробка вимог, що висуваються до постредагування текстів інформативного характеру у межах конкретної мовної пари в залежності від типу постредагування.

**Мета та завдання дослідження.** Мета праці полягає у встановленні специфіки постредагування текстів МП.

У якості завдань праці виокремлюємо такі:

- проаналізувати та класифікувати проблеми, які виникають під час МП;
- встановити релевантність застосування постредагування до результатів МП;
- розробити вимоги, що висуваються до постредагування текстів інформативного характеру у межах англо-української мовної пари в залежності від типу постредагування.

**Матеріали та методи дослідження.** Матеріалом дослідження слугували англomовні тексти офіційно-ділового стилю (офіційні звіти ОБСЄ 2019 року), перекладені українською мовою за допомогою програми машинного перекладу Google Translate. Загальна кількість вихідних текстів дорівнює 1,5 друкованим аркушам. У якості методів дослідження використовувалися лінгвістичний аналіз, метод кількісних підрахунків, зіставний аналіз, перекладознавчо-трансформаційний метод, метод лінгвістичної інтерпретації одержаних даних.

**Результати та дискусія.** Асоціація TAUS (Translation Automation User Society) — асоціація гравців ринку перекладацьких послуг і засо-

бів автоматизації перекладу (<https://www.taus.net/>) дає таке визначення постредагування: «постредагування машинного перекладу — це процес поліпшення результату машинного перекладу за допомогою мінімальних зусиль». Важливою складовою цього визначення є уточнення «за допомогою мінімальних зусиль» та «за мінімально можливою кількістю часу», в іншому випадку ефективність використання машинного перекладу значно знижується і виникає питання, а чи потрібно було його використовувати взагалі.

Зауважимо, що сучасне перекладознавство має дескриптивний характер. Результат перекладацької діяльності, зумовлений скопосом, тобто метою перекладу, відкриває широкі можливості перетворень вихідного тексту, включаючи і такий параметр оцінки якості перекладу, як коректність мовного оформлення перекладеного тексту, який постає критерієм, що передбачає опції.

Враховуючи вищевикладене, сучасна індустрія перекладацьких послуг виділяє два рівня якості перекладу, яких можна досягти в результаті постредагування МП:

- «Задовільна» якість («good enough» quality) передбачає відтворення основної інформації вихідного повідомлення, припускає наявність стилістичних, синтаксичних та морфологічних огріхів.
- Якість перекладу, наближена або рівна якості перекладу, виконаного людиною (quality similar or equal to human translation) — «парадний» варіант відредагованого тексту, придатний для публікації.

Слід зауважити, що у фокусі процесу постредагування знаходиться не сам переклад, а етап, що слідує за ним. Внаслідок цього об'єктом подальшої роботи виступає вже перекладений текст, а початковий текст постає в якості фону і матеріалу для зіставлення.

Згідно з типологією текстів, запропонованою К. Райс, тексти, до яких можна застосувати МП з постредагуванням, відносяться до інформативних текстів (текстів, орієнтованих на інформування), мірою їхньої еквівалентності є інваріантність на рівні плану змісту. Підхід до трансляції МП може бути охарактеризований як функціональний і повністю адекватний «комунікативній прийнятності» (Honig, 1998).

Огляд систем МП дозволяє стверджувати, що найпотужнішою системою МП на сьогоднішній день є Google Neural Machine Translation (GNMT), яка використовує найсучасніші методи навчання для досягнення найбільших удосконалень. В основі системи лежить прин-

цип роботи рекурентних двоспрямованих нейронних мереж, які працюють з матричними обчисленнями ймовірності. Система обчислює значення слова або фрази на основі попередніх значень в послідовності. Саме це дозволяє системі враховувати контекст і правильно обирати серед різних варіантів перекладу. Якщо машинний переклад на основі фрази розбиває вхідне речення на слова і фрази, які в основному перекладаються незалежно, нейронний машинний переклад розглядає все вихідне речення як одиницю перекладу, що значно підвищує якість перекладу.

