

**Бірюков А.В.**  
**КОМП'ЮТЕРНИЙ ПЕРЕКЛАД ТА ОЦІНКА ЙОГО ЯКОСТІ**

УДК 81'322.4

Оцінка якості комп'ютерного (машинного) перекладу прямо або опосередковано цікавить досить широке коло осіб: користувачів систем, споживачів продуктів подібних систем, всіх тих, хто працює над розробкою та покращенням системи, значною мірою і тих, хто вирішує, чи застосовувати йому такі системи в їх сфері діяльності. Оцінка, звичайно, є цікавою і для перекладачів-практиків – незалежно від їх мотивів – чи то бажання бути поінформованим про стан та перспективи галузі з позиції своєї власної діяльності, чи намірів щодо майбутньої діяльності в галузі машинного перекладу – наприклад, у ролі постредактора. Безумовно, поняття „оцінка” є важливим й для науковців, особливо для тих, хто займається науковою діяльністю в галузі перекладу і професійна діяльність яких майже щоденно пов'язана з проблематикою оцінки та критики перекладу.

Існує чимала кількість експериментів з оцінки якості, проведених на основі тієї чи іншої системи перекладу (Брокман, 1995; Канічев, 1998; Кехлінг, 1993; Москович, 1970; Осієвський, 2004; Пилипенко, 2002; Суханова, 1997; Тчеке, 1996). Проте, слід констатувати, що всі вони містять різні методологічні підходи до питання оцінки якості, або такі підходи відсутні взагалі, а оцінка здійснюється на основі механічного підрахунку помилок. Питання систематизації та критичного огляду методів оцінки якості машинного перекладу, їх узагальнення залишається й досі не вирішеним.

Поділ описаних нижче методів оцінки здійснено в межах двох основних рівнів, що розрізняються загалом у фаховій літературі: макрооцінки, або адекватної, чи декларативної оцінки та мікрооцінки, або діагностичної, чи типологічної оцінки. Надалі ми використовуватимемо для обох рівнів терміни макро- та мікрооцінка.

Лінгвістичну оцінку або мікрооцінку, що полягає у виявленні помилок та пошуку можливостей їх усунення, бажано починати після визначення низки суттєвих питань, таких як вид та частота мовних явищ у тексті, необхідність та обсяг використання термінології та передредагування. Мікрооцінка складається з п'яти рівнів аналізу: граматичного, формального та каузального рівня, рівня „можливостей покращення” і рівня „реального покращення” та надає факти, які характеризують з різних точок зору роботу системи і якість виконаних перекладів.

На *граматичному* рівні виявляється частота помилок граматичного, синтаксичного та семантичного характеру, що були виправлені під час постредагування. Для цього помилки класифікуються на типи (наприклад „помилка у підметі” або „помилка у присудку”) та порівнюються із загальною кількістю підметів, присудків тощо. Деякі оцінювачі додатково класифікують помилки за рівнем їх „ваги”.

На *формальному* рівні визначається обсяг роботи з постредагування: наприклад вилучення певних слів або словосполучень, заміни, доповнення, перестановки. Вимір обсягу постредагування є корисним критерієм, який надає важливу інформацію для визначення макрооціночних показників швидкості редагування та зрозумілості. Проте, слід пам'ятати, що на цьому формальному рівні досліджується власне не система комп'ютерного перекладу, а техніка виправлень під час коригування.