Під час аналізу текстів, перекладених за допомогою Google Translate, були констатовані такі проблеми:

1. Спотворення референційного змісту вихідного повідомлення:

*Regular restrictions related to disengagement areas and mines/UHO: .....*

*Регулярні обмеження, пов'язані з зонами відключення та шахтами / НЛЮ:.....*

*(UHO (Unexploded Ordnance) — НЛЮ (замість невибухлі боєприпаси).*

*At the eastern entrance of Shevchenko (non-government-controlled, 69 km south of Donetsk), two armed members of the armed formations prevented the SMM from proceeding west citing “kinetic activities in the area”.*

*На східному в'їзді Шевченка (підконтрольний урядові, 69 км на південь від Донецька) двоє озброєних членів збройних формувань перешкодили СММ рухатись на захід, посиляючись на «кінетичну діяльність в цьому районі».*

*They have also agreed that the JCCC should contribute to such response and co-ordinate mine clearance.*

*Вони також погодилися з тим, що СККК повинен сприяти такому реагуванню та координувати розмітання. — процес розмінування*

2. Неправильний добір варіантних відповідностей, відсутність уніфікованості термінів:

*The SMM continued to monitor the withdrawal of weapons in implementation of the Memorandum and the Package of Measures and its Addendum.*

*СММ продовжувала стежити за вилученням зброї під час виконання Меморандуму та пакету заходів та додатку до нього. — відведення*

*On the same day, beside the road between Bohdanivka (government-controlled, 41km southwest of Donetsk) and Viktorivka (government-controlled, 42km south-west of Donetsk)...*

*Того ж дня, поруч із дорогою між Богданівкою (підконтрольний уряду, 41 км на південний захід від Донецька) та Вікторівкою (під контролем уряду, 42 км на південний захід від Донецька)...*

*The SMM continued to monitor the situation at the Petrivske disengagement area.*

*СММ продовжувала стежити за ситуацією в зоні відсторонення Петрівського. — на ділянці розведення сил.*

3. Невідтворення аббревіатур:

*On 18 November, the SMM saw two members of the State Emergency Service (SES) of Ukraine driving south on road T-1316 towards the railway bridge in a white vehicle with the SES logo on the side.*

*18 листопада СММ побачила двох членів Державної служби з надзвичайних ситуацій (ДСНС) України, які їхали на південь по дорозі Т-1316 до залізничного мосту у білому транспортному засобі з логотипом SES збоку.*

*...and 15 personnel of the State Emergency Service (SES) of Ukraine...*

*...та 15 чоловік особового складу Державної служби з надзвичайних ситуацій (СЕС) України, ... — Державної служби України з надзвичайних ситуацій (ДСНС)*

*On the evening of 14 November, the SMM camera in Zolote recorded two explosions assessed as airbursts about 3–5 km north-east and a projectile in flight about 2–4 km north-east.*

*Увечері 14 листопада камера СММ у Золоті зафіксувала два вибухи, оцінені як повітряні вибухи приблизно на відстані 3–5 км на північний схід та снаряд у польоті приблизно на 2–4 км на північний схід. — СММ (Служба Моніторингової Місії)*

4. Неузгодженість мовних одиниць за особами, числами і відмінками:

*On the same day, positioned on the southern edge of the disengagement area, the SMM saw 17 workers and seven deminers from non-government-controlled areas with metal detectors, and three members of the armed formations (wearing armbands with “JCCC” written on them) entering the disengagement area.*

*Того ж дня, розмістившись на південному краю зони відключення, СММ побачила 17 робітників та семеро розмівальників з підконтрольних державам районів із металодетекторами та троє членів збройних формувань (носили пов'язки з написом «JCCC») в'їжджаючи в зону відключення.*

On 15 November, at the latter hospital in Donetsk city, the SMM saw a man (54 years old) with four bandages — two on each thigh.

15 листопада в останній лікарні міста Донецька СММ побачила чоловіка (54 роки) з чотирма пов'язками — по два на кожне стегно. — **по дві**

On 15 November, positioned about 2km north of Petrivske (non-government-controlled, 41km south of Donetsk), the SMM heard three explosions about 2km south-west, assessed as controlled detonations, inside the disengagement area near Petrivske.

15 листопада, розташований приблизно в 2 км на північ від Петрівського (непідконтрольний уряду, 41 км на південь від Донецька), СММ почула три вибухи приблизно в 2 км на південний захід, оцінені як контрольовані детонації, у зоні відключення поблизу Петрівське. — **розташована**

5. Неправильний добір функціонального аналога при відтворенні абсолютних, герундіальних та дієприкметникових конструкцій:

On 19 November, inside the disengagement area near Stanytsia Luhanska (governmentcontrolled, 16 km north-east of Luhansk), the SMM observed at least ten workers from government-controlled areas painting sections of the new bridge and removing scaffolding.

19 листопада у зоні відводу поблизу Станиці Луганської (підконтрольний уряду, 16 км на північний схід від Луганська) СММ спостерігала щонайменше десять робітників з підконтрольних урядам районів, фарбуючи ділянки нового мосту та знімаючи риштування. — **які фарбували секції нового моста та знімали будівельні ліси**

On the same day, inside the disengagement area, the SMM saw 17 workers from non-government-controlled areas clearing vegetation east and west of road T-1316, about 700 m north of the area's southern edge.

Того ж дня у зоні відсторонення СММ побачило 17 робітників з недержавних підконтрольних районів очищення рослинності на схід і захід від дороги Т-1316, приблизно в 700 м на північ від південного краю району. — **як 17 працівників** із непідконтрольних урядові районів **видаляли рослинність** на схід та захід

Mission saw at least four journalists with cameras recording the construction works on the bridge and five workers from government-controlled areas installing four stationary radiation monitoring systems, two at each end of the EECP.

Місія побачила щонайменше чотирьох журналістів із камерами, що фіксують будівельні роботи на мосту, та п'ятьох працівників з облас-

тей, що контролюються державою, встановлюючи чотири стаціонарні системи радіаційного моніторингу, по два на кожному кінці ЕЕС. — **які встановлювали**

6. Буквальний переклад словосполучень, відсутність застосування трансформацій граматичної структури вихідного повідомлення (додавань, перестановок):

In Donetsk region, the SMM recorded more ceasefire violations, including, however, fewer explosions (about 95), compared with the previous reporting period (about 175 explosions). The majority of ceasefire violations, including the majority of explosions were recorded in areas south-east and east-south-east of Svitlodarsk (government-controlled, 57 km north-east of Donetsk) and north and north-north-east of Shyrokyne (government-controlled, 100 km south of Donetsk).

У Донецькій області СММ зафіксувала більше порушень режиму припинення вогню, включаючи, однак, меншу кількість вибухів (близько 95) порівняно з попереднім звітним періодом (близько 175 вибухів). Більшість порушень режиму припинення вогню, включаючи більшість вибухів, було зафіксовано в районах південний схід та схід-південний схід Світлодарська (підконтрольне уряду 57 км на північний схід від Донецька) та північ і північ-північний схід від Широкине (уряд-контрольовано, 100 км на південь від Донецька). **на південний схід та східно-південний схід від Світлодарська**

The majority of ceasefire violations were recorded in areas south-west of Kalynove-Borshchuvate (non-government-controlled, 61 km west of Luhansk) and close to the disengagement area near Zolote (government-controlled, 60km west of Luhansk), where the majority of explosions were recorded.

Більшість порушень режиму припинення вогню було зафіксовано в південному заході Калинове-Борщуваті (непідконтрольне уряду, 61 км на захід від Луганська) та поблизу району відключення поблизу Золоте (підконтрольне уряду, 60 км на захід від Луганська), де більшість зафіксовано вибухи. — **де було зафіксовано більшість вибухів**

On 15 November, positioned about 2km north of Petrivske (non-government-controlled, 41 km south of Donetsk), the SMM heard three explosions about 2 km south-west, assessed as controlled detonations, inside the disengagement area near Petrivske.