На *каузальному* рівні помилки перекладу підпорядковуються певним етапам роботи системи, які є ймовірною причиною їх виникнення – наприклад, етапу аналізу, трансферу або синтезу. Оскільки аналіз помилок здійснюється тут лише на основі неправильних перекладів, класифікація відбувається лише в рамках припущень, що за оцінки методом „чорної скриньки” є досить складною та сміливою процедурою. Доцільнішим видається спостереження за роботою системи або хоча б контроль результатів окремих кроків програми – звичайно, наскільки це можливо в рамках інтерактивного спілкування із системою. Ф. Шефер дещо розширює рамки каузального рівня і відносить сюди всі методи, пов'язані з класифікацією та аналізом помилок машинного перекладу на основі будь-якого критерію. Такі методи він називає діагностичними [7, с. 185]. Щодо перебігу процесу оцінювання на основі таких критеріїв, то єдиної думки серед науковців немає. Так, Дж. Лербергер і Л. Борбю в „лінгвістичній оцінці, яка виконується користувачем”, наполягають на детальному визначенні окремих помилок з їх подальшою класифікацією на основі певної типології [6, с. 150]. М. Васконселос, навпаки, не визнає формального визначення помилок перекладу, оскільки вважає, що класифікація помилок за граматичними критеріями сама по собі не надає жодних даних щодо причин їх виникнення. А оскільки мікрооцінка має на меті детальне дослідження причин виникнення помилок та пошук можливостей їх усунення, ці критерії є абсолютною непридатними [9, с. 205]. Замість цього вона пропонує визначення помилок на каузальному рівні, оскільки таким чином можна провести хоча б приблизну діагностику джерел і причин виникнення помилок та зробити перші висновки щодо можливостей покращення системи. На наш погляд, можна погодитися з обома підходами, проте, без беззаперечного наслідування жодного з них. М. Васконселос, з одного боку, безперечно, має рацію, стверджуючи, що формально-граматична категоризація помилок не відповідає етапам роботи системи, які оцінюються, і таким чином не надає жодних даних про причини їх виникнення. З іншого боку, дослідниця заходить надто далеко, пропонуючи зовсім вилучити лінгвістичну типологію помилок, яку застосовують Дж. Лербергер та Л. Борбю та використовувати замість цього лише категоризацію на основі вірогідних причин помилок. Такий підхід залишає поза увагою той факт, що робочим матеріалом будь-якої системи перекладу є тексти, тобто форми застосування природної мови. Тому, на наш погляд, послідовним було б врахування цього аспекту в методологічних положеннях до мікрооцінки. Нарешті, лише врахування лінгвістичних феноменів, досить часто вживаних у певному виді тексту, дозволяє зробити

висновки щодо можливостей опрацювання цих феноменів системою, яка оцінюється. І саме це може надати відповідь на дуже важливі питання будь-якої оцінки:

1. Чи підходить взагалі система, що оцінюється, для перекладу конкретних текстів? Якщо так, то наскільки?

2. Якщо підходить, то які компоненти системи при перекладі цих текстів викликають найсуттєвіші помилки і внаслідок цього мають бути покращені в першу чергу?

Четвертим є рівень *“можливості покращення”* системи. Припущені помилки на цьому рівні розглядаються в контексті їх усунення – наприклад, за рахунок поповнення словника, або – у комерційних системах таке можливо лише інколи – завдяки покращенню алгоритмів лінгвістичного апарату. Докладне дослідження методів оцінки якості на рівні теоретичних можливостей покращення системи виконали також Д. Брокман [3] та Г. Каспарі [4]. На нашу думку, подібна оцінка є дуже корисною, оскільки вона надає можливість опрацювання концептів вирішення певних проблем перекладу, а також визначає обсяг роботи, який для цього потрібен. Зрештою, таким чином має бути винайдена загальна стратегія покращення оцінюваної системи.

Останнім є рівень *„реального покращення”* системи. На цьому рівні після спроби усунення помилок текст перекладається вдруге, після чого знову проводиться аналіз для виявлення модифікацій, які дійсно призвели до покращення якості перекладу. Оцінка якості перекладу після проведення робіт з покращення системи вважається також важливим методом, оскільки таким чином отримуємо дані про покращення ефективності роботи системи. Недоліком цього методу, на нашу думку, є його висока вартість і той факт, що перевірка внесених до системи змін на основі ідентичних текстів є не дуже об'єктивною, оскільки отримані таким чином висновки стосуються лише цього конкретного текстового матеріалу і не завжди можуть бути використані як аргумент на користь загального покращення якості комп'ютерного перекладу.

На закінчення розгляду рівнів мікрооцінки слід зазначити, що лінгвістична оцінка, яка одночасно складається з симптоматики, діагностики, прогнозу та терапії помилок перекладу на основі практичного досвіду, надає критиці машинного перекладу солідний інструментарій щодо критеріїв постредування взагалі та видів коригування зокрема, особливо якщо вони практично співвідносяться з пропозиціями щодо покращення роботи системи.