15 листопада, розташований приблизно в 2 км на північ від Петрівського (непідконтрольний уряду, 41 км на південь від Донецька), СММ

почула три вибухи приблизно в 2 км на південний захід, оцінені як контрольовані детонації, у зоні відключення поблизу Петрівське. — на відстані 2 км

З урахуванням класифікованих проблем машинного перекладу, а також рівней якості, які висуваються до постредагування, було проведено постредагування текстів, перекладених засобами машинного перекладу, та сформульовані вимоги до постредагування результатів машинного перекладу.

Під час проведення першого етапу постредагування, спрямованого на досягнення «задовільної» якості, були внесені такі корективи у тексти машинного перекладу:

1. виправлення спотвореного референційного змісту вихідного повідомлення:

*At the eastern entrance of Shevchenko (non-government-controlled, 69km south of Donetsk), two armed members of the armed formations prevented the SMM from proceeding west citing “kinetic activities in the area”.*

*На східному в'їзді Шевченка (підконтрольний урядові, 69 км на південь від Донецька) двоє озброєних членів збройних формувань перешкодили СММ рухатись на захід, посилаючись на «кінетичну діяльність в цьому районі».*

*На східному в'їзді до п.м.т. Шевченка (непідконтрольний урядові, у 69 км на південь від Донецька) двоє озброєних членів збройних формувань перешкодили СММ рухатись на захід, посилаючись на «кінетичну діяльність в цьому районі».*

*The SMM continued to monitor the withdrawal of weapons in implementation of the Memorandum and the Package of Measures and its Addendum.*

*СММ продовжувала стежити за вилученням зброї під час виконання Меморандуму та пакету заходів та додатку до нього. — відведення*

*СММ продовжувала стежити за відведенням зброї під час виконання Меморандуму та комплексу заходів та додатку до нього.*

2. Пошук коректних варіантних відповідностей, уніфікація термінів:

*On 15 November, inside the disengagement area near Stanytsia Luhanska (government-controlled, 16 km north-east of Luhansk), the SMM observed ten workers from government-controlled areas painting the Stanytsia Luhanska bridge (15 km north-east of Luhansk) and compacting earth below the bridge.*

*15 листопада у зоні відключення поблизу Станиці Луганської (підконтрольний урядові, 16 км на північний схід від Луганська) СММ спостерігала за десятьма робітниками з підконтрольних уряду територій, які малювали Станицький Луганський міст (15 км на північний схід від Луганська) та ущільнювали землю нижче мосту.*

*15 листопада у зоні відключення поблизу Станиці Луганської (підконтрольний урядові, 16 км на північний схід від Луганська) СММ спостерігала за десятьма робітниками з підконтрольних уряду територій, які фарбували Станицький Луганський міст (15 км на північний схід від Луганська) та утрамбовували землю нижче мосту.*

3. Експлікація абревіатур під час першого вживання у межах документу:

*Compared with the previous reporting period, the SMM recorded more ceasefire violations both in Donetsk and Luhansk regions.*

*Порівняно з попереднім звітним періодом, Спеціальна Моніторингова Місія (СММ) зафіксувала більше порушень режиму припинення вогню як у Донецькій, так і в Луганській областях.*

4. Узгодження за особами, числами і відмінками

*Ceasefire violations recorded by more than one patrol and assessed to be the same are entered only once.*

*Порушення припинення вогню зафіксовано більш ніж одним патруль і оцінені як однакові вводяться лише один раз.*

*Порушення припинення вогню, зафіксовані більш ніж одним патрулем і однаково оцінені, фіксуються лише один раз.*

5. Коректний добір функціонального аналога при відтворенні абсолютних, герундіальних та дієприкметникових конструкцій:

*On the same day, inside the disengagement area, the SMM saw 17 workers from non-government-controlled areas clearing vegetation east and west of road T-1316, about 700 m north of the area's southern edge.*

*Того ж дня у зоні відсторонення СММ побачило 17 робітників з недержавних підконтрольних районів очищення рослинності на схід і захід від дороги Т-1316, приблизно в 700 м на північ від південного краю району.*

*Того ж дня на ділянці розведення спостерігачі бачили, як 17 працівників із непідконтрольних урядові районів видаляли рослинність на схід та захід від автошляху Т1316, орієнтовно за 700 м на північ від південного краю ділянки розведення.*

6. Трансформація граматичної структури вихідного висловлення у разі спотворення тема-рематичної структури вихідного повідомлення внаслідок буквального відтворення, додавання окремих елементів у структуру словосполучення у разі відсутності узгодженості між компонентами внаслідок копіювання вихідної структури:

*A member of medical staff told the SMM that a bullet was removed from the man's thigh.*

*Член медичного персоналу повідомив СММ, що куля була вилучена з стегна чоловіка.*

*Член медичного персоналу повідомив СММ... що зі стегна чоловіка було вилучено кулю.*

*On 14 November, medical staff at Hospital No. 14 in Donetsk told the SMM that a man (in his fifties) had been admitted on the afternoon of 13 November with injuries to his thigh, close to the artery, caused by small-arms fire and was transferred to the Regional Hospital in Donetsk city for further treatment.*

*14 листопада медичний персонал лікарні № 14 у Донецьку повідомив СММ, що чоловіка (у його п'ятдесяти роки) було прийнято в другій половині дня 13 листопада з травмами стегна, близько до артерії, спричиненими вогнем стрілецької зброї. та переведений до обласної лікарні міста Донецька для подальшого лікування.*

*14 листопада медичний персонал лікарні № 14 у Донецьку повідомив спостерігачам про те, що у другій половині 13 листопада до них доправили чоловіка (у віці від 50 до 59 років) з пораненням стегна (поруч з артерією), отриманим в результаті вогню зі стрілецької зброї.*

Задля досягнення «задовільної» якості результату перекладу з урахуванням постредагування були сформульовані такі вимоги до постредагування результатів машинного перекладу:

- Не додавати і не видаляти інформацію з тексту.
- Дбати про семантичний інваріант перекладу.
- Здійснювати прагматичну адаптацію явищ, що є незрозумілими потенційному реципієнту тексту перекладу.
- Не редагувати синтаксичні структури речень.
- Не редагувати стилістичні огріхи.
- Терміни залишати як є (за винятком випадків дослівного машинного перекладу чи включення до тексту перекладу термінів в оригінальній формі).

• Редагувати лише очевидні помилки в перекладі словосполучень.

• Стежити за одноманітністю вживання термінів, одиниць виміру.

• Редагувати граматичні помилки, особливо звернути увагу на узгодження за особами, числами і відмінками, відтворення абсолютних, герундіальних та дієприкметникових конструкцій.

Під час проведення другого етапу постредагування, спрямованого на досягнення якості постредагування, наближеної до якості перекладів, виконаних людиною, виправлення, внесені у рамках першого етапу, були доповнені такими додатками:

1. Додавання, зумовлені стилістичними вимогами мови перекладу у межах офіційно-ділового стилю:

*15 листопада, розташований приблизно в 2 км на північ від Петрівського (непідконтрольний уряду, 41 км на південь від Донецька), СММ почула три вибухи приблизно в 2 км на південний захід, оцінені як контрольовані детонації, у зоні відключення поблизу Петрівське.*

*15 листопада, розташований приблизно в 2 км на північ від Петрівського (непідконтрольний уряду, 41 км на південь від Донецька), СММ почула три вибухи приблизно на відстані 2 км на південний захід, оцінені як контрольовані детонації, у зоні відключення поблизу Петрівського.*

*The SMM's monitoring and freedom of movement are restricted by security hazards and threats, including risks posed by mines, ...*

*Моніторинг СММ та свобода пересування обмежуються небезпеками безпеки та загрозами, включаючи ризики, спричинені мінами,...*

*Моніторинг СММ та свобода пересування обмежені загрозами у сфері безпеки; включаючи ризики, пов'язані з наявністю мін...*