Представлена нами лінгвістична оцінка або мікрооцінка є лише частиною оцінки системи, що проводиться в рамках орієнтованої на користувача макрооцінки. Завдання макрооцінки полягає в тому, щоб виявити, чи відповідають ї, якщо так, то якою мірою, виконані системою переклади відповідним вимогам користувача. Йдеться про специфічне професійне середовище користувача системи та такі поняття, як його робоче оточення, технічне обладнання й процес створення перекладу. У застосуванні макрооцінки розрізняють також декілька рівнів дослідження: когнітивний, економічний, лінгвістичний та операційний. Д. Брокман пропонує розрізнити ще технічний рівень, який складається з необхідних передумов для встановлення певної СМП, оцінки зручності користування та визначення фінансових переваг/недоліків від використання СМП. Макрооцінка Д. Арнольда й А. Тчеке складається з операційного, декларативного та лінгвістичного рівнів [1, с. 5; 8, с. 29]. Знову спостерігається розбіжність у термінології, проте по суті мова йде про ті самі методи.

Традиційно до критеріїв *когнітивного* рівня відносять зрозумілість, правильність перекладу, ступінь його корисності та прийнятності. В. Гатчінз та Х. Сомерс вважають за доцільне віднести до цього рівня також критерій стилістичної коректності, якому вони надають суттєвої ваги. Проте, на нашу думку, у такому перекладі не втрачається ані зрозумілість перекладу, ані (за великим рахунком) його правильність. Тому додавання до когнітивного рівня (в рамках якого йдеться лише про прийнятність або неприйнятність перекладу для користувача) критерію стилістичної коректності нам видається недоречним.

Зрозумілість тексту визначається його прозорістю та чіткістю для середньостатистичного читача з обмеженими фаховими знаннями без доступу до оригіналу на основі безпосередньо отриманої інформації. Правильність перекладу визначає обсяг інформації оригіналу, що відображена в тексті перекладу. З методичної точки зору встановлення правильності перекладу потребує залучення оригіналу, а в ідеалі – його ретельний аналіз змісту. Під придатністю розуміють практичне застосування або конкретну користь від комп'ютерного перекладу для кінцевого користувача. Цей критерій якості може бути виявлений шляхом безпосереднього опитування споживача, чії враження фіксують на основі каталогу питань. Критерій прийнятності характеризує готовність користувача визнати машинний переклад як такий, що відповідає його вимогам. Цей критерій оцінки тісно переплітається з критерієм придатності, тому деякі дослідники не розрізняють їх і розглядають разом. Підсумовуючи критерії когнітивного рівня, зазначимо, що його суб'єктивність полягає у виборі питань та пробандів, оскільки не всі люди усвідомлюють однакову кількість інформації та не ту саму інформацію з певного тексту за певний проміжок часу. До того ж, дослідження на когнітивному рівні дуже важко реалізувати та систематизувати, оскільки вони потребують багато часу та коштів, часто є незрозумілими і майже завжди необ'єктивними.

У рамках *економічного* рівня визначається час перекладу, читання та редагування, і залишається без уваги оцінка коштів на придбання системи, технічне обслуговування та актуалізацію, оскільки вони не стосуються безпосередньо якості продукту. Під часом перекладу розуміють часовий відрізок, необхідний для створення перекладу, починаючи від отримання оригіналу до передачі цільового тексту замовнику. Під часом, необхідним на коригування, розуміють проміжок часу між отриманням комп'ютерного перекладу постредактором та закінченням процесу редагування. Час редагування співвідноситься з видом та кількістю виправлень машинного тексту, що може включати сюди також залучення оригіналу та пошук необхідної термінології. Час прочитання – це швидкість, з якою може бути прочитаний і усвідомлений машинний

переклад. Усі три чинники вважаються важливими критеріями оцінки якості, оскільки час, необхідний для виправлення помилок та прочитання всього тексту, дозволяє зробити досить достовірні висновки щодо якості перекладу, а час перекладу є важливим моментом, який намагається зрівноважити високу швидкість машинного перекладу з якістю мануального. Підсумовуючи, можна констатувати, що всі критерії економічного рівня є відносно об'єктивними, проте поводитися з ними – особливо з таким, як швидкість перекладу – слід дуже обережно.

На *лінгвістичному* рівні, який полягає в підрахунку та класифікації помилок, має бути досліджений характер подібності оригіналу в лексичному, граматичному, синтаксичному і семантичному планах. Відмінність мікрооцінки від лінгвістичного рівня макрооцінки полягає в намаганні не лише підрахувати та класифікувати помилки, але й встановити причини їх виникнення та можливості усунення. Доцільність оцінки якості машинного перекладу на основі визначення лінгвістичної типології помилок та їх частоти визнається не всіма дослідниками, оскільки вважається, що мовна якість комп'ютерного перекладу для комерційно орієнтованих користувачів матиме скоріше за все другорядне значення, якщо вона не ставить під сумнів збереження інформації оригіналу. Якщо не брати до уваги фактор суб'єктивності, то основною проблемою орієнтованої на користувача лінгвістичної методики оцінки в рамках як мікро- так і макрооцінки є те, що вона не може бути проведена „середньостатистичним“ користувачем, проте надає цінну й необхідну інформацію щодо очікуваної якості перекладів та можливостей розширення і покращення системи. Виходом із становища може бути створення незалежної комісії з оцінки якості, до складу якої входили б потенційні користувачі, незалежні лінгвістичні експерти та представники установ, чії системи оцінюються. У рамках цієї групи могла б бути організована, наприклад, така співпраця: користувачі могли б надати експертам у формі репрезентативного текстового корпусу свої специфічні лінгвістичні вимоги. Потім експерти утворюють відповідні тестові збірки, на які в подальшому під час оцінки тієї чи іншої системи спираються всі зацікавлені сторони. Представники розробників, у свою чергу, оцінюють результати дослідження для відповідного розширення та покращення системи. Але сьогодні ми ще досить далекі від такого підходу до справи. Проте, саме діалог між названими групами міг би усунути або врешті-решт хоча б скоротити ті непорозуміння, які виникають внаслідок появи в користувачів нереальних сподівань на основі неточної інтерпретації обіцянок розробника.

Останнім рівнем макрооцінки є *операційний* рівень. На ньому визначається справжній перебіг процесу перекладу в системі та рівень її продуктивності. Відбуватися це може, наприклад, таким чином: текст, перекладений за допомогою системи машинного перекладу, потім опрацьовується програмою ідентифікації природної мови, яка підраховує кількість речень, що були визначені нею як „правильні“.

Поза макрооцінкою, яка розглядається у фаховій літературі, залишаються суттєві критерії оцінки, що орієнтовані на користувача. У зв'язку з цим В. Гатчінз та Х. Сомерс виділяють у рамках макрооцінки ще технічну оцінку, оцінку зручності користування, а також визначення коштів на придбання та експлуатацію системи [5, с. 17]. У рамках *технічної* оцінки визначають спочатку особливості програмного забезпечення та апаратні засоби відповідної системи (наприклад, на яких процесорах та на основі яких операційних систем може бути встановлена СМП). Проте, важливішими для технічної оцінки є критерії програмного забезпечення, такі, наприклад, як обмежений обсяг словника, сумісність системи з іншими програмами обробки текстової інформації тощо.

Для теорії машинного перекладу найактуальнішим питанням сучасних та подальших наукових розвідок залишається оцінка його якості, ще більшою мірою – дослідження його хиб та недоліків з метою винайдення нових, нестандартних підходів до вирішення чинних проблем. На нашу думку, створення єдиної, загальновизнаної концепції оцінки якості машинного перекладу є навряд чи можливим. Урахування всіх цих вимог до якості призводитиме до створення занадто обтяжливої оцінки, яка, для певних кіл зацікавлених осіб може виявитися взагалі зайвою. З іншого боку, навіть найповніша оцінка не позбавлена небезпеки відсутності якоїсь специфічної вимоги, властивої для конкретного користувача в його професійному оточенні. Разом з тим великий досвід, накопичений в цій галузі, значна кількість практично проведених оцінок та теоретичних методик можуть надати суб'єктам оцінювання цінні поради. Проте вибір певної методики в кожному окремому випадку має робитися з урахуванням своїх специфічних вимог, які, ймовірно, досить часто не будуть збігатися, що, власне, і доводить аналіз всіх здійснених оцінок, які нам відомі.