2. Заміна еквівалентів синонімічними відповідниками задля усунення повторів:

*The Mission continued to facilitate repairs to and the operation of critical civilian infrastructure on both sides of the contact line.*

*Restrictions of the SMM's freedom of movement continued, including at checkpoints near non-government-controlled Verkhnohyrovivske and Kruta Balka.*

*СММ продовжувала сприяти проведенню ремонтних робіт та забезпеченню функціонування цивільних об'єктів життєзабезпечення з обох сторін лінії зіткнення.*

Обмеження свободи пересування **спостерігачів** тривали, зокрема на блокпостах біля непідконтрольних урядові Верхньошироківського та Крутої Балки.

3. Комплексне перетворення синтаксичної структури вихідного висловлення:

*The Mission saw a truck leaving the area loaded with cut wood, and thereafter re-entering it empty.*

*Команда СММ бачила завантажену зрубаною деревиною вантажівку, що виїхала з ділянки та пізніше повернулася порожньою.*

4. Ретельна перевірка коректності відтворення термінів з урахуванням загальноприйнятого варіанту перекладу:

*The SMM continued to monitor the situation at the Petrivske **disengagement area**.*

*Місія продовжила спостерігати за ситуацією на **ділянці розведення сил і засобів** у районі Петрівського.*

Для досягнення якості перекладу, близької до «людської», вимоги, що сформульовані до задовільного рівня якості, були доповнені рекомендаціями, що стосуються граматики, стилістики, рекомендаціями з перевірки термінів, вимогами дотримання формальних правил щодо оформлення:

- Не додавати і не видаляти інформацію з тексту.
- Дбати про семантичний інваріант перекладу.
- Здійснювати прагматичну адаптацію явищ, що є незрозумілими потенційному реципієнту тексту перекладу.
- Редагувати очевидні стилістичні огріхи.
- Редагувати синтаксичні структури речень (у разі потреби).
- Перевіряти вживання термінів.
- Перевіряти переклад словосполучень.
- Стежити за одноманітністю вживання термінів, одиниць виміру.
- Редагувати граматичні помилки, особливо звернути увагу на узгодження за особами, числами і відмінками, відтворення абсолютних, герундіальних та дієприкметникових конструкцій.
- Стежити за пунктуацією.
- Стежити за форматуванням тексту.

**Висновки.** Проведений аналіз труднощів машинного перекладу дозволяє констатувати таке: найскладнішим для МП виявляється узгодження мовних одиниць за особами, числами і відмінками та не-

коректний добір однієї з варіантних відповідностей (майже 25 % усіх зафіксованих помилок), що пояснюється відсутністю та/або неспівпадінням зазначених категорій у межах досліджуваної мовної пари (англійська-українська). Труднощі у доборі варіантного відповідника пов'язані із багатоаспектністю майже будь-якого вислову та необхідністю конкретизування закладеного змісту у межах широкого та вузького контексту.

Друге місце (13–14 %) займають спотворення змісту вихідного повідомлення, буквальний переклад та некоректний добір функціонального аналога граматичним конструкціям вихідної мови. Ці труднощі пояснюємо значними розбіжностями граматичних структур неблизькоспоріднених мов у межах мовної пари.

Наймеша кількість помилок пов'язана з невідтворенням аббревіатур.

Вимоги, які висуваються до результатів машинного перекладу, звичайно, відрізняються від перекладу, який виконує людина. Переклад, виконаний комп'ютером, далекий від ідеального, але текст, отриманий в результаті такого опрацювання, дозволяє у більшості випадків зрозуміти суть перекладеного тексту. Незважаючи на значний розвиток систем МП, текст МП потребує доопрацювання людиною.

Сформульовані вимоги до постредагування результатів машинного перекладу урахують бажану якість кінцевого результату перекладу.

У якості перспектив дослідження вбачаємо проведення оцінки ефективності запропонованих вимог у порівнянні з постредагуванням, виконаним без будь-яких рекомендацій.