#### Джерела та література:

1. Arnold D. Evaluation: an Assessment / D. Arnold, L. Sadler, R. L. Humphreys // Machine Translation: An Introductory Guide. – 1993. – № 8. – P. 1-24.
2. Blatt A. Computer und Übersetzen. Eine Einführung / A. Blatt, K.-H. Freigang, K.-D. Schmitz, G. Thome. – Hildesheim, Zürich, N. Y. : Georg Olms Verlag, 1985. – 331 s.
3. Brockmann D. Was kann Logos? Linguistische Bewertung eines kommerziellen MÜ-System unter Berücksichtigung weiterer markt- und forschungsorientierter Systeme. Untersuchte Sprachrichtung: Deutsch-Englisch / D. Brockmann // Saarbrücker Studien zu Sprachdatenverarbeitung und Übersetzen. – 1995. – Bd. 4. – S. 1-174.
4. Caspari G. Untersuchungen zu Bewertungskriterien für maschinell erstellte Übersetzungen / G. Caspari. – Saarbrücken : IAI, 1999. – 108 s.
5. Hutchins W. J. An Introduction to machine translation / W. J. Hutchins, H. L. Somers. – London : Academic Press, 1992. – 159 p.

6. Lehrberger J. Machine Translation – Linguistic Characteristics of MT Systems and General Methodology of Evaluation / J. Lehrberger, L. Bourbeau. – Amsterdam, Philadelphia : J. Benjamins Publ. Company, 1988. – 214 p.
7. Schäfer F. Die maschinelle Übersetzung von Wirtschaftstexten: eine Evaluierung anhand des MÜ-Systems der EU-Kommission, SYSTRAN, im Sprachenpaar Französisch-Deutsch: Dissertation zur Erlangung des akademischen Grades eines Doktors der Philosophie der Philosophischen Fakultät der Uni Saarbrücken / F. Schäfer. – Frankfurt am Main : Peter Lang, 2002. – 393 s.
8. Tschöke A. Eine linguistische Evaluation der maschinellen Übersetzungssysteme T 1 Professional 3.0 und PT plus 98 anhand der Rohübersetzung von Nominalphrasen / A. Tschöke // Saarbrücker Studien zu Sprachdatenverarbeitung und Übersetzen. – 1996. – Bd. 5. – S. 26-34.
9. Vasconcellos M. Factors in the Evaluation of MT: Formal vs. Functional Approaches / M. Vasconcellos // Technology as Translation Strategy. – 1988. – Vol. 11. – P. 203-213.

**Брідко Т.В.**

**УДК 811.112**

## **ВАРІАНТИ ВИМОВИ НІМЕЦЬКИХ ГОЛОСНИХ [i:] ТА [ɪ] (НА МЕТЕРІАЛІ ЗАПИСІВ МОВЛЕННЯ РОСІЙСЬКОМОВНИХ ПЕРЕСЕЛЕНЦІВ)**

Сучасна соціолінгвістика звертає посилену увагу до питань територіальної та соціальної варіативності мовлення. Соціолінгвістичні розвідки останніх років характеризуються зростанням зацікавленості вчених проблемами соціальної й територіальної диференціації мови (М. П. Кочерган, О. Д. Петренко, Л. І. Прокопова, О. І. Стеріополо, W. Labov, H. Löffler, H. Menge, A. Mihm, P. Trudgill), гендерної та вікової варіативності мовлення (О. І. Горшко, А. В. Кириліна, А. П. Мартинюк, N. Coupland, J. Peters, L. Push, S. Romaine, D. Tannen, S. Trömel-Plötz), білінгвізму й мовних контактів (О. Р. Валігура, Ю. М. Захарова, Н. Frank, A. Goldbach, F. Grosjean, J. Hammers, E. Haugen, C. Myers-Scotton, B. Pabst, A. Reynolds, W. Weinreich). З розширенням міграційних процесів залишається актуальним дослідження особливостей мовлення іммігрантів, у тому числі мігрантів з колишнього СРСР.

Питання взаємодії німецької та російської мов серед російських німців починає привертати увагу лінгвістів з 1990 року у зв'язку зі збільшенням кількості російськомовних мігрантів у Німеччину. Одними з перших у цій галузі були розвідки німецької дослідниці Х. Франк, спрямовані на вивчення розвитку німецьких національних меншин у Росії та в СРСР [Frank 1992]. У подальшому в Німеччині виходить ціла серія досліджень, присвячених аналізу історичних, психологічних і мовних особливостей російських німців, що проживають у країнах колишнього СРСР та ФРН [Frank 1992; Meng 2005; Baur 1999]. Цікавими, зокрема, є роботи німецької дослідниці К. Менг, яка протягом шести років спостерігала за впливом німецької мови на російську в родинах російськомовних переселенців, що протягом декілька років проживали у Німеччині [Meng 2005]. У центрі уваги цих студій стало вивчення багатьох змін і розвитку домінуючої мови в сім'ях російськомовних мігрантів (зміна М1 (російська мова) на М2 (німецька мова)), а також комплексу труднощів, з якими стикалися російськомовні переселенці у ході оволодіння фонетичним і граматичним устроєм німецької мови [Meng 2001].

Ця стаття є частиною комплексного дослідження соціолінгвістичних особливостей німецької вимови переселенців з колишнього СРСР. Джерелом матеріалу є 81 запис інтерв'ю, основна частина яких здійснена в період з 2005 до 2009 року на території Рурської області землі Північний Рейн-Вестфалія (Федеративна Республіка Німеччина) на цифровий диктофон Olympus VN-240PC. Сумарна тривалість звучання аудіоматеріалу складає 32 години 31 хвилину. Матеріалом фонетичного дослідження слугували сегментні одиниці фонетичного рівня німецької мови (30537 фонетичних реалізацій) в підготовленому і в невідповідному мовленні. Цільова група дослідження особливостей вимови російськомовних мігрантів Німеччини складалася з 27 інформантів (12 чоловіків і 15 жінок) різного віку (від 15 до 57 років).

**Метою статті** є дослідження особливостей реалізації голосних переднього ряду високого підняття [i:] та [ɪ] зафіксованих у німецькому мовленні російськомовних переселенців залежно від позамовних факторів (віку та гендеру мовців, ситуації спілкування). Поставлена мета передбачила на сам перед розв'язання таких завдань: виділити диференційні ознаки фонем, що потенційно реалізуються в німецькому мовленні російськомовних переселенців і утворюють їх фонетичну систему; встановити гендерні та вікові особливості німецького мовлення інформантів. **Об'єктом** є німецьке мовлення російськомовних мігрантів у Федеративній Республіці Німеччина. **Предметом дослідження** є соціофонетична варіативність голосних переднього ряду високого підняття [i:] та [ɪ] у німецькому мовленні російськомовних переселенців із колишнього СРСР.

У результаті аудіоаналізу німецької вимови досліджуваних інформантів у трьох ситуаціях спілкування встановлено, що довгий закритий голосний переднього ряду високого підняття [i:] нарівні з випадками стандартної вимови зазнав у мовленні інформантів переважно кількісної редукції (подовження чи скорочення його вимови залежно від позиції в слові, темпу мовлення та ситуації спілкування). Довгота фонем [i:] збільшувалася за наявності буквосполучення <ie> у слові, особливо під час повільного темпу розмови у ситуації неофіційного спілкування. Таких змін зазнав насамперед довгий закритий голосний переднього ряду високого підняття [i:] в таких словах: sie [zi:], diese ['di:ze], Krieg [kri:k<sup>h</sup>], studieren [ʃtu:'di:ren], Miete ['mi:t<sup>h</sup>e]. В інших випадках, переважно при швидкому темпі розмови в ситуації неофіційного спілкування незалежно від віку та статі мовців, спостерігалася тенденція до скорочення довгого голосного [i:]: ihn [ɪn], wir [vɪr]. Можливо, частково ці фонетичні особливості в мовленні мігрантів