Зауважимо, що вимоги та рекомендації повинні бути складені для кожної окремої мовної пари з урахуванням особливостей конкретної системи МП і текстів, що перекладаються.

## ЛІТЕРАТУРА

- Андрієнко Л. О. Проблеми розвитку машинного перекладу на сучасному етапі. *Гуманітарний вісник*. Серія: Іноземна філологія. Черкаси : ЧДТУ, 2005. № 27. С. 348–351.
- Нечаева Н. В., Светова С. Ю. Постредактирование машинного перевода как актуальное направление подготовки переводчиков в вузах. *Вопросы методики преподавания в вузе*. 2018. Т. 7, № 25. С. 64–72.
- Осіпа Л. В. Комп'ютерний переклад тексту за допомогою системи машинного перекладу PRAGMA. *Інформатика та інформаційні технології в навчальних закладах*. 2008. № 1 (13). С. 14–19.

Стахмич Ю. С. Адекватність та еквівалентність перекладу в контексті комп'ютерної лінгвістики. *Вісник Житомирського державного університету імені Івана Франка. Житомир*. 2013. Вип. 66. С. 235–238.

Honig H. G. Positions, Power and Practice: Functionalist Approaches and Translation Quality Assessment. *Translation and Quality* / C. Schaffner (ed.). Philadelphia: Multilingual Matters, 1998. P. 6–34.

## ПОСТРЕДАКТИРОВАНИЕ ПРИ МАШИННОМ ПЕРЕВОДЕ

### **Татьяна Королева**

доктор филологических наук, профессор, заведующая кафедрой перевода и теоретической и прикладной лингвистики Государственного учреждения

«Южноукраинский национальный педагогический университет имени К. Д. Ушинского»

Одесса, Украина

e-mail: kortami863@gmail.com

ORCID ID <https://orcid.org/0000-0003-3441-196X>

### **Наталья Жмаева**

кандидат филологических наук, доцент кафедры перевода и теоретической и прикладной лингвистики Государственного учреждения «Южноукраинский национальный педагогический университет имени К. Д. Ушинского»

Одесса, Украина

e-mail: zhmaeva@gmail.com

ORCID ID <https://orcid.org/0000-0003-3382-0155>

### **Юлия Колчаг**

магистр филологии Государственного учреждения «Южноукраинский национальный педагогический университет имени К. Д. Ушинского»

Одесса, Украина

e-mail: juliaigorivna@ukr.net

### **АННОТАЦИЯ**

Современная индустрия переводческих услуг выделяет два уровня качества перевода, которых можно достичь в результате постредактирования МП (машинного перевода): «удовлетворительное» качество (“good enough quality”) предусматривает воссоздание основной информации исходного сообщения, допускает наличие стилистических, синтаксических и морфологических огрехов; качество перевода, приближенное или равное качеству перевода, выполненного человеком (“quality similar or equal to human translation”) — «парадный» вариант отредактированного текста, пригодный для публикации.

Обзор систем МП позволяет утверждать, что самой мощной системой МП на сегодняшний день является Google Neural Machine Translation (GNMT), которая использует самые современные методы обучения для достижения максимальных усовершенствований.

При анализе текстов, переведенных с помощью Google Translate, были констатированы следующие проблемы: искажение референциального содержания исходного сообщения, неправильный подбор вариантных соответствий, отсутствие унифицированности терминов, отсутствие передачи аббревиатур, несогласованность языковых единиц по лицам, числам и падежам, неправильный подбор функционального аналога при воспроизведении абсолютных, герундиальных и причастных конструкций, буквальный перевод словосочетаний, отсутствие трансформаций грамматической структуры исходного сообщения (добавлений, перестановок).

Учитывая классифицированные проблемы машинного перевода, а также уровни качества, предъявляемые к постредактированию, проведено постредактирование текстов, переведенных посредством машинного перевода, и сформулированы требования и рекомендации к постредактированию результатов машинного перевода в пределах исследуемой языковой пары с учетом особенностей конкретной системы МП и типа переводимых текстов.

**Ключевые слова:** машинный перевод, постредактирование, информативный текст, Google Neural Machine Translation, уровень качества перевода, тип постредактирования.



## POSTEDITING IN MACHINE TRANSLATION

**Tetiana Korolova**

Doctor of Philology, Professor, Head of the Department of Translation, Theoretical and Applied Linguistics, State Institution “South Ukrainian National Pedagogical University named after K. D. Ushynsky”  
Odesa, Ukraine  
e-mail: kortami863@gmail.com  
ORCID ID <https://orcid.org/0000–0003–3441–196X>

**Natalya Zhmayeva**

Ph.D. in Philology, Associate Professor at the Department of Translation and Theoretical and Applied Linguistics, State Institution «South Ukrainian National Pedagogical University named after K. D. Ushynsky»  
Odesa, Ukraine  
e-mail: zhmaeva@gmail.com  
ORCID ID <https://orcid.org/0000–0003–3382–0155>

**Yulia Kolchah**

Master of Arts in Philology, State Institution «South Ukrainian National Pedagogical University named after K. D. Ushynsky»  
Odesa, Ukraine  
e-mail: juliaigorivna@ukr.net

### SUMMARY

*Modern industry of translation services singles out two translation quality levels that can be reached as a result of machine translation (MT) post-editing: good enough quality foresees rendering the main information of the source message, admitting stylistic, syntactic and morphological flaws while quality similar or equal to human translation is a full dress version of a post-edited text, ready to be published.*

*The overview of MT systems enables us to consider Google Neural Machine Translation (GNMT) which is based on the most modern methods of training to reach maximum improvements the most powerful one.*

*When analyzing texts translated by means of Google Translate the following problems were identified: distortion of the referential meaning of the source message, incorrect choice of variant equivalences, lack of terms harmonization, lack of abbreviations rendering, inconformity of linguistic units in persons, numbers and cases, incorrect choice of functional correspondings when rendering absolute constructions, gerund and participial constructions, literal translation of phrases, lack of transformations of the grammatical structure of the source message (additions, rearrangements).*

*Taking into account the classified issues of machine translation as well as the levels of post-editing quality post-editing of the texts translated by means of MT*

*is carried out, demands and recommendations applicable to post-editing results of MT within the language pair under analysis with respect to peculiarities of the specific MT system and the type of translated texts are provided.*

**Key words:** machine translation, informative text, Google Neural Machine Translation, translation quality level, type of post-editing.

## REFERENCES

- Andriienko, L. O. (2005) Problemy rozvytku mashynoho perekladu na suchasnomu etapi [The Issues of Machine Translation Development at the Present Stage], Humanitarnyi visnyk. Seriia: Inozemna filolohiia Cherkasy, 348–351 [in Ukrainian].
- Nechaeva, N. V., Svetova, S. Yu. (2018) Postredaktirovanie mashinnogo perevoda kak aktualnoe napravlenie podgotovki perevodchikov v vuzah [Post-Editing Machine Translation as a New Activity for Teaching Translation at Universities], Teaching Methodology in Higher Education, 7, 25, 64–72 [in Russian].
- Osipa, L. V. (2008) Kompiuternyi pereklad tekstu za dopomohoiu systemy mashynoho perekladu RRAGMA [Computer Assisted Translation by Means of Pragama], Informatyka ta informatsiini tekhnolohii v navchalnykh zakladakh, 1 (13), 14–19 [in Ukrainian].
- Stakhmych, Yu. S. (2013) Adekvatnist ta ekvivalentnist perekladu v konteksti kompiuternoi linhvistyky [Adequacy and Equivalence of Translation in the Context of Computer Linguistics], Visnyk Zhytomyrskoho derzhavnogo universytetu imeni Ivana Franka, 66, 235–238 [in Ukrainian].
- Honig H. G. (1998) Positions, Power and Practice: Functionalist Approaches and Translation Quality Assessment // Translation and Quality / C. Schaffner (ed.). — Philadelphia: Multilingual Matters, 6–34.

*Стаття надійшла до редакції 18.11.2019